

Ю. К. БЕЛЯЕВ

ВЕРОЯТНОСТНЫЕ
МЕТОДЫ
ВЫБОРОЧНОГО
КОНТРОЛЯ



ТЕОРИЯ ВЕРОЯТНОСТЕЙ
И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Ю. К. БЕЛЯЕВ

ВЕРОЯТНОСТНЫЕ
МЕТОДЫ
ВЫБОРОЧНОГО
КОНТРОЛЯ



ИЗДАТЕЛЬСТВО «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
МОСКВА 1975

517.8

Б.43

УДК 519.21

Вероятностные методы выборочного контроля, Беляев Ю. К.,
Главная редакция физико-математической литературы издательства
«Наука», 1975 г.

Методы выборочного обследования применяются при изучении широкого круга проблем. Примерами могут служить проблемы оценки качества промышленной продукции, планирования получения и обработки экономических и социологических данных, анализ популяций в биологии, анализ текстов в лингвистике и т. д.

Основное внимание в книге уделяется случаю, когда обследуемые объекты принадлежат одному из двух или нескольких классов, т. е. классифицируются по альтернативному или качественному признаку. Заключительная глава посвящена обследованию совокупностей объектов, классифицируемых по количественному признаку. Возможность использования вероятностных методов является следствием случайного отбора объектов для обследования. При этом на обследуемую совокупность накладываются минимальные ограничения.

У читателя предполагаются лишь самые начальные знания по теории вероятностей. Математический аппарат в основном ограничивается комбинаторными методами.

Книга будет полезна всем, кто интересуется приложениями методов теории вероятностей и математической статистики.

Библ. — 266 назв., рис. 31, табл. 12.

ОГЛАВЛЕНИЕ

Предисловие	7
Глава 1. Введение	11
§ 1.1. Роль исходных математических моделей	11
О контроле по количественному и альтернативному признаку (11). Примеры математических моделей (13).	
§ 1.2. Планы бездефектного контроля	19
Вероятность принять совокупность (19). Последующие оценки с учетом результатов разбраковки (28).	
§ 1.3. Гипергеометрическое распределение.	33
Основные соотношения (33). Моментные характеристики (37). Аппроксимация другими распределениями (40). Расчет параметров одноступенчатых планов (44).	
§ 1.4. Последующие оценки для одноступенчатых планов	49
Последующие оценки с учетом результатов разбраковки (49). Пуассоновское приближение (52). Последующие оценки для планов с двумя уровнями (54). Контроль с проверкой дополнительного объекта (57). Последующие оценки для планов с одним уровнем (59).	
Глава 2. Планы контроля по альтернативному признаку и их характеристики	63
§ 2.1. Вспомогательные сведения	63
Об основных понятиях теории вероятностей (63). Статистические выводы (72).	
§ 2.2. Планы первого вхождения и их характеристики	81
Основные определения (81). Вероятности результатов контроля (85). Рекуррентное вычисление вероятностных характеристик (91). Оперативная характеристика и связанные с ней показатели (93).	

§ 2.3. Априорные распределения числа дефектных объектов	98
Смеси биномиальных распределений (98). Схема Пуассона (103). Предельная теорема для смеси биномиальных распределений (106). Моментные характеристики (109).	
§ 2.4. Байесовский подход	112
Средняя доля дефектных объектов в непроконтролируемой части совокупности (112). Одноступенчатые планы контроля (116).	
Глава 3. Статистические задачи	125
§ 3.1. Последующие точечные оценки для планов первого вхождения	125
Несмещенные оценки функций числа дефектных объектов в контролируемой совокупности (125). Полнота планов контроля и единственность несмещенных оценок (132). Несмещенные оценки для предъявленного и принятого брака (138).	
§ 3.2. Последующие интервальные оценки	145
Контроль с биномиальными выборками (145). Метод, основанный на асимптотической нормальности распределений оценок (151).	
§ 3.3. Оценка априорного распределения числа дефектных объектов в совокупности	156
Основные соотношения при использовании планов первого вхождения (156). Применение общих формул. Примеры (161). Постановки задач (164).	
§ 3.4. Выборочный контроль по качественному признаку	167
Основные вероятностные модели (167). Задачи статистического оценивания (176).	
Глава 4. Системы планов приемочного контроля по альтернативному признаку	181
§ 4.1. Основные принципы построения систем планов приемочного контроля	181
Системы планов контроля (181). Нормальный ход производства и разладки (183). Таблицы планов контроля (187). Таблицы планов с заданным ПСВК (189).	

§ 4.2. Системы планов контроля и их характеристики. . .	194
Система Доджа и Ромига, МИЛ-СТД—105 (194). Переключающие правила (198). Статистическое моде- лирование на ЭВМ процесса производства и приемоч- ного контроля (206).	
§ 4.3. Система экономичных планов (СЭП)	210
Расходы на контроль (210). Рациональные экономич- ные планы (215). Основные принципы построения СЭП (227).	
Глава 5. О разных задачах выборочного контроля	233
§ 5.1. Контроль непрерывного потока продукции	233
Планы Доджа выборочного контроля (233). Критич- ные планы выборочного контроля (243).	
§ 5.2. Планы направленного контроля.	251
Выборочный контроль совокупностей, содержащих серии дефектных объектов (251). Марковская мо- дель производства (253). Альтернирующие после- довательности годных и дефектных объектов (259).	
§ 5.3. Планы испытаний, гарантирующие минимум предель- ного числа критических событий	263
Постановка задачи (263). Двухэтапные планы (пуас- соновское приближение) (266). Многоэтапные планы (270).	
§ 5.4. Оценка объема совокупности	274
Двухступенчатые планы (274). Последовательные планы (277).	
Глава 6. Выборки из конечных совокупностей	289
§ 6.1. Оценка среднего конечной совокупности.	289
Случайная выборка без возвращения (289). Асимп- тотическая нормальность выборочного среднего (293). Выборка с возвращением (303).	
§ 6.2. Непараметрические статистические задачи	308
Порядковые статистики выборки (308). Гипергео- метрический процесс (312). Неравенство для момен- тов четвертого порядка (320). Локальная предель- ная теорема (322). Предельные теоремы для функ- ционалов от траекторий гипергеометрических процес- сов (325). О выявлении результатов воздействий (336).	

§ 6.3. Планы выборочного обследования	343
Основные определения (343). Допустимость оценок и планов (346). Оценка разности средних при воздействиях (350). Расслоение (353). Вариационная диаграмма (355). О некоторых задачах и результатах (357).	
Приложение	360
Таблица I. Значения λ_L , ρ_c для расчета предела среднего уровня выходного качества при использовании одноступенчатых планов	360
Таблица II. Планы $(n, c)_{23}$, обеспечивающие минимум инспекции при заданном предельном выходном качестве	360
Таблица III. Значения $\lambda_L(x)$, $\rho_{0,1}(x)$ для расчета предела среднего уровня выходного качества при использовании планов двухступенчатого контроля с приемочными числами 0 и 1	361
Таблица IV. Значения $\lambda_L(x)$, $\rho_{0,2}(x)$ для расчета предела среднего уровня выходного качества при использовании планов двухступенчатого контроля с приемочными числами 0 и 2	362
Таблица V. Наименьшие значения $M_{0,1}^a$, $M_{0,1}^E$, $M_{0,1}^q$	363
Таблица VI. Параметры одноступенчатых планов СЭП	365
Таблица VII. Параметры двухступенчатых планов СЭП	366
Таблица VIII. Двухэтапные планы, гарантирующие минимум предельного числа критических отказов	368
Таблица IX. Трехэтапные планы, гарантирующие минимум предельного числа критических отказов	370
Таблица X. Значения критических уровней $\beta_{0,9}(n, N)$	372
Таблица XI. Значения функций $\omega_1(x, y)$ и $\omega_2(x, y)$	373
Таблица XII. Уровни оперативной характеристики распределения Пуассона	377
Литература	388
Предметный указатель	404
Список обозначений	407

ПРЕДИСЛОВИЕ АВТОРА

Эта книга посвящена специальному разделу математической статистики. В ней рассматриваются вопросы планирования и обработки результатов наблюдений, полученных на основе выборочного обследования. Основной замысел, которому старался следовать автор, состоит в том, что вероятностные методы используются при выборочном обследовании в первую очередь потому, что отбор объектов для обследования производится случайным образом. Таким образом, сама обследуемая совокупность объектов не обязательно должна обладать вероятностными свойствами и накладываемые на нее ограничения являются минимальными. Образно говоря, вероятностные методы дают возможность изучать совокупности так же, как, например, в физических исследованиях рентгеновские лучи дают возможность судить о свойствах облучаемых материалов. Минимальность исходных допущений весьма существенна для расширения области возможных приложений. Примерами приложений могут служить методы расчета уровня качества и надежности промышленных изделий, методы обработки результатов экономических и социологических обследований, методы анализа данных о биологических популяциях и т. д. Удивительно, что при слабых ограничениях на обследуемые совокупности все же удается построить достаточно богатую фактами теорию.

Название книги, вообще говоря, шире ее содержания в настоящем издании. Значительная часть материала связана с так называемой классификацией по альтернативному или качественному признаку, когда по результатам обследования объекты включаются в один из двух или нескольких классов. Лишь в последней

главе рассмотрены задачи выборочного обследования совокупностей, содержащих конечное число объектов, каждый из которых характеризуется скалярной величиной, которая может принимать любые вещественные значения. Такой подход накладывает определенные ограничения. В книге не рассматриваются вопросы выборочного обследования объектов, имеющих геометрическую форму (нитей, пленок, объемных материалов и т. п.) без дробления их на дискретные однотипные участки.

В окружающей нас действительности многие явления можно рассматривать как случайные. Однако далеко не во всех явлениях факторы случайности являются доминирующими. Так будет, если мы изучаем совокупность, содержащую конечное число объектов, параметры которых стабильны во времени. За конечный отрезок времени мы можем измерить интересующие нас характеристики этих объектов и тем самым получим описание совокупности. Однако часто такой подход может быть нецелесообразен в связи с большими экономическими затратами и длительностью времени, необходимыми для сплошного обследования всех без исключения объектов. Иногда сплошное обследование невозможно потому, что измерение либо связано с разрушением объектов (испытания элементов на надежность), либо существенно меняет их характеристики. Например, если мы хотим оценить производительность труда сотрудников какого-то предприятия, то можно объявить, что через столько-то дней и в течение такого-то интервала времени будет измеряться индивидуальная производительность труда и ее изменения в течение дня для каждого члена коллектива. Заблаговременное оповещение может вызвать существенные изменения производительности. В ситуациях, когда массовые измерения могут дать результаты, не согласующиеся со сводными типичными показателями, предпочтительными будут методы выборочного обследования, основанные на вероятностных методах. При этом измерения будут проводиться только для части случайным образом отобранных для контроля объектов. В социологических обследованиях случайно будет выбираться и момент обследования (летучий контроль).

В книге частично использована терминология, применяемая при контроле качества. Это связано с тем, что основные приложения выборочных методов были использованы автором для решения задач статистического контроля качества и надежности. Если вместо слов *изделие* и *партия* в книге используются (с целью расширения области приложений при обследовании экономических и биологических объектов) более широкие термины *объект* и *совокупность*, то при классификации по альтернативному признаку сохранены термины *годный* и *дефектный* объекты. В более широком понимании термин *годный* (*дефектный*) объект следует заменить более широким термином *объект* первого (второго) класса.

В результате работы над книгой у автора возникло глубокое убеждение, что в теории выборочных методов можно получить содержательные аналоги большинства основных утверждений теории вероятностей и математической статистики, которые к настоящему времени найдены в предположении взаимной независимости результатов наблюдений. В особенности это относится к такому важному разделу математической статистики, как непараметрические методы. В связи с этим хочется обратить внимание читателя на содержание § 6.2, где вводится аналог известного непараметрического критерия А. Н. Колмогорова.

Недостаточная разработанность вопросов анализа по результатам выборочного обследования взаимосвязи компонент признаков у объектов, характеризующихся векторными величинами, явилась причиной того, что важные статистические задачи регрессионного и факторного анализа не включены в содержание книги. Но следует отметить, что здесь также по-видимому удастся получить некоторые интересные результаты. Эта область исследований для развития выборочных методов весьма перспективна.

Проблемами, связанными с выборочными методами решения задач статистического приемочного контроля качества, автор начал заниматься в Лаборатории статистических методов МГУ по инициативе А. Н. Колмогорова. Исходные задачи поиска экономичных планов приемочного контроля были поставлены автору

академиком А. Н. Колмогоровым еще в 1960 г. В течение этого времени автор имел постоянные контакты с инженерами, работающими над решением проблем контроля качества и надежности. В книге частично использован материал лекций, прочитанных зимой 1973 г. для инженеров в Политехническом музее в Москве.

Внутри параграфов одной главы используется нумерация формул, в которой первая цифра соответствует номеру параграфа, а вторая — номеру формулы внутри параграфа. При ссылках на формулы других глав слева дописывается цифра, соответствующая номеру главы; например (3.4.5) означает формулу (4.5) в главе 3.

Библиография, приведенная в конце книги, не является полной, однако автор старался включить в нее работы, связанные с материалом книги.

В работе над книгой большую помощь оказала Л. В. Рыкова, результаты которой были включены в § 5.2, § 6.2. Существенные замечания были сделаны Л. Н. Большевым. Учет этих замечаний позволил улучшить и сделать более информативными часть таблиц в приложении к книге. Е. В. Чепурин и Я. П. Лумельский внимательно прочли рукопись. Сотрудники Лаборатории статистических методов Т. Н. Дугина, В. А. Красавкина и Г. К. Степерман оказали существенную помощь при вычислении таблиц и подготовке рукописи книги к изданию. Всем им автор глубоко благодарен.

Май 1974 г.

Ю. К. Беляев

ГЛАВА I

ВВЕДЕНИЕ

§ 1.1. Роль исходных математических моделей

О контроле по количественному и альтернативному признаку. Степень пригодности объектов к дальнейшему использованию, т. е. качество объектов, можно определять различными способами. Можно, например, регистрировать точные численные значения параметров, определяющих качество. Такой контроль носит название **контроля по количественному признаку**. Иногда по результатам измерений можно отмечать лишь категории, классы, которым принадлежит контролируемое изделие. В этом случае контроль называется **контролем по качественному признаку**. Если же контролируемые объекты разбиваются на годные или дефектные, то такой контроль носит название **контроля по альтернативному признаку**. Таким образом, контроль по альтернативному признаку является частным случаем контроля по качественному признаку, когда общее число различных классов равно двум.

Контроль по качественному признаку обладает рядом преимуществ по сравнению с контролем по количественному признаку. Можно заготовить простые шаблоны или автоматические измерительные устройства, которые будут разбраковывать изделия. Методы обработки результатов измерений при контроле по качественному признаку, как правило, проще и, соответственно, объем вычислений меньше, чем при контроле по количественному признаку. И, наконец, возможно, самое важное преимущество состоит в том, что методы, основанные на контроле по качественному признаку, не зависят от вида вероятностных распределений

замеряемых параметров, тогда как в большинстве случаев при контроле по количественному признаку предполагается, что измеряемые параметры подчиняются одному из стандартных типов распределений. Чаще всего в качестве такого распределения выбирается нормальное (гауссовское) распределение.

При контроле по качественному признаку используется лишь часть информации, содержащаяся в наблюдениях. Это приводит к тому, что для достижения одинаковых показателей эффективности при контроле по качественному признаку требуется проконтролировать большее количество объектов, чем при контроле по количественному признаку. Однако если исходные допущения о виде распределения измеряемого признака не выполняются, то контроль по количественному признаку может привести к ошибочным решениям. В связи с этим следует отметить ставший классическим пример Бернштейна [1].

Чтобы можно было сделать точное количественное сравнение различных методов выборочного контроля, нужно сформулировать все необходимые для таких расчетов предположения. Другими словами, нужно определить соответствующую *математическую модель процесса контроля*. Например, если контролируется совокупность объектов $\mathfrak{F} = \{O_i\}$, $i = 1, \dots, N$, то прежде всего нужно определить возможные результаты измерений. Если контроль проводится по количественному признаку и результаты измерений являются скалярными величинами, то часто предполагается, что результаты наблюдений являются взаимно независимыми случайными величинами с функцией распределения $F(x)$. При этом предположении результаты контроля n объектов образуют последовательность $(x_1, \dots, x_n)$ n значений взаимно независимых одинаково распределенных случайных величин с функцией распределения $F(x)$. Допустим, что по измеряемым параметрам контроль является альтернативным. В этом случае задано некоторое множество A_Q допустимых значений измеряемого параметра x . Если для объекта O_i значение $x_i \in A_Q$ (x_i принадлежит A_Q), то объект классифицируется как годный, если же $x_i \notin A_Q$ (x_i не принадлежит A_Q), то как дефектный. В этом случае результаты сплошного

контроля N объектов можно было бы представить в виде последовательности $(\varepsilon_1, \dots, \varepsilon_N)$, где $\varepsilon_i = 0$, если объект O_i годный, и $\varepsilon_i = 1$, если дефектный. Общее число дефектных изделий $D = \sum_{1 \leq i \leq N} \varepsilon_i$ в совокупности $\mathfrak{F}$ нам до измерений неизвестно. Если при альтернативном контроле принять гипотезу взаимной независимости x_i с функцией распределения (ф. р.) $F(x)$, то значения ε_i будут взаимно независимыми и $P\{\varepsilon_i = 1\} = P\{x_i \in A_Q\} = q$, $\bar{q} = 1 - q$, $i = 1, \dots, N$. Следовательно, значения ε_i , $i = 1, \dots, N$, образуют биномиальную последовательность. Но тогда*) и общее число D_ω событий $\{\varepsilon_i = 1\}$ является случайной величиной, имеющей биномиальное распределение с параметрами q и N :

$$P\{D_\omega = D\} = \binom{N}{D} q^D \bar{q}^{N-D}. \quad (1.1)$$

Чтобы подчеркнуть случайность D_ω — числа дефектных объектов в $\mathfrak{F}$, мы дописываем подстрочный индекс ω . Однако ни откуда не следует, что D_ω всегда имеет функцию распределения $F(D_0) = P\{D_\omega \leq D_0\}$, совпадающую с биномиальным распределением, т. е. с функцией

$$F(D_0) = \sum_{D=0}^{D_0} \binom{N}{D} q^D \bar{q}^K, \quad K = N - D. \quad (1.2)$$

Таким образом, если распределение числа дефектных объектов в контролируемой совокупности нельзя считать биномиальным, то не выполняется и предположение о независимости результатов контроля по количественному признаку. Более того, можно поставить под сомнение и предположение о том, что число дефектных объектов в $\mathfrak{F}$ вообще является случайной величиной. В такой ситуации D надо рассматривать как неизвестный целочисленный параметр.

Примеры математических моделей. Для иллюстрации зависимости количественных показателей выборочного контроля от исходной математической модели рассмотрим несколько примеров.

*) См., например, учебник Гнеденко [1], стр. 71.

Определение 1.1.1. Случайным выбором без возвращения называется такой отбор объектов на контроль, когда объекты из $\mathfrak{F}$ отбираются последовательно один за другим, уже отобранные объекты в дальнейшем отборе не участвуют и для всех еще не отобранных объектов вероятность отбора одинакова. Общее число объектов, отобранных для контроля, называют **объемом выборки**.

Определение 1.1.2. Случайной выборкой объема n называется такой план контроля, при котором все наборы из n различных объектов имеют равные вероятности.

Случайную выборку объема n можно получить в результате случайного выбора без возвращения n объектов.

Модель I. Контроль совокупности $\mathfrak{F}$ объема N проводится по альтернативному признаку. Общее число дефектных объектов D рассматривается как неизвестный параметр. Из $\mathfrak{F}$ берется случайная выборка без возвращения объема n . Если в выборке встретилось хотя бы одно дефектное изделие, то принимается **решение о забраковании** $\mathfrak{F}$, в противном случае $\mathfrak{F}$ **«принимается»**. При заданном D вероятность того, что первый объект годный, равна $(N-D)/N$, второй годный — $(N-D-1)/(N-1)$, ..., n -й годный — $(N-D-n+1)/(N-n+1)$. Следовательно, вероятность принять $\mathfrak{F}$ равна

$$\pi_1(q) = \frac{N-D}{N} \cdot \frac{N-D-1}{N-1} \cdot \dots \cdot \frac{N-D-n+1}{N-n+1} = \frac{\binom{N-D}{n}}{\binom{N}{n}}, \quad (1.3)$$

где $q = D/N$.

Модель II. Отличие от модели I состоит в том, что D предполагается биномиально распределенной случайной величиной, вероятности значений которой определяются в соответствии с (1.1). По формуле полной вероятности из (1.1) и (1.3) находим, что вероятность не обнаружить дефектных объектов среди n , отобранных для контроля, равна

$$\pi_2(q) = \sum_{D=0}^N \binom{N}{D} q^D \bar{q}^{N-D} \left[\frac{\binom{N-D}{n}}{\binom{N}{n}} \right] = \bar{q}^n. \quad (1.4)$$

Модель III. Объекты O_i совокупности $\mathfrak{F}$ характеризуются вещественными числами x_i , $i = 1, \dots, N$, которые становятся известными после проведения контроля. Если $x_i \in [-3, +3]$, то O_i классифицируется как годный. Если же $x_i < -3$, или $x_i > 3$, то O_i считается дефектным. Предполагается, что выполнена гипотеза независимости значений x_i с $F(x) = \Phi(x - \mu)$, где μ — вещественный параметр, равный математическому ожиданию случайных величин x_i ,

$$\Phi(z) = \int_{-\infty}^z \varphi(z) dz, \quad \varphi(z) = (2\pi)^{-1/2} \exp(-z^2/2),$$

т. е. предполагается, что измеряемые значения x_i являются взаимно независимыми нормально распределенными случайными величинами с единичной дисперсией и математическим ожиданием μ . Таким образом, общее число дефектных объектов D в $\mathfrak{F}$ является биномиально распределенной случайной величиной. При этом среднее значение доли дефектных объектов q в $\mathfrak{F}$ равно

$$q = 1 - \int_{-3}^{+3} \varphi(x - \mu) dx = 1 - [\Phi(3 - \mu) - \Phi(-3 - \mu)]. \quad (1.5)$$

Если при решении вопроса о возможном значении q использовать всю информацию, получаемую в результате контроля, то, в отличие от модели II, нужно учитывать фактические значения x_i , а не только лишь принадлежность x_i к интервалу $[-3, +3]$ допустимых значений. Из общей теории математической статистики (Дунин-Барковский и Смирнов [1], Леман [1], Шор [1]) следует, что решение о приемлемости значений q нужно принимать по значениям статистики $x(n) = \sum_{1 \leq i \leq n} x_i$.

Как сумма независимых нормально распределенных случайных величин, эта статистика имеет нормальное распределение со средним $n\mu$ и дисперсией n . В симметричном случае, когда одинаково нежелательны выходы значений параметра как за верхнюю, так и за нижнюю границы интервала, совокупность $\mathfrak{F}$ будем принимать, только если $|x(n)| \leq c$. Константа c подби-

рается таким образом, чтобы при заданном значении $q = q_n$ вероятность принятия решения о приемлемости совокупности со значением q была бы равна заданной величине $1 - \alpha$. Пусть $\mu = \mu(q) > 0$ — положительное решение уравнения (1.5), которое существует при $0 \leq q \leq 0,975$. Отрицательное решение равно $(-\mu(q))$. Так как $(x(n) - n\mu)/\sqrt{n}$ — стандартная нормальная случайная величина с функцией распределения $\Phi(x)$, то вероятность принять совокупность со значением q равна

$$\pi_3(q) = P(|x(n)| < c) = \\ = \Phi\left(\frac{c - n\mu(q)}{\sqrt{n}}\right) - \Phi\left(\frac{-c - n\mu(q)}{\sqrt{n}}\right). \quad (1.6)$$

Значение $c = c_\alpha$ находится из уравнения

$$\Phi\left(\frac{c_\alpha - n\mu(q_n)}{\sqrt{n}}\right) - \Phi\left(\frac{-c_\alpha - n\mu(q_n)}{\sqrt{n}}\right) = 1 - \alpha. \quad (1.7)$$

Таким образом, при $q = q_n$ из (1.5) — (1.7) следует, что $\pi_3(q_n) = 1 - \alpha$.

Модель IV. Предположим, как и в модели III, что контроль проводится по количественному признаку. Классификация объектов на годные и дефектные аналогична модели III. Контролируемые значения x_i также взаимно независимые случайные величины с ненулевым математическим ожиданием μ и единичной дисперсией, принимающие с равными вероятностями два значения, т. е. $P\{x_i = 1 + \mu\} = P\{x_i = -1 + \mu\} = 1/2$. Следовательно, функция распределения x_i существенно отличается от нормальной. Подсчитаем, какова будет вероятность принять совокупность, если мы воспользуемся решающей функцией модели III. На первый взгляд результаты должны быть близки, так как первые два момента (среднее и дисперсия) совпадают, а $x(n)$ при больших n приближенно имеет такое же, как и в модели III, нормальное распределение, что следует из центральной предельной теоремы теории вероятностей (Гнеденко [1], стр. 258). Введем случайные величины $\delta_i = (x_i - \mu + 1)/2$. Так как $P\{\delta_i = 0\} = P\{\delta_i = 1\} = 1/2$, то $\delta(n) = \sum_{i=1}^n \delta_i$ имеет биномиальное

распределение с параметрами $1/2$ и n , т. е. $P(\delta(n) = k) = \binom{n}{k} \left(\frac{1}{2}\right)^n$. При изменении значений μ доля q засоренности $\mathfrak{F}$ дефектными объектами меняется скачкообразно. Если $-2 \leq \mu \leq 2$, то $q \equiv 0$, если же $2 < \mu \leq 4$ или $-4 \leq \mu < -2$, то $q = 1/2$; если $|\mu| > 4$, то $q \equiv 1$. Вероятность принять совокупность как функция μ равна

$$\begin{aligned} \pi_4(\mu) &= P\{|x(n)| \leq c_\alpha\} = \\ &= P\left\{\frac{-c_\alpha + n(1-\mu)}{2} \leq \delta(n) \leq \frac{c_\alpha + n(1-\mu)}{2}\right\} = \\ &= \sum' \binom{n}{k} \left(\frac{1}{2}\right)^n, \quad (1.8) \end{aligned}$$

где $\sum'$ берется по всем целым неотрицательным k из интервала $\left[\frac{-c_\alpha + n(1-\mu)}{2}, \frac{c_\alpha + n(1-\mu)}{2}\right]$.

Обсудим теперь возможные последствия того, что одна из этих четырех моделей принята в качестве основы для расчета показателей контроля. Чтобы соответствующие выводы можно было проиллюстрировать численно, для определенности будем считать объем совокупности $\mathfrak{F}$ равным $N = 1000$, а объем выборки $n = 20$. В моделях II—IV результаты вычислений не зависят от N . Принципиальное отличие модели I от моделей II—IV заключается в том, что фактическое число дефектных объектов D не предполагается случайной величиной, а рассматривается как неизвестный параметр. Таким образом, рассчитанные характеристики контроля для модели I будут столь хорошо соответствовать реальному положению вещей, сколь хорошо нам удастся организовать случайность отбора объектов в выборку. Модели II и III отличаются тем, что при контроле по количественному признаку в модели III существенно полнее используется информация о результатах контроля. В итоге мы будем иметь меньше ошибки ложного бракования $\mathfrak{F}$ при малых q и ошибочного принятия $\mathfrak{F}$ при больших q . Этот факт отражен в таблице 1.1.1 (для модели I $q = D/N$). Подчеркнем, что в таблице 1.1.1 $\pi_4(\mu)$ есть функция только параметра μ . Для моделей II, III величина $\pi_4(\mu)$ зависит также от q .

связаны соотношением (1.5). Если $q_n = 0,01$, $\alpha = 0,1$, то для значений $q \leq 0,01$ вероятность бракования совокупности не превышает 0,1 в модели III и 0,182 в модели II. Если значение $q = 0,05$ считается недопустимым, то при использовании процедуры модели II почти в одной трети случаев совокупности будут ошибочно приниматься, тогда как для модели III вероятность ошибочного принятия совокупности была бы равна всего лишь 0,078.

Таблица 1.1.1

Характеристики контроля как функции q и μ

	Значение q									
	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09	0,10
$\pi_1(q)$	0,816	0,665	0,541	0,438	0,355	0,284	0,231	0,186	0,149	0,119
$\pi_2(q)$	0,818	0,668	0,544	0,442	0,359	0,290	0,234	0,189	0,152	0,122
μ	0,670	0,950	1,120	1,250	1,360	1,450	1,520	1,600	1,660	1,720
$\pi_3(q)$	0,900	0,656	0,358	0,173	0,078	0,034	0,015	0,006	0,003	0,001
$\pi_4(\mu)$	0,885	0,587	0,412	0,132	0,058	0,020	0,002	0,001	0,000	0,000

Предположим, однако, что выполнены условия модели IV, а мы ошибочно используем решающую процедуру модели III. Пусть для определенности $\mu = 1,25$. Тогда в соответствии с (1.8) получаем $\pi_4(1,25) = 0,132$, т. е. с большой вероятностью совокупность будет забракована. Однако в модели IV для $\mu = 1,25$ все значения x_i лежат в интервале $[-3, +3]$, дефектных объектов нет, $q = 0$. Таким образом, совокупность ошибочно бракуется с вероятностью 0,868! При контроле по альтернативному признаку и процедуре модели II ошибок не было бы, так как дефектные объекты не были бы обнаружены. Этот пример вскрывает одну из основных проблем математической статистики — проблему устойчивости статистических процедур относительно изменений исходных предпосылок.

Итак, мы видим, что от выбора правильной модели зависит почти все. Если информации для выбора модели мало, то, по-видимому, лучше использовать не оптимальные, но зато устойчивые процедуры, такие,

как в модели I или в модели II. Если же есть уверенность в предполагаемой форме модели, то статистические процедуры усложняются, но характеристики контроля существенно улучшаются. Еще более важны вопросы статистической устойчивости применительно к задачам выборочного контроля показателей надежности.

Следует иметь в виду, что использование обоснованных исходных математических моделей еще не гарантирует соответствия расчетных характеристик реальному положению дел. В частности, несоответствие может возникнуть при нарушении правил контроля. Примеры нарушений правил контроля даны в книге Гнеденко, Беляев, Соловьев [1] (гл. 7).

§ 1.2. Планы бездефектного контроля

Вероятность принять совокупность. Изучение методов выборочного контроля начнем с рассмотрения случая, когда требования к качеству контролируемой совокупности столь велики, что наличие в ней хотя бы одного дефектного объекта недопустимо. В таких условиях обнаружение даже одного дефектного объекта при выборочном контроле означает, что должно быть принято решение о забраковании всей совокупности. Решение о забраковании трактуется в зависимости от конкретных условий. Если контрольная операция носит *разрушительный* характер, то при обнаружении одного дефектного объекта часть совокупности, не включенная в выборку, отвергается без дальнейшего контроля. Если же контрольная операция не меняет свойств объектов и ее стоимость невелика, то забракование совокупности означает *сплошную проверку* всех объектов этой совокупности с целью выявления и изъятия (или исправления) всех дефектных объектов.

Приведем элементарные экономические расчеты (см. Колмогоров [5]). Пусть контроль не разрушающий, стоимость контроля одного объекта b , а ущерб от принятия одного дефектного объекта g . Сплошной контроль целесообразен, когда $bN < gD$ или $Nq_{кр} < D$, где D — общее число дефектных объектов, $q_{кр} = b/g$ — критический уровень качества. Поэтому при $Nq_{кр} < 1$

сплошной контроль целесообразен уже при появлении в совокупности хотя бы одного дефектного объекта.

Пусть n — объем выборки, т. е. число объектов, отбираемых для обследования. В общем случае можно задать приемочное число c , т. е. такое число, что при обнаружении в выборке d дефектных объектов при $d \leq c$ совокупность принимается, а при $d > c$ ее забраковывают. Недопустимость дефектных объектов означает, что приемочное число $c = 0$. Прочитируем вывод, взятый из указанной работы Колмогорова [5]: «По-видимому, можно с достаточной уверенностью дать... рекомендацию: если $Nq_{кр} < 1$, то целесообразно при любом ходе производства употреблять только методы контроля с $c = 0$ ». Кратко планы выборочного контроля с фиксированным объемом выборки n и приемочным числом $c = 0$ будем называть **планами бездефектного контроля**.

Основным показателем планов бездефектного контроля является функция $\pi(N, n, D)$, равная вероятности принять совокупность $\mathfrak{F}$ объема N , когда в результате выборочного обследования n объектов из $\mathfrak{F}$ не было обнаружено ни одного дефектного объекта. Пусть $K = N - D$. Тогда в соответствии с (1.3) имеем

$$\begin{aligned} \pi(N, n, D) &= \prod_{j=1}^n \left(1 - \frac{D}{N-j+1}\right) = \frac{K^{(n)}}{N^{(n)}} = \frac{(N-n)! K!}{N! (K-n)!} = \\ &= \frac{(N-n)^{(D)}}{N^{(D)}} = \prod_{j=1}^D \left(1 - \frac{n}{N-j+1}\right). \quad (2.1) \end{aligned}$$

Здесь и всюду ниже используется обозначение $a^{(m)} = a(a-1)\dots(a-m+1)$, где $a > 0$, а $m = 1, 2, \dots$ — целое число. Вычисление $\pi(N, n, D)$ по формуле (2.1) трудоемко. При перемножении сомножителей ошибки округления могут существенно исказить результат, поэтому сомножители надо вычислять с большой точностью. При вычислении значений факториалов в (2.1) можно воспользоваться таблицей 7.6 книги Большева и Смирнова [1], где значения $m!$ приведены для $m \leq 1000$. Затабулировать достаточно полные таблицы значений $\pi(N, n, D)$ не представляется возможным, так как $\pi(N, n, D)$ есть функция трех переменных. В работе

Колмогорова [5] даются приближенные методы вычисления $\pi(N, n, D)$.

Если $\max\left(\frac{n}{N}, \frac{D}{N}\right) \ll 1$, то из (2.1) можно получить простейшую приближенную формулу

$$\pi(N, n, D) \approx \exp\{-nD/N\}. \quad (2.2)$$

Заметим, что nD/N равно математическому ожиданию числа дефектных объектов, попадающих в выборку объема n .

Если $v = n/N \ll 1$, а $q = D/N$ существенно больше v , то из левой части равенств (2.1) находим

$$\pi(N, n, D) \approx (1 - D/N)^n. \quad (2.3)$$

Если же, наоборот, $q = D/N \ll 1$, а v существенно больше q , то из правой части равенств (2.1) получаем

$$\pi(N, n, D) \approx (1 - n/N)^D. \quad (2.4)$$

Шор [4] предложил использовать для аппроксимации (2.1), когда $n/N \ll 1$ или $D/N \ll 1$, сомножители, равные соответственно $\left(1 - \frac{D}{N - (n-1)/2}\right)$ и $\left(1 - \frac{n}{N - (D-1)/2}\right)$.

Поэтому вместо (2.3) и (2.4) Шор [4] рекомендует формулы

$$\pi(N, n, D) \approx \left(1 - \frac{D}{N - (n-1)/2}\right)^n, \quad (2.5)$$

$$\pi(N, n, D) \approx \left(1 - \frac{n}{N - (D-1)/2}\right)^D. \quad (2.6)$$

Формула (2.5) рекомендуется при $n \leq D$, а формула (2.6) при $n \geq D$. Для упрощения расчетов по формулам (2.5), (2.6) удобно пользоваться значениями функции $\omega = \omega(x)$, определяемой формулой

$$\omega = (-1/x) \ln(1 - x), \quad 0 < x \leq 1. \quad (2.7)$$

С учетом (2.7) соотношения (2.5), (2.6) принимают вид

$$\pi(N, n, D) \approx \exp\{-\omega(x_1) x_1 n\}, \quad (2.8)$$

$$\pi(N, n, D) \approx \exp\{-\omega(x_2) x_2 D\}, \quad (2.9)$$

где $x_1 = \frac{D}{N - (n-1)/2}$, $x_2 = \frac{n}{N - (D-1)/2}$. Таблицы функции $\omega(x)$ приведены в работах Колмогорова [5]

и Шора [4]. Таблица 1.2.1 является частью таблицы, содержащейся в статье Шора [4].

Т а б л и ц а 1.2.1

Значения $\omega = (-1/x) \ln(1-x)$

x	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09	0,10
ω	1,005	1,010	1,015	1,020	1,026	1,031	1,037	1,042	1,043	1,054
x	0,11	0,12	0,13	0,14	0,15	0,16	0,17	0,18	0,19	0,20
ω	1,059	1,065	1,071	1,077	1,083	1,090	1,096	1,102	1,109	1,116
x	0,21	0,22	0,23	0,24	0,25	0,26	0,27	0,28	0,29	0,30
ω	1,122	1,129	1,136	1,143	1,151	1,158	1,166	1,178	1,181	1,189
x	0,31	0,32	0,33	0,34	0,35	0,36	0,37	0,38	0,39	0,40
ω	1,197	1,205	1,214	1,122	1,231	1,240	1,249	1,258	1,267	1,277
x	0,41	0,42	0,43	0,44	0,45	0,46	0,47	0,48	0,49	0,50
ω	1,287	1,297	1,307	1,318	1,329	1,340	1,351	1,362	1,374	1,386

При использовании табличных значений $\omega(x)$ надо вычислить показатель экспоненты в (2.8) или (2.9) и воспользоваться таблицей функции e^{-y} для значений $y = \omega(x_1)x_1n$ или $y = \omega(x_2)x_2D$. Таблица $\omega(x)$ делает вычисления по формулам (2.2), (2.5) или (2.6), (2.8) или (2.9) почти одинаковыми по трудоемкости. Однако результаты, полученные на основе формул (2.8) или (2.9), существенно ближе к значениям, рассчитанным по точной формуле (2.1). Для сравнения результаты расчетов по формулам (2.1), (2.3), (2.4) и (2.6) при

$N = 10$, $n = 5$; $N = 25$, $n = 20$; $N = 50$, $n = 20$, $D = 1(1)4$ представлены в сводной таблице 1.2.2. Основная часть таблицы 1.2.2, за исключением последнего столбца, также взята из работы Шора [4]. Расчеты проводятся следующим образом.

Таблица 1.2.2

Результаты расчета $\pi(N, n, D)$ по точной
и приближенным формулам

			Вероятность $\pi(N, n, D)$, рассчитанная по формулам				
N	n	D	(2.1)	(2.3)	(2.4)	(2.6)	(2.15)
10	5	1	0,500	0,590	0,500	0,500	0,500
10	5	2	0,222	0,328	0,250	0,224	0,222
10	5	3	0,083	0,168	0,125	0,088	0,083
10	5	4	0,024	0,078	0,062	0,029	0,024
25	20	1	0,200	0,442	0,200	0,200	0,200
25	20	2	0,033	0,189	0,040	0,034	0,033
25	20	3	0,004	0,078	0,008	0,005	0,004
25	20	4	0,0 ³ 4*)	0,031	0,002	0,0 ³ 5*)	0,0 ³ 3
50	20	1	0,600	0,668	0,600	0,600	0,600
50	20	2	0,355	0,442	0,360	0,353	0,355
50	20	3	0,207	0,290	0,216	0,208	0,207
50	20	4	0,119	0,189	0,129	0,119	0,119

*) 0,0³4 = 0,0004, 0,0³5 = 0,0005 и т. д.

Пусть для $N = 10$, $n = 5$ и $D = 4$ требуется найти вероятность $\pi(10, 5, 4)$. Имеем

$$x = \frac{n}{N - (D - 1)/2} = 0,588, \quad \omega(x) \approx -(1/x) \ln(1-x) = 1,511.$$

По формуле (2.6) получаем $\pi(10, 5, 4) \approx \exp(-1,511 \cdot 4 \cdot 0,588) = 0,029$. Расчет по формуле (2.1) дает значение 0,024.

Немного усложнив алгоритм вычислений, мы можем несколько повысить точность вычислений. Для этого воспользуемся представлением $\pi(N, n, D)$ в виде отношения

факториалов (см. (2.1)) и формулой Стирлинга (см. Феллер [1], стр. 67)

$$m! \sim \sqrt{2\pi} \dots m^{m+1/2} e^{-m+1/(12m)}. \quad (2.10)$$

Из (2.1) и (2.10) в предположении $D+n < N$ получаем

$$\begin{aligned} \pi(N, n, D) &\approx \sqrt{\frac{(N-D)(N-n)}{N(N-(D+n))}} \cdot \frac{(N-D)^{N-D} (N-n)^{N-n}}{N^N (N-(D+n))^{N-(D+n)}} \times \\ &\times \exp \left\{ \frac{1}{12} \left(\frac{1}{N-D} + \frac{1}{N-n} - \frac{1}{N} - \frac{1}{N-(D+n)} \right) \right\} = \\ &= \sqrt{\frac{(1-D/N)(1-n/N)}{\left(1-\frac{D+n}{N}\right)}} \left[\frac{(1-D/N)(1-n/N)}{1-\frac{n+D}{N}} \right]^{(N/D)(N/n)(nD/N)} \times \\ &\times \left(1 - \frac{n/N}{1-D/N} \right)^{(N/n)(Dn/N)} \left(1 - \frac{D/N}{1-n/N} \right)^{(N/D)(nD/N)} \times \\ &\times \exp \left\{ -\frac{1}{N^2} \left[\frac{2 - \frac{D+n}{N}}{12 \left(1 - \frac{D+n}{N} \right) (1-n/N) (1-D/N)} \right] nD/N \right\}. \end{aligned} \quad (2.11)$$

Если ввести обозначения

$$\omega_1(x, y) = \frac{1}{2} \ln \frac{(1-x)(1-y)}{1-(x+y)}, \quad (2.12)$$

$$\begin{aligned} \omega_2(x, y) &= -(1/xy) [\ln(1-x) + \ln(1-y) - \ln(1-(x+y))] - \\ &- (1/x) \ln(1-x/(1-y)) - (1/y) \ln(1-y/(1-x)), \end{aligned} \quad (2.13)$$

$$\omega_3(x, y) = \frac{2-(x+y)}{12(1-(x+y))(1-x)(1-y)}, \quad (2.14)$$

то (2.11) перепишется в виде

$$\pi(N, n, D) \approx \exp \{ \omega_1(n/N, D/N) - [\omega_2(n/N, D/N) + (1/N^2) \omega_3(n/N, D/N)] \cdot nD/N \}. \quad (2.15)$$

Если $n+D > N$, то $\pi(N, n, D) = 0$. Предложенная асимптотическая формула (2.15) была выведена в предположениях $n+D < N$. В пограничном случае $n+D = N$, а также при $n+D = N-k$, где $k = 1, 2, \dots$ — небольшие числа, нужно использовать асимптотическую формулу Стирлинга (2.10) в (2.1)

лишь для $N!(N-D)!$ и $(N-n)!$. В итоге мы получим формулу

$$\pi(N, n, D) \approx (1/k!) \sqrt{2 \cdot 3,1415 \dots N(1-n/N)(1-D/N)} \times \\ \times \exp \{ -N [\omega_2^*(n/N, D/N) + (1/(12N^2)) ((N^2/nD) - 1)] \},$$

где $\omega_2^*(x, y) = -(1-x) \ln(1-x) - (1-y) \ln(1-y)$. Расчет по этой формуле для $N = 10$, $n = 5$ и $D = 5$ ($k=0$) дает значение $\pi = 0,00387$. Точное значение $\pi = 0,00397$. Расчет по формуле (2.6) дает существенно большую ошибку; при этом $\pi = 0,00742$. Функции $\omega_1(x, y)$ и $\omega_2(x, y)$ в (2.15) вычисляются просто. При малых x, y ($x+y < 0,5$) можно считать $\omega_1(x, y) = \frac{xy}{2(1-(x+y))} + \left[\frac{xy}{12(1-(x+y))} \right]^2$. При больших N ($N > 50$) вычислением ω_3 можно пренебречь, так как при $x+y < 0,5$ имеем $(1/N^2)\omega_3(x, y) < 2/(3N^2)$. Таким образом, если заранее рассчитать таблицу значений $\omega_1(x, y)$ и $\omega_2(x, y)$, то вычисления $\pi(N, n, D)$ по формуле (2.15) несложны. Функции $\omega_i(x, y)$, $i = 1, 2$ симметричны по x, y . Поэтому в таблице XI даны значения пар функций $\omega_2(x, y)$, $\omega_1(x, y)$ только для $x \leq y$, $x = 0,00(0,01)0,25^*$, $y - x = 0,00(0,01)0,25$. Например, для $x = 0,05$, $y = 0,11$ ($y - x = 0,06$) из таблицы XI находим $\omega_2 = 1,0885$, $\omega_1 = 0,0033$.

Проиллюстрируем высокую точность вычислений по формуле (2.15). Пусть $N = 250$, $n = 25$. Вычисления по точной формуле (2.1) дают результаты, приведенные в первой строке таблицы 1.2.3. Расчеты, проведенные с использованием формулы (2.15) и таблицы XI для тех же значений D , даны во второй строке той же таблицы 1.2.3 (см. стр. 26).

Точно так же получаем хорошее совпадение для примеров таблицы 1.2.2.

Планы бездефектного контроля полностью характеризуются объемом выборки n . Подбор подходящего значения n может быть произведен различными способами. Если основными являются требования потребителя, то

*) Обозначение $a(b)c$ соответствует ряду значений $a, a+b, a+2b, \dots, c-b, c$, т. е. $0,00(0,01)0,25$ эквивалентно последовательности $0,00, 0,01, 0,02, \dots, 0,19, 0,24, 0,25$.

задается **предельное значение** $D_{кр}$ для **среднего** числа дефектных объектов. Если фактическое число дефектных объектов $D \geq D_{кр}$, то совокупность следует браковать. Требования к жесткости контроля следует формулировать в вероятностных терминах. Например, мы можем потребовать, чтобы при $D \geq D_{кр}$ значения $\pi(N, n, D)$ были бы не более заданного числа β . Поскольку $\pi(N, n, D)$ убывает с ростом D , то условие выбора плана бездефектного контроля, гарантирующего потребителю вероятность приемки не большую β , когда $D \geq D_{кр}$, имеет вид

$$\pi(N, n, D_{кр}) \leq \beta. \quad (2.16)$$

При этом мы выбираем наименьшее из целых чисел n , удовлетворяющих (2.16), которое обозначим n_β . Для плана бездефектного контроля с объемом n_β в среднем будет принято не более $\beta \cdot 100\%$ совокупностей, содержащих $D_{кр}$, или больше, дефектных объектов.

Таблица 1.2.3

Значения $\pi(250, 25, D)$,
рассчитанные по формулам (2.1) и (2.15)

D формула <th>1</th> <th>2</th> <th>3</th> <th>5</th> <th>7</th> <th>10</th>	1	2	3	5	7	10
(2.1)	0,9000	0,8096	0,7280	0,5878	0,4738	0,3416
(2.15)	0,9009	0,8095	0,7287	0,5883	0,4739	0,3418

D формула <th>12</th> <th>15</th> <th>20</th> <th>25</th> <th>30</th>	12	15	20	25	30
(2.1)	0,2740	0,1961	0,1112	0,0622	0,0343
(2.15)	0,2744	0,1958	0,1114	0,0623	0,0344

В тех случаях, когда качество высокое, а ущерб от проникновения дефектных объектов невелик, можно исходить из учета **стоимости контроля**. Контроль не бу-

дет давать жесткой гарантии потребителю, что данная совокупность не содержит большого числа дефектных объектов. Однако получаемая по результатам контроля многих совокупностей информация поможет сделать выводы о приемлемости данного производства с точки зрения числа дефектных объектов, проникающих к потребителю. Если считать, что при $D \leq D_n$ совокупности бракуются напрасно, то условие выбора n может иметь вид

$$\pi(N, n, D_n) \geq 1 - \alpha. \quad (2.17)$$

При этом отбирается наибольшее целое число n_α , удовлетворяющее (2.17). План бездефектного контроля с объемом выборки n_α гарантирует производителю, что при $D \leq D_n$ в среднем не более $\alpha \cdot 100\%$ предъявленных для выборочного контроля совокупностей будет принято.

Для стандартных наборов значений α и β можно составить таблицы соответствующих значений $n_\alpha = n_\alpha(D_n, N)$, $n_\beta = n_\beta(D_{кр}, N)$. Такого рода таблицы приведены в работе Шора и Дуброва [1]. Можно предложить методы упрощенного расчета n_α и n_β с использованием вспомогательных таблиц аналогично тому, как это было сделано выше при выводе приближенных формул $\pi(N, n, D)$. В следующем параграфе будет приведена одна из подобных возможных рекомендаций для более общего класса одноступенчатых планов контроля.

Для защиты интересов потребителя можно исходить из значений **предела среднего выходного качества**, который для плана бездефектного контроля задается формулой

$$D_L = \max_{0 \leq D \leq N} D\pi(N, n, D). \quad (2.18)$$

Так как совокупность, содержащая D дефектных объектов, по результатам выборочного обследования принимается с вероятностью $\pi(N, n, D)$, то $D\pi(N, n, D)$ равно среднему числу дефектных объектов, проникающих к потребителю при бездефектном контроле. Таким образом, предел среднего выходного качества D_L равен максимально возможному значению среднего числа принимаемых дефектных объектов. На этот показатель можно ориентироваться в тех случаях, когда при

решении о забраковании совокупности все ее объекты проходят проверку, при которой дефектные объекты заменяются годными. При экспоненциальном приближении (2.2) для $\pi(N, n, D)$ имеем

$$D_L \approx \max_D [D \exp \{-nD/N\}] \approx \approx (N/n) \max_{x \geq 0} x \cdot e^{-x} = (N/n) \cdot e^{-1}. \quad (2.19)$$

Если же использовать приближение (2.4), то

$$D_L \approx \max_D [D \exp \{-\omega(n/N) nD/N\}] \approx \approx \frac{N}{n\omega(N/n)} \max_{x \geq 0} [x \cdot e^{-x}] = \frac{N}{n\omega(N/n)e}. \quad (2.20)$$

Пусть в интересах потребителя необходим такой контроль, при котором в среднем на одну совокупность, прошедшую контроль, приходилось бы не более D_L дефектных объектов при любом входном качестве. Тогда из (2.19) находим, что соответствующее значение объема выборки равно

$$n_L \approx N/(eD_L). \quad (2.21)$$

Поскольку в (2.19) максимум достигается при $x = nD/N = 1$, то в среднем наибольшее число дефектных объектов проникает к потребителю при бездефектном контроле, когда $D = D_L^* \approx (N/n)$. Следует иметь в виду, что планирование объема выборки n_L , обеспечивающее заданное значение предела среднего выходного качества, рассчитано на предъявление совокупностей со значениями D , близкими к D_L^* . Можно ожидать, что часто D существенно меньше D_L^* . При этом среднее число дефектных объектов, принимаемых в одной совокупности, будет существенно меньше D_L . Объем контроля оказывается завышенным.

Последующие оценки с учетом результатов разбраковки. Число обнаруженных дефектных объектов при плане бездефектного контроля равно 0 с вероятностью $\pi(N, n, D)$ и D с вероятностью $\pi(N, n, D)$, так как проводится сплошная проверка контролируемой совокупности при обнаружении хотя бы одного дефектного объекта в выборке. При таком плане контроля легко получить несмещенные оценки для предъявленного и принятого чисел дефектных объектов. Под *оценкой*

мы понимаем такую функцию результатов контроля, которая вычисляется по результатам наблюдений. Оценка является случайной величиной. Основными характеристиками оценки обычно считают математическое ожидание и дисперсию. Если математическое ожидание оценки совпадает с математическим ожиданием оцениваемой величины, то оценка называется **несмещенной**. В тех случаях, когда оцениваемая величина не случайна, несмещенность означает, что математическое ожидание оценки совпадает с оцениваемой величиной. Точные определения несмещенности оценок будут даны в § 2.1 второй главы.

Пусть контролируется совокупность $\mathfrak{F}$, содержащая N объектов, среди которых D дефектных. Кратко такую совокупность обозначаем $\mathfrak{F} \asymp [N, D]$. Используется план бездефектного контроля объема n со сплошной разбраковкой при обнаружении дефектных объектов. Пусть d — число дефектных объектов, обнаруженных в выборке. Запишем результат контроля x в следующем виде:

$$x = \begin{cases} 0, & d = 0, \\ D, & d > 0, \end{cases} \quad (2.22)$$

учитывая, что при забраковании все объекты контролируются. В соответствии с (2.22) число D''' дефектных объектов, принимаемых без контроля (**пропущенный брак**), случайно и равно

$$D''' = \begin{cases} D, & d = 0, \\ 0, & d > 0. \end{cases} \quad (2.23)$$

Найдем выражение для несмещенных оценок $\varphi = \varphi(x)$ и $\varphi''' = \varphi'''(x)$ для **предъявленного брака** D и пропущенного брака D''' соответственно. Так как $P\{x=0\} = \pi = \pi(N, n, D)$, $P\{x=D\} = \bar{\pi} = 1 - \pi$, то согласно (2.22) и (2.23) условия несмещенности оценок φ , φ''' имеют вид тождеств

$$\varphi(0)\pi + \varphi(D)\bar{\pi} \equiv D, \quad (2.24)$$

$$\varphi'''(0)\pi + \varphi'''(D)\bar{\pi} \equiv D\pi, \quad (2.25)$$

где $D = 0, \dots, N$. Левые части тождеств (2.24), (2.25) есть математические ожидания оценок φ , φ''' , а правые равны оцениваемому параметру D в (2.24) и математи-

ческому ожиданию D''' в (2.25). Полагая в тождествах (2.24), (2.25) $D = 0$, находим, что $\varphi(0) = \varphi'''(0) = 0$, так как при $D = 0$ имеем $\pi = 1$, $\bar{\pi} = 0$. Итак, первые слагаемые в (2.24), (2.25) отсутствуют, поэтому оценки $\varphi(x)$, $\varphi'''(x)$, удовлетворяющие (2.24), (2.25), имеют вид

$$\varphi(x) = \begin{cases} 0, & x = 0, \\ \frac{D}{\bar{\pi}(N, n, D)}, & x = D, \end{cases} \quad \varphi'''(x) = \begin{cases} 0 & x = 0, \\ \frac{D\pi(N, n, D)}{\bar{\pi}(N, n, D)}, & x = D. \end{cases} \quad (2.26)$$

Заметим, что при вычислении несмещенных оценок предъявленного и пропущенного брака используются значения $\pi(N, n, D)$. Для расчета $\pi(N, n, D)$ можно исходить из приближенных формул, приведенных в первой части этого параграфа.

Из (2.26) находим выражения для дисперсий несмещенных оценок φ , φ''' , равные

$$V(\varphi) = E(\varphi)^2 - (E\varphi)^2 = (D/\bar{\pi})^2 \bar{\pi} - D^2 = D^2(\pi/\bar{\pi}), \quad (2.27)$$

$$V(\varphi''') = E(\varphi''')^2 - (E\varphi''')^2 = \\ = (D^2\pi^2/\bar{\pi}^2) \bar{\pi} - D^2\pi^2 = D^2\pi^3/\bar{\pi}. \quad (2.28)$$

Заметим, что вместо дисперсии $V(\varphi''')$ (2.28) целесообразнее рассчитать дисперсию уклонения оценки φ''' от оцениваемой величины D''' . Для такой характеристики получаем следующее выражение:

$$V(\varphi''' - D''') = E(\varphi''' - D''')^2 = \\ = D^2\pi + (D\pi/\bar{\pi})^2 \bar{\pi} = D^2\pi/\bar{\pi}. \quad (2.29)$$

Сравнивая (2.27) и (2.29), получаем, что дисперсия несмещенной оценки φ предъявленного брака и дисперсия уклонения несмещенной оценки φ''' от оцениваемого пропущенного брака совпадают. Значения характеристики $V(\varphi) = V(\varphi''' - D''')$ выражаются через D , которое нам неизвестно при $d = 0$. Поэтому можно поставить задачу получения несмещенной оценки $\psi = \psi(x)$ для $D^2\pi/\bar{\pi}$. Условие несмещенности имеет вид

$$\psi(0)\pi + \psi(D)\bar{\pi} = D^2\pi/\bar{\pi}. \quad (2.30)$$

Как и выше, полагая сначала $D = 0$, получаем $\psi(0) = 0$,

а затем, разрешая уравнение (2.30) относительно $\psi(D)$, находим, что

$$\psi(x) = \begin{cases} 0, & x=0, \\ D^2\pi/\bar{\pi}^2, & x=D. \end{cases} \quad (2.31)$$

Простые вычисления показывают, что дисперсия оценки (2.31) равна

$$V(\psi) = D^4\pi^3/\bar{\pi}^3. \quad (2.32)$$

Можно вычислить максимальные значения дисперсий (2.27) и (2.32), равные

$$\left. \begin{aligned} V_1 &= \max_{0 \leq D \leq 1} \frac{D^2\pi(N, n, D)}{1 - \pi(N, n, D)}, \\ V_2 &= \max_{0 \leq D \leq 1} \frac{D^4\pi^3(N, n, D)}{(1 - \pi(N, n, D))^3}. \end{aligned} \right\} \quad (2.33)$$

Если воспользоваться приближением (2.4), то

$$\pi(N, n, D) \approx \exp\{-\omega(n/N)nD/N\}, \quad (2.34)$$

где $\omega(x)$ определяется соотношением (2.7). Подставляя (2.34) в (2.33), находим, что

$$\begin{aligned} V_1 &\approx \max_{1 \leq D \leq N} \frac{(D\omega(n/N)(n/N))^2 \exp\{-\omega(n/N)(n/N)D\}}{(\omega(n/N)n/N)^2 (1 - \exp\{-\omega(n/N)(n/N)D\})} \approx \\ &\approx (\omega(n/N)n/N)^{-2} \max_u \frac{u^2}{e^u - 1} \approx 0,65 (\omega(n/N)n/N)^{-2}. \end{aligned} \quad (2.35)$$

Аналогичные вычисления дают приближенное выражение для V_2 :

$$V_2 \approx 0,23 (\omega(n/N)n/N)^{-4}. \quad (2.36)$$

Приближенные оценки (2.35) и (2.36) были получены А. Н. Колмогоровым [5].

Приведенные выше формулы можно использовать для оценки эффективности бездефектного контроля по результатам разбраковки многих совокупностей. Поэтому найденные выше оценки называют **последующими**. Если контролировались совокупности $\mathfrak{F}_i \propto [N_i, D_i]$, $i = 1, \dots, m$, то можно ожидать, что для суммарного

предъявленного брака $\sum_{1 \leq i \leq m} D_i$ и суммарного пропущенного брака $\sum_{1 \leq i \leq m} D_i'''$ случайные величины

$$\xi_1 = \frac{\sum_{i=1}^m D_i - \sum_{i=1}^m \varphi_i}{\sqrt{\sum_{i=1}^m V(\varphi_i)}}, \quad \xi_2 = \frac{\sum_{i=1}^m D_i''' - \sum_{i=1}^m \varphi_i'''}{\sqrt{\sum_{i=1}^m V(\varphi_i)}}, \quad (2.37)$$

где φ_i, φ_i''' находятся по формулам (2.26), а $V\varphi_i$ — по (2.27), асимптотически при $m \rightarrow \infty$ являются нормально распределенными случайными величинами с нулевым средним и единичной дисперсией. Такое предположение основывается на допущении, что $D_i - \varphi_i$ ($D_i''' - \varphi_i'''$) есть взаимно независимые по i случайные величины, для которых выполняются условия центральной предельной теоремы о сходимости распределений суммы взаимно независимых случайных величин к нормальному распределению (см. Гнеденко [1], стр. 258). Точность последующих оценок можно оценить средним квадратическим отклонением, т. е.

$$\sum_{i=1}^m D_i \approx \sum_{i=1}^m \varphi_i \pm \sqrt{\sum_{i=1}^m V(\varphi_i)}, \quad (2.38)$$

$$\sum_{i=1}^m D_i''' \approx \sum_{i=1}^m \varphi_i''' \pm \sqrt{\sum_{i=1}^m V(\varphi_i)}. \quad (2.39)$$

В (2.38) и (2.39) значения, стоящие под корнем, нам неизвестны, поэтому вместо них можно использовать верхние границы для $V(\varphi_i)$, определяемые (2.33), или их приближенные выражения (2.35). Другой подход состоит в использовании оценок ψ_i и $V_{2,i}$, определяемых формулами (2.31) и (2.33). При таком подходе

$$\sum_{i=1}^m D_i \approx \sum_{i=1}^m \varphi_i \pm \sqrt{\sum_{i=1}^m \psi_i + \sum_{i=1}^m V_{2,i}}, \quad (2.40)$$

$$\sum_{i=1}^m D_i''' \approx \sum_{i=1}^m \varphi_i''' \pm \sqrt{\sum_{i=1}^m \psi_i + \sum_{i=1}^m V_{2,i}}. \quad (2.41)$$

В работе Колмогорова [5] отмечается, что вторые слагаемые в формулах (2.40), (2.41) есть приемлемый суррогат среднего квадратического отклонения.

§ 1.3. Гипергеометрическое распределение

Основные соотношения. Пусть $\mathfrak{P} = \{O_1, \dots, O_N\}$ — контролируемая совокупность объема N , объекты которой классифицируются по альтернативному признаку. Контроль проводим, используя случайную выборку без возвращения объема n . Выведем формулу для вероятности $h_{N,D}^{n,d}$ обнаружения в выборке d дефектных объектов при условии, что общее число дефектных объектов в $\mathfrak{P}$ равно D ($\mathfrak{P} \asymp [N, D]$). Эту вероятность можно подсчитать различными способами. В первом способе исходим из того, что все выборки по n объектов из N равновероятны. Число таких выборок равно $\binom{N}{n} = \frac{N!}{n!(N-n)!} = N^{(n)}/n!, n=0, \dots, N, \binom{N}{n}=0$, если $n < 0$ или $n > N$. При этом число выборок из D дефектных объектов по d дефектных объектов равно $\binom{D}{d}$, а число выборок из $K = N - D$, годных по $k = n - d$ годных объектов, равно $\binom{K}{k}$. Общее число событий, соответствующих появлению d дефектных объектов в выборке, равно произведению $\binom{D}{d}\binom{K}{k}$. Искомую вероятность $h_{N,D}^{n,d}$ получаем как отношение числа благоприятных исходов к общему числу исходов, т. е.

$$h_{N,D}^{n,d} = \frac{\binom{D}{d}\binom{K}{k}}{\binom{N}{n}}, \quad k + d = n, \quad K + D = N, \quad (3.1)$$

при $\max(0, D + n - N) \leq d \leq \min(n, D)$; в противном случае полагаем $h_{N,D}^{n,d} = 0$.

Легко проверить, что $\sum_{0 \leq d \leq n} h_{N,D}^{n,d} = 1$. Распределение, соответствующее вероятностям (3.1), называется *гипергеометрическим распределением*. Это одно из

основных распределений в теории вероятностей и его обычно кратко упоминают почти во всех учебниках по теории вероятностей, см., например, книгу Феллера [1]. Такое название распределению с вероятностями (3.1) дано потому, что производящая функция $\sum_d h_{N,D}^{n,d} z^d$ входит в класс гипергеометрических производящих функций (Бейтмен, Эрдейи [1], Кендалл и Стьюарт [1], стр. 189). Поскольку в задачах контроля качества это распределение играет особую роль, мы подробно рассмотрим его свойства в настоящем параграфе.

Воспользуемся соотношением $m^{(r)} = \frac{m!}{(m-r)!}$. Тогда формулу (3.1) можно переписать в следующем виде:

$$h_{N,D}^{n,d} = \frac{D^{(d)} n! K^{(k)}}{d! N^{(n)} k!} = \binom{n}{d} \frac{D^{(d)} K^{(k)}}{N^{(n)}}. \quad (3.2)$$

Так как при $k + d = n$ имеем $\binom{n}{d} = \binom{n}{k}$, то из (3.2) находим, что

$$h_{N,D}^{n,d} = h_{N,K}^{n,k}. \quad (3.3)$$

Далее, замечая, что $\binom{n}{d} = n^{(d)} / d!$, $D^{(d)} / d! = \binom{D}{d}$, из (3.2) получаем

$$h_{N,D}^{n,d} = \frac{D^{(d)} n^{(d)} K^{(k)}}{d! N^{(n)}} = \binom{D}{d} \frac{n^{(d)} K^{(k)}}{N^{(n)}} = h_{N,n}^{D,d}. \quad (3.4)$$

Формулу (3.4) можно было бы вывести и непосредственно. Пусть N — номера мест, занимаемых объектами, среди которых указано n мест, образующих выборку. Рассматриваем всевозможные размещения D дефектных объектов по N местам. Число таких размещений $\binom{N}{D}$. Событию, при котором в выборке обнаружено ровно d дефектных объектов, благоприятствуют размещения d дефектных объектов на n указанных для выборки местах, и $D - d$ дефектных объектов на $N - n$ местах, не включенных в выборку. Общее число благоприятных исходов равно $\binom{n}{d} \binom{N-n}{D-d}$. Искомую вероят-

ность $h_{N,D}^{n,d}$ получаем как отношение числа благоприятных исходов к общему числу исходов. т. е.

$$h_{N,D}^{n,d} = \frac{\binom{n}{d} \binom{N-n}{D-d}}{\binom{N}{D}} = h_{N,n}^{D,d}. \quad (3.5)$$

Формула (3.5) совпадает с (3.4). Используя (3.3) — (3.5), мы можем выписать и другие возможные тождества:

$$h_{N,D}^{n,d} = h_{N,n}^{D,d} = h_{N,K}^{n,k} = h_{N,n}^{K,k} = h_{N,N-n}^{D,D-d} = h_{N,D}^{N-n,D-d}, \quad (3.6)$$

где $N \geq D$, $N \geq n$, $D \geq d$, $n \geq d$.

Для функции гипергеометрического распределения будем использовать обозначение

$$H_{N,D}^{n,d} = \sum_{j=0}^d h_{N,D}^{n,j}, \quad \bar{H}_{N,D}^{n,d} = \sum_{j=d+1}^n h_{N,D}^{n,j}. \quad (3.7)$$

На основе соотношений (3.6) легко выводятся тождества для $H_{N,D}^{n,d}$, $\bar{H}_{N,n}^{D,d}$, например, $H_{N,D}^{n,d} = H_{N,n}^{D,d} = \bar{H}_{N,D}^{N-n,D-d-1}$ и т. д.

Гипергеометрическое распределение определяется заданием четырех величин, N , D , n , d . Поэтому таблицы гипергеометрических вероятностей занимают очень много места. Небольшие таблицы гипергеометрического распределения включены в широко распространенные сборники таблиц математической статистики, см. Бولшев и Смирнов [1], Оуэн [1]. Наиболее полные таблицы имеются в специальной книге Либермана и Оуэна [1], объем которой превышает 600 страниц. В этой книге приведены вероятности $H = H_{N,D}^{n,d}$ для следующих наборов параметров:

$$\begin{aligned} N=2(1)25, & \quad 1 \leq D \leq n \leq N-1, \quad 0 \leq d \leq D, \\ N=26(1)50, & \quad 1 \leq D \leq (N/2), \quad 0 \leq d \leq D, \\ N=60(10)100, & \quad 1 \leq D \leq (N/2), \quad 0 \leq d \leq D, \\ N=1000, n=500, & \quad 1 \leq D \leq 50, \quad 0 \leq d \leq (D/2), \quad (H > 10^{-6}), \\ N=100(100)2000, & \quad n=(N/2), \quad D=n, \quad D=n-1, \\ & \quad 0 \leq d \leq (D/2), \quad (H > 10^{-6}). \end{aligned}$$

Мы видим, что в таблицы включены не все возможные комбинации значений N, D, n, d . Пользуясь соотношениями (3.6), можно расширить область параметров. Покажем, что всегда достаточно иметь таблицы, в которых табулированы значения $H_{N,D}^{n,d}$ в условиях

$$D \leq n \leq N - n \leq N - D = K. \quad (3.8)$$

Действительно, если $D > K$, то можно воспользоваться соотношением $h_{N,D}^{n,d} = h_{N,K}^{n,k}$. Если $n < D \leq K$, то используем соотношение $h_{N,D}^{n,d} = h_{N,n}^{D,d} (h_{N,K}^{n,k} = h_{N,n}^{K,k})$ и т. д. Проиллюстрируем эту простую технику преобразований следующим примером. Пусть нам потребуется вычислить $H_{200,110}^{30,10}$, тогда как имеются лишь таблицы, удовлетворяющие (3.8). В примере $N=200, D=110, n=30, d=10$, т. е. (3.8) не выполнено, $D > K$. Сделав замену согласно (3.6), находим, что $H_{200,110}^{30,10} = H_{200,30}^{110,10}$. Однако в новом наборе $N=200, D=30, n=110, d=10$, соотношение (3.8) все еще нарушено, $N - n < n$. Поэтому надо еще раз воспользоваться (3.6) ($h_{N,D}^{n,d} = h_{N,D}^{N-n,D-d}$), после чего получаем $H_{200,30}^{110,10} = \bar{H}_{200,30}^{90,19} = 1 - H_{200,30}^{90,19}$. Здесь имеем $N=200, D=30, n=90, d=19$. Соотношение (3.8) для $H_{200,30}^{90,19}$ выполнено.

Для оценки $\max_d h_{N,D}^{n,d} = h_{N,D}^{n,d_0}$ — максимального значения вероятности гипергеометрического распределения, — заметим, что из (3.2) легко выводится соотношение

$$h_{N,D}^{n,d+1} / h_{N,D}^{n,d} = \frac{n-d}{d+1} \cdot \frac{D-d}{(N-D)-(n-d)+1}.$$

Правая часть этого соотношения убывает с ростом d . Поэтому максимальное значение достигается только при одном или двух соседних значениях d_0 . Величина d_0 находится из неравенств

$$\begin{aligned} \frac{(n-d_0)(D-d_0)}{(d_0+1)(N-D-(n-d_0)+1)} &\leq 1 \leq \\ &\leq \frac{(n-(d_0-1))(D-d_0+1)}{d_0(N-D-(n-d_0))}. \end{aligned}$$

Если дробь справа для всех d_0 меньше 1, то $d_0 = 0$.

В остальных случаях, разрешая эти неравенства относительно d_0 , получим

$$\frac{(D+1)(n+1)}{N+2} - 1 \leq d_0 \leq \frac{(D+1)(n+1)}{N+2}.$$

Лемма 1.3.1. Если $d \geq (D+1)(n+1)/(N+2)$, то

$$\bar{H}_{N,D}^{n,d-1} \leq h_{N,D}^{n,d} \frac{(d+1)(N-D-(n-d)+1)}{(d+1)(N-D-(n-d)+1) - (n-d)(D-d)} \leq \\ \leq (n-d) h_{N,D}^{n,d};$$

если $d \leq \frac{(D+1)(n+1)}{N+1} - 1$, то

$$H_{N,D}^{n,d} \leq h_{N,D}^{n,d} \frac{(n-d-1)(D-d+1)}{(n-d-1)(D-d+1) - d(N-D-n+d)} \leq \\ \leq d h_{N,D}^{n,d}.$$

Доказательство. Ограничимся выводом второго неравенства. Так как при $m \leq \frac{(D+1)(n+1)}{N+2} - 1$ выполнено $h_{N,D}^{n,m} < h_{N,D}^{n,m+1}$, то очевидно, что $H_{N,D}^{n,d} \leq d h_{N,D}^{n,d}$. Более точную оценку найдем следующим образом. Имеем

$$H_{N,D}^{n,d} = \sum_{m=0}^d \prod_{r=m}^{d-1} \frac{h_{N,D}^{n,r}}{h_{N,D}^{n,r+1}} h_{N,D}^{n,d} \leq \\ \leq \sum_{m=0}^d (h_{N,D}^{n,d-1} / h_{N,D}^{n,d})^{d-m} h_{N,D}^{n,d} \leq \{1 - (h_{N,D}^{n,d-1} / h_{N,D}^{n,d})\}^{-1} h_{N,D}^{n,d} = \\ = \frac{(n-d+1)(D-d)}{(n-d+1)(D-d) - d(N-D-n+d)} h_{N,D}^{n,d},$$

где коэффициент при $h_{N,D}^{n,d}$ при $r=d$ полагаем равным единице. Аналогичным образом, используя свойства монотонности отношения $h_{N,D}^{n,d+1} / h_{N,D}^{n,d}$, выводим неравенство для $\bar{H}_{N,D}^{n,d-1}$.

Моментные характеристики. Можно легко вычислить математическое ожидание и дисперсию гипергеометрического распределения. В наших задачах контроля это соответствует математическому ожиданию и дисперсии числа обнаруженных в выборке дефектных объектов.

Используя соотношение $\binom{r}{s} = \frac{r}{s} \binom{r-1}{s-1}$ и (3.5), получаем

$$h_{N,D}^{n,d} = \frac{n \binom{n-1}{d-1} \binom{N-n}{D-d}}{d \frac{N}{D} \binom{N-1}{D-1}} = \left(\frac{nD}{Nd} \right) h_{N-1,D-1}^{n-1,d-1}. \quad (3.9)$$

Из (3.9) находим выражение для среднего случайной величины δ_n — числа дефектных объектов, обнаруженных в выборке объема n ,

$$\begin{aligned} \lambda = E\delta_n &= \sum_{d=0}^n d h_{N,D}^{n,d} = \sum_{d=1}^{n-1} d (nD/Nd) h_{N-1,D-1}^{n-1,d-1} = \\ &= \frac{nD}{N} \sum_{d=0}^{n-1} h_{N-1,D-1}^{n-1,d} = \frac{nD}{N}. \end{aligned} \quad (3.10)$$

Также легко вычисляется и дисперсия. Здесь, дважды используя (3.9), находим

$$\begin{aligned} E\delta_n (\delta_n - 1) &= \sum_{d=2}^n d(d-1) h_{N,D}^{n,d} = \\ &= \sum_{d=2}^{n-2} \frac{n(n-1)D(D-1)}{N(N-1)} \cdot h_{N-2,D-2}^{n-2,d-2} = \frac{n^{(2)}D^{(2)}}{N^{(2)}}. \end{aligned} \quad (3.11)$$

Из (3.10) и (3.11) получаем, что искомая дисперсия равна

$$\begin{aligned} \sigma^2 &= E(\delta_n - \lambda)^2 = E\delta_n^2 - \lambda^2 = E\delta_n(\delta_n - 1) + E\delta_n - \lambda^2 = \\ &= D^{(2)}(n^{(2)}/N^{(2)}) + (n/N)D - (Dn/N)^2 = \tau^2 N/(N-1), \end{aligned} \quad (3.12)$$

где

$$\tau^2 = N(D/N)(K/N)(n/N)(1 - n/N). \quad (3.13)$$

Общий результат о моментах гипергеометрического распределения сформулируем в виде следующего утверждения.

Теорема 1.3.1. Для смешанных факториальных моментов случайных чисел δ_n , обнаруженных в выборке

дефектных объектов, и $(D - \delta_n)$ — дефектных объектов, не попавших в выборку, выполнено соотношение

$$E\delta_n^{(r)}(D - \delta_n)^{(s)} = \frac{n^{(r)}(N - n)^{(s)}D^{(r+s)}}{N^{(r+s)}}. \quad (3.14)$$

Доказательство. Легко проверить следующие простые тождества:

$$\begin{aligned} d^{(r)} \binom{n}{d} &= n^{(r)} \binom{n-r}{d-r}, \quad (D - d)^{(s)} D^{(d)} = D^{(d+s)} = \\ &= D^{(r+s)} [D - (r + s)]^{(d-r)}, \quad N^{(n)}(N - n)^{(s)} = N^{(n+s)} = \\ &= N^{(r+s)} [N - (r + s)]^{(n-r)}. \end{aligned}$$

Последовательно используя эти тождества, получаем

$$\begin{aligned} E\delta_n^{(r)}(D - \delta_n)^{(s)} &= \sum_{d=0}^n d^{(r)}(D - d)^{(s)} \binom{n}{d} \frac{K^{(k)}D^{(d)}}{N^{(n)}} = \\ &= (N - n)^{(s)} \sum_{d=r}^{n-r} d^{(r)} \frac{n^{(r)}}{d^{(r)}} \binom{n-r}{d-r} \frac{K^{(k)}D^{(d+s)}}{N^{(n+s)}} = \\ &= n^{(r)}(N - n)^{(s)} \sum_{d=r}^{n-r} \binom{n-r}{d-r} \frac{K^{(k)}D^{(t)}(D - t)^{(d-r)}}{N^{(t)}(N - t)^{(n-r)}} = \\ &= \frac{n^{(r)}(N - n)^{(s)}D^{(t)}}{N^{(t)}} \sum_{d=r}^{n-r} \binom{n-r}{d-r} \frac{(N - t - [D - t])^{(k)}(D - t)^{(d-r)}}{(N - t)^{(n-r)}} = \\ &= \frac{n^{(r)}(N - n)^{(s)}D^{(t)}}{N^{(t)}} \sum_{d'=0}^{n-r} h_{N-t, D-t}^{n-r, d'} \end{aligned}$$

где $t = r + s$. Последняя сумма равна 1, поэтому (3.14) выполнено.

Соотношение (3.14) мы используем позднее при изучении априорных распределений для числа D дефектных объектов. Из (3.14) после элементарных преобразований можно найти выражения для третьего и четвертого центральных моментов δ_n , которые имеют

следующий вид:

$$E(\delta_n - \lambda)^3 = \sigma^2 \frac{(K - D)}{N} \cdot \frac{(N - 2n)}{N - 2},$$

$$E(\delta_n - \lambda)^4 = \frac{\sigma^2}{(N - 2)(N - 3)} \{N(N + 1) - 6n(N - n) + \\ + 3(D/N)(K/N)[N^2(n - 2) - Nn^2 - 6n(N - n)]\}.$$

Аналогично выводу формулы (3.14), используя тождества (3.6), можно получить другие формулы моментных характеристик. Например,

$$E\delta_n^{(r)}(n - \delta_n)^{(s)} = \sum_{d=0}^n d^{(r)}(n - d)^{(s)} h_{N,D}^{n,d} = \\ = \sum_{d=0}^n d^{(r)}(n - d)^{(s)} h_{N,n}^{D,d} = \frac{D^{(r)}(N - D)^{(s)} n^{(r+s)}}{N^{(r+s)}}.$$

Аппроксимация другими распределениями. Поскольку таблицы для гипергеометрического распределения при больших значениях параметров N, n, D, d отсутствуют, то естественно поставить вопрос об отыскании предельных распределений для δ_n — числа дефектных объектов в выборке, когда $N \rightarrow \infty$. В силу сказанного выше о возможности замены значений аргументов будем предполагать выполненным соотношение (3.8). Заметим, что из (3.8) следует, что для $n \leq N/2$ имеет место неравенство $1/4 \leq \tau^2/\lambda \leq 1$, поскольку $\tau^2/\lambda = (1 - n/N)(1 - D/N)$. Оказывается, что возможными предельными распределениями для δ_n могут быть только распределения одного из следующих четырех типов:

- (1) вырожденное (D_0), когда $P\{\delta_n = 0\} \rightarrow 1$,
- (2) пуассоновское (P_λ), когда $P\{\delta_n = d\} \rightarrow e^{-\lambda} \lambda^d / d!$,
- (3) v — биномиальное ($B(v)$), когда $P\{\delta_n = d\} \rightarrow \binom{D}{d} v^d (1 - v)^{D-d}$,
- (4) нормальное $\Phi_{\lambda, \tau}$, когда $P\left\{\frac{\delta_n - \lambda}{\tau} \leq x\right\} \rightarrow \Phi(x)$,

$$\left(\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp(-u^2/2) du \right).$$

Сходимость к нормальному распределению нормированной и центрированной случайной величины $\frac{\delta_n - \lambda}{\tau}$ будет тогда и только тогда, когда $nD/N \rightarrow \infty$. Доказательство во многом аналогично приведенному в книге Феллера [1], стр. 185, где рассмотрен аналогичный вопрос аппроксимации биномиального распределения нормальным. Сходимость распределения δ_n к биномиальному ($B(v)$) будет тогда и только тогда, когда $nD/N \rightarrow \lambda < \infty$, $n/N \rightarrow v \neq 0$. Для сходимости к пуассоновскому распределению P_λ необходимо и достаточно, чтобы $nD/N \rightarrow \lambda < \infty$, а $n/N \rightarrow 0$. Наконец, вырожденное распределение D_0 будет только в том случае, когда $nD/N \rightarrow 0$. Указание об аппроксимации гипергеометрического распределения нормальным дано в книге Хальда [1], сходимость к пуассоновскому распределению изучал Колмогоров [6]. Полное рассмотрение всех возможных распределений изложено в статье Ван Еедена и Ранненбурга [1]. Обычно из сходимости функции распределения δ_n к одному из распределений P_λ , $B(v)$ или $\Phi_{\lambda, \tau}$ делают вывод, что можно воспользоваться одним из следующих соотношений:

$$\begin{aligned}
 P\{\delta_n \leq d\} &= H_{N, D}^{n, d} \approx P_\lambda(d) = \sum_{r=0}^d (\lambda^r / r!) e^{-\lambda}, \\
 \lambda &= nD/N, \\
 P\{\delta_n \leq d\} &= H_{N, D}^{n, d} \approx B_D^d(v) = \sum_{r=0}^d \binom{D}{r} v^r (1-v)^{D-r}, \quad (3.15) \\
 v &= n/N, \\
 P\{\delta_n \leq d\} &= H_{N, D}^{n, d} \approx \Phi\left(\frac{d - \lambda}{\tau}\right), \\
 \lambda &= nD/N, \quad \tau^2 = \lambda(K/N)(1 - n/N).
 \end{aligned}$$

Однако такое «непосредственное» использование предельных теорем может давать существенные погрешности. В книге Большева и Смирнова [1] в разделе, посвященном гипергеометрическому распределению, есть указания о необходимости получения более точных аппроксимаций. Недавно интересные результаты были получены Молёнаром [1], который, применив современные

вычислительные средства, вывел сравнительно простые формулы, позволяющие довольно хорошо аппроксимировать пуассоновское, биномиальное и гипергеометрическое распределения. В книге Моленара [1] содержится большое число примеров, показывающих преимущества и недостатки различных аппроксимирующих формул.

В самых общих чертах замысел получения асимптотически более точных аппроксимирующих формул состоит в следующем. Пусть при $m \rightarrow \infty$ последовательность функций распределений $F_m(x) \rightarrow F(x)$ для каждого фиксированного x . Идеальным решением было бы отыскание такой функции $u_m(x)$, чтобы $F(u_m(x)) = F_m(x)$. Однако найти точное выражение для $u_m(x)$ столь же сложно, сколь сложно решение исходной задачи. Если мы будем считать возможной замену $u_m(x)$ на x , что является самым начальным приближением, то мы приходим к соотношению $F_m(x) \approx F(x)$. В нашей задаче аппроксимации гипергеометрического распределения такая замена соответствует соотношениям (3.15). Однако иногда удается получить асимптотическое разложение $u_m(x) = u_m^1(x) + u_m^2(x) + \dots$ такое, что аппроксимации $F(u_m^1(x)) \approx F(x)$, $F(u_m^1(x) + u_m^2(x)) \approx F(x)$ и т. д. будут все более и более точными. Успех будет в том случае, если выражения для $u_m^i(x)$ не слишком сложны, а выигрыш в точности аппроксимации значителен. Именно так и обстоит дело с гипергеометрическим распределением.

Не останавливаясь на технике вывода приводимых ниже рекомендаций Моленара [1], так как это заняло бы слишком много места, мы дадим лишь только их формулировки. Прежде всего укажем, что при значениях $H_{N,D}^n = P\{\delta_n \leq d\}$ в интервалах (0,005, 0,05) и (0,95, 0,995) точность аппроксимации распределением $P_a\{\delta_n \leq d\}$ целесообразно оценивать по величине отношений «хвостов», т. е. по величине отношений

$$\frac{P_a(\delta_n \leq d) - P(\delta_n \leq d)}{P(\delta_n \leq d)} \cdot 100\%, \quad \text{если } P(\delta_n \leq d) \leq 0,5,$$

$$\frac{P_a(\delta_n \geq d) - P(\delta_n \geq d)}{P(\delta_n \geq d)} \cdot 100\%, \quad \text{если } P(\delta_n \geq d) > 0,5.$$

Моленар показал, что гипергеометрическое распределение очень хорошо аппроксимируется биномиальным. Поэтому при наличии таблиц биномиального распределения $B_D^d(v)$ можно использовать соотношение

$$H_{N,D}^{n,d} \approx B_D^d(\bar{v}), \quad (3.16)$$

где

$$\bar{v} = \frac{2n-d}{2N-D+1} \quad (3.17)$$

для быстрых расчетов и

$$\bar{v} = \frac{2n-d}{2N-D+1} - \frac{2D(k+1/2-nD/N)}{3(2N-D+1)^2} \quad (3.18)$$

для более точной аппроксимации. При наличии только таблиц пуассоновского распределения $P_\lambda(d)$

$$H_{N,D}^{n,d} \approx P_\lambda(d), \quad (3.19)$$

где нужно положить при $(D/N) \leq (n/N) \leq 0,1$

$$\bar{\lambda} = \frac{(2D-d)(2n-d)}{2(2N-D-n+1)}, \quad (3.20)$$

а при $(n/N) > 0,1$ нужно считать

$$\bar{\lambda} = (nD/N) + (1/(3N))((nD/N) - d)(2n-D+10Dn/N). \quad (3.21)$$

При аппроксимации гипергеометрического распределения нормальным Моленар рекомендует следующие формулы:

для значений $H_{N,D}^{n,d}$ в интервалах (0,005, 0,05) и (0,95, 0,995)

$$H_{N,D}^{n,d} \approx \Phi\left(\frac{2}{\sqrt{N-1}}\{\sqrt{(d+1)(N-D-k+1)} - \sqrt{(D-d)k}\}\right), \quad (3.22)$$

для значений $H_{N,D}^{n,d}$ в интервале (0,05, 0,95)

$$H_{N,D}^{n,d} \approx \Phi\left\{(2/\sqrt{N})\sqrt{(d+3/4)(N-D-k+3/4)} - \sqrt{(D-d-1/4)(k-1/4)}\right\}, \quad (3.23)$$

где $k = n - d$.

Сравнительную точность различных аппроксимаций проиллюстрируем численным примером. Пусть $D = n = 20$, $N = 200$. В приводимой ниже таблице 1.3.1

во втором столбце даны значения $H_{200, 20}^{20, d}$, $d = 0(1)4$. В последующих вертикальных столбцах даны относительные ошибки, т. е. отношения «хвостов» распределений, выраженные в %, когда вычисления проводились по формулам, номера которых указаны в первой строчке каждого столбца.

Таблица 1.3.1

Сравнительная точность аппроксимаций гипергеометрического распределения (Моленар [1])

d	$H_{N, D}^{n, d}$	(3.22)	(3.23)	(3.15)	(3.20)	(3.21)	(3.17)	(3.18)
0	0,1085	+38,16	-1,01	+24,68	+9,12	+0,43	+0,23	-0,08
1	0,3782	+17,30	+3,83	+7,35	+2,66	-0,08	+0,06	-0,01
2	0,6787	-12,90	-1,96	+0,62	+0,62	+0,62	+0,08	-0,00
3	0,8778	-9,17	+4,75	+16,88	+7,21	+1,96	+0,43	+0,23
4	0,9655	+3,17	+21,39	+52,46	+19,94	+4,37	+1,16	+0,70

Результаты таблицы 1.3.1 показывают, что аппроксимация биномиальным распределением дает очень хорошие результаты. Заметное увеличение точности по сравнению с (3.15) получаем при пуассоновской аппроксимации (3.20), (3.21), когда приближаемые вероятности малы или близки к единице.

Расчет параметров одноступенчатых планов. Рассмотрим применение гипергеометрического распределения в задаче планирования выборочного контроля. В практике статистического контроля качества широко используются так называемые одноступенчатые планы.

Определение 1.3.1. Одноступенчатым планом контроля с параметрами (n, c) называется такое правило обследования совокупности объектов $\mathfrak{F}$, при котором из $\mathfrak{F}$ берется случайная выборка без возвращения объема n . Если число d обнаруженных в выборке дефектных объектов больше c , то совокупность $\mathfrak{F}$ бракуется, если же $d \leq c$, то совокупность $\mathfrak{F}$ принимается. Число c называется **приемочным**.

Определение 1.3.2. Оперативной характеристикой одноступенчатого плана контроля называется вероятность принятия контролируемой совокупности $\mathfrak{F}$,

рассматриваемая как функция числа D дефектных объектов в $\mathfrak{F}$.

Оперативная характеристика одноступенчатого плана контроля с параметрами (n, c) равна $\pi(D) = H_{N,D}^{n,c}$.

Часто при планировании контроля на оперативную характеристику искомого плана накладываются ограничения.

Типичной является следующая задача планирования выборочного контроля. Требуется найти значение объема плана n и приемочное число c , чтобы для заданных чисел дефектных объектов $D_1 < D_2$ вероятность принятия $\mathfrak{P} \asymp [N, D_1]$ была бы не менее $1 - \alpha$, а вероятность принятия $\mathfrak{P} \asymp [N, D_2]$ — не более β . Числа α, β называют **ошибками поставщика и потребителя (приемщика)**, или **ошибками первого и второго рода**. В книге Шора и Кузьмина [1] приведены таблицы для приближенного решения подобных задач. По-видимому, для малых значений α, β более точные аппроксимации $H_{N,D}^{n,c}$ в духе приведенных выше рекомендаций Моленара дадут хорошие результаты. Если исходить из пуассоновской аппроксимации (3.20), (3.21), то объем выборки n и приемочное число c одноступенчатого плана с ошибками первого и второго рода, не превышающими α, β , можно найти следующим способом. Из таблицы XII находим процентные точки $\lambda_q = \lambda_q(d)$, для которых $P_{\lambda_q(d)}(d) = q$, $d = 0, 1, \dots, q = \alpha, \beta$. Для значений (N, D_2, c) из уравнения

$$\lambda_\beta(c) = \frac{(2D_2 - c)(2n_c - c)}{2(2N - D_2 - n_c + 1)} \quad (3.24)$$

находим число n_c , которое после округления до ближайшего целого будет соответствовать объему выборочного плана. Если $(n_c/N) > 0,1$, то производим пересчет n_c , исходя из уравнения

$$\lambda_\beta(c) = \frac{n_c D_2}{N} + \frac{1}{3N} \left(\frac{n_c D_2}{N} - c \right) \left(2n_c - D_2 + 10 \frac{n_c D_2}{N} \right). \quad (3.25)$$

При этом мы должны следить за выполнением условий (3.8) ($d = c$). Если условия (3.8) для набора (N, D_2, n_c, c) нарушены, то делаются очевидные замены аргументов, для того чтобы (3.8) выполнялось. После этого

расчет проводится по формулам (3.24) или (3.25). Исходя из найденного числа n_c , вычисляем

$$\bar{\lambda}(c) = \begin{cases} \frac{(2D_1 - c)(2n_c - c)}{2(2N - D_1 - n_c + 1)}, & (n_c/N) \leq 0,1, \\ \frac{n_c D_1}{N} + \frac{1}{3N} \left(\frac{n_c D_1}{N} - c \right) \left(2n_c - D_1 + 10 \frac{n_c D_1}{N} \right), & (n_c/N) > 0,1. \end{cases} \quad (3.26)$$

Если $\bar{\lambda}(c) \leq \lambda_{1-\alpha}(c)$, то план с объемом выборки n_c и приемочным числом c — искомый. Если же $\bar{\lambda}(c) > \lambda_{1-\alpha}(c)$, то расчет повторяется со значением $c' = c + 1$ и т. д. до тех пор, пока впервые не будет выполнено неравенство $\bar{\lambda}(c') \leq \lambda_{1-\alpha}(c')$. Так как метод расчета приближенный, то требование $\bar{\lambda}(c) < \lambda_{1-\alpha}(c)$ можно несколько ослабить, отбирая план (n_c, c) , если $\bar{\lambda}(c) > \lambda_{1-\alpha}(c)$, но значение $P_{\bar{\lambda}(c)}(c)$ не намного меньше $1 - \alpha$, тогда как переход к следующему значению $c + 1$ существенно увеличивает объем выборки.

Проиллюстрируем этот метод двумя примерами. Пусть $N = 15$, $D_1 = 2$, $D_2 = 8$, $\alpha = \beta = 0,1$. Значения процентных точек $\lambda_{0,1}(1) = 3,890$, $\lambda_{0,9}(1) = 0,532$. Из (3.24) для $D_2 = 8$ и $c = 1$ находим $n_c = 5,13$. Округляя до целого, получаем $n_c = 5$. Нам нужно рассчитать теперь значение оперативной характеристики $H_{15,2}^{5,1}$. Поэтому, замечая, что $(n_c/N) > 0,1$, расчет $\bar{\lambda}(1)$ проводим по нижней формуле (3.26). Получаем значение $\bar{\lambda}(1) = 0,558 > \lambda_{0,9}(1) = 0,532$. Так как $H_{15,2}^{5,1} \approx P_1(\bar{\lambda}(1)) = 0,892$, т. е. близко к 0,9, то расчет для $c = 2$ не проводим. Отобранный план имеет объем выборки $n = 5$, $c = 1$. Точный расчет дает значения его оперативной характеристики $H_{15,2}^{5,1} = 0,9047$, $H_{15,8}^{5,1} = 0,1002$.

Во втором примере $N = 10\,000$, $D_1 = 25$, $D_2 = 78$, $\alpha = \beta = 0,1$. Числа $c = 0, 1, 2, 3$ не дают положительных результатов. Проводим расчет для $c = 4$. Здесь $\lambda_{0,1}(4) = 7,994$, $\lambda_{0,9}(4) = 2,433$. Из (3.24) находим $n_c = 997,29$ или, округляя, $n_c = 997$. Так как $(n_c/N) \leq 0,1$, то $\bar{\lambda}(4)$ рассчитываем по верхней формуле (3.26) и находим $\bar{\lambda}(4) = 2,412 < \lambda_{0,9}(4)$. Таким образом, у искомого плана $n = 997$, $c = 4$. Точное решение, получен-

ное путем расчетов на ЭВМ, дает план $n = 1000$, $c = 4$, $H_{10000,25}^{1000,4} = 0,9022$, $H_{10000,78}^{1000,4} = 0,0986$. Найденное приближение достаточно хорошее.

Приведенный выше метод, конечно, не может рассматриваться как опробованная рекомендация, а скорее призван привлечь внимание к проблеме получения сравнительно простых, но достаточно точных методов планирования выборочного контроля, основанных на аппроксимациях гипергеометрического распределения. Заметим, что вопросы планирования выборочного контроля рассматривал Хальд [9].

Завершая этот раздел, еще раз напомним, что мы предполагаем выполненными соотношения (3.8). Поэтому, если мы хотим воспользоваться предложенными выше рекомендациями в тех случаях, когда (3.8) нарушается, надо сначала сделать переход к набору параметров, соответствующих условиям (3.8). Например, при приближенном вычислении $H_{500,50}^{30,3}$ из (3.6) сначала находим, что $H_{500,50}^{30,3} = H_{500,30}^{50,3}$, а далее используем приведенные выше рекомендации, так как для $N = 500$, $D = 30$, $n = 50$, $d = 3$ соотношения (3.8) выполнены.

В статье Дугиной [1] приведены алгольные программы вычисления процентных точек гипергеометрического распределения. Программы используют асимптотические свойства гамма-функции, и потому время счета не зависит от N . Так как для гипергеометрического распределения выполнены неравенства

$$H_{N,D}^{n,d} \geq H_{N,D+1}^{n,d} \quad \text{и} \quad H_{N,D}^{n,d} \leq H_{N+1,D}^{n,d},$$

то при заданном значении q программа вычисляет пару значений D_q^* , $D_q^* + 1$ таких, что

$$H_{N,D_q^*+1}^{n,d} \leq q \leq H_{N,D_q^*}^{n,d}.$$

В окончательных результатах вычислений указывается число $D_q = D_q^* + 1$ или $D_q = D_q^*$ в зависимости от того, минимально значение $\Delta' = q - H_{N,D_q^*+1}^{n,d}$ или же $\Delta = H_{N,D_q^*}^{n,d} - q$. При заданных значениях N , n , d процентные точки D_q выдаются для выбранного заранее набора

значений q . Таким образом, числа $\{H_{N,D}^{n,d}\}$ соответствуют точкам оперативной характеристики одноступенчатого плана (n, d) . В таблице 1.3.2 приведен пример расчета оперативной характеристики одноступенчатого плана $n=100$, $c=3$ по программе Дугиной [1], когда $N=1000$.

Т а б л и ц а 1.3.2

Оперативная характеристика одноступенчатого плана
(гипергеометрическое распределение) при $N=1000$,
 $n=100$, $c=3$

q	0,9900	0,9500	0,9000	0,8300	0,5000	0,2000	0,1000	0,0500
Δ	0,0020	0,0070	0,0036	0,0092	0,0064	0,0029	0,0057	0,0015
D_q	9	15	18	23	36	53	64	73
Δ'_q	0,0023	0,0041	0,0130	0,0124	0,0160	0,0092	0,0013	0,0021

Если $\Delta < \Delta'$, то D_q соответствует Δ , если же $\Delta > \Delta'$, то D_q соответствует Δ' . Таким образом, при $q=0,5000$ в таблице 1.3.2 $D_q=36$, $H_{N,D}^{n,c} = H_{1000,36}^{100,3} = 0,5000 + \Delta = 0,5064$. Для $q=0,1000$, $\Delta > \Delta'$, поэтому $H_{1000,64}^{300,3} = 0,1 - \Delta' = 0,0987$.

Для небольших значений $d=0, 1, 2$ вычисление $H_{N,D}^{n,d}$ можно проводить по формуле (2.15). Например, для $d=1$ имеем

$$H_{N,D}^{n,1} = \left(1 + \frac{nD}{N-D-n+1}\right) h_{N,D}^{n,0}. \quad (3.27)$$

Вычисление $h_{N,D}^{n,0}$ проводится по формуле (2.15) с использованием таблицы XI приложения.

Итак, мы сравнительно подробно ознакомились с гипергеометрическим распределением. Оперативная характеристика одноступенчатого плана объема n с приемочным числом c при контроле совокупности $\mathfrak{B} \asymp [N, D]$ равна $H_{N,D}^{n,c}$, т. е. непосредственно выражается через гипергеометрическое распределение. Гипергеометрические вероятности входят также в выражения оперативных характеристик многоступенчатых планов контроля. Ниже, в § 2.2, мы кратко рассмотрим некоторые типы

планов контроля, характеристики которых также связаны с гипергеометрическим распределением. Там же будет дано определение отрицательного гипергеометрического распределения.

§ 1.4. Последующие оценки для одноступенчатых планов

Последующие оценки с учетом результатов разбраковки. Для правильной организации выборочного контроля поступающих на контроль совокупностей необходимо уметь оценивать эффективность различных систем правил выборочного контроля. Наиболее объективные оценки эффективности контроля можно получить в результате обработки накопленных данных в ходе контроля.

Эффективность системы контроля существенным образом зависит от принимаемых решений. При контроле с одноступенчатыми планами обычно принимается одно из следующих трех решений:

$\mathcal{D}_1$ — отвергнуть оставшуюся непроверенной часть совокупности без контроля,

$\mathcal{D}_2$ — провести сплошную проверку всех объектов, не попавших в выборку,

$\mathcal{D}_3$ — принять оставшуюся непроверенной часть совокупности без контроля.

При контроле качества промышленной продукции возможны и другие решения, такие, как снижение сортности, изъятие отдельных частей продукции для последующей переработки и т. д.

Если d — число дефектных объектов, обнаруженных в выборке объема n , а c — приемочное число, то при $d > c$ принимается решение забраковать совокупность (т. е. принимается решение $\mathcal{D}_1$ или $\mathcal{D}_2$), а при $d \leq c$ — решение принять совокупность (т. е. принимается решение $\mathcal{D}_2$ или $\mathcal{D}_3$). Одноступенчатые планы контроля с параметрами (n, c) будем обозначать $(n, c)_{i, j}$, если забракование соответствует решению $\mathcal{D}_i$, а принятие — решению $\mathcal{D}_j$.

Возможны следующие типы планов контроля $(n, c)_{12}$, $(n, c)_{13}$, $(n, c)_{23}$. Например, при плане $(120, 1)_{23}$ из совокупности берут случайную выборку без возвращения объема 120. Если среди отобранных 120 объектов не

более одного дефектного объекта, то совокупность принимают без дальнейшего контроля. Если же обнаружено два или более дефектных объектов, то проводят сплошную проверку всех объектов, не попавших в выборку. Планы $(n, c)_{12}$ используют в тех случаях, когда контроль не разрушительный, стоимость его велика, требования к качеству высокие. Планы типа $(n, c)_{13}$ используют главным образом при разрушительном контроле или в тех случаях, когда стоимость контроля очень высокая. Наконец планы $(n, c)_{23}$ применяют как при разрушительном, так и при неразрушительном контроле, когда представляет интерес получение дополнительной информации о качестве продукции, а стоимость контроля невелика. При неразрушительном контроле выявленные дефектные объекты могут заменяться годными.

Более широкий класс образуют **одноступенчатые планы контроля с двумя уровнями**. Эти планы определяются объемом n случайной выборки без возвращения и двумя целыми числами c_1, c_2 ($c_1 < c_2$). Если в выборке обнаружено d дефектных объектов и $d \leq c_1$, то принимается решение $\mathcal{D}_3$, если $c_1 < d < c_2$, то решение $\mathcal{D}_2$, и, наконец, при $d \geq c_2$ принимается решение $\mathcal{D}_1$. Кратко такие планы обозначаем (n, c_1, c_2) . Одноступенчатые планы контроля с двумя уровнями могут оказаться экономически целесообразными, если контроль не разрушительный, но дорогой, и сравнительно часто встречаются совокупности, сильно засоренные дефектными объектами.

В соответствии с принимаемыми решениями контролируемая совокупность объема N , содержащая D дефектных объектов, разбивается на три части, причем

$$N = N' + N'' + N''', \quad D = D' + D'' + D''', \quad (4.1)$$

где N' — число объектов, отвергнутых без контроля, среди которых содержится D' дефектных объектов, N'' — число проконтролированных объектов, D'' — число дефектных объектов, обнаруженных среди N'' , N''' — число объектов, принятых без контроля, D''' — число дефектных объектов среди принятых без контроля. При одноступенчатых планах известны значения N', N'', N''', D'' , тогда как значения D, D', D''' не известны. Ниже в этом параграфе даны оценки для D, D', D''' , которые,

конечно, различны для различных одноступенчатых планов контроля. При этом мы опираемся на результаты работ Колмогорова [4], Сираждинова [1], [5], Сираждинова и Эйдельманта [1], Стандарты статистического контроля (проект РТМ) [1]. Вопрос об оценке априорных вероятностей для возможных значений D рассмотрен в главе 3. Полезно заметить, что при учете результатов сплошной разбраковки, когда используется решение $\mathcal{D}_2$, нужно быть уверенным, что сплошной контроль проводится по тем же правилам, что и контроль объектов в выборке. И, наконец, еще раз подчеркнем, что классификация объектов на годные и дефектные предполагается безошибочной.

При плане $(n, c)_{23}$ число обнаруженных дефектных объектов равно

$$D'' = \begin{cases} d, & d \leq c, \\ D, & d > c. \end{cases} \quad (4.2)$$

Несмещенная оценка существует для любой функции $f(D)$, $D = 0, \dots, N$. Найдем эту оценку. В соответствии с (4.2) уравнение для несмещенной оценки $\varphi_f(D'')$ для $f(D)$ имеет вид

$$\sum_{d \leq c} \varphi_f(d) h_{N,D}^{n,d} + \varphi_f(D) \bar{H}_{N,D}^{n,c} = f(D). \quad (4.3)$$

Подставляя в (4.3) значения $D = 0, \dots, c$ и учитывая, что при этом $\bar{H}_{N,D}^{n,c} = 0$, находим $c + 1$ уравнение для $\varphi_f(d)$, $d = 0, \dots, c$. Эти уравнения имеют вид

$$\sum_{d \leq c} \varphi_f(d) h_{N,l}^{n,d} = f(l), \quad l = 0, \dots, c, \quad (4.4)$$

причем система (4.4) такая, что значения $\varphi_f(d)$ находятся последовательно одно за другим. После того как эти значения $\varphi_f(d)$ найдены, из (4.3) получаем

$$\varphi_f(D) = \frac{f(D) - \sum_{d \leq c} \varphi_f(d) h_{N,D}^{n,d}}{\bar{H}_{N,D}^{n,c}} \quad (4.5)$$

— значение оценки для $D'' = D$. В частности, при $f(D) = D$ получаем, что φ_f есть несмещенная оценка для предъяв-

ленного брака. Если же $f(D) = \sum_{d \leq c} (D - d) h_{N,D}^{n,d}$, то φ_f есть несмещенная оценка среднего пропущенного брака D''' .

Пуассоновское приближение. В качестве второго примера рассмотрим план $(n, c)_{13}$. Здесь $N'' = n$, $D'' = d$, D' и D''' — неизвестные случайные величины, для которых найдем несмещенные оценки. Используя оценку для D' , можно оценить число отвергнутых годных объектов, равное $N' - D'$.

Так как для плана $(n, c)_{13}$ имеем

$$D''' = \begin{cases} D - d, & d \leq c, \\ 0, & d > c, \end{cases} \quad (4.6)$$

то $ED''' = \sum_{0 \leq d \leq c} (D - d) h_{N,D}^{n,d}$. Поэтому уравнение для несмещенной оценки φ''' для D''' имеет вид

$$\sum_{d=0}^n \varphi'''(d) h_{N,D}^{n,d} = \sum_{d=0}^c (D - d) h_{N,D}^{n,d}. \quad (4.7)$$

Уравнение (4.7) противоречиво. Это следует из того факта, что $h_{N,D}^{n,d}$ является многочленом n -й степени относительно D . Следовательно, левая часть (4.7) является многочленом n -й степени по D , тогда как правая часть есть многочлен $n + 1$ -й степени. Таким образом, несмещенной оценки для D''' нет!

Найдем несмещенную оценку для D''' при контроле по плану $(n, c)_{13}$ в том асимптотическом случае, когда гипергеометрические вероятности $h_{N,D}^{n,d} \approx e^{-\lambda} \lambda^d / d! = p_\lambda(d)$, $\lambda = nD/N$. Используя соотношение $\lambda p_\lambda(d) = (d+1)p_\lambda(d+1)$, уравнение (4.7) переписываем в виде

$$\begin{aligned} \sum_{d=0}^n \varphi'''(d) p_\lambda(d) &= D \sum_{d=0}^c p_\lambda(d) - \sum_{d=0}^c d p_\lambda(d) = \\ &= \frac{N}{n} \sum_{d=0}^c \lambda p_\lambda(d) - \sum_{d=0}^c d p_\lambda(d) = \frac{N}{n} \sum_{d=1}^{c+1} d p_\lambda(d) - \sum_{d=1}^c d p_\lambda(d). \end{aligned} \quad (4.8)$$

Уравнение (4.8) мы рассматриваем как тождество по λ . Поэтому, приравнявая коэффициенты при $p_\lambda(d)$ слева и справа, находим выражение для несмещенной оценки D :

$$\varphi''' = \begin{cases} \frac{N}{n} d - d, & d \leq c, \\ \frac{N}{n} d, & d = c + 1, \\ 0, & d > c + 1. \end{cases} \quad (4.9)$$

Несмещенная оценка D для плана $(n, c)_{13}$ равна

$$\varphi = (N/n) d. \quad (4.10)$$

Полезно иметь оценки дисперсий оценок φ и φ''' , определяемых уравнениями (4.10) и (4.9). Если считать d случайной величиной, имеющей пуассоновское распределение с параметром $\lambda = nD/N$, то дисперсия оценки φ равна

$$\begin{aligned} V\varphi &= E\varphi^2 - (E\varphi)^2 = \\ &= (N/n)^2 \left\{ \sum_{d=0}^c d^2 p_\lambda(d) - \lambda^2 \right\} = (N/n)^2 \lambda. \end{aligned} \quad (4.11)$$

Так как несмещенная оценка для D имеет вид (4.10), то несмещенная оценка для λ равна d . Следовательно, согласно (4.11) находим, что оценка φ_2 дисперсии оценки (4.10) имеет вид

$$\varphi_2 = (N/n)^2 d. \quad (4.12)$$

Так как φ''' является несмещенной оценкой для случайной величины D''' , то представляет интерес получение несмещенной оценки для дисперсии разности $D''' - \varphi'''$. Предварительно заметим, что

$$(D''' - \varphi''')^2 = \begin{cases} (N/n)^2 (\lambda - d)^2, & d \leq c, \\ (N/n)^2 (c + 1)^2, & d = c + 1, \\ 0, & d > c + 1. \end{cases} \quad (4.13)$$

Учитывая (4.13), находим, что

$$\begin{aligned}
 E(D''' - \varphi''')^2 &= \\
 &= (N/n)^2 \left\{ \sum_{d=0}^c (\lambda - d)^2 p_\lambda(d) + (c+1)^2 p_\lambda(c+1) \right\} = \\
 &= (N/n)^2 \left\{ \sum_{d=0}^{c+1} d^2 p_\lambda(d) - 2\lambda \sum_{d=0}^c d p_\lambda(d) + \lambda^2 \sum_{d=0}^c p_\lambda(d) \right\} = \\
 &= (N/n)^2 \left\{ \sum_{d=1}^{c+1} d p_\lambda(d) + (c+1)(c+2) p_\lambda(c+2) \right\}. \quad (4.14)
 \end{aligned}$$

Уравнение, которому должна удовлетворять несмещенная оценка φ_2''' для $E(D''' - \varphi''')^2$, в соответствии с (4.14) имеет вид

$$\begin{aligned}
 \sum_{d=0}^c \varphi_2'''(d) p_\lambda(d) &= \\
 &= (N/n)^2 \left\{ \sum_{d=1}^{c+1} d p_\lambda(d) + (c+1)(c+2) p_\lambda(c+2) \right\}. \quad (4.15)
 \end{aligned}$$

Приравнявая в (4.15) коэффициенты при одинаковых функциях $p_\lambda(d)$, находим, что

$$\varphi_2''' = \begin{cases} (N/n)^2 d, & d \leq c+1, \\ (N/n)^2 (c+1)(c+2), & d = c+2, \\ 0, & d > c+2, \end{cases}$$

есть несмещенная оценка для дисперсии разности $D''' - \varphi'''$.

Последующие оценки для планов с двумя уровнями. Рассмотрим теперь одноступенчатые планы с двумя уровнями с параметрами (n, c_1, c_2) . Последующие оценки для таких планов были найдены Сираждиновым и Эйдельмантом [1]. Планы (n, c_1, c_2) можно рассматривать как обобщение планов типа $(n, c)_{i,j}$. Действительно, $(n, c)_{12} = (n, c_1, c_2)$, если $c_1 = -1$, $c_2 = c+1$; $(n, c)_{13} = (n, c_1, c_2)$, $c_1 = c$, $c_2 = c+1$; $(n, c)_{23} = (n, c_1, c_2)$, если $c_1 = c$, $c_2 = n+1$. В таблице 1.4.1 приведены сводные данные о значениях N' , N'' , N''' и D' , D'' , D''' . Наша очередная задача состоит в выводе несмещенных оценок φ' и φ''' для случайных величин D' , D''' .

Таблица 1.4.1

Число объектов, которые по результатам контроля		
отвергаются без дальнейшего контроля	подвергаются контролю	принимаются без контроля
$N' = \begin{cases} 0, & d < c_2, \\ N-n, & d \geq c_2. \end{cases}$	$N'' = \begin{cases} n, & d \leq c_1 \text{ или } d \geq c_2, \\ N, & c_1 < d < c_2. \end{cases}$	$N''' = \begin{cases} N-n, & d \leq c_1, \\ 0, & d > c_1. \end{cases}$
$D' = \begin{cases} 0, & d < c_2, \\ D-d, & d \geq c_2. \end{cases}$	$D'' = \begin{cases} d, & d \leq c_1 \text{ или } d \geq c_2, \\ D, & c_1 < d < c_2. \end{cases}$	$D''' = \begin{cases} D-d, & d \leq c_1, \\ 0, & d > c_1. \end{cases}$

Несмещенную оценку φ для D можно затем получить из соотношения

$$\varphi = \varphi' + D'' + \varphi''' \quad (4.16)$$

Оценки, удовлетворяющие (4.16), называются **согласованными**. Так как вывод формул несмещенных оценок φ' и φ''' несложен, то мы проведем его только для φ' .

D' — случайная величина, математическое ожидание которой равно $\sum_{c_2 \leq d \leq n} (D-d) h_{N,D}^{n,d}$. Поэтому уравнение для искомой несмещенной оценки имеет вид

$$E \varphi' = \sum_{d=c_2}^n (D-d) h_{N,D}^{n,d}. \quad (4.17)$$

Беря значение D'' из таблицы 1.4.1, перепишем (4.17) в виде

$$\begin{aligned} \sum_{d=0}^{c_1} \varphi'(d) h_{N,D}^{n,d} + \varphi'(D) \sum_{d=c_1+1}^{c_2-1} h_{N,D}^{n,d} + \sum_{d=c_2}^n \varphi'(d) h_{N,D}^{n,d} = \\ = \sum_{d=c_2}^n (D-d) h_{N,D}^{n,d}. \end{aligned} \quad (4.18)$$

Пусть $c_2 - c_1 > 1$. Решение φ' , удовлетворяющее (4.18), можно получить, приравнявая коэффициенты при одинаковых степенях D . Еще проще получить решение, последовательно полагая $D = 0, \dots, c_1$. Так как $h_{N,D}^{n,d} = 0$ при

$d > c$, то из (4.18) находим, что $\varphi'(d) = 0$, $d = 0, \dots, c$. Полагая затем последовательно $D = N, N-1, \dots, N-n-c_2+1$, найдем, что $\varphi'(N) = \left(\frac{N-n}{n}\right)n$, $\varphi'(N-1) = \frac{N-n}{n}(n-1)$ и т. д., т. е. что $\varphi'(d) = \frac{N-n}{n}d$, если $d \geq c_2$. Подставляя полученные значения $\varphi'(d)$ в (4.18), приходим к уравнению

$$\varphi'(D) \sum_{d=c_1+1}^{c_2-1} h_{N,D}^{n,d} + \frac{N-n}{n} \sum_{d=c_1}^n d h_{N,D}^{n,d} = \sum_{d=c_2}^n (D-d) h_{N,D}^{n,d}. \quad (4.19)$$

Разрешая (4.19) относительно $\varphi'(D)$, найдем, что

$$\varphi'(D) = \frac{D \bar{H}_{N,D}^{n,c_2-1} - \frac{N}{n} \sum_{d=c_2}^n d h_{N,D}^{n,d}}{\bar{H}_{N,D}^{n,c_1} - \bar{H}_{N,D}^{n,c_2-1}}. \quad (4.20)$$

Точно так же находится несмещенная оценка φ''' для

Таблица 1.4.2

Несмещенные оценки для плана (n, c_1, c_2)

d	$d \leq c_1$	$c_1 < d < c_2$	$d \geq c_2$
D''	d	D	d
φ'	0	$\frac{D \bar{H}_{N,D}^{n,c_2-1} - \frac{N}{n} \sum_{d=c_2}^n d h_{N,D}^{n,d}}{\bar{H}_{N,D}^{n,c_1} - \bar{H}_{N,D}^{n,c_2-1}}$	$\frac{N-n}{n} d$
φ'''	$\frac{N-n}{n} d$	$\frac{D H_{N,D}^{n,c_1} - \frac{N}{n} \sum_{d=0}^{c_1} d h_{N,D}^{n,d}}{\bar{H}_{N,D}^{n,c_1} - \bar{H}_{N,D}^{n,c_2-1}}$	0
φ	$\frac{N}{n} d$	$\frac{\frac{N}{n} \sum_{d=c_1+1}^{c_2-1} d h_{N,D}^{n,d}}{\bar{H}_{N,D}^{n,c_1} - \bar{H}_{N,D}^{n,c_2-1}}$	$\frac{N}{n} d$

D''' . Получаемые формулы для несмещенных оценок даны в виде сводной таблицы 1.4.2.

Контроль с проверкой дополнительного объекта. В работе Сираждинова и Эйдельманта [1] рассматривается контроль с проверкой $(n+1)$ -го объекта. Правило проверки $(n+1)$ -го объекта при плане (n, c_1, c_2) состоит в следующем. По результатам контроля n объектов, составляющих выборку, решение о забраковании или принятии совокупности производят в соответствии с планом (n, c_1, c_2) . Однако если окажется, что обнаружено d дефектных объектов и $d \leq c_1$ или $d \geq c_2$, то случайным образом отбирают для контроля $(n+1)$ -й объект. Результаты контроля не влияют на принятое ранее решение. Пусть $\varepsilon_{n+1} = 1$, когда $(n+1)$ -й объект дефектный, и $\varepsilon_{n+1} = 0$, когда он годный.

Для этого правила имеем

$$\begin{aligned} N' &= \begin{cases} 0, & d < c_2, \\ N - n - 1, & d \geq c_2, \end{cases} \\ N'' &= \begin{cases} n + 1, & d \leq c_1 \text{ или } d \geq c_2, \\ N, & c_1 < d < c_2, \end{cases} \\ N''' &= \begin{cases} N - n - 1, & d \leq c_1, \\ 0, & d > c_1, \end{cases} \\ D'' &= \begin{cases} d + \varepsilon_{n+1}, & d \leq c_1 \text{ или } d \geq c_2, \\ D, & c_1 < d < c_2. \end{cases} \end{aligned} \quad (4.21)$$

Заметим, что $D'' = d + \varepsilon_{n+1} = c_1 + 1$ только тогда, когда $d(n) = c_1$, $\varepsilon_{n+1} = 1$; соответственно, $D'' = d + \varepsilon_{n+1} = c_2$ только тогда, когда $d = c_2$, $\varepsilon_{n+1} = 0$. Общее число дефектных объектов, обнаруженных в выборке, обозначаем через $d' = d + \varepsilon_{n+1}$. При выводе формул для несмещенных оценок полезно учитывать тождества

$$\sum_{d=0}^c (D-d) h_{N,D}^{n,d} = \frac{N-n}{n+1} \sum_{d'=0}^{c+1} d' h_{N,D}^{n+1,d'}, \quad (4.22)$$

и

$$\sum_{d=c}^n (D-d) h_{N,D}^{n,d} = \frac{N-n}{n+1} \sum_{d'=c+1}^{n+1} d' h_{N,D}^{n+1,d'}. \quad (4.23)$$

Эти тождества являются прямым следствием легко проверяемого соотношения

$$(D-d)h_{N,D}^{n,d} = \frac{N-n}{n+1} (d+1) h_{N,D}^{n+1,d+1}. \quad (4.24)$$

Ограничимся выводом несмещенной оценки для числа дефектных объектов, принятых без контроля, которое при плане (n, c_1, c_2) с проверкой $(n+1)$ -го объекта равно

$$D_{\pi}''' = \begin{cases} D - (d + \varepsilon_{n+1}), & d \leq c_1, \\ 0, & d > c_1. \end{cases} \quad (4.25)$$

Согласно (4.25) и (4.22) находим, что

$$\begin{aligned} ED_{\pi}''' &= \sum_{d=0}^{c_1} (D-d) h_{N,D}^{n,d} h_{N-n, D-d}^{1,0} + \\ &+ \sum_{d=0}^{c_1} (D-d-1) h_{N,D}^{n,d} h_{N-n, D-d}^{1,1} = \\ &= \frac{N-(n+1)}{N-n} \sum_{d=0}^{c_1} (D-d) h_{N,D}^{n,d} = \frac{N-(n+1)}{n+1} \sum_{d'=0}^{c_1+1} d' h_{N,D}^{n+1,d'}. \end{aligned} \quad (4.26)$$

При правиле проверки $(n+1)$ -го объекта уравнение несмещенной оценки φ_{π}''' для числа D_{π}''' , определяемого (4.25), с учетом (4.26) принимает вид

$$\begin{aligned} \sum_{d'=0}^{c_1} \varphi'''(d') h_{N,D}^{n+1,d'} + \varphi'''(c_1+1) h_{N,D}^{n,c_1} h_{N-n, D-c_1}^{1,1} + \\ + \varphi'''(D) \sum_{d=c_1+1}^{c_2-1} h_{N,D}^{n,d} + \varphi'''(c_2) h_{N,D}^{n,c_2} h_{N-n, D-d}^{1,0} + \\ + \sum_{d'=c_2+1}^{n+1} \varphi'''(d') h_{N,D}^{n+1,d'} = \frac{N-(n+1)}{n+1} \sum_{d'=0}^{c_1+1} d' h_{N,D}^{n+1,d'}. \end{aligned} \quad (4.27)$$

Согласно (4.24) $h_{N,D}^{n,c_1} h_{N-n, D-c_1}^{1,1} = \frac{c_1+1}{n+1} h_{N,D}^{n+1,c_1+1}$. Поэтому, приравнявая коэффициенты при $h_{N,D}^{n,d}$, находим, что

несмещенная оценка D_d''' имеет вид

$$\varphi_d''' = \begin{cases} \frac{N - (n + 1)}{n + 1} (d + \varepsilon_{n+1}), & d + \varepsilon_{n+1} \leq c_1, \\ N - (n + 1), & d = c_1, \varepsilon_{n+1} = 1, \\ 0, & d > c_1. \end{cases} \quad (4.28)$$

Аналогичным способом можно найти выражение для несмещенной оценки D_d' — числа дефектных объектов в части совокупности, отвергнутой без контроля. Оцениваемая величина D_d' имеет вид

$$D_d' = \begin{cases} D - (d + \varepsilon_{n+1}), & d \geq c_2, \\ 0, & d < c_2. \end{cases} \quad (4.29)$$

Окончательный вид несмещенной оценки φ_d' для D' следующий:

$$\varphi_d' = \begin{cases} \frac{N - (n + 1)}{n + 1} (d + \varepsilon_{n+1}), & d + \varepsilon_{n+1} \geq c_2 + 1, \\ 0, & d < c_2 \text{ или } d = c_2, \varepsilon_{n+1} = 0. \end{cases} \quad (4.30)$$

Последующие оценки для планов с одним уровнем. Исходя из полученных результатов, можно дать набор формул для несмещенных оценок при одноступенчатых планах.

П л а н $(n, c)_{12}$. Одноступенчатый план $(n, c)_{12}$ соответствует плану с двумя уровнями (n, c_1, c_2) , когда $c_1 = -1$, $c_2 = c + 1$. Из таблицы 1.4.2 находим, что несмещенная оценка для D' имеет вид

$$\varphi' = \begin{cases} \frac{D\bar{H}_{N,D}^{n,c} - \frac{N}{n} \sum_{d=c+1}^n dh_{N,D}^{n,d}}{H_{N,D}^{n,c}}, & d \leq c, \\ \frac{N-n}{n} d, & d \geq c + 1. \end{cases} \quad (4.31)$$

Вопрос о вычислении несмещенной оценки D''' не возникает, так как здесь $D''' \equiv 0$. Если проводится

дополнительная проверка $(n+1)$ -го объекта, то несмещенная оценка D'_d равна

$$\varphi'_d = \begin{cases} 0, & d < c+1 \text{ или } d = c+1, \quad e_{n+1} = 0, \\ \frac{N - (n+1)}{n+1} (d + e_{n+1}), & d + e_{n+1} \geq c+2. \end{cases} \quad (4.32)$$

Легко проверить, что несмещенной оценкой для D' при контроле по плану $(n, c)_{12}$ без проверки $(n+1)$ -го объекта будет

$$\varphi'_d = \begin{cases} N(d/n) - D, & d \leq c, \\ \frac{N-n}{n} d, & d > c. \end{cases} \quad (4.33)$$

В работе Сираждинова и Эйдельманта [1] указывается, что оценка (4.33) имеет меньшую точность, чем оценка (4.31). Однако вычисление оценки (4.33) проводится существенно проще, чем оценки (4.31).

П л а н $(n, c)_{23}$. Для планов этого типа $D' = 0$, поэтому представляет интерес получение только несмещенной оценки для D''' . Замечая, что план $(n, c)_{23}$ эквивалентен плану (n, c_1, c_2) , $c_1 = c$, $c_2 = n+1$, из таблицы 1.4.2 находим, что несмещенная оценка для D''' равна

$$\varphi''' = \begin{cases} \frac{N-n}{N} d, & d \leq c, \\ \frac{DH_{N,D}^{n,c} - \frac{N}{n} \sum_{d=0}^c dh_{N,D}^{n,d}}{\bar{H}_{N,D}^{n,c}}, & d > c. \end{cases} \quad (4.34)$$

При использовании правила проверки $(n+1)$ -го объекта получим, что несмещенная оценка для D'''_d имеет вид

$$\varphi'''_d = \begin{cases} \frac{N - (n+1)}{n+1} (d + e_{n+1}), & d + e_{n+1} \leq c, \\ N - (n+1), & d = c, \quad e_{n+1} = 1, \\ 0, & d > c. \end{cases} \quad (4.35)$$

Если правило проверки $(n+1)$ -го объекта не используется, то менее точная, чем (4.34), но зато простая

несмещенная оценка D''' имеет вид

$$\Phi''' = \begin{cases} \frac{N-n}{n} d, & d \leq c, \\ \frac{N}{n} d - D, & d > c. \end{cases} \quad (4.36)$$

П л а н $(n, c)_{13}$. При плане $(n, c)_{13}$ сплошная разбраковка не производится. Несмещенных оценок для D' , D''' не существует. Однако если применить правило проверки $(n+1)$ -го объекта, то соответствующие несмещенные оценки Φ'_d и Φ'''_d для D'_d и D'''_d существуют и имеют вид ($c_1 = c_2 - 1 = c$):

$$\Phi'_d = \begin{cases} \frac{N-(n+1)}{n+1} (d + \varepsilon_{n+1}), & d + \varepsilon_{n+1} > c + 1, \\ 0, & d \leq c \text{ или } d = c + 1, \varepsilon_{n+1} = 0, \end{cases} \quad (4.37)$$

$$\Phi'''_d = \begin{cases} \frac{N-(n+1)}{n+1} (d + \varepsilon_{n+1}), & d + \varepsilon_{n+1} \leq c, \\ N - (n + 1), & d = c, \varepsilon_{n+1} = 1, \\ 0, & d > c. \end{cases} \quad (4.38)$$

На основе последующих оценок можно выяснить фактическую эффективность правил приемочного контроля последовательностей совокупностей $\mathfrak{P}_i \propto [N, D_i], i=1, \dots, m$. Для оценки суммарного предъявленного брака $\sum_{1 \leq i \leq m} D_i$ используем

$$\sum_{1 \leq i \leq m} \Phi(D'_i) \quad (\Phi(D'') = \Phi'(D'') + D'' + \Phi'''(D'')),$$

для оценки принятого брака $\sum_{1 \leq i \leq m} D''_i$ используем

$\sum_{1 \leq i \leq m} \Phi'''(D'_i)$. Как видно из приведенного выше материала, существуют различные несмещенные оценки для D , D' , D''' . Показателем эффективности правил приемочного контроля может служить разность оценок суммарного предъявленного и пропущенного брака, т. е. оценка суммарного числа дефектных объектов, изъятых при контроле. Для планов контроля с решениями о сплошной разбраковке можно применять не-

смещенные оценки указанных характеристик только на основе данных выборочного обследования либо с учетом данных о сплошной разбраковке. Одновременное вычисление двух типов оценок может быть использовано для проверки адекватности операции выборочного контроля и контроля со сплошной проверкой. Если такое сопоставление выявляет существенное различие в оценке суммарного предъявленного или пропущенного брака, то следует искать причину несоответствий. Выявленное несоответствие может быть следствием нарушений правил контроля.

Выше мы ограничились только выводом несмещенных оценок. Естественно возникает вопрос о точности этих оценок. В § 1.2, посвященном планам бездефектного контроля, были получены оценки дисперсий оценок. Аналогично можно найти оценки дисперсии несмещенных оценок, полученных выше. Заметим, что для получения оценок дисперсий несмещенных оценок при правиле проверки $(n + 1)$ -го объекта нужно проверить еще $(n + 2)$ -й. При этом предполагаем, что результат проверки $(n + 2)$ -го объекта также не влияет на принятое ранее решение. Такие оценки легко найти, используя теорию получения несмещенных оценок для общих планов типа первого вхождения, изложенную ниже в § 3.1.

ГЛАВА 2

ПЛАНЫ КОНТРОЛЯ ПО АЛЬТЕРНАТИВНОМУ ПРИЗНАКУ И ИХ ХАРАКТЕРИСТИКИ

§ 2.1. Вспомогательные сведения

Об основных понятиях теории вероятностей. В этом параграфе дается краткий перечень основных понятий теории вероятностей и математической статистики, которые используются в дальнейшем в книге непосредственно либо косвенно и могут помочь более полному пониманию рассматриваемых проблем. К настоящему времени имеются хорошо известные книги по теории вероятностей, рассчитанные на читателей, имеющих различную математическую подготовку, например, Мостеллер, Рурке, Томас [1], Феллер, т. 1, т. 2 [1], [2], Гнеденко [1], Боровков [1] и др. Несколько сложнее положение с математической статистикой. Уже написанные обстоятельные монографии по математической статистике главным образом рассчитаны на читателей-математиков, например, книги Линника [1], Рао [1], Крамера [2], Лемана [1], Закса [1] и др.

В большинстве математических моделей, рассматриваемых в книге, используются дискретные пространства исходов испытаний, результатов наблюдений, эксперимента и т. п., содержащие конечное или счетное множество результатов наблюдений. Пусть Π — план, по которому получают исходы наблюдения x , возможные значения которых образуют некоторое множество $X_\Pi = \{x\}$. Рассмотрим сначала такие планы, при которых X_Π содержит конечное или счетное множество исходов. Элементы X_Π можно снабдить порядковыми номерами $i = 1, 2, \dots$. В итоге мы получим $X_\Pi = \{x_1, x_2, \dots, x_N\}$, $N \leq \infty$. Говорят, что в пространстве

X_{Π} задана **вероятностная мера** P , если каждому исходу x_i сопоставлено неотрицательное число $p(x_i) = p_i \geq 0$, $\sum_{1 \leq i \leq N} p_i = 1$. Любое подмножество исходов

$A = \{x_{i_1}, \dots, x_{i_k}\}$ называется **событием**. **Вероятностью** $P(A)$ **события** A называется сумма вероятностей составляющих A исходов, т. е. $P(A) = \sum_{x \in A} p_i$. Пустому мно-

жеству, не содержащему ни одного исхода, сопоставляется вероятность, равная нулю, $P\{\emptyset\} = 0$. Множество исходов и заданная на нем вероятностная мера P образуют **вероятностное пространство** $[X_{\Pi}, P]$.

Например, если все исходы x имеют вид $x = (\varepsilon_1, \dots, \varepsilon_n)$, где $\varepsilon_i = 0$ или 1, то множество возможных исходов X_{Π} есть подмножество узлов целочисленной решетки в n -мерном пространстве. Общее число возможных исходов равно $N = 2^n$. Пусть $\delta_n = \delta_n(x) = \sum_{1 \leq i \leq n} \varepsilon_i$, $q > 0$, $\bar{q} = 1 - q$.

Тогда сумма неотрицательных чисел $P(x) = q^{\delta_n} \bar{q}^{n-\delta_n}$ равна единице, так как $\sum_{x \in X_{\Pi}} q^{\delta_n(x)} \bar{q}^{n-\delta_n(x)} = (q + \bar{q})^n = 1$.

Вероятностное пространство $[X_{\Pi}, P]$, у которого исходам $x = (\varepsilon_1, \dots, \varepsilon_n)$ сопоставляются вероятности $P(x) = q^{\delta_n} \bar{q}^{n-\delta_n}$, соответствует так называемой **схеме испытаний Бернулли**.

Функция исходов испытаний $f = f(x)$, принимающая вещественные значения, называется **случайной величиной**. Так как исходы можно занумеровать, то случайная величина, определяемая $f(x)$, также принимает не более счетного числа значений.

Атомом $A(f, f_0)$ случайной величины, соответствующим значению (уровню) f_0 , называется такое множество исходов, для которых $f(x) = f_0$. Другими словами, атом, соответствующий значению (уровню) f_0 , определяется соотношением $A(f, f_0) = \{x: f(x) = f_0\}$. За двоеточием в фигурных скобках стоит условие, которому должны удовлетворять все x , составляющие определяемое множество.

Для любых событий A, B совокупность исходов, входящих одновременно и в A и в B , обозначается AB , т. е. $AB = \{x: x \in A, x \in B\}$. Операцию, сопостав-

ляющую двум событиям A, B событие AB , называют **пересечением** событий A и B . Если $AB = \emptyset$, т. е. пустое множество, то A, B называются **непересекающимися**. **Разностью** $A \setminus B$ называется подмножество всех точек x , принадлежащих A и не принадлежащих B .

Дополнением A называется множество $\bar{A} = X_{\Pi} \setminus A$.

Объединение событий A, B есть событие $A \cup B = \{x: x \text{ принадлежит } A \text{ либо } B, \text{ либо одновременно } A \text{ и } B\}$. Символ $B \subseteq A$ означает, что B является частью A .

Для набора событий $A_1, \dots, A_n$ пересечение $A_1 \dots A_n$ определяется как результат последовательного выполнения операции пересечения $A_1 A_2$, затем $(A_1 A_2) A_3$ и т. д. $(A_1 \dots A_{n-1}) A_n$. Конечный результат не зависит от порядка множеств $A_1, \dots, A_n$. Аналогично определяется объединение событий $A_1 \cup \dots \cup A_n$ и т. д. Для системы взаимно непересекающихся множеств $\{A_k\}$ имеем $P\left(\bigcup_k A_k\right) = \sum_k P(A_k)$.

События $A_1, \dots, A_n$ называются **взаимно независимыми**, если для любого подмножества индексов $i_1, \dots, i_k$, $i_j = 1, \dots, n$, $j = 1, \dots, k$, $P(A_{i_1} \dots A_{i_k}) = P(A_{i_1}) \cdot \dots \cdot P(A_{i_k})$. **Случайные величины** $f_1, \dots, f_n$ называются **взаимно независимыми**, если для любых наборов уровней $f_{i_1,0} \dots f_{i_k,0}$ атомы $A(f_{i_1}, f_{i_1,0}), \dots, A(f_{i_k}, f_{i_k,0})$ являются взаимно независимыми событиями.

Например, в схеме испытаний Бернулли определим случайные величины $f_k(x) = e_k$, $f_l(x) = e_l$. $e_k, e_l = 0, 1$. Четыре атома $A(f_k, 0)$, $A(f_k, 1)$, $A(f_l, 0)$, $A(f_l, 1)$ образованы такими исходами $x = (e_1, \dots, e_n)$, у которых на k -м и l -м местах стоят нули или единицы. Легко проверить, что

$$P(A(f_k, 0)) = P(A(f_l, 0)) = \bar{q},$$

$$P(A(f_k, 1)) = P(A(f_l, 1)) = q,$$

$$P(A(f_k, i) A(f_l, j)) = P(A(f_k, i)) P(A(f_l, j)), \quad i, j = 0, 1,$$

т. е. случайные величины $f_k(x)$ и $f_l(x)$, равные значениям e_k и e_l исходов в «моменты» k и l соответственно, являются взаимно независимыми случайными величинами. Точно так же легко убедиться, что случайные

величины $f_1(x) = \varepsilon_1, \dots, f_n(x) = \varepsilon_n$ взаимно независимы в совокупности.

Пусть событие $A \subset X_{\Pi}$ имеет положительную вероятностную меру $P(A) > 0$. Тогда каждому событию B можно сопоставить число $P(B|A) = \sum_{x_i \in BA} p_i/P(A)$, кото-

рое называется **условной вероятностью** события B при условии события A . В частности, если $B = x_i$, т. е. состоит из одного элемента, то x_i сопоставляется вероятность $p_i/P(A)$, когда $x_i \in A$, и сопоставляется 0, когда $x_i \notin A$. Таким образом, в $A \subset X_{\Pi}$ возникает новая вероятностная мера, которая называется **условной вероятностной мерой** относительно события A . Если события A и B взаимно независимы, то $P(B|A) = P(B)$. В общем случае $P(B|A) = P(AB)/P(A)$ или $P(AB) = P(A)P(B|A)$.

Если в схеме испытаний Бернулли рассмотреть случайную величину $f(x) = \delta_n$, то атом $A(f, k) = \{x: \delta_n = k\}$ имеет вероятность $P(A(f, k)) = \binom{n}{k} q^k \bar{q}^{n-k} = b_n^k(q)$. Поэтому точкам $x \in A(f, k)$ для выбранной функции $f(x)$ сопоставляется условная вероятностная мера, равная

$$p_k(x) = \frac{q^k \bar{q}^{(n-k)}}{b_n^k(q)} = \binom{n}{k}^{-1}. \quad (1.1)$$

Заметим, что значения условной вероятностной меры (1.1), определенной на $A(f, k)$, не зависят от величины параметра q .

Математическим ожиданием (средним) случайной величины $f(x)$ называется

$$Ef = \sum_{x_i \in X_{\Pi}} f(x_i) p_i, \quad (1.2)$$

где $p_i = P(x_i)$.

В случае $N = \infty$ считаем, что сумма в (1.2) сходится абсолютно, т. е. $\sum_{x_i \in X_{\Pi}} |f(x_i)| p_i < \infty$. Если $\chi_A = \chi_A(x) = 1$, когда $x \in A$, $\chi_A(x) = 0$, когда $x \notin A$, то χ_A называется **индикатором множества A** . Из (1.2) следует, что $E\chi_A = P(A)$, для двух множеств A_1, A_2 $E\chi_{A_1}\chi_{A_2} = P(A_1A_2)$.

Если рассматривать математическое ожидание по условной вероятностной мере относительно события A , то мы приходим к определению **условного матема-**

математического ожидания случайной величины f относительно события A , полагая его равным

$$E(f|A) = \sum_{x_i \in A} \frac{f(x_i) p_i}{P(A)}. \quad (1.3)$$

Если в (1.3) в качестве событий A рассмотреть атомы $A(g, g_0)$ случайной величины $g(x)$, то соотношение

$$E(f|g = g_0) = \sum_{x_i \in A(g, g_0)} \frac{f(x_i) p_i}{P(A(g, g_0))}, \quad (1.4)$$

рассматриваемое как функция возможных значений g_0 , при $P\{A(g, g_0)\} > 0$ и $E(f|g = g_0) = 0$ при $P\{A(g, g_0)\} = 0$ называется **условным математическим ожиданием случайной величины f относительно случайной величины g** .

Представим $\sum_{x_i \in A(g, g_0)} f(x_i) p_i$ в виде

$$\begin{aligned} \sum_{x_i \in A(g, g_0)} f(x_i) p_i &= \sum_{f_0} \sum_{x_i \in G(f_0, g_0)} f_0 p_i = \\ &= \sum_{f_0} f_0 P(A(f, f_0) A(g, g_0)), \end{aligned}$$

где

$$\begin{aligned} G(f_0, g_0) &= \{x_i: x_i \in A(g, g_0), f(x_i) = f_0\} = \\ &= \{x_i: x \in A(g, g_0) A(f, f_0)\}. \end{aligned}$$

Тогда из (1.4) легко следует, что если f и g — взаимно независимые случайные величины, то $E(f|g = g_0) \equiv Ef$, т. е. условное математическое ожидание совпадает с безусловным.

Точно так же легко получаем, что если f и h — случайные величины и f постоянна на атомах $A(g, g_0)$, т. е. $f(x) \equiv f_{g_0}$ на $A(g, g_0)$, то

$$E(fh|g = g_0) = f_{g_0} E(h|g = g_0). \quad (1.5)$$

Очень часто оказывается полезной формула полного условного математического ожидания. Так как (1.4) есть функция значений g_0 , то $E(f|g = g_0)$ зависит от исхода испытаний x не непосредственно, а через значение $g_0 = g(x)$. Поэтому, рассматривая $E(f|g = g_0)$ как функцию условия $g_0 = g(x)$, мы можем считать $E(f|g = g_0)$ случайной величиной, для которой

используем запись $E(f|g)$. Покажем, что имеет место следующее тождество:

$$E(E(f|g)) = Ef, \quad (1.6)$$

т. е. математическое ожидание от условного математического ожидания f при условии g совпадает с математическим ожиданием f . Тождество (1.6) кратко будем называть **формулой полного условного математического ожидания**. Частным случаем (1.6) является формула полной вероятности. Для доказательства (1.6) воспользуемся определением математического ожидания. В соответствии с (1.3), (1.4) имеем

$$\begin{aligned} E(E(f|g)) &= \sum_{g_0} \left(\sum_{x_i \in A(g, g_0)} f(x_i) \frac{p_i}{P(A(g, g_0))} \right) P(A(g, g_0)) = \\ &= \sum_{g_0, x_i \in A(g, g_0)} f(x_i) p_i = \sum_{x_i \in X_\Pi} f(x_i) p_i = Ef, \end{aligned}$$

что и требовалось доказать. Здесь мы воспользовались тем фактом, что атомы $A(g, g_0)$ случайной величины g — взаимно непересекающиеся множества, образующие **разбиение** пространства исходов. Другими словами, пространство исходов X_Π представляется в виде объединения взаимно непересекающихся атомов $A(g, g_0)$, т. е. $X_\Pi = \bigcup_{g_0} A(g, g_0)$. Так как каждому

уровню g_0 соответствует один атом $A(g, g_0)$, вероятность которого равна $p(g_0) = P(A(g, g_0))$, то $E(f|g)$ мы можем рассматривать как случайную величину, определенную на пространстве, «точками» которого являются атомы $A(g, g_0)$ с вероятностями $P(A(g, g_0))$.

Пусть $A_1, \dots, A_m$, $m \leq N$, — взаимно непересекающиеся события, что кратко записывается $A_i A_j = \emptyset$, $i \neq j$, $A_i \subset X_\Pi$, $i, j = 1, \dots, m$ (индексы i, j принимают значения от 1 до m). Для события A определим индикаторную функцию $\chi_A(x)$ и функцию $\tilde{g}(x)$, равную g_i на множестве A_i , $g_i \neq g_j$, $i \neq j$, $\tilde{g}(x) = \sum_i g_i \chi_{A_i}(x)$. Для функций $f(x) = \chi_A(x)$, $g(x) = \tilde{g}(x)$ формула (1.6) принимает вид

$$\sum_{i=1}^m P(A|A_i) P(A_i) = P(A). \quad (1.7)$$

Тождество (1.7) называется *формулой полной вероятности*.

Существенно более сложной оказывается теория для тех пространств исходов X_{Π} , которые содержат несчетное множество элементов (исходов). Так будет, если X_{Π} является отрезком прямой, областью n -мерного евклидова пространства или функциональным пространством, когда исходы есть функции некоторого аргумента. Здесь уже нельзя каждое подмножество X_{Π} называть событием. Обычно поступают следующим образом. Некоторые типы подмножеств X_{Π} , не более чем в счетном количестве, объявляются *событиями*. Пусть эти подмножества $A_1, A_2, \dots$. *Событиями* объявляются также такие подмножества, которые можно получить из исходного набора $A_1, A_2, \dots$ с помощью теоретико-множественных операций объединения, пересечения и дополнения, взятых не более чем счетное число раз. Операция дополнения сопоставляет множеству A множество $\bar{A}$ всех элементов, не входящих в A . В итоге получается класс множеств $\mathfrak{B}_{X_{\Pi}}$, называемый *σ -алгеброй событий* $\mathfrak{B}_{X_{\Pi}}$. Пара $[X_{\Pi}, \mathfrak{B}_{X_{\Pi}}]$ называется *измеримым пространством*. Все подмножества σ -алгебры $\mathfrak{B}_{X_{\Pi}}$ называются *событиями*. Если A — событие ($A \in \mathfrak{B}_{X_{\Pi}}$), то ему сопоставляется неотрицательное число $P(A) \geq 0$, $P(X_{\Pi}) = 1$, $P(\emptyset) = 0$. Если $P(A)$ есть функция множеств $A \in \mathfrak{B}_{X_{\Pi}}$ такая, что для любой не более чем счетной последовательности непересекающихся событий $\{A_i\}$

$$P\left(\bigcup_i A_i\right) = \sum_i P(A_i),$$

то эта функция P называется *вероятностной мерой*. Число $P(A)$ называется *вероятностью события* A . Случайная величина f уже не любая функция, а только такая, для которой множества $\{x: f(x) \leq f_0\} \in \mathfrak{B}_{X_{\Pi}}$ при любом f_0 . Математическое ожидание Ef определяется как интеграл Лебега по мере P . Теоретико-множественные основы теории вероятностей были даны Колмогоровым [1]. Мы не будем далее проводить построение теории в этом более общем случае. Читатель может ознакомиться с ним по любому из учебников теории вероятностей, например, Неве [1], Лозв [1] и др. Для

основной части материала книги достаточно изложенных выше элементов теории вероятностей. Однако если мы рассматриваем бесконечные последовательности случайных величин, то общая схема с введением σ -алгебры событий становится необходимой. Действительно, ведь даже в схеме Бернулли, когда $n = \infty$, будет континуум исходов $(\varepsilon_1, \varepsilon_2, \dots)$, $\varepsilon_i = 0, 1$.

Если $X_{\Pi} = R^1 = (-\infty, +\infty)$ — прямая, то событиями считаются полуинтервалы, $(-\infty, a] = \{x: x \leq a\}$. Функция $F(y) = P\{(-\infty, y]\}$ называется **функцией распределения**. Если функция распределения дифференцируема по y , то $f(y) = dF(y)/dy$ называется **плотностью вероятностей распределения** $F(y)$. Для случайной величины g функция $P\{x: g(x) \leq y\} = F(y)$ называется функцией распределения случайной величины g . **k -й момент функции распределения** (случайной величины) определяется как интеграл

$$\alpha_k = E g^k = \int_{-\infty}^{+\infty} x^k dF(x) = \int_{-\infty}^{+\infty} x^k f(x) dx. \quad (1.8)$$

Правое равенство в (1.8) используется тогда, когда у $F(x)$ существует плотность $f(x)$.

Некоторые распределения играют существенную роль в теории вероятностей и математической статистике. Мы здесь отметим **бэ́та-распределение**, плотность которого отлична от нуля только на интервале $[0, 1]$ и имеет вид

$$f(x) = \frac{\Gamma(p_1 + p_2)}{\Gamma(p_1)\Gamma(p_2)} x^{p_1-1} (1-x)^{p_2-1}, \quad 0 \leq x \leq 1, \quad (1.9)$$

где $p_1 > 0$, $p_2 > 0$, $\Gamma(p) = \int_0^{\infty} x^{p-1} e^{-x} dx$ — гамма-функция,

$p > 0$. r -й момент бэ́та-распределения имеет вид

$$\alpha_r = \frac{\Gamma(p_1 + p_2) \Gamma(p_1 + r)}{\Gamma(p_1 + p_2 + r) \Gamma(p_1)} = \frac{(p_1 + r - 1)^{(r)}}{(p_1 + p_2 + r - 1)^{(r)}}. \quad (1.10)$$

Гамма-распределение имеет плотность вероятностей, равную

$$f(x) = \begin{cases} \frac{1}{\Gamma(p)} x^{p-1} e^{-x}, & x > 0, \\ 0, & x \leq 0. \end{cases}$$

Хи-квадрат распределение с n степенями свободы имеет плотность

$$f(x) = \begin{cases} \frac{1}{2^{n/2} \Gamma(n/2)} x^{(n/2)-1} e^{-x/2}, & x > 0, \\ 0, & x \leq 0. \end{cases}$$

Говорят, что последовательность функций распределения $\{F_n(x), n=1, 2, \dots\}$ **сходится в основном** к ф. р. $F(x)$, если для каждой точки x , в которой $F(x)$ непрерывна, $F_n(x) \rightarrow F(x), n \rightarrow \infty$. Пусть у ф. р. $F(x), F_n(x)$

существуют моменты всех порядков $m_k = \int_{-\infty}^{+\infty} x^k dF(x)$,

$m_{k,n} = \int_{-\infty}^{+\infty} x^k dF_n(x), k, n=1, 2, \dots$ и ф. р. $F(x)$ однозначно определяется своими моментами. Тогда из сходимости моментов $m_{k,n} \rightarrow m_k, n \rightarrow \infty$, следует, что $\{F_n(x), n=1, 2, \dots\}$ сходится в основном к ф. р. $F(x)$ (см. Лоев [1], стр. 198).

Важным разделом теории вероятностей является теория суммирования случайных величин. Пусть $\{\xi_n, n=1, 2, \dots\}$ — последовательность взаимно независимых случайных величин, $a_k = E\xi_k, b_k^2 = E\xi_k^2 - (E\xi_k)^2, B_n^2 = \sum_{1 \leq k \leq n} b_k^2$.

Следующая теорема называется **центральной предельной теоремой теории суммирования**.

Если определенная выше последовательность $\{\xi_n, n=1, 2, \dots\}$ удовлетворяет условию Линденберга, т. е. при любом $\tau > 0$

$$\lim_{n \rightarrow \infty} \frac{1}{B_n^2} \sum_{1 \leq k \leq n} \int_{|x-a_k| > \tau B_n} (x-a_k)^2 dF_k(x) = 0, \quad (1.11)$$

то при $n \rightarrow \infty$ равномерно по x

$$\lim_{n \rightarrow \infty} P \left\{ \frac{1}{B} \sum_{k=1}^n (\xi_k - a_k) \leq x \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-u^2/2} du. \quad (1.12)$$

Теория суммирования подробно изложена в учебнике Гнеденко [1].

Последовательность случайных величин $\{\eta_n, n=1, 2, \dots\}$ называется *сходящейся* к случайной величине η **по вероятности**, если для любого $\varepsilon > 0$, $P\{|\eta_n - \eta| > \varepsilon\} \rightarrow 0, n \rightarrow \infty$. Последовательность $\{\eta_n, n=1, 2, \dots\}$ называется *сходящейся к η с вероятностью единица*, если для любого $\varepsilon > 0$ с вероятностью единица найдется такое $n_\varepsilon = n_\varepsilon(\omega)$, что для всех $n > n_\varepsilon$ выполнено $|\eta - \eta_n| < \varepsilon$.

Последовательности взаимно независимых одинаково распределенных случайных величин $\{\xi_k\}$, $k=1, 2, \dots$, ..., $\xi_k > 0$ сопоставим случайные величины $\tau_1 = \xi_1$, $\tau_2 = \xi_1 + \xi_2$, ..., $\tau_k = \sum_{1 \leq i \leq k} \xi_i$, $\tau_1 < \tau_2 < \dots < \tau_k < \dots$.

Последовательность точек τ_k на полупрямой $R_+^1 = (0, +\infty]$ называется **процессом восстановления**. Математическое ожидание числа точек τ_k на полуинтервале $(0, x]$, обозначаемое через $H(x)$, называется **функцией восстановления**. Функция восстановления удовлетворяет интегральному уравнению

$$H(x) = F(x) + \int_0^x H(x-t) dF(t),$$

где $F(t)$ — функция распределения случайных величин ξ_k , $k=1, 2, \dots$. **Альтернирующим процессом восстановления** называется такой процесс, у которого $P\{\xi_{2k+1} \leq x\} = F(x)$, $P\{\xi_{2k} \leq x\} = G(x)$. Примерами процессов восстановления могут служить последовательности моментов τ_k появлений определенных комбинаций из нулей и единиц, например, серий нулей длины i в схеме испытаний Бернулли. Серии нулей и серии единиц в схеме испытаний Бернулли образуют альтернирующий процесс восстановления. С элементами теории восстановления можно ознакомиться по книге Кокса и Смита [1].

Статистические выводы. Общая схема математической модели в статистике отличается от рассматриваемых моделей в теории вероятностей тем, что вместо одной вероятностной меры P в нее включается целое семейство $\mathcal{P}_\Pi$ вероятностных мер. При этом считается неизвестным, с какой вероятностной мерой $P \in \mathcal{P}_\Pi$ реализован наблюдаемый в результате эксперимента Π исход $x \in X_\Pi$. Таким образом, **статистическая математическая модель** определяется заданием плана Π , со-

ответствующего ему пространства исходов X_Π , набора событий (σ -алгебры событий) $\mathfrak{B}_\Pi$, семейства вероятностных мер $\mathcal{P} = \mathcal{P}_\Pi$. Кратко такую модель обозначим $[\Pi, X_\Pi, \mathfrak{B}_\Pi, \mathcal{P}]$. В рассмотренном выше примере схемы испытаний Бернулли каждому значению q , $0 \leq q \leq 1$, соответствует своя вероятностная мера $\mathbf{P} = \mathbf{P}_q$. Поэтому можно положить $\mathcal{P} = \{\mathbf{P}_q, 0 \leq q \leq 1\}$.

Статистикой называется функция исхода эксперимента, которая является случайной величиной и не зависит от неизвестных параметров.

Особую роль играют достаточные статистики. Статистика $t(x)$, $x \in X_\Pi$ называется **достаточной**, если при любом уровне t_0 условное распределение на атомах $A(t, t_0) = \{x: t(x) = t_0\}$ не зависит от меры $\mathbf{P} \in \mathcal{P}$. В рассмотренной выше схеме испытаний Бернулли достаточной статистикой в соответствии с (1.1), в частности, является $t(x) = \delta_n$. Для пространств исходов X_Π , содержащих не более счетного числа исходов x , основные результаты, относящиеся к достаточным статистикам, выводятся несложно. Поэтому мы их приводим ниже с доказательствами.

Если число исходов $x \in X_\Pi$ не более чем счетно, то каждому исходу x соответствует вероятность $p(x)$, которая определяется выбором вероятностной меры $\mathbf{P} \in \mathcal{P}$. Если фиксировать x и менять $\mathbf{P}$, то значения $p(x)$, вообще говоря, будут меняться, образуя функцию $x(\mathbf{P})$, $\mathbf{P} \in \mathcal{P}$, которая называется **функцией правдоподобия**. Функция правдоподобия зависит от исхода x , однако, возможно, что разным исходам x соответствуют одинаковые функции правдоподобия. Например, в схеме n испытаний Бернулли исходам x с одинаковыми значениями δ_n соответствуют одинаковые функции правдоподобия $q^{\delta_n} \bar{q}^{n-\delta_n}$, $0 \leq q \leq 1$. Два исхода, x_1 и x_2 , будем относить к одному **классу правдоподобия**, если для некоторой $\mathbf{P}_0 \in \mathcal{P}$, $\mathbf{P}_0(x) > 0$, нормированные функции правдоподобия $x(\mathbf{P})/\mathbf{P}_0(x)$, соответствующие значениям x , равным x_1 и x_2 , совпадают. Классы правдоподобия образуют разбиение пространства исходов. В схеме Бернулли классы соответствуют значениям $\delta_n = 0, 1, \dots, n$. Следующая теорема, полученная Нейманом [1] и Фишером [1], фактически утверждает, что

атомы достаточных статистик не могут содержать элементы разных классов правдоподобия.

Теорема 2.1.1 (критерий факторизации). Статистика $t(x)$ является достаточной тогда и только тогда, когда функция правдоподобия представима в виде произведения двух сомножителей

$$x(\mathbf{P}) = g(t(x), \mathbf{P}) h(x), \quad (1.13)$$

где первый сомножитель зависит от $\mathbf{P} \in \mathcal{P}$ и от значения исхода x только через значение статистики $t(x)$, а второй сомножитель зависит от x , но не зависит от $\mathbf{P} \in \mathcal{P}$.

Доказательство. Покажем сначала необходимость условия (1.13). Пусть $t(x)$ — достаточная статистика. Вероятности атомов статистики $t(x)$ можно рассматривать как функцию $g(t_0, \mathbf{P}) = \mathbf{P}\{A(t, t_0)\}$ двух переменных — уровня t_0 и вероятностной меры $\mathbf{P}$. По определению достаточной статистики условная вероятность исхода x при условии $t(x) = t_0$ не зависит от $\mathbf{P}$, т. е. $\mathbf{P}\{x | t(x) = t_0\} = p(x) / \mathbf{P}(A(t, t_0)) = h(x)$, где $h(x)$ не зависит от $\mathbf{P} \in \mathcal{P}$. Таким образом получаем, что тождество

$$p(x) = \mathbf{P}(A(t, t_0)) \times p(x) / \mathbf{P}(A(t, t_0)),$$

рассматриваемое как функция $\mathbf{P}$ при любом фиксированном x , эквивалентно (1.13). Необходимость доказана. Если (1.13) рассматривать при фиксированной вероятностной мере $\mathbf{P}$, то условная вероятность исхода x при условии $t(x) = t_0$ равна

$$\frac{p(x)}{\mathbf{P}(A(t, t_0))} = \frac{g(t_0, \mathbf{P})}{\mathbf{P}(A(t, t_0))} \cdot h(x). \quad (1.14)$$

Просуммируем (1.14) по $x \in A(t, t_0)$. Суммирование правых частей в (1.14) дает единицу, поэтому получаем

$$1 = \frac{g(t_0, \mathbf{P})}{\mathbf{P}(A(t, t_0))} \sum_{x_i \in A(t, t_0)} h(x_i). \quad (1.15)$$

Находя из (1.15) выражение для $g(t_0, \mathbf{P})$ и используя его в (1.14), приходим к выражению

$$\frac{p(x)}{\mathbf{P}(A(t, t_0))} = \frac{h(x)}{\sum_{x' \in A(t, t_0)} h(x')}. \quad (1.16)$$

Так как $h(x)$ не зависит от P , то из (1.16) следует, что условное распределение на атомах статистики $t(x)$ не зависит от P . Следовательно, статистика $t(x)$ является достаточной.

Критерий факторизации (1.13) может быть обобщен на модели, в которых X_{Π} содержит несчетное множество исходов. В этом случае надо потребовать, чтобы все меры $P \in \mathcal{P}$ были бы представимы в виде интегралов $P(A) = \int_{A \in \mathcal{B}_{\Pi}} p(x) \lambda(dx)$ относительно некоторой

меры λ . В этом случае $p(x)$ есть плотность вероятностей по мере λ . Если при фиксированном x менять P , то значения $p(x)$ образуют функцию $x(P)$, которая называется функцией правдоподобия и т. д. Теория строится аналогично рассмотренному выше дискретному случаю, но при этом приходится привлекать глубокие факты из теории меры. Изложение этого подхода читатель может найти в книгах Закса [1] и Лемана [1].

Основной задачей статистики является построение **решающей функции** $\delta = \delta(x)$, дающей отображение пространства исходов X_{Π} в пространство D возможных решений. Примером решающих функций являются точечные оценки. Каждой вероятностной мере $P \in \mathcal{P}$ мы можем сопоставить число или вектор с вещественными компонентами $\psi(P)$. Таким образом ψ есть функционал, заданный на вероятностных мерах семейства $\mathcal{P}$. В схеме биномиальных испытаний $\psi(P) = q$ является примером функционала. Пусть $D_{\psi} = \{\psi(P) : P \in \mathcal{P}\}$ — множество возможных значений функционала ψ .

Точечной оценкой $\psi(P)$ будем называть статистику $t(x)$, отображающую X_{Π} в D_{ψ} или в некоторое множество $D \subset D_{\psi}$.

Оценку $t(x)$ называют **несмещенной** для $\psi(P)$, если

$$Et(x) = \psi(P), \quad P \in \mathcal{P}. \quad (1.17)$$

При векторных функциях равенство (1.17) понимается как равенство математических ожиданий компонент векторной статистики $t(x)$ компонентам вектора $\psi(P)$. Как легко проверить, несмещенной оценкой для параметра q в схеме Бернулли является статистика $\hat{q} = \sum \delta_n(x)/n$.

Точечные оценки можно сравнивать между собой различными способами. Например, их можно сравнивать по величине **среднеквадратического уклонения**

$$V(t) = E[t(x) - \psi(P)]^2, \quad (1.18)$$

которое равно дисперсии, если оценка $t(x)$ — несмещенная. Для получения несмещенных оценок с меньшей дисперсией можно использовать **метод усреднения по достаточной статистике**, который основан на известной теореме Блекуэлла — Рао; см. Рао [1], Колмогоров [4]. Эту теорему мы для простоты докажем для не более чем счетного пространства X_{Π} , однако сформулируем теорему чуть в более общей форме. Для этого аналогично (1.17) статистику $t(x)$ будем называть несмещенной оценкой для случайной функции $\psi(x, P)$, зависящей от исхода x и неизвестной вероятностной меры P , если для любой $P \in \mathcal{P}$

$$Et(x) = E\psi(x, P)^*.$$

Теорема 2.1.2. Пусть $t(x)$ есть несмещенная оценка случайной функции $\psi(x, P)$, а $s(x)$ — достаточная статистика для семейства $\mathcal{P}$. Тогда статистика

$$t_s(x) = E(t | s), \quad (1.19)$$

равная условному математическому ожиданию t при условии s , также является несмещенной оценкой для $\psi(x, P)$. При этом дисперсия t_s не более дисперсии исходной оценки t .

Доказательство. На каждом из атомов $A(s, u)$ зададим некоторую «систему координат». Результат наблюдения $x \in X_{\Pi}$ запишем в виде пары $x = (u, y_u)$, если значение статистики $s(x) = u$. Здесь y_u соответствует координатам точек атома $A(s, u)$ достаточной статистики $s(x)$. Из критерия факторизации следует, что вероятность наблюдения $x = (u, y_u)$ имеет вид

$$p(x) = g(u, P) h(u, y_u), \quad (1.20)$$

где без ограничения общности можно считать, что $\sum_{(u, y_u) \in A(s, u)} h(u, y_u) = 1$, т. е. $h(u, y_u)$ есть условная

*) Здесь E соответствует вероятностной мере P .

вероятность наблюдения исхода $x = (u, y_u)$ при условии, что значение достаточной статистики $s(x) = u$. В этих обозначениях мы можем записать выражение для оценки t_s , которая с учетом (1.20) равна

$$\begin{aligned} t_s = E(t | s(x) = u) &= \sum_{x=(u, y_u) \in A(s, u)} \frac{t(x) p(x)}{g(u, P)} = \\ &= \sum_{(u, y_u) \in A(s, u)} t(u, y_u) h(u, y_u). \end{aligned} \quad (1.21)$$

Из (1.21) следует, что t_s зависит от исхода только через достаточную статистику $s(x)$. Несмещенность оценки t_s в (1.21) легко проверяется на основе формулы полного условного математического ожидания:

$$E t_s = E(E(t | s)) = E t(x) = E \psi(x, P).$$

Далее из (1.21) и свойства (1.5) имеем

$$\begin{aligned} E(t(x) - t_s)(t_s - E\psi) &= \\ &= E\{E(t_s - E\psi)(t(x) - t_s) | s(x) = u\} = \\ &= E\{(t_s - E\psi) E[(t(x) - t_s) | s(x) = u]\} = 0, \end{aligned} \quad (1.22)$$

так как $E[(t(x) - t_s) | s(x) = u] = E[t(x) | s(x) = u] - t_s = 0$. Учитывая (1.22), получаем, что дисперсия оценки $t(x)$ равна

$$\begin{aligned} V(t) = E(t(x) - E\psi)^2 &= E[(t(x) - t_s) + (t_s - E\psi)]^2 = \\ &= E(t(x) - t_s)^2 + 2E(t(x) - t_s)(t_s - E\psi) + \\ &\quad + V(t_s) \geq V(t_s). \end{aligned} \quad (1.23)$$

Неравенство (1.23) соответствует утверждению теоремы, которая остается верной и для пространств X с бесконечным числом исходов.

Во многих моделях математической статистики семейства вероятностных мер $\mathcal{P}$ задаются в параметрической форме. При этом вероятность исхода x или плотность вероятности исхода имеет вид $p(x) = p(x, \theta)$, где $p(x, \theta)$ — определяемая в модели функция двух переменных: результата наблюдения x и неизвестного параметра $\theta \in \Theta$. Обычно множество значений параметра Θ является подмножеством евклидова пространства. Например, в схеме испытаний Бернулли исходу

$x = (\varepsilon_1, \dots, \varepsilon_n)$ соответствует вероятность $p(x, q) = q^{\delta_n} \bar{q}^{n-\delta_n}$, $\theta = q \in \Theta = [0, 1]$. Методы получения оценок параметров разнообразны. Широко используется **метод максимума правдоподобия**, при котором исходу x сопоставляется оценка $\theta(x)$, удовлетворяющая условию

$$\max_{\theta \in \Theta} p(x, \theta) = p(x, \hat{\theta}(x)).$$

Для получения оценок применяются также **методы моментов** и **квантилей**; см. Гнеденко, Беляев, Соловьев [1], стр. 172.

Важным классом решающих функций являются доверительные множества и соответствующие им интервальные оценки. Пусть значениями решающей функции $\delta(x)$ являются подмножества Θ_x некоторого множества $\Theta = \{\theta\}$, $\theta = \theta(P)$, $P \in \mathcal{P}$, т. е. $\delta(x) = \Theta_x \subset \Theta$ образует отображение пространства исходов X_{Π} в подмножества Θ . Будем предполагать, что для любого элемента $d \in \Theta$ множество $\{x: \Theta_x \ni d\}$ (множество тех $x \in X_{\Pi}$, для которых Θ_x содержит d) является событием. Кратко будем говорить, что такие **множества** являются **случайными** и образуют **систему случайных множеств** $\{\Theta_x\}$. Если $\{\Theta_x\}$ — система случайных множеств, то для любой $P \in \mathcal{P}$ и $\theta = \theta(P)$ определены вероятности $P\{\Theta_x \ni \theta\}$. Нижняя грань этих вероятностей

$$\gamma = \inf_{P \in \mathcal{P}} P\{\Theta_x \ni \theta(P)\} \quad (1.24)$$

называется **коэффициентом доверия системы** $\{\Theta_x\}$, а сами случайные множества Θ_x называются **γ -доверительными**.

Рассмотрим некоторую функцию $\psi(x, d)$, $x \in X_{\Pi}$, $d \in D$. Предположим, что в D определена система γ -доверительных множеств $\{D_x\}$, причем функции

$$\underline{\psi}(x) = \inf_{d \in D_x} \psi(x, d) \quad \text{и} \quad \bar{\psi}(x) = \sup_{d \in D_x} \psi(x, d) \quad (1.25)$$

являются статистиками. Так как событие $\{D_x \ni d\}$ влечет за собой событие $\{\psi(x) \leq \psi(x, d) \leq \bar{\psi}(x)\}$, то его вероятность не менее $P\{\underline{D}_x \ni d\} \geq \gamma$. Таким образом, имеем

$$P\{\underline{\psi}(x) \leq \psi(x, d) \leq \bar{\psi}(x)\} \geq \gamma. \quad (1.26)$$

Из (1.26) следует, что система γ -доверительных множеств для функции $\psi(x, d)$ есть система случайных интервалов $[\underline{\psi}(x), \bar{\psi}(x)]$, концы которых определяются соотношениями (1.25).

Примеры построения γ -доверительных интервалов в задачах теории надежности имеются в работах Большева и Логинова [1], Беляева [1], [2], [4], Павлова [1], [2], Судакова [1] и др.

Статистической гипотезой H называется любое утверждение о виде вероятностной меры P , в соответствии с которой получен наблюдаемый исход x . Например, в определенной выше схеме испытаний Бернулли гипотезами будут утверждения, что значение $q \leq q_0$ или что $q = q_0$, и т. п. Таким образом, гипотезам соответствуют определенные подмножества семейства вероятностных мер $\mathcal{P}$. Если гипотезе H соответствует одна определенная мера P , то эта гипотеза называется **простой**. В противном случае гипотеза называется **сложной**. В математической статистике подробно исследован случай проверки двух гипотез. Одна из них называется **основной** или **нулевой** (H_0), а другая — **конкурирующей** (H_k). При проверке двух гипотез строится решающая функция $\delta(x)$, принимающая два значения, 0 и 1. Если $\delta(x) = 0$, то считают, что исход x согласуется с гипотезой H_0 , или что гипотеза H_0 подтверждается. Если же $\delta(x) = 1$, то считают, что гипотеза H_0 отвергается. Если функция $\delta(x)$ определена, то говорят, что задан **критерий проверки** гипотез. Основной числовой характеристикой решающей функции $\delta(x)$ при проверке гипотез является **функция мощности**

$$\omega_\delta(P) = P\{\delta(x) = 1\}, \quad P \in \mathcal{P}. \quad (1.27)$$

Если $P \in \mathcal{P}_0$, где $\mathcal{P}_0$ — совокупность вероятностных мер, соответствующих гипотезе H_0 , то $\omega_\delta(P)$ соответствует **вероятности напрасного отвержения верной гипотезы** H_0 . Величина

$$\alpha_\delta = \sup_{P \in \mathcal{P}_0} \omega_\delta(P) \quad (1.28)$$

называется **размером критерия** или **ошибкой первого рода**. Для $P \in \mathcal{P}_1$, где $\mathcal{P}_1$ — семейство вероятностных мер, образующих конкурирующую гипотезу H_k , с вероятностью $1 - \omega_\delta(P)$ гипотеза H_k будет **ошибочно**

отвергнута. Поэтому желательно при заданной величине ошибки α_δ использовать критерии с возможно большими значениями функции мощности $w_\delta(P)$, $P \in \mathcal{P}_1$ или, как принято говорить, **более мощные критерии**. Техника построения таких критериев подробно изложена в книге Лемана [1]. Если проверка гипотез проводится на основе значений некоторой статистики $t(x)$, то она называется в таких случаях **тестовой статистикой**. При этом выделяется некоторое множество T_0 такое, что при $t(x) \in T_0$ гипотеза H_0 принимается, а при $t(x) \notin T_0$ отвергается. Часто тестовые статистики строятся на основе отношения функций правдоподобия $x(P_0)/x(P_1)$, где $P_0 \in \mathcal{P}_0$, а $P_1 \in \mathcal{P}_1$. Полученные при таком способе критерии носят название **критериев отношения правдоподобия**.

Если критерий построен с помощью тестовой статистики $t(x)$, то область отвержения исходной гипотезы $R = \{x: t(x) \notin T_0\}$ называется **критической областью**. Иногда функция распределения $F_t(z) = P\{t(x) \leq z\}$ статистики $t(x)$ одна и та же для всех $P \in \mathcal{P}_0$. Если при конкурирующих гипотезах становятся вероятными большие значения $t(x)$, то критическое множество имеет вид $\{x: t(x) > c\}$, где наибольшее c подбирают так, чтобы $P\{t(x) > c\} \leq \alpha$, $P \in \mathcal{P}_0$. При такой критической области ошибка первого рода не превысит α или будет равна α при непрерывности $F_t(x)$. Величина $F_t(t(x))$ называется **критическим уровнем значимости** критерия, соответствующим значению x . Полезно иметь в виду, что если при $P \in \mathcal{P}_0$ ф.р. $F_t(x)$ одна и та же и непрерывна по x , то критический уровень значимости является случайной величиной, равномерно распределенной на интервале $(0, 1)$.

Сравнение различных решающих функций как в задачах оценки, так и при проверке гипотез можно проводить, задавая функции потерь. **Функция потерь** $L(P, d)$ определяется для вероятностных мер $P \in \mathcal{P}$ и решений $d \in D$. Среднее значение потерь

$$R(P) = EL(P, \delta(x)),$$

рассматриваемое как функция $P \in \mathcal{P}$, называется **функцией риска**, соответствующей решающей функции $\delta(x)$. Если решающим функциям $\delta'(x)$, $\delta''(x)$ соответствуют функции риска $R'(P)$, $R''(P)$, причем $R'(P) \leq R''(P)$ и

хотя бы для одной меры $P_0 \in \mathcal{P}$, $R'(P_0) < R''(P_0)$, то решающая функция $\delta''(x)$ называется **недопустимой** (при ущербе $L(P, d)$). В противном случае решающая функция δ'' называется **допустимой относительно δ'** . Функция δ называется **допустимой относительно класса** решающих функций $\mathcal{C}$, если она допустима относительно любой $\delta' \in \mathcal{C}$. Общая теория решающих функций была развита Вальдом [2].

В заключение этого параграфа еще раз напомним, что краткий перечень некоторых из основных понятий теории вероятностей и математической статистики был приведен здесь только для удобства чтения последующего текста и не может заменить основных руководств по указанным дисциплинам.

§ 2.2. Планы первого вхождения и их характеристики

Основные определения. В первой главе были рассмотрены различные характеристики планов контроля совокупности объектов с фиксированным объемом выборки. В практике статистического контроля качества широко применяются также планы, у которых объем выборки является случайной величиной. Особо популярны планы **последовательного контроля**, предложенные Вальдом [1]. Выборочный контроль совокупности можно проводить и на основе других планов выборочного контроля. Каждый план выборочного контроля определяется **системой правил отбора объектов на контроль, прекращения отбора и принятия соответствующего решения**. При этом естественно строить теорию для достаточно широкого класса планов выборочного контроля, включающего в себя одноступенчатые и двухступенчатые планы, а также планы Вальда последовательного контроля. В этой и следующей главах рассматривается определяемый ниже широкий класс планов первого вхождения. К необходимости изучения планов первого вхождения приходят при отыскании оптимальных планов контроля, минимизирующих некоторые заранее определяемые классы функций риска, связанные с расходом средств на проведение контроля и с ущербом от принятия тех или иных ошибочных решений.

решения. Это может означать, что по результатам контроля принимается решение забраковать или принять контролируемую совокупность $\mathfrak{F}$. В конкретных условиях производства решение о забраковании в одних случаях может означать сплошную проверку оставшейся непроконтролированной части совокупности, в других — уценку стоимости объектов. Возможны случаи, когда забракование означает уничтожение совокупности. Для определенности, как и выше в § 1.4, ограничимся таким контролем, по результатам которого принимается одно из трех решений: $\mathcal{D}_1$, $\mathcal{D}_2$, $\mathcal{D}_3$. Таким образом, каждый путь $\varepsilon_1, \dots, \varepsilon_{k+d}$ заканчивается в результате принятия решения об остановке выборочного контроля. После этого по результатам контроля принимается одно из возможных решений $\mathcal{D}_i$. В общем случае решения $\mathcal{D}_i$ могут приниматься при одном и том же пути $\varepsilon_1, \dots, \varepsilon_{k+d}$ с некоторыми вероятностями. Мы ограничимся классом планов контроля типа первого вхождения, в котором решение $\mathcal{D}_i$ будет однозначно определяться координатами точки остановки.

Определение 2.2.1. Пусть на множестве точек решетки $\mathbb{R}^2 = \{(k, d), k, d = 0, 1, \dots\}$ задано подмножество граничных точек Γ , которое в свою очередь разбито на взаимно, непересекающиеся части Γ_i ($\Gamma = \bigcup_i \Gamma_i$). Контроль прекращается, как только путь впервые достигает граничной точки $(k, d) \in \Gamma$. При этом если граничная точка (k, d) принадлежит подмножеству Γ_i , то принимается решение $\mathcal{D}_i$. Определяемый таким образом план называется **планом первого вхождения**.

План с двумя решениями $\mathcal{D}_i$, $\mathcal{D}_j$ обозначается $\Pi_{i,j}$. На рис. 2.2.1 приведен пример плана первого вхождения. Граница Γ разбита на три участка. Если путь выходит на участок A_1, B_1 , то принимается решение $\mathcal{D}_1$, т. е. остаток совокупности бракуется без контроля, на участках A_2B_2 , A_3B_3 соответственно, принимаются решения $\mathcal{D}_2$ и $\mathcal{D}_3$. На рис. 2.2.1 показан путь, выходящий на участок границы, где после остановки принимается решение $\mathcal{D}_2$.

Класс планов первого вхождения будем обозначать Π . Это достаточно широкий класс планов контроля. Он

включает в себя планы типа одноступенчатой выборки, когда граница Γ образована точками (k, d) , $k + d = n$, n — объем выборки (рис. 2.2.2, а). Выше уровня $d = c$ принимается решение о браковке ($\mathcal{D}_1$ или $\mathcal{D}_2$); для $d \leq c$ — решение о приемке ($\mathcal{D}_2$ или $\mathcal{D}_3$).

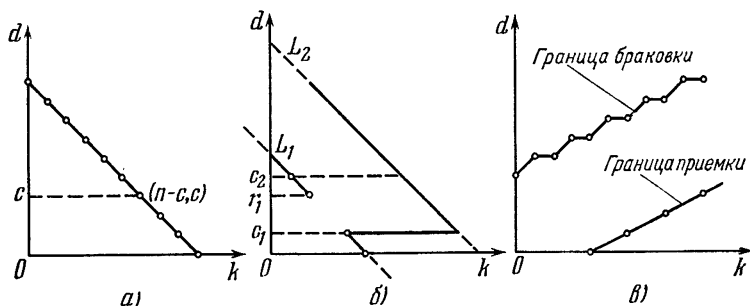


Рис. 2.2.2. Планы первого вхождения: а) одноступенчатый, б) двухступенчатый, в) последовательный.

Планы типа двухступенчатой выборки также являются планами первого вхождения. Такие планы определяются следующим образом. Задаются два приемочных числа c_1 , c_2 и браковочное число r_1 . Сначала проводится контроль n_1 объектов. Если при этом обнаружено d_1 дефектных объектов, то при $d_1 \leq c_1$ совокупность принимается, а при $d_1 \geq r_1$ бракуется. В случае $c_1 < d_1 < r_1$ дополнительно проверяется еще n_2 объектов, в результате чего выявляется d_2 дефектных объектов. Если окажется, что $d_1 + d_2 \leq c_2$, то совокупность принимается. Если же $d_1 + d_2 > c_2$, то бракуется. Таким образом, граница Γ образована интервалами точек (k, d) , лежащих на прямых L_1 , L_2 , для которых $k + d = n_1$, $k + d = n_2$ соответственно. На рис. 2.2.2, б приведен пример плана типа двухступенчатой выборки.

Наконец, в связи с некоторыми свойствами оптимальности широко используются планы последовательного типа, границы которых представлены в виде параллельных ломаных, аппроксимируемых параллельными прямыми. Примерный вид границ таких планов показан на рис. 2.2.2, в.

Каждый из указанных частных типов планов первого вхождения обладает рядом преимуществ и недостатков. Отметим здесь, что планы типа одноступенчатой выборки значительно проще организовать, чем, скажем, планы последовательного типа. Однако, используя планы двухступенчатые и последовательные, можно (при том же среднем объеме контроля при нормальном ходе производства и при той же трудоемкости контроля, что и при одноступенчатых планах) получить большие гарантии браковки совокупностей, сильно засоренных дефектными объектами.

Аналогично одноступенчатым планам, планы первого вхождения с решениями $\mathcal{D}_1$, $\mathcal{D}_2$ могут быть использованы в тех случаях, когда контроль не разрушительный, стоимость контроля каждого объекта велика и требования к качеству высокие. Планы с решениями $\mathcal{D}_1$, $\mathcal{D}_3$ применяются при разрушительном контроле, а также в тех случаях, когда стоимость контроля объектов велика. Наконец, планы с решениями $\mathcal{D}_2$, $\mathcal{D}_3$ применяются как при разрушительном, так и при неразрушительном контроле, когда важно получить дополнительную информацию о качестве продукции. У разных планов контроля различны их количественные характеристики. Их значениями мы и руководствуемся при выборе планов для контроля.

Вероятности результатов контроля. Для того чтобы правильно рассчитать характеристики планов контроля, нужно сначала четко определить правило отбора объектов на контроль. Основным методом отбора объектов на контроль является случайная выборка без возвращения, определение которой было дано выше в § 1.1. Введем случайные величины $\delta_n = \sum_{1 \leq i \leq n} \varepsilon_i$ и $\bar{\delta}_n = \sum_{1 \leq i \leq n} \bar{\varepsilon}_i$, где $\varepsilon_i = 0, 1$, $\bar{\varepsilon}_i = 1 - \varepsilon_i$. Очевидна справедливость равенства $\delta_n + \bar{\delta}_n = n$. Подсчитаем вероятность появления пути $\varepsilon_1, \dots, \varepsilon_n$. Пусть контролируемая совокупность $\mathfrak{F}$ содержит D дефектных объектов и $K = N - D$ годных, т. е. $\mathfrak{F} \asymp [N, D]$. При этом предположении вероятность наблюдения ε_1 при контроле первого случайным образом отобранного объекта равна $\frac{D \varepsilon_1 \cdot K \bar{\varepsilon}_1}{N}$, т. е. $P\{\varepsilon_1 = 1\} = D/N$ и $P\{\varepsilon_1 = 0\} = K/N$. После

контроля первого объекта совокупность содержит $N - 1$ объектов, из которых $D - e_1$ дефектных и $K - \bar{e}_1$ годных. Следовательно, условная вероятность появления e_2 при фиксированном значении e_1 равна $\frac{(D - e_1)^{e_2} (K - \bar{e}_1)^{\bar{e}_2}}{N - 1}$. Вероятность появления пары e_1, e_2 равна произведению вычисленных вероятностей, т. е. равна

$$\frac{D^{e_1} (D - e_1)^{e_2} K^{\bar{e}_1} (K - \bar{e}_1)^{\bar{e}_2}}{N(N - 1)}. \quad (2.1)$$

После контроля двух объектов с результатами контроля e_1, e_2 остается совокупность из $N - 2$ объектов, среди которых $D - (e_1 + e_2)$ дефектных. В обозначениях $M^{(m)} = M(M - 1) \cdot \dots \cdot (M - m + 1)$ формула (2.1) переписывается в виде

$$P\{e_1, e_2\} = \frac{D^{(e_1 + e_2)} \cdot K^{(\bar{e}_1 + \bar{e}_2)}}{N^{(2)}}. \quad (2.2)$$

Используя метод математической индукции, легко получить общую формулу для вероятности пути $e_1, \dots, e_n$

$$P\{e_1, \dots, e_n\} = \frac{D^{(\delta_n)} \cdot K^{(\bar{\delta}_n)}}{N^{(n)}}. \quad (2.3)$$

Действительно, после контроля n объектов остается совокупность $\mathfrak{B} \propto [N - n, D - \delta_n]$. Поэтому условная вероятность регистрации e_{n+1} при фиксированных значениях $e_1, \dots, e_n$ равна

$$\begin{aligned} P(e_{n+1} | e_1, \dots, e_n) &= \frac{(D - \delta_n)^{e_{n+1}} (K - \bar{\delta}_n)^{\bar{e}_{n+1}}}{N - n} = \\ &= \frac{(D - \delta_{n+1} + 1)^{e_{n+1}} (K - \bar{\delta}_{n+1} + 1)^{\bar{e}_{n+1}}}{N - (n + 1) + 1}. \end{aligned} \quad (2.4)$$

Из (2.3) и (2.4) получаем

$$\begin{aligned} P(e_1, \dots, e_{n+1}) &= P(e_1, \dots, e_n) \cdot P(e_{n+1} | e_1, \dots, e_n) = \\ &= \frac{D^{(\delta_n)} (D - \delta_{n+1} + 1)^{e_{n+1}} K^{(\bar{\delta}_n)} (K - \bar{\delta}_{n+1} + 1)^{\bar{e}_{n+1}}}{N^{(n)} [N - (n + 1) + 1]} = \\ &= \frac{D^{(\delta_{n+1})} K^{(\bar{\delta}_{n+1})}}{N^{(n+1)}}, \end{aligned}$$

что и требовалось доказать. Таким образом, справедлива

Теорема 2.2.1. При случайной выборке без возвращения вероятность пути $e_1, \dots, e_n$, $\sum_{1 \leq i \leq n} e_i = \delta_n$, $\sum_{1 \leq i \leq n} \bar{e}_i = \bar{\delta}_n$ вычисляется по формуле (2.3).

Возможны другие методы отбора объектов для контроля. Например, в статье чешского математика Гаека [1] рассмотрен следующий метод отбора.

Определение 2.2.2. Пусть $\mathfrak{P} = \{O_1, \dots, O_N\}$ — совокупность N объектов, а $\gamma_1, \dots, \gamma_N$ — биномиальная последовательность, т. е. $\gamma_i = 0, 1$ взаимно независимы и $P\{\gamma_i = 1\} = v$, $i = 1, \dots, N$. Объект O_i включается в выборку тогда и только тогда, когда $\gamma_i = 1$. Образованную таким способом выборку будем называть **биномиальной** *).

В соответствии с определением 2.2.2 объем биномиальной выборки, т. е. число составляющих ее объектов, есть случайная величина η , для которой

$$P\{\eta = n\} = P\left\{\sum_{i=1}^N \gamma_i = n\right\} = \binom{N}{n} v^n \bar{v}^{N-n}.$$

Заметим, что биномиальную выборку можно реализовать следующим образом. Сначала получаем с помощью таблиц случайных чисел значение биномиальной случайной величины η , а затем из $\mathfrak{P}$ берем случайную неповторную выборку объема η .

Получим один частный полезный результат о распределении числа δ дефектных объектов, выявленных при биномиальной выборке. По формуле полных вероятностей для $n = k + d$ имеем

$$\begin{aligned} P\{\delta = d\} &= \sum_{n=d}^N P\left\{\sum_{i=1}^n e_i = d \mid \sum_{i=1}^N e_i = D\right\} \binom{N}{n} v^n \bar{v}^{N-n} = \\ &= \sum_{k=0}^K \binom{n}{d} \frac{K^{(k)} D^{(d)}}{N^{(n)}} \cdot \frac{N^{(n)}}{n!} v^n \bar{v}^{N-n} = \\ &= \sum_{k=0}^K \frac{K^{(k)}}{k!} v^k \bar{v}^{K-k} \frac{D^{(d)}}{d!} v^d \bar{v}^{D-d} = \frac{D^{(d)}}{d!} v^d \bar{v}^{D-d} = \binom{D}{d} v^d \bar{v}^{D-d}. \end{aligned} \quad (2.5)$$

*) В статье Гаека [1] такой способ отбора называется **пуассоновским**.

Здесь мы воспользовались тем фактом, что при $K \leq N - d$

$$\sum_{k=0}^{N-d} \frac{K^{(k)}}{k!} v^k \bar{v}^{K-k} = \sum_{k=0}^K \binom{K}{k} v^k \bar{v}^{K-k} \equiv 1.$$

Таким образом, при биномиальной выборке число δ обнаруженных дефектных объектов есть биномиальная случайная величина с параметрами D и v .

Вернемся к методу случайной бесповторной выборки. Пусть план контроля $P \in \Pi$, т. е. P есть план первого вхождения с границей Γ . Контроль прекращается, когда путь контроля достигает границы. Точки решетки $\mathbb{N}^2$ можно классифицировать следующим образом.

Определение 2.2.3. При случайной бесповторной выборке и плане первого вхождения точка $x = (k, d) \in \mathbb{N}^2$ называется **достижимой**, если при некотором D с положительной вероятностью существует хотя бы один путь, который проходит через точку x или оканчивается в точке x . Если такого пути нет, то точка x называется **недостижимой**. Достижимая точка, не являющаяся граничной, называется **проходимой**.

Вероятность $p_D(k, d)$ того, что путь контроля подойдет к достижимой точке (k, d) , легко вычисляется по числу $L(k, d)$ всех возможных путей (т. е. всех путей, имеющих положительные вероятности), соединяющих точку $(0, 0)$ начала контроля с точкой (k, d) . В соответствии с формулой (2.3) все пути $(\varepsilon_1, \dots, \varepsilon_n)$, у которых $\delta_n = d$, $\bar{\delta}_n = k$, $k + d = n$, имеют равные вероятности. Следовательно,

$$p_D(k, d) = L(k, d) \frac{D^{(d)} K^{(k)}}{N^{(k+d)}}. \quad (2.6)$$

В § 2.1 была отмечена та особая роль, которую при построении оценок играют достаточные статистики.

В исследуемых нами задачах неизвестным параметром, однозначно определяющим вероятностную меру P , является число D дефектных объектов. Результаты наблюдения x соответствуют пути $(\varepsilon_1, \dots, \varepsilon_{k+d})$.

Теорема 2.2.2. При контроле совокупности $\mathfrak{P} \times [N, D]$ с помощью случайной выборки без возвра-

щения с использованием плана первого входа с границей Γ достаточной статистикой являются координаты точки остановки $x = (k, d) \in \Gamma$.

Доказательство является простым следствием полученных соотношений (2.3) и (2.6). Условная вероятность пути равна

$$P(e_1, \dots, e_n | x = (k, d) \in \Gamma, k + d = n) = \\ = \frac{D^{(d)} K^{(k)}}{N^{(n)}} \bigg| p_D(k, d) = \frac{1}{L(k, d)},$$

т. е. не зависит от неизвестного значения D . Следовательно, координаты точки остановки образуют достаточную статистику.

Еще раз отметим, что достаточность статистики $x = (k, d) \in \Gamma$ означает, что никакой информации о неизвестном значении D в порядке обнаружения годных и дефектных объектов в процессе контроля с помощью случайной выборки без возвращения и плана $\Pi \in \Pi$ нет. Вся информация о D «сконцентрирована» в координатах точки остановки $x = (k, d) \in \Gamma$.

Аналогичным образом (как и выше) легко показать, что $p_D((k_1, d_1), (k_2, d_2))$ — вероятность того, что путь пройдет через (k_1, d_1) и достигнет (k_2, d_2) , $k_1 \leq k_2$, $d_1 \leq d_2$, равна

$$p_D(k_1, d_1; k_2, d_2) = \\ = L(k_1, d_1) \frac{D^{(d_1)} K^{(k_1)}}{N^{(d_1+k_1)}} L(k_1, d_1; k_2, d_2) \frac{(D - d_1)^{(d_2)} (K - k_1)^{(k_2)}}{(N - (d_1 + k_1))^{(k_2+d_2)}} = \\ = L(k_1, d_1) L(k_1, d_1; k_2, d_2) \frac{D^{(d_1+d_2)} K^{(k_1+k_2)}}{N^{(k_1+k_2+d_1+d_2)}}, \quad (2.7)$$

где $L(k_1, d_1; k_2, d_2)$ — число путей, соединяющих точку (k_1, d_1) с точкой (k_2, d_2) .

Заметим, что для одноступенчатого плана объема n число путей $L(k, d)$, соединяющих граничную точку (k, d) , $k + d = n$, с точкой $(0, 0)$, равно числу различных выборок из n объектов по k (или по d) различных объектов, т. е. $L(k, d) = \binom{n}{k} = \binom{n}{d}$. Однако для планов $\Pi \in \Pi$ общего вида подсчет числа $L(k, d)$ представляет некоторые трудности.

Сумму координат точки $x = (k, d)$ будем называть, следуя Гиршику, Мостеллеру и Сэвиджу [1], **индексом точки**, $\text{ind } x = k + d$. **Индексом плана** $\Pi \in \Pi$ с границей Γ назовем максимум индексов его граничных точек, т. е. $\text{ind } \Pi = \max \{k + d, (k, d) \in \Gamma\} = \text{ind } \Gamma$, соответственно $\text{ind } \Gamma_i = \max \{k + d, (k, d) \in \Gamma_i\}$, $\Gamma_i \subseteq \Gamma$. Предположим, что мы хотим подсчитать $L(k, d)$, $\text{ind}(k, d) = n$. Для этого точке (k, d) сопоставим метку 1, а всем другим

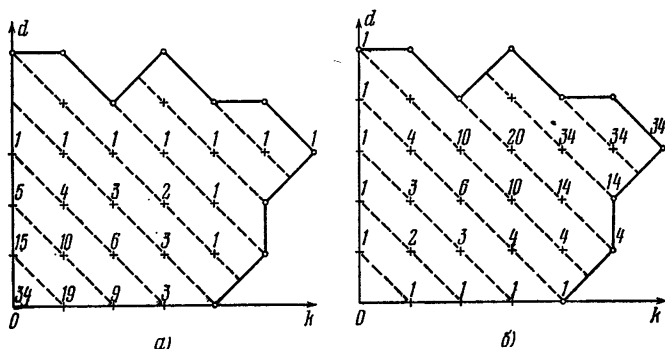


Рис. 2.2.3. Методы расчета числа путей: а) метод обратного счета, б) метод прямого счета.

достижимым точкам индекса (k', d') , $k' + d' = n$ сопоставим метку 0. Метку 0 сопоставим также всем граничным точкам независимо от их индекса. Расчет проводим сначала для точек индекса $n - 1$. Для этого каждой точке (k', d') с индексом $n - 1$ сопоставим сумму меток двух точек индексов n , достижимых из этой точки. Точно так же поступаем при расчете меток для точек индекса $m < n - 1$. Метку для (k'', d'') , $k'' + d'' = m$ получаем, складывая значения меток, достижимых за один шаг из точки (k'', d'') . В конце концов мы приходим к вычислению метки точки $(0, 0)$. Полученное значение метки точки $(0, 0)$ и есть искомое число $L(k, d)$ путей, соединяющих $(0, 0)$ с (k, d) . Этот простой метод проиллюстрирован на рис. 2.2.3, а. Значения меток выписаны над точками первого квадранта. Например, точке $(2, 2)$ соответствует метка 3. Точки границы соединены жирной линией, а точки с равными значениями ин-

дексов — наклонными пунктирными линиями. Кратко этот метод можно называть **методом обратного счета**.

Подсчет числа путей, соединяющих $(0, 0)$ с точкой (k, d) , можно проводить и в прямом порядке. Точкам $(1, 0)$ и $(0, 1)$ ставим метки, равные 1. Если поставлены метки всем достижимым точкам индекса n , то каждой достижимой точке (k, d) индекса $n+1$ сопоставляется сумма меток проходимых точек $(k-1, d)$ и $(k, d-1)$. Если точка $(k-1, d)$ ($(k, d-1)$) — граничная, то метка точки (k, d) совпадает с меткой $(k, d-1)$ ($(k-1, d)$). Одна из точек $(k-1, d)$ или $(k, d-1)$ должна быть проходимой, иначе бы точка (k, d) не могла быть достижимой. Полученное таким способом значение меток равно искомому числу путей $L(k, d)$. Этот метод можно назвать **методом прямого счета**. На рис. 2.2.3, б показан пример расчета $L(6, 3) = 34$ по методу прямого счета для того же плана Π , что и на рис. 2.2.3, а.

Рекуррентное вычисление вероятностных характеристик. Для планов $\Pi \in \Pi$ можно привести простые рекуррентные соотношения для вероятностей принятия определенного решения. Пусть $P_D(\Gamma_i)$ — вероятность принятия решения $\mathcal{D}_i$ при случайной неповторной выборке по плану Π , т. е. вероятность выхода пути в одну из точек $(k, d) \in \Gamma$. Пусть $P_D(x, \Gamma_i)$ — вероятность выхода на Γ_i при условиях, что путь начинается из точки $x = (k, d)$ и контролируется совокупность $\mathfrak{B} \asymp [N - (k + d), D - d]$. Если $x = (k, d) \in \Gamma$, то

$$P_D(x, \Gamma_i) = \begin{cases} 1, & x = (k, d) \in \Gamma_i, \quad K \geq k, \quad D \geq d, \\ 0, & \text{в противном случае.} \end{cases} \quad (2.8)$$

Эти граничные условия соответствуют тому, что после достижения граничной точки контроль прекращается. Расчет ведем от точек с индексом $n+1$ к точкам с индексом n . Итак, пусть $P_D(x, \Gamma_i)$ подсчитаны для всех x , $\text{ind } x = n+1$. Заметим, что при проверке $(n+1)$ -го объекта при условии, что к моменту проверки n -го объекта путь проходит через $x = (k, d)$, ведется контроль совокупности $\mathfrak{B} \asymp [N - (k + d), D - d]$. Следовательно, с вероятностью $(K - k) / (N - (k + d))$ $(n+1)$ -й объект окажется годным, а с вероятностью $(D - d) / (N - (k + d))$ — дефектным. Поэтому по формуле

полных вероятностей имеем

$$P_D(k, d; \Gamma_i) = \frac{(K-k)P_D(k+1, d; \Gamma_i)}{N-(k+d)} + \frac{(D-d)P_D(k, d+1; \Gamma_i)}{N-(k+d)}. \quad (2.9)$$

Последовательно уменьшая значения индекса, мы приходим к точке $(0, 0)$. При этом получим, что $P_D(0, 0; \Gamma_i) = P_D(\Gamma_i)$ является искомой вероятностью. Таким образом, $P_D(\Gamma_i)$ рассчитывается на основе граничных условий (2.8) и рекуррентных соотношений (2.9).

Точно так же выводятся рекуррентные соотношения для среднего значения объема контроля до достижения границы

$$m_D(\Gamma) = \sum_{(k, d) \in \Gamma} (k+d)P_D(k, d). \quad (2.10)$$

Для этого надо ввести $m_D(x, \Gamma)$ — средний объем контроля до достижения границы Γ при условии, что контроль начинается из точки $x = (k, d)$ и контролируется совокупность $\mathfrak{B} \propto [N-(k+d), D-d]$. Граничные условия имеют вид

$$m_D(k, d; \Gamma) = 0, \quad (k, d) \in \Gamma. \quad (2.11)$$

Рекуррентные уравнения равны

$$m_D(k, d; \Gamma) = 1 + \frac{K-k}{N-(k+d)} m_D(k+1, d; \Gamma) + \frac{D-d}{N-(k+d)} m_D(k, d+1; \Gamma), \quad (k, d) \notin \Gamma. \quad (2.12)$$

Если при контроле возможно решение $\mathcal{D}_2$, то начальное условие (2.11) на участке границы Γ_2 меняется. При расчете среднего значения числа контролируемых объектов здесь надо положить

$$m_D(k, d; \Gamma) = \begin{cases} N-(k+d), & (k, d) \in \Gamma_2, \\ 0, & (k, d) \in \Gamma \setminus \Gamma_2, \end{cases} \quad (2.13)$$

где $\Gamma \setminus \Gamma_2$ — часть границы, за исключением точек $(k, d) \in \Gamma_2$.

Часто отбор планов для «жесткого» контроля проводят на основе предельного выходного качества. Эта

характеристика используется для планов P_{23} , когда при решении $\mathcal{D}_2$ о сплошной разбраковке все выявленные дефектные объекты заменяются на годные. Как и ранее, пусть Γ_3 — участок границы, на котором принимается решение $\mathcal{D}_3$. Тогда среднее число $a_D(\Gamma)$ принятых дефектных объектов равно

$$a_D(\Gamma) = \sum_{(k, d) \in \Gamma_3} (D - d) P_D(k, d).$$

Предельное выходное качество D_L определяется как максимум этого выражения по D ; т. е.

$$D_L = \max_D \sum_{(k, d) \in \Gamma_3} (D - d) P_D(k, d). \quad (2.14)$$

То значение D , при котором достигается максимум в (2.14), обозначим D_L^* . При $D = D_L^*$ потребитель в среднем получает максимальное число дефектных объектов.

Вычисление $a_D(\Gamma)$ можно проводить, используя рекуррентные соотношения. Для этого введем функцию $a_D(k, d; \Gamma)$, граничные значения которой равны

$$a_D(k, d; \Gamma) = \begin{cases} D - d, & (k, d) \in \Gamma_3, \\ 0, & (k, d) \notin \Gamma_3. \end{cases} \quad (2.15)$$

С учетом вероятностей возможных переходов для проходимых точек $(k, d) \in \Gamma$ легко получить соотношения

$$a_D(k, d; \Gamma) = \frac{K - k}{N - (k + d)} a_D(k + 1, d; \Gamma) + \frac{D - d}{N - (k + d)} a_D(k, d + 1; \Gamma). \quad (2.16)$$

Если в правой части (2.16) одна из точек $(k, d + 1)$ или $(k + 1, d)$ принадлежит Γ , то ее значение находится из граничных условий (2.15). Последовательно вычисляя значения $a_D(k, d, \Gamma)$ для точек с уменьшающимися значениями их индексов, придем к начальной точке $(0, 0)$. Искомое значение $a_D(\Gamma) = a_D(0, 0; \Gamma)$.

Оперативная характеристика и связанные с ней показатели. Использование методов выборочного контроля связано с риском напрасной браковки хороших партий и риском принять партии, недопустимо засоренные дефектными объектами. Возможно, что при случайном

отборе объектов в выборку в ней окажется значительное число дефектных, тогда как в совокупности в целом их доля допустима. В этом случае будет принято ложное решение о браковке приемлемой совокупности. Ошибки такого типа носят название **ошибок первого рода**, или **ошибок поставщика**. С другой стороны, может случиться и так, что при сильной засоренности совокупности дефектными объектами в выборке окажется небольшое количество дефектных объектов, вследствие чего плохая совокупность будет принята. Такие ошибки называются **ошибками второго рода**, или **ошибками потребителя**. Ошибки первого и второго рода неизбежно возникают в результате применения методов выборочного контроля. Вопрос правильной организации выборочного контроля, в частности, состоит в том, чтобы такие ошибочные решения появлялись редко, т. е. были бы маловероятными событиями. Для этого надо правильно выбирать планы контроля.

Одной из наиболее важных вероятностных характеристик плана контроля с двумя решениями является так называемая оперативная характеристика.

Определение 2.2.4. **Оперативной характеристикой** плана контроля $\Pi_{i,j}$ с двумя решениями $\mathcal{D}_i, \mathcal{D}_j$, $i < j$, называется функция $\pi(D)$, равная вероятности принять решение $\mathcal{D}_i$ о приемке контролируемой совокупности, содержащей D дефектных объектов.

Как правило, оперативная характеристика $\pi(D)$ является монотонно убывающей функцией D . Примерный вид оперативных характеристик показан на рис. 2.2.4, а, б. Для наглядности оперативные характеристики обычно показывают в виде сплошных линий, хотя на самом деле имеют смысл только их значения в целых точках.

Используя понятие оперативной характеристики, можно задавать различные численные показатели планов. Из экономических или каких-либо других соображений может оказаться, что совокупности, имеющие $D < D_{кр}$, считаются хорошими. Такие совокупности следует принимать, а совокупности с $D \geq D_{кр}$ надо браковать как плохие. В этих случаях желательно было бы использовать план контроля с «идеальной» оперативной характеристикой, которая равна 1 для $D \leq D_{кр}$ и равна

нулю при $D \geq D_{кр}$; см. рис. 2.2.4, а. Однако все планы Π , кроме плана сплошного контроля всей совокупности, имеют оперативные характеристики, отличные от «идеальной».

Не всегда можно так категорично, исходя лишь из одного числа $D_{кр}$, делить совокупности на хорошие и плохие. Часто возникает зона неопределенности, когда мы не можем точно сказать, плохи эти совокупности или нет. В таких случаях исходят из двух чисел D_1 и D_2 .

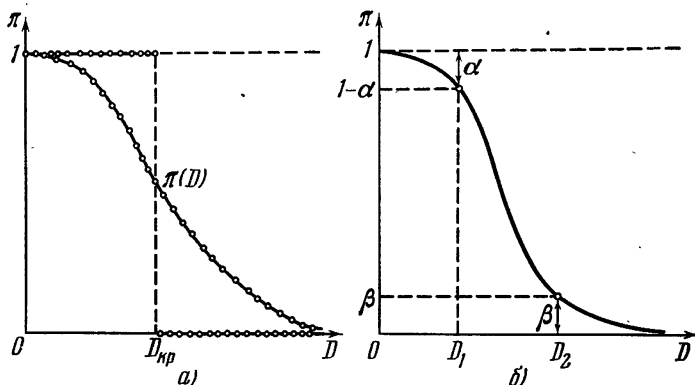


Рис. 2.2.4. Оперативные характеристики планов контроля: а) «идеальная» и «реальная» оперативные характеристики, б) оперативная характеристика и риски поставщика и потребителя.

Совокупности считают хорошими при $D \leq D_1$ и плохими при $D \geq D_2$ ($D_1 < D_2$). Поскольку оперативная характеристика планов монотонно убывает, то неизбежны ошибочные решения. Поэтому желательно так подобрать план, чтобы вероятности этих ошибочных решений были бы невелики. Обычно к плану Π предъявляются требования, чтобы $\pi(D) \geq 1 - \alpha$ для $D \leq D_1$ и $\pi(D) \leq \beta$ для $D \geq D_2$. Здесь $1 - \pi(D_1)$ — вероятность ошибки первого рода, или риск поставщика, $\pi(D_2)$ — вероятность ошибки второго рода, или риск потребителя. D_1 называют **приемлемым числом дефектных**, а D_2 — **предельно допустимым числом дефектных** объектов в контролируемой совокупности. Итак, требования к плану выборочного контроля могут состоять в том, что вероятности ошибок, т. е. риск производителя и риск

потребителя, должны быть не более α и β ; см. рис. 2.2.4, б. Часто значение D_1 берут несколько больше среднего значения засоренности дефектными объектами при нормальном ходе производства, чем и гарантируют, что почти все совокупности (партии изделий), выпущенные при соблюдении основных требований к технологии производства, будут приняты.

Число $D_{1/2}$ такое, что $\pi(D_{1/2}) \geq 1/2 > \pi(D_{1/2} + 1)$ называется **безразличным качеством** для плана с оперативной характеристикой $\pi(D)$. Если план контроля выбран, а доля засоренности совокупностей дефектными объектами совпадает с $D_{1/2}$, то в среднем половина таких совокупностей с $D = D_{1/2}$ будет принята, а половина забракована.

Возможно задание других числовых характеристик планов контроля по альтернативному признаку. Часто используются такие относительные характеристики, как **уровень качества** $q = D/N$, **уровень предельного выходного качества** $q_L = D_L/N$ и т. д. Терминология, связанная с методами контроля уровня качества, выпущена в виде ГОСТа [2], [3].

В заключение заметим, что оперативные характеристики многих типов планов легко выражаются через гипергеометрические вероятности $h_{N, D}^{n, d}$ и их суммы $H_{N, D}^{n, d}$.

Например, оперативная характеристика двухступенчатого плана с объемами первой и второй выборки n_1 и n_2 , приемочными числами c_1 , c_2 и браковочным числом r_1 , равна

$$\pi(D) = H_{N, D}^{n_1, c_1} + \sum_{d_1=c_1+1}^{r_1-1} h_{N, D}^{n_1, d_1} H_{N-n_1, D-d_1}^{n_2, c_2-d_1}. \quad (2.17)$$

В (2.17) первое слагаемое соответствует вероятности принять совокупность в результате того, что число дефектных объектов δ_{n_1} в первой выборке не превысит c_1 . Второе слагаемое соответствует тем случаям, когда $c_1 < \delta_{n_1} < r_1$. При этом берется вторая выборка объема n_2 и совокупность принимается, если число δ_{n_2} дефектных объектов, обнаруженных во второй выборке, не более $c_2 - d_1$ ($\delta_{n_1} = d_1$).

Через гипергеометрическое распределение выражаются различные вероятности для планов первого вхождения с невозрастающей границей, которая дана

на рис. 2.2.5. Для точек границы Γ введем понятие **порядка**. Скажем, что точка $x = (k_x, d_x) \in \Gamma$ «**предшествует**» точке $y = (k_y, d_y) \in \Gamma$, ($x < y$), если либо $k_x < k_y$, либо $k_x = k_y$, но $d_x > d_y$. Для планов с невозрастающей границей все точки можно разбить на два класса. В первый класс включаем точки $x = (k_x, d_x) \in \Gamma$ такие, что $(k_x - 1, d_x)$ являются проходимыми. Во второй класс включаем все остальные точки. На рис. 2.2.5 x — точка первого класса, y — точка второго класса. Пусть $x^* = (k^*, d^*)$ — координаты случайной точки выхода на границу в результате выборочного контроля совокупности $\mathfrak{P} \propto [N, D]$. Если x — точка первого класса, то событие $\{x^* \leq x\}$, т. е. событие, что точка остановки контроля предшествует x или совпадает с $x \in \Gamma$, эквивалентно прохождению пути контроля через одну из точек отрезка $[A, B]$, $A = (0, d_x)$, $B = (k_x - 1, d_x)$. Так как $\text{ind } B = k_x + d_x - 1$, то прохождение пути через $[A, B]$ эквивалентно тому, что после контроля $k_x + d_x - 1$ объектов будет обнаружено не менее d_x дефектных. Следовательно, вероятность про-

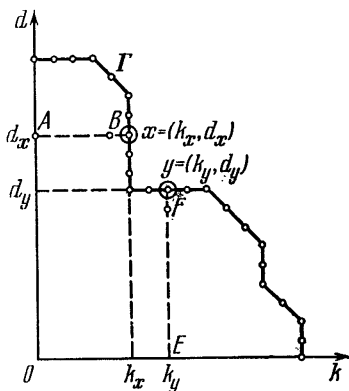


Рис. 2.2.5. План первого входа с невозрастающей границей.

хождения через $[A, B]$ равна $\sum_{d=d_x}^{\text{ind } B} h_{N, D}^{k_x + d_x - 1, d} = \bar{H}_{N, D}^{k_x + d_x - 1, d_x - 1}$. Итак, мы нашли, что вероятность выйти в точки границы, предшествующие или совпадающие с x , равна

$$P\{x^* \leq x \in \Gamma\} = \bar{H}_{N, D}^{k_x + d_x - 1, d_x - 1}. \quad (2.18)$$

Если же $y \in \Gamma$ есть точка второго класса, то аналогичный метод, но уже с использованием вертикального отрезка $[E, F]$, дает, что

$$P\{x^* \geq y \in \Gamma\} = H_{N, D}^{k_y + d_y - 1, d_y - 1}. \quad (2.19)$$

Соотношения (2.18), (2.19) показывают, что для планов с невозрастающей границей вероятности выхода на любом участке границы выражаются через гипергеометрическое распределение. Здесь же отметим, что если контроль ведется до первого обнаружения m дефектных объектов, то вероятность проверки $m + k$ объектов равна $\frac{m}{m+k} h_{N,D}^{m+k,m}$. Распределение с такими вероятностями называется *отрицательным гипергеометрическим распределением*.

§ 2.3. Априорные распределения числа дефектных объектов

Смеси биномиальных распределений. Всюду выше мы рассматривали число D дефектных объектов в совокупности $\mathfrak{F}$ как неизвестный параметр. В тех случаях, когда контролируется последовательность совокупностей, могут оказаться полезными выводы, полученные на основе изучения математических моделей, в которых число дефектных объектов в $\mathfrak{F}$ предполагается случайной величиной D_ω с функцией распределения $F_N(D) = P\{D_\omega \leq D\}$. Априорные вероятности событий $D_\omega = D$ обозначим $f_N(D) = P\{D_\omega = D\}$. Распределение $F_N(D)$ будем называть *априорным*. Вид априорного распределения самым существенным образом влияет на эффективность использования любой системы планов контроля. Это положение легко иллюстрируется следующим примером А. Н. Колмогорова. Пусть контроль последовательности совокупностей $\{\mathfrak{F}\}$ проводится с одноступенчатыми планами. Контроль предполагается разрушительным. Если в выборке обнаружен хотя бы один дефектный объект, то $\mathfrak{F}$ бракуется. Если же дефектных объектов в выборке нет, то $\mathfrak{F}$ принимается. Кажется естественным, что такая система контроля никогда не может ухудшить качества принимаемой продукции. Предположим, что априорное распределение $F_N(D) \neq 0$ только для $D = 0, 1$, т. е. в каждой из совокупностей либо все объекты годные, либо дефектен только один объект. При таком априорном распределении будут браковаться совокупности, в выборках из которых выявлен один возможный дефектный объект.

Ясно, что доля дефектных в принятых совокупностях может оказаться больше, чем если бы контроль не проводился. При сохранении контроля доля дефектных объектов в принятых совокупностях уменьшится, если $\mathfrak{F}$ с обнаруженным дефектным объектом будут приняты. Конечно, есть много других априорных распределений $F_N(D)$, при которых планы с приемочным числом $c=0$ будут вполне эффективными.

Исследование свойств выборочного контроля совокупности $\mathfrak{F} \propto [N, D_\omega]$ в предположении существования априорных распределений $F_N(D)$ для D_ω начнем со следующей модели. Предположим, что (1) вероятность объекта $O_i \in \mathfrak{F}$ быть дефектным равна q , т. е. одинакова для всех объектов, (2) любые события, состоящие в том, что объекты с любыми заданными номерами $i_1, \dots, i_l$ дефектны, взаимно независимы. Таким образом, совокупности $\mathfrak{F} = \{O_1, \dots, O_N\}$ сопоставляется последовательность взаимно независимых случайных величин $(\varepsilon_1, \dots, \varepsilon_N)$, $\varepsilon_i = 1$ соответствует событию « O_i дефектен», $P\{\varepsilon_i = 1\} = q$. При таких предположениях $D_\omega = \sum_{1 \leq i \leq N} \varepsilon_i$ есть биномиально распределенная случайная величина

$$F_N(D) = B_N^D(q) = \sum_{D'=0}^N \binom{N}{D'} q^{D'} (1-q)^{N-D'}. \quad (3.1)$$

Теорема 2.3.1. Если априорное распределение D_ω является биномиальным с функцией распределения (3.1), то результаты выборочного контроля с применением случайной выборки без возвращения образуют последовательность Бернулли с вероятностью q появления дефектного объекта.

Доказательство. Пусть $(\varepsilon_1, \dots, \varepsilon_n)$ есть результаты контроля n последовательно случайным образом отобранных объектов. Случайные величины $(\varepsilon_1, \dots, \varepsilon_n)$ образуют последовательность Бернулли, если ε_i взаимно независимы и $P\{\varepsilon_i = 1\} = q$, т. е. $P\{\varepsilon_1 = e_1, \dots, \varepsilon_n = e_n\} = q^d \bar{q}^k$, где значения $e_i = 0, 1$, $\sum_{1 \leq i \leq n} e_i = d$, $k = n - d$. Если $D_\omega = D$, то согласно (2.6) вероятность результатов контроля $(e_1, \dots, e_n)$ равна

$\frac{K^{(k)}D^{(d)}}{N^{(n)}}$, $K = N - D$. По формуле полных вероятностей с учетом (3.1) имеем

$$\begin{aligned} P\{e_1 = e_1, \dots, e_n = e_n\} &= \\ &= \sum_{D=0}^N P\{e_1 = e_1, \dots, e_n = e_n | D_\omega = D\} P\{D_\omega = D\} = \\ &= \sum_{D=0}^N \frac{K^{(k)}D^{(d)}}{N^{(n)}} \binom{N}{D} q^D \bar{q}^K = \\ &= q^d \bar{q}^k \sum_{D=d}^{N-n} \binom{N-n}{D-d} q^{D-d} \bar{q}^{K-k} = q^d \bar{q}^k, \end{aligned}$$

что и требовалось доказать.

Как следствие теоремы 2.3.1, получаем, что если априорное распределение $F_N(D)$ является биномиальным, то число δ_n дефектных объектов, обнаруженных в выборке объема n , также является биномиально распределенной случайной величиной,

$$P\{\delta_n = d\} = b_d^n(q) = \binom{n}{d} q^d \bar{q}^k.$$

Априорное распределение D_ω может быть гипергеометрическим. Действительно, пусть имеется исходная совокупность $\mathfrak{P}_\Sigma \asymp [N_\Sigma, D_\Sigma]$, где D_Σ — некоторое целое число. Из совокупности $\mathfrak{P}_\Sigma$ отбираем случайным образом N объектов. В итоге выборка образует совокупность $\mathfrak{P} \asymp [N, D_\omega]$, в которой число дефектных объектов D_ω будет случайной величиной, имеющей гипергеометрическое распределение

$$f_N(D) = P\{D_\omega = D\} = h_{N, D_\Sigma}^{N, D}. \quad (3.2)$$

Теорема 2.3.2. Если априорное распределение D_ω в совокупности $\mathfrak{P} \asymp [N, D_\omega]$ есть гипергеометрическое с вероятностями (3.2), то результаты выборочного контроля с применением случайной выборки без возвращения из $\mathfrak{P}$ соответствуют результатам случайной выборки без возвращения из $\mathfrak{P}_\Sigma \asymp [N_\Sigma, D_\Sigma]$.

Доказательство. Пусть результаты случайной бесповторной выборки из $\mathfrak{P}$ есть $e_1 = e_1, \dots, e_n = e_n$,

$e_i = 0, 1, \sum_{1 \leq i \leq n} e_i = d$. Тогда из (2.6), (3.2) и тождества $\binom{N_\Sigma}{D_\Sigma} = \binom{N_\Sigma - n}{D_\Sigma - d} N_\Sigma^{(n)} / (D_\Sigma^{(d)} K_\Sigma^{(k)})$ получаем

$$\begin{aligned} P\{\varepsilon_1 = e_1, \dots, \varepsilon_n = e_n\} &= \\ &= \sum_{D=0}^N \frac{D^{(d)} K^{(k)}}{N^{(n)}} h_{N_\Sigma, D_\Sigma}^{N, D} = \sum_{D=0}^n \binom{n}{d}^{-1} h_{N, D}^{n, d} h_{N_\Sigma, D_\Sigma}^{N, D} = \\ &= \sum_{D=d}^{N-n} h_{N_\Sigma - n, D_\Sigma - d}^{N-n, D-d} \frac{D_\Sigma^{(d)} K_\Sigma^{(k)}}{N_\Sigma^{(n)}} = \frac{D_\Sigma^{(d)} K_\Sigma^{(k)}}{N_\Sigma^{(n)}}. \end{aligned}$$

Полученное равенство в соответствии с (2.6) означает, что вероятность пути $\varepsilon_1 = e_1, \dots, \varepsilon_n = e_n$ при выборе из $\mathfrak{F}$ совпадает с вероятностью этого пути при случайной выборке без возвращения из $\mathfrak{F}_\Sigma$, что и требовалось доказать.

Следствие 2.3.1. Условное математическое ожидание доли $\frac{D_\omega - d}{N - n}$ дефектных объектов в непроконтролированной части совокупности при одноступенчатом плане контроля, случайной выборке без возвращения и априорном распределении (3.2) при условии, что обнаружено d дефектных объектов, равно

$$q_n(d) = E\left\{ \frac{D_\omega - d}{N - n} \mid \delta_n = d \right\} = \frac{D_\Sigma - d}{N_\Sigma - n}. \quad (3.3)$$

Если же априорное распределение биномиальное (3.1), то

$$q_n(d) = E\left\{ \frac{D_\omega - d}{N - n} \mid \delta_n = d \right\} = q. \quad (3.4)$$

Доказательство. Из теоремы 2.3.2 следует, что образование совокупности $\mathfrak{F}$ можно представить в два этапа. Сначала из $\mathfrak{F}_\Sigma$ отбирается выборка объема n , результаты которой есть $\varepsilon_1, \dots, \varepsilon_n, \sum_{1 \leq i \leq n} \varepsilon_i = d$, а затем из оставшейся совокупности $\mathfrak{F}'_\Sigma \times [N_\Sigma - n, D_\Sigma - d]$ добывается оставшая часть $\mathfrak{F}$. Здесь результаты $\varepsilon_{n+1}, \dots, \varepsilon_N$ соответствуют случайной выборке объема $N - n$ из $\mathfrak{F}'_\Sigma$,

$D_\omega - \delta_n = \sum_{n+1 \leq i \leq N} \varepsilon_i$, а потому $E \left\{ \sum_{n+1 \leq i \leq N} \varepsilon_i \mid \delta_n = d \right\} = (N-n)(D_\Sigma - d)/(N_\Sigma - n)$. Отсюда, для $E \left\{ \sum_{n+1 \leq i \leq N} \varepsilon_i \mid \delta_n = d \right\}$ на $N-n$, получаем (3.3). При априорном распределении (3.1) $\varepsilon_{n+1}, \dots, \varepsilon_N$ соответствуют схеме Бернулли с $P\{\varepsilon_i = 1\} = q$, а потому $E \left(\sum_{n+1 \leq i \leq N} \varepsilon_i \mid \delta_n = d \right) = E \left(\sum_{n+1 \leq i \leq N} \varepsilon_i \right) = (N-n)q$, разделив которое на $N-n$, получаем (3.4).

Из теорем 2.3.1 и 2.3.2, в частности, следует, что априорные биномиальное и гипергеометрическое распределения снова приводят к биномиальному и гипергеометрическому распределениям для числа дефектных объектов, обнаруженных в выборке при одноступенчатом плане контроля. В терминологии Хальда [2] это свойство сохранения типа распределения называется **воспроизводимостью**. В этой же работе Хальд [2] показал, что свойство воспроизводимости в некотором смысле является характеристическим для следующего класса априорных биномиальных распределений со случайным значением вероятности q появления дефектных объектов.

Определение 2.3.1. *Априорное распределение D_ω называется биномиальным со случайным значением параметра q , если для некоторой функции распределения $G(q)$, $0 \leq q \leq 1$,*

$$f_N(D) = \int_{q \in [0, 1]} b_N^D(q) dG(q). \quad (3.5)$$

Кратко такие распределения будем называть ***G-биномиальными***. Итак, для G -биномиального распределения вероятность появления дефектного объекта является случайной величиной q_ω , $P\{q_\omega \leq q\} = G(q)$. Класс всех возможных G -биномиальных распределений обозначим через $\mathcal{B}$. Таким образом, запись $F_N \in \mathcal{B}$ означает, что априорное распределение $F_N(D)$ является биномиальным со случайным значением параметра, т. е. определяется по формуле (3.5).

Следствие 2.3.2. *Если априорное распределение является G -биномиальным, то вероятность результатов*

контроля $\varepsilon_1 = e_1, \dots, \varepsilon_n = e_n$, $\sum_{1 \leq i \leq n} e_i = d$ при случайной бесповторной выборке равна

$$P\{\varepsilon_1 = e_1, \dots, \varepsilon_n = e_n\} = \int_{q \in [0, 1]} q^d \bar{q}^k dG(q). \quad (3.6)$$

Доказательство. Вероятность события $\{\varepsilon_1 = e_1, \dots, \varepsilon_n = e_n\}$ при $q_\omega = q$ по теореме 2.3.1 равна $q^d \bar{q}^k$. Соотношение (3.6) получается усреднением по распределению q_ω .

Свойство воспроизводимости для априорных распределений $F_N \in \mathcal{B}$ легко получаем из (3.5) и простого соотношения

$$h_{N,D}^{n,d} \binom{N}{D} = \binom{n}{d} \binom{N-n}{D-d}.$$

Здесь имеем

$$\begin{aligned} P\{\delta_n = d\} &= \sum_{D=0}^N h_{N,D}^{n,d} \int_{q \in [0, 1]} \binom{N}{D} q^D \bar{q}^{N-D} dG(q) = \\ &= \binom{n}{d} \int_{q \in [0, 1]} q^d \bar{q}^{n-d} \left(\sum_{D=0}^N \binom{N-n}{D-d} q^{D-d} \bar{q}^{N-D} \right) dG(q) = \\ &= \binom{n}{d} \int_{q \in [0, 1]} q^d \bar{q}^{n-d} dG(q). \quad (3.7) \end{aligned}$$

Таким образом, распределение δ_n — числа дефектных объектов в выборке объема n из $\mathfrak{B}$ с априорным G -биномиальным распределением снова есть G -биномиальное распределение, определяемое (3.7).

Доказательство того, что из воспроизводимости и некоторых дополнительных требований при $N \rightarrow \infty$ следует принадлежность априорного распределения к классу $\mathcal{B}$, существенно сложнее. Это доказательство приведено в уже отмечавшейся статье Хальда [2].

Схема Пойа. Рассмотрим G -биномиальное априорное распределение, когда априорное распределение имеет плотность

$$g(q) = \frac{dG(q)}{dq} = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} q^{a-1} \bar{q}^{b-1}, \quad a > 0, \quad b > 0, \quad (3.8)$$

т. е. когда $G(\lambda)$ есть бэта-распределение. Здесь $\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt$ — гамма-функция. В этом случае имеем

$$\begin{aligned} f_N(D) &= \binom{N}{D} \int_{q \in [0, 1]} q^D \bar{q}^K \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} q^{a-1} \bar{q}^{b-1} dq = \\ &= \binom{N}{D} \frac{\Gamma(D+a)}{\Gamma(a)} \cdot \frac{\Gamma(K+b)}{\Gamma(b)} \cdot \frac{\Gamma(a+b)}{\Gamma(N+a+b)}. \end{aligned} \quad (3.9)$$

Для любого a и целого $x \geq 1$ будем использовать сокращенное обозначение

$$\prod_{k=0}^{x-1} (a + kc) = a^{(x, c)}. \quad (3.10)$$

Из рекуррентного соотношения $\Gamma(y+1) = y\Gamma(y)$ получаем, что $\Gamma(a+b)/\Gamma(a) = a^{(b, 1)}$. Поэтому (3.9) в обозначениях (3.10) можно записать в виде

$$f_N(D) = \binom{N}{D} \frac{a^{(D, 1)} b^{(K, 1)}}{(a+b)^{(N, 1)}}. \quad (3.11)$$

Распределение D_∞ с вероятностями (3.11) называется **распределением Пойа**. В общем случае распределение Пойа возникает в следующей математической модели, называемой **схемой Пойа**. Пусть в процессе производства объекты выпускаются последовательно один за другим. Если уже произведено n объектов, среди которых k годных и d дефектных ($k+d=n$), то вероятности произвести $(n+1)$ -й объект годным или дефектным равны соответственно

$$q_{n+1} = \frac{q + dg}{1 + ng}, \quad \bar{q}_{n+1} = \frac{\bar{q} + kg}{1 + ng}, \quad (3.12)$$

где $0 \leq q \leq 1$, $\bar{q} = 1 - q$, $\min(q + dg, \bar{q} + kg) \geq 0$. Таким образом, для $g > 0$ в этой математической модели при «производстве» дефектного объекта вероятность последующего появления дефектного объекта возрастает, тогда как при производстве годного объекта она уменьшается. Пусть, как и ранее, $\varepsilon_i = 1$ соответствует

дефектности i -го объекта, а $e_i = 0$ — его годности. В рассматриваемой модели Пойа

$$\begin{aligned} P\{e_1 = e_1, \dots, e_N = e_N\} = \\ = \frac{q(q+g) \dots \left(q + \left(\sum_{i=1}^N e_i - 1\right)g\right)}{(1+g) \dots (1+(N-1)g)} \times \\ \times \bar{q}(\bar{q}+g) \dots \left(\bar{q} + \left(N - \sum_{i=1}^N e_i - 1\right)g\right) = \frac{q^{(D, g)} \bar{q}^{(K, g)}}{1^{(N, g)}}. \end{aligned} \quad (3.13)$$

Предположим, что $g > 0$. Сделаем в (3.13) замену переменных, положив $g = (a+b)^{-1}$, $q = a(a+b)^{-1}$, $\bar{q} = b(a+b)^{-1}$, соответственно $a = qg^{-1}$, $b = \bar{q}g^{-1}$. В итоге (3.13) переписывается в виде

$$P\{e_1 = e_1, \dots, e_N = e_N\} = \frac{a^{(D, 1)} b^{(K, 1)}}{(a+b)^{(N, 1)}}, \quad (3.14)$$

где $D = \sum_{1 \leq i \leq N} e_i$, $K = N - D$. Для получения вероятности $f_N(D)$ того, что общее число произведенных дефектных объектов $D_\omega = D$, достаточно заметить, что число исходов $(e_1, \dots, e_N)$, у которых $\sum_{1 \leq i \leq N} e_i = D$, равно $\binom{N}{D}$.

Таким образом, умножая вероятность в (3.14) на $\binom{N}{D}$ получаем, что $f_N(D)$ определяется формулой (3.11). Следовательно, справедливо утверждение

Теорема 2.3.3. При $g > 0$ распределение числа D_ω дефектных объектов в схеме Пойа есть G -биномиальное распределение, в котором $G(\lambda)$ является бэтараспределением с плотностью вероятностей (3.8) и параметрами $a = qg^{-1}$, $b = \bar{q}g^{-1}$.

Схему Пойа можно также интерпретировать и как процесс последовательного отбора объектов из совокупности объема N , при котором на каждом шаге вероятности отбора для всех оставшихся невыбранными объектов одинаковы. Если отобранный из совокупности объект оказывается годным (дефектным), то к совокупности добавляется $s = gN$ годных (дефектных) объектов. Здесь мы, конечно, предполагаем, что $s = gN$ есть

целое число, а $q = D/N$, где D — число дефектных объектов в совокупности до начала контроля. В этих предположениях общее число обнаруженных дефектных объектов в выборке объема n равно $d = 0, \dots, n$ с вероятностями

$$f_n(d) = \binom{n}{d} \frac{D^{(d, c)} K^{(k, c)}}{N^{(n, c)}}.$$

Схема Пойа при $c = -1$ ($g = -1/N$) соответствует случайной выборке без возвращения. Если же $c = 0$, то схема Пойа соответствует выборке с возвращением.

Заметим, что при $a = b = 1$ из (3.11) получаем, что

$$f_N(D) = 1/(N+1), \quad 0 \leq D \leq N. \quad (3.15)$$

Таким образом, равномерное априорное распределение также соответствует частному случаю схемы Пойа. При этом $g(q) = 1$, т. е. равновероятному распределению D_ω соответствует G -биномиальное априорное распределение, у которого параметр q равномерно распределен на интервале $[0, 1]$.

Предельная теорема для смеси биномиальных распределений. Рассмотрим следующую общую схему задания вероятностей априорных распределений $f_N(D)$ для D_ω . Пусть $C_{[0, 1]}^+$ — класс всех непрерывных, неотрицательных функций $f(q)$, $0 \leq q \leq 1$, таких, что $\int_{q \in [0, 1]} f(q) dq = 1$.

Для любой функции $f(q)$ из класса $C_{[0, 1]}^+$ ($f(q) \in C_{[0, 1]}^+$) положим

$$f_N(D) = (1/N) f(D/N) + \varepsilon_{D, N}, \quad (3.16)$$

где поправочные числа $\varepsilon_{D, N}$ введены для того, чтобы $\sum_{0 \leq D \leq N} f_N(D) = 1$. Так как при $N \rightarrow \infty$ имеет место сходимость $\sum_{0 \leq D \leq N} (1/N) f(D/N) \rightarrow \int_{q \in [0, 1]} f(q) dq$, а $\sum_{0 \leq D \leq N} \varepsilon_{D, N} = 1 - (1/N) \sum_{0 \leq D \leq N} f(D/N) \rightarrow 0$, то можно считать, что $N \varepsilon_{D, N} \rightarrow 0$,

Как показывает формулируемая ниже локальная предельная теорема, только что описанная схема задания априорных вероятностей может быть реализована при $N \rightarrow \infty$ с помощью априорных биномиальных распределений со случайным значением параметра q .

Теорема 2.3.4. Пусть $f(q) \in C_{[0, 1]}^+$. Тогда при $N \rightarrow \infty$ и $(D/N) \rightarrow q (q \neq 0, 1)$ имеем

$$f_N(D) = \int_{q \in [0, 1]} \binom{N}{D} q^D \bar{q}^K f(q) dq = (1/N) f(D/N) + o(1/N). \quad (3.17)$$

Доказательство. Слагаемое в (3.17), обозначенное $o(1/N)$, имеет при $N \rightarrow \infty$ порядок малости более высокий, чем $1/N$, и поэтому становится сколь угодно малым по сравнению с первым слагаемым $(1/N)f(D/N)$, если только $f(q) \neq 0$. Интегральное представление для $f_N(D)$ в (3.17) следует из (3.5) при $dG(q)/dq = f(q)$. Под знаком интеграла сделаем замену переменных $v = (q - D/N)\sqrt{N}$. В результате получаем

$$f_N(D) = (1/N) \int_{-\sqrt{N} D/N}^{(1-D/N)\sqrt{N}} H_N(v) f((D/N) + v/\sqrt{N}) dv, \quad (3.18)$$

где

$$H_N(v) = \sqrt{N} \binom{N}{D} ((D/N) + v/\sqrt{N})^D ((K/N) - v/\sqrt{N})^K. \quad (3.19)$$

Функция $H_N(v)$ имеет один максимум в точке $v=0$. В этом легко убедиться простым дифференцированием. Пусть $v_e = v_e(N) > 0$ такое, что $v_e/\sqrt{N} \rightarrow 0$, $N \rightarrow \infty$. Тогда для $|v| \leq v_e$, используя формулу Стирлинга $n! \sim \sqrt{2\pi} n^{n+1/2} e^{-n}$ и разложение $\ln(1+x) = x - (x^2/2) + (x^3/3) - \dots$, находим, что

$$\ln H_N(v) = -\frac{v^2}{2q_1 \bar{q}_1} - \ln(\sqrt{2\pi q_1 \bar{q}_1}) + o(v/\sqrt{N}), \quad (3.20)$$

где $q_1 = D/N$, $\bar{q}_1 = K/N$. Таким образом, при достаточно большом числе c_e и $v_e = c_e \sqrt{\ln N}$ функция $H_N(v)$ в (3.19) на интервале $[-c_e \sqrt{\ln N}, c_e \sqrt{\ln N}]$ эквивалентна

$(2\pi q_1 \bar{q}_1)^{-1/2} \exp \{-v^2/(2q_1 \bar{q}_1)\}$, а вне этого интервала стремится к нулю. Отсюда, разбивая в (3.18) область интегрирования на три части, находим, что

$$\begin{aligned} f_N(D) &= \\ &= (1/N) \left(\int_{-v_e}^{v_e} + \int_{-\sqrt{N} D/N}^{-v_e} + \int_{v_e}^{(1-D/N)\sqrt{N}} \right) H_N(v) f((D/N) + \\ &+ v/\sqrt{N}) dv = (1/N) \int_{-v_e}^{v_e} (2\pi q_1 \bar{q}_1)^{-1/2} \exp \{-v^2/2q_1 \bar{q}_1\} f((D/N) + \\ &+ v/\sqrt{N}) dv + o(1/N) = \\ &= (1/N) f(D/N) \left[\frac{1}{\sqrt{2\pi}} \int_{-(v_e/\sqrt{q_1 \bar{q}_1})}^{+(v_e/\sqrt{q_1 \bar{q}_1})} e^{-x^2/2} dx \right] + o(1/N) = \\ &= (1/N) f(D/N) + o(1/N), \end{aligned}$$

что и требовалось доказать. Заметим, что это доказательство отличается от данного Хальдом в [2].

Таким образом, если априорное распределение для D_ω имеет вероятности $f_N(D) \approx (1/N) f(D/N)$, где $f(q)$ — непрерывная неотрицательная функция, $\int_{q \in [0, 1]} f(q) dq = 1$,

то приближенно можно считать, что это априорное распределение есть биномиальное со случайным параметром q , имеющим плотность вероятностей $f(q)$. Следовательно, класс В биномиальных распределений со случайным параметром является весьма широким классом априорных распределений. Поэтому было бы полезно решить вопросы, связанные с оценками точности аппроксимации различных априорных распределений распределениями класса В. В этом отношении теорема 2.3.4 может рассматриваться лишь как самый начальный результат.

Поскольку мы рассматриваем случай, когда объем совокупности неограниченно возрастает ($N \rightarrow \infty$), то посмотрим, как при этом растет дисперсия σ_F^2 для числа D_ω дефектных объектов в совокупности $\mathfrak{F}$. Так как $\mathfrak{F}$

характеризуется последовательностью $(\varepsilon_1, \dots, \varepsilon_N)$, где ε_i равно 0 или 1, когда O_i годен или дефектен соответственно, то для вычисления дисперсии достаточно знать $q_i = P(\varepsilon_i = 1)$ и $q_{i,j} = P(\varepsilon_i = \varepsilon_j = 1)$, $i, j = 1, \dots, N$. В этом случае имеем

$$ED_\omega = \sum_{i=1}^N q_i, \quad ED_\omega(D_\omega - 1) = \sum_{i \neq j} q_{i,j}. \quad (3.21)$$

Если априорное распределение является G -биномиальным, то $q_i = q_G = \int_{q \in [0, 1]} q dG(q)$, $q_{i,j} = \int_{q \in [0, 1]} q^2 dG(q)$. Следовательно, из (3.21) находим, что дисперсия

$$\begin{aligned} \sigma_F^2 &= ED_\omega^2 - (ED_\omega)^2 = \sum_{i \neq j} q_{i,j} + \sum_{1 \leq i \leq N} q_i - \left(\sum_{1 \leq i \leq N} q_i \right)^2 = \\ &= N(N-1) \int_{q \in [0, 1]} (q - q_G)^2 dG(q) + Nq_G \bar{q}_G. \end{aligned} \quad (3.22)$$

Таким образом, если априорное распределение является G -биномиальным, то дисперсия числа дефектных объектов в совокупности растет пропорционально квадрату объема совокупности. Исключением является случай «чисто» биномиального априорного распределения, когда при некотором q_0 ф. р. $G(q) = 0$ для $q < q_0$ и $G(q) = 1$ для $q \geq q_0$. В этом случае $q_G = q_0$,

$\int_{q \in [0, 1]} (q - q_G)^2 dG(q) = 0$, и дисперсия σ_F^2 растет линейно

с ростом N . Заметим также, что если $q_{i,j}^F = q_G^2 + r_{i,j}$, где $r_{i,j} \rightarrow 0$ при $|i - j| \rightarrow 0$, то дисперсия σ_F^2 , вообще говоря, растет медленнее N^2 . Эти замечания следует иметь в виду при выборе класса априорных распределений для числа дефектных объектов в контролируемых совокупностях.

Моментные характеристики. Совокупность N объектов, среди которых D_ω дефектных, причем D_ω — случайная величина, для которой $F_N(D) = P\{D_\omega \leq D\} = \sum_{D_0 \leq D} f_N(D_0)$,

кратко обозначим $\mathfrak{P}_{N,F}(\mathfrak{P}_{N,F} \asymp [N, D_\omega])$. Пусть, как и раньше, δ_n равно числу дефектных объектов, обнаруженных в случайной выборке объема n . Используя формулу (1.3.14) для смешанных факториальных моментов, получим выражение для дисперсии δ_n и ковариации δ_n с числом дефектных объектов $D_\omega - \delta_n$, не попавших в выборку и, следовательно, не обнаруженных при выборочном контроле.

Пусть $q_{N,F}$ — среднее значение доли D_ω/N дефектных объектов в $\mathfrak{P}_{N,F}$, а $\sigma_{N,F}^2$ — дисперсия D_ω ,

$$q_{N,F} = E(D_\omega/N),$$

$$\sigma_{N,F}^2 = ED_\omega^2 - (ED_\omega)^2.$$

Усредняя (1.3.14) по D_ω , получаем, что

$$E\delta_n^{(r)}(D_\omega - \delta_n)^{(s)} = \frac{n^{(r)}(N-n)^{(s)}}{N^{(r+s)}} ED_\omega^{(r+s)}. \quad (3.23)$$

Для $r, s = 0, 1, 2$ из (3.23) находим

$$E\delta_n = nq_{N,F}, \quad E(D_\omega - \delta_n) = (N-n)q_{N,F}, \quad (3.24)$$

$$\begin{aligned} E\delta_n(D_\omega - \delta_n) &= \frac{n(N-n)}{N(N-1)} ED_\omega^{(2)} = \\ &= \frac{n(N-n)}{N(N-1)} [\sigma_{N,F}^2 + (Nq_{N,F})^2 - Nq_{N,F}], \end{aligned} \quad (3.25)$$

$$E\delta_n(\delta_n - 1) = \frac{n(n-1)}{N(N-1)} [\sigma_{N,F}^2 + (Nq_{N,F})^2 - Nq_{N,F}], \quad (3.26)$$

$$\begin{aligned} E(D_\omega - \delta_n)(D_\omega - \delta_n - 1) &= \\ &= \frac{(N-n)(N-n-1)}{N(N-1)} [\sigma_{N,F}^2 + (Nq_{N,F})^2 - Nq_{N,F}]. \end{aligned} \quad (3.27)$$

Из (3.24) — (3.27) путем элементарных преобразований получаем выражение для $\sigma_{n,F}^2$ — дисперсии δ_n и $c_{n,F} = E\delta_n(D_\omega - \delta_n) - E\delta_n E(D_\omega - \delta_n)$, т. е. ковариации чисел обнаруженных и не обнаруженных дефектных объектов.

Здесь имеем

$$\begin{aligned}\sigma_{n,F}^2 &= E\delta_n^2 - (E\delta_n)^2 = E\delta_n^{(2)} + E\delta_n - (E\delta_n)^2 = \\ &= \frac{n(n-1)}{N(N-1)} [\sigma_{N,F}^2 + (Nq_{N,F})^2 - Nq_{N,F}] + nq_{N,F} - (nq_{N,F})^2 = \\ &= \frac{nq_{N,F}\bar{q}_{N,F}(N-n)}{N-1} \left[1 + \frac{n-1}{N-n} \cdot \frac{\sigma_{N,F}^2}{Nq_{N,F}\bar{q}_{N,F}} \right], \quad (3.28)\end{aligned}$$

$$\begin{aligned}c_{n,F} &= \frac{n(N-n)}{N(N-1)} [\sigma_{N,F}^2 + (Nq_{N,F})^2 - Nq_{N,F}] - n(N-n)q_{N,F}^2 = \\ &= \frac{n(N-n)}{N(N-1)} [\sigma_{N,F}^2 - Nq_{N,F}\bar{q}_{N,F}], \quad (3.29)\end{aligned}$$

$$\begin{aligned}\bar{\sigma}_{n,F}^2 &= E(D_\omega - \delta_n)^2 - \{E(D_\omega - \delta_n)\}^2 = \\ &= \frac{(N-n)(N-n-1)}{N(N-1)} [\sigma_{N,F}^2 + (Nq_{N,F})^2 - Nq_{N,F}] + \\ &\quad + (N-n)q_{N,F} - (N-n)^2q_{N,F}^2 = \\ &= (N-n)q_{N,F}\bar{q}_{N,F} \frac{n}{N-1} \left[1 + \frac{N-n-1}{n} \frac{\sigma_{N,F}^2}{Nq_{N,F} - \bar{q}_{N,F}} \right]. \quad (3.30)\end{aligned}$$

Соотношения (3.28) — (3.30) записаны таким образом, что они дают представление об изменениях этих характеристик при отклонениях априорного распределения F от биномиального (3.1), где $q = q_F$. Если априорное распределение F есть биномиальное с параметром q , то $q = q_{N,F}$. В этом случае из (3.28) — (3.30) находим, что $\sigma_{n,F}^2 = nq\bar{q}$, $\bar{\sigma}_{n,F}^2 = (N-n)q\bar{q}$, $c_{n,F} = 0$.

Определение 2.3.2. Априорное распределение F для числа дефектных объектов в совокупности назовем **суббиномиальным (гипербиномиальным)***, если $\sigma_{N,F}^2$ — дисперсия числа дефектных объектов — меньше (больше) дисперсии биномиального априорного распределения с одинаковым значением средней доли $q_{N,F}$ дефектных объектов. Если дисперсии совпадают, но априорное распределение отличается от биномиального, то такое распределение назовем **эквибиномиальным**.

*) Хальд [2] использует термин **субнормальное (гипернормальное)** распределение.

Введем коэффициент $\Delta_{N,F}$ по формуле

$$\Delta_{N,F} = \frac{\sigma_{N,F}^2}{Nq_{N,F}\bar{q}_{N,F}} - 1. \quad (3.31)$$

Таким образом при суббиномиальном априорном распределении $\Delta_{N,F} < 0$, при гипербиномиальном $\Delta_{N,F} > 0$, при эквибиномиальном априорном распределении $\Delta_{N,F} = 0$. Формулы (3.28)–(3.30) с использованием коэффициента $\Delta_{N,F}$ выражаются следующим образом:

$$\begin{cases} \sigma_{n,F}^2 = nq_{N,F}\bar{q}_{N,F} \left(1 + \frac{n-1}{N-1} \Delta_{N,F} \right), \\ c_{n,F} = nq_{N,F}\bar{q}_{N,F} \frac{N-n}{N-1} \Delta_{N,F}, \\ \bar{\sigma}_{n,F}^2 = (N-n)q_{N,F}\bar{q}_{N,F} \left(1 + \frac{N-n-1}{N-1} \Delta_{N,F} \right). \end{cases} \quad (3.32)$$

Во введенной терминологии из (3.32) следует, что свойство суббиномиальности, гипербиномиальности или эквибиномиальности воспроизводится для распределений чисел δ_n и $D_\omega - \delta_n$. Если априорное распределение суббиномиальное, то числа обнаруженных дефектных объектов δ_n и не обнаруженных ($D_\omega - \delta_n$) — отрицательно коррелированы. Положительная корреляция будет только при гипербиномиальном априорном распределении. При эквибиномиальном априорном распределении корреляция равна нулю. Эти результаты были получены Мудом [1], они приведены в статье Хальда [2].

Отметим, что рассматривавшееся выше гипергеометрическое априорное распределение (3.2) является суббиномиальным. Для этого распределения $\Delta_{N,F} = -(N-1)/(N_\Sigma - 1)$. Во введенной выше схеме Пойа распределение числа дефектных объектов будет гипербиномиальным при $g > 0$ и суббиномиальным при $g < 0$. Как это следует из (3.22), G -биномиальное априорное распределение является гипербиномиальным, если только оно не совпадает с биномиальным распределением.

§ 2.4. Байесовский подход

Средняя доля дефектных объектов в непроконтролированной части совокупности. При байесовском подходе к решению задач статистического контроля качества

предполагается, что неизвестные параметры сами являются случайными величинами. В частном случае контроля по альтернативному признаку предполагается, что число дефектных объектов в контролируемой совокупности $\mathfrak{F}$ является случайной величиной D_ω с известной (или неизвестной) функцией распределения $F_N(D)$. На основе результатов § 2.3 легко сделать вывод, что вид априорного распределения $F_N(D)$ самым существенным образом влияет на показатели эффективности применения любой системы контроля.

Начнем с простейшего вопроса о выборе решения при контроле по плану первого вхождения Π с границей Γ . Пусть в результате выборочного контроля путь контроля приводит в граничную точку $x = (k, d) \in \Gamma$. Какое решение следует принять о контролируемой совокупности $\mathfrak{F}$? Следует забраковать ее или признать приемлемой? Если априорное распределение $F_N(D)$ задано, то ответ можно было бы получить следующим образом. Нужно вычислить для всех $x \in \Gamma$ значения $q(x) = E \left\{ \left(\frac{D_\omega - d}{N - (k + d)} \right) \middle| \delta_n = d \right\}$ — средней доли дефектных объектов в непроконтролированной части $\mathfrak{F}$. Затем выбрать некоторый уровень $q_{кр}$ и для всех $x \in \Gamma$ таких, что $q(x) < q_{кр}$, принимать решение о приемке $\mathfrak{F}$, а для x таких, что $q(x) \geq q_{кр}$, принимать решение о браковке. Существенное осложнение заключается в том, что $F_N(D)$ нам неизвестна. Более того, в условиях производства изделий массовой промышленной продукции вид функции $F_N(D)$ может изменяться с течением времени. Однако можно взять сначала какое-то стандартное правило принятия решения и затем по результатам контроля большого числа партий убедиться, вычисляя последующие оценки для предъявленного и пропущенного брака, в степени эффективности принятого правила.

Сформулируем поставленную задачу до класса В априорных распределений. Вопрос стоит, в частности, так. Целесообразно ли браковать совокупность, если при контроле по одноступенчатому плану число обнаруженных дефектных объектов в выборке превысило некоторый приемочный уровень, т. е. целесообразно ли использовать планы типа (n, c) ? В интересной работе Большева

[5] приведены примеры, когда использование планов (n, c) явно нецелесообразно. Однако наш вывод будет положительным, так как мы исходим из предположения, что априорное распределение D_ω является G -биномиальным. Как уже было показано в § 2.3, класс В является довольно общим, хотя, конечно, не включает в себя многие априорные распределения (например, распределения в схеме Пойа (3.13) при $g < 0$).

Теорема 2.4.1. Если априорное распределение D_ω есть G -биномиальное, то для точек $x_0 = (k, d)$, $x_1 = (k, d + 1)$, $x_2 = (k + 1, d)$ выполнены неравенства

$$q(x_2) \leq q(x_0) \leq q(x_1). \quad (4.1)$$

Доказательство. Так как априорное распределение есть G -биномиальное, то вероятность к моменту проверки n -го объекта ($k + d = n$) оказаться в точке $x_0 = (k, d)$ в соответствии с теоремой 2.3.1 будет равна

$$P(\delta_n = d) = \int_0^1 L(x_0) q^d \bar{q}^k dG(q), \quad (4.2)$$

где $L(x)$ — число путей, соединяющих $(0, 0)$ с x_0 . Аналогично, находим, что

$$P(D_\omega = D, \delta_n = d) = \int_0^1 L(x_0) \binom{N-d}{D-d} q^D \bar{q}^K dG(q). \quad (4.3)$$

Таким образом, из (4.2) и (4.3) получаем

$$q(x_0) = \sum_{D=d}^{N-n} \frac{(D-d) P(D_\omega = D, \delta_n = d)}{(N-n) P(\delta_n = d)} = \frac{\int_0^1 q^{d+1} \bar{q}^k dG(q)}{\int_0^1 q^d \bar{q}^k dG(q)}. \quad (4.4)$$

Выражение для $q(x_1)$ получается заменой в (4.4) d на $d + 1$, а $q(x_2)$ — заменой в (4.4) k на $k + 1$. Вывод правого и левого неравенства проводится одинаковым методом, поэтому мы ограничимся выводом только левого

неравенства $q(x_2) \leq q(x_0)$, которое с учетом (4.4) эквивалентно неравенству

$$\frac{\int_0^1 q^{d+1} \bar{q}^{k+1} dG(q)}{\int_0^1 q^d \bar{q}^{k+1} dG(q)} \leq \frac{\int_0^1 q^{d+1} \bar{q}^k dG(q)}{\int_0^1 q^d \bar{q}^k dG(q)}. \quad (4.5)$$

Для доказательства неравенства (4.5) достаточно убедиться в неотрицательности разности

$$\begin{aligned} J &= \int_0^1 q_1^{d+1} \bar{q}_1^k dG(q_1) \int_0^1 q_2^d \bar{q}_2^{k+1} dG(q_2) - \\ &\quad - \int_0^1 q_1^d \bar{q}_1^k dG(q_1) \int_0^1 q_2^{d+1} \bar{q}_2^{k+1} dG(q_2) = \\ &= \int_0^1 \int_0^1 q_1^d \bar{q}_1^k q_2^d \bar{q}_2^k (q_1 \bar{q}_2 - q_2 \bar{q}_1) dG(q_1) dG(q_2). \end{aligned} \quad (4.6)$$

Функция $q_1^d \bar{q}_1^k q_2^d \bar{q}_2^k$ в (4.6) симметрична и не меняется при замене (q_1, q_2) на (q_2, q_1) . Следовательно, представляя интеграл J в виде суммы двух интегралов по областям $q_1 \geq q_2$ и $q_2 \geq q_1$ и замечая, что при $q_1 = q_2$ интеграл $J = 0$, получим при замене во втором интеграле (q_1, q_2) на (q_2, q_1) , что

$$\begin{aligned} J &= \int_{q_1 \geq q_2} \int_{q_1 \geq q_2} q_1^d \bar{q}_1^k q_2^d \bar{q}_2^k (q_1 \bar{q}_2 - q_2 \bar{q}_1 + q_2 \bar{q}_1 - q_1 \bar{q}_1) dG(q_1) dG(q_2) = \\ &= \int_{q_1 \geq q_2} \int_{q_1 \geq q_2} q_1^d \bar{q}_1^k q_2^d \bar{q}_2^k (q_1 - q_2)^2 dG(q_1) dG(q_2) \geq 0. \end{aligned}$$

Таким образом, (4.5), а следовательно, и левое неравенство в (4.1) выполнены. Правое неравенство (4.1) получается аналогично. Теорема доказана.

Следствие 2.4.1. Если априорное распределение G -биномиальное, то для $x_1 = (k_1, d_1)$, $x_2 = (k_2, d_2)$, $k_1 + d_1 = k_2 + d_2$ и $d_1 < d_2$, выполняется неравенство

$$q(x_1) \leq q(x_2). \quad (4.7)$$

Доказательство. Неравенство (4.7) непосредственно следует из (4.1), если $d_2 = d_1 + 1$. Если же $d_2 = d_1 + l$, $l > 1$, то (4.7) получаем как результат последовательных неравенств для точек $x = (k, d)$ со значениями d , отличающимися на 1. Здесь мы, конечно, предполагаем, что промежуточные все точки с индексом $k_1 + d_1$ достижимы.

Для скачкообразных функций $G(q)$ неравенство (4.7) было доказано Хальдом [2].

Итак, если априорное распределение является биномиальным со случайным значением вероятности появления дефектных объектов, то среднее значение доли дефектных объектов в непроконтролированной части совокупности тем больше, чем больше число дефектных объектов, обнаруженных в выборке. Аналогичное утверждение справедливо и для широко используемых двухступенчатых и многоступенчатых планов.

Одноступенчатые планы контроля. Исследование свойств $q(x)$ дает возможность построить систему последующих оценок, которые можно рассматривать как развитие эмпирического байесовского подхода в задачах статистического контроля качества. Существенные результаты в этом направлении для планов одноступенчатого типа были получены Большевым [5].

Начнем с вывода тождеств для факториальных моментов доли дефектных объектов $(D_\omega - \delta_n)/(N - n)$ в оставшейся непроконтролированной части совокупности, когда контроль проводится по одноступенчатому плану объема n . Пусть априорные вероятности равны $f_N(D)$. Тогда имеем

$$\begin{aligned} q_n^F(d, l) &= E \left\{ \frac{(D_\omega - \delta_n)^{(l)}}{(N - n)^{(l)}} \mid \delta_n = d \right\} = \\ &= \frac{1}{p_n^F(d)} \sum_{D=0}^N \frac{(D - d)^{(l)}}{(N - n)^{(l)}} h_{N, D}^{n, d} f_N(D), \quad (4.8) \end{aligned}$$

где

$$p_n^F(d) = P \{ \delta_n = d \} = \sum_{D=0}^N h_{N, D}^{n, d} f_N(D) \quad (4.9)$$

есть вероятность обнаружить d дефектных объектов в выборке объема n . Для слагаемых в (4.8) можем

записать

$$\begin{aligned} \frac{(D-d)^{(l)}}{(N-n)^{(l)}} h_{N,D}^{n,d} &= \frac{(D-d)^{(l)}}{(N-n)^{(l)}} \binom{n}{d} \frac{D^{(d)} K^{(k)}}{N^{(n)}} = \\ &= \frac{(d+l)^{(l)}}{(n+l)^{(l)}} \binom{n+l}{d+l} \frac{D^{(d+l)} K^{(k)}}{N^{(n+l)}} = \frac{(d+l)^{(l)}}{(n+l)^{(l)}} h_{N,D}^{n+l,d+l}. \end{aligned} \quad (4.10)$$

Из (4.8), учитывая (4.10), находим, что

$$\begin{aligned} q_n^F(d, l) &= \frac{1}{p_n^F(d)} \sum_{D=0}^N \frac{(d+l)^{(l)}}{(n+l)^{(l)}} h_{N,D}^{n+l,d+l} f_N(D) = \\ &= \frac{(d+l)^{(l)} \mathbf{P}\{\delta_{n+l}=d+l\}}{(n+l)^{(l)} p_n^F(d)} = \frac{(d+l)^{(l)} p_{n+l}^F(d+l)}{(n+l)^{(l)} p_n^F(d)}. \end{aligned} \quad (4.11)$$

Полезно получить еще одно представление для $q_n^F(d, l)$. Введем обозначения $p_n^F(d, e_1, \dots, e_l) = \mathbf{P}\{\delta_n = d, \varepsilon_{n+1} = e_1, \dots, \varepsilon_{n+l} = e_l\}$. Здесь событие $\{\delta_n = d, \varepsilon_{n+1} = e_1, \dots, \varepsilon_{n+l} = e_l\}$ означает, что в выборке объема n из совокупности $\mathfrak{B}_{N,F}$ было обнаружено d дефектных объектов, после чего было отобрано еще l объектов, результаты контроля которых дали значения $e_1, \dots, e_l, e_i = 0, 1, i = 1, \dots, l$. Замечая теперь, что $(D-d)^{(l)} / (N-n)^{(l)}$ есть вероятность того, что $(n+1)$ -й, $(n+2)$ -й, $\dots$, $(n+l)$ -й объекты окажутся дефектными, из (4.8) находим

$$\begin{aligned} q_n^F(d, l) &= \\ &= \frac{1}{p_n^F(d)} \sum_{D=0}^N \mathbf{P}(\varepsilon_{n+1} = 1, \dots, \varepsilon_{n+l} = 1 \mid \delta_n = d, D_\omega = D) \times \\ &\quad \times h_{N,D}^{n,d} f_N(D) = \frac{p_n^F(d, \overbrace{1, \dots, 1}^{l \text{ раз}})}{p_n^F(d)}. \end{aligned} \quad (4.12)$$

В частном случае, когда $l = 1, 2$, из (4.11) и (4.12) получаем

$$q_n^F(d) = q_n^F(d, 1) = \frac{d+1}{n+1} \frac{p_{n+1}^F(d+1)}{p_n^F(d)} = \frac{p_n^F(d, 1)}{p_n^F(d)}, \quad (4.13)$$

$$q_n^F(d, 2) = \frac{(d+1)(d+2)}{(n+1)(n+2)} \frac{p_{n+2}^F(d+2)}{p_n^F(d)} = \frac{p_n^F(d, 1, 1)}{p_n^F(d)}. \quad (4.14)$$

Используя формулы (4.13), (4.14), легко найти выражение для условной дисперсии $V_n^F(d)$ доли $(D_\omega - \delta_n)/(N - n)$ при условии, что $\delta_n = d$. Здесь имеем

$$\begin{aligned}
 V_n^F(d) &= \\
 &= E \left\{ \left(\frac{D_\omega - \delta_n}{N - n} \right)^2 \middle| \delta_n = d \right\} - \left(E \left\{ \frac{D_\omega - \delta_n}{N - n} \middle| \delta_n = d \right\} \right)^2 = \\
 &= \frac{N - n - 1}{N - n} E \left\{ \frac{(D_\omega - \delta_n)^{(2)}}{(N - n)^{(2)}} \middle| \delta_n = d \right\} + \\
 &+ \frac{1}{N - n} E \left\{ \frac{D_\omega - \delta_n}{N - n} \middle| \delta_n = d \right\} - \left(E \left\{ \frac{D_\omega - \delta_n}{N - n} \middle| \delta_n = d \right\} \right)^2 = \\
 &= \frac{N - n - 1}{N - n} \frac{(d + 1)(d + 2)}{(n + 1)(n + 2)} \frac{p_{n+2}^F(d + 2)}{p_n^F(d)} + \\
 &+ \frac{d + 1}{(N - n)(n + 1)} \frac{p_{n+1}^F(d + 1)}{p_n^F(d)} - \left(\frac{d + 1}{n + 1} \right)^2 \left(\frac{p_n^F(d + 1)}{p_n^F(d)} \right)^2 = \\
 &= \left(1 - \frac{1}{N - n} \right) \frac{p_n^F(d, 1, 1)}{p_n^F(d)} + \\
 &+ \frac{1}{N - n} \frac{p_n^F(d, 1)}{p_n^F(d)} - \left(\frac{p_n^F(d, 1)}{p_n^F(d)} \right)^2. \quad (4.15)
 \end{aligned}$$

Таким образом, и в выражении (4.13) для средней доли $q_n^F(d)$ и в выражении (4.14) для условной дисперсии доли дефектных изделий в оставшейся непроконтролированной части совокупности входят вероятности событий того, что при контроле $(n + 1)$ -го, $(n + 2)$ -го объектов они оба оказались дефектными. В тех случаях, когда контроль дополнительных объектов нежелателен, нужны такие оценки для $q_n^F(d)$ и дисперсии $V_n^F(d)$, которые выражаются только через вероятности $p_n^F(d)$, связанные с выборкой объема n . Для получения таких оценок можно использовать следующее

неравенство:

$$\begin{aligned}
 p_{n+1}^F(d+1) &= \sum_{D=0}^N \left(\frac{D-d}{N-n} \right) \binom{n}{d} \frac{K^{(k)} D^{(d)}}{N^{(n)}} f_N(D) + \\
 &+ \sum_{D=0}^N \frac{K-k+1}{N-n} \binom{n}{d+1} \frac{K^{(k-1)} D^{(d+1)}}{N^{(n)}} f_N(D) = \\
 &= \sum_{D=0}^N \frac{K-k+1}{N-n} \left(1 + \frac{d+1}{n-d} \right) \binom{n}{d+1} \frac{K^{(k-1)} D^{(d+1)}}{N^{(n)}} f_N(D) \leqslant \\
 &\leqslant \left[\frac{n+1}{n-d} p_n^F(d+1) \right], \quad (4.16)
 \end{aligned}$$

где ${}_1[b] = 1$, если $b > 1$, и ${}_1[b] = b$, $0 \leqslant b \leqslant 1$. Точно так же можно вывести, что $p_{n+1}^F(d+1) \leqslant \frac{d+1}{n+1} p_n^F(d)$, что вместе с (4.16) дает верхнюю оценку

$$\begin{aligned}
 p_{n+1}^F(d+1) &\leqslant a_{n+1}^F(d+1) = \\
 &= \min \left(\frac{n+1}{n-d} p_n^F(d+1), \frac{d+1}{n+1} p_n^F(d), 1 \right). \quad (4.17)
 \end{aligned}$$

Соответственно (4.17) находим верхнюю оценку для $p_{n+2}^F(d+2)$:

$$\begin{aligned}
 p_{n+2}^F(d+2) &\leqslant a_{n+2}^F(d+2) = \\
 &= \min \left(\frac{n+2}{n-d} a_{n+1}^F(d+2), \frac{d+2}{n+2} a_{n+1}^F(d+1), 1 \right). \quad (4.18)
 \end{aligned}$$

Таким образом, верхняя оценка для $V_n^F(d)$ получается подстановкой в первые два слагаемых (4.15) значений $a_{n+1}^F(d+1)$ и $a_{n+2}^F(d+2)$, найденных по формулам (4.17) и (4.18), вместо $p_{n+1}^F(d+1)$ и $p_{n+2}^F(d+2)$. Последнее слагаемое в (4.15) отрицательно, поэтому значение $p_{n+1}^F(d+1)$ заменяем оценкой снизу:

$$p_{n+1}^F(d+1) \geqslant b_n^F(d+1) = \frac{1}{N-n} \frac{n+1}{d+1} p_n^F(d). \quad (4.19)$$

Поскольку числа $p_n^F(d)$ являются определяющими для

многих важных характеристик, то найдем рекуррентные формулы, на основе которых их можно последовательно вычислять. Вероятность обнаружить среди $n+1$ контролируемых объектов d дефектных при контроле $\mathfrak{P} \propto [N, D]$ равна

$$h_{N,D}^{n+1,d} = h_{N,D}^{n,d} \left(1 - \frac{D-d}{N-n}\right) + h_{N,D}^{n,d-1} \frac{D-d+1}{N-n}. \quad (4.20)$$

Умножая обе части равенства (4.20) на $f_N(D)$ и суммируя по D , с учетом (4.9) находим, что

$$\begin{aligned} p_{n+1}^F(d) &= p_n^F(d) - \sum_{D=0}^N \frac{d+1}{n+1} \binom{n+1}{d+1} \frac{K^{(k)} D^{(d+1)}}{N^{(n+1)}} f_N(D) + \\ &+ \sum_{D=0}^N \frac{d}{n+1} \binom{n+1}{d} \frac{K^{(k+1)} D^{(d)}}{N^{(n+1)}} f_N(D) = \\ &= p_n^F(d) - \frac{d+1}{n+1} p_{n+1}^F(d+1) + \frac{d}{n+1} p_{n+1}^F(d). \quad (4.21) \end{aligned}$$

Полученное рекуррентное соотношение (4.21) удобно записать в виде

$$p_n^F(d) = p_{n+1}^F(d) \left[1 - \frac{d}{n+1}\right] + p_{n+1}^F(d+1) \frac{d+1}{n+1}. \quad (4.22)$$

Из (4.22) с учетом «начальных» условий

$$p_N^F(D) = f_N(D) \quad (4.23)$$

можем, последовательно уменьшая индекс $n=N, N-1, \dots$, найти любое значение $p_n^F(d)$.

Выведем еще несколько соотношений. Стандартными преобразованиями получаем

$$\begin{aligned} p_n^F(d, 1) &= \sum_{D=0}^N \frac{D-d}{N-n} \binom{n}{d} \frac{K^{(k)} D^{(d)}}{N^{(k+d)}} f_N(D) = \\ &= \sum_{D=0}^n \frac{(K-k)(d+1)}{(N-n)(n-d)} \binom{n}{d+1} \frac{K^{(k-1)} D^{(d+1)}}{N^{(n)}} f_N(D) = \\ &= \frac{d+1}{n-d} p_n^F(d+1, 0). \quad (4.24) \end{aligned}$$

Из (4.13) и (4.24) находим, что

$$\begin{aligned} q_n^F(d) &= \frac{p_n^F(d, 1)}{p_n^F(d)} = \frac{d+1}{n-d} \frac{p_n^F(d+1, 0)}{p_n^F(d)} = \\ &= \frac{d+1}{n-d} \frac{p_n^F(d+1) - p_n^F(d+1, 1)}{p_n^F(d)} = \\ &= \frac{d+1}{n-d} \frac{p_n^F(d, 1)}{p_n^F(d)} [1 - q_n^F(d+1)], \quad d=0, \dots, n-1. \end{aligned} \quad (4.25)$$

Система рекуррентных соотношений (4.25) получена в работе Большева [5]. Если значения $p_n^F(d)$ считать заданными, то (4.25) есть система n уравнений, содержащая $n+1$ неизвестных значений $q_n^F(d)$, $d=0, \dots, n$. Следовательно, возможные решения (4.25) образуют отрезок прямой, пересекающей единичный куб, так как $0 \leq q_n^F(d) \leq 1$. Как указывается в работе Большева [5], (4.25) вместе с неравенствами $0 \leq q_n^F(d) \leq 1$ можно использовать для получения более точных неравенств для $q_n^F(d)$. Так как $0 \leq q_n^F(0) \leq 1$, то из (4.25) находим

$$1 - n \frac{p_n^F(0)}{p_n^F(1)} \leq q_n^F(1) \leq 1,$$

из последнего неравенства и (4.25) находим неравенство для $q_n^F(2)$ и т. д. Если исходить из неравенства $0 \leq q_n^F(n) \leq 1$, то из (4.25) получим неравенство для $q_n^F(n-1)$, затем неравенство для $q_n^F(n-2)$ и т. д. В качестве нижних границ для $q_n^F(d)$ выбираем максимум из полученных таким способом двух нижних оценок. В качестве верхних границ берем соответственно минимум. Большев [5] приводит примеры, когда найденные этим методом оценки $q_n^F(d)$ оказываются достаточными для того, чтобы на их основе можно было бы выбирать план контроля.

Выше было показано, что для G -биномиальных априорных распределений $q_n^F(d)$ не убывают по d , поэтому рационально использование одноступенчатого плана (n, c) с приемочным числом c .

Пусть

$$q_{\text{пр}}^F(c) = \sum_{d=0}^c \sum_{D=0}^N \frac{D-d}{N-n} h_{N, D}^{n, d} f_N(D) = \sum_{d=0}^c p_n^F(d, 1)$$

среднее значение доли дефектных объектов в непроконтролированной части партии, когда принимается решение принять совокупность по результатам контроля по плану (n, c) , т. е. когда $\delta_n \leq c$.

Теорема 2.4.2. $q_{\text{пр}}^F(c)$ находится из соотношения

$$q_{\text{пр}}^F(c) = \sum_{d=0}^c (d/n) p_n^F(d) + \frac{c+1}{n} p_n^F(c+1, 0). \quad (4.26)$$

Доказательство. Используем метод математической индукции. При $c=0$ из определения $q_{\text{пр}}^F(0)$ и (4.24) находим $q_{\text{пр}}^F(0) = p_n^F(0, 1) = (1/n) p_n^F(1, 0)$, что соответствует (4.26) при $c=0$. Пусть (4.26) верно для всех $c=0, 1, \dots, c_0$. Покажем, что (4.26) выполнено для $c=c_0+1$. В этом случае из (4.24) и (4.26) имеем

$$\begin{aligned} q_{\text{пр}}^F(c_0+1) &= \sum_{d=0}^{c_0+1} p_n^F(d, 1) = q_{\text{пр}}^F(c_0) + p_n^F(c_0+1, 1) = \\ &= \sum_{d=0}^{c_0} \frac{d}{n} p_n^F(d) + \frac{c_0+1}{n} p_n^F(c_0+1, 0) + p_n^F(c_0+1, 1) = \\ &= \sum_{d=0}^{c_0} \frac{d}{n} p_n^F(d) + \frac{c_0+1}{n} p_n^F(c_0+1) + \left[1 - \frac{c_0+1}{n}\right] p_n^F(c_0+1, 1) = \\ &= \sum_{d=0}^{c_0+1} \frac{d}{n} p_n^F(d) + \frac{n-c_0-1}{n} \frac{c_0+2}{n-c_0-1} p_n^F(c_0+2, 0), \end{aligned}$$

что соответствует (4.26) при $c=c_0+1$. Теорема доказана.

Заметим, что, используя (4.24), мы можем (4.26) переписать в виде

$$q_{\text{пр}}^F(c) = \sum_{d=0}^c \frac{d}{n} p_n^F(d) + \frac{n-c}{n} p_n^F(c, 1). \quad (4.27)$$

Так как $p_n^F(c+r, 0) \leq p_n^F(c+r)$ при любом целом r , то для $q_{np}^F(c)$ выполнены неравенства

$$\sum_{d=0}^c (d/n) p_n^F(d) \leq q_{np}^F(c) \leq \sum_{d=0}^{c+1} (d/n) p_n^F(d). \quad (4.28)$$

Найденные выше соотношения можно использовать для получения последующих оценок в условиях существования априорного распределения F . Для этого, конечно, нужно предположить, что частотные оценки соответствующих вероятностей сходятся к этим вероятностям. Пусть m — общее число совокупностей $\mathfrak{P}_N$, контролируемых по плану (n, c) , из которых у $v_{n,m}(d)$ числа совокупностей было $\delta_n = d$. Пусть проводился контроль дополнительных объектов. Обозначим $v_{n,m}(d, e_1)$ — число случаев, когда $\delta_n = d$, а $\varepsilon_{n+1} = e_1$, $v_{n,m}(d, e_1, e_2)$ — число случаев, когда $\delta_n = d$, $\varepsilon_{n+1} = e_1$, $\varepsilon_{n+2} = e_2$ и т. п. Предположение о сходимости означает, что при $m \rightarrow \infty$

$$\frac{v_{n,m}(d)}{m} \rightarrow p_n^F(d), \quad \frac{v_{n,m}(d, e_1)}{m} \rightarrow p_n^F(d, e_1) \text{ и т. д.}$$

Таким образом, не зная вида априорного распределения F , можно оценивать необходимые вероятности. Такой метод носит название **эмпирического байесовского подхода**; он был предложен Роббинсом [1], [2]. В соответствии с этим подходом из (4.27) находим, что оценка числа дефектных объектов в принятых партиях при контроле по плану $(n, c)_{1,3}$ с проверкой одного дополнительного объекта будет равна

$$m(N-n) \left[\sum_{d=0}^c \frac{d}{n} \frac{v_{n,m}(d)}{m} + \frac{n-c}{n} \frac{v_{n,m}(c, 1)}{m} \right]. \quad (4.29)$$

Если бы расчет проводился с использованием несмещенных оценок по формуле (1.4.28), то несмещенная оценка числа принятых дефектных объектов имела бы вид

$$[N - (n+1)] \left[\sum_{d=0}^c \frac{d}{n+1} v_{n+1,m}(d) + v_{n,m}(c, 1) \right].$$

Таким образом, мы видим, что при больших n и малых c ($c/n \ll 1$) оба подхода дают близкие результаты.

Аналогичным образом легко найти эмпирические байесовские оценки для $q_n^F(d)$. Заменяя в (4.13) вероятности наблюдаемыми частотами, получим

$$\hat{q}_n^F(d) = \frac{d+1}{n+1} \frac{v_{n+1, m}(d+1)}{v_{n, m}(d)}, \quad \hat{q}_n'^F(d) = \frac{v_{n, m}(d, 1)}{v_{n, m}(d)}.$$

Таким же способом находятся оценки дисперсии полученных оценок, которые в свою очередь можно использовать для построения приближенных доверительных интервалов и т. д.

Байесовский подход следует развить и для двухступенчатых планов контроля. Здесь также можно получить соответствующие рекуррентные соотношения для средней доли $(D_\omega - \delta')/(N - n')$ дефектных изделий, (где δ' — число обнаруженных дефектных объектов, а n' — число проконтролированных объектов), можно найти оценки дисперсий этих оценок и т. д. Некоторые начальные результаты для двухступенчатых планов были получены Юсасом [1].

ГЛАВА 3

СТАТИСТИЧЕСКИЕ ЗАДАЧИ

§ 3.1. Последующие точечные оценки для планов первого вхождения

Несмещенные оценки функций числа дефектных объектов в контролируемой совокупности. В этой главе рассматриваются методы получения точечных и интервальных оценок. На основе таких оценок можно делать выводы о степени эффективности используемых систем правил (стандартов) приемочного контроля. Для одноступенчатых планов контроля последующие оценки предъявленного и пропущенного брака были даны выше в § 1.4. Теперь мы рассмотрим более общую задачу построения несмещенных оценок для широкого класса планов контроля первого вхождения.

Пусть задана коллекция планов выборочного приемочного контроля $\Pi = \{P\}$ и система правил использования этой коллекции для контроля последовательности совокупностей (партий продукции) $\mathfrak{P}_i \asymp [N_i, D_i]$, $i = 1, 2, \dots$. Для контроля $\mathfrak{P}_i$ применяется план $P_i \in \Pi$. Будем предполагать, что при любом D вероятность выхода пути контроля на границу равна единице, т. е. что план P_i *замкнутый*. Результаты контроля, т. е. числа обнаруженных при контроле годных и дефектных объектов, регистрируются в журнале контроля, пробиваются на перфокартах или записываются в запоминающих устройствах ЭВМ (на дисках, магнитных лентах и т. п.). Как уже указывалось, по результатам контроля принимаются решения принять или забраковать $\mathfrak{P}_i$. В зависимости от принимаемых решений ставятся задачи вычисления тех или иных числовых показателей фактической эффективности приемочного контроля.

Наше исследование начнем с планов Π_{13} , т. е. планов, в которых оставшаяся непроконтролированной часть $\mathfrak{F}$ отвергается без контроля (решение $\mathcal{D}_1$) либо принимается без контроля (решение $\mathcal{D}_3$). Как и в § 1.4, представим D в виде

$$D = D' + D'' + D''', \quad (1.1)$$

где D' — число дефектных объектов в непроконтролированной части забракованной совокупности $\mathfrak{F}$, D'' — число выявленных дефектных объектов при выборочном контроле, а D''' — число дефектных объектов, содержащихся в непроконтролированной части принятой совокупности $\mathfrak{F}$. Если проведен контроль m совокупностей $\mathfrak{F}_1, \dots, \mathfrak{F}_m$, то его эффективность можно оценить, сравнивая суммарное число дефектных объектов $\sum_{1 \leq i \leq m} D_i$ в $\bigcup_{1 \leq i \leq m} \mathfrak{F}_i$ (предъявленный брак) и суммарное число дефектных объектов $\sum_{1 \leq i \leq m} D_i''$ в принятых $\mathfrak{F}_i$ (принятый брак). Аналогично (1.1), для годных объектов можно записать

$$K = K' + K'' + K''', \quad (1.2)$$

где K' — число годных объектов в непроконтролированной части $\mathfrak{F}$ при браковании, K'' — число обнаруженных при выборочном контроле годных объектов, K''' — число годных объектов, содержащихся в той части $\mathfrak{F}$, которая принимается без контроля. Соответственно

$$N = N' + N'' + N''', \quad (1.3)$$

где $N' = K' + D'$, $N'' = K'' + D''$, $N''' = K''' + D'''$. Как $\sum_{1 \leq i \leq m} D_i$, так и $\sum_{1 \leq i \leq m} D_i''$ не известны, поэтому рассмотрим методы получения точечных (а в следующем параграфе и интервальных) статистических оценок этих величин.

Напомним, что в математической статистике оценками называют функции от результатов наблюдений. Если бы речь шла об оценке D или D''' по результатам контроля одной отдельной совокупности $\mathfrak{F}$, то при отборе среди многих возможных оценок наилучшей можно было бы исходить из задания определенной функции риска. После этого мы искали бы оценки, минимизирующие этот риск. При контроле последовательностей $\mathfrak{F}_i$

естественно исходить из суммарных значений используемых оценок, т. е. если $\hat{D}_i$ и $\hat{D}'_i$ — оценки для D_i и D'_i , то оценками для $\sum_{1 \leq i \leq m} D_i$ и $\sum_{1 \leq i \leq m} D'_i$ будут $\sum_{1 \leq i \leq m} \hat{D}_i$ и $\sum_{1 \leq i \leq m} \hat{D}'_i$. Наиболее типичной ситуацией является такая, в которой $\hat{D}_i, \hat{D}'_i$ являются взаимно независимыми по i и поэтому можно использовать центральную предельную теорему, т. е. можно считать $\sum_{1 \leq i \leq m} \hat{D}_i$ и $\sum_{1 \leq i \leq m} \hat{D}'_i$ приближенно нормально распределенными случайными величинами. Тогда суммы уклонений $\sum_{1 \leq i \leq m} (\hat{D}_i - D_i)$ и $\sum_{1 \leq i \leq m} (\hat{D}'_i - D'_i)$ достаточно полно характеризуются математическими ожиданиями и дисперсиями. При не слишком жестких допущениях математические ожидания дисперсии сумм уклонений растут при увеличении m пропорционально m , тогда как зона вероятных значений растет пропорционально $\sqrt{m}$. Зона вероятных значений соответствует интервалам, вероятность попадания в которые постоянна или стремится при $m \rightarrow \infty$ к некоторому числу, отличному от 0 и 1. Это говорит в пользу того, что при отборе оценок прежде всего надо добиваться несмещенности, а затем уже среди несмещенных оценок искать оценки с наименьшими дисперсиями. Напомним, что оценки называются **несмещенными**, если при любых возможных значениях неизвестных параметров, определяющих вероятности исходов наблюдений, математические ожидания оценок равны математическим ожиданиям оцениваемых величин. В рассматриваемом случае для несмещенных оценок $\hat{D}$ и $\hat{D}'$ соответственно имеем

$$E\hat{D} \equiv D, \quad E\hat{D}' \equiv D'. \quad (1.4)$$

Здесь D — не случайная величина, а неизвестный параметр, поэтому $E\hat{D} = D$, тогда как оцениваемая величина D' является случайной.

Целесообразность использования несмещенных оценок проиллюстрируем следующим элементарным примером. Пусть заданы две последовательности взаимно независимых нормально распределенных случайных

величин $\{\xi_i\}$, $\{\eta_i\}$, $i = 1, 2, \dots$, $E\xi_i = 1$, $E\eta_i = 0$, дисперсии которых равны $V\xi_i = 1$, $V\eta_i = 4$. Случайные величины ξ_i , η_i будем рассматривать как возможные оценки нуля. Если оценивание производится по одному наблюдению, то ξ_1 во многих отношениях лучше η_1 . Например, интервал, соответствующий одинаковым значениям вероятностей попадания, построенный по корню из дисперсии, умноженному на 3, для ξ_1 будет $[-2, +4]$, а для η_1 — соответственно $[-6, +6]$. Положение изменится, если мы оценивание будем производить по 100 наблюдениям. Здесь $E\left(\sum_{1 \leq i \leq 100} \xi_i\right) = 100$, $E\left(\sum_{1 \leq i \leq 100} \eta_i\right) = 0$, $V\left(\sum_{1 \leq i \leq 100} \xi_i\right) = 100$, $V\left(\sum_{1 \leq i \leq 100} \eta_i\right) = 400$. Поэтому интервалы, соответствующие трем корням из дисперсии для $\frac{1}{100} \sum_{1 \leq i \leq 100} \xi_i$, есть $[0,7, 1,3]$, а для

$\frac{1}{100} \sum_{1 \leq i \leq 100} \eta_i$ есть $[-0,6, +0,6]$. Таким образом, 0 будет при 100 наблюдениях лучше оценен по последовательности $\{\eta_i\}$, чем по последовательности $\{\xi_i\}$.

В основу построения несмещенных оценок положим следующий метод. Сначала найдем хотя бы какую-нибудь несмещенную оценку, а затем будем ее усреднять по достаточной статистике, в результате чего на основе найденной исходной несмещенной оценки построим новую несмещенную оценку, имеющую меньшую дисперсию. Построение улучшенных оценок базируется на теореме 2.1.2, которая была сформулирована в удобной для наших целей форме в § 2.1. Именно такой подход был использован Колмогоровым [4] для получения последующих несмещенных оценок для одноступенчатых планов контроля.

В рассматриваемой задаче контроля совокупности $\mathfrak{F} \times [N, D]$ по плану $P \in \Pi$ с границей Γ результаты контроля записываются в виде последовательности $(e_1, \dots, e_{k+d})$, где $(k, d) \in \Gamma$, $\sum_{1 \leq i \leq k+d} e_i = d$, $\sum_{1 \leq i \leq k+d} \bar{e}_i = k$.

Начнем с построения несмещенной оценки $\bar{D}$ для предъявленного брака D при контроле по плану Π_{13} . Для выбора исходной оценки ξ нам достаточно заме-

тить, что вероятность того, что $\varepsilon_1 = 1$, т. е. вероятность того, что первый проконтролированный объект окажется дефектным, равна D/N . Следовательно, среднее случайной величины $\xi = N \cdot \varepsilon_1$ равно

$$E\xi = N \cdot P(\varepsilon_1 = 1) = N \cdot \frac{D}{N} = D.$$

Таким образом, оценка ξ есть несмещенная оценка для D . Однако это очень грубая оценка. Постараемся ее «улучшить», используя теорему 2.1.2. Как уже было показано выше в § 2.2, достаточной статистикой является $S = x = (k, d) \in \Gamma$, где (k, d) — координаты точки прекращения контроля. Улучшенная оценка имеет вид

$$\begin{aligned} \hat{D} &= E\{\xi | S = (k, d)\} = \\ &= N \cdot P\{\varepsilon_1 = 1, x = (k, d)\} / P\{x = (k, d)\}. \end{aligned} \quad (1.5)$$

Знаменатель в (1.5) вычисляется по формуле (2.2.6). Для вычисления числителя достаточно заметить, что $P\{x = (k, d) | \varepsilon_1 = 1\}$ равна вероятности выхода на границу в точке x , когда процесс контроля начинается из точки $(0, 1)$ и контролируется совокупность $\mathfrak{P} \asymp [N-1, D-1]$. Таким образом, мы снова можем воспользоваться формулой (2.2.6) с заменой N, D, d на $N-1, D-1, d-1$ и $L(x)$ на $L((0, 1), x)$ — число путей, соединяющих $(0, 1)$ с точкой $x \in \Gamma$. Если $(0, 1) \in \Gamma$, то следует положить $L((0, 1), (0, 1)) = 1$, $L((0, 1), x) = 0$, $x \neq (0, 1)$. Итак, имеем

$$P\{\varepsilon_1 = 1, x = (k, d)\} = \frac{D}{N} L((0, 1), x) \frac{(D-1)^{(d-1)} K^{(k)}}{(N-1)^{(k+d-1)}}. \quad (1.6)$$

Из (1.5), (2.2.6) и (1.6) окончательно находим

$$\hat{D} = N \frac{D}{N} \frac{L((0, 1), x) (D-1)^{(d-1)} K^{(k)}}{P\{x = (k, d)\} (N-1)^{(k+d-1)}} = N \frac{L((0, 1), x)}{L(x)}. \quad (1.7)$$

Полученный результат сформулируем в виде следующей теоремы.

Теорема 3.1.1. Если контроль совокупности $\mathfrak{P} \asymp [N, D]$ проводится по плану $\Pi_{13} \in \Pi$, то несмещенная оценка $\hat{D}$ для предъявленного брака D определяется по формуле (1.7). Другими словами, оценка $\hat{D}$ равна объему контролируемой совокупности, умножен-

ному на отношение числа путей, соединяющих точку $(0, 1)$ с точкой $x = (k, d)$ остановки контроля, к числу путей, соединяющих начальную точку $(0, 0)$ с точкой x .

Если Π_{13} — одноступенчатый план с объемом выборки n , то $L(x) = \binom{n}{d}$, $L((0, 1)x) = \binom{n-1}{d-1}$. Поэтому здесь оценка

$$\hat{D} = N \binom{n-1}{d-1} / \binom{n}{d} = N \frac{d}{n}. \quad (1.8)$$

Формула (1.8) согласуется с результатами § 1.4.

Для вычисления оценки дисперсии $\sqrt{D}$ нужно найти несмещенную оценку $\hat{D}$ для D^2 . Действительно, если такая оценка была бы найдена, то

$$V(\hat{D}) = E(\hat{D})^2 - D^2 = E\{(\hat{D})^2 - \hat{U}\}.$$

Поэтому несмещенная оценка дисперсии оценки $\hat{D}$ для предъявленного брака была бы равна

$$V(\hat{D}) = (\hat{D})^2 - \hat{U}. \quad (1.9)$$

Для этой цели обобщим теорему 3.1.1 на случай оценок произведений $D^{(m)}K^{(l)}$, $l + m \leq \text{ind } \Pi_{13}$. Если $e_1, \dots, e_{k+d}$ — результаты контроля, а точка $z = (l, m)$ — достижимая, то в качестве исходной оценки ξ возьмем

$$\xi = N^{(l+m)} \prod_{1 \leq j \leq m} e_{i_j} \prod_{1 \leq r \leq l} \bar{e}_{i_r}, \quad i_j, i_r \in (1, 2, \dots, l+m), \quad (1.10)$$

где последовательность $e_{i_j} = 1$, $j = 1, \dots, m$, $e_{i_r} = 0$, $i_r \neq i_j$, $r = 1, \dots, l$, $i_j, i_r \in (1, 2, \dots, l+m)$, образует путь, соединяющий $(0, 0)$ с (l, m) . Применяя теорему 2.1.2, получаем $\hat{M}(m, l)$ — несмещенную оценку $D^{(m)}K^{(l)}$, $\hat{M}(m, l) =$

$$= \frac{N^{(l+m)}}{P\{x = (k, d)\}} P\{e_{i_j} = 1, j = 1, \dots, m, e_{i_r} = 0, i_r \neq i_j, \\ r = 1, \dots, l, x = (k, d)\}. \quad (1.11)$$

Числитель и знаменатель в (1.11) снова вычисляются с использованием формулы (2.2.6) и учетом того, что после прохождения точки (l, m) ведется контроль сово-

купности $\mathfrak{P}' \asymp [N - (l + m), D - m]$. В итоге мы получаем

$$\begin{aligned} \hat{M}(m, l) &= N^{(l+m)} \frac{D^{(m)} K^{(l)} L(z, x) (D - m)^{(d-m)} (K - l)^{(k-l)}}{N^{(m+l)} \mathbf{P}\{x\} (N - (m + l))^{(k+d-l-m)}} = \\ &= N^{(l+m)} L(z, x) / L(x), \quad x = (k, d). \end{aligned} \quad (1.12)$$

Если точка x недостижима из точки z , т. е. нет пути, соединяющего z с x , то полагаем $L(z, x) = 0$. Сформулируем полученный результат.

Теорема 3.1.2. Пусть контроль совокупности $\mathfrak{P} \asymp [N, D]$ проводится по плану $\Pi_{13} \in \Pi$, для которого точка $z = (l, m)$ является проходимой (достижимой). Тогда несмещенная оценка $D^{(m)} K^{(l)}$ определяется по формуле (1.12), т. е. она равна $N^{(m+l)}$, умноженному на отношение числа путей, соединяющих точку z с точкой остановки контроля x , к числу путей из $(0, 0)$ в x .

Воспользуемся этой общей теоремой для вывода несмещенной оценки $\hat{V}(\hat{D})$ дисперсии $V(\hat{D})$. Замечая, что $D^2 = D^{(2)} + D$, из (1.12) имеем

$$\hat{U} = N^{(2)} \frac{L((0, 2), x)}{L(x)} + N \frac{L((0, 1), x)}{L(x)}, \quad (1.13)$$

если, конечно, $(0, 2)$ — достижимая точка. Из (1.13) и (1.9) в случае достижимости $(1, 1)$ получаем выражение искомой несмещенной оценки $\hat{V}(\hat{D})$ для $V(\hat{D})$:

$$\begin{aligned} \hat{V}(\hat{D}) &= N^2 \left(\frac{L((0, 1), x)}{L(x)} \right)^2 - N^{(2)} \frac{L((0, 2), x)}{L(x)} - N \frac{L((0, 1), x)}{L(x)} = \\ &= \frac{N^2 [L((0, 1), x)^2 - L((0, 2), x) L(x)] - NL((1, 1), x) L(x)}{L(x)^2}. \end{aligned} \quad (1.14)$$

Если же точка $(0, 1) \in \Gamma$, то $(0, 2)$ недостижима и поэтому пользоваться формулой (1.14) нельзя. В предположении $\text{ind } \Pi_{13} \geq 2$ имеем $(1, 0) \in \Gamma$. Следовательно, точка $(1, 1)$ достижима. В этом случае можем воспользоваться соотношением

$$D^2 = (N - K)D = ND - KD. \quad (1.15)$$

Из (1.15) и теоремы 3.1.2 получаем

$$\hat{U}' = N^2 \frac{L((0, 1), x)}{L(x)} - N^{(2)} \frac{L((1, 1), x)}{L(x)}. \quad (1.16)$$

Из (1.9), (1.7) и (1.12) находим искомую несмещенную оценку $V(\hat{D})$, которая равна

$$\hat{V}(\hat{D}) = N^2 \left(\frac{L((0, 1), x)}{L(x)} \right)^2 + N^{(2)} \frac{L((1, 1), x)}{L(x)} - N^2 \frac{L((0, 1), x)}{L(x)}. \quad (1.17)$$

В частном случае плана контроля одноступенчатой выборки объема n имеем

$$L(x) = \binom{n}{d}, \quad L((0, 2), x) = \binom{n-2}{d-2}, \\ L((0, 1), x) = \binom{n-1}{d-1}, \quad L((1, 1), x) = \binom{n-2}{d-1}.$$

Подставляя эти выражения в (1.14), находим, что

$$\hat{V}(\hat{D}) = N(N-n)d(n-d)/n^2(n-1). \quad (1.18)$$

Формула (1.18) согласуется с известной формулой для оценки дисперсии доли дефектных изделий в выборке объема n .

Полнота планов контроля и единственность несмещенных оценок. Итак, мы умеем выводить несмещенные оценки, являющиеся функциями от координат точки остановки контроля. Когда же эти оценки будут единственными в классе несмещенных оценок, зависящих от координат точки остановки? Этот вопрос интересен хотя бы потому, что в случае единственности в силу теоремы 2.1.2 найденные оценки имеют в классе несмещенных оценок наименьшие дисперсии. Ниже будет дано решение этого вопроса. Основным здесь является важное понятие о полноте планов контроля.

Определение 3.1.1. План контроля $\Pi \in \Pi$ называется **полным**, если для функции $\varphi = \varphi(k, d)$, $(k, d) \in \Gamma$, из выполнения равенства

$$E\varphi = \sum_{(k, d) \in \Gamma} \varphi(k, d) p_D(k, d) \equiv 0, \quad (1.19)$$

при любых $D = 0, \dots, N$, следует, что $\varphi(k, d) \equiv 0$ для достижимых (k, d) . Другими словами, полные планы — такие, для которых не существует нетривиальных (т. е. отличных от нуля) оценок нуля.

Лемма 3.1.1. Пусть план контроля Π — полный. Если для функции $f((k, d), D)$ существует несмещенная

оценка, зависящая только от координат остановки $(k, d) \in \Gamma$, то эта оценка — единственная.

. Доказательство. Предположим, что существуют две несмещенные оценки $\varphi_i(k, d)$, $i = 1, 2$, функции f , т. е. при любом $D = 0, \dots, N$

$$\sum_{(k, d) \in \Gamma} \varphi_i(k, d) p_D(k, d) \equiv \sum_{(k, d) \in \Gamma} f((k, d), D) p_D(k, d).$$

Для разности $\varphi_1(k, d) - \varphi_2(k, d)$ соответственно имеем

$$\sum_{(k, d) \in \Gamma} [\varphi_1(k, d) - \varphi_2(k, d)] p_D(k, d) \equiv 0.$$

Так как план Π — полный, то отсюда следует, что $\varphi_1(k, d) \equiv \varphi_2(k, d)$. Таким образом, несмещенная оценка — единственна.

Верно и обратное утверждение. Если план Π — неполный, то для любой $f((k, d), D)$, допускающей несмещенное оценивание, существует бесконечное множество несмещенных оценок. Для этого достаточно заметить, что из неполноты следует существование несмещенной оценки нуля $\varphi_0(k, d) \not\equiv 0$. Следовательно, если $\varphi(k, d)$ — несмещенная оценка для $f((k, d), D)$, то $\varphi(k, d) + \varphi_0(k, d)$ — также несмещенная оценка для $f((k, d), D)$.

Поскольку для неполных планов существует много различных несмещенных оценок, оценивающих одну и ту же функцию, то возникают задачи отбора из этих оценок наилучших. Эти задачи еще полностью не решены. Можно отметить интересные результаты Кагана, Линника и Рао [1], стр. 310—312. Что же касается вопроса о полноте, то мы получим его окончательное решение. Введем следующее важное понятие простоты плана.

Определение 3.1.2. План первого входа $\Pi \in \Pi$ называется **простым**, если все проходимые точки с одинаковым значением индекса образуют отрезки, т. е. не разделены граничными или недостижимыми точками. В противном случае план называется **непростым**.

Рассмотрим примеры простых и непростых планов первого входа. На рис. 3.1.1, а дан простой план с границей общего типа. Отрезки соединяют внутренние проходимые точки. На рис. 3.1.1, б дан непростой план общего вида. Пунктирной прямой L на рис. 3.1.1, б со-

ответствует значение индекса, при котором проходимые точки образуют несвязные отрезки, помеченные буквами A , B .

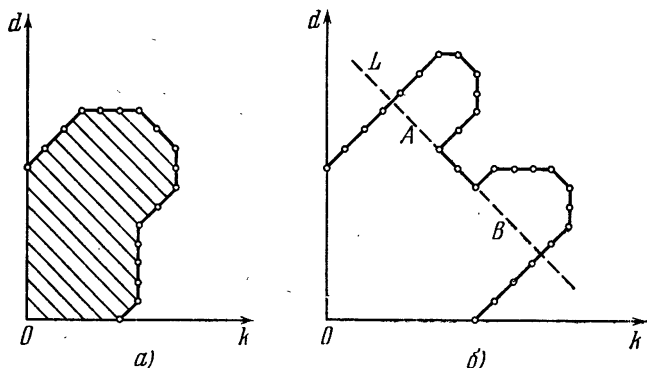


Рис. 3.1.1. Примеры планов первого вхождения: а) простой план, б) непростой план.

Основным результатом о полноте плана является следующая теорема.

Теорема 3.1.3. План контроля $P \in \Pi$ является полным тогда и только тогда, когда он простой.

Доказательство необходимости. Итак, пусть план P не является простым. Это означает, что найдется хотя бы одна граничная точка такая, что при том же значении индекса вдоль прямой по обе стороны от нее имеются проходимые точки. Часть границы непростого плана показана на рис. 3.1.2, где белыми кружками отмечены проходимые точки, разделяемые граничными точками, которые показаны черными кружками. Кроме того, крестиками выделены две достижимые точки A_1 , A_2 . Пусть координаты граничных

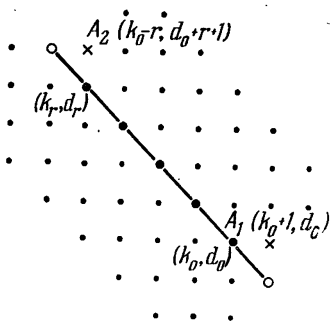


Рис. 3.1.2. Часть границы непростого плана контроля.

кружками. Кроме того, крестиками выделены две достижимые точки A_1 , A_2 . Пусть координаты граничных

точек равны (k_0, d_0) , (k_1, d_1) , ..., (k_r, d_r) , $k_i = k_0 - i$, $d_i = d_0 + i$, $i = 0, \dots, r$, $A_1 = (k_0 + 1, d_0)$, $A_2 = (k_0 - r, d_0 + r + 1)$. Построим нетривиальную (не тождественно равную нулю) оценку нуля. Основным элементом в построении будет следующее комбинаторное тождество:

$$S_r = \sum_{i=0}^r (-1)^i \frac{D^{(d_0+i)} K^{(k_0-i)}}{N^{(k_0+d_0)}} = \\ = \frac{D^{(d_0)} K^{(k_0+1)}}{N^{(k_0+d_0+1)}} + (-1)^r \frac{D^{(d_0+r+1)} K^{(k_0-r)}}{N^{(k_0+d_0+1)}}, \quad (1.20)$$

Вывод этого тождества проведем по индукции. Действительно, производя элементарные преобразования, имеем

$$\frac{D^{(d_0)} K^{(k_0+1)}}{N^{(k_0+d_0+1)}} + \frac{D^{(d_0+1)} K^{(k_0)}}{N^{(k_0+d_0+1)}} = \\ = \frac{D^{(d_0)} K^{(k_0)}}{N^{(k_0+d_0)}} \frac{(K - k_0) + (D - d_0)}{N - k_0 - d_0} = \frac{D^{(d_0)} K^{(k_0)}}{N^{(k_0+d_0)}},$$

что соответствует (1.20) при $r = 0$. Пусть для числа r тождество (1.20) справедливо. Покажем, что оно справедливо и для числа $r + 1$. Рассмотрим разность

$$S_{r+1} - S_r = (-1)^{r+1} \frac{D^{(d_0+r+1)} K^{(k_0-r-1)}}{N^{(k_0+d_0)}} = \\ = (-1)^{r+1} \frac{D^{(d_0+r+1)} K^{(k_0-r-1)}}{N^{(k_0+d_0)}} \frac{(D - d_0 - r - 1) + (K - k_0 + r + 1)}{N - (k_0 + d_0)} = \\ = (-1)^{r+1} \frac{D^{(d_0+r+2)} K^{(k_0-r-1)}}{N^{(k_0+d_0+1)}} - (-1)^r \frac{D^{(d_0+r+1)} K^{(k_0-r)}}{N^{(k_0+d_0+1)}}. \quad (1.21)$$

Выражая в (1.21) S_{r+1} через S_r и используя для S_r тождество (1.20), находим, что S_{r+1} также удовлетворяет (1.20) с заменой r на $r + 1$. Таким образом, тождество (1.20) справедливо при любом r .

Построим нетривиальную несмещенную оценку нуля. Для этого определим функцию $\varphi(k, d)$, полагая

$$\varphi(k, d) = \begin{cases} (-1)^r \{L(k, d)\}^{-1}, & \text{если } k = k_i, d = d_i, 0 \leq i \leq r \\ & \text{или если } (k, d) = A_j, j = 1, 2, \\ 0, & \text{в противном случае.} \end{cases}$$

Вычисляя математическое ожидание оценки $\varphi(k, d)$, получаем

$$\begin{aligned} E\varphi &= \sum_{0 \leq i \leq r} (-1)^i \{L(k_i, d_i)\}^{-1} L(k_i, d_i) \frac{K^{(k_i)} D^{(d_i)}}{N^{(k_i+d_i)}} + \\ &+ \sum_{(k, d) \in \Gamma^*} (-1)^i \{L(A_1)\}^{-1} L(k_0+1, d_0) \frac{D^{(d_0)} K^{(k_0+1)}}{N^{(d_0+k_0+1)}} Z_1(k, d) + \\ &+ \sum_{(k, d) \in \Gamma^*} (-1)^i \{L(A_2)\} L(k_0-r, d_0+r+1) \times \\ &\quad \times \frac{D^{(d_0+r+1)} K^{(k, -r)}}{N^{(k_0+d_0+1)}} Z_2(k, d) = \\ &= S_r - \frac{D^{(d_0)} K^{(k_0+1)}}{N^{(k_0+d_0+1)}} - (-1)^{r+1} \frac{D^{(d_0+r+1)} K^{(k_0-r)}}{N^{(d_0+k_0+1)}} = 0, \end{aligned}$$

где

$$\begin{aligned} \Gamma^* &= \{(k, d): (k, d) \in \Gamma, (k, d) \neq (k_i, d_i), i = 0, \dots, r\}, \\ Z_1(k, d) &= \\ &= L((k_0+1, d_0), (k, d)) \frac{(D-d_0)^{(d-d_0)} (K-k_0-1)^{(k-k_0-1)}}{(N-d_0-k_0)^{(d+k-d_0-k_0-1)}}, \\ Z_2(k, d) &= L((k_0-r, d_0+r+1), (k, d)) \times \\ &\quad \times \frac{(D-d_0-r-1)^{(d-d_0-r-1)} (K-k_0+r)^{(k-k_0+r)}}{(N-k_0-d_0-1)^{(k+d-k_0-d_0-1)}}. \end{aligned}$$

Здесь при суммировании по $(k, d) \in \Gamma$ мы использовали тот факт, что вероятность выхода на границу из точек A_j , $j = 1, 2$, при любом D тождественно равна единице. Далее использовалось тождество (1.20). Таким образом, найдена нетривиальная несмещенная оценка нуля. Однако φ не есть функция достаточной статистики, если A_j — не граничные точки. В таком случае надо воспользоваться теоремой 2.1.2 и улучшить эту оценку по формуле (2.1.21). Улучшенная оценка нуля не станет тривиальной, так как ее не нулевые значения в граничных точках (k_i, d_i) , $i = 0, \dots, r$, при этом останутся равными значениям $\varphi(k_i, d_i)$. Следовательно, простота плана необходима для его полноты, так как в про-

тивном случае существуют нетривиальные несмещенные оценки нуля.

Доказательство достаточности. Из предположения о простоте плана Π следует, что все точки границы Γ мы можем последовательно разбить на непересекающиеся множества $\Gamma^{(i)}$. В $\Gamma^{(1)}$ заключаем те точки Γ , для которых при $k=0, \dots, k_{(1)}$ на каждой из вертикалей $k=\text{const}$ есть только одна точка (k, d_k) границы Γ . Аналогично выделяется множество $\Gamma^{(2)}$ (рис. 3.1.3). Для этого берутся значения $d=0, \dots$

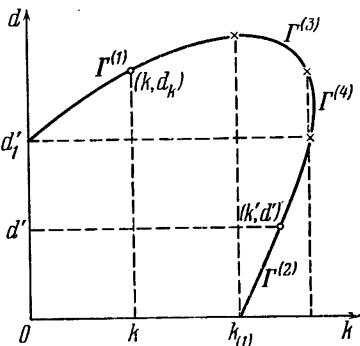


Рис. 3.1.3. Разбиение границы простого плана контроля.

..., d'_1 , для которых каждому d соответствует одна точка $(k_d, d) \in \Gamma$. После выделения $\Gamma^{(1)}$ и $\Gamma^{(2)}$ разбиение оставшейся части границы $\Gamma \setminus (\Gamma^{(1)} \cup \Gamma^{(2)})$ проводится аналогичным образом. Покажем, что для простого плана система уравнений (1.19) имеет тривиальное решение $\varphi(k, d) \equiv 0$. Так как $p_D(k, d) = L(k, d) D^{(d)} K^{(k)} / N^{(d+k)}$, то, полагая $K=0$ ($D=N$), получаем, что $p_D(k, d) = 0$ для всех точек $(k, d) \in \Gamma^{(1)}$, кроме точки $(0, d_0) \in \Gamma^{(1)}$, для которой $p_N(0, d_0) = 1$ (на рис. 3.1.3 $d_0 = d'_1$). Следовательно, система (1.19) сводится к уравнению $\varphi(0, d_0) \cdot 1 = 0$, т. е. $\varphi(0, d_0) = 0$. Положим теперь $K=1$. Тогда лишь вероятности $p_{N-1}(0, d_0)$ и $p_{N-1}(1, d_1)$ точек $(0, d_0), (1, d_1) \in \Gamma^{(1)}$ отличны от нуля. Поэтому система (1.19) имеет вид

$$\begin{aligned} \varphi(0, d_0) p_{N-1}(0, d_0) + \varphi(1, d_1) p_{N-1}(1, d_1) = \\ = \varphi(1, d_1) p_{N-1}(1, d_1) = 0. \end{aligned}$$

Таким образом, находим, что и $\varphi(1, d_1) = 0$. Последовательно полагая далее $K=2, \dots, k_{(1)}$, находим, что $\varphi(k, d_k) = 0$ для всех $(k, d) \in \Gamma^{(1)}$. Для участка границы $\Gamma^{(2)}$ точно также последовательно полагая $D=0, 1, \dots, d'_1$, найдем, что $\varphi(k, d) = 0$, если $(k, d) \in \Gamma^{(2)}$. Для оставшейся части $\Gamma \setminus (\Gamma^{(1)} \cup \Gamma^{(2)})$ поступаем точно так

же, последовательно полагая значения $K = k_{(1)} + 1, k_{(1)} + 2, \dots$ до исчерпания границы $\Gamma^{(3)}$ и аналогично для $\Gamma^{(4)}$ и т. д. Таким образом, все $\varphi(k, d) \equiv 0$, что и соответствует тривиальности несмещенной оценки нуля. Теорема доказана.

Отметим, что в работах Линника и Романовского [1], [2] была сформулирована интересная теорема о том, что полнота плана не зависит от того, используется случайный отбор объектов на контроль с возвращением или без возвращения. Теория несмещенного оценивания для случая повторной выборки, точнее, для схемы однородных случайных блужданий, была развита в работе Гиршика, Мостеллера и Сэвиджа [1], а некоторые дополнительные результаты были получены Де Гроотом [1].

Несмещенные оценки для предъявленного и принятого брака. Дадим рекомендации по несмещенному оцениванию D''' — числа дефектных объектов в части совокупности $\mathfrak{F}$, принимаемой без контроля. Напоминаем, что рассматриваются планы контроля $\Pi_{13} \in \Pi$, для которых граница принятия решения $\Gamma = \Gamma_1 \cup \Gamma_3$. Так как при выходе в точку $x = (k, d) \in \Gamma_3$ остаток $\mathfrak{F}$ принимается без контроля, то

$$D''' = \begin{cases} D - d, & x = (k, d) \in \Gamma_3, \\ 0, & x = (k, d) \in \Gamma_1. \end{cases} \quad (1.22)$$

Математическое ожидание D''' равно

$$ED''' = \sum_{(k, d) \in \Gamma_3} (D - d) p_D(k, d). \quad (1.23)$$

Если $\text{ind } \Gamma = \text{ind } \Gamma_3$, то ED''' есть полином по D степени на единицу большей, чем любой полином вида $\sum_{(k, d) \in \Gamma} \varphi(k, d) p_D(k, d)$. Поэтому для таких планов не существует несмещенных оценок для принятого брака. Заметим, что как для одноступенчатых, так и для двухступенчатых планов $\text{ind } \Gamma = \text{ind } \Gamma_3$. Следовательно, для важных классов планов контроля отсутствуют несмещенные оценки D''' . Однако лишь немного изменив план контроля, можно получить несмещенные оценки для

D''' . Для этого перепишем (1.23) в следующем виде:

$$\begin{aligned} ED''' &= \sum_{(k, d) \in \Gamma_3} [N - (k + d)] \frac{D - d}{N - (k + d)} p_D(k, d) = \\ &= \sum_{(k, d) \in \Gamma_3} [N - (k + d)] p_D(k, d) \times \\ &\quad \times p(\varepsilon_{k+d+1} = 1 | x = (k, d)). \quad (1.24) \end{aligned}$$

Очевидно, что выражение (1.24) можно интерпретировать как математическое ожидание случайной величины ξ , которая при достижении границы Γ_1 полагается равной нулю, а при достижении Γ_3 проводится выбор еще одного дополнительного объекта. При этом если $(k + d + 1)$ -й объект оказывается дефектным, то эта случайная величина ξ полагается равной $N - (k + d)$, в противном случае — равной нулю. Таким образом, видоизмененный план контроля, при котором существует несмещенная оценка ξ для D''' , состоит в следующем: при достижении точки $(k, d) \in \Gamma_1$ совокупность $\mathfrak{P}$ бракуется, при достижении точки $(k, d) \in \Gamma_3$ совокупность $\mathfrak{P}$ принимается, но проводится контроль еще одного $(k + d + 1)$ -го объекта, который, независимо от того, годен он или дефектен, включается в часть $\mathfrak{P}$, принимаемую без контроля. Так видоизмененный план Π будем называть **планом с дополнительным контролем** и обозначать Π_d . Достаточной статистикой для плана Π_d с учетом результатов возможной проверки дополнительного объекта являются координаты (k', d') , где $k' = k + \varepsilon_{k+d+1}$, $d' = d + \varepsilon_{k+d+1}$ при $(k, d) \in \Gamma_3$, $k' = k$, $d' = d$ при $(k, d) \in \Gamma_1$. Оценку ξ можно улучшить, взяв ее среднее при фиксированном значении (k', d') . Такое усреднение существенно, если только $(k' - 1, d')$ и $(k', d' - 1)$ обе одновременно входят в Γ_3 . Так как при этом в (k', d') можно попасть из этих двух точек, то усредненное значение оценки ξ равно $[N - (k' + d' - 1)] \times \times \frac{L(k', d' - 1)}{L(k', d' - 1) + L(k' - 1, d')}$. В остальных случаях известно, каким способом мы попали в точку (k', d') , поэтому усреднение не меняет значения ξ . Итак, мы получаем, что при плане Π_d (с дополнительным контролем)

несмещенная оценка равна

$$\hat{D}''' = \begin{cases} [N - (k' + d') + 1] \frac{L(k', d' - 1)}{L(k', d' - 1) + L(k' - 1, d')}, & \text{если } \varepsilon_{k'+d'} = 1, (k' - 1, d'), (k', d' - 1) \in \Gamma_3, \\ N - (k' + d') + 1, & \text{если } (k', d' - 1) \in \Gamma_3, \\ & (k' - 1, d') \in \Gamma_3, \varepsilon_{k'+d'} = 1, \\ 0, & \text{если либо } (k', d') \in \Gamma_1, \\ & \text{либо } (k' - 1, d') \in \Gamma_3, \text{ но } \varepsilon_{k'+d'} = 0. \end{cases} \quad (1.25)$$

Заметим, что в частном случае одноступенчатого плана контроля объема n , когда $\Gamma_3 = \{(n - k, k), 0 \leq k \leq c\}$, где c — приемочное число, формула (1.25) принимает вид

$$\hat{D}''' = \begin{cases} (N - n) \frac{d'}{n + 1}, & 0 \leq d' \leq c, \\ N - n, & d' = c + 1, \varepsilon_{n+1} = 1, \\ 0, & d' > c + 1. \end{cases} \quad (1.26)$$

Оценки для D''' в случае планов типа одноступенчатой выборки были даны в работах Сираждинова [6], Сираждинова и Эйдельманта [1]. Как видно из изложенного выше, оценки, аналогичные (1.26), могут быть получены для любых планов первого вхождения.

Включение дефектного объекта, выявленного при дополнительном контроле, в принимаемую без контроля часть $\mathfrak{P}$ неестественно. Если такое включение не производить, то число принятых дефектных объектов при плане Π_d равно

$$D_d''' = \begin{cases} D - d - \varepsilon_{k+d+1}, & (k, d) \in \Gamma_3, \\ 0, & \text{если } (k, d) \in \Gamma_1. \end{cases} \quad (1.27)$$

Для математического ожидания D_d''' выводим соотношение

$$\begin{aligned} ED_d''' &= \sum_{(k, d) \in \Gamma_3} \left\{ (D - d) p_D(k, d) \frac{K - k}{N - (k + d)} + \right. \\ &\quad \left. + (D - d - 1) p_D(k, d) \frac{D - d}{N - (k + d)} \right\} = \\ &= \sum_{(k, d) \in \Gamma_3} [N - (k + d) - 1] p_D(k, d) \frac{D - d}{N - (k + d)}. \end{aligned} \quad (1.28)$$

Следовательно, несмещенная оценка ξ_d для D_d''' равна $N - (k + d) - 1$ только когда $(k, d) \in \Gamma_3$ и $\varepsilon_{k+d+1} = 1$, в противном случае $\xi_d = 0$. Как и ранее, усредняя ξ_d при фиксированных значениях $k' = k + \varepsilon_{k+d+1}$, $d' = d + \varepsilon_{k+d+1}$, $(k, d) \in \Gamma_3$, получим, что несмещенная оценка D_d''' равна

$$\hat{D}_d''' = \begin{cases} [N - (k' + d')] \frac{L(k', d' + 1)}{L(k', d' - 1) + L(k' - 1, d')}, & \text{при } (k', d' - 1) \text{ и } (k' - 1, d') \in \Gamma_3, \\ \varepsilon_{k'+d'} = 1, & \\ N - (k' + d') \text{ при } (k', d' - 1) \in \Gamma_3, & \\ (k' - 1, d') \in \Gamma_3, \varepsilon_{k'+d'} = 1, & \\ 0 \text{ в остальных случаях.} & \end{cases} \quad (1.29)$$

Для уменьшения дисперсий оценки $\hat{D}_d'''$ целесообразно изменить правило контроля, отбирая после принятия решения дополнительно для контроля не один, а несколько объектов. Для планов типа одноступенчатой выборки такие оценки были получены Лумельским [4]. Можно получить такие оценки и оценки дисперсий этих оценок для общих планов первого вхождения, однако в связи с ограниченностью объема книги мы их здесь не приводим.

Рассмотрим теперь вывод несмещенных оценок для предъявленного и принятого брака при использовании планов P_{23} . Для плана P_{23} координаты точки достижения границы $x \in \Gamma_2$ не являются достаточной статистикой, когда принимается решение $\mathcal{D}_2$. Достаточной статистикой будет

$$S = \begin{cases} (k, d), & x = (k, d) \in \Gamma_3, \\ (K, D), & x \in \Gamma_2, \end{cases}$$

где, напомним, Γ_i есть часть границы Γ , при достижении которой принимается решение $\mathcal{D}_i$. Хотя координаты точек достижения границы и не являются достаточной статистикой, их можно использовать для построения несмещенных оценок, которые затем улучшаются путем усреднения по достаточной статистике S . Итак, для несмещенного оценивания D можно в качестве исходной

оценки взять $\hat{D}$, определяемую соотношением (1.7). Заметим, что при $(k, d) \in \Gamma_2$ величина D известна. Улучшенная оценка для предъявленного брака D при контроле по плану Π_{23} имеет вид

$$\hat{D} = \begin{cases} N \frac{\sum_{(k, d) \in \Gamma_2} L((0, 1), (k, d)) p_D(k, d) / L(k, d)}{\sum_{(k, d) \in \Gamma_2} p_D(k, d)}, & \text{при } (k, d) \in \Gamma_2, \\ NL((0, 1), (k, d)) / L(k, d), & \text{при } (k, d) \in \Gamma_3. \end{cases} \quad (1.30)$$

Для оценки дисперсии оценки $\hat{D}$ можно использовать соотношение (1.9), где $\hat{D}$ вычисляется по формуле (1.30), а $\hat{O} = \hat{M}(2, 0) + \hat{D}$ получается на основе аналогичного улучшения оценки $\hat{M}(2, 0)$, вычисляемой по формуле (1.12).

Оценка, найденная по формуле (1.30) с учетом результатов сплошной разбраковки, имеет меньшую дисперсию, чем несмещенная оценка (1.7), однако она существенно сложнее. Расчет по формуле (1.30) практически возможен лишь с использованием ЭВМ. Программы, реализующие вычисление (1.30), не сложны. Если возможностей оперативного использования ЭВМ нет, то оценку предъявленного брака для планов Π_{23} можно проводить по формуле (1.7).

Такое же положение при планах Π_{23} и с трудностями вычисления оценок для принятого брака. В связи с этим приведем одно простое правило получения несмещенных оценок $\hat{D}'''$ на основе несмещенных оценок $\hat{D}$. Излагаемый ниже прием привлекателен еще и тем, что оценки дисперсий $\hat{D}''' - D'''$ совпадают с оценками дисперсий $\hat{D}$. Так как D''' при планах Π_{23} определяется по формуле, аналогичной (1.22) с очевидной заменой Γ_1 на Γ_2 , то для оценки, задаваемой соотношением

$$\hat{D}''' = \begin{cases} \hat{D} - d, & (k, d) \in \Gamma_3, \\ \hat{D} - D, & (k, d) \in \Gamma_2, \end{cases} \quad (1.31)$$

легко проверяется выполнение тождества

$$\hat{D}''' - D''' = \hat{D} - D. \quad (1.32)$$

Так как $\hat{D}$ — несмещенная оценка для D , то из (1.32) следует, что

$$E\hat{D}''' = ED'''.$$

Таким образом, D''' , определяемая в (1.31), является несмещенной оценкой для принятого брака. Если желательно использовать простую оценку, то $\hat{D}$ вычисляем по формуле (1.7). Если сложность вычисления не является препятствием, то используем улучшенную оценку (1.30).

Заметим, что в случае одноступенчатого плана контроля объема n , у которого приемочное число равно c , можно в (1.31) взять $\hat{D} = N \cdot (d/n)$. Тогда оценка дисперсии оценки D''' определяется соотношением (1.18). Это следует из тождества (1.32).

В § 2.4 было показано, что при некоторых априорных распределениях использование одноступенчатых планов с заданием приемочного числа c нецелесообразно. В таких случаях методы получения несмещенных оценок для D''' можно применить для оптимизации разбиения граничных точек на подмножества Γ_i , в которых принимается решение $\mathcal{D}_i$. Для определенности предположим, что контролируются совокупности $\mathfrak{F}_i \propto [N, D_i], i = 1, 2, \dots, m$, и для контроля используется план Π_{13} с границей Γ . Если при контроле $\mathfrak{F}$ путь достигает граничной точки $x_i = (k_i, d_i) \in \Gamma$, то проводится контроль дополнительного $n(i) = (k_i + d_i + 1)$ -го объекта. Результат дополнительного контроля $\varepsilon_{n(i)} = 1$, если объект дефектный, и $\varepsilon_{n(i)} = 0$ в противном случае. При любом разбиении границы $\Gamma = \Gamma_1 \cup \Gamma_3$ по формуле (1.29) можно найти $\hat{D}_{d,i}'''(\Gamma_1, \Gamma_3)$ — несмещенную оценку $D_{d,i}'''(\Gamma_1, \Gamma_3)$ — числа дефектных объектов, принимаемых потребителем. $D_{d,i}'''(\Gamma_1, \Gamma_3)$ определяется соотношением (1.27). При этом несмещенная оценка суммарного принятого брака равна

$$\hat{D}_d'''(\Gamma_1, \Gamma_3) = \sum_{1 \leq i \leq m} \hat{D}_{d,i}'''(\Gamma_1, \Gamma_3).$$

«Наилучшее» разбиение $\Gamma^* = \Gamma_1^* \cup \Gamma_3^*$ соответствует минимуму по всем разбиениям Γ , т. е.

$$\hat{D}_d'''(\Gamma_1^*, \Gamma_3^*) = \min_{\Gamma_1, \Gamma_3} \hat{D}_d'''(\Gamma_1, \Gamma_3).$$

Возможно, что найденное разбиение Γ_1^* , Γ_3^* будет согласовываться с выбором решений на основе эмпирического байесовского подхода, о котором упоминалось в конце § 2.4. Подчеркнем, что разбиение Γ_1^* , Γ_3^* ищется без предположения существования априорного распределения для числа дефектных объектов в контролируемых совокупностях.

В тех случаях, когда оправданным является предположение о существовании априорного распределения для D — числа дефектных объектов в $\mathfrak{F}$, представляет интерес построение несмещенных оценок для этого априорного распределения. Если соответствующие оценки найдены, то через них можно выразить несмещенные оценки для многих функционалов от априорного распределения. Такие задачи будут рассмотрены ниже в § 3.3.

В заключение отметим еще один возможный подход к задаче получения оценок для функций $f(x, D)$ от результата контроля $x = (k, d) \in \Gamma$ и неизвестного параметра D .

Определение 3.1.3. Статистики $f^+(x)$ и $f^-(x)$ называются *верхней и нижней оценками* для $f(x, D)$, если при любом значении D

$$Ef^-(x) \leq Ef(x, D) \leq Ef^+(x). \quad (1.33)$$

В основополагающей работе Колмогорова [4] отмечается целесообразность использования верхних и нижних оценок в задачах выборочного контроля. Рассмотрим общий метод построения верхних и нижних оценок для тех задач, в которых $Ef(x, D) = a_l D^l + a_{l-1} D^{l-1} + \dots + a_0$ — полином степени $l = l + \text{ind } \Gamma = l + \max_{(k, d) \in \Gamma} (k, d)$.

Именно такая ситуация возникает при оценивании D''' . Пусть $Q_l(D) = D^l + b_{l-1} D^{l-1} + \dots + b_0$ — полином степени l с единичным коэффициентом при D^l , а $Q_l^*(D)$ — полином с единичным коэффициентом при D^l , наименее уклоняющийся от нуля, т. е.

$$Q_l^* = \max_{D=0, \dots, N} |Q_l^*(D)| = \min_{\{b_0, \dots, b_{l-1}\}} \max_{D=0, \dots, N} |Q_l(D)|.$$

Тогда разность

$$\sum_{0 \leq k \leq l} a_k D^k - \sum_{(k, d) \in \Gamma} f^*(k, d) p_D(k, d) = a_l Q_l^*(D), \quad (1.34)$$

рассматриваемая как функция D , дает систему $l+1$ линейных уравнений относительно $f^*(k, d)$. Эту систему можно решить, если число точек $(k, d) \in \Gamma$ равно $l+1$ (план, соответствующий такой границе Γ , будет простым). В тех случаях, когда число точек Γ больше $l+1$, можно уменьшить число неизвестных $f^*(k, d)$, полагая их значения совпадающими на некоторых точках одинакового индекса. Если числа $f^*(k, d)$, $(k, d) \in \Gamma$ найдены, то верхняя и нижняя оценки с наименьшим смещением будут

$$f^+(k, d) = f^*(k, d) + a_l Q_l^*, \quad f^-(k, d) = f^*(k, d) - a_l Q_l^*.$$

Это утверждение непосредственно следует из (1.33) и (1.34).

В интересной работе Сираждинова [2] аналогичный метод используется для построения верхних и нижних оценок, когда рассматривается одноступенчатый план контроля объема n , а распределение числа d дефектных объектов, обнаруженных в выборке, является биномиальным распределением.

§ 3.2. Последующие интервальные оценки

Контроль с биномиальными выборками. Точечные несмещенные оценки, методы получения которых были подробно рассмотрены в предыдущем параграфе, недостаточно полно характеризуют оцениваемые величины. Это связано с тем, что значения точечных оценок являются случайными и иногда могут существенно отличаться от оцениваемых величин. Такие нежелательные отклонения могут встречаться часто, если оценивание производится по небольшому количеству данных, т. е. по небольшому числу проконтролированных объектов или совокупностей. В этих случаях можно использовать для оценки одномерного параметра доверительные интервалы, а для оценки векторного параметра — доверительные множества, определения которых были даны в § 2.1.

Результаты контроля последовательности совокупностей $\mathfrak{P}_i \propto [N_i, D_i]$ по планам первого вхождения $\Pi^{(i)}$ с границами точек остановки $\Gamma^{(i)}$, $1 \leq i \leq m$, представимы

в виде $\mathbf{x} = (k_1, d_1; \dots; k_m, d_m)$, $(k_i, d_i) \in \Gamma^{(i)}$. Если при плане $\Pi^{(i)}$ принимается решение $\mathcal{D}_2$, то соответствующие координаты (k_i, d_i) в $\mathbf{x}$ заменяются значениями (K_i, D_i) , где $K_i = N_i - D_i$ — число годных объектов в $\mathfrak{F}_i$. Задача состоит в построении таких статистик $\varphi_n(\mathbf{x})$, $\varphi_v(\mathbf{x})$ от результатов наблюдений $\mathbf{x}$, $\varphi_n(\mathbf{x}) \leq \varphi_v(\mathbf{x})$, что для заданной функции $\varphi(\mathbf{x}, D_1, \dots, D_m)$ при любых $D_1, \dots, D_m$ выполнено соотношение

$$P\{\varphi_n(\mathbf{x}) \leq \varphi(\mathbf{x}, D_1, \dots, D_m) \leq \varphi_v(\mathbf{x})\} \geq \gamma. \quad (2.1)$$

При выполнении (2.1) система случайных интервалов $[\varphi_n(\mathbf{x}), \varphi_v(\mathbf{x})]$ имеет коэффициент доверия не меньший γ . В задаче интервального оценивания предъявленного брака имеем

$$\varphi(\mathbf{x}, D_1, \dots, D_m) = D_\Sigma = \sum_{1 \leq i \leq m} D_i. \quad (2.2)$$

При интервальном оценивании пропущенного брака имеем

$$\varphi(\mathbf{x}, D_1, \dots, D_m) = D'_\Sigma = \sum_{1 \leq i \leq m} D'_i. \quad (2.3)$$

В (2.3) числа D'_i при планах $\Pi_{13}^{(i)}$ определяются по формуле (1.22), а при планах $\Pi_{13д}^{(i)}$ (с проверкой дополнительного объекта при уже принятом решении) — по формуле (1.27).

Итак, при заданном уровне доверия γ требуется дать рекомендации для получения γ -доверительных интервалов $[D_n, D_v]$, $[D'_n, D'_v]$ для D_Σ и D'_Σ , т. е. надо найти такие функции результатов контроля D_n, D_v, D'_n, D'_v , чтобы при любых возможных значениях $D_1, \dots, D_m$ были выполнены соотношения

$$P\left\{D_n \leq \sum_{1 \leq i \leq m} D_i \leq D_v\right\} \geq \gamma, \quad (2.4)$$

$$P\left\{D'_n \leq \sum_{1 \leq i \leq m} D'_i \leq D'_v\right\} \geq \gamma. \quad (2.5)$$

При этом, конечно, желательно получить такие γ -доверительные интервалы, которые были бы по возможности более узкими. Точное решение этих задач, по-видимому, является очень сложным. К настоящему времени имеются в основном лишь асимптотические рекомендации по-

строения γ -доверительных интервалов для функций (2.2), (2.3) в условиях, когда $m \rightarrow \infty$.

Краткое изложение возможных методов решения указанных задач начнем с рассмотрения примера, в котором γ -доверительный интервал для (2.2) при конечных значениях N_Σ , d_Σ находится просто. Для этого предположим, что планы Π контроля совокупностей $\mathfrak{F}_i$ являются биномиальными выборками (см. определение 2.2.2) с одним и тем же значением вероятности отбора ν , $0 < \nu < 1$. При больших значениях N_i контроль по биномиальным выборкам с заданным значением ν приближенно соответствует одноступенчатым планам, у которых объемы выборок $n_i \approx N_i \nu$. Как было показано в § 2.2, см. (2.2.5), d_i — числа обнаруженных дефектных объектов при использовании указанных биномиальных выборок, являются биномиально распределенными случайными величинами с параметрами ν и D_i , $1 \leq i \leq m$. Известно, что при сложении биномиально распределенных взаимно независимых случайных величин с одинаковыми значениями вероятностей «успехов» ν сумма $d_\Sigma = \sum_{1 \leq i \leq m} d_i$ будет также биномиально распределенной случайной величиной с параметрами $D_\Sigma = \sum_{1 \leq i \leq m} D_i$ и ν , т. е.

$$P\{d_\Sigma = d\} = \binom{D_\Sigma}{d} \nu^d \bar{\nu}^{D_\Sigma - d}. \quad (2.6)$$

Для краткости записи будем пользоваться обозначениями

$$B_D^d(\nu) = \sum_{0 \leq l \leq d} b_D^l(\nu), \quad \bar{B}_D^d(\nu) = 1 - B_D^d(\nu), \quad (2.7)$$

где $b_D^d(\nu) = \binom{D}{d} \nu^d \bar{\nu}^{D-d}$.

В математической статистике, см., например, Ван дер Варден [1], стр. 50, традиционным является изложение так называемого метода Клоппера — Пирсона построения γ -доверительного интервала для параметра ν биномиального распределения. При этом известным считается «число испытаний» D_Σ и «число успехов» d_Σ . Мы же в (2.6) пришли к специфической задаче построения γ -доверительного интервала для неизвестного

«числа испытаний» D_{Σ} при известной «вероятности успеха» ν и известном «числе успехов» d_{Σ} (Беляев, Лумельский [1]).

На решетке $\mathfrak{R}^2 = \{(D, d)\}$, $D, d = 0, 1, 2, \dots$, рассмотрим множества H_D , образованные точками (D, d) по следующему правилу: $(D, d) \in H_D$ тогда и только тогда, когда

$$B_D^d(\nu) \geq \varepsilon \text{ и } \bar{B}_D^{d-1}(\nu) \geq \varepsilon. \quad (2.8)$$

Здесь $\varepsilon = (1 - \gamma)/2$. Из соотношений (2.8) следует, что при $D_{\Sigma} = D$ вероятность того, что d_{Σ} порождает точку

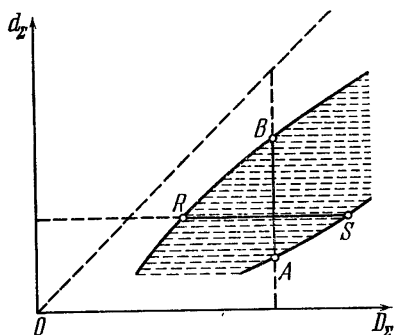


Рис. 3.2.1. Построение γ -доверительного интервала для предъявленного брака.

$(D, d_{\Sigma}) \in H_D$, не менее $\gamma = 1 - 2\varepsilon$, так как d_{Σ} есть биномиально распределенная случайная величина с параметрами D и ν . На рис. 3.2.1 множеству H_D соответствует вертикальный отрезок AB . Объединение всех множеств H_D дает множество $H(\gamma) = \bigcup_D H_D$. Объединение берется по всем возможным значениям $D_{\Sigma} = 0, 1, \dots$,

$$N_{\Sigma}, N_{\Sigma} = \sum_{1 \leq i \leq m} N_i. \text{ Из не-}$$

равенств $B_D^d(\nu) \geq B_{D+1}^d(\nu)$ следует что нижняя граница $H(\gamma)$ есть неубывающая функция D . Соответственно верхняя граница $H(\gamma)$ — также неубывающая функция D , что следует из неравенств $\bar{B}_D^d(\nu) \leq \bar{B}_{D+1}^d(\nu)$. Таким образом, горизонтальные сечения $H(\gamma)$ при фиксированном значении $d_{\Sigma} = d$ образуют горизонтальные отрезки RS ; см. рис. 3.2.1. Пусть (D_n, d) — левая точка, а (D_b, d) — правая точка горизонтального отрезка RS . Следующие события эквивалентны:

$$\{D_n \leq D_{\Sigma} \leq D_b\} = \{(D_{\Sigma}, d) \in H(\gamma)\} = \{(D_{\Sigma}, d) \in H_{D_{\Sigma}}\}. \quad (2.9)$$

Вероятность правого события в (2.9) не менее $\gamma = 1 - 2\varepsilon$ по построению. Эквивалентные события

имеют одинаковые вероятности. Следовательно, выполнено (2.4) и D_H , D_B образуют границы γ -доверительного интервала. Остается лишь указать, что D_H и D_B находятся как решения уравнений

$$B_{D_B+1}^{d_\Sigma}(\nu) < \varepsilon \leq B_{D_B}^{d_\Sigma}(\nu), \quad (2.10)$$

$$\bar{B}_{D_H-1}^{d_\Sigma-1}(\nu) < \varepsilon \leq \bar{B}_{D_H}^{d_\Sigma-1}(\nu). \quad (2.11)$$

Значения D_H , D_B , удовлетворяющие этим неравенствам, можно найти по таблицам неполной бэта-функции; см. Пирсон [1], Большев, Гладков, Щеглова [1]. Решения неравенств (2.10), (2.11) легко найти численными методами. В качестве примера расчетов приведем таблицу 3.2.1, в которой взяты значения $\nu = 0,1$, $\gamma = 0,95$ и $D_\Sigma = 0(5)100$. Для значений d_Σ , отличных от табличных, можно пользоваться линейной интерполяцией. Возникающие при этом ошибки не превышают 1.

Таблица 3.2.1

γ — доверительные границы для D_Σ при контроле
с биномиальными выборками ($\gamma=0,95$, $\nu=0,1$)

d_Σ	D_H	D_B	d_Σ	D_H	D_B
0	0	53	55	421	708
5	18	113	60	465	764
10	51	180	65	509	819
15	87	243	70	554	875
20	126	303	75	598	931
25	166	363	80	643	988
30	207	422	85	688	1041
35	249	480	90	733	1096
40	292	537	95	778	1151
45	334	595	100	823	1205
50	378	651			

Например, пусть в результате контроля с биномиальными выборками со значением $\nu = 0,1$ (т. е. в среднем контролировалось 10% объектов) было обнаружено всего $d_\Sigma = 72$ дефектных объекта. Требуется построить γ -доверительный интервал для суммарного числа дефектных объектов D_Σ , предъявленных на контроль,

когда $\gamma = 0,95$. Из таблицы 3.2.1 находим $D_n = 554 + (598 - 554) \cdot ^2/5 \approx 572$, $D_v = 875 + (931 - 875) \cdot ^2/5 \approx 897$. Следовательно, с вероятностью не меньшей 0,95 справедливы неравенства $572 \leq D_\Sigma \leq 897$.

Если функция распределения исхода наблюдения монотонно зависит от неизвестного параметра, то можно предложить сравнительно простой общий метод построения γ -доверительных интервалов. Заметим, что именно так обстоит дело в рассмотренном выше примере, так как вероятности $B_{D_\Sigma}^{d_\Sigma}(\nu)$ монотонно убывают с ростом D_Σ при фиксированных значениях d_Σ . Пусть случайная величина ξ принимает вещественные значения и $P\{\xi < x\} = G(x, \theta)$ есть непрерывная функция x и θ , где θ — вещественный параметр, значения которого лежат в интервале (θ_1, θ_2) , $-\infty < \theta_1 < \theta_2 < +\infty$. Если $G(x, \theta)$ монотонно возрастает или убывает с ростом θ , то верхняя $\theta_v(x)$ и нижняя $\theta_n(x)$ границы γ -доверительного интервала находятся как решения уравнений

$$G(x, \theta_v) = \varepsilon_v, \quad G(x, \theta_n) = 1 - \varepsilon_n$$

при монотонном росте $G(x, \theta)$, или уравнений

$$G(x, \theta_n) = \varepsilon_n, \quad G(x, \theta_v) = 1 - \varepsilon_v$$

при монотонном убывании $G(x, \theta)$, $\gamma = 1 - \varepsilon_n - \varepsilon_v$. Этот метод, предложенный Фишером [2], легко следует из простого факта, что случайная величина $G(\xi, \theta)$ имеет равномерное распределение на интервале $[0, 1]$, т. е. $P\{G(\xi, \theta) < y\} = y$.

Непосредственно использовать метод Фишера в рассматриваемых нами задачах нельзя, так как основные интересующие нас случайные величины принимают дискретные значения и соответствующие им распределения не являются непрерывными. Метод Фишера был обобщен на произвольные распределения Большевым [3]. В основе обобщения лежит следующая лемма Большева.

Лемма 3.2.1. Если для случайной вещественной величины ξ определена $G(x) = P\{\xi < x\}$, то для любого $y \in [0, 1]$ выполняются неравенства

$$P\{G(\xi + 0) \leq y\} \leq y \leq P\{G(\xi) < y\}. \quad (2.12)$$

Доказательство. Пусть $x_y = \inf \{x: G(x) \geq y\}$; тогда $G(x_y) \leq y \leq G(x_y + 0)$. Функция $G(x)$ непрерывна слева, поэтому $P\{G(\xi) < y\} = P\{\xi \leq x_y\} = G(x_y + 0) \geq y$. Аналогичным образом находим $P\{G(\xi + 0) \leq y\} = P\{\xi < x_y\} = G(x_y) \leq y$.

Так как рассматриваются распределения общего вида, то границы $\theta_n(x)$, $\theta_b(x)$ для γ -доверительных интервалов не всегда следует включать в γ -доверительный интервал, т. е. иногда целесообразно рассматривать открытые интервалы $(\theta_n(x), \theta_b(x))$ или открытые с одной стороны и замкнутые с другой, например, $[\theta_n(x), \theta_b(x))$ и т. п. Метод построения γ -доверительных интервалов, предложенный Большевым [3], содержится в формулировке следующего утверждения.

Теорема 3.2.1. *Нижняя (верхняя) граница γ -доверительного интервала $\theta_n(x)$ ($\theta_b(x)$) с коэффициентом доверия $\gamma = 1 - \varepsilon_n - \varepsilon_b$ находится как верхняя грань значений $\theta \in [\theta_1, \theta_2]$, удовлетворяющих неравенству*

$$G(x, \theta) \geq 1 - \varepsilon_n, \quad (2.13)$$

если при всех x функция $G(x, \theta)$ по θ монотонно убывает (возрастает), и неравенству

$$G(x + 0, \theta) \leq \varepsilon_b, \quad (2.14)$$

если при всех x функция $G(x, \theta)$ по θ монотонно возрастает (убывает).

Если $\theta_n(x)$ ($\theta_b(x)$) соответствующему неравенству не удовлетворяет, то $\theta_n(x)$ ($\theta_b(x)$) включается в доверительный интервал, а в остальных случаях не включается. Если решение соответствующего неравенства не существует, то следует положить $\theta_n(x) = \theta_1$ ($\theta_b(x) = \theta_2$).

Доказательство теоремы основано на использовании леммы 3.2.1. Его мы не приводим.

Метод, основанный на асимптотической нормальности распределений оценок. Так как выбор объектов из каждой совокупности $\mathcal{F}_i$, $1 \leq i \leq m$, производится независимо, то несмещенные оценки (2.2) и (2.3) являются суммами m взаимно независимых случайных величин. Основное допущение состоит в том, что мы будем предполагать нормальность распределений несмещенных

оценок (2.2) и (2.3), или, точнее, будем предполагать, что нормально распределены центрированные и нормированные величины

$$\xi = \frac{\sum_i \hat{D}_i - \sum_i D_i}{\sqrt{\sum_i v(\hat{D}_i)}}, \quad \xi''' = \frac{\sum_i \hat{D}_i''' - \sum_i D_i'''}{\sqrt{\sum_i v(\hat{D}_i''' - D_i''')}}. \quad (2.15)$$

Необходимо, конечно, каким-либо способом проверить гипотезу о нормальности функции распределения случайных величин ξ , ξ''' . Числители и знаменатели в (2.15) будут известными функциями D_i , выражения для которых можно было бы получить, усредняя соответствующие несмещенные оценки. Из сделанного допущения о нормальности распределений ξ , ξ''' следует, что при $\gamma = 1 - 2\varepsilon$ для u_ε , $\Phi(-u_\varepsilon) = \varepsilon$, имеем

$$P\{|\xi| \leq u_\varepsilon\} \approx \gamma, \quad P\{|\xi'''| \leq u_\varepsilon\} \approx \gamma. \quad (2.16)$$

При заданных значениях $D_1, \dots, D_m$ неравенства

$$\left| \frac{\sum_i \hat{D}_i - \sum_i D_i}{\sqrt{\sum_i v(\hat{D}_i)}} \right| \leq u_\varepsilon, \quad \left| \frac{\sum_i \hat{D}_i''' - \sum_i D_i'''}{\sqrt{\sum_i v(\hat{D}_i''' - D_i''')}} \right| \leq u_\varepsilon \quad (2.17)$$

выполняются только при некоторых результатах контроля $x = (k_1, d_1; \dots, k_m, d_m)$. Всю совокупность x , удовлетворяющих левому или правому неравенствам (2.17), обозначим H_D , H_D''' соответственно, где $D = (D_1, \dots, D_m)$.

Множества H_D , H_D''' являются аналогами вертикальных отрезков в примере контроля с биномиальными выборками. При фиксированных исходах x неравенства (2.17) можно рассматривать относительно наборов неизвестных нам значений $D_1, \dots, D_m$. Все наборы $D_1, \dots, D_m$, удовлетворяющие (2.17), образуют при фиксированных наблюдениях x множества H_x , H_x''' , которые являются аналогами горизонтальных отрезков в упомянутом примере. Так как события $(D_1, \dots, D_m) \in H_x (H_x''')$ эквивалентны событиям $x \in H_D (H_D''')$, то

$$P\{(D_1, \dots, D_m) \in H_x\} \approx \gamma, \quad (2.18)$$

$$P\{(D_1, \dots, D_m) \in H_x'''\} \approx \gamma. \quad (2.19)$$

Рассмотрим экстремальные значения функций D_{Σ} , D_{Σ}''' (2.2), (2.3) на множествах H_x , H_x''' :

$$D_H = \min_{(D_1, \dots, D_m) \in H_x} \sum_i D_i, \quad D_B = \max_{(D_1, \dots, D_m) \in H_x} \sum_i D_i, \quad (2.20)$$

$$D_H''' = \min_{(D_1, \dots, D_m) \in H_x'''} \sum_i D_i'', \quad D_B''' = \max_{(D_1, \dots, D_m) \in H_x'''} \sum_i D_i''. \quad (2.21)$$

Из (2.20) и (2.21) следует, что если $(D_1, \dots, D_m) \in H_x$ или $(D_1, \dots, D_m) \in H_x'''$, то заведомо выполнены неравенства

$$D_H \leq D_{\Sigma} \leq D_B, \quad D_H''' \leq D_{\Sigma}''' \leq D_B'''. \quad (2.22)$$

Следовательно, вероятности выполнения (2.22) не менее вероятностей (2.18) и (2.19). Таким образом, (2.20) и (2.21) определяют значения верхних и нижних границ γ -доверительных интервалов для D_{Σ} и D_{Σ}''' .

Итак, мы видим, что интересующая нас задача построения γ -доверительных интервалов для D_{Σ} и D_{Σ}''' сводится к задаче поиска экстремума по $(D_1, \dots, D_m)$ функций (2.2) и (2.3) на множествах наборов $(D_1, \dots, D_m)$, определяемых неравенствами (2.17). Это весьма сложная задача, решить которую в общем случае можно, лишь реализуя соответствующие алгоритмы на ЭВМ.

Для иллюстрации рассмотрим частный случай, когда контроль совокупностей $\mathfrak{F}_i \asymp [N_i, D_i]$ проводится по одноступенчатому плану $(n_i, c_i)_{13}$. Пусть $d_1, \dots, d_m$ — числа дефектных объектов, обнаруженных при таком контроле, $k_i = n_i - d_i$, $1 \leq i \leq m$. При построении доверительного интервала для предъявленного брака (2.2) используем несмещенные оценки $\hat{D}_i = (N_i/n_i)d_i$. Как следует из (1.3.12), дисперсия $\hat{D}_i$ равна

$$V(\hat{D}_i) = \left(\frac{N_i}{n_i}\right)^2 V(\delta_{n_i}) = D_i \left[\frac{N_i(N_i - n_i)}{n_i(N_i - 1)} - \frac{N_i - n_i}{n_i(N_i - 1)} D_i \right]. \quad (2.23)$$

Если исходить из допущения, что дробь ξ в (2.15) имеет распределение, близкое к стандартному нормальному при больших m , то из (2.15) и (2.23) получаем, что

γ -доверительные множества образованы наборами $(D_1, \dots, D_m)$, удовлетворяющими неравенствам

$$\begin{aligned} \sum_i D_i - u_\varepsilon \sqrt{\sum_i D_i \left[\frac{N_i(N_i - n_i)}{n_i(N_i - 1)} - \frac{N_i - n_i}{n_i(N_i - 1)} D_i \right]} &\leq \\ &\leq \sum_i (N_i/n_i) d_i \leq \sum_i D_i + \\ &+ u_\varepsilon \sqrt{\sum_i D_i \left[\frac{N_i(N_i - n_i)}{n_i(N_i - 1)} - \frac{N_i - n_i}{n_i(N_i - 1)} D_i \right]}. \end{aligned} \quad (2.24)$$

Согласно (2.20) нужно найти максимум и минимум оцениваемой функции $\varphi(D_1, \dots, D_m) = \sum_i D_i$ по области значений $(D_1, \dots, D_m)$, удовлетворяющих (2.24). Для этого можно воспользоваться следующим результатом Лумельского [5].

Теорема 3.2.2. Пусть дана линейная функция $f = \sum_i c_i \theta_i$, вещественных переменных θ_i , и множество значений $\theta_i \geq 0$, удовлетворяющих соотношениям

$$\sum_i c_i \theta_i - u_\varepsilon \sqrt{\sum_i \theta_i (\alpha_i - \beta_i \theta_i)} \leq \sigma, \quad (2.25)$$

$$\sum_i c_i \theta_i + u_\varepsilon \sqrt{\sum_i \theta_i (\alpha_i - \beta_i \theta_i)} \geq \sigma. \quad (2.26)$$

Коэффициенты u_ε , σ , c_i , α_i , β_i — положительные, $1 \leq i \leq m$. Если выполнено условие

$$\sum_i \alpha_i c_i / \beta_i \geq \sigma, \quad (2.27)$$

то максимум функции f достигается на границе, соответствующей равенству в (2.25), и значение его равно

$$f_{\text{в}} = \frac{\sigma + B u_\varepsilon^2 + u_\varepsilon \sqrt{2B\sigma + B^2 u_\varepsilon^2 - \sigma^2 E + A + A E u_\varepsilon^2}}{1 + E u_\varepsilon^2}, \quad (2.28)$$

а минимум f равен

$$f_{\text{н}} = \max \left\{ 0, \frac{\sigma + B u_\varepsilon^2 - u_\varepsilon \sqrt{2B\sigma + B^2 u_\varepsilon^2 - \sigma^2 E + A + A E u_\varepsilon^2}}{1 + E u_\varepsilon^2} \right\}. \quad (2.29)$$

Здесь

$$E = \left(\sum_i c_i^2 / \beta_i \right)^{-1}, \quad A = 1/4 \left[\sum_i \frac{\alpha_i^2}{\beta_i} - \left(\sum_i \frac{\alpha_i c_i}{\beta_i} \right)^2 E \right],$$

$$B = \frac{E}{2} \sum_i \frac{\alpha_i c_i}{\beta_i}.$$

Если дополнительно потребовать выполнения условия

$$\sigma > u_e \sqrt{A}, \quad (2.30)$$

то f достигает минимума на границе, определяемой равенством в (2.26), а значение $f_H > 0$ и равно второй компоненте под знаком максимума в (2.29).

В рассматриваемой нами задаче построения γ -доверительного интервала для предъявленного брака θ_i соответствуют $D_i, c_i = 1$,

$$\alpha_i = N_i(N_i - n_i)/(n_i(N_i - 1)), \quad \beta_i = (N_i - n_i)/(n_i(N_i - 1)),$$

$$\sigma = \sum_i \frac{N_i d_i}{n_i} \leq \sum_i \frac{\alpha_i}{\beta_i} = \sum_i N_i,$$

т. е. условие (2.27) выполнено. Следовательно, приближенные значения верхней и нижней доверительных границ для предъявленного брака можно найти по формулам (2.28), (2.29). Заметим, что в теореме 3.2.1 используется более широкая область значений, так как θ_i не предполагаются целыми. Кроме того, в нашем случае следовало бы рассматривать более узкую область, учитывая очевидные неравенства $D_i \geq d_i$. Следует иметь в виду, что при увеличении A ширина доверительного интервала также растет.

В частном случае $N_i = N, n_i = n, 1 \leq i \leq m$, приближенные γ -доверительные границы для $\sum_i D_i$ имеют вид

$$\frac{N}{n} d_\Sigma + \frac{u_e^2}{2} \frac{N - n}{N - 1} \pm u_e \left(d_\Sigma \frac{N - n}{N - 1} + \frac{u_e^2}{4} \left(\frac{N - n}{N - 1} \right)^2 - \frac{d_\Sigma (N - n)}{nm(N - 1)} \right)^{1/2}.$$

$$1 + \frac{u_e^2 (N - n)}{nm(N - 1)}$$

где $d_\Sigma = \sum_i d_i$, $(+u_e)$ соответствует верхней доверительной границе, а $(-u_e)$ — нижней ($\gamma = 1 - 2\varepsilon$).

Тем же самым способом можно получить приближенные выражения для границ γ -доверительных интервалов для пропущенного брака.

Напомним, что другой общий подход к получению приближенных формул для границ доверительных интервалов был предложен Колмогоровым [4], [5]. Этот подход был изложен выше в конце § 1.2. Приложение метода Колмогорова в задачах приемочного контроля имеется в работе Сираждинова [1].

Читатель, конечно, видит, что изложенная в этом параграфе теория нуждается в дальнейшем развитии. Неясно, например, сколь велики ошибки от допущения о нормальности распределений величин ξ , ξ''' , определяемых (2.15). Поэтому необходимо прежде всего найти четкие рекомендации возможности использования процедуры построения доверительных интервалов, основанной на нормальности распределения случайных величин (2.15). Быть может, ответ удастся получить, развив теорию аппроксимации распределений $\sum_i \hat{D}_i$, $\sum_i \hat{D}_i'''$ распределениями, близкими к нормальным. При таком подходе вместо «непосредственного» использования предельного распределения следует искать аппроксимации распределений $\sum_i \hat{D}_i$, $\sum_i \hat{D}_i'''$ соответствующим образом «деформированными» нормальными распределениями. Примером успеха такого подхода в задаче аппроксимации гипергеометрического распределения нормальным могут служить упоминавшиеся в § 1.3 результаты Моленара [1].

§ 3.3. Оценка априорного распределения числа дефектных объектов в совокупности

Основные соотношения при использовании планов первого вхождения. Изложенная выше в главе 2 теория байесовского подхода показывает, что выбор оптимального плана контроля существенным образом зависит от вида априорного распределения числа дефектных объектов в контролируемой совокупности. Рассматриваемая ниже в главе 4 система экономических планов контроля также исходит из предположения о существовании не-

которого априорного распределения, соответствующего «нормальному» ходу производства. Многие числовые характеристики контроля, в том числе и основные показатели его эффективности, могут быть представлены в виде функций от априорного распределения. Поэтому оценка априорного распределения необходима для обоснованного планирования приемочного контроля. Связанные с этим задачи еще очень далеки от окончательного решения. В этом параграфе будут использованы результаты Лумельского [7].

Итак, мы предположим, что в контролируемых совокупностях $\mathfrak{F}_i \asymp [N_i, D_i]$ числа дефектных объектов D_i являются реализациями случайных величин $D_{i\omega}$. Прежде всего следует заметить, что задача оценки априорного распределения становится существенно сложнее, если объемы N_i контролируемых совокупностей различны. В такой ситуации приходится учитывать, что априорные вероятности $f_N(D) = \mathbf{P}\{D_\omega = D\}$ зависят от объема N контролируемой совокупности. Если по результатам контроля совокупностей с объемами N_i требуется оценить $f_N(D)$, когда $N \neq N_i$, то необходимо исходить из определенного вида зависимости f_N от N . Это дополнительное предположение даст возможность «пересчитать» статистические данные контроля на совокупности, имеющие другие объемы. К сожалению, сейчас мы не можем привести содержательных рекомендаций по выбору вида зависимости f_N от N и поэтому ограничимся лишь контролем совокупностей одинакового объема N . При контроле совокупностей, у которых объемы принимают значения из множества, состоящего из небольшого набора целых чисел, приводимые ниже результаты могут быть применены отдельно для каждой из групп совокупностей, имеющих равные объемы. Как правило, успешное использование статистических оценок возможно тогда, когда с ростом количества статистических данных точность оценок также увеличивается. В связи с этим придется наложить ограничения на последовательность случайных величин $\{D_{i\omega}\}$. Прежде всего будем предполагать, что вероятности $f_N(D) = \mathbf{P}\{D_{i\omega} = D\}$ одинаковы для всех контролируемых совокупностей $\mathfrak{F}_i$, $1 \leq i \leq m$. Для того чтобы точность оценок величин $f_N(D)$ росла с увеличением m , только

что сформулированное предположение о стационарности одномерных распределений надо дополнить еще одним допущением об убывании зависимости между случайными величинами $D_{k\omega}$, $D_{l\omega}$ при увеличении $|k-l|$. Естественно, самым простым является допущение, что $D_{i\omega}$ — взаимно независимые и одинаково распределенные величины. Будем предполагать, что это допущение выполнено.

Дадим теперь точную формулировку решаемой задачи. Контролируется последовательность совокупностей $\mathfrak{F}_i \propto [N, D_{i\omega}]$, $1 \leq i \leq m$, в которой числа дефектных объектов $D_{i\omega}$ являются взаимно независимыми, одинаково распределенными случайными величинами. По результатам выборочного контроля требуется найти оценки для априорных вероятностей $f_N(D) = P\{D_{i\omega} = D\}$. Пусть для выборочного контроля используется план первого вхождения Π с границей Γ , $\text{ind } \Gamma = n$. Предположим далее, что семейство вероятностных мер $\{p_{D_j}(k, d)\}$, образованное вероятностями $p_{D_j}(k, d)$ достижения граничных точек $(k, d) \in \Gamma$, когда $D_{i\omega} = D_j$, является полным относительно любого заданного набора значений D_j , $0 \leq j \leq n$. Это означает, что из тождества

$$\sum_{(k, d) \in \Gamma} a(k, d) p_{D_j}(k, d) = 0, \quad 0 \leq j \leq n, \quad (3.1)$$

следует, что $a(k, d) \equiv 0$. Очевидно, что такие планы должны быть простыми. Типичным примером полного плана, удовлетворяющего (3.1), является одноступенчатый план объема n .

С учетом априорных вероятностей $f_N(D)$ получим, что вероятности $p(x)$ прекращения контроля в точке $x = (k, d) \in \Gamma$ удовлетворяют системе уравнений

$$p(x) = \sum_D p_D(x) f_N(D). \quad (3.2)$$

Система (3.2) является основной для определения априорных вероятностей $f_N(D)$. Значения $p_D(x)$ легко рассчитать по формуле (2.2.6), а величины $p(x)$ можно заменить значениями оценок $\hat{p}(x) = m(x)/m$, где $m(x)$ равно числу совокупностей, при контроле которых остановка произошла в точке x . Однако таким способом

нельзя получить оценки для $f_N(D)$, не накладывая дополнительных ограничений на априорные вероятности, так как число неизвестных значений $f_N(D)$ равно $N + 1$ ($0 \leq D \leq N$), тогда как число уравнений в (3.2) равно $n + 1$. Предположим, что $f_N(D) \geq 0$ только для заданных значений D_j , $0 \leq j \leq n$. При таком предположении искомыми неизвестными становятся числа $f_j = f_N(D_j) \geq 0$, $0 \leq j \leq n$, образующие вектор строку $\mathbf{f} = (f_0, \dots, f_n)$, $\sum_{0 \leq j \leq n} f_j = 1$. Из полноты плана Π следует, что его граница Γ состоит из $n + 1$ точки. Занумеруем возможные исходы $x \in \Gamma$ в виде последовательности $x_0, \dots, x_n$. При этом мы можем обходить границу Γ «справа налево», начав с точки x_0 , соответствующей нулевому числу дефектных объектов в выборке. Числа $a_{k,l} = p_{D_k}(x_l)$, $k, l = 0, \dots, n$, образуют матрицу, которую мы будем обозначать A . Числа $p_k = p(x_k)$ образуют вектор $\mathbf{p} = (p_0, \dots, p_n)$. При сделанных обозначениях уравнение (3.2) запишется в виде

$$\mathbf{p} = \mathbf{f}A. \quad (3.3)$$

Полнота плана Π согласно (3.1) означает, что строки матрицы линейно независимы, т. е. A — матрица полного ранга $n + 1$. Поэтому существует обратная матрица A^{-1} и решение уравнения (3.3) имеет вид

$$\mathbf{f} = \mathbf{p} \cdot A^{-1}. \quad (3.4)$$

Вектор $\hat{\mathbf{p}} = (\hat{p}_0, \dots, \hat{p}_n)$, $\hat{p}_l = m(x_l)/m$ есть несмещенная оценка для $\mathbf{p} = (p_0, \dots, p_n)$. Поэтому из уравнения (3.4) немедленно следует, что несмещенная оценка $\hat{\mathbf{f}}$ вектора $\mathbf{f}$ имеет вид

$$\hat{\mathbf{f}} = \hat{\mathbf{p}}A^{-1}. \quad (3.5)$$

Согласно (3.5) ковариационная матрица $V(\hat{\mathbf{f}}) = E(\hat{\mathbf{f}} - \mathbf{f})(\hat{\mathbf{f}} - \mathbf{f})' = E(\hat{\mathbf{f}}_i - f_i)(\hat{\mathbf{f}}_j - f_j)'$, $0 \leq i, j \leq n$, равна

$$V(\hat{\mathbf{f}}) = E(A^{-1})' \hat{\mathbf{p}}' \hat{\mathbf{p}} A^{-1} - (A^{-1})' \mathbf{p}' \mathbf{p} A^{-1}. \quad (3.6)$$

О точности найденных оценок $\hat{\mathbf{f}}$ можно было бы судить по значениям элементов ковариационной матрицы (3.6), которые, однако, неизвестны. Поэтому найдем несмещенную оценку $V(\hat{\mathbf{f}})$. Так как мы предполагаем, что случай-

ные величины $D_{i\omega}$ являются взаимно независимыми, то числа $m(x_i)$ наблюдений остановок контроля в точках x_i , $0 \leq i \leq n$, имеют полиномиальное распределение

$$P\{m(x_i) = m_i, 0 \leq i \leq n\} = \frac{m!}{m_0! \dots m_n!} p_0^{m_0} \dots p_n^{m_n}. \quad (3.7)$$

Непосредственной проверкой, исходя из (3.7), легко показать, что несмещенными оценками для p_i^2 , $p_i p_j$ являются

$$m_{i,i} = \frac{m(x_i)(m(x_i) - 1)}{m(m-1)}, \quad m_{i,j} = \frac{m(x_i)m(x_j)}{m(m-1)}, \\ i \neq j, \quad 0 \leq i, j \leq n.$$

Таким образом, матрица $M = \|m_{i,j}\|$ будет несмещенной оценкой для матрицы $p'p = \|p_i p_j\|$, а матрица $(A^{-1})' M A^{-1}$ будет несмещенной оценкой для $(A^{-1})' p' p A^{-1}$. Как обычно, несмещенность означает, что математические ожидания элементов матрицы оценок совпадают с соответствующими элементами оцениваемой матрицы. Таким образом, несмещенная оценка $V(\hat{f})$, как это следует из (3.5), имеет вид

$$\hat{V}(\hat{f}) = (A^{-1})' \{\hat{p}' \hat{p} - M\} A^{-1}. \quad (3.8)$$

Суммируем полученные результаты.

Теорема 3.3.1. Пусть в результате контроля последовательности совокупностей $\mathfrak{P}_i \asymp [N, D_{i\omega}]$ по плану первого вхождения Π с границей Γ , обладающему полной (3.1) относительно значений D_j таких, что $f_N(D) = 0$, $D \neq D_j$, m_j раз наблюдается значение $x_j \in \Gamma$, $0 \leq j \leq n$. Тогда вектор $\hat{f}$ несмещенных оценок компонент вектора $f = (f_0, \dots, f_n)$ априорных вероятностей $f_j = f_N(D_j)$ определяется соотношением (3.5), а несмещенной оценкой матрицы ковариаций компонент $\hat{f}$ является $\hat{V}(\hat{f})$, определяемая (3.8).

Используя этот результат, можно легко найти несмещенную оценку для линейной функции

$$b(f) = \sum_{0 \leq j \leq n} b_j f_j = b f', \quad (3.9)$$

которая согласно (3.5) равна

$$\hat{b}(f) = \sum_{0 \leq j \leq n} b_{j|f} \hat{f}_j = b \hat{f}' = b(A^{-1})' \hat{p}'. \quad (3.10)$$

Здесь $b = (b_0, \dots, b_n)$. Из (3.10) и (3.8) можно получить формулу для несмещенной оценки $\hat{V}(\hat{b})$ дисперсии оценки $\hat{b} = \hat{b}(f)$, которая равна

$$\hat{V}(\hat{b}) = b \hat{V}(\hat{f}) b'. \quad (3.11)$$

Заметим, что поиск несмещенных оценок для широкого класса линейных функций (3.9) можно было бы проводить и без ограничительного допущения о том, что $f_N(D) = 0$ для $D \neq D_j$, $0 \leq j \leq n$. Изложение соответствующей общей теории имеется в книгах Закса [1], глава 3, Сирли [1], главы 2, 3, Рао [1], глава 3.

Применение общих формул. Примеры. Математическое ожидание суммарного числа дефектных объектов в проконтролированных совокупностях равно $m \sum_{0 \leq j \leq n} D_{j|f} \hat{f}_j$, а несмещенная оценка, соответственно, равна $m \sum_{0 \leq j \leq n} D_{j|f} \hat{f}_j$.

Еще одним примером линейных несмещенных оценок могут служить оценки для $p_r^F(d)$ — вероятностей обнаружения d дефектных при контроле r случайным образом отобранных из $\mathfrak{F}$ объектов. Как мы уже видели в § 2.4, эти вероятности играют существенную роль при планировании приемочного контроля на основе байесовского подхода. В частности, при одноступенчатом плане контроля объема n условное математическое ожидание числа дефектных объектов в непроконтролированной части совокупности при условии, что в выборке обнаружено d дефектных объектов, в соответствии с (2.4.13) равно

$$D(d) = (N - n) \frac{d+1}{n+1} p_{n+1}^F(d+1) / p_n^F(d). \quad (3.12)$$

По формуле (2.4.9) найдем, что

$$p_n^F(d) = \sum_{0 \leq j \leq n} h_{N, D_j}^{n, d} \hat{f}_j, \quad p_{n+1}^F(d+1) = \sum_{0 \leq j \leq n} h_{N, D_j}^{n+1, d+1} \hat{f}_j. \quad (3.13)$$

Согласно (3.10) несмещенные оценки для вероятностей в (3.13) равны

$$\begin{aligned}\hat{p}_n^F(d) &= \sum_{0 \leq j \leq n} h_{N, D_j}^{n, d} \hat{f}_j, \\ \hat{p}_{n+1}^F(d+1) &= \sum_{0 \leq j \leq n} h_{N, D_j}^{n+1, d+1} \hat{f}_j.\end{aligned}\quad (3.14)$$

В условиях сформулированного выше предположения о взаимной независимости $D_{i\omega}$ оценки (3.14) по вероятности сходятся при $m \rightarrow \infty$ к оцениваемым величинам (3.13). Если $p_n^F(d) \neq 0$, то состоятельной при $m \rightarrow \infty$ оценкой для величины $D(d)$, определяемой (3.12), будет

$$\hat{D}(d) = (N - n) \frac{d+1}{n+1} \hat{p}_{n+1}^F(d+1) / \hat{p}_n^F(d). \quad (3.15)$$

Статистику (3.15) можно использовать при выборочном контроле совокупностей $\mathfrak{F}_i$ с отсроченным решением, когда сначала проводится выборочный контроль, а затем по результатам обработки всех данных контроля принимается общее решение о том, какие $\mathfrak{F}_i$ бракуются и какие принимаются. Исходя из статистики (3.15), следует браковать те совокупности, которым соответствуют наибольшие (и недопустимые) значения $\hat{D}(d)$. Подчеркнем, что существенным моментом в предлагаемом подходе является то, что при построении эмпирических байесовских оценок $\hat{D}(d)$ не требуется контролировать дополнительный $n+1$ -й объект. Этот подход, предложенный Лумельским [7], может быть использован, конечно, только при выполнении исходных допущений, наиболее ограничительным из которых является требование $f_N(D) = 0$, $D \neq D_j$, $0 \leq j \leq n$.

Приведем пример (Лумельский [7]). Пусть контролируются совокупности $\mathfrak{F}_i \propto [N, D_{i\omega}]$, $1 \leq i \leq 10^4$, $N = 6$. Из каждой совокупности берется случайная выборка без возвращения объема $n = 3$. Результаты контроля следующие: $m_0 = 8154$, $m_1 = 877$, $m_2 = 871$, $m_3 = 98$. Здесь m_d — число совокупностей, в выборках из которых обнаружено d дефектных объектов. Предположим, что $f_6(4) = f_6(5) = f_6(6) = 0$, $D_0 = 0$, $D_1 = 1$, $D_2 = 2$, $D_3 = 3$. Элементы матрицы A : $a_{i,j} = h_{6,i}^{3,j}$, $j \leq i$,

$a_{i,j} = 0, j > i$. Поэтому получаем

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0,5 & 0,5 & 0 & 0 \\ 0,2 & 0,6 & 0,2 & 0 \\ 0,05 & 0,45 & 0,45 & 0,05 \end{pmatrix}, \quad A^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 2 & -6 & 5 & 0 \\ -10 & 36 & -45 & 20 \end{pmatrix}. \quad (3.16)$$

Из (3.16) и (3.5) имеем

$$\hat{f} = \hat{p} \cdot A^{-1} = (0,8039, 0,0056, -0,0055, 0,1960). \quad (3.17)$$

По формулам (3.14) и (3.15), используя (3.17), найдем, что

$$\hat{D}(0) = 0,0368, \quad \hat{D}(1) = 1,9738, \quad \hat{D}(2) = 1,0126, \quad \hat{D}(3) \approx 0.$$

Из полученных результатов следует, что при обнаружении в выборке одного дефектного объекта соответствующие совокупности надо подвергнуть сплошной разбраковке в первую очередь, так как в непроконтролированной части таких партий в среднем содержится около двух дефектных объектов. Совокупности, в выборках из которых обнаружено три дефектных объекта, следует принять без контроля. Такой вывод в рассматриваемом примере согласуется с исходным предположением, что число дефектных объектов в партиях не может превышать трех.

При обработке реальных данных необходимо убедиться в правомерности предположения $f_N(D) = 0$ для $D > n$. В рассмотренном иллюстративном примере данные были получены с помощью таблиц случайных чисел, причем $f_6(D) = 0$ для $D = 4, 5, 6$.

При контроле по планам с решениями $\mathcal{D}_2$ и $\mathcal{D}_3$ оценка априорных вероятностей $f_N(D)$ может быть получена без ограничений на число значений D , при которых $f_N(D)$ положительны. Например, при контроле по плану $(n, c)_{23}$ результаты контроля представимы значениями $x = (k, d)$, если $d \leq c$, $k + d = n$, или $x = (N - D, D)$, если $d > c$. Здесь, как и выше, d равно числу дефектных объектов, обнаруженных в выборке объема n . Исходными уравнениями для определения $f_N(D)$ являются по-прежнему уравнения системы (3.2). Число уравнений равно числу неизвестных $f_N(D)$, т. е.

равно $N + 1$. Пусть X — фактическое число дефектных объектов, обнаруженных при контроле $\mathfrak{F}$ (включая сплошную разбраковку, если таковая производится), т. е.

$$X = \begin{cases} d, & d \leq c, \\ D, & d > c. \end{cases}$$

Элементы матрицы A равны $a_{i,j} = P\{X=j | D=i\}$. Для рассматриваемых планов контроля

$$a_{i,j} = h_{N,i}^{n,j}, \text{ если } 0 \leq i \leq N - n + c, j \leq c;$$

$$a_{i,i} = \bar{h}_{N,i}^{n,c}, \text{ если } c + 1 \leq i \leq N - n + c;$$

$$a_{i,i} = 1, \text{ если } i > N - n + c;$$

$$a_{i,j} = 0, \text{ если } i < j \text{ или } i > j, \text{ но либо } j \geq c + 1, \text{ либо } i > N - n + c, j \leq c.$$

Матрица A с такими элементами является треугольной, а следовательно, невырожденной. Поэтому вектор несмещенных оценок $\hat{f} = (\hat{f}_0, \dots, \hat{f}_N)$ для априорных вероятностей $f_N(0), \dots, f_N(N)$ находится по формуле (3.5). Инамов [1], [2] нашел в явном виде элементы обратной матрицы A^{-1} и назвал их оценочными функциями. Случай с приемочным числом $c = 0$ подробно рассмотрен в статье Эйдельманта и Инамова [1], [2], где приведен числовой пример расчета основных показателей контроля на основе оценок априорного распределения.

В работах Инамова [1], [2] и Эйдельманта, Инамова [1], [2] решалась также задача построения несмещенных оценок для чисел $m(D)$ предъявленных на контроль совокупностей $\mathfrak{F}_i$, $1 \leq i \leq m$, в которых содержалось D дефектных объектов, $\sum_D m(D) = m$.

Постановки задач. Закончим этот небольшой параграф постановками задач, решение которых в некотором смысле дает возможность избежать предварительного задания тех D , для которых $f_N(D) = 0$. Как правило, с ростом объема N контролируемой совокупности используются планы $P \in \Pi$, у которых $\text{ind } \Gamma = n$ также увеличивается. В связи с этим рассмотрим постановки задач оценки априорных распределений при $N \rightarrow \infty$, $n \rightarrow \infty$.

Множество $S(f) = \{D: f_N(D) > 0\}$ называется **носителем априорного распределения**, а число входящих

в $S(f)$ точек — **мощностью носителя**. Пусть $F(N, n)$ — множество всех возможных априорных распределений $\hat{f} = (f_N(0), \dots, f_N(N))$, мощности носителей которых не превышают $n + 1$. Заметим, что $F = F(N, N)$ соответствует классу всех априорных распределений. Рассмотрим следующие классы априорных распределений:

- (1) $F_1(N, n) = \{f: S(f) \subseteq \{D: D = [N/n]d, \quad 0 \leq d \leq n\}\},$
- (2) $F_q(N, n) = \{f: S(f) \subseteq \{0, 1, \dots, [nq]\} \cup$
 $\cup \{D: D = [nq] + [N/n]d, \quad d = 0, \dots, [(1 - q)n] +$
 $+ \chi_A \{(1 - q)n - [(1 - q)n]\}\}\}, \quad A = (0, \infty),$
- (3) $F^*(N, l) = \{f: \hat{f} \in F(N, l), \quad f_N(D) =$
 $= r/(l + 1), \quad r \in [0, l + 1], \quad l \leq n\}.$

Для класса $F_1(N, n)$ носитель может быть любым подмножеством арифметической прогрессии с шагом $[N/n]$. Для класса $F_q(N, n)$ возможные значения носителя заполняют целые точки интервала $[0, [nq]]$, а затем следуют арифметической прогрессии с шагом $[N/n]$. В частности, при $q = 1$ носитель $\hat{f} \in F_1(N, n)$ сосредоточен на множестве $0 \leq D \leq n$. Априорные распределения класса F_q можно использовать для аппроксимации распределений, соответствующих так называемому «нормальному» ходу производства. Вероятностные распределения класса $F^*(N, l)$ могут иметь произвольные носители, мощности которых не превышает $l + 1$, однако вероятности возможных значений D для них кратны $1/(l + 1)$.

Ограничившись априорными распределениями $\hat{f} \in F_q(N, n)$, можно использовать изложенные в этом параграфе линейные методы получения несмещенных оценок распределений. Если же исходное распределение $\hat{f} \in F_q(N, n)$, то найденные такими методами оценки $\hat{f}$ уже не будут несмещенными оценками для $\hat{f}$. Однако в некотором смысле найденные оценки $\hat{f}$ будут «близки» к истинным. Например, если последовательность векторов $b_N = (b_{N1}, \dots, b_{NN})$ определяет линейные функционалы $b_N \hat{f}'$, обладающие свойством «гладкости»

$$\max_{\hat{f} \in F} \min_{\tilde{f} \in F_q(N, n)} |b_N \hat{f}' - b_N \tilde{f}'| \rightarrow 0, \quad N \rightarrow \infty, \quad n \rightarrow \infty,$$

денное таким способом априорное распределение и будет искомым $f \in F^*(N, l)$. При статистическом подходе вместо $p(x)$ используются оценки $\hat{p}(x) = m(x)/m$, $x = (k, d) \in \Gamma$. Если исходное распределение $f \in F^*(N, l)$ и вместо $p(x)$ используются оценки $\hat{p}(x)$, то, конечно, корни полинома (3.21) не обязательно целые положительные числа, заключенные в интервале $[0, N]$. По-видимому, могут быть и отрицательные и комплексные корни. Как в таких случаях следует строить приближения для априорного распределения f ? Вместо уравнений (3.19) можно выписать уравнения для смешанных факториальных моментов $E(k^{(l)}d^{(r)})$ координат точки остановки $x = (k, d)$, которые потом заменяются оценками. После этого задача аналогичным способом сводится к отысканию корней полинома со случайными коэффициентами.

Возникает вопрос, является ли такой подход к оценке априорных распределений рациональным?

§ 3.4. Выборочный контроль по качественному признаку

Основные вероятностные модели. О контроле по качественному признаку было кратко сказано в начале § 1.1. Напомним, что при контроле по качественному признаку каждый объект совокупности $\mathfrak{P}$ включается только в один из r классов $\mathcal{K}_1, \dots, \mathcal{K}_r$. В частности, к этой схеме приходим в том случае, когда каждый из объектов $O_i \in \mathfrak{P}$ характеризуется скалярной величиной X_i , значение которой становится известной в результате контроля O_i . Объект O_i включается в класс $\mathcal{K}_j$, если

$$X_i \in G_j, \bigcup_{j=1}^r G_j = R^1 = (-\infty, +\infty), G_i \cap G_j = \emptyset. \text{ Типичным}$$

является случай группирования данных контроля, когда $G_j = [a_{j-1}, a_j]$, $a_{j-1} < a_j$, $1 \leq j \leq r$, $a_0 = -\infty$, $a_r = +\infty$. Аналогичным образом контроль по качественному признаку может проводиться при характеристизации объектов $O_i \in \mathfrak{P}$ векторными величинами $X_i \in R^m$. В этом случае множества G_j будут подмножествами в m -мерном пространстве. В некоторых моделях теории надежности объекты классифицируются по типам отказов: объект $O_i \in \mathcal{K}_1$, если в результате контроля

(испытаний на надежность) наблюдается только обрыв; $O_i \in \mathcal{K}_2$, если наблюдается только замыкание; $O_i \in \mathcal{K}_3$, если имеет место только постепенный уход параметров за зону допустимых значений. Если возможны сочетания отказов разных типов, то выделяется класс $\mathcal{K}_4$ отказов с одновременным обрывом и замыканием, но без постепенного ухода параметров за зону допустимых отказов и т. д.

Итак, считаем, что классы $\mathcal{K}_1, \dots, \mathcal{K}_r$ определены и в контролируемой совокупности $\mathfrak{P}$ имеется N_j объектов, принадлежащих классу $\mathcal{K}_j$. Кратко такую контролируемую совокупность обозначим $\mathfrak{P} \asymp [N, N_1, \dots, N_{r-1}]$, $N = N_1 + \dots + N_r$. Числа N_j нам неизвестны, и всю информацию об их значениях мы можем получить в результате выборочного контроля. Эта схема обобщает подробно рассмотренный выше случай контроля по альтернативному признаку, в котором $r = 2$, $N_1 = D$, $N_2 = K -$ числа дефектных и годных объектов в $\mathfrak{P}$. Ниже будет показано, что многие факты, доказанные для контроля по альтернативному признаку, имеют прямые аналоги для контроля по качественному признаку.

Рассмотрим целочисленную решетку $\mathfrak{R}^r = \{(n_1, \dots, n_r) : n_j = 0, 1, 2, \dots, 1 \leq j \leq r\}$. Результатам контроля объекта $O_i \in \mathfrak{P}$ будем сопоставлять вектор $e_{(j)} = (e_1, \dots, e_r)$, у которого при включении O_i в класс $\mathcal{K}_j$ только $e_j = 1$, а все $e_k = 0$, $k \neq j$. Результаты последовательного контроля объектов $\mathfrak{P} \asymp [N, N_1, \dots, N_{r-1}]$ представимы в виде пути по точкам $\mathfrak{R}^r$, когда переход в каждую следующую точку соответствует сдвигу на вектор e , определяемый по результатам контроля. Основным рассматриваемым здесь методом отбора объектов на контроль по качественному признаку является случайная выборка без возвращения. Определение плана первого вхождения аналогично определению 2.2.1 с очевидным изменением, что множество граничных точек Γ задано в $\mathfrak{R}^r$. При одноступенчатом плане контроля объем выборки фиксирован, поэтому для таких планов $\Gamma = \{(n_1, \dots, n_r) : n_1 + \dots + n_r = n, n_j \in [0, n]\}$. Схематически граница Γ этого плана изображена на рис. 3.4.1, а, где $r = 3$. Двухступенчатым планам соответствуют две гиперплоскости: $\Gamma^{(1)} = \{(n_1, \dots, n_r) :$

$n_1 + \dots + n_r = n^{(1)}$ — гиперплоскость результатов первой выборки объема $n^{(1)}$, и $\Gamma^{(2)} = \{(n_1, \dots, n_r) : n_1 + \dots + n_r = n^{(1)} + n^{(2)}\}$ — гиперплоскость результатов двух выборок. В $\Gamma^{(1)}$ выделено подмножество $\Gamma_{\text{пр}}^{(1)}$ проходимых точек, при попадании в которые необходимо взять вторую выборку объема $n^{(2)}$. Двухступенчатый план для $r=3$ схематически изображен на рис. 3.4.1, б,

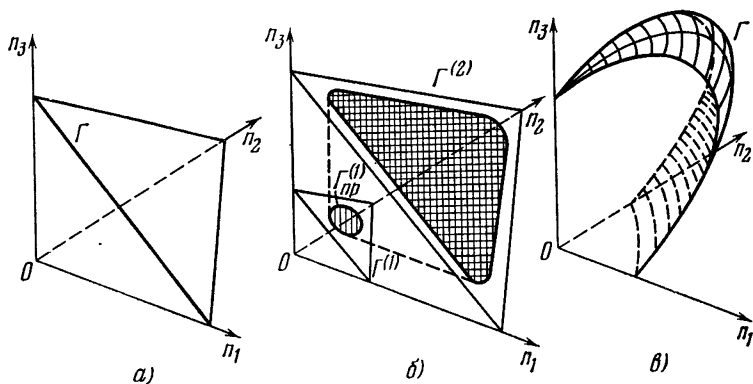


Рис. 3.4.1. Примеры границ планов первого вхождения при контроле по качественному признаку: а) граница одноступенчатого плана, б) граница двухступенчатого плана, в) граница общего плана.

где заштрихованы области проходимых точек границы $\Gamma_{\text{пр}}^{(1)}$ и достижимых точек границы $\Gamma^{(2)}$. В общем случае можно представлять, что плану Π первого вхождения ($\Pi \in \Pi$) соответствует некоторая «поверхность» Γ точек остановки (см. рис. 3.4.1, в).

Пусть $e_1, \dots, e_n$ — результаты контроля последовательно извлекаемых объектов согласно случайной выборке без возвращения из $\mathfrak{P} \asymp [N, N_1, N_2, \dots, N_{r-1}]$. Точно так же как и в § 2.2, можно получить, что вероятность результатов контроля равна

$$P(e_1, \dots, e_n) = \frac{N_1^{(n_1)} \dots N_r^{(n_r)}}{N^{(n)}}, \quad (4.1)$$

где $n = n_1 + \dots + n_r$, n_j — число объектов, отнесенных к классу $\mathcal{K}_j$, $1 \leq j \leq r$. Формула (4.1) обобщает (2.2.3). В терминах вероятностей путей точки из $\mathfrak{R}$ классифицируются на **достижимые**, **границные** и **прохо-**

димые точно так же, как это сделано в определении 2.2.3. Пусть $L_\Gamma(x)$ — число путей, соединяющих точку $0 = (0, \dots, 0)$ с достижимой (граничной) точкой $x = (n_1, \dots, n_r)$ и не задевающих других граничных точек. Из (4.1) находим, что вероятность прохождения (достижения граничной) точки x равна

$$p_\Gamma(x) = L_\Gamma(x) \frac{N_1^{(n_1)} \dots N_r^{(n_r)}}{N^{(n)}}. \quad (4.2)$$

Для одноступенчатого плана объема n число путей, соединяющих 0 с $x = (n_1, \dots, n_r)$ равно $\frac{n!}{n_1! \dots n_r!}$. Таким образом, из (4.2) находим, что вероятность обнаружить $n_1, \dots, n_r$ объектов r классов при контроле совокупности $\mathfrak{P} \asymp [N, N_1, \dots, N_{r-1}]$ по одноступенчатому плану объема n равна

$$h_{N, N_1, \dots, N_r}^{n, n_1, \dots, n_r} = \frac{n!}{n_1! \dots n_r!} \frac{N_1^{(n_1)} \dots N_r^{(n_r)}}{N^{(n)}}, \quad (4.3)$$

где $n_j \in [0, n]$, $n_1 + \dots + n_r = n$.

Распределение, вероятности которого определяются по формуле (4.3), называется **многомерным гипергеометрическим распределением**. Практически все факты, связанные с гипергеометрическим распределением и рассмотренные в § 1.3, можно обобщить на многомерное гипергеометрическое распределение. Обозначим $\eta_{n,j}$ случайное число объектов класса $\mathcal{K}_j$, обнаруженных при контроле по одноступенчатому плану объема $\eta_{n,1} + \dots + \eta_{n,r} = n$, а $(N_j - \eta_{n,j})$ — число объектов j -го класса в непроконтролированной части $\mathfrak{P}_{N-n} \asymp [N - n, N_1 - \eta_{n,1}, \dots, N_{r-1} - \eta_{n,r-1}]$. Легко выводится следующая формула для смешанных факториальных моментов (см. Абдурахманов [1]):

$$\begin{aligned} E \prod_{j=1}^r \eta_{n,j}^{(l_j)} (N_j - \eta_{n,j})^{(m_j)} &= \\ &= \frac{n^{(l)} (N - n)^{(m)}}{N^{(l+m)}} N_1^{(l_1+m_1)} \dots N_r^{(l_r+m_r)}, \end{aligned} \quad (4.4)$$

$$l = \sum_j l_j, \quad m = \sum_j m_j, \quad 1 \leq j \leq r,$$

которая является обобщением (1.3.14),

Вычисление вероятностей (4.2) весьма трудоемко. Поэтому для расчета вероятностных показателей планов первого вхождения можно пользоваться рекуррентными соотношениями. Например, для $p_{\Gamma}(z, \Gamma_i)$ — вероятностей того, что путь из точки z достигает граничной точки из подмножества $\Gamma_i \subset \Gamma$ (при этом будет принято решение $\mathcal{D}_i$), рекуррентные соотношения имеют вид

$$p_{\Gamma}(z, \Gamma_i) = \sum_{1 \leq j \leq r} p_{\Gamma}(x^{(j)}, \Gamma_i) \frac{N_j - n_j}{N - n}, \quad (4.5)$$

$$x^{(j)} = z + e_{(j)},$$

которые вместе с граничными условиями

$$p_{\Gamma}(z, \Gamma_i) = \begin{cases} 1, & z \in \Gamma_i, \\ 0, & z \in \Gamma \setminus \Gamma_i, \end{cases} \quad (4.6)$$

дают возможность найти $p_{\Gamma}(\Gamma_i) = p_{\Gamma}(0, \Gamma_i)$ — вероятность выхода пути контроля в граничные точки подмножества $\Gamma_i \subset \Gamma$, $0 = (0, \dots, 0)$. Рекуррентные соотношения (4.5) фактически являются вариантами формулы полной вероятности, в которой взаимно непересекающиеся события соответствуют извлечению на контроль объектов разных типов. Легко выписать рекуррентные соотношения и граничные условия, аналогичные (4.5), (4.6), и для других характеристик плана контроля, например, для среднего объема контроля. Точно так же можно использовать методы обратного и прямого счета с метками (см. рис. 2.2.3) для расчета числа возможных путей $L_{\Gamma}(x)$ или для расчета $L_{\Gamma}(x^{(1)}, x^{(2)})$ — числа возможных путей, выходящих из точки $x^{(1)}$ и достигающих точки $x^{(2)}$.

Расчет $p_{\Gamma}(z, \Gamma_i)$ можно провести, суммируя вероятности (4.2) по точкам области Γ . В. Лаборатории статистических методов МГУ Филимоновой С. С. составлены программы для счета вероятностей $p_{\Gamma}(z, \Gamma_i)$ для одноступенчатых, двухступенчатых планов контроля, а также для некоторых общих типов планов первого вхождения, когда число классов равно трем.

Используем биномиальную выборку (определение 2.2.2) для контроля совокупности $\mathfrak{F} \asymp [N, N_1, \dots, N_{r-1}]$.

Как уже указывалось, биномиальная выборка может быть реализована в два этапа: на первом этапе получается значение случайного объема η выборки, $P\{\eta=n\} = \binom{N}{n} v^n \bar{v}^{N-n}$, а затем на втором этапе из $\mathfrak{B}$ берется случайная выборка без возвращения объема n . В соответствии с (4.3) находим, что вероятность обнаружить n_j объектов класса $\mathcal{K}_j$, $1 \leq j \leq r-1$, при биномиальной выборке равна

$$\begin{aligned} p(n_1, \dots, n_{r-1}) &= \sum_{0 \leq n \leq N} \binom{N}{n} v^n \bar{v}^{N-n} h_{N, N_1, \dots, N_{r-1}}^{n, n_1, \dots, n_{r-1}} = \\ &= \frac{N_1^{(n_1)}}{n_1!} \dots \frac{N_{r-1}^{(n_{r-1})}}{n_{r-1}!} v^{n_1 + \dots + n_{r-1}} \bar{v}^{N_1 + \dots + N_{r-1} - (n_1 + \dots + n_{r-1})} \times \\ &\quad \times \sum_{n_r = n - (n_1 + \dots + n_{r-1}) \geq 0} \frac{N_r^{(n_r)}}{n_r!} v^{n_r} \bar{v}^{N_r - n_r} = \\ &= \prod_{j=1}^{r-1} \binom{N_j}{n_j} v^{n_j} \bar{v}^{N_j - n_j}. \quad (4.7) \end{aligned}$$

Таким образом, при биномиальной выборке числа обнаруженных объектов из классов $\mathcal{K}_1, \dots, \mathcal{K}_{r-1}$ есть взаимно независимые биномиальные случайные величины с вероятностью «успеха» v и числами «испытаний», равными числам объектов этих классов в контролируемой совокупности. Этот факт обобщает результат (2.2.5). Его можно использовать для построения интервальных последующих оценок N_j и критериев проверки гипотез об изменениях значений N_j в последовательности контролируемых совокупностей.

Точкам гиперплоскости $\{(k_1, \dots, k_r): k_1 + \dots + k_r = k, k_j \in [0, k]\}$ сопоставим вероятности

$$m_q^k = m_{q_1, \dots, q_r}^{k_1, \dots, k_r} = \frac{k!}{k_1! \dots k_r!} q_1^{k_1} \dots q_r^{k_r}, \quad (4.8)$$

где $0 \leq q_j \leq 1$, $\sum_j q_j = 1$, $q_j^0 \equiv 1$. Соответствующее вероятностям (4.8) распределение называется **полиномиальным**. **Случайное блуждание** по точкам решетки $\mathfrak{R}^r$ называется **полиномиальным**, если вероятности переходов из точки

$k = (k_1, \dots, k_{j-1}, k_j, k_{j+1}, \dots, k_r)$ в точку $k_{(j)} = (k_1, \dots, k_{j-1}, k_j + 1, k_{j+1}, \dots, k_r)$ равны q_j , $\sum_j q_j = 1$, независимо от предыдущих переходов. При полиномиальном блуждании вероятность результатов $(e_1, \dots, e_n)$ равна $q_1^{n_1} \dots q_r^{n_r}$, где n_j — число переходов вдоль j -й координатной оси.

Пусть контролируется совокупность $\mathfrak{F} \asymp [N, N_{1\omega}, \dots, N_{r\omega}]$, в которой числа объектов разных классов $N_{j\omega}$ являются случайными величинами. Вероятность значений этих случайных величин считаем равными $P\{N_{j\omega} = N_j, 1 \leq j \leq r\} = m_{q_1, \dots, q_r}^{N_1, \dots, N_r}$. Другими словами, априорное распределение $N_{j\omega}, 1 \leq j \leq r$, есть полиномиальное. Из соотношений (4.1), (4.8) находим

$$\begin{aligned} P\{e_1, \dots, e_n\} &= \\ &= \sum_{\substack{N_j, N_1 + \dots + N_r = N \\ 1 \leq j \leq r}} m_{q_1, \dots, q_r}^{N_1, \dots, N_r} \frac{N_1^{(n_1)} \dots N_r^{(n_r)}}{N^{(n)}} = \\ &= q_1^{n_1} \dots q_r^{n_r} \sum_{\substack{N_j \geq n_j, \\ N_1 + \dots + N_r = N, \\ 1 \leq j \leq r}} \left[\frac{(N - n)!}{(N_1 - n_1)! \dots (N_r - n_r)!} \times \right. \\ &\quad \left. \times q_1^{N_1 - n_1} \dots q_r^{N_r - n_r} \right] = q_1^{n_1} \dots q_r^{n_r}. \end{aligned}$$

Следовательно, при априорном полиномиальном распределении и случайной выборке без возвращения результаты контроля соответствуют полиномиальному блужданию.

Широкий класс априорных распределений образуют **смеси полиномиальных распределений**, для которых значения $q_j = q_{j\omega}$ считаются случайными величинами с ф. р. $G(q_1, \dots, q_r) = P\{q_{1\omega} \leq q_1, \dots, q_{r\omega} \leq q_r\}$. Для смесей полиномиальных распределений можно доказать утверждения, аналогичные изложенным выше в § 2.2.3 для G -биномиальных априорных распределений. Некоторые предельные теоремы для смесей полиномиальных распределений были получены Сираждиновым и Абдурахмановым [1].

Обобщение схемы Пойа на контроль по качественному признаку содержится в работе Джанардана и Патила [1]*). В ней рассматривается схема контроля совокупности $\mathfrak{F} \asymp [N, N_1, \dots, N_{r-1}]$, которую ниже мы будем называть *обобщенной схемой Пойа*. Отбор объектов на контроль является случайным, т. е. каждый из объектов может быть извлечен для контроля с равной вероятностью. При извлечении объекта из класса $\mathcal{K}_j$ в совокупность $\mathfrak{F}$ добавляется $c+1$ объектов того же класса $\mathcal{K}_j$. Для компактной формулировки результатов снова используем обозначение

$$a^{(x, c)} = a(a+c) \dots (a+(x-1)c).$$

Можно показать, что в обобщенной схеме Пойа вероятность обнаружить в результате контроля n объектов n_j объектов из класса $\mathcal{K}_j$ равна

$$p(n_1, \dots, n_j, c, N_1, \dots, N_r) = \frac{n!}{n_1! \dots n_r!} \prod_{j=1}^r \frac{N_j^{(n_j, c)}}{N^{(n, c)}}. \quad (4.9)$$

Распределение с вероятностями (4.9) называется *многомерным распределением Пойа*. Кратко такое распределение будем обозначать $MP(N, N_1, \dots, N_{r-1}, c, n)$. При $c=0$ обобщенная схема Пойа соответствует случайной выборке с возвращением, при $c=-1$ — случайной выборке без возвращения. Распределение (4.9) соответствует результатам контроля $\mathfrak{F} \asymp [N, N_{1\omega}, \dots, N_{r-1, \omega}]$, когда $N_{j\omega}$ имеют априорное распределение с вероятностями

$$\begin{aligned} P\{N_{j\omega} = N_j, 1 \leq j \leq r\} = \\ = \frac{N! \Gamma(\omega)}{N_1! \dots N_r! \prod_{j=1}^r \Gamma(\omega_j)} \int q_1^{\omega_1-1} \dots q_r^{\omega_r-1} dq_1 \dots dq_r, \end{aligned} \quad (4.10)$$

где интегрирование ведется по области $\sum_{1 \leq j \leq r} q_j = 1$, $0 \leq q_j \leq 1$. Другими словами, (4.10) означает, что априорное распределение является смесью полиномиальных распределений, когда плотность вероятностей зна-

*) На эту работу обратил мое внимание Линник Ю. В.

чений $q_{j\omega} = q_j$ имеет вид $\frac{\Gamma(\omega)}{\prod_{j=1}^r \Gamma(\omega_j)} q_1^{\omega_1-1} \dots q_r^{\omega_r-1}$. Рас-

пределение с плотностью вероятностей такого вида называется **распределением Дирихле**.

Многомерное распределение Пойа обладает интересными свойствами. Отметим лишь некоторые из них. Если вместо классов $\mathcal{K}_1, \dots, \mathcal{K}_r$ рассмотреть классы $\mathcal{K}'_1, \dots, \mathcal{K}'_s, s < r$, где $\mathcal{K}'_l, 1 \leq l \leq s$, получены в результате объединения в непересекающиеся подмножества классов $\mathcal{K}_j, 1 \leq j \leq r$, то $MP(N, N_1, \dots, N_{r-1}, c, n)$ преобразуется в $MP(N, N'_1, \dots, N'_{s-1}, c, n)$ -распределение. Здесь N'_l равно сумме тех N_j , которым соответствуют классы $\mathcal{K}_j$, включаемые в $\mathcal{K}'_l$. Пусть в обобщенной схеме Пойа результаты контроля $(n_1, \dots, n_r)$ имеют $MP(N, N_1, \dots, N_{r-1}, c, n)$ -распределение, где $n = n_1 + \dots + n_r$. Тогда условное распределение $n_1, \dots, n_k$ при фиксированных значениях $n_{k+1}, \dots, n_r$ есть $MP(N - \sum_{k+1 \leq j \leq r} N_j, N_1, \dots, N_{k-1}, c, n - \sum_{k+1 \leq j \leq r} n_j)$ -распределение. Условное распределение $n_1, \dots, n_l$ при условии, что $n_1 + \dots + n_l = m$, также будет MP -распределением. Условное математическое ожидание n_{r-1} при заданных значениях $n_1, n_2, \dots, n_{r-2}$ равно

$$E\{n_{r-1} | n_1, \dots, n_{r-2}\} = np_{r-1} - \frac{p_{r-1}}{p_r + p_{r-1}} \left\{ \sum_{j=1}^{r-2} (n_j - np_j) \right\},$$

где $p_j = N_j/N, 1 \leq j \leq r$.

Перечисленные выше свойства многомерного распределения Пойа приведены в работе Джанардана и Патила [1]. Там же рассмотрены простейшие предельные теоремы, когда параметры N и n стремятся к ∞ .

Так же как и в § 2.4, байесовский подход целесообразно использовать для выбора рациональных планов контроля по качественному признаку. Например, ответ на вопрос о выборе решения при выходе пути контроля в граничную точку $(n_1, \dots, n_r) \in \Gamma$ можно получить из значений условных математических ожиданий $q_i^F(n_1, \dots, n_r) = E\left\{ \frac{N_{j\omega} - \eta_j}{N - n} \mid \eta_k = n_k, 1 \leq k \leq r \right\}$, равных доле объектов класса $\mathcal{K}_j$ в непроконтролированной части $\mathfrak{P}$,

Здесь через η_j обозначены числа объектов из класса $\mathcal{K}_j$, обнаруженных к моменту выхода пути контроля на границу Γ . При вычислении $q_j^F(n_1, \dots, n_k)$ усреднение проводится с учетом априорного распределения F случайных величин $N_{1\omega}, \dots, N_{r\omega}$. В тех случаях, когда априорное распределение неизвестно, следует использовать эмпирический байесовский подход. Для одноступенчатых планов контроля можно получить необходимые соотношения, аналогичные рассмотренным в § 2.4.

Задачи статистического оценивания. Кратко наметим пути возможного обобщения результатов статистического оценивания при контроле по альтернативному признаку на контроль по качественному признаку. Прежде всего заметим, что из формулы (4.2) следует, что при контроле по плану первого вхождения с границей Γ достаточной статистикой являются координаты точки останова. План первого вхождения будем называть **замкнутым**, если вероятность выхода на границу точек останова тождественно равна единице при любых наборах чисел $N_1, \dots, N_r$, $N_1 + \dots + N_r = N$, объектов классов $\mathcal{K}_j$, $1 \leq j \leq r$, составляющих контролируемую совокупность $\mathfrak{B}$.

Для замкнутых планов первого вхождения легко получить обобщение результатов § 3.1. Предположим, что нам требуется найти несмещенную оценку для функции $\Phi(N_1, \dots, N_r) = N_1^{(k_1)} \dots N_r^{(k_r)}$, причем для простоты предположим, что точка $y = (k_1, \dots, k_r)$, $k_1 + \dots + k_r = k$ является проходимой. Проходимость точки означает, что она достижима и не является граничной (аналогично определению 2.2.3). Пусть $x = (n_1, \dots, n_r) \in \Gamma$ — координаты останова, а $L_\Gamma(y, x)$ — число возможных путей, соединяющих точку y с точкой x . Из (4.2) и тождества

$$\begin{aligned} N^{(k)} \sum_{x \in \Gamma} \frac{L_\Gamma(y, x)}{L_\Gamma(x)} p_\Gamma(x) &= \sum_{x \in \Gamma} L_\Gamma(y, x) \frac{N_1^{(n_1)} \dots N_r^{(n_r)}}{(N - k)^{(n - k)}} = \\ &= N_1^{(k_1)} \dots N_r^{(k_r)} \times \\ &\times \sum_{x - y \in \Gamma - y} L_{\Gamma - y}(x - y) \frac{(N_1 - k_1)^{(n_1 - k_1)} \dots (N_r - k_r)^{(n_r - k_r)}}{(N - k)^{(n - k)}} = \\ &= N_1^{(k_1)} \dots N_r^{(k_r)}, \quad (4.11) \end{aligned}$$

следует, что несмещенная оценка $\hat{\phi}$ для произведения $N_1^{(k_1)} \dots N_r^{(k_r)}$ имеет вид

$$\hat{\phi} = N^{(k)} \frac{L_{\Gamma}(y, x)}{L_{\Gamma}(x)}. \quad (4.12)$$

Заметим, что при выводе (4.11) была использована замкнутость плана со сдвинутой на вектор y границей $\Gamma - y$, которая является следствием замкнутости исходного плана с границей Γ . Используя (4.12), можно легко вывести формулы несмещенных оценок для N_j , оценки дисперсий и ковариаций этих оценок и т. д.

Рассмотрим в заключение важный вопрос о полноте семейства вероятностных мер, которое образовано вероятностями выхода пути контроля на точки границы Γ плана первого вхождения. Необходимость решения задачи о полноте планов первого вхождения при контроле по качественному признаку отмечается в работах Линника и Романовского [1], [2]. Напомним, что Линником и Романовским [1] была получена теорема об эквивалентности полноты плана первого вхождения при контроле по качественному признаку со случайным выбором с возвращением и при контроле со случайным выбором без возвращения. Как уже видел читатель в § 3.1, ответ на вопрос о полноте при контроле со случайным выбором без возвращения по альтернативному признаку получен. Положение дел при контроле по качественному признаку, когда число классов не менее трех, существенно сложнее. Приводимые ниже достаточные условия полноты отличаются от достаточных условий, данных в работах Линника и Романовского [1], [2].

Вернемся к методу доказательства достаточности простоты плана с границей Γ для полноты этого плана при контроле по альтернативному признаку с использованием случайной выборки без возвращения. Точке $M = (k, d)$ линии L , $k + d = \text{ind } L \geq \text{ind } \Gamma$, сопоставим прямоугольник $K(M)$, стороны которого совпадают с осями значений k и d , а вершина которого лежит на L ; см. рис. 3.4.2, а. Каждой точке $(k, d) \in \Gamma$ поставим в соответствие два таких прямоугольника, с вершинами в точках $M_k = (k, \text{ind } L - k)$, $M_d = (\text{ind } L - d, d)$. На

рис. 3.4.2, а точкам $x' = (k', d')$, $x'' = (k'', d'')$ соответствуют заштрихованные прямоугольники $K(M_{d'})$, $K(M_{k''})$. Введем следующую **операцию исключения граничных точек**. Исключение точки $x = (k, d) \in \Gamma$ считается допустимым в том и только в том случае, когда хотя бы в одном из подмножеств $\Gamma \cap K(M_d)$ или $\Gamma \cap K(M_k)$ все отличные от x граничные точки уже исключены. Исключение граничных точек $(k, 0)$, $(0, d)$ допустимо. Для простого плана допустимо исключение

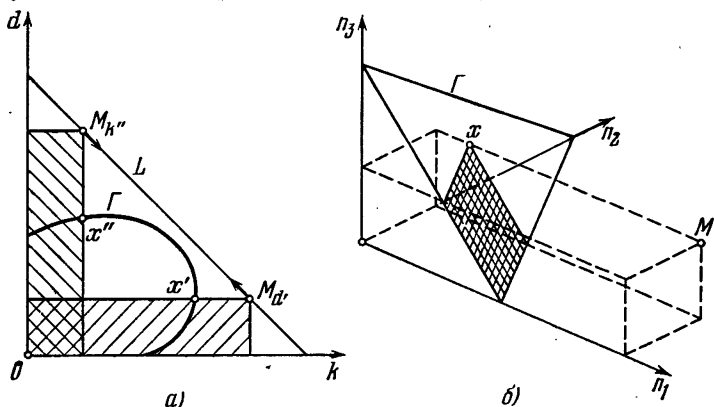


Рис. 3.4.2. Исключение точек а) при контроле по альтернативному признаку, б) при контроле по качественному признаку.

всех граничных точек. Для этого нужно вершины $M_{d'}$, $M_{k''}$ прямоугольников сдвигать от осей навстречу друг другу. Можно показать, что эта геометрическая конструкция эквивалентна методу, рассмотренному в § 3.1 при доказательстве достаточности простоты для полноты плана.

Операция исключения точек очевидным образом обобщается на планы первого вхождения при контроле по качественному признаку с любым числом классов r . Отличие состоит в том, что точке $x = (n_1, \dots, n_r) \in \Gamma$ сопоставляется r прямоугольников с гранями, образованными координатными осями и вершинами в точках

$$M_j = (n_1, \dots, n_{j-1}, \text{ind } L - \sum_{k \neq j} n_k, n_{j+1}, \dots, n_r),$$

$$L = \{(k_1, \dots, k_r): \text{ind } k = k_1 + \dots + k_r = N\}, \text{ ind } L = N.$$

Исключение точки $x \in \Gamma$ допустимо, если хотя бы в одном из подмножеств $\Gamma \cap K(M_{nj})$ все граничные точки, отличные от x , уже исключены. Используя понятие исключения граничных точек, приходим к следующему утверждению.

Теорема 3.4.1. *Если план Π первого вхождения допускает исключение всех его граничных точек, то он полный.*

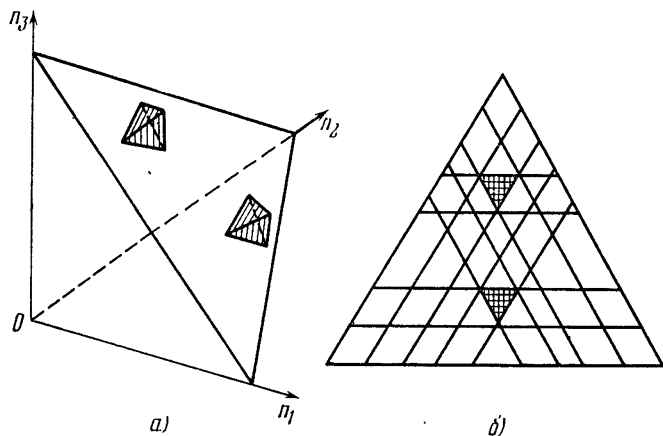


Рис. 3.4.3. Пример полного плана, предложенный А. Н. Мальковым.

Это достаточное условие полноты легко проверяется для многих классов планов первого вхождения. Например, легко убедиться, что одноступенчатый план контроля допускает исключение всех его граничных точек; см. рис. 3.4.2, б. Возникает вопрос, а не является ли допустимость исключения всех граничных точек и необходимым условием? Конечно, несмотря на сравнительную несложность метода исключения, хотелось бы найти какие-то простые геометрические характеристики границы плана первого вхождения, которые были бы либо критерием простоты либо давали достаточные условия полноты (или, наоборот, неполноты). Хотелось бы считать, что аналогом простоты может служить связность множества проходимых точек. Однако это не так. А. Н. Мальков предложил автору несложный пример

полного плана, допускающего исключение всех его граничных точек, для которого проходимые точки с одинаковым значением индекса образуют несвязные множества, т. е. разделены граничными и недостижимыми точками. В этом примере граничные точки образованы гиперплоскостью точек x , $n_1 + \dots + n_r = n$, и кусками двух параллелепипедов, лишь частично выходящих за гиперплоскость; см. рис. 3.4.3, а. Эти параллелепипеды размещены в соответствии с рис. 3.4.3, б, т. е. не «зате-няют» друг друга при просмотре вдоль любой из координатных осей.

ГЛАВА 4

СИСТЕМЫ ПЛАНОВ ПРИЕМОЧНОГО КОНТРОЛЯ ПО АЛЬТЕРНАТИВНОМУ ПРИЗНАКУ

§ 4.1. Основные принципы построения систем планов приемочного контроля

Системы планов контроля. Выше были подробно рассмотрены вопросы контроля отдельно взятой совокупности объектов с использованием конкретного плана выборочного контроля. Результаты контроля служили основой для принятия соответствующего решения. В промышленности при приемочном контроле качества партий (совокупностей) готовой продукции или при входном приемочном контроле партий комплектующих изделий представляют интерес не столько результаты контроля одной партии (совокупности) изделий, сколько результаты контроля последовательности партий. Из общих задач управления качеством продукции следует, что правила контроля отдельных совокупностей должны быть частью системы правил контроля, т. е. **стандарта статистического приемочного контроля**. Использование одинаковых стандартов статистического приемочного контроля на различных предприятиях, выпускающих однотипные виды продукции, дает возможность объективного сравнения уровней качества продукции этих предприятий.

Итак, мы рассматриваем задачу построения системы правил контроля последовательности совокупностей объектов $\mathfrak{P}_i, i = 1, 2, \dots$, предъявляемых на контроль. Предполагается, что объекты по результатам контрольной операции классифицируются на годные и дефектные, т. е. $\mathfrak{P}_i \asymp [N_i, D_i]$. Система правил контроля указывает, какой план контроля Π_i должен быть использо-

ван для совокупности $\mathfrak{F}_i$. При этом можно (а вообще говоря, и следует) учитывать результаты контроля предшествующих совокупностей $\mathfrak{F}_j$, $j < i$. Естественно, что система правил должна учитывать возможные или ожидаемые изменения в последовательности $\{D_i\}$. В первом приближении можно считать, что последовательность $\{[N_i, D_i]\}$ соответствует «нормальному ходу» процесса производства, который изредка прерывается интервалами «разладок». При разладках на контроль предъявляются совокупности $\mathfrak{F}_i$, содержащие повышенное количество дефектных объектов. Система правил контроля должна включать: 1) правила перехода на более жесткий контроль в тех случаях, когда результаты контроля свидетельствуют о возможно появившейся разладке, 2) правила возвращения к нормальному контролю. В некоторых стандартах допускается при устойчивом достижении высокого качества переход к облегченному контролю. Наконец, в систему правил контроля должны быть включены методы оценки рекомендуемых показателей эффективности системы контроля, равно как и последующие оценки уровня качества контролируемого процесса производства. Таким образом, основными элементами системы правил контроля являются следующие: (1) таблицы планов выборочного контроля, используемые в условиях нормального хода производства, (2) таблицы планов контроля, используемые в специальных случаях (в условиях разладок для усиленного контроля, при достижении высокого качества для облегченного контроля и т. п.), (3) правила выбора планов из таблиц с учетом конкретных условий проведения контроля, (4) правила переключения на усиленный контроль и обратного перехода на контроль при нормальном ходе производства и т. п., (5) методы вычисления последующих оценок уровня и других показателей качества контролируемого процесса, а также (6) методы оценки показателей эффективности используемой системы правил контроля на основе полученных данных о ходе контроля. С общей точки зрения, охватывающей всю систему правил контроля последовательности совокупностей $\mathfrak{F}_i$, числовые характеристики отдельных планов контроля (например, оперативные характеристики) рассматриваются как «локальные», тогда

как система в целом приобретает новые качества адаптации, которые можно характеризовать общими («интегральными») показателями. Примеры общих числовых показателей систем планов контроля будут рассмотрены ниже в следующем параграфе.

Нормальный ход производства и разладки. Контролируемая последовательность совокупностей $\mathfrak{F}_i$ при контроле по альтернативному признаку полностью характеризуется значениями чисел D_i дефектных объектов. При стационарном характере производства, когда стремятся создавать однотипные изделия, числа D_i можно рассматривать как последовательность случайных величин. При этом каждой совокупности $\mathfrak{F}_i \propto [N_i, D_i]$ сопоставляется функция распределения $F_{N_i}(D) = P(D_i \leq D) = \sum_{D_j \leq D} f_{N_i}(D_j)$, $f_{N_i}(D) = P(D_i = D)$. Проблемы

оценки априорного распределения на основе результатов выборочного контроля были подробно рассмотрены выше в § 3.3. Эти проблемы не простые. Тем не менее решать их нужно, так как для правильной организации приемочного контроля весьма важно знать вид априорного распределения числа дефектных объектов в контролируемых совокупностях. В некоторых случаях на начальных этапах организации контроля такую информацию можно получить путем сплошного контроля некоторого числа совокупностей. В этом случае оценками вероятностей $f_{N_i}(D)$ будут частоты $\frac{m(D, N_i)}{m(N_i)}$, где $m(N_i)$ — число совокупностей с объемом N_i , прошедших сплошной контроль, а $m(D, N_i)$ — число тех совокупностей, в которых было обнаружено D дефектных объектов, $\sum_D m(D, N_i) = m(N_i)$.

Проиллюстрируем сказанное примером. Пусть сплошному контролю были подвергнуты 108 совокупностей, объемы которых равны $N_i = 1000$. Результаты контроля для значений $D \leq 15$ показаны на рис. 4.1.1. Число горизонтальных черточек над D соответствует числу совокупностей, в которых было выявлено ровно D дефектных объектов. Таким образом, наличие четырех черточек в столбце с $D = 0$ означает, что из 108 совокупностей только в четырех не было дефектных объектов,

16 черточек в столбце с $D = 1$ означает, что один дефектный объект был обнаружен в 16 совокупностях и т. д. В рассматриваемом примере

$$\begin{aligned} m(1000) &= 108, & m(0, 1000) &= 4, & m(1, 1000) &= 16, \\ m(2, 1000) &= 21, & m(3, 1000) &= 23, & m(4, 1000) &= 17, \\ m(5, 1000) &= 9, & m(6, 1000) &= 6, & m(7, 1000) &= \\ & & & & &= m(12, 1000) = 2, \\ m(D, 1000) &= 1, \text{ для } D = 8, 10, 13, 14, 29, 41, 80, 89. \end{aligned}$$

Частотные оценки $m(D, 1000)/m(1000)$ для $D \leq 6$, по-видимому, близки к оцениваемым вероятностям $f_{N_i}(D)$.

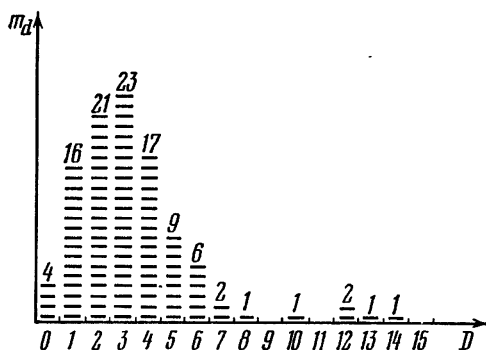


Рис. 4.1.1. Результаты сплошного контроля 108 совокупностей ($N_i = 1000$, $D \leq 15$).

Наличие совокупностей с большими D , которые наблюдаются всего лишь несколько раз (один-два раза), говорит о том, что в соответствующей зоне значений D , $f_{N_i}(D) > 0$. Ограничимся интервалом значений D , $0 \leq D \leq 9$, в который попадает основная часть контролируемых совокупностей, 99 из 108. Распределение вероятностей, соответствующее интервалу $[0, 9]$, следует считать типичным для данного процесса производства, тогда как совокупности, содержащие $D \geq 10$, можно отнести к выпущенным в условиях разладок. Такой вывод был бы более объективным, если до приемочного контроля были бы измерены основные параметры тех-

нологического процесса. Под нормальным ходом технологического процесса будем понимать такое его состояние, когда основные требования технологии соблюдены. Предположим, что о каждой из 108 совокупностей известно, изготовлена она при нормальном ходе производства или при существенных отклонениях параметров технологического процесса от нормы. Пусть в примере оказалось, что совокупности, содержащие до 8 дефектных объектов, изготовлены при нормальном ходе производства. Совокупности, содержащие $D = 10, 13, 14$, изготовлены при наличии незначительных отклонений параметров, а совокупности с $D = 29, 41, 80, 89$ выпущены при серьезных нарушениях хода производства. Из рис. 4.1.1 видно, что при условии нормального хода производства число дефектных объектов D в контролируемых совокупностях можно считать случайной величиной, имеющей функцию распределения $F_{n, N_i}(D) \approx \approx \sum_{d \leq D} [m(d, N_i)/m(N_i)]$. Значения этой случайной величины в основном попадают в интервал $[0, 9]$. В рассматриваемом примере 9 из 108 совокупностей имеют $D > 9$. Мы говорим «в основном», так как изредка и при нормальном ходе производства могут быть изготовлены совокупности с $D > 9$. Однако доля таких совокупностей будет близка к $9/108$. Итак, мы приходим к понятию априорного распределения $F_{n, N}(D)$, $N_i = N$, числа дефектных объектов, выпущенных при нормальном ходе производства.

Если технологический процесс хорошо отлажен и дефектные объекты составляют малую долю от общего числа произведенных объектов, то имеются определенные предпосылки считать распределение $F_{n, N}(D)$ близким к биномиальному. Точно биномиальное распределение будет в том случае, когда дефектные объекты производятся независимо от предшествующего хода процесса с одной и той же вероятностью; см. § 2.3. Если гипотеза о принадлежности априорного распределения $F_{n, N}(D)$ к классу биномиальных отвергается, то следует рассмотреть возможность аппроксимации $F_{n, N}(D)$ биномиальными распределениями со случайными весами, т. е. возможность аппроксимации G -биномиальными распределениями, введенными в § 2.3.

В общей модели производства априорное распределение $F_N(D)$ числа D дефектных объектов в контролируемой совокупности представимо в виде суперпозиции двух распределений

$$F_N(D) = w F_{н, N}(D) + (1 - w) F_{р, N}(D), \quad (1.1)$$

где $F_{н, N}(D)$ — априорное распределение, введенное выше, а $F_{р, N}(D)$ — априорное распределение D в совокупности, выпущенной при разладках в процессе производства. Число w равно доле совокупностей, выпущенных при нормальном ходе производства. Для хорошо отлаженного производства разладки наступают редко и w близко к единице. Если $F_{н, N}(D)$ можно считать стабильным для данного процесса производства, то распределение $F_{р, N}(D)$ скорее всего существенно зависит от вида возникающих разладок. Поэтому о виде $F_{р, N}(D)$ трудно делать какие-либо предположения без тщательного анализа имеющихся данных. Пусть значения D , типичные для нормального хода производства, удовлетворяют потребителя. Если при значениях D_i , для которых $F_{н, N}(D) \approx 1$, $F_{р, N}(D)$ еще мало, то задача организации приемочного контроля заключается в следующем. Надо так организовать контроль, чтобы основная часть совокупностей, изготавливаемых при нормальном ходе производства, принималась, тогда как совокупности, выпущенные в условиях разладок, браковались. Кратко назовем такое свойство контроля **разделимостью**. При организации контроля, обладающего свойством разделимости, можно исходить из конкретного вида распределения $F_{н, N}(D)$. Однако свойство разделимости должно сохраняться по отношению к возможно более широкому классу функции распределения $F_{р, N}(D)$. Требования «нечувствительности» к виду распределения $F_{р, N}(D)$ объясняется возможной сильной зависимостью $F_{р, N}(D)$ от вида разладок.

После того как определено распределение $F_{н, N}(D)$, необходимо исследовать характер возникновения разладок. Правила переключения на усиленный контроль и обратного перехода на нормальный контроль должны быть основаны на данных о возникающих разладках. Если разладка технологического процесса сказывается

на увеличении числа дефектных объектов в нескольких последовательно выпускаемых совокупностях или в течение сравнительно длительного периода имеет место недопустимое снижение доли совокупностей, содержащих долю дефектных объектов в зоне, характерной для нормального хода процесса, то в этих случаях целесообразен переход на усиленный контроль. Усиленный контроль должен продолжаться в течение действия разладки. Продолжение усиленного контроля должно быть веским аргументом для организации наладки хода технологического процесса. Если же совокупности, сильно засоренные дефектными объектами, появляются случайным образом, не образуя серий, и частота появлений невелика, то переход на усиленный контроль нецелесообразен. В некоторых условиях могут наблюдаться систематические изменения вида распределения $F_{n,N}(D)$. Например, в результате износа оборудования распределение $F_{n,N}(D)$ постепенно смещается вправо. В других же условиях постоянное совершенствование параметров технологического процесса будет уменьшать значения D , характерные для нормального хода производства. Может оказаться, что изменение $F_{n,N}(D)$ должно повлечь изменение способа выбора планов контроля из соответствующих таблиц, так как изменяются определяющие порядок выбора параметры, связанные с $F_{n,N}(D)$. Для того чтобы своевременно выявить необходимость изменения выбора планов из таблиц, нужно использовать соответствующие последующие оценки характеристик $F_N(D)$.

Таблицы планов контроля. При выборе любой системы правил контроля нужно учитывать определенную противоречивость интересов поставщика и потребителя (приемщика). Поставщик, примерно зная засоренность совокупностей дефектными объектами, может исходить из выбора таких планов контроля, при которых с заданной достаточно высокой вероятностью принимаются совокупности, выпущенные при нормальном ходе производства. Основным показателем при таком подходе является *приемочный уровень качества* (ПРУК), под которым понимают предельно допустимое значение доли дефектных объектов в контролируемых совокупностях при нормальном ходе производства. Если q_1 — значение

ПРУК, то поставщик может настаивать на использовании планов, у которых значения оперативной характеристики — вероятности приемки совокупностей, содержащих долю дефектных объектов q_1 , не менее $1 - \alpha$, где α — риск поставщика. В некоторых стандартах величина α зависит от объема N контролируемой совокупности, убывая с ростом N . В примере, результаты которого приведены на рис. 4.1.1, в качестве ПРУК можно взять $q_1 = 0,01$. Точно так же приемщик при планировании контроля может исходить из **браковочного уровня качества** (БРУК), под которым понимается такое значение доли дефектных объектов в контролируемой совокупности, начиная с которой совокупности должны браковаться. Обычно значение q , соответствующее БРУК, жестко связывается с вероятностью $1 - \beta$ забракования совокупностей с долей q дефектных изделий. При этом чаще всего считают $\beta = 0,05$ или $0,1$. При таком подходе БРУК — это то значение доли дефектных объектов, при которой совокупность будет при данном в стандарте плане забракована с вероятностью $0,95$ или $0,90$. Заметим, что в некоторых стандартах под ПРУК понимается то значение доли дефектных объектов в контролируемой совокупности, которому соответствует при использовании плана, указанного в стандарте, вероятность приемки $0,95$, $0,9$ или какое-то другое значение.

В таблицах, соответствующих усиленному контролю, основными могут быть показатели, отличные от БРУК и ПРУК. Так как усиленный контроль применяется при разладках, когда засоренность совокупностей дефектными объектами возрастает, то приемщик может настаивать на использовании таблиц планов усиленного контроля, содержащих планы с заданными значениями предельного среднего выходного качества D_L (определение см. в § 2.2), или уровня предельного среднего выходного качества (ПСВК) $q_L = D_L/N$. Показатель ПСВК вводится для планов, в которых забракование означает сплошную разбраковку, контрольная операция — неразрушительная, а выявленные дефектные объекты заменяются на годные. В некоторых стандартах таблицы планов усиленного и облегченного контроля могут быть

связаны с таблицами планов нормального контроля простыми соотношениями. Например, если при нормальном ходе производства используется одноступенчатый план с объемом выборки n и приемочным числом c , то переход к усиленному контролю означает уменьшение приемочного числа c без изменения объема n . Если же $c = 0$, то стандартным образом увеличивается объем выборки, а нулевое значение приемочного числа c сохраняется.

Если контроль не разрушительный и используются планы контроля с решениями $\mathcal{D}_2, \mathcal{D}_3$, то поставщик может исходить из планирования среднего числа объектов, контролируемых при нормальном ходе производства. Расчет этого показателя, как правило, проводится в предположении, что доля дефектных объектов в совокупностях постоянна и равна q_n — средней засоренности при нормальном ходе производства. Для данных контроля, показанных на рис. 4.1.1, соответствующее среднее значение $q_n \approx 0,003$.

Итак, в таблицы планов контроля включаются планы, имеющие те значения показателей, которые определяются поставщиком или приемщиком. Использование таких таблиц гарантирует, что для каждой отдельной совокупности будет выбран план с заданными показателями. Кратко планы с одинаковыми значениями выбранных показателей будем называть *гарантийными*.

Таблицы планов с заданным ПСВК. Рассмотрим один из простейших примеров построения таблицы планов приемочного контроля. Будем предполагать, что контрольная операция не является разрушительной и что все выявленные дефектные объекты заменяются годными. Для простоты ограничимся планами типа однократной выборки $(n, c)_{23}$. Допустим, что приемщик готовой продукции настаивает на введении такого контроля, при котором среднее значение доли дефектных объектов в принятых совокупностях не превышает заданной величины. Для плана $(n, c)_{23}$ значение ПСВК определяется соотношением

$$q_L = \max_{0 \leq D \leq N} (1/N) \sum_{0 \leq d \leq c} (D - d) h_{N, D}^{n, d}. \quad (1.2)$$

Используя пуассоновское приближение $h_{N,D}^{n,d} \approx \frac{\lambda^d}{d!} e^{-\lambda}$, $\lambda = nD/N$ при $\max(n/N, D/N) \ll 1$, $c \ll n$, из (1.2) получим

$$q_L \approx \max_{\lambda \geq 0} (\lambda/n) \sum_{0 \leq d \leq c} e^{-\lambda} \cdot \lambda^d/d! = \rho_c/n, \quad (1.3)$$

где $\rho_c = \max_{\lambda \geq 0} \lambda \sum_{0 \leq d \leq c} e^{-\lambda} \lambda^d/d!$, который достигается при $\lambda = \lambda_L$. Таблица значений ρ_c , λ_L для $c = 0(1)20$ дана в приложении (таблица I). При заданных значениях q_L и c объем плана можно рассчитать по формуле (1.3). Хотя отношение n/N предполагается априори малым, можно предложить уточнение, учитывающее конечность объема контролируемой совокупности. Из (1.2) получаем

$$\begin{aligned} q_L &\approx \max_{0 \leq q \leq 1} (1/N) \sum_{0 \leq d \leq c} (Nq - d) \frac{(nq)^d}{d!} e^{-nq} \approx \\ &\approx \max_{nq \geq 0} q \sum_{0 \leq d \leq c} \frac{(nq)^d}{d!} e^{-nq} [1 - n/N] = \frac{\rho_c(1 - n/N)}{n}. \end{aligned} \quad (1.4)$$

Таким образом, если значение приемочного числа известно, то объем выборки n , найденный из соотношения (1.4), при заданных q_L и c равен

$$n \approx N \left(\frac{\rho_c}{q_L N + \rho_c} \right). \quad (1.5)$$

Поскольку числа q_L задаются приемщиком, то в таблице параметров планов они могут следовать определенной последовательности предпочтительных чисел, приемочные числа $c = 0, 1, 2, \dots$. В итоге мы получаем таблицу планов $(n, c)_{23}$, имеющих заданные значения показателя ПСВК. Обычно q_L берут не менее чем в $1,5 \div 2$ раза превышающим средний уровень качества при нормальном ходе производства, так как и в противном случае объем контрольной работы составляет заметную долю от количества объектов, предъявляемых на контроль.

Более сложна задача построения таблицы одноступенчатых планов приемочного контроля, обеспечивающих минимальное значение среднего объема инспекции при уровне качества q_n и имеющих заданное значение q_L . При этом для контроля партии объема N нужно определить одноступенчатый план $(n, c)_{23}$. Пусть D_n —

среднее число дефектных объектов в совокупностях, выпущенных при нормальном ходе производства. Вероятность забраковать такую совокупность равна

$\sum_{c+1 \leq d \leq n} h_{N, D_H}^{n, d}$. При забраковании объем последующей контрольной работы равен $N - n$, так как остаток совокупности при решении $\mathcal{D}_2$ надо проконтролировать полностью. Таким образом, средний объем инспекции при $D = D_H$ с учетом выборочного контроля равен

$$M(D_H) = (N - n) \left(1 - \sum_{0 \leq d \leq c} h_{N, D_H}^{n, d} \right) + n. \quad (1.6)$$

При более полном учете априорного распределения $F_{n, N}(D_0) = \sum_{D \leq D_0} f_{n, N}(D)$ значение $M(D_H)$ нужно усреднить, т. е. определить средний объем M_H при нормальном ходе производства из соотношения

$$M_H = \sum_D M(D) f_{n, N}(D) + n. \quad (1.7)$$

Пуассоновская аппроксимация $h_{N, D_H}^{n, d}$ в (1.6) и априорных вероятностей $f_{n, N}(D)$ в (1.7) в предположении биномиальности априорного распределения $(f_{n, N}(D) = \binom{N}{D} \left(\frac{D_H}{N}\right)^D \times \left(1 - \frac{D_H}{N}\right)^{N-D})$ дает соотношение

$$M_H \approx (N - n) \left[1 - \sum_{0 \leq d \leq c} \left(\frac{n D_H}{N}\right)^d e^{-n D_H / N} \right] + n. \quad (1.8)$$

Решение поставленной задачи можно производить на основе точных соотношений (1.2), (1.6), (1.7). Для этого при заданном значении q_L и целом числе $c = 0, 1, 2, \dots$ нужно найти наименьшее n_c , при котором правая часть (1.2) окажется не больше q_L . Затем для плана $(n_c, c)_{23}$ подсчитывается по формуле (1.7) средний объем инспекции $M_H = M_H(q_L, c)$. Последовательно меняя значение c , отбираем в таблицу тот план (n_{c^*}, c^*) , для которого средний объем инспекции минимальный, т. е. $M_H(q_L, c^*) = \min_c M_H(q_L, c)$. Если априори предпо-

ложить, что значение n_{c^*} для плана, минимизирующего средний объем инспекции M_H , составляет малую долю от объема контролируемой совокупности, то можно использовать пуассоновское приближение. В этом случае

решение n_c уравнения (1.2) находится по формуле (1.5), а соответствующее значение $M_n(q_L, c)$ — по формуле (1.8). Реализуя необходимые вычисления на ЭВМ, можно получить компактную таблицу для определения оптимального значения c . Такая таблица приведена в приложении (таблица II).

Пусть $q_n = D_n/N$, $k_L = q_L/q_n$. В таблице II для значений $k_L = 1,25(0,25)3(0,5)6(1)10, 12, 15$ даны θ_c — граничные значения D_n для выбора приемочного числа c . Если $D_n \leq \theta_0$, то при заданном значении k_L приемочное число c^* надо положить равным нулю. Если же $\theta_{l-1} < D_n \leq \theta_l$, то полагаем $c^* = l$. После того как значение c^* найдено, необходимый объем выборки для плана $(n^*, c^*)_{23}$, минимизирующего средний объем инспекции, находится по формуле (1.5), в которой полагаем $c = c^*$. Теперь нужно проверить исходное предположение о том, что n^*/N мало ($n^*/N < 0,1$), так как только для планов с малой долей объема объектов, отбираемых в выборку, можно использовать пуассоновскую аппроксимацию.

Приведем пример. Пусть контроль неразрушительный, а выявленные дефектные объекты заменяются годными. Потребитель настаивает на использовании планов со значением ПСВК $q_L = 0,02$. Среднее значение доли дефектных объектов в совокупностях, выпущенных при нормальном ходе производства, равно $q_n = 0,01$. Поставщик предпочитает планы, которые при $q = q_n$ дают минимум среднего объема инспекции. Предположим, что на контроль предъявлена совокупность объема $N = 750$. Так как $k_L = q_L/q_n = 2$, $D_n = N \cdot q_n = 750 \cdot 0,01 = 7,5$, то из столбца таблицы II, соответствующего $k_L = 2$, находим $\theta_1 = 6,7$, $\theta_2 = 17,8$. Учитывая $6,7 < 7,5 < 17,8$, получаем оптимальное значение приемочного числа $c^* = 2$. Из таблицы I находим $\rho_2 = 1,37$. Объем $n^* = 62,77$ выборки искомого плана вычисляем по формуле (1.5). Округлив до ближайшего целого числа, получим, что искомым одноступенчатым планом контроля является $(63,2)_{23}$. Отношение $63/750 < 0,1$.

Для двухступенчатых планов с решениями $\mathcal{D}_2, \mathcal{D}_3$ вычисление значения ПСВК трудоемко. Пусть n_1, n_2 — объемы выборки, c_1, c_2 — приемочные числа, а r_1 — бра-

ковочное число. Из общей формулы (2.2.14) при контроле совокупности объема N находим, что предельное выходное качество равно

$$D_L = \max_{0 \leq D \leq N} \left\{ \sum_{0 \leq d_1 \leq c_1} (D - d_1) h_{N, D}^{n_1, d_1} + \right. \\ \left. + \sum_{c_1 < d_1 < r_1} \sum_{0 \leq d_2 \leq c_2 - d_1} (D - d_1 - d_2) h_{N, D}^{n_1, d_1} h_{N - n_1, D - d_1}^{n_2, d_2} \right\} \quad (1.9)$$

Если $\max \left(\frac{n_1 + n_2}{N}, \frac{D}{N} \right) \ll 1$, то, используя пуассоновское приближение $h_{N, D}^{n_1, d_1} \approx (\lambda_1^{d_1} / d_1!) e^{-\lambda_1}$, $h_{N - n_1, D - d_1}^{n_2, d_2} \approx ((\lambda_1 \kappa)^{d_2} / d_2!) \times \times e^{-\lambda_1 \kappa}$, $\kappa = \frac{n_2}{n_1}$, из (1.9) получим

$$D_L \approx \frac{N}{n} \max_{\lambda \geq 0} \rho_{c_1, c_2, r_1}(\lambda, \kappa), \quad (1.10)$$

где

$$\rho_{c_1, c_2, r_1}(\lambda, \kappa) = \\ = \lambda \left\{ \sum_{0 \leq d_1 \leq c_1} \frac{\lambda^{d_1}}{d_1!} e^{-\lambda} + \sum_{c_1 < d_1 < r_1} \sum_{0 \leq d_2 \leq c_2 - d_1} \frac{\lambda^{d_1 + d_2}}{d_1! d_2!} \kappa^{d_2} e^{-\lambda(1 + \kappa)} \right\}.$$

Для двухступенчатых планов со значениями $c_1 = 0$, $r_1 = 1 + j$, $c_2 = j$, $j = 1, 2$ и $\kappa = 0(0, 1)10$ нами были рассчитаны таблицы функций

$$\rho_{0, 1}(\kappa) = \max_{\lambda > 0} \rho_{0, 1, 2}(\lambda, \kappa), \\ \rho_{0, 2}(\kappa) = \max_{\lambda > 0} \rho_{0, 2, 3}(\lambda, \kappa), \quad (1.11)$$

и соответствующих им значений $\lambda_L(\kappa)$, при которых функции $\rho_{0, 1, 2}(\lambda, \kappa)$ и $\rho_{0, 2, 3}(\lambda, \kappa)$ достигают максимума по λ . Величины $\lambda_L(\kappa)$ и $\rho_{0, j}(\kappa)$, $j = 1, 2$, даны в таблицах III, IV приложения.

Для двухступенчатых планов, имеющих $c_1 = 0$, $c_2 = j$, $r_1 = j + 1$, в соответствии с (1.9) — (1.11) находим значение ПСВК

$$q_L = D_L / N \approx \rho_{0, j}(\kappa) / n_1.$$

Если число дефектных объектов в контролируемой совокупности равно

$$D_L^* \approx \lambda_L(\kappa) N / n_1,$$

то в среднем принимается наибольшее число дефектных объектов, равное D_L .

Приведенные выше примеры показывают специфику составления таблиц планов приемочного контроля. Характерной чертой таких таблиц является то, что часть переменных меняется по некоторой выбранной заранее последовательности чисел, на которую нужно ориентироваться при задании требований к планам контроля. Другие же переменные, вообще говоря, должны были бы принимать все значения, но в таблицы выносятся только некоторые значения, на основе которых можно либо производить интерполяцию, либо приближенно считать фактическое значение параметра равным табличному.

§ 4.2. Системы планов контроля и их характеристики

Система Доджа и Ромига, МИЛ-СТД — 105. Первая наиболее обстоятельно разработанная система планов приемочного контроля по альтернативному признаку была предложена Доджем и Ромигом еще в конце двадцатых годов. В широко известной монографии Доджа и Ромига [1], которая явилась определенным этапом в развитии теории приемочного контроля, опубликованы таблицы одноступенчатых и двухступенчатых планов контроля по альтернативному признаку. В этой монографии дано описание принципов приемочного контроля по альтернативному признаку. Изложена методика, на основе которой были составлены таблицы планов приемочного контроля. Планы, представленные в книге Доджа и Ромига [1], разбиваются на две группы. Одни обеспечивают заданные значения вероятности 0,1 для приемки совокупностей, содержащих долю $q_{0,1}$ дефектных объектов, другие — заданное значение q_L величины ПСВК. В таблицах приведены планы, дающие минимум среднего значения инспекции в условиях принятия решений $\mathcal{D}_2$, $\mathcal{D}_3$ при заданном значении средней доли засоренности дефектными объектами контролируемых совокупностей. Значительную часть книги занимают графики оперативных характеристик одноступенчатых и двухступенчатых планов, указанных в таблицах этой монографии. Приведем последовательность

действий, рекомендуемую Доджем и Ромигом [1]. Эта последовательность действий вместе с таблицами и составляет систему планов Доджа и Ромига. (1) Надо уяснить, какие параметры объектов надо контролировать и какие объекты по результатам контроля считать дефектными. (2) Уточнить понятие контролируемой совокупности и подсчитать ее объем N . (3) Выбрать вид защиты от принятия совокупностей, содержащих большое число дефектных объектов. В данном случае делается выбор между двумя показателями $q_{0,1}$ и q_L . (4) Решить вопрос о типе используемого плана, т. е. выбрать одноступенчатый или двухступенчатый план. При этом во внимание могут быть приняты результаты пункта (7). (5) Провести оценку $q_{\text{ср}}$ — средней доли засоренности совокупностей дефектными объектами. (6) По таблицам найти параметры плана контроля. Выбор плана контроля из таблиц проводится на основе задания значений $q_{0,1}$ или q_L и значений $q_{\text{ср}}$ и N . (7) Найти чертеж со значением оперативной характеристики, в зависимости от вида которой могут быть внесены изменения в выбор плана. (8) Провести контроль совокупности в соответствии с выбранным планом. (9) Результаты контроля зарегистрировать для последующего использования при оценке качества продукции.

Заметим, что найденные с помощью таблицы II одноступенчатые планы, обеспечивающие заданное значение q_L и минимизирующие средний объем инспекции при $q_{\text{ср}} = q_n$, практически совпадают или очень близки к соответствующим одноступенчатым планам Доджа и Ромига, определяемым на основе q_L , $q_{\text{ср}}$ и N .

Мы видим, что таблицы содержат не все планы, а только те, которые удовлетворяют определенным требованиям заказчика и поставщика. Тем не менее ответ на вопрос о том, хорош ли отдельный план выборочного контроля или плох, будет неполон, если мы не рассматриваем всю систему правил контроля — стандарт в целом. В настоящее время наибольшее распространение получил американский стандарт МИЛ-СТД—105Д [3]. В СССР утвержден государственный стандарт ГОСТ 18247—72 [4], который близок к МИЛ-СТД—105Д. В основу МИЛ-СТД—105Д положено понятие *приемочного уровня качества* ПРУК, определяемого

как максимальная доля дефектных объектов в контролируемых совокупностях, которую потребитель рассматривает как допустимый уровень качества на все время поставок. Вероятность забракования совокупностей с долей дефектных объектов, равной ПРУК, при использовании планов из таблиц МИЛ-СТД—105Д невелика и убывает с ростом объема совокупности, предъявляемой на контроль. Для основной части планов вероятность забракования не превышает 0,05. При таком подходе планы контроля имеют медленно убывающие оперативные характеристики. Поэтому система МИЛ-СТД—105Д ориентирована на условия стабильного производства, когда значение доли дефектных объектов в контролируемых совокупностях, как правило, не превышает ПРУК. Для того чтобы защитить потребителя от ошибочной приемки совокупностей в условиях разладок производства, в МИЛ-СТД—105Д предусмотрен усиленный контроль. Переход на усиленный контроль осуществляется на основе правил переключения, о которых будет сказано ниже. Допускается также переход на сокращенный контроль. Такой переход происходит при последовательной приемке большего числа совокупностей в условиях нормального контроля. Концепция трех уровней контроля (нормального, усиленного и сокращенного) дает возможность оказывать давление на поставщика при ухудшении качества контролируемой продукции. Таблицы МИЛ-СТД—105Д содержат параметры одноступенчатых, двухступенчатых и многоступенчатых (семиступенчатых) планов контроля. Для двухступенчатых и многоступенчатых планов объемы последовательных выборок совпадают. Например, если объем N контролируемой совокупности заключен в интервале $[501, 1200]$, то рекомендуемый МИЛ-СТД—105Д (таблица IV—А в МИЛ-СТД—105Д [3]) семиступенчатый план состоит в последовательном извлечении выборок по 20 объектов. Принятие решения определяется выбранным значением ПРУК и числами $d_1, d_2, \dots, d_7$ дефектных объектов, обнаруженных при контроле в последовательных выборках. Если ПРУК равен 0,01, то при $d_1 \geq 2$ совокупность бракуется. Если же $d_1 < 2$, то берется вторая выборка, причем при $d_1 + d_2 \geq 3$ совокупность бра-

куется, а при $d_1 + d_2 = 0$ принимается. В случае $0 < d_1 + d_2 < 3$ берется третья выборка. Если $d_1 + d_2 + d_3 \geq 3$, то совокупность бракуется. Если же $0 < d_1 + d_2 + d_3 < 3$, то берется четвертая выборка и т. д. В общем случае правила прекращения контроля и принятия решения задаются последовательностью приемочных чисел c_i и браковочных чисел r_i . В рассматриваемом примере $r_1 = 2$, $r_2 = r_3 = 3$, $r_4 = r_5 = 4$, $r_6 = r_7 = 5$, $c_2 = 0$, $c_4 = 1$, $c_5 = 2$, $c_6 = 3$, $c_7 = 4$. Решение о приемке по результатам первой и третьей выборок не производятся. При увеличении ПРУК увеличиваются значения приемочных и браковочных чисел, но объем последовательных выборок не меняется. Если в условиях только что рассмотренного примера ПРУК равен 0,04, то рекомендуемый объем выборок — 20, а браковочные и приемочные числа равны $r_1 = 4$, $r_2 = 6$, $r_3 = 8$, $r_4 = 10$, $r_5 = 11$, $r_6 = 12$, $r_7 = 14$, $c_1 = 0$, $c_2 = 1$, $c_3 = 3$, $c_4 = 5$, $c_5 = 7$, $c_6 = 10$, $c_7 = 13$.

В главе 2 для планов первого вхождения показано, что различным границам остановки соответствуют планы с отличающимися оперативными характеристиками, средними объемами выборок и т. д. Одноступенчатые планы наиболее просты с точки зрения организации контроля на производстве. Напомним, что двухступенчатые и многоступенчатые планы позволяют уменьшить среднее число контролируемых объектов, не увеличивая вероятностей ошибочного принятия совокупностей с большой долей дефектных объектов. В МИЛ-СТД—105Д двухступенчатые и многоступенчатые планы подобраны таким образом, что их оперативные характеристики близки к оперативным характеристикам одноступенчатых планов. Таким образом в рамках системы МИЛ-СТД—105Д при выбранном значении ПРУК одноступенчатым, двухступенчатым или многоступенчатым планам соответствуют приближенно одинаковые вероятности принятия решений о браковании или приемке, а среднее число объектов, контролируемых до принятия решения, существенно различно. Указаний о том, как конкретно проводился отбор планов и что взято за критерий близости оперативных характеристик, в тексте МИЛ-СТД—105Д [3] нет. В таблице IX этого стандарта даны графики среднего числа

контролируемых объектов для двухступенчатых и многоступенчатых планов в долях от объема выборки n соответствующего одноступенчатого плана (n, c). По оси абсцисс отложены значения nq — среднего числа дефектных объектов в выборке объема n . На рис. 4.2.1, *a* и *б* приведены два графика из таблицы IX МИЛ-СТД—105Д. Рис. 4.2.1, *a* соответствует приемочному числу $c = 1$, рис. 4.2.1, *б* — числу $c = 3$. Стрелками

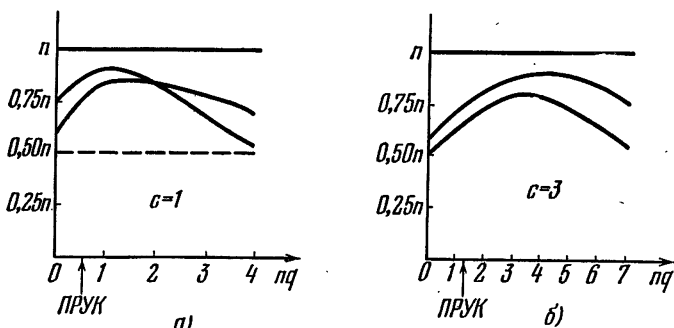


Рис. 4.2.1. Графики среднего числа объектов, контролируемых до принятия решения при использовании двухступенчатых и семиступенчатых планов из МИЛ-СТД—105Д, оперативные характеристики которых близки к оперативным характеристикам одноступенчатых планов (n, c): *a*) $c = 1$, *б*) $c = 3$.

на оси абсцисс указаны значения ПРУК. Рис. 4.2.1 иллюстрирует тот факт, что переход к более сложным двухступенчатым и многоступенчатым планам позволяет существенно уменьшить средний объем контрольной работы. В МИЛ-СТД—105Д включены таблицы планов сокращенного и усиленного контроля, содержащие одноступенчатые, двухступенчатые и многоступенчатые планы.

Переключающие правила. Введение различных уровней контроля основано на том предположении, что возможные отклонения от «нормального хода производства» носят длительный характер и сказываются на уровне качества в серии совокупностей, последовательно предъявляемых на контроль. Для простоты ограничимся системами контроля с двумя уровнями: нор-

мальным и усиленным. Длительное проведение усиленного контроля предполагает вмешательство в процесс производства с целью устранения разладки. В стандартах статистического приемочного контроля даются **переключающие правила**, регулирующие переход от нормального контроля к усиленному и обратно. Переключение с одного уровня на другой производится на основе результатов контроля не только очередной контролируемой совокупности, но и некоторого числа предыдущих совокупностей. Результаты контроля могут использоваться полностью либо частично. В более сложных вариантах правил переключения используются данные о числе выявленных при контроле дефектных объектов. В упрощенных правилах переключения используются лишь данные о числе принятых и забракованных совокупностей. В стандарте МИЛ-СТД—105С [2], предшествующем МИЛ-СТД—105Д, переключение на усиленный контроль проводится на основе суммарного числа дефектных объектов, выявленных при контроле 10 последних совокупностей. В этом стандарте дана таблица предельно допустимых суммарных чисел дефектных объектов. Если суммарное число дефектных объектов, выявленных при контроле последних 10 совокупностей, превышает предельно допустимое, то производится переход к усиленному контролю. В стандарте МИЛ-СТД—105Д переключающие правила были упрощены и сформулированы следующим образом. (1) Переход от нормального контроля к усиленному производится как только из пяти последних совокупностей две будут забракованы. (2) Переход от усиленного контроля к нормальному производится как только пять последних совокупностей, проконтролированных в режиме усиленного контроля, будут приняты.

Эти переключающие правила можно изобразить в виде графа, у которого H_0 соответствует началу контроля либо такому состоянию, когда пять предыдущих совокупностей были приняты. Состояния $(H_1, 1)$, $(H_1, 2)$, ..., $(H_1, 5)$ соответствуют нормальному контролю, в котором среди пяти последних партий только одна забракована. Аналогичным образом интерпретируются состояния при усиленном контроле $(U, 1)$, ..., $(U, 5)$. Схема переходов, соответствующая переключению

чающим правилам МИЛ-СТД—105Д, дана на рис. 4.2.2. При приемке очередной совокупности происходит переход вдоль пути «пр», при браковке — вдоль пути «бр».

Чтобы выбрать оптимальный вариант переключающих правил или хотя бы сравнить два варианта переключающих правил между собой, нужно задать коли-

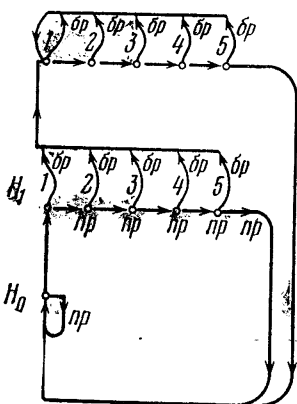


Рис. 4.2.2. Граф состояний, соответствующий переключающим правилам для нормального и усиленного контроля в МИЛ-СТД — 105Д.

чественные показатели эффективности этих правил. Такие показатели можно ввести, определив уровни производства, при которых необходим нормальный или усиленный приемочный контроль. При этом **ложными переключениями на усиленный контроль** будут такие, в момент которых уровень производства — нормальный. **Ошибочными будут переходы на нормальный контроль** и тогда, когда состояние производства предполагает усиленный контроль. Наконец, важно учесть **запаздывание в переключении** с нормального контроля на усиленный при разладке производства. Вероятностными характеристиками

переключающих правил могли бы быть следующие: **вероятность ложного переключения** (от нормального контроля к усиленному), **среднее запаздывание** (математическое ожидание числа совокупностей, проконтролированных в режиме нормального контроля, тогда как требовался режим усиленного контроля), **вероятность ложного перехода от усиленного контроля к нормальному**. Однако рассчитать эти характеристики и, соответственно, поставить задачу поиска оптимальной системы правил можно только в рамках определенной математической модели процесса производства. В. К. Олшанский предложил для расчетов следующую модель производства, которое может находиться только в двух состояниях: в нормальном (0) и разлаженном (1). (а) При нормальном ходе производства и нормальном

контроле процесс переходит из нормального состояния в разлаженное с постоянной вероятностью p_{01} . Таким образом, момент появления разладки имеет геометрическое распределение. (б) После появления разладки при нормальном уровне контроля нормальное состояние процесса производства не восстанавливается. (в) После правильного перехода к усиленному контролю процесс может восстановить нормальное состояние при выпуске очередной совокупности с постоянной вероятностью p_{10} . Таким образом число совокупностей, выпущенных в условиях разладки от момента перехода к усиленному контролю до устранения, имеет геометрическое распределение. (г) Если после восстановления нормального состояния производства продолжается усиленный контроль, то процесс производства разладиться не может. Точно так же предполагается, что разладка не может наступить при усиленном контроле, если переход к усиленному контролю был ложным. (д) Аналогично предполагается, что при ложном переходе от усиленного контроля к нормальному нормальное состояние производства не восстанавливается.

Возможны и другие математические модели производства. Сформулированная выше модель представляется естественной. В такой модели переключающие правила описываются цепью Маркова с конечным числом состояний. Поэтому нетрудно рассчитывать различные вероятностные характеристики эффективности как переключающих правил, так и системы в целом. Можно поставить более сложную задачу поиска оптимальных правил переключения, которые, например, при заданной вероятности ложного переключения минимизируют среднее время запаздывания. Здесь мы приходим к типичным задачам обнаружения разладки, которые исследовались Ширяевым [1]. Общая теория оптимальных последовательных методов обнаружения разладок дает эффективные решения для моделей появления разладок, описываемых с помощью марковских процессов. При этом оптимальные переключающие правила основаны на значениях апостериорных вероятностей появления разладок. Для апостериорных вероятностей разладок выводятся рекуррентные уравнения, позволяющие последовательно вычислять их значения по результатам

контроля совокупностей. Ольшанский [1] провел подробное численное исследование правил переключения, основанных на апостериорных вероятностях разладок. Анализ переключающих правил, используемых в стандартах, посвящен ряд работ, среди которых следует отметить работу Пабста [1], определившего количественные характеристики переключающих правил. На Международной конференции по вопросам управления качеством в Токио (1969 г.) проблемы выбора рациональных переключающих правил обсуждались в сообщениях Коямы [1], Омаэ и Суга [1], Йоко [1]. Любые количественные характеристики системы планов приемочного контроля определяются не только значениями параметров планов контроля, включенных в таблицы, но зависят также от выбранных переключающих правил и исходной математической модели производства. В простейшей математической модели производства распределение чисел дефектных объектов, содержащихся в предъявляемых на контроль совокупностях, предполагается биномиальным со средней долей дефектных объектов q и «числами испытаний», равными объемам совокупностей. Пусть величина q не меняется от совокупности к совокупности. В таких предположениях можно рассчитать вероятность $\pi_n(q)$ ($\pi_y(q)$, $\pi_c(q)$) того, что в установившемся режиме контроля очередная совокупность пройдет нормальный (усиленный, сокращенный) контроль, $\pi_n(q) + \pi_y(q) + \pi_c(q) = 1$. Если сокращенный контроль не производится, то $\pi_c(q) = 0$. Значения $\pi_n(q)$ зависят от вида переключающих правил. На рис. 4.2.3, а даны графики $\pi'_n(q)$, $\pi''_n(q)$ для двух различных наборов переключающих правил. Переключающие правила, соответствующие пунктирной кривой, представляются предпочтительными по сравнению с переключающими правилами, соответствующими графику со сплошной линией, так как при доле q , равной БРУК, им соответствуют меньшие вероятности использования нормального контроля. Такое заключение является естественным следствием требования, предъявляемого потребителем к переключающим правилам и состоящего в том, что при значениях q , больших БРУК, должен применяться усиленный контроль. С другой стороны поставщик заинтересован в том, чтобы переключающие пра-

вила при значениях q , меньших или равных ПРУК, с вероятностью, близкой к единице, держали контроль на нормальном уровне.

Пусть $\pi_1(q)$ — вероятность приемки совокупности при нормальном контроле, а $\pi_2(q)$ ($\pi_3(q)$) — вероятность приемки при усиленном (сокращенном) контроле. Тогда

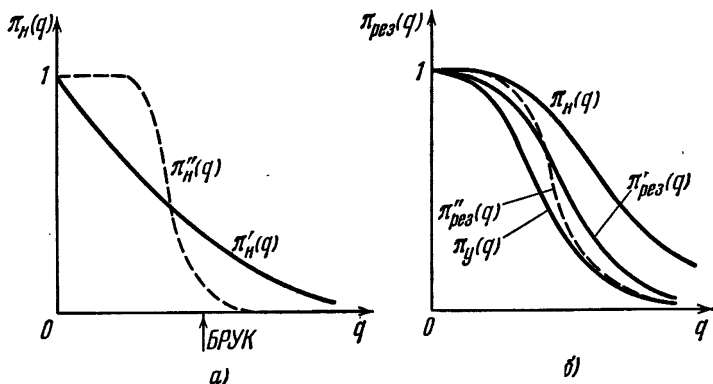


Рис. 4.2.3. а) Вероятности нормального контроля совокупностей с долей дефектных объектов q для двух типов переключающих правил, б) графики результирующих оперативных характеристик.

вероятность приемки контролируемой совокупности как функция q равна

$$\pi_{рез}(q) = \pi_1(q) \pi_n(q) + \pi_2(q) \pi_y(q) + \pi_3(q) \pi_c(q).$$

Функция $\pi_{рез}(q)$ называется **результирующей (составной) оперативной характеристикой системы правил контроля**. На рис. 4.2.3, б приведены графики результирующих оперативных характеристик, $\pi'_{рез}(q)$ и $\pi''_{рез}(q)$, соответствующие переключающим правилам рис. 4.2.3, а. Напомним, что в этом примере сокращенный контроль не используется, т. е. $\pi_c(q) = 0$. Из рис. 4.2.3, б видно, что при малых значениях q результирующая оперативная характеристика близка к оперативной характеристике планов при нормальном контроле. При больших значениях q результирующая оперативная характеристика приближается к оперативной характеристике планов усиленного контроля.

Однако результирующие оперативные характеристики не содержат информации о частоте переключений при контроле. Поэтому представляет интерес исследование переключающих правил с точки зрения устойчивости режима контроля. Устойчивость нормального контроля можно характеризовать вероятностями $\pi_{н,у}(q, m)$, $\pi_{н,с}(q, m)$, того, что в результате контроля не более чем m совокупностей произойдет переход от

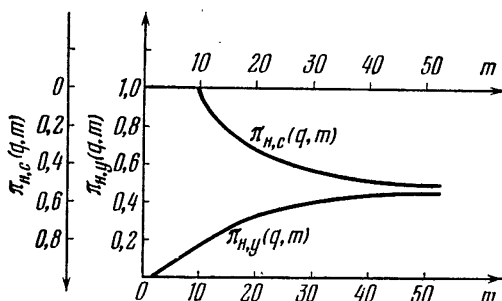


Рис. 4.2.4. Динамические характеристики перехода от нормального к сокращенному или усиленному контролю для одноступенчатого плана (50,1) при $q = 0,01$ и переключающих правилах МИЛ-СТД — 105Д (Кояма [1]).

режима нормального контроля к усиленному или, соответственно, к сокращенному контролю при доле q дефектных объектов в каждой из контролируемых совокупностей. Вероятность того, что режим нормального контроля не изменится в результате приемочного контроля m совокупностей, равна $1 - \pi_{н,у}(q, m) - \pi_{н,с}(q, m)$. Вероятности $\pi_{н,у}(q, m)$, $\pi_{н,с}(q, m)$, рассматриваемые как функции m при фиксированной доле дефектных объектов q , называются **динамическими характеристиками**. На рис. 4.2.4, взятом из работы Коямы [1], приведены динамические оперативные характеристики $\pi_{н,у}(q, m)$, $\pi_{н,с}(q, m)$ одноступенчатого плана контроля (n, c) , $n = 50$, $c = 1$, для $q = 0,01$ и переключающих правил МИЛ-СТД—105Д. Значения $\pi_{н,у}(q, m)$, $\pi_{н,с}(q, m)$ считываются по шкалам в левой части рис. 4.2.4. Если $q = 0,01$ соответствует ПРУК, то переходы на усиленный и сокращенный контроль ошибочны.

Из графиков рис. 4.2.4 видно, что после контроля 15 совокупностей режим нормального контроля будет прерван ошибочным переключением на режим усиленного или сокращенного контроля с вероятностью, большей 0,5.

Расчет динамических характеристик можно проводить на основе статистического моделирования процесса производства годных и дефектных объектов и процесса выборочного контроля. Для широкого класса моделей производства задача вычисления $\pi_{н,у}(q, m)$, $\pi_{н,с}(q, m)$ сводится к подсчету вероятностей состояний в некоторой цепи Маркова с поглощающими состояниями. Примеры даны в статьях Коямы [1] и Йоко [1]. Основные элементы необходимой теории цепей Маркова читатель может найти в книге Феллера [1].

В связи с широким использованием системы МИЛ-СТД—105Д в практике приемочного контроля были исследованы ее динамические характеристики. В работе Омаэ и Суга [1] отмечается, что для системы МИЛ-СТД—105Д характерны неоправданно большие вероятности ошибочного перехода на усиленный и сокращенный объем в условиях, когда следует применять нормальный контроль. Напомним, что в МИЛ-СТД—105Д переход от нормального контроля к усиленному производится при забраковании двух совокупностей из последних пяти предъявленных на приемочный контроль. Переход от нормального контроля к сокращенному производится в том случае, когда все 10 последних совокупностей, проконтролированных при нормальном контроле, приняты и общее число дефектных объектов в выборках из 10 последних совокупностей не более соответствующего табличного числа. Омаэ и Суга [1] указывают, что переключающие правила в предшествующем МИЛ-СТД—105Д варианте МИЛ-СТД—105С обладали более приемлемыми характеристиками. Приведем некоторые из требований, которым, по мнению Омаэ и Суга [1], должны удовлетворять количественные характеристики переключающих правил. При уровне качества q , равном ПРУК, с вероятностью не менее 0,9 должны проходить нормальный контроль 50 совокупностей, т. е. $1 - \pi_{н,у}(q, 50) - \pi_{н,с}(q, 50) \geq 0,9$. Вероятность бракования совокупности с уровнем качества, равном ПРУК, при усиленном контроле должна быть не

более 0,1. Вопросы выбора рациональных переключающих правил еще далеки от окончательного решения.

Статистическое моделирование на ЭВМ процесса производства и приемочного контроля. Стандарты контроля качества, основанные на статистических методах, должны проходить всестороннюю проверку до того, как они становятся обязательными для использования на производстве. Одним из возможных путей ускорения такой проверки является разработка соответствующих математических моделей различных режимов производства и последующая проверка правил, записанных в стандартах путем статистического моделирования на ЭВМ. Ниже мы рассмотрим круг вопросов, возникающих при анализе числовых характеристик стандартов приемочного контроля по альтернативному признаку объектов (изделий) промышленной продукции. Пусть по-прежнему N_i — объем совокупности с номером i , содержащей D_i дефектных объектов.

Аналитические расчеты эффективности стандартов приемочного контроля качества в предположениях, достаточно точно отражающих реальные условия производства, выполнить очень трудно. Однако такие расчеты с приемлемой степенью точности можно провести путем статистического моделирования соответствующей математической модели производства. Меняя на модели режимы производства, можно отбирать наиболее эффективные стандарты, а также определять целесообразность различных типов правил корректировки и т. п. Предварительное опробование стандартов на математических моделях особенно важно при разработке систем управления качеством с использованием ЭВМ.

В математической модели заданию режима производства соответствует определение последовательности (N_i, D_i) как случайного процесса. Если отбор объектов из совокупности на контроль производится случайным образом, то для оценки показателей контроля достаточно определить D_i как случайную величину с функцией распределения $F_i(D)$. В ситуациях, когда возможны отличные от случайного способы отбора объектов в выборку, иногда возникает необходимость определения номеров изделий, которые окажутся дефектными. В таких случаях совокупность объема N

должна описываться случайным вектором $(\varepsilon_1, \dots, \varepsilon_N)$, где $\varepsilon_t = 0$, если объект с номером t годный, и $\varepsilon_t = 1$, если он дефектный. Методы случайного отбора и методы специальным образом организованного отбора, как правило, имеют существенно отличающиеся показатели эффективности.

Если используется случайный отбор объектов для контроля, то достаточно широкий класс математических моделей производства можно получить следующим образом. Введем конечное число состояний производства $S = (H_1, \dots, H_l, R_1, \dots, R_k)$ и соответствующих им ф. р. $F_{H_1}(x), \dots, F_{H_k}(x), F_{R_1}(x), \dots, F_{R_k}(x)$, $x \in [0, 1]$. В простейшем случае, когда $k = l = 1$, H_1 — состояние, соответствующее нормальному ходу производства, а R_1 — состояние, соответствующее разладке. При переходе от контроля совокупности $\mathfrak{P}_i$ (в «момент i ») к контролю $\mathfrak{P}_{i+1}$ (в «момент $i + 1$ ») смена состояний S описывается цепью Маркова с матрицей переходных вероятностей $P_r = \|p_{s,t}^r\|$,

$s, t \in S$, r — режим контроля в «момент i ». Таким образом изменение состояний производства может зависеть от вида контроля. Если в «момент i » состояние производства $s = s(i)$, то случайная величина, определяющая число дефектных объектов D_i в $\mathfrak{P}$, реализуется в два этапа. Сначала получают реализацию случайной величины ξ с ф. р. $F_s(x)$. По значению $\xi = x$ число D_i находим как число «успехов» в N_i взаимно независимых испытаниях с вероятностью «успеха» x . При составлении программы для ЭВМ можно использовать аппроксимации, облегчающие вычисления. Например, можно считать, что каждая $F_s(x)$ является **взвесью** нескольких равномерных распределений, т. е. каждому $s \in S$ сопоставляется набор чисел $\alpha_{s,j}$, $0 \leq \alpha_{s,j} \leq 1$, $\sum_j \alpha_{s,j} = 1$, и интервалов $\Delta_{s,j} \in$

$[0, 1]$, $1 \leq j \leq m_s$. Значение ξ получается так: (1) реализуется случайная величина η , $P\{\eta = j\} = \alpha_{s,j}$, (2) если $\eta = j$, то значение $\xi = \delta$, где δ — равномерно распределенная в интервале $\Delta_{s,j}$ случайная величина. В условиях проведения выборочного контроля $D_i = d_i + D_i^*$, где d_i — число дефектных объектов, обнаруженных среди n_i объектов, отобранных на контроль, а D_i^* — число дефектных объектов среди $N_i - n_i$, не отобранных в выборку. Если

$\xi = q$, то D_i^* при больших значениях $(N_i - n_i)q$ можно приближенно считать нормально распределенной случайной величиной со средним $(N_i - n_i)q$ и дисперсией $(N_i - n_i)q(1 - q)$. Поэтому можно использовать соотношение $D_i^* = [(N_i - n_i)q + \gamma \sqrt{(N_i - n_i)q(1 - q)} + 0,5]$, где γ — стандартная нормально распределенная случайная величина с нулевым средним и единичной дисперсией, $[a]$ — целая часть числа a . Если значение $(N_i - n_i)q < K$ ($K = 10 \div 30$), то в качестве приближения при моделировании D_i^* можно считать, что D_i^* имеет пуассоновское распределение со средним $(N_i - n_i)q$. Следует иметь в виду, что существуют более точные аппроксимации биномиальных распределений; см. Моленар [1]. Так как n_i , как правило, не велико, то d_i можно получить как сумму n_i взаимно независимых случайных величин $e_t = 0, 1$, $P\{e_t = 1\} = q$. При соответствующем подборе параметров $P_r = \|p_{s,t}^r\|$, $\alpha_{s,j}$, $\Delta_{s,j}$ можно моделировать различные типы разладок и нормального хода производства. В частности, можно получать стабильные режимы нормального хода производства с редкими спонтанно возникающими разладками или постепенные переходы с одного уровня качества на другой и т. д.

Изменение величин N_i можно задавать разными способами. Например, можно получить m -периодические реализации взаимно независимых случайных величин, ф. р. которых в «момент i » равны $W_s(N)$, если $i \equiv s \pmod{m}$, $s = 0, \dots, m-1$, $W_0(N), \dots, W_{m-1}(N)$ — фиксированные распределения, $N = 1, 2, \dots$. Можно комбинировать такой способ со способом марковского блуждания по индексам m законов $W_s(N)$ путем моделирования цепи Маркова $s(i)$ на состояниях $0, \dots, m-1$. При этом в «момент i » выбираем ф. р. $W_{s(i)}(N)$.

Если в «момент i » на контроль поступает партия объема N_i с соответствующим значением средней доли засоренности дефектными изделиями $\xi = q$, то на основе текущей (или заданной априори) информации из таблиц приемочного контроля исследуемой системы планов выбирается рекомендуемый план. В простейшей ситуации планы Π и их параметры могут содержаться в виде массива в памяти ЭВМ. Если параметры плана Π определены, то далее моделируется процесс выбороч-

ного контроля. Например, если $\Pi = \Pi(n_1, n_2, c_1, r_1, c_2)$ — план двухступенчатой выборки, то моделируется биномиальная случайная величина δ_i , равная числу «успехов» в n_i независимых испытаниях с вероятностью успеха q . Если $\delta_1 = d_1 \leq c_1$, то партия t принимается, если $\delta_1 = d_1 \geq r_1$, то бракуется, при $c_1 < d_1 < r_1$ производится повторная выборка объема n_2 и т. д. Если n — фактическое число объектов, проверенных при двухступенчатой выборке, т. е. $n = n_1$ или $n_1 + n_2$, d_i — число обнаруженных при контроле дефектных изделий, то далее по описанной выше схеме моделируется D_i^* . Если забракование партии означает сплошную проверку $N_i - n_i$ оставшихся непроверенных объектов, то D_i^* считается известным. Если же партия принята, то D_i^* включается в число дефектных изделий, принятых потребителем. В соответствии с правилами обработки информации о ходе контроля, записанными в стандарте, строятся последующие оценки. По мере накопления информации может возникнуть необходимость в корректировке процесса приемочного контроля. Соответствующие правила должны быть описаны в стандарте. Корректировка должна определяться характером изменений режимов работы. В простейшем случае изменяется скачком только значение q_n — средней доли засоренности партий дефектными изделиями при нормальном ходе производства. В некоторых случаях возможны постепенные изменения q_n . Достаточно широкий класс перекрывающих правил можно получить, задавая три числа (M, g, h) . В соответствии с выборочными значениями M, g, h вопрос о необходимости изменения уровня контроля выясняется сразу после контроля g очередных партий, когда среди них содержится h забракованных партий, а также каждый раз после контроля M партий. В момент корректировки производится вычисление необходимых статистических оценок, доверительных интервалов и т. п. Иногда целесообразно более полно учитывать текущую информацию о результатах контроля. Например, момент корректировки можно определять не только по числу h забракованных партий, но и в зависимости от случайного числа обнаруженных дефектных объектов в l последних проверяемых партиях и т. п.

Для того чтобы избежать частых ложных переключений режимов контроля, можно использовать метод доверительных интервалов. Например, если рассчитанный в момент корректировки доверительный интервал для среднего уровня входного качества при нормальном ходе производства содержит используемое значение q_n , то это значение q_n сохраняется до следующего момента корректировки. Очевидно, что основным вопросом при таком подходе является выбор соответствующего коэффициента доверия, от величины которого будет зависеть число ложных переходов на другие уровни контроля и длительность запаздывания в необходимых переключениях при существенных изменениях значений q_n .

§ 4.3. Система экономических планов (СЭП)

Расходы на контроль. Приемочный контроль качества массовой промышленной продукции связан с определенными затратами на оплату труда контролеров, содержание помещений и оборудования отделов контроля и т. п. Если почти все контролируемые совокупности принимаются, то основная задача контроля состоит в получении объективной информации о возможных колебаниях уровня качества. Естественно требовать, чтобы получение информации было бы не слишком дорогим. Потребитель и поставщик заинтересованы в выделении обоснованной суммы средств на проведение контроля. Расходы на контроль непосредственным образом связаны с «жесткостью» контроля. Вопрос об уровне расходов решается путем соглашения между потребителем и поставщиком и зависит от конкретных условий производства (стоимости контрольных операций, уровня качества и т. п.). При таком подходе основным параметром является уровень расходов на проведение контроля. Увеличение или уменьшение уровня расходов будет означать соответственное увеличение или уменьшение жесткости контроля. Если поставщиком выделены определенные средства на проведение контроля, то потребитель заинтересован, чтобы эти средства были наилучшим образом использованы для его защиты от возможной приемки совокупностей, содержащих недопус-

тимо большое число дефектных объектов. В некоторых специальных случаях, когда требования к качеству высоки, исходной может быть позиция потребителя, требующего использования контроля с заданной степенью жесткости. В таких случаях поставщик будет стремиться использовать наиболее экономичные планы контроля, выбирая их из множества планов, удовлетворяющих потребителя. Разработке систем планов контроля с учетом экономических показателей уделяется значительное внимание. В Японии был подготовлен стандарт приемочного контроля с учетом экономических показателей. В Дании была издана небольшая книга Хальда и Тирегода [1], в которой были приведены таблицы планов приемочного контроля, минимизирующих суммарные расходы на контроль и ущерб от приемки дефектных объектов в условиях, когда известны априорные распределения числа дефектных объектов в контролируемых совокупностях. Вопросам построения оптимальных систем планов приемочного контроля с учетом расходов посвящены многочисленные статьи, из которых отметим работы Ван дер Вардена [1], Сираждинова [6], Ситтига [1], Хальда [2]—[11], Хамакера [1], Хилла [1], Хорснелла [1], Эйдельманта, Бектаева [1].

По-видимому, универсальной методики построения системы экономических планов контроля, пригодной во всех случаях, не существует. Трудность возникающих при этом проблем состоит в правильном учете конкретных особенностей производства. Поэтому разрабатываемая система экономических планов должна быть достаточно гибкой, допускающей, в частности, возможность ее использования не только в случаях неразрушительного, но и разрушительного контроля. В этом параграфе будет дано изложение основ методики построения системы экономических планов, разработанной в Лаборатории статистических методов МГУ (см. Методика СЭП [1]).

Решение о приемке или браковке совокупности $\mathfrak{F} \times [N, D]$ связано с ущербом (напрасными расходами) двоякого рода. А именно, решив забраковать $\mathfrak{F}$, мы напрасно бракуем годные объекты, число которых равно $K = N - D$. Пусть $E_{\text{БГ}}(D)$ — ущерб от браковки

годных объектов. Если бы эта же совокупность $\mathfrak{P}$ была принята, то последующее использование дефектных объектов также привело бы к некоторому ущербу $E_{\text{ПД}}(D)$. Естественно предположить, что $E_{\text{БГ}}(D)$ монотонно убывает, а $E_{\text{ПД}}(D)$ монотонно возрастает с ростом D . То значение $D_{\text{кр}}$, при котором выполняются неравенства $E_{\text{БГ}}(D_{\text{кр}}) \geq E_{\text{ПД}}(D_{\text{кр}})$, $E_{\text{БГ}}(D_{\text{кр}} + 1) < E_{\text{ПД}}(D_{\text{кр}} + 1)$, будем называть **критическим числом**

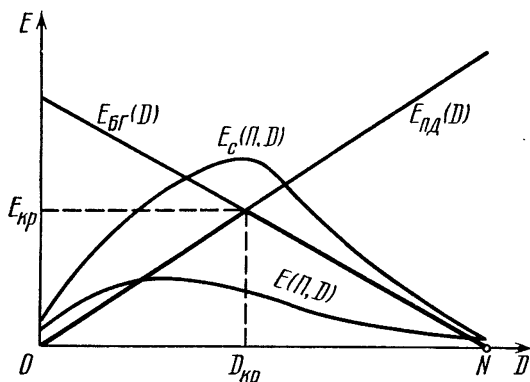


Рис. 4.3.1. Зависимость расходов на контроль от числа дефектных объектов в контролируемой совокупности.

дефектных объектов. Если $D > D_{\text{кр}}$, то данную совокупность выгоднее забраковать без контроля. Если же $D < D_{\text{кр}}$, то ее выгоднее принять без контроля. Таким образом, при известном значении D наименьшие расходы совпадают с $\min(E_{\text{БГ}}(D), E_{\text{ПД}}(D))$. На рис. 4.3.1 наименьшие расходы совпадают с линией, проходящей через $(0, 0)$, $(D_{\text{кр}}, E_{\text{кр}})$, $(N, 0)$. Однако, реальные расходы будут больше, так как фактическое значение D нам неизвестно. Чтобы получить информацию о величине D , нужно использовать некоторый план Π выборочного контроля, что будет связано с некоторыми средними расходами $E(\Pi, D)$. Суммарные средние расходы $E_c(\Pi, D)$ будут включать в себя как средние расходы $E(\Pi, D)$ на проведение контроля, так и средний ущерб от браковки годных и приемки дефектных объектов. Конечно, $E_c(\Pi, D) > \min(E_{\text{БГ}}(D), E_{\text{ПД}}(D))$, однако

$E_c(P, D)$ при правильно организованном контроле будет в среднем существенно меньше расходов при волевых решениях о приемке или браковке совокупности $\mathfrak{F}$ без контроля. Если в нашем распоряжении имеется большое число данных, позволяющих оценить априорное распределение $F_N(D)$ для числа дефектных объектов в контролируемых совокупностях, то средний суммарный ущерб равен

$$E_c(P) = \sum_D E_c(P, D) f_N(D). \quad (3.1)$$

Если ориентироваться на общие средние расходы (3.1), то мы приходим к задаче поиска плана P , минимизирующего эти расходы. Задача эта далеко не простая, а решение ее существенным образом будет зависеть как от вида априорного распределения $F_N(D)$, так и от $E_{пд}(D)$. Если потребители разные и величина ущерба от приемки дефектных объектов для них различна, то при минимизации среднего ущерба (3.1) один из потребителей может получить преимущество перед другим. Возможно, что в таких случаях задачу минимизации (3.1) следует ставить отдельно для каждого потребителя, что фактически соответствует целевому изготовлению совокупности в самом процессе производства (например, экспортная продукция.) Задача еще более усложнится, если при минимизации (3.1) ищутся не только планы P , но и переключающие правила. Однако при таком расширении поставленной задачи необходимо знать не только априорное распределение $F_N(D)$, но и распределения числа дефектных объектов в режимах нормального хода производства и при разладах. Решить столь общую задачу в форме, пригодной для стандартизации, затруднительно. Поэтому мы упростим исходную задачу, рассматривая средние расходы при нормальном ходе производства и линейных функциях потерь.

Пусть a — ущерб от бракования годного объекта, а b — стоимость контроля одного объекта вне зависимости от того, годный он или дефектный. Считаем, что расходы растут пропорционально числу объектов, например, $E_{бг}(D) = a(N - D)$. Если для контроля используется план первого вхождения P с границей

$\Gamma = \Gamma_{\text{пр}} \cup \Gamma_{\text{бр}}$, где $\Gamma_{\text{бр}}(\Gamma_{\text{пр}})$ — часть границы, образованная точками (k, d) , при достижении которых совокупность бракуется (принимается), то средние расходы, связанные с проведением контроля, равны

$$E_N(\Pi, D) = a \sum_{(k, d) \in \Gamma_{\text{бр}}} (K - k) p_D(k, d) + b \sum_{(k, d) \in \Gamma} (k + d) p_D(k, d). \quad (3.2)$$

Первая сумма в (3.2) равна среднему ущербу от забракования $K - k$ годных объектов в непроконтролированной части совокупности, вторая сумма равна средним расходам на контроль объектов, отбираемых в выборку. Заметим, что вторая сумма равна среднему объему выборки при контроле совокупности $\mathfrak{F} \propto [N, D]$ по плану Π с границей Γ . Если нормальному ходу производства соответствует априорное распределение с вероятностями $f_{n, N}(D)$, то средние расходы, связанные с проведением контроля при нормальном ходе производства, равны

$$E_{n, N}(\Pi) = \sum_D E_N(\Pi, D) f_{n, N}(D). \quad (3.3)$$

Заметим, что $\sum_D \left(\sum_{(k, d) \in \Gamma} (k + d) p_D(k, d) \right) f_{n, N}(D) = n_n$ — средний объем выборки при нормальном ходе производства, а $\sum_D \sum_{(k, d) \in \Gamma_{\text{бр}}} p_D(k, d) f_{n, N}(D) = \alpha_n$ — вероят-

ность забракования совокупности, выпущенной при нормальном ходе производства, т. е. фактически вероятность напрасного бракования. Доля дефектных объектов в совокупностях, выпущенных при нормальном ходе производства, как правило, невелика. Поэтому, заменяя $K - k$ на N , получаем

$$\begin{aligned} \sum_D \sum_{(k, d) \in \Gamma_{\text{бр}}} (K - k) p_D(k, d) f_{n, N}(D) &\approx \\ &\approx N \sum_D \sum_{(k, d) \in \Gamma_{\text{бр}}} p_D(k, d) f_{n, N}(D) = N \alpha_n. \end{aligned} \quad (3.4)$$

С учетом сказанного, из (3.3) и (3.4) имеем

$$E_{n, N}(\Pi) \approx a N \alpha_n + b n_n. \quad (3.5)$$

Удобно определить средний относительный уровень расходов

$$E(\Pi) = \frac{E_{n,N}(\Pi)}{aN} \approx \alpha_n + \frac{\mu_n}{\theta} = \alpha_n + \gamma \mu_n. \quad (3.6)$$

Формула (3.6) непосредственно следует из (3.5). В ней $\gamma = \frac{b}{aNq_n}$, $\theta = \frac{aNq_n}{b}$ — стоимостные коэффициенты, $\mu_n = n_n q_n$ — среднее нормированное число объектов в выборке, а $q_n = N^{-1} \sum_D D f_{n,N}(D)$ — среднее значение доли дефектных объектов в контролируемой совокупности при нормальном ходе производства. Переход к приближенной формуле (3.6) дает возможность характеризовать различные планы Π двумя показателями, α_n и μ_n . Однако переход от точных формул (3.2), (3.3) к приближенным формулам (3.5), (3.6) возможен лишь для малого относительного среднего объема выборки. Это требование выполнено для планов Π , у которых $\text{ind } \Gamma/N \ll 1$. Другим показателем возможности использования приближенных соотношений (3.5), (3.6) может служить величина среднего уровня относительных расходов. Если $E < 0,1$, то, как правило, приближение получается вполне удовлетворительным.

Если значения среднего объема выборки соизмеримы с N , а q_n мало, то, заменяя в (3.2) $K - k = N - (k + d) - (D - d)$ на $N - (k + d)$ из (3.3), получим

$$E_{n,N}(\Pi) \approx a(N - n'_n) \alpha_n + b n_n. \quad (3.7)$$

В (3.7) n'_n — условное математическое ожидание объема выборки при условии, что принимается решение о забраковании контролируемой совокупности, выпущенной в условиях нормального хода производства. Полученное соотношение является уточнением (3.5), однако оно сложнее, так как даже при переходе к относительному уровню расходов $E(\Pi)$ приходится оперировать с тремя параметрами α_n , n_n и n'_n .

Рациональные экономичные планы. Будем исходить из предположения, что основным показателем для поставщика являются средние расходы на контроль, определяемые (3.3). Таким образом, поставщик рассматривает любые два плана, Π_1 , Π_2 , имеющие равные средние

расходы $E_{н, N}(P_1) = E_{н, N}(P_2)$, как равноценные. Если поставщик исходит из определенной суммы расходов $E_{н}$, то все планы контроля, средние расходы которых равны $E_{н}$, будем называть **эквивалентными**. Задача состоит в том, чтобы среди множества эквивалентных планов, соответствующих $E_{н}$, отобрать такие, которые наилучшим образом защищали бы потребителя от приемки совокупностей, сильно засоренных дефектными объектами. Количественные показатели, на основе которых должен проводиться такой отбор, могут существенным образом зависеть от условий и характера контроля. Например, если контроль неразрушительный с решениями $\mathcal{D}_2, \mathcal{D}_3$, а по условиям поставок все выявленные при контроле дефектные объекты заменяются на годные, то основным показателем «фильтрующей» способности планов контроля P может служить $A(D, P)$ — среднее число принимаемых дефектных объектов, рассматриваемое как функция от D . Если окажется, что для двух эквивалентных планов P_1 и P_2 в широкой зоне возможных значений D выполнено неравенство $A(D, P_1) < A(D, P_2)$, то для потребителя план P_1 будет предпочтительным. При плане P_1 потребитель будет в среднем принимать меньше дефектных объектов, чем при плане P_2 , тогда как для поставщика, ориентирующегося на расходы $E_{н}$ при нормальном ходе производства, оба плана, P_1 и P_2 , представляются равноценными. При отыскании оптимальных планов множество эквивалентных планов $\Pi(E_{н})$, соответствующих расходам $E_{н}$, разбивается на классы (подмножества), соответствующие различным типам планов. Например, выделяются классы одноступенчатых планов $\Pi^1(E_{н})$, двухступенчатых планов $\Pi^2(E_{н})$, последовательных планов $\Pi^p(E_{н})$ и т. д. Оптимальные эквивалентные планы отыскиваются отдельно для каждого выбранного **класса эквивалентных планов**. При сформулированном выше подходе к отбору планов контроля на основе функций $A(D, P)$ идеальным решением задачи было бы отыскание такого плана P^* , у которого $A(D, P^*) = \inf_{P \in \Pi(E_{н})} A(D, P)$. Это означало бы,

что у плана P^* функция $A(D, P^*)$ совпадает с нижней огибающей семейства функций, соответствующих планам $P \in \Pi(E_{н})$. Однако такая задача неразрешима,

так как нижняя огибающая не совпадает ни с одной из функций $A(D, P)$. Все функции $A(D, P)$ либо лишь частично проходят «вдоль» огибающей, совпадая с ней на некоторых участках, либо идут выше.

Отбор планов среди эквивалентных можно проводить на основе значений их оперативных характеристик $\pi(D, P)$, равных вероятностям принятия контролируемой совокупности, содержащей D дефектных объектов. Из двух эквивалентных планов $P_1, P_2 \in \Pi(E_n)$ в зоне больших значений D считаем допустимым P_1 , если $\pi(D, P_1) < \pi(D, P_2)$. Идеальным решением задачи было бы отыскание такого плана P , оперативная характеристика которого совпадала бы с нижней огибающей $\inf_{P \in \Pi(E_n)} \pi(D, P)$ семейства всех оперативных характе-

ристик эквивалентных планов. Как и предыдущая, эта задача неразрешима. Однако, как показывают результаты конкретных расчетов, в широкой зоне параметров удается найти планы, близкие к «идеальным».

Предположим, что при нормальном ходе производства априорное распределение числа дефектных объек-

тов является биномиальным $\left(f_{n, N}(D) = \binom{N}{D} q_n^D \bar{q}_n^{N-D} \right)$.

Ограничимся классом $\Pi^1(E_n)$ эквивалентных одноступенчатых планов. Из (3.2), (3.3) получаем, что средние расходы на контроль при плане $P = (n, c)$ равны

$$E_{n, N}(n, c) = a(N - n) \bar{B}_n^c(q_n) + bn. \quad (3.8)$$

Если считать, что $q_n \ll 1$, $n \gg 1$, $nq_n = O(1)$, то, используя пуассоновское распределение вместо биномиального в (3.8) для относительного уровня расходов, получим

$$E(n, c) \approx (1 - v) [1 - L_c(\lambda_n)] + \frac{b}{a} v, \quad (3.9)$$

где $\lambda_n = nq_n$ — среднее число дефектных объектов в выборке, $v = n/N$ — относительный объем выборки,

а $L_c(\lambda) = \sum_{0 \leq d \leq c} \frac{\lambda^d}{d!} e^{-\lambda}$. В следующем разделе этого

параграфа будут рассмотрены возможность пуассоновской аппроксимации и погрешности, возникающие при

переходе от (3.8) к (3.9). Уравнение (3.9) можно переписать в следующем виде:

$$E = \left(1 - \frac{\lambda_n}{D_n}\right)(1 - L_c(\lambda_n)) + \frac{b}{a} \frac{\lambda_n}{D_n}, \quad (3.10)$$

где $D_n = Nq_n$ — среднее число дефектных объектов в контролируемой совокупности, выпущенной при нормальном ходе производства. Для упрощения вычислений будем предполагать, что $\lambda_n = nq_n$ и соответственно n может принимать любые значения. Если значение E_n задано, то уровень расходов $E = E_n/aN$. Каждому корню $\lambda_n = \lambda_{n,c}$ уравнения (3.10) соответствует план из $\Pi^1(E_n)$, $c = 0, 1, 2, \dots$. Так как $\lambda_n/D_n = n/N \leq 1$, то имеет смысл искать корни только в интервале $0 \leq \lambda_{n,c} \leq D_n$. Правая часть (3.10), рассматриваемая как функция λ , равна

$$f(\lambda) = \left(1 - \frac{\lambda}{D_n}\right)[1 - L_c(\lambda)] + \frac{b}{a} \frac{\lambda}{D_n}.$$

Непосредственной проверкой легко убедиться, что вторая производная $f(\lambda)$ в интервале $[0, D_n]$ лишь один раз меняет свой знак. Так как $f(D_n) = b/a$, то в интервале $[0, D_n]$ $f(\lambda)$ является монотонно возрастающей функцией либо имеет в этом интервале одну экстремальную точку (максимум), в которой значение $f(\lambda)$ больше b/a . При сплошном контроле расходы равны bN , поэтому с экономической точки зрения контроль целесообразен только в том случае, когда $E < b/a$. Поэтому мы будем предполагать, что $E < b/a$. Учитывая сделанные выше замечания относительно характера поведения $f(\lambda)$, получаем, что для каждого целого c уравнение $f(\lambda) = E$ в интервале $[0, D_n]$ имеет единственное решение, удовлетворяющее условию $E \leq b/a$. Планы $\Pi(n_c, c)$, $n_c = [\lambda_{n,c}/q_n]$, являются искомыми допустимыми из класса $\Pi^1(E_n)$. Заметим, что при $\lambda/D_n \ll 1$ уравнение (3.10) можно заменить приближенным уравнением

$$E = 1 - L_c(\lambda_{n,c}) + \gamma \lambda_{n,c}, \quad (3.11)$$

где $\gamma = b(aD_n)^{-1}$. Уравнение (3.11) проще, так как его решения являются функциями только двух параметров, E и γ , тогда как решения (3.10) зависят от трех параметров, b/a , E , D_n . Легко показать, что корни (3.11) меньше корней (3.10). Вычисления показывают, что

если решение (3.11) удовлетворяет неравенству $(\lambda_{н,с}/D_n) < 0,1$, то оно близко к соответствующему решению $\lambda_{н,с}$ уравнения (3.10). Для выяснения возможности использования (3.11) полезно заметить, что решения (3.11) удовлетворяют неравенствам

$$\frac{\lambda_{н,с}}{D_n} < \frac{\lambda_{н,с+1}}{D_n} < \frac{Ea}{b}, \quad c=0, 1, \dots$$

На конкретных примерах рассмотрим возможные методы отбора планов из множества допустимых планов $P^1(E_n)$.

Пусть каждая забракованная совокупность, содержащая D дефектных объектов, заменяется совокупностью, выпущенной при нормальном ходе производства. Если ограничиться пуассоновской аппроксимацией, то число D'' дефектных объектов, принимаемых при плане (n, c) , $n = \lambda_{н,с} q_n^{-1}$, $n/N \ll 1$, равно

$$D'' = \begin{cases} D - d, & \text{с вероятностью } \frac{\lambda^d}{d!} e^{-\lambda}, \\ D_\omega - d, & \text{с вероятностью } \frac{\lambda_n^d}{d!} e^{-\lambda_n} P_c(\lambda, \lambda_n), \end{cases} \quad (3.12)$$

где $d = 0, \dots, c$, $\lambda = nD/N$, $\lambda_n = nD_n/N$, $P_c(\lambda, \lambda_n) = [1 - L_c(\lambda)][L_c(\lambda_n)]^{-1}$, а D_ω — число дефектных объектов в совокупности, выпущенной при нормальном ходе производства, $ED_\omega = D_n$. Введем **функции фильтрации** $k_{пр,с}(k) = ED''/D_n$ — отношение среднего числа принятых объектов к среднему числу дефектных объектов в совокупности, выпущенной при нормальном ходе производства, $k = D/D_n$. Из (3.12), учитывая предположение $(n/N) \ll 1$, находим, что функции фильтрации равны

$$k_{пр,с}(k) \approx (k-1)L_c(\lambda_{н,с}k) + 1. \quad (3.13)$$

Легко проверить, что $k_{пр,с}(1) = 1$, $k_{пр,с}(k) \downarrow 1$ при $k \rightarrow \infty$. На рис. 4.3.2 приведены графики функций фильтрации семейства эквивалентных одноступенчатых планов, соответствующих $E = 0,01$, $\gamma = 0,0025$. Как видно из рис. 4.3.2, практически равномерно наилучшей фильтрацией обладает план, соответствующий $c = 9$, $\lambda_{н,с} = 3,24$ — корню (3.11) для $E = 0,01$, $\gamma = 0,0025$. Функция фильтрации для этого плана проходит близко от нижней огибающей функций фильтрации всего се-

мейства эквивалентных одноступенчатых планов. Таким образом в рассматриваемом примере можно рекомендовать план (n, c) , $c = 9$, $n = [3,24 \cdot q_n^{-1}]$.

Как уже отмечалось выше, при отборе планов из семейства эквивалентных можно исходить из оперативных

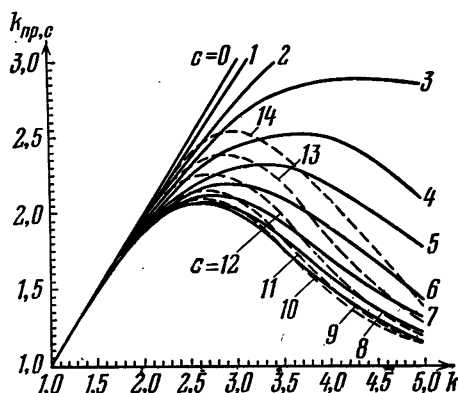


Рис. 4.3.2. Графики среднего выходного качества для эквивалентных одноступенчатых планов с различными приемочными числами.

характеристик. На рис. 4.3.3 даны оперативные характеристики семейства эквивалентных планов, удовлетворяющих уравнению (3.10), когда $b/a = 0,063$, $E = 0,35$, $D_n = 10$. В этом случае $v = n/N > 0,1$, поэтому для расчета $\pi(D, \Pi)$ использовалась формула v -биномиального приближения $\pi(D, \Pi) = \sum_{0 \leq d \leq c} \binom{D}{d} v^d (1-v)^{D-d}$.

Из рис. 4.3.3 видно, что для различных значений $k = D/D_n > 1$ минимум вероятности приемки достигается на различных оперативных характеристиках. В частности, для интервала $1,55 \leq k \leq 1,82$ наилучшим будет план, соответствующий $c = 4$, $\lambda_{n,c} = 320$, а в интервале $1,82 \leq k \leq 2,06$ — план, соответствующий $c = 5$, $\lambda_{n,c} = 3,77$. Из рис. 4.3.3 видно, что в широкой зоне значений k оперативная характеристика эквивалентного плана с приемочным числом $c = 5$ близка к нижней огибающей. Численные подсчеты показали, что отбор планов, оперативные характеристики которых близки к нижней огибающей (кратко такие планы мы назы-

ваем *рациональными*) можно проводить путем сравнения площадей под их оперативными характеристиками.

Поиск рациональных планов проводится на ЭВМ. Рассмотрим задачу поиска рациональных двухступенчатых планов. Пусть n_1, n_2 — объемы первой и второй

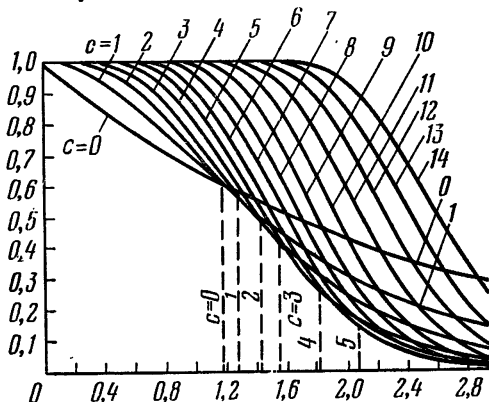


Рис. 4.3.3. Семейство оперативных характеристик для эквивалентных одноступенчатых планов контроля как функции $k = D/D_n$.

выборок, c_1, c_2 — приемочные числа, $r_1 = c_2 + 1$. Введем «временной» параметр t , считая, что $t = 1$ соответствует числу проверяемых объектов $n_1 = q_n^{-1}$. При этом имеем $t_1 = n_1 q_n$, $t_2 = n_2 q_n$. Если априорное распределение для числа дефектных объектов — биномиальное, то в соответствии с (3.2) и (3.3) имеем

$$E_{n,N}(\Pi) = a \{ (N - n_1) \bar{B}_{n_1}^{c_2}(q_n) + \\ + (N - n_1 - n_2) \sum_{c_1 < d \leq c_2} b_{n_1}^d(q_n) \bar{B}_{n_2}^{c_2-d}(q_n) \} + \\ + b \{ n_1 [B_{n_1}^{c_1}(q_n) + \bar{B}_{n_1}^{c_2}(q_n)] + (n_1 + n_2) \sum_{c_1 < d \leq c_2} b_{n_1}^d(q_n) \}. \quad (3.14)$$

Если ограничиться двухступенчатыми планами с малыми относительными объемами $(n_1/N) \ll 1$, $(n_2/N) \ll 1$, то для относительного уровня расходов из (3.14) (используя пуассоновскую аппроксимацию) приходим к уравнению (3.6), в котором $\alpha_n = \alpha$, $\mu_n = \mu$, где

$$\alpha = 1 - L_{c_2}(t_1) + \sum_{c_1 < d \leq c_2} \frac{t_1^d}{d!} e^{-t_1} [1 - L_{c_2-d}(t_2)] \quad (3.15)$$

— вероятность забракования при нормальном ходе производства, а

$$\mu = t_1 + t_2 \sum_{c_1 < d \leq c_2} \frac{t_1^d}{d!} e^{-t_1} \quad (3.16)$$

— средний объем инспекции, выраженный во введенных временных единицах. При фиксированных значениях μ и α существует лишь конечное число эквивалентных двухступенчатых планов, удовлетворяющих (3.6). Если задать значения μ , α и пару приемочных чисел c_1, c_2 , то из (3.16) можно выразить t_2 как функцию t_1 :

$$t_2(t_1) = (\mu_1 - t_1) \left(\sum_{c_1 < d \leq c_2} \frac{t_1^d}{d!} e^{-t_1} \right)^{-1}. \quad (3.17)$$

Значение t_1 находится как решение трансцендентного уравнения

$$1 - \alpha = L_{c_1}(t_1) + \sum_{d_1=c_1+1}^{c_2} \frac{t_1^{d_1}}{d_1!} e^{-t_1} L_{c_2-d_1}(t_2(t_1)), \quad (3.18)$$

которое получается из условия равенства риска поставщика выбранному значению α . Можно показать, что это трансцендентное уравнение имеет либо одно решение t_1 , либо два решения t_1' и $t_1'' > t_1'$, либо совсем не имеет решений (при выбранных значениях μ , α , c_1 , c_2). Отыскав значение t_1 , значение t_2 находим из (3.17). Таким образом, окончательно получаем, что планом, удовлетворяющим условию равенства риска поставщика α и имеющим средний объем инспекции μq_n^{-1} , является $\Pi([t_1 q_n^{-1}], [t_2 q_n^{-1}], c_1, r_1 = c_2 + 1, c_2)$. Установлено, что при наличии двух корней t_1' , t_1'' наилучшие результаты дают планы с использованием большего корня $t_1'' > t_1'$. Оперативные характеристики таких планов имеют меньшие значения в области больших значений засоренностей партий дефектными изделиями. Обозначим правую часть (3.18) $f(t_1, c_1, c_2)$. Легко непосредственным сравнением слагаемых убедиться в справедливости неравенств

$$L_{c_1}(t_1) < f(t_1, c_1, c_2) < f(t_1, c_1, c_2 + 1), \quad 0 \leq t_1 \leq \mu. \quad (3.19)$$

Таким образом, если $L_{c_1}(t_1) > 1 - \alpha$, то уравнение (3.18) не имеет корней. Так как для любого α найдется такое c_1 , что $L_{c_1}(t_1) > 1 - \alpha$ для всех $0 \leq t_1 \leq \mu$, то число планов, имеющих заданные значения μ и α , конечно, и если нет планов для c_1, c_2 , то их нет и для значений $c_1, c_2 + k, k = 1, 2, \dots$. Простой проверкой можно убедиться также в справедливости неравенства

$$f(t_1, c_1, c_1 + 1) < f(t_1, c_1 + 1, c_1 + 2), \quad 0 \leq t \leq \mu. \quad (3.20)$$

Из этого неравенства следует, что если нет планов для значений $c_1, c_1 + 1$, то их нет и для всех значений $c'_1, c'_2, c'_1 > c_1, c'_2 > c_1 + 1$. Все эти результаты суммирует

Теорема 4.3.1. *Существует конечное число двухступенчатых планов, имеющих заданные значения риска поставщика α и среднего объема инспекции μ до принятия решения ($\mu \ll D_H = Nq_H$). Пары приемочных чисел c_1, c_2 ($c_2 > c_1$), для которых существуют такие планы, образуют множество, содержащее точки вида c_1, c_2 , где $c_1 = 0, 1, \dots, c_1(\mu, \alpha)$, а $c_2 = c_1 + 1, \dots, c_2(\mu, \alpha, c_1)$.*

Таблица 4.3.1

Число корней уравнения (3.18)

$c_2 \backslash c_1$	0	1	2	3	$c_2 \backslash c_1$	0	1	2	3
1	1	—	—	—	6	2	2	—	—
2	1	1	—	—	7	2	2	—	—
3	1	1	1	—	8	2	2	—	—
4	2	2	2	—	9	—	—	—	—
5	2	2	2	—					

В таблице 4.3.1 указаны пары значений c_1, c_2 ($c_2 > c_1$), для которых существуют двухступенчатые планы с параметрами $n_1, n_2, c_1, r_1 = c_2 + 1, c_2$, когда $\alpha = 0,1, \mu = 1$. При этом цифрой 1 отмечено наличие одного корня t_1 , цифрой 2 — двух корней t'_1, t''_1 уравнения (3.18). Прочерк означает, что при заданных значениях c_1, c_2 корней нет.

Так как $\alpha = 0,1$, а $\mu = 1$, то средний уровень расходов E определяется по формуле (3.6) после задания

стоимостного коэффициента γ . При этом все эквивалентные планы, соответствующие корням (3.18), будут иметь один и тот же средний относительный уровень расходов. Оперативные характеристики рассматриваемых двухступенчатых планов Π при пуассоновской аппроксимации равны

$$\pi(k) = \pi(kD_n, \Pi) \approx L_{c_1}(kt_1) + \sum_{c_1 < d \leq c_2} \frac{(kt_1)^d}{d!} e^{-kt_1} L_{c_2-d}(kt_2). \quad (3.21)$$

Оперативные характеристики (3.21) приведены на рис. 4.3.4. Они соответствуют двухступенчатым планам

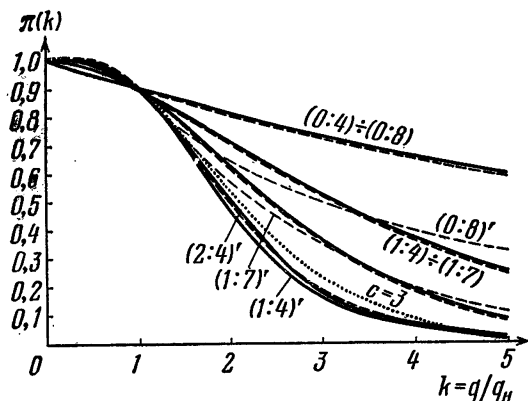


Рис. 4.3.4. Семейство оперативных характеристик двухступенчатых планов контроля с заданными значениями вероятности ошибочного бракования и среднего объема контроля при $k = 1$.

с вероятностью ошибки поставщика $\alpha = 0,1$ и средним объемом инспекции $\mu = 1$. При этом оперативные характеристики планов с приемочными числами c_1, c_2 и наибольшим значением корня t_1'' уравнения (3.18) обозначены $(c_1:c_2)'$, а с наименьшим значением корня обозначены $(c_1:c_2)$. Точечным пунктиром обозначена оперативная характеристика одноступенчатого плана с приемочным числом $c = 3$ и нормированным объемом $\mu = nq_n = 1,745$, для которого $\alpha = 0,1$. Объем контрольной работы при одноступенчатом плане суще-

ственно больше. Из рис. 4.3.4 видно, что как и при одноступенчатых планах, показателем степени рациональности плана является площадь S под оперативной характеристикой. Формула для площади S под оперативной характеристикой двухступенчатого плана с параметрами $t_1 q_n^{-1}$, $t_2 q_n^{-1}$, c_1 , c_2 в условиях применимости пуассоновского приближения имеет вид

$$S = \frac{c_1 + 1}{t_1} + \sum_{d_1=c_1+1}^{c_2} \sum_{d_2=0}^{c_2-d_1} \frac{t_1^{d_1} t_2^{d_2} (d_1 + d_2)!}{d_1! d_2! (t_1 + t_2)^{d_1+d_2+1}}. \quad (3.22)$$

Отбор планов только по этому показателю, вообще говоря, недостаточен. Так, при $\alpha = 0,0025$, $\mu = 2,0$ наименьшее значение S имеет план, у которого $t_1 = 0,948$, $t_2 = 4,297$, $c_1 = 1$, $c_2 = 9$; при этом значение $S_1 = 2,512$. Однако существует план со значением $S_2 = 2,546$, лишь незначительно большим S_1 , который характеризуется параметрами $t_1 = 0,849$, $t_2 = 2,013$, $c_1 = 0$, $c_2 = 6$. Несмотря на несколько большую площадь, второй план в некоторых случаях может оказаться более привлекательным.

Дело в том, что у второго плана максимальный объем выборки равен $(t_1 + t_2) q_n^{-1} = 2,862 \cdot q_n^{-1}$, тогда как у первого плана он равен $5,058 \cdot q_n^{-1}$, т. е. почти в два раза больше, хотя напоминаем, что при нормальном ходе производства в среднем объем инспекции одинаков для обоих планов и равен μq_n^{-1} . Использование планов с большим отношением $t_2 : t_1 = n_2 : n_1$ может быть затруднительным в связи с ограниченной пропускной способностью контрольных постов. Кроме того, такие планы могут привести к неравномерной загрузке контролеров. Итак, при заданных значениях E и γ , меняя значения α и μ в соответствии с уравнением (3.6), мы можем отобрать несколько планов, имеющих относительно небольшие площади S и умеренные значения $t_1 + t_2$. Использование двухступенчатых планов хотя и сопряжено с некоторыми неудобствами, однако дает выигрыш в большей чувствительности к партиям, сильно засоренным дефектными изделиями. При сочетании высоких требований к стабильности качества

принятой продукции с высокой стоимостью контрольной операции этот выигрыш становится существенным.

Следует иметь в виду, что применение при контроле двухступенчатых планов само по себе еще не гарантирует преимуществ по сравнению с одноступенчатыми планами. Из рис. 4.3.4 видно, что двухступенчатые планы могут иметь плохие показатели. Так, для планов с приемочными числами $c_1 = 0$, $c_2 = 4, 5, 6, 7, 8$ вероятности ошибочной приемки партии с большими

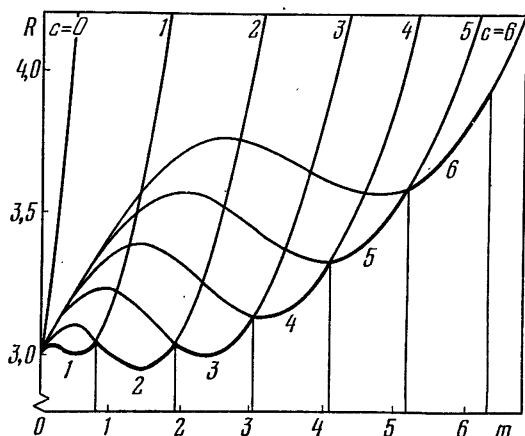


Рис. 4.3.5. Нижняя граница расходов на контроль для различных значений приемочных чисел (Хальд, Тирегод [1]).

значениями $k = q/q_n(D/D_n)$ недопустимо велики. Однако в классе $\Pi^2(E_n)$ можно найти планы, которые защищают потребителя существенно эффективнее, чем одноступенчатые планы. Поиск таких планов в широкой зоне параметров E , γ связан с большим объемом вычислений. Объем вычислений существенно возрастает при поиске рациональных последовательных планов. Осложняющим обстоятельством в такого рода расчетах является структура целевых функций. Нижняя огибающая семейства оперативных характеристик, нижняя огибающая площадей и многие другие функции параметров E , γ образованы из кусков, соответствующих разным планам контроля, что затрудняет применение

градиентных методов при поиске экстремума. При этом может оказаться, что эти функции имеют большое число локальных минимумов. Такое положение, по-видимому, типично для задач, связанных с построением систем экономических планов. Сказанное иллюстрирует рис. 4.3.5, на котором суммарные расходы R на контроль представлены как функции от объема выборки и приемочного числа одноступенчатого плана, когда априорное распределение числа дефектных объектов в совокупности является пуассоновским со случайным значением параметра. Рис. 4.3.5 взят из книги Хальда и Тиревода [1]. Выделенная жирной линией нижняя огибающая имеет несколько локальных минимумов.

Основные принципы построения СЭП. В 1972 г. Государственным комитетом стандартов была издана подготовленная в Лаборатории статистических методов МГУ при содействии Всесоюзного Научно-исследовательского института (ВНИИС) Методика СЭП [1]. В методике СЭП [1] даны планы контроля, соответствующие планируемому уровню расходов и обеспечивающие с вероятностью, близкой к максимально возможной, забракование совокупностей, в которых число дефектных объектов существенно больше, чем при нормальном ходе производства. Одно из условий эффективного использования планов СЭП состоит в том, что уровень качества при нормальном ходе производства является приемлемым для потребителя, тогда как при разладах доля дефектных объектов существенно увеличивается.

Основным управляющим параметром СЭП является средний относительный уровень расходов на контроль E . Приведенные в методике планы контроля и их характеристики рассчитаны на условия, в которых можно использовать пуассоновское приближение, а доля контролируемых объектов невелика. Вторым параметром, определяющим выбор плана в таблицах СЭП, является стоимостной коэффициент $\gamma = b/(a n q_n)$. Если известно значение γ и выбран управляющий параметр E , то характеристики α_n , μ_n эквивалентных планов связаны соотношением (3.6). При этом средние ожидаемые расходы на контроль совокупности объема N , выпущенной в условиях нормального хода производства, приближенно

равны aNE . Было решено ограничиться поиском рациональных планов трех типов: одноступенчатых, двухступенчатых и последовательных. Из эквивалентных планов отбирались такие, у которых площадь Sq_n под оперативной характеристикой являлась минимальной в соответствующем классе эквивалентных планов. Для этого были разработаны необходимые программы.

Исходное предположение о возможности использования пуассоновского приближения выполнено не всегда и поэтому требует специальной проверки. Этот вопрос был исследован для одноступенчатых планов. Каждой паре значений (E, γ) соответствует рациональный одноступенчатый план с параметрами: приемочным числом c и средним числом дефектных объектов nq_n в выборке при нормальном ходе производства. При фиксированном значении γ каждому значению q_n соответствует объем $N = b(a\gamma q_n)^{-1}$. Так как при заданных значениях E, γ величина nq_n не меняется, то при $q_n \downarrow 0$ N неограниченно возрастает и соответственно растет точность пуассоновской аппроксимации. В условиях нормального хода производства при малых значениях q_n априорное распределение $F_{n,N}(D)$ естественно считать биномиальным. Так как $q_n = (M\gamma)^{-1}$, $M = aNb^{-1}$, то точность пуассоновской аппроксимации будет определяться величиной M . Пусть E^* — относительный средний уровень расходов, α^* — вероятность бракования, q^* — уровень качества, при котором вероятность приемки равна 0,1, рассчитаны на основе биномиальных распределений для одноступенчатого плана СЭП, соответствующего E, γ и полученного при пуассоновской аппроксимации. В таблице V для некоторых значений E и интервалов γ , в которых рациональные планы имеют одинаковые значения приемочных чисел c , указаны минимальные числа $M_{0,1}^a, M_{0,1}^E, M_{0,1}^q$ такие, что для всех M , больших этих чисел, относительные ошибки

$$\left| \frac{\alpha^*}{\alpha} - 1 \right| < 0,1, \quad \left| \frac{E^*}{E} - 1 \right| < 0,1, \quad \left| \frac{q^*}{q_{0,1}} - 1 \right| < 0,1,$$

т. е. не превышают 10%. Здесь α — вероятность бракования совокупности, которая дается в таблицах планов СЭП и которая рассчитана на основе пуассоновского приближения, $q_{0,1}$ — доля дефектных объектов,

найденная из таблиц СЭП, при которой вероятность приемки равна 0,1. Предположим, что выбранный уровень расходов $E = 0,04$, $\gamma = 0,13 \in [0,94; 0,133]$. Тогда для рационального плана c равно единице. Параметры этого плана α , E , $q_{0,1} = k_{0,1}q_n$ будут не более чем на 10% отклоняться от более точных значений, найденных исходя из биномиального распределения, если величина M больше соответствующих этим параметрам чисел, найденных линейной интерполяцией по γ на основе граничных значений [741, 1093], [281, 688], [358, 567]. Таблица V рассчитана для значений $\gamma \leq 1$. Для $c = 12$ и $E = 0,1$ числа M приведены лишь для начального значения $\gamma = 0,01$. Таблица V существенно отличается от аналогичных, но менее точных таблиц 2.8—2.9 Методики СЭП [1]. Анализ полученных результатов показывает, что пуассоновским приближением можно пользоваться для совокупностей с небольшими объемами N , если E и γ не слишком малы. Выбор относительной ошибки в 10% является условным. Если допустить относительную ошибку в 20%, то, как показали расчеты, с достаточной степенью точности значения M , приведенные в таблице V, можно уменьшить вдвое.

Так как E выбирается на основе соглашения между поставщиком и потребителем, то при составлении таблиц можно ограничиться дискретным набором предпочтительных значений E . Оказалось целесообразным эти значения задавать в виде ряда, близкого к геометрической прогрессии, в частности, в виде ряда $E = 0,1; 0,063; 0,04; 0,025; 0,016; 0,01$. В отличие от E , при практическом использовании СЭП могут встретиться, вообще говоря, произвольные значения стоимостного коэффициента γ , поскольку q_n может принимать различные значения. Исследование зависимости характеристик оптимальных планов контроля от γ показало, что здесь также можно использовать ряды геометрических прогрессий. Выбор большого числа значений γ нецелесообразен, так как при статистическом оценивании параметра q_n в этом случае пришлось бы излишне часто менять планы контроля за счет малых колебаний при оценке q_n . При расчете оптимальных планов СЭП был выбран следующий набор значений γ : 1,0; 0,8; 0,63; 0,5;

0,4; 0,32; 0,25; 0,2; 0,16; 0,13; 0,1; 0,08; 0,063; 0,05; 0,04; 0,032; 0,025; 0,02; 0,016; 0,013; 0,01.

В Методике СЭП [1] каждой паре значений E , γ соответствуют три плана, Π^1 , Π^2 и Π^3 . Для облегчения выбора типа плана приведены графики оперативных характеристик и графики, показывающие изменение

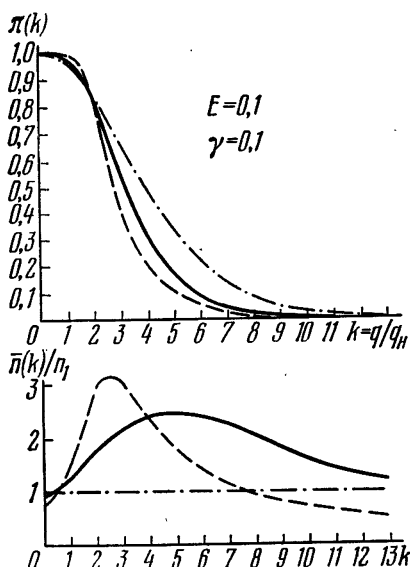


Рис. 4.3.6. Оперативные характеристики и относительные объемы инспекции для планов СЭП ($E = 0,1$, $\gamma = 0,1$).

среднего количества проверяемых до принятия решения объектов при изменениях уровня качества $k = q/q_n$. На рис. 4.3.6 указанные графики даны для значений $E = 0,1$, $\gamma = 0,1$. Штрих-пунктиром даны кривые одноступенчатых планов, сплошной линией — двухступенчатых и пунктиром — кривые последовательных планов. Заметим, что в зоне значений $k \in [2, 4]$ средний объем инспекции при последовательных планах часто существенно больше, чем для одноступенчатых и двухступенчатых планов. В методике СЭП для каждой пары E , γ дана подробная информация о поведении оперативных характеристик в зоне значений k , близких

к нормальному ходу производства ($k = 1$). Приводятся подробные данные о параметрах рациональных планов Π^1 , Π^2 , Π^p , вероятностях ошибок α , средних объемах инспекции при нормальном ходе производства, площадях под оперативными характеристиками.

Подсчеты показали, что в широкой зоне значений E и γ планы Π^2 имеют существенные преимущества перед планами Π^1 . При определенных условиях планы Π^p могут быть более предпочтительными, так как риски ложной забраковки у планов Π^p меньше, чем у планов Π^1 и Π^2 , а оперативные характеристики более круто падают с ростом доли дефектных. Однако планы Π^p могут быть более трудоемки как по среднему количеству изделий, проверяемых до принятия решения при нормальном ходе производства, так и по методам организации контроля и последующей оценке его эффективности. Например, для $E = 0,1$, $\gamma = 0,1$, для рациональных планов Π^1 $\mu_n = nq_n = 0,68$, для Π^2 $\mu_n = 0,87$, а для плана Π^p $\mu_n = 0,96$.

В таблицах VI, VII приложения даны параметры рациональных одноступенчатых и двухступенчатых планов контроля, оперативные характеристики которых близки к нижним огибающим для множеств допустимых одноступенчатых и двухступенчатых планов, соответствующих указанным значениям E , γ . Эти таблицы содержат 126 планов Π^1 и Π^2 . Для пользования таблицами VI, VII найденное значение $\gamma = b(aNq_n)^{-1}$ нужно округлить до «ближайшего» указанного в таблицах значения. При округлении фактических значений γ можно использовать последовательность разбивающих чисел: 1,14; 0,9; 0,71; 0,56; 0,45; 0,36; 0,28; 0,22; 0,18; 0,145; 0,114; 0,09; 0,071; 0,056; 0,045; 0,036; 0,028; 0,022; 0,018; 0,0145; 0,0114; 0,009. Если фактическое значение $\gamma \in (0,9; 1,14]$, то берутся табличные данные, соответствующие $\gamma \in 1,0$; если $\gamma \in (0,71; 0,9]$, то — данные, соответствующие $\gamma = 0,8$ и т. д. Если задано значение E и вычислено соответствующее значение γ , то в таблице VI находим значение приемочного числа c и величину nq_n . Проиллюстрируем применение таблиц примером. Пусть контроль неразрушительный, $a = b$, $q_n = 0,005$ и для контроля совокупности объема $N = 2000$ выбран относительный уровень расходов $E = 0,1$. При этих условиях стоимостный пара-

метр $\gamma = 0,1$. Паре значений $E = 0,1$, $\gamma = 0,1$ в таблице VI соответствует одноступенчатый план с приемочным числом $c = 2$. Так как табличное значение $nq_n = 0,681$, то полагаем $n = 136$. Для тех же значений $E = 0,1$, $\gamma = 0,1$ по таблице VII находим параметры рационального двухступенчатого плана: $c_1 = 1$, $c_2 = 5$, $n_1 = 134$, $n_2 = 276$. Средние расходы на проведение контроля при нормальном ходе производства практически одинаковы для обоих найденных планов. Однако, как видно из рис. 4.3.6, вероятности приемки совокупностей, содержащих число дефектных объектов в несколько раз большее D_n , для двухступенчатого плана будут существенно меньше.

В условиях длительного применения приемочного контроля целесообразно заготовить рабочую таблицу, где определяющими план контроля являются интервалы для объемов контролируемых совокупностей. В таблице 4.3.2 приведен фрагмент рабочей таблицы из Методики СЭП [1], соответствующей $q_n = 0,0025$, $b/a = 0,5$, $E = 0,063$.

Таблица 4.3.2

Объем N	Тип плана	Параметры
от 282 до 357	P^2	$n_1 = 32$ $n_2 = 100$ $c_1 = 0$, $c_2 = 2$, $r_1 = 3$
от 358 до 444	P^2	$n_1 = 38$, $n_2 = 107$ $c_1 = 0$, $c_2 = 2$, $r_1 = 3$
от 445 до 556	P^1	$n = 55$, $c = 1$
от 557 до 714	P^1	$n = 65$, $c = 1$
от 715 до 909	P^1	$n = 76$, $c = 1$
от 909 до 1111	P^1	$n = 86$, $c = 1$

Интервалы объемов N получаются на основе указанных выше рекомендаций по округлению чисел γ . Для подсчитанных значений E , γ в рабочую таблицу включается один из рациональных планов P^1 , P^2 или P^p .

На этом мы закончим краткое знакомство лишь с некоторыми частями Методики СЭП [1].

ГЛАВА 5

О РАЗНЫХ ЗАДАЧАХ ВЫБОРОЧНОГО КОНТРОЛЯ

§ 5.1. Контроль непрерывного потока продукции

Планы Доджа выборочного контроля. В предыдущих главах подробно был рассмотрен вопрос о выборочном контроле как отдельных совокупностей, так и последовательностей совокупностей. Однако деление контролируемых объектов на совокупности иногда может быть весьма условным. Например, так будет в случае непрерывного массового производства однородных видов штучной продукции. Здесь трудно определить, где начинается и где кончается та часть продукции, которую следует забраковать, если результаты контроля говорят об ухудшении качества. При рассмотрении задач организации контроля непрерывного потока объектов O_t , $t = 0, \pm 1, \pm 2, \dots$, ограничимся случаем классификации объектов по альтернативному признаку, сопоставляя объекту O_t величину $\varepsilon_t = 0$ или 1 в соответствии с тем, годный объект O_t или нет. Рассмотрим планы контроля, приведенные в работе Доджа [1]. Контроль начинается со сплошной проверки i объектов, следующих один за другим. Если все проконтролированные объекты окажутся годными, то переходят к случайному контролю с вероятностью f , $0 < f < 1$. При этом для каждого очередного объекта с вероятностью f производят контроль O_t , с вероятностью $1 - f$ принимают O_t без контроля. Если O_t принят без контроля либо проконтролирован и $\varepsilon_t = 0$, то с той же вероятностью f независимо от остальной информации в прошлом подвергают контролю O_{t+1} . С вероятностью $1 - f$ объект O_{t+1} принимают без контроля и т. д. Если при

случайном выборочном контроле обнаружен дефектный объект, то переходят к сплошному контролю, который ведут до тех пор, пока не будут приняты как годные подряд i объектов. Другими словами, для перехода от сплошного контроля к выборочному необходимо, чтобы после очередного обнаруженного дефектного объекта в результате сплошного контроля было бы выявлено i годных объектов. При этом можно различать два вида контроля: при первом все обнаруженные дефектные объекты исправляются или заменяются на годные, при втором такой замены не производится, а все обнаруженные дефектные объекты изымаются. Кратко будем обозначать введенные выше планы контроля непрерывного потока объектов НВК-1 или НВК-2 в зависимости от вида контроля. После исследований Доджа [1] предлагались различные планы контроля непрерывного потока объектов; см. Вальд и Вольфовиц [1]. Например, были исследованы планы с несколькими уровнями контроля, при котором задают пары чисел (f_k, i_k) , $k = 0, \dots, l$, $f_l < f_{l-1} < \dots < f_0 = 1$, i_k — целые числа. Если контроль ведется на уровне k , то объекты отбираются на контроль с вероятностью f_k . При последовательном обнаружении i_k годных объектов переходят к контролю на уровне $k+1$. При обнаружении дефектного объекта переходят к «нулевому» уровню контроля, т. е. начинается сплошной контроль, который ведется до тех пор, пока не будет обнаружена серия из i_0 годных объектов. После этого переходят к выборочному контролю на уровне $k=1$ и т. д. По поводу планов с несколькими уровнями см. статьи Либермана и Соломона [1], Дермана, Джонса, Либермана [1], Дермана, Литтауэра, Соломона [1]. Обзор работ по контролю непрерывного потока объектов дан в статье Боукера [1].

Вероятностные характеристики планов контроля определяются свойствами случайной последовательности e_t . Как и в большинстве статей, посвященных контролю непрерывного потока объектов, будем предполагать, что e_t есть биномиальная последовательность, т. е. e_t — взаимно независимые случайные величины $P\{e_t = 1\} = 1 - P\{e_t = 0\} = q$. Выведем формулу для средней доли $f(q)$ объектов, контролируемых по планам Доджа с параметрами f и i . Так как значения e_t взаимно

независимы и выборочный контроль производится независимо от объекта к объекту, то номера объектов, проверяемых при выборочном контроле, являются моментами регенерации, т. е. «будущее» процесса после момента регенерации не зависит от прошлого. Последовательность моментов регенерации образует процесс восстановления (определение см. в § 2.1). Средняя длительность интервала между двумя последовательными моментами регенерации подсчитывается следующим способом. Во-первых, замечаем, что каждый интервал регенерации заканчивается серией объектов, принимаемых без контроля. Без контроля пропускается k объектов с вероятностью $(1-f)^k f$. Поэтому среднее число объектов в интервале регенерации, которые пропускаются без контроля, равно

$$\sum_{k=1}^{\infty} k(1-f)^k f = f^{-1} - 1. \quad (1.1)$$

С вероятностью q в момент регенерации будет обнаружен дефектный объект. После этого контроль ведется до момента, когда впервые будет обнаружена серия из i годных объектов. Как известно (см. § 1.5, Феллер [1], стр. 325), математическое ожидание числа объектов в биномиальной последовательности e_i до первого появления серии из i нулей равно $(1-\bar{q}^i)(q\bar{q}^i)^{-1}$. Поэтому в среднем контроль проходят

$$1 + q(1-\bar{q}^i)(q\bar{q}^i)^{-1} = (\bar{q}^i)^{-1} \quad (1.2)$$

объектов. Таким образом, из (1.1) и (1.2) следует, что среднее число объектов в интервале от момента регенерации до следующего момента регенерации равно

$$1 + q(1-\bar{q}^i)(q\bar{q}^i)^{-1} + f^{-1} - 1 = (f\bar{q}^i)^{-1} [f + (1-f)\bar{q}^i]. \quad (1.3)$$

Средняя доля контролируемых объектов $f(q)$ равна отношению (1.2) к (1.3). Поэтому имеем

$$f(q) = (\bar{q}^i)^{-1} \frac{f\bar{q}^i}{f + (1-f)\bar{q}^i} = \frac{f}{f + (1-f)(1-q)^i}. \quad (1.4)$$

Другим важным показателем плана контроля является **средний уровень выходного качества (СВК)**,

равный средней доле дефектных объектов, оставшихся в непрерывном потоке объектов после контроля. При контроле по плану НВК-1 величина СВК равна $L_1(q)$ — вероятности того, что дефектный объект принят без контроля. События, состоящие в том, что объект дефектный и что он принят без контроля, взаимно независимы и их вероятности равны q и $1 - f(q)$. Таким образом при контроле по плану НВК-1

$$L_1(q) = q[1 - f(q)]. \quad (1.5)$$

При контроле по плану НВК-2 величина СВК равна $L_2(q)$ — условной вероятности того, что объект окажется дефектным при условии, что он принят. Контролируемые дефектные объекты исключаются, вероятность чего равна $qf(q)$. Соответственно, вероятность того, что объект принят, равна $1 - qf(q)$. Таким образом, искомая условная вероятность равна

$$L_2(q) = \frac{q[1 - f(q)]}{1 - qf(q)}. \quad (1.6)$$

Заметим, что в (1.5) и (1.6) $f(q)$ вычисляется по формуле (1.4). Поэтому находим, что

$$L_j(q) = \frac{q(1-f)(1-q)^{i-j+1}}{f + (1-f)(1-q)^{i-j+1}}, \quad j = 1, 2. \quad (1.7)$$

Таким образом, СВК при контроле по плану НВК-1, т. е. с заменой или исправлением дефектных объектов, с параметрами плана f и i , совпадает с СВК при контроле по плану НВК-2 с параметрами f и $i - 1$.

Важной характеристикой является **уровень предельного среднего выходного качества** (ПСВК), равный

$$L_j^* = \max_{0 \leq q \leq 1} L_j(q). \quad (1.8)$$

Пусть q_j^* — величина входного уровня качества, при котором $L_j(q_j^*) = L_j^*$, $j = 1, 2$. Если входной уровень качества $q = q_j^*$, то при контроле по плану НВК- j к потребителю в среднем проникает наибольшее число дефектных объектов. Числа q_j^* и L_j^* связаны простыми соотношениями. Приравнявая к нулю производную $L_j(q)$ по q ,

из (1.7) находим, что

$$\frac{dL_j(q)}{dq} = (1-f) \{ [(1-q)^{i-j+1} - (i-j+1)q(1-q)^{i-j}] \times \\ \times [f + (1-f)(1-q)^{i-j+1}]^{-1} + (1-f)(i-j+1) \times \\ \times (1-q)^{i-j+1} q(1-q)^{i-j} \} \{ f + (1-f)(1-q)^{i-j+1} \}^{-2} = 0. \quad (1.9)$$

Из (1.9) получаем

$$(1-q_j^*)^{i-j+1} = \frac{f[(i-j+2)q_j^* - 1]}{(1-f)(1-q)}. \quad (1.10)$$

Подставляя (1.10) в (1.7) при $q = q_j^*$, имеем

$$L_j^* = \frac{(i-j+2)q_j^* - 1}{i-j+1} = \frac{1-f}{f(i-j+1)} (1-q_j^*)^{i-j+2}. \quad (1.11)$$

Из уравнений (1.11) при заданных значениях L_j^* и q_j^* можно найти параметры f и $(i-j+1)$ планов НВК- j

$$i-j+1 = \frac{1-q_j^*}{q_j^* - L_j^*}, \quad f = \frac{(1-q_j^*)^{i-j+2}}{(i-j+1)L_j^* + (1-q_j^*)^{i-j+2}}. \quad (1.12)$$

Возможные наборы q_j^* , L_j^* должны быть таковыми, чтобы левая дробь в (1.12) давала целые числа.

На основе введенных характеристик $f(q)$, $L_j(q)$ и L_j^* можно определять параметры планов контроля, т. е. вероятность f случайного отбора и длину серии i из годных объектов, необходимую для перехода при плане НВК- j к выборочному случайному контролю с вероятностью f . В работе Мерфи [1] при заданном среднем уровне входного качества q_n указывается план НВК-1, имеющий заданную трудоемкость $f(q_n)$ и заданный ПСВК L_1^* . В статьях Шора и Пахомова [1], [2], [3] детально разрабатываются методы построения систем планов контроля типов НВК- j . В статье Шора и Пахомова [3] отмечается необходимость разработки отечественных стандартов контроля непрерывного потока объектов по альтернативному признаку ($\varepsilon_i = 0,1$). В книге Хэнсена [1] описанию зарубежных стандартов контроля непрерывного потока объектов посвящена специальная глава (глава 13).

Если при нормальном ходе производства задать определенный уровень трудоемкости γ , то параметры f , i планов НВК- j должны удовлетворять соотношению

$$f(q_n) = \frac{f}{f + (1-f)(1-q_n)^i} = \gamma. \quad (1.13)$$

Использование плана НВК- j , удовлетворяющего (1.13), означает, что при нормальном ходе производства в среднем будет проконтролировано $100 \cdot \gamma\%$ объектов в непрерывном потоке. На рис. 5.1.1, *а* для $\gamma = 0,05$, $q_n = 0,01$ даны кривые $1 - f(q)$, равные той доле объектов, которые принимаются без контроля. Иногда по аналогии с приемочным контролем отдельных совокупностей кривые $1 - f(q)$ называют оперативными характеристиками планов контроля. На рис. 5.1.1, *б* для тех же значений параметров планов, что и на рис. 5.1.1, *а*, даны графики функций $L_1(q)$. Все планы, удовлетворяющие (1.13), имеют одинаковую трудоемкость при нормальном ходе производства. Если потребитель ориентируется на максимальное значение возможного брака, т. е. на величину ПСВК, то из рис. 5.1.1, *б* следует, что для потребителя среди планов, удовлетворяющих (1.13), более предпочтительны планы с меньшими значениями f . Однако это не совсем так. Дело в том, что планы с малыми значениями f будут слабо реагировать на увеличение величины q — доли брака. Меньшие значения L_1^* получаются для таких планов потому, что при их использовании при любом фиксированном уровне q будут встречаться продолжительные интервалы сплошного контроля, между которыми также в течение длительного времени будет проводиться случайный выборочный контроль с малой вероятностью f .

Для того чтобы исключить планы НВК- j с малыми значениями f , можно использовать следующий подход Доджа [1]. Пусть уровень качества $q \geq q_1$ ($q_1 > q_n$) считается недопустимым. Потребуем, чтобы при случайном выборочном контроле с вероятностью отбора объектов на контроль, равной f , среди N следующих друг за другом объектов, с вероятностью не меньшей $1 - \alpha$ был обнаружен хотя бы один дефектный объект. Так как указанная вероятность равна $1 - (1 - fq_1)^N$, то фактически требуется, чтобы при заданных N , q_1 и α

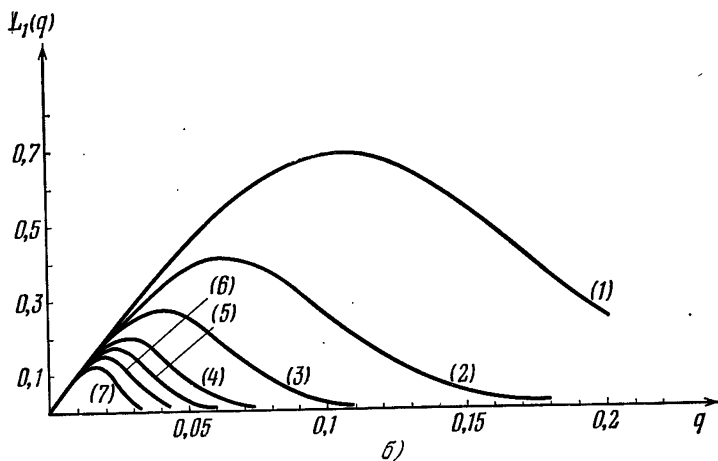
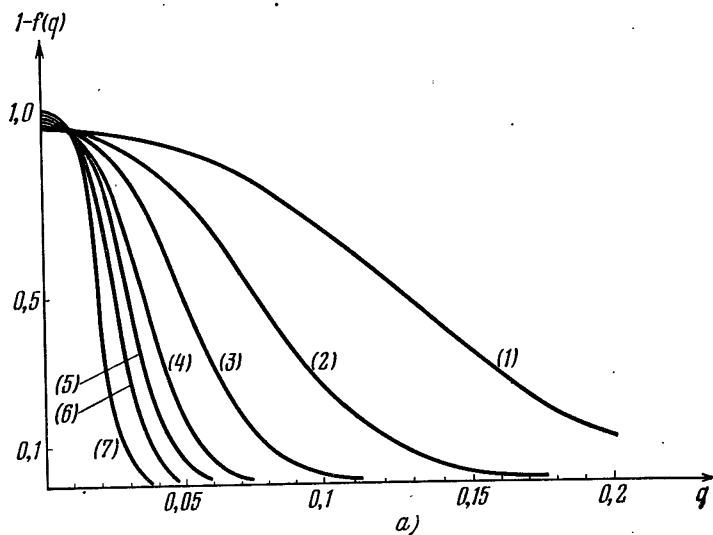


Рис. 5.1.1. Графики а) оперативных характеристик $1 - f(q)$, б) среднего выходного качества $L_1(q)$ для планов НВК-1 со значениями параметров: (1) $f = 0,04$, $i = 23$, (2) $f = 0,0333$, $i = 42$, (3) $f = 0,025$, $i = 71$, (4) $f = 0,0167$, $i = 112$, (5) $f = 0,0125$, $i = 141$, (6) $f = 0,00833$, $i = 182$, (7) $f = 0,004$, $i = 256$.

вероятность f удовлетворяла неравенству

$$(1 - f q_n)^N \leq \alpha. \quad (1.14)$$

Додж [1] предлагал рассматривать в (1.14) $N = 1000$, $\alpha = 0,1$.

Отбор планов НВК- j может быть основан на гарантируемом потребителю ПСВК, равном L_j^* . В этом случае производитель среди всех планов НВК- j , имеющих заданное значение L_j^* , отберет план, у которого при $q = q_n$ будет минимальная трудоемкость. В работе Шора и Пахомова [3] получено следующее утверждение.

Теорема 5.1.1. Пусть $\Pi(L^*)$ — множество планов контроля, имеющих заданное значение ПСВК, равное L^* , а функции СВК $L(q)$ для планов из $\Pi(L^*)$ имеют только один максимум. Тогда плану контроля из $\Pi(L^*)$, для которого L^* — максимум СВК, достигается при $q = q_n$, соответствует наименьшая по сравнению с другими планами трудоемкость контроля $f(q_n)$.

Доказательство. Обозначим $L_\Pi(q)$ и $f_\Pi(q)$ СВК и трудоемкость плана Π . Для планов $\Pi \in \Pi(L^*)$ имеем $\max L_\Pi(q) = L^*$. В теореме утверждается, что

$$\min_{\Pi \in \Pi(L^*)} f_\Pi(q_n) = f_{\Pi_1}(q_n), \quad (1.15)$$

где Π_1 — план, у которого $L_{\Pi_1}(q_n) = \max L_{\Pi_1}(q) = L^*$.

Так как для планов с заменой обнаруженных дефектных объектов годными $L_\Pi(q) = q[1 - f_\Pi(q)]$, то для любого плана Π_2 такого, что $L_{\Pi_2}(q_n) < L^*$ (максимум достигается при $q = q_n$), имеем

$$q_n[1 - f_{\Pi_2}(q_n)] < q_n[1 - f_{\Pi_1}(q_n)].$$

Следовательно, $f_{\Pi_2}(q_n) > f_{\Pi_1}(q_n)$.

Для планов без замены обнаруженных дефектных объектов СВК равно $L_\Pi(q) = q \frac{1 - f(q)}{1 - qf(q)}$. Поэтому, если для Π_2 имеем $L_{\Pi_2}(q_n) < L^*$, а для Π_1 справедливо равенство $L_{\Pi_1}(q_n) = L^*$, то

$$q_n \frac{1 - f_{\Pi_2}(q_n)}{1 - q_n f_{\Pi_2}(q_n)} < q_n \frac{1 - f_{\Pi_1}(q_n)}{1 - f_{\Pi_1}(q_n)}. \quad (1.16)$$

Из (1.16) получаем $f_{\Pi_2}(q_n) > f_{\Pi_1}(q_n)$, что и завершает доказательство теоремы.

Теорема 5.1.1 дает возможность по заданным значениям уровня качества при нормальном ходе производства $q = q_n$ и по величине L — ПСВК найти параметры ПВК- j . Отбор планов с минимальной трудоемкостью при заданном q , имеющих заданное значение ПСВК, рассматривается в работе Резникова [1].

Таблица 5.1.1

**Показатели эффективности планов НВК-1
и одноступенчатых планов с одинаковыми значениями
уровня предельного среднего выходного качества**

$L_1 = 2\%$		Входной уровень качества q в %			
		0,25	0,5	2	4
Параметры планов	$(n, c)_{23}$	$n = 40, c = 1$	$n = 40, c = 1$	$n = 90, c = 3$	$n = 90, c = 3$
	(i, f)	$i = 105, f = 1/45$	$i = 105, f = 1/45$	$i = 95, f = 1/35$	$i = 95, f = 1/35$
Трудоемкость плана с параметрами	$(n, c)_{23}$	4,43%	5,69%	19,07%	53,2%
	(i, f)	2,98%	3,7%	16,85%	59,7%
СВК плана с параметрами	$(n, c)_{23}$	0,24%	0,47%	1,62%	1,87%
	(i, f)	0,24%	0,48%	1,66%	1,61%
Значение q_1 в (1,14) $N = 1000, \alpha = 0,1$		9,6%	9,6%	7,4%	7,4%

Приведем примеры сравнительного анализа эффективности планов НВК-1 Доджа с планами типа одноступенчатой выборки, когда непрерывный поток объектов искусственно разбивается на равные по объему

совокупности, к которым применяются одноступенчатые планы, имеющие заданное значение ПСВК. Результаты вычислений сведены в таблицу 5.1.1, которая взята из работы Шора и Пахомова [1]. Параметры одноступенчатых планов получены на основе таблиц Доджа и Ромига [1].

В таблице 5.1.1 $(n, c)_{23}$ соответствует параметрам одноступенчатых планов с объемом n и приемочным числом c . Из таблицы 5.1.1 видно, что планы НВК-1 могут иметь лучшие показатели, чем одноступенчатые планы. Однако основное преимущество планов НВК-1 в том, что они полнее учитывают особенности непрерывного процесса производства. В книге Хэнсена [1] приводятся примеры, когда использование планов НВК-1 дало заметный экономический эффект в результате ликвидации промежуточных пунктов хранения продукции, уменьшения внутризаводской транспортировки и т. п., необходимых при приемочном контроле.

В статье Хиллиера [1] вводится показатель эффективности планов НВК-1, отличный от $f(q)$, $L(q)$. Этот показатель назван *средним превышением предела дефектности* СППД (Average Extra Defectives Limit — AEDL). Этот показатель определяется следующим способом. Пусть m_0 — момент разладки, когда входной уровень качества скачком возрастает от значения q_0 до q_1 . Пусть $X_m = 1$, если объект O_m — дефектный ($\varepsilon_m = 1$), но он не контролируется. В противном случае полагаем $X_m = 0$. Если значение ПСВК равно L^* , то $EX_m - L^*$ характеризует отклонение от предельного значения. Значение C характеристики СППД определяется по формуле

$$C = \sup_{0 \leq q_0 \leq q_1 \leq 1} \sup_N \sum_{m=m_0+1}^{m=m_0+N} (EX_m - L^*). \quad (1.17)$$

Хиллиер [1] получил явное выражение для величины C , которую следует интерпретировать как максимально возможное среднее значение разности числа дефектных объектов, принимаемых потребителем при скачкообразном увеличении доли брака, и числа принимаемых дефектных объектов при стационарном режиме, когда доля дефектных не меняется. Для плана НВК- j с пара-

метрами f, i величина СППД равна

$$C = \begin{cases} (1-f)f^{-1}[1-(1-f)^{i^*}] - i^*L^*, & \text{если } i^* > 0, \\ 0, & \text{если } i^* \leq 0, \end{cases} \quad (1.18)$$

где $i^* = \ln \left[\frac{-fL^*}{(1-f)\ln(1-f)} \right] [\ln(1-f)]^{-1}$. При выводе (1.18) Хиллиер существенно использовал тот факт, что после момента разладки процесс e_t остается биномиальным. Меняется лишь вероятность появления дефектных объектов. При других типах разладок показатель C в (1.17) не будет выражаться формулой (1.18). Величина СППД растет с уменьшением f , что отражает замедленную реакцию планов НВК- j с малыми f при разладках со скачкообразным увеличением доли дефектных.

Критичные планы выборочного контроля. В связи с отмеченными выше трудностями, возникающими при отборе оптимальных значений параметров f, i планов НВК- j , представляется обоснованным исследование других планов выборочного контроля непрерывного потока продукции. Рассмотрим планы контроля, в которых так же, как и для планов НВК- j , контроль ведется в двух режимах. В режиме случайного отбора каждый очередной объект отбирается на контроль независимо с вероятностью f . При обнаружении дефектного объекта переходят на режим сплошного контроля. Объем последующей контрольной работы при обнаружении каждого очередного дефектного объекта возрастает на i единиц. После окончания периода сплошного контроля начинается режим случайного отбора с вероятностью f , $0 < f < 1$ и т. д. Кратко такие планы назовем **критичными** и будем обозначать через $\Pi[f, i]$. Типичная реализация хода контроля с использованием критичного плана $\Pi[f, i]$ приведена на рис. 5.1.2, где дан график функции $\xi(t)$, равной минимально возможному числу объектов, которые необходимо проконтролировать в режиме сплошного контроля до перехода на режим выборочного контроля с вероятностью случайного отбора на контроль, равной f . Крестиками отмечены номера контролируемых объектов. Объект с номером t_1 — годный и проконтролирован в режиме случайного отбора,

$[t_2, t_3]$ и $[t_4, t_6]$ — интервалы сплошного контроля. В моменты t_2, t_4, t_5 были обнаружены дефектные объекты. Для плана, соответствующего рис. 5.1.2, значение $i = 3$.

При расчете основных характеристик критичных планов контроля, как и выше, будем предполагать, что сопоставляемая непрерывному потоку объектов последовательность e_t образует биномиальный процесс, $P\{e_t = 1\} = q$.

Интервал сплошного контроля начинается с проверки i объектов в момент, следующий за обнаружением дефектного объекта. Если среди i объектов дефектных не обнаружено, то осуществляется переход

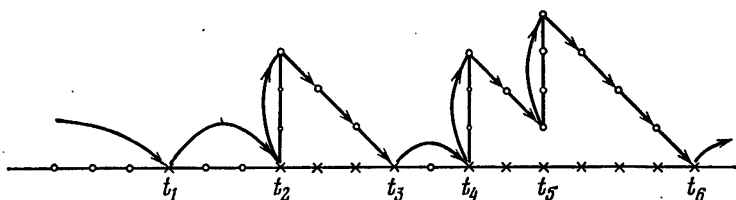


Рис. 5.1.2. Ход контроля при использовании плана $\Pi[f, i]$.

к случайному отбору с вероятностью f . Пусть обнаружено v дефектных объектов, $v = 1, \dots, i$. Длительность сплошного контроля не изменится, если мы будем считать каждый из этих v дефектных объектов начинающим «свой» период сплошного контроля. Пусть β_k — случайные величины, равные числу проконтролированных объектов при сплошном контроле в результате обнаружения k -го дефектного объекта на начальном этапе контроля i объектов, $k = 1, \dots, v$. При этом k -й дефектный объект не включается в β_k . Так как e_t — взаимно независимые случайные величины, то и β_k — взаимно независимые случайные величины. Во введенных обозначениях β — общее число проконтролированных объектов до перехода к случайному отбору, равно

$$\beta = i + \beta_1 + \dots + \beta_v. \quad (1.19)$$

В число β не включается дефектный объект, «начинающий» период сплошного контроля. Случайные величины $\beta, \beta_k, k = 1, \dots, i$, имеют одинаковое распределение ве-

роятностей. Поэтому их производящие функции совпадают. Соотношение (1.19) не противоречит этому факту, так как v есть случайная величина, $P\{v=0\} > 0$ при $q < 1$. Обозначим производящую функцию случайных величин β, β_k через $B(z) = Ez^\beta = Ez^{\beta_k}$. Из взаимной независимости β_k следует, что

$$Ez^{\beta_1 + \dots + \beta_k} = [B(z)]^k. \quad (1.20)$$

Из (1.19) и (1.20), используя формулу полных условных математических ожиданий, находим, что

$$\begin{aligned} B(z) &= Ez^{i+\beta_1 + \dots + \beta_v} = E\{E[z^{i+\beta_1 + \dots + \beta_k} | v=k]\} = \\ &= \sum_{k=1}^i \{z^i [B(z)]^k\} P\{v=k\}. \end{aligned} \quad (1.21)$$

Так как e_t — биномиальный процесс, то $v = \sum_{t=t_0}^{t_0+i} e_t$ имеет биномиальное распределение. Следовательно, $P\{v=k\} = \binom{i}{k} q^k \bar{q}^{i-k}$ и, продолжая вычисления, из (1.21) получаем соотношение

$$\begin{aligned} B(z) &= \sum_{k=1}^i z^i [B(z)]^k \binom{i}{k} q^k \bar{q}^{i-k} = (z\bar{q})^i \sum_{k=1}^i \binom{i}{k} \left[\frac{qB(z)}{\bar{q}} \right]^k = \\ &= (z\bar{q})^i \left[1 + \frac{qB(z)}{\bar{q}} \right]^i = [z\bar{q} + zqB(z)]^i. \end{aligned} \quad (1.22)$$

Функциональное уравнение (1.22) является основным. Исследуя его, получим интересные нас характеристики. Заметим, что вывод (1.22) аналогичен выводу функциональных соотношений для периода занятости в системах массового обслуживания (см. Карлин [1], стр. 497).

Значение $B(1) = \pi$ есть вероятность того, что за период сплошного контроля будет проконтролировано конечное число объектов. Из (1.22) находим, что π является решением уравнения

$$\pi = (\bar{q} + q\pi)^i. \quad (1.23)$$

Так как $g(\pi) = (\bar{q} + q\pi)^i$ — выпуклая вниз функция π , то на интервале $(0, 1)$ не будет решений, отличных от

$\pi = 1$, если $\left. \frac{dg(\pi)}{d\pi} \right|_{\pi=1} = iq \leq 1$. Если $iq < 1$, то средний объем периода сплошного контроля равен $\mu_B = E\beta = \left. \frac{dB(z)}{dz} \right|_{z=1}$. Дифференцируя обе части (1.22) по z , получаем

$$\frac{dB(z)}{dz} = i(z\bar{q} + zqB(z))^{i-1} \left(\bar{q} + qB(z) + zq \frac{dB(z)}{dz} \right). \quad (1.24)$$

Полагая в (1.24) $z = 1$, находим, что

$$\mu_B = i\{1 + q\mu_B\}, \quad \mu_B = \frac{i}{1 - iq}. \quad (1.25)$$

Таким образом, доказано следующее утверждение:

Лемма 5.1.1. Если последовательность ε_t образует биномиальный процесс, $P\{\varepsilon_t = 1\} = q$, $iq < 1$, то за каждый период сплошного контроля по плану $\Pi[f, i]$ с вероятностью 1 контролируется конечное число объектов. Средний объем μ_B периода сплошного контроля находится по формуле (1.25).

Рассмотрим следующий альтернирующий процесс восстановления. Моменты регенерации определим как номера объектов, контролируемых при случайном отборе. Если в момент регенерации t значение $\varepsilon_t = 0$, то далее следует серия неконтролируемых объектов, средний объем которой равен $f^{-1} - 1$. Если же окажется, что в момент регенерации t значение $\varepsilon_t = 1$, то далее сначала следует период сплошного контроля со средней длительностью μ_B , а затем серия неконтролируемых объектов, средняя длительность которой $f^{-1} - 1$. Таким образом, средняя длительность интервала введенного процесса восстановления с учетом (1.25) равна

$$1 + q \frac{i}{1 - iq} + \left(\frac{1}{f} - 1 \right) = \frac{1 - (1 - f)iq}{f(1 - iq)}. \quad (1.26)$$

Из эргодичности введенного процесса восстановления находим, что $f(q)$ — средняя доля контролируемых объектов (уровень контроля) при плане контроля $\Pi[f, i]$ равна отношению среднего объема периода сплошного контроля к средней длительности интервала в процессе восстановления, т. е. равна

$$f(q) = \frac{1 + iq(1 - iq)^{-1}}{[1 - (1 - f)iq] \{f(1 - iq)\}^{-1}} = \frac{f}{1 - (1 - f)iq}. \quad (1.27)$$

Из (1.27) следует, что $f(q) \uparrow 1$ при $q \uparrow i^{-1}$, т. е. контроль по плану $\Pi[f, i]$ переходит в сплошной 100% контроль непрерывного потока объектов.

Формула (1.27) справедлива только при $iq < 1$, поэтому необходимо специальное исследование трудоемкости контроля при $iq > 1$. Если $iq > 1$, то $\left. \frac{dg(\pi)}{d\pi} \right|_{\pi=1} > 1$ и, следовательно, существует значение $\pi < 1$, являющееся решением (1.23). Таким образом, положительная вероятность $\bar{\pi} = 1 - \pi$ того, что период сплошного контроля окажется бесконечным, т. е. будет осуществлен переход на сплошной контроль. Конечные объемы интервалов сплошного контроля — взаимно независимые случайные величины. Следовательно, вероятность того, что при $iq > 1$ на первых k интервалах сплошного контроля не будет осуществлен переход к 100% контролю непрерывного потока объектов, равна π^k . Так как $\pi^k \rightarrow 0$, $k \rightarrow \infty$, то с вероятностью 1 осуществляется переход к 100% контролю.

Мы не рассматриваем правил прекращения 100% контроля, так как вводимый ниже показатель не будет существенным образом зависеть от них, когда сплошной контроль проводится достаточно долго.

Уровень контроля плана $\Pi[f, i]$ при нормальном ходе производства равен $f(q_n)$. Это один из наиболее существенных показателей плана для поставщика. Потребителя интересуют различные показатели того, как использование плана $\Pi[f, i]$ защищает от проникновения брака. Только что было показано, что если q является недопустимым уровнем качества, причем $iq > 1$, то с вероятностью 1 будет лишь конечное число интервалов выборочного случайного контроля. Показателем защиты потребителя от принятия большого числа дефектных объектов является $D(q)$, равное среднему числу дефектных объектов, принятых без контроля, начиная с момента разладки, когда уровень качества стал равным q ($q > q_n$, $iq > 1$). Для простоты будем рассчитывать $D(q)$ при условии, что разладка наступает в момент начала проведения выборочного контроля. Это позволяет избежать вычисления вероятностей состояний процесса контроля в переходных режимах. С момента разладки начинается серия объектов, пропускаемых без контроля.

Ее длина равна k с вероятностью $(1-f)^k \cdot f$. Средняя длина этой серии равна $f^{-1} - 1$, а среднее число принятых дефектных объектов $q(f^{-1} - 1)$. Если объект контролируется, то с вероятностью $\bar{q}$ он оказывается годным, после чего снова следует серия объектов, принимаемых без контроля, средняя длина которой равна $(f^{-1} - 1)$. Если при контроле объект оказался дефектным, то следует переход на сплошной контроль, который с вероятностью $\pi < 1$ завершается переходом на выборочный контроль со средней продолжительностью $(f^{-1} - 1)$. С вероятностью $q\bar{\pi}$ происходит переход на 100% контроль. Таким образом, с вероятностью $(\bar{q} + q\pi)^k q\bar{\pi}$ будет k интервалов выборочного контроля объектов. При этом среднее число дефектных объектов, содержащихся в этих интервалах, равно

$$\sum_{k=1}^{\infty} qk(f^{-1} - 1)(\bar{q} + q\pi)^k q\bar{\pi} = (f^{-1} - 1)(1 - q\pi)/\bar{\pi}.$$

Следовательно, искомая характеристика определяется соотношением

$$D(q) = q(f^{-1} - 1) + (f^{-1} - 1) \frac{1 - q\bar{\pi}}{\bar{\pi}} = \bar{\pi}^{-1}(f^{-1} - 1). \quad (1.28)$$

Характеристика $D(q)$ не имеет прямого аналога для планов Доджа, так как для этих планов при всех $q < 1$ нет перехода на 100% контроль.

Используя показатели $f(q)$ и $D(q)$, можно выбирать параметры планов контроля $\Pi[f, i]$. Например, если уровень $q = q_n$ соответствует нормальному ходу производства, то надо контролировать небольшую долю объектов γ . Задавая $\gamma = f(q_n)$, из (1.27) получаем конечный набор пар величин i и f_i , где i — целые числа. Для $i(iq_n < 1)$ из (1.27) находим, что

$$f_i = \frac{\gamma(1 - iq_n)}{1 - \gamma iq_n}. \quad (1.29)$$

С ростом i вероятность f_i согласно (1.29) убывает. Все планы имеют одинаковую трудоемкость при $q = q_n$. Однако они существенно различаются при $q > q_n$. Если значение $q_{кр} > q_n$ является недопустимым для потребителя, то естественно потребовать, чтобы использовались планы $\Pi[f_i, i]$, у которых $iq_{кр} > 1$. При этом потреби-

тель отбирает тот план контроля, который имеет меньшие значения $D(q)$ при $q \geq q_{кр}$. Например, можно исходить из максимального значения

$$\bar{D}(f, i) = \max_{q \geq q_{кр}} D(q) = D(q_{кр}), \quad (1.30)$$

где $D(q)$ определяется соотношением (1.28). В случае использования характеристики (1.30) потребитель отбирает план $\Pi[f_i, i_0]$ у которого $\bar{D}(f_{i_0}, i_0) = \min_{i, iq_{кр} > 1} \bar{D}(f_i, i)$.

Проиллюстрируем изложенный выше подход на следующем примере. Пусть $q_n = 0,01$, $\gamma = 0,05$, т.е. при уровне входного качества, когда в среднем на сто объектов приходится один дефектный, предполагается контролировать в среднем 5% производимых объектов. С точки зрения поставщика при $q = 0,01$ эквивалентны все планы $\Pi[f_i, i]$, удовлетворяющие (1.29) для $\gamma = 0,05$. Потребитель считает, что повышение доли дефектных до $q \geq q_{кр} = 0,05$ для него неприемлемо. В этом случае целесообразен переход на сплошной контроль. Таким образом, $iq > 1$ при $q > 0,05$, тогда как $iq_n = 0,01i < 1$. Поэтому у планов, удовлетворяющих (1.29), возможны значения $i = 21(1)99$. Характеристики f_i, i некоторых из этих возможных планов приведены в таблице 5.1.2.

Таблица 5.1.2

Параметры критичных планов с уровнем контроля
 $\gamma=0,05$ при $q_n=0,01$

i	25	30	35	40	50	60	80	90	95
f	0,0380	0,0355	0,0331	0,0306	0,0256	0,0206	0,0104	0,0052	0,0026

Из планов, имеющих одинаковый уровень контроля $\gamma = 0,05$ при $q_n = 0,01$, потребитель заинтересован выбрать такой, у которого наименьшее значение $D(q_{кр})$. На рис. 5.1.3 представлены графики функций $D(q)$ только для трех из 79 возможных планов: $f = 0,0380$, $i = 25$; $f = 0,0306$, $i = 40$; $f = 0,0256$, $i = 50$. Вычисления показали, что при $q = 0,05$ наименьшее значение $D(q)$ среди всех 79 возможных планов имеет план

с $i = 40$. Для него $D(0,05) = 39,25$, тогда как, например, для планов с $i = 25$, $D(0,05) = 65,43$, $i = 50$, $D(0,05) = 42,23$. Таким образом, потребитель может отобрать план $\Pi[f = 0,0306, i = 40]$. Конечно, в реальных условиях вместо $f = 0,0306$ можно взять $f = 0,03$, что лишь немного увеличит значение $D(0,05)$ и немного уменьшит уровень контроля, но облегчит реализацию случайного

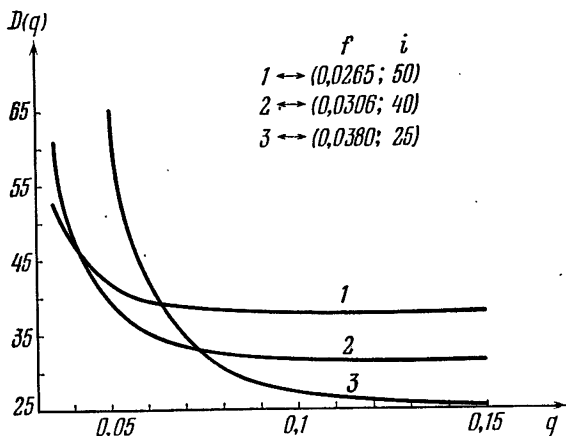


Рис. 5.1.3. Графики $D(q)$ -среднего числа дефектных объектов, проверяемых после разладки до перехода на сплошной контроль при контроле по планам $\Pi[f, i]$.

контроля с помощью датчиков случайных чисел. Поставщику следует иметь в виду, что, вообще говоря, не все возможные планы для него равнозначны. Если $i q_n$ близко к 1, то небольшое увеличение доли дефектных в контролируемой последовательности может привести к резкому увеличению уровня контроля. В рассматриваемом примере резкое увеличение уровня контроля произойдет для отобранного плана с $i = 40$, когда q , возрастающая от $q_n = 0,01$, будет приближаться к значению $q = 0,025$.

Изложенный в этом параграфе материал показывает, что можно организовать контроль непрерывного потока объектов на основе планов, отличающихся от планов Доджа. Преимущества или недостатки каждого способа

контроля существенно зависят от типа возможных разладок. Такие характеристики, как уровень контроля или предельное выходное качество для планов Доджа вычисляются без учета разладок. Рассмотренный выше показатель $D(q)$ рассчитан для разладок, в которых доля дефектных объектов скачком изменяется от значения q_n до значения $q > q_{кр}$. После разладки контролируется биномиальный процесс с параметром q , Беляев [10]. Если типичными являются разладки типа «сбоев», когда качество временно ухудшается, а затем без учета результатов контроля снова повышается до исходного уровня, то возможно, что в этих случаях определяющими будут другие количественные характеристики планов контроля непрерывного потока объектов.

§ 5.2. Планы направленного контроля

Выборочный контроль совокупностей, содержащих серии дефектных объектов. Основным методом отбора объектов из совокупности в выборку является метод случайного отбора, когда объекты на каждом шаге контроля имеют равную вероятность быть отобранными в выборку. Соответствующие планы контроля были подробно рассмотрены выше. Методы случайного отбора являются фактически единственно возможными, когда неизвестен порядок изготовления объектов. Если объекты предъявляются на контроль в порядке их изготовления или имеют индивидуальные номера, по которым этот порядок можно восстановить, то в результате разладок производства дефектные объекты могут следовать друг за другом, образуя серии. Ниже будут приведены некоторые результаты из работы Рыковой [1], которая исследовала методы отбора объектов в выборку, отличные от случайных.

Если в контролируемых совокупностях появляются серии дефектных объектов, то вместо случайного отбора объектов при обнаружении очередного дефектного объекта можно рекомендовать проверку объектов, изготовленных перед и после найденного дефектного объекта. В общем случае формирование выборки целесообразно проводить в зависимости от предыдущих результатов контроля и номеров контролируемых объектов.

Контролируемой по альтернативному признаку совокупности $\mathfrak{F}$ объема N сопоставляется последовательность ε_t , $t = 1, \dots, N$, в которой $\varepsilon_t = 1$, если объект $O_t \in \mathfrak{F}$ дефектный, $\varepsilon_t = 0$, если O_t годный. Предполагается, что каждый объект O имеет свой номер t , по которому его легко выбрать из $\mathfrak{F}$. При **направленном контроле** (НК) номер каждого последующего объекта, отбираемого на контроль, зависит от результатов контроля предыдущих объектов. Если ε_t рассматривать как случайный процесс на ограниченном интервале $1 \leq t \leq N$, для которого известны вероятности всех реализаций, то на контроль можно отбирать тот объект O_t , для которого значение условной вероятности события $\{\varepsilon_t = 1\}$ при фиксированных результатах контроля предыдущих проконтролированных объектов окажется максимальным. Следует ожидать, что так организованный **направленный контроль** будет иметь заметное преимущество перед контролем со случайным отбором объектов, когда в процессе НК для значительной доли объектов O_t указанные условные вероятности $\varepsilon_t = 1$ существенно меньше максимума из этих вероятностей. При обнаружении дефектного объекта может использоваться **процедура окружения**, при которой событие $\varepsilon_t = 1$ влечет последовательную проверку объектов O_{t-1} , O_{t-2} , $\dots$, O_{t+1} , O_{t+2} , $\dots$ до первого обнаружения слева и справа от O_t годных объектов. Если при контроле объекта O_N имеем $\varepsilon_N = 1$, то последовательно проверяются объекты O_{N-1} , O_{N-2} , $\dots$ до тех пор, пока впервые не будет обнаружен годный объект. Возможны различные способы прекращения контроля: (1) контроль прекращается после проверки n объектов; (2) контроль прекращается, если $(n-1)$ -й и n -й контролируемые объекты годные; если же один из них дефектный, то контроль прекращается одновременно с окончанием процедур окружения; (3) контроль прекращается после n этапов, где под **этапом** понимается контроль годного объекта или обнаружение дефектного объекта вместе с процедурой окружения. При способах (2), (3) общее число проверяемых объектов есть случайные величины n_2 , n_3 .

Считаем, что нормальному ходу производства соответствуют биномиальные последовательности ε_t со сред-

ней долей дефектных объектов, равной $q_n = 1 - \alpha_0$ (гипотеза H_0). Под **разлаженным ходом производства** будем понимать такое состояние процесса производства, при котором в контролируемых совокупностях по сравнению с нормальным ходом производства в последовательностях e_t увеличивается либо средняя длина серий из единиц, либо частота появлений серий из единиц, либо то и другое одновременно (гипотеза H_1).

Если d — число дефектных объектов, обнаруженных в процессе контроля, проводимого в соответствии с одним из описанных выше способов отбора объектов на контроль, то при заданном заранее приемочном числе c и $d > c$ ($d \leq c$) принимается решение о забраковании (приемке) совокупности. Сравнительный анализ различных планов контроля проведем на основе их оперативных характеристик. Значения оперативных характеристик следует учитывать как при нормальном, так и при разлаженном ходе производства.

Марковская модель производства. Пусть контролируется совокупность $\mathfrak{F}$, для которой последовательность e_t является марковской последовательностью с матрицей переходных вероятностей

$$\mathcal{P} = \begin{pmatrix} P_{0,0} & P_{0,1} \\ P_{1,0} & P_{1,1} \end{pmatrix} = \begin{pmatrix} 1 - \alpha_1 & \alpha_1 \\ \alpha_0 & 1 - \alpha_0 \end{pmatrix}. \quad (2.1)$$

Для определенности будем считать, что для объекта O_1 $P\{e_1 = 0\} = P_{0,0} = 1 - \alpha_1$. В соответствии с общей теорией цепей Маркова (см., например, Чжун Кай-Лай [1], Романовский [1]), матрица $\mathcal{P}^{(s)}$ переходных вероятностей $P_{i,j}^{(s)} = P\{e_{t+s} = j | e_t = i\}$, $i, j = 0, 1$, равна s -й степени матрицы $\mathcal{P}$ в (2.1), т. е.

$$\mathcal{P}^{(s)} = \begin{pmatrix} P_{0,0}^{(s)} & P_{0,1}^{(s)} \\ P_{1,0}^{(s)} & P_{1,1}^{(s)} \end{pmatrix} = \frac{1}{\alpha_0 + \alpha_1} \begin{pmatrix} \alpha_0 & \alpha_1 \\ \alpha_0 & \alpha_1 \end{pmatrix} + \frac{(1 - \alpha_0 - \alpha_1)^s}{\alpha_0 + \alpha_1} \begin{pmatrix} \alpha_1 & -\alpha_1 \\ -\alpha_0 & \alpha_0 \end{pmatrix}. \quad (2.2)$$

Ограничимся случаем $\alpha_0 \leq 1 - \alpha_1$, что соответствует модели, при которой вероятность производства годного объекта вслед за дефектным $P_{1,0}$ не превосходит вероятности производства годного объекта вслед за годным $P_{0,0}$. При $\alpha_0 + \alpha_1 > 1$ НК существенно отличается от НК в предположении $\alpha_0 + \alpha_1 \leq 1$. Обозначим $P_{g,i,k}^{e,g,i,k}$ условную

вероятность того, что при контроле объекта O_i получим значение e_i при условии, что результатами проверки объектов O_g и O_k являются значения e_g и e_k соответственно. Вероятности $P_{g, i, k}^{e_g e_i e_k}$ вычисляются по формуле

$$P_{g, i, k}^{e_g e_i e_k} = \frac{P_{e_g, e_i}^{(i-g)} P_{e_i, e_k}^{(k-i)}}{P_{e_g, e_k}^{(k-g)}}. \quad (2.3)$$

Рассчитаем по этой формуле вероятности $P_{g, i, k}^{010}$, $P_{g, i, k}^{011}$, $P_{g, i, k}^{110}$ и найдем максимумы из них и максимум вероятности $P_{01}^{(i)}$ по i , $g < i < k$, где g и k — фиксированные числа, $g, k, i = 0, \dots, N$. Например, найденное число i_0 , при котором $P_{g, i, k}^{010} = \max_{g < i < k} P_{g, i, k}^{010}$, будет означать номер объекта, который с наибольшей вероятностью будет дефектным при условии, что на g -м и k -м местах контролируемые объекты оказались годными. Легко проверить, что при $\lambda = 1 - (\alpha_0 + \alpha_1) \geq 0$

$$\max_{1 \leq i \leq N} P_{0,1}^{(i)} = P_{0,1}^{(N)} = \frac{\alpha_1}{\alpha_0 + \alpha_1} (1 - \lambda^N), \quad (2.4)$$

$$\max_{g < i < k} P_{g, i, k}^{011} = P_{g, k-1, k}^{011} = \frac{(1 - \alpha_0)(1 - \lambda^{k-g-1})}{1 - \lambda^{k-g}}, \quad (2.5)$$

$$\max_{g < i < k} P_{g, i, k}^{110} = P_{g, g+1, k}^{110} = P_{g, k-1, k}^{011}, \quad (2.6)$$

$$\max_{g < i < k} P_{g, i, k}^{010} = \begin{cases} P_{g, \frac{g+k}{2}, k}^{010} = B \left(1 - \lambda^{\frac{k-g}{2}} \right)^2, \\ \text{если } (k+g) \text{ — четное,} \\ P_{g, \left[\frac{g+k}{2} \right], k}^{010} = P_{g, \left[\frac{g+k}{2} \right] + 1, k}^{010} = \\ = B \left(1 - \lambda^{\frac{k-g+1}{2}} \right) \left(1 - \lambda^{\frac{k-g-1}{2}} \right), \\ \text{если } (k+g) \text{ — нечетное,} \end{cases} \quad (2.7)$$

где $[a]$ — целая часть числа a , $B = \alpha_0 \alpha_1 / ((\alpha_0 + \alpha_1) \times (\alpha_0 + \alpha_1 \lambda^{k-g}))$.

НК организуем следующим образом. На каждом шаге будем отбирать для контроля объект, для кото-

рого условная вероятность оказаться дефектным будет наибольшей. Если максимальное значение этой вероятности соответствует нескольким объектам, то для контроля с равной вероятностью выбирается любой из них. Из (2.4) — (2.7) следует, что сначала проверяется N -й объект. Если $\varepsilon_N = 0$, то проверяется $[N/2]$ -й (или $[N/2] + 1$ -й) объект. Если же $\varepsilon_N = 1$, то проверяется O_{N-1} и т. д. до обнаружения первого годного объекта. Если ближайшие проконтролированные объекты — O_g, O_k , то контроль объектов $O_i, g < i < k$ проводится аналогично. А именно, при $\varepsilon_g = \varepsilon_k = 0$ контролируется O_i , когда $(k - g) - \text{четное}$, или один из объектов O_i, O_{i+1} , когда $(k - g) - \text{нечетное}$, где $i = g + \left[\frac{k - g}{2} \right]$. При обнаружении дефектного объекта проводится операция окружения.

Таким образом, если ε_t — марковская последовательность, а НК проводится по максимуму условной вероятности события $\varepsilon_t = 1$, то формулы (2.4) — (2.7) приводят к операции окружения, независимо введенной выше. Из (2.2) и (2.3) следует, что при $\alpha_0 + \alpha_1 = 1 (\lambda = 0)$ все значения $P_{0,1}^{(i)}$ и $P_{g,i,k}^{g\varepsilon_i\varepsilon_k}$ одинаковы. Поэтому НК совпадает со случайным отбором объектов в выборку. При $\alpha_0 + \alpha_1 \leq 1$ описанные выше правила НК не зависят от фактических значений α_0, α_1 .

Пусть $\pi(n, c), \pi_j(n, c) (j = 1, 2, 3)$ есть оперативные характеристики планов контроля при случайном отборе n объектов и при НК с одним из определенных выше правил j — прекращения контроля $j = (1), (2), (3)$. Здесь c — приемочное число, а n определяется в зависимости от правил остановки. Точные выражения для оперативных характеристик как функций α_0, α_1 сложны, поэтому ниже приводятся их асимптотические выражения для $c = 1, 2$, когда $N \rightarrow \infty, n = 2' + k \leq M < \infty$. В этих условиях имеем

$$\begin{aligned} \pi_2(n, 1) &= \left(\frac{1}{1 + \gamma} \right)^n [1 + \alpha_0 \gamma (1 + \gamma) + \alpha_0^2 \gamma (1 + \gamma)^2 (n - 3) + \\ &\quad + \alpha_0^2 \gamma (2 + \gamma)] + o(1), \\ \pi_1(n, 1) &= \left(\frac{1}{1 + \gamma} \right)^n [1 + 2\alpha_0 \gamma (1 + \gamma) + \alpha_0^2 \gamma (1 + \gamma)^2 \times \\ &\quad \times (n - 3) + \gamma] + o(1), \quad (2.8) \end{aligned}$$

$$\pi_1(n, 2) = \left(\frac{1}{1+\gamma}\right)^n \left\{ 1 + 2\alpha_0\gamma(1+\gamma) + \alpha_0^2\gamma(1+\gamma)^2(n-3) + \gamma + \left[\frac{3\alpha_0(1-\alpha_0)(1+\gamma)^2}{\gamma} + (2n-9)\alpha_0^3(1+\gamma)^3 + (n-4)\alpha_0^2(1+\gamma)^2 + (2n-8)\frac{\alpha_0^2(1+\gamma)^3(1-\alpha_0)}{\gamma} + \frac{(1-\alpha_0)(1+\gamma)}{\gamma} + \alpha_0(1+\gamma) + \alpha_0^2(1+\gamma)^2 + \binom{n-5}{2}\alpha_0^4(1+\gamma)^4\right]\gamma^2 \right\} + o(1),$$

где $\gamma = \alpha_1/\alpha_0$.

При одном и том же c выполняется соотношение $\pi_3(n, c) \leq \pi_2(n, c) \leq \pi_1(n, c)$. Мы рассмотрим лишь оперативные характеристики $\pi_1(n, c)$ с целью их сравнения с оперативными характеристиками $\pi(n, c)$. В условиях гипотезы $H_0(\lambda=0)$ оперативные характеристики $\pi(n, c)$ и $\pi_1(n, c)$ совпадают и равны

$$\pi_1(n, 1) = \pi(n, 1) = \left(\frac{1}{1+\gamma}\right)^n [1 + n\gamma], \quad (2.9)$$

$$\pi_1(n, 2) = \pi(n, 2) = \left(\frac{1}{1+\gamma}\right)^n [1 + n\gamma + \binom{n}{2}\gamma^2] \text{ и т. п.}$$

Таким образом, в условиях нормального хода производства совпадают и риски поставщика.

Обозначим $D_k^{n,c}$ множество всех тех (α_0, α_1) , $\alpha_0 + \alpha_1 \leq 1$, для которых верно

$$\pi(n, c) \geq k\pi_1(n, c). \quad (2.10)$$

При $c=1$ области $D_k^{n,1}$ находятся следующим образом. Подставляя (2.8) и (2.9) в (2.10), получим

$$\alpha_0^2[k\gamma(1+\gamma)^2(n-3)] + \alpha_0[2k\gamma(1+\gamma)] - [1 + (n-k)\gamma - k] \geq 0.$$

При любом $\gamma > \gamma_k^{n,1} = (k-1)/(n-1)$ это неравенство будет верно при $0 < \alpha_0 < \alpha_k^{n,1}(\gamma)$, где

$$\alpha_k^{n,1}(\gamma) = \frac{-k\gamma + \sqrt{k^2\gamma^2 + k\gamma(n-3)[(n-k)\gamma - k + 1]}}{k\gamma(1+\gamma)(n-3)}. \quad (2.11)$$

Из (2.11) следует, что неравенство $\pi(n, 1) \geq \pi_1(n, 1)$ верно для любого $\gamma > 0$ при $\alpha_0 + \alpha_1 < 1$, а неравенство

$\pi(n, 1) \geq 2\pi_1(n, 1)$ для любого $\gamma > \gamma_2^{n,1} = 1/(n-2)$, когда

$$0 < \alpha_0 < \alpha_2^{n,1}(\gamma) = \frac{-2\gamma + \sqrt{4\gamma^2 + 2\gamma(n-3)[(n-2)\gamma - 1]}}{2\gamma(1+\gamma)},$$

что и определяет искомую область $D_2^{n,c}$ при $c=1$ и т. д.

Проиллюстрируем двумя примерами преимущества НК перед контролем со случайным отбором объектов на контроль в условиях разладок.

Пусть на контроль предъявляются совокупности объема $N = 2^m$, каждой из которых соответствует последовательность e_t марковского типа с матрицей переходных вероятностей (2.1). Чтобы исключить влияние операции окружения, положим приемочное число $c=0$. Для плана НК при $n = 2^l + k$, $0 \leq k < 2^l$, оперативная характеристика равна

$$\begin{aligned} \pi_1(n, 0) &= \\ &= P_{0,0}^{(N)} \left\{ \prod_{s=1}^l [P_{0, N \cdot 2^{-s}, N \cdot 2^{-s}+1}^{0,0,0}]^{2^{s-1}} \right\} (P_{0, N \cdot 2^{-l}-1, N \cdot 2^{-l}}^{0,0,0})^k. \end{aligned} \quad (2.12)$$

При случайном отборе вероятность не обнаружить ни одного дефектного объекта в выборке объема n из совокупности N , содержащей D_ω дефектных объектов, равна $h_{N, D_\omega}^{n,0}$. Так как e_t — марковская последовательность, то

$D_\omega = \sum_{1 \leq t \leq N} e_t$ является случайной величиной. Пусть $P(N, D) = P\{D_\omega = D\}$. Введем вероятности $P_\varepsilon(N, D) = P\{D_\omega = D, e_N = \varepsilon\}$, $\varepsilon = 0, 1$, которые вычисляются по рекуррентным формулам

$$\begin{aligned} P_0(N, D) &= P_0(N-1, D)(1-\alpha_1) + P_1(N-1, D)\alpha_0, \\ P_1(N, D) &= P_0(N-1, D-1)\alpha_1 + P_1(N-1, D-1)(1-\alpha_0) \end{aligned} \quad (2.13)$$

с начальными условиями $P_0(0, 0) = 1$, $P_0(0, 1) = P_1(0, 1) = P_1(0, 0) = 0$. Таким образом, оперативная характеристика плана со случайным отбором объектов с учетом априорных вероятностей $P(N, D) = P_0(N, D) + P_1(N, D)$ равна

$$\pi(n, 0) = \sum_{D=0}^N h_{N, D}^{n,0} P(N, D). \quad (2.14)$$

Вычисления, проведенные по формулам (2.12) — (2.14) для $N = 32$, $n = 8$ и доле брака при нормальном ходе производства $q_n = 0,0125$, дают значения оперативных характеристик $\pi_1(8,0) \approx \pi(8,0) = 0,905$. При разладке, когда $\alpha_0 = 0,3$, $\alpha_1 = 0,13$ ($q \approx 0,3$), значения оперативных характеристик равны $\pi_1(8,0) = 0,067$, $\pi(8,0) = 0,127$. Таким образом, при одном и том же риске поставщика $\alpha = 0,095$ использование НК позволяет уменьшить риск потребителя почти в два раза (при равном объеме контрольной работы).

При $N \rightarrow \infty$ и $c \neq 0$ сравнительные данные НК и контроля со случайным отбором объектов для $n = 90$ и $c = 2$ даны в таблице 5.2.1, где $q = 0,01$ соответствует нормальному ходу производства, а $q = 0,06$ получается при разладках с различными значениями α_0 .

Таблица 5.2.1

Значения оперативных характеристик $\pi_1(90,2)$ и $\pi(90,2)$ при различных (α_0, α_1) ,

$$\left(q = \frac{\alpha_1}{\alpha_0 + \alpha_1}, \gamma = \frac{\alpha_1}{\alpha_0} \right)$$

	H_0	H_1			
$\gamma(q)$	$\frac{1}{99} (q=0,01)$	$\frac{3}{47} (q=0,06)$			
α_0	0,99	0,8	0,6	0,47	0,35
π_1	0,943	0,0582	0,0310	0,020	0,0129
π	0,943	0,088			

При доле брака $q_n = 0,01$ в условиях H_0 оперативные характеристики случайного и направленного контроля с $n = 90$ и $c = 2$ равны 0,943. Если тот же НК проводить в условиях H_1 с $q = 0,06$, когда $\{e_i\}$ — марковская последовательность, то в зависимости от характера разладки (α_0, α_1) риск потребителя уменьшается в несколько раз. Например, при $\alpha_0 = 0,47$ ($\alpha_1 = \gamma\alpha_0 = 0,03$) имеем $\pi_1(90,2) < \frac{1}{4} \pi(90,2)$. Аналогично можно показать, что при заданных значениях рисков α (поставщика) и β (потребителя), q_n и $q > q_n$ применение НК

позволяет существенно уменьшить объем выборки по сравнению с планами со случайным отбором объектов, гарантирующими те же значения α , β .

Альтернирующие последовательности годных и дефектных объектов. Предположим, что ε_t есть альтернирующая последовательность, в которой серии нулей и единиц есть взаимно независимые случайные величины с функциями распределения длительностей серий нулей и единиц соответственно равными $F_0(x)$ и $F_1(x)$. Рассмотренные выше марковские последовательности являются альтернирующими с геометрически распределенными длинами серий из нулей и единиц (см. Кокс и Смит [1], гл. 7). При переходе к общим альтернирующим последовательностям ε_t максимум условных вероятностей $P_{g,i,k}^{0,1,0} = P\{\varepsilon_i = 1 | \varepsilon_g = \varepsilon_k = 0\}$ не всегда достигается при значениях $i_0 = \left\lfloor \frac{g+k}{2} \right\rfloor$ или $i_0 + 1$. Приведем пример. Пусть для альтернирующей последовательности, начинающейся серией из нулей, функция распределения $F_0(x) = 1 - p^x$, $x = 0, 1, \dots$, а серии из единиц имеют постоянную длину r . Для условной вероятности $P_0^0(i, g) = P_0^0(i - g)$ того, что $\varepsilon_i = 0$, если $\varepsilon_g = 0$, имеем

$$P_0^0(i) = \begin{cases} 0, & i < 0, \\ p^i, & 0 \leq i \leq r, \\ P_0^0(i-1)p + P_0^0(i-r-1)(1-p), & i > r, \end{cases} \quad (2.15)$$

$$P_{0,i,g}^{0,1,0} = (1-p) \{P_0^0(g)\}^{-1} \left[\sum_{s=1}^r P_0^0(i-s) P_0^0(g-i-r+s-1) \right]. \quad (2.16)$$

Вычисления, проведенные по (2.15) и (2.16) при $p=0,8$, $r=5$ и $g=40$, $g=20$, $g=10$ показали, что имеет место соотношение $\max_{0 \leq i \leq N} P_{0,i,g}^{0,1,0} = P_{0,r,g}^{0,1,0} = P_{0,g-r,g}^{0,1,0}$, т. е. что

$\max_{0 \leq i \leq N} P_{0,i,g}^{0,1,0}$ не достигается при $i = [g/2]$. Это иллюстрируется таблицей 5.2.2., соответствующей значениям $g=10$, $p=0,8$ и $r=2(1)6$.

Таблица 5.2.2

Значения $P_{0,i,r}^{0,1,0}$, $g=10$, $p=0,8$

$r \backslash i$	2	3	4	5	6
1	0,205	0,129	0,106	0,082	0,102
2	0,257	0,233	0,171	0,164	0,205
3	0,205	0,311	0,237	0,246	0,307
4	0,195	0,234	0,302	0,328	0,410
5	0,208	0,208	0,262	0,410	0,410

Заметим, однако, что $P_{0,i,r}^{0,1,0}$ для $r < i < g - r$ несущественно отличается от $P_{0,r,g}^{0,1,0}$. При этом чем меньше g (интервал между номерами последних проверочных качественных изделий), тем ближе номер изделия, при котором достигается $\max_{0 < i < g} P_{0,i,g}^{0,1,0}$, к номеру $[g/2]$. Таким образом, как и в случае марковской последовательности e_t , в качестве следующего проверяемого изделия можно выбирать изделие, номер которого равен $[g/2]([g/2] + 1)$. Можно ожидать, что НК с окружением для рассматриваемой последовательности приближенно будет иметь те же преимущества, что и для марковской последовательности.

Обозначим через $\tilde{p}_l$ вероятность того, что серия из единиц имеет длину l , $\tilde{p}_l = F_1(l) - F_1(l-1)$, μ_0 и μ_1 — средние длины серий из нулей и из единиц. Проведем сравнение оперативных характеристик НК с остановкой (3)-го типа и контроля со случайным отбором, когда номера i_k контролируемых объектов (без операции окружения, если они годные, или дефектных объектов, начинающих операцию округления), отстоят далеко друг от друга, $i_1 \ll i_2 \ll \dots$. Точнее, мы рассмотрим асимптотический случай, когда для НК число n фиксировано и при $N \rightarrow \infty$

$$\min_{1 \leq k \leq n} (i_k - i_{k-1}) \rightarrow \infty. \quad (2.17)$$

При выполнении условий (2.17) числа дефектных изделий, обнаруженные при НК с окружением при проверке i_k -го изделия, — случайные величины v_k , асимптотиче-

ски взаимно независимые ($k = 1, \dots, n$) и одинаково распределенные при $k = 1, \dots, n-1$. Случайная величина $v^* = v_N$, соответствующая числу дефектных изделий, обнаруженных при центре окружения $i = N$, имеет распределение, отличное от распределения величин v_k , $k = 1, \dots, n-1$. Из основных фактов теории восстановления (см. Кокс и Смит [1], гл. 7) получаем

$$\left. \begin{aligned} p_0 &= P\{v_k = 0\} = \frac{\mu_0}{\mu_1 + \mu_0}, \quad p_0^* = P\{v^* = 0\} = \frac{\mu_0}{\mu_1 + \mu_0}, \\ p_l &= P\{v_k = l\} = \frac{l\bar{p}_l}{\mu_1 + \mu_0}, \quad p_l^* = P\{v^* = l\} = \frac{\sum_{l \leq l \leq \infty} \bar{p}_l}{\mu_1 + \mu_0}, \end{aligned} \right\} \quad (2.18)$$

где $\mu_1 = \sum l\bar{p}_l$, $l = 1, 2, \dots$. Производящая функция $\Pi_d(z)$ для $d = \sum_{1 \leq k \leq n-1} v_k + v^*$ равна

$$\begin{aligned} \Pi_d(z) &= [\Pi v_k(z)]^{n-1} \Pi v^*(z) = \left(\sum_{0 \leq l \leq \infty} p_l z^l \right)^{n-1} \left(\sum_{0 \leq l \leq \infty} p_l^* z^l \right) = \\ &= \sum_{0 \leq s \leq \infty} P\{d = s\} z^s = \sum_{0 \leq s \leq \infty} \bar{p}_s z^s. \end{aligned} \quad (2.19)$$

Так как $\pi_3(n, c) = P\{d < c\} = \sum_{0 \leq s \leq c} \bar{p}_s$, то, используя в (2.19) выражения (2.18), найдем выражения для $\bar{p}_s$. В результате вычислений получаем в условиях (2.17) асимптотические выражения для оперативных характеристик

$$\pi_3(n, 1) \approx \left(\frac{\mu_0}{\mu_1 + \mu_0} \right)^n \left[1 + \frac{(n-1)\bar{p}_1 + 1}{\mu_0} \right], \quad (2.20)$$

$$\begin{aligned} \pi_3(n, 2) \approx \left(\frac{\mu_0}{\mu_1 + \mu_0} \right)^n \left[1 + \frac{(n-2)\bar{p}_1 + 2(n-1)\bar{p}_2 + 2}{\mu_0} + \right. \\ \left. + \frac{(n-1)\bar{p}_1 + \frac{n(n-1)}{2}\bar{p}_1^2}{\mu_0^2} \right]. \end{aligned} \quad (2.21)$$

Аналогичным способом можно получить формулы для $\pi_3(n, c)$, $c = 3, 4, \dots$. Вид формул усложняется с ростом c . Сравним значения оперативных характеристик $\pi_3(n, c)$ и $\pi'(n, c)$, соответствующих НК и контролю со случайным отбором, дополненным операцией окружения, для значений $c = 1, 2$. Заметим, что в условиях нормального хода производства (гипотеза H_0) ϵ_i есть

биномиальная последовательность, поэтому $\pi_3(n, c) = \pi'(n, c)$. Ограничимся тем случаем, когда ε_t есть альтернирующая последовательность с тем же значением средней доли брака q , что и при нормальном ходе производства (гипотеза H_1). Пусть μ_j^i, μ_1^i — средние длины серий нулей и единиц при гипотезе $H_i, i = 0, 1$. Сохранение доли брака $q = \mu_1/(\mu_0 + \mu_1)$ при переходе от H_0 к H_1 эквивалентно равенству отношений средних

$$\frac{\mu_0^1}{\mu_0^0} = \frac{\mu_1^1}{\mu_1^0} = \beta. \quad (2.22)$$

При нормальном ходе производства H_0 имеем

$$\mu_0^0 = (1 - \alpha_0)^{-1}, \quad P\{L_0 = l\} = \alpha_0^{l-1} (1 - \alpha_0),$$

$$\mu_1^0 = \alpha_0^{-1}, \quad P\{L_1 = l\} = \alpha_0 (1 - \alpha_0)^{l-1}, \quad l = 1, 2, \dots$$

$$\pi_3(n, 1 | H_0) = \left(\frac{\mu_0^0}{\mu_1^0 + \mu_0^0} \right)^n \left[1 + \frac{(n-1)\alpha_0 + 1}{\mu_0^0} \right], \quad (2.23)$$

$$\pi_3(n, 2 | H_0) =$$

$$\begin{aligned} &= \left(\frac{\mu_0^0}{\mu_1^0 + \mu_0^0} \right)^n \left[1 + \frac{2 + (n-2)\alpha_0 + 2\alpha_0(1 - \alpha_0)(n-1)}{\mu_0^0} + \right. \\ &\quad \left. + \frac{(n-1)\alpha_0 + \binom{n-1}{2}\alpha_0^2(1 - \alpha_0)}{(\mu_0^0)^2} \right]. \quad (2.24) \end{aligned}$$

Если (2.22) выполнено, то сравнение (2.20) и (2.23) с учетом приведенного выше совпадения оперативных характеристик при НК и при случайном отборе с окружением в условиях нормального хода производства показывает, что при H_1 $\pi_3(n, 1) < \pi'(n, 1)$, как только

$$\tilde{\rho}_1 \leq \frac{\beta}{\mu_1^0} + \frac{\beta - 1}{n - 1}. \quad (2.25)$$

В частности, при $\mu_1^1 = \mu_1^0$ и $\mu_0^1 = \mu_0^0 (\beta = 1)$ неравенство (2.25) принимает вид $\tilde{\rho}_1 \leq (\mu_1^0)^{-1} = \alpha_0$, что соответствует уменьшению доли серий единичной длины при постоянной средней длине μ_1 серий из единиц. При

этом возрастает вероятность появления серий длины больше единичной.

Значительно труднее указать при выполнении соотношений (2.22) множество всех тех $\tilde{p}_1$ и $\tilde{p}_2$, при которых

$$\pi_3(n, 2/H_1) \leq \pi'(n, 2).$$

Формальное сравнение (2.21) и (2.24) приводит к неравенству, которое верно, в частности, если имеют место неравенства

$$\beta \geq 1, \quad \tilde{p}_a \leq \alpha_0, \quad \tilde{p}_2 \leq (1 - \alpha_0) \alpha_0. \quad (2.26)$$

Система неравенств (2.26) выполняется при уменьшении доли серий, состоящих из одной или двух единиц.

Итак, направленный контроль может иметь преимущества перед контролем со случайным отбором объектов. Однако следует подчеркнуть, что решения, принимаемые при планах контроля со случайным отбором, не зависят от априорной информации, если D — число дефектных объектов в контролируемой совокупности, — рассматривается как параметр. При НК мы исходили из определенных допущений. Так, при контроле марковской последовательности предполагалось, что $\lambda = \alpha_0 + \alpha_1 \leq 1$. Если же, например, при контроле марковской последовательности $\lambda = \alpha_0 + \alpha_1 > 1$, то НК будет иметь худшие показатели, чем контроль со случайным отбором.

§ 5.3. Планы испытаний, гарантирующие минимум предельного числа критических событий

Постановка задачи. Рассматриваемая в этом параграфе частная задача связана с направлением исследований оптимального планирования расширенных испытаний. В связи с этим мы кратко остановимся на формулировке некоторых задач из этого класса, а затем дадим полное решение одной из них. Задачи планирования расширенных испытаний возникают тогда, когда качество испытываемых объектов либо не известно совсем, либо о нем имеется лишь неполная информация, когда время испытаний ограничено, а появление при испытаниях определенного типа отказов, называемых ниже критическими, крайне нежелательно. Появление крити-

ческих отказов может быть связано с большим экономическим ущербом или даже катастрофическими последствиями. Таким образом, появление критических отказов свидетельствует о полной непригодности данного типа объектов, либо о необходимости их конструктивной доработки. При появлении критических отказов испытания прекращаются. В связи со сказанным представляется предпочтительной программа испытаний, в которой число испытываемых объектов нарастает определенным образом по мере того, как накапливаемая информация подтверждает высокое качество испытываемых образцов изделий. В первом приближении планы расширенных испытаний можно разбить на этапные и непрерывные. При **этапных планах** на испытания каждого объекта затрачивается определенное время t_0 , обнаружение критического отказа относится к концу этого временного интервала. При **непрерывных планах** время испытания каждого объекта жестко не фиксируется и может быть продолжено в зависимости от накопленной информации. Этапные планы характерны для испытаний новых образцов промышленных изделий, рассчитанных на единичное применение, лекарств, новых видов пищи и т. п. Непрерывные планы могут служить упрощенными моделями внедрения новых сложных и дорогостоящих образцов изделий, изготовление которых нарастающими темпами продолжается при успешной эксплуатации уже изготовленных.

Если для проведения этапных испытаний выделено N изделий, и задано предельное время проведения испытаний T , то число этапов не может превышать m , $mt_0 \leq T < (m+1)t_0$. В этих условиях план этапных испытаний определяется набором целых чисел $n_1, \dots, n_m$, $n_1 + \dots + n_m = N$, где n_i — число случайным образом отобранных объектов, испытываемых на i -м этапе. При отсутствии отказов до i -го этапа включительно переходят к $i+1$ -му этапу. В противном случае испытания прекращаются. Кратко такие планы будем обозначать $\Pi(n_1, \dots, n_{m-1}; N)$. Так как каждый объект за время испытаний t_0 либо отказывает либо успешно проходит испытания, то фактически объекты классифицируются по альтернативному признаку. Объект годный, если он проходит испытания успешно, объект — дефектный, если

он отказывает. Такие отказы будем называть **критическими**, поскольку появление хотя бы одного из них влечет прекращение всей программы испытаний. Обозначим D число объектов, которые дают при испытаниях критические отказы. В условиях полного отсутствия информации $0 \leq D \leq N$. Если известна максимально возможная доля дефектных изделий κ , $0 < \kappa \leq 1$, то $D \leq \kappa N$. При фиксированном значении D и плане $\Pi(n_1, \dots, n_{m-1}; N)$ среднее наблюдаемое число критических отказов равно

$$f_D(n_1, \dots, n_{m-1}; N) = n_1 \frac{D}{N} + h_{N, D}^{n_1, 0} \left\{ n_2 \frac{D}{N - n_1} + \right. \\ \left. + h_{N - n_1, D}^{n_2, 0} \left\{ \dots + h_{N - n_1 - \dots - n_{m-2}, D}^{n_{m-1}, 0} \right\} \dots \right\}, \quad (3.1)$$

где $h_{N, D}^{n, 0}$ — вероятность гипергеометрического распределения. В частности, для двухэтапных планов ($m = 2$)

$$f_D(n_1; N) = n_1 \frac{D}{N} + h_{N, D}^{n_1, 0} D. \quad (3.2)$$

Для трехэтапных планов ($m = 3$)

$$f_D(n_1, n_2, N) = n_1 \frac{D}{N} + h_{N, D}^{n_1, 0} \left[n_2 \frac{D}{N - n_1} + h_{N - n_1, D}^{n_2, 0} \cdot D \right]. \quad (3.3)$$

Предельным числом критических отказов $\bar{f}(n_1, \dots, n_{m-1}; N, \kappa)$ называется максимально возможное значение среднего числа критических отказов, т. е.

$$\bar{f}(n_1, \dots, n_{m-1}; N, \kappa) = \max_{0 \leq D \leq N\kappa} f_D(n_1, \dots, n_{m-1}; N). \quad (3.4)$$

Значение $\bar{D}$, при котором достигается этот максимум, назовем **неблагоприятным**. Если показатель (3.4) выбран в качестве основного, то план этапных испытаний, дающий минимум предельного числа критических отказов, будем считать оптимальным. Таким образом, для **оптимального плана этапных испытаний** предельное число критических отказов равно

$$\bar{f} = \min_{\substack{n_1, \dots, n_m \\ n_1 + \dots + n_m = N}} \max_{0 \leq D \leq N\kappa} f_D(n_1, \dots, n_{m-1}; N). \quad (3.5)$$

Ниже будут рассмотрены задачи отыскания таких оптимальных планов.

Для двухступенчатых и трехэтапных планов для специальным образом отобранных значений N и $\kappa = 1$ (полное отсутствие информации о возможной величине D), $\kappa = 0,5$, $\kappa = 0,1$, $\kappa = 0,05$, $\kappa = 0,01$ были найдены параметры оптимальных планов и соответствующие им значения предельных чисел критических отказов $\bar{f}$; см. таблицы VIII, IX приложения. Эти таблицы были рассчитаны на ЭВМ Лаборатории статистических методов МГУ на основе программ, разработанных Т. Н. Дугиной.

Ниже мы ограничимся только этапными планами испытаний. Относительно непрерывных планов отметим здесь лишь то, что они могут быть заданы многими способами. Например, для испытаний N объектов может быть задано предельное время проведения испытаний T и минимальное время суммарной наработки S . При этом требуется таким образом выбрать времена $0 \leq t_1 \leq t_2 \leq \dots \leq t_N \leq T$ подключения объектов на испытания, чтобы при заданной гипотезе о законе распределения времени безотказной работы к моменту первого критического отказа максимум среднего числа испытываемых объектов, взятый по всем возможным функциям распределения, был бы минимальным. При успешных испытаниях без критических отказов наработка должна быть не меньше S .

Содержание настоящего параграфа в основном следует статье автора (см. Беляев [5]), однако таблицы VIII и IX были пересчитаны, что позволяет более точно находить параметры оптимальных планов для значений N , отличных от табулированных. На целесообразность такого пересчета обратил внимание Большев Л. Н.

Двухэтапные планы (пуассоновское приближение). Предполагая, что оптимальный двухэтапный план надо искать среди планов с малым относительным объемом n_1 первого этапа ($(n_1/N) \ll 1$) и малой относительной долей $q = (D/N) \ll 1$ изделий с критическими отказами, можно для функции (3.2) использовать пуассоновское приближение

$$f_D(n_1; N) = \nu\lambda + e^{-\nu\lambda}\lambda, \quad (3.6)$$

где ν и λ удовлетворяют соотношениям $[N\nu] = n_1$, $[\lambda] = D$, $[x]$ — целая часть x . В соответствии с (3.4) и (3.6)

предельное число критических отказов приближенно равно

$$\max_{0 \leq \lambda \leq N\kappa} [\lambda v + e^{-\lambda v} \lambda].$$

Задача состоит в отыскании плана, у которого предельное число критических отказов минимально, т. е. равно

$$\bar{f} = \min_{0 \leq v \leq 1} \max_{0 \leq \lambda \leq N\kappa} [\lambda v + e^{-\lambda v} \lambda]. \quad (3.7)$$

Дифференцируя $\lambda v + e^{-\lambda v} \lambda$ по λ , получаем, что при $v < e^{-2}$ эта функция имеет локальный максимум, который соответствует минимальному положительному корню $\bar{\lambda} = \bar{\lambda}(v)$ уравнения

$$(\bar{\lambda}v - 1)e^{-(\bar{\lambda}v-1)} - ev = 0. \quad (3.8)$$

Введем обозначение $\bar{f}(v) = \bar{\lambda}v + e^{-\bar{\lambda}v}\bar{\lambda}$. Исследуем зависимость $\bar{\lambda} = \bar{\lambda}(v)$ и $\bar{f}(v)$ от v . Из (3.8) следует, что

$$1 < \bar{\lambda}v < 2, \quad v < e^{-2}. \quad (3.9)$$

Беря полный дифференциал по dv в (3.8) и используя (3.8) для исключения $e^{-(\bar{\lambda}v-1)}$, получаем

$$\frac{d(\bar{\lambda}v)}{dv} = \frac{e^{\bar{\lambda}v}}{2 - \bar{\lambda}v}, \quad \frac{d\bar{\lambda}}{dv} = \frac{(\bar{\lambda}v)^2 - \bar{\lambda}v - 1}{v^2(2 - \bar{\lambda}v)}. \quad (3.10)$$

Из (3.10) следует, что с ростом v произведение $\bar{\lambda}v$ растет, а $\bar{\lambda}$ как функция v , сначала убывает, а затем начинает расти. Смена убывания $\bar{\lambda} = \bar{\lambda}(v)$ на рост при увеличении v происходит при значении $\bar{\lambda}v = \frac{1 + \sqrt{5}}{2}$. Используя соотношения (3.8) — (3.10), получаем, что при $v < e^{-2}$

$$\begin{aligned} \frac{d\bar{f}(v)}{dv} &= (1 - \bar{\lambda}e^{-\bar{\lambda}v}) \frac{d(\bar{\lambda}v)}{dv} + e^{-\bar{\lambda}v} \frac{d\bar{\lambda}}{dv} = \\ &= -\frac{1}{v(2 - \bar{\lambda}v)} + \frac{(\bar{\lambda}v)^2 - \bar{\lambda}v - 1}{v(\bar{\lambda}v - 1)(2 - \bar{\lambda}v)} = -\frac{\bar{\lambda}v}{v(\bar{\lambda}v - 1)} < 0. \end{aligned} \quad (3.11)$$

Следовательно, с ростом v функция $\bar{f}(v)$ монотонно убывает.

Для решения задачи поиска оптимального плана надо учесть граничные значения функции $\lambda v + e^{-\lambda v} \lambda$ для $\lambda = N\kappa$. Функция $g(v) = N\kappa v + e^{-N\kappa v} N\kappa$ сначала быстро убывает при увеличении v до значения $v_0 = \ln(N\kappa)/(N\kappa)$, а затем начинает расти, приближаясь к своей асимптоте $N\kappa v$. Если $v_0 = 0,112$ ($N\kappa > 28,5$), то $g(v_0) = \ln(N\kappa) + 1 < \bar{f}(v_0)$. Так как $\bar{f} = \bar{f}(\bar{v})$ находится как минимальное по v значение максимума из двух функций $g(v)$ и $\bar{f}(v)$, то $\bar{v}$ находим как решение уравнения $\bar{f}(\bar{v}) = g(\bar{v})$ или, подробнее,

$$\bar{\lambda} \bar{v} + e^{-\bar{\lambda} \bar{v}} \bar{\lambda} = \bar{v} N\kappa + e^{-\bar{v} N\kappa} N\kappa, \quad (3.12)$$

где $\bar{\lambda} = \bar{\lambda}(v)$ — решение (3.8) при $v = \bar{v}$.

Полученный результат дает возможность вывести простейшую асимптотическую формулу при условии, что $N\kappa \rightarrow \infty$. При этом $\bar{v} \downarrow 0$, поэтому из (3.8) находим, что

$$\bar{\lambda} \bar{v} - 1 = e\bar{v} + \frac{3}{2}(e\bar{v})^2 + o((e\bar{v})^2). \quad (3.13)$$

Используя (3.13), получаем

$$\bar{f}(\bar{v}) = 1/(e\bar{v}) + 1 + o(1). \quad (3.14)$$

Пренебрегая вторым слагаемым в (3.12), из (3.14) имеем, что при $N\kappa \rightarrow \infty$

$$1/(e\bar{v}) + 1 \approx \bar{v} N\kappa. \quad (3.15)$$

Из (3.15) получаем асимптотическую формулу:

$$\bar{v} = 1/\sqrt{eN\kappa} + (2N\kappa)^{-1}. \quad (3.16)$$

В соответствии с (3.16) число $\bar{n}_1$ изделий, испытываемых на первом этапе, при оптимальном планировании приближенно равно

$$\bar{n}_1 \approx \sqrt{\frac{N}{e\kappa}} + \frac{1}{2\kappa}. \quad (3.17)$$

Как показали результаты численных расчетов, формула (3.17) дает вполне удовлетворительное приближение, когда $N\kappa > 28,5$ и $\kappa \geq 0,05$. Для того чтобы учесть случай малых κ , заменим (3.15) более точным соотношением

$$\frac{1}{e\bar{v}} + 1 \approx \bar{v} N\kappa + \exp\left(-\sqrt{\frac{N\kappa}{e}}\right) N\kappa, \quad (3.18)$$

где в соответствии с (3.16) добавлено еще одно слагаемое, приближенно равное последнему слагаемому правой части (3.12). Из (3.18), учитывая существенные слагаемые, получаем

$$\bar{v} \approx \frac{1}{\sqrt{Nek}} \left\{ 1 + \frac{e}{2\sqrt{Nek}} - \frac{1}{2} \exp\left(-\sqrt{\frac{Nk}{e}} - \frac{1}{2}\right) \sqrt{Nek} \right\}. \quad (3.19)$$

Из (3.19) находим, что оптимальное число испытываемых на первом этапе изделий приближенно равно

$$n'_1 \approx \sqrt{\frac{N}{ek}} + \frac{1}{2k} - \frac{N}{2} \exp\left(-\sqrt{\frac{Nk}{e}} - \frac{1}{2}\right). \quad (3.20)$$

Вычислив правую часть (3.20), в качестве приближения n'_1 к n_1 возьмем ближайшее целое число.

Эта формула дает существенно лучшие результаты при малых k , когда $Nk > 25$. Например, для $k = 0,01$ и $N = 3000$ точное значение $n_1 = 339$. По формуле (3.17) получаем $n_1 \approx 332,23 + 50 \approx 382$, тогда как из (3.20) находим $n_1 \approx 332,23 + 50 - 33,15 = 348,85 \approx 349$. Таблица 5.3.1 показывает, как неожиданно близки к оптимальным значения n'_1 , рассчитанные по формуле (3.20), когда $k = 0,05$.

Таблица 5.3.1

Оптимальные значения n_1 и приближенные значения n'_1 , рассчитанные по формуле (3.20) для двухэтапных планов

N	k=1,0		k=0,5		k=0,1		k=0,05		k=0,01	
	n_1	n'_1	n_1	n'_1	n_1	n'_1	n_1	n'_1	n_1	n'_1
80	6	6	8	8	22	18	34	27	—	—
100	6	6	9	9	23	20	37	29	—	—
150	8	8	11	11	27	24	41	35	—	—
250	10	10	14	15	32	32	50	44	125	117
400	12	13	18	18	41	41	59	56	148	136
600	15	15	22	22	51	50	70	70	180	157
1 000	19	20	28	28	65	65	92	92	225	197
3 000	33	34	48	48	109	110	157	158	339	349
10 000	61	61	86	87	196	197	280	281	645	652

Многоэтапные планы. Аналитическое исследование задачи поиска оптимальных многоэтапных планов испытаний существенно сложнее, чем двухэтапных. Для случая трехэтапных планов $f_D(n_1, n_2; N)$ определяется по формуле (3.3).

Здесь, так же как и для двухэтапных планов, можно детально исследовать случай пуассоновского приближения, когда в соответствии с (3.3)

$$f_D(n_1, n_2; N) \approx v_1 \lambda + e^{-v_1 \lambda} \{v_2 \lambda + e^{-v_2 \lambda} \lambda\}, \quad (3.21)$$

где $n_1 = [Nv_1]$, $n_2 = [(N - n_1)v_2]$.

Для широкой зоны параметров v_1, v_2 выражение в правой части (3.21) как функция λ имеет два локальных максимума. Если значения $f_D(n_1, n_2; N)$ в этих локальных максимумах обозначить $\bar{f}_1(v_1, v_2)$ и $\bar{f}_2(v_1, v_2)$ соответственно, то приближенные уравнения, аналогичные (3.12), имеют вид

$$\bar{f}_1(\bar{v}_1, \bar{v}_2) = \bar{f}_2(\bar{v}_1, \bar{v}_2) = \bar{v}_1 N \kappa + e^{-\bar{v}_1 N \kappa} \{\bar{v}_2 N \kappa + e^{-\bar{v}_2 N \kappa} N \kappa\}. \quad (3.22)$$

Так же как и для двухэтапных планов, можно допустить, что

$$N \kappa \rightarrow \infty, \quad \bar{v}_1 \rightarrow 0, \quad \frac{\bar{v}_1}{\bar{v}_2} \rightarrow 0, \quad \frac{1}{\bar{v}_1} \exp\left(-\frac{\bar{v}_2}{\bar{v}_1}\right) \rightarrow 0, \\ N e^{-\bar{v}_1 N \kappa} \rightarrow 0.$$

В этих предположениях в точках $\bar{\lambda}_1$ и $\bar{\lambda}_2$ первого и второго локальных максимумов (3.21) имеют место приближенные соотношения $((\bar{v}_1 + \bar{v}_2) \bar{\lambda}_1 \approx 1, \bar{v}_1 \bar{v}_2 \approx 1)$:

$$\bar{f}_1(\bar{v}_1, \bar{v}_2) \sim 1/(\bar{v}_2 e), \quad (3.23)$$

$$\bar{f}_2(\bar{v}_1, \bar{v}_2) \sim \bar{v}_2/(\bar{v}_1 e). \quad (3.24)$$

Учитывая в правой части (3.22) только первое слагаемое, в соответствии с (3.23), (3.24) получаем

$$1/(\bar{v}_2 e) = \bar{v}_2/(\bar{v}_1 e) = \bar{v}_1 N \kappa. \quad (3.25)$$

Решение уравнений (3.25) имеет вид

$$\bar{v}_1 = (e N \kappa)^{-2/3}, \quad \bar{v}_2 = (e N \kappa)^{-1/3}. \quad (3.26)$$

Таким образом, для оптимального трехэтапного плана из (3.26) находим

$$n'_1 = (e\kappa)^{-2/3} N^{1/3} + o(N^{1/3}), \quad n'_2 = (e\kappa)^{-1/3} N^{2/3} + o(N^{2/3}), \quad (3.27)$$

когда $N\kappa \rightarrow \infty$.

Численные примеры показывают, что формулы (3.27) можно использовать для приближенного вычисления значений параметров планов n_1 и n_2 , когда $\kappa \geq 0,5$. Однако здесь остается нерешенной задача получения поправочных членов, зависящих от κ , которые так же, как это было в случае двухэтапных планов, существенно улучшили бы степень приближения.

Эвристический метод вывода формул, аналогичных (3.27), можно использовать для получения асимптотических характеристик m — этапных оптимальных планов испытаний. При этом пуассоновская аппроксимация функции (3.1) имеет вид

$$\begin{aligned} & v_1 \lambda + v_2 \lambda e^{-v_1 \lambda} + v_3 \lambda e^{-(v_1 + v_2) \lambda} + \dots \\ & \dots + v_{m-1} \lambda e^{-(v_1 + \dots + v_{m-2}) \lambda} + \lambda e^{-(v_1 + \dots + v_{m-1}) \lambda}, \end{aligned} \quad (3.28)$$

$$v_i = \frac{n_i}{N - n_1 - \dots - n_{i-1}}.$$

Снова предполагая, что оптимальный план при $N\kappa \rightarrow \infty$ находится в зоне значений $v_1 \ll v_2 \ll \dots \ll v_{m-1} \ll 1$, вычисляем оптимальные значения $\bar{v}_1, \dots, \bar{v}_{m-1}$, уравнивая с $\bar{v}_1 N\kappa$ величины локальных максимумов функции (3.28) в точках $\bar{\lambda}_1, \dots, \bar{\lambda}_{m-1}$ этих локальных максимумов. Аналогично (3.25), получаем систему уравнений

$$\frac{\bar{v}_2}{\bar{v}_1 e} = \frac{\bar{v}_3}{\bar{v}_2 e} = \dots = \frac{1}{\bar{v}_{m-1} e} = \bar{v}_1 N\kappa. \quad (3.29)$$

Решение системы (3.29) имеет вид

$$\bar{v}_1 = (eN\kappa)^{-1+1/m}, \quad \bar{v}_2 = (eN\kappa)^{-1+2/m}, \quad \dots, \quad \bar{v}_{m-1} = (eN\kappa)^{-1/m}. \quad (3.30)$$

Пользоваться этими асимптотическими формулами, по видимому, можно только при очень больших значениях $N\kappa$, когда $(eN\kappa)^{1/m} \gg 1$.

Поскольку в таблицы VIII, IX приложения включены специальным образом отобранные значения N , то необходимо пояснить рекомендуемые правила расчета параметров оптимальных планов для значений N , не включенных в эти таблицы.

Заметим, что при $D = 1$ разбиение N испытываемых объектов по двум этапам излишне, так как среднее число наблюдаемых критических отказов не зависит от n . Если $D = 2$, то несложные вычисления показывают, что оптимальное значение $n = [N/2]$. Поэтому при заданном κ следует табулировать значения N , для которых $N\kappa > 2$.

В таблице VIII для $\kappa = 1,0$ даны такие граничные значения N , что для промежуточных N число объектов n , испытываемых на первом этапе, такое же, как и для граничных N . Например, в интервале от $N = 16$ до $N = 31$ число $n = 3$. В интервале от $N = 79$ до $N = 110$ число $n = 6$ и т. д. Таким образом, из таблицы VIII для $N = 100$ при $\kappa = 1,0$ находим, что на первом этапе следует испытать $n = 6$ объектов. Если критических отказов не было, то на втором этапе испытываем 94 объекта. Предельное число отказов $\bar{f}$ заключено в интервале от 6,00 до 6,98. Точный расчет показывает, что $\bar{f} = 6,42$, что близко к значению $\bar{f} = 6,63$, получаемому линейной интерполяцией по N . Заметим, что начальное значение $N = 10$ и конечное $N = 1000$ не являются граничными. Для $\kappa = 0,5$ в связи с уменьшением длин интервалов с равными n граничные значения N в таблице VIII даны только до $N = 388$. В интервале от $N = 345$ до $N = 388$ число $n = 17$. Далее приведены лишь шесть значений $N = 500(100)1000$. Вычисления показали, что для $\kappa = 0,1, 0,05$ и $0,01$ наблюдается нарушение монотонности роста n . При увеличении N число n может скачком уменьшаться, а затем снова начинает расти до следующего убывания и т. д. С ростом N величины таких скачков имеют тенденцию к убыванию. Для $\kappa = 0,1$ в таблице VIII даны значения от $N = 25$ до $N = 1000$. Первое нарушение монотонности происходит при переходе от $N = 30$ к $N = 31$, когда n скачком убывает от 15 до 13. После этого в интервале от $N = 31$ до $N = 40$ n начинает возрастать от 13 до 17. При переходе от $N = 40$ к $N = 41$ снова n скачком убывает от 17 до 15 и т. д.

Последнее скачкообразное убывание для $\kappa = 0,1$ дано при переходе от $N = 140$ к $N = 141$, когда n убывает всего лишь на 1 от $n = 26$ до $n = 25$. Далее табулированы значения $N = 150(50)300(100)1000$. При $\kappa = 0,05$ величина скачка возрастает до 4. Последний скачок в таблице VIII дан для $\kappa = 0,05$ при переходе от $N = 260$ к $N = 261$, когда n меняется от 52 до 49. Далее дан конец очередного интервала $N = 280$ и значения $N = 300(100)1000$. Для $\kappa = 0,01$ даны только два случая скачкообразного уменьшения n при переходе от $N = 299$ к $N = 300$, когда n изменяется от 149 до 127, т. е. скачком уменьшается на 22. Затем n увеличивается до 169 при $N = 399$ и снова скачком при $N = 400$ убывает до 148 и т. д. Кроме указанных значений, приведены значения $N = 250, 350, 500, 600, 800, 1000$.

В таблице IX даны значения объемов n_1, n_2 объектов, испытываемых на первом и втором этапах для трехэтапных планов, имеющих минимальное значение предельного числа критических отказов. Для небольших значений N (при $\kappa = 1,0$ до $N = 88$, при $\kappa = 0,5$ до $N = 75$, при $\kappa = 0,1$ до $N = 58$, при $\kappa = 0,05$ до $N = 60$) указаны граничные значения N , между которыми объемы выборок n_1 и n_2 не меняются. Например, при $\kappa = 1,0$ для всех N , $34 \leq N \leq 40$, $n_1 = 2$, $n_2 = 8$. Остальные данные соответствуют некоторому набору значений N . Отметим, что при вычислениях иногда наблюдались уменьшения n_2 при росте N . Например, для $\kappa = 0,5$, $N = 66$ параметры у рекомендуемого в таблице IX плана $n_1 = 3$, $n_2 = 12$, а для $N = 67$ параметры плана $n_1 = 4$, $n_2 = 11$.

В заключение проиллюстрируем изложенную выше теорию. Предположим, что необходимо провести испытания 30 объектов. Ограничения на время проведения испытаний таковы, что возможны лишь двухэтапные испытания. Качество объектов неизвестно ($\kappa = 1$). Для $N = 30$ и $\kappa = 1$ в таблице VIII находим значения $n_1 = 3$, $n_2 = 30 - 3 = 27$ — чисел объектов, которые должны испытываться на первом и втором этапах. Если бы было известно, что число объектов, которые могут дать критические отказы, не более 10% ($\kappa = 0,1$), то параметры оптимального плана были бы $n_1 = n_2 = 15$. Испытания проходили бы более равномерно. При $\kappa = 1$ значение предельного числа возможных критических

отказов $\bar{f} = 3,8$. Если это число считается недопустимо большим, то его можно уменьшить, увеличив число этапов испытаний. Для $N = 30$ и $k = 1$ в таблице IX находим параметры оптимального плана трехэтапных испытаний $n_1 = 2$, $n_2 = 6$, $n_3 = 22$. Здесь $\bar{f} \approx 2$, т. е. предельное число критических отказов существенно уменьшается. Таким образом, увеличивая в полтора раза время испытаний (с двух этапов до трех), мы почти вдвое уменьшаем среднее значение предельного числа возможных критических отказов.

§ 5.4. Оценка объема совокупности

Двухступенчатые планы. Статистические методы в задачах контроля качества, как правило, ориентированы на определение числа дефектных объектов в совокупности при контроле по альтернативному признаку или на определение чисел объектов в некоторых классах при контроле по качественному признаку. В биологии, в социологических исследованиях, при оценке состава и количества популяций рыб и животных возникают задачи, в которых неизвестным параметром является объем совокупности, исследуемой выборочными методами, а иногда и число классов совокупности. В этом параграфе мы ограничиваемся простейшими из указанных задач, которые наиболее близки к задачам контроля качества.

Рассмотрим задачу оценки неизвестного объема N совокупности $\mathfrak{F}$ методом двухступенчатой выборки. Из $\mathfrak{F}$ отбирают n_1 объектов, которые маркируют метками и возвращают обратно в $\mathfrak{F}$. Таким образом, в результате маркировки объектов первой выборки совокупность $\mathfrak{F}$ состоит из двух классов объектов: n_1 объектов имеют метки (первый класс), $N - n_1$ объектов меток не имеют (второй класс). После этого из $\mathfrak{F}$ берется простая случайная выборка без возвращения объема n_2 , среди объектов которой n_{12} имеют метки, $n_{12} = 0, \dots, \min(n_1, n_2)$. Случайная величина n_{12} имеет гипергеометрическое распределение

$$P\{n_{12} = d\} = h_{N, n_1}^{n_2, d}, \quad (4.1)$$

у которого N — неизвестный параметр. Эта задача является предельно упрощенным вариантом широкого

класса задач оценки числа рыб в замкнутых водоемах. Наиболее ограничительным является предположение о случайном (равновероятном) характере отлова рыб в выборку. Если в популяции рыбы имеют разные вероятности отлова, то фактические результаты могут существенно отличаться от тех, которые дает изложенная выше математическая модель двухступенчатой выборки, изучаемая в работах Бейли [1], Чепмана [1] и Виттеса [1].

Оценка максимального правдоподобия $\hat{N}_1$ находится из соотношений

$$h_{\hat{N}_1-1, n_1}^{n_2, n_{12}} \leq h_{\hat{N}_1, n_1}^{n_2, n_{12}}, \quad h_{\hat{N}_1+1, n_1}^{n_2, n_{12}} \leq h_{\hat{N}_1, n_1}^{n_2, n_{12}}, \quad (4.2)$$

при выполнении которых $h_{\hat{N}_1, n_1}^{n_2, n_{12}} = \max_N h_{N, n_1}^{n_2, n_{12}}$. Так же, как и в § 1.3, находим, что неравенства (4.2) приводят к неравенствам

$$\left[\frac{n_1 n_2}{n_{12}} \right] \leq \hat{N}_1 \leq \left[\frac{n_1 n_2}{n_{12}} \right] + 1. \quad (4.3)$$

Если $n_1 + n_2 \leq N$, то с положительной вероятностью $n_{12} = 0$, что приводит к оценке $\hat{N}_1 = \infty$. Следовательно, оценка $\hat{N}_1$ является смещенной, и математическое ожидание ее равно бесконечности. Для устранения этого неприятного свойства оценки $\hat{N}_1$ Чепман [1] предложил использовать оценку

$$\hat{N} = \frac{(n_1 + 1)(n_2 + 1)}{n_{12} + 1} - 1. \quad (4.4)$$

В работе Виттеса [1] было показано, что оценка (4.4) при $n_1 + n_2 \geq N$ является несмещенной. Для доказательства этого факта воспользуемся следующим легко проверяемым тождеством

$$\frac{1}{d+1} h_{N, d}^{n_2, d} = \frac{N+1}{(n+1)(D+1)} h_{N+1, D+1}^{n_2+1, d+1}. \quad (4.5)$$

Используя (4.5), находим, что

$$\begin{aligned} E \left\{ \frac{1}{n_{12} + 1} \right\} &= \sum_{d=\max(n_1+n_2-N, 0)}^{\min(n_1, n_2)} \frac{1}{d+1} h_{N, n_1}^{n_2, d} = \\ &= \frac{N+1}{(n_1+1)(n_2+1)} \sum_{d+1=\max(n_1+n_2-N, 0)+1}^{\min(n_1, n_2)} h_{N+1, n_1+1}^{n_2+1, d+1}. \end{aligned} \quad (4.6)$$

Заметим теперь, что если $n_1 + n_2 - N \geq 0$, то $\max(n_1 + n_2 - N, 0) + 1 = \max(n_1 + n_2 - N + 1, 0)$, поэтому сумма в правой части (4.6) проводится по всем возможным значениям $d + 1$ и, следовательно, равна единице. Если же $n_1 + n_2 < N$, то до единицы не хватает одного слагаемого $h_{N+1, n_1+1}^{n_2+1, 0}$. Таким образом, имеем

$$\begin{aligned} E \left\{ \frac{1}{n_{12} + 1} \right\} &= \\ &= \begin{cases} \frac{N + 1}{(n_1 + 1)(n_2 + 1)}, & n_1 + n_2 \geq N, \\ \frac{N + 1}{(n_1 + 1)(n_2 + 1)} (1 - h_{N+1, n_1+1}^{n_2+1, 0}), & n_1 + n_2 < N. \end{cases} \end{aligned} \quad (4.7)$$

Точно так же можно получить и более общую формулу:

$$\begin{aligned} E \left\{ \frac{1}{(n_{12} + k)^{(k)}} \right\} &= \\ &= \begin{cases} \frac{(N + k)^{(k)}}{(n_1 + k)^{(k)}(n_2 + k)^{(k)}}, & n_1 + n_2 \geq N, \\ \frac{(N + k)^{(k)}}{(n_1 + k)^{(k)}(n_2 + k)^{(k)}} (1 - H_{N+k, n_1+k}^{n_2+k, k-1}), & n_1 + n_2 < N. \end{cases} \end{aligned} \quad (4.8)$$

Из (4.7) находим, что при $n_1 + n_2 \geq N$,

$$E\hat{N} = (n_1 + 1)(n_2 + 1) E \left\{ \frac{1}{n_{12} + 1} \right\} - 1 = N, \quad (4.9)$$

т. е. оценка (4.4) несмещенная. Ее смещение при $n_1 + n_2 < N$ равно

$$b = -(N + 1) h_{N+1, n_1+1}^{n_2+1, 0} = -\frac{(N - n_1)! (N - n_2)!}{N! (N - n_1 - n_2 - 1)!}. \quad (4.10)$$

Следующая таблица, взятая из работы Виттеса [1], иллюстрирует величину смещения b для совокупности объема $N = 100$.

Т а б л и ц а 5.4.1

Величина смещения b оценки $\hat{N}$ (4.4)
для $N = 100$

$n_1 \backslash n_2$	1	5	10	20
1	97,02	89,30	80,10	63,20
5	89,30	69,263	49,619	23,948
10	80,10	49,619	26,438	6,658
20	63,20	23,948	6,658	0,396
30	48,30	10,449	1,375	0,015
40	35,40	3,990	0,218	0,000
50	24,50	1,266	0,024	0,000
60	15,60	0,306	0,001	0,000
70	8,70	0,047	0,000	0,000
80	3,80	0,003	0,000	0,000
90	0,90	0,000	0,000	0,000

Если $n_1 + n_2 \geq N$, то несмещенной оценкой для дисперсии $V(\hat{N})$ оценки $\hat{N}$ (4.4) является

$$\hat{\sigma} = \frac{(n_1 + 1)(n_2 + 1)(n_1 - n_{12})}{(n_{12} + 1)^2(n_{12} + 2)}. \quad (4.11)$$

Действительно, оценка (4.11) представима в виде

$$\hat{\sigma} = (\hat{N})^2 + 3\hat{N} + 2 - \frac{(n_1 + 2)^{(2)}(n_2 + 2)^{(2)}}{(n_{12} + 2)^{(2)}}. \quad (4.12)$$

Из (4.12) находим

$$\begin{aligned} E\hat{\sigma} &= E(\hat{N})^2 + 3N + 2 - (N + 1)(N + 2) = \\ &= E(\hat{N})^2 - N^2 = V(\hat{N}). \end{aligned}$$

Последовательные планы. Рассмотрим последовательные планы, в которых объекты извлекаются из $\mathfrak{P}$ один за другим. При каждом извлечении вероятность того, что объект будет взят из совокупности $\mathfrak{P}$, одна и та же и равна $1/N$. На извлеченном объекте ставится метка, после чего он снова возвращается в $\mathfrak{P}$. Если на извлеченном объекте уже есть метка, то новая метка не ставится. В начале выборочного обследования все N объектов $\mathfrak{P}$ не имеют меток. Можно определить различные правила прекращения выборочного обследования.

Например, можно заранее указать, что общее число извлечений объектов равно A или что выбор и маркировка объектов из $\mathfrak{F}$ продолжаются до тех пор, пока не будет зарегистрировано B повторно извлекаемых объектов, т. е. имеющих метки и т. п. По результатам выборочного обследования требуется дать оценку для объема N совокупности $\mathfrak{F}$. Такая задача рассматривалась в работах Гудмана [1], Реньи [1], Дарлинг и Роббинса [1], Самюэля [1], [2]. В этих работах исследовались различные свойства оценок N для определенного набора правил остановки выбора объектов. В работе Самюэля [1] указаны следующие правила остановки выбора.

П л а н Π_A . Задается целое число $A > 0$. Общее число извлечений $n_A = A$.

П л а н Π_B . Задается целое число B . Выборочное обследование прекращается, когда число извлекаемых объектов с метками станет равным B .

П л а н Π_C . Задается число $C > 0$. Выбор прекращается, когда впервые отношение числа извлеченных объектов, имеющих метки, к числу извлеченных объектов без меток превысит C . Пусть $k = k(n)$ — число извлеченных объектов, не имеющих меток, а $d = d(n)$ — число извлеченных объектов, имеющих метки, когда общее число извлечений равно n ($n = k + d$). В этих обозначениях момент остановки n_C по плану Π_C равен

$$n_C = \inf [n: d(n) \geq Ck(n)].$$

П л а н Π_D . Выбирается число D , $-\infty < D < \infty$. Момент прекращения отбора объектов считается равным

$$n_D = \inf \{n: d(n) \geq \max(1, k(n) \ln k(n) + k(n) D)\}.$$

П л а н Π_E . Пусть D_l — последовательность чисел такая, что $\lim_{l \rightarrow \infty} D_l = \infty$. Момент прекращения отбора объектов равен

$$n_E = \inf [n: d(n) \geq \max(1, k(n) \ln k(n) + k(n) D_{k(n)})].$$

Оценки объема совокупности $\mathfrak{F}$ при этих правилах остановки исследовались в цитированных выше работах. В работе Дарлинг и Роббинса [1], в частности, было показано, что для любого числа α , $0 < \alpha < 1$, можно найти такое D , что при использовании плана Π_D равно-

мерно по N вероятность того, что в момент остановки общее число обнаруженных объектов без меток $k(n_D) = N$, не меньше $1 - \alpha$. При $N \rightarrow \infty$ для указанных правил остановки были получены приводимые ниже утверждения (библиографические указания и некоторые результаты см. Реньи [1], Самюэль [1]).

Пусть P_N обозначает вероятности событий, связанных с выборочным обследованием $\mathfrak{F}$ объема N , а λ_N , a_N , b_N , c_N — последовательности чисел. Для плана Π_A и $n_A = n$ справедливы следующие утверждения:

(1) Если $\lambda_N \rightarrow 0$, то при $n = \sqrt{N\lambda_N}$

$$P_N \{d(n) = 0\} \rightarrow 1 \quad (N \rightarrow \infty).$$

(2) Если $\lambda_N \rightarrow \lambda$, $0 < \lambda < \infty$, то при $n = \sqrt{2N\lambda_N}$

$$P_N \{d(n) = k\} \rightarrow \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, \dots, \quad (N \rightarrow \infty).$$

(3) Если $n = a_N N$, где $b_N \cdot N^{-1/2} < a_N < \ln N - c_N$, $b_N \rightarrow \infty$, $c_N \rightarrow \infty$, то

$$P_N \left\{ \frac{k(n) - E k(n)}{\sqrt{N k(n)}} \leq x \right\} \rightarrow \Phi(x), \quad (N \rightarrow \infty).$$

(4) Если $n = N \ln N + N a_N$, где $a_N \rightarrow a$, $-\infty < a < +\infty$, то

$$P_N \{N - k(n) = m\} \rightarrow \frac{\lambda^m}{m!} e^{-\lambda},$$

$$m = 0, 1, \dots, \lambda = e^{-a}, \quad (N \rightarrow \infty).$$

(5) Если $n = N \ln N + N a_N$, где $a_N \rightarrow \infty$, то при $N \rightarrow \infty$.

$$P_N \{k(n) = N\} \rightarrow 1.$$

Мы не приводим здесь доказательства этих утверждений, отсылая читателя к работам Реньи [1] и Самюэля [1]. Самюэль [2] исследовал при $N \rightarrow \infty$ асимптотические свойства оценок максимального правдоподобия для планов типов Π_A , Π_B , Π_C , Π_D , Π_E . Наша цель другая. Покажем, что можно построить теорию статистического оценивания, которая во многом аналогична

рассмотренной выше теории для планов первого вхождения при классификации по альтернативному признаку.

Процессу последовательного случайного отбора и маркировки объектов без меток можно сопоставить следующую схему блужданий по точкам (k, d) , $k, d = 0, 1, 2, \dots$ целочисленной решетки $\mathbb{N}^2$. Блуждание начинается из точки $(0, 0)$. При извлечении объекта без меток сдвигаемся вправо из точки (k, d) в $(k+1, d)$,

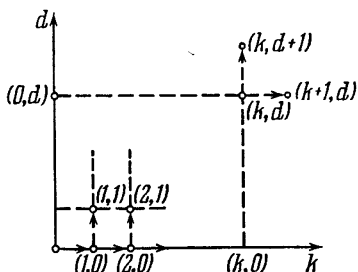


Рис. 5.4.1. Процесс последовательного отбора и маркировки объектов.

при извлечении объекта с меткой — вверх из точки (k, d) в $(k, d+1)$ (рис. 5.4.1). Так как в начальный момент все объекты в $\mathfrak{F}$ не маркированы, то первый извлеченный объект не имеет метки. Поэтому за точкой $(0, 0)$ следует $(1, 0)$. На втором шаге с вероятностью $(N-1)/N$ (N — неизвестно!) отбирается объект без метки, а с вероятностью $1/N$ с меткой. Этим исходам контроля

соответствуют переходы из $(1, 0)$ в $(2, 0)$ или $(1, 1)$. Если в результате отбора n объектов всего было обнаружено k объектов без меток и $d = n - k$ с метками, то такому состоянию контроля соответствует точка (k, d) . В результате маркировки к моменту отбора $(n+1)$ -го объекта в контролируемой совокупности k объектов имеют метки. Следовательно, вероятность перехода из (k, d) в $(k+1, d)$ равна $(N-k)/N$, а вероятность перехода в точку $(k, d+1)$ равна k/N . Как только k станет равным N , начнется детерминированное движение вверх. Интуиция подсказывает, что при больших значениях N на начальном этапе контроля будет наблюдаться преимущественное движение вправо, нарушаемое редкими скачками вверх. На заключительном этапе, когда почти все объекты уже имеют метки, будет наблюдаться преимущественное движение вверх со все более редкими скачками вправо. Теперь графически легко представить введенные выше планы Π_A , Π_B и т. д. Определим класс планов первого вхождения, для процессов маркировки задавая на решетке $\mathbb{N}^2$ точек (k, d) множество Γ точек

остановки. Для плана первого вхождения процесс контроля прекращается, как только траектория процесса контроля, образованная в результате переходов по точкам $\mathfrak{M}^2$, впервые достигнет границы Γ . В Γ включаются только достижимые точки. На рис. 5.4.2 изображены точки границы Γ для планов P_A (прямая $k+d=A$), P_B (прямая $d=B$) и граница плана первого вхождения общего вида (ломаная (в)). При соответствующем подборе функций $d=d(k)$ можно получить границы планов P_C , P_D , P_E , которые, таким образом, также являются частными примерами планов первого вхождения.

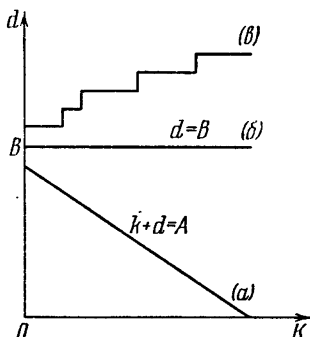


Рис. 5.4.2. Границы остановки процесса отбора и маркировки а) при плане P_A , б) при плане P_B , в) при плане общего вида.

Определение 5.4.1. План первого вхождения для процесса отбора с маркировкой называется **замкнутым**, если при любом N вероятность выхода на точки границы Γ равна единице.

Обозначим $d_\Gamma = d_\Gamma(k) = \max \{d: (k, d) \in \Gamma\}$, $d'_\Gamma(k) = d_\Gamma - 1$, если все точки $(k, d_\Gamma - j) \in \Gamma$, $j = 0, 1, \dots, l$.

Лемма 5.4.1. Пусть для всех $k = 0, 1, \dots$ значения $d_\Gamma(k) < \infty$ и $d_\Gamma(k) \geq d'_\Gamma(k-1) - 1$. Тогда план первого вхождения с границей Γ является замкнутым.

Доказательство. При любом конечном N возможны скачки только по точкам (k, d) , у которых значения $k \leq N$. В условиях леммы траектории не могут уходить вверх за точки $d_\Gamma(k)$, что и соответствует замкнутости.

Пусть результаты контроля n отобранных объектов изображаются траекторией s_0^n , последовательно проходящей через точки (k_i, d_i) , $i = 0, \dots, n$, где $k_0 = d_0 = 0$, $k_1 = 1, d_1 = 0$, $k_i \leq k_{i+1}$, $d_i \leq d_{i+1}$. Вероятность перехода из (k_i, d_i) в (k_{i+1}, d_{i+1}) равна $\left(\frac{k_i}{N}\right)^{d_{i+1}-d_i} \cdot \left(\frac{N-k_i}{N}\right)^{k_{i+1}-k_i}$, $i = 1, \dots, n$. Поэтому вероятность появления траектории s_0^n

равна

$$P_N\{s_0^n\} = \prod_{i=1}^{n-1} \left(\frac{k_i}{N}\right)^{d_{i+1}-d_i} \left(\frac{N-k_i}{N}\right)^{k_{i+1}-k_i} = \\ = \frac{N^{(k_n)}}{N^{k_n+d_n}} \prod_{i=1}^{n-1} k_i^{d_{i+1}-d_i}. \quad (4.13)$$

Если (k, d) — достижимая точка при плане первого вхождения Π_Γ , то вероятность того, что траектория достигнет (k, d) , равна в соответствии с (4.13)

$$p_\Gamma(k, d) = \frac{N^{(k)}}{N^{k+d}} S_\Gamma(k, d). \quad (4.14)$$

Здесь

$$S_\Gamma(k, d) = \sum_{(k_i, d_i)} \prod_{i=1}^{n-1} k_i^{d_{i+1}-d_i}, \quad (4.15)$$

и суммирование ведется по всем траекториям, соединяющим $(0, 0)$ с (k, d) и не «задевающим» граничных точек.

Лемма 5.4.2. Для плана первого вхождения при процессе маркировки достаточной статистикой являются координаты точки остановки.

Доказательство. Заметим, что в (4.13)

$\prod_{i=1}^n k_i^{d_{i+1}-d_i}$ не зависит от неизвестного параметра N . Следовательно, из (4.13), где $k_n = k$, $d_n = d$, находим, что вероятность любой траектории s_0^{k+d} , соединяющей $(0, 0)$ с $(k, d) \in \Gamma$, представима в виде произведения двух сомножителей. Первый сомножитель зависит от неизвестного параметра N и координат (k, d) — точки остановки, а второй является функцией координат самой траектории. Таким образом, утверждение леммы является следствием критерия факторизации (см. § 2.1).

Функция $S_\Gamma(k, d)$, определяемая (4.15), играет важную роль в построении оценок для неизвестного объема N совокупности $\mathfrak{F}$. Непосредственное вычисление $S_\Gamma(k, d)$ по формуле (4.15) сложно. Однако легко получить про-

стые рекуррентные соотношения. По формуле полных вероятностей находим

$$p_{\Gamma}(k, d) = \frac{k}{N} p_{\Gamma}(k, d-1) \delta_{\Gamma}(k, d-1) + \\ + \frac{N-k+1}{N} p_{\Gamma}(k-1, d) \delta_{\Gamma}(k-1, d), \quad (4.16)$$

где $\delta_{\Gamma}(k, d) = 0$, если $(k, d) \notin \Gamma$, $\delta_{\Gamma}(k, d) = 1$, если $(k, d) \in \Gamma$. Подставляя в (4.16) выражения для $p_{\Gamma}(k, d)$, найденные по формуле (4.14), получим

$$S_{\Gamma}(k, d) = k S_{\Gamma}(k, d-1) \delta_{\Gamma}(k, d-1) + \\ + S_{\Gamma}(k-1, d) \delta_{\Gamma}(k-1, d). \quad (4.17)$$

Рекуррентные соотношения (4.17) дают возможность быстро вычислить необходимые значения $S_{\Gamma}(k, d)$. Для планов Π_A, Π_B значения $S_{\Gamma}(k, d)$ легко выражаются через числа Стирлинга второго рода.

Параллельным сдвигом вверх границы Γ на единицу назовем множество $\Gamma_1 = \{(k, d+1) : (k, d) \in \Gamma\}$. Границе Γ_1 соответствует новый план первого вхождения.

Теорема 5.4.1. Если параллельный сдвиг вверх на единицу границы Γ плана первого вхождения Π_{Γ} дает замкнутый план Π_{Γ_1} , то статистика

$$\hat{N} = \frac{S_{\Gamma_1}(k, d+1)}{S_{\Gamma}(k, d)}, \quad (k, d) \in \Gamma, \quad (4.18)$$

является несмещенной оценкой N .

Доказательство. Используя (4.14) и тот факт, что $(k, d) \in \Gamma$ эквивалентно $(k, d+1) \in \Gamma_1$, получаем

$$E\hat{N} = \sum_{(k, d) \in \Gamma} \frac{S_{\Gamma_1}(k, d+1)}{S_{\Gamma}(k, d)} p_{\Gamma}(k, d) = \\ = \sum_{(k, d) \in \Gamma} S_{\Gamma_1}(k, d+1) \frac{N^{(k)}}{N^{k+d}} = N \sum_{(k, d+1) \in \Gamma_1} p_{\Gamma_1}(k, d+1) = N.$$

Проиллюстрируем эту простую, но полезную теорему следующим примером. Пусть контроль проводится по плану Π_B , $B = 1$, т. е. контроль прекращается при первом выборе объекта с меткой. Если объект с меткой был обнаружен при извлечении n -го объекта, то точка остановки

$(n-1, 1)$. Легко подсчитать, что здесь $p_{\Gamma}(n-1, 1) = \frac{N^{(n-1)}}{N^n} (n-1)$. Для плана Π_{Γ} прекращение контроля наступает в момент обнаружения второго объекта с меткой. При этом вероятность того, что $(n+1)$ -й объект будет вторым объектом с меткой, равна $p_{\Gamma_1}(n-1, 2) = \frac{N^{(n-1)}}{N^n} (n-1) [1 + 2 + \dots + n-1]$. Таким образом, $S_{\Gamma}(n-1, 1) = n-1$, $S_{\Gamma_1}(n-1, 2) = (n-1)^2 n/2$. В соответствии с (4.18) несмещенной оценкой $\hat{N}$ при плане Π_B , $B=1$, является $\hat{N} = (n-1)n/2$. Для плана Π_B , $B=2$, находим, что в соответствии с (4.18) несмещенной оценкой N является $\hat{N} = \frac{(n-1)}{12} [3n^2 - n - 2]$, $n > 1$.

Рассмотрим вопрос о существовании и единственности несмещенных оценок N . В условие теоремы 5.4.1 входит требование замкнутости плана Π_{Γ} , полученного параллельным сдвигом вверх на единицу границы Γ . План Π_{Γ} назовем **неограниченным**, если $\sup\{(k+d): (k,d) \in \Gamma\} = \infty$. В противном случае план Π_{Γ} называется **ограниченным**. План Π_A , $A < \infty$, является ограниченным, а план Π_B , $B > 0$ — неограниченным. Следующая элементарная лемма дает свойство границы Γ , необходимое для выполнения несмещенности.

Лемма 5.4.3. *Для того чтобы оценка $\hat{N}$, определяемая по формуле (4.18), была несмещенной, необходимо, чтобы план Π_{Γ} был неограниченным.*

Доказательство. Из доказательства несмещенности оценки (4.18) следует необходимость замкнутости плана Π_{Γ} . Пусть план Π_{Γ} — ограниченный. Тогда на прямой $L = \{(k, 0): k = 0, 1, \dots\}$ имеется одна достижимая точка $(k', 0) \in \Gamma$, $k' > 0$. Случай $k' = 1$ дает тривиальный план. При переходе к плану Π_{Γ} с параллельно сдвинутой вверх границей точка $(k', 0)$ становится проходимой. Правее $(k', 0)$ других точек на границе Γ_1 нет, так как для $(k, d) \in \Gamma_1$, $d \geq 1$. Следовательно, при достаточно большом $N > k'$, $p_{\Gamma_1}(k', 0) > 0$ и

$\sum_{(k, d+1) \in \Gamma_1} p_{\Gamma_1}(k, d+1) \leq 1 - p_{\Gamma_1}(k', 0) < 1$. Таким образом, оценка (4.18) смещенная, что и требовалось доказать.

Так же как и при выборочном контроле $\mathfrak{P}$ по альтернативному признаку, вопрос о единственности несмещенных оценок является следствием полноты семейства вероятностных мер $\mathcal{P}_\Gamma(N) = p_\Gamma\{(k, d), (k, d) \in \Gamma\}$, $N = 1, 2, \dots$. Если семейство $\mathcal{P}_\Gamma(N)$, $N = 1, 2, \dots$, полное, то из тождества

$$E\varphi(k_\tau, d_\tau) \equiv 0, \quad (4.19)$$

справедливого при любом N , следует, что $\varphi(k_\tau, d_\tau) \equiv 0$. Здесь (k_τ, d_τ) — координаты точки выхода траектории на границу Γ .

Определение 5.4.2. Граница Γ плана первого вхождения Π_Γ называется *однозначной*, если при каждом k существует не более одной точки $(k, d_k) \in \Gamma$.

Теорема 5.4.2. Для того, чтобы у замкнутого плана Π_Γ семейство вероятностных мер $\mathcal{P}_\Gamma(N)$, $N = 1, 2, \dots$, было полным, необходимо и достаточно, чтобы граница Γ была однозначной.

Доказательство. Пусть Π_Γ — замкнутый план с однозначной границей Γ , а $\varphi(k, d)$ удовлетворяет (4.19). Так как $p_\Gamma(1, d_1) = 1$ при $N = 1$, то из условия $0 = E\varphi(k_\tau, d_\tau) = \varphi(1, d_1)p_\Gamma(1, d_1)$ находим $\varphi(1, d_1) = 0$. Далее доказательство проводим по индукции. Пусть $\varphi(k, d_k) = 0$, $k \leq K_0$. При $N = K_0 + 1$ вероятность $p_\Gamma(K_0 + 1, d_{K_0+1}) \neq 0$, $p_\Gamma(K_0 + l, d_{K_0+l}) = 0$, $l = 2, 3, \dots$. Таким образом, из (4.19) имеем $\sum_k \varphi(k, d)p_\Gamma(k, d) = \varphi(K_0 + 1, d_{K_0+1})p_\Gamma(K_0 + 1, d_{K_0+1}) = 0$. Следовательно, $\varphi(K_0 + 1, d_{K_0+1}) = 0$. Итак, из (4.19) вытекает, что $\varphi(k, d) \equiv 0$, $(k, d) \in \Gamma$. Достаточность однозначности границы Γ для полноты семейства вероятностных мер $\mathcal{P}_\Gamma(N)$, $N = 1, 2, \dots$, доказана.

Если граница Γ не является однозначной, то найдется такое K_1 , что при некоторых $d' \neq d''$ имеем $(K_1, d') \in \Gamma$, $(K_1, d'') \in \Gamma$. Построим функцию $\varphi(k, d) \neq 0$, $(k, d) \in \Gamma$, которая удовлетворяет (4.19). Для всех граничных точек (k, d) , $k < K_1$, положим $\varphi(k, d) \equiv 0$. В точках (K_1, d') , (K_1, d'') имеем $p_\Gamma(K_1, d') = p' \neq 0$, $p_\Gamma(K_1, d'') = p'' \neq 0$. Положим $\varphi(K_1, d') = p''$, $\varphi(K_1, d'') = -p'$. В других граничных точках (K_1, d) , если таковые имеются, считаем $\varphi(K_1, d) = 0$. Таким образом, для $N \leq K_1$ условие (4.19)

выполняется. Далее построение $\varphi(k, d)$ проводим следующим образом. Пусть d_k^* — максимальное значение координат d точек $(k, d) \in \Gamma$. Если $(k, d) \in \Gamma$, $k > K_1$, $d < d_k^*$, то полагаем $\varphi(k, d) = 0$. Значения в точках (k, d_k^*) получаем последовательно, переходя от k к значению $k+1$, $k > K_1$. При $k = K_1 + 1$, полагая $N = K_1 + 1$, находим, что (4.19) можно переписать в следующем виде:

$$p'' p_{\Gamma}(K_1, d') - p' p_{\Gamma}(K_1, d'') + \\ + \varphi(K_1 + 1, d_{K_1+1}^*) p_{\Gamma}(K_1 + 1, d_{K_1+1}^*) = 0. \quad (4.20)$$

Значение $\varphi(K_1 + 1, d_{K_1+1}^*)$ находим, решая уравнение (4.20), в котором $p_{\Gamma}(K_1, d')$, $p_{\Gamma}(K_1, d'')$ и $p_{\Gamma}(K_1 + 1, d_{K_1+1}^*)$ — вероятности выхода в точки (K_1, d') , (K_1, d'') и $(K_1 + 1, d_{K_1+1}^*)$ при $N = K_1 + 1$. Точно так же из (4.19) находим при $N = K_1 + 2$

$$p'' p_{\Gamma}(K_1, d') - p' p_{\Gamma}(K_1, d'') + \\ + \varphi(K_1 + 1, d_{K_1+1}^*) p_{\Gamma}(K_1 + 1, d_{K_1+1}^*) + \\ + \varphi(K_1 + 2, d_{K_1+2}^*) p_{\Gamma}(K_1 + 2, d_{K_1+2}^*) = 0. \quad (4.21)$$

Значение $\varphi(K_1 + 2, d_{K_1+2}^*)$ находим, решая уравнение (4.21). Далее поступаем аналогичным образом. Построенная таким способом функция $\varphi_{\Gamma}(k, d)$ удовлетворяет (4.19) при любом N и в то же время не тождественно равна нулю. Таким образом, необходимость однозначности границы Γ доказана.

Из полученных выше результатов легко вытекает следующее утверждение.

Следствие 5.4.1. Если у плана первого вхождения граница Γ является однозначной и выполнены условия теоремы 5.4.1, то несмещенная оценка $\hat{N}$, определяемая по формуле (4.18), является единственной.

Заметим, что у планов P_B , P_C , P_D , P_E семейства $\mathcal{P}_{\Gamma}(\hat{N})$ — полные, $N = 1, 2, \dots$, а единственные несмещенные оценки N определяются формулой (4.18).

Можно было бы продолжить изложение этой небольшой теории, рассмотрев задачи построения оценок для дисперсий $\hat{N}$ и т. д. Однако мы ограничимся изложенными выше фактами. В уже отмечавшейся статье Гуд-

мана [1] помимо последовательного отбора объектов по планам Π_A , Π_B рассматриваются последовательные выборки, содержащие, вообще говоря, различные числа объектов. Полученные выше результаты можно обобщить и на такие схемы отбора. Отметим также возможность изучения асимптотических свойств оценок (4.18), когда $N \rightarrow \infty$ и может соответствующим образом меняться граница плана. В работе Гудмана [1], в частности, показано, что для планов Π_B случайная величина $(k_\tau + d_\tau)^2/N$ при $N \rightarrow \infty$ имеет χ^2 -распределение с $2B$ степенями свободы. Здесь (k_τ, d_τ) — координаты точки остановки, $d_\tau = B$. Таблица 5.4.2, взятая из работы Гудмана [1], иллюстрирует тот факт, что для плана Π_B , $B = 1$, процентные точки функции распределения $\frac{(k_\tau + d_\tau)^2}{N}$, $d_\tau = 1$, достаточно хорошо соответствуют χ^2 -распределению с двумя степенями свободы.

Таблица 5.4.2

Процентные точки χ^2 -распределения с двумя степенями свободы и распределения квадрата числа наблюдений, деленного на N , при последовательном отборе объектов с маркировкой до первого обнаружения объекта с меткой, $N = 365$

Уровень в %	1	2	5	10	20	30	50
$(k_\tau + d_\tau)^2/N$	0,02 0,02	0,04 0,04	0,10 0,10	0,21 0,22	0,45 0,46	0,71 0,70	1,39 1,33
Уровень в %	70	80	90	95	98	99	
$(k_\tau + d_\tau)^2/N$	2,41 2,30	3,22 3,17	4,61 4,38	5,99 5,80	7,82 7,41	9,21 8,59	

Например, медиана χ^2 -распределения с двумя степенями свободы равна 1.39, а для статистики $(k_\tau + d_\tau)^2/N$ ее значение равно 1.33. Асимптотические свойства распределений некоторых оценок N для частных планов Π_A , Π_B , Π_C , Π_D , Π_E исследовались в работах Самюэля [1], [2].

В заключение этого параграфа отметим еще две возможные области приложений. В работе Гудмана [1] планы P_B используются для оценки числа дней в древних календарях по датам, регистрируемым на основе археологических исследований. В частности, для этих целей Гудманом и была рассчитана таблица 5.4.2, в которой $N = 365$ соответствует числу дней в году. Другая и, по-видимому, более содержательная область приложений связана с определением числа классов локусов хромосом. Одна из таких задач рассматривалась в работе Левонтина и Праута [1]. В связи с этой работой отметим, что определение числа различных классов исходов при равной вероятности каждого исхода в схеме независимых полиномиальных испытаний приводит к задачам, рассмотренным во второй части этого параграфа.

ГЛАВА 6

ВЫБОРКИ ИЗ КОНЕЧНЫХ СОВОКУПНОСТЕЙ

§ 6.1. Оценка среднего конечной совокупности

Случайная выборка без возвращения. Выше мы рассматривали разнообразные статистические и вероятностные задачи для совокупностей $\mathfrak{P}$, объекты которых характеризовались по альтернативному или по качественному признаку. Предположим теперь, что объекты $O_i \in \mathfrak{P}$ характеризуются скалярными величинами X_i , $i = 1, \dots, N$. На основе планов выборочного обследования требуется оценить некоторые показатели совокупности $\mathfrak{P}$, определяемые набором значений $(X_1, \dots, X_N)$. Часто в качестве такого показателя выбирается общее среднее $\bar{X} = N^{-1} \sum_{1 \leq i \leq N} X_i$. Если, например, $\mathfrak{P}$ есть со-

вокупность (партия) электрических ламп, то $N \cdot \bar{X}$ будет соответствовать общей продолжительности работы ламп, когда они будут находиться в одинаковых условиях работы. Необходимо, конечно, исходить из того, что выборочная проверка ламп на срок свечения проводится в условиях, идентичных режиму работы. Можно привести и другие примеры, когда требуется оценить общее среднее $\bar{X}$.

К сожалению, перейдя к обследованию совокупностей объектов, характеризующихся скалярными величинами, мы вынуждены ограничиться простыми планами отбора объектов в выборку. Будем исходить из результатов, получаемых на основе случайной выборки без возвращения фиксированного объема. Напоминаем, что при этом вероятность отбора для всех подмножеств объектов, составляющих выборку, одинакова. Мы не рассматриваем таких планов контроля, при которых продолжение

отбора зависит от наблюдаемых значений X_i . Типичным примером планов такого типа мог бы служить двухступенчатый план, при котором на первом этапе берется простая выборка без возвращения объема n_1 , измеряется среднее $\bar{X}^{(1)}$ для n_1 объектов в выборке, при $\bar{X}^{(1)} > X_B$ или $\bar{X}^{(1)} < X_H$ контроль прекращается, при $X_H \leq \bar{X}^{(1)} \leq X_B$ берется вторая простая выборка без возвращения (X_H, X_B — заданные заранее числа). Для такого двухступенчатого плана вероятность извлечения объекта в выборку для контроля зависит от величин $(X_1, \dots, X_N)$. Разработка статистических методов для подобных планов контроля конечных совокупностей является важной задачей. При случайной выборке без возвращения объема n выборка $x = (x_1, \dots, x_n)$ соответствует объектам $O_{i_1}(x_1 = X_{i_1}), \dots, O_{i_n}(x_n = X_{i_n})$. Все наборы номеров $(i_1, \dots, i_n)$ объектов, извлеченных в выборку, равновероятны и вероятность их извлечения равна $\binom{N}{n}^{-1}$. Под-

черкнем, что случайность наблюдений $x_1, \dots, x_n$ определяется случайным характером отбора объектов в выборку, тогда как значения $(X_1, \dots, X_N)$ не предполагаются случайными.

Основными параметрами конечной совокупности $\mathfrak{P}$ обычно считают *общее среднее* и *общую дисперсию*

$$\bar{X} = N^{-1} \sum_{1 \leq i \leq N} X_i, \quad \bar{V} = N^{-1} \sum_{1 \leq i \leq N} (X_i - \bar{X})^2. \quad (1.1)$$

Это, по-видимому, объясняется неявным предположением, что $X_i, i = 1, \dots, N$, могут рассматриваться как реализация N взаимно независимых нормально распределенных случайных величин или как реализация случайных величин, распределение которых достаточно полно характеризуется первыми двумя моментами. На основе результатов x выборки без возвращения легко получить несмещенные оценки для $\bar{X}$ и $\bar{V}$. При i -м извлечении с равной вероятностью может быть извлечен любой из объектов, поэтому $P\{x_i = X_k\} = N^{-1}, k = 1, \dots, N$. Аналогично, учитывая, что число возможных исходов на i -м и j -м извлечении равно $N(N-1)$ и все

исходы равновероятны, получаем $P\{x_i = X_k, x_j = X_l\} = [N(N-1)]^{-1}$. В общем случае имеем

$$P\{x_{i_r} = X_{k_r}, r = 1, \dots, l\} = (N^{(l)})^{-1}, \quad l \leq n. \quad (1.2)$$

Пусть среднее по выборке равно

$$\bar{x}_n = n^{-1} \sum_{1 \leq i \leq n} x_i. \quad (1.3)$$

Из (1.2) находим, что математическое ожидание выборочного среднего равно

$$E\bar{x}_n = n^{-1} \sum_{1 \leq i \leq n} E x_i = n^{-1} \sum_{1 \leq i \leq n} \sum_{1 \leq k \leq N} X_k P\{x_i = X_k\} = \bar{X}.$$

Таким образом, $\bar{x}_n$ является несмещенной оценкой для общего среднего $\bar{X}$. Используя (1.2), можно найти выражение для дисперсии оценки $\bar{x}_n$:

$$\begin{aligned} V(\bar{x}_n) &= E(\bar{x}_n)^2 - (\bar{X})^2 = \frac{1}{n^2} \left(\sum_{i=1}^n E x_i^2 + \sum_{\substack{i, j=1 \\ i \neq j}}^n E x_i x_j \right) - (\bar{X})^2 = \\ &= \frac{1}{n^2} \left[\sum_{i=1}^n \sum_{k=1}^N X_k^2 P\{x_i = X_k\} + \right. \\ &\quad \left. + \sum_{\substack{i, j=1 \\ i \neq j}}^n \sum_{\substack{k_1, k_2=1 \\ k_1 \neq k_2}}^N X_{k_1} X_{k_2} P\{x_i = X_{k_1}, x_j = X_{k_2}\} \right] - (\bar{X})^2 = \\ &= \frac{1}{nN} \sum_{i=1}^N X_i^2 + \frac{n-1}{n} \frac{1}{N(N-1)} \sum_{\substack{k_1, k_2=1 \\ k_1 \neq k_2}}^N X_{k_1} X_{k_2} - (\bar{X})^2 = \\ &= \frac{1}{n} \frac{N-n}{N-1} \bar{V}, \quad (N \geq 2), \quad (1.4) \end{aligned}$$

где $\bar{V}$ — общая дисперсия, определяемая (1.1). Несмещенной оценкой для общей дисперсии является

$$\begin{aligned} \bar{v}_n &= \frac{1}{N} \frac{N-1}{n-1} \sum_{1 \leq i \leq n} (x_i - \bar{x}_n)^2 = \\ &= \frac{N-1}{N} \frac{n}{n-1} \left\{ \frac{1}{n} \sum_{1 \leq i \leq n} x_i^2 - (\bar{x}_n)^2 \right\}. \quad (1.5) \end{aligned}$$

Действительно, используя промежуточные формулы в (1.4), находим, что

$$E\bar{v}_n = \frac{N-1}{N} \frac{n}{n-1} \left\{ \left(\frac{1}{N} \sum_{1 \leq i \leq N} X_i^2 \right) - \frac{1}{nN} \sum_{1 \leq i \leq N} X_i^2 - \right. \\ \left. - \frac{n-1}{nN(N-1)} \sum_{\substack{k, l=1 \\ k \neq l}}^N X_k X_l \right\} = \bar{V}.$$

Для получения несмещенных оценок общих моментов старших порядков вида $N^{-1} \sum_{1 \leq i \leq N} X_i^m$ может оказаться полезным подход, изложение которого дано в книге Уилкса ([1], стр. 232). Метод состоит в следующем. Пусть $g(x_1, \dots, x_r)$ — функция r переменных, $r \leq n$. Образуем симметрическую функцию n переменных

$$q(x_1, \dots, x_n) = \frac{1}{n^{(r)}} \sum' g(x_{i_1}, \dots, x_{i_r}), \quad (1.6)$$

где в $\sum'$ индексы $i_1, \dots, i_r$ образованы всевозможными наборами по r индексов из n . Общее число таких наборов равно $n^{(r)}$. Если в (1.6) считать, что значения $x_1, \dots, x_n$ есть значения случайной выборки без возвращения, то

$$Eq(x_1, \dots, x_n) = \frac{1}{N^{(r)}} \sum'' g(X_{i_1}, \dots, X_{i_r}), \quad (1.7)$$

где сумма $\sum''$ берется по всем наборам $(i_1, \dots, i_r)$ индексов $i_k \neq i_l$, $k, l = 1, \dots, r$, $i_k, i_r = 1, \dots, N$. Таким образом, справедливо следующее утверждение.

Лемма 6.1.1. *Симметрическая функция результатов выборки, задаваемая (1.6), есть несмещенная оценка для симметрической функции*

$$Q(X_1, \dots, X_N) = \frac{1}{N^{(r)}} \sum'' g(X_{i_1}, \dots, X_{i_r}). \quad (1.8)$$

Из леммы 6.1.1, например, следует, что $q(x_1, \dots, x_n) = [n(n-1)]^{-1} \sum_{i \neq j} x_i x_j$ является несмещенной оценкой для $Q(X_1, \dots, X_N) = [N(N-1)]^{-1} \sum_{i \neq j, i, j=1, \dots, N} X_i X_j$. Если искомую оцениваемую симметрическую функцию от

$X_1, \dots, X_N$ записать в виде линейной комбинации стандартных симметрических функций $\sum_{1 \leq i \leq N} X_i$, $\sum_{i \neq j} X_i X_j$, $\sum_{i \neq j \neq l} X_i X_j X_l$, ..., то по лемме 6.1.1 можно легко найти выражения для ее несмещенной оценки.

Выше мы ограничились проверкой несмещенности оценки (1.5) для общей дисперсии, т. е. мы как бы «угадали» оценку, а затем проверили ее несмещенность. Чтобы применить лемму 6.1.1, надо расписать $\bar{V}$ в виде суммы стандартных симметрических функций типа (1.8). Так как общая дисперсия равна

$$\begin{aligned} \bar{V} &= \frac{1}{N} \sum_{1 \leq i \leq N} (X_i - \bar{X})^2 = \frac{1}{N} \sum_{1 \leq i \leq N} X_i^2 - (\bar{X})^2 = \\ &= \frac{N-1}{N} \left[\frac{1}{N} \sum_{1 \leq i \leq N} X_i^2 - \frac{1}{N(N-1)} \sum_{i \neq j, i, j=1, \dots, N} X_i X_j \right], \quad (1.9) \end{aligned}$$

то несмещенная оценка для $\bar{V}$ имеет вид

$$\begin{aligned} \frac{N-1}{N} \left[\frac{1}{n} \sum_{1 \leq i \leq n} x_i^2 - \frac{1}{n(n-1)} \sum_{i, j=1, \dots, n, i \neq j} x_i x_j \right] &= \\ &= \frac{N-1}{N} \frac{n}{n-1} \left[\frac{1}{n} \sum_{1 \leq i \leq n} x_i^2 - (\bar{x}_n)^2 \right] = \bar{v}_n, \end{aligned}$$

т. е. совпадает с оценкой (1.5).

Асимптотическая нормальность выборочного среднего. Очень часто считают, что выборочное среднее $\bar{x}_n$ является нормально распределенной случайной величиной. Однако это не всегда так. Можно доказать общую предельную теорему, дающую критерий асимптотической нормальности $\bar{x}_n$ при одновременном неограниченном росте N и n . Наиболее интересный результат был получен Гаеком [1], который обобщил доказанные ранее Эрдемем и Реньи [1] предельные теоремы.

Исследование вопроса о предельном поведении выборочного среднего существенно упрощается, если воспользоваться свойствами «близости» простой случайной выборки без возвращения и специальным образом

заданной биномиальной выборки. Напомним определение биномиальной выборки, введенной выше в главе 2. При биномиальной выборке каждый объект может быть включен в выборку независимо от других с одной и той же вероятностью v . Положим $v = n/N$. Тогда объем биномиальной выборки есть случайная величина n' , для среднего и дисперсии которой имеем

$$En' = n, \quad V(n') = n(1 - (n/N)). \quad (1.10)$$

При этом вероятность отбора объектов $O_{i_1}, \dots, O_{i_k}$ равна

$$v^k (1 - v)^{N-k} = \binom{N}{k} v^k (1 - v)^{N-k} \binom{N}{k}^{-1}, \quad (1.11)$$

Соотношение (1.11) соответствует тому, что биномиальную выборку можно получить в два этапа. Сначала получаем реализацию биномиально распределенной случайной величины n' , $P\{n' = k\} = \binom{N}{k} v^k (1 - v)^{N-k}$, а

затем из $\mathfrak{B}$ берем случайную выборку без возвращения (сл. в. б. в.) объема k . Обозначим $s(n)$, $s'(k)$ соответственно множества номеров объектов, отобранных при сл. в. б. в. объема n и биномиальной выборке объема $n' = k$. Отбор объектов из $\mathfrak{B}$ организуем следующим образом. Если $n' = n$, то из $\mathfrak{B}$ в соответствии со сл. в. б. в. отбираем n объектов и считаем, что биномиальная выборка и сл. в. б. в. дают совпадающие результаты $s(n) = s'(n)$. Если окажется $n' = k > n$, то в соответствии со сл. в. б. в. объема k из $\mathfrak{B}$ отбираем k объектов $s'(k)$, а затем из $s'(k)$ в соответствии со сл. в. б. в. отбираем n объектов $s(n) \subset s'(k)$. Если же окажется, что $n' = k < n$, то сначала отбираем в соответствии со сл. в. б. в. n объектов $s(n)$, а затем из $s(n)$ отбираем в соответствии со сл. в. б. в. объема k объекты, образующие $s'(k) \subset s(n)$. В итоге мы получаем два набора объектов $(s(n), s'(n'))$, причем при $n' \leq n$, $s'(n') \subseteq s(n)$, а при $n' > n$ $s'(n') \supset s(n)$. Введем вспомогательные случайные величины:

$$\eta(n) = \sum_{j \in s(n)} (X_j - \bar{X}), \quad \eta'(n') = \sum_{j \in s'(n')} (X_j - \bar{X}). \quad (1.12)$$

Легко проверить, что $E\eta(n) = E\eta'(n') = 0$. Из определения $(s(n), s'(n'))$ и (1.12) следует, что

$$\eta(n) - \eta'(n') = \begin{cases} \sum_{j \in s(n) \setminus s'(n')} (X_j - \bar{X}), & n' < n, \\ 0, & n' = n, \\ - \sum_{j \in s'(n') \setminus s(n)} (X_j - \bar{X}), & n' > n. \end{cases} \quad (1.13)$$

Лемма 6.1.2. (Гаек [1]). Для отношения дисперсий случайных величин $\eta(n) - \eta'(n')$, $\eta'(n')$ справедливо неравенство

$$[V(\eta(n) - \eta'(n'))/V\eta'(n')] \leq \sqrt{\frac{1}{n} + \frac{1}{N-n}}. \quad (1.14)$$

Доказательство. Оценим значение условного математического ожидания квадрата разности $\eta(n) - \eta'(n')$ при условии, что $n' = k$. Учитывая, что $|\eta(n) - \eta'(k)|$ при таком условии соответствует сл. в. б. в. объема $|n - k|$, находим

$$\begin{aligned} E\{(\eta(n) - \eta'(n'))^2 | n' = k\} &= E\left\{\sum_{i \in s(|n-k|)} (X_i - \bar{X})\right\}^2 = \\ &= \frac{|n-k|}{N} \sum_{1 \leq i \leq N} (X_i - \bar{X})^2 + \\ &+ \frac{|n-k|(|n-k|-1)}{N(N-1)} \sum_{i \neq j, i, j=1, \dots, N} (X_i - \bar{X})(X_j - \bar{X}) \leq \\ &\leq \frac{|n-k|(N-|n-k|)}{N(N-1)} \sum_{1 \leq i \leq N} (X_i - \bar{X})^2 \leq |n-k|\bar{V}. \end{aligned} \quad (1.15)$$

Из (1.15) и (1.10) получаем

$$\begin{aligned} V(\eta(n) - \eta'(n')) &= E(\eta(n) - \eta'(n'))^2 = \\ &= E\{E[(\eta(n) - \eta'(n'))^2 | n' = k]\} \leq E\{|n' - n|\bar{V}\} \leq \\ &\leq \sqrt{E(n' - n)^2 \bar{V}^2} = \sqrt{n(1 - n/N)\bar{V}}. \end{aligned} \quad (1.16)$$

Пусть γ_i — взаимно независимые случайные величины, равные 1 или 0 в зависимости от того, включается объект O_i в биномиальную выборку или нет,

$$P\{\gamma_i = 1\} = n/N, \quad P\{\gamma_i = 0\} = 1 - (n/N).$$

Тогда $\eta'(n')$ можно представить в виде

$$\eta'(n') = \sum_{1 \leq j \leq N} \xi_j, \quad \xi_j = (X_j - \bar{X}) \gamma_j, \quad (1.17)$$

откуда очевидно

$$V\eta'(n') = \sum_{1 \leq j \leq N} (X_j - \bar{X})^2 V(\gamma_j) = n(1 - (n/N)) \bar{V}. \quad (1.18)$$

Утверждение леммы непосредственно следует из (1.16) и (1.18).

Рассмотрим последовательность совокупностей $\{\mathfrak{F}_i\}$, $i = 1, 2, \dots$ с объемами N_i , из которых берутся согласно описанному выше правилу выборки $(s(n_i), s'(n'_i))$, $n_i \rightarrow \infty$, $N_i - n_i \rightarrow \infty$, $i \rightarrow \infty$. Из (1.14) в этих условиях имеем

$$\lim_{i \rightarrow \infty} \frac{V(\eta(n_i) - \eta'(n'_i))}{V(\eta'(n'_i))} = 0. \quad (1.19)$$

Из (1.19) получаем

Следствие 6.1.1. *Предельные при $i \rightarrow \infty$ дисперсии и распределения случайных величин*

$$A_i + B_i \eta(n_i) \quad \text{и} \quad A_i + B_i \eta'(n'_i), \quad B_i^{-2} = V(\eta'(n'_i))$$

существуют и совпадают при одинаковых условиях, накладываемых на значения параметров объектов $(X_{i,1}, \dots, X_{i,N_i})$ совокупностей $\mathfrak{F}_i$ $i = 1, 2, \dots$

Доказательство. Для любого x и $\varepsilon > 0$ имеем

$$\{A_i + B_i \eta'(n'_i) \leq x\} \subseteq \{A_i + B_i \eta(n_i) \leq x + \varepsilon\} \cup \{B_i(\eta(n_i) - \eta'(n'_i)) \geq \varepsilon\}.$$

Пусть $F(x) = \lim_{i \rightarrow \infty} P\{A_i + B_i \eta(n_i) \leq x\}$. Тогда из соотношения (1.19), используя неравенство Чебышева, имеем

$$\begin{aligned} \overline{\lim}_{i \rightarrow \infty} P\{A_i + B_i \eta'(n'_i) \leq x\} &\leq F(x + \varepsilon) + \\ &+ \overline{\lim}_{i \rightarrow \infty} \frac{V(\eta(n_i) - \eta'(n'_i))}{\varepsilon^2 V(\eta'(n'_i))} \cdot B_i^2 V(\eta'(n'_i)) = F(x + \varepsilon). \end{aligned}$$

Так как $F(x)$ непрерывна справа, то

$$\overline{\lim}_{i \rightarrow \infty} P\{A_i + B_i \eta'(n'_i) \leq x\} \leq F(x).$$

Аналогичные рассуждения дают неравенство

$$F(x-0) \leq \lim_{i \rightarrow \infty} P \{A_i + B_i \eta'(n'_i) \leq x\},$$

откуда и вытекает утверждение следствия.

Используя следствие 6.1.1, можно найти условия, при которых будут асимптотически нормально распределены случайные величины $A_i + B_i \eta'(n'_i)$. При этих же условиях асимптотически нормально будут распределены случайные величины $A_i + B_i \eta(n_i)$. Изучение предельного распределения для случайных величин $\eta'(n')$ существенно проще, чем изучение предельного распределения для $\eta(n)$, так как $\eta'(n')$ в соответствии с (1.17) есть сумма взаимно независимых одинаково распределенных случайных величин ξ_j , $j = 1, \dots, N$.

Для конечной совокупности объектов $O_{i,j} \in \mathfrak{F}_i$, характеризующихся скалярными величинами $X_{i,j}$, $j = 1, \dots, N_i$, $i = 1, 2, \dots$, определим подмножества номеров объектов

$$\mathfrak{N}(i, \tau) = \{j: |X_{i,j} - \bar{X}_i| > \tau \sqrt{V(\eta_i(n_i))}\}, \quad (1.20)$$

где $\tau > 0$, а $\eta_i(n_i)$ определяется из (1.12), $n = n_i$ — объем выборки из $\mathfrak{F}_i$. Выборочное среднее получается делением на n_i случайной величины

$$\xi_i = \sum_{j \in s(n_i)} X_{i,j}, \quad (1.21)$$

где $s(n_i)$ — набор номеров объектов, включенных в выборку. Так как $E \xi_i = n_i \bar{X}_i$, то согласно (1.12) случайная величина $\eta_i(n_i) = \xi_i - E \xi_i$, а потому $V(\eta_i(n_i)) = V \xi_i$. Следующая теорема дает критерий асимптотической нормальности последовательности $\{\xi_i\}$. Она была получена Гаеком [1], который заметил, что приведенные ранее в работе Эрдеша и Реньи [1] достаточные условия являются также необходимыми.

Теорема 6.1.1. Случайные величины ξ_i имеют асимптотически нормальное распределение с параметрами $(E \xi_i, V \xi_i)$ при $i \rightarrow \infty$ тогда и только тогда, когда для любого $\tau > 0$

$$\lim_{i \rightarrow \infty} \frac{\sum_{j \in \mathfrak{N}(i, \tau)} (X_{i,j} - \bar{X}_i)^2}{\sum_{1 \leq j \leq N_i} (X_{i,j} - \bar{X}_i)^2} = 0. \quad (1.22)$$

Доказательство. Так как центрирование ξ_i дает $\eta_i(n_i)$, то из следствия 6.1.1 вытекает, что вместо случайных величин $\eta_i(n_i) = \xi_i - E\xi_i$ можно рассматривать случайные величины $\eta'_i(n'_i)$. Предположим сначала, что случайные величины $\zeta_{i,j} = (X_{i,j} - \bar{X}_i) \gamma_{i,j}$, аналогичные введенным в (1.17), являются бесконечно малыми, т. е.

$$\lim_{i \rightarrow \infty} \max_{1 \leq j \leq N_i} \frac{V((X_{i,j} - \bar{X}_i) \gamma_{i,j})}{\sum_{1 \leq j \leq N_i} V((X_{i,j} - \bar{X}_i) \gamma_{i,j})} = 0. \quad (1.23)$$

Из (1.18) и определения $\gamma_{i,j}$ находим, что (1.23) эквивалентно соотношению

$$\lim_{i \rightarrow \infty} \frac{\max_{1 \leq j \leq N_i} (X_{i,j} - \bar{X}_i)^2}{\sum_{1 \leq j \leq N_i} (X_{i,j} - \bar{X}_i)^2} = 0. \quad (1.24)$$

Соотношение (1.24) вытекает из (1.22). При выполнении предположения (1.23) $\eta'_i(n'_i)$ является суммой бесконечно малых взаимно независимых случайных величин. В теории суммирования случайных величин хорошо известно (см. (2.1.11), Гнеденко [1], Феллер [2]), что в условиях бесконечной малости слагаемых $\zeta_{i,j}$ критерий асимптотической нормальности имеет вид:

$$\lim_{i \rightarrow \infty} \frac{1}{V\eta'_i(n'_i)} \times \times \sum_{1 \leq j \leq N_i} \int_{|x| > \tau \sqrt{V\eta'_i(n'_i)}} x^2 dP\{\zeta_{i,j} - E\zeta_{i,j} \leq x\} \rightarrow 0, \quad (1.25)$$

для любого $\tau > 0$. В нашем случае величины $\zeta_{i,j} - E\zeta_{i,j}$ принимают значения $(X_{i,j} - \bar{X}_i)(1 - n_i/N_i)$ и $[-(X_{i,j} - \bar{X}_i) \times \times n_i/N_i]$ с вероятностями, равными n_i/N_i и $(1 - n_i/N_i)$ соответственно. Следовательно, интегралы в сумме (1.25)

будут равны

$$\int_{|x| > \tau \sqrt{v\eta'_i(n')}} x^2 dP \{ \xi_{i,j} - E\xi_{i,j} \leq x \} =$$

$$= \frac{n_i}{N_i} \left(1 - \frac{n_i}{N_i}\right)^2 (X_{i,j} - \bar{X}_i)^2 I'_{i,j,\tau} +$$

$$+ \left(1 - \frac{n_i}{N_i}\right) \left(\frac{n_i}{N_i}\right)^2 (X_{i,j} - \bar{X}_i)^2 I''_{i,j,\tau}, \quad (1.26)$$

где $I'_{i,j,\tau}$ и $I''_{i,j,\tau}$ равны 1 или 0 в зависимости от того, принадлежит или нет индекс j множествам:

$$\left. \begin{aligned} c'(i, \tau) &= \left\{ j: |X_{i,j} - \bar{X}_i| \left(1 - \frac{n_i}{N_i}\right) > \right. \\ &\quad \left. > \tau \sqrt{\frac{n_i}{N_i} \left(1 - \frac{n_i}{N_i}\right) \sum_{1 \leq l \leq N_i} (X_{i,l} - \bar{X}_i)^2} \right\}, \\ c''(i, \tau) &= \left\{ j: |X_{i,j} - \bar{X}_i| \frac{n_i}{N_i} > \right. \\ &\quad \left. > \tau \sqrt{\frac{n_i}{N_i} \left(1 - \frac{n_i}{N_i}\right) \sum_{1 \leq l \leq N_i} (X_{i,l} - \bar{X}_i)^2} \right\}. \end{aligned} \right\} \quad (1.27)$$

Так как $\xi_i = n_i \bar{x}_{i,n}$, то из (1.4) получаем

$$V(\xi_i) = \frac{n_i(N_i - n_i)}{N_i(N_i - 1)} \sum_{1 \leq j \leq N_i} (X_{i,j} - \bar{X}_i)^2. \quad (1.28)$$

Из (1.27) и (1.28) следует, что

$$\left. \begin{aligned} c'(i, \tau) &= \mathfrak{N} \left(i, \tau \sqrt{\frac{N_i(N_i - 1)}{(N_i - n_i)^2}} \right), \\ c''(i, \tau) &= \mathfrak{N} \left(i, \tau \sqrt{\frac{N_i(N_i - 1)}{n_i^2}} \right). \end{aligned} \right\} \quad (1.29)$$

В обозначениях (1.27) и (1.28) критерий (1.25) имеет вид

$$\lim_{i \rightarrow \infty} \left\{ \frac{\frac{n_i}{N_i} \left(1 - \frac{n_i}{N_i}\right)^2 \sum' + \left(1 - \frac{n_i}{N_i}\right) \left(\frac{n_i}{N_i}\right)^2 \sum''}{\frac{n_i}{N_i} \left(1 - \frac{n_i}{N_i}\right) \sum_{1 \leq j \leq N_i} (X_{i,j} - \bar{X}_i)^2} \right\}^2 =$$

$$= \lim_{i \rightarrow \infty} \left\{ \frac{\left(1 - \frac{n_i}{N_i}\right) \sum' + \frac{n_i}{N_i} \sum''}{\sum_{1 \leq j \leq N_i} (X_{i,j} - \bar{X}_i)^2} \right\} = 0, \quad (1.30)$$

где $\sum' = \sum_{j \in c'(i, \tau)} (X_{i,j} - \bar{X}_i)^2$, $\sum'' = \sum_{j \in c''(i, \tau)} (X_{i,j} - \bar{X}_i)^2$.

Соотношение (1.30) эквивалентно (1.22). Действительно, если (1.22) выполнено, то из (1.29) имеем $c'(i, \tau) \cup c''(i, \tau) \subseteq \mathfrak{N}(i, \tau)$. Следовательно, дробь под знаком предела в (1.30) не превосходит дроби в (1.22). Таким образом, из (1.22) вытекает (1.30). Обратное утверждение доказываем следующим способом. При любом i либо $1 - (n_i/N_i) \geq 1/2$ либо $(n_i/N_i) \geq 1/2$. Если $1 - (n_i/N_i) \geq 1/2$, то $c'(i, \tau/2) \supseteq \mathfrak{N}(i, \tau)$, если же $(n_i/N_i) \geq 1/2$, то $c''(i, \tau/2) \supseteq \mathfrak{N}(i, \tau)$. Следовательно, имеем

$$\left(1 - \frac{n_i}{N_i}\right) \sum_{j \in c'(i, \tau/2)} (X_{i,j} - \bar{X}_i)^2 + \frac{n_i}{N_i} \sum_{j \in c''(i, \tau/2)} (X_{i,j} - \bar{X}_i)^2 \geq$$

$$\geq 2^{-1} \sum_{j \in \mathfrak{N}(i, \tau)} (X_{i,j} - \bar{X}_i)^2.$$

Таким образом, из (1.30) следует (1.22). Итак, если выполнено (1.24), то соотношение (1.22) является критерием нормальности.

Если последовательность n_i/N_i строго отделена от 0 и 1, т. е. если существует $\varepsilon > 0$ такое, что

$$\varepsilon < \lim_{i \rightarrow \infty} (n_i/N_i) \leq \overline{\lim}_{i \rightarrow \infty} (n_i/N_i) < 1 - \varepsilon, \quad (1.31)$$

то (1.22) и (1.24) эквивалентны. Это следует из того, что для всех достаточно больших i введенные множества $\mathfrak{N}(i, \tau)$ — пустые. Очевидно и обратное, если (1.24) не выполняется, то $\mathfrak{N}(i, \varepsilon)$ не пусто и (1.22) не выполняется.

Итак, нам осталось показать, что при нарушении (1.24) случайные величины $\eta'_i(n'_i)$ не могут иметь нормального распределения с параметрами $(E\eta'_i(n'), V(\eta'_i(n')))$. Для определенности будем считать, что $n_i \leq (N/2)$ и что объекты занумерованы в порядке убывания уклонений от общего среднего их скалярных характеристик, т. е.

$$|X_{i,1} - \bar{X}_i| \geq |X_{i,2} - \bar{X}_i| \geq \dots \geq |X_{i,N_i} - \bar{X}_i|.$$

Если (1.24) нарушается, то найдется такое $\delta \neq 0$, что

$$\lim_{i \rightarrow \infty} \frac{X_{i,1} - \bar{X}_i}{\sqrt{\sum_{1 \leq j \leq N_i} (X_{i,j} - \bar{X}_i)^2}} = \delta. \quad (1.32)$$

Пусть верхний предел в (1.32) достигается по некоторой подпоследовательности индексов i . Из этой подпоследовательности мы выберем такую подпоследовательность $\{i_k\}$, что существует

$$\lim_{k \rightarrow \infty} n_{i_k} / N_{i_k} = m \leq 1/2. \quad (1.33)$$

Далее подпоследовательность $\{i_k\}$ будем обозначать $\{l\}$. Из (1.32) и (1.33) выводим, что для $\xi_{l,1} = (X_{l,1} - \bar{X}_l) / \sqrt{V(\eta'_l(n'_l))}$

$$\left. \begin{aligned} P \left\{ \frac{\xi_{l,1}}{\sqrt{V(\eta'_l(n'_l))}} = \right. \\ \left. = \frac{X_{l,1} - \bar{X}_l}{\sqrt{\frac{n_l}{N_l} \left(1 - \frac{n_l}{N_l}\right) \sum_{1 \leq j \leq N_l} (X_{l,j} - \bar{X}_l)^2}} \right\} = \frac{n_l}{N_l}, \\ P \left\{ \frac{\xi_{l,1}}{\sqrt{V(\eta'_l(n'_l))}} = 0 \right\} = 1 - (n_l / N_l). \end{aligned} \right\} \quad (1.34)$$

Из (1.33) при $m > 0$ следует, что $\xi_{l,1} / \sqrt{V(\eta'_l(n'_l))}$ в пределе сходится к случайной величине, принимающей два значения 0 и $\delta / \sqrt{m(1-m)}$ с вероятностями $1-m$ и m

соответственно. Так как выполнено тождество

$$\frac{\eta'_i(n'_i)}{\sqrt{V(\eta'_i(n'_i))}} = \frac{\xi_{i,1}}{\sqrt{V(\eta'_i(n'_i))}} + \frac{\sum_{2 \leq j \leq N_i} \xi_{i,j}}{\sqrt{V(\eta'_i(n'_i))}}, \quad (1.35)$$

где слагаемые взаимно независимы, то $\eta'_i(n'_i)/\sqrt{V(\eta'_i(n'_i))}$ не может иметь нормального распределения $N(0, 1)$. В противном случае мы имели бы разложение нормально распределенной случайной величины на сумму двух взаимно независимых случайных величин, одна из которых принимает только два значения и, следовательно, не нормально распределена. Известно, однако (Крамер [1]), что нормально распределенная случайная величина может быть представлена только в виде суммы двух взаимно независимых нормально распределенных случайных величин. Таким образом, при $m > 0$ условие (1.32) приводит к противоречию. Если же $m = 0$, то предельное распределение $\xi_{i,1}/\sqrt{V\eta'_i(n'_i)}$ — вырожденное (вероятность значения 0 равна в пределе 1). Однако из (1.32) имеем $\lim_{i \rightarrow \infty} V(\xi_{i,1}/\sqrt{V(\eta'_i(n'_i))}) = \delta^2 \neq 0$. Поэтому в соответствии с (1.35) предельное распределение $\eta'_i(n'_i)/\sqrt{V(\eta'_i(n'_i))}$ будет иметь дисперсию, не большую $1 - \delta^2$, что также противоречит исходному предположению. Теорема доказана.

Отметим важный частный случай применения теоремы 6.1.1. Пусть $\bar{x}_{i,n}$ — выборочное среднее (1.3), а $\bar{X}_i$ и $\bar{V}_i$ — общее среднее и дисперсия совокупности $\mathfrak{F}_i$. Предельное распределение $(\bar{x}_{i,n} - \bar{X}_i)/\sqrt{\frac{N-n}{n(N-1)\bar{V}_i}}$ будет асимптотически нормальным, если значения $X_{i,j}$ равномерно ограничены по i, j , а

$$\sum_{1 \leq j \leq N_i} (X_{i,j} - \bar{X}_i)^2 \rightarrow \infty, \quad i \rightarrow \infty.$$

Асимптотическая нормальность $\bar{x}_n$ не всегда имеет место. В статье Гаека [1] приведены условия, при которых сумма выборочных значений (1.21) сходится к некоторому классу безгранично делимых законов. Из более

поздних работ отметим результат Розена [1], который получил решение общей задачи о предельном поведении накопленных сумм

$$S_i(t) = \sum_{1 \leq j \leq [N_i t]} X_{i,j},$$

образованных реализацией случайных выборок $(X_{i,1}, X_{i,2}, \dots, X_{i,[N_i t]})$ без возвращения объемов $[N_i t]$, $0 < t < 1$, из конечных совокупностей $\mathfrak{F}_i$ с объемами N_i , $i = 1, 2, \dots$. При выполнении условий $\sum_{1 \leq j \leq N_i} X_{i,j} = 0$, $\sum_{1 \leq j \leq N_i} X_{i,j}^2 \rightarrow \sigma^2$, $0 < \sigma^2 < \infty$, $\max_{1 \leq j \leq N_i} |X_{i,j}| \rightarrow 0$, $i \rightarrow \infty$,

последовательность сумм $S_i(t)$, рассматриваемая как последовательность случайных процессов с «временным» параметром t , слабо сходится к условному винеровскому процессу. Близкие задачи исследованы в работах Сена [2] и Калленберга [1]. Впоследствии Хагберг [1] исследовал аналогичную задачу без предположения о том, что

$$\max_{1 \leq j \leq N_i} |X_{i,j}| \rightarrow 0, \quad i \rightarrow \infty.$$

Интересные результаты об асимптотической нормальности линейных статистик для некоторых схем выборок, отличных от сл. в. б. в., были найдены Колчиным и Чистяковым [1], [2].

Выборка с возвращением. При выборке с возвращением (в. в.) извлеченный для контроля объект O_i после определения X_i возвращается в совокупность $\mathfrak{F}$ и снова может быть включен в выборку при последующих извлечениях. Точнее, при каждом извлечении для каждого объекта $O_i \in \mathfrak{F}$, $i = 1, \dots, N$, вероятность извлечения равна $1/N$. Таким образом, если $(x'_1, \dots, x'_n)$ — результаты в. в., то

$$P\{x'_i = X_{i,j}, i = 1, \dots, n, j_i \in [1, N]\} = N^{-n}. \quad (1.36)$$

Легко проверить, что выборочное среднее

$$\bar{x}'_n = n^{-1} \sum_{1 \leq i \leq n} x'_i, \quad (1.37)$$

является несмещенной оценкой для общего среднего $\bar{X}$. Учитывая (1.36), находим дисперсию оценки (1.37):

$$\begin{aligned} V(\bar{x}'_n) &= E(\bar{x}'_n)^2 - (E\bar{x}'_n)^2 = \\ &= \frac{1}{n^2} \sum_{i=1}^n E(x'_i)^2 + \frac{1}{n^2} \sum_{i,j=1, i \neq j}^n E x'_i x'_j - (\bar{X})^2 = \\ &= \frac{1}{n} \left(\frac{1}{N} \sum_{i=1}^N X_i^2 \right) + \frac{n-1}{n} \left(\frac{1}{N} \sum_{1 \leq i \leq N} X_i \right)^2 - (\bar{X})^2 = \\ &= \frac{1}{n} \frac{1}{N} \sum_{1 \leq i \leq N} (X_i - \bar{X})^2 = \frac{1}{n} \bar{V}. \quad (1.38) \end{aligned}$$

Сравнивая (1.38) и (1.4), получаем, что дисперсия выборочного среднего при сл. в. б. в. меньше дисперсии выборочного среднего при в. в. Дисперсия оценки общего среднего при контроле с возвращением уменьшится, если выборочное среднее будет оцениваться только на основе неповторяющихся наблюдений, т. е. будет использоваться оценка

$$\bar{x}''_v = v_n^{-1} \sum_{1 \leq i \leq v_n} x_{j_i}, \quad j_i \in [1, n], \quad (1.39)$$

где v_n — число различных объектов, включенных в выборку при в. в. Оценку (1.39) можно интерпретировать следующим образом. Сначала реализуется случайная величина v_n — число несовпадающих значений, а затем из $\mathfrak{F}$ берется сл. в. б. в. объема v_n . Так как при каждом значении $v_n = k$ оценка $\bar{x}''_v = \bar{x}_k$ является несмещенной для $\bar{X}$, то и оценка $\bar{x}''_v$ — также несмещенная для $\bar{X}$. Ее дисперсия в соответствии с (1.4) равна

$$V(\bar{x}''_v) = E\{E[\bar{x}_k - \bar{X}]^2 | v_n = k\} = E\left(\frac{1}{v_n} \frac{N - v_n}{N - 1}\right) \bar{V}. \quad (1.40)$$

Функция $\frac{1}{x} \frac{N - x}{N - 1} = \frac{N}{N - 1} \cdot \frac{1}{x} - \frac{1}{N - 1}$ является на отрезке $[1, N]$ выпуклой вниз функцией. Поэтому, применяя неравенство Йенсена, получим

$$E \frac{1}{v_n} \frac{N - v_n}{N - 1} = \frac{N}{N - 1} E \frac{1}{v_n} - \frac{1}{N - 1} \geq \frac{1}{E v_n} \frac{N - E v_n}{N - 1}. \quad (1.41)$$

Из соотношений (1.40) и (1.41) следует

Лемма 6.1.3. Если в. в. имеет среднее число неповторяющихся объектов $E v_n = m$, то оценка общего среднего на основе сл. в. б. в. будет иметь меньшую дисперсию, чем оценка (1.39), если только ее объем не меньше m .

В работе Корвара и Серфлинга [1] получены оценки для $E(1/v_n)$. В частности, показано, что выполняется неравенство

$$\frac{1}{n} + \frac{1}{2N} + \frac{n-1}{12N^2} - \frac{1}{720N} \leq E \frac{1}{v_n} \leq \frac{1}{n} + \frac{1}{2N} + \frac{n-1}{12N^2}. \quad (1.42)$$

Для сл. в. б. в. и в. в. легко находится расстояние по вариации между вероятностными мерами, индуцированными этими способами отбора на множествах значений $(x_1, \dots, x_n)$ характеристик объектов, включенных в выборку. В соответствии с (1.2) и (1.36) вероятности получения выборки $(x_1, \dots, x_n)$ из совокупности $\mathfrak{F} \propto [X_1, \dots, X_N]$ равны: для сл. в. б. в.

$$p_1(x_1, \dots, x_n) = \begin{cases} 1/N^{(n)}, & \text{если для } i \neq j, x_i \neq x_j, \\ i, j = 1, \dots, n, x_i \in [X_1, \dots, X_N], \\ 0, & \text{если для некоторых } i, j \text{ либо} \\ x_i = x_j \text{ либо } x_i \notin [X_1, \dots, X_N], \end{cases} \quad (1.43)$$

а для в. в.

$$p_2(x_1, \dots, x_n) = \begin{cases} 1/N^n, & x_i \in [X_1, \dots, X_N], \\ & i = 1, \dots, n, \\ 0, & \text{если для некоторых } i \\ & x_i \notin [X_1, \dots, X_N]. \end{cases} \quad (1.44)$$

Расстояние по вариации $\rho(p_1, p_2)$ между двумя распределениями вероятностей p_1, p_2 , сосредоточенными на конечном или счетном множестве точек, определяется как сумма абсолютных величин разностей их вероятностей. В нашем случае имеем

$$\rho(p_1, p_2) = \sum_{(x_1, \dots, x_n)} |p_1(x_1, \dots, x_n) - p_2(x_1, \dots, x_n)|,$$

где сумма берется по всем возможным наборам по n чисел из множества $\{X_1, \dots, X_N\}$. Эта сумма разбивается на две части. В первой суммирование ($\sum_1$) ведется по таким наборам $(x_1, \dots, x_n)$, которые не содержат

повторяющихся значений x_i , во второй ($\sum_2$) суммирование ведется по остальным наборам значений x_i . Таким образом находим,

$$\begin{aligned} d(p_1, p_2) &= \\ &= \sum_1 |p_1(x_1, \dots, x_n) - p_2(x_1, \dots, x_n)| + \sum_2 p_2(x_1, \dots, x_n) = \\ &= N^{(n)}[(1/N^{(n)}) - (1/N^n)] + [N^n - N^{(n)}]N^{-n} = \\ &= 2[1 - (N^{(n)}/N^n)] = 2\left[1 - \prod_{1 \leq j \leq n-1} (1 - (j/N))\right] \leq \\ &\leq 2 \sum_{1 \leq j \leq n-1} j/N = n(n-1)/N \leq n^2/N. \quad (1.45) \end{aligned}$$

Оценка (1.45) приведена в работе Тага [1]. Из (1.45) следует, что при $n_i^2/N_i \rightarrow 0$ ($N_i \rightarrow \infty$) для последовательности выборок объемов n_i из совокупностей $\mathfrak{F}_i$ расстояние по вариации $\rho(p_1, p_2) \rightarrow 0$.

Лемма 6.1.4. Пусть $T(x_1, \dots, x_n)$ — некоторая статистика, определенная на возможных значениях выборок, а $F_i(t) = P_i\{T(x_1, \dots, x_n) \leq t\}$ — функции распределения этой статистики при выборках с возвращением ($i = 1$) и без возвращения ($i = 2$); тогда

$$|F_1(t) - F_2(t)| \leq \rho(p_1, p_2) \leq n^2/N. \quad (1.46)$$

Доказательство является прямым следствием определения $\rho(p_1, p_2)$ и того факта, что при любом t функции $F_1(t)$ и $F_2(t)$ представимы в виде сумм вероятностей $F_i(t) = \sum p_i(x_1, \dots, x_n)$, по значениям $(x_1, \dots, \dots, x_n)$, для которых $T(x_1, \dots, x_n) \leq t$.

Теми же методами, которые были использованы в теореме 6.1.1., можно исследовать асимптотические свойства распределений широкого класса статистик $T(x_1, \dots, x_n)$ для выборки с возвращением. В условиях $n^2/N \rightarrow 0$ можно показать, что распределения таких статистик для выборки без возвращения и с возвращением совпадают. Именно этот подход используется для оценки скорости сходимости функции распределения выборочного среднего к функции распределения нормального закона в работе Тага [1].

В настоящем параграфе мы ограничиваемся только задачей оценки общего среднего и рассматриваем только самые простые планы выборочного контроля. В послед-

нем параграфе этой главы будет рассмотрен вопрос о «допустимости» различных оценок общего среднего. Точные определения будут даны ниже в § 6.3. Однако интуитивно ясно, что при использовании случайной выборки с возвращением объема n оценка $\bar{x}'_n$ хуже, чем $\bar{x}''_n$, в том смысле, что ее дисперсия больше. Поэтому, если несмещенная оценка среднего выбрана, важно знать, а есть ли другие оценки $\bar{X}$, у которых дисперсия меньше? Можно сравнивать также оценки, полученные при различных, но в некотором смысле эквивалентных (например, по стоимости) планах контроля.

В заключение кратко рассмотрим возможные обобщения задач на выборки из конечных совокупностей, объекты которых характеризуются векторными величинами. Пусть $\mathfrak{P}$ — конечная совокупность объектов $O_j \in \mathfrak{P}$, $j = 1, \dots, N$. Объекту O_i сопоставляется m -мерный вектор $X_j = (X_{j,1}, \dots, X_{j,m})$. Вектор *общего среднего* $\bar{X}$, компоненты *общей ковариационной матрицы* $V = \|\bar{V}_{r,s}\|$ определяются по формулам

$$\left. \begin{aligned} \bar{X} &= N^{-1} \sum_{1 \leq j \leq N} X_j, \\ \bar{V}_{r,s} &= N^{-1} \sum_{1 \leq j \leq N} (X_{j,r} - \bar{X}_r)(X_{j,s} - \bar{X}_s) = \\ &= N^{-1} \sum_{1 \leq j \leq N} X_{j,r} X_{j,s} - \bar{X}_r \bar{X}_s, \end{aligned} \right\} \quad (1.47)$$

где $\bar{X}_i = N^{-1} \sum_{1 \leq j \leq N} X_{j,i}$, $i = r, s \in [1, m]$. Коэффициенты корреляции компонент векторов объектов совокупности $\mathfrak{P}$ равны

$$\rho_{r,s} = \bar{V}_{r,s} / \sqrt{\bar{V}_{r,r} \bar{V}_{s,s}}, \quad r, s = 1, \dots, m. \quad (1.48)$$

Для оценки общего среднего при простой случайной выборке без возвращения можно использовать выборочное среднее

$$\bar{x}_n = n^{-1} \sum_{1 \leq j \leq n} x_j, \quad (1.49)$$

где $(x_1, \dots, x_n)$ — значения векторных характеристик объектов, включенных в выборку. Оценка (1.49) — несмещенная, так как каждая ее компонента есть несмещенная оценка для общего среднего по этой компоненте.

Элементы ковариационной матрицы $V(\bar{x}_n)$ оценки (1.49) равны

$$\begin{aligned} V_{n,r,s} &= E \left(n^{-1} \sum_{1 \leq i \leq n} x_{i,r} - \bar{X}_r \right) \left(n^{-1} \sum_{1 \leq i \leq n} x_{i,s} - \bar{X}_s \right) = \\ &= n^{-2} E \sum_{i, i'=1, \dots, n} x_{i,r} x_{i',s} - \bar{X}_r \bar{X}_s = \\ &= n^{-1} E x_{1,r} x_{1,s} + (n-1) n^{-1} E x_{1,r} x_{2,s} - \bar{X}_r \bar{X}_s = \\ &= \frac{1}{nN} \sum_{1 \leq j \leq N} X_{j,r} X_{j,s} + \frac{n-1}{nN(N-1)} \sum_{\substack{j_1 \neq j_2, \\ j_1, j_2=1, \dots, N}} X_{j_1,r} X_{j_2,s} - \bar{X}_r \bar{X}_s = \\ &= \frac{N-n}{n(N-1)} \bar{V}_{r,s}. \end{aligned}$$

В матричной форме полученный результат сформулируем в виде следующей леммы.

Лемма 6.1.5. *Ковариационная матрица вектора выборочного среднего при простой случайной выборке без возвращения равна*

$$V(\bar{x}_n) = \frac{N-n}{n(N-1)} V,$$

где элементы $V = \|\bar{V}_{r,s}\|$ определяются (1.47).

Для совокупностей с векторными характеристиками можно решить аналогичные одномерному случаю задачи о предельном распределении $\bar{x}$, когда $N \rightarrow \infty$, $n \rightarrow \infty$.

§ 6.2. Непараметрические статистические задачи

Порядковые статистики выборки. Известно, что методы непараметрической статистики занимают важное место в теории математической статистики и широко используются в приложениях. В значительной мере это объясняется тем, что ограничения, накладываемые на ф. р. наблюдаемых случайных величин, минимальны. Так, например, от функции распределения случайных величин, реализации которых образуют выборку, требуется лишь только непрерывность. Однако в большинстве исследований предполагается, что наблюдаемые случайные величины взаимно независимы и одинаково распределены. Такое ограничение весьма существенно и его далеко не всегда можно считать выполненным. В связи с этим представляет интерес развитие методов

математической статистики, которые можно было бы применять для анализа совокупностей данных без обязательного предположения взаимной независимости и одинаковой распределенности измеряемых значений. В настоящем параграфе мы проиллюстрируем возможность развития таких методов применительно к анализу конечных совокупностей.

Конечно, чем больше мы знаем о статистических свойствах исследуемой совокупности объектов, тем более эффективные методы могут быть использованы для оценки ее характеристик или проверки соответствующих гипотез. Классическим примером могут служить методы анализа выборок, когда их компоненты являются реализациями нормально распределенных взаимно независимых случайных величин. В этом случае любая статистическая задача формулируется в терминах неизвестных параметров (среднего и дисперсии). Для многих таких задач найдены оптимальные статистические процедуры. Если, однако, нет уверенности в правильности исходных допущений о нормальности распределений, то пользоваться «оптимальными» для нормальных распределений процедурами надо осторожно. Действительно, существуют многочисленные примеры, показывающие неустойчивость оптимальных процедур. Интересный пример неустойчивости процедуры оценивания на основе максимума функции правдоподобия при замене логнормального распределения распределением Вейбулла приведен в статье Чепурина и Дугиной [1].

Итак, пусть $\mathfrak{F}$ — конечная совокупность объектов O_j , характеризующихся скалярными величинами X_j , $j = 1, \dots, N$. Всюду ниже предполагается, что никакие два объекта из $\mathfrak{F}$ не имеют совпадающих значений X_j , т. е. $X_{j'} \neq X_{j''}$, $j' \neq j''$. Пусть $X_{(1)} < \dots < X_{(N)}$ — **вариационный ряд**, образованный из значений X_j , $j = 1, \dots, N$. Без ограничения общности будем считать объекты из $\mathfrak{F}$ занумерованными таким образом, что O_j соответствует значению $X_{(j)}$. Пусть, далее, $x_1, \dots, x_n$ — наблюдаемые значения характеристик объектов, отобранных при простой случайной выборке без возвращения объема n , а $x_{(1)} < \dots < x_{(n)}$ — вариационный ряд, соответствующий $x_1, \dots, x_n$.

В задачах приемочного контроля при анализе результатов испытаний на надежность и т. д. может представлять интерес получение оценок для чисел $\bar{N}(k)$, $\bar{N}(k)$, $N(k_1, k_2)$ объектов, характеристики которых соответственно не более $x_{(k)}$, не менее $x_{(k)}$, заключены в интервале $[x_{(k_1)}, x_{(k_2)}]$. Осуществление события $\{x_{(k)} = X_{(l)}\}$ происходит, если в выборку включаются объект со значением $X_{(l)}$, $k-1$ объектов, номера которых не превышают $l-1$ и $n-k$ объектов, номера которых больше l . Так как все сл. в. б. в. равновероятны, то

$$P\{x_{(k)} = X_{(l)}\} = \frac{\binom{l-1}{k-1} \binom{N-l}{n-k}}{\binom{N}{n}} = \frac{k}{l} h_{N,l}^{n,k}, \quad (2.1)$$

где $h_{N,l}^{n,k}$ — вероятности гипергеометрического распределения (см. § 1.3 гл. 1). Аналогичным образом находим, что для $k_1 \leq l_1 < l_2$, $k_2 \leq l_2$, $k_1 < k_2$,

$$\begin{aligned} P\{x_{(k_1)} = X_{(l_1)}, x_{(k_2)} = X_{(l_2)}\} &= \\ &= \binom{l_1-1}{k_1-1} \binom{l_2-l_1-1}{k_2-k_1-1} \binom{N-l_2}{n-k_2} \binom{N}{n}^{-1} = \\ &= \frac{k_1}{l_1} h_{N,l_1}^{n,k_1} \cdot \frac{k_2-k_1}{l_2-l_1} h_{N-l_1,l_2-l_1}^{n-k_1,k_2-k_1}. \end{aligned} \quad (2.2)$$

Распределения вероятностей для величин $\bar{N}(k)$, $\bar{N}(k)$ и $N(k_1, k_2)$, введенных выше, легко получить из (2.1) и (2.2). Так, например,

$$P\{\bar{N}(k) \leq m\} = \sum_{k \leq l \leq m} \frac{k}{l} h_{N,l}^{n,k}, \quad 1 \leq m \leq N. \quad (2.3)$$

Если обозначим $m(\gamma) = m(\gamma, N, n, k)$ — наименьшее целое число такое, что $\sum_{k \leq l \leq m(\gamma)} \frac{k}{l} h_{N,l}^{n,k} \geq \gamma$, то из (2.3)

получим

$$P\{\bar{N}(k) \leq m(\gamma)\} \geq \gamma. \quad (2.4)$$

Таким образом, $m(\gamma)$ есть верхняя доверительная граница для $\bar{N}(k)$ с коэффициентом доверия, не меньшим γ . Из (2.1) можно найти общие выражения для факториальных моментов старших порядков для случайного номера

$l_\omega = l_\omega(k)$, при котором $x_{(k)} = X_{(l_\omega)}$, $l_\omega(k) = \underline{N}(k)$. Как и при выводе (1.3.14), имеем

$$\begin{aligned} E(l_\omega + r - 1)^{(r)} &= \sum_{k \leq l \leq N - (n - k)} (l + r - 1)^{(r)} \frac{k}{l} h_{N, l}^{n, k} = \\ &= \sum_{k \leq l \leq N - (n - k)} (l + r - 1)^{(r)} \frac{(N + r)^{(r)} (k + r - 1)^{(r)} (k + r)}{(n + r)^{(r)} (l + r - 1)^{(r)} (l + r)} h_{N + r, l + r}^{n + r, k + r} = \\ &= \frac{(N + r)^{(r)} (k + r - 1)^{(r)}}{(n + r)^{(r)}}. \quad (2.5) \end{aligned}$$

Из (2.5) легко получить выражения для среднего и дисперсии

$$E l_\omega = \frac{k(N + 1)}{n + 1}, \quad V l_\omega = \frac{k(N + 1)(N - n)(n - k + 1)}{(n + 1)^2(n + 2)}. \quad (2.6)$$

Точно так же, используя формулу (2.2), можно найти моменты $N(k_1, k_2)$.

Случайная величина $\underline{N}(k)/N$ равна доле объектов контролируемой совокупности, для которых значения характеристик не превышают $x_{(k)}$. При $N \rightarrow \infty$ для любых ограниченных n из (2.5) находим, что

$$\lim_{N \rightarrow \infty} E \left(\frac{\underline{N}(k)}{N} \right)^r = \lim_{N \rightarrow \infty} E \frac{(l_\omega + r - 1)^{(r)}}{N^r} = \frac{(k + r - 1)^{(r)}}{(n + r)^{(r)}}. \quad (2.7)$$

Правая часть (2.7) равна k -му моменту случайной величины, имеющей бэ́та-распределение (2.1.10), где $p_1 = k$, $p_2 = n - k + 1$. Из сходимости моментов (см. § 2.1) следует, что по распределению случайные величины $\underline{N}(k)/N$ сходятся при $N \rightarrow \infty$ к бэ́та-распределению. Таким образом, если y_γ есть квантиль уровня γ бэ́та-распределения, т. е.

$$\frac{\Gamma(n + 1)}{\Gamma(k)\Gamma(n - k + 1)} \int_0^{y_\gamma} x^{k-1} (1 - x)^{n-k} dx = \gamma, \quad (2.8)$$

то

$$\lim_{N \rightarrow \infty} P \left\{ \frac{\underline{N}(k)}{N} \leq y_\gamma \right\} = \gamma. \quad (2.9)$$

Так же с помощью сходимости моментов можно найти при $N \rightarrow \infty$ асимптотический вид распределений для

$N(k_i, k_{i+1})/N$, где $0 < k_1 < \dots < k_m < n$ — фиксированные целые числа.

Из формулы (2.2) следующим образом выводятся уравнения для построения γ -доверительного интервала для значения $X_{(l)}$. Для любых целых чисел $k_1, k_2, k_1 \leq l, k_1 \leq k_2 \leq n$ из (2.2) находим

$$\begin{aligned} P\{x_{(k_1)} \leq X_{(l)} \leq x_{(k_2)}\} &= \sum_{m \geq 0, r \geq 0} P\{x_{(k_1)} = X_{(l-m)}, x_{(k_2)} = X_{(l+r)}\} = \\ &= \sum_{m \geq 0, r \geq 0} \frac{k_1}{l-m} h_{N, l-m}^{n, k_1} \cdot \frac{k_2 - k_1}{r+m} h_{N-l+m, r+m}^{n-k_1, k_2-k_1}. \quad (2.10) \end{aligned}$$

Уменьшение k_1 и увеличение k_2 ведет к росту суммы в (2.10). Если $P\{x_{(1)} \leq X_{(l)} \leq x_{(n)}\} \geq \gamma$, то найдутся такие наборы чисел $\underline{k}_1 = k, \underline{k}_2 = \bar{k}$, что

$$P\{x_{(\underline{k}_1)} \leq X_{(l)} \leq x_{(\underline{k}_2)}\} \geq \gamma, \quad (2.11)$$

тогда как для всех наборов (k_1, k_2) , у которых либо $k_1 > \underline{k}, k_2 = \bar{k}$ либо $k_1 = \underline{k}, k_2 < \bar{k}$, неравенство (2.11) нарушается. Для значений γ , близких к 1, сумма (2.10) может оказаться меньше γ при любых $k_1 \geq 1, k_2 \leq n$. В этом случае нетривиальный двухсторонний γ -доверительный интервал (2.11) построить нельзя. Однако если в таких случаях либо $P\{X_{(l)} \leq x_{(n)}\} \geq \gamma$, либо $P\{x_{(1)} \leq X_{(l)}\} \geq \gamma$, то можно построить односторонние γ -доверительные интервалы. Если же ни одно из этих неравенств не выполнено, то нельзя построить нетривиальных γ -доверительных интервалов.

Гипергеометрический процесс. Приведенные выше примеры оценок характеристик обследуемой совокупности $\mathfrak{F}$ не предполагают наличия каких-либо ограничений, накладываемых на значения X_i , и в этом смысле являются прямыми аналогами непараметрических методов, применяемых тогда, когда наблюдаемые значения выборки $x_1, \dots, x_n$ предполагаются взаимно независимыми с непрерывной функцией распределения $F(x)$. В непараметрической статистике существенное место занимают работы, посвященные критериям согласия (критерии Колмогорова, Колмогорова — Смирнова и др.) см. Колмогоров [2], [3], Смирнов [1], Большев и

Смирнов [1], Гнеденко, Королюк [1]. Представляет интерес получение аналогов этих критериев для выборочного обследования конечных совокупностей в условиях, когда не делается никаких специальных предположений о вероятностных распределениях значений $X_1, \dots, X_N$, соответствующих объектам совокупности.

Конечную совокупность $\mathfrak{F}$ будем характеризовать **общей функцией «распределения»** $F_N(x) = l/N$, где $l = l(x)$ — число объектов O_i , для которых $X_i \leq x$. Основные числовые характеристики $\mathfrak{F}$ выражаются в виде функционалов от $F_N(x)$. Например, общее среднее

$$\bar{X} = N^{-1} \sum_{1 \leq i \leq N} X_i = \int_{-\infty}^{+\infty} x dF_N(x)$$

представимо в виде интеграла Стильтьеса от x по функции $F_N(x)$. Общая дисперсия V может быть представлена в виде интеграла

$$V = \int_{-\infty}^{+\infty} (x - \bar{X})^2 dF_N(x).$$

Сопоставим результатам $x_1, \dots, x_n$ случайной выборки без возвращения **эмпирическую функцию распределения** (э. ф. р.) $F_n(x) = k/n$, где $k = k(x)$ — число величин $x_i \leq x$. Основными объектами исследования будут различные показатели степени соответствия $F_n(x)$ и $F_N(x)$. Для случая взаимно независимых значений $x_1, \dots, x_n$ с ф. р. $F(x)$ Колмогоров [2] исследовал предельное поведение $D_n = \sup_x |F_n(x) - F(x)|$ и показал, что

$$\lim_{n \rightarrow \infty} P\{\sqrt{n} D_n < y\} = K(y) = \sum_{-\infty < k < +\infty} (-1)^k e^{-2k^2 y^2}. \quad (2.12)$$

Распределение статистики D_n можно исследовать и для параметрических семейств ф. р. Так, в работе Тюрина [1] изучалось предельное распределение D_n , когда

$F(x) = F_0\left(\frac{x - \mu}{\sigma}\right)$, где F_0 — известная функция, μ и σ — неизвестные параметры, оцениваемые по выборке. Отметим также обзор Винце [1].

Как и в классическом случае независимых наблюдений, мы найдем статистики, распределения которых не будут зависеть от фактических значений $X_1, \dots, X_n$. Затем будут исследованы асимптотические свойства распределений таких статистик, когда $N \rightarrow \infty$, $n \rightarrow \infty$. Часть приводимых ниже результатов была опубликована в статье Беляева и Рыковой [1]. В конце параграфа будет приведен результат, полученный Рыковой [3]. Этот результат можно использовать для обобщения критерия Колмогорова — Смирнова на случай трех выборок.

Итак, пусть из конечной совокупности $\mathfrak{P}$ взята выборка без возвращения объема n . Полученные выборочные значения соответствуют некоторым объектам из $\mathfrak{P}$, т. е. $x_{(1)} = X_{(l_1)} < x_{(2)} = X_{(l_2)} < \dots < x_{(n)} = X_{(l_n)}$, $1 \leq l_1 < l_2 < \dots < l_n \leq N$. Введем случайный процесс $\kappa(l)$, $l = 1, \dots, N$, равный числу $x_i \leq X_{(l)}$, $\kappa(0) = 0$. Заметим, что введение процесса $\kappa(l)$ аналогично переходу к равномерно распределенным на интервале $[0, 1]$ взаимно независимым случайным величинам в классическом случае независимых наблюдений. При любых фиксированных l_1 и l_2 , $l_i = 0, \dots, N$, $\kappa(l_1, l_2) = \kappa(l_2) - \kappa(l_1)$ имеет гипергеометрическое распределение, т. е.

$$P\{\kappa(l_1, l_2) = d\} = h_{N, l_1, l_2}^{n, d} \quad (2.13)$$

Соотношение (2.13) становится очевидным, если представить совокупность $\mathfrak{P}$ состоящей из объектов двух типов. К первому типу отнесем объекты $X_{(j)}$, $l_1 + 1 \leq j \leq l_2$, число которых $l_2 - l_1$, а ко второму типу — все остальные. Тогда $\kappa(l_1, l_2)$ будет равно числу объектов первого типа, попавших в выборку и, следовательно, будет выполнено (2.13). Точно так же легко показать, что совместное распределение величин $\kappa(l_i) - \kappa(l_{i-1})$ для последовательности $0 < l_{i-1} < l_i < N$, $i = 2, \dots, m$, есть сумма произведений соответствующих гипергеометрических вероятностей. Здесь имеем

$$\begin{aligned} P\{\kappa(l_i) - \kappa(l_{i-1}) = k_i, 2 \leq i \leq m\} = \\ = h_{N, l_1}^{n, k_1} h_{N-l_1, l_2-l_1}^{n-k_1, k_2} \dots h_{N-l_{m-1}, l_m-l_{m-1}}^{n-(k_1+\dots+k_{m-1}), k_m}. \end{aligned}$$

В соответствии с этим свойством процесс $\kappa(l)$ будем называть **гипергеометрическим процессом**. Из свойств гипергеометрического распределения (см. § 1.3) находим, что математическое ожидание, дисперсия, третий и четвертый центральные моменты равны соответственно

$$E\kappa(l_1, l_2) = pn, \quad V(\kappa(l_1, l_2)) = \frac{np\bar{p}(N-n)}{N-1}, \quad (2.14)$$

$$\mu_3(\kappa(l_1, l_2)) = \frac{np\bar{p}(\bar{p}-p)(N-n)(N-2n)}{(N-1)(N-2)}, \quad (2.15)$$

$$\mu_4(\kappa(l_1, l_2)) = \frac{np\bar{p}(N-n)}{(N-1)(N-2)(N-3)} [N(N+1) - 6n(N-n) + 3p\bar{p}\{N^2(n-2) - Nn^2 + 6n(N-n)\}], \quad (2.16)$$

где $p = (l_2 - l_1)/N$, $\bar{p} = 1 - p$.

Теорема 6.2.1. *Гипергеометрический процесс $\kappa(l)$, $0 \leq l \leq N-1$, является цепью Маркова с вероятностями переходов*

$$\begin{aligned} P\{\kappa(l+1) = k+1 | \kappa(l) = k\} &= \frac{n-k}{N-l}, \\ P\{\kappa(l+1) = k | \kappa(l) = k\} &= 1 - \frac{n-k}{N-l}. \end{aligned} \quad (2.17)$$

Доказательство. По определению процесса $\kappa(l)$ каждой случайной выборке объема n соответствует траектория процесса $\kappa(l)$, $0 \leq l \leq N$, со скачками при тех значениях l , которым соответствуют $X_{(l)}$, попавшие в выборку. Все выборки равновероятны и общее их число равно $\binom{N}{n}$. Поэтому вероятность появления траектории со скачками в заранее указанных точках $1 \leq l_1 < l_2 < \dots < l_n \leq N$ равна $\binom{N}{n}^{-1}$.

Рассмотрим теперь марковский процесс $\tilde{\kappa}(l)$ с вероятностями переходов (2.17), $\tilde{\kappa}(0) = 0$, и вычислим для него вероятность появления траектории со скачками в точках $1 \leq l_1 < l_2 < \dots < l_n \leq N$. Введем случайные величины ε_l , $0 \leq l \leq N$, полагая $\varepsilon_l = 1$, если $\tilde{\kappa}(l) - \tilde{\kappa}(l-1) = 1$, и $\varepsilon_l = 0$ в противном случае. Тогда $\tilde{\kappa}(l) = \sum_{1 \leq i \leq l} \varepsilon_i$. Из марковости процесса $\tilde{\kappa}(l)$ находим,

что вероятность заданной траектории равна

$$\begin{aligned} P\{\varepsilon_{l_i}=1, \varepsilon_j=0, j \neq l_i, 1 \leq i \leq n, 1 \leq j \leq N\} = \\ = \prod_{j=0}^{N-1} P\{\varepsilon_{j+1} | \varepsilon_j\} = \prod_{j=1}^{N-1} \left(1 - \frac{n - \sum_{1 \leq i \leq j} \varepsilon_i}{N-j+1}\right)^{1-\varepsilon_j} \times \\ \times \left(\frac{n - \sum_{1 \leq i \leq j} \varepsilon_i + 1}{N-j+1}\right)^{\varepsilon_j} = \\ = \frac{(N-n)(N-n+1) \dots (N-l_n+1) n(n-1) \dots 2 \cdot 1}{N(N-1) \dots (N-l_n+1)} = \frac{1}{\binom{N}{n}}, \end{aligned}$$

где $P\{\varepsilon_0=0\}=1$.

Утверждение теоремы вытекает из совпадения вероятностей для траекторий процессов $\kappa(l)$ и $\tilde{\kappa}(l)$.

Введем статистику

$$\rho_{n,N} = \max_x |F_n(x) - F_N(x)|, \quad (2.18)$$

равную максимуму отклонения между эмпирической функцией распределения выборки $F_n(x)$ и общей функцией «распределения» совокупности $F_N(x)$.

Лемма 6.2.1. *Выполнено соотношение*

$$P\{\rho_{n,N} \leq y\} = P\left\{\max_{1 \leq l \leq N} \left|\frac{\kappa(l)}{n} - \frac{l}{N}\right| \leq y\right\}. \quad (2.19)$$

Доказательство немедленно следует из того, что максимум модуля разности $F_n(x)$ и $F_N(x)$ достигается в точках $X_{(l)}$, для которых $F_N(X_{(l)}) = l/N$, а $F_n(X_{(l)}) = \kappa(l)/n$.

Используя свойства марковости процесса $\kappa(l)$, можно предложить алгоритмы вычисления процентных точек функции распределения $\rho_{n,N}$. Длительность счета по таким алгоритмам быстро растет с ростом N и n . Имея в виду предельную сходимость результатов вычислений при больших N и n , удобнее вычисления проводить

для нормированной случайной величины $\sqrt{\frac{nN}{N-n}} \rho_{n,N}$.

В работе Рыковой [2] был предложен алгоритм вычисления γ -процентных точек (квантилей уровня γ) статистики $\sqrt{\frac{nN}{N-n}} \rho_{n,N}$ и рассчитаны таблицы для значе-

ний $\gamma = 0,9, 0,95, 0,99$, $n \leq N/2$, $N = 10, 12, 14, 16, 20, 25, 30$, $n = 3(1)15$. Метод, положенный в основу этого алгоритма, состоит в следующем. Каждой выборке s соответствует траектория $\kappa_s(l)$ марковского процесса $\kappa(l)$, $1 \leq l \leq N$. В «момент» l каждой траектории поставим в соответствие две величины

$$\varphi_s(l) = \frac{\kappa_s(l)}{n} - \frac{l}{N} \quad \text{и} \quad \rho_s(l) = \max_{1 \leq i \leq l} \left| \frac{\kappa_s(i)}{n} - \frac{i}{N} \right|.$$

На множестве траекторий введем отношение эквивалентности. Две траектории, соответствующие s_1 и s_2 , назовем *l-эквивалентными*, если в момент l имеем $\varphi_{s_1}(l) = \varphi_{s_2}(l)$, $\rho_{s_1}(l) = \rho_{s_2}(l)$. Пусть $j = j(l)$ — номера различных *эквивалентных классов*, на которые разбивается все множество траекторий на интервале $[0, l]$, а $k_j(l)$ — число различных траекторий, входящих в класс с номером j .

Каждый класс *l-эквивалентных траекторий* характеризуется набором $(\varphi_j(l), \rho_j(l), k_j(l))$. В последний момент $l = N$ все значения $\varphi_j(N) = 0$, $\rho_j(N) = \rho_{n,N}$, а эквивалентные классы оказываются классами траекторий (выборок), имеющих одинаковые значения $\rho_{n,N}$. При этом $k_j(N)$ равно числу траекторий, имеющих одинаковое значение $\rho_{n,N} = \rho_j(N)$. Алгоритм дает правила рекуррентного пересчета значений $\varphi_j(l)$, $\rho_j(l)$, $k_j(l)$ при переходе от l к $l+1$. Так как все траектории равновероятны, то вероятности появления значений $\rho_j(N)$ найдутся как отношения $k_j(N)/\binom{N}{n}$. Полученное распределение вероятностей является дискретным, поэтому указывается наименьшее из чисел $\beta_\gamma(n, N)$ таких, что

$$P \left\{ \sqrt{\frac{nN}{N-n}} \rho_{n,N} \leq \beta_\gamma(n, N) \right\} \geq \gamma. \quad (2.20)$$

Так, например, при $N = 30$, $n = 10$, $\beta_{0,9}(10, 30) = 1,162$. Фактическое значение вероятности в (2.20) при этом оказывается равным 0,938, т. е. больше 0,9. В приложении приведена таблица X чисел $\beta_\gamma(n, N)$ для $\gamma = 0,9$, которая взята из работы Рыковой [2]. В каждой паре чисел этой таблицы верхнее значение соответствует $\beta_\gamma(n, N)$, а нижнее — фактическому значению вероятности (2.20). В последнем вертикальном столбце приведены значения точек уровня 0,9 для классического

распределения Колмогорова $K_n(y)$ при тех же объемах выборки n , соответствующие выбору при $N = \infty$. Таблицы точек уровня $\beta_\gamma(n, N)$ можно использовать для построения γ -доверительных полос для неизвестной общей функции «распределения» $F_N(x)$. Из определения статистики $\rho_{n, N}$ (см. (2.18)) и (2.20) следует, что с вероятностью, не меньшей γ , $F_N(x)$ при любом x заключена между $\underline{F}_\gamma(x)$ и $\bar{F}_\gamma(x)$, где

$$\begin{aligned}\underline{F}_\gamma(x) &= F_n(x) - \sqrt{\frac{N-n}{N \cdot n}} \beta_\gamma(n, N), \\ \bar{F}_\gamma(x) &= F_n(x) + \sqrt{\frac{N-n}{N \cdot n}} \beta_\gamma(n, N).\end{aligned}\quad (2.21)$$

Легко проверить тождество

$$F_n(x) - F_{N-n}(x) = \frac{N-n}{N} [F_n(x) - F_{N-n}(x)],$$

где $F_{N-n}(x) = (l(x) - k(x)) / (N - n)$ — ф. р., соответствующая объектам, не включенным в выборку. Значения статистики $\rho_{n, N}$ зависят лишь от номеров объектов, включенных в выборку. Поэтому величины X_i , $i = 1, \dots, N$, можно рассматривать как случайные. Предполагается, конечно, что случайные величины X_i и случайный набор номеров, образующих выборку, являются взаимно независимыми. Это предположение выполняется, когда X_i — взаимно независимые одинаково распределенные случайные величины с непрерывной ф. р. В этом случае $F_n(x)$ и $F_{N-n}(x)$ взаимно независимы. Таким образом,

$$P\{\rho_{n, N} \leq y\} = P^* \left\{ \max_x |F_{n_1}(x) - F_{n_2}(x)| \leq \frac{N \cdot y}{N - n} \right\},$$

где P^* соответствует схеме Н. В. Смирнова двух взаимно независимых выборок $n_1 = n$, $n_2 = N - n$. Соответствующие таблицы критических уровней (аналогичные таблице X) для $n_1, n_2 \leq 50$ были рассчитаны Боровковым, Марковой, Сычевой [1].

При $N \rightarrow \infty$, $n \rightarrow \infty$ объектом нашего исследования будет случайный процесс

$$\beta_{n, N} \left(\frac{l}{N} \right) = \sqrt{\frac{nN}{N-n}} \left[\frac{x(l)}{n} - \frac{l}{N} \right], \quad (2.22)$$

получающийся из $\kappa(l)$ путем линейных преобразований. Этот процесс можно доопределить для значений $t \in \left(\frac{l}{N}, \frac{l+1}{N}\right)$ по формуле

$$\beta_{n,N}(t) = \left[1 - N\left(t - \frac{l}{N}\right)\right] \beta_{n,N}\left(\frac{l}{N}\right) + N\left(t - \frac{l}{N}\right) \cdot \beta_{n,N}\left(\frac{l+1}{N}\right). \quad (2.23)$$

Так как $\kappa(l)$ и его приращения имеют гипергеометрические распределения, то нетрудно найти все моментные характеристики процесса $\beta_{n,N}(t)$. В частности, легко проверить, что

$$\begin{aligned} E\beta_{n,N}(t) &\equiv 0, \quad E(\beta_{n,N}(t))^2 = t(1-t) \frac{N}{N-1}, \\ t &= \frac{l}{N}, \quad 1 \leq l \leq N. \end{aligned} \quad (2.24)$$

Просто вычисляются среднее и дисперсия при условии $\beta_{n,N}(t) = x$. Здесь для $t = l/N$, $h = m/N$ имеем

$$\begin{aligned} E\{\beta_{n,N}(t+h) - \beta_{n,N}(t) | \beta_{n,N}(t) = x\} &= \\ &= \sqrt{\frac{nN}{N-n}} E\left\{ \frac{\kappa(l+m) - \kappa(l)}{n} - \frac{m}{n} \middle| \kappa(l) = \right. \\ &= n\left(x \sqrt{\frac{N-n}{nN}} + \frac{l}{N}\right)\Big\} = \\ &= \sqrt{\frac{nN}{N-n}} \left\{ \frac{n - n\left(x \sqrt{\frac{N-n}{nN}} + \frac{l}{N}\right)}{n(N-l)} m - \frac{m}{N} \right\} = \\ &= -\frac{xh}{1-t}, \end{aligned} \quad (2.25)$$

$$\begin{aligned} E\left\{ \left[\beta_{n,N}(t+h) - \beta_{n,N}(t) + \frac{xh}{1-t} \right] \middle| \beta_{n,N}(t) = x \right\} &= \\ &= \frac{h}{1-t} \left(1 - \frac{h}{1-t}\right) + o(1) \end{aligned} \quad (2.26)$$

при $\min(n, N-n) \rightarrow \infty$, $|x| < c < \infty$, $0 < \underline{c} \leq h \leq \bar{c} < 1$.

Соотношения (2.24)–(2.26) говорят о том, что случайный процесс $\beta_{n,N}(t)$, $t = l/N$, является дискретным аналогом условного винеровского процесса с закрепленными концами. Напомним, что **винеровский процесс** $\beta(t)$,

$t \geq 0$, $\beta(0) = 0$, есть гауссовский процесс с независимыми приращениями, $E\beta(t) = 0$, $E\beta^2(t) = t$, имеющий ковариационную функцию $E\beta(s)\beta(t) = \min(s, t)$. **Условный винеровский процесс $\beta_0(t)$ с двумя закрепленными концами** («броуновский мост») получается, если перейти к условным распределениям $\beta(t)$, предположив, что $\beta(1) = \beta(0) = 0$. Его ковариационная функция равна $E\beta_0(s)\beta_0(t) = s(1-t)$, $s < t$, $E\beta_0(t) = 0$. Локальные характеристики этого процесса имеют вид

$$E\{\beta_0(t+h) - \beta_0(t) \mid \beta_0(t) = x\} = -\frac{xh}{1-t},$$

$$E\left\{\left[\beta_0(t+h) - \beta_0(t) + \frac{xh}{1-t}\right]^2 \mid \beta_0(t) = x\right\} =$$

$$= \frac{h}{1-t} \left(1 - \frac{h}{1-t}\right),$$

что согласуется с приведенными выше формулами (2.25), (2.26).

Неравенство для моментов четвертого порядка. Для моментов четвертого порядка приращений процесса $\beta_{n,N}(t)$ справедливо следующее утверждение (Рыкова [3]).

Теорема 6.2.2. Для любых t_1, t_2 из $[0, 1]$, всех v , $0 < v_1 \leq v = \frac{n}{N} \leq v_2 < 1$ и $N > 3$ существует $H = H(v_1, v_2)$ такое, что для $\Delta_{n,N}(t_1, t_2) = \beta_{n,N}(t_2) - \beta_{n,N}(t_1)$ выполнено неравенство

$$\mu_4 = E[\Delta_{n,N}(t_1, t_2)]^4 \leq H |t_2 - t_1|^2. \quad (2.27)$$

Доказательство. Пусть $t_i = \frac{l_i}{N}$, $i = 1, 2$; тогда из (2.22) находим, что

$$\Delta_{n,N}(t_1, t_2) = \frac{1}{\sqrt{n(1-v)}} \left[\chi(l_1, l_2) - \frac{n(l_2 - l_1)}{N} \right]. \quad (2.28)$$

Четвертый момент (2.28) в соответствии с (2.16) при $p = \Delta t$, $\bar{p} = 1 - \Delta t$, $\Delta t = \frac{l_2 - l_1}{N}$, $l_1 \neq l_2$, равен

$$\mu_4 = \Delta t (1 - \Delta t) \left[\left(1 + \frac{1}{N}\right) - 6v(1-v) + 3\Delta t(1 - \Delta t) \times \right.$$

$$\times \{(n-2) - nv + 6v(1-v)\} \times$$

$$\times \left[n(1-v) \left(1 - \frac{1}{N}\right) \left(1 - \frac{2}{N}\right) \left(1 - \frac{3}{N}\right) \right]^{-1}. \quad (2.29)$$

Учитывая неравенства $1 + \frac{1}{N} \leq 2$, $\frac{1}{n} = \frac{1}{vN} \leq \frac{1}{v} \Delta t$, для (2.29) заключаем, что

$$\mu_4 \leq \frac{(\Delta t)^2 (1 - \Delta t) [2 + 3(1 - \Delta t)v(1 - v)]}{v(1 - v)(1 - 3/N)^3} \leq H_1(v_1, v_2)(\Delta t)^2, \quad (2.30)$$

где $H_1 = H_1(v_1, v_2) = \frac{1}{(1 - 3/N)^3} \left[\frac{2}{v_1(1 - v_2)} + 3 \right] < \infty$.

Пусть теперь $t_i \in \left[\frac{l_i}{N}, \frac{l_i + 1}{N} \right]$, $p_i = 1 - N \left(t_i - \frac{l_i}{N} \right)$, $i = 1, 2$. В этих обозначениях согласно (2.23) имеем

$$\begin{aligned} \Delta_{n, N}(t_1, t_2) &= p_2 \Delta_{n, N} \left(\frac{l_1}{N}, \frac{l_2}{N} \right) + \\ &+ (1 - p_2) \Delta_{n, N} \left(\frac{l_1 + 1}{N}, \frac{l_2 + 1}{N} \right) + \\ &+ (p_1 - p_2) \Delta_{n, N} \left(\frac{l_1}{N}, \frac{l_1 + 1}{N} \right). \end{aligned} \quad (2.31)$$

Для получения оценки сверху для μ_4 воспользуемся неравенством $(a + b + c)^4 \leq 27(a^4 + b^4 + c^4)$. Тогда из (2.31) и (2.30) получаем

$$\begin{aligned} \mu_4 &\leq 27 \left(p_2^4 H_1 \left(\frac{l_2 - l_1}{N} \right)^2 + (1 - p_2)^4 H_1 \left(\frac{l_2 - l_1}{N} \right)^2 + \right. \\ &\quad \left. + (p_1 - p_2)^4 H_1 \frac{1}{N^2} \right). \end{aligned} \quad (2.32)$$

В том случае, когда t_1, t_2 принадлежат одному интервалу, имеем $t_2 - t_1 \leq 1/N$ или $N(t_2 - t_1) \leq 1$, поэтому

$$(p_1 - p_2)^4 = N^4 (t_2 - t_1)^4 \leq N^2 (t_2 - t_1)^2. \quad (2.33)$$

Из (2.32), (2.33) и равенства $l_1 = l_2$ находим, что

$$\mu_4 \leq 27 H_1 (t_2 - t_1)^2. \quad (2.34)$$

Если t_1, t_2 берутся из разных интервалов ($l_2 > l_1$) и $\Delta t = t_2 - t_1 \geq 1/N$, то с учетом неравенств

$$\begin{aligned} \left| \frac{l_2}{N} - \frac{l_1}{N} - \Delta t \right| &\leq \frac{1}{N} \leq \Delta t, \\ \left(\frac{l_2}{N} - \frac{l_1}{N} \right)^2 &\leq 2 \left(\left| \frac{l_2}{N} - \frac{l_1}{N} - \Delta t \right|^2 + (\Delta t)^2 \right) \leq 4(\Delta t)^2 \end{aligned}$$

из (2.32) получаем

$$\begin{aligned}\mu_4 &\leq 27H_1(4p_2^4(\Delta t)^2 + (1-p_2)^4 4(\Delta t)^2 + (p_1-p_2)^4(\Delta t)^2) \leq \\ &\leq 27 \cdot 5H_1(\Delta t)^2 = 135 \cdot H_1(\Delta t)^2.\end{aligned}\quad (2.35)$$

Осталось рассмотреть только случай, когда $l_2 = l_1 + 1$ и $t_2 - t_1 < \frac{1}{N}$. В этом случае (2.31) переписывается в следующем виде:

$$\begin{aligned}\Delta_{n,N}(t_2, t_1) &= p_1 \Delta_{n,N}\left(\frac{l_1}{N}, \frac{l_1+1}{N}\right) + \\ &+ (1-p_2) \Delta_{n,N}\left(\frac{l_1+1}{N}, \frac{l_1+2}{N}\right),\end{aligned}\quad (2.36)$$

где $p_1 = 1 - Nt_1 + l_1$, $p_2 = 1 - Nt_2 + l_1 + 1$.

Из (2.30), (2.35) и неравенств $(a+b)^4 \leq 8(a^4 + b^4)$, $p_1^4 + (1-p_2)^4 \leq (p_1 + 1 - p_2)^2 = [N(t_2 - t_1)]^2$, получим

$$\begin{aligned}\mu_4 &\leq 8 \left[p_1^4 H_1 \frac{1}{N^2} + (1-p_2)^4 H_1 \frac{1}{N^2} \right] = \\ &= 8H_1 \frac{1}{N^2} [p_1^4 + (1-p_2)^4] \leq 8H_1(\Delta t)^2.\end{aligned}\quad (2.37)$$

Из полученных неравенств (2.30), (2.34), (2.35) и (2.37) следует утверждение теоремы для $H = 135H_1$.

Оценку (2.30) можно заменить оценкой

$$\mu_4 \leq 4,2 \frac{\Delta t}{n(1-v)} + 10,8 \frac{\Delta t^2}{1-v}, \quad (n \geq 6), \quad (2.38)$$

которая получается из (2.29) с учетом неравенств $6v(1-v) \leq nv$ и $\left(1 - \frac{1}{N}\right)\left(1 - \frac{2}{N}\right)\left(1 - \frac{3}{N}\right) \geq \frac{5}{18}$, $1 + \frac{1}{N} \leq \frac{7}{6}$.

Теорема 6.2.2 может быть использована для получения результатов о сходимости по распределениям функционалов, определенных на траекториях $\beta_{n,N}(t)$, когда $n, N \rightarrow \infty$.

Локальная предельная теорема. Пусть N, M, n, m — целые числа, $m \leq M \leq N$, $m \leq n \leq N$, $0 < \underline{c} \leq \Delta \leq \bar{c} < 1$,

$b(n, N) = \frac{n(N-n)}{N}$, и выполнены условия

$$w = \min(n, N-n) \rightarrow \infty, \quad \left| \frac{M}{N} - \Delta \right| \leq \frac{C}{N}, \quad (2.39)$$

$$\frac{\left(m - n \frac{M}{N}\right)^3}{b(n, N)^2} = \alpha_{N, M}^{n, m} \rightarrow 0 \quad (w \rightarrow \infty). \quad (2.40)$$

Докажем важный вспомогательный результат.

Теорема 6.2.3 (Рыкова [3]). Если выполнены условия (2.39) и (2.40), то для любого $\varepsilon > 0$ и всех достаточно больших w имеем

$$h_{N, M}^{n, m} = \frac{1}{\sqrt{2\pi\Delta(1-\Delta)b(n, N)}} \exp\left\{-\frac{z^2}{2\Delta(1-\Delta)b(n, N)}\right\} \times \\ \times (1 + \theta_{N, M}^{n, m}), \quad (2.41)$$

где $z = m - n \frac{M}{N}$, а

$$|\theta_{N, M}^{n, m}| \leq \frac{\gamma_{N, M}^{n, m}}{[\Delta(1-\Delta)(1-\varepsilon)]^2}, \quad \gamma_{N, M}^{n, m} = \left(\alpha_{N, M}^{n, m} + \frac{1}{\sqrt[3]{w}}\right). \quad (2.42)$$

Доказательство. Пусть κ — случайная величина $P\{\kappa = m\} = h_{N, M}^{n, m}$. Из свойств гипергеометрического распределения имеем

$$E\kappa = n(M/N), \quad V\kappa = (M/N)(1-M/N) \frac{n(N-n)}{N} \cdot \frac{N}{N-1}. \quad (2.43)$$

Согласно формуле Стирлинга (Феллер [1] стр. 67)

$$k! = \sqrt{2\pi} k^{k+1/2} e^{-k+\theta_k/(12k)}, \quad (2.44)$$

где $|\theta_k| \leq \theta_+ < 2$, т. е. θ_k равномерно ограничены по k . Расписывая выражение для $h_{N, M}^{n, m}$ через факториалы и применяя (2.44), получим

$$h_{N, M}^{n, m} = \frac{1}{\sqrt{2\pi}} \left[\frac{m(n-m)(M-m)(N-n-M+m)N}{n(N-n)M(N-M)} \right]^{-1/2} \times \\ \times \left(\frac{nM/N}{m}\right)^m \left(\frac{n \frac{N-M}{N}}{n-m}\right)^{n-m} \left(\frac{M(N-n)}{(M-m)N}\right)^{M-m} \times \\ \times \left(\frac{(N-n)(N-M)}{N(N-M-n+m)}\right)^{N-M-n+m} \exp\{-A_{N, M}^{n, m}\}, \quad (2.45)$$

где

$$A_{N,M}^{n,m} = \frac{\theta_n}{n} + \frac{\theta_{N-n}}{N-n} + \frac{\theta_M}{M} + \frac{\theta_{N-M}}{N-M} - \frac{\theta_m}{m} - \frac{\theta_{n-m}}{n-m} - \\ - \frac{\theta_{M-m}}{M-m} - \frac{\theta_N}{N} - \frac{\theta_{N-M-n+m}}{N-M-n+m}. \quad (2.46)$$

Из (2.39), (2.40) находим оценки $N-n \geq w$, $M \geq w\Delta(1-\varepsilon)$, $N-M \geq w(1-\Delta)(1-\varepsilon)$, ..., $N-M-(n-m) \geq w(1-\Delta)(1-\varepsilon)$ для всех достаточно больших w . Из (2.46) и этих оценок получаем

$$|A_{N,M}^{n,m}| \leq \frac{50+}{\Delta(1-\Delta)(1-\varepsilon)w}. \quad (2.47)$$

Полагая в (2.45) $m = n(M/N) + z$, после элементарных преобразований находим, что

$$h_{N,M}^{n,m} = \frac{1}{\sqrt{2\pi}} [(M/N)(1-M/N)b(n, N)B_{N,M}^{n,m}]^{-\frac{1}{2}} \times \\ \times C_{N,M}^{n,m} \exp\{-A_{N,M}^{n,m}\}, \quad (2.48)$$

где

$$\left. \begin{aligned} B_{N,M}^{n,m} &= \left(1 + \frac{z}{nM/N}\right) \left(1 - \frac{z}{(N-n)M/N}\right) \times \\ &\times \left(1 - \frac{z}{n(1-M/N)}\right) \left(1 + \frac{z}{(N-n)(1-M/N)}\right), \\ C_{N,M}^{n,m} &= \left(1 + \frac{z}{nM/N}\right)^{n(M/N)+z} \times \\ &\times \left(1 - \frac{z}{(N-n)M/N}\right)^{(N-n)(M/N)-z} \times \\ &\times \left(1 - \frac{z}{n(1-M/N)}\right)^{n(1-M/N)-z} \times \\ &\times \left(1 + \frac{z}{(N-n)(1-M/N)}\right)^{(N-n)(1-M/N)+z}. \end{aligned} \right\} \quad (2.49)$$

Из (2.49) следует, что для достаточно больших w

$$(B_{N,M}^{n,m})^{-\frac{1}{2}} = 1 + \tilde{B}_{N,M}^{n,m}, \quad |\tilde{B}_{N,M}^{n,m}| \leq \frac{2}{\Delta(1-\Delta)(1-\varepsilon)\sqrt[3]{w}}. \quad (2.50)$$

Вторые слагаемые в каждом из сомножителей в (2.49) стремятся к нулю при выполнении (2.39). Используя неравенства $\alpha - \alpha^2/2 < \ln(1+\alpha) < \alpha - \alpha^2/2 + \alpha^3/3$ при $0 <$

$\alpha < 1$ и $\alpha - (\alpha^2/2) + 2\alpha^3/3 < \ln(1 + \alpha) < \alpha - \alpha^2/2$ при $-1/2 < \alpha < 0$ имеем

$$\ln C_{N,M}^{n,m} = -(z/2)^2 [b(n; N)(M/N)(1 - M/N)]^{-1} + G_{N,M}^{n,m}, \quad (2.51)$$

где $G_{N,M}^{n,m}$ равномерно ограничена по n, m, N, M при выполнении (2.39), (2.40), и для достаточно больших w

$$|G_{N,M}^{n,m}| \leq \frac{\alpha_{N,M}^{n,m}}{[\Delta(1 - \Delta)(1 - \epsilon)]^2}. \quad (2.52)$$

Объединяя полученные оценки (2.47), (2.50), (2.52), приходим к искомому утверждению теоремы.

Теорему 6.2.3 естественно называть *локальной предельной теоремой для гипергеометрического распределения*.

Предельные теоремы для функционалов от траекторий гипергеометрических процессов. Введенный выше условный винеровский процесс $\beta_0(t)$ следует рассматривать как предельный для последовательности гипергеометрических процессов $\beta_{n,N}(t)$ (Беляев, Рыкова [1]).

Теорема 6.2.4. При $n \rightarrow \infty$, $\lim_{N \rightarrow \infty} n/N = v_+ < 1$

конечномерные распределения процессов $\beta_{n,N}(t)$ сходятся к конечномерным распределениям процесса $\beta_0(t)$, $t \in [0, 1]$.

Доказательство. Надо показать, что для любого набора интервалов $A_i = (a_i, b_i)$, $a_i < b_i$ и моментов времени t_i , $i = 1, \dots, r$, $0 < t_1 < t_2 < \dots < t_r < 1$,

$$\lim_{n \rightarrow \infty} P \{ \beta_{n,N}(t_i) \in A_i, 1 \leq i \leq r \} = P \{ \beta_0(t_i) \in A_i, 1 \leq i \leq r \}.$$

Из определяющих соотношений (2.22) и (2.23) следует, что при $n \rightarrow \infty$ для $t \in [l/N, (l+1)/N]$ с вероятностью

$$\text{единица} \quad | \beta_{n,N}(t) - \beta_{n,N}(l/N) | \leq \max \left(\sqrt{\frac{n}{N(N-n)}}, \right.$$

$$\left. \sqrt{\frac{N-n}{Nn}} \right) \rightarrow 0. \text{ Поэтому предельные распределения}$$

$\beta_{n,N}(l_i/N)$ и $\beta_{n,N}(t_i)$ при $| (l_i/N) - t_i | \leq 1/N$ совпадают.

В связи с этим мы будем изучать предельное поведение вероятностей $P \{ \beta_{n,N}(l_i/N) \in A_i, 1 \leq i \leq r \}$, где l_i при каждом N выбираются таким образом, что $| (l_i/N) - t_i | =$

$$= \min_{1 \leq i \leq N} | (l/N) - t_i | \leq (2N)^{-1}.$$

Совместная плотность вероятностей случайных величин $\beta_0(t_i)$, $1 \leq i \leq r$, имеет вид

$$p_0(x_1, \dots, x_r) = p_0(x_1) p_0(x_2 | x_1) \dots p_0(x_r | x_{r-1}), \quad (2.53)$$

где

$$p_0(x_{i+1} | x_i) = \frac{1}{\sqrt{2\pi\Delta_i \left(1 - \frac{\Delta_i}{1-t_i}\right)}} \times \\ \times \exp \left\{ -\frac{\left[x_{i+1} - x_i + \frac{x_i \Delta_i}{1-t_i}\right]^2}{2\Delta_i \left(1 - \frac{\Delta_i}{1-t_i}\right)} \right\}, \quad (2.54)$$

$$\Delta_i = t_{i+1} - t_i, \quad p_0(x_1) = p_0(x_1 | 0), \quad t_0 = 0.$$

Так как $\beta_{n,N}(l_i/N)$ принимает конечное число значений x_i , то

$$P\{\beta_{n,N}(l_i/N) \in A_i, 1 \leq i \leq r\} = \\ = \sum_{(x_1, \dots, x_r) \in W} P\{\beta_{n,N}(l_i/N) = x_i, 1 \leq i \leq r\}, \quad (2.55)$$

где $W = \{(x'_1, \dots, x'_r): x'_i \in A_i, 1 \leq i \leq r\}$.

Используя свойство марковости процесса $\beta_{n,N}(t)$, можем записать

$$P_{n,N}(x_1, \dots, x_r) = P\{\beta_{n,N}(l_i/N) = x_i, 1 \leq i \leq r\} = \\ = P_{n,N}(x_1) \prod_{i=1}^{r-1} P_{n,N}(x_{i+1} | x_i), \quad (2.56)$$

где

$$P_{n,N}(x_{i+1} | x_i) = h_{N-l_i, l_{i+1}-l_i}^{n-x_i, x_{i+1}-x_i}, \quad (2.57)$$

$$P_{n,N}(x_1) = P_{n,N}(x_1 | 0),$$

$l_0 = 0$, $x_0 = 0$. Числа x_j , x_j , l_j связаны соотношениями ($j = i, i+1$)

$$\beta_{n,N}(l_j/N) = x_j = \sqrt{\frac{nN}{N-n}} \left(\frac{x_j}{n} - \frac{l_j}{N} \right). \quad (2.58)$$

Проверим выполнение условий (2.39), (2.40) локальной предельной теоремы 6.2.3. Левое соотношение (2.39) выполнено, так как $w \geq \min(n, N(1-\lambda_+)) \rightarrow \infty$, правое

выполнено, так как по выбору l_i , $\left| \frac{l_{i+1} - l_i}{N} - (t_{i+1} - t_i) \right| \leq \leq \frac{1}{N}$. Условие (2.40) в нашем случае принимает вид

$$\alpha' = \frac{\left[x_{i+1} - x_i - (n - x_i) \frac{l_{i+1} - l_i}{N - l_i} \right]^3}{\left[\frac{(n - x_i)(N - l_i - (n - x_i))}{N - l_i} \right]^2} \rightarrow 0. \quad (2.59)$$

Из (2.58) находим, что

$$\left. \begin{aligned} \frac{n - x_i}{n} &= 1 - l_i/N - \sqrt{\frac{N - n}{Nn}} x_i, \\ \frac{N - l_i - (n - x_i)}{N - n} &= 1 - (l_i/N) + \sqrt{\frac{n}{N(N - n)}} x_i. \end{aligned} \right\} \quad (2.60)$$

Из (2.60) и $|x_i| < c < \infty$ следует, что знаменатель в (2.59) эквивалентен $(b(n, N))^2 (1 - l_i/N)^2$, тогда как числитель равен $\left\{ \sqrt{\frac{n(N - n)}{N}} \left[x_{i+1} - x_i \left(1 - \frac{l_{i+1} - l_i}{N} \right) \right] \right\}^3$ и поэтому (2.59) выполнено, причем $\theta_{N, M}^{n, m}$ в (2.41) имеет порядок $O(1/\sqrt[3]{n})$, т. к. $\alpha' = O(1/\sqrt[3]{n})$.

Таким образом, в соответствии с (2.41)

$$\begin{aligned} h_{N-l_i, l_{i+1}-l_i}^{n-x_i, x_{i+1}-x_i} &= \\ &= \frac{1}{\sqrt{2\pi(n-x_i) \frac{l_{i+1}-l_i}{N-l_i} \left(1 - \frac{l_{i+1}-l_i}{N-l_i}\right) \left(1 - \frac{n-x_i}{N-l_i}\right)}} \times \\ &\times \exp \left\{ - \frac{\left[x_{i+1} - x_i - (n - x_i) \frac{l_{i+1} - l_i}{N - l_i} \right]^2}{2 \frac{l_{i+1} - l_i}{N - l_i} \left(1 - \frac{l_{i+1} - l_i}{N - l_i}\right) \frac{(n - x_i)(N - l_i - n + x_i)}{N - l_i}} \right\} \times \\ &\times (1 + O(1/\sqrt[3]{n})) = \frac{1}{\sqrt{2\pi\Delta_i \left(1 - \frac{\Delta_i}{1-t_i}\right) b(n, N)}} \times \\ &\times \exp \left\{ - \frac{\left(x_{i+1} - x_i + x_i \frac{\Delta_i}{1-t_i} \right)^2}{2\Delta_i \left(1 - \frac{\Delta_i}{1-t_i}\right)} \right\} (1 + O(1/\sqrt[3]{n})), \quad (2.61) \end{aligned}$$

где $\sqrt[3]{n} \cdot O(1/\sqrt[3]{n})$ — равномерно ограниченная величина, когда все x_i изменяются в ограниченной области. Используя (2.61) для каждого из множителей в (2.56), получим

$$P_{n, N}(x_1, \dots, x_r) = [b(n, N)]^{-r/2} p_0(x_1, \dots, x_r) (1 + O(1/\sqrt[3]{n})). \quad (2.62)$$

Заметим теперь, что число возможных значений $x_i \in A_i$ эквивалентно $\sqrt{b(n, N)} |b_i - a_i|$. Это следует из того, что множество значений случайной величины $\beta_{n, N}(l_i/N)$ является арифметической прогрессией с разностью $b(n, N)^{-1/2}$. Общее количество наборов $(x_1, \dots, x_r)$, по которым берется сумма в (2.55), эквивалентно $b(n, N)^{r/2} \times \prod_{i=1}^r (b_i - a_i)$. Таким образом, с учетом (2.55) и (2.62), находим

$$P \left\{ \beta_{n, N} \left(\frac{l_i}{N} \right) \in A_i, 1 \leq i \leq r \right\} = \sum_{(x_1, \dots, x_r) \in \mathbb{W}} p_0(x_1, \dots, x_r) (b(n, N))^{-r/2} + O(1/\sqrt[3]{n}). \quad (2.63)$$

Правая часть (2.63) есть интегральная сумма для $\int_{\{y_i \in A_i, 1 \leq i \leq r\}} p_0(y_1, \dots, y_r) dy_1, \dots, dy_r$, если каждую точку $\mathbf{x} = (x_1, \dots, x_r)$ рассматривать как центр r -мерного куба

$K(\mathbf{x})$ со сторонами $[x_i - 1/(2\sqrt{b(n, N)}), x_i + 1/(2\sqrt{b(n, N)})]$. По теореме о среднем существует такое значение $\mathbf{x}^0 = (x_1^0, \dots, x_r^0) \in K(\mathbf{x})$, для которого выполнено

$$\int_{K(\mathbf{x})} p_0(y_1, \dots, y_r) dy_1, \dots, dy_r = (b(n, N))^{-r/2} p_0(x_1^0, \dots, x_r^0).$$

Отсюда получаем

$$(b(n, N))^{-r/2} p_0(x_1, \dots, x_r) = \int_{K(\mathbf{x})} p_0(y_1, \dots, y_r) dy_1 \dots dy_r \cdot \frac{p_0(x_1, \dots, x_r)}{p_0(x_1^0, \dots, x_r^0)}. \quad (2.64)$$

Из неравенств $|x_i^0 - x_i| \leq b(n, N)^{-1/2}$ и ограниченности x_i , $a_i \leq x_i \leq b_i$, $1 \leq i \leq r$, имеем

$$\frac{p_0(x_{i+1} | x_i)}{p_0(x_{i+1}^0 | x_i^0)} = \exp \left\{ - \frac{(x_{i+1} - x_i^0) + (x_i^0 - x_i) + \frac{x_i - x_i^0}{1 - t_i} \cdot \Delta_i}{2\Delta_i \left(1 - \frac{\Delta_i}{1 - t_i}\right)} \right\} \times \\ \times [x_{i+1} + x_{i+1}^0 - x_i - x_i^0 + (x_i + x_i^0) \Delta_i / (1 - t_i)] = \\ = 1 + O(1/\sqrt{n}). \quad (2.65)$$

Таким образом, из (2.56) и (2.65) выводим

$$\frac{p_0(x_1, \dots, x_r)}{p_0(x_1^0, \dots, x_r^0)} = 1 + O(1/\sqrt{n}). \quad (2.66)$$

Преобразуя слагаемые в (2.63) с учетом (2.64) и (2.66), находим, что

$$P \left\{ \beta_{n, N} \left(\frac{l_i}{N} \right) \in A_i, \quad 1 \leq i \leq r \right\} = \\ = \int_{\{y_i \in A_i, \quad 1 \leq i \leq r\}} p_0(y_1, \dots, y_r) dy_1 \dots dy_r + O(1/\sqrt{n}). \quad (2.67)$$

Утверждение теоремы 6.2.4 является непосредственным следствием (2.67).

Из сходимости всех конечномерных распределений $\beta_{n, N}(t)$ к соответствующим распределениям $\beta_0(t)$ еще, вообще говоря, нельзя заключить, что предельное при $N \rightarrow \infty$ распределение $\max_{0 \leq t \leq 1} |\beta_{n, N}(t)|$ совпадает с распределением $\max_{0 \leq t \leq 1} |\beta_0(t)|$. Если предположить, что при

$n \rightarrow \infty$ $0 < \lim_{n \rightarrow \infty} \frac{n}{N} \leq \overline{\lim}_{n \rightarrow \infty} \frac{n}{N} < 1$, то можно, используя

неравенство для моментов (2.27), получить искомое утверждение о сходимости распределений $\max_{0 \leq t \leq 1} |\beta_{n, N}(t)|$

как следствие теоремы 2 на стр. 485 книги Гихмана и Скорохода [1]. Однако ограничение $\lim_{n \rightarrow \infty} \frac{n}{N} > 0$ является

излишним. Для доказательства основной предельной теоремы нам потребуется следующее утверждение.

Теорема 6.2.5. Для любых сколь угодно малых $\varepsilon > 0$, $\delta > 0$ и всех достаточно больших n , N , $\frac{n}{N} \leq \lambda_+ < 1$, найдется такое достаточно малое $h(\varepsilon, \delta)$, что при $h \leq h(\varepsilon, \delta)$

$$P \left\{ \max_{\substack{|t_1 - t_2| \leq h, \\ 0 \leq t_i \leq 1, \\ i=1, 2}} |\beta_{n, N}(t_1) - \beta_{n, N}(t_2)| > \varepsilon \right\} < \delta. \quad (2.68)$$

Доказательство. Введем обозначение

$$A_\varepsilon(s_1, s_2) = \left\{ \max_{\substack{|t_1 - t_2| \leq h, \\ s_1 \leq t_i \leq s_2, \\ i=1, 2}} |\beta_{n, N}(t_1) - \beta_{n, N}(t_2)| > \varepsilon \right\}.$$

В (2.68) дана оценка $P(A_\varepsilon(0, 1))$. При любом h , $0 < h < 1$, имеем

$$A_\varepsilon(0, 1) \subseteq \bigcup_{k=1}^{k_0} A_\varepsilon\left(\frac{(k-1)h}{2}, \frac{(k+1)h}{2}\right) \cup A_\varepsilon\left(\frac{k_0+1}{2}h, 1\right), \quad (2.69)$$

где k_0 — целое и такое, что $\frac{k_0+1}{2}h < 1 < \frac{k_0+2}{2}h$. Определим l'_k и l''_k соотношениями $l'_k = \left[\frac{(k-1)}{2}hN \right]$, $l''_k = l'_k + [Nh] + 1$, где $[a]$, как обычно, целая часть числа a , $k = 1, \dots, k_0$. Так как при $[s, t] \subseteq [s', t']$ верно $A_\varepsilon(s, t) \subseteq A_\varepsilon(s', t')$, то из (2.69) получим

$$P\{A_\varepsilon(0, 1)\} \leq \sum_{1 \leq k \leq k_0} P\left\{A_\varepsilon\left(\frac{l'_k}{N}, \frac{l''_k}{N}\right)\right\} + P\left\{A_\varepsilon\left(\frac{1}{N}\left[\frac{k_0Nh}{2}\right], 1\right)\right\}. \quad (2.70)$$

Для любого случайного процесса $\xi(t)$, $a < b$ и $t_1, t_2 \in [a, b]$

$$\left\{ \max_{|t_1 - t_2| \leq h} |\xi(t_1) - \xi(t_2)| > \varepsilon \right\} \subseteq \left\{ \max_{a \leq t \leq b} |\xi(t) - \xi(a)| > \frac{\varepsilon}{2} \right\}.$$

Поэтому для любых целых l_1, l_2 , $0 \leq l_1 < l_2 \leq N$, с учетом (2.22), найдем, что

$$\begin{aligned} A_\varepsilon\left(\frac{l_1}{N}, \frac{l_2}{N}\right) &\equiv \\ &\equiv \left\{ \max_{|l_1 - l_2| \leq [Nh] + 1} \left| \kappa(\tilde{l}_2) - \kappa(\tilde{l}_1) - \frac{l_2 - l_1}{N} n \right| > \varepsilon \sqrt{n} \sqrt{1 - \lambda_+} \right\} \equiv \\ &\equiv B_c(l_1, l_2) = \left\{ \max_{l_1 \leq l \leq l_2} \left| \kappa(l) - \kappa(l_1) - \frac{l - l_1}{N} n \right| > c \sqrt{n} \right\}, \end{aligned} \quad (2.71)$$

где $c = \varepsilon \sqrt{1 - \lambda_+}$, $l_1 \leq \tilde{l}_i \leq l_2$, $i = 1, 2$. Так как разность $\kappa(l) - \kappa(l_1)$ соответствует числу объектов в выборке из совокупности с номерами от $l_1 + 1$ до l , то вероятности событий $B_c(l'_k, l''_k)$ одинаковы, так как $l''_k - l'_k = [Nh] + 1$. Поэтому из (2.70) и (2.71) получаем оценку

$$\mathbf{P}\{A_\varepsilon(0, 1)\} \leq \frac{2}{h} \mathbf{P}\{B_c(0, [Nh] + 1)\}. \quad (2.72)$$

Итак, задача сведена к оценке вероятности того, что гипергеометрический процесс на интервале времени $\Delta_h = [0, [Nh] + 1]$ уклонится от своего математического ожидания более чем на $c \sqrt{n}$. Искомая оценка будет получена, если при $n \rightarrow \infty$, $N \rightarrow \infty$ ($\frac{n}{N} \leq \lambda_+ < 1$) выполнено соотношение

$$h^{-1} \mathbf{P}\{B_c(0, [Nh] + 1)\} \rightarrow 0 \quad (2.73)$$

равномерно по $h \geq \frac{c \sqrt{n} - 1}{N}$, тогда как при $h < \frac{c \sqrt{n} - 1}{N}$ $\mathbf{P}\{B_c(0, [Nh] + 1)\} = 0$.

Выше было показано, что гипергеометрический процесс $\kappa(l)$ является марковским с вероятностями переходов (2.17). Для множества точек (k, l) , $\left| k - \frac{l}{N} n \right| \leq c \sqrt{n}$ минимальное и максимальное значения вероятностей скачков ($\kappa(l) = k$, $\kappa(l + 1) = k + 1$) лежат внутри интервала $[p_1, p_2]$, где

$$p_1 = \frac{n(1 - h - 1/N) - c \sqrt{n}}{N(1 - h)}, \quad p_2 = \frac{n(1 - h) + c \sqrt{n}}{N(1 - h - 1/N)}. \quad (2.74)$$

Введем для $l \in \Delta_h$ два марковских процесса, $\kappa_1(l)$ и $\kappa_2(l)$, с вероятностями скачков (2.74). Так как p_i не зависит от значений $\kappa_i(l)$, то $\kappa_i(l) = \sum_{j=1}^l \kappa_{j,i}$, где $\kappa_{j,i} = 0, 1$ — взаимно независимые по j случайные величины, $P\{\kappa_{j,i} = 1\} = p_i$, $i = 1, 2$, $1 \leq j \leq l$. Для всех $l \in \Delta_h$ и любого как угодно малого $\varepsilon_1 > 0$ начиная с некоторого n , для $h < \frac{\varepsilon_1}{\varepsilon_1 + c}$ выполнено

$$|E\kappa(l) - E\kappa_i(l)| < \varepsilon_1 \sqrt{n}, \quad i = 1, 2.$$

Поэтому для достаточно больших N и n

$$P\{B_c(0, [Nh] + 1)\} \leq \sum_{i=1,2} P\left\{\max_{l \in \Delta_h} \left| \sum_{1 \leq j \leq l} (\kappa_{j,i} - p_i) \right| > (c - \varepsilon_1) \sqrt{n}\right\}. \quad (2.75)$$

Используя неравенство Чебышева и очевидное неравенство $V(\kappa_i(l)) \leq (Nh + 1)p_i(1 - p_i)$, имеем

$$\begin{aligned} P\left\{\left| \sum_{1 \leq j \leq Nh+1} (\kappa_{j,i} - p_i) \right| \leq \frac{c - \varepsilon_1}{2} \sqrt{n}\right\} &\geq \\ &\geq 1 - \frac{4(Nh + 1)}{(c - \varepsilon_1)^2 n} \cdot \frac{n(1 - h) + c\sqrt{n}}{N\left(1 - h - \frac{1}{N}\right)} \times \\ &\times \left(1 - \frac{n\left(1 - h - \frac{1}{N}\right) - c\sqrt{n}}{N(1 - h)}\right) \geq \frac{1}{2}, \end{aligned} \quad (2.76)$$

если только n и N выбраны достаточно большими, а $h \leq (c - \varepsilon_1)^2 / [16 + 2(c - \varepsilon_1)^2]$.

Напомним неравенство Оттавиани, аналогичное неравенству Колмогорова (см. Ито [1], стр. 33).

Если ξ_j — взаимно независимые случайные величины,

$$E\xi_i = 0, \quad P\left\{\left| \sum_{j=1}^m \xi_j \right| \leq a\right\} \geq \frac{1}{2}, \quad 1 \leq l \leq m - 1, \quad \text{то}$$

$$P\left\{\max_{1 \leq l \leq m} \left| \sum_{j=1}^l \xi_j \right| > 2a\right\} \leq 2P\left\{\left| \sum_{j=1}^m \xi_j \right| > a\right\}. \quad (2.77)$$

Так как (2.76) выполнено, то для оценки правой части (2.75) можно использовать неравенство (2.77). Тогда

$$\begin{aligned} P\{B_c(0, [Nh] + 1)\} &\leq \\ &\leq 2 \sum_{i=1, 2} P\left\{\left|\sum_{j=1}^{[Nh]+1} (\kappa_{j, i} - p_i)\right| > \frac{c - \varepsilon_1}{2} \sqrt{n}\right\}, \quad (2.78) \end{aligned}$$

где $\sum_{j=1}^{[Nh]+1} \kappa_{j, i}$ — биномиально распределенные случайные величины с вероятностями появления «успеха» p_i и числом испытаний $[Nh] + 1$. Как известно, четвертый центральный момент $\mu_{4, i}$, $i = 1, 2$, биномиального распределения с параметром p_i и числом испытаний $[Nh] + 1$ равен

$$\mu_{4, i} = 3([Nh] + 1)^2 p_i^2 \bar{p}_i^2 + ([Nh] + 1) p_i \bar{p}_i (1 - 6p_i \bar{p}_i).$$

Учитывая $[Nh] + 1 \leq 2Nh$, получим оценку

$$\begin{aligned} \mu_{4, i} &\leq 12N^2 h^2 \left(\frac{n(1-h) + c\sqrt{n}}{N(1-h-1/N)} \right)^2 + 2Nh \frac{n(1-h) + c\sqrt{n}}{N(1-h-1/N)} = \\ &= 12h^2 n^2 \left(\frac{1-h + c/\sqrt{n}}{1-h-1/N} \right)^2 + 2nh \frac{1-h + c/\sqrt{n}}{1-h-1/N}. \quad (2.79) \end{aligned}$$

По неравенству Чебышева с четвертым центральным моментом, учитывая (2.78) и (2.79), имеем

$$\begin{aligned} h^{-1} P\{B_c(0, [Nh] + 1)\} &\leq \\ &\leq \frac{32}{(c - \varepsilon_1)^4} \left[12h \left(\frac{1-h + c/\sqrt{n}}{1-h-1/N} \right)^2 + \frac{2}{n} \left(\frac{1-h + c/\sqrt{n}}{1-h-1/N} \right) \right]. \quad (2.80) \end{aligned}$$

Таким образом, (2.73) выполнено. Следовательно, утверждение (2.68) теоремы 6.2.5 справедливо.

Для условного винеровского процесса $\beta_0(t)$ исследованы распределения различных функционалов. Например, Колмогоров [2], [3] показал, что

$$P\left\{\max_{0 \leq t \leq 1} |\beta_0(t)| \leq y\right\} = K(y) = \sum_{m=-\infty}^{+\infty} (-1)^m e^{-2m^2 y^2}. \quad (2.81)$$

Известно также (см. Смирнов [1], Большев и Смирнов [1]), что

$$P\{\max_{0 \leq t \leq 1} \beta_0(t) \leq y\} = 1 - e^{-2y^2}, \quad (2.82)$$

$$P\left\{\int_0^1 (\beta_0(t))^2 dt \leq y\right\} = a_1(y), \quad (2.83)$$

где

$$a_1(y) = \frac{1}{\sqrt{2y}} \sum_{j=0}^{\infty} \frac{\Gamma(j+1/2)}{\Gamma(1/2)\Gamma(j+1)} \sqrt{4j+1} \exp\left\{-\frac{(4j+1)^2}{16y}\right\} \times \\ \times \left\{I_{-\frac{1}{4}}\left[\frac{(4j+1)^2}{16y}\right] - I_{\frac{1}{4}}\left[\frac{(4j+1)^2}{16y}\right]\right\}, \quad (2.84)$$

$I_{-\frac{1}{4}}, I_{\frac{1}{4}}$ — функции Бесселя.

Функционалы в (2.81)–(2.83) равны соответственно

$$f_1(g) = \max_{0 \leq t \leq 1} |g(t)|,$$

$$f_2(t) = \max_{0 \leq t \leq 1} g(t),$$

$$f_3(t) = \int_0^1 [g(t)]^2 dt.$$

Все эти функционалы непрерывно зависят от функций $g(t)$ в равномерной метрике, т. е. для любого $\delta > 0$ найдется такое $\varepsilon > 0$, что из $\max_{0 \leq t \leq 1} |g_1(t) - g_2(t)| < \varepsilon$ следует $|f_j(g_1) - f_j(g_2)| < \delta$. В нашем случае аналогами указанных функционалов являются

$$f_1(\beta_{n,N}(t)) = \max_x \sqrt{\frac{nN}{N-n}} |F_n(x) - F_N(x)|,$$

$$f_2(\beta_{n,N}(t)) = \max_x \sqrt{\frac{nN}{N-n}} [F_n(x) - F_N(x)],$$

$$f_3(\beta_{n,N}(t)) = \sqrt{\frac{nN}{N-n}} \int_{-\infty}^{+\infty} [F_n(x) - F_N(x)]^2 dF_N(x).$$

Здесь в качестве аргумента функционала записана соответствующая траектория процесса $\beta_{n,N}(t)$, определяе-

мого соотношениями (2.22), (2.23), к которой мы приходим при отображении $X_{(l)} \rightarrow \frac{l}{N}$, $1 \leq l \leq N$.

Теорема 6.2.6. Для любого функционала f , определенного на непрерывных функциях $g(t)$, $0 \leq t \leq 1$, и непрерывно зависящего от $g(t)$ в равномерной метрике при $n \rightarrow \infty$, $\overline{\lim} \frac{n}{N} = \lambda_+ < 1$,

$$P\{f(\beta_{n,N}(t)) \leq y\} \rightarrow P\{f(\beta_0(t)) \leq y\}. \quad (2.85)$$

Для доказательства этого утверждения нужно использовать теоремы 6.2.4 и 6.2.5. Далее доказательство этого утверждения фактически следует доказательству достаточности предельной теоремы 1 на стр. 482 книги Гихмана и Скорохода [1], поэтому мы его не приводим.

Следствие 6.2.1. Для общей функции распределения совокупности $F_N(x)$ и эмпирической функции распределения $F_n(x)$ при условии $\overline{\lim}_{N \rightarrow \infty} \frac{n}{N} \leq \lambda_+ < 1$ выполнены предельные соотношения

$$\lim_{n \rightarrow \infty} P\left\{\sqrt{\frac{nN}{N-n}} \max_x |F_n(x) - F_N(x)| \leq y\right\} = K(y), \quad (2.86)$$

$$\lim_{n \rightarrow \infty} P\left\{\sqrt{\frac{nN}{N-n}} \max_x [F_n(x) - F_N(x)] \leq y\right\} = 1 - e^{-2y^2}, \quad (2.87)$$

$$\lim_{n \rightarrow \infty} P\left\{\sqrt{\frac{nN}{N-n}} \int_{-\infty}^{+\infty} [F_n(x) - F_N(x)]^2 dF_N(x) \leq y\right\} = a_1(y), \quad (2.88)$$

где $K(y)$ и $a_1(y)$ определяются соотношением (2.12) и (2.84).

Доказательство. Соотношения (2.86) и (2.87) непосредственно вытекают из непрерывности рассматриваемых функционалов и (2.85) теоремы 6.2.6. Соотношение (2.88) получаем в два этапа. Сначала, используя отображение $X_{(l)} \rightarrow \frac{l}{N}$, получаем функционал на траекториях процесса $\beta_{n,N}(t)$ при $t = \frac{l}{N}$, $1 \leq l \leq N$. Потом переходим к непрерывному времени согласно (2.23).

После этого убеждаемся, что замена $\frac{1}{N} \sum \sqrt{\frac{nN}{N-n}} \times$
 $\times \left[\frac{\kappa(l)}{N} - \frac{l}{N} n \right]^2$ на $\int_0^1 [\beta_{n, N}(t)]^2 dt$ не влияет на предель-
 ное распределение. Для функционала $\int_0^1 [\beta_{n, N}(t)]^2 dt$ ис-
 пользуем теорему 6.2.6.

Предельные теоремы можно было бы дополнить слу-
 чаем, когда при $N \rightarrow \infty$ значение n фиксировано. Анало-
 гично (2.86) можно показать, что

$$\lim_{N \rightarrow \infty} P \left\{ \sqrt{\frac{nN}{N-n}} \max_x |F_n(x) - F_N(x)| \leq y \right\} = K_n(y),$$

где

$$K_n(y) = P \left\{ \sqrt{n} \max_x |F_n(x) - F(x)| \leq y \right\},$$

$F_n(x)$ — эмпирическая функция распределения, соответ-
 ствующая значениям n взаимно независимых одинаково
 распределенных случайных величин с функцией рас-
 пределения $F(x)$.

О выявлении результатов воздействий. Пусть $\mathfrak{F} =$
 $= \{O_1, \dots, O_N\}$ — конечная совокупность объектов, кото-
 рые могут подвергаться воздействиям двух типов, при-
 чем каждый объект O_i может быть подвергнут воздей-
 ствию только одного типа. Предположим, что нас
 интересует вопрос о выявлении различий в степени воз-
 действия разных типов на некоторую скалярную харак-
 теристику объектов из $\mathfrak{F}$, значение которой становится
 известным в результате контроля воздействия на объ-
 екты. Если все объекты были бы проконтролированы
 после воздействий первого (второго) типа, то наблюда-
 лись бы величины $X_1^{(1)}, \dots, X_N^{(1)}$ ($X_1^{(2)}, \dots, X_N^{(2)}$); $X_i^{(j)}$ соот-
 ветствует значению скалярной характеристики объекта
 $O_i \in \mathfrak{F}$ в условиях воздействия j -го типа, $j = 1, 2$. Так
 как каждый объект нельзя проконтролировать при двух
 типах воздействий, то в плане контроля должны быть
 определены номера объектов, подвергаемых воздей-
 ствиям первого типа, и номера объектов, подвергаемых

воздействиям второго типа. Положим $\gamma_i = 0$, если O_i не контролируется, $\gamma_i = j$, если O_i контролируется после воздействия j -го типа, в результате чего становится известным $X_i^{(j)}$. План контроля Π , таким образом, определяется системой правил получения набора $(\gamma_1, \dots, \gamma_N)$. Пусть $s_j = (i_1^{(j)}, \dots, i_{n_j}^{(j)})$ — номера объектов, отбираемых для контроля по плану Π с воздействием j -го типа. В этом случае результаты контроля представимы двумя последовательностями значений

$$X_{i_1^{(1)}}^{(1)}, \dots, X_{i_{n_1}^{(1)}}^{(1)}; X_{i_1^{(2)}}^{(2)}, \dots, X_{i_{n_2}^{(2)}}^{(2)}, \quad n_1 + n_2 \leq N. \quad (2.89)$$

Задача состоит в проверке значимости различия воздействий на основе данных (2.89).

В качестве исходной гипотезы H_0 сформулируем гипотезу отсутствия различий в степени воздействий разных типов, которая в наших обозначениях эквивалентна утверждению о том, что $X_i^{(1)} = X_i^{(2)}$, $1 \leq i \leq N$. Требуется построить критерий, который бы ошибочно отвергал гипотезу H_0 , если она верна, с вероятностью, не превышающей заранее заданного числа α вне зависимости от фактических значений $\{X_i^{(j)}\}$.

Из возможных практических интерпретаций поставленной задачи прежде всего отметим, что она является типичной при обработке результатов разрушающих испытаний на надежность. Если испытываемый объект может быть реализован в одном из двух вариантов, то длительность безотказной работы $X_i^{(j)}$ объекта O_i , реализованного в варианте (j) , соответствует в описанной выше задаче воздействию j -го типа. Например, j -вариант (воздействие) соответствует определенному типу припоя для соединения контактов, а $X_i^{(j)}$ — число циклов нагрузки при испытаниях с вибрацией до обрыва соединения. Близкие задачи могут возникнуть при сравнительном анализе результатов лечения двумя типами лекарств или, вообще, при сравнении двух методов лечения. При этом объектами O_i , образующими $\mathfrak{X}$, будут пациенты — возможные кандидаты для лечения. Предполагается, что окончательные итоги лечения можно характеризовать скалярными величинами, например, временем $X_i^{(j)}$

от начала лечения до выздоровления. В других случаях $X_i^{(j)}$ может соответствовать степени восстановления функционирования пораженного органа и т. д. Известный американский ученый Ю. Нейман (J. Neuman), один из создателей современных методов математической статистики, в своих докладах (см. Нейман [2]) подчеркивал необходимость разработки статистических методов в задачах выявления влияния воздействия распыления йодистого серебра в атмосфере на количество выпадаемых осадков. Здесь $X_i^{(j)}$ — количество выпадаемых осадков, без распыления ($j = 1$) и с распылением ($j = 2$). Перечень возможных интерпретаций поставленной выше математической задачи можно было бы продолжить.

В общем случае план получения данных о воздействиях определяется заданием последовательности случайных величин $\gamma = (\gamma_1, \dots, \gamma_N)$, где, как и выше, $\gamma_i = 0, 1, 2$, если объект $O_i \in \mathfrak{F}$ не подвергается воздействию вообще или на него оказывается воздействие j -го типа, $s_j(\gamma) = \{i: \gamma_i = j, i = 1, \dots, N\}$, $j = 1, 2$. Пусть $p(\gamma)$ — вероятность того, что при плане Π будет использована последовательность $\gamma = (\gamma_1, \dots, \gamma_N)$. Число M различных последовательностей γ , для которых $p(\gamma) > 0$, не превышает 3^N . Для только что рассмотренного плана двух случайных выборок случайная последовательность γ получалась в два этапа: сначала бралась случайная выборка без возвращения объема n_1 , а затем вторая случайная выборка объема n_2 . Объекты, попавшие в j -ю выборку, подвергались воздействию j -го типа ($j = 1, 2$). Для плана двух случайных выборок $p(\gamma) = \frac{n_0! n_1! n_2!}{N!}$.

Схема выявления воздействий при заданном плане Π состоит в следующем. Исходя из заданных значений вероятностей $\{p(\gamma)\}$ с помощью случайного механизма (таблиц или датчиков случайных чисел) реализуется случайная последовательность $\gamma_\omega = (\gamma_{1\omega}, \dots, \gamma_{N\omega})$, $P\{\gamma_\omega = \gamma\} = p(\gamma)$. Объекты с номерами $i \in s_j(\gamma_\omega)$ подвергаются воздействию j -го типа. Окончательные результаты измерений имеют вид

$$\mathfrak{X} = \{X_i^{(1)}, i \in s_1(\gamma_\omega), X_l^{(2)}, l \in s_2(\gamma_\omega)\}. \quad (2.90)$$

Если $s_1(\gamma_\omega) \cup s_2(\gamma_\omega) = \{1, 2, \dots, N\}$, т. е. все объекты подвергаются воздействиям, то критерий проверки гипотезы H_0 можно получить, выбрав подходящую статистику $T(\gamma) \equiv T(\mathcal{X})$ от данных (2.90). При этом в условиях гипотезы H_0 для каждого объекта O_i известно значение характеристики $X_i^{(\gamma_i)} = X_i, i = 1, \dots, N$.

Пусть $\gamma^1, \dots, \gamma^M$ — все последовательности γ , которые при плане Π имеют положительные вероятности. Каждому значению γ^k сопоставим $T_k = T(\gamma^k)$, относя значение X_i в массив значений $X_i^{(1)}$ или $X_i^{(2)}$, в зависимости от того, будет $i \in s_1(\gamma^k)$ или $i \in s_2(\gamma^k)$. Если несколько последовательностей γ^k дают одинаковые значения статистики $T(\gamma^k) = T'$, то значению T' приписываем вероятность $P(T')$, равную сумме вероятностей этих последовательностей. Учитывая тенденцию изменений статистики T при появлении воздействий, когда $|X_i^{(1)} - X_i^{(2)}|$ возрастает, выделяем в множестве значений $\{T_k\}$ критическую область R_T такую, что сумма вероятностей $P(T_k), T_k \in R_T$, не превышает заданной величины α -вероятности ошибки первого рода. Если исходная последовательность γ_ω , в соответствии с которой были проведены испытания, дает значение $T(\gamma_\omega) \in R_T$, то гипотезу H_0 следует отвергнуть.

Указанный подход к построению критериев выявления воздействий весьма привлекателен, так как исходные ограничения минимальны. Аргументацией использования некоторых статистик T для построения критериев на начальной стадии исследования может служить тот факт, что эти статистики имеют прямые аналоги в схеме независимых испытаний и обладают при этом определенными свойствами оптимальности.

В тех случаях, когда не все объекты подвергаются воздействиям, задача построения критериев выявления воздействий становится существенно сложнее.

По результатам измерений $\mathcal{X}$, определенным соотношением (2.90), образуем вариационный ряд $Y_{(1)} < Y_{(2)} < \dots < Y_{(n_1+n_2)}$, где n_j — числа объектов в $s_j(\gamma_\omega)$, $j = 1, 2$. Если значение $Y_{(l)}$ соответствует объекту $O_{i(l)}$, подвергнутому воздействию $\varepsilon(l) = 1, 2$, то $Y_{(l)} = X_{i(l)}^{(\varepsilon(l))}$, $l = 1, \dots, n_1 + n_2$. По результатам измерений $\mathcal{X}$ можно

получить информацию о рангах

$$\rho(\gamma_\omega) = \begin{pmatrix} 1 & 2 & \dots & n_1 + n_2 \\ \varepsilon_1 & \varepsilon_2 & \dots & \varepsilon_1 + \varepsilon_2 \end{pmatrix}.$$

Статистики, являющиеся функциями от $\rho(\gamma_\omega)$, будем называть **наблюдаемыми ранговыми статистиками**. Отметим, что при $n_1 + n_2 < N$ значения рангов объектов, включенных в выборки $s_j(\gamma_\omega)$, $j = 1, 2$, относительно всей совокупности $\mathfrak{F}$, неизвестны. В случае $n_1 + n_2 = N$ распределения наблюдаемых ранговых статистик при выполнении гипотезы H_0 отсутствия воздействий можно рассчитать по значениям вероятностей $p(\gamma)$ плана получения данных о воздействиях. Однако для некоторых планов распределения наблюдаемых ранговых статистик можно рассчитать и при $n_1 + n_2 < N$. Так будет для плана двух случайных выборок, рассмотренного Рыковой [3]. Пусть $F_{n_j}(x)$ — э. ф. р., построенные по значениям $X_i \in s_j(\gamma_\omega)$. Статистика

$$\rho_{n_1, n_2, N}(\gamma) = \max_x |F_{n_1}(x) - F_{n_2}(x)|$$

является наблюдаемой ранговой статистикой для плана двух случайных выборок объемов n_1, n_2 . Распределение статистики $\rho_{n_1, n_2, N}(\gamma)$ не зависит от фактических значений X_i . В частности, распределение $\rho_{n_1, n_2, N}(\gamma)$ не изменится, если X_i считать взаимно независимыми, одинаково распределенными случайными величинами. Однако при таком предположении значения X_i , включенные в выборку, являются взаимно независимыми и одинаково распределенными. Поэтому ф. р. статистики $\rho_{n_1, n_2, N}(\gamma)$ не зависит от N и совпадает с ф. р. в схеме двух взаимно независимых выборок Н. В. Смирнова. Аналогичная редукция к схеме двух выборок Н. В. Смирнова имеет место и для других ф. р. наблюдаемых ранговых статистик.

Заметим, что рассчитать до испытаний ф. р. наблюдаемых ранговых статистик можно далеко не для всех планов получения данных о воздействиях. План должен обладать определенными свойствами симметричности.

Определение 6.2.1. План Π получения данных назовем **симметричным**, если он реализуется следующим

образом. Сначала по некоторому правилу отбираются объекты из $\mathfrak{F}$ в общую выборку объема n , а затем из отобранных объектов берется случайная выборка объема n_1 , объекты которой подвергаются воздействиям первого типа, а оставшиеся $n_2 = n - n_1$ объектов общей выборки подвергаются воздействиям второго типа. Числа n_1 , n_2 , n могут быть случайными.

Например, отбор объектов в общую выборку можно организовать в соответствии со значениями цепи Маркова $\eta_t = 0, 1$, $t = 1, \dots, N$. Значения t , при которых $\eta_t = 1$, соответствуют номерам объектов, отбираемых в общую выборку. Имея в виду другие возможные методы обработки результатов измерений, отбор в общую выборку следует организовать так, чтобы для любого объекта $O_i \in \mathfrak{F}$ вероятность включения в общую выборку была бы положительна. Условные ф. р. наблюдаемых ранговых статистик при фиксированных значениях n_1 , n_2 для симметричных планов совпадают с ф. р. эквивалентным им наблюдаемым ранговым статистикам планов двух случайных выборок объемов n_1 , n_2 . Если ф. р. наблюдаемой ранговой статистики известна, то, определив критические уровни, можно эту статистику использовать как тестовую для построения критериев выявления воздействий. Заметим, что изложенный выше подход к выявлению воздействий на примере статистики $\rho_{n_1, n_2, N}(\gamma)$ был предложен Рыковой [3]. В ее статье фактически рассматривается математическая модель, в которой считается известной информация о значениях X_i для объектов, не подвергаемых воздействиям, а данные измерений имеют вид

$$\mathfrak{X}_{n_1, n_2} = \{X_i^{(1)}, i \in s_1(\gamma_\omega), X_i^{(0)}, i \in s_1(\gamma_\omega) \cup s_2(\gamma_\omega)\}.$$

Такая математическая модель допускает несколько интерпретаций. Можно представить, что данные $\mathfrak{X}_{n_1, n_2}$ получены в результате трех контрольных проверок совокупности хранимых объектов. При проверках последовательно реализуются три случайные выборки без возвращения объемов n_1 , n_2 , $N - n_1 - n_2$. Аналогичная ситуация возникает, когда проверяются воздействия одного фактора на трех уровнях: нулевом, первом и втором. Эту модель можно также интерпретировать как

одновременное сравнение трех методов лечения, испытаний на надежность трех вариантов элементов и т. п.

В условиях отсутствия воздействий Рыковой [3] были исследованы предельные распределения наблюдаемых ранговых статистик, определяемых на траекториях процессов $F_{n_i}(x) - F_N(x)$, где $F_N(x)$ — ф. р., соответствующая всей совокупности. Используемые при этом методы близки к рассмотренным выше в этом параграфе, однако их реализация существенно сложнее, так как исследуется векторный процесс

$$\beta_{n_1, n_2, N}(t) = \{\beta_{n_1, n_2, N}^{(1)}(t), \beta_{n_1, n_2, N}^{(2)}(t)\}, \quad 0 \leq t \leq 1,$$

где

$$\beta_{n_1, n_2, N}^{(j)}\left(\frac{l}{N}\right) = \sqrt{\frac{N-1}{N} \cdot \frac{n_1 n_2}{n_1 + n_2}} \left[\frac{\kappa_j(l)}{n_j} - \frac{l}{N} \right], \quad (2.91)$$

$$j = 1, 2,$$

для значений $t = l/N$, а для промежуточных значений $t \in \left(\frac{l}{N}, \frac{l+1}{N}\right)$, $0 \leq l \leq N-1$, $\beta_{n_1, n_2, N}(t)$ доопределяется путем линейной интерполяции. В (2.91) $\kappa_j(l)$ равно числу $X_i^{(j)} \leq x$, $i \in s_j$. При $n_j \rightarrow \infty$, $v_j = \lim_{n_j \rightarrow \infty} \frac{n_j}{N} < 1$, $j = 1, 2$, все конечномерные распределения $\beta_{n_1, n_2, N}(t)$ сходятся к распределениям двумерного гауссовского процесса $(\beta_0^{(1)}(t), \beta_0^{(2)}(t))$ с нулевым средним и матрицей ковариаций $E\beta_0^{(j_1)}(t)\beta_0^{(j_2)}(s)$, $t \leq s$, $j_1, j_2 = 1, 2$, равной

$$\left\| \begin{array}{cc} \frac{v_2(1-v_1)t(1-s)}{v_1+v_2} & -\frac{v_1v_2t(1-s)}{v_1+v_2} \\ -\frac{v_1v_2t(1-s)}{v_1+v_2} & \frac{v_1(1-v_2)t(1-s)}{v_1+v_2} \end{array} \right\|. \quad (2.92)$$

Из (2.92) следует, что конечномерные распределения $\beta_{n_1, n_2, N}^{(1)}(t) - \beta_{n_1, n_2, N}^{(2)}(t)$ сходятся к распределениям условного винеровского процесса $\beta_0(t)$. Так как для компонент $\beta_{n_1, n_2, N}(t)$ справедливы утверждения теорем 6.2.2 и 6.2.4, то распределения всех непрерывных функционалов, определенных на реализациях $(F_{n_1}(x), F_{n_2}(x))$ и инвариантных относительно монотонных преобразований

переменной x , сходятся к распределениям соответствующих функционалов, заданных на реализациях $(\beta_j^{(1)}(t), \beta_j^{(2)}(t))$. Заметим, что из (2.92) легко выводится соотношение $\beta_0^{(2)}(t) - \beta_0^{(1)}(t) = \beta_0(t)$, где $\beta_0(t)$ — «броуновский мост». Таким образом, распределение статистики

$$\sqrt{\frac{n_1 n_2}{n_1 + n_2}} \max_x |F_{n_1}(x) - F_{n_2}(x)|$$

при $n_i \rightarrow \infty$ сходится к распределению Колмогорова. Однако этот факт можно получить и в рамках одномерной теории, изложенной в начале этого параграфа.

§ 6.3. Планы выборочного обследования

Основные определения. Рассматривая оценку общего среднего конечной совокупности $\mathfrak{F}$ в § 6.1 и методы построения доверительных полос для общей функции распределения $F_N(x)$ в § 6.2, мы ограничивались случайной выборкой без возвращения фиксированного объема n . Для тех же целей можно использовать и другие планы выборочного обследования, в которых объем n является случайной величиной, а вероятности отбора объектов $O_i \in \mathfrak{F}$ зависят от i . В этом параграфе, исходя из результатов Годамба и Джоши, мы сформулируем основные понятия для широкого класса планов выборочного обследования и докажем некоторые результаты о возможности использования различных оценок. Материал параграфа, однако, не содержит обзора всей области исследований по выборкам из конечных совокупностей.

Пусть $\mathfrak{F} = [O_1, \dots, O_N]$ — конечная совокупность, объекты O_i которой характеризуются скалярными величинами X_i , $1 \leq i \leq N$. Вектор $X = (X_1, \dots, X_N)$ можно рассматривать как точку в N -мерном евклидовом пространстве R^N . Любое подмножество $s \subset I_N = [1, \dots, N]$ отождествим с выборкой из $\mathfrak{F}$ элементов O_i , $i \in s$. Объем выборки $n(s)$ — это число ее элементов. Пусть S — множество всех подмножеств I_N . Тогда S соответствует множеству всех возможных выборок из $\mathfrak{F}$. На S можно задать вероятностную меру P , сопоставив каждому s число $p(s) \geq 0$, $\sum_{s \in S} p(s) = 1$. Таким образом, вероят-

ностное пространство $[S, P]$ определяет выборочный план $\Pi = \{p(s)\}$. Для простой случайной выборки без возвращения объема n вероятности $p(s) = 0$, если $n(s) \neq n$, тогда как для всех s таких, что $n(s) = n$, $p(s) = \binom{N}{n}^{-1}$. Для заданного плана Π вероятность того, что объект O_i или пара O_i, O_j попадут в выборку, равны соответственно

$$\pi_i = \sum_{s \ni O_i} p(s), \quad \pi_{i,j} = \sum_{s \ni O_i, O_j} p(s). \quad (3.1)$$

Здесь суммирование ведется по всем выборкам s , содержащим O_i в первой сумме и пару объектов O_i, O_j — во второй. Величины $\pi_i, \pi_{i,j}$ называются **вероятностями включения**. Так как для случайной выборки без возвращения объема n число выборок, содержащих O_i , равно $\binom{N-1}{n-1}$, то согласно (3.1) вероятность включения $\pi_i = n/N$. Заметим, что средний объем выборки для плана $\Pi = \{p(s)\}$ равен

$$n_\Pi = \sum_{s \in S} n(s) p(s) = \sum_{i \in I} \pi_i. \quad (3.2)$$

Определение 6.3.1. *Оценкой при выборочном обследовании $\mathfrak{F}$ называется функция $e(s, X)$, определенная на прямом произведении $S \times R^N$, которая зависит только от значений $X_i, i \in s$.*

В работах по выборкам из конечных совокупностей большое внимание уделяется оценкам общей суммы $T(X) = \sum_{i \in I} X_i$. Оценки вида

$$e(s, X) = \sum_{i \in s} b(s, i) X_i, \quad (3.3)$$

где $b(s, i)$ зависит от набора s номеров объектов, составляющих выборку, $b(s, i) = 0, i \notin s$, называются **линейными оценками**. Пусть $\mathfrak{L}$ — класс всех линейных оценок (3.3).

Класс несмещенных оценок $T(X)$ обозначим $\mathfrak{U}_T$, класс несмещенных линейных оценок T обозначим $\mathfrak{U}_T \mathfrak{L}$. Условие несмещенности оценки e общей суммы $T(X)$ имеет вид

$$\sum_{s \in S} e(s, X) p(s) = T(X). \quad (3.4)$$

Лемма 6.3.1 (Горвиц, Томпсон [1], Годамб, Джоши [1]). Если для плана Π все $\pi_i > 0$, то линейная оценка вида

$$e(s, X) = \sum_{i \in s} X_i / \pi_i \quad (3.5)$$

является несмещенной оценкой $T(X)$. Ее дисперсия равна

$$V(e, X) = \sum_{1 \leq i \leq N} \frac{X_i^2}{\pi_i} + 2 \sum_{1 \leq i < j \leq N} \pi_{i,j} \frac{X_i X_j}{\pi_i \pi_j} - T^2(X). \quad (3.6)$$

Если все $\pi_{i,j} > 0$ ($i \neq j$), то несмещенная оценка для дисперсии (3.6) имеет вид

$$\begin{aligned} \hat{V}(e, X) = \\ = \sum_{i \leq s} \frac{X_i^2}{\pi_i} \left(\frac{1}{\pi_i} - 1 \right) + \sum_{i \neq j} \frac{X_i X_j}{\pi_{i,j}} \left(\frac{\pi_{i,j}}{\pi_i \pi_j} - 1 \right). \end{aligned} \quad (3.7)$$

Доказательство. Несмещенность получаем, учитывая (3.1), прямым вычислением среднего

$$E e(s, X) = \sum_{s \ni S} \sum_{i \in s} \frac{X_i}{\pi_i} p(s) = \sum_{1 \leq i \leq N} \frac{X_i}{\pi_i} \sum_{s \ni O_i} p(s) = T(X).$$

Дисперсия оценки $e(s, X)$ равна

$$\begin{aligned} V(e, X) &= E \left(\sum_{i \in s} X_i / \pi_i \right)^2 - T^2(X) = \\ &= \sum_{s \ni S} \left(\sum_{i \in s} X_i^2 / \pi_i^2 \right) p(s) + \\ &+ 2 \sum_{s \ni S} \left(\sum_{\substack{i, j \in s \\ i \neq j}} X_i X_j / \pi_i \pi_j \right) p(s) - T^2(X) = \\ &= \sum_{1 \leq i \leq N} (X_i^2 / \pi_i^2) \sum_{s \ni O_i} p(s) + \\ &+ 2 \sum_{1 \leq i < j \leq N} (X_i X_j / \pi_i \pi_j) \sum_{s \ni O_i, O_j} p(s) - T^2(X), \end{aligned}$$

что с учетом (3.1) соответствует (3.6). Несмещенность оценки (3.7) для дисперсии (3.6) легко проверить непосредственными вычислениями.

Несмещенные оценки (3.5) были введены в работе Горвица — Томпсона [1]. Заметим, что при случайной вы-

борке без возвращения объема n выборочное среднее $\bar{x}_n = n^{-1} \sum_{i \in s} X_i$ есть оценка Горвица — Томпсона для общего среднего $N^{-1}T(X)$.

Выборочное обследование конечной совокупности $\mathfrak{P}$ можно планировать разными способами. Для этого надо определить вероятности $p(s)$ плана Π . Если план Π задан, то для оценивания характеристик совокупности $\mathfrak{P}$, например, $T(X)$, можно использовать различные оценки.

Центральными проблемами выборочного обследования являются отыскание рационального плана выборочного контроля и отбор при заданном плане наилучших оценок.

Допустимость оценок и планов. Показателем качества оценки e при плане обследования Π можно считать среднеквадратичное уклонение

$$V_{\Pi}(e, X) = \sum_{s \in S} [e(s, X) - T(X)]^2 p(s), \quad (3.8)$$

которое в случае несмещенности оценки e совпадает с ее дисперсией.

Определение 6.3.2. Оценка e называется **недопустимой при плане Π относительно класса оценок C** , если найдется оценка $e_1 \in C$ такая, что для всех $X \in R^N$

$$V_{\Pi}(e_1, X) \leq V_{\Pi}(e, X), \quad (3.9)$$

причем хотя бы при одном X имеет место строгое неравенство. В противном случае оценка e называется **равномерно (по X) допустимой при плане Π относительно класса C** .

В качестве класса C можно взять класс $\mathfrak{A}_T$ всех несмещенных оценок или класс $\mathfrak{A}_{T\mathfrak{P}}$ линейных несмещенных оценок. Следующая теорема принадлежит Годамбу и Джоши [1].

Теорема 6.3.1. Оценка Горвица — Томпсона (3.5) является допустимой относительно класса $\mathfrak{A}_T$ при любом плане Π , для которого все $\pi_i > 0$.

Доказательство. Пусть оценка e , определяемая (3.5), недопустима. Следовательно, существует такая оценка $e_1(s, X) \in \mathfrak{A}_T$, не совпадающая с $e(s, X)$, что для всех $X \in R^N$ выполнено (3.9) со строгим неравенством

хотя бы для одного $X \in R^N$. Тождество

$$e_1(s, X) = e(s, X) + [e_1(s, X) - e(s, X)] = e(s, X) + h(s, X)$$

определяет функцию $h(s, X)$, которая зависит только от X_i , $i \in s$. Из несмещенности оценок $e(s, X)$, $e_1(s, X)$ следует, что

$$\sum_{s \in S} h(s, X) p(s) = 0. \quad (3.10)$$

Так как для несмещенной оценки из (3.8) вытекает, что

$$V_{\Pi}(s, X) = \sum_{s \in S} e^2(s, X) p(s) - T^2(X), \quad (3.11)$$

то из (3.11) получаем, что неравенство (3.9) эквивалентно

$$\sum_{s \in S} e_1^2(s, X) p(s) \leq \sum_{s \in S} e^2(s, X) p(s), \quad (3.12)$$

с выполнением строгого неравенства хотя бы при одном $X \in R^N$. Подставляя в (3.12) выражение для $e_1(s, X)$, находим

$$\sum_{s \in S} h^2(s, X) p(s) \leq -2 \sum_{s \in S} h(s, X) e(s, X) p(s). \quad (3.13)$$

Пусть $R^N(K) \subset R^N$ образовано точками X , у которых не более, чем K координат ($K < N$) отличны от нуля, $R^N(N) = R^N$.

Лемма 6.3.2. Если $h(s, X) p(s) = 0$ для всех (s, X) таких, что $s \in S$, $X \in R^N(K)$, то $h(s, X) p(s) = 0$ для всех (s, X) , $s \in S$, $X \in R^N(K+1)$.

Доказательство. Пусть S_j состоит из тех и только тех выборок s , в которые входят ровно j объектов, имеющих $X_i \neq 0$. Так как при $X^* = (X_1^*, \dots, X_N^*) \in R^N(K+1)$ очевидно $\bigcup_{1 \leq j \leq K+1} S_j = S$, то соотношения (3.10) и (3.12) принимают вид

$$\sum_{0 \leq j \leq K+1} \sum_{s \in S_j} h(s, X^*) p(s) = 0, \quad (3.14)$$

$$\begin{aligned} \sum_{0 \leq j \leq K+1} \sum_{s \in S_j} h^2(s, X^*) p(s) &\leq \\ &\leq -2 \sum_{0 \leq j \leq K+1} \sum_{s \in S_j} h(s, X) e(s, X) p(s). \end{aligned} \quad (3.15)$$

Из (3.14), (3.15) и условий леммы имеем

$$\sum_{s \in S_{K+1}} h(s, X^*) p(s) = 0, \quad (3.16)$$

$$\begin{aligned} \sum_{s \in S_{K+1}} h^2(s, X^*) p(s) &\leq \\ &\leq -2 \sum_{s \in S_{K+1}} h(s, X^*) e(s, X^*) p(s). \end{aligned} \quad (3.17)$$

Так как у X^* не более $K+1$ координат отличны от нуля, то для всех $s \in S_{K+1}$ несмещенная оценка $e(s, X^*)$ есть постоянная, равная $\sum_{1 \leq i \leq N} X_i^* / \pi_i$. Учитывая это, из (3.10), (3.17) и (3.16) находим

$$\begin{aligned} \sum_{s \in S_{K+1}} h^2(s, X^*) p(s) &\leq \\ &\leq -2 \left(\sum_{1 \leq i \leq N} X_i^* / \pi_i \right) \sum_{s \in S_{K+1}} h(s, X^*) p(s) = 0. \end{aligned} \quad (3.18)$$

Величины $h^2(s, X^*) \geq 0$, $p(s) \geq 0$, поэтому из (3.18) следует, что для всех $s \in S_{K+1}$, $h(s, X^*) p(s) = 0$.

Так как $e(s, X) = 0$ при $X \in R^N(0)$ ($X = 0$), то из (3.13) вытекает, что $h(s, X) p(s) = 0$ для всех $s \in S$ и $X \in R^N(0)$. Следовательно, из леммы 6.3.2 находим, что $h(s, X) p(s) = 0$ для всех $s \in S$ и $X \in R^N$, $K+1 = N$. Таким образом, $e(s, X) = e_1(s, X)$, что противоречит исходному предположению. Теорема доказана.

Приведем пример, взятый из работы Годамба и Джоши [1], иллюстрирующий понятия плана и допустимости оценки. Пусть $\mathfrak{P} = [O_1, O_2, O_3]$, возможные выборки $s_i = \{i\}$, $i = 1, 2, 3$, $s_4 = \{O_1, O_2\}$, $s_5 = \{O_1, O_3\}$, $s_6 = \{O_2, O_3\}$ имеют одинаковые вероятности $p(s_i) = 1/6$, $1 \leq i \leq 6$. В качестве оценки общей суммы $T(X) = X_1 + X_2 + X_3$ возьмем линейную несмещенную оценку

$$e(s, X) = \sum_{1 \leq i \leq 3} b(s, i) X_i, \quad (3.19)$$

для которой $b(s_1, 2) = b(s_5, 1) = b(s_6, 3) = 0$, $b(s_4, 1) = 6 - b(s_1, 1)$, $b(s_5, 3) = 6 - b(s_3, 3)$, $b(s_6, 2) = 6 - b(s_2, 2)$. В соответствии с (3.6) имеем

$$\begin{aligned} V(e, X) &= 1/6 \sum_{1 \leq i \leq 3} X_i^2 \{b^2(s_i, i) + [6 - b(s_i, i)]^2\} - T^2(X) = \\ &= 1/6 \sum_{1 \leq i \leq 3} X_i^2 \{[b(s_i, i) - 3]^2 + 9\} - T^2(X). \end{aligned} \quad (3.20)$$

Наименьшее значение V при любом X достигается при $b(s_i, i) = 3$, $1 \leq i \leq 3$, что легко следует из (3.20). Таким образом, при $b(s_i, O_i) \neq 3$, $1 \leq i \leq 3$, оценка (3.19) недопустима. В рассматриваемом примере $\pi_i = 1/2$, $1 \leq i \leq 3$. Поэтому согласно доказанной теореме 6.3.1 допустимой оценкой будет несмещенная оценка Горвица — Томпсона

$$e(s, X) = 2 \sum_{i \in s} X_i.$$

В работе Джоши [1] доказано, что оценка

$$e(s, X) = \frac{N}{n(s)} \sum_{i \in s} X_i \quad (3.21)$$

допустима при любом плане Π относительно класса всех оценок. Заметим, что при случайной выборке без возвращения (3.21) является оценкой Горвица — Томпсона и совпадает с обычно рекомендуемой оценкой общей суммы.

В работе Годамба [1] дано определение одновременной допустимости оценки и плана. Пусть заданы два плана $\Pi = \{p(s)\}$ и $\Pi' = \{p'(s)\}$. При плане Π используется оценка $e(s, X)$ для $T(X)$, а при плане Π' — оценка $e'(s, X)$. Пара (Π', e') называется **равномерно предпочтительнее пары** (Π, e) , если для всех $X \in R^N$

$$\begin{aligned} \sum_{s \in S} [e'(s, X) - T(X)]^2 p'(s) &\leq \\ &\leq \sum_{s \in S} [e(s, X) - T(X)]^2 p(s), \end{aligned} \quad (3.22)$$

причем хотя бы для одного $X \in R^N$ в (3.22) имеет место строгое неравенство.

Определение 6.3.3. Пара (Π, e) называется **допустимой в классе пар** C , если не существует пары $(\Pi', e') \in C$ такой, что (Π', e') равномерно предпочтительнее (Π, e) .

В качестве множества планов Π в парах (Π, e) Годамб [1] рассматривает планы с фиксированным средним объемом выборки, т. е. такие $\Pi = \{p(s)\}$, что $\sum_{s \in S} n(s) p(s) = n$. Кратко класс пар (Π, e) планов Π со

средним объемом n и оценок e обозначим C_n . Годамб показал, что в классе C_n допустимы пары с оценками

$$e(s, X) = \sum_{i \in s} X_i + \sum_{i \notin s} \lambda_i, \quad (3.23)$$

где λ_i — любые числа, заранее сопоставленные всем объектам $O_i \in \mathfrak{F}$, $1 \leq i \leq N$. Эриксон [1] рассмотрел допустимость пар (Π, e) в более узком классе D_n планов фиксированного объема n , т. е. планов $\Pi = \{p(s)\}$, $p(s) = 0$, если $n(s) \neq n$. Эриксон показал, что в классе D_n допустимы пары, в которых $e(s, X) = \alpha_n \sum_{i \in s} X_i + \beta_n$, где $1 < \alpha_n < N/n$, β_n — любые фиксированные числа.

Джоши [1] ввел понятие допустимости плана $\Pi = \{p(s)\}$, когда в качестве оценок $T(X)$ берутся оценки Горвица — Томпсона (3.5). Пусть n_Π — средний объем плана Π . Обозначим R_+^N множество X с положительными координатами. При использовании оценки (3.5) план Π' называется **равномерно предпочтительнее плана Π** , если: (1) $n_{\Pi'} \leq n_\Pi$, (2) $V_{\Pi'}(e, X) \leq V_\Pi(e, X)$ для всех $X \in R_+^N$; при этом строгое неравенство либо в (1) либо в (2) выполняется хотя бы для одного $X \in R_+^N$.

Определение 6.3.4. План Π **допустим**, если не существует плана Π' , который **равномерно предпочтительнее Π** при использовании оценок (3.5).

Джоши [3] показал, что план Π недопустим, если существует пара выборок s_1, s_2 , $p(s_1) > 0$, $p(s_2) > 0$, таких, что $s_1 \subset s_2$, $n(s_1) \leq n(s_2) - 2$.

Интересные качественные результаты о допустимости пар (план, оценка) содержатся в статье Скотта [1]. В статье Секкаппана и Томпсона [1] при слабых ограничениях показана допустимость линейных оценок (3.3),

у которых $b(s, i) = b_i > 1$, $i = 1, 2, \dots, N$, $\sum_{i=1}^N b_i^{-1} = n_\Pi$, а также допустимость соответствующих пар (план, оценка).

Оценка разности средних при воздействиях. В конце § 6.2 была сформулирована задача выявления различия воздействий разных типов на величину скалярной характеристики объектов некоторой совокупности $\mathfrak{F}$. Если такое различие имеется, то представляют интерес задачи оценки разности средних, которые соответствуют

двум типам воздействий. Ниже оценку получим для широкого класса планов выборочного контроля.

Итак, пусть объектам $O_i \in \mathfrak{F}$ соответствуют скалярные величины $X_i^{(j)}$, если на объект O_i оказывается воздействие $j = 1, 2$. Значения $X_i^{(j)}$ становятся известны в результате контрольной операции. Напомним существенное ограничение, заключающееся в том, что объект может быть подвергнут воздействию либо только первого типа ($j = 1$), либо только второго ($j = 2$). Рассматриваемая нами задача состоит в оценке разности общих средних $\bar{X}^{(1)} - \bar{X}^{(2)}$, $\bar{X}^{(j)} = N^{-1} \sum_{1 \leq i \leq N} X_i^{(j)}$, $j = 1, 2$.

Такие задачи возникают при оценке эффективности воздействий (изменений) на среднее значение времени безотказной работы (элементов, блоков и т. п.) при конструировании и последующих испытаниях на безотказность.

План Π выборочного контроля определяется заданием вероятностей $P(\gamma)$ векторов $\gamma = (\gamma_1, \dots, \gamma_N)$, $\gamma_i = 0, 1, 2$. Как и выше в § 6.2, $\gamma_i = j$ соответствует при $j = 1, 2$ отбору на контроль объекта O_i , подвергаемого воздействию j -го типа; при $\gamma_i = 0$ объект не контролируется. Таким образом, наборы индексов i , определяемые соотношениями $s_j = \{i: \gamma_i = j\}$, $j = 1, 2$ соответствуют номерам объектов, отбираемых для контроля с воздействием j -го типа. Событие $\{s_j \ni i\}$ означает, что объект O_i включен в выборку s_j . Для вероятностей включений используем обозначения $\pi_i(j) = P\{s_j \ni i\}$, $\pi_{k,l}(m, r) = P\{s_m \ni k, s_r \ni l\}$, $i, k, l = 1, \dots, N$; $j, m, r = 1, 2$. В этих обозначениях невозможность одновременного наблюдения пар значений $(X_i^{(1)}, X_i^{(2)})$ означает, что $\pi_{i,i}(1, 2) = 0$, $1 \leq i \leq N$, т. е. что выборки s_1, s_2 не содержат общих элементов ($s_1 \cap s_2 = \emptyset$). Если для плана Π вероятности $\pi_i(j) > 0$, $j = 1, 2$, $i = 1, \dots, N$, то из леммы 6.3.1 получаем

Следствие 6.3.1. Несмещенными оценками общих средних $\bar{X}^{(1)}, \bar{X}^{(2)}$ при воздействиях первого и второго типов и использовании плана Π , для которого $\pi_i(j) > 0$, будут

$$\hat{\bar{X}}^{(j)} = N^{-1} \sum_{i \in s_j} X_i \pi_i^{-1}(j), \quad j = 1, 2. \quad (3.24)$$

Соответственно несмещенная оценка для $\bar{X}^{(1)} - \bar{X}^{(2)}$ равна $\hat{\bar{X}}^{(1)} - \hat{\bar{X}}^{(2)}$. Если все $\pi_{k,l}(j, j) > 0$, то несмещенная оценка для дисперсии $V(\hat{\bar{X}}^{(j)})$ оценки $\hat{\bar{X}}^{(j)}$ имеет вид

$$\hat{V}^{(j)} = \frac{1}{N^2} \left[\sum_{i \in s_j} \frac{(X_i^{(j)})^2}{\pi_i(j)} \left(\frac{1}{\pi_i(j)} - 1 \right) + \sum_{\substack{k \in s_j, l \in s_j \\ k \neq l}} \frac{X_k^{(j)} X_l^{(j)}}{\pi_{k,l}(j, j)} \left(\frac{\pi_{k,l}(j, j)}{\pi_k(j) \pi_l(j)} - 1 \right) \right]. \quad (3.25)$$

Для оценки значимости разности $\hat{\bar{X}}^{(1)} - \hat{\bar{X}}^{(2)}$ желательно иметь оценку для ее дисперсии $V(\hat{\bar{X}}^{(1)} - \hat{\bar{X}}^{(2)})$. Однако получить несмещенную оценку для $V(\hat{\bar{X}}^{(1)} - \hat{\bar{X}}^{(2)})$ не удастся. Поэтому мы ограничимся следующей рекомендацией. При любой корреляции статистик $\hat{\bar{X}}^{(1)}, \hat{\bar{X}}^{(2)}$ имеем

$$V(\hat{\bar{X}}^{(1)} - \hat{\bar{X}}^{(2)}) \leq [(V(\hat{\bar{X}}^{(1)}))^{1/2} + (V(\hat{\bar{X}}^{(2)}))^{1/2}]^2. \quad (3.26)$$

В качестве оценки правой части (3.26) можно взять

$$v^2(s_1, s_2, X) = [(\hat{V}^{(1)})^{1/2} + (\hat{V}^{(2)})^{1/2}]^2, \quad (3.27)$$

где $\hat{V}^{(j)}$ — несмещенные оценки для $V(\hat{\bar{X}}^{(j)})$, которые находятся по формуле (3.25).

Можно высказать гипотезу, что при не слишком ограничительных предположениях разность $\hat{\bar{X}}^{(1)} - \hat{\bar{X}}^{(2)}$ имеет нормальное распределение со средним $\bar{X}^{(1)} - \bar{X}^{(2)}$ и дисперсией $V(\hat{\bar{X}}^{(1)} - \hat{\bar{X}}^{(2)})$. Тогда приближенные выражения для доверительных интервалов разности $\bar{X}^{(1)} - \bar{X}^{(2)}$, имеющих нижнюю границу коэффициента доверия, не меньшую $\gamma = 1 - \varepsilon$, записываются в следующем виде:

$$\begin{aligned} \hat{\bar{X}}^{(1)} - \hat{\bar{X}}^{(2)} - u_{\varepsilon/2} \sqrt{v(s_1, s_2, X)} &\leq \bar{X}^{(1)} - \bar{X}^{(2)} \leq \\ &\leq \hat{\bar{X}}^{(1)} - \hat{\bar{X}}^{(2)} + u_{\varepsilon/2} \sqrt{v(s_1, s_2, X)}, \end{aligned}$$

где $u_{\varepsilon/2}$ — квантиль стандартного нормального распределения, соответствующий уровню $\varepsilon/2$.

Заметим, что для плана Π , у которого s_1, s_2 получаются путем последовательного случайного выбора без возвращения n_1 и n_2 объектов из N ($n_1 + n_2 \leq N$), вероятности включения равны:

$$\pi_i(j) = n_j/N, \quad \pi_{k,l}(m, r) = \frac{n_m n_r}{N(N-1)}, \quad m \neq r,$$

$$\pi_{k,l}(m, m) = \frac{n_m(n_m - 1)}{N(N-1)}.$$

Расслоение. Еще одно важное направление исследований связано с поиском информативных параметров. Мы ограничимся лишь частным случаем использования расслоения исследуемой совокупности $\mathfrak{F}$ с целью уменьшения дисперсии оценок.

Определение 6.3.5. Разбиение совокупности $\mathfrak{F}$ на непересекающиеся подмножества объектов $\mathfrak{F}_l$, $\mathfrak{F} = \bigcup \mathfrak{F}_l$, $\mathfrak{F}_l \cap \mathfrak{F}_j = \emptyset$, $l \neq j$, $l, j = 1, \dots, r$, называется **расслоением**, а $\mathfrak{F}_l$ — **слоями**. План $\Pi = (\Pi_1, \dots, \Pi_r)$ соответствует расслоению $\mathfrak{F} = \bigcup_{1 \leq l \leq r} \mathfrak{F}_l$, если выборочный контроль слоя $\mathfrak{F}_l$ проводится на основе плана Π_l . При этом отбор объектов для контроля из каждого слоя производится независимо.

Расслоение выборки $\mathfrak{F}$ проводят, используя дополнительную информацию. Приведем одну из общих схем расслоения. Пусть каждый объект $O_i \in \mathfrak{F}$ характеризуется вектором (X_i, Y_i) , где значения Y_i — известны, а X_i могут быть измерены лишь на основе выборочного контроля. Разбивая множество значений Y_i на непересекающиеся подмножества G_j , $1 \leq j \leq r$, образуем слои $\mathfrak{F}_j = \{O_i: Y_i \in G_j\}$. Расслоение может дать существенный выигрыш в точности оценки общего среднего, если значения X_i для объектов O_i , входящих в один слой, значительно отличаются друг от друга.

Проиллюстрируем выигрыш от расслоения следующим простым примером. Пусть объекты, составляющие $\mathfrak{F}$, имеют скалярные характеристики X_i . При этом величины X_i могут быть равны только либо a_1 либо a_2 . Общий объем совокупности $\mathfrak{F}$ равен $N = N_1 + N_2$, где N_j — число O_i , имеющих $X_i = a_j$, $j = 1, 2$. Оценку общего среднего $\bar{X} = N^{-1} \sum_{1 \leq i \leq N} X_i$ можно получить,

используя случайную выборку без возвращения объема n . При этом согласно (3.5) несмещенная оценка $\hat{X} = N^{-1} \sum_{i \in s} X_i (n/N)^{-1} = n^{-1} \sum_{i \in s} X_i$, где s — номера объектов, отобранных для контроля, $n(s) = n$. На основе (3.6) и (3.7) легко подсчитать величину дисперсии этой оценки и найти оценку дисперсии оценки. Оценку общего среднего можно получить другим способом. Пусть нам для каждого объекта O_i известно значение Y_i . Предположим, что при $Y_i < 0$ значение $X_i = a_1$, а при $Y_i \geq 0$ значение $X_i = a_2$. Образуем два слоя: $\mathfrak{P}_1 = \{O_i: Y_i < 0\}$, $\mathfrak{P}_2 = \{O_i: Y_i \geq 0\}$. В качестве плана, соответствующего расслоению $\mathfrak{P} = \mathfrak{P}_1 \cup \mathfrak{P}_2$, возьмем случайные выборки без возвращения объемов n_1 , n_2 , $n_1 + n_2 = n$. Пусть s_j — номера объектов, включенных в j -ю выборку. Тогда несмещенные оценки для $X^{(j)} = \sum_{O_i \in \mathfrak{P}_j} X_i$ — общих сумм слоев, равны $\hat{X}^{(j)} = (N_j/n_j) \sum_{i \in s_j} X_i$. Следовательно, несмещенная оценка для общего среднего $\bar{X}$ имеет вид

$$\hat{\bar{X}} = \frac{N_1}{Nn_1} \sum_{i \in s_1} X_i + \frac{N_2}{Nn_2} \sum_{i \in s_2} X_i. \quad (3.28)$$

Так как для $i \in s_j$ значения $X_i = a_j$, то согласно (3.28) находим, что $\hat{\bar{X}} = (1/N) (N_1 a_1 + N_2 a_2)$. Таким образом, полученная оценка совпадает с оцениваемым значением общего среднего, а ее дисперсия равна нулю. Успех от применения расслоения в этом нарочито упрощенном примере — абсолютный. Выигрыш от введения расслоения является следствием того, что в слоях колебания значений характеристик X_i уменьшаются.

В практическом плане, по-видимому, расслоение надо вести в два этапа. Сначала надо взять предварительную выборку объема n_0 , в результате которой мы получаем информацию вида (X_i, Y_i) , $1 \leq i \leq n_0$. Если по этой информации мы приходим к выводу о целесообразности расслоения на основе значений Y_i , то далее, измеряя значения Y_i для всех остальных $N - n_0$ объектов, проводим расслоение. Например, если результаты предварительной выборки объема n_0 имеют вид, пока-

занный на рис. 6.3.1, а, то можно ожидать, что расслоение на три слоя даст существенный выигрыш. Если же данные имеют такой вид, как на рис. 6.3.1, б, то, по-видимому, расслоение, быть может, даже ухудшит точность оценки среднего.

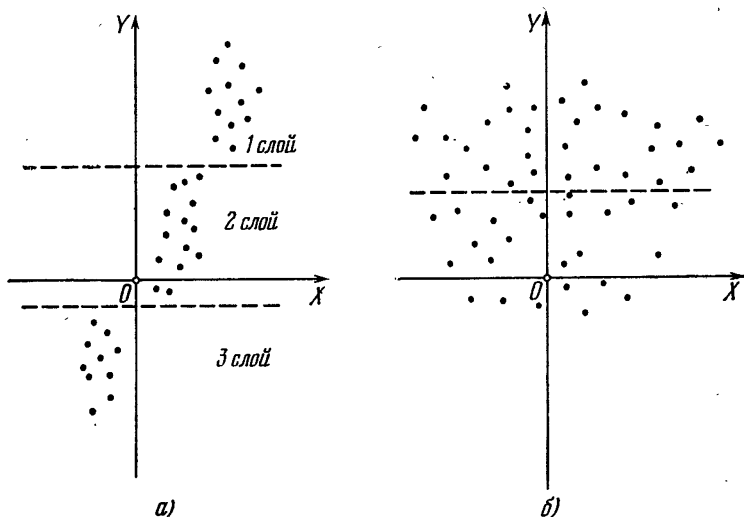


Рис. 6.3.1. Примеры расслоения исследуемой совокупности.

Вариационная диаграмма. В общей проблематике поиска информативных признаков важную роль играют методы исключения «шумовой» информации, т. е. исключение таких данных, которые не связаны с изучаемыми характеристиками объектов. При этом можно использовать вариационную диаграмму, сопоставляемую ряду значений

$$(X_1, Y_1), \dots, (X_N, Y_N), \quad X_i \neq X_j, \quad Y_i \neq Y_j, \quad i \neq j.$$

Для этого образуем вариационные ряды $X_{(1)} < X_{(2)} < \dots < X_{(N)}$, и $Y_{(1)} < Y_{(2)} < \dots < Y_{(N)}$. Если $X_i = X_{(k)}$, $Y_i = Y_{(l)}$, то точке плоскости (X_i, Y_i) сопоставим точку (k, l) целочисленной решетки $\mathbb{R}^2$. При этом множество N несовпадающих точек $\{(X_i, Y_i)\}$ отображается на подмножество $\mathcal{Q}(\mathbb{R})$ точек решетки $\mathbb{R}^2$ квадрата $\mathcal{Q}_N$

со стороны N . $\mathfrak{Q}(\mathfrak{P})$ называется **вариационной диаграммой**. На рис. 6.3.2, *а*, *б*, *в* показаны примеры подмножеств $\mathfrak{Q}(\mathfrak{P})$. Рис. 6.3.2, *а* соответствует прямой монотонной зависимости значений X_i, Y_i , т. е. увеличению X_i при увеличении Y_i . Вариационная диаграмма рис. 6.3.2, *б* получается, когда для половины объектов совокупности $\mathfrak{P}$

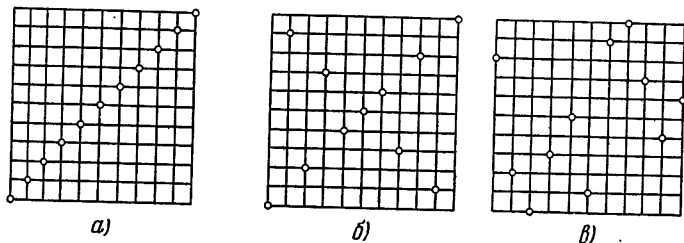


Рис. 6.3.2. Примеры вариационных диаграмм: *а*) при монотонной зависимости характеристик, *б*) для смеси объектов с двумя типами монотонной зависимости, *в*) при взаимной независимости характеристик объектов.

имеет место прямая монотонная зависимость, а для другой половины объектов — обратная монотонная зависимость X от Y .

На рис. 6.3.2, *в* дано изображение диаграммы, полученной реализацией взаимно независимых величин (X_i, Y_i) . Характеристическим свойством вариационных диаграмм является то, что при любом фиксированном значении $k = k_0$ ($l = l_0$) имеется ровно одна точка $(k_0, l) \in \mathfrak{Q}(\mathfrak{P})$ ($(k, l_0) \in \mathfrak{Q}(\mathfrak{P})$). Если $Y_1, \dots, Y_N$ есть последовательность взаимно независимых одинаково распределенных случайных величин, то для любых наборов $X_1, \dots, X_N$ несовпадающих между собой значений все возможные реализации подмножеств $\mathfrak{Q}(\mathfrak{P})$ будут равновероятны. Другими словами, при переходе от «столбца» точек (k, l) к «столбцу» точек $(k+1, l)$, $l \in (1, 2, \dots, N)$, выбор точки, включаемой в $\mathfrak{Q}(\mathfrak{P})$, происходит случайным образом, т. е. с равной вероятностью выбирается одно из еще не «извлекавшихся» значений l . Можно построить критерии проверки случайности вариационной диаграммы. Аналогом гипергеометрического процесса, введенного в § 6.2, будет случайное поле $\kappa(l_1, l_2)$, $(l_1, l_2) \in \mathfrak{Q}(\mathfrak{P})$, равное числу точек $(k, l) \in$

$\in \mathfrak{D}(\mathfrak{P})$ вариационной диаграммы, попавших в прямоугольник, $k \leq l_1$, $l \leq l_2$. Полезно заметить, что при $N \rightarrow \infty$, $(l_i/N) \rightarrow t_i$, случайное поле $\beta_N(l_1/N, l_2/N) = \sqrt{N} \left[\frac{\kappa(l_1, l_2)}{N} - \frac{l_1 l_2}{N^2} \right]$ сходится по распределениям к условному винеровскому полю $\beta(t_1, t_2)$ на плоскости, вырождающемуся в нуль на сторонах квадрата $0 \leq t_1 \leq 1$, $0 \leq t_2 \leq 1$. Ковариационная функция условного винеровского поля равна

$$E\beta(t_1, s_1)\beta(t_2, s_2) = t_1(1-t_2)s_1(1-s_2), \quad t_1 < t_2, \quad s_1 < s_2.$$

Таким образом, можно построить обобщение теории, изложенной в § 6.2, на случай векторных характеристик объектов. К сожалению, ограниченный объем книги и недостаточная проработанность этих вопросов не позволяют сейчас дать полное их изложение.

О некоторых задачах и результатах. Результаты, приведенные в этой главе, показывают, что теория выборок из конечных совокупностей является содержательным разделом математической статистики. Еще раз подчеркнем тот факт, что у обследуемых совокупностей, как правило, не предполагается наличие каких-либо вероятностных свойств. Вероятностные методы используются потому, что отбор объектов из $\mathfrak{P}$ для обследования проводится случайным образом в соответствии с вероятностными характеристиками плана обследования. Обратим внимание на то, что при выборочном обследовании конечных совокупностей часто имеют место утверждения (теоремы, неравенства и т. п.), которые можно рассматривать как обобщения соответствующих утверждений для взаимно независимых случайных величин. Например, для случайных выборок без возвращения получены аналоги известного неравенства Колмогорова (Гаек, Шидак [1], стр. 236, Сен [1]). Получены аналоги и других неравенств для взаимно независимых случайных величин, Карлин [2]. Так в работе Серфлинга [1] приведено неравенство, соответствующее неравенству Хеффдинга [1].

Вообще, рассматривая математическую модель выборочного обследования конечной совокупности $\mathfrak{P}_N$, можно исследовать свойства статистических оценок или некоторых решающих процедур. Например, для плана

$\Pi = \{p(s)\}$ с вероятностями включения $\pi_i > 0$ несмещенная оценка — для общей эмпирической функции $F_N(x) = k(x)/N$ имеет вид

$$\hat{F}_N(x) = \sum_{i \in s} \frac{g_x(X_i)}{\pi_i \cdot N},$$

$$g_x(y) = 1, \quad y \leq x, \quad g_x(y) = 0, \quad y > x,$$

где $k(x)$ — число характеристик X_i объектов из $\mathfrak{F}_N$, для которых $X_i \leq x$. Можно поставить задачи исследования свойств оценок $\hat{F}_N(x)$, когда $N \rightarrow \infty$. Задачи можно сформулировать и таким образом, что при изучении разности $F_N(x) - \hat{F}_N(x)$ станут корректными проблемы, аналогичные усиленному закону больших чисел, закону повторного логарифма и др. При этом надо рассматривать процесс выборочного обследования неограниченной совокупности $\mathfrak{F} = \{O_1, O_2, \dots, O_N, O_{N+1}, \dots\}$, отбор из которой осуществляется в соответствии с реализацией случайной последовательности $\gamma_1, \gamma_2, \dots, \gamma_N, \gamma_{N+1}, \dots$, $\gamma_i = 0, 1$. N -ому шагу ($N = 1, 2, \dots$) соответствует конечная совокупность $\mathfrak{F}_N = \{O_1, O_2, \dots, O_N\} \subset \mathfrak{F}_{N+1} \subset \dots \subset \mathfrak{F}$. Объект O_N обследуется (включается в выборку), тогда когда $\gamma_N = 1$.

Правдоподобно, что параллельно с теорией непараметрической статистики взаимно независимых случайных величин можно построить ее аналог для выборок из конечных совокупностей. Именно в этом плане можно рассматривать непараметрические критерии, обоснованию которых посвящен § 6.2.

Важной областью исследований являются задачи статистического анализа многомерных данных. Методы регрессионного анализа, факторного анализа, планирования эксперимента и т. д. имеют существенные приложения. В связи с этим возникает вопрос о построении аналогов методов многомерного статистического анализа данных для выборочного обследования конечных совокупностей, объекты которых характеризуются векторными величинами. Пусть $O_i \in \mathfrak{F}$ характеризуется набором $(X_i, Y_{i1}, \dots, Y_{ip})$, причем величины Y_{ij} известны для всех $1 \leq i \leq N$, $1 \leq j \leq p$, а X_i известны лишь для объектов, отобранных в выборку. В этих условиях можно

поставить задачу поиска улучшенных оценок для $\sum_{1 \leq i \leq N} X_i$, получаемых на основе данных Y_{ij} , $1 \leq i \leq N$, $1 \leq j \leq p$. Близкая по постановке задача рассмотрена в работе Шриваставы [1].

Этот раздел закончим замечанием о необходимости расширения класса планов выборочного обследования конечных совокупностей, который был введен в начале этого параграфа. Составляющие этот класс планы $\Pi = \{p(s)\}$ имеют вероятности $p(s)$ выборок s , не зависящие от фактически наблюдаемой последовательности значений X_i , $i \in s$. Такое ограничение весьма существенно, так как даже при классификации объектов по альтернативному признаку для планов первого вхождения, рассмотренных выше в главах 2 и 3, реализация выборки s определяется не только методом отбора, но и последовательностью наблюдаемых значений $X_{i_1}, X_{i_2}, \dots$

ПРИЛОЖЕНИЕ

Таблица I

Значения λ_L , ρ_c для расчета предела среднего уровня
выходного качества при использовании одноступенчатых планов
(Гнеденко, Беляев, Соловьев [1])

c	λ_L	ρ_c	c	λ_L	ρ_c
0	1,0000	0,3679	11	8,8194	7,2334
1	1,6180	0,8400	12	9,5900	7,9476
2	2,2695	1,3711	13	10,3660	8,6595
3	2,9452	1,9424	14	11,1471	9,3984
4	3,6395	2,5435	15	11,9328	10,1338
5	4,3490	3,1682	16	12,7228	10,8751
6	5,0712	3,8120	17	13,5169	11,6219
7	5,8041	4,4720	18	14,3147	12,3738
8	6,5464	5,1457	19	15,1160	13,1305
9	7,2970	5,8314	20	15,9206	13,8917
10	8,0549	6,5277			

Таблица II

Планы $(n, c)_{23}$, обеспечивающие минимум инспекции
при заданном предельном выходном качестве (значения θ_c)
(Гнеденко, Беляев, Соловьев [1])

$k_L \backslash c$	0	1	2	3	4
1,25	1,9	6,0	12,8	22,8	36,2
1,5	1,8	6,2	14,3	27,3	46,5
1,75	1,7	6,4	16,0	32,7	60,1
2,0	1,7	6,7	17,8	39,3	77,7
2,25	1,6	7,0	19,9	47,0	100,0
2,5	1,6	7,3	22,2	56,0	127,8
2,75	1,6	7,6	24,6	66,4	162,0
3,0	1,6	8,0	27,2	78,2	203,6
3,5	1,5	8,6	33,0	106,5	312,8
4,0	1,5	9,3	39,4	141,8	464,2
4,5	1,5	10,0	46,6	184,7	667,7
5,0	1,5	10,7	54,4	236,0	934,2
5,5	1,4	11,4	62,9	296,5	1275,7
6,0	1,4	12,1	72,1	366,9	1705,2
7,0	1,4	13,6	92,5	540,6	2885,7
8,0	1,4	15,1	115,5	763,1	4601,0
9,0	1,4	16,5	141,3	1040,4	6994,3
10,0	1,4	18,0	169,7	1378,7	10223,2
12,0	1,4	21,0	234,5	2262,3	19917,0
15,0	1,3	25,4	351,8	4196,8	45696,8

Таблица III

Значения $\lambda_L(\kappa)$, $\theta_{0,1}(\kappa)$ для расчета предела среднего уровня выходного качества при использовании планов двухступенчатого контроля с приемочными числами 0 и 1 (Методика СЭП [1])

[illegible]

Таблица V

Наименьшие значения $M_{0,1}^a$, $M_{0,1}^E$, $M_{0,1}^q$

E c	0,1				0,063			
	γ	$M_{0,1}^a$	$M_{0,1}^E$	$M_{0,1}^q$	γ	$M_{0,1}^a$	$M_{0,1}^E$	$M_{0,1}^q$
0	1,0	198	198	218	1,0	315	315	347
	0,88	211	211	232	0,93	327	327	360
1	0,87	297	117	149	0,92	477	181	230
	0,18	414	276	242	0,157	684	421	369
2	0,179	531	173	185	0,156	862	247	282
	0,0845	647	356	259	0,070	1071	536	390
3	0,084	779	244	219	0,069	1288	332	332
	0,053	898	419	279	0,042	1472	631	421
4	0,052	1041	301	251	0,041	1744	436	380
	0,037	1144	439	306	0,029	1916	730	457
5	0,036	1318	388	285	0,028	2213	544	427
	0,028	1419	574	332	0,022	2361	824	490
6	0,027	1588	477	305	0,021	2674	649	453
	0,023	1657	616	338	0,017	2828	961	514
7	0,022	1852	530	318	0,016	3166	812	488
	0,019	1923	685	350	0,014	3293	1061	531
8	0,018	2124	632	350	0,013	3667	948	536
	0,016	2200	772	379	0,012	3755	1129	565
9	0,015	2414	727	371	0,011	4162	1077	560
	0,014	2461	835	389	0,01	4264	1304	597
10	0,013	2690	814	387				
	0,012	2747	950	410				
11	0,011	2996	970	414				
12	0,01	3263	1031	438				

Продолжение табл. V

$E \backslash c$	0,04				0,025			
	γ	$M_{0,1}^a$	$M_{0,1}^E$	$M_{0,1}^q$	γ	$M_{0,1}^a$	$M_{0,1}^E$	$M_{0,1}^q$
0	1,0	498	498	548	1,0	798	798	878
	0,95	511	511	562	0,97	810	810	891
1	0,94	741	281	358	0,96	1 176	446	568
	0,133	1 093	688	567	0,12	1 643	974	852
2	0,132	1 378	352	433	0,11	2 246	551	679
	0,056	1 706	8'6	594	0,043	2 800	1 284	934
3	0,055	2 107	478	506	0,042	3 438	706	794
	0,032	2 438	968	646	0,024	3 972	1 511	1 007
4	0,031	2 851	611	582	0,023	4 733	911	911
	0,022	3 150	1 108	693	0,016	5 263	1 737	1 086
5	0,021	3 640	770	651	0,015	6 107	1 166	1 026
	0,0'6	3 917	1 272	756	0,0'2	6 492	1 833	1 152
6	0,015	4 484	1 007	702	0,011	7 486	1 413	1 083
	0,013	4 639	1 326	763	0,01	7 621	1 732	1 139
7	0,012	5 268	1 097	731				
	0,01	5 548	1 634	817				

$E \backslash c$	0,016				0,01			
	γ	$M_{0,1}^a$	$M_{0,1}^E$	$M_{0,1}^q$	γ	$M_{0,1}^a$	$M_{0,1}^E$	$M_{0,1}^q$
0	1,0	1 248	1 248	1 373	1,0	1 998	1 998	2 198
	0,98	1 261	1 261	1 387	0,99	2 008	2 008	2 209
1	0,97	1 891	694	883	0,98	3 016	1 106	1 408
	0,092	2 734	1 660	1 367	0,075	4 323	2 625	2 162
2	0,091	3 530	785	1 046	0,074	5 692	1 242	1 656
	0,034	4 440	1 954	1 421	0,023	7 121	3 072	2 234
3	0,033	5 468	1 013	1 215	0,025	8 950	1 492	1 918
	0,018	6 446	2 321	1 547	0,014	10 234	3 280	2 362
4	0,017	7 642	1 332	1 403	0,013	12 415	1 852	2 179
	0,012	8 397	2 552	1 647	0,01	13 095	3 004	2 403
5	0,011	9 862	1 715	1 573				
	0,01	10 070	2 089	1 642				

Таблица VI

Параметры одноступенчатых планов СЭП
(Методика СЭП [1])

γ	Обозначения параметров плана	Значения параметров для E					
		0,1	0,063	0,04	0,025	0,016	0,01
1,0	nq_H	0,051	0,032	0,020	0,013	0,008	0,005
	c	0	0	0	0	0	0
0,8	nq_H	0,117	0,075	0,049	0,031	0,020	0,012
	c	1	1	1	1	1	1
0,63	nq_H	0,144	0,093	0,061	0,039	0,025	0,016
	c	1	1	1	1	1	1
0,5	nq_H	0,173	0,114	0,075	0,048	0,031	0,020
	c	1	1	1	1	1	1
0,4	nq_H	0,204	0,136	0,090	0,058	0,038	0,024
	c	1	1	1	1	1	1
0,32	nq_H	0,237	0,161	0,108	0,071	0,047	0,030
	c	1	1	1	1	1	1
0,25	nq_H	0,274	0,189	0,129	0,086	0,058	0,037
	c	1	1	1	1	1	1
0,2	nq_H	0,307	0,215	0,149	0,101	0,069	0,045
	c	1	1	1	1	1	1
0,16	nq_H	0,524	0,240	0,170	0,117	0,081	0,054
	c	2	1	1	1	1	1
0,13	nq_H	0,595	0,417	0,284	0,131	0,092	0,063
	c	2	2	2	1	1	1
0,1	nq_H	0,681	0,492	0,346	0,232	0,107	0,074
	c	2	2	2	2	1	1
0,08	nq_H	1,007	0,553	0,400	0,277	0,188	0,084
	c	3	2	2	2	2	1
0,063	nq_H	1,136	0,834	0,456	0,325	0,228	0,151
	c	3	3	2	2	2	2
0,05	nq_H	1,564	0,945	0,690	0,371	0,268	0,182
	c	4	3	3	2	2	2
0,04	nq_H	1,723	1,305	0,786	0,560	0,305	0,215
	c	4	4	3	3	2	2
0,032	nq_H	2,258	1,454	0,872	0,646	0,460	0,247
	c	5	4	3	3	3	2
0,025	nq_H	2,883	1,940	1,243	0,732	0,545	0,375
	c	6	5	4	3	3	3
0,02	nq_H	3,555	2,475	1,651	1,037	0,616	0,443
	c	7	6	5	4	3	3
0,016	nq_H	4,291	3,080	1,819	1,152	0,872	0,509
	c	8	7	5	4	4	3
0,013	nq_H	5,577	3,734	2,322	1,538	0,976	0,706
	c	10	8	6	5	4	4
0,01	nq_H	7,081	4,527	2,938	2,021	1,341	0,832
	c	12	9	7	6	5	4

Таблица VII

Параметры двухступенчатых планов СЭП
(Методика СЭП [1])

γ	Обозначения параметров плана	Значения параметров для E					
		0,1	0,363	0,04	0,025	0,016	0,01
1,0	$n_1 q_H$	0,078	0,052	0,034	0,022	0,015	0,009
	$n_2 q_H$	0,132	0,102	0,076	0,120	0,092	0,070
	c_1	0	0	0	0	0	0
0,8	c_2	1	1	1	2	2	2
	$n_1 q_H$	0,094	0,063	0,042	0,027	0,018	0,012
	$n_2 q_H$	0,142	0,222	0,174	0,138	0,107	0,079
0,63	c_1	0	0	0	0	0	0
	c_2	1	2	2	2	2	2
	$n_1 q_H$	0,115	0,078	0,052	0,034	0,022	0,014
0,5	$n_2 q_H$	0,313	0,248	0,196	0,150	0,118	0,089
	c_1	0	0	0	0	0	0
	c_2	2	2	2	2	2	2
0,4	$n_1 q_H$	0,140	0,094	0,064	0,042	0,028	0,018
	$n_2 q_H$	0,334	0,266	0,213	0,165	0,132	0,102
	c_1	0	0	0	0	0	0
0,32	c_2	2	2	2	2	2	2
	$n_1 q_H$	0,167	0,113	0,077	0,051	0,034	0,022
	$n_2 q_H$	0,358	0,290	0,231	0,185	0,144	0,112
0,25	c_1	0	0	0	0	0	0
	c_2	2	2	2	2	2	2
	$n_1 q_H$	0,197	0,135	0,093	0,062	0,042	0,027
0,2	$n_2 q_H$	0,380	0,311	0,250	0,200	0,162	0,124
	c_1	0	0	0	0	0	0
	c_2	2	2	2	2	2	2
0,16	$n_1 q_H$	0,235	0,164	0,113	0,076	0,052	0,034
	$n_2 q_H$	0,398	0,324	0,270	0,214	0,174	0,136
	c_1	0	0	0	0	0	0
0,13	c_2	3	3	3	2	2	2
	$n_1 q_H$	0,273	0,190	0,132	0,091	0,062	0,041
	$n_2 q_H$	0,688	0,564	0,467	0,227	0,187	0,147
0,1	c_1	0	0	0	0	0	0
	c_2	3	3	3	3	3	3
	$n_1 q_H$	0,320	0,226	0,157	0,107	0,074	0,050
0,08	$n_2 q_H$	0,720	0,598	0,501	0,407	0,328	0,158
	c_1	0	0	0	0	0	0
	c_2	3	3	3	3	3	3
0,06	$n_1 q_H$	0,367	0,262	0,185	0,127	0,088	0,059
	$n_2 q_H$	0,753	0,632	0,524	0,431	0,356	0,287
	c_1	0	0	0	0	0	0
0,05	c_2	3	3	3	3	3	3

Продолжение табл. VII

γ	Обозначения параметров плана	Значения параметров для E					
		0,1	0,063	0,04	0,025	0,016	0,01
0,1	$n_1 q_H$	0,670	0,479	0,225	0,156	0,109	0,074
	$n_2 q_H$	1,382	1,229	0,555	0,460	0,388	0,312
	c_1	1	1	0	0	0	0
	c_2	5	5	3	3	3	3
0,08	$n_1 q_H$	0,772	0,562	0,256	0,184	0,130	0,089
	$n_2 q_H$	1,305	1,237	0,849	0,487	0,409	0,337
	c_1	1	1	0	0	0	0
	c_2	5	5	4	3	3	3
0,063	$n_1 q_H$	0,885	0,653	0,474	0,214	0,151	0,108
	$n_2 q_H$	1,874	1,704	1,491	0,751	0,633	0,355
	c_1	1	1	1	0	0	0
	c_2	6	6	6	4	4	3
0,05	$n_1 q_H$	1,322	0,757	0,558	0,394	0,181	0,125
	$n_2 q_H$	2,680	1,722	1,553	1,362	0,679	0,560
	c_1	2	1	1	1	0	0
	c_2	8	6	6	6	4	4
0,04	$n_1 q_H$	1,488	0,857	0,638	0,465	0,214	0,149
	$n_2 q_H$	3,243	2,220	2,001	1,405	0,707	0,600
	c_1	2	1	1	1	0	0
	c_2	9	7	7	6	4	4
0,032	$n_1 q_H$	1,662	1,278	0,738	0,538	0,393	0,177
	$n_2 q_H$	3,802	3,464	2,019	1,797	1,278	0,625
	c_1	2	2	1	1	1	0
	c_2	10	10	7	7	6	4
0,025	$n_1 q_H$	2,294	1,454	1,123	0,624	0,465	0,325
	$n_2 q_H$	5,209	4,088	3,218	2,272	1,692	1,462
	c_1	3	2	2	1	1	1
	c_2	13	11	10	8	7	7
0,02	$n_1 q_H$	2,993	2,018	1,269	0,721	0,537	0,386
	$n_2 q_H$	6,563	4,916	3,857	2,322	2,076	1,557
	c_1	4	3	2	1	1	1
	c_2	16	13	11	8	8	7
0,016	$n_1 q_H$	3,278	2,228	1,426	1,091	0,623	0,449
	$n_2 q_H$	7,378	6,167	4,369	3,528	2,138	1,952
	c_1	4	3	2	2	1	1
	c_2	17	15	12	11	8	8
0,013	$n_1 q_H$	4,072	2,891	1,958	1,224	0,700	0,521
	$n_2 q_H$	9,488	7,747	5,679	4,093	2,616	1,982
	c_1	5	4	3	2	1	1
	c_2	21	18	15	12	9	8
0,01	$n_1 q_H$	5,598	3,706	2,613	1,742	1,092	0,611
	$n_2 q_H$	13,033	9,844	7,744	5,346	3,933	2,419
	c_1	7	5	4	3	2	1
	c_2	27	22	19	15	12	9

Таблица VIII
Двухэтапные планы, гарантирующие минимум предельного числа критических отказов

$\kappa=1,0$			$\kappa=0,5$			$\kappa=0,1$			$\kappa=0,05$			$\kappa=0,01$		
N	n	$\bar{f}$	N	n	$\bar{f}$	N	n	$\bar{f}$	N	n	$\bar{f}$	N	n	$\bar{f}$
10	2	2,13	10	3	1,92	25	12	1,48	50	25	1,49			
15	2	2,86	14	3	2,23	30	15	1,48	60	30	1,49			
16	3	3,00	15	4	2,23	31	13	1,80	61	26	1,82			
31	3	3,95	22	4	2,58	40	17	1,81	80	34	1,83			
32	4	4,00	23	5	2,65	41	15	2,05	81	30	2,08			
52	4	4,97	34	5	3,07	50	18	2,06	100	37	2,09			
53	5	5,00	35	6	3,11	51	17	2,26	101	33	2,29			
78	5	5,95	48	6	3,52	60	20	2,27	120	39	2,30			
79	6	6,00	49	7	3,56	61	18	2,43	121	36	2,47			
110	6	6,98	66	7	4,04	70	21	2,44	140	42	2,47			
111	7	7,00	67	8	4,03	71	19	2,58						
147	7	7,97	86	8	4,52	80	22	2,58						
148	8	8,00	87	9	4,51	81	20	2,71						
190	8	8,99	108	9	4,99	90	23	2,72						
191	9	9,00	109	10	4,99	91	21	2,82						
238	9	9,99	134	10	5,50	100	23	2,83						

239	10	10,00	135	11	5,48	101	22	2,93	141	39	2,62	200	100	1,50
291	10	10,97	162	11	5,99	110	24	2,94	160	44	2,62	250	125	1,50
292	11	11,00	163	12	5,98	111	23	3,02	161	41	2,75	299	149	1,50
350	11	11,98	192	12	6,46	120	25	3,03	180	46	2,76	300	127	1,84
351	12	12,00	193	13	6,47	121	24	3,11	181	43	2,87	350	148	1,84
415	12	13,00	226	13	6,97	130	26	3,12	200	47	2,88	399	169	1,84
416	13	13,00	227	14	6,97	131	24	3,19	201	47	2,99	400	148	2,10
484	13	13,97	262	14	7,46	140	26	3,20	220	49	2,98	500	165	2,32
485	14	14,00	263	15	7,47	141	25	3,27	221	47	3,08	600	180	2,50
560	14	15,00	302	15	7,98	150	27	3,27	240	50	3,08	800	205	2,79
561	15	15,00	303	16	7,98	200	29	3,58	241	48	3,17	1000	225	3,02
640	15	15,98	344	16	8,48	250	32	3,83	260	52	3,17			
641	16	16,00	345	17	8,48	300	35	4,04	261	49	3,25			
726	16	16,98	388	17	8,97	400	41	4,46	280	53	3,25			
727	17	17,00	500	20	10,00	500	46	4,85	300	54	3,33			
818	17	17,99	600	22	11,00	600	51	5,25	400	59	3,65			
819	18	18,00	700	23	11,77	700	54	5,60	500	63	3,90			
915	18	18,99	800	25	12,50	800	58	5,90	600	70	4,11			
916	19	19,00	900	26	13,31	900	62	6,24	800	82	4,51			
1000	19	19,66	1000	28	14,00	1000	65	6,52	1000	92	4,89			

Таблица IX
Трехэтапные планы, гарантирующие минимум предельного числа критических отказов

$\kappa=1,0$				$\kappa=0,5$				$\kappa=0,1$				$\kappa=0,05$				$\kappa=0,01$			
N	n_1	n_2	$\bar{f}$	N	n_1	n_2	$\bar{f}$	N	n_1	n_2	$\bar{f}$	N	n_1	n_2	$\bar{f}$	N	n_1	n_2	$\bar{f}$
10	1	3	1,50	10	2	3	1,44												
12	1	3	1,63	13	2	3	1,52												
13	1	4	1,65	14	2	4	1,55												
17	1	4	1,83	17	2	4	1,64												
18	1	5	1,88	18	2	5	1,67												
22	1	5	1,98	21	2	5	1,73												
23	2	4	2,00	22	3	4	1,77												
24	2	4	2,00	23	3	4	1,78												
25	2	5	2,00	24	3	5	1,82	25	8	9	1,31								
27	2	5	2,00	28	3	5	1,87	27	8	9	1,31								
28	2	6	2,00	29	3	6	1,86	28	9	10	1,31								
30	2	6	2,00	32	3	6	1,91	34	9	10	1,48								
31	2	7	2,00	33	3	7	1,92	35	9	11	1,48								
33	2	7	2,01	37	3	7	1,99	36	9	11	1,49								
34	2	8	2,04	38	3	8	2,00	37	10	11	1,49								
40	2	8	2,14	42	3	8	2,05	38	10	12	1,49								

41	2	9	2,17	43	3	9	2,06	40	11	12	1,49	50	16	17	1,32	200	66	67	1,33
48	2	9	2,28	47	3	9	2,10	45	11	12	1,61	52	17	18	1,32	250	83	84	1,33
49	2	10	2,30	48	3	10	2,12	46	11	13	1,61	54	18	19	1,32	300	83	92	1,52
57	2	10	2,32	53	3	10	2,17	48	11	13	1,61	57	19	20	1,32	400	88	115	1,66
58	2	11	2,44	54	3	11	2,19	49	11	14	1,61	59	20	21	1,32	500	95	134	1,75
66	2	11	2,56	59	3	11	2,24	55	11	14	1,70	60	21	22	1,32	600	100	150	1,83
67	2	12	2,57	60	3	12	2,26	56	11	15	1,70	61	22	23	1,32	800	109	177	1,95
76	2	12	2,58	66	3	12	2,32	58	11	15	1,70	62	23	24	1,32	1000	114	199	2,04
77	2	13	2,71	67	4	11	2,32	60	11	16	1,70	63	24	25	1,32				
87	2	13	2,83	70	4	11	2,36	70	11	17	1,78	64	25	26	1,32				
88	2	14	2,85	71	4	12	2,35	80	12	18	1,84	65	26	27	1,32				
100	2	15	2,99	75	4	12	2,38	90	12	20	1,88	66	27	28	1,32				
120	3	17	3,00	100	4	17	2,52	100	12	21	1,93	67	28	29	1,32				
150	3	21	3,09	150	5	22	2,83	150	13	28	2,09	68	29	30	1,32				
200	3	24	3,46	200	5	28	2,97	200	14	36	2,25	69	30	31	1,32				
250	3	28	3,80	250	5	33	3,22	250	15	42	2,37	70	31	32	1,32				
300	3	30	4,09	300	6	38	3,29	300	16	48	2,49	71	32	33	1,32				
400	4	39	4,17	400	6	45	3,63	400	18	59	2,67	72	33	34	1,32				
500	4	45	4,58	500	7	55	3,74	500	19	70	2,83	73	34	35	1,32				
600	4	49	4,92	600	7	61	4,02	600	21	80	2,96	74	35	36	1,32				
800	5	63	5,14	800	8	75	4,32	800	23	100	3,17	75	36	37	1,32				
1000	5	71	5,67	1000	9	88	4,58	1000	24	119	3,33	76	37	38	1,32				

Т а б л и ц а X

Значения критических уровней $\beta_{0,9}(n, N)$
(Рыкова [2])

$n \backslash N$	10	12	14	16	20	25	30	$y_{0,9}(n)$
3	1,035 0,933	1,000 0,909	1,116 0,945	1,081 0,929	1,052 0,903	1,108 0,922	1,095 0,916	1,102
4	1,033 0,905	1,021 0,915	1,099 0,916	1,155 0,952	1,118 0,930	1,091 0,924	1,096 0,902	1,130
5	0,949 0,921	1,122 0,934	1,076 0,914	1,146 0,926	1,162 0,948	1,100 0,915	1,143 0,929	1,139
6		1,155 0,974	1,080 0,907	1,097 0,908	1,122 0,918	1,124 0,918	1,095 0,907	1,146
7			1,069 0,947	1,102 0,902	1,149 0,906	1,122 0,906	1,118 0,901	1,154
8				1,000 0,913	1,095 0,909	1,097 0,909	1,123 0,908	1,159
9					1,124 0,911	1,125 0,933	1,147 0,911	1,162
10					1,118 0,948	1,143 0,923	1,162 0,938	1,166
12						1,105 0,917	1,118 0,938	1,171
14							1,0247 0,903	1,176
15							1,095 0,922	1,177

Таблица XI

Значения функций $\omega_1(x, y)$ и $\omega_2(x, y)$

$y-x$		0,00	0,01	0,02	0,03	0,04	0,05	0,06
x								
0,01	ω_2	1,0101	1,0153	1,0205	1,0258	1,0312	1,0366	1,0421
	ω_1	0,0001	0,0001	0,0002	0,0002	0,0003	0,0003	0,0004
0,02	ω_2	1,0205	1,0258	1,0311	1,0365	1,0420	1,0476	1,0533
	ω_1	0,0002	0,0003	0,0004	0,0005	0,0007	0,0008	0,0009
0,03	ω_2	1,0311	1,0365	1,0420	1,0476	1,0532	1,0589	1,0647
	ω_1	0,0005	0,0006	0,0008	0,0010	0,0012	0,0013	0,0015
0,04	ω_2	1,0420	1,0475	1,0531	1,0588	1,0646	1,0705	1,0765
	ω_1	0,0009	0,0011	0,0013	0,0016	0,0018	0,0021	0,0023
0,05	ω_2	1,0531	1,0588	1,0646	1,0704	1,0764	1,0824	1,0885
	ω_1	0,0014	0,0017	0,0020	0,0023	0,0026	0,0029	0,0033
0,06	ω_2	1,0646	1,0704	1,0763	1,0823	1,0884	1,0946	1,1009
	ω_1	0,0020	0,0024	0,0028	0,0032	0,0036	0,0040	0,0044
0,07	ω_2	1,0763	1,0823	1,0884	1,0945	1,1008	1,1072	1,1136
	ω_1	0,0028	0,0033	0,0037	0,0042	0,0047	0,0052	0,0057
0,08	ω_2	1,0883	1,0945	1,1007	1,1071	1,1135	1,1200	1,1267
	ω_1	0,0038	0,0043	0,0049	0,0054	0,0060	0,0065	0,0071
0,09	ω_2	1,1007	1,1070	1,1134	1,1199	1,1266	1,1333	1,1401
	ω_1	0,0049	0,0055	0,0061	0,0068	0,0074	0,0081	0,0088
0,10	ω_2	1,1134	1,1199	1,1266	1,1332	1,1400	1,1469	1,1540
	ω_1	0,0062	0,0069	0,0076	0,0084	0,0091	0,0099	0,0107
0,11	ω_2	1,1265	1,1332	1,1399	1,1468	1,1538	1,1610	1,1682
	ω_1	0,0077	0,0085	0,0093	0,0102	0,0110	0,0119	0,0128
0,12	ω_2	1,1399	1,1468	1,1538	1,1609	1,1681	1,1754	1,1829
	ω_1	0,0094	0,0103	0,0112	0,0122	0,0132	0,0142	0,0152
0,13	ω_2	1,1537	1,1608	1,1680	1,1753	1,1828	1,1903	1,1980
	ω_1	0,0113	0,0123	0,0134	0,0144	0,0155	0,0167	0,0178
0,14	ω_2	1,1680	1,1753	1,1827	1,1902	1,1979	1,2057	1,2137
	ω_1	0,0134	0,0146	0,0157	0,0170	0,0182	0,0195	0,0208
0,15	ω_2	1,1827	1,1902	1,1978	1,2056	1,2135	1,2216	1,2298
	ω_1	0,0158	0,0171	0,0184	0,0198	0,0211	0,0226	0,0240
0,16	ω_2	1,1978	1,2055	1,2134	1,2214	1,2296	1,2380	1,2464
	ω_1	0,0185	0,0199	0,0214	0,0229	0,0244	0,0260	0,0276
0,17	ω_2	1,2134	1,2214	1,2295	1,2378	1,2463	1,2549	1,2637
	ω_1	0,0214	0,0230	0,0246	0,0263	0,0280	0,0298	0,0316
0,18	ω_2	1,2295	1,2378	1,2462	1,2548	1,2635	1,2724	1,2815
	ω_1	0,0247	0,0264	0,0282	0,0301	0,0320	0,0339	0,0359
0,19	ω_2	1,2461	1,2547	1,2634	1,2723	1,2813	1,2905	1,2999
	ω_1	0,0283	0,0302	0,0322	0,0342	0,0363	0,0385	0,0407
0,20	ω_2	1,2634	1,2722	1,2812	1,2904	1,2998	1,3093	1,3191
	ω_1	0,0323	0,0344	0,0366	0,0388	0,0411	0,0435	0,0460
0,21	ω_2	1,2812	1,2903	1,2997	1,3092	1,3189	1,3288	1,3389
	ω_1	0,0366	0,0390	0,0414	0,0438	0,0464	0,0490	0,0517
0,22	ω_2	1,2996	1,3091	1,3188	1,3287	1,3387	1,3490	1,3596
	ω_1	0,0414	0,0440	0,0466	0,0494	0,0522	0,0551	0,0581
0,23	ω_2	1,3187	1,3286	1,3386	1,3489	1,3594	1,3701	1,3810
	ω_1	0,0467	0,0495	0,0524	0,0554	0,0585	0,0617	0,0651
0,24	ω_2	1,3386	1,3488	1,3592	1,3699	1,3808	1,3919	1,4034
	ω_1	0,0525	0,0556	0,0588	0,0621	0,0655	0,0690	0,0727
0,25	ω_2	1,3592	1,3698	1,3807	1,3918	1,4031	1,4147	1,4266
	ω_1	0,0589	0,0623	0,0658	0,0694	0,0732	0,0771	0,0811

Продолжение табл. XI

$y-x$ x		0,07	0,08	0,09	0,10	0,11	0,12	0,13
0,01	ω_2	1,0477	1,0534	1,0592	1,0651	1,0710	1,0770	1,0832
	ω_1	0,0004	0,0005	0,0006	0,0006	0,0007	0,0008	0,0008
0,02	ω_2	1,0590	1,0649	1,0708	1,0768	1,0829	1,0891	1,0954
	ω_1	0,0010	0,0011	0,0013	0,0014	0,0015	0,0017	0,0018
0,03	ω_2	1,0706	1,0766	1,0827	1,0889	1,0951	1,1015	1,1080
	ω_1	0,0017	0,0019	0,0021	0,0023	0,0025	0,0027	0,0030
0,04	ω_2	1,0825	1,0887	1,0949	1,1013	1,1077	1,1142	1,1209
	ω_1	0,0026	0,0028	0,0031	0,0034	0,0037	0,0040	0,0043
0,05	ω_2	1,0947	1,1011	1,1075	1,1140	1,1206	1,1273	1,1342
	ω_1	0,0036	0,0039	0,0043	0,0047	0,0050	0,0054	0,0058
0,06	ω_2	1,1073	1,1138	1,1204	1,1271	1,1339	1,1408	1,1478
	ω_1	0,0048	0,0052	0,0057	0,0061	0,0066	0,0071	0,0075
0,07	ω_2	1,1202	1,1269	1,1336	1,1405	1,1475	1,1547	1,1619
	ω_1	0,0062	0,0067	0,0072	0,0078	0,0083	0,0089	0,0095
0,08	ω_2	1,1334	1,1403	1,1473	1,1544	1,1616	1,1689	1,1764
	ω_1	0,0077	0,0084	0,0090	0,0096	0,0103	0,0110	0,0117
0,09	ω_2	1,1471	1,1542	1,1613	1,1686	1,1761	1,1836	1,1913
	ω_1	0,0095	0,0102	0,0110	0,0117	0,0125	0,0133	0,0141
0,10	ω_2	1,1611	1,1684	1,1758	1,1833	1,1910	1,1988	1,2068
	ω_1	0,0115	0,0123	0,0132	0,0141	0,0150	0,0159	0,0169
0,11	ω_2	1,1756	1,1831	1,1907	1,1985	1,2064	1,2145	1,2227
	ω_1	0,0138	0,0147	0,0157	0,0167	0,0177	0,0188	0,0199
0,12	ω_2	1,1905	1,1982	1,2061	1,2141	1,2223	1,2306	1,2391
	ω_1	0,0163	0,0173	0,0185	0,0196	0,0208	0,0220	0,0233
0,13	ω_2	1,2059	1,2139	1,2220	1,2303	1,2387	1,2473	1,2561
	ω_1	0,0190	0,0203	0,0215	0,0228	0,0242	0,0255	0,0270
0,14	ω_2	1,2217	1,2300	1,2384	1,2469	1,2557	1,2646	1,2736
	ω_1	0,0221	0,0235	0,0249	0,0264	0,0279	0,0294	0,0310
0,15	ω_2	1,2381	1,2467	1,2553	1,2642	1,2732	1,2824	1,2918
	ω_1	0,0255	0,0271	0,0287	0,0303	0,0320	0,0337	0,0355
0,16	ω_2	1,2551	1,2639	1,2729	1,2820	1,2914	1,3009	1,3107
	ω_1	0,0293	0,0310	0,0328	0,0346	0,0365	0,0385	0,0405
0,17	ω_2	1,2726	1,2817	1,2910	1,3005	1,3102	1,3201	1,3302
	ω_1	0,0334	0,0354	0,0373	0,0394	0,0415	0,0437	0,0459
0,18	ω_2	1,2907	1,3002	1,3098	1,3197	1,3297	1,3400	1,3505
	ω_1	0,0380	0,0401	0,0423	0,0446	0,0470	0,0494	0,0519
0,19	ω_2	1,3095	1,3193	1,3294	1,3396	1,3500	1,3607	1,3716
	ω_1	0,0430	0,0454	0,0478	0,0504	0,0530	0,0557	0,0585
0,20	ω_2	1,3290	1,3392	1,3496	1,3602	1,3711	1,3822	1,3935
	ω_1	0,0485	0,0511	0,0539	0,0567	0,0596	0,0626	0,0657
0,21	ω_2	1,3493	1,3599	1,3707	1,3817	1,3930	1,4046	1,4164
	ω_1	0,0546	0,0575	0,0605	0,0636	0,0668	0,0702	0,0736
0,22	ω_2	1,3703	1,3813	1,3926	1,4041	1,4159	1,4279	1,4403
	ω_1	0,0612	0,0644	0,0677	0,0712	0,0748	0,0785	0,0824
0,23	ω_2	1,3922	1,4037	1,4154	1,4274	1,4397	1,4523	1,4652
	ω_1	0,0685	0,0721	0,0757	0,0796	0,0835	0,0877	0,0920
0,24	ω_2	1,4150	1,4270	1,4392	1,4517	1,4646	1,4778	1,4913
	ω_1	0,0765	0,0804	0,0845	0,0888	0,0932	0,0978	0,1026
0,25	ω_2	1,4388	1,4513	1,4641	1,4772	1,4906	1,5045	1,5186
	ω_1	0,0853	0,0897	0,0942	0,0989	0,1038	0,1089	0,1143

Продолжение табл. XI

$y-x$ x		0,14	0,15	0,16	0,17	0,18	0,19	0,20
0,01	ω_2	1,0894	1,0957	1,1021	1,1086	1,1153	1,1220	1,1289
	ω_1	0,0009	0,0010	0,0010	0,0011	0,0012	0,0013	0,0013
0,02	ω_2	1,1018	1,1083	1,1149	1,1216	1,1284	1,1353	1,1424
	ω_1	0,0019	0,0021	0,0022	0,0024	0,0026	0,0027	0,0029
0,03	ω_2	1,1145	1,1212	1,1280	1,1349	1,1419	1,1491	1,1563
	ω_1	0,0032	0,0034	0,0036	0,0039	0,0041	0,0044	0,0046
0,04	ω_2	1,1276	1,1345	1,1415	1,1486	1,1558	1,1632	1,1706
	ω_1	0,0046	0,0049	0,0052	0,0056	0,0059	0,0063	0,0066
0,05	ω_2	1,1411	1,1482	1,1554	1,1627	1,1701	1,1777	1,1854
	ω_1	0,0062	0,0066	0,0070	0,0075	0,0079	0,0084	0,0088
0,06	ω_2	1,1550	1,1623	1,1697	1,1772	1,1849	1,1927	1,2007
	ω_1	0,0080	0,0086	0,0091	0,0096	0,0102	0,0108	0,0113
0,07	ω_2	1,1693	1,1768	1,1844	1,1922	1,2001	1,2082	1,2164
	ω_1	0,0101	0,0107	0,0114	0,0120	0,0127	0,0134	0,0141
0,08	ω_2	1,1840	1,1917	1,1996	1,2076	1,2158	1,2241	1,2326
	ω_1	0,0124	0,0132	0,0139	0,0147	0,0155	0,0163	0,0172
0,09	ω_2	1,1992	1,2072	1,2153	1,2236	1,2320	1,2406	1,2494
	ω_1	0,0150	0,0159	0,0168	0,0177	0,0186	0,0196	0,0206
0,10	ω_2	1,2148	1,2231	1,2315	1,2400	1,2488	1,2576	1,2667
	ω_1	0,0179	0,0189	0,0199	0,0210	0,0221	0,0232	0,0244
0,11	ω_2	1,2310	1,2395	1,2482	1,2571	1,2661	1,2753	1,2847
	ω_1	0,0210	0,0222	0,0234	0,0246	0,0259	0,0272	0,0286
0,12	ω_2	1,2477	1,2565	1,2655	1,2747	1,2840	1,2935	1,3033
	ω_1	0,0245	0,0259	0,0272	0,0287	0,0301	0,0316	0,0332
0,13	ω_2	1,2650	1,2741	1,2834	1,2929	1,3026	1,3125	1,3226
	ω_1	0,0284	0,0299	0,0315	0,0331	0,0347	0,0365	0,0382
0,14	ω_2	1,2829	1,2923	1,3019	1,3118	1,3218	1,3321	1,3426
	ω_1	0,0327	0,0344	0,0362	0,0380	0,0399	0,0418	0,0438
0,15	ω_2	1,3014	1,3112	1,3212	1,3314	1,3418	1,3525	1,3634
	ω_1	0,0374	0,0393	0,0413	0,0433	0,0455	0,0477	0,0499
0,16	ω_2	1,3206	1,3308	1,3411	1,3517	1,3626	1,3736	1,3850
	ω_1	0,0426	0,0447	0,0470	0,0493	0,0516	0,0541	0,0567
0,17	ω_2	1,3405	1,3511	1,3619	1,3729	1,3842	1,3957	1,4075
	ω_1	0,0483	0,0507	0,0532	0,0558	0,0584	0,0612	0,0641
0,18	ω_2	1,3612	1,3722	1,3834	1,3949	1,4066	1,4187	1,4310
	ω_1	0,0545	0,0572	0,0600	0,0629	0,0659	0,0690	0,0722
0,19	ω_2	1,3828	1,3942	1,4059	1,4178	1,4301	1,4426	1,4555
	ω_1	0,0614	0,0644	0,0675	0,0707	0,0741	0,0776	0,0812
0,20	ω_2	1,4052	1,4171	1,4293	1,4418	1,4546	1,4677	1,4812
	ω_1	0,0689	0,0723	0,0758	0,0794	0,0832	0,0871	0,0912
0,21	ω_2	1,4285	1,4410	1,4537	1,4668	1,4802	1,4939	1,5080
	ω_1	0,0772	0,0810	0,0849	0,0889	0,0931	0,0975	0,1021
0,22	ω_2	1,4529	1,4659	1,4793	1,4929	1,5070	1,5214	1,5362
	ω_1	0,0864	0,0906	0,0949	0,0994	0,1041	0,1091	0,1142
0,23	ω_2	1,4785	1,4921	1,5060	1,5204	1,5351	1,5503	1,5659
	ω_1	0,0964	0,1011	0,1060	0,1110	0,1163	0,1219	0,1277
0,24	ω_2	1,5052	1,5194	1,5341	1,5492	1,5647	1,5806	1,5971
	ω_1	0,1076	0,1128	0,1182	0,1239	0,1298	0,1361	0,1426
0,25	ω_2	1,5332	1,5482	1,5636	1,5795	1,5958	1,6127	1,6300
	ω_1	0,1198	0,1257	0,1317	0,1381	0,1448	0,1518	0,1592

Продолжение табл. XI

$y-x$ x		0,21	0,22	0,23	0,24	0,25
0,01	ω_2	1,1358	1,1429	1,1501	1,1574	1,1649
	ω_1	0,0014	0,0015	0,0016	0,0017	0,0018
0,02	ω_2	1,1496	1,1569	1,1643	1,1718	1,1795
	ω_1	0,0031	0,0032	0,0034	0,0036	0,0038
0,03	ω_2	1,1637	1,1712	1,1789	1,1866	1,1946
	ω_1	0,0049	0,0052	0,0055	0,0058	0,0061
0,04	ω_2	1,1783	1,1860	1,1939	1,2019	1,2101
	ω_1	0,0070	0,0074	0,0078	0,0082	0,0086
0,05	ω_2	1,1933	1,2013	1,2094	1,2177	1,2261
	ω_1	0,0093	0,0098	0,0103	0,0109	0,0114
0,06	ω_2	1,2088	1,2170	1,2254	1,2340	1,2427
	ω_1	0,0119	0,0126	0,0132	0,0139	0,0145
0,07	ω_2	1,2247	1,2333	1,2419	1,2508	1,2598
	ω_1	0,0149	0,0156	0,0164	0,0172	0,0180
0,08	ω_2	1,2412	1,2501	1,2590	1,2682	1,2776
	ω_1	0,0181	0,0190	0,0199	0,0209	0,0219
0,09	ω_2	1,2583	1,2674	1,2767	1,2862	1,2959
	ω_1	0,0217	0,0227	0,0238	0,0250	0,0261
0,10	ω_2	1,2760	1,2854	1,2951	1,3049	1,3149
	ω_1	0,0256	0,0269	0,0281	0,0295	0,0308
0,11	ω_2	1,2943	1,3040	1,3140	1,3242	1,3347
	ω_1	0,0300	0,0314	0,0329	0,0344	0,0360
0,12	ω_2	1,3132	1,3234	1,3337	1,3443	1,3552
	ω_1	0,0348	0,0364	0,0381	0,0399	0,0417
0,13	ω_2	1,3329	1,3434	1,3542	1,3652	1,3765
	ω_1	0,0401	0,0419	0,0439	0,0459	0,0480
0,14	ω_2	1,3533	1,3642	1,3755	1,3869	1,3987
	ω_1	0,0459	0,0480	0,0502	0,0526	0,0550
0,15	ω_2	1,3745	1,3859	1,3976	1,4095	1,4218
	ω_1	0,0523	0,0547	0,0572	0,0599	0,0626
0,16	ω_2	1,3966	1,4085	1,4207	1,4331	1,4459
	ω_1	0,0593	0,0621	0,0649	0,0679	0,0710
0,17	ω_2	1,4196	1,4320	1,4447	1,4578	1,4711
	ω_1	0,0671	0,0702	0,0734	0,0768	0,0803
0,18	ω_2	1,4436	1,4566	1,4699	1,4835	1,4975
	ω_1	0,0756	0,0791	0,0828	0,0866	0,0905
0,19	ω_2	1,4687	1,4823	1,4962	1,5105	1,5252
	ω_1	0,0850	0,0890	0,0931	0,0974	0,1019
0,20	ω_2	1,4950	1,5092	1,5238	1,5388	1,5543
	ω_1	0,0954	0,0999	0,1045	0,1093	0,1144
0,21	ω_2	1,5226	1,5375	1,5528	1,5686	1,5849
	ω_1	0,1069	0,1119	0,1171	0,1226	0,1284
0,22	ω_2	1,5515	1,5672	1,5833	1,6000	1,6172
	ω_1	0,1196	0,1253	0,1312	0,1374	0,1439
0,23	ω_2	1,5819	1,5985	1,6155	1,6332	1,6514
	ω_1	0,1337	0,1401	0,1468	0,1539	0,1613
0,24	ω_2	1,6140	1,6315	1,6496	1,6683	1,6876
	ω_1	0,1495	0,1567	0,1643	0,1723	0,1808
0,25	ω_2	1,6480	1,6665	1,6857	1,7055	1,7261
	ω_1	0,1670	0,1752	0,1839	0,1930	0,2027

Таблица XII

Уровни оперативной характеристики распределения Пуассона

(Дугина [1])

 $c = 0$

	0	1	2	3	4	5	6	7	8	9
0.99	0,01005	0,00904	0,00803	0,00702	0,00602	0,00501	0,00401	0,00300	0,00200	0,00100
0.98	0,02020	0,01918	0,01816	0,01715	0,01613	0,01511	0,01410	0,01309	0,01207	0,01106
0.97	0,03046	0,02943	0,02840	0,02737	0,02634	0,02532	0,02429	0,02327	0,02225	0,02122
0.96	0,04082	0,03978	0,03874	0,03770	0,03666	0,03563	0,03459	0,03356	0,03252	0,03149
0.95	0,05129	0,05024	0,04919	0,04814	0,04709	0,04604	0,04500	0,04395	0,04291	0,04186
0.94	0,06188	0,06081	0,05975	0,05869	0,05763	0,05657	0,05551	0,05446	0,05340	0,05235
0.93	0,07257	0,07150	0,07042	0,06935	0,06828	0,06721	0,06614	0,06507	0,06401	0,06294
0.92	0,08338	0,08230	0,08121	0,08013	0,07904	0,07796	0,07688	0,07580	0,07472	0,07365
0.91	0,09431	0,09321	0,09212	0,09102	0,08992	0,08883	0,08774	0,08665	0,08556	0,08447
0.90	0,10536	0,10425	0,10314	0,10203	0,10093	0,09982	0,09872	0,09761	0,09651	0,09541
0.8	0,22314	0,21072	0,19845	0,18633	0,17435	0,16252	0,15082	0,13926	0,12783	0,11653
0.7	0,35667	0,34249	0,32850	0,31471	0,30111	0,28768	0,27444	0,26136	0,24846	0,23572
0.6	0,51083	0,49430	0,47804	0,46204	0,44629	0,43078	0,41552	0,40048	0,38566	0,37106
0.5	0,69315	0,67334	0,65393	0,63488	0,61619	0,59784	0,57982	0,56212	0,54473	0,52763
0.4	0,91629	0,89160	0,86750	0,84397	0,82098	0,79851	0,77653	0,75502	0,73397	0,71335
0.3	1,20397	1,17118	1,13943	1,10866	1,07881	1,04982	1,02165	0,99425	0,96758	0,94161
0.2	1,60944	1,56065	1,51413	1,46968	1,42712	1,38629	1,34707	1,30933	1,27297	1,23787
0.1	2,30259	2,20727	2,12026	2,04022	1,96611	1,89712	1,83258	1,77196	1,71480	1,66073
0.0	—	4,60517	3,91202	3,50656	3,21888	2,99573	2,81341	2,65926	2,52573	2,40795

Продолжение табл. XII

c = 1

	c	1	2	3	4	5	6	7	8	9
0,99	0,14855	0,14056	0,13216	0,12326	0,11376	0,10349	0,09222	0,07953	0,06462	0,04540
0,98	0,21470	0,20886	0,20290	0,19678	0,19052	0,18408	0,17745	0,17061	0,16354	0,15620
0,97	0,26753	0,26261	0,25762	0,25256	0,24743	0,24221	0,23690	0,23151	0,22601	0,22041
0,96	0,31357	0,30919	0,30476	0,30029	0,29577	0,29120	0,28658	0,28190	0,27717	0,27238
0,95	0,35536	0,35133	0,34727	0,34318	0,33906	0,33490	0,33071	0,32648	0,32222	0,31792
0,94	0,39419	0,39041	0,38662	0,38280	0,37896	0,37509	0,37120	0,36728	0,36333	0,35936
0,93	0,43081	0,42724	0,42364	0,42002	0,41639	0,41274	0,40907	0,40538	0,40167	0,39794
0,92	0,46575	0,46232	0,45888	0,45542	0,45195	0,44847	0,44497	0,44145	0,43792	0,43438
0,91	0,49933	0,49602	0,49271	0,48938	0,48604	0,48269	0,47932	0,47595	0,47256	0,46916
0,90	0,53181	0,52861	0,52539	0,52217	0,51894	0,51569	0,51244	0,50918	0,50590	0,50262
0,8	0,82439	0,79664	0,76869	0,74050	0,71204	0,68324	0,65405	0,62440	0,59421	0,56339
0,7	1,09735	1,07008	1,04285	1,01565	0,98847	0,96128	0,93406	0,90678	0,87943	0,85197
0,6	1,37642	1,34776	1,31930	1,29104	1,26296	1,23504	1,20727	1,17963	1,15211	1,12469
0,5	1,67835	1,64663	1,61531	1,58435	1,55375	1,52347	1,49351	1,46384	1,43445	1,40531
0,4	2,02231	1,98530	1,94895	1,91323	1,87812	1,84357	1,80955	1,77605	1,74303	1,71047
0,3	2,43922	2,39285	2,34771	2,30371	2,26077	2,21884	2,17786	2,13776	2,09850	2,06003
0,2	2,99431	2,92904	2,86647	2,80635	2,74846	2,69263	2,63870	2,58651	2,53595	2,48688
0,1	3,88972	3,76952	3,65908	3,55685	3,46163	3,37244	3,28852	3,20924	3,13406	3,06254
0,0	—	6,63835	5,83392	5,35595	5,01276	4,74386	4,52218	4,33321	4,16826	4,02172

c = 2

	0	1	2	3	4	5	6	7	8	9
0.99	0.43605	0.41927	0.40137	0.38209	0.36109	0.33786	0.31163	0.28101	0.24320	0.19053
0.98	0.56721	0.55309	0.54464	0.53283	0.52063	0.50798	0.49485	0.48117	0.46687	0.45186
0.97	0.66480	0.65593	0.64689	0.63767	0.62827	0.61867	0.60886	0.59883	0.58855	0.57802
0.96	0.74618	0.73855	0.73033	0.72300	0.71506	0.70701	0.69884	0.69053	0.68210	0.67353
0.95	0.81769	0.81088	0.80400	0.79705	0.79003	0.78293	0.77575	0.76849	0.76115	0.75371
0.94	0.88246	0.87623	0.86995	0.86361	0.85723	0.85078	0.84429	0.83773	0.83111	0.82443
0.93	0.94230	0.93550	0.92879	0.92479	0.91887	0.91291	0.90691	0.90087	0.89478	0.88864
0.92	0.99836	0.99290	0.98741	0.98189	0.97633	0.97075	0.96513	0.95947	0.95378	0.94806
0.91	1.05143	1.04624	1.04102	1.03578	1.03052	1.02523	1.01991	1.01457	1.00919	1.00379
0.90	1.10207	1.09710	1.09211	1.08710	1.08207	1.07702	1.07195	1.06685	1.06173	1.05659
0.8	1.53504	1.49540	1.45522	1.41443	1.37293	1.33064	1.28742	1.24314	1.19763	1.15069
0.7	1.91378	1.87672	1.83953	1.80232	1.76490	1.72730	1.68947	1.65138	1.61298	1.57422
0.6	2.28508	2.24752	2.21013	2.17286	2.13570	2.09863	2.06163	2.02467	1.98772	1.95077
0.5	2.67403	2.63371	2.59375	2.55415	2.51489	2.47594	2.43728	2.39889	2.36074	2.32281
0.4	3.10538	3.05947	3.01427	2.96976	2.92588	2.88260	2.83989	2.79771	2.75603	2.71482
0.3	3.61557	3.55940	3.50458	3.45103	3.39835	3.34738	3.29715	3.24788	3.19954	3.15205
0.2	4.27903	4.20180	4.12758	4.05509	3.98710	3.92040	3.85582	3.79318	3.73234	3.67318
0.1	5.32232	5.18381	5.05517	4.93768	4.82699	4.72305	4.62499	4.53210	4.44379	4.35957
0.0	—	8.40595	7.51660	6.93381	6.59891	6.29579	6.04479	5.82996	5.64175	5.47395

Продолжение табл. XII

с = 3

	0	1	2	3	4	5	6	7	8	9
0.99	0.82325	0.79791	0.77067	0.74108	0.70855	0.67221	0.63058	0.58118	0.51876	0.42855
0.98	1.01624	1.00019	0.98362	0.96845	0.94865	0.92013	0.91082	0.89061	0.86939	0.84699
0.97	1.15504	1.14255	1.12981	1.11679	1.10349	1.08987	1.07591	1.06160	1.04691	1.03180
0.96	1.26832	1.25779	1.24711	1.23626	1.22524	1.21404	1.20266	1.19108	1.17929	1.16728
0.95	1.36632	1.35704	1.34766	1.33817	1.32857	1.31885	1.30901	1.29904	1.28894	1.27871
0.94	1.45398	1.44559	1.43712	1.42857	1.41994	1.41123	1.40243	1.39355	1.38457	1.37549
0.93	1.53414	1.52640	1.51861	1.51076	1.50284	1.49487	1.48683	1.47872	1.47054	1.46230
0.92	1.60858	1.60135	1.59408	1.58677	1.57940	1.57199	1.56452	1.55701	1.54944	1.54182
0.91	1.67850	1.67169	1.66483	1.65794	1.65101	1.64404	1.63703	1.62998	1.62289	1.61576
0.90	1.74477	1.73829	1.73178	1.72523	1.71866	1.71205	1.70541	1.69873	1.69202	1.68528
0.8	2.29579	2.24717	2.19570	2.14529	2.09281	2.03910	1.98399	1.92729	1.86873	1.80802
0.7	2.76371	2.71854	2.67316	2.62753	2.58160	2.53532	2.48865	2.44152	2.39389	2.34566
0.6	3.21132	3.16643	3.12165	3.07694	3.03228	2.98764	2.94300	2.89832	2.85356	2.80871
0.5	3.67206	3.62460	3.57753	3.53081	3.48441	3.43832	3.39249	3.34690	3.30153	3.25635
0.4	4.17526	4.12203	4.06956	4.01780	3.96671	3.91625	3.86638	3.81706	3.76826	3.71994
0.3	4.76223	4.69799	4.63521	4.57379	4.51364	4.45468	4.39683	4.34003	4.28421	4.22930
0.2	5.51505	5.42794	5.34410	5.26323	5.18508	5.10943	5.03605	4.96481	4.89552	4.82803
0.1	6.68078	6.52708	6.38516	6.25319	6.12970	6.01354	5.90377	5.79962	5.70047	5.60576
0.0	—	10.04512	9.08412	8.50525	8.08539	7.75366	7.47817	7.24178	7.03420	6.84873

Продолжение табл. XII

c = 4

	0	1	2	3	4	5	6	7	8	9
0.99	1,27911	1,24572	1,20966	1,17033	1,12685	1,07793	1,02149	0,95384	0,86723	0,73937
0.98	1,52953	1,50894	1,48764	1,46553	1,44255	1,41859	1,39354	1,36726	1,33958	1,31028
0.97	1,70603	1,69027	1,67415	1,65766	1,64078	1,62349	1,60574	1,58751	1,56876	1,54945
0.96	1,84827	1,83511	1,82175	1,80817	1,79436	1,78031	1,76601	1,75145	1,73661	1,72148
0.95	1,97015	1,95866	1,94703	1,93525	1,92332	1,91124	1,89900	1,88659	1,87400	1,86123
0.94	2,07836	2,06803	2,05761	2,04707	2,03643	2,02568	2,01482	2,00384	1,99274	1,98151
0.93	2,17670	2,16723	2,15769	2,14807	2,13837	2,12859	2,11873	2,10877	2,09873	2,08859
0.92	2,26752	2,25873	2,24987	2,24096	2,23198	2,22294	2,21383	2,20465	2,19540	2,18609
0.91	2,35245	2,34418	2,33588	2,32752	2,31911	2,31064	2,30213	2,29355	2,28494	2,27626
0.90	2,43259	2,42477	2,41690	2,40899	2,40105	2,39306	2,38503	2,37695	2,36883	2,36066
0.8	3,08954	3,03116	2,97168	2,91094	2,84878	2,78503	2,71945	2,65178	2,58169	2,50879
0.7	3,63361	3,58136	3,52878	3,47583	3,42246	3,36860	3,31419	3,25916	3,20344	3,14693
0.6	4,14774	4,09646	4,04524	3,99405	3,94285	3,89162	3,84032	3,78890	3,73734	3,68559
0.5	4,67091	4,61727	4,56400	4,51109	4,45849	4,40618	4,35411	4,30227	4,25061	4,19911
0.4	5,23662	5,17702	5,11822	5,06017	5,00281	4,94611	4,89001	4,83448	4,77948	4,72497
0.3	5,89036	5,81910	5,74939	5,68113	5,61422	5,54857	5,48411	5,42075	5,35842	5,29707
0.2	6,72098	6,62526	6,53305	6,44401	6,35789	6,27443	6,19342	6,11468	6,03802	5,96329
0.1	7,99359	7,82659	7,67222	7,52847	7,39379	7,26697	7,14699	7,03303	6,92441	6,82055
0.0	—	11,60463	10,58038	9,96096	9,51037	9,15352	8,85655	8,60129	8,37674	8,17580

Продолжение табл. XII

с=5

	0	1	2	3	4	5	6	7	8	9
0,99	1,78528	1,74435	1,70003	1,65152	1,59771	1,53691	1,46641	1,38136	1,27152	1,10710
0,98	2,03914	2,06436	2,03867	2,01193	1,98420	1,95518	1,92480	1,89286	1,85915	1,82340
0,97	2,30045	2,28166	2,26244	2,24275	2,22259	2,20189	2,18054	2,15879	2,13630	2,11310
0,96	2,46927	2,45371	2,43788	2,42179	2,40542	2,38876	2,37179	2,35449	2,33684	2,31884
0,95	2,61301	2,59949	2,58580	2,57194	2,55788	2,54364	2,52920	2,51455	2,49969	2,48460
0,94	2,73999	2,72790	2,71558	2,70333	2,69086	2,67825	2,66550	2,65261	2,63957	2,62637
0,93	2,85488	2,84384	2,83271	2,82149	2,81016	2,79873	2,78720	2,77557	2,76382	2,75196
0,92	2,96082	2,95040	2,94010	2,92973	2,91928	2,90876	2,89815	2,88746	2,87669	2,86583
0,91	3,05916	3,04959	3,03993	3,03027	3,02051	3,01069	3,00081	2,99087	2,98085	2,97077
0,90	3,15190	3,14285	3,13376	3,12462	3,11543	3,10619	3,09689	3,08754	3,07814	3,06868
0,8	3,90366	3,83740	3,76977	3,70032	3,62974	3,55692	3,48187	3,40429	3,32377	3,23984
0,7	4,51714	4,45852	4,39948	4,33995	4,27989	4,21921	4,15784	4,09570	4,03269	3,96872
0,6	5,09099	5,03397	4,97698	4,91997	4,86291	4,80576	4,74847	4,69101	4,63334	4,57539
0,5	5,67016	5,61097	5,55216	5,49369	5,43553	5,37764	5,31998	5,26252	5,20522	5,14806
0,4	6,29192	6,22662	6,16214	6,09844	6,03547	5,97316	5,91148	5,85039	5,78983	5,72977
0,3	7,00555	6,92799	6,85203	6,77767	6,70469	6,63305	6,56265	6,49341	6,42525	6,35811
0,2	7,90599	7,80255	7,70281	7,60545	7,51316	7,42270	7,33484	7,24936	7,16610	7,08487
0,1	9,27467	9,09572	8,93013	8,77579	8,63107	8,49465	8,36549	8,24271	8,12560	8,01353
0,0	—	13,10848	12,02698	11,37088	10,89255	10,51303	10,19672	9,92442	9,68459	9,46973

Продолжение табл. XII

c = 6

	0	1	2	3	4	5	6	7	8	9
0,99	2,33021	2,28219	2,23008	2,17293	2,10937	2,03734	1,95351	1,85191	1,71989	1,52034
0,98	2,68410	2,65539	2,62561	2,59464	2,56236	2,52862	2,49324	2,45600	2,41665	2,37486
0,97	2,92782	2,90621	2,88410	2,86144	2,83821	2,81436	2,78985	2,76463	2,73864	2,71182
0,96	3,12132	3,10352	3,08542	3,06700	3,04825	3,02916	3,00970	2,98986	2,96961	2,94894
0,95	3,28532	3,26992	3,25432	3,23852	3,22249	3,20625	3,18977	3,17305	3,15607	3,13883
0,94	3,42964	3,41592	3,40205	3,38803	3,37386	3,35953	3,34503	3,33037	3,31554	3,30052
0,93	3,55984	3,54734	3,53474	3,52203	3,50920	3,49626	3,48319	3,47000	3,45668	3,44323
0,92	3,67933	3,66779	3,65617	3,64445	3,63265	3,62076	3,60877	3,59669	3,58450	3,57222
0,91	3,79043	3,77965	3,76880	3,75788	3,74689	3,73592	3,72468	3,71346	3,70216	3,69079
0,90	3,89477	3,88460	3,87438	3,86410	3,85376	3,84336	3,83291	3,82238	3,81180	3,80115
0,8	4,73366	4,66016	4,58506	4,50819	4,42930	4,34815	4,26441	4,17771	4,08760	3,99352
0,7	5,41074	5,34629	5,28132	5,21578	5,14958	5,08266	5,01491	4,94625	4,87657	4,80575
0,6	6,03924	5,97698	5,91471	5,85238	5,78995	5,72738	5,66462	5,60163	5,53835	5,47474
0,5	6,66964	6,60538	6,54149	6,47794	6,41469	6,35170	6,28892	6,22633	6,16388	6,10153
0,4	7,34265	7,27213	7,20246	7,13360	7,06548	6,99806	6,93128	6,86509	6,79945	6,73431
0,3	8,11105	8,02772	7,94611	7,86611	7,78759	7,71046	7,63463	7,56001	7,48652	7,41409
0,2	9,07539	8,96487	8,85824	8,75517	8,65533	8,55847	8,46433	8,37270	8,28339	8,19622
0,1	10,53207	10,34216	10,16629	10,00224	9,84831	9,70312	9,56555	9,43470	9,30980	9,19021
0,0	—	14,57062	13,43638	12,74656	12,24274	11,84240	11,50830	11,22038	10,96654	10,73890

Продолжение табл. XII

c = 7

	0	1	2	3	4	5	6	7	8	9
0,99	2,90611	2,85138	2,79192	2,72659	2,65379	2,57110	2,47461	2,35726	2,20407	1,97081
0,98	3,30712	3,27473	3,24110	3,20610	3,16959	3,13140	3,09131	3,04908	3,00441	2,95691
0,97	3,58126	3,55702	3,53219	3,50675	3,48064	3,45383	3,42626	3,39787	3,36860	3,33838
0,96	3,79788	3,77799	3,75776	3,73716	3,71619	3,69482	3,67303	3,65081	3,62812	3,60495
0,95	3,98082	3,96367	3,94629	3,92867	3,91081	3,89269	3,87431	3,85565	3,83670	3,81745
0,94	4,14137	4,12612	4,11070	4,09512	4,07936	4,06342	4,04730	4,03099	4,01447	3,99775
0,93	4,28584	4,27199	4,25802	4,24392	4,22969	4,21533	4,20083	4,18618	4,17140	4,15646
0,92	4,41817	4,40540	4,39253	4,37957	4,36650	4,35333	4,34006	4,32667	4,31318	4,29957
0,91	4,54098	4,52907	4,51708	4,50501	4,49287	4,48063	4,46832	4,45591	4,44342	4,43084
0,90	4,65512	4,64491	4,63363	4,62229	4,61089	4,59941	4,58787	4,57626	4,56457	4,55281
0,8	5,57606	5,49583	5,41380	5,32974	5,24341	5,15451	5,06268	4,96751	4,86847	4,76494
0,7	6,31217	6,24231	6,17185	6,10072	6,02884	5,95611	5,88244	5,80773	5,73185	5,65468
0,6	6,99137	6,92424	6,85707	6,78981	6,72240	6,65480	6,58697	6,51885	6,45038	6,38151
0,5	7,66925	7,60029	7,53170	7,46344	7,39547	7,32775	7,26023	7,19288	7,12564	7,05849
0,4	8,33977	8,31441	8,23994	8,16629	8,09341	8,02124	7,94973	7,87882	7,80847	7,73863
0,3	9,20895	9,12028	9,03340	8,94819	8,86453	8,78231	8,70145	8,62185	8,54342	8,46608
0,2	10,23254	10,11546	10,00245	9,89315	9,78724	9,68443	9,58447	9,48713	9,39222	9,29954
0,1	11,77091	11,57083	11,38541	11,21235	11,04987	10,89653	10,75116	10,61281	10,48069	10,35412
0,0	—	15,99996	14,81659	14,09537	13,56782	13,14811	12,79750	12,49506	12,22818	11,98868

с = 8

ТАБЛИЦЫ

385

	0	1	2	3	4	5	6	7	8	9
0,99	3,50746	3,44637	3,37992	3,30682	3,22524	3,13240	3,02384	2,89147	2,71803	2,45242
0,98	3,95311	3,91723	3,87996	3,84115	3,80064	3,75823	3,71369	3,66673	3,61701	3,56410
0,97	4,25599	4,22926	4,20188	4,17380	4,14499	4,11537	4,08491	4,05353	4,02116	3,98772
0,96	4,49443	4,47257	4,45032	4,42767	4,40460	4,38108	4,35710	4,33263	4,30764	4,28211
0,95	4,69523	4,67642	4,65736	4,63804	4,61844	4,59856	4,57838	4,55789	4,53708	4,51593
0,94	4,87104	4,85435	4,83749	4,82043	4,80318	4,78573	4,76807	4,75020	4,73211	4,71379
0,93	5,02895	5,01383	4,99856	4,98316	4,96761	4,95192	4,93606	4,92006	4,90389	4,88755
0,92	5,17335	5,15943	5,14540	5,13126	5,11700	5,10263	5,08815	5,07354	5,05880	5,04394
0,91	5,30717	5,29420	5,28115	5,26800	5,25477	5,24144	5,22802	5,21450	5,20089	5,18717
0,90	5,43247	5,42028	5,40801	5,39568	5,38327	5,37078	5,35822	5,34558	5,33286	5,32006
0,8	6,42848	6,34194	6,25340	6,16261	6,06930	5,97313	5,87371	5,77057	5,66316	5,55075
0,7	7,21993	7,14500	7,06939	6,99303	6,91581	6,83765	6,75843	6,67804	6,59635	6,51322
0,6	7,94661	7,87493	7,80317	7,73128	7,65921	7,58691	7,51432	7,44139	7,36805	7,29426
0,5	8,66895	8,59559	8,52260	8,44993	8,37754	8,30539	8,23343	8,16162	8,08990	8,01825
0,4	9,43395	9,35406	9,27509	9,19696	9,11962	9,04301	8,96706	8,89174	8,81698	8,74273
0,3	10,30068	10,20700	10,11519	10,02510	9,93663	9,84966	9,76408	9,67981	9,59675	9,51482
0,2	11,37977	11,25654	11,13755	11,02243	10,91082	10,80244	10,69703	10,59435	10,49419	10,39635
0,1	12,99471	12,78508	12,59071	12,40921	12,23872	12,07774	11,92506	11,77968	11,64079	11,50769
0,0	—	17,40265	16,17308	15,42236	14,87253	14,43465	14,06852	13,75246	13,47337	13,22274

с — 9

	0	1	2	3	4	5	6	7	8	9
0,99	4,13020	4,06306	3,98994	3,90942	3,81945	3,71692	3,59682	3,45006	3,25721	2,96052
0,98	4,61835	4,57916	4,53842	4,49599	4,45167	4,40525	4,35646	4,30500	4,25048	4,19241
0,97	4,94854	4,91945	4,88964	4,85906	4,82766	4,79539	4,76218	4,72795	4,69264	4,65614
0,96	5,20768	5,18395	5,15979	5,13519	5,11013	5,08458	5,05851	5,03191	5,00473	4,97696
0,95	5,42541	5,40503	5,38438	5,36344	5,34220	5,32064	5,29876	5,27654	5,25396	5,23101
0,94	5,61569	5,59764	5,57940	5,56095	5,54228	5,52340	5,50428	5,48494	5,46535	5,44551
0,93	5,78633	5,77000	5,75351	5,73687	5,72008	5,70311	5,68598	5,66868	5,65120	5,63354
0,92	5,94216	5,92714	5,91201	5,89675	5,88137	5,86587	5,85023	5,83446	5,81856	5,80252
0,91	6,08640	6,07243	6,05836	6,04420	6,02993	6,01557	6,00110	5,98653	5,97185	5,95706
0,90	6,22130	6,20818	6,19498	6,18171	6,16835	6,15490	6,14138	6,12777	6,11407	6,10028
0,8	7,28922	7,19673	7,10204	7,00489	6,90497	6,80193	6,69534	6,58468	6,46933	6,34853
0,7	8,13293	8,05321	7,97275	7,89143	7,80918	7,72589	7,64143	7,55568	7,46851	7,37975
0,6	8,90441	8,82844	8,75235	8,67610	8,59963	8,52288	8,44581	8,36834	8,29042	8,21197
0,5	9,66871	9,59120	9,51405	9,43722	9,36067	9,28434	9,20819	9,13217	9,05623	8,98033
0,4	10,47568	10,39152	10,30830	10,22594	10,14439	10,06359	9,98347	9,90397	9,82505	9,74665
0,3	11,38727	11,28888	11,19241	11,09772	11,00471	10,91324	10,82322	10,73454	10,64711	10,56086
0,2	12,51875	12,38972	12,26508	12,14445	12,02748	11,91385	11,80329	11,69556	11,59045	11,48774
0,1	14,20599	13,98734	13,78452	13,59505	13,41699	13,24878	13,08921	12,93720	12,79192	12,65265
0,0	—	18,78312	17,50981	16,73119	16,16028	15,70522	16,32443	14,99549	14,70485	14,44371

Продолжение табл. XII

c = 10

	0	1	2	3	4	5	6	7	8	9
0,99	4,77125	4,69830	4,61881	4,53119	4,43317	4,32136	4,19019	4,02961	3,81810	3,49148
0,98	5,30001	5,25766	5,21361	5,16772	5,11976	5,06951	5,01667	4,96090	4,90178	4,83878
0,97	5,65627	5,62492	5,59279	5,55983	5,52597	5,49116	5,45533	5,41839	5,38026	5,34084
0,96	5,93515	5,90963	5,88365	5,85720	5,83023	5,80274	5,77469	5,74605	5,71679	5,68688
0,95	6,16901	6,14714	6,12498	6,10249	6,07968	6,05653	6,03302	6,00915	5,98489	5,96023
0,94	6,37307	6,35373	6,33418	6,31439	6,29438	6,27413	6,25363	6,23288	6,21187	6,19058
0,93	6,55583	6,53835	6,52070	6,50289	6,48490	6,46674	6,44839	6,42985	6,41113	6,39220
0,92	6,72254	6,70649	6,69030	6,67398	6,65753	6,64095	6,62422	6,60735	6,59033	6,57316
0,91	6,87670	6,86177	6,84674	6,83161	6,81637	6,80102	6,78556	6,76998	6,75429	6,73848
0,90	7,02075	7,00674	6,99265	6,97848	6,96421	6,94986	6,93542	6,92088	6,90625	6,89152
0,8	8,15702	8,05887	7,95833	7,85514	7,74895	7,63938	7,52596	7,40815	7,28527	7,15649
0,7	9,05036	8,96610	8,88102	8,79501	8,70798	8,61981	8,53038	8,43954	8,34715	8,25304
0,6	9,86440	9,78434	9,70414	9,62375	9,54310	9,46214	9,38080	9,29902	9,21674	9,13388
0,5	10,66852	10,58706	10,50597	10,42520	10,34469	10,26439	10,18426	10,10425	10,02430	9,94436
0,4	11,51533	11,42711	11,33985	11,25349	11,16794	11,08316	10,99907	10,91562	10,83275	10,75040
0,3	12,46951	12,36663	12,26574	12,16669	12,06936	11,97363	11,87939	11,78653	11,69496	11,60459
0,2	13,65073	13,51618	13,38618	13,26033	13,13825	13,01963	12,90419	12,79167	12,68185	12,57453
0,1	15,40664	15,17943	14,96858	14,77153	14,58628	14,41123	14,24509	14,08679	13,93544	13,79030
0,0	—	20,14468	18,82975	18,02458	17,43363	16,96222	16,56750	16,22631	15,92470	15,65357

ЛИТЕРАТУРА

- AMS — Annals of Mathematical Statistics
 AS — Applied Statistics
 B — Biometrics
 BSTJ — Bell System Technical Journal
 BISI — Bulletin de l'Institut international de Statistique
 IQC — Industrial Quality Control
 JASA — Journal American Statistical Association
 JIE — Journal of Industrial Engineering
 JQT — Journal Quality Technology
 JRSS — Journal Royal Statistical Society
 P3BS (P5BS) — Proceedings of the third (fifth) Berkeley Symposium on Mathematical Statistics and Probability, California
 S — Sankhya
 SN — Statistika Neerlandica
 T — Technometrics
 TB — Теория вероятностей и ее применения
 TK — Известия АН СССР Техническая кибернетика
 НК — Надежность и контроль качества

Абдурахманов Т.

1. Статистический приемочный контроль по многим признакам, сб. «Теория вероятностей и математическая статистика», Ташкент, «ФАН» (1964), 5—12.
2. Об одном плане непрерывного статистического приемочного контроля, ТашГУ, Ферганский гос. пед. ин-т, Научные труды, вып. 350(2) (1969), 101—111.

Азаров Б. Г., Ольшанский В. К., Чернобровка Л. К.

1. Корректируемые планы статистического приемочного контроля по альтернативному признаку, «Стандарты и качество», 11 (1973), 66—68.

Азгальдов Г. Г., Гличев А. В. и др.

1. Квалиметрия — научная область об оценке качества продукции, сб. «Измерение качества продукции», Изд-во стандартов, 1971, 11—26.

Анскомб (Anscombe F. J.)

1. Tables of sequential inspection schemes to control fraction defective, JRSS, ser. A. CXII, part II (1949), 180—206.
2. Rectifying inspection of a continuous output, JASA 53 (1958), 702—719.

Барнард (Barnard G. A.)

1. Sampling inspection and statistical decisions, JRSS, ser. B 16 (1954), 151—171.

Бейтглер (Beightler C. S.)

1. Design of an optimal sequence of interrelated sampling plans, JASA 59, 305 (1964), 96—104.

Бейли (Bailey N. T. J.)

1. On estimating the size of mobil populations from recapture data, Biometrika 38 (1951), 293—306.

Бейтмен Г., Эрдейи А.

1. Вышие трансцендентные функции (гипергеометрическая функция, функции Лежандра), «Наука», 1973.

Беляев Ю. К.

1. Доверительные интервалы для функций от многих неизвестных параметров, ДАН 169, 4 (1966), 755—758.
2. On confidence intervals and sets for various statistical models, P5BS 1 (1966), 51—58.
3. О стандарте экономичных планов приемочного контроля, «Надежность и долговечность приборов», НИИМАШ, СХІ, 7 (1967), 125—132.
4. Об упрощенных методах построения доверительных границ для надежности систем по результатам испытания компонент, ТК 5 (1968), 96—101.
5. Планы испытаний, гарантирующие минимум предельного числа критических событий, ТК 6 (1971), 60—67.
6. Система экономичных планов приемочного контроля качества продукции по альтернативному признаку, сб. «Статистические методы обеспечения качества и их стандартизация», Материалы XV конф. ЕОКК, V сессия, Изд-во стандартов, 1972, 43—48.
7. Математические модели анализа стандартов приемочного контроля качества, сб. статей «Статистические методы в теории надежности и контроле качества», МЛСМ, Изд-во МГУ, 43 (1973), 38—46.
8. Приемочный контроль по альтернативному признаку, «Знание», вып. I, 1973.
9. Приемочный контроль по альтернативному признаку, «Знание», вып. II, 1973.
10. Критичные планы контроля, ТК 6 (1973), 69—73.

Беляев Ю. К., Колмогоров А. Н.

1. Экономичные планы приемочного контроля, программа IV Всесоюзного математического съезда, Ленинград, 1960.

Беляев Ю. К., Лумельский Я. П.

1. К вопросу об оценке эффективности приемочного контроля при однократной выборке, «Заводская лаборатория» 5 (1968), 571—576.
2. Планы статистического приемочного альтернативного контроля, сб. «Вопросы управления качеством продукции», Изд-во стандартов, 1970, 41—62.

Беляев Ю. К., Рыкова Л. В.

1. Непараметрический критерий Колмогорова для выборок из конечных совокупностей, ДАН 210, 6 (1973), 1261—1264.

Бернштейн С. Н.

1. О «доверительных» вероятностях Фишера, Изв. АН СССР, сер. матем., 5 (1941), 85—94.

Большев Л. Н.

1. Об оценках вероятностей, ТВ 5, 4 (1960), 453—456.
2. Распределения, родственные гипергеометрическому, ТВ 9, 4 (1964), 687—691.
3. О построении доверительных пределов, ТВ 10, 1 (1965), 187—191.
4. О характеристике распределения Пуассона и ее статистических приложениях, ТВ 10, 3 (1965), 488—498.
5. Приложения эмпирического байесовского подхода, сб. «Международный конгресс математиков в Ницце, 1970», «Наука», 1972, 48—55.

Большев Л. Н., Гладков Б. В., Щеглова М. В.

1. Таблицы для вычисления функций B и Z распределений, ТВ 6, 4 (1961), 446—455.

Большев Л. Н., Логинов Э. А.

1. Интервальные оценки при наличии мешающих параметров, ТВ 11, 1 (1966), 94—106.

Большев Л. Н., Смирнов Н. В.

1. Таблицы математической статистики, ВЦ АН СССР, вып. 46, 1968.

Боровков А. А., Маркова Н. П., Сычева Н. М.

1. Таблицы для критериев Н. В. Смирнова однородности двух выборок, Новосибирск, АН СССР, 1964.

Бош (Bosch A. J.)

1. The Polya distribution, SN 17, 3 (1963), 202—213.

Боукер (Bowker A. H.)

1. Continuous sampling plans, P3BS 5 (1956), 75—85.

Вадсворт (Wadsworth Harrison M.)

1. The effects of class inspection under MIL STD-105D, JQT 2, 4 (1970), 181—185.

Вайт (White Leon S.)

1. Markovian decision models for the evaluation of a large class of continuous sampling inspection plans, AMS 36 (1956), 1408—1460.

Валлис (Allen Wallis W.)

1. Techniques of statistical analysis, N. Y., 1947.

Вальд А.

1. Последовательный анализ (перев. с англ.), Физматгиз, 1960.
2. Статистические решающие функции, сб. «Позиционные игры» (перев. с англ.), «Наука», 1967, 300—522.

Вальд, Вольфовиц (Wald A., Wolfowitz I.)

1. Sampling inspection plans for continuous production which insure a prescribed limit of the outgoing quality, AMS 16 (1945), 30—49.

Ван дер Варден Б. Л.

1. Математическая статистика (перев. с немецкого), ИЛ, 1960.
2. Sampling inspection as a minimum loss problem, AMS 31, 2 (1960), 369—384.

Ван Ееден, Ранненбург (van Eeden C., Runnenburg I. T.)

1. Conditional limit distributions for two entries in a 2×2 table, SN 14 (1960), 111—126.

Виттес (Wittes J. T.)

1. On the bias and estimated variance of Chapman's two — sample Capture-recapture population estimate, В 28, 2 (1972), 592—597.

Винце (Vincze I.)

1. On some results and problems in connection with statistics of the Kolmogorov — Smirnov type, P6BS 1 (1972), 459—470.

Гаек Я. (Hajek J.)

1. Limiting distributions in simple random sampling from a finite population, Publ. Math. Inst. Hung. Acad. Sci. 5 (1960), 361—374.

Гаек, Шидак

1. Теория ранговых критериев (перев. с англ.), «Наука», 1967.

Гиршик, Мостеллер, Сэвидж (Girshick M. A., Mosteller F., Savage L. I.)

1. Unbiased estimates for certain binomial sampling problems with applications, AMS 17 (1946), 13—23.

Гихман И. И., Скороход А. В.

1. Теория случайных процессов, т. I, «Наука», 1971.

Гнеденко Б. В.

1. Курс теории вероятностей, Физматгиз, 1961.

Гнеденко Б. В., Корольук В. С.

1. О максимуме отклонения между двумя эмпирическими функциями распределения, ДАН 80 (1951), 525—528.

Гнеденко Б. В., Беляев Ю. К., Соловьев А. Д.

1. Математические методы в теории надежности, «Наука», 1965.

Говиндараюлу (Z. Govindarajulu)

1. Normal approximations to the classical discrete distributions, S, ser. A 27 (1965), 2—4.

Годамб (Godambe V. P.)

1. Admissibility and Bayes estimation in sampling finite population, AMS 40 (1969), 672—676.

Годамб, Джоши (Godambe V. P., Joshi V. M.)

1. Admissibility and Bayes estimation in sampling finite populations I, AMS 36 (1965), 1707—1722.

Горвиц, Томпсон (Horvitz P. G., Thompson D. I.)

1. A generalization of sampling without replacement from a finite universe, JASA 47 (1952), 663—685.

ГОСТы

1. ГОСТ 15895—70. Качество продукции. Термины.
2. ГОСТ 15467—70. Качество продукции. Термины.
3. ГОСТ 16431—70. Качество продукции. Показатели качества и методы оценки уровня качества продукции. Термины и определения.
4. ГОСТ 18247—72. Качество продукции. Статистический приемочный контроль по альтернативному признаку. Одноступенчатые и двухступенчатые корректируемые планы контроля.

Градштейн И. С., Рыжик И. М.

1. Таблицы интегралов, сумм, рядов и произведений, Физматгиз, 1962.

Группа исследований Колумбийского университета.

Statistical Research Group Columbia University, Sampling inspection, N. Y., 1948.

Гудман (Goodman L. A.)

1. Sequential sampling tagging for population size problems, AMS 24 (1953), 56—69.

Гюнтер (Guenter W. C.)

1. Use of binomial, hypergeometric and Poisson tables to obtain sampling plans, JQT 1, 2 (1969).

Дарлинг, Роббинс (Darling D. A., Robbins H.)

1. Finding the size of a finite population, AMS 38 (1967), 1392—1398.

Дворецкий, Кифер, Вольфовиц (Dvoretzky A., Kiefer J., Wolfowitz I.)

1. Sequential decision problems for processes with continuous time parameter, testing hypotheses, AMS 24 (1953), 254—264.

Де Гроот (De Groot M. H.)

1. Unbiased sequential estimation for binomial populations, AMS 30 (1959), 80—101.

Дерман, Джонс, Либерман (Derman C., Johns M. V. J., Lieberman G. J.)

1. Continuous sampling procedures without control, AMS 30 (1959), 1175—1191.

Дерман, Литтауэр, Соломон (Derman C., Littauer S., Solomon H.)

1. Tightened multi-level continuous sampling plans, AMS 28 (1957), 395—404.

Джанардан, Патил (Janardan K. G., Patil G. P.)

1. A unified approach for a class of multivariate hypergeometric models, S, ser. A 34, 4 (1972), 363—376.

Джоши (Joshi V. M.)

1. Admissibility and Bayes estimation in sampling finite population II, AMS 36 (1965), 1723—1729.
2. Admissibility and Bayes estimation in sampling finite populations III, AMS 36 (1965), 1730—1742.
3. A note on admissible sampling designs for a finite population, AMS 42 (1971), 1425—1428.

Додж (Dodge H. F.)

1. A sampling inspection plan for continuous production, AMS 14 (1943), 264—279.
2. A general procedure for sampling inspection by attributes — based on the AQL concept, ASQG Annual Convention Transactions (1963), 7—19.
3. Notes on evaluation of acceptance sampling plans, JQT 1, 2 (1969); 2, 1 (1970), 1—8.

Додж, Ромиг (Dodge H. F., Romig H. G.)

1. Sampling Inspection tables single and double sampling, N. Y., 1944; 2 ed., 1959.

Додж, Стефенс (Dodge H. F., Stephens K. S.)

1. Some new chain sampling inspection plans, IQC 23, 2 (1966).

Додж, Шиллинг (Dodge H. F., Schilling E. G.)

1. Procedures and tables for evaluating dependent mixed acceptance sampling plans, T 11, 2 (1969), 341—372.

Додж, Торрей (Dodge H. F., Torrey M. N.)

1. Additional continuous sampling inspection plans, IQC 14, 11 (1951).

Дружинин Н. К.

1. Выборочный метод и его применение в социально-экономических исследованиях, «Статистика», 1970.

Дугина Т. Н.

1. Программы счета основных характеристик планов выборочного контроля, сб. статей «Статистические методы в теории надежности и контроле качества», МЛСМ, Изд-во МГУ, вып. 43 (1973), 59—117.

Дунин-Барковский И. В., Смирнов Н. В.

1. Теория вероятностей и математическая статистика в технике (общая часть), Гостехиздат, 1955.

Зайдман Р. А., Линник Ю. В., Романовский И. В.

1. Планы последовательного оценивания и марковские моменты остановки, ДАН 185, 6 (1969), 1222—1225.

Закс Ш.

1. Теория статистических выводов (перев. с англ.), «Мир», 1975.

Инамов И.

1. О несмещенных оценках распределения основных показателей статконтроля, Научные труды ТашГУ, Математика, вып. 228, 1963.
2. Оценка результатов контроля, проводимого по планам D_1 , сб. «Теория вероятностей и математическая статистика», Ин-т матем. АН Уз. ССР, 1 (1964), 26—47.
3. Об оценках показателей статконтроля без учета результатов сплошной разбраковки, Учен. записки Ферганского гос. пед. ин-та, Сер. матем., вып. I (1965), 67—74.

Инамов И., Эйдельмант М. И.

1. Об одном применении метода несмещенных оценок к статистическому приемочному контролю, Ташкентский ин-т народного хозяйства. Научные записки «Математика в приложениях», вып. 34 (1970), 14—18.

Ито К.

1. Вероятностные процессы, вып. 1 (перев. с япон.), ИЛ, 1960.

Ишикава, Омае (Ishikawa K. and Ohmae Y.)

1. Study on Severity of Inspection in MIL—STD—105D, Rep. Stat. Appl. Res. JUSE 13, 1 (1966), 1—16.

Йоко Т.

1. Теоретическая основа усовершенствования выборочных планов стандарта MIL—STD—105D, сб. «Выборочный контроль качества», Изд-во стандартов, 1973, 69—79.

Йейтс Ф.

1. Выборочный метод в переписях и обследованиях (перев. с англ.), «Статистика», 1965.

Каган А. М.

1. Об эмпирическом байесовском подходе к задаче оценивания, ДАН 147, 5 (1962), 1020—1021.
2. О схеме Роббинса, ДАН 150, 4 (1963), 733—735.

Каган А. М., Линник Ю. В., Рао С.

1. Характеризационные задачи математической статистики, «Наука», 1972.

- Калленберг (Kallenberg O.)
1. Canonical representations and convergence criteria for processes with interchangeable increments. *Z. Wahrschein. und verw Gebiet* 27, 1 (1973), 23—36.
- Каплан, Макдональд (Kaplan A., MacDonald E.)
1. Instantaneous switching procedure for MIL—STD—105D, *JQT* 1, 3 (1969), 172—174.
- Карлин С. (Karlin S.)
1. Основы теории случайных процессов, «Мир», 1971.
 2. Inequalities for symmetric sampling I, *Annals of statistics* 2, 6 (1974), 1065—1094.
- Кендалл М. Д., Стьюарт А.
1. Теория распределений (перев. с англ.), «Наука», 1966.
- Кифе (Keefe G. J.)
1. Attribute sampling MIL—STD—105D, *IQC* 19 (1961), 7—12.
- Кнайт (Knight W.)
1. A method of sequential estimation applicable to the hypergeometric, binomial, Poisson and exponential distribution, *AMS* 36 (1965), 1494—1503.
- Князева Л. И.
1. Контроль по многим признакам, Ташкентский ин-т народного хозяйства. Научные записки «Математика в приложениях», вып. 55 (1969), 55—67.
- Кокка (Cossa O. A.)
1. MIL—STD—105D — An international standard for attribute sampling, *IQC* 21, 5 (1964).
- Козырева Е. Д.
1. Прогнозирующая несмещенная оценка пропущенной дефектности, сб. «Математические методы контроля качества и надежности», Ташкент, «ФАН» (1969), 43—50.
- Кокс Д., Смит В.
1. Теория восстановления (пер. с англ.), «Советское радио», 1967.
- Колмогоров А. Н.
1. Основные понятия теории вероятностей, «Наука», 1974.
 2. Sulla determinazione di una legge di distribuzione, *Giorn. dell'Institut. Ital. degli Attuari* 4 (1933), 83—91.
 3. Confidence limits for an unknown distribution function, *AMS* 12 (1941), 461—463.
 4. Несмещенные оценки, *Изв. АН СССР, сер. матем.* 14, 4 (1950), 303—326.
 5. Статистический приемочный контроль при допустимом числе дефектных изделий, равно нулю, Дом н.-т. пропаганды, Ленинград, 1951.
 6. Обобщение формулы Пуассона на случай выборки из конечной совокупности, *УМН* 6, 3 (1951), 133—134.
- Колчин В. Ф., Чистяков В. П.
1. Об одной комбинаторной предельной теореме, *ТВ* 18, 4 (1973), 767—777.
 2. О предельных распределениях одной статистики, *ТВ* 19, 2 (1974), 367—374.
- Корвар, Серфлинг (Korwar R. M., Serfling R. J.)
1. On averaging over distinct units in sampling with replacement, *AMS* 41 (1970), 2132—2134.

Коуден Д.

1. Статистические методы контроля качества (перев. с англ.), Физматгиз, 1961.

Кояма Т.

1. Анализ динамических характеристик регулирования уровня контроля на основе военного стандарта США MIL—STD—105D, сб. «Выборочный контроль качества», Изд-во стандартов, 1973, 51—60.

Крамер Г.

1. Случайные величины и распределения вероятностей (перев. с англ.), ГИИЛ, 1947.
2. Математические методы статистики (перев. с англ.), ИЛ, 1948.

Левонтин, Праут (Lewontin R. C., Praut T.)

1. Estimation of the number of different classes in a population, В 12 (1956), 211—223.

Леман Э.

1. Проверка статистических гипотез (перев. с англ.), «Наука», 1964.

Либерман (Lieberman G. J.)

1. A note on Dodge's continuous inspection plan, AMS 24 (1953), 480—484.

Либерман, Оуэн (Lieberman G. J., Owen D. B.)

1. Tables of the hypergeometric probability distribution, Stanford Univ. Press. Stanford (California), IV, 1961.

Либерман, Соломон (Lieberman G. J., Solomon H.)

1. Multi-level continuous sampling plans, AMS 26 (1955), 686—704.

Линдли, Барнет (Lindley D. V., Barnett B. N.)

1. Sequential sampling: two decision problems with linear losses for binomial and normal random variables, В 52 3, 4 (1965), 507—532.

Линник Ю. В.

1. Статистические задачи с мешающими параметрами, «Наука», 1966.

Линник Ю. В., Романовский И. В.

1. К теории последовательного оценивания, ДАН 194, 2 (1970), 270—272.
2. Some new results in sequential estimation theory, P6BS 1 (1971), 85—96.

Линник Ю. В., Рухин А. Л.

1. Выпуклые функции потерь в теории несмещенного оценивания, ДАН 198, 3 (1971), 527—529.

Лоэв М.

1. Теория вероятностей (перев. с англ.), ИЛ, 1962.

Лумельский Я. П.

1. Несмещенные оценки для планов типа последовательного анализа, Учен. записки Пермского гос. ун-та, Матем., вып. 131 (1966), 140—150.
2. К вопросу о несмещенных оценках, Учен. записки Пермского гос. ун-та, Матем., вып. 131 (1966), 98—103.
3. Несмещенные достаточные оценки вероятностей в случае многомерного нормального закона, Вестник МГУ, Матем. 6 (1968), 14—17.

4. Несмещенные оценки пропущенного брака по увеличенной выборке. Сб. «Математические методы контроля качества и надежности», Ташкент, «ФАН» (1969), 29—31.
 5. Доверительные границы для линейных функций от неизвестных параметров, ТВ 14, 2 (1969), 381—384.
 6. Оценка эффективности статистического приемочного альтернативного контроля, сб. «Вопросы управления качеством продукции», Изд-во стандартов (1970), 63—78.
 7. Несмещенные оценки априорного распределения и их применение, ТК 1 (1973), 80—89.
 8. Случайные блуждания, отвечающие обобщенным урновым схемам, ДАН 209, 6 (1973), 1281—1284.
- Лумельский Я. П., Сапожников П. Н.
1. Несмещенные оценки для плотностей распределений, ТВ 14, 2 (1969), 372—380.
- Лундберг (Lundberg O.)
1. On random processes and their application to sickness and accident statistics, Almqvist and Wiksell, Uppsala, 1940.
- Мерфи (Murphy R. B.)
1. Stopping rules with CSP-1 sampling inspection plans in continuous production, IQC 1 (XVI), 5 (1959), 5—12.
- Методика СЭП
1. Методика по разработке стандартов на статистический приемочный контроль качества продукции по альтернативному признаку с учетом экономических показателей, Изд-во стандартов, 1972, 1—232.
- МИЛ — STD — 105 (MIL — STD — 105).
1. Military Standard 105A, sampling procedures and tables for inspection by attributes, US Government Printing Office, Washington, 1950.
 2. Military Standard 105C, sampling procedures and tables for inspection by attributes, US Government Printing Office, Washington, 1961.
 3. Military Standard 105D, sampling procedures and tables for inspection by attributes, US Government Printing Office, Washington, 1963.
- Моленар (Molenaar W.)
1. Approximations to the Poisson, binomial and hypergeometric distribution functions, Math. Centre Tracts 31, Amsterdam, 1970.
- Мостеллер Ф., Рурке Р., Томас Д.
1. Вероятность (перев. с англ.), «Мир», 1969.
- Муд (Mood A. M.)
1. On the dependence of sampling inspection plans upon population distributions, AMS 14 (1943), 415—425.
- Невё Ж.
1. Математические основы теории вероятностей (перев. с франц.), «Мир», 1969.
- Нейман (Neyman T.)
1. Sur un teorems concerente le considette statistiche sufficienti, Inst. Ital. Atti. Giorn. 6 (1935), 320—334.
 2. Опыты, выясняющие возможность активного воздействия на осадки, и некоторые связанные с ними задачи математической

- статистики, Изв. АН Уз. ССР, Сер. физико-математическая, 1 (1974), 15—27.
- О л ь ш а н с к и й В. К.
1. Оптимальный метод статистического регулирования технологического процесса. Автоматика и телемеханика, 4 (1973), 195—197.
 2. Система экономических планов выборочного приемочного контроля качества, НК 1 (1973), 61—69.
- О м а э И., С у г а Р.
1. Разработка усовершенствованной схемы выборочного контроля на основе стандарта MIL — STD — 105D, сб. «Выборочный контроль качества», Изд-во стандартов, 1973, 61—69.
- О у э н Д. Б.
1. Сборник статистических таблиц (перев. с англ.), Изд-во ВЦ АН СССР, 1966.
- П а б с т (W. R. Pabst)
1. «MIL — STD — 105D», IQC, 20, 5 (1963), 4—9.
 2. Sampling inspection plans by attributes with severity adjustment, JQT 3 (1971), 89—94.
- П а в л о в И. В.
1. Доверительная оценка надежности системы по результатам испытаний компонент, ТК 1 (1973), 51—61.
 2. Вычисление доверительных границ для функций многих неизвестных параметров, ТК 2 (1972), 89—92.
- П а т и л, Д ж о ш и (Patil G. P., Joshi S. W.)
1. A dictionary and bibliography of discrete distributions, Publishers Oliver and Poyd, Edinburgh, 1968.
- П и р с о н (Pearson K.)
1. Tables of the incomplete Beta-function, London, 1934.
- Р а о С. Р.
1. Линейные статистические методы и их применения (перев. с англ.), «Наука», 1968.
- Р е з н и к о в (Resnikoff G. J.)
1. Minimum average fraction inspected for a continuous sampling plan, JIE 11, 3 (1960), 208—209.
- Р е к о м е н д а ц и я СЭВ, РС 966—67, Статистический контроль качества, термины, определения и обозначения.
- Р е н ь и (Renyi A.)
1. Three new proofs and generalization of a theorem of Irving Weiss. Publ. Math. Inst. Hung. Acad. Sciences, Ser. A. 7 (1962), 203—214.
- Р о б б и н с (Robbins H.)
1. The empirical bayes approach to statistical decision problems, AMS 35 (1964), 1—20.
 2. Эмпирический байесовский подход в статистике, сборник переводов, «Математика» 8, 2 (1964), 133—140.
- Р о з э н (Rosén B.)
1. Limit theorems for sampling from finite populations, Ark. Mat. 5 (1964), 383—424.
- Р о м а н о в с к и й В. И.
1. Дискретные цепи Маркова, Гостехиздат, 1949.
- Р о н ж и н В. И.
1. О критерии эффективности приемочного статистического контроля, Труды Ин-та матем. и механики АН Уз. ССР, вып. 17 (1956).

2. Дисперсия несмещенных оценок распределений предъявленного и пропущенного брака. Ташкентский ин-т народного хозяйства. Научные записки «Математика в приложениях», Ташкент (1969), 68—78.
 3. Несмещенные оценки распределения предъявленного и пропущенного брака по результатам статистического контроля, сб. «Вероятностные модели и статистический контроль», Ташкент, «ФАН», (1968).
 4. Приложение несмещенных оценок распределений предъявленного и пропущенного брака к выбору плана статистического контроля. Ташкентский ин-т народного хозяйства. Научные записки «Математика в приложениях», вып. 34 (1970), 112—121.
- Рутерфорд, Крючков** (Rutherford J. K., Krutchkoff R. G.)
1. Some empirical Bayes techniques in point estimation, *Biometrika* 56, 1 (1969), 133—137.
- Рыкова Л. В.**
1. Выборочный контроль по альтернативному признаку с учетом возможности появления серий дефектных изделий, *ТК* 2 (1973), 66—74.
 2. Построение доверительных границ для функции распределения конечной совокупности, сб. статей «Статистические методы в теории надежности и контроле качества», *МЛСМ*, Изд-во МГУ, вып. 43 (1973), 47—58.
 3. Непараметрический критерий Колмогорова — Смирнова для выборок из конечных совокупностей, *ДАН* 214 (1974), 43—46.
- Самуэль** (Samuel E.)
1. Sequential maximum likelihood estimation of the size of population, *AMS* 39 (1968), 1057—1068.
 2. Comparison of sequential rules for estimation of the size of a population, *В* 25, 3 (1969), 517—527.
- Севастьянов Б. А.**
1. Доверительные пределы для средней вероятности в схеме Пуассона, *ТВ* 6, 2 (1961), 242—244.
- Севастьянов Б. А., Сираждинов С. Х.**
1. Математическая статистика и контроль промышленной продукции, «Природа», 8 (1955), 28—34.
- Секкаппан, Томпсон** (Sekkappan Rm., Thompson M. E.)
1. On a class of uniformly admissible estimators for finite populations, *Annals of Statistics* 3, 2 (1975), 492—499.
- Сен** (Sen P. K.)
1. The Najeck — Renyi inequality for sampling from a finite population, *S*, ser. A 32, part 2 (1970), 181—188.
 2. Finite population sampling and weak convergence to a brownian Bridge. *S*, ser. A 34, 1 (1972), 85—90.
- Серфлинг** (Serfling R. J.)
1. Probability inequalities for the sum in sampling without replacement *Annals of Statistics* 2 (1974), 39—48.
- Сираждинов С. Х.**
1. Одинарный статистический контроль, *Труды ин-та матем. и механики АН Уз. ССР*, вып. 15 (1955), 41—56.

2. Об оценках с наименьшим смещением при биномиальном распределении, ТВ 1, 1 (1956), 168—174.
 3. Методы статистического контроля оценки качества продукции, «Стандартизация», 1 (1956).
 4. Несмещенные оценки при распределении по закону Пуассона и некоторые применения, ДАН Уз. ССР 6 (1956), 3—5.
 5. Несмещенные оценки доли пропущенного брака при методе: однократной выборки, Труды Ин-та матем. и механики АН Уз. ССР, вып. 20 (1957), 89—100.
 6. К вопросу рационального выбора плана статистического приемочного контроля, Труды ТашГУ, вып. 189 (1961), 79—88.
- Сираждинов С. Х., Абдурахманов Т.
1. Статистический приемочный контроль по многим признакам для $N \rightarrow \infty$, сб. «Теория вероятностей и математическая статистика», вып. 1, Ташкент (1964), 13—25.
- Сираждинов С. Х., Эйдельмант М. И.
1. К вопросу об оценках качества продукции по результатам выборочного контроля, Тр. Ин-та матем. АН Уз. ССР, Ташкент, 22 (1961), 135—145.
- Сирли (Searle S. R.)
1. Linear models, N. Y., 1971.
- Ситтиг (Sittig J.)
1. The Economic choice of sampling systems in Acceptance sampling, BISI 33 (1951), 51—84.
- Стандарты статистического контроля (проект РТМ), Ташкент, Ин-т математики им. Романовского В. И. АН, Уз. ССР (1961).
- Скотт (Scott A. J.)
1. On admissibility and uniform admissibility infinite population sampling, Annals of Statistics 3, 2 (1975), 489—491.
- Смирнов Н. В.
1. Теория вероятностей и математическая статистика, Избр. труды, «Наука», 1970.
- Стефенс, Ларсон (Stephens K. S., Larson K. E.)
1. An evaluation of the MIL—STD—105D system of sampling plans, IQC 23, 7 (1967), 310—319.
- Субрахманья (Subrahmanya M. T.)
1. Error areas under the operating characteristic curves, S, ser. A, 28, 1 (1966), 71—86.
- Судаков Р. С.
1. К вопросу об интервальном оценивании показателя надежности последовательной системы, ТК 3 (1974), 86—90.
- Сэвидж (Savage I. R.)
1. A production model and continuous sampling plan, JASA 54 (1959), 231—247.
- Тага (Taga Yasushi)
1. A note on the degree of normal approximation to the distribution function of the mean of samples from finite populations, Annals of the Inst. of Statist. Math. 3 (1964), 427—430.
- Типпетт (Tippett L. H. C.)
1. Guide to acceptance sampling, AS 7 (1958), 133—48.

Томлинсон (Tomlinson P. K.)

1. Some sampling problems in fishery work, В 27, 3 (1971), 631—641.

Тюрин Ю. Н.

1. О проверке параметрических гипотез непараметрическими методами, ТВ 15, 4 (1970), 745—750.

Уилкс С.

1. Математическая статистика (перев. с англ.), «Наука», 1967.

Феллер В.

1. Введение в теорию вероятностей и ее приложения, том 1 (перев. с англ.), «Мир», 1967.
2. Введение в теорию вероятностей и ее приложения, том 2 (перев. с англ.) «Мир», 1967.

Фишер (Fisher R. A.)

1. On the mathematical foundation of theoretical statistics, Phil. Trans. Roy. Soc., Ser. A, 222 (1921), 309—368.
2. The fiducial argument in statistical inference, Annals of Eugenics 5 (1935), 391—398.

Фриман, Вейс (Freeman David, Weiss Lionel)

1. Sampling plans which approximately minimize the maximum expected sample size, JASA (1964), 67—88.

Фукао (Fukao Kichiji)

1. Several problems in designing sequential sampling inspection plans by control point and relative slope, Rep. Stat. Appl. Res., JUSE 10, 4 (1963), 31—54.

Хагберг (Hagberg J.)

1. Approximation of the summation process obtained by sampling from a finite population, ТВ 18, 4 (1973), 790—803.

Хальд (Hald A.)

1. Математическая статистика с техническими приложениями (перев. с англ.), ИЛ, 1956.
2. The compound hypergeometric distribution and a system of single sampling inspection plans based on prior distribution and costs, Т 2, 3 (1960), 275—340, 370—372.
3. Some limit theorems for the Dodge—Romig LTPD single sampling inspection plans, Т 4 (1962), 497—513.
4. Efficiency of sampling inspection plans for attributes, BISI 40 (1964), 681—697.
5. Bayesian single sampling attribute plans for discrete prior distributions. Mat. Fys. Skr. Dan. Vid. Selsk. 3, 2 (1956), 88.
6. Single sampling inspection plans with specified acceptance probability and minimum costs, Skandinavisk Aktuarietidskrift (1966), 48.
7. On the theory of single sampling inspection by attributes based on two quality levels, Rev. Int. Statist. Inst. 35, (1967), 17—45.
8. Asymptotic properties of Bayesian single sampling plans, JRSS, B. 29 (1967), 162—173 (corrigenda in 29, 3, p. 586).
9. The determination of single sampling attribute plans with given producer's and consumer's risk, Т 9, 3 (1967), 401—441.

10. Bayesian single sampling attribute plans for continuous prior distributions, Т 10 (1968), 667—683.
 11. The mixed binomial distribution and the posterior distribution of p for a continuous prior distribution, JRSS B. 30 (1968), 359—367.
- Хальд, Коузгард (Hald A., Kousgaard E.)
1. Some limit theorems for the Dodge—Romig AOQL single sampling inspection plans, S, ser., A 25 (1963), 255—268.
- Хальд, Тирегод (Hald A., Thyregod P.)
1. Bayesian single sampling plans based on linear costs and the poisson distribution, København, 1971, 1—100.
- Хамакер (Hamaker H. C.)
1. Economic principles in industrial sampling problems: a general introduction, BISI, Inst. 33 (1951), 105—122.
 2. Some basic principles of sampling inspection by attributes, AS 7 (1958), 149—159.
 3. Attribute sampling in operation, BISI 37 (1960), 265—281.
- Хеффдинг (Hoeffding W.)
1. Probability inequalities for sums of bounded random variables, JASA 58 (1963), 13—30.
- Хилл (Hill J. D.)
1. The Economic Incentive provided by sampling inspection AS 9, 2 (1960), 69—81.
 2. Sampling inspection and defence specification, DEF—131 (with discussion), JRSS. Ser A 125, 1 (1962), 31—87.
- Хиллиер (Hillier F. S.)
1. New criteria for selecting continuous sampling plans Т 6, 2 (1964), 161—178.
- Хорснелл (Horsnell G.)
1. The determination of single—sample schemes for percentage defectives, AS 3 (1954), 150—158.
 2. Economical acceptance sampling schemes, JRSS, ser. A 120 (1957), 148—201.
- Хэнсен Б.
1. Контроль качества, «Прогресс», 1968.
- Чамперноун (Champernowne D. G.)
1. The economics of sequential sampling Procedures for defectives, AS 2 (1953), 118—130.
- Чепман (Chapman D. G.)
1. Some properties of the hypergeometric distribution with applications to zoological censuses, Univ. Calif. Publ. Statist. 1 (1951), 131—160.
- Чепурин Е. В., Дугина Т. Н.
1. Об одном классе устойчивых статистических оценок в теории надежности, ТК 1 (1970), 72—79.
- Чжун Кай Лай
1. Однородные цепи Маркова, «Мир», 1964.
- Шиндовский Э., Шюрц О.
1. Статистические методы контроля производства (перев. с нем.), Изд-во Комитета стандартов, 1969.
- Ширяев А. Н.
1. Статистический последовательный анализ, «Наука», 1969.

Шор Я. Б.

1. Статистические методы анализа и контроля качества и надежности, «Советское радио», 1962.
2. Статистические методы контроля качества и надежности промышленной продукции, Изд-во «Знание», вып. 1—4, 1968.
3. Методы комплексной оценки качества продукции, Изд-во «Знание», 1971.
4. Приближенные формулы для расчетов при гипергеометрическом распределении, НК 2 (1972), 20—28.
5. Стандарт на статистический приемочный контроль по альтернативному признаку с последующими оценками, сб. «Статистические методы обеспечения качества и их стандартизация», Материалы XV конференции ЕОКК, сессия V, Изд-во стандартов, 1972, 21—27.

Шор Я. Б., Дубров А. М.

1. Таблицы для планирования эксперимента при гипергеометрическом распределении, НК 6 (1971), 60—67.

Шор Я. Б., Кузьмин Ф. И.

1. Таблицы для анализа и контроля надежности, «Советское радио», 1968.

Шор Я. Б., Пахомов А. А.

1. Непрерывный выборочный контроль, НК 11 (1969), 3—17.
2. Выбор плана одностадийного статистического контроля непрерывного потока продукции, НК 8 (1970), 25—31.
3. Методы выбора плана непрерывного статистического контроля качества по альтернативному признаку, НК 1 (1973), 10—22.

Шривастава (Srivastava S.)

1. A generalized estimator for the mean of a finite population using multi-anxiliary information, JASA 66, 334 (1971), 404—407.

Эйдельмант М. И.

1. К вопросу о рациональных методах контроля продукции, Труды Ин-та матем. и механики им. В. И. Романовского АН Уз. ССР, вып. 15 (1955), 57—68.
2. Эффективность статистического контроля типа Д. Труды Ин-та матем. и механики им. В. И. Романовского АН Уз. ССР, вып. 17 (1956), 25—44.
3. Приближенные формулы для гипергеометрического распределения, Изв. АН Уз. ССР, сер. физ.-матем. Н., 5 (1958), 79—92.

Эйдельмант М. И., Бектаев К. Б.

1. Экономическая эффективность приемочного статистического контроля, Тр. Ин-та матем. и механики им. В. И. Романовского АН Уз. ССР, вып. 20 (1957), 101—126.

Эйдельмант М. И., Инамов И.

1. К вопросу о применении метода несмещенных оценок дискретных распределений для контроля качества продукции (тезисы доклада), VII Всесоюзное совещание по теории вероятностей и математической статистике, Ташкент (1963), 1—9.

2. О практическом применении оценочных функций для анализа результатов статистического контроля качества продукции, сб. «Теория вероятностей и математическая статистика», Ин-т матем. и механики АН Уз. ССР, 1 Ташкент (1964), 48—57.
- Э р д е ш, Р е н ь и (Erdős P., Renyi A.)
1. On a central limit theorem from a finite population, Publ. Math. Inst. Hung. Acad. Sci. 4 (1959), 49—61.
- Э р и к с о н (Ericson W. A.)
1. On a class of uniformly admissible estimators of a finite population total, AMS 41 (1970), 1369—1372.
- Ю с а с (Jusas Jonas)
1. Atrankines kontroles efektyvumo ivertinimas, Mokslas ir technika, 11 (1972).
 2. О распределении максимальной частоты в полиномиальной схеме, ТВ XVII, 4 (1972), 755—761.
- Я н к о Я.
1. Математико-статистические таблицы, Госстатиздат, 1961.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

- Альтернирующий процесс восстановления 72, 246
- Аппроксимация гипергеометрического распределения 43
- Асимптотическая нормальность выборочного среднего 293
 - гипергеометрического распределения 41, 323
 - последующих оценок 32, 152
- Атом случайной величины 64

- Байесовский подход 112
 - эмпирический 123
- Биномиальная выборка 87, 145, 149, 172, 294
- Блуждание случайное полиномиальное 172
- Брак предьявленный 29, 52, 129, 138, 146, 153
 - суммарный 32
 - пропущенный 29, 52, 138, 146, 156
 - суммарный 32
 - средний принятый 93

- Вариационная диаграмма 356
- Вариационный ряд 309
- Вероятностная мера 64, 69
 - условная 66
- Вероятностное пространство 64
- Вероятность априорная 157
 - бракования 18
 - включения 344
 - ложного переключения 200
 - напрасного отвержения 79
 - приемки 26
 - пути 88
 - события 64, 69
 - условная 66
- Взаимно независимые случайные величины 15, 65, 32
 - события 65
- Воспроизводимость распределения 102

- Гипотеза (статистическая) конкурирующая 79
 - независимости 15
 - нулевая (основная) 79
 - простая (сложная) 79
- Граница γ -доверительная 149, 151
 - однозначная 285
 - плана 97, 177, 285

- Динамические характеристики 204
- Дисперсия 15
 - оценки 30, 37, 345
- Допустимая оценка 346
 - пара (план, оценка) 349
 - равномерно предпочтительная 349

- Индекс плана 90
 - точки 90
- Индикатор множества 66
- Интервалы γ -доверительные 146
- Исключение граничных точек 179

- Качество безразличное 96
- Класс планов одноступенчатых 27
 - первого вхождения 83
 - эквивалентных 216, 219
- Классы правдоподобия 73
 - решающих функций 81
 - эквивалентных траекторий 317
- Контроль направленный 252
 - неразрушительный 188
 - нормальный (сокращенный, усиленный) 199
 - по признаку альтернативному 11, 14
 - — качественному 11, 167
 - — количественному 11
 - последовательный 81
 - разрушающий 19
 - с проверкой дополнительного объема 57
 - сплошной 19, 243
- Коэффициент доверия системы множеств 78, 151
- Критерий отношения правдоподобия 80
 - проверки гипотез 79
 - факторизации 74
- Критическая область 80
- Критический уровень значимости критерия 80
 - качества 19
- Критическое число дефектных объектов 212

- Математическая модель процесса контроля 12
 - — производства 200, 206
 - — — Марковская 253
 - — статистическая 72

Математическое ожидание 15, 21, 37, 66
 — условное 67, 101
 Метод максимума правдоподобия 78
 — моментов и квантилей 78
 — прямого (обратного) счета числа путей 91
 — усреднения по достаточной статистике 76
 Множества p -доверительные 78
 Моменты случайной величины 70
 — факториальные 116
 — смешанные 38, 110, 170
 — центральные 40, 320
 Мощность носителя априорного распределения 164, 165

Независимые случайные величины 127
 Нормальный ход производства 183, 207, 219, 222, 238, 261

Общая дисперсия 290
 — ковариационная матрица 307
 — общее среднее 290, 307, 345, 351
 Объект годный (дефектный) 9, 12, 14, 264
 Объем выборки 14, 20, 28, 190, 274
 — — средний 344
 — контроля средний 92, 223
 — совокупности 274
 Оперативная характеристика 94, 240
 — направленного контроля 255
 — — одноступенчатого плана 44, 48
 — — результирующая 203
 Оптимальный план этапных испытаний 265
 Отказы критические 264, 265
 Оценка 28, 344
 — Верхняя (нижняя) 144
 — Горвица — Томпсона 345, 346
 — дисперсии 53, 54, 131
 — линейная 344
 — несмещенная 29, 51, 56, 58, 75, 127, 138, 176, 283, 286, 345, 351, 358
 — — априорных вероятностей 159
 — общего среднего совокупности 289, 305
 — точечная 75
 — функции 51, 176
 Оценки последующие 31, 49, 54, 59
 — — интервальные 145
 — согласованные 55
 Ошибка второго рода (поставщика) 45, 94, 226
 — первого рода (потребителя) 45, 79, 94, 224

Переключающие правила 199, 202

Планы 63
 — бездефектного контроля 20, 25, 27
 — выборочного контроля 81
 — гарантийные 189
 — двухступенчатые 84, 196, 221, 274
 — двухэтапные 266
 — допустимые 350
 — замкнутые 125, 176, 281
 — критичные 243

Планы многоступенчатые 196
 — многоступенчатые 270
 — направленного контроля 251
 — НВК-1, НВК-2 (выборочного контроля непрерывного потока объектов) 234
 — непрерывные 264
 — ограниченные (неограниченные) 284
 — одноступенчатые с двумя уровнями 50, 54
 — — с одним уровнем 44, 49, 59, 84, 89, 116, 153, 196, 229
 — первого вхождения 83, 160, 177, 197, 282
 — — с двумя решениями 83, 94, 138, 189
 — — с невозрастающей границей 97
 — полные 132, 177
 — последовательные 81, 84, 230, 277
 — простые (непростые) 133, 177
 — равномерно предпочтительные 350
 — рациональные 221, 228, 231
 — с дополнительным контролем 57, 139
 — семиступенчатые 196
 — симметричные 340
 — эквивалентные 216
 — — одноступенчатые 217
 — — двухступенчатые 222
 — этапные 264
 — — оптимальные 265
 Плотность вероятностей 70
 Полнота плана 132
 — семейства вероятностных мер 177, 285
 Порядок точек границы 97
 Последовательность альтернирующая 259
 — Бернулли 99
 — марковская 253
 Предел среднего выходного качества 27
 Предельно допустимое число дефектных объектов 95
 Предельное значение среднего числа дефектных объектов 26
 — число критических отказов 265
 Приемлемое число дефектных объектов 95
 Приемочное число 20, 44, 190, 221, 255
 Приемочный уровень качества 187, 195
 Пространство измеримое 69
 Процедура окружения 252
 Процентные точки 45, 47
 Процесс биномиальный 246
 — винеровский 319
 — — условный 320
 — восстановления 72, 235
 — — альтернирующий 72
 — гипергеометрический 315
 Путь контроля 82

Разбиение пространства исходов 68
 — совокупности 353
 Разладки производства 183
 Разлаженный ход производства 253
 Размер критерия 79

- Распределение априорное 39, 98, 102, 114, 157, 183, 191, 221
 — биномиальное 13, 14, 99, 102, 147, 217, 220, 245
 — бэта 70, 104, 311
 — вырожденное 40, 302
 — гамма 70, 104
 — гипербиномиальное 111
 — гипергеометрическое 33, 100, 190, 265, 274, 310
 — — многомерное 170
 — — отрицательное 98
 — G -биномиальное 102, 185
 — Дирихле 175
 — Колмогорова 333
 — нормальное 15, 40, 152, 293
 — Пойа 104
 — — многомерное 174
 — полиномиальное 172
 — пуассоновское 40, 190, 217, 266
 — равномерное 207
 — суббиномиальное 111
 — хи-квадрат 71
 — эквибиномиальное 111
 Расслоение 353
 Расстояние по вариации 305
 Решающая функция (процедура) 18, 75
 — — допустимая (недопустимая) 81
 Решение о забраковании совокупности 14
 — о приемлемости совокупности 16
 Решения $\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3$ 49, 85, 258
 Риск поставщика (потребителя) 95, 258
- Система γ -доверительных интервалов 146
 — планов 181, 205
 — — Доджа — Ромига 194
 — — экономичных 210
 — случайных множеств 78
 Слон совокупности 353
 Случайная величина 13, 64
 — выборка 14
 — — без возвращения (бесповторная) 87, 91, 99, 100, 289, 306
 — — с возвращением 303
 — функция 76
 Случайные множества 78
 Случайный выбор без возвращения 14, 87, 91, 99, 289, 306, 353
 — отбор с маркировкой 280
 Смеси полиномиальных распределений 173
 Событие 64, 69
 Совокупность объектов 9, 12
 Среднее запаздывание 200
 — превышение предела дефектности 242
 Среднеквадратическое отклонение 32, 76
 Средний уровень выходного качества 235
 — — дефектных объектов 15
 — — расходов на контроль (относительный) 227
- Средняя доля контролируемых объектов 235
 Стандарты статистического приемочного контроля 181
 Статистика 15, 73, 306
 — достаточная 73, 89, 282
 — наблюдаемая ранговая 340
 — тестовая 80
 Стирлинга формула 24
 Стоимость контроля 26
 Схема испытаний Бернулли 64, 77
 — Пойа 104
 — — обобщенная 174
 Сходимость в основном 71
 — по вероятности 72
 — с вероятностью единица 72
- Точка граничная 83, 169
 — достижимая 88, 130, 169
 — недостижимая 88
 — остановки 89, 130, 282
 — проходимая 88, 131, 169, 176
 Траектории процесса 315
 — l -эквивалентные 317
- Уровень значимости критерия 80
 — качества 96
 — — браковочный 188
 — — приемочный 188, 195
 — контроля нормального (сокращенного, усиленного) 199
 — предельного выходного качества 96
 — — среднего выходного качества 188, 236
- Формула полного условного математического ожидания 68
 — полной вероятности 14, 69
 Функция восстановления 72
 — мощности 79
 — потеря 80
 — правдоподобия 73
 — производящая 34, 261
 — распределения 12, 16, 70, 306
 — — эмпирическая 313, 335, 358
 — решающая 75
 — риска 80
 — симметрии 292
 — фильтрации 219
- Центральная предельная теорема 71
 Цепь Маркова 207, 315
- Число браковочное 197
 — выборок 33
 — приемочное 197, 223
 — путей 176
- Этап 252
 Эффективность планов 241

СПИСОК ОБОЗНАЧЕНИЙ

$A = \{x : G\}$ — множество элементов x , удовлетворяющих условию G .

$P(A)$ — вероятность события.

Ef — математическое ожидание случайной величины.

$E(f|g)$ — условное математическое ожидание f относительно g .

$\hat{h}$ — оценка функции h .

$V(\hat{h})$ — дисперсия оценки $\hat{h}$.

$a(x, c) = a(a+c) \dots (a+(x-1)c)$.

$a^{(b)} = a(a-1) \dots (a-b+1)$, $b=1, 2, \dots$

$\binom{n}{d} = n^{(d)}/d!$

$$h_{N,D}^{n,d} = \binom{n}{d} \frac{D^{(d)}(N-D)^{(n-d)}}{N^{(n)}}, \quad H_{N,D}^{n,d} = \sum_{r=0}^d h_{N,D}^{n,r},$$

$$b_D^d(v) = \binom{D}{d} v^d (1-v)^{D-d}, \quad B_D^d(v) = \sum_{r=0}^d b_D^r(v),$$

$$p_\lambda(d) = (\lambda^d/d!) e^{-\lambda}, \quad P_\lambda(d) = \sum_{r=0}^d p_\lambda(r).$$

ф. р. — функция распределения.

э. ф. р. — эмпирическая функция распределения.

в. в. — выборка с возвращением.

сл. в. б. в. — случайная выборка без возвращения.

ПРУК — приемочный уровень качества.

БРУК — браковочный уровень качества.

СВК — средний уровень выходного качества.

ПСВК — уровень предельного среднего выходного качества.

СЭП — система экономичных планов.

НВК-1 — непрерывный выборочный контроль с заменой обнаруженных дефектных объектов на годные.

НВК-2 — непрерывный выборочный контроль с исключением обнаруженных дефектных объектов без последующей замены на годные.

НК — направленный контроль.

$\mathcal{D}_1$ — решение отвергнуть оставшуюся непроверенной часть совокупности без контроля.

$\mathcal{D}_2$ — решение провести сплошную проверку всех объектов, не попавших в выборку.

$\mathcal{D}_3$ — решение принять оставшуюся непроверенной часть совокупности без контроля.

Π_{ij} — план контроля с решениями $\mathcal{D}_i, \mathcal{D}_j$.

Юрий Константинович Беляев

**ВЕРОЯТНОСТНЫЕ МЕТОДЫ
ВЫБОРОЧНОГО КОНТРОЛЯ**

(Серия: «Теория вероятностей
и математическая статистика»)

М., 1975 г., 408 стр. с илл.

Редактор *Л. В. Рыкова*

Технический редактор *В. Н. Кондакова*

Корректоры *Е. А. Белицкая, Н. Б. Румянцева*

Сдано в набор 11/II 1975 г. Подписано к печати 16/IX 1975 г. Бумага 84×108 $\frac{1}{32}$ № 2. Физ. печ. л. 12,75. Условн. печ. л. 21,42. Уч.-изд. л. 22,73. Тираж 9000 экз. Т-14892. Цена книги 1 р. 62 к. Заказ № 542.

Издательство «Наука»

Главная редакция
физико-математической литературы

117071, Москва, В-71, Ленинский проспект, 15

Ордена Трудового Красного Знамени
Ленинградская типография № 2
имени Евгении Соколовой Союзполиграфпрома
при Государственном комитете Совета
Министров СССР по делам издательств,
полиграфии и книжной торговли
198052, Ленинград, Л-52,
Измайловский проспект, 29

