

Дж. Клейнор



СТАТИСТИЧЕСКИЕ
МЕТОДЫ
В ИМИТАЦИОННОМ
МОДЕЛИРОВАНИИ

STATISTICAL TECHNIQUES IN SIMULATION

(in two parts)

Jack P. C. Kleijnen

Katholieke Hogeschool
Tilburg, The Netherlands

Part 1

Дж. Клейнен

СТАТИСТИЧЕСКИЕ МЕТОДЫ
В ИМИТАЦИОННОМ
МОДЕЛИРОВАНИИ

В ы п у с к 1

Перевод с английского Ю. П. АДЛЕРА,
К. Д. АРГУНОВОЙ, В. Н. ВАРЫГИНА, А. М. ТАЛАЛАЯ

Под редакцией и с предисловием
Ю. П. АДЛЕРА и В. Н. ВАРЫГИНА



Москва «Статистика» 1978

Илк 22.172

К18

МАТЕМАТИКО-СТАТИСТИЧЕСКИЕ
МЕТОДЫ ЗА РУБЕЖОМ

ВЫШЛИ ИЗ ПЕЧАТИ

1. Ли Ц., Джадж Д., Зельнер А. Оценивание параметров марковских моделей по агрегированным временным рядам.

2. Г. Райфа, Р. Шлейфер. Прикладная теория статистических решений.

ГОТОВЯТСЯ К ПЕЧАТИ

1. Клейпен Дж. Статистические методы в имитационном моделировании. Вып. 2.

2. Болч Бен У., Юань Клифф Дж. Многомерные статистические методы для экономики.

3. Бард Я. Нелинейное оценивание параметров.

Редколлегия: А. Г. Аганбегян, Ю. П. Адлер, Ю. Н. Благовещенский, А. Я. Боярский, Н. К. Дружинин, Э. Б. Ершов, Т. В. Рябушкин, Э. М. Четыркин

К $\frac{10805^1-096}{008(01)-78}$ 36-78

¹ Второй индекс 20204.

© Перевод на русский язык, предисловие и указатели, «Статистика», 1978

© 1974, Marcel Dekker, Inc.

● ПЛАНИРОВАНИЕ ИМИТАЦИОННЫХ ЭКСПЕРИМЕНТОВ

Русское название книги, которая лежит перед вами: «Статистические методы в имитационном моделировании», содержит не совсем обычное сочетание вполне привычных терминов. Действительно, статистическими методами теперь трудно удивить, да и с моделированием связана, можно сказать, вся история человеческих исканий. А вот сочетание «имитационное моделирование» требует пояснений. Слова «имитация» и «моделирование» имеют латинское происхождение соответственно от корней «imito» и «simulo». Для гнезд, образованных этими корнями, «Большой латинский словарь» дает такие варианты русских эквивалентов: подражание, видимость, образ, изображение, копия, подделка; подделывать(ся), уподобляться, притворяться, представлять, выражать, обнаруживать, высказывать, заменять — в первом случае, образ, изображение, статуя, подобие, видимость, тень, призрак, привидение, сновидение, характеристика, описание свойств, отображение, форма, познавательный образ, видимость, притворство, лицемерие, симулирование, волшебство; делать похожим, уподоблять, изображать, воспроизводить, представлять — во втором. Отсюда на первый взгляд кажется, что «имитационное моделирование» — это тавтология. Более подробное рассмотрение приведенных полей значений этих терминов, однако, показывает, что эпитет «имитационное» придает более привычному понятию «моделирование» ряд важных в нашем случае оттенков. Он говорит как бы о претензии на большую близость к реальному моделируемому объекту, о готовности воспроизвести многие его характеристики, о непрерывности (или многократности) моделирования, его неокончателности и эмпиричности. Он вычленяет некоторые игровые и диалоговые аспекты такого моделирования. Так что же мы понимаем под этим термином? «Имитационное моделирование» — это *исследование модели сложной системы, направленное на получение информации о самой системе.*

Если принять это определение, то остается уточнить ряд моментов. А именно, о каких системах идет речь, что за модели мы рассматриваем и что значит «исследование модели».

Итак, сложные системы — предмет нашего интереса, — оказываются, весьма распространены. Сложность — это форма отношения исследователя к объекту. Она формируется целями и потребностями исследователя, а также его возможностями (научными и техническими). Так, домарковские модели экономических систем были достаточно просты, хотя вряд ли экономисты сомневались в сложности своего объекта. Просто они не располагали возможностями. Современные задачи привели к такому уровню развития науки и техники, который обеспечил возможности исследования многих сложных систем. В качестве приме-

ров моделей, которые использовались в имитационном моделировании, можно назвать модели фирм, отраслей промышленности, крупных проектов и даже денежного обращения в США. (Эти примеры взяты из книги Т. Нейлора. Машинные имитационные эксперименты с моделями экономических систем. М., «Мир», 1975. Работы этого автора сыграли огромную роль в становлении имитационного моделирования и в его приложении ко многим сложным экономическим проблемам.) Хорошо известны модели сложных технических систем, таких, как прокатный стан, химический реактор и т. п.

Чтобы воспользоваться моделью, ее надо сначала построить. А объект сложен. На помощь приходят такие приемы моделирования, которые позволяют представлять модель не в аналитическом виде, как это было принято раньше, а в виде машинной программы.

Когда модель задана аналитически, ее исследование — дело математика. А если у него возникают технические трудности, то на помощь можно позвать ЭВМ и численные методы анализа. Для моделей, заданных программно, первый путь заказан. Их приходится исследовать эмпирически, «проигрывать» на ЭВМ различные варианты. Но такое «проигрывание» вариантов должно быть целенаправленным, организованным и оптимальным (хотя бы в смысле экономии машинного времени). Именно для этих целей в обычном эксперименте была разработана математическая теория планирования эксперимента.

Планирование эксперимента начинает играть в имитации роль первой скрипки. В этом главное значение книги Клейнена, которую мы обсуждаем.

Инструментом исследования служит ЭВМ. Поначалу развитие ЭВМ связывали лишь с революцией чисто количественной. Но постепенно начали понимать, что открылась возможность качественно новой постановки задач. Одним из первых примеров такого рода явился метод Монте-Карло, имеющий прямое отношение к нашей теме. Этот метод ввел в обиход такие процедуры, о свойствах которых было трудно что-либо сказать. Некоторая потеря достоверности и чистоты с лихвой компенсировалась решительным расширением класса задач, которые стали теперь доступны для решения.

Этот метод возник из простой идеи сравнения частот случайных попаданий в разные области, поскольку такие частоты пропорциональны площадям (объемам), ограниченными этими областями.

Фактически «Монте-Карло» оказался первым методом экспериментирования на математических моделях с помощью ЭВМ. Довольно рано было замечено, что и эта разновидность эксперимента допускает применение методологии планирования для организации процедуры и минимизации числа прогонов. Правда, эта мысль лишь спорадически возникает в публикациях. Четкое оформление идея сочетания метода Монте-Карло с планированием эксперимента нашла в последнее время.

В основе реализации метода Монте-Карло лежит процедура получения случайных чисел или случайных последовательностей. Сначала думали, что для генерирования таких чисел надо пользоваться естественными источниками, вроде распадающегося изотопа или электронной лампы, работа которой сопровождается шумовым эффектом. Жизнь

показала, что подобные источники случайных чисел во многих отношениях неудобны и теперь используют исключительно числа, порожденные программным способом. Причем существует множество алгоритмов и программ, предназначенных для этой цели. Это привело к сложнейшим проблемам оценки качества таких псевдослучайных чисел, которые вылились в кардинальный вопрос о том, *что* такое случайность и *как* ее измерить. Постановка вопросов ясно выражена у Ст. Бира (Кибернетика и управление производством. М., «Наука», 1965, с. 258—266), а наиболее последовательная попытка ответа на них содержится в алгоритмической теории сложности. Можно поставить в упрек Клейнену то, что он не поднялся до рассмотрения всего этого круга вопросов, ограничившись чисто прагматическим подходом.

Из генерированных случайных чисел можно строить случайные величины с требуемыми законами распределения или случайные последовательности заданной структуры, что делает случайные числа важными не только в задачах метода Монте-Карло, но и во многих других случаях, в том числе в имитационном моделировании.

Вслед за «Монте-Карло» начал формироваться другой эмпирический подход, сложившийся в теорию случайного поиска, ставшую незаметно разделом математической теории планирования эксперимента. Случайный поиск внес важную статистическую идею рандомизации в целый ряд численных процедур, считавшихся ранее детерминированными. Это содействовало росту эффективности многих вычислительных схем и расширению класса функций, для которых стали возможными практически просчеты, дающие приемлемые результаты.

Идея всякого моделирования, в том числе и имитационного, предполагает наличие модели, заданной в том или ином виде. Причем, как правило, речь идет о моделях далеко не тривиальных систем. Их надо где-то брать. Поэтому следующий шаг в нашем рассмотрении — методы получения моделей. Получением моделей занимаются, конечно, наряду со всеми конкретными науками такие тесно переплетенные и трудно различимые направления исследований, как системотехника и исследование операций.

В очень сложных случаях для создания моделей приходится разрабатывать специальные приемы, такие, как, например, метод ПАТТЕРН. Ясно, что в той мере, в какой при построении модели используется прямой эксперимент, вся концепция связана с методологией планирования эксперимента. Иногда при моделировании сложных систем используются различные комбинации натуральных и имитационных экспериментов.

Среди частных моделей, наиболее широко применяемых в имитации, необходимо прежде всего назвать модели теории массового обслуживания. Их выдающаяся роль объясняется просто тем, что многие моделируемые системы удобно рассматривать как системы, обслуживающие случайный поток «клиентов». Такими клиентами могут быть, например, покупатели в магазине, пассажиры на автобусной остановке или программы в вычислительной машине, работающие в режиме мультипрограммирования или в системе распределения времени между задачами по приоритетам с учетом некоторых ограничений.

Пока имитационная модель не введена в машину, она бесполезна для наших целей. Поэтому возникает задача алгоритмического описания модели и ее программирования. В первое направление большой вклад внесли работы недавно умершего Н. П. Буслупко. Замечательным результатом попыток программной реализации сложных моделей было уяснение того факта, что алгоритмически можно описывать даже такие модели, которые, в силу их сложности или громоздкости, не допускают аналитического описания. Это обстоятельство резко расширило класс объектов, доступных для имитационного моделирования.

Что касается языков моделирования, то здесь еще нет ясности. Современное состояние дел никого не удовлетворяет. Обзор языков можно найти в приложении к упомянутой книге Т. Нейлора. Подчеркнем, что под влиянием отчасти идей имитации само программирование начали рассматривать как инженерную задачу. Наконец, даже организация вычислительного процесса в машине может рассматриваться методами имитационного моделирования. То же можно сказать и о вычислительных сетях, системах машин.

Пока складывалась концепция статистического моделирования, шло развитие вычислительной техники. Теперь между человеком и ЭВМ стал возможен диалог. Это радикально изменило ситуацию. Если до этого говорилось лишь об итеративных методах решения задачи, то теперь можно говорить об итеративных методах построения моделей. Причем таких методах, которые способны использовать эвристические процедуры и личный опыт пользователя. В моделировании проникла идея непрерывности, и моделирование стало эволюционным. Это, конечно, снизило ценность каждой текущей модели и сделало акцент на самом процессе моделирования. Но зато удалось добиться соответствия с общей схемой процессов познания. Интересно, что параллельно в планировании эксперимента (даже с некоторым опережением) возникли и получили распространение идеи эволюционного планирования, имеющие ту же логическую основу. Диалог открыл новые возможности и перед автоматизацией эксперимента, включив имитацию в число средств, усиливающих интеллект экспериментатора.

Наличие имитационной модели и диалога позволяет интерпретировать ситуацию работы с моделью еще и как игру. Игровой подход делает имитацию прекрасным средством для обучения многих категорий людей, нуждающихся в быстром и эффективном приобретении сложных навыков. Модель становится тренажером. Это возобновляет интерес к программированному обучению с его идеей индивидуализации обучения при сохранении коллективной формы. Кроме того, игровой подход подчеркивает психологические аспекты имитационных задач, возникшие сразу, как только появился диалог.

Где игры — там и решения. Модели вообще нужны для организации рациональных процедур принятия решений, а имитационные модели — тем более. Чтобы решения были реалистичными, они, как правило, должны использовать информацию из трех источников: из эксперимента, из моделирования и из экспертных суждений. С первыми двумя все ясно, а вот включение в систему экспертной информации и ее обработка нуждаются в специальном обеспечении. Теоретиче-

скую основу многих экспертных процедур образуют методы парных сравнений. С принятием решений тесно связана задача прогнозирования, в которой тоже возможно применение имитационных подходов.

Если собрать вместе все, что было упомянуто выше, да учесть еще особенности реальных вычислительных систем и терминалов, да не забыть специфику каждой конкретной задачи, то у нас как раз и получится то, что называется *имитационным моделированием*.

Существуют такие области человеческой деятельности, которые без имитационного моделирования, хотя бы в неполной форме, просто не могут существовать. Пример такого рода — прежде всего космические исследования.

Пожалуй, наиболее стимулирующую роль имитационные модели играют сейчас в экономике. Мы уже упоминали книгу Т. Нейлора — наиболее полную и яркую иллюстрацию этого тезиса. Добавим еще, что работы этого автора и связанного с ним коллектива были первыми в области внедрения методов планирования эксперимента в имитацию.

Нам представляется важным подчеркнуть здесь, что существует множество областей исследования, которые сегодня еще не могут похвастаться большими успехами в области имитационного моделирования, но для которых перспектива такого моделирования выглядит весьма обнадеживающе. Во всяком случае, некоторые попытки в социологии и демографии, в экологии и ландшафтоведении, в искусствоведении и астроархеологии, да и в ряде других областей заслуживают упоминания. В области технических систем имитация используется пока мало. Но и здесь есть некоторые прорывы. Так, при идентификации систем управления имитация служит стандартным средством сравнения вариантов. Теперь нам осталось выяснить отношения имитации с математикой, и в частности с математической статистикой. Дух эмпиризма уже заразил математику, и это неизбежно приведет к имитационным постановкам задач. В ближайшее время особенно полезной имитация будет в статистике. Уже наметился ряд задач такого рода. Это прежде всего проблема устойчивости (робастности) результатов к нарушению предпосылок или проблема выбора критерия, обладающего наибольшей мощностью в данных условиях. О моделировании законов распределения мы уж и не говорим.

Нам представляется, что имитационное моделирование сулит настоящий переворот в методологии и практике статистических методов обработки экспериментальных данных. Основой для этого может послужить рассмотрение вычислительного процесса как эксперимента, приспособляющегося к особенностям каждой конкретной задачи и каждого конкретного алгоритма. Из этого следует, что надо отказаться от надежды на получение «оптимальных» программ для обработки данных на все случаи жизни. Вместо этого нас ждет каждодневная адаптация. Из этого следует, что надо отказаться от иллюзии однозначности ответа: при обработке данных всегда можно получить множество нетождественных ответов и некоторые процедуры это уже предусматривают, например процедура «ветвящаяся стратегия» в одном из методов планирования эксперимента — методе случайного баланса. Из этого следует, что одни и те же данные можно обрабатывать сколь

угодно долго. Из этого следует еще и многое другое. Можно надеяться, что первые диалоговые имитационные процедуры обработки данных появятся достаточно скоро, особенно в регрессионном анализе, где к этому все готово и где этого очень ждут.

Заключая наше описание имитации, подчеркнем, что появление ее— это появление нового инструмента, который должен как-то вписаться в общую схему познавательной деятельности. Если раньше обычно строили такую схему: эксперимент — анализ — принятие решений — эксперимент — ..., то теперь в ней должно появиться новое звено — *имитация*, видимо, между анализом и принятием решений.

И вот в этом направлении, важность, разносторонность и перспективность которого мы пытались показать, появляется монография голландского статистика Дж. Клейнена «Статистические методы в имитационном моделировании». Что же отличает ее от других работ и почему она выходит в русском переводе? Попробуем ответить на эти вопросы, не стремясь, конечно, пересказать ее содержание или подменить его. Хотя имитационное моделирование всегда связывало себя со статистическими процедурами, эта связь отнюдь не всегда и не везде была достаточно последовательной. До проверок статистических гипотез относительно результатов эксперимента на ЭВМ дело часто не доходило. Про планирование эксперимента и говорить не приходится. В книге Дж. Клейнена чрезвычайно высоко оценивается значение планирования эксперимента в имитационных задачах. Она просто пропитана духом статистики. Нам кажется, что это весьма важно. И дело вовсе не в претенциозности, а в прямом экономическом эффекте. Мы все время помним, что имитационные модели стоят не дешево и что планирование, помимо прочих благ, сулит экономию времени счета на каждом этапе моделирования, а проверка гипотезы обеспечивает получение результатов с требуемым уровнем риска.

Следует подчеркнуть четкую практическую направленность данной книги. Автор систематично, подробно и аккуратно рассматривает каждую проблему, связанную с его предметом. Сосредоточившись на статистических аспектах имитации, автор оставил без внимания многие другие вопросы, такие, например, как искусство построения моделей или языки моделирования. В книге предполагается, что модель построена и дальше блок-схем алгоритмов автор не идет. Впрочем, «нельзя объять необъятное». Хорошим естественным дополнением для читателя будет служить упоминавшаяся книга Т. Нейлора. Некоторые важные термины в переводах этих двух книг не совпадают. Все особенности нашего перевода терминов читатель найдет в указателе терминов в конце книги.

Книга выходит в двух выпусках — по три главы в каждом.

В главе I вводятся и раскрываются основные понятия имитационного моделирования. Значительно облегчают усвоение текста приложения, в которых на конкретных примерах иллюстрируется сущность метода Монте-Карло, идея выборочного распределения, совершенные генераторы случайных чисел и т. п.

Координирующей все дальнейшее изложение является глава II. В ней рассматриваются статистические аспекты имитационного модели-

рования. Назначение этой главы состоит в том, чтобы показать место статистических методов в имитационном моделировании и их возможности. Речь идет, в частности, о статистическом анализе входных и выходных сигналов, об идеях методов понижения дисперсии в выборочном процессе, о роли и возможностях методов планирования эксперимента и о проверке статистических гипотез в имитационных задачах.

Центральное место в этом выпуске, несомненно, занимает глава III, в которой методы понижения дисперсии выборки, весьма важные для экономии времени счета и повышения надежности результатов, рассматриваются с подробностью и систематичностью, беспрецедентной для имеющихся работ. Все методы понижения дисперсии основаны на рациональном конструировании выборки. Автор подчеркивает мысль о целесообразности совместного применения разных методов в рамках одной и той же задачи. И в этой главе весьма полезны приложения, вносящие в теоретическое изложение практический акцент. Они способствуют переходу от восприятия основных идей к их непосредственному использованию без привлечения дополнительной литературы.

Если говорить не об этом выпуске, а о всей книге, то центральное место в ней по новизне занимает глава IV, повествующая о планировании имитационных экспериментов. Глава V содержит тщательное исследование объемов выборок и связанных с ними проверок статистических гипотез. А в главе VI приведен подробный пример, охватывающий весь предыдущий материал. Краткая оценка этих глав будет предпослана второму выпуску.

Данная книга ясно показывает, какую роль играют в этой важной проблеме статистические методы, и в частности методы планирования эксперимента. Эта книга будет полезна всем тем, кто в своей деятельности применяет или хочет применять имитационное моделирование и заинтересован в повышении экономичности, надежности и корректности своей деятельности. Более того, она должна заинтересовать и тех, кто занимается применением статистических методов в любых других областях исследования, где возникают, например, задачи определения объема выборки или используется метод Монте-Карло. Наконец, ее не должны пропустить исследователи, занимающиеся методологией и приложениями методов планирования эксперимента.

Автором сформулированы и эмпирически рассмотрены некоторые задачи, не нашедшие еще строгих формальных решений. Поэтому и математики могут найти здесь область приложения своих сил.

Рост интересов к имитационному моделированию носит взрывообразный характер. Это дает надежду на то, что предлагаемая книга внесет свой вклад в общее дело повышения эффективности научных исследований и практической деятельности. Ее будет удачно дополнять одновременно выходящая в русском переводе книга: Р. Шеннон. Имитационное моделирование систем — искусство и наука. М., «Мир», 1978.

Ю. АДЛЕР, кандидат технических наук
В. ВАРЫГИН, доктор технических наук

● ПРЕДИСЛОВИЕ

В этой книге обсуждаются статистическое планирование и анализ имитационных экспериментов и экспериментов по методу Монте-Карло, особенно цифровое моделирование с моделями управленческих и экономических систем. Имитация — всегда некий выборочный эксперимент, как только модель содержит одну или несколько случайных величин (хотя это и особый тип выборочного эксперимента, поскольку имитация ведется на математической модели, а не на настоящем объекте). В любом статистическом эксперименте желательны тщательное планирование и анализ. Анализ призван извлечь максимум информации, содержащейся в данном эксперименте, и, кроме того, показать ограниченность выводов, основанных на выборочных данных. Планирование же должно обеспечить такой эксперимент, который содержит максимально возможную информацию.

Хотя имитационное моделирование — это выборочный эксперимент, оно обычно используется не статистиками, а инженерами, социологами и т. д. (поскольку оно всегда требует построения модели). Основательный статистический анализ и планирование эксперимента отсутствуют в большинстве работ по имитации. Мы надеемся, что благодаря этой книге пользователи станут более осведомленными в статистических аспектах моделирования. Цель этой книги — дать читателю практические знания о сложных статистических методах. Так как имеется слишком много методов, относящихся к данной теме, мы ограничились наше внимание избранными процедурами (но глава II дает общую картину статистических аспектов имитационного моделирования). Мы попытались сделать все главы более или менее независимыми друг от друга. Таким образом, практик, интересующийся конкретным вопросом (например, определением времени счета), может прочесть только соответствующую часть книги. Поскольку мы стремимся дать практические знания определенных методов, для дальнейшего изучения соответствующих тем даны ссылки. (В конце каждой главы дана большая библиография, во многих главах есть раздел, где рекомендуется еще конкретная литература в качестве руководства для дальнейшего изучения. Мы были бы весьма благодарны, если бы читатели проинформировали нас о любой подходящей литературе, которую мы пропустили. Особенно это относится к ссылкам на публикации, появившиеся в последнее время.) Для чтения этой книги требуется только знание основ теории вероятностей и математической статистики. Мы предполагаем, что большинство экономистов и специалистов по управлению знают о зависимых событиях, t -критерии, простом

регрессионном анализе и т. д. Перед тем, как дать краткий обзор содержания всей книги, отметим, что каждая глава начинается с раздела, дающего ее более детальный обзор. В главе I описываются системы, модели, а также методы моделирования и метод Монте-Карло. Затем обсуждаются генерирование случайных чисел и выбор случайных величин. Целью этой главы является не описание новых идей, а определение основных понятий. Она также содержит добавления, в которых представлены простые приложения методов Монте-Карло и имитации. В главе II дан обзор статистических аспектов имитационного моделирования. Она показывает, как последующие главы укладываются в общую картину. В главе III представлены методы понижения дисперсии, которые могут широко использоваться в сложных имитационных экспериментах. Это методы стратификации (расслоения), селективной выборки, контрольных величин, значимой выборки, дополняющих величин и общих случайных чисел. Результаты этой главы — распространение, ограничение и коррекция существующих методов. В ней также приведен алгоритм для оптимального распределения машинного времени при использовании дополняющих и общих случайных чисел. В главе IV описаны подходящие для имитации экспериментальные планы, детально обсуждаются полные факторные эксперименты (и их дисперсионный и регрессионный анализы), планы типа 2^{k-p} , рандомизированные, сверхнасыщенные и группового отсеивания. Затем приведена библиография по концепции поверхности отклика. Глава V охватывает влияние объема выборки на надежность основанных на ней выводов. Эта глава состоит из трех независимых частей. В части А обсуждаются вычисления для единичной совокупности или для сравнения только двух совокупностей (или систем). Объем выборки может быть фиксирован или подобран так, чтобы получить заданную надежность выводов. Часть В включает методы множественных сравнений, т. е. методы, работающие когда объемы выборок фиксированы и $k \geq 2$ систем сравниваются друг с другом. В ней также обсуждаются различные типы долей ошибок (например, аналогичные экспериментальным), выбор подмножества, содержащего наилучшую совокупность, эффективность и робастность (устойчивость) процедуры имитационного моделирования. В части С приводятся методы множественного ранжирования, т. е. решается вопрос о том, сколько наблюдений нужно от каждой из $k \geq 2$ систем, чтобы выделить наилучшую систему (используется подход с так называемой «зоной безразличия»). Также обсуждаются эффективность и робастность этих процедур. Наконец, в главе VI показано на примере, как пользоваться различными методами предыдущих глав. В примере используется метод Монте-Карло с процедурой множественного ранжирования Бехгофера (Bechhofer) и Блюментала (Blumenthal). Исследуется робастность этой процедуры. Все главы содержат множество простых упражнений и много ссылок на литературу для дальнейших упражнений. (Отметим, что случайные величины в этой книге набраны жирным шрифтом, в отличие от детерминированных.)

1.1. ВВЕДЕНИЕ И РЕЗЮМЕ

Поскольку изложение относится главным образом к цифровым имитационным моделям коммерческих и экономических систем, мы должны в первую очередь обсудить понятия системы, модели, метода Монте-Карло и собственно моделирования. Это обсуждение кажется нам необходимым, так как в литературе нет единой терминологии, относящейся к этим понятиям (см. [Naylor et al., 1967a, p. 2—3]). Мы не будем стремиться давать четкие определения, но постараемся объяснить, что мы понимаем под такими терминами, как моделирование, метод Монте-Карло и т. д.

Кроме терминологии мы приведем еще некоторые примеры применения моделирования и метода Монте-Карло. Три приложения в конце главы как раз и содержат эти простые примеры. Они будут использоваться в последующих главах. В 1.6.4 мы обсудим генерирование случайных чисел и выбор случайных величин — входов программ имитационных моделей.

Читатели, знакомые с имитационным моделированием, могут пропустить эту главу, за исключением глоссария (словаря специальных терминов) в конце главы (и, возможно, приложений, на которые мы будем ссылаться в последующих главах; стоит обратить внимание и на предыдущее замечание о случайных числах и выборах).

1.2. СИСТЕМЫ

Следуя Нельсону [Nelson, 1966, p. 1], ван Диксхорну и Лайзену [van Dixhoorn and Lyesen, 1968, p. 15], мы определим понятие системы следующим образом (см. также [Mihram, 1972, p. 213—222]): система — это множество *элементов*, называемых также компонентами. Например, промышленное предприятие (завод) можно рассматривать как систему, элементами которой будут машины и люди; национальная экономика — система, компонентами которой являются производители и потребители. Элементы имеют определенные характеристики, или *«признаки»*, которые могут принимать непрерывные или дискретные значения. На заводе в качестве признака машины может быть использована ее занятость: «работает — не работает»; в системе национальной экономики признаком потребителя может быть размер его потребности в определенном продукте. Элементы системы находятся в *отношениях* между собой и, следовательно, взаимодействуют. Например,

Если нет оператора, то машину некому обслужить. Кроме так называемых внутренних отношений существуют внешние отношения. Внешние отношения связывают элементы системы с *окружающей средой*, с внешним миром системы. Например, существует определенный способ поступления требований в систему. Представим систему в виде диаграммы, изображенной на рис. 1: сама система представлена в виде квадрата или «ящика». Внешняя среда оказывает воздействие на систему через *вход*. Входные воздействия *преобразуются* или *перерабатываются* в *выход* системы.

Для рассмотрения динамического поведения систем необходимо ввести понятия состояния и устойчивости. *Состояние* системы определяется непрерывными или дискретными значениями признаков эле-

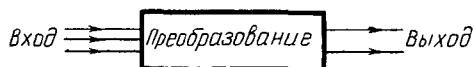


Рис. 1. Графическое представление системы

ментов системы. Например, если на заводе меняется длина очереди, то изменяется и состояние системы; если меняются предпочтения (вкусы) потребителей, то экономическая система переходит в новое состояние. Мы говорим, что система находится в установившемся режиме, если вероятность нахождения ее в некотором состоянии не меняется со временем.

Заметим, что если система находится в установившемся режиме, то она может легко переходить из одного состояния в другое. Это происходит, когда система близка к равновесию, а распределение вероятностей переходов из одного состояния в другое фиксировано. (В теории марковских процессов это фиксированное распределение называется инвариантным, или стационарным, распределением.) Если система не находится в установившемся режиме, то ее состояние называется *неустановившимся (переходным) режимом*. Фиксированные вероятности установившегося режима — это пределы вероятностей, достигаемые после долгого периода времени. Эти предельные вероятности предполагаются *независимыми* от *начального состояния* системы.

Система называется *устойчивой*, если она возвращается к установившемуся режиму после внешнего возмущающего воздействия на систему, выводящего ее из этого режима. Таким возмущающим воздействием может быть, скажем, временное увеличение скорости поступления требований в систему (см. [Copway, 1963, р. 48—49], [Kiviat, 1967, р. 17], [Nelson, 1966, р. 2]). Заметим, что в статье Эмери [Emery, 1969] по системному подходу установившийся режим определяется иначе, так как там рассматриваются детерминированные системы.

Мы отсылаем* к литературе¹ для дальнейшего обсуждения концепции систем и для их многообразных классификаций на реальные — абстрактные, черного ящика — идентифицированные, открытые — за-

* Здесь и далее цифры обозначают отсылку к примечаниям, расположенным в конце глав. — *Прим. ред.*

крытые, адаптивные — неадаптивные, с обратными связями, с предварением и разомкнутые (неуправляемые), статические — динамические, стохастические — детерминированные, непрерывные, дискретные и смешанные, технико-организационные и абстрактные. Эмери [Emery, 1969] и Роуз [Rose, 1970] написали несколько интересных статей по системному подходу к проблемам различных дисциплин, например к теории организации, социологии и биологии.

1.3. МОДЕЛИ

В очень общем виде можно определить модель как нечто, представленное через что-то другое. В свете такого определения статус модели приобретают описания жизни отдельных людей, пьесы, изображающие исторические события, и т. д. Очевидно, существует очень много типов моделей. Черчмен [Churchman et al., 1959, p. 155—162] и Кайвиет [Kiviat, 1967, p. 2] рассматривают модели-изображения, символные и аналоговые модели. Форрестер [Forrester, 1961, p. 49—52] предложил свою детально разработанную классификацию.

Михрам [Mihram, 1971, p. 1—17] обсуждал модели и их отношение к моделированию. Наша монография посвящена *абстрактным моделям*, т. е. моделям, состоящим из *математических символов* или *блок-схем программ*. Если модель состоит из одного уравнения, например $z = x + y + c$, то математическая формула и блок-схема этой зависимости будут, по существу, одни и те же. Однако если мы имеем дело со сложной системой, то гораздо легче построить модель в виде блок-схемы, чем написать множество уравнений и неравенств. В качестве примера приведем модель «производство — запасы» [Neulog et al., 1967a, p. 174], где блок-схема показывает отношение между компонентами системы. Далее под словом модель будем всегда иметь в виду этот тип абстрактных моделей (кроме случаев, когда этот термин определен как-то иначе, скажем, как в I.6.1).

Объект, представленный моделью, — это система. Как заметили ван Диксхорн и Лайзен [van Dixhoorn and Lyesen, 1968, p. 42], абстрактная модель системы — тоже система. Когда мы употребляем в книге слово *система*, мы хотим подчеркнуть, что элементы системы взаимодействуют между собой и с окружающей средой. Слово *модель* означает, что мы изучаем не интересующий нас объект, а лишь его описание, представление. Предполагаем, что это описание ведет себя сходно с объектом.

В экономике принято *переменные* модели делить на эндогенные (внутренние) и экзогенные (внешние). Значения экзогенных переменных не определяются моделью, их называют также независимыми переменными. Следуя терминологии теории систем, независимые переменные можно разделить на управляемые и неуправляемые. Неуправляемые переменные, например внешнеторговый спрос, служат входными переменными системы. Управляемые переменные, например государственные расходы, это переменные, на которые можно влиять с помощью определенных компонент системы. Значения эндогенных переменных зависят от модели. В системной терминологии эти зависи-

мые переменные можно разделить на промежуточные переменные, или переменные, описывающие состояние системы, и на выходные. Кроме переменных мы различаем еще параметры модели. *Параметры* квантифицируют влияние эндогенных переменных и в отличие от экзогенных переменных являются константами (в 1.6.3 мы определим, «как долго» параметры остаются постоянными в имитационной модели).

Отношения описывают связи между переменными и параметрами. Эти отношения можно разделить на идентифицирующие, или определяющие, и операционные характеристики. Кроме того, мы можем изучить разные модели, которые различаются между собой экзогенными переменными, параметрами или отношениями. То, что меняется в модели от одного ее варианта к другому, — экзогенная переменная, параметр или отношение — называется *фактором* с точки зрения статистической теории планирования эксперимента.

Если мы изменим один либо более факторов, выходные переменные также изменятся. В терминологии планирования эксперимента выходная переменная называется *откликом* (см. [Emshoff and Sisson, 1971, p. 52], [Hanken and Buijs, 1971, p. 10—11] и [Naylor et al., 1967a, p. 10—15, 322]).

1.4. МЕТОДЫ РЕШЕНИЯ

Чтобы ответить на интересующие нас вопросы, касающиеся объекта исследования, представленного моделью, мы должны «решить» модель. Существуют *аналитические* и численные методы решения. Для аналитического решения используются методы анализа — теории дифференциального и интегрального исчисления. Решение при этом дается в виде формулы, которая содержит все возможные значения независимых переменных и параметров. Аналитические решения для простых задач теории запасов и теории массового обслуживания можно найти у Черчмена [Churchman et al., 1959]; Тейл и Бут [Theil and Boot, 1962] рассматривают аналитическое решение эконометрических моделей, представляющих собой совместную систему линейных дифференциальных уравнений n -го порядка.

При *численном* решении вместо независимых переменных и параметров подставляются числа, и решение получается на основе операций с этими числами. Многие численные методы итеративны, т. е. на каждом шаге дают решение лучшее, чем на предыдущем шаге, результаты которого используются в решении на новом этапе; примерами могут служить метод линейного программирования и метод Ньютона для приближенного отыскания корней уравнений. К численным методам можно отнести методы Монте-Карло и имитационного моделирования. Мы обсудим их более подробно ниже. Заметим также, что при решении сложных моделей можно комбинировать различные методы.

1.5. МЕТОД МОНТЕ-КАРЛО

Следуя Хеммерсли и Хендсcombe [Hammersley and Handscomb, 1964, p. 2], мы определим *метод Монте-Карло* в широком смысле как любой способ решения модели, в котором используются случайные или псев-

дослучайные числа. Обсудим понятие «случайные числа» более подробно.

Случайные числа — это случайные величины, которые *равномерно* распределены на интервале $[0,1]$ и (стохастически) независимы. (Случайные, или переменные, величины — это вероятностные переменные; они введены по контрасту с переменными, значения которых можно установить заранее. Случайные числа (независимые) — равномерные случайные величины.)

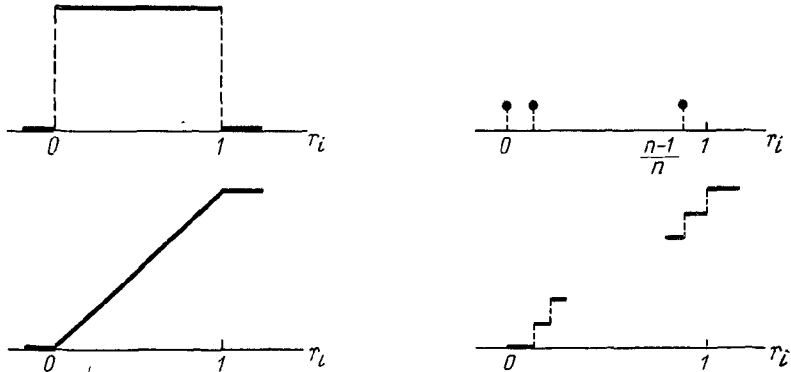


Рис. 2. Функции плотности и накопленные (кумулятивные) распределения непрерывных и дискретных случайных чисел

Случайные числа имеют следующие две характеристики:

1. Если r_i ($i = 1, 2, 3, \dots$) есть случайные числа, то их кумулятивное распределение, обозначим его F , удовлетворяет уравнению (1) для всех i (см. рис. 2).

$$\begin{aligned} F(r_i) &= P(r_i < r_i), \\ &= r_i \quad \text{для } 0 \leq r_i \leq 1, \\ &= 0 \quad \text{для } r_i < 0, \\ &= 1 \quad \text{для } r_i > 1. \end{aligned} \quad (1)$$

Заметим, что теоретически случайные числа должны быть *непрерывными* переменными с функцией плотности f , определяемой следующим выражением:

$$\begin{aligned} f(r_i) &= 1 \quad \text{для } 0 \leq r_i \leq 1, \\ &= 0 \quad \text{для всех других значений.} \end{aligned} \quad (2)$$

На практике, однако, мы имеем только *дискретные* измерения. Пусть мы измеряем с точностью до n десятичных знаков. Тогда, если $n=2$, мы наблюдаем значения 0,00; 0,01; ...; 0,98; 0,99. Дискретная функция плотности записывается в виде (3):

$$\begin{aligned} P(r_i = r_i) &= 10^{-n} \quad \text{для } r_i \in S, \\ &= 0 \quad \text{для всех других значений,} \end{aligned} \quad (3)$$

где S включает числа $(j \cdot 10^{-n})$ с $j = 0, 1, \dots, (10^n - 1)$. Следовательно, если $n = 2$, то S содержит числа $0,00; 0,01; \dots; 0,99$ (см. также [Haitsma and Oosterhoff, 1964, p. 5—6], [Halton, 1970, p. 29] и [Molenaar, 1968, p. 104—105]). Кумулятивное распределение просто следует из (3). Это распределение есть ступенчатая аппроксимация непрерывного распределения (1). На рис. 2 показаны разные функции. Если при моделировании на ЭВМ n велико или, наоборот, 10^{-n} мало, то мы можем пренебречь различием между непрерывным и дискретным распределениями. Непрерывное распределение из (1) более удобно в работе.

2. Следуя Фишу [Fisz, 1967, p. 55], мы назовем переменные $r_1, r_2, \dots, r_i, \dots, r_n$ независимыми, если их совместное кумулятивное распределение (обозначим его G) можно представить в виде произведения индивидуальных функций распределения:

$$G(r_1, r_2, \dots, r_n) = F_1(r_1) F_2(r_2) \dots F_n(r_n). \quad (4)$$

Так как все n случайных чисел имеют одинаковые распределения определенные в (1), мы можем переписать (4) в следующем виде:

$$G(r_1, r_2, \dots, r_n) = F(r_1) F(r_2) \dots F(r_n). \quad (5)$$

Независимость случайных чисел имеет своим следствием тот факт, что знание r_1, r_2, \dots, r_{i-1} не дает нам никакой информации об r_i . Физические процессы, генерирующие случайные числа, описаны в [Shreider, 1964, p. 220—238], [Teichrow, 1965, p. 40] и [Tocher, 1963, p. 50—67]. Мы приведем в качестве примеров очень простых физических процессов 10-гранную игральную кость, генерирующую числа между 0 и 9, и монету — для построения двоичных чисел, состоящих из нулей и единиц. Из дальнейшего будет видно, что эти чисто случайные последовательности менее привлекательны для практики по сравнению с псевдослучайными последовательностями чисел.

Псевдослучайные числа генерируются с помощью *детерминированной* алгебраической формулы; порождаемые числа для практических целей можно считать случайными, равномерно распределенными в интервале $(0,1)$ и независимыми. Такая независимость предполагает, что длина «цикла» будет достаточно большой. Длиной цикла считается число псевдослучайных чисел, полученных до того, как генерируемые числа начнут повторяться. Часто применяется формула (6); этот метод называется *мультипликативно-конгруэнтным*, или методом степенных вычетов.

$$\begin{aligned} x_i &= ax_{i-1} \pmod{m} \quad (i = 1, 2, \dots), \\ x_0 &= b. \end{aligned} \quad (6)$$

Уравнение использует предыдущее число x_{i-1} , умножая его на константу a и деля результат на m , остаток берется как новое число x_i ; b служит начальным значением, «затравкой» для x_i . Соответствующее псевдослучайное число r_i , которое должно быть между 0 и 1, получается от деления x_i на m . Псевдослучайные числа следует *проверить* на равномерность и независимость. Длину цикла можно определить методами теории чисел. Мы вернемся к псевдослучайным чис-

лам в пункте I.6.4, чтобы обсудить некоторые технические аспекты и обширную литературу, посвященную генерированию псевдослучайных чисел.

Преимущество псевдослучайных чисел над чисто случайными числами в том, что они генерируются самой ЭВМ с использованием формулы, подобной (6). В таком случае нет необходимости *хранить* в памяти машины большие таблицы случайных чисел и тратить время на их чтение. В главе III мы увидим также важное статистическое преимущество, которое состоит в том, что можно повторить последовательность псевдослучайных чисел. Это видно, например, из формулы (6): если мы возьмем те же самые b , a и m , то получим ту же последовательность. Так как мы обычно используем псевдослучайные числа вместо чисто случайных, в дальнейшем будем пользоваться термином «случайные числа» и для псевдослучайных чисел.

Мы подробно обсудили понятие случайного числа, так как это основное понятие в методе Монте-Карло. Более того, ниже мы будем вынуждены в некоторых отношениях отклоняться от сказанного. Далее рассмотрим три области применения метода Монте-Карло.

1. Первой областью является решение *детерминированных* задач с помощью метода Монте-Карло (в широком смысле), т.е. с применением случайных чисел. Моргенталер [Morgenthaler, 1961, p. 368—370] считает, что термин «Монте-Карло» был предложен фон Нейманом (von Neumann) и Уламом (Ulam) в конце 1940-х годов как раз для способа решения детерминированных задач с помощью случайных чисел, случайных величин. Детерминированная задача решается методом Монте-Карло, если она *формально* записывается как некоторый случайный процесс. Например, рассмотрим интеграл

$$\int_0^{\infty} \frac{1}{x} \lambda e^{-\lambda x} dx \quad (v, \lambda > 0). \quad (7)$$

Величину интеграла можно оценить, если учесть, что $\lambda \exp(-\lambda x)$ для $x \geq 0$ есть экспоненциальная плотность распределения. Следовательно, можно сделать выборку x из совокупности с этой функцией плотности и подставить выборочное значение x в $g(x)$, определенное выражением (8):

$$g(x) = 0, \text{ если } x < v, \\ = \frac{1}{x}, \text{ если } x \geq v. \quad (8)$$

Математическое ожидание $g(x)$ равно интегралу (7). Этот простой пример используется в приложении I.1. Другой пример — вычисление Моргенталером [Morgenthaler, 1961, p. 370] площади нерегулярной поверхности. Маккракен [McCracken, 1955] рассматривает возможность экспериментального вычисления числа π . Кроме того, можно привести примеры решения многомерных интегралов, дифференциальных уравнений с граничными условиями, линейных уравнений, нахождения

собственных значений с помощью метода Монте-Карло. Все это можно найти в [Milgram, 1972, p. 186—199], [Newman and Odell, 1971, p. 53—57, 67—74] и [Shreider, 1964]. Хэлтон [Halton, 1970] дает обширный обзор (с 251 ссылкой!) использования метода Монте-Карло для решения задач из математики, физики и т. д.

2. Другая область применения метода Монте-Карло — это *выборочные распределения*, которые Моргенталер [Morgenthaler, 1961, p. 370] называет моделями выборок. Выборочное распределение используется в *математической статистике* несколько десятков лет. Целью его является определение распределения или некоторых параметров распределения случайной величины. Эта случайная величина, которую мы называем выходной переменной, — известная функция одной или большего числа других — входных случайных величин, имеющих известные распределения. Чтобы оценить распределение выходной переменной, мы берем значения каждой из входных переменных из их распределений и вычисляем значение выходной переменной. Такая выборка повторяется много раз, и затем оценивается искомое распределение. В приложении I.2 мы приводим простой пример выборочного распределения для оценки вероятности того, что случайная величина x меньше некоторой константы a или

$$p = P(x < a). \quad (9)$$

Мы предполагаем, что выходная переменная x есть функция (10) входных переменных x_1 и x_2 :

$$x = \min(x_1, x_2), \quad (10)$$

где x_1 и x_2 имеют известные распределения. Заметим, что вместо *процентилей* в (9) нас могут интересовать «квантили», т. е. p фиксировано (например, 5%), и мы хотим определить соответствующую этому p величину a (см. [Lewis, 1972, p. 9—10]).

Другими примерами применения метода выборочных распределений с помощью метода Монте-Карло служат исследования «робастности» статистик. Статистика называется более робастной, если она менее чувствительна к нарушению исходных предположений. Рассмотрим, например, хорошо известную t -статистику:

$$t = \frac{\bar{x} - \mu}{s/\sqrt{N}}. \quad (11)$$

Мы можем использовать метод Монте-Карло, чтобы изучить распределение t , если \bar{x} и s основаны на наблюдениях, не подчиненных нормальному распределению. Эндриус [Andrews et al., 1972] исследовал методом Монте-Карло 70 различных оценок положения. Невье и Гренджер [Neave and Granger, 1968] применили метод Монте-Карло для изучения чувствительности различных критериев сравнения двух средних. Дональдсон [Donaldson, 1966] изучил робастность F -критерия, Хойтс и Ренс [Heuts and Rens, 1972] сравнивали мощность двух критериев согласия. В главе VI мы применим метод Монте-Карло к изучению робастности одного из методов многомерного ранжирова-

ния. Много других исследований до 1953 г. упомянуты в книге [Teichroew, 1965, p. 35]; Льюис [Lewis, 1972] дал блестящий обзор вычислительных аспектов получения выборочных распределений, включающий более 50 ссылок.

К примерам работ по выборочным распределениям можно отнести многочисленные исследования последних лет, посвященные применению различных методов *регрессионного анализа* для эконометрических моделей. В этих исследованиях модель служит для генерирования данных, а к данным применяются различные методы регрессионного анализа для оценки параметров предполагаемой модели; распределения оценок параметров в итоге сравниваются между собой. Джонстон [Johnston, 1963, p. 275—295] описывает несколько исследований по методу Монте-Карло, посвященных обычному методу наименьших квадратов, двухступенчатому методу наименьших квадратов, оценкам максимума правдоподобия с полной и ограниченной информацией, применяемым в эконометрических моделях. Аналогичные исследования осуществлены Кмента и Гильбертом [Kmenta and Gilbert, 1968], Михайлом [Mikhail, 1972], Нилменом [Neeleman, 1973], Сэссером [Sasser, 1969], Шинком и Чу [Schink and Chi, 1966], см. также обзор [Meier et al., 1969, p. 138—141]. Позднее в этой главе мы вернемся к выборочному распределению, чтобы указать сходство и различия этих выборочных экспериментов и имитации².

3. Третья область применения методов Монте-Карло — это *имитационное моделирование*. Мы определим имитационное моделирование в следующем параграфе. Здесь мы только заметим, что многие имитационные исследования используют последовательности случайных чисел и, следовательно, являются формой метода Монте-Карло. Мы считаем, что необходимо отличать имитационное моделирование от других применений этого метода. В дальнейшем мы будем говорить о методе Монте-Карло в узком смысле, если речь не идет об имитационном моделировании. Далее всюду термин «Монте-Карло» употребляется в узком смысле, включает в себя все те методы, которые используют случайные числа, исключая имитационное моделирование (см. рис. 5 ниже). Это означает, что мы в своем определении не следуем определениям Хеммерсли и Хендскома [Hammersley and Handscomb, 1964], в которых «методы Монте-Карло» включают в себя и «методы понижения дисперсии». Эти методы будут обсуждаться в главе III.

1.6. ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ

1.6.1. Описание метода моделирования

Нейлор [Naylor et al., 1967a, p. 2] и Точер [Tocher, 1966, p. 693—695] отмечают, что не существует всеми принятого определения моделирования. Следуя определению Нейлора [Naylor et al., 1967a, p. 3], мы определяем моделирование в широком смысле как «экспериментирование с моделью во времени». Рассмотрим это определение подробнее.

Моделирование предполагает *экспериментирование*, но экспериментирование не с реальным объектом, а с его моделью. Поведение объекта рассматривается *во времени*. Такое определение требует двух замечаний. Первое: объектом исследования при моделировании служит система. Так, Кайвиет [Kiviat, 1967, p. 5] указывает, что очень часто термины «*моделируемые системы*» и «*моделирование*» используются как синонимы. Второе: в разделе 1.3 мы определили «модель» так, что ограничились только абстрактными моделями. В определении моделирования мы отбросим это ограничение, включая в это понятие как абстрактные, так и физические модели. Примеры моделирования с *физическими моделями* — проведение экспериментов с масштабными моделями судов в искусственных бассейнах, с моделями самолетов в аэродинамической трубе, с моделями для изучения влияния морских течений и т. д. Некоторые исследования с физическими моделями могут проходить с участием человека, например на тренажерах (для пилотов) и испытательных автотреках. Примерами моделирования на абстрактных моделях и с участием людей являются *деловые и военные игры*. Такой вид моделирования Моргенталер [Morgenthaler, 1961, p. 37] и Нейлор [Naylor et al., 1967a, p. 3] называют *операционными играми* с риском. Литературу по теории игр можно найти в библиографии к [Naylor, 1969], [Shubik, 1960, 1970], [Shubik and Wolf, 1972]. (См. также [Emshoff and Sisson, 1971, p. 245—249] и [Meier et al., 1969, p. 179—213].) Соотношение между математической теорией игр и имитационными играми обсуждается в [Shubik, 1972]. Моделирование с участием человека с реальными или абстрактными моделями называют *человеко-машинным моделированием*; см., например, [Geisler, 1960] и [Meier et al., 1969, p. 287—289]. В этой монографии мы будем рассматривать только имитационное моделирование с абстрактными моделями и без участия человека. Итак, мы определяем моделирование в широком смысле как экспериментирование с абстрактными моделями во времени.

Это определение моделирования не обязательно требует случайных чисел. *Детерминистическое моделирование* очень широко применяется в экономике. Экономические модели состоят из регрессионных уравнений, которые содержат стохастические возмущения (или ошибки). При оценке параметров регрессионных уравнений эти стохастические составляющие надо учитывать. Однако, когда параметры модели оценены, возмущающие воздействия устранены, траектории *эндогенных* переменных вычисляются на основе *экзогенных* переменных и обратных связей без использования случайных чисел. Служить примерами такого моделирования могут модель Коэна для кожевенной и обувной промышленности [Cohen, 1960] и квартальная эконометрическая модель Голландской экономики Дрихейса [Driehuis, 1972]. Хаури [Howrey, 1966] показал, что этот подход может привести к ложным заключениям (см. также [Howrey, 1972, p. 26], [Kleijnen, 1970, p. 14—18], [Naylor, 1967b, p. 1314—1319], [Naylor et al., 1968, p. 188]). Особый класс моделей — детерминистические модели Форрестера [Forrester, 1961], основанные на методе *промышленной динамики*, который мы кратко обсудим в 1.6.2. Другими примерами являются модель Канадской

угольной промышленности Клуг [Clough et al., 1965, p. 126] и модель Моленара [Molenaar, 1968, p. 91—96], представляющие собой модель очереди. Мы рекомендуем [Emshoff and Sisson, 1971, p. 169—170] для выбора между детерминистическим и стохастическим моделированием. Однако большинство имитационных моделей использует случайные числа.

Мы определили *моделирование в узком смысле* как эксперимент с абстрактными моделями во времени; этот эксперимент включает выборку значений случайных величин из их распределений, поэтому такое моделирование называется *стохастическим*. Метод выборки случайных величин с помощью *случайных чисел* будет обсужден в 1.6.4. Так как при этом используются случайные числа, то иногда говорят о *моделировании Монте-Карло*. Мы обсудим пример стохастического моделирования, который понадобится и в последующих главах.

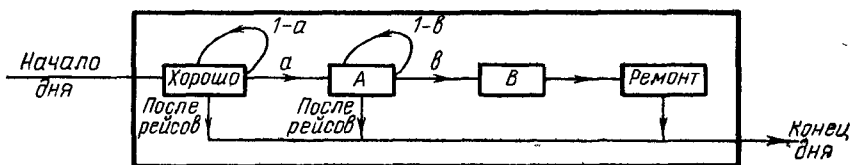


Рис. 3. Эксплуатационная политика β

Пусть мы должны выбрать между двумя вариантами *эксплуатации* автобуса. По графику автобус должен делать N рейсов ежедневно. Каждый день автобус начинает маршрут в хорошем состоянии. Если автобус начинает определенный рейс в хорошем состоянии, то он имеет некоторую вероятность закончить его в «ухудшенном» состоянии А.

В состоянии А автобус может продолжать свой путь; необходимый ремонт занял бы время одного рейса. Если, однако, автобус в состоянии А продолжает маршрут без ремонта, вероятность того, что он закончит рейс в фазе β , равна b . Если он достигает состояния В, все оставшиеся рейсы текущего дня отменяются. Допустим, нам нужно выбирать между стратегией α , когда автобус ремонтируется в фазе А, и стратегией β , когда автобус не ремонтируется, пока его состояние не достигнет фазы В. Стратегия β представлена на рис. 3. Для моделирования этой системы достаточно иметь одно случайное число r для каждого рейса. Если рейс начинается при хорошем состоянии автобуса, мы проверяем, выполняется ли неравенство $r < a$. Если выполняется $r < a$, то маршрут заканчивается в состоянии А. Поскольку автобус находится в состоянии А, то проверяется, меньше ли следующее случайное число, чем b . Если $r < b$, то автобус переходит в состояние В; все оставшиеся маршруты отменяются, так как автобус подлежит ремонту. Моделирование других дней производится аналогично. Этот пример разбирается в приложении 1.3.

Примером, который часто приводится в литературе, служит моделирование системы массового обслуживания с одной станцией обслуживания; обслуживание в порядке поступления, время обслуживания

и время между поступлениями распределены по экспоненциальному закону. Если мы возьмем некоторый интервал времени между поступлениями i -го и $(i + 1)$ -го требований и обозначим его AT_i , а время обслуживания i -го требования — ST_i , то время ожидания (WT_i) и простой станции (IT) можно изобразить графически, как это сделано на рис. 4. Мы предлагаем читателю самому нарисовать блок-схему моделирования этой системы.

Имеется большое количество примеров стохастического моделирования. Из области науки об управлении упомянем сложные системы массового обслуживания с приоритетами, несколькими станциями обслуживания и т. п., такие, как регулирование уличного движения,

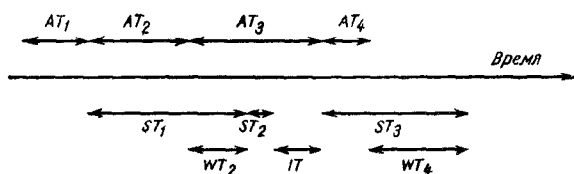


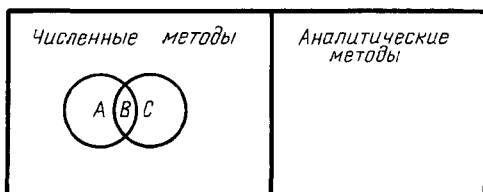
Рис. 4. Одноканальная система массового обслуживания с обслуживанием в порядке поступления требований

аэропорты, сталелитейные заводы и др. [Naylor et al., 1967a], [Schmidt, Taylor, 1970, p. 325—403], [Tocher, 1963]. Михрам [Mihram, 1972, p. 200—206] предлагает классификацию систем массового обслуживания, основанную на особенностях поступления и обслуживания и правилах очередности.

Обзор по моделированию планирования работы магазина предварительных заказов, дан в [Moore and Wilson, 1967] и [Naylor, 1971, p. 59]. Моделирование сетей ЭВМ обсуждается в [Bovet, 1972], [Dewan et al., 1972], [Hutter, 1970], [Maisel and Gnugnoli, 1972], [Pomerantz, 1970] и [Rechtschaffen, 1972, p. 265—268]; много других примеров содержится в [Lucas, 1971] и [Maguire, 1972]. Моделям запасов уделяют внимание Нейлор [Naylor, 1971, p. 53], Шмидт и Тейлор [Schmidt, Taylor, 1970, p. 404—475]. Обзоры моделей массового обслуживания и моделей запасов можно найти в [Smith, 1968, p. 67—69 и 1970, p. 770—772]. В [Mann, 1970] обсуждаются моделирование (и другие подходы) для систем надежности. Вычисление критических путей в сетях исследовано в [Burt, 1972], [Burt et al., 1970], [Klingel, 1966] и [Van Slyke, 1963]. Подробное обсуждение моделирования рынка (с библиографией) есть в книге [Elton and Rosenhead, 1971]. Имитационные модели финансовых операций (с учетом статей движения капитала в бюджете) предложены Нейлором [Naylor, 1971, p. 61—68]; (см. также [Benecke and Westgaard, 1972], [Carter, 1971], [Salazar and Sen, 1968]). Много примеров можно найти в учебниках и трудах конференций по моделированию. Многочисленные рефераты публикаций по применению и методологии моделирования можно найти в журналах *Computing Reviews*, *Operations Research/Management Science* (ежемесячный) и *International Abstracts in Operations Research* (выходит раз в два месяца). Голландский периодический журнал *Mededelingen Operationele*

Research (ежемесячный) также содержит рефераты (в большинстве на английском языке) по моделированию. Не удивительно, что в обзорах, например, [Shannon and Biles, 1970, p. 744] и [Turban, 1972, p. 718] утверждается, что моделирование — наиболее распространенный на практике метод исследования операций.

Другая область применения моделирования — экономика. Имитационные модели фирм были построены Форрестером на основе его промышленной динамики; Бонини (Bonini) применил междисциплинарный подход; моделирование использовалось также Корпорацией системных исследований (*Systems Development Corporation*) и т. д. Сайертом (Syert) и Марчем (March)



$A+B$ = метод Монте-Карло в широком смысле
 $B+C$ = моделирование в широком смысле
 A = метод Монте-Карло в узком смысле = собственно „Монте-Карло“
 B = моделирование в узком смысле

Рис. 5. Различные методы решения

были предложены модели для монополии двух и многих конкурирующих компаний. Модель лесоразработок была построена Балдерстоуном (Balderston) и Хогеттом (Hoggatt); как мы уже упоминали, для кожевенной и обувной промышленности модель была построена Коэном [Cohen, 1960]. Хорн [Horn, 1968] изучал циклические изменения в снабжении свиной Нидерландов, используя эконометрические регрессионные и имитационные модели. Модели экономики в целом включают Большую Брукингскую модель — SSRC для США, модель Оркатта и др. [Orcutt et al., 1961], модель Холленда (Holland) и Гиллеспи (Gillespie) для Индии. Эти и ряд других экономических моделей обсуждаются в [Driehuis, 1972], [Dutton and Starbuck, 1971a, p. 465—588], [Meier et al., 1969, p. 118—138], [Naylor et al., 1967a, p. 192—233]. Мы вернемся кратко к обсуждению моделирования в экономике в I.6.2.

Многие крупномасштабные модели построены в области военных исследований (см. [Koorman, 1970], [Morgenthaler, 1961, p. 393—395], [Smith, 1968, p. 50—57] и массу рефератов в *Mededelingen Operationele Research*).

Модели эвристических путей решения задач, искусственного интеллекта, мышления и т. п. есть в работах [Dutton and Starbuck, 1971a], [Meier et al., 1969, p. 147—178]. Различные области применения моделирования, особенно в науках о поведении, обсуждаются в [Emshoff and Sisson, 1971, p. 243—264]. В книге [Guetzkow et al., 1972] обсуждаются также и некоторые другие, кроме вышеуказанных, области применения моделирования. Моделирование в политике обширно представлено в [Corlin, 1968] и [Guetzkow, 1971]. К другим областям применения относятся демография, биология, метеорология и вообще любые науки, использующие количественные отношения,

На рис. 5 показано отношение между некоторыми понятиями. Заметим, что моделирование в узком смысле использует случайные числа и, следовательно, образует подмножество методов Монте-Карло в широком смысле. Некоторые авторы используют термин Монте-Карло для определения того, что мы называем моделированием (в узком смысле) [Churchman et al., 1959, p. 174, 407], [Jessop, 1956]. В последнее время в литературе в основном принят термин моделирование [Mihram, 1972, p. 207—208], [Naylor et al., 1967a], [Tocher, 1963].

1.6.2. Моделирование и системы

Поскольку моделирование — это эксперимент во времени, то его методы можно использовать для изучения динамического поведения систем. Особенно если модель содержит *случайные величины*, тогда даже простая модель может стать непреодолимой для аналитического исследования.

Мы считаем, что сказанное более справедливо для изучения *переходных* характеристик системы, а не *установившихся режимов*. В последнем случае можно применять предельные теоремы. Может случиться, что нас интересует не динамическое поведение системы, а только *конец* траектории, например прибыль за три месяца. Но даже в этом случае моделирование может оказаться необходимым, если переменные системы взаимодействуют *сложным* образом. Заметим: мы теперь осознали, что систему надо изучать как единое целое, не фиксируя внимание на отдельных ее частях, так как это ведет лишь к частной оптимизации. К сожалению, поиск глобального оптимума слишком усложняет модель. Такая модель может стать аналитически неразрешимой, особенно если есть *случайные величины*.

Моделирование используется для сложных систем. Часто мы знаем больше о поведении частных *компонент* такой системы, чем о ее поведении *в целом*. Например, мы знаем, когда требование покидает станцию после обслуживания, затем следующий заказчик занимает освободившееся место; однако мы ничего не знаем об отношениях между «конкретными» средним временем обслуживания и средним временем ожидания. *Блок-схема* показывает сами элементы системы, их реакции и взаимное влияние. Математические описания для каждого элемента обычно очень *просты*; не обязательно это линейные уравнения, могут быть и простые нелинейные уравнения или неравенства. Каждый элемент воздействует на другие элементы; взаимодействия и делают сложной эту модель. Если модель основана на знании о поведении *компонент* системы, а не всей системы в целом, мы говорим о модульной или *блочной* конструкции или методе декомпозиции. Общую модель можно решить с помощью *моделирования*. Как отмечает Конвэй и др. [Conway et al., 1959, p. 93—94], моделирование возможно, если мы знаем (быть может, вероятностное) поведение *элементов* системы.

В *экономике* большинство моделей не относится к моделям блочного типа. Традиционные эконометрические модели представляют собой систему линейных регрессионных уравнений и используются для предсказания значений эндогенных переменных в заданном периоде.

Сколько недавно эти модели удалось адаптировать так, чтобы можно было включать дополнительные условия и нелинейности, учитывающие априорные сведения о системе. Эти модели можно решать методами моделирования, в результате мы получаем полные временные траектории (в отличие от предсказания только для заданного периода [Naylor, 1971]). Однако модульные модели были бы возможны, если бы только нам удалось использовать априорные теоретические знания об индивидуальных предпочтениях и производимых, чтобы объяснить макроэкономические отношения. Так, мы не пытаемся объяснить непосредственно поведение системы (как говорят, отношения между макроэкономическими переменными), а наоборот, мы развиваем каждую макроэкономическую переменную на множество микроэкономических переменных. Например, национальное потребление разлагается на потребления индивидов 1, 2 и т. д. Такой подход применялся Оркаттом [Orcutt et al., 1961] в исследовании, которое ставилось как экономическое, а на проверку оказалось демографическим. Различия традиционных эконометрических моделей и имитационных моделей можно найти в [Solon, 1960], [Kleijnen, 1970] и [Reith, 1969]. Хаури и Келсей [Howrey and Kelsey, 1969] написали прекрасную статью, в которой сравнивают аналитические решения для экономических моделей с моделированием. Они очень скептически относятся к использованию моделирования в экономике. Их заключение, однако, не относится к моделям, которые, по нашему мнению, имеют наиболее важные применения, к *стохастическому* моделированию нелинейных систем. В последнее время экономисты проявляют возрастающий интерес к промышленной динамике Форрестера. Как мы заметили в I.6.1, этот подход (в большей части) является детерминистическим. Он допускает моделирование подсистем, которые связаны между собой. Для этого подхода характерно использование дифференциальных уравнений (система развивается в течение фиксированных интервалов времени), а также высокая степень агрегирования. Для дальнейшего обсуждения индустриальной динамики и ее отличий от методов моделирования, используемых в исследовании операций, мы отсылаем к книге [Meier et al., 1969, p. 80—117, 277—282], а также к [Niedereichholz, 1971] и [Smith, 1968, p. 27—32].

Вернемся к выборочному распределению, обсуждавшемуся еще в разделе I.5. Тейкроу [Teichrow, 1965, p. 43] предлагает следующие признаки, которые отличают метод *выборочного распределения* от *моделирования*.

1. Выходная переменная в выборочном распределении — относительно *простая* функция входных переменных и параметров. При моделировании же это очень сложная функция, которую нельзя сформулировать иначе, как только с помощью всей *вычислительной программы*!

2. Функция в выборочном распределении — *статическая*, а при моделировании — отношения *динамические*. Заметим, что исследования методами регрессионного анализа можно считать примерами выборочного распределения. Здесь ищется распределение оценок параметров эконометрических моделей. Такие исследования имеют ту же

цель, что и иные примеры выборочного распределения, но, как и при имитационном моделировании, они могут использовать сложные и динамические отношения.

Мы видим, что большинство имитационных моделей основывается на наших знаниях о поведении элементов системы. Поведение отдельных элементов часто можно выразить с помощью *простых* операций. (Систему делают сложной многочисленные взаимодействия.) Имитационное моделирование — эксперимент с выполнением этих простых операций много раз. Так, например, поступление и обслуживание требования производится огромное число раз. Поскольку при моделировании необходимо повторение большого количества операций, оно возможно только при использовании *цифровых вычислительных машин*. Проблемы программирования при цифровом моделировании описаны Конвэем [Conway et al., 1959, p. 95—103] и Моргенталером [Morgenthaler, 1961, p. 405—411], несколько интересных замечаний по программированию дает Хаузер [Hauser et al., 1966, p. 84—85]. Обсуждение цифрового и аналогового моделирования *непрерывных* систем можно найти в книгах [Bauknecht, 1971, p. 157—160], [Emshoff and Sisson, 1971, p. 9], [Mihram, 1972, p. 226—228], [Morgenthaler, 1961, p. 376—378, 384—385], [Nisenfeld, 1971], [Smith, 1968, p. 15—26] и [van Dixhoorn and Lyesen, 1968, p. 4—6, 35—39]. Многие вопросы программирования были решены с развитием специальных *языков моделирования*. Обзор этих языков с многочисленными ссылками можно найти в [Bauknecht, 1971], [Emshoff and Sisson, 1971, p. 115—158], [Kay, 1972], [Maguire, 1972], [Maisel and Gnugnoli, 1972], [Meier et al., 1969], [Mihram, 1972, p. 238—244, 493—495], [Naylor et al., 1967a], [Naylor, 1971] и [Rytz, 1971].

1.6.3. Проблемы моделирования

Моделирование имеет свои *трудности* и проблемы. Даже на современной вычислительной машине может потребоваться слишком много времени для выполнения одного «прогона», т. е. одного прохождения системы от начала до конца моделируемого периода времени. Кроме того, этот эксперимент дает значение выхода только для некоторых частных значений параметров и независимых переменных системы, установленных именно в этом опыте. (Параметры системы — это те характеристики, которые остаются постоянными в течение одного опыта-прогона.) Моделирование не дает *функциональной* связи выхода с независимыми переменными и параметрами. Поэтому *«анализ чувствительности»* параметров требует новых прогонов с одним или несколькими наборами новых значений параметров. Эксперименты на «чувствительность» нужны для определения силы влияния вариации какого-то параметра на выход системы. (Если же нужна более точная информация о значении этого параметра, то соответственно требуются дополнительные усилия и деньги для ее получения.)

Приведенное выше соображение означает, что *оптимальное* решение для системы можно получить только экспериментально, т. е. мы должны «попробовать» различные наборы параметров для приближе-

ний к оптимальному набору. При моделировании в узком смысле *стохастические* возмущения создают дополнительные трудности в определении оптимального решения. К счастью, есть методы систематического поиска оптимума, так называемые «методы поверхности отклика», которые будут обсуждены в последующих главах. (Существуют, однако, задачи, которые не «по зубам» и для этих методов. Некоторые ситуации имеют комбинаторный характер (сравни модели расписаний.) Мы можем попытаться найти решение задач подобного рода с помощью *эвристических методов* (см. [Meier et al., 1969, p. 147—178]).

Вглядимся в *стохастическое* моделирование. В этом типе моделирования мы выбираем значения случайных величин в модели. Следовательно, это фактически *статистический выборочный* эксперимент с моделью системы. Поэтому стохастическое моделирование имеет и недостатки, свойственные вообще выборочным методам. Эта выборка содержит в себе все проблемы *статистического планирования и анализа эксперимента*. (Планирование и анализ часто усложняются спецификой имитационного моделирования, приводящей, например, к автокоррелированным данным.) Статистические аспекты моделирования будут рассмотрены в главе II и более подробно обсуждены в последующих главах.

Для дальнейшего обсуждения недостатков (и достоинств) моделирования в узком и в широком смысле мы отсылаем к литературе, например, [Dutton and Starbuck, 1971a, p. 705], [Hillier and Lieberman, 1968, p. 470—471] и [Naylor et al., 1967a, p. 4—9]. Несмотря на все свои недостатки, моделирование, и в узком и в широком смысле, остается весьма важным методом исследования. Моргенталер [Morgenthaler, 1961, p. 366—367, 372—375] перечисляет восемнадцать причин, по которым целесообразно применение моделирования. Здесь мы укажем только, что исследователи конструируют *все более сложные модели*. Такие модели необходимы, как мы уже отметили выше, поскольку лишь при рассмотрении глобальных моделей с включением всех взаимодействий можно избежать локально-оптимальных решений, а решить подобные модели часто можно только моделированием. Мощные современные вычислительные машины и ориентированные на моделирование языки программирования делают эти методы все более доступными для широкого использования. Поэтому не удивительно утверждение Кайвиета [Kiviat, 1967, p. 5] о том, что «в настоящее время комплексная, законченная программа обучения инженеров и менеджеров должна включать в себя имитационное моделирование». Многие упомянутые выше примеры моделирования могут послужить дальнейшими иллюстрациями широкого применения этих методов.

1.6.4. Генерирование случайных чисел и величин

В разделе 1.5 мы кратко рассмотрели псевдослучайные числа и упоминали в качестве примера мультипликативный конгруэнтный генератор. Другие типы генераторов (например, аддитивные, смешанные мультипликативные, среднеквадратичные) можно найти в литературе (см. [Mihram, 1972, p. 44—57], [Naylor et al., 1967a, p. 45—47] и

[Tocher, 1963, p. 72—84] — на первый случай). Выбор значений параметров имеет несколько аспектов. Для выбора таких значений параметров, которые дают достаточную длину цикла, можно использовать теорию чисел. Математическая статистика предлагает методы выбора параметров, которые минимизируют *серийную корреляцию первого порядка* между случайными числами (т. е. корреляцию между r_i и r_{i+1}). Эти методы, однако, обеспечивают только необходимые, но не достаточные условия; более того, они не распространяются на серийные коэффициенты более высокого порядка. Методы выбора параметров можно найти в [J. M. Chambers, 1970] и [R. P. Chambers, 1967], [Jansson, 1966, p. 39—68], [Knuth, 1969, p. 9—33], [Newman and Odell, 1971, p. 7—17]; в учебниках по моделированию, например, [Mihram, 1972], [Naylor et al., 1967a], [Naylor, 1971], [Schmidt and Taylor, 1970] и [Tocher, 1963]. Заметим, что Марсалья [Marsaglia, 1972] отбрасывает все прежние рекомендации и дает новую таблицу рекомендуемых значений мультипликативной постоянной « a » в уравнении (6), для модуля m , равного 2^{32} , 2^{36} или 2^{36} . (Удобно принимать m равным степени 2, быть может, ± 1 , — для двоичных вычислительных машин и степени 10 — для десятичных вычислительных машин.) Рекомендуемые параметры не гарантируют действительно случайных чисел. Поэтому необходимо проверять выход генератора на равномерность и независимость. В рекомендуемой выше литературе обсуждается ряд *статистических критериев*, пригодных для подобных проверок (см. также [Newman and Odell, 1971, p. 75—81]). К сожалению, эти критерии проверяют зависимость только низких порядков, и, как отметили Льюис [Lewis, 1972] и Марсалья [Marsaglia, 1972], зависимости высоких размерностей могут представить серьезную проблему. Эти *высокоразмерные зависимости* можно уменьшить предварительным *таблицированием* случайных чисел (см. [Andrews et al., 1972, p. 63, 306], [Lewis, 1972], [Marsaglia, 1972], [Marsaglia et al., 1972], [Marsaglia and Bray, 1968]). Кавью и Макферсон [Coveyou and MacPherson, 1967] применили *анализ Фурье* для получения априорных данных о статистическом поведении различных генераторов (см. также [Knuth, 1969, p. 82—97] и [Mihram, 1972, p. 473—477]). Анализ Фурье, или спектральный анализ, будет подробнее рассмотрен в II.9. Здесь мы заметим, что все названные авторы не применяют статистические критерии к экспериментальным выходам генераторов, а находят математические отклонения статистических характеристик генераторов. Кавью и Макферсон [Coveyou and MacPherson, 1967, p. 119] пришли к выводу, что «в настоящее время не существует методов генерирования псевдослучайных последовательностей лучших, чем простой мультипликативный конгруэнтный метод с удачно выбранным множителем». Льюис с сотрудниками [Lewis et al., 1969] предложили такой генератор, который действительно выдержал все их проверки и оказался весьма эффективным для системы ИВМ-360. Вскоре был построен очень простой вариант их генератора, который и описан в приложении I.4. Марсалья [Marsaglia, 1972], однако, настойчиво отвергает анализ Кавью и Макферсона и показывает, что Льюис с сотрудниками использовали множитель, который не удовлетворяет проверкам Марсалья; как мы уже упомина-

ли выше, Марсалья дал собственную таблицу рекомендованных множителей.

Мы можем сделать вывод, что после многих лет исследований в этой области все еще нет высококачественного генератора псевдослучайных чисел. Наилучший в настоящее время, по-видимому, *мультипликативный* генератор, однако он требует тщательного выбора *множителя*, иначе результат может быть неудовлетворительным; как указывают Льюис [Lewis et al., 1972, p. 5—8] и Марсалья [Marsaglia, 1972], многие широко распространенные генераторы фактически непригодны. Льюис [Lewis, 1972] рекомендует генератор Льюиса—Гудмена — Миллера; Марсалья предложил свою таблицу множителей; результат можно еще улучшить предварительным *тасованием* чисел. Для дальнейшего изучения лучше всего взять книги [Lewis, 1972], [Marsaglia, 1972] и библиографию, приведенную в [Halton, 1970, p. 32—47] и в [Nance and Overstreet, 1972]. Последняя библиография содержит 491 ссылку на литературу по генераторам случайных чисел, их проверке и использованию для построения выборок с различными распределениями. Выборке случайных величин с произвольным распределением посвящена вторая часть этого пункта.

Выборочные значения случайных величин, использующие случайные числа, — основа стохастического моделирования, и в дальнейшем мы воспользуемся некоторыми из приведенных ниже соотношений. Весьма простое изложение существа процесса формирования выборки имеется у Флэгла [Flagle, 1960, p. 428—435]. Мы рассмотрим широко применимую процедуру.

Обозначим кумулятивную функцию распределения случайной величины x через $F(x)$, т. е.

$$P(x < x) = F(x). \quad (12)$$

Если $F(x)$ — непрерывная возрастающая функция³, то существует обратная функция, обозначенная $F^{-1}(x)$. (Сравните: если y — непрерывная возрастающая функция от z , скажем, $y = g(z)$, то мы можем выразить и z как функцию от y с помощью уравнения $z = g^{-1}(y)$, т. е. $y = g(z)$ разрешимо относительно z .) Теперь мы покажем, что значение x можно выбрать из его распределения $F(x)$ с помощью обратной функции F^{-1} и случайного числа r , а именно

$$x = F^{-1}(r). \quad (13)$$

Из (13) следует, что

$$P(x < x) = P[F^{-1}(r) < x] = P[r < F(x)] = F(x), \quad (14)$$

где последнее равенство следует из (1). Сравнение (14) с (12) показывает, что (13) дает желаемый результат. Проиллюстрируем метод обращения примером (который используется также и в дальнейшем).

Рассмотрим экспоненциальное распределение, которое часто встречается в моделировании. Переменная x распределена по *экспоненци-*

альному закону, если ее плотность распределения $f(x)$ задана следующим выражением:

$$\begin{aligned} f(x) &= \lambda e^{-\lambda x} \text{ для } x \geq 0, \\ &= 0 \text{ для } x < 0, \end{aligned} \quad (15)$$

где $\lambda > 0$. Следовательно, кумулятивная функция распределения вероятностей дается выражением

$$F(x) = \int_0^x \lambda e^{-\lambda t} dt = 1 - e^{-\lambda x} (x \geq 0). \quad (16)$$

Применение (13) означает, что мы принимаем $F(x)$ равным случайному числу r и решаем относительно x , т. е.

$$1 - e^{-\lambda x} = r \quad (17)$$

или

$$x = -\frac{1}{\lambda} \ln(1-r) = -\ln(1-r)/\lambda. \quad (18)$$

Так, уравнение (18) можно использовать для генерации экспоненциальных случайных переменных. Обычно применяется несколько иная формула. Для ее вывода нам нужно показать, что распределение $(1-r)$ — то же самое, что и для r . Это можно сделать, опираясь на общую формулу для плотности распределения преобразованной случайной величины (см. например, [Fisz, 1967, p.39]). Теперь можно достаточно просто показать искомое. Обозначим $(1-r)$ через r . Тогда справедливо (19)

$$\begin{aligned} P(z < z) &= P(1-r < r) = P(r > 1-z) = \\ &= 1 - P(r < 1-z) = 1 - (1-z) = z, \end{aligned} \quad (19)$$

где $0 \leq z \leq 1$.

Сравнение (19) с (1) показывает, что $z (= 1-r)$ имеет то же самое равномерное распределение, что и r . Следовательно, в (18) мы можем заменить $(1-r)$ на r .

Тогда

$$x = -\ln(r)/\lambda. \quad (20)$$

Выражение (20) предпочтительнее, чем (18), так как (18) содержит одно лишнее вычитание, что соответственно увеличит машинное время счета. В главе III тот факт, что r и $(1-r)$ имеют одинаковые распределения, будет снова использован для построения «дополняющего» метода уменьшения дисперсии.

Метод инверсии можно применить и для других распределений. Но для некоторых распределений или необходимо, или более эффективно применение *других методов*. Например, в приложении I.2 нормально

распределенная переменная генерируется с помощью центральной предельной теоремы. Широкое обсуждение метода обращения и других методов можно найти в [Chambers, 1970, p. 6—10], [Halton, 1970, p. 29—32], [Knuth, 1969, p. 100—121], [Kohlas, 1971, p. 20—25], [Mihram, 1972, p. 37—39, 94—146], [Naylor et al., 1967a, p. 68—122], [Newman and Odell, 1971, p. 18—52], [Schmidt and Taylor, 1970, p. 258—324], [Tocher, 1963, p. 6—42]. В этих публикациях рассматривается формирование выборки для самых различных распределений: нормального, χ^2 , F , t , β , γ , Пуассона, логарифмически нормального, многомерного нормального, Уишарта, временного ряда. В некоторых книгах, например, [Mihram, 1972] и [Naylor et al., 1967a], приводятся рекомендации по областям целесообразного применения этих законов распределения. Специальные виды распределений рассматриваются в следующих книгах: в [Ahrens and Dieter, 1972] и [Marsaglia et al., 1972] — нормальное и экспоненциальное распределения; в [Andrews et al., 1972, p. 56—57] — выборка x вида z/y , где z имеет нормальное, а соотвествующий выбор y дает различные распределения; в [van Doeland and van Daal, 1972] — экспоненциальные распределения с произвольными параметрами (см. также [van Daal and van Doeland, 1972], [Ten Broeke, 1972], [van Daal, 1972]). В приложении I.5 мы рассмотрим выборку двух коррелированных переменных, так как это может пригодиться при моделировании капиталовложений (см. [Benecke and Westgaard, 1972]). В разделе VI.3 мы будем генерировать линейные комбинации переменных с экспоненциальными законами распределения для получения распределений с желаемыми характеристиками.

1.7. ЛИТЕРАТУРА

Из всей литературы, приведенной в конце главы, в качестве наилучшего учебника, вводящего во все аспекты моделирования, мы порекомендовали бы книгу [Naylor et al., 1967a]. Хорошие книги по моделированию также [Emshoff and Sisson, 1971], [Meier et al., 1969] и [Tocher, 1963]. В качестве превосходных введений в моделирование, не таких, правда, обширных, как названные выше, можно рекомендовать книги [Flagle, 1960], [Hillier and Lieberman, 1968], [Morgenthaler, 1961], [Nelson, 1966] и [Sisson, 1969]. Среди других ссылок есть также некоторые отличные монографии, которые, однако, посвящены специальным вопросам и поэтому нами здесь не упоминаются.

Для глубокого изучения моделирования мы отсылаем к библиографии по моделированию и играм [Dutton and Starbuck, 1971a, p. 9—102, 693—699, 701—708], дополненной также в книгах [Dutton and Starbuck, 1971b] и [Naylor, 1969]. Это наиболее современные и полные библиографии (свыше 2000 наименований в [Dutton and Starbuck, 1971b]). Для полноты картины мы приведем некоторые библиографии, которые у них не упоминаются, а именно: исследовательского центра А. Д. Р. в Нидерландах [Netherlands A. D. P. Research Center, 1967] и [Shubic, 1970]; недавно Михрам [Mihram, 1971, p. 43—44] представил библиографию работ по моделированию.

ГЛОССАРИЙ (СЛОВАРЬ СПЕЦИАЛЬНЫХ ТЕРМИНОВ)

Этот глоссарий содержит только краткое описание тех понятий, которые необходимы в дальнейшем. Полный и современный глоссарий можно найти в [Mihgam, 1971].

Выборочное распределение	Применение метода Монте-Карло для статистической оценки распределения или некоторых параметров распределения случайной величины.
Модель	Математические символы или блок-схема, представляющие реальную систему.
Метод Монте-Карло в узком смысле	Метод Монте-Карло в широком смысле без моделирования.
Метод Монте-Карло в широком смысле	Любой метод решения модели, использующий <i>случайные</i> (или псевдослучайные) <i>числа</i>
(Псевдо) случайные числа	Равномерно распределенные на интервале $[0, 1]$ независимые случайные величины.
Моделирование в узком смысле	Эксперимент с моделью во времени с использованием случайных величин.
Моделирование в широком смысле	Эксперимент с моделью во времени.
Установившийся режим	Система находится в установившемся режиме, если ее вероятностное распределение не меняется во времени.
Система	Множество элементов, имеющих определенные характеристики или свойства, заданные в числовой или логической форме; заданы отношения между элементами в самой системе и с окружающей средой.

ПРИЛОЖЕНИЯ К ГЛАВЕ I

ПРИЛОЖЕНИЕ 1.1. ОЦЕНИВАНИЕ ИНТЕГРАЛА МЕТОДОМ МОНТЕ-КАРЛО

В этом приложении мы дадим пример использования метода Монте-Карло для решения *детерминированной* задачи. Пусть мы хотим определить значение интеграла $\xi(\lambda, \nu)$, заданного следующим выражением:

$$\xi(\lambda, \nu) = \int_{\nu}^{\infty} \frac{1}{x} \lambda e^{-\lambda x} dx \quad (\lambda, \nu > 0). \quad (1.1)$$

Интеграл (1.1) выглядит очень просто, но он не разрешим непосредственно ни интегрированием по частям, ни разложением в ряд. Можно показать, что справедливо следующее выражение⁴:

$$\xi(\lambda, \nu) = \lambda \left[-c - \ln(\lambda \nu) + \sum_{i=1}^{\infty} \left\{ (-1)^{i+1} \frac{(\lambda \nu)^i}{i! i} \right\} \right], \quad (1.2)$$

где c — постоянная Эйлера. Мы полагаем, что многие нематематики имели бы серьезные затруднения с решением (1.2), а метод Монте-Карло позволяет очень просто оценить (1.1). Это можно сделать так.

1. Выбираем значение x из экспоненциального распределения

$$\begin{aligned} f(x) &= \lambda e^{-\lambda x} \text{ для } x \geq 0, \\ &= 0 \text{ для } x < 0. \end{aligned} \quad (1.3)$$

Для этого применяем выражение (20).

2. Подставляем выборочное значение x в следующее выражение $g(x)$:

$$\begin{aligned} g(x) &= 0, \text{ если } x < v, \\ &= \frac{1}{x}, \text{ если } x \geq v. \end{aligned} \quad (1.4)$$

Математическое ожидание $g(x)$ определяется выражением (1.5):

$$E[g(x)] = \int_{-\infty}^{\infty} g(x) f(x) dx = 0 + \int_v^{\infty} \frac{1}{x} \lambda e^{-\lambda x} dx \xi(\lambda, v). \quad (1.5)$$

3. Многократно повторяем первые два шага для разных случайных чисел. Если через x_i обозначить наблюдение x , произведенное на i -м повторе ($i = 1, 2, \dots, n$), то ξ можно оценить как $\hat{\xi}$ из выражения (1.6):

$$\hat{\xi} = \frac{1}{n} \sum_{i=1}^n g(x_i), \quad (1.6)$$

так как $\hat{\xi}$ есть несмещенная оценка ξ . С увеличением n уменьшается дисперсия этой оценки и возрастает вероятность верного оценивания ξ . В главе III мы вернемся к этому примеру, чтобы показать, как усовершенствованную выборку можно использовать для уменьшения дисперсии в так называемом методе значимой выборки⁶.

Отметим, что интеграл (1.2) может подойти для описания некоторых систем хранения запасов. Согласно Крайензу и де Леве [Kriens and de Leve, 1966, p. 71], можно показать, что для оптимизации некоторой системы запасов надо решить относительно v уравнение (1.7):

$$\int_0^v f(x) dx + v \int_v^{\infty} \frac{1}{x} f(x) dx = \frac{C_2}{C_1 + C_2}, \quad (1.7)$$

где v — начальный запас после получения определенного количества единиц хранения, x — случайный спрос в единицу времени, $f(x)$ — плотность вероятности спроса в единицу времени, C_1 — стоимость единицы хранения в единицу времени, C_2 — стоимость нехватки (дефицита) единицы хранения в единицу времени.

Если $f(x)$ — экспоненциальная плотность вероятности, то решение (1.7) включает в себя определение $\xi(\lambda, v)$ согласно (1.1)⁶. Причем в этой задаче формулировка интеграла (1.1) имеет вероятностную интерпретацию. Но, вообще говоря, мы можем использовать описанную

ише процедуру Монте-Карло для интегралов с подынтегральной функцией $h(x)$; она представима в виде $h_1(x)$, $h_2(x)$, где $h_2(x)$ — функция распределения, для которой осуществляется выборка x и подставляется в $h_1(x)$.

ПРИЛОЖЕНИЕ 1.2. ПРИМЕР ВЫБОРОЧНОГО РАСПРЕДЕЛЕНИЯ

В этом приложении мы будем использовать метод Монте-Карло для оценки вероятности p , определяемой выражениями:

$$p = P(x < a), \quad (2.1)$$

$$x = \min(x_1, x_2), \quad (2.2)$$

где x_1 и x_2 по предположению — независимые случайные величины с нормальным распределением, со средними и дисперсиями, соответственно равными 100, 400 — для x_1 и 90, 100 — для x_2 , т. е.

$$x_1 : N(100, 400)^*, \quad (2.3)$$

$$x_2 : N(90, 100). \quad (2.4)$$

Проблема определения p становится труднее, если мы имеем произведение, состоящее из двух составляющих, распределения которых определены соответственно (2.3) и (2.4). Распределение произведения x определяется выражением (2.2). Мы хотим знать вероятность того, что произведения меньше, чем некоторое заданное число, скажем, a , т. е. мы хотим знать p , определенное выражением (2.1). Физическую интерпретацию этой задачи можно найти у Черчмена [Churchman et al., 1959, p. 174].

Для применения метода Монте-Карло удобно ввести переменную y , определенную выражением (2.5):

$$\begin{aligned} y &= 1, \text{ если } x < a, \\ &= 0, \text{ если } x \geq a. \end{aligned} \quad (2.5)$$

Математическое ожидание y равно p (2.1), так как

$$\begin{aligned} E(y) &= 1 \cdot P(x < a) + 0 \cdot P(x \geq a), \\ &= P(x < a) = p. \end{aligned} \quad (2.6)$$

Здесь мы не станем генерировать нормально распределенную переменную методом обращения (13), а вместо этого используем центральную предельную теорему: сумма «большого» числа независимых одинаково распределенных случайных величин с конечными средними и среднеквадратическими отклонениями имеет приблизительно нормальное распределение (см. [Fisz, 1967, p. 197]). Мы суммировали двенадцать случайных чисел, предполагая, что двенадцать — это достаточно большое число для получения (приближенно) нормально распределенной переменной. Заметим, что случайное число r в (2.7)

* В отечественной литературе это обозначение имеет вид $x_1 \sim N(100, 400)$. — Прим. перев.

имеет среднее 0,5 и дисперсию 1/12. Следовательно, z , определенное выражением (2.7), имеет среднее 0 и дисперсию 1, и если справедлива центральная предельная теорема, то z имеет нормальное распределение.

$$z = \sum_{j=1}^{12} r_j - 6. \quad (2.7)$$

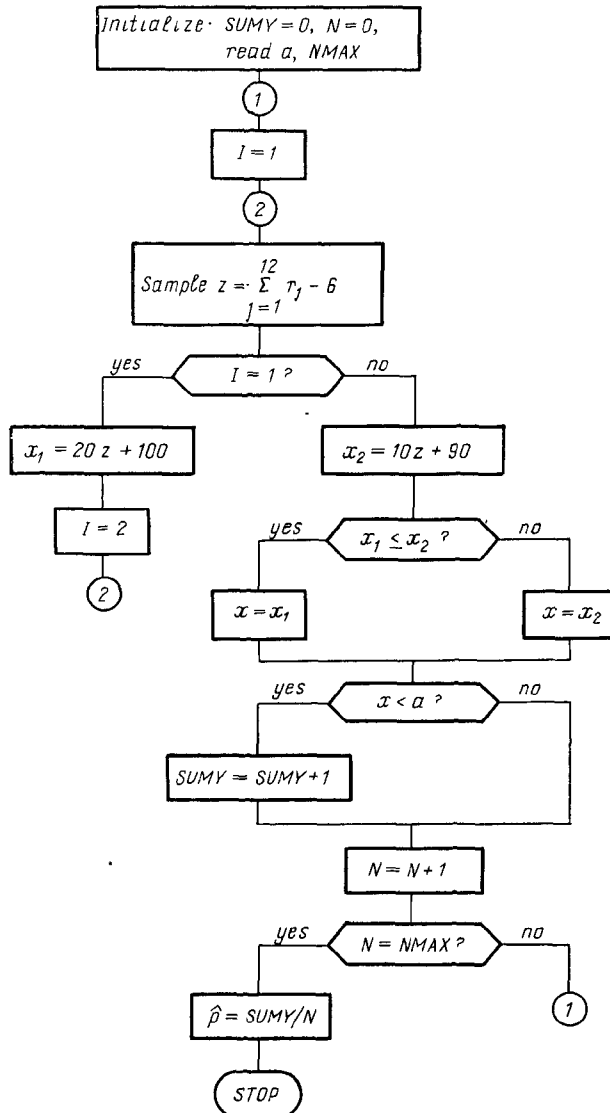


Рис. 6. Оценивание вероятности p методом Монте-Карло

Если z имеет распределение $N(0, 1)$, то $(\sigma z + \mu)$ имеет распределение $N(\mu, \sigma^2)$. Итак, для генерации x_1 и x_2 , определенных (2.3) и (2.4), мы используем выражения (2.8) и (2.9) соответственно:

$$x_1 = 20z + 100, \quad (2.8)$$

$$x_2 = 10z + 90. \quad (2.9)$$

Заметим, что для независимости x_1 и x_2 мы должны использовать различные значения z в (2.8) и (2.9). Новое независимое значение требуется генерации новых двенадцати случайных чисел в (2.7). Для этого мы вводим переключатель (см. переменную I на рис. 6). На рис. 6 показана вся процедура выборки. Согласно (2.6), мы оцениваем p из выражения

$$\hat{p} = \bar{y} = \sum_{j=1}^N y_j / N, \quad (2.10)$$

где $\sum y_j$ обозначено знаком SUMY на блок-схеме⁷. Мы предполагаем, что число значений y , которое должно содержаться в выборке, равно NMAX. В главе V мы покажем, как можно определить NMAX, чтобы оценка с 95%-ной надежностью имела ошибку меньше чем 10%.

Отметим, что для оценки p методом Монте-Карло решающим является знание распределений x_1 и x_2 , однако совсем не обязательно, чтобы вид этих распределений был такой же, как в нашем примере. Два независимых нормальных распределения удобны только для простоты аналитического определения p , которое используется при контроле оценки, найденной методом Монте-Карло.

ПРИЛОЖЕНИЕ 1.3. МОДЕЛИРОВАНИЕ СТРАТЕГИИ ОБСЛУЖИВАНИЯ

Пример, который мы будем рассматривать в этом приложении, относится к выбору одной политики обслуживания из двух возможных (последовательности ремонта). Выше эта проблема была нами сформулирована (см. рис. 3 и соответствующий ему текст). Для удобства читателя мы повторим постановку задачи.

Предположим, автобус должен делать N рейсов в день. Автобус стареет и изнашивается, так что ремонт необходим. Когда автобус находится в *состоянии А*, он еще может работать, но в конце концов потребует ремонта. Если автобус будет продолжать работать без ремонта после того, как он пришел в состояние А, он может перейти в *состояние В*. В состоянии В автобус уже немедленно должен встать на «прикол». *Вероятности перехода* определяются следующими условиями: если автобус уходит в рейс в хорошем состоянии, то он имеет вероятность (a) окончить этот рейс в состоянии А и вероятность $(1 - a)$ остаться в первоначальном состоянии. Если автобус начинает рейс в состоянии А, то он имеет вероятность b закончить рейс в состоянии В и вероятность $(1 - b)$ остаться в состоянии А. *Продолжительность* ремонта разная. Так, если автобус попадает в состояние А, то для его ремонта нужно время, равное времени одного рейса, и в этом случае один рейс придется отменить. Если же автобус придет в состояние В, то все оставшиеся рейсы этого дня должны быть аннулированы — ав-

тобус не пригоден к работе до следующего дня. Предположим, что мы хотим выбрать одну из двух следующих стратегий обслуживания:

стратегия α: как только автобус переходит в состояние А, он ремонтируется;

стратегия β: автобус работает до тех пор, пока не перейдет в состояние В.

Лучшей считается та стратегия, которая дает наибольшее среднее число рейсов в день. Наконец, мы предполагаем, что каждый день автобус выходит на линию в хорошем состоянии, т.е. при стратегии α или β автобус, заканчивающий N рейсов в состоянии А или В, ночью ремонтируется.

Мы сформулировали задачу обслуживания автобуса, но очевидно, что ее можно применить и к обслуживанию любых других машин. Впервые эта модель была построена Морзом [Morse, 1962, p. 98—100], но мы употребляем ее в формулировке Крайенза [Kriens, 1964, p. 31]; в обоих случаях для решения была использована теория *марковских цепей*. Оба автора получили следующие формулы для среднего числа рейсов в день при стратегиях α и β:

$$\mu_\alpha = \frac{N}{1+a} + a \frac{1-(-a)^N}{(1+a)^2}, \quad (3.1)$$

$$\mu_\beta = \frac{a^2 [1-(1-b)^N] - b^2 [1-(1-a)^N]}{ab(a-b)}. \quad (3.2)$$

Мы покажем теперь, как легко можно решить эту задачу с помощью моделирования.

Таблица. 1.3.1. Выборка состояний автобуса в конце рейса

Автобус выезжает в состоянии	Автобус приезжает в состоянии		
	хорошее	А	В
Хорошее А	$a \leq r < 1$ —	$0 \leq r < a$ $b \leq r < 1$	— $0 \leq r < b$

При моделировании мы производим выборку случайного числа r и с помощью табл. 1.3.1 имитируем состояние, в котором находится автобус в конце рейса. Пусть, например, автобус начинает рейс в состоянии А. (Это возможно только при стратегии β.) Если полученное значение r удовлетворяет неравенству $r \geq b$, то автобус и заканчивает рейс в состоянии А. При этом

$$P(r \geq b) = 1 - P(r < b) = 1 - b. \quad (3.3)$$

Выражение (3.3) показывает, что принятая выборочная процедура имеет нужные для данной модели вероятности перехода.

На рис. 7 приведена блок-схема данной имитационной модели. Заметим, что можно проводить моделирование M дней для стратегии α и затем M дней для β. На рис. 7, однако, вначале совершается один рейс при стратегии α и затем при стратегии β; затем совершается следующий рейс, опять вначале при стратегии α, затем при стратегии

β} и т. д. Данная процедура была принята, чтобы простейшим образом использовать метод «общих случайных чисел». Этот метод будет изложен в главе III, где мы снова вернемся к задаче обслуживания.

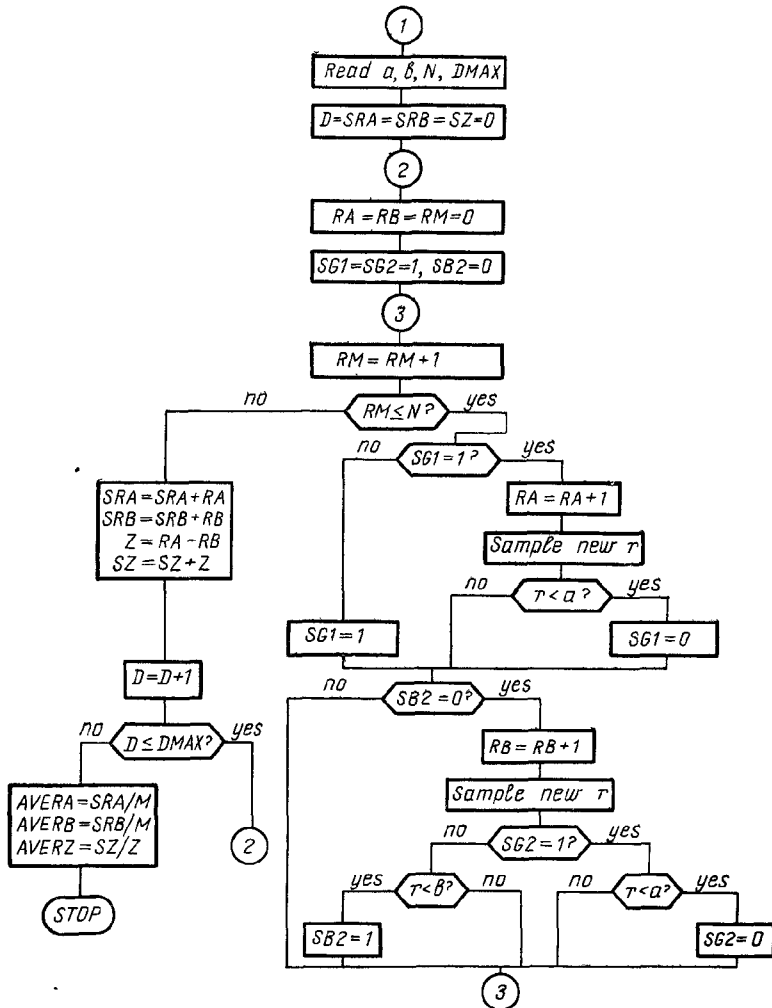


Рис 7 Моделирование двух стратегий обслуживания

На схеме рис. 7 употребляются следующие, не определенные ранее мнемонические обозначения:

SG1

логическая переменная «хорошее состояние при 1-й стратегии»; принимает значение 1, если автобус находится в хорошем состоянии при стратегии α ; и значение 0, если автобус находится в состоянии A (также при стратегии α);

$\hat{S}G2$:	логическая переменная «хорошее состояние при 2-й стратегии»; равна 1, если автобус в хорошем состоянии при стратегии β , и 0, если автобус в состоянии А;
$SB2$:	переменная «состояние В при 2-й стратегии»; равна 1, если автобус в состоянии В при стратегии β , в иных случаях равна 0;
$DMAX$:	число дней, которые мы хотим моделировать;
D :	текущее число дней при моделировании;
RM :	число моделируемых рейсов в день, включая отдельные рейсы;
RA, RB :	число действительно выполненных рейсов в течение дня при стратегиях α и β соответственно;
SRA, SRB :	общее число рейсов в течение D дней при стратегиях α и β соответственно;
Z :	разность между числом действительно выполненных рейсов при стратегиях α и β за определенный день, т. е. $Z = RA - RB$;
SZ :	сумма значений Z за D дней;
$AVERA, AVERB$:	среднее число рейсов в день при стратегиях α и β соответственно после M дней;
$AVERZ$:	$AVERA - AVERB$.

Заметим, что на рис. 7 часть блок-схемы между точками 2 и 3 используется только один раз за каждый моделируемый день; в начале каждого дня еще ни одного рейса не выполнено ($RA = RB = 0$) или замоделировано ($RM = 0$) и автобус начинает рейс в хорошем состоянии при обеих стратегиях ($SG1 = 1, SG2 = 1, SB2 = 0$). После точки 3 RM сравнивается с N для проверки условия, не выполнено ли уже максимальное число рейсов на день. Если при стратегии α автобус начинает рейс в хорошем состоянии ($SG1 = 1$), то этот рейс выполняется ($RA = RA + 1$), и производится выборка состояния в конце этого рейса, т. е. если $r < a$, то автобус уже не в хорошем состоянии ($SG1 = 0$); если $r \geq a$, то никаких изменений не произошло. Если же автобус находится не в хорошем состоянии уже в начале рейса ($SG1 = 0$), то рейс отменяется ($RA \neq RA + 1$), и автобус ремонтируется ($SG1 = 1$). Нет необходимости в выборке случайного числа; после ремонта автобус в любом случае переходит в хорошее состояние. Затем мы определяем с помощью табл. 1.3.1, в каком состоянии закончит автобус рейс при стратегии β . После имитации целого дня ($RM = N$) подсчитывается SRA, SRB и SZ и имитируется следующий день ($D = D + 1$), если еще не достигнуто требуемого числа имитируемых дней ($DMAX$). После имитации $DMAX$ дней подсчитываются среднее число рейсов в день при стратегиях α и β и их разность.

Заметим, что метод имитации позволяет производить изменения в модели простым изменением блок-схемы. Предположим, например что мы отбросили предположение, что при стратегии β автобус всегда начинает день в хорошем состоянии, т. е. мы рассматриваем стратегию γ , при которой автобус, закончивший день в состоянии А, не ремонтируется. Предоставляем читателю произвести необходимые изменения в блок-схеме на рис. 7.

```

*****
* LEWIS-LEARNMOUTH UNIFORM RANDOM NUMBER GENERATOR
* NAVAL POSTGRADUATE SCHOOL MONTEREY CALIFORNIA
*
*****

```

```

SUBROUTINE: RANDOM (ADDITIONAL ENTRY POINT: OVFLOW)

USAGE: RANDOM IS A PSEUDO-RANDOM GENERATOR FOR UNIFORMLY DISTRIBUTED
VARIATES. RANDOM RETURNS BOTH INTEGER*4 VARIATES AND REAL*4
VARIATES. FOR INTEGER VARIATES, RANDOM RETURNS ONE VARIATE
ON THE CLOSED INTERVAL 1 TO 2147483647 PER CALL. FOR REAL*4
VARIATES, RANDOM RETURNS AS MANY VARIATES AS DESIRED (SEE
CALLING SEQUENCE R) ON THE OPEN INTERVAL 0.0 TO 1.0. TO
AVOID THE TIME CONSUMING PROCESS OF DIVISION BY THE MODULUS,
RANDOM REQUIRES A SPECIAL CALL TO ITS ENTRY POINT OVFLOW TO
INTERCEPT FIXED-POINT OVERFLOWS AND TO IMPLEMENT A FASTER
DIVISION SIMULATION ALGORITHM (SEE REFERENCE 2.)

```

```

CALLING SEQUENCE:

1.) ONCE, AT THE BEGINNING OF THE PROGRAM, A CALL MUST BE
MADE TO OVFLOW AS FOLLOWS:
CALL OVFLOW
NO ARGUMENTS ARE REQUIRED FOR THIS CALL.

```

```

2.) TO RECEIVE RANDOM VARIATES, A CALL TO RANDOM IS MADE AS
FOLLOWS:
CALL RANDOM(IX,A,N)
WHERE IX
IS AN INTEGER*4 VARIABLE PRIOR TO CALLING
RANDOM FOR THIS FIRST TIME, IX SHOULD BE
INITIALIZED TO ANY INTEGER VALUE IN THE RANGE
1 TO 2147483647. UPON RETURN FROM RANDOM IX IS
REPLACED BY THE NEXT INTEGER VARIATE IN THE
SEQUENCE. IX SHOULD NOT BE CHANGED BY THE USER
SINCE IT WOULD ALTER THE PERFORMANCE OF RANDOM
AND RECEIVE THE REAL*4 VARIATE(S) WHICH WILL
BE INITIALIZED AND MAY BE CHANGED BY THE USER
IS AN INTEGER*4 VARIABLE WHICH DETERMINES
THE NUMBER OF REAL*4 VARIATES TO BE PLACED
IN THE VECTOR A. IF ONLY INTEGER*4 VARIATES
ARE DESIRED (I.E. THE RETURNED VALUE OF IX)
THEN N MUST BE 1; OTHERWISE THE PROGRAM MAY
ABEND.

```

- RAND0001
- RAND0002
- RAND0003
- RAND0004
- RAND0005
- RAND0010
- RAND0020
- RAND0030
- RAND0040
- RAND0050
- RAND0060
- RAND0070
- RAND0080
- RAND0090
- RAND0100
- RAND0110
- RAND0120
- RAND0130
- RAND0140
- RAND0150
- RAND0160
- RAND0170
- RAND0180
- RAND0190
- RAND0200
- RAND0210
- RAND0220
- RAND0230
- RAND0240
- RAND0250
- RAND0260
- RAND0270
- RAND0280
- RAND0290
- RAND0300
- RAND0310
- RAND0320
- RAND0330
- RAND0340
- RAND0350
- RAND0360
- RAND0370
- RAND0380
- RAND0390
- RAND0400
- RAND0410


```

** REFERENCES:
**
** 1.) LEWIS, P. A. W., A. S. GOODMAN, AND J. M. MILLER, "A PSEUDO-RANDOM
** NUMBER GENERATOR FOR THE SYSTEM/360.", IBM SYSTEMS
** JOURNAL, VOL. 8, NUMBER 2, 1969, PP. 136-146.
**
** 2.) PAYNE, H. M., J. R. RABUNG, AND T. P. BOGYO, "CODING THE LEHMER
** PSEUDO-RANDOM NUMBER GENERATOR, COMMUNICATIONS OF THE
** ACM, VOL. 12, NUMBER 2, FEBRUARY 1969, PP. 85-86.
**
** CATEGORY V1: RANDOM NUMBER GENERATORS
**
** CHECKED OUT AT NPS ON IBM 360/67 BY G. LEARNMOUTH, APRIL 1972.
**
** *****
** EJECT *****
** OVFLOW *****
** ENTRY RANDOM *****
** USING OVFLOW, R12 *****
** LR R14, R15, * *****
** ST R12, R15 *****
** ST R15, SA+4 *****
** LR R2, R13 *****
** LA R13, SA *****
** ST R13+8(, R2) *****
**
** *****
** I.S.S.U.E S.P.I.E M.A.C.R.O T.O F.I.E.L.D I.N.T.E.R.R.U.P.T.S. *****
**
** SPIE XIT, (8,9,11,12,13,15) *****
**
** ST R1,PICA *****
** L R13, SA+4 *****
** RETURN (14,12) *****
**
** *****
** HERE ON INTERRUPT *****
**
** XIT *****
** VM 7(8,11), X, F7, *****
** BC 5, FORT *****
** CLC 17(13, R1), ARANDOM+1, DOES BASE POINT TO RANDOM? *****
** BNE 0(, R14) *****

```

```

RAND0860
RAND0850
RAND0860
RAND0870
RAND0880
RAND0890
RAND0900
RAND0910
RAND0920
RAND0930
RAND0940
RAND0950
RAND0960
RAND0970
RAND0980
RAND0990
RAND1000
RAND1010
RAND1020
RAND1030
RAND1040
RAND1050
RAND1060
RAND1070
RAND1080
RAND1090
RAND1100
RAND1110
RAND1120
RAND1130
RAND1140
RAND1150
RAND1160
RAND1170
RAND1180
RAND1190
RAND1200
RAND1210
RAND1220
RAND1230
RAND1240
RAND1250
RAND1260

```



```

* * LOOP FOR N RANDOM NUMBERS
* * LOOP
MR R4,R9
SLLA R4,1
SRL R5,1
AR R4,R5
LR R5,R4
SRL R4,7
OR R4,R10
ST R4,R7,R6)
CR R4,R11
BNL O(R12)
LE FR2,O(R7,R6)
AER FR2,FRO
STE FR2,O(R7,R6)
BXLE R7,R2,LOOP

NONORM
ST R4,O(R1)
ST R5,O(R4)
R13,SA+4
RETURN (14,12)
EJECT

* * CONSTANTS AND STORAGE
DC F'0'
ARANDOM OC V(RANDOM)
PM2 OC F'2147483645' PM2=2**31-3
A75 DC F'16807'
DC X'40000001'
DC X'40100000'
DS 18F
SA 0
R0 EQU 0
R1 EQU 1
R2 EQU 2
R3 EQU 3
R4 EQU 4
R5 EQU 5
R6 EQU 6
R7 EQU 7
R8 EQU 8
R9 EQU 9
R10 EQU 10
R11 EQU 11
R12 EQU 12

(R5=IX)*(R9=A)
DIVIDE BY P FOR REMAINDER
R4=REMAINDER; R5=QUOTIENT
Q+R = NEXT IX
R5 = IX FOR NEXT GO AROUND
CONVERT R4 TO FLOATING POINT
OR ON AN EXPONENT FIELD
STORE IN X INDEXED BY R7
NEED NORMALIZATION?
NO, BRANCH AROUND
X(R7) TO FLOATING POINT REG. 2
NORMALIZE BY ADDITION OF 0
STORE BACK NORMALIZED

GET ADDRESS OF IX AGAIN
STORE AS INTEGER INTO IX
RESTORE CALLER'S R13

```

```

RAND1620
RAND1630
RAND1640
RAND1650
RAND1660
RAND1670
RAND1680
RAND1690
RAND1700
RAND1710
RAND1720
RAND1730
RAND1740
RAND1750
RAND1760
RAND1770
RAND1780
RAND1790
RAND1800
RAND1810
RAND1820
RAND1830
RAND1840
RAND1850
RAND1860
RAND1870
RAND1880
RAND1890
RAND1900
RAND1910
RAND1920
RAND1930
RAND1940
RAND1950
RAND1960
RAND1970
RAND1980
RAND1990
RAND2000
RAND2010
RAND2020
RAND2030
RAND2040
RAND2050
RAND2060
RAND2070

```

```

R13
R14
RAS
FR0
FR2
END

```

JOB COMPLETE

```

RANDZ080
RANDZ090
RAND2100
RAND2110
RAND2120
RAND2130

```

```

IEE2851
IEF2851
IEF2851
IEE2851
IEF2851
IEF2851
IEF2851
IEF3751
IEF3751
IEF3751
IEF3761
JOB PARAMETERS: CORE=
                    5 SEC. CPU TIME, CLASS=A, PRTY= 7.

```

```

SYS72334.T093206.SV000.80X213$L.R0000001  SYSOUT
VOL SER NOS= SPOOL2
SYS72334.T093206.SV000.80X213$L.R0000002  SYSOUT
VOL SER NOS= SPOOL3
SYS1.PROCLIB
VOL SER NOS= MVTRES
SYS72334.T093206.RV000.80X213$L.S0000003  SYSIN
VOL SER NOS= SPOOL3
SYS72334.T093206.RV000.80X213$L.S0000003  DELETED
VOL SER NOS= SPOOL3
STEP /PRINT / START 72334.0934
STEP /PRINT / STOP 72334.0935 CPU OMIN 00.76SEC MAIN 54K LCS OK
JOB /BOX213$L/ START 72334.0934
JOB /BOX213$L/ STOP 72334.0935 CPU OMIN 00.76SEC
JOB /BOX213$L/ OK COMPLETE & 60K GO, DISK TRKS= 120 SYSOUT & 0 SYSDA,

```

```

54K LCS
OK

```


ПРИЛОЖЕНИЕ 1.5. ВЫБОРКА ДВУХ КОРРЕЛИРОВАННЫХ ВЕЛИЧИН

1. *Общая схема.* Если x и y должны иметь коэффициент корреляции ρ (а средние и дисперсии, соответственно равные $\mu_1, \sigma_1^2, \mu_2, \sigma_2^2$), то найдем

$$y = ax + b + u \quad (5.1)$$

при

$$a = \rho \frac{\sigma_2}{\sigma_1}, \quad b = \mu_2 - \rho \frac{\sigma_2}{\sigma_1} \mu_1$$

$$u = (1 - \rho^2)^{1/2} \sigma_2 z, \quad (5.2)$$

где u — переменная, независимая от x и z , имеет среднее 0 и дисперсию 1.

Доказательство:

$$E(y) = aE(x) + bE(u) = \left(\rho \frac{\sigma_2}{\sigma_1} \mu_1\right) + \left(\mu_2 - \rho \frac{\sigma_2}{\sigma_1} \mu_1\right) + 0 = \mu_2, \quad (5.3)$$

$$\text{var}(y) = a^2 \sigma_1^2 + \sigma_u^2 = \rho^2 \frac{\sigma_2^2}{\sigma_1^2} \sigma_1^2 + (1 - \rho^2) \sigma_2^2 = \sigma_2^2, \quad (5.4)$$

$$E(x, y) = E[ax^2 + bx + ux] = \rho \frac{\sigma_2}{\sigma_1} (\sigma_1^2 + \mu_1^2) + \left(\mu_2 - \rho \frac{\sigma_2}{\sigma_1} \mu_1\right) \mu_1 +$$

$$+ E(u)E(x) = \rho \sigma_2 \sigma_1 + \rho \frac{\sigma_2}{\sigma_1} \mu_1^2 + \mu_1 \mu_2 - \rho \frac{\sigma_2}{\sigma_1} \mu_1^2 + 0 =$$

$$= \rho \sigma_2 \sigma_1 + \mu_1 \mu_2, \quad (5.5)$$

$$\text{Corr}(x, y) = \frac{E(x, y) - E(x)E(y)}{\sigma_1 \sigma_2} = \frac{\rho \sigma_2 \sigma_1 + \mu_1 \mu_2 - \mu_1 \mu_2}{\sigma_1 \sigma_2} = \rho. \quad (5.6)$$

2. Более строгое условие, что x и y должны иметь особые *распределения* (кроме удовлетворения условиям на средние и дисперсии). Тогда приведенная выше схема может работать, если правильно выбрать u . Например, пусть x и y должны иметь нормальное распределение. Сумма нормально распределенных переменных также имеет нормальное распределение.

Из (5.1) мы получим

$$y = \rho \frac{\sigma_2}{\sigma_1} x + \mu_2 - \rho \frac{\sigma_2}{\sigma_1} \mu_1 + (1 - \rho^2)^{1/2} \sigma_2 z, \quad (5.7)$$

где $z : N(0, 1)$. Этот частный случай (5.1) соответствует результату, полученному другим путем Нейлором [Naylor et al., 1967a, p. 99]. Для других распределений задача не так проста. Приближенное решение можно получить как показано далее.

3. *Максимальная зависимость.* Возможен случай, когда корреляция или, в более общем случае, зависимость, должна не иметь определенное значение, а быть максимальной.

1. *Максимальная линейная корреляция:* $\rho = +1$ или $\rho = -1$. Уравнение (5.1) принимает вид:

$$y = \alpha x + \beta \quad (5.8)$$

при

$$\alpha = -\frac{\sigma_2}{\sigma_1}, \quad \beta = \mu_2 - \frac{\sigma_2}{\sigma_1} \mu_1 \quad \text{для } \rho = +1 \quad (5.9)$$

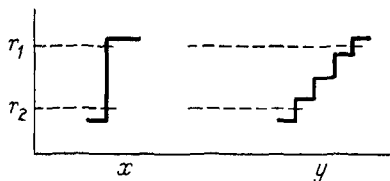
и

$$\alpha = -\frac{\sigma_2}{\sigma_1}, \quad \beta = \mu_2 + \frac{\sigma_2}{\sigma_1} \mu_1 \quad \text{для } \rho = -1. \quad (5.10)$$

2. *Максимальная положительная зависимость.* Мы можем использовать одни и те же случайные числа для получения x и y .

а) *Непрерывные, возрастающие распределения для x и y .* Если

$$x = F_1^{-1}(r), \quad (5.11)$$



то

$$y = F_2^{-1}(r) = F_2^{-1}[F_1(x)] = g(x). \quad (5.12)$$

Так что

Рис. 8. Выборка x и y с общими случайными числами

$$\text{var}(y|x=x) = 0,$$

$$E(y|x=x) = g(x). \quad (5.13)$$

Заметим, что $g(x)$ может быть линейным, как в (5.8), т. е. для экспоненциальных x и y мы имеем

$$y = -\frac{1}{\lambda_2} \ln(r) = -\frac{1}{\lambda_2} \ln(e^{-\lambda_1 x}) = \frac{\lambda_1}{\lambda_2} x. \quad (5.14)$$

б) *Дискретное распределение.* На рис. 8 r_1 и r_2 дают одинаковый x , но разные y . Так как $\text{var}(y|x=x) > 0$, y не полностью зависит от x . Во всяком случае, можно ожидать между x и y положительной корреляции. Если распределения симметричны, то их средние и медианы совпадают; отсюда ясно, что x и y положительно коррелированы (Сравни: если $r > 0,5$, то $x > \mu_1$ и $y > \mu_2$.)

3. *Максимальная отрицательная зависимость.* Возьмем r и его дополнение $(1-r)$ для x и y соответственно. Если распределение x и y одинаково ($F_1 = F_2 = F$), то, как показал Андреассон [Andréasson, 1971, р. 5], эта процедура минимизирует корреляцию ρ , причем если F симметрично, то $\rho = -1$.

У П Р А Ж Н Е Н И Я

1. Почему длительность цикла генератора умножения в (б) может быть не более чем m ? Почему длительность цикла в следующем генераторе (так называемом генераторе Фибоначчи) не больше m : $x_i = x_{i-1} + x_{i-2}$ (по модулю m)?

2. Как использовать метод Монте-Карло для того, чтобы оценить интеграл

$$\int_a^{\infty} \exp\left\{-\frac{1}{2}\left(\frac{x-b}{c}\right)^2\right\} \ln x \, dx$$

для известных a , b и c^2

3. Создайте программу для системы массового обслуживания с дисциплиной очереди «первый пришел — первый обслужен», имеющей одно обслуживающее устройство на рис. 4.

4. Дайте аналитический вывод для равенства $p = P(x < a)$; см. (2.1), используя (2.4). Каково значение p при $a = 83$?

5. Настроить программу на рис. 7 для того, чтобы смоделировать стратегию γ , определенную в конце приложения 1.3.

6. Рассмотрим равенство $P(x = x_i) = p_i$ ($i = 1, \dots, n$). Каково преимущество перестановки величин x_i в порядке увеличения p_i при отборе x ?

7. Рассмотрим экономическую модель $y_t = \alpha + \beta x_t + \beta \varepsilon_t$ при $\varepsilon_t : N(0, \sigma^2)$. Предположим, даны (x_t, y_t) при $t = 1, \dots, n$. Из какого распределения вы отберете значения ошибки в модели? Будете ли вы также отбрасывать значения оценок параметров $\hat{\alpha}$ и $\hat{\beta}$?

ПРИМЕЧАНИЯ

¹ Например, [Ackoff, 1971], [Emery, 1969], [Emshoff and Sisson, 1971, p.250—251], [Eukhoff et al., 1966], [van Dixhoorn and Lyesen, 1968].

² Другой способ для оценивания статистического метода (например, простого метода наименьших квадратов) прилагает этот же метод к *части* имеющихся данных, что приводит к оцениванию параметров. Эти оцениваемые параметры применяются для того, чтобы предсказать имеющиеся данные, которые не были использованы в оценке параметров. «Предсказанные» данные можно сравнить с действительными имеющимися данными [Graver, 1969].

³ Вспомним, что « $y = g(z)$ является функцией», которая ставит в соответствие каждому значению z не более чем одно значение y . Если $F(x)$ непрерывная, нубывающая функция (интегральное распределение)* или дискретная функция (распределение), то метод инверсии остается работоспособным.

За точным выводом мы отсылаем читателя к работе [Haitsma and Oosterhoff, 1964, p.12—13]; практически особых проблем не возникает, как следует из выборочной процедуры для дискретных функций распределения в [Naylor et al., 1967a, p.115].

⁴ Мы обязаны этому решению покойному Л. Липцу (Katholieke Hogeschool Tilburg).

⁵ Другая процедура, предложенная нам Г. Ломберсом (H. Lombaers — Technische Hogeschool Delft), который уменьшил машинное время следующим образом. Из (20) видно, что $x < v$, если $r > \exp(-\lambda v)$. Для таких значений r мы не вычисляем x , поскольку (1.4) показывает, что $g(x) = 0$ в любом случае. Таким образом, мы избегаем затрат времени на вычисления логарифма. Так, величина ξ оценивается выражением

$$\hat{\xi} = \frac{-\lambda}{n} \sum_j \frac{1}{\ln(r_j)},$$

где r_j — такие случайные числа среди « n » случайных чисел, которые не превышают константу $\exp(-\lambda v)$.

⁶ В (1.1) λ должно быть положительным, поскольку это параметр экспоненциальной функции плотности распределения. В [Kleijnen, 1968, p.180] показано, что v также должно быть положительным, так как иначе интеграл в (1.1) будет расходящимся.

⁷ Можно аргументировать то, что переменные в программе должны быть подчеркнуты для того, чтобы показать их стохастический характер. На самом деле эти переменные являются символическими адресами памяти машины. Поэтому мы не подчеркиваем переменные в программах.

⁸ Запись этого генератора воспроизведена с доброго согласия П. Льюиса.

⁹ Значительно больше упражнений можно найти в учебниках, например в [Meier et al., 1969], [Mihram, 1972], [Naylor et al., 1967a] и [Schmidt and Taylor, 1970].

1. Ackoff R. L. (1971). Towards a system of systems concepts. *Management Sci.*, 17, 661—671.
2. Ahrens J. H. and Dieter U. (1972). Computer methods for sampling from the exponential and normal distributions. *Commun. ACM*, 15, 873—882.
3. Andréasson I. J. (1971). On the generation of negatively correlated random numbers. Report Na 71.32, Department of Information Processing, The Royal Institute of Technology, Stockholm.
4. Andrews D. F., Bickel P. J., Hampel F. R., Huber P. J., Rogers W. H. and Tukey J. W. (1972). *Robust Estimates of Location*. Princeton University Press, Princeton, N. Y.
5. Bauknecht K. (1971). Programmiersprachen für Simulationen auf Rechenautomaten (Programming languages for simulations on computers). — In: *Digitale Simulation* (K. Bauknecht and W. Nef, eds.), Springer, Berlin.
6. Benecke R. W. and Westgaard T. P. (1972). On the Use of Risk Analysis in Business. University of Nebraska, Omaha, Neb.
7. Bovee D. P. (1972). On the Use of Models Employing Both Simulation and Analytical Solutions for Scheduling Computing Centers. — In: *Working Papers, vol. 2; Symposium Computer Simulation versus Analytical Solutions for Business and Economic Models*, Graduate School of Business Administration, Gothenburg (Sweden).
8. Burt J. M. (1972). Resource Allocation in Stochastic Project Net-Works: A Simulation and Programming Approach. Graduate School of Management, University of California, Los Angeles, Calif.
9. Burt J. M., Graver D. P. and Perlas M. (1970). Simple stochastic networks: some problems and procedures. *Naval Res. Logistics Quart.*, 17, 439—459.
10. Carter E. E. (1971). A simulation approach to investment decision. *California Management Rev.*, 13, 18—26.
11. Chambers J. M. (1970). Computers in statistical research: simulation and computer-aided mathematics. *Technometrics*, 12, 1—15.
12. Chambers R. P. (1967). Random-number generation. *IEEE Spectrum*, 4, 48—56.
13. Churchman C. W., Ackoff R. L. and Arnoff E. L. (1959). *Introduction to Operations Research*, Wiley, New York.
14. Clough D. J., Levine J. B., Mowbray G. and Walter J. R. (1965). A simulation model for subsidy policy determination in the Canadian uranium mining industry. *Can. Operational Res. Soc. J.*, 3, 115—128.
15. Cohen K. J. (1960). *Computer Models of the Shoe, Leather, Hide Sequence*. Prentice-Hall, Englewood Cliffs, N. Y.
16. *Computing Reviews*. Association for Computing Machinery, New York.
17. Conway R. W. (1963). Some tactical problems in digital simulation. *Management Sci.*, 10, 47—61.
18. Conway R. W., Johnson B. M. and Maxwell W. L. (1959). Some problems of digital systems simulation. *Management Sci.*, 6, 92—110.
19. Coplin W. D. (ed) (1968). *Simulation in the Study of Politics*. Markham, Chicago, Ill.
20. Coveyou R. R. and MacPherson R. D. (1967). Fourier analysis of uniform random number generators. *J. Assoc. Computing Machinery*, 14, 100—119.
21. Dewan P. B., Donaghey C. E. and Wyatt J. B. (1972). OSSL — A specialized language for simulating computer systems. — In: *Proc. AFIPS 1972 Spring Joint Computer Conf.*, AFIPS Press, Montvale, N. J.
22. Donaldson T. S. (1966). Power of the F-Test for Nonnormal Distribution and Unequal Error Variances. RM-5072-PR, The Rand Corporation, Santa Monica, Calif.
23. Driehuis W. (1972). *Fluctuations and Growth in a Near Full Employment Economy*. Universitaire Pers, Rotterdam.
24. Dutton J. M. and Starbuck W. H. (1971a). *Computer Simulation of Human Behavior*, Wiley, New York.

25. Dutton J. M. and Starbuck W. H. (1971b). Computer simulation models of human behavior: A history of an intellectual technology. *IEEE Trans: Systems, Man, Cybernetics, SMC—1*, 128—171.
26. Elton M. and Rosenhead J. (1971). Micro-simulation of markets. *Operational Res. Quart.*, 22, 117—144.
27. Emery F. E. (ed.) (1969). *Systems Thinking*, Penguin, Harmondsworth, England.
28. Emshoff J. R. and Sisson R. L. (1971). *Design and Use of Computer Simulation Models*, Macmillan, New York.
29. Eykhoff P., van der Grinten P. M. E. M., Kwakernaak H. and Veltman B. P. T. (1966). Systems modelling and identification. — In: *Automatic and Remote Control, III, Proc. 3rd Congr. IFAC, United Kingdom Automation Council*, London, 1966.
30. Fisz M. (1967). *Probability Theory and Mathematical Statistics*, 3rd ed., Wiley, New York.
31. Flagle C. D. (1960). Simulation techniques. — In: *Operations Research and Systems Engineering* (C. D. Flagle, W. H. Huggins and R. H. Roy, eds.), The Johns Hopkins Press, Baltimore, Md.
32. Forrester J. (1961). *Industrial Dynamics*. M. I. T. Press, Cambridge, Mass. Русский перевод: Форрестер Дж. *Промышленная динамика*. М., «Прогресс», 1971.
33. Geisler M. A. (1960). The use of man-machine simulation for support planning. *Naval Res. Logistics Quart.*, 7, 421—428.
34. Graver C. A. (1969). *Historical Simulation: A Procedure for the Evaluation of Estimating Procedures*, vol. 1, Contract DACH15-68-C-0364, General Research Corporation, Santa Barbara, Calif.
35. Guetzkow H. (1971). Simulations in the consolidation and utilization of knowledge about international relations. — In: *Cybernetics, Simulation and Conflict Resolution* (D. E. Knight, H. W. Curtis and L. J. Fogel, eds.), Spartan Books, New York.
36. Guetzkow H., Kotler P. and Schultz R. L. (1972). *Simulation in Social and Administrative Science*, Prentice-Hall, Englewood Cliffs, N. J.
37. Haitsma A. H. and Oosterhoff J. (1964). *Monte Carlo Methoden (Monte Carlo methods)*. Rapport S 265 (C13), Leergang Besliskunde, Hoofdstuk XVI, Stichting Mathematisch Centrum, Amsterdam.
38. Halton J. H. (1970). A retrospective and prospective survey of the Monte Carlo method. *SIAM Rev.*, 12, 1—63.
39. Hammersley J. M. and Handscomb D. C. (1964). *Monte Carlo Methods*. Wiley, New York; Methuen, London.
40. Hanken A. F. G. and Buys B. G. F. (1971). Systems analysis and business models. *Ann. Systems Res.*, 1, 9—16.
41. Hauser N., Barish N. N. and Ehrenfeld S. (1966). Design problems in a process control simulation. *J. Ind. Eng.*, 17, 79—86.
42. Heuts R. M. J. and Rens P. J. (1972). A Monte Carlo Study of the Kuiper Test Statistic for Testing Exponentiality (Two Different Approaches). R. C.-notitie nr. 13, Rekencentrum Katholieke Hogeschool, Tilburg, Netherlands.
43. Hillier F. S. and Lieberman G. J. (1968). *Introduction to Operations Research*. Holden-Day, San Francisco, Calif., Chapter 14.
44. Horn W. (1968). *Het Aanbod van Varkens in Nederland (The supply of hogs in the Netherlands)*. Centrum voor Landbouwpublicaties en Landbouwdocumentatie, Wageningen, Netherlands.
45. Howrey E. P. (1972). Selection and Evaluation of Econometric Models. — In: *Working Papers*, vol. 4, Symposium Computer Simulation versus Analytical Solutions for Business and Economic Models, Graduate School of Business Administration, Gothenburg, Sweden.
46. Howrey E. P. (1966). *Stabilization Policy in Linear Stochastic Systems*. Econometric Research Program, Princeton University, Princeton, N. J.
47. Howrey P. and Kelejian H. H. (1969). Simulation versus analytical solutions. — In: *The Design of Computer Simulation Experiments*. T. H. Naylor (ed.), Duke University Press, Durham, N. C. (Reprinted in: T. H. Naylor.

- Computer Simulation Experiments with Models of Economic Systems, Wiley, New York, 1971.)
48. Hutter R. (1970). Simulation eines Informationssystems mit Benutzerprioritäten (Simulation of an information processing system with users' priorities) *Ablauf Planungsforsch*, 11, 238—243.
 49. International Abstracts in Operations Research. International Federation of Operations Research Societies (c/o N. Miller, 428 E. Preston St., Baltimore, Md.)
 50. Jansson B. (1966). Random Number Generators. Viktor Pettersons Bokindustri Aktiebolag, Stockholm.
 51. Jessop W. N. (1956). Monte Carlo methods and industrial problems. *Appl. Stat.*, 5, 158—165
 52. Johnston J. (1963). *Econometric Methods*. McGraw-Hill, New York.
 53. Kay I. M. (1972). An over-the-shoulder look at discrete simulation languages. — In: Proc. AFIPS 1972 Spring Joint Computer Conf., AFIPS Press, Montvale, N. J.
 54. Kiviat P. J. (1967). *Digital Computer Simulation: Modeling Concepts RM-5378-PR*, The Rand Corporation, Santa Monica, Calif.
 55. Kleijnen J. P. C. (1970). *Simulation Terminology and Application Areas*, Working Paper, Katholieke Hogeschool, Tilburg, Netherlands.
 56. Kleijnen J. P. C. (1968). Een toepassing van „importance sampling“ (An application of importance sampling). *Stat. Neerl.*, 22, 179—198. (Also published as report EIT, no. 2, of the Economisch Instituut Tilburg, Econometrische Afdeling, Tilburg, Netherlands).
 57. Klingel A. R. (1966). Bias in PERT project completion time calculations for a real network. *Management Sci.*, 13, 194—201.
 58. Kmenta J. and Gilbert R. F. (1968). Small sample properties of alternative estimators of seemingly unrelated regressions. *J. Amer. Stat. Assoc.*, 13, 1180—1200
 59. Knuth D. E. (1969). *The Art of Computer Programming. Vol. 2, Semi-Numerical Algorithms*, Addison-Wesley, Reading, Pa.
Русский перевод: Кнут Д. Искусство программирования для ЭВМ. Т. 2. М., «Мир», 1977.
 60. Kohlas J. (1971). Die Monte Carlo Methode (The Monte Carlo method). In: *Digitale Simulation* (K. Bauknecht and W. Nef, eds.), Springer, Berlin
 61. Koorman B. O. (1970). A study on the logical basis of combat simulation. *Operations Res.*, 18, 855—882.
 62. Kriens J. (1964). *Markovketens (Markov chains)*. Katholieke Hogeschool, Tilburg, Netherlands.
 63. Kriens J. and de Leve G. (1966). *Inleiding tot de Mathematische Besliskunde (Introduction to operations research)*. *Leergang Besliskunde*, 1.5, Mathematisch Centrum, Amsterdam.
 64. Lewis P. A. W. (1972). Large-Scale Computer-Aided Statistical Mathematics. Naval Postgraduate School, Monterey, Calif. — In: *Proc. Computer Science and Statistics: 6th Annual Symp. Interface*, Western Periodical Co, Hollywood, Calif.
 65. Lewis P. A. W., Goodman A. S. and Miller J. M. (1969). A pseudo-random number generator for the system/360. *IBM Systems J.*, 8, 136—147.
 66. Lucas H. C. (1971). Performance evaluation and monitoring. *Computing Surveys*, 3, 79—91.
 67. McCracken D. D. (1955). The Monte Carlo method. *Scientific American*, 192, 90—97.
 68. Maguire J. N. (1972). Discrete computer simulation-technology and applications — the next ten years. Proc. AFIPS 1972 Spring Joint Computer Conf., AFIPS Press, Montvale, N. J.
 69. Maisel H. and Gnugnoli G. (1972). *Simulation of Discrete Stochastic Systems*. Science Research Associates, Palo Alto, Calif.
 70. Mann N. R. (1970). Computer-aided selection of prior distributions for generating Monte-Carlo confidence bounds on system reliability. *Naval Res. Logistics Quart.*, 17, 41—54.

71. Marsaglia G. (1972). The structure of linear congruential sequences — In: Applications of Number Theory to Numerical Analysis, S. K. Zaremba (ed.), Academic, New York.
72. Marsaglia G. and Bray T. A. (1968). One-line random number generators and their use in combinations. *Commun. ACM*, 11, 757—759.
73. Marsaglia G., Ananthanarayanan K. and Paul N. (ca. 1972). How to Use the McGill Random Number Package "SUPER—DUPER", School of Computer Science, McGill University, Montreal.
74. Mededelingen Operationele Research (Information on Operations Research), Vereniging voor Statistiek (c/o J. Meinardi, Ant. Duyckstraat 102, The Hague).
75. Meier R. C., Newell W. T. and Pazer H. L. (1969). *Simulation in Business and Economics*. Prentice-Hall, Englewood Cliffs, N. J.
76. Mihran G. A. (1972). *Simulation: Statistical Foundations and Methodology*. Academic, New York.
77. Mihran G. A. (1971). A glossary of simulation terminology. *J. Stat. Computation Simulation*, 1, 35—44.
78. Mikhail W. M. (1972). Simulating the small-sample properties of econometric estimators. *J. Amer. Stat. Assoc.*, 67, 620—624.
79. Molenaar W. (1968). Simulatie (Simulation). — In: *Minimaxmethode, Network Planning, Simulatie* (J. Kriens, F. Gobel and W. Molenaar, eds.). Leergang Besliskunde, 1.8, Mathematisch Centrum, Amsterdam.
80. Moore J. M. and Wilson R. C. (1967). A review of simulation research in job shop scheduling. *Production Inventory Management*, 8, 1—10.
81. Morgenthaler G. W. (1961). The theory and application of simulation in operations research. — In: *Progress in Operations Research* (R. L. Ackoff, ed.), Wiley, New York.
82. Morse P. M. (1962). Markov processes. — In: *Notes on Operations Research*, assembled by Operations Research Center, M. I. T., The M. I. T. Press, Cambridge, Mass.
83. Nance R. E. and Overstreet C. (1972). A bibliography on random number generation. *Computing Rev.*, 13, 495—508.
84. Naylor T. H. (1971). *Computer Simulation Experiments with Models of Economic Systems*, Wiley, New York. Русский перевод: Нейлор Т. Машинные имитационные эксперименты с моделями экономических систем. М., «Мир», 1975.
85. Naylor T. H. (1969). Bibliography 19; Simulation and gaming. *Computing Rev.*, 10, 61—69.
86. Naylor T. H., Balintfy J. L., Burdick D. S. and Chu K. (1967a). *Computer Simulation Techniques*. Wiley, New York.
87. Naylor T. H., Burdick D. S. and Sasser W. E. (1967b). Computer simulation experiments with economic systems: The problem of experimental design. *J. Amer. Stat. Assoc.*, 62, 1315—1337.
88. Naylor T. H., Wertz K. and Wonnacott T. (1968). Some methods for evaluating the effects of economic policies using simulation experiments. *Rev. Intern. Stat. Inst.*, 36, 184—200.
89. Neave H. R. and Granger C. W. J. (1968). A Monte Carlo study comparing various two-sample tests for differences in mean. *Technometrics*, 10, 509—522.
90. Neeleman D. (1973). *Multicollinearity in Linear Economic Models*, Tilburg University Press, Tilburg, Netherlands.
100. Nelson R. T. (ca. 1966). *Systems, Models and Simulation*. Mimeographed notes, Graduate School of Business Administration, University of California at Los Angeles, Los Angeles, Calif.
101. Netherlands A. D. P. Research Center (1967). *Simulation. Series special bibliographies no. SB 39*, Amsterdam.
102. Newman T. G. and Odell P. L. (1971). *The Generation of Random Variates*. Griffin, London.
103. Niedereichholz J. (1971). Business simulation and management decisions. *Management Intern. Rev.*, 11, 47—52.

104. Nisenfeld A. E. (1971). Analog and digital simulation: A comparison. *Instrument Control Systems*, 44, 91—93.
105. Operations Research-Management Science. Executive Sciences Institute, Inc., Whippany, N. J.
106. Orcutt G. H., Greenberger M., Korbcl J. and Rivlin A. M. (1961). *Microanalysis of Socioeconomic Systems: A Simulation Study*. Harper and Row, New York.
107. Pomerantz A. G. (1970). Predict your system's fortune: use simulation's crystal ball. *Computer Decisions*, 2, 16—19.
108. Rechtschaffen R. N. (1972). Queuing simulation using a random number generator. *IBM Systems J.*, 11, 255—271.
109. Reith P. F. (1969). Computers in de econometrie (Computers in econometrics). *IBM Kwartaalschrift*, 6, 49—53.
110. Rose J. (1970). *Progress of Cybernetics*. Vols. 1, 2, 3, Gordon and Breach, London.
111. Rytz R. (1971). Sim — ein neues Simulationskonzept (SIM — A new simulation concept). — In: *Digitale Simulation* (K. Bauknecht and W. Nef. eds.), Springer, Berlin.
112. Salazar R. C. and Sen S. K. (1968). A simulation model of capital budgeting under uncertainty. *Management Sci., Application Ser.*, 15, 161—179.
113. Sasser W. E. (1969). A Causal Relationship between a Model's Characteristics and the Performances of the Estimators of the Model's Parameters: A Pilot Study. Graduate School of Business Administration. Harvard University, Boston, Mass.
114. Schink W. A. and Chiu J. S. Y. (1966). A simulation study of effects of multi-collinearity and autocorrelation on estimates of parameters. *J. Financial Quantitative Analysis*, 1, 36—67.
115. Schmidt J. W. and Taylor R. E. (1970). *Simulation and Analysis of Industrial Systems*. Richard D. Irwin, Inc., Homewood.
116. Shannon R. E. and Biles W. E. (1970). The utility of certain curriculum topics to operations research practitioners. *Operations Res.*, 18, 741—745.
117. Бусленко П. Н., Голенко Д. И., Соболев И. М., Срагович В. Г., Шрейдер Ю. А. *Метод статистических испытаний (метод Монте-Карло)*. Под ред. Ю. А. Шрейдера. М., Физматгиз, 1962.
118. Shubik M. (1972). On gaming and game Theory. *Management Sci., Professional Series*, 18, 37—53.
119. Shubik M. (1970). A Preliminary Bibliography on Gaming. Department of Administrative Sciences, Yale University, New Haven, Conn.
120. Shubik M. (1960). Bibliography on simulation, gaming, artificial intelligence and allied topics. *J. Amer. Stat. Assoc.*, 55, 736—751.
121. Shubik M. and Wolf G. (1972). A List of Publications Relevant to the Game and Its Analysis. Appendix B of Gaming Newsletter, Department of Administrative Sciences, Yale University, New Haven, Conn.
122. Sisson R. L. (1969). Simulation: Uses: — In: *Progress in Operations Research*, vol. 3 (Y. S. Aronofsky, ed.), Wiley, New York.
123. Smith J. (1968). *Computer Simulation Models*. Griffin, London.
124. Smith J. U. M. (1970). Computer simulation of industrial operations. — In: *Progress of Cybernetics*, vol. 2 (J. Rose, ed.), Gordon and Breach, London.
125. Teichrow D. (1965). A history of distribution sampling prior to the era of the computer and its relevance to simulation. *J. Amer. Stat. Assoc.*, 60, 27—49.
126. Ten Broeke A. M. (1972). Commentaar op artikel Simulating inter-arrival times with time-dependent arrival rates (Comment on article Simulating inter-arrival times with time-dependent arrival rates). *Mededelingen Operationele Res.*, 11, 174—177.
127. Theil H. and Boot J. C. G. (1962). The final form of econometric equation systems. *Rev. Intern. Stat. Inst.*, 30, 136—152.
128. Tocher K. D. (1963). *The Art of Simulation*. English Universities Press, London.

129. Tocher K. D. (1966). The state of the art of simulation — A survey. Proc. 4th Intern. Conf. Operational Res.
130. Turban E. (1972). A sample survey of operations-research activities at the corporate level. *Operations Res.*, 20, 708—721.
131. Van Daal J. (1972). Antwoord (Reply). *Mededelingen Operationele Res.*, 11, 208.
132. Van Daal J. and van Doeland F. (1972). Waiting for the Bridge. Report 7208, Econometric Institute, Netherlands School of Economics, Rotterdam.
133. Van Dixhoorn J. J. and Lyesen D. P. (1968). Simulatie: Methodologische en Systemtheoretische Aspecten, Kenmerk 2410330, Dhn/Sks, Afdeling der Electrotechniek, Technische Hogeschool Twente, Enschede, Netherlands.
134. Van Doeland F. and van Daal J. (1972). Simulating inter-arrival times with time-dependent arrival rates. *Mededelingen Operationele Res.*, 11, 107—119.
135. Van Slyke R. M. (1963). Monte Carlo methods and the PERT problem. *Operations Res.*, 11, 839—860.

Глава II ● СТАТИСТИЧЕСКИЕ АСПЕКТЫ МОДЕЛИРОВАНИЯ¹

II.1. ВВЕДЕНИЕ И РЕЗЮМЕ

В этой главе мы дадим обзор статистических аспектов моделирования. Особое внимание будет уделено анализу входных данных, начальным условиям, методам уменьшения дисперсии, обоснованности, экспериментальным планам (включая планы отсеивающих экспериментов и планы для анализа поверхности отклика), процедурам множественного сравнения и ранжирования, объему выборки в ограниченных и неограниченных системах и продолжительности опытов при моделировании установившегося состояния системы. Невозможно обсудить все статистические аспекты моделирования достаточно подробно. Поэтому в последующих главах будут рассмотрены только некоторые аспекты моделирования. В настоящей главе даны ссылки, позволяющие читателю (в той степени, в какой это возможно) изучить другие аспекты этой темы, представляющие для него интерес.

По определению, данному в главе I, под *моделированием* подразумевается экспериментирование с математической моделью во времени. Большинство имитационных моделей содержит случайные величины, например случайное время обслуживания. Мы ограничимся рассмотрением *таких* моделей, или, иначе говоря, моделированием по методу Монте-Карло. Стохастическое моделирование на самом деле просто *выборочный эксперимент* (хотя существуют эксперименты специального типа, которые выполняются на модели взамен испытаний на реальном объекте). Следовательно, по самой своей природе моделирование имеет много статистических аспектов.

В любом статистическом эксперименте необходимы тщательное планирование и анализ. При *статистическом анализе* эксперимента требуется: (1) извлечь всю имеющуюся в эксперименте информацию и (2) выявить ограничения в выводах, которые основаны на статистических данных. Однако из эксперимента нельзя получить больше информации, чем ее в ней содержится. Задача *статистического планирования эксперимента* — получить как можно больше полезной информации при данной стоимости эксперимента. Хотя моделирование — это эксперимент над выборкой, методы моделирования используются обычно не специалистами по статистике, а инженерами, социологами и т. д., поскольку моделирование требует еще построения модели. В связи с этим многие пользователи не обладают специальными знаниями математической статистики. Кроме того, мы думаем, что экспериментаторы стремятся использовать основное время на построение модели и по-

следующую разработку программы для вычислительной машины. Исключение из правила составляет моделирование работы пожарного депо, проделанное в [Carter and Ignall, 1970, p. 18]; авторы затратили на разработку модели от 0,5 до 1 человеко-года, а на планирование и анализ — еще больше времени. (Важность развития техники планирования и статистического анализа экспериментов для моделирования также подчеркивается Вагнером [Wagner, 1971, p. 1274]) Мы надеемся, что эта глава даст возможность читателю осознать статистические аспекты моделирования и снабдить его ссылками на литературу для дальнейшего изучения. К сожалению, исследование этих аспектов началось совсем недавно, поэтому некоторые заключения и рекомендации предварительны и требуют дальнейших разработок.

Уделим более пристальное внимание методам моделирования, применяемым *учеными-социологами*. До сих пор при моделировании они в общем не могли ставить *управляемые* эксперименты. Теперь же стало возможным использовать по крайней мере некоторые формы управляемых экспериментов. Однако при этом ученые сталкиваются с новыми для себя проблемами, связанными с планированием экспериментов и их анализом. В отличие от студентов технического профиля большинство студентов в области социальных наук знакомятся только с введением в методы планирования и анализа экспериментов, поскольку считается, что эти методы им не понадобятся. Поэтому социологи находятся в менее благоприятной ситуации (см. [Naylor et al., 1967b, p. 1315—1316] и [Naylor et al., 1968, p. 184—185]). (Последние из названных авторов сравнили различные типы данных, которые могут использоваться *экономистами*: данные по аналогии, экспериментальные данные с реального объекта и данные экспериментов с моделью.) Мы обсудим статистические аспекты моделирования, следуя различным *этапам* численного моделирования, как это описано Нейлором [Naylor et al., 1967a, p. 23—41]. Мы отсылаем к литературе², где описан процесс планирования, начинающийся с формулировки задачи, которую необходимо смоделировать, и кончающийся анализом результатов моделирования. Там же дано описание различных взаимосвязей в этом процессе. Мы обсудим *только* статистические аспекты различных этапов этого процесса, объединяя и перестраивая различные этапы из [Naylor et al., 1967a].

II.2. ФОРМУЛИРОВКА ЗАДАЧИ

В работе [Naylor et al., 1967a, p. 26] в качестве первого этапа планирования имитационного эксперимента приведена *формулировка задачи*. Этот этап содержит два статистических аспекта.

1. *Определение выхода*. Мы определяем ту выходную переменную или переменные, которые нас интересуют, т. е. говорят, что именно нужно измерить. В работе [Fishman and Kiviat, 1967, p. 24] выделены две цели моделирования: *сравнение* откликов для различных правил управления и определение *функциональных связей* между откликом и входными факторами. Нейлор и др. [Naylor et al., 1967a, p. 338] и [Naylor et al., 1967b, p. 1333] выделяют в качестве целей определение

отношения между комбинациями факторов и *общее* исследование зависимости между этими факторами. Очевидно, что цели, выделенные всеми этими авторами, не противоречат друг другу. Позже мы увидим, что статистические методы пригодны для определенных целей при изучении объекта с помощью моделирования (например, методы изучения поверхности отклика для определения оптимальной комбинации факторов). Оверхолт [Overholt, 1968, p. 64, 66] говорит, что при крупномасштабном моделировании выход модели может быть интересен на *нескольких* уровнях управления, каждый из которых интересен сам по себе. В этом случае трудно определить действительную цель моделирования.

2. *Статистическая надежность*. Нам надо узнать статистическую «правильность», с которой должно наблюдаться выход. При *проверке гипотез* мы можем потребовать, чтобы вероятность ошибочного отбрасывания правильной гипотезы была α , а вероятность ошибочного принятия (точнее не отбрасывания) ложной гипотезы была β . В задачах *оценивания* мы можем потребовать, чтобы выборочное среднее \bar{x} отличалось не более чем на δ единиц от математического ожидания μ с вероятностью $(1 - \alpha)$. (Напомним, что случайные величины, вроде выборочного среднего, набраны жирным шрифтом.)

Отметим: такая формулировка статистической надежности говорит, что мы используем *классическую* статистическую теорию, а не статистическую *теорию принятия решений*. В теории принятия решений не фиксируются в качестве статистических мер надежности ошибки α и β . Вместо этого строят функции потерь и часто применяют априорную информацию — так называемый байесовский подход (см., например, [Schlaifer, 1959] или резюме в [Wetherill, 1966, p. 85—110]). Дальше основное внимание мы уделим применению методов параметрической статистики.

11.3. АНАЛИЗ ВХОДНЫХ ДАННЫХ

Второй этап в планировании имитационного эксперимента — это «сбор и обработка данных с реального объекта» (см. [Naylor et al., 1967a, p. 27] и [Emshoff and Sisson, 1971, p. 55—56]). При анализе входных данных можно предположить, что время обслуживания подчинено экспоненциальному закону распределения. После того как мы выдвинули гипотезу об определенном виде закона распределения входной переменной (используя входные данные или теоретические соображения), можно применить *критерий согласия* распределения данных и гипотетического теоретического распределения. Критерии согласия обсуждаются в [Fishman and Kiviat, 1967, p. 20—21] и в [Naylor and Finger, 1967, p. 98]. Применение их при моделировании приведено в [Pegels, 1969, p. 223]. Кроме определения *вида* функциональных связей мы еще оцениваем значения *параметров* этих зависимостей. Подробные примеры содержатся в [Schmidt and Taylor, 1970, p. 481—498].

В некоторых задачах моделирования (например, вычисление критического пути, распределение статей движения капитала в *бюджете*) во входные данные вводят *субъективные* вероятностные оценки

(получаемые от управляющих и т. д.). В [Kotler, 1970] обсуждаются методы получения таких субъективных оценок. Отметим, что Чен [Chan, 1971] использовал энтропийный критерий для вывода закона распределения в случае отсутствия каких-либо знаний о входе, кроме диапазона его значений.

Может случиться, что с точки зрения критерия согласия приемлемы (или, лучше сказать, не отвергаются) сразу несколько видов распределений, особенно если входных данных мало. Тогда можно с помощью анализа чувствительности определить, чувствителен ли выход к вариациям вида распределения входных данных. Если выход чувствителен, то важно получить больше информации о входе. Полезные ссылки — это [Chambers and Mullick, 1968, p. 105—107], (см. также [Clough et al., 1965, p. 127—128], [Marks, 1964, p. 255—256]). Добавим, что иногда даже лучше, когда нет дополнительной информации. В этом случае может подойти метод из [Cohen, 1960, p. 36]. Коэн нашел значения некоторых параметров путем прогонки модели для разных значений параметров и выбрал те из них, которые минимизируют сумму квадратов отклонений между реальным и моделируемым выходом (см. также [Emshoff and Sisson, 1971, p. 203, 216, 278]).

Нейлор [Naylor et al., 1967a, p. 34, 36] замечает, что кроме оценивания значений параметров надо еще проверять их *значимость*. Более того, при использовании регрессионной модели надо проверять основные предпосылки, чтобы оценить, не нарушаются ли они наличием неоднородности дисперсии, автокорреляцией и т. д. Отметим, что такая проверка фактически относится к любому эконометрическим регрессионным моделям. Полное описание методов оценивания и проверки для таких моделей можно найти в [Johnston, 1963]. Однако, как было упомянуто в 1.6.2, типичные имитационные модели основываются не на регрессионных уравнениях, а на блок-схемах, описывающих соотношения между различными компонентами системы. А сами эти соотношения основаны на априорных знаниях. В регрессионных моделях все переменные связаны линейным уравнением (или системой уравнений), и мы проверяем, действительно ли нужно включать в это уравнение те или иные переменные, т. е. мы проверяем, значимо ли оцениваемый параметр отличается от нуля. Однако рассмотрим модель, предназначенную для моделирования в простой задаче массового обслуживания. Параметр, представляющий для нас интерес, — это параметр экспоненциального распределения, которому (как было предположено) подчиняется время обслуживания. Он должен быть положительным, поэтому мы не проверяем, значимо ли его отличие от нуля.

Мы видели, как можно использовать реальные данные для уточнения модели. Вместо того, чтобы извлекать из старой информации только представления о теоретическом распределении, мы можем взять эти данные просто как вход модели³. Поэтому отметим следующие возможности:

1. Реальные данные используются только для того, чтобы сформулировать вид *теоретического* распределения. Выборочные значения, отвечающие этому распределению, получаются затем подходящим преобразованием случайных чисел, как мы видели в 1.6.4.

2. Реальные данные дают *эмпирическое распределение*, из которого мы отбираем входные данные для моделирующего эксперимента. Отбор ведем снова с помощью случайных чисел.

3. *Старые данные* используются в моделировании в том же *порядке*, как они возникли в реальном процессе. Поэтому вместо внутреннего генератора входных данных, дающего случайные числа, пользуются уже готовыми входными данными для ввода в вычислительную машину.

Сравним эти три возможности. Использование старых данных в соответствии с (3) рекомендуется для «*обоснования*» модели. При обосновании модели с применением старых данных мы имеем эти данные, обработанные с помощью моделирующей программы, и полученный в результате моделирования выход. Далее сравниваем этот смоделированный выход с выходом, полученным для старых данных, и определяем, реалистична ли модель. Мы вернемся к обоснованию модели в II.5. Пусть мы имеем обоснованную модель и будем ее использовать для предсказания отклика при определенных вариантах системы. Тогда нас больше не интересует «*воспроизведение прошлого*». Старые данные в (2) и (3) — это только *выборка* из всей совокупности данных. Предполагается, что всю совокупность адекватно представляет теоретическое распределение. Таким образом, после обоснования модели мы имеем выборку из генеральной совокупности, представленной теоретическим распределением, как в случае (1).

II.4. МОДЕЛЬ И ПРОГРАММА ДЛЯ ВЫЧИСЛИТЕЛЬНОЙ МАШИНЫ

Третий этап в планировании имитационного эксперимента — *формирование модели*. Элементы модели уже были построены на втором этапе (II.3), где выбирались распределения входных переменных. Очевидно, у нас еще есть некоторые степени свободы для окончательного уточнения модели. Итак, реальные системы можно представить *различными* моделями. Если несколько моделей отвечают проверке на соответствие объекту, то следует выбрать *самую простую* из них. Простую модель можно выразить в рекурсивной форме, а не в виде совместных уравнений. Нейлор [Naylor et al., 1967a, p. 33] отмечает, что рекурсивная форма модели ведет к уменьшению времени счета и упрощает статистическую оценку параметров.

Теперь, после формирования модели, можно перейти к *программе для вычислительной машины*. Для моделирования типична работа с моделью во времени. Следовательно, нужно определить *начальные условия* моделируемой системы. Выделим две ситуации: моделирование устойчивого состояния и моделирование переходного состояния.

1. *Моделирование устойчивого состояния*. Часто при моделировании исследуют работу системы в *устойчивом состоянии*. Мы будем считать, что система находится в устойчивом состоянии, если вероятность ее пребывания в одном из таких состояний определяется *постоянной* функцией вероятностей. В противном случае система находится в переходном состоянии. (Отметим, что в устойчивом, стационарном,

или равновесном, «состоянии», или «условиях», система тоже может переходить из одного состояния в другое; например, длина очереди может измениться от состояния с тремя требованиями к состоянию с четырьмя требованиями.) Эти неизменные вероятности суть предельные вероятности, которые имеют место при реализации в течение длительного периода времени, и они не должны *зависеть* от *начального* состояния системы. Так, характеристики устойчивого состояния определяются независимо от начальных условий имитационного эксперимента⁴.

Мы могли бы взять любые начальные условия, поскольку при любых начальных условиях устойчивое состояние будет тем же самым. Однако все же уместно выбрать «хорошие» начальные условия. В [Nelson, 1966, p. 14 — 15] отмечено, что устойчивое состояние — это предельный случай, который в реальном имитационном эксперименте никогда точно не реализуется, поэтому реальные начальные условия *смещают* результаты. Тем не менее можно надеяться, что при достаточной длительности моделирования это смещение становится несущественным. Если мы интересуемся только характеристиками устойчивого состояния, то обычно не рассматриваем наблюдения, сделанные в *переходном состоянии*, т. е. когда процесс находится между началом моделирования и устойчивым состоянием. (Однако позже мы обсудим процедуру Крейна — Иглхарта (Crane — Iglehart), которая использует все наблюдения.) Это приводит к необходимости определения начала и конца переходного процесса. Если мы возьмем большую длительность переходного процесса, чем она есть на самом деле, то мы потеряем информацию об устойчивом состоянии, что увеличит дисперсии оценок. (Время работы машины при этом теряется.) С другой стороны, если мы возьмем меньшую длительность переходного процесса, чем на самом деле, то наблюдения над переходным процессом будут восприняты как наблюдения над устойчивым состоянием. При этом окажется смещенной оценка отклика в устойчивом состоянии. Мы не знаем статистического метода определения окончания переходного процесса [Conway et al., 1959, p. 108]. *Эмпирические правила* есть в [Conway, 1963, p. 49], [Emshoff and Sisson, 1971, p. 192] и [Tocher, 1963, p. 1746]. Следуя работе [Conway, 1963], можно не рассматривать начальные наблюдения до тех пор, пока не появится увеличение устойчивости (или ее уменьшение, в зависимости от задачи)⁵. В ней нет совета использовать накопленные наблюдения, поскольку накопленные данные имеют задержки, а это приводит к переоцениванию переходного процесса. Тем не менее на практике часто используют накопленные наблюдения [Gürtler, 1969, p. 94 — 96]. В [van Daal and van Doeland, 1972, p. 22] выбирается длительность переходного процесса и проверяется, соответствуют ли результаты моделирования знаниям, аналитическому решению. (Затем авторы изменили модель, чтобы сделать более реальным представление систем. См. также об обоснованности в II.5.) Лучше, если общее время моделирования велико и можно предположить, что смещение, вызванное наблюдениями переходного процесса, пренебрежимо мало. Недавно Бломквист [Blomqvist, 1970] получил очень интересный результат. Он рассмотрел одноканальную

систему массового обслуживания с независимыми поступлениями требований и независимым временем обслуживания, с обслуживанием в порядке поступления и «пустым» начальным состоянием (т. е. когда в системе нет ни одного требования). Он показал, что средняя квадратичная ошибка оцениваемого по методу наименьших квадратов среднего времени ожидания в устойчивом состоянии будет минимальной, если усреднить *все* наблюдения (и для переходного и для устойчивого состояний). Этот результат основан на двух допущениях: а) число прогонов велико; б) коэффициент вариации σ/μ — меньше единицы. (Условие б), как было доказано, соблюдается для многих простых систем массового обслуживания.) Различные методы, исключая начальные наблюдения, разработаны в [Fishman, 1971, p. 30]; совсем другой подход развит, как мы вскоре увидим, в [Crane and Iglehart, 1972 b].

Если исключить наблюдения за переходным процессом, то время, необходимое для перехода через это состояние, станет излишним. Можно попытаться минимизировать это *холостое время счета*, выбирая адекватные начальные условия; в [Meier et al., 1969, p. 297] есть пример. Все-таки «пустое» состояние часто выбирают в качестве начального условия. Тогда процесс продолжается в переходном состоянии. Правда, с другой стороны, это упрощает программу моделирования. Начальные условия выбирались в [Copway, 1963, p. 51], а в [Nelson, 1966, p. 15] рекомендуется брать одни и те же начальные условия для каждого моделируемого варианта системы, чтобы получить хорошие сравнения; в [Meier et al., 1969, p. 307 — 308] обсуждены отклонения от этого общего правила. Про начальные условия можно также прочесть в [Fetter and Thomson, 1965, p. 692], [Hillier and Lieberman, 1968, p. 462 — 463], [Huisman, 1969, p. 9].

В [Crane and Iglehart, 1972 b] предложена процедура, в которой не надо знать длину переходного периода. В качестве примера рассмотрим одноканальную систему массового обслуживания. Начало моделирования — «пустое» состояние и среднее для всех наблюдений, затем идет переходный процесс и устойчивое состояние до тех пор, пока система не *возвратится* в свое «пустое» состояние. Далее снова проводят очередные наблюдения, независимые от предыстории. Таким образом формируются блоки наблюдений, при этом размер (случайного) блока зависит от возвращения в «пустое» состояние. (Точное определение размера блока дано в VA.) Важное достоинство этого подхода — независимость блоков, поэтому анализ упрощается (см. также II.9). В то же время этот подход позволяет накапливать наблюдения с самого начала моделирования!

2. *Моделирование переходного состояния.* В [Nelson, 1966, p. 14] сказано, что до сих пор большинство исследований посвящено анализу устойчивого состояния, в то время как в действительности большинство задач относится к *нестационарным*. Многие системы вообще никогда не достигают устойчивого состояния, поскольку они прекращают работать в конце заданного периода времени либо подвержены разного рода возмущениям; примерами служат: бюро обслуживания,

которое заканчивает работу в конце каждого дня, или экономика страны соответственно.

Большое количество примеров переходных процессов можно найти в литературе⁶. Особенно рекомендуем [Emshoff and Sisson, 1971, p. 189 — 202].

Влияние начальных условий на отклик системы в переходном состоянии совершенно отличается от их влияния в устойчивом состоянии. Мы можем взять «старые» начальные условия подобно «пустому» состоянию для системы, которая возобновляет работу, или взять условия последнего года, для которого данные собраны при воздействии на систему возмущений [Orgutt et al., 1961, p. 131]. Начальные условия и сами могут быть предметом изучения. Тогда мы исследуем, как система реагирует на различные начальные условия. Как и Нельсон [Nelson, 1966], мы ожидаем увеличения числа задач по моделированию переходного процесса. Мы вернемся к рассмотрению начальных условий в II.8 и II.9, где рассмотрим влияние начальных условий на определение объема выборки и на анализ выхода.

Укажем еще специфические аспекты моделирования. Фактически экспериментирование при моделировании — это прогон программы для вычислительной машины. Отличие моделирования от реального эксперимента состоит в том, что при моделировании мы можем *манипулировать* выборкой, не вводя смещения в отклик, которым мы интересуемся. Эти манипуляции прodelьваются с помощью *методов уменьшения дисперсии*, также называемых методами Монте-Карло. Эти методы берут начало вне моделирования (например, от выборочных обследований и оценивания интегралов). В главе III обсуждаются шесть методов, которые можно использовать при сложном моделировании. Это: 1) *расслоенная выборка*: отклики в (повторных) моделируемых опытах имеют различные веса. Эти веса известны и зависят от слоя, к которому относятся случайные числа; 2) *селективная выборка*: входные переменные отбираются так, что их частоты согласуются с теоретическими, ожидаемыми частотами; 3) *контрольные величины*, или *регрессионный выбор*: средние значения входных переменных сравниваются с их математическими ожиданиями и отклик корректируется на отклонение между эмпирическими данными и математическими ожиданиями; 4) *значимая выборка*: исходные распределения входных переменных заменяются новыми (так, что тем входным величинам, которые приводят к наиболее важным значениям отклика, присваиваются большие вероятности), а отклик в связи с искажениями корректируют; 5) *дополняющие величины*: повторная реализация генерируется с помощью дополняющих случайных чисел $(1 - r)$. В результате отрицательная корреляция между двумя опытами уменьшает дисперсию отклика, усредненную по опытам; 6) *общие случайные числа*: два или более разных варианта системы моделируются с использованием одних и тех же случайных чисел. Для этого метода уменьшения дисперсии можно достигнуть за счет увеличения *объема выборки* или эффективного выбора *комбинаций* уровней факторов в эксперименте. Два последних метода будут обсуждены в II.8 и II.6 соответственно.

II.5. ОБОСНОВАННОСТЬ МОДЕЛИ

После указанных выше этапов можно перейти к обоснованию модели. Как мы объяснили в II.3, можно взять старые данные, обработать с помощью программы для вычислительной машины, *смоделировать выход* и сравнить его со «старым» выходом. Для обоснования необходимо, чтобы реально существовал хотя бы один вариант моделируемой системы и были доступны данные. Отметим, что мы можем моделировать существующую систему не ради смоделированного выхода, а только для обоснования модели. Как только модель обоснована, мы изменяем ее так, чтобы можно было надеяться, что она реалистично представляет интересующую нас систему (см., однако, замечания в [Schmidt and Taylor, 1970, p. 499 — 501]), обсуждение в более методологическом плане дано в [Naylor et al., 1967a, p. 310 — 320] и в [Naylor and Finger, 1967]; (см. также [Gregg and Simon, 1967]). В [Naylor and Finger, 1967] перечислено несколько возможных мер и методов, которые можно использовать для обоснования модели. В качестве мер они упомянули средние моделируемой траектории системы, саму траекторию, амплитуду, локальные экстремумы и т. д. Как в [Naylor et al., 1967a, p. 318], так и в [Naylor and Finger, 1967] приведено множество методов для проверки того, насколько хорошо согласуется моделируемый выход со старым выходом. Они упомянули хорошо известный критерий *хи-квадрат* и критерий Колмогорова — Смирнова для сравнения реальных частот с теоретическими частотами. В [Fishman and Kiviat, 1967, p. 20] обсужден критерий Кокрена как альтернатива по отношению к критерию хи-квадрат. Для моделирования выхода можно применить *факторный* и *спектральный анализ*, чтобы узнать, являются ли факторные нагрузки и спектр для модельного выхода теми же, что и для реального выхода. (Мы вернемся к краткому описанию спектрального анализа в II.9.) Можно использовать простой *регрессионный анализ* для описания связи реального и моделируемого выхода и для проверки того, равен ли свободный член нулю и равен ли тангенс угла наклона единице. В [Aigner, 1972], однако, эта процедура подвержена сомнению. В [Wolff, 1967, p. 338 — 342] показано, как можно вычислить «коэффициент неравенства» Тейла или его варианты⁷. Критерий Тьюринга применяется при моделировании решения задач человеком⁸. В [Migram, 1972a, p. 250 — 254, 307 — 310; 1972 b] обсуждены эти и некоторые другие меры и методы, например проверка среднего значения моделируемого выхода по отношению к реальному выходу с помощью *параметрических* и *непараметрических* критериев. В [Howrey, 1972, p. 24 — 33] упоминаются задачи с *функциями потерь* для оценивания модели. Там перечислены и некоторые другие меры и достаточно полно обсуждено измерение локальных экстремумов (часто употребляемое в эконометрическом моделировании), а также описан метод спектрального анализа. В [Hanna, 1971] введена *теоретико-информационная* мера оценивания моделей для имитации. Здесь сконцентрировано внимание на моделировании искусственного интеллекта (см. также [Starbuck, 1971]). Должно быть ясно то, что многие аспекты обоснования модели вовсе не

специфичны для моделей, используемых в имитационном моделировании. Они также характерны и для других моделей. Специфично для имитационного моделирования то, что оно дает полную временную траекторию, состоящую из *автокоррелированных* данных.

Как мы видели, критерии согласия, подобные критерию «хи-квадрат», могут также употребляться для сравнения реальных и теоретических *входных данных*. В [Naylor and Finger, 1967, p. 100] отмечено, что последняя проверка полезна, ибо иначе можно сконструировать нереалистичную модель (если ее входные распределения нереалистичны). Не делая проверки входных распределений, мы смогли бы обнаружить нереалистичность модели на стадии ее обоснования. Таким образом, мы бы затратили основное время на обоснование, применение и проверку этой нереалистичской модели! В [Fishman and Kiviat, 1967, p. 3, 11—19] обсуждены методы «верификации» (проверки), относящиеся к обоснованию и проверке входных данных. Описаны методы проверки того, ведет ли себя модель и особенно ее *части* так, как мы *предполагали* (в то время как обоснование модели означает, что мы проверяем, дает ли модель такой же выход, что и реальная система). Мы можем проверить, дает ли генератор независимые случайные числа, как мы предполагали. Мы можем также проверить полную модель (вместо ее частей) путем рассмотрения упрощенной, возможно детерминированной, задачи, для которой существует *аналитическое* решение. Для этой упрощенной задачи можно также провести моделирование и сравнить полученный выход с аналитическим решением (см. [Emshoff and Sisson, 1971, p. 170], [Mihram, 1972a, p. 245—250], [Schmidt and Taylor, 1970, p. 502—507]). (Эта упрощенная модель затем может быть использована для предварительного анализа чувствительности.) Проверка или «отладка» программы для вычислительной машины возможна, если сделать некоторые вычисления вручную.

К верификации также относятся *проверка чувствительности* и эволюционное (непрерывное) построение модели. Если модель чувствительна к точности задания входа, можно собрать дополнительную информацию для конструирования подходящей модели. Если обнаружилось, что модель несовершенна, можно начать следующий цикл моделирования процесса. Поскольку блок построения подходит к большинству этапов моделирования, имеет смысл обосновать *подмодели* для компонент системы. Несколько хороших обзоров этих аспектов можно найти в литературе⁹.

Очевидно, что модель или часть модели может быть *ошибочно* отвергнута при обосновании или на стадии проверки, если *объем выборки* невелик. Поэтому возможна большая ошибка выборки. В связи с этим имеют место так называемые ошибка 1-го рода — α и ошибка второго рода — β , т. е. вероятность того, что ложная модель или соответственно часть модели будет принята. В [Orcutt et al., 1961, p. 32] детально обсуждены отклонения между реальным и моделируемым выходом, возникшие из-за ошибки определения модели и ошибки выборки. Последние типы ошибок относятся к ошибкам α и β (см. также [Stash, 1968]). Дополнительные ссылки по обоснованию моделей

имеются в [Howtaу, 1972], где дан интересный обзор обоснования линейных эконометрических моделей (см. также [Dutton and Starbuck, 1971a, p. 706, «априорное и апостериорное использование данных») и т. д.¹⁰ Примеры с подробным рассмотрением обоснования модели есть в [Gürtler, 1969, p. 106—109], [Pegels, 1969, p. 228] и в [Van Horn, 1971, p. 253—256]. В последних двух случаях формальные статистические критерии для сравнения выхода модели и «старого» выхода не использовались.

11.6. ПЛАН ЭКСПЕРИМЕНТА

Если модель для имитации на вычислительной машине не отвергается на стадии обоснования, мы можем приступить к эксплуатации программы на ЭВМ, чтобы осуществить моделирование различных систем, представляющих интерес. Эти системы (или варианты систем) отличаются друг от друга, поскольку у них отличаются значения или типы системных параметров, входные переменные и соотношения, определяющие поведение систем (также называемые операционными характеристиками). Эти различные параметры, переменные и соотношения в терминологии статистического планирования эксперимента называются *факторами*. Для того чтобы определить эффект фактора, этот фактор нужно варьировать, или, по терминологии планирования эксперимента, рассматривать на нескольких уровнях. Очевидно, что *количественные* факторы могут изменяться на многих уровнях. Примером количественного фактора служит параметр λ экспоненциального распределения в модели массового обслуживания.

Качественный фактор может изменяться только на ограниченном числе уровней. «Правило приоритета» в задаче массового обслуживания является примером такого фактора, который мы можем изменять на двух уровнях, говоря «обслуживание в порядке поступления» и «случайный порядок». Отметим, что числа для уровней качественного фактора — это только удобный способ краткого обозначения, не используемый для построения математической функции, связывающей уровни факторов и отклик.

Если мы хотим исследовать k факторов и i -й ($i = 1, \dots, k$) фактор имеет L_i уровней, то *число комбинаций* уровней факторов будет $L_1 \times L_2 \times \dots \times L_k$. Проблема состоит в том, что даже в случае немногих факторов число полученных комбинаций велико. Так, например, если имеется только семь факторов и каждый фактор имеет минимальное число уровней (т. е. $L_i = 2$), то число комбинаций будет $2^7 = 128$! Попытаемся ограничить число комбинаций, которые будут реально исследоваться. (Это исследование означает, что отобранная комбинация факторов определяет один вариант системы, которая моделируется во времени, в результате чего и образуется временная траектория для такого частного варианта системы.) Это — область *планирования эксперимента*. Мы находим удобным выделить следующие разделы планирования экспериментов.

1. *Определение важных факторов, или «отсеивающий эксперимент».* Как видно из изложенного выше, число комбинаций факторов быстро

растет с ростом числа факторов в эксперименте. Следовательно, нужно сконцентрировать внимание на наиболее важных факторах. Необходимо *предварительное* отсеивание наиболее важных факторов. Особенно подходящими для исследования большого числа факторов планами при относительно малом числе наблюдений являются следующие.

Планы типа 2^{k-p} . Все k факторов имеют два уровня, и только часть всех комбинаций (эта часть 2^{-p}) используется. В зависимости от выбранной части (или реплики) сохраняется возможность оценить главные эффекты факторов и взаимодействия низкого порядка. Если k велико, число комбинаций все-таки остается большим (т. е. 2^{k-p}), и тогда более подходящими будут планы отсеивающих экспериментов.

Случайные планы. Комбинации уровней факторов случайно отбираются среди всех возможных комбинаций. Число комбинаций N может быть определено независимо от числа факторов и уровней. Так, N можно выбрать даже меньше k . Недостаток состоит в том, что случайный отбор может не привести к комбинациям, которые позволяют получить «хорошие» оценки индивидуальных эффектов. (Выражаясь техническим языком: степень ортогональности столбцов матрицы независимых переменных не управляема, а случайна.) Затем были разработаны следующие типы планов.

Сверхнасыщенные планы. Число комбинаций (N) меньше, чем число факторов (k), и комбинации отбираются так, чтобы (для данных N и k) оценки эффектов факторов были достаточно «хорошими». (Минимизируется максимальная неортогональность столбцов матриц плана.) Если имеется очень много факторов, то более подходящими будут следующие планы.

Планы группового отсеивания. k факторов разбиваются на g групп ($g \ll k$) и эти g групп факторов испытываются в плане типа 2^{k-p} или в сверхнасыщенном плане. Та группа факторов, которая оказалась наиболее важной, разбивается на несколько групп меньшей размерности (в конечном счете до групп размерности 1).

В главе IV эти четыре типа планов детально обсуждаются и сравниваются между собой.

2. *Дальнейшее исследование наиболее важных факторов.* Пусть мы уже определили, какие факторы, какие уровни и какие комбинации уровней будут изучаться в эксперименте. Выбор факторов может быть основан на упомянутой стадии отсеивания, если подозреваемых важных факторов много. Число уровней стоит ограничить, чтобы число комбинаций оказалось небольшим. Часто берут два или три уровня. Если в результате число комбинаций оказывается небольшим, то используют *полный факторный* эксперимент, т. е. все возможные комбинации уровней. Вместо этого можно применить неполные планы. Рассмотрим следующие два простых типа таких планов.

Планы типа 2^{k-p} . Эти планы уже обсуждались как отсеивающие. Они могут быть также использованы для более детального изучения факторов, поскольку меньшая дробность эксперимента позволяет дальше изучить влияние k факторов. Так, при большой дробности эксперимента можно найти только главные эффекты (смещенные, если есть определенные взаимодействия). При малой дробности эксперимен-

та можно оценить главные эффекты и взаимодействия не более двух различных факторов (исключенные в случае отсутствия взаимодействия трех и более факторов).

Планы для анализа поверхности отклика. Это комбинации планов типа $2^k - p$ и дополнительных комбинаций, позволяющих оценить отклик как функцию (полином первого или второго порядка) независимых переменных. (Все k факторов должны быть количественными.)

Оба типа планов можно просто применить *последовательно*, т. е. сначала получить наблюдения для нескольких комбинаций уровней факторов, затем произвести анализ этих наблюдений и только после этого анализа решать, для какой комбинации (старой или новой) произвести дополнительные наблюдения. Эти новые наблюдения снова анализируются (обычно вместе со старыми), прежде чем решить, какие наблюдения еще потребуются, и т. д. В планах типа $2^k - p$ можно сначала использовать малую часть эксперимента, проанализировать наблюдения и если этот анализ показывает, что данная часть эксперимента слишком мала для оценки всех возможных эффектов, то расширить эксперимент так, чтобы он позволил оценить все эти эффекты. Планами для анализа поверхности отклика пользуются в *методологии анализа поверхности отклика (RSM)*, которая состоит в отыскании *оптимальной* комбинации уровней k количественных факторов. Она включает следующие этапы:

а) *Планирование эксперимента*, т. е. выбор комбинаций уровней факторов. Эта методика использует полные и дробные типа $2^k - p$ планы эксперимента и некоторые специальные планы.

б) *Подбор уравнения регрессии* по наблюдениям. Это регрессионное уравнение* называется поверхностью отклика; оно выражает отклик как функцию независимых переменных.

в) «*Влезание*» по поверхности отклика к вершине. Для нахождения направления увеличения отклика обычно служит метод *крутого восхождения*. В направлении, где ожидается увеличение отклика, этапы а, б и в повторяются до тех пор, пока не будет достигнута область максимума.

г) *Канонический анализ* в области максимума функции отклика. Этот анализ выявляет, имеется ли один максимум, несколько максимумов, седловая точка или гребень.

В главе IV дано дальнейшее описание планов типа $2^k - p$, их упорядочение и анализ. Библиография методологии поверхности отклика приведена в главе IV и в обзорах [Hill and Hunter, 1966] и [Herzberg and Cox, 1969]. Подчеркнем, что планирование эксперимента, описанное выше, — только часть планов экспериментов для имитационного моделирования на ЭВМ, которые включают в себя еще определение начальных условий, методы уменьшения дисперсии, определение объема выборки и т. д.

* Вернее, соответствующая ему поверхность. — *Прим. перев.*

II.7. АНАЛИЗ НЕСКОЛЬКИХ ВАРИАНТОВ СИСТЕМЫ

Анализ и планирование эксперимента можно было бы выделить отдельно, но обычно этого не делается. Как видно из II.6, мы предполагаем *последовательное* планирование и анализ имитационных экспериментов. *Чисто* последовательный подход дал бы возможность анализа после *каждого* следующего наблюдения. На практике, однако, анализ выполняется обычно после получения нескольких наблюдений. Это называется *шаговым* подходом. Последовательные планы более важны. Но большинство экспериментальных планов возникло при решении задач в области *сельского хозяйства*, где последовательное экспериментирование трудно осуществить, так как приходится ждать целый год, пока станут доступными следующие планируемые наблюдения. В последние годы экспериментальные планы также используются в промышленности, где период времени, необходимый для получения наблюдений, короче. Последовательные планы развиты и применяются в значительной степени в *химической промышленности*. *Эксперименты имитационного моделирования на ЭВМ* по своей природе очень отвечают идее последовательного экспериментирования, поскольку ЭВМ работает последовательно! Существует много последовательных планов. Ссылки даны в [Naylor et al., 1967a, p.338], [Naylor et al., 1967b, p.1329—1330] и [Wetherill, 1966].

После выявления связей между планированием и анализом мы в конце главы остановимся на анализе экспериментов. Обсудим некоторые методы, которые можно применить при сравнении *нескольких* вариантов систем. Пусть мы хотим изучить k вариантов систем, или, короче, k систем. Каждой системе соответствует одна частная комбинация уровней факторов, варьируемых в эксперименте. Каждая комбинация определяет одну совокупность, из которой производится выборка. Мы обсудим такие ситуации, при которых число наблюдений, входящих на совокупность, не задается, а определяется во время эксперимента, и такие ситуации, в которых число наблюдений задано.

1. Методы *множественного ранжирования*. Эти методы призваны определить *число наблюдений*, которое надо взять от каждой из k (≥ 2) совокупностей, чтобы выбрать *наилучшую* совокупность. Наилучшая — это обычно такая совокупность, которая имеет наибольшее или (в зависимости от задачи) наименьшее среднее. Некоторые методы развиты и для других критериев отбора, отличных от среднего, например для дисперсии или для других формулировок задачи, иных, чем задача выбора наилучшей совокупности. В частности, для задачи ранжирования всей совокупности. Методы полного или неполного ранжирования называются методами множественного ранжирования. В литературе они также называются методами множественной селекции или принятия решений. Большинство из них основано на подходе, использующем «зону безразличия», т. е. для выделения наилучшей совокупности требуются большие выборки, если средние совокупностей мало отличаются друг от друга. В то же время потери от неправильного выбора были бы невелики. Следовательно, процедура гарантирует корректное выделение с вероятностью не меньше P^* (или $1 - \alpha$), *только* если среднее наи-

лучшей совокупности не меньше, чем на δ^* , отличается от следующего за ним среднего (P^* и δ^* уточняются в эксперименте). Большинство процедур ранжирования последовательно. Раздел VC содержит обзор существующих методов, обоснование их эффективности и робастности (устойчивости), а также некоторые эвристические процедуры. (Робастность процедуры — это ее чувствительность к отклонению от допущений, таких, как нормальность распределения, независимость и т. д.)

2. *Методология анализа поверхности отклика.* Если все факторы количественные, то этот подход можно использовать для отыскания наилучшей системы (см. II.6).

3. *Методы множественных сравнений.* Число наблюдений над совокупностью может быть задано потому, что мы осуществляем лабораторный эксперимент, или потому, что машинное время ограничено. Мы можем произвести различные типы *сравнений* между средними совокупностей, — например, сравнение со «стандартным» средним, которое соответствует имеющейся системе (или $\mu_i - \mu_0$, $i = 1, \dots, k - 1$); все попарные сравнения ($\mu_i - \mu_j$, $i \neq j$), сравнения отклика, усредненного по $(k - 1)$ экспериментальным системам, со стандартной системой (или, в более общем случае, линейные контрасты

$\sum_{i=1}^k c_i \mu_i$ при $\sum_{i=1}^k c_i = 0$), — изучить средние μ_i сами по себе, отобрать под-

группу, содержащую наилучшую совокупность, или, в других ситуациях, подгруппу, содержащую все совокупности, лучшие, чем стандартная совокупность. (Совокупности в такой подгруппе можно потом изучить с помощью дополнительных экспериментов.) В методах множественного сравнения вероятность того, что *все* выводы, основанные на единичном эксперименте, верны, равна $(1 - \alpha)$. Например, все $k(k - 1)/2$ доверительных интервалов для парных сравнений корректны с вероятностью не меньше $1 - \alpha$ (α суть так называемая доля ошибки эксперимента). Отметим, что если все факторы количественные, то регрессионный анализ более эффективен, чем методы множественных сравнений. Мы отсылаем читателя к разделу VB, где обсуждаются методы множественных сравнений для различных ситуаций, их эффективность, робастность и типы ошибок.

II.8. ОБЪЕМ ВЫБОРКИ

В II.7 мы уже обсудили определение объема выборки, т. е. числа наблюдений над совокупностью. В этом и следующем параграфах мы исследуем некоторые аспекты, связанные с объемом выборки, которые *специфичны для моделирования*. Но сначала дадим несколько формулировок. Объем выборки влияет на «*статистическую надежность*» оцениваемого отклика моделируемой системы. Эту надежность можно измерять с помощью стандартного отклонения среднего отклика для вариантов систем, которые моделируются. Надежность можно увеличить, взяв большую выборку. Предположим, нас интересует среднее значение отклика μ , и мы имеем n независимых наблюдений над этим средним x_1, x_2, \dots, x_n . Хорошо известно, что стандартное отклонение \bar{x} —

среднего n независимых наблюдений — равно σ/\sqrt{n} , где σ — стандартное отклонение единичных наблюдений x_i ($i = 1, \dots, n$). Значит, чтобы уменьшить наполовину стандартное отклонение этого среднего, надо взять выборку, в четыре раза большую, чем исходная. Это показывает, что нужно большое увеличение объема выборки для уменьшения влияния случайных отклонений (как было упомянуто в II.4, можно пользоваться методами уменьшения дисперсии, чтобы не слишком увеличивать объем выборки). Упомянутые соображения нужно скорректировать при моделировании в случае *зависимых* наблюдений, как будет показано позже.

Существует несколько возможностей для увеличения объема выборки в моделирующем эксперименте. Эти возможности изменяются в зависимости от следующих двух дихотомий, введенных в [Gafarian and Anker, 1966].

1. *Такты по времени или такты по событиям в программе.* Такты по времени означают, что мы разделяем моделируемый период времени на малые интервалы. В конце каждого такого интервала мы определяем, изменилась ли и как изменилась система. Если изменения системы непрерывны, то можно оперировать только техникой программирования. Многие системы, однако, изменяют свое состояние лишь в определенных моменты времени. Тогда мы можем использовать либо временные интервалы, либо *такты по событиям*. В последнем случае осуществляется перескок от одного момента времени к такому следующему моменту, с которого система изменяет свое состояние. Например, в простой задаче массового обслуживания следующий момент времени, когда что-то происходит в моделируемой системе, — это момент, когда либо в систему поступает требование, либо заканчивается обслуживание требования (см. также пример в [Naylor et al., 1967a, p.126—136]). Подробное обсуждение дано в [Emshoff and Sisson, 1971, p.159—169] и в [Mihram, 1972 a, p.228—231]. Отметим, что в [Nance, 1971] определены тонкие различия между этими двумя подходами к программированию и произведено сравнение эффективности результирующих (четырёх) процедур. В I.6.2 мы видели, что подход промышленной динамики Форрестера — это также моделирование с временными интервалами, но с упором на высоко агрегированные и детерминированные модели (см. [Meier et al., 1969, p. 80—117], [Niedereichholz 1971] и [Smith, 1970, p.769—770]).

2. *Системы с остановом или системы без останова.* В системах с *остановом* моделирование заканчивается, если произошло определенное событие. В [Gafarian and Anker, 1966, p.27] даны в качестве примеров модель дуэли и модель отказа оборудования, когда время моделирования заканчивается при убийстве одного или двух дуэлянтов или когда оборудование выходит из строя. В системах *без останова* нет критического события, которое останавливает процесс моделирования; ср. моделирование телефонной станции (коммутатора). Мы отсылаем читателя к [Gafarian and Anker, 1966] для дальнейшего обсуждения связей между фиксированным и переменным приращением времени при программировании систем с остановами и без остановов, дискретных

и непрерывных. Однако конец этого параграфа изложен иначе, чем названная статья. Заметим, что один «пробег» при моделировании — это одно движение системы от начала к концу моделируемого периода времени.

Существуют системы, которые физически не имеют остановов, но ведут себя так же, как системы с остановом, и поэтому связаны с объемом выборки. Мы имеем в виду три важных типа таких систем. Во-первых, это системы, которые регулярно заканчивают работу. Например, банк закрывается в конце дня, последние клиенты обслуживаются, и на следующий день банк открывается снова. Тогда один прогон — это один день. Во-вторых, рассмотрим системы, для которых в заданные моменты времени надо оценить отклик для варьирования стратегий (эти стратегии вносят в систему возмущения). Примером служит фирма, для которой мы хотим найти выгоду в ближайшие три месяца при различных вариантах капиталовложений. Таким «определенным событием», из-за которого заканчивается моделирование, является наступление конца *планируемого периода*. Третий тип охватывает системы без физических остановов, у которых мы хотим изучить *переходный* процесс (например, как функцию начальных условий). Как только система достигает устойчивого состояния, моделирование кончается. Поскольку мы классифицировали эти три типа систем, физически не имеющих остановов, как системы с остановами, то ясно, почему мы будем полагать, что в *системах без остановов* интересны только *устойчивые состояния* (см. также II.4).

Мы обсудим три пути увеличения объема выборки. Мейер [Meier et al., 1969, p.300] различает больше возможностей, но основные варианты входят в нашу классификацию: (1) уменьшение *тактов по времени*, (2) *повторение* прогонов при моделировании, (3) *продление* имитирующих экспериментов. Очевидно, что первая возможность реализуема только в случае программирования с тактами по времени. Уменьшение длины такта приводит к тому, что система наблюдается чаще за тот же период времени. В [Gafarian and Anker, 1966, p.41] сделан вывод, что в большинстве случаев этот путь наименее эффективен. Мы могли бы добавить, что кроме *неэффективности* манипулирования с тактами этот путь еще и *непрактичен*. В [Fishman and Kiviat, 1967, p.25—26] показано, что такты могут в той или иной степени определять вид работы системы (см. также [Meier et al., 1969, p.300]). Программирование, использующее такты по времени, — не лучший путь, если система легче программируется тактами по событиям или если при этом моделирование системы требуется меньше машинного времени. Пример такой ситуации — простая задача массового обслуживания. Пример противоположного типа — свободный обмен алмазами, который трудно запрограммировать с помощью тактов по событиям, поскольку трудно задать события. И наконец, программирование с тактами по времени невозможно, если мы хотим имитировать на *подходящем языке* моделирования, который основан на последовательности событий. Так что далее мы не будем смотреть на длину тактов по времени как на средство увеличения объема выборки.

Для систем с *остановами* повторение опытов — простейшее сред-

ство, поскольку длительностью опыта нельзя манипулировать. Если мы используем набор случайных чисел для каждой реализации, то полученные наблюдения будут *независимыми*, поэтому можно использовать традиционные статистические методы анализа модельного эксперимента. (Повторение опытов так же просто применять, как методы понижения дисперсии, см. главу III.) Отметим, что в исследовании Бруцелиуса [Brucelius, 1972a] не удалось выявить различий между системами с остановами и без остановов. Он моделировал банк в течение $83 \frac{1}{3}$ часа без прерывания. В разделе VA мы обсудим подходящий метод *анализа* для определения объема выборки и построения доверительного интервала заранее заданной длины или проверку с заранее заданными ошибками α и β , для одного среднего или для различий между двумя средними.

В системах *без* останова можно использовать как *повторные* реализации, так и продление реализаций. В [Gafarian and Anker, 1966, p.41] сделан вывод, что *повторение* реализаций более *эффективно*, так как обеспечивает положительную сериальную корреляцию между отдельными наблюдениями (это разумное допущение для многих случаев моделирования). Если мы *допустим*, что система была в устойчивом состоянии с самого начала моделирования, то можно легко показать, что повторение опытов приводит к уменьшению дисперсии среднего значения отклика (см. упражнения 1 и 2). На самом деле в литературе можно найти лишь один случай, когда система находилась в устойчивом состоянии с самого начала. Это замкнутая система управления с обратной связью, обсуждаемая в [Hauseg et al., 1966, p.79]. В общем случае моделируемая система проходит через *переходное* состояние прежде, чем достигает устойчивого состояния. Соответственно, если нас интересует только устойчивое состояние, мы обычно не используем наблюдения за переходным состоянием. (Однако, как было видно в II.4, метод Крейна—Иглекхарта использует все наблюдения.) Если мы повторяем опыты, то устраняем наблюдения за переходным процессом в начале каждого опыта. Этот момент также обсуждается в [Conway, 1963, p. 55], [Conway et al., 1959, p.109—110] и [Tocher, 1963, p.176—177]. К сожалению, в настоящее время нет однозначного критерия для выбора из этих способов: что лучше — повторить опыты или удлинить их, *уменьшить дисперсию* или иметь потери *машинного времени* в *переходном* состоянии (или смещение в переходном состоянии). Более того, в [Mihram, 1972 a, p. 448—450] отмечено, что оценивание моментов из одной удлиненной реализации может дать некорректные оценки соответствующих моментов совокупности в том случае, если моделируемый временной ряд не «эргодичен»¹¹. Существует еще фактор, подходящий для выбора между повторными реализациями и удлиненными реализациями, — это *анализ* выхода модели. Анализ повторных реализаций очень прост, поскольку каждая реализация дает одно независимое наблюдение. Удлиненная реализация дает наблюдения, которые автокоррелированы, а это значительно усложняет анализ. Эта задача анализа будет обсуждена в II.9. Прежде чем к нему перейти, мы резюмируем аргументы в пользу повторных реализаций (см. также [Mihram, 1972a, p.449]):

- а) многие системы имеют остановы и поэтому нельзя удлинять моделирующие реализации;
- б) независимые реализации допускают классические методы анализа;
- в) устраняется проблема эргодичности;
- г) повторение реализаций обеспечивает меньшую дисперсию без предварительных усилий (или при небольших усилиях).

II.9. АНАЛИЗ ВЫХОДА В УСТАНОВИВШЕМСЯ СОСТОЯНИИ И ДЛИНА РЕАЛИЗАЦИИ

Здесь мы обсудим анализ выхода систем *без остановов*, для которых мы хотим оценить *ожидаемый* отклик в *установившемся состоянии*. Мы допустим, что не используется метод Крейна — Иггхарта, и поэтому можем отбросить наблюдения за переходным состоянием, так что оставшаяся временная траектория содержит только наблюдения установившегося состояния¹². Остается задача анализа выхода этого состояния. Мы обсудим отдельно повторные реализации и удлиненные реализации.

1. *Повторные реализации*. Мы можем принять решение повторять реализации, несмотря на возможность потери машинного времени в переходном состоянии. Каждая реализация дает одно *независимо* оцениваемое среднее \bar{x} . Центральная предельная теорема для стационарных r -зависимых реализаций, сформулированная в [Mechanic and McKay, 1966, p.4] и в разделе VA, показывает, что среднее каждой реализации (приближенно) имеет нормальный закон распределения (при отсутствии периодичностей). Следовательно, можно использовать статистику Стьюдента для построения доверительного интервала в виде $(\bar{x} \pm t_{\alpha} s_{\bar{x}} / \sqrt{n})$, где \bar{x} — общее среднее всех средних \bar{x}_i и $s_{\bar{x}}$ — стандартное отклонение, вычисленное по \bar{x}_i ($i = 1, \dots, n$).

2. *Удлиненная реализация*. Существуют три подхода к анализу удлиненной реализации.

1-й подход. «Независимые» отрезки реализации. Разделим всю длину удлиненной реализации на n отрезков (после исключения переходного процесса на начальной стадии реализации). Обозначим средние этих отрезков по всей реализации $\bar{x}_1, \dots, \bar{x}_n$. Допустим, что сериальная корреляция уменьшается как «лаг», т. е. с увеличением расстояния между отдельными наблюдениями x . (Из этого допущения следует отсутствие периодичностей.) Соответственно только «несколько» первых наблюдений i -го отрезка будут коррелированы с последними наблюдениями $(i - 1)$ -го отрезка и, следовательно, средние \bar{x}_i и \bar{x}_{i-1} будут слабо коррелированы. Если длины отрезков реализаций достаточно велики, то корреляцией между их средними с практической точки зрения можно пренебречь. Доверительные интервалы, основанные на t -статистике, подобны доверительным интервалам, обсужденным в связи с повторными реализациями. В [Mechanic and McKay, 1966] разработан метод определения длины отрезков, обеспечивающей приблизительную неза-

висимость средних этих реализаций. Этот метод *итеративен*. Он оценивает корреляцию между средними частных реализаций и затем удлиняет эти реализации до тех пор, пока указанная корреляция не станет достаточно «малой». Малость коэффициента корреляции обоснована с помощью выведенного экспериментального критерия, данного в [Mechanic and McKay, 1966, p.20]. Они применили эти правила к нескольким задачам массового обслуживания и получили некоторые теоретические соотношения. Метод дает удовлетворительные результаты.

2-й подход. Оценивание сериальных корреляций. Этот подход не ограничивает отрезки реализации, а оценивает дисперсию \bar{x} — среднего всей реализации. В выражении для этой дисперсии есть коэффициент сериальной корреляции¹³. Поэтому нужна оценка корреляции между отдельными наблюдениями. Различные формулы для такого оценивания употреблялись в [Fishman, 1967] и в [Hauser et al., 1966].

3-й подход. Независимые блоки. В II.4 мы коротко описали подход, использованный в [Crane and Iglehart, 1972 b). Ограниченные независимые блоки (или туры) предложены в [Kabak, 1968]. Анализ этих блоков дан в [Crane and Iglehart 1972a, 1972b] и в [Fishman, 1972].

В разделе VA дается более детальное обоснование этих и других методов определения надежности оцениваемого среднего отклика в установившемся состоянии. Оцененную дисперсию можно применить для определения (последовательно) того, является ли (общая) реализация достаточно длинной, чтобы удовлетворить требуемой надежности оцениваемого при моделировании отклика, аналогично анализу систем с остановом.

Может оказаться так, что кроме среднего отклика и дисперсии мы захотим узнать динамические характеристики отклика в устойчивом состоянии (например, периодичности). В [Kiviat, 1967, p.7] и [Mihram, 1972a, p.261—262] отмечено, что при этом можно выделить: (1) *статические меры*, или меры среднего выхода, подобные средним, стандартным отклонениям и гистограммам; (2) *динамические меры*, подобные автокорреляции или спектру. Коэффициенты автокорреляции были получены в [Hoggatt and Holtbrugge, 1966] при моделировании рынка. В [Fishman and Kiviat, 1967, p.15—16] отмечено, что спектр является преобразованием Фурье для автокорреляции, поэтому он дает ту же самую информацию, но при этом с ним удобнее работать. Для более подробного знакомства со спектральным анализом и дополнительной литературой мы отсылаем читателя к [Anderson, 1971], [Fishman, 1967, 1968], [Fishman and Kiviat, 1967], [Mihram, 1972a, p.175—179, 467—483], [Naylor, 1971, p.35, 247—268, 312—317].

Поскольку *спектральный анализ* выделен в публикации [Fishman and Kiviat, 1967], — авторитетных в моделировании авторов, — видимо, полезно выявить *ограничения*, которые мы видим в этом методе¹⁴. Как следует из работы [Fishman, 1968, p.17], спектральный анализ применяется к наблюдениям *установившегося состояния*. Если моделирование выхода удовлетворяет этим условиям, то спектральный анализ в большинстве случаев будет полезным методом (когда мы хотим произвести более сложный анализ, использующий как статические, так и динамические свойства системы). Отметим, что временные ряды (возможно,

генерируемые имитирующим экспериментом) можно сделать стационарными с помощью подходящего преобразования (например, взяв приращения или скользящие средние), которое устраняет дрейф среднего или дисперсии (см. [Mihram, 1972a, p.445—459], а для более подробного изучения см. [Anderson, 1971]). Однако в II.4 мы отметили, что ожидаем увеличение числа работ по моделированию *переходных* процессов. Мы ожидаем также, что методы традиционного анализа, основанные на *независимых* наблюдениях, останутся полезными для анализа выхода моделируемой системы. Более того, эти методы необходимы для планирования и анализа при моделировании *нескольких* систем. При моделировании мы можем интересоваться несколькими вариантами системы. Решение о том, какие варианты мы станем изучать, приводят к планированию эксперимента, что обсуждалось в II.6. Анализ выхода нескольких систем можно сделать с использованием дисперсионного анализа, множественных сравнений, множественного ранжирования, методологии анализа поверхности отклика. Все эти методы предполагают, что наблюдения независимы.

II.10. ПРОЧИЕ СТАТИСТИЧЕСКИЕ АСПЕКТЫ

Остались прочие статистические аспекты моделирования. Например, планирование и анализ эксперимента усложняется, если имеется более чем одна выходная переменная, представляющая интерес. В [Naylor et al., 1967a, p. 339—440] отмечают, что *многомерные отклики* расширяют область статистического планирования и анализа и что практики решают эту проблему, рассматривая эксперименты с многомерным откликом как многократное повторение эксперимента с единственным откликом. (В главе VI мы приведем пример анализа, где также есть многомерный отклик. Анализируется метод множественного ранжирования для нескольких уровней надежности P^* , каждому из которых ставится в соответствие отклик.) Теоретически можно изучать при моделировании несколько систем для того, чтобы отобрать оптимальную систему; отбор при этом основан на использовании некоторого критерия. Этот *критерий* — функция, которая объединяет несколько откликов в одно выражение так, чтобы можно было осуществить выбор и таким образом избавиться от проблемы многомерного отклика. В [Naylor, 1971, p. 28] упомянута функция полезности Фромма для сравнения нескольких экономических стратегий. На *практике* может не оказаться готовности к определению критериальной функции или возможностей для этого. При *предварительном* исследовании, когда мы только хотим прозондировать возможные соотношения, критериальная функция также часто отсутствует. Для обсуждения проблемы многомерного критерия мы отсылаем читателя к [Bakes et al., 1970, p.47—50]. Нам на самом деле придется иметь дело с многомерным откликом. Мы будем кратко возвращаться к нему в следующих главах.

Существуют и другие статистические аспекты моделирования кроме тех, которые мы уже обсуждали. Например, в [Copway, 1963, p.59—60] дано несколько правил принятия решения о том, *какие* конкретные «мгновения» надо «останавливать» после получения выборки объема N ,

когда, например, надо наблюдать требования (т. е. наблюдать ли первые N требований, которые поступили после некоторого момента времени, или первые N требований, которые покинули систему после этого же момента времени). В [Conway et al., 1963, p. 54] также обсуждается измерение работы «постоянных элементов» системы, подобных обслуживающим каналам. В [Orcutt et al., 1961, p.377] также упомянута возможность предсказания отклика системы не непосредственно по результатам моделирования. Вместо этого можно использовать *регрессионное уравнение* для предсказания выхода, включающее переменные, применяемые при моделировании выхода, либо включающие и эти переменные и некоторые другие данные, например суммарные экономические данные, собираемые правительством. В [Nelson, 1966, p.19—20] отмечено, что фактически надо наметить план и анализ перед тем, как выполнять эксперимент. Однако часто заранее неизвестно, какие данные окажутся необходимыми. Лучше иметь очень много данных, чем очень мало. Следовательно, ЭВМ может выводить данные на перфоленту, а эта перфолента может использоваться позднее для ввода в ЭВМ в программе анализа выхода.

Наконец, отметим, что статистическую надежность следует отделить от *численной точности*. Ломбайрз [Lombaers, 1968, p.250] отметил следующие факторы точности:

1) число *значимых цифр*, используемых в вычислениях. Как и Хаузер [Hayser et al., 1966, p.84], замечаем, что потерю значимых цифр можно уменьшить при вычислении дисперсии для величины $(x_i - a)$ вместо x_i (где a — подходящим образом выбранная константа), поскольку дисперсия для x_i и для $(x_i - a)$ одинакова;

2) *интерполяция* по таблицам. Как отмечено в [Naylor et al., 1967a, p. 250], некоторые языки моделирования, например GPSS, используют интерполяцию по таблицам при наличии выборочных случайных величин; такая интерполяция также применяется в [Orcutt et al., 1961, p. 123—124].

II.11. ЛИТЕРАТУРА

Мы надеемся, что эта глава поможет читателю, интересующемуся определенными статистическими аспектами моделирования, найти подходящую литературу. Если его интересуют вообще статистические аспекты моделирования, то мы рекомендуем работы [Naylor et al., 1967a] и [Naylor et al., 1967b]. Эти две публикации дают введение в различные аспекты планирования и анализа имитационных экспериментов на ЭВМ, а также содержат значительно большую библиографию. Обширный список литературы, которую мы относим к категории «формальные и статистические методы», отобран в библиографии, данной в [Dutton and Starbuck, 1971b, p.164—167] (см. также [Dutton and Starbuck, 1971a, p.589—594, 693—699, 705—706] и [Meier et al., 1969, p.299—330]). Мы надеемся, что наш собственный анализ также поможет разъяснить статистическую подоплеку моделирования. Для такого анализа мы попытались охватить много статистических аспектов моделирования с некоторыми подробностями. Остальные главы содержат

большее количество ссылок. В этих главах сосредоточено внимание на отдельных статистических техниках. Они дают как обзор имеющихся методов, так и некоторые новые результаты. Глава III охватывает методы понижения дисперсии. В главе IV обсуждаются планы экспериментов. В главе V показано влияние объема выборки на надежность (включая множественные сравнения и процедуры ранжирования). В главе VI дается анализ робастности процедуры ранжирования применительно к нескольким методам, описанным в предыдущих главах.

У П Р А Ж Н Е Н И Я ¹⁶

1. Пусть определенные наблюдения в моделируемой реализации положительно коррелированы и при этом нет переходного процесса. Докажите, что более эффективно генерировать две независимые реализации по n наблюдений в каждой, чем удлинить реализацию до числа наблюдений $2n$.

2. Как можно было бы улучшить эффективность указанной выше процедуры генерирования двух реализаций взамен одной удлиненной реализации? (см. II.4).

3. Пусть надо смоделировать два варианта системы с положительной корреляцией. Как это можно сделать? Каковы при этом преимущества? Как можно проверить различия между средними значениями откликов системы?

4. Моделируются четыре системы, что дает средние $\hat{\mu}_i$ ($i = 1, \dots, 4$), затем используется t -статистика Стьюдента для того, чтобы построить доверительные интервалы для $\mu_i - \mu_{i'}$ ($i \neq i'$). Какова вероятность того, что доверительный интервал для $\mu_1 - \mu_2$ накрывает доверительный интервал для $\mu_3 - \mu_4$. Содержит ли такой совместный интервал также доверительные интервалы для $\mu_1 - \mu_2$ и $\mu_1 - \mu_3$?

5. Пусть моделируется система запасов, при этом входной фактор «требование» отбирается из распределения с математическим ожиданием η . Среднее выборочное значение числа требований в реализации будет \bar{y} . Среднее значение выхода оценивается не с помощью $\hat{\mu}$, а с помощью выражения $\hat{\mu} - a(\bar{y} - \eta)$ (так называемый метод контрольных величин). Определите оптимальное значение постоянной a .

6. В [Naylor et al., 1967b, p.705] использована формула установившегося состояния для оценки ожидаемого общего дохода в течение всего планируемого периода (т. е. включая доход в переходной стадии в начале планируемого периода). Корректно ли применять эту формулу?

7. Смоделируйте простую систему массового обслуживания. Оцените среднее время ожидания в установившемся состоянии, один раз используя все наблюдения по Бломквисту, а другой раз исключая наблюдения переходной стадии применив эмпирическое правило Конвея (см. II.4).

8. Проверьте, даст ли имитационная модель для простой системы массового обслуживания значимое различие среднего времени ожидания при замене экспоненциального распределения времени ожидания на распределение Гаусса с тем же математическим ожиданием.

9. Выберите модель с известным аналитическим решением, например простую систему массового обслуживания или модель запасов, или дифференциальные уравнения эконометрики. Смоделируйте систему и проверьте имитационную модель.

10. Покажите, что применение дополняющих переменных в системах массового обслуживания с одним обслуживающим устройством и обслуживанием в порядке поступления приводит к отрицательной корреляции между средними значениями времени ожидания в двух реализациях. В одной реализации используются r_1, r_2, \dots , а в другой $(1 - r_1), (1 - r_2), \dots$.

П Р И М Е Ч А Н И Я

¹ Обзор этой главы опубликован в (Management Informatics); см. [Kleijnen, 1972].

² См. [Chambers and Mullick, 1968, p. 100], [Dutton and Starbuck, 1971 a, p. 705], [Emshoff and Sisson, 1971, p.49—59], [Kiviat, 1967, p.19—22], [Mihram, 1972a, p. 209—260], [Naylor et al., 1967a, p.23—41].

³ См. [Chambers and Mullick, 1968, p.106], [Emshoff and Sisson, 1971, p.228], [Harling, 1958, p.17], [Hillier and Lieberman, 1968, p.444], [Kiviat, 1967, p.31], [Mihram, 1972a, p. 236], [Naylor et al., 1967a, p.69].

⁴ В [Emshoff and Sisson, 1971, p.190] дано иное определение установившегося состояния (в связи с возможностью циклического поведения).

⁵ В [Blomqvist, 1970] для простых систем массового обслуживания доказано что времена ожидания в переходной фазе имеют монотонно увеличивающиеся, средние и средние квадратичные отклонения.

⁶ См. [Chambers and Mullick, 1968, p.101], [Emshoff and Sisson, 1971, p.190, 192, 239], [Marks, 1964, p.256], [Mihram, 1970, p.89], [Saleeb and Hartley, 1968], [Schmidt and Taylor, 1970, p.346, 436], [Smith and Shan, 1964, p.219], [van Horn, 1971, p.252].

⁷ Коэффициент Тейла (Theil) имеет вид:

$$\left\{ \sum_{t=1}^m \sum_{i=1}^n (y_{it} - \tilde{y}_{it})^2 / mn \right\}^{1/2} / \{ \Sigma \Sigma y_{it}^2 / mn \}^{1/2},$$

где y_{it} и \tilde{y}_{it} обозначают фактическую и предсказанную переменную i соответственно за период времени t (см. [Driehuis, 1972, p. 185]).

⁸ Некая «третья сторона» должна определить, какие ответы дает машина и какие человек. Если она не может различить ответы, то моделирование обоснованно (см. [Mitroff, 1969, p. 636]).

⁹ См. [Elton and Rosenhead, 1971, p. 122—124, 128—219], [Emshoff and Sisson, 1971, p. 204—206], [Schmidt and Taylor, 1970, p. 498—507], [van Horn, 1971]. Дополнительные ссылки [Clarkson, 1968], [Hanna, 1971], [Lovell, 1969, p. 10], [Mihram, 1970, p. 83], [Starbuck, 1971, p. 678—679].

¹⁰ См. [Bruzelius, 1972b], [Fagerstedt and Petterson, 1972], [Fain et al., 1970], [Gafarian and Walsh, 1966], [Kiviat, 1967, p. 48—49], [Maisei and Gnugnoli, 1972], [McKenney, 1967], [Meier et al., 1969, p. 294—296], [Mihram, 1972b], [Naylor, 1971, p. 119, 153—164], [Nelson, 1966, p. 19], [Schrank and Holt, 1967], [Smith, 1970].

¹¹ Пример, взятый из [Mihram, 1972a, p.448]: пусть характеристики генеральной совокупности: $y(0) = +1,0$ или -1 с равной вероятностью (1/3) и $y(t) = y(0)$ при $t = 1, 2, \dots$ Тогда единичная реализация приводит к оценкам: $\hat{y}_t = +1,0$ или -1 [в зависимости от первоначального выборочного значения $y(0)$] и $\text{var}(y_t) = 0$. Обсуждение эргодичности приведено у Парзена [Parzen, 1962, p.72—76].

¹² Следуя Бломквисту [Blomqvist, 1970], мы могли бы использовать все наблюдения переходного и установившегося состояния, допуская, что эта процедура приводит к более эффективной оценке. Ограничение анализа рассмотрением только наблюдений установившегося состояния привело бы к меньшей надежности оценки. Другой подход состоит в том, что мы могли бы считать наблюдения переходной стадии относящимися к установившемуся состоянию, допуская, что это не влияет существенно на анализ.

¹³ $\text{var}(x) = \frac{\sigma^2}{N} \left\{ 1 + 2 \sum_{s=1}^N \left(1 - \frac{s}{N} \right) \rho_s \right\}$, где $\bar{x} = \sum_{i=1}^N x_i / N$, $\text{var}(x_i) = \sigma^2$ и ρ_s — сериальный коэффициент корреляции при лаге s .

¹⁴ Сравните [Dear, 1961, p.18, 20], [Emshoff and Sisson, 1971, p. 202], [Van Horn, 1971, p. 252].

¹⁵ Много примеров можно найти в [Schmidt and Taylor, 1970, p. 507—515].

БИБЛИОГРАФИЯ

1. Aigner D. J. (1972). A note on verification of computer simulation models, *Management Sci.* 18, 615—619.
2. Anderson T. W. (1971). *The Statistical Analysis of Time Series*. Wiley, New York. Русский перевод: Андерсон Т. Статистический анализ временных рядов. М., «Мир», 1976.
3. Bakes M. D., Bramson M. J., Freckleton S., Roberts P. C. and Ryan D. (1970). Stochastic-network reduction and sensitivity techniques in a cost effectiveness study of a military communication system, *Operational Res. Quart.* 21 (Spec. Conf. Iss.), 45—67.
4. Blomqvist N. (1970). On the transient behaviour of the GI/G/I waiting-times. *Skand. Aktuar.* 118—129.
5. Bruzelius L. H. (1972a). Estimating Endogenous Parameters in a Dynamic Stimulation Model; in *Working Papers, vol. 1, Symposium Computer simulation versus analytical solutions for business and economic models*, Graduate School of Business Administration, Gothenburg (Sweden).
6. Bruzelius L. H. (1972b). A Proposal for a General Method for Validating a Simulation Model. — In: *Working Papers, vol. 1, Symposium Computer simulation versus analytical solutions for business and economic models*, Graduate School of Business Administration, Gothenburg (Sweden).
7. Carter G. and Ignall E. (1970). *A Simulation Model of Fire Department Operations: Design and Preliminary Results*. R-632-NYC, The New York City Rand Institute, New York.
8. Chambers J. C. and Mullick S. K. (1968). Strategic new product planning models for dynamic situations. *IEEE Trans. Eng. Management*, 15, 100—108.
9. Chan M. M. W. (1971). System simulation and maximum entropy. *Operations Res.*, 19, 1751—1753.
10. Clarkson G. P. E. (1968). Letter to the editor. *Management Sci.*, 14, 548—550.
11. Glough D. J., Levine J. B., Mowbray G. and Walter J. R. (1965). A simulation model for subsidy policy determination in the Canadian uranium mining industry. *Can. Operational Res. Soc.*, 3, 115—128.
12. Cohen K. J. (1960). *Computer Models of the Shoe, Leather, Hide Sequence*, Prentice-Hall, Englewood Cliffs, N. J.
13. Conway R. W. (1963). Some tactical problems in digital simulation. *Management Sci.* 10, 47—61.
14. Conway R. W., Johnson B. M. and Maxwell W. L. (1959). Some problems of digital systems simulation. *Management Sci.*, 6, 92—110.
15. Crane M. A. and Iglehart D. L. (1972a). A New Approach to Simulating Stable Stochastic Systems: I — General Multi-Server Queues, Technical Report No. 86—1, Control Analysis Corp., Palo Alto, Calif.
16. Crane M. A. and Iglehart D. L. (1972b). Simulating Stable Stochastic Systems: II — Markov Chains. Technical Report No. 86—3, Control Analysis Corp., Palo Alto, Calif.
17. Dear R. E. (1961). Multivariate Analyses of Variance and Covariance for Simulation Studies Involving Normal Time Series. Field note 5644, System Development Corp., Santa Monica, Calif.
18. De Wolff P. (1967). Macro-economic forecasting. — In: *Forecasting on a Scientific Basis (Proceedings international summer institute 1966, sponsored by the NATO Science Committee and the Gulbenkian Foundation)*, Centro de economia e finanças, Lisbon.
19. Driehuis W. (1972). *Fluctuations and Growth in a Near Full Employment Economy*. Universitaire Pres. Rotterdam.
20. Dutton J. M. and Starbuck W. H. (1971a). *Computer Simulation of Human Behavior*, Wiley, New York.
21. Dutton J. M. and Starbuck W. H. (1971b). Computer simulation models of human behavior: a history of an intellectual technology. *IEEE Trans. Systems, Man, Cybernetics, SMC—1*, 128—171.

22. Elton M. and Rosenhead J. (1971). Micro-simulation of markets. *Operational Res. Quart.*, 22, 117—144.
23. Emshoff J. R. and Sisson R. L. (1971). *Design and Use of Computer Simulation Models*. Macmillan, New York.
24. Fagerstedt L. and Patterson S. (1972). Validation of Simulation Models — Some Methodological Views. — In: Working Papers, vol. 1, Symposium Computer simulation versus analytical solutions for business and economic models; Graduate School of Business Administration, Gothenburg (Sweden).
25. Fain W. W., Fain J. B., Feldman L. and Simon S. (1970). Validation of Combat Models Against Historical Data. Professional paper No. 27, Center for Naval Analyses, Arlington, Va (obtainable at Clearinghouse, Springfield, Va.).
26. Fetter R. B. and Thompson J. D. (1965). The simulation of hospital systems. *Operations Res.*, 13, 689—711.
27. Fishman G. S. (1967). Digital Computer Simulation: The Allocation of Computer Time in Comparing Simulation Experiments RM-5288-1-PR, The Rand Corp., Santa Monica, Calif. (Also published in *Operations Res.*, 16, 280—295, 1968).
28. Fishman G. S. (1967). Digital Computer Simulation: The Allocation The Rand Corp., Santa Monica, Calif.
29. Fishman G. S. (1971). Estimating sample size in computing simulation experiments. *Management Sci.*, 18, 21—38.
30. Fishman G. S. (1972). Estimation in Multiserver Queuing Simulations. Technical Report 58, Department of Administrative Sciences, Yale University, New Haven, Conn.
- 31, 32. Fishman G. S. and Kiviat P. J. (1967). Digital Computer Simulation: Statistical Considerations. RM-5387-PR, The Rand Corp., Santa Monica, Calif., Nov. 1967. (Published as: The statistics of discrete event simulation, *Sci. Simulation*, 10, 185—195, 1968. Reprinted in: J. M. Dutton, W. H. Starbuck, *Computer Simulation of Human Behavior*, Wiley, New York, 1971.)
33. Gafarian A. V. and Anker C. J. (1966). Mean value estimation from digital computer simulation. *Operations Res.*, 14, 25—44.
34. Gafarian A. V. and Walsh J. E. (1966). Statistical approach for validating simulation models by comparison with operational systems. 4th Int. Conf. Operational Res. (S. B. Hertz and J. Melese, eds.) Wiley-Interscience, New York.
35. Gregg L. W. and Simon H. A. (1967). Process models and stochastic theories of simple concept formation. *J. Math. Psychol.*, 4, 246—276. (Reprinted in: J. M. Dutton and W. H. Starbuck, *Computer Simulation of Human Behavior*. Wiley, New York, 1971).
36. Gürtler H. (1969). Quantitative Modelle zur Optimierung des Schalterverkehrs in einem Postamt. (Quantitative models for optimizing traffic at counters in a post office.) Doctoral dissertation, Wilhelms-Universität, Münster, Germany.
37. Hanna J. F. (1971). Information-theoretic techniques for evaluating simulation models. — In: J. M. Dutton and W. H. Starbuck, eds. *Computer Simulation of Human Behavior*, Wiley, New York.
38. Harling J. (1958). Simulation techniques in operations research. *Operational Res. Quart.*, 9, 9—21.
39. Hauser N., Barish N. N. and Ehrenfeld S. (1966). Design problems in a process control simulation. *J. Ind. Eng.*, 17, 79—86.
40. Herzberg A. M. and Cox D. R. (1969). Recent work on the design of experiments: a bibliography and a review. *J. Roy. Stat. Soc. Ser. A*, 132, Part 1, 29—67.
41. Hill W. J. and Hunter W. G. (1966). A review of response surface methodology: a literature survey. *Technometrics*, 8, 571—590.
42. Hillier F. S. and Lieberman G. J. (1968). *Introduction to Operations Research*. Holden—Day, San Francisco, Calif., Chapter 14.
43. Hoggatt A. C. and Holbrugge B. J. (1966). Statistical techniques for the computer analysis of simulation models. — In: *Studies in a Simulated*

- Market (L. F. Preston and N. R. Collins, eds.). Institute of Business and Economic Research, University of California, Berkeley, Calif.
44. Howrey E. P. (1972). Selection and Evaluation of Econometric Models. — In: Working Papers, vol. 4, Symposium Computer simulation versus analytical solutions for business and economic models, Graduate School of Business Administration, Gothenburg (Sweden).
 45. Huismann F. (1969). Statistische Aspecten van Simulatie (Statistical aspects of simulation). Report no. 2, Afdeling Werktuigbouwkunde, Technische Hogeschool Twente, Enschede (Netherlands).
 46. Johnston J. (1963). *Econometric Methods*. McGraw-Hill, New York.
 47. Kabak I. W. (1968). Stopping rules for queuing simulations. *Operations Res.*, 16, 431—437.
 48. Kiviat P. J. (1967). *Digital Computer Simulation: Modeling Concepts*. RM-5378-PR, The Rand Corp., Santa Monica, Calif.
 49. Kleijnen J. P. C. (1972). The statistical design and analysis of digital simulation: a survey. *Management Informatics*, 1, 57—66.
 50. Kotler P. (1970). A guide to gathering expert estimates. *Business Horizons*, 13, 79—87.
 51. Lombaers H. J. M. (1968). Enige statistische aspecten van simulatie (Some statistical aspects of simulation). *Statistica Neerl.*, 22, 249—255.
 52. Lovell C. C. (1969). *Simulation in Field Testing*. P-4152. The Rand Corp., Santa Monica, Calif.
 53. McKenney J. L. (1967). Critique of: "verification of computer simulation models". *Management Sci.*, 14, 102—103.
 54. Meisel H. and Gnugnoli G. (1972). *Simulation of Discrete Stochastic Systems*. Science Research Assoc., Palo Alto, Calif.
 55. Marks B. L. (1964). Digital simulations of runway utilization. *Operational Res. Quart.*, 15, 249—259.
 56. Mechanic H. and McKay W. (1966). Confidence Intervals for Averages of Dependent Data in Simulations II. Technical Report 17—202, IBM Advanced Systems Development Division, Yorktown Heights, N. Y.
 57. Meier R. C., Newell W. T. and Pazer H. L. (1969). *Simulation in Business and Economics*. Prentice-Hall, Englewood Cliffs, N. Y.
 58. Mihram G. A. (1970). A cost-effectiveness study for strategic airlift. *Transportation Sci.*, 4, 79—96.
 59. Mihram G. A. (1972a). *Simulation: Statistical Foundations and Methodology*. Academic, New York.
 60. Mihram G. A. (1972b). Some practical aspects of the verification and validation of simulation models. *Operational Res. Quart.*, 23, 17—29.
 61. Mitroff I. I. (1969). Fundamental issues in the simulation of human behavior: A case in the strategy of behavioral science. *Management Sci., Application Ser.*, 15, 635—649.
 62. Nance R. E. (1971). On time flow mechanisms for discrete system simulation. *Management Sci.*, 18, 59—73.
 63. Naylor T. H. (1971). *Computer Simulation Experiments with Models of Economic Systems*. Wiley, New York.
 64. Naylor T. H., Balintfy J. L., Burdick D. S. and Chu K. (1967a). *Computer Simulation Techniques*. Wiley, New York.
 65. Naylor T. H., Burdick D. S. and Sasser W. E. (1967b). Computer simulation experiments with economic systems: the problem of experimental design. *J. Amer. Stat. Assoc.*, 62, 1315—1337.
 66. Naylor T. H. and Finger J. M. (1967). Verification of computer simulation models. *Management Sci.*, 14, 92—101.
 67. Naylor T. H., Wertz K. and Wonnacott T. (1968). Some methods for evaluating the effects of economic policies using simulation experiments. *Rev. Int. Stat. Inst.*, 36, 184—200.
 68. Nelson R. T. (1966). *Systems, Models and Simulation*. Mimeographed notes, Graduate School of Business Administration, University of California, Los Angeles, Calif.
 69. Niedereichholz J. (1971). Business simulation and management decisions. *Management Int. Rev.*, 11, 47—52.

70. Orcutt G. H., Greenberger M., Korbel J. and Rivlin A. M. (1961). *Microanalysis of Socioeconomic Systems: A Simulation Study*. Harper, New York.
71. Overholt J. L. (1969). The problem of factor selection. — In: *The Design of Computer Simulation Experiments* (T. H. Naylor, ed.), Duke University Press, Durham.
72. Parzen E. (1962). *Stochastic Processes*. Holden-Day, San Francisco, Calif.
73. Pegels C. C. (1969). Simulation and the optimal design of a production process. *Int. J. Prod. Res.*, 7, 219—231.
74. Saleeb S. and Hartley M. G. (1968). Simulation of traffic behaviour through a linked-pair of intersections. *Transportation Res.*, 2, 51—61.
75. Schlaifer R. (1959). *Probability and Statistics for Business Decisions*, McGraw-Hill, New York.
76. Schmidt J. W. and Taylor R. E. (1970). *Simulation and Analysis of Industrial Systems*. Richard D. Irwin, Inc., Homewood, Ill.
77. Schrank W. E. and Holt C. C. (1967). Critique of "verification of computer simulation models". *Management Sci.*, 14, 104—106.
78. Smith J. U. M. (1970). Computer simulation of industrial operations. — In: *Progress of Cybernetics*, vol. 2 (J. Rose, ed.), Gordon and Breach, London.
79. Smith W. P. and Shah J. C. (1964). Design and development of a manufacturing systems simulator. *J. Ind. Eng.*, 15, 214—220.
80. Starbuck W. H. (1971). Testing case-descriptive models. — In: *Computer Simulation of Human Behavior* (J. M. Dutton and Starbuck W. H., eds.), Wiley, New York.
81. Stash S. F. (1968). Some aspects of Multidimensional Verification of Simulations. Presented at the Symposium on the design of computer simulation experiments, Duke University, Durham, North Carolina, 14—16 Oct. 1968. (An abstract is published in: *The Design of Computer Simulation Experiments* (T. H. Naylor, ed.), Duke University Press, Durham, N. C. 1969.)
82. Tocher K. D. (1963). *The Art of Simulation*. English Universities Press, London.
83. Van Daal J. and van Doel and F. (1972). Waiting for the Bridge. Report 7208, Econometric Institute, Netherlands School of Economics, Rotterdam.
84. Van Horn R. L. (1971). Validation of simulation results. *Management Sci.*, 17, 247—258.
85. Wagner H. M. (1971). The ABC's of OR. *Operations Res.*, 19, 1259—1281.
86. Wetherill G. B. (1966). *Sequential Methods in Statistics*. Methuen, London, and Wiley, New York.

III.1. ВВЕДЕНИЕ И РЕЗЮМЕ (КРАТКОЕ ИЗЛОЖЕНИЕ)

В этой главе мы рассмотрим некоторые методы понижения дисперсии (МПД). В II.4 мы упоминали о том, что МПД основаны на замене простой случайной выборки более совершенной выборкой. Точнее, МПД понижают дисперсию оценки с помощью замены исходной выборочной процедуры новой процедурой, которая приводит к оценке с тем же математическим ожиданием, но с меньшей дисперсией. Некоторые МПД приводят к полной замене исходного выборочного процесса. Например, в III.5 мы рассмотрим метод значимой выборки, при котором исходное распределение, из которого производится выборка, заменяется совершенно новым распределением. При других МПД используется тот же случайный отбор, но потом вместо простого среднего \bar{x} берется более сложная оценка. Примерами служат метод стратифицированной выборки и метод контрольных переменных, которые будут рассмотрены в III.2.4 и III.4 соответственно. Некоторые методы модифицируют выборочный процесс очень сложным путем, в этом мы убедимся в III.6 и III.7, где обсудим метод дополняющих переменных и метод общих случайных чисел.

В литературе по МПД обычно освещается только применение этих методов к оценкам *среднего* или математического ожидания. Следовательно, мы также будем полагать, что цель имитационного эксперимента — это оценка среднего значения моделируемой системы. Таким откликом могут быть время ожидания требования в установившемся режиме, общий доход фирмы в течение всего планируемого периода и т. п.

Мы не будем исследовать, можно ли и если можно, то как применять МПД для улучшения оценок дисперсии, квантилей или сериальных коэффициентов корреляции².

Заметим, что целью некоторых имитационных исследований является не оценка среднего, а оценка вероятности того, что отклик превзойдет некоторую фиксированную величину, т. е. процентиля. В качестве примеров можно привести систему, моделирующую систему массового обслуживания, в которой мы должны оценить вероятность того, что время ожидания будет больше некоторого фиксированного числа, или экономическую систему, для которой мы хотели бы вычислить вероятность того, что дефицит платежного баланса превысит некоторую фиксированную сумму. Однако оценку такой вероятности очень просто сформулировать как оценку среднего. Как мы уже видели (уравнение (2.5), приложение 1.2), можно ввести новую переменную, скажем u ,

с математическим ожиданием p , где p — вероятность того, что интересующая нас переменная, скажем x , превысит некоторое фиксированное значение.

В литературе эффективность МПД обычно измеряется уменьшением дисперсии оценки среднего³. На первый взгляд непонятно, почему берется дисперсия, а не стандартное отклонение. Надежность оценки может быть определена с помощью доверительного интервала вида

$$\bar{x} \pm a (\sigma/\sqrt{n}), \quad (1)$$

где \bar{x} — среднее n независимых наблюдений, т. е. независимых имитационных опытов (прогонов), σ — стандартное отклонение отдельного наблюдения и a — некоторая константа. Однако измерение уменьшения дисперсии также имеет смысл. Как мы увидим в главе V, число наблюдений n_w , необходимое для достижения заданной надежности, можно выразить как функцию дисперсии:

$$n_w = b\sigma^2, \quad (2)$$

где b — подходящая константа. Итак, по нашему мнению, при применении МПД в случае, когда объем выборки фиксирован, целесообразно измерять уменьшение стандартного отклонения, так как изменение стандартного отклонения пропорционально длине доверительного интервала. Если предположить, что желаемая надежность задана заранее, то целесообразно брать дисперсию, так как ее уменьшение пропорционально числу наблюдений, необходимых для достижения заданной надежности. Естественно, понижение дисперсии влечет за собой понижение и стандартного отклонения, и наоборот. Заметим, что мы определяем понижение дисперсии (в процентах) как $(\sigma_0^2 - \sigma_1^2)/\sigma_0^2 (\times 100\%)$, где σ_0^2 и σ_1^2 — старая и новая дисперсии соответственно. Радема [Radema, 1969] и Андреассон [Andreasson, 1971b, p.13—14] рассмотрели стандартную ошибку оцениваемого понижения дисперсии. Радж [Raj, 1968, p.190] дает формулу коэффициента вариации оценки дисперсии, т. е. $\{\text{var}(\hat{\sigma}^2)\}^{1/2}/E(\hat{\sigma}^2)$. Робастные доверительные интервалы для оценок дисперсии можно получить с помощью метода «складного ножа» [Gray and Schucany, 1972, p.176—182]; метод «складного ножа» описан далее в этой главе и в приложении 1. Помимо понижения стандартного отклонения или дисперсии мы должны рассчитать дополнительное машинное время, которое может потребоваться для МПД. К сожалению, машинное время очень зависит от эффективности частных программ и от типа машины. Тем не менее мы приведем некоторые результаты о дополнительном машинном времени, необходимом для МПД. Рекомендуем также работы [Handscomb, 1968, p.2], [Lewis, 1972, p.11] и [Moу, 1965, p.8—11]. Хендскомб говорит, что МПД заслуживают внимания, «если выгода от понижения дисперсии превышает дополнительные усилия, необходимые для выполнения этой работы». Помимо дополнительного машинного времени мы также должны учитывать время на собственное программирование. Далее в этой главе мы увидим, что в отличие от некоторых сложных методов МПД с дополняющими переменными и общими случайными числами не требуют какого-либо дополнительного времени на программирование.

Впервые МПД применены не в имитационном моделировании, а в исследованиях по методу Монте-Карло⁴ (мы понимаем метод Монте-Карло в том смысле, который пояснен нами в I.5). Серьезный обзор многих МПД, применяемых в исследованиях по методу Монте-Карло, дан в [Hammersley and Handscomb, 1964]. Впоследствии МПД были также применены в имитационном моделировании. Точер [Toscher, 1963, p.171] предложил прибегать к МПД при моделировании сложных промышленных систем. Тем не менее, следуя Максвеллу [Maxwell, 1965, p.2] и Мою [Moу, 1965, p.21], мы приходим к выводу, что МПД не нашли практического применения в имитационном моделировании. Маршалл [Marshall, 1958] и Майер [Meier, 1969, p.274—275] обсуждают причины, которые могли бы объяснить этот факт. Применение МПД при моделировании по сравнению с методом Монте-Карло вызывает ряд новых проблем. Поэтому, прежде чем говорить о МПД, укажем на существенные различия между имитационным моделированием и методом Монте-Карло в двух аспектах (см. I.6.2); за подробным описанием мы отсылаем к [Garman, 1971, p.3—6; 1973, p.5—7].

1. Наблюдения в методе Монте-Карло независимы. В имитационном моделировании процесс протекает во времени; как правило, между наблюдениями есть *серийная корреляция*. Это различие хорошо видно на примерах в приложении к главе I. В приложении I.1 мы оценивали интеграл по n наблюдениям $g(x_i)$ (см. (1.6)). Каждое наблюдение $g(x_i)$ независимо. В приложении I.2 мы оценивали p (вероятность того, что x меньше фиксированной величины a) из N наблюдений y_j (см. (2.10)). В модели из приложения I.3 N рейсов одного и того же дня серийно коррелируют между собой; рейсы в разные дни независимы для стратегий α и β , а для стратегии γ рейсы, относящиеся к разным дням, тоже зависимы между собой. Заметим, что в теории очередей время ожидания требований зависимо. Итак, в общем случае в имитационном моделировании наблюдения серийно коррелируют. (Конечно, мы можем повторить имитационный опыт, используя новую последовательность случайных чисел, тогда как наблюдения, относящиеся к разным опытам, будут независимы.)

2. В задачах Монте-Карло можно выразить отклик как довольно простую функцию от случайных входных переменных. В имитационном моделировании отклик — обычно очень сложная функция от входных переменных и выражается явно только с помощью самой *вычислительной программы* [Garman, 1971, p.5; 1973].

Из-за указанных различий МПД, которые применяются в исследованиях Монте-Карло, нужно приспособлять для имитационного моделирования. Описание такой адаптации можно найти в малоизвестных публикациях [Moу, 1965, 1966], где показано, как различные методы — стратификации, значимой и регрессионной выборки — можно усовершенствовать для имитационного моделирования всех видов систем. Мы в дальнейшем обсудим еще применение в имитационном моделировании селективной выборки, выборки с дополняющими переменными и общих случайных чисел. Как мы увидим далее, все эти шесть методов настолько усовершенствованы, что их можно применять при имитационном моделировании сложных систем. В литературе встре-

чается множество описаний МПД, применение которых в имитационном моделировании ограничено⁵.

Итак, позднее мы сосредоточимся на этих шести МПД, которые применимы в имитационном моделировании *широкого* класса систем. Эти МПД могут также применяться и в различных задачах Монте-Карло. Некоторые МПД, особенно методы стратифицированной и значимой выборки, имеют варианты, упрощенные для задач Монте-Карло. Эти простые варианты будут описаны прежде, чем мы перейдем к обсуждению МПД в имитационном моделировании.

Далее даются *описание и критическая оценка* всех шести методов. Описаны также ограничения, поправки и расширения существующих процедур. Например, показано, что некорректно применять стратифицированную процедуру Мюя при моделировании со случайной длиной прогона; с другой стороны, мы распространим стратифицированную выборку на моделирование со стратификацией после выборки. Метод селективной выборки Бреннера (Brenner) дает смещенную оценку и не применяется в имитационном моделировании со случайной длиной прогона. Отмечено применение статистики «складного ножа» с управляемыми переменными. Мы приведем пример, когда в задачах Монте-Карло удалось понизить дисперсию более чем на 99% с помощью метода значимой выборки (новая функция плотности была выбрана так, чтобы минимизировать размах оценки). Случай многомерного отклика может привести к серьезным трудностям для значимой выборки. Показаны причины, почему метод « 2^k -факторного эксперимента» Точера не подходит для метода дополняющих переменных; для генерации дополняющих случайных чисел представлены несколько процедур.

Обычно в литературе игнорируется *анализ экспериментов*, в которых применялись МПД. Поскольку анализ экспериментов в этом случае усложняется, мы рассмотрим также и этот аспект. Все указанные методы, за исключением метода с дополняющими переменными, приводят к осложнениям при определении доверительных интервалов.

И наконец, мы рассмотрим возможность *совместного* применения двух самых простых методов — с *дополняющими* переменными и общими случайными числами. Мы увидим, что такое применение не может быть оптимальным из-за нежелательных корреляционных эффектов. Поэтому процедура будет усовершенствована так, чтобы обеспечить наибольшее понижение дисперсии. В то же время она будет эффективной в смысле расхода *машинного времени*. Тем, кто интересуется МПД только из практических соображений, мы советуем перейти к III.6 и III.7, в которых рассматриваются методы с дополняющими переменными и общими случайными числами.

III.2. СТРАТИФИЦИРОВАННАЯ ВЫБОРКА

III.2.1. Основы стратифицированной выборки

Стратификация выборки — давно известный прием в *выборочных обследованиях*. Прежде, чем обсуждать применение стратификации в методе Монте-Карло и имитационном моделировании, сформулируем основные принципы стратификации. Предположим, что мы хотим оце-

нить μ — величину среднего потребления некоторой группы населения. Чтобы оценить среднее, мы должны взять выборку n индивидуумов из этой группы. Хорошо известной *простой* оценкой среднего значения μ служит среднее \bar{x} , определенное формулой

$$\bar{x} = \sum_{i=1}^n x_i/n, \quad (3)$$

где x_i — потребление i -го ($i = 1, \dots, n$) случайно выбранного индивидуума. Как известно, \bar{x} представляет собой несмещенную оценку с дисперсией σ^2/n , где σ^2 — дисперсия индивидуального наблюдения x_i .

В *стратифицированной выборке* мы измеряем помимо интересующей нас переменной x_i дополнительную переменную, скажем, y_i , для каждого ($i = 1, \dots, n$) элемента выборки. Эта дополнительная переменная называется *стратифицирующей переменной*. Она служит для того, чтобы отнести каждого отобранного индивидуума к одному из, скажем, K классов, или слоев. Эти слои формируются так, чтобы разделить область стратифицирующей переменной на K непересекающихся классов, исчерпывающих всю совокупность. Обозначим k -й слой S_k ($k = 1, \dots, K$). Допустим, что в нашем примере стратифицирующая переменная — это доход. Положим, что индивидуум с потреблением, равным x , принадлежит классу k , если его доход y принадлежит слою S_k . Для стратификации нам надо знать вероятность того, что индивидуум принадлежит к k -му классу. Обозначим эту вероятность через p_k . Пусть нам известно⁶

$$p_k = P(y \in S_k). \quad (4)$$

Мы можем выбрать n_k индивидуумов с доходами, которые принадлежат слою k . Затем оценить среднее μ с помощью стратифицированной оценки \bar{x}_{ST} , определенной выражением

$$\bar{x}_{ST} = \sum_{k=1}^K p_k \bar{x}_k, \quad (5)$$

где \bar{x}_k — оценка среднего в k -м слое,

$$\bar{x}_k = \sum_{j=1}^{n_k} x_{kj}/n_k, \quad (6)$$

где, в свою очередь, x_{kj} — это j -е наблюдение ($j = 1, \dots, n_k$) в слое k . Стратифицированная оценка будет *несмещенной*, как видно из (7), где мы пользуемся определением (5) и обозначаем среднее в слое k через μ_k :

$$E(\bar{x}_{ST}) = \sum_{k=1}^K p_k E(\bar{x}_k) = \sum_{k=1}^K p_k \mu_k = \mu. \quad (7)$$

Последнее равенство (7) должно быть очевидно читателю, а если нет, то отсылаем к [Fisz, 1967, p.84]. Дисперсия стратифицированной

оценки выводятся из (8) с помощью (5) и (6):

$$\text{var}(\bar{x}_{\text{СТ}}) = \sum_{k=1}^K p_k^2 \text{var}(\bar{x}_k) = \sum_{k=1}^K p_k^2 \frac{\sigma_k^2}{n_k}, \quad (8)$$

где σ_k^2 — дисперсия в слое k , т. е. σ_k^2 — дисперсия x_{kj} из (7), определяемая (9)⁷:

$$\sigma_k^2 = E[(x - \mu_k)^2 | x \in S_k]. \quad (9)$$

Перед тем, как исследовать, приводит ли стратификация к понижению дисперсии, определим *доверительные интервалы* для μ . Очевидно, что σ_k^2 можно оценить как

$$s_k^2 = \sum_{j=1}^{n_k} (x_{kj} - \bar{x}_k)^2 / (n_k - 1). \quad (10)$$

Эту оценку подставим в (8) и получим несмещенную оценку

$$s^2(\bar{x}_{\text{СТ}}) = \sum_{k=1}^K p_k^2 \frac{s_k^2}{n_k}. \quad (11)$$

Далее, следуя Кокрену [Cochran, 1966, p.94—95], определим для заданной надежности $(1 - \alpha)$ *доверительный интервал* среднего μ :

$$\bar{x}_{\text{СТ}} \pm z^{\alpha/2} s(\bar{x}_{\text{СТ}}), \quad (12)$$

где $s(\bar{x}_{\text{СТ}})$ — квадратный корень из $s^2(\bar{x}_{\text{СТ}})$. Значение $z^{\alpha/2}$ можно найти в таблицах *нормального* распределения, если мы полагаем, что $\bar{x}_{\text{СТ}}$ нормально распределена, а $s(\bar{x}_{\text{СТ}})$ хорошо определена. Относительно условия нормальности можно считать, что среднее слоя асимптотически нормально на основании центральной предельной теоремы. Взвешенная сумма независимых нормально распределенных переменных также распределена нормально [Fisz, 1967, p.147, 150]. Итак, мы можем предполагать, что $\bar{x}_{\text{СТ}}$ нормально распределена, по крайней мере приближенно. Однако, если мы имеем не очень много наблюдений в каждом из слоев, асимптотическая нормальность может не выполняться, и при таких условиях $s(\bar{x}_{\text{СТ}})$ становится плохой оценкой. Следуя Кокрену, мы можем определить $z^{\alpha/2}$ в (12) из таблиц *t-статистики Стьюдента*. Укажем при этом на Шеффе [Scheffé, 1964, p.335, 346], который нашел, что *t*-статистика не очень чувствительна к нарушению предположения о нормальности. Более серьезная проблема состоит в том, что $s^2(\bar{x}_{\text{СТ}})$ выражается не как обычно $\sum (x_i - \bar{x})^2 / (n - 1)$, а более сложной функцией, как видно из (11). Кокрен [Cochran, 1966, p.94] советует применять в таком случае аппроксимацию Саттерзвайта [Satterthwaite, 1946], см. также [Raj, 1968, p.193—194]. Эта аппроксимация $z^{\alpha/2}$ снова, как и (12), использует таблицы Стьюдента, но при скорректированном числе степеней свободы⁸ n_c .

$$n_c = \frac{\{s^2(\bar{x}_{\text{СТ}})\}^2}{\sum_{k=1}^K (p_k s_k)^4 / (n_k^3 - n_k^2)}. \quad (13)$$

Как замечает Кокрен, аппроксимация Саттерзвэйта предполагает, что *индивидуальные* наблюдения x_{kj} распределены *нормально*. Предположим, что нас интересует средний отклик системы в установившемся состоянии. На основании центральной предельной теоремы для стационарных r -зависимых реализаций, приведенной в [Fraser, 1957, p.219] и в главе V, можно предположить, что средний отклик такой стационарной системы асимптотически нормален. Однако если мы изучаем нестационарную систему, то эту теорему применить нельзя и доверительный интервал (12) становится еще более грубым. Наконец, общий объем выборки может быть не фиксированным, а зависимым от variability оценки среднего совокупности, оценка которой производится в ходе получения выборки. Тогда число наблюдений в каждом из слоев становится *случайным*. Стохастический характер слоя усложняет определение доверительных пределов (12), так как они основаны на фиксированном объеме выборки. (Мы еще вернемся к этому в III.8 и в главе V). Подводя итоги, можно сказать, что *доверительные пределы* для оценок стратифицированной выборки вычисляются только *приблизженно*. (Здесь возможно применение метода «складного ножа», но это требует дополнительных исследований.)

Вернемся к дисперсии стратифицированной оценки \bar{x}_{ST} и сравним ее с дисперсией простой оценки \bar{x} . Из (8) следует, что дисперсия \bar{x}_{ST} зависит от выбора n_k , числа наблюдений в слое k . Ясно, что надо взять 10% от выборки в слой k , если 10% совокупности попадает в этот слой. Вообще n_k выбирается в соответствии с

$$n_k = p_k n. \quad (14)$$

Если мы выберем n_k согласно (14), то получим так называемую *пропорциональную стратифицированную выборку*. Подстановка (14) в (8) дает дисперсию пропорционально стратифицированной оценки \bar{x}_{PS} .

$$\text{var}(\bar{x}_{PS}) = \frac{1}{n} \sum_{k=1}^K p_k \sigma_k^2. \quad (15)$$

Отношение между дисперсией простой оценки \bar{x} (3) и дисперсией \bar{x}_{PS} , вычисленной для пропорционально стратифицированной выборки, определяется выражением (16); это отношение доказано для дискретных переменных Кокреном [Cochran, 1966, p.98—99], а для непрерывных — Точером [Tocher, 1963, p.106—107]:

$$\text{var}(\bar{x}) = \text{var}(\bar{x}_{PS}) + \sum_{k=1}^K p_k (\mu_k - \mu)^2 / n. \quad (16)$$

Итак, из уравнения (16) видно, что стратификация полезна, если μ_k — средние различных слоев — не все равны (т. е. не равны μ). Средние μ_k отличаются, если интересующая нас переменная x зависит от переменной y , по которой идет стратификация. В нашем примере мы можем предполагать, что x , потребление, коррелирует с y , доходом (здесь мы имеем положительную корреляцию, так как повышенные доходы приводят к росту потребления). Следовательно, стратификация

эффективна, если существует *стратифицирующая переменная*, которая сильно коррелирует с интересующей нас переменной. Очевидно, для стратификации нам надо знать вероятности p_h ; т. е. мы должны знать распределение стратифицирующей переменной, как показано в (4).

Мы можем попытаться определить *оптимальное число* наблюдений в каждом из слоев вместо использования принципа пропорционального распределения (14). Оптимальное размещение можно найти в [Cochran, 1966, p.95—97], где учитывается, существенно ли различие стоимостей выборок каждого из слоев. Однако это реализуемо только в том случае, если мы знаем дисперсию в каждом из слоев σ_h^2 , как определено в (9). Поскольку обычно σ_h^2 неизвестны, мы не будем обсуждать проблему оптимальной выборки.

Обсуждение стратификации можно найти также в учебнике Кокрена [Cochran, 1966], в исследованиях по МПД Моея [Моей, 1965 и 1966]. Кокрен обсуждает выбор оптимального числа наблюдений по слоям для многомерного отклика, применение двух стратифицирующих переменных вместо одной, оптимальный выбор границ и числа слоев, оценивание эффективности стратификации, оценивание дисперсии, если некоторый слой содержит только одно наблюдение, так что знаменатель в (10) обращается в нуль, и т. д.; см. также [Henzler, 1970] и [Raj, 1968, p.61—84, 151—152, 206—209]. В дальнейшем мы обсудим результат Моея [Моей, 1965, p.83—84], который показывает, что число слоев должно быть небольшим, скажем, между 3 и 10. Будет рассмотрен также вариант стратификации, который лишь упомянут Кокреном [Cochran, 1966, p.135—136] и совсем не упоминается у Моея. Как мы увидим в III.2.4, этот вариант, который мы называем стратификацией после выборки, может быть очень важным в имитационном моделировании.

III.2.2. Стратификация после выборки

Предположим, что выборочный процесс представляет собой *простой*, т. е. нестратифицированный отбор из n наблюдений. Однако вместо оценивания выборочного среднего \bar{x} по формуле (3) мы *стратифицируем* эту выборку *после того*, как процесс выбора закончен. В отличие от пропорционально стратифицированной выборки число наблюдений в каждом из слоев не предопределено, а зависит от самой выборки. Новая оценка среднего, назовем ее \bar{x}_{SA} , определяется по формуле

$$\bar{x}_{SA} = \sum_{k=1}^K p_k \bar{z}_k, \quad (17)$$

где

$$\bar{z}_k = \sum_{j=1}^{n_k} x_{kj} / n_k \quad (18)$$

и

$$\sum_{k=1}^K n_k = n. \quad (19)$$

Сравним оценку \bar{x}_{SA} с *простым* выборочным средним \bar{x} . Перепишем выражение для \bar{x} в виде

$$\begin{aligned} \bar{x} &= n^{-1} \sum_{i=1}^n x_i = n^{-1} \sum_{k=1}^K \sum_{j=1}^{n_k} x_{kj} = \\ &= \sum_{k=1}^K \frac{n_k}{n} \sum_{j=1}^{n_k} \frac{x_{kj}}{n_k} = \sum_{k=1}^K \frac{n_k}{n} \bar{z}_k, \end{aligned} \quad (20)$$

где \bar{z}_k определено по формуле (18). Сравнение (20) с (17) показывает, что в оценке среднего при стратификации после выборки *эмпирические* веса n_k/n заменяются на «истинные» или *теоретические* веса p_k .

Далее мы рассмотрим обычную стратифицированную оценку с *предопределенным* числом наблюдений в каждом из слоев. В случае *пропорциональной* стратификации число наблюдений в слое равно ($p_k n$) по (14). Сравнение (18) с (6) показывает, что при стратификации после выборки число наблюдений нефиксировано и зависит от результатов выборки, т. е. становится *случайной* величиной⁹.

Кокрен [Cochran, 1966, p.135—136] показывает, что стратификация *после* выборки совпадает с *пропорциональным* стратифицированным выбором при условии, что число наблюдений в слое довольно велико, скажем, больше 20. Заметим, интуитивно разумно предположить, что по эффективности стратификация после выбора совпадает с пропорциональной стратификацией, так как *математическое ожидание* случайной величины n_k равно размеру слоя в пропорциональной выборке. При стратификации *после* выборки мы имеем

$$E(n_k) = p_k n, \quad (21)$$

а это равно размеру слоя для случая пропорциональной стратификации (14). Стратификация *после* выборки представляет интерес, если невозможно или трудно фиксировать число наблюдений в слое; следовательно, определение слоя, к которому принадлежит наблюдение, происходит лишь *после* того, как выборка произведена.

Как мы только что видели из утверждения Кокрена, эффективность стратификации *после* выборки зависит от условия: объем подвыборки «достаточно большой, скажем, больше 20». Это условие не будет ограничением для больших слоев (когда p_k не очень мало) и для больших объемов выборки n . В противном случае может оказаться, что случайная величина n_k очень мала или даже равна нулю. Если $n_k = 0$, т. е. k -й слой пуст, то мы не сможем оценить \bar{x}_{SA} , так как в формуле (18) мы должны делить на нуль. Однако если окажется, что n_k равно нулю, то мы должны внести уточнение в определение стратификации *после* выборки, данное формулами (17)—(19). Одно возможное уточнение было предложено Кокреном [Cochran, 1966, p.136], но не было им детально разработано. Он предложил объединять два или более слоев, если оказывается, что одно или более значений n_k равны нулю. Мы формализуем предложение Кокрена в виде (22), где дано определение

\bar{x}_{SAC} — оценки, основанной на стратификации после выборки с *объединением* слоев, если это необходимо:

$$\bar{x}_{\text{SAC}} = \sum_{k=1}^K p_k \bar{w}_k, \quad (22)$$

причем

$$\bar{w}_k = \sum_{j=1}^{n_k} x_{kj} / n_k, \text{ если } n_k > 0, \quad (23a)$$

$$\bar{w}_k = \bar{w}_g = \sum_{j=1}^{n_g} x_{gj} / n_g, \text{ если } n_k = 0, n_g > 0, \quad (23b)$$

и
$$\sum_{k=1}^K n_k = n. \quad (24)$$

В приложении III.2 мы покажем, что в общем случае предложение Кокрена приводит к *смещенным* оценкам. Если это смещение и дисперсия \bar{x}_{SAC} невелики, то \bar{x}_{SAC} как оценка еще представляет интерес. Смещение и дисперсия вместе измеряются *средним квадратичным отклонением*, или кратко СКО. Рассмотрим кратко СКО.

СКО определено в [Goldberger, 1964, p.126] как математическое ожидание квадрата отклонения оценки от истинного значения параметра совокупности¹⁰. В нашем случае оценивается μ с помощью \bar{x}_{SAC} и, следовательно, СКО определяется следующей формулой:

$$\text{СКО}(\bar{x}_{\text{SAC}}) = E[(\bar{x}_{\text{SAC}} - \mu)^2]. \quad (25)$$

В приложении III.3 мы докажем хорошо известное соотношение, а именно: СКО равно *дисперсии* оценки плюс квадрат *смещения* оценки. Так, в нашем случае справедливо следующее выражение:

$$\text{СКО}(\bar{x}_{\text{SAC}}) = \text{var}(\bar{x}_{\text{SAC}}) + [E(\bar{x}_{\text{SAC}}) - \mu]^2, \quad (26)$$

где

$$\text{var}(\bar{x}_{\text{SAC}}) = E\{[\bar{x}_{\text{SAC}} - E(\bar{x}_{\text{SAC}})]^2\}. \quad (27)$$

Далее в приложении III.3 мы покажем, что несмещенную оценку СКО \bar{x}_{SAC} можно вычислить, если есть N повторных выборок, т. е. нам известны N оценок \bar{x}_{SAC_i} ($i = 1, \dots, N$), а также известен параметр μ совокупности (в работах по методу Монте-Карло для оценки эффективности различных статистик мы должны знать μ). *Несмещенная оценка* СКО (25) определяется следующим выражением:

$$\hat{\text{СКО}}(\bar{x}_{\text{SAC}}) = \sum_{i=1}^N (\bar{x}_{\text{SAC}_i} - \mu)^2 / N. \quad (28)$$

Можно оценить СКО (28), по крайней мере приближенно, даже если значение генерального среднего неизвестно [Raj, 1968, p.88—93, 105, 149]. После краткого обсуждения СКО вернемся вновь к смещенной оценке \bar{x}_{SAC} .

Поскольку СКО измеряет дисперсию и смещение одновременно, эта оценка служит полезным критерием для выбора той оценки, которая наименее смещена (если некоторая оценка несмещенная, то ее СКО уменьшается до величины дисперсии). К сожалению, вычислить значение СКО для \bar{x}_{SAC} не так просто (см. также примечание 11). Заметим, что смещение \bar{x}_{SAC} зависит от той совокупности, из которой производится выбор. Именно для этой совокупности определяется смещение, которое возникает в результате замены \bar{w}_h на \bar{w}_g в (22).

Далее мы рассмотрим другой способ коррекции оценки, основанной на стратификации после выборки, для случая, когда есть пустой слой. Точер [Tocher, 1963, p.105] предполагает *продолжить* выборку до тех пор, пока каждый из слоев не будет содержать по крайней мере одно наблюдение. В этом случае общий объем выборки нельзя фиксировать и приравнять к n . При пустом слое мы должны продолжать (простой) случайный отбор. Формализация предложения Точера содержится в выражениях (29)—(32). Эти уравнения определяют \bar{x}_{SAP} — оценку, основанную на стратификации после выборки с продолжением отбора до тех пор, пока каждый из слоев становится непустым. \square

$$\bar{x}_{SAP} = \sum_{k=1}^K p_k \bar{v}_k, \quad (29)$$

где

$$\bar{v}_k = \sum_{j=1}^{n_k} x_{kj} / n_k. \quad (30)$$

$$n_k > 0 \text{ для всех } k, \quad (31)$$

$$\sum_{k=1}^K n_k = n. \quad (32)$$

Несложно доказать, что это *несмещенная* оценка. Трудным представляется сравнение дисперсии \bar{x}_{SAP} с СКО \bar{x}_{SAC} или с дисперсией \bar{x}_{PS} и \bar{x} . Необходимы дополнительные исследования в этой области (см. [Williams, 1964, p.1060—1061]).

Мы видели ранее, как можно определить *приближенные доверительные пределы* (12) для среднего совокупности из стратифицированной выборки с *фиксированным* числом наблюдений в каждом из слоев. Когда мы производим стратификацию *после* выборки, эти пределы могут стать более грубыми, потому что при определении средних по слоям \bar{z}_k , \bar{w}_k и \bar{v}_k для оценок \bar{x}_{SA} , \bar{x}_{SAC} и \bar{x}_{SAP} соответственно число наблюдений в слое n_k будет *случайным*. Оценку стандартного отклонения для доверительного интервала (12) можно получить следующим образом: в приложении III.4 мы доказываем, что несмещенная оценка дисперсии \bar{v}_k дается выражением

$$\hat{\text{var}}(\bar{v}_k) = \frac{\sum_{j=1}^{n_k} (x_{kj} - \bar{v}_k)^2}{(n_k - 1)} \cdot \frac{1}{n_k} \quad (33)$$

с заменой начального условия (31) условием (34), чтобы знаменатель выражения (33) не обращался в нуль;

$$n_k > 1 \text{ для всех } k. \quad (34)$$

Следовательно, несмещенная оценка дисперсии оценки x_{SAP} определяется следующим выражением:

$$\widehat{\text{var}}(\bar{x}_{\text{SAP}}) = \sum_{k=1}^K p_k^2 \widehat{\text{var}}(\bar{v}_k). \quad (35)$$

Существует аналогичная оценка дисперсии \bar{x}_{SA} для условия $n_k > 1$ в (18). С оценкой x_{SAC} связано много проблем, так как это смещенная оценка и представляется сложным вывести несмещенную оценку для ее дисперсии¹¹. Так, определение доверительных пределов для среднего μ становится более трудным, чем для других стратифицированных оценок.

В заключение заметим, что простые способы получения *неслучайного* числа наблюдений можно найти в [Handscorn, 1968, p.6], их идея сводится к следующему. Определяем желаемое число наблюдений в слое, берем одно наблюдение и смотрим, к какому слою оно принадлежит. Если слой уже заполнен, надо *отбросить* это наблюдение, если нет, то взять, и т. д. Эта процедура отбрасывания занимает много времени, если и выборка, и установление, к какому слою отнести данное наблюдение, требуют много времени. В случае стратифицированной выборки отличие состоит в том (как обсуждалось уже в III.2.1), что предполагается известным заранее, к какому слою принадлежит данное наблюдение. Мы увидим в III.2.4, как это можно узнать при имитационном моделировании.

III.2.3. Стратифицированная выборка в методе Монте-Карло

В своей хорошо известной книге по имитационному моделированию Точер [Tocher, 1963] приводит примеры стратифицированной выборки. В действительности примеры относятся не к имитационному моделированию вообще, а к методам Монте-Карло, а более точно — к *выборочному распределению*, которое мы обсуждали в I.5. Точер [Tocher, 1963, p.100] обсуждает оценивание ожидаемого *размаха*. Как известно, размах R выборки из m наблюдений для x_i ($i = 1, \dots, m$) определяется выражением

$$R = \max_i x_i - \min_i x_i. \quad (36)$$

Точер [Tocher, 1963, p.103] в качестве возможной стратифицирующей переменной указывает на выборочное среднее, определяемое следующим выражением:

$$\bar{x} = \sum_{i=1}^m x_i / m. \quad (37)$$

В экспериментах Монте-Карло мы имеем m выборочных значений x с некоторым заданным распределением. Эти m наблюдений дают оценку размаха, а также оценку стратифицирующей переменной — выборочной средней. Повторение такой выборки n раз дает оценки R_1, \dots, R_n . Далее, образуем два слоя S_1 и S_2 , например, такие, что x_j (и, следовательно, R_j) ($j = 1, \dots, n$) попадает в слой S_1 , если $\bar{x}_j \leq \mu$, и \bar{x}_j попадает в слой S_2 , если $\bar{x}_j > \mu$, где μ — известное среднее совокупности, из которой извлекается выборка. Если известно, что плотность распределения, из которой производится выборка, симметрична, то мы знаем веса слоев p_h ($h=1, 2$). Как утверждает Точер [Tocher, 1963, p.103], если нет скачка вероятности при $x = \mu$, то из симметричности функции плотности распределения \bar{x} следует:

$$p_1 = P(\bar{x}_j \in S_1) = P(\bar{x}_j \leq \mu) = \frac{1}{2} \quad (38)$$

и

$$p_2 = P(\bar{x}_j \in S_2) = P(\bar{x}_j > \mu) = \frac{1}{2}. \quad (39)$$

Итак, стратифицированная оценка размаха при m наблюдениях, взятых из некоторого симметричного распределения, определяется выражением

$$\begin{aligned} \bar{R} &= \sum_{k=1}^2 p_k \bar{R}_k = \\ &= \frac{1}{2} \sum_{g=1}^{n_1} R_{1g}/n_1 + \frac{1}{2} \sum_{h=1}^{n_2} R_{2h}/n_2, \end{aligned} \quad (40)$$

где R_{1g} и R_{2h} обозначают области средних значений соответственно для слоев 1 и 2. То, что выборочное среднее \bar{x} можно взять в качестве стратифицирующей переменной, так как она *коррелирует* с переменной R , видно из следующего. Если некоторое значение x_i среди m наблюдений принимает очень большое значение, то обе величины \bar{x} и R увеличиваются сравнительно с их истинными средними, т. е. \bar{x} и R положительно коррелируют между собой. Точер [Tocher, 1963, p.104—105] обсуждает и другие переменные, по которым можно проводить стратификацию для оценки размаха.

В [Hillier and Lieberman, 1968, p. 455—458] также рассматривается стратификация в имитационном моделировании. Однако на самом деле в приводимом там примере методом Монте-Карло оценивается среднее экспоненциального распределения. Используемая в этой работе стратификация для вычисления интеграла описана и в [Hammersley and Handscomb, 1964, p.55—57], [Gaver, 1969], [Molenaar, 1968, p.120], [Newman and Odell, 1971, p. 57—62] (см. также наше примечание 13 к данной главе).

III.2.4. Стратифицированная выборка в имитационном моделировании

Стратифицирующей переменной в исследованиях Монте-Карло для размаха служит выборочное среднее. Обе статистики коррелируют, если они генерируются из одной и той же последовательности случайных чисел. Так же можно генерировать два зависимых отклика в имитационном моделировании. В качестве примера рассмотрим задачу эксплуатации автобусов из приложения I.3. Предположим, что функцию распределения вероятностей отклика для стратегии β можно определить аналитически, а среднее для стратегии γ надо оценить с помощью моделирования. Тогда мы можем моделировать оба варианта β и γ , даже если β предполагается *известным*. Если обе стратегии моделируются с помощью одной и той же последовательности случайных чисел, то оба отклика коррелируют между собой. (Мы вернемся к обсуждению этой корреляции в III.7, где рассмотрим работу одной и той же последовательности случайных чисел для моделирования нескольких систем.) Поскольку вариант β имеет аналитическое решение, мы можем вычислить веса слоев p_k и воспользоваться откликами стратегии β в качестве стратифицирующей переменной. Эта процедура имеет два очевидных недостатка:

1) должна быть подходящая *система* с известным *аналитическим* решением;

2) нужно *дополнительное время* для программирования и прогонов этой системы, дающей стратифицирующую переменную (в примере из III.2.3 для метода Монте-Карло время, необходимое для вычисления стратифицирующей переменной x , мало по сравнению со временем, которое требуется на генерирование x_i и вычисления¹² R).

В [Ehrenfeld and Ben-Tuvia, 1962, p.261—263, 264—265] предлагается другая процедура стратификации для простых систем массового обслуживания. К сожалению, она неприменима для общего случая в моделировании (ср. с [Clark, 1959], [Price, 1972] и [Ringer, 1965]). Итак, стратификация может применяться в исследованиях по методу Монте-Карло, однако мы согласны с Моем [Моу, 1965, p.85), что все процедуры, предложенные для стратификации в имитационных моделях, не обладают свойством общности. Не удивительно, что Точер [Tocher, 1963, p.171] сомневается, может ли вообще стратификация применяться при моделировании сложных систем. Однако Мой избрал процедуру, которая может найти широкое применение в моделировании. В этой процедуре стратифицирующая переменная не служит откликом соответствующей системы, но определяется прямо из последовательности случайных чисел в моделировании интересующей нас системы¹³. Далее мы изучим эту процедуру.

Мой [Моу, 1965, p.90; 1966, p.17] предлагает *стратифицирующую* переменную определять таким образом:

$$y = \sum_{i=1}^m t_i, \quad (41)$$

где

$$\begin{aligned} t_i &= 1, \text{ если } r_i \geq c, \\ &= 0, \text{ если } r_i < c, \end{aligned} \quad (42)$$

и где c — константа, удовлетворяющая неравенству $0 < c < 1$, а r_i представляет собой i -е случайное число ($i = 1, \dots, m$) в последовательности случайных чисел для генерирования одного значения отклика x . Заметим, что у Моя общее число случайных чисел в одном имитационном опыте предполагается известной константой. Последнее условие, например, встречается при моделировании одноканальной системы массового обслуживания для числа требований, равного m_1 . Каждому требованию соответствуют два случайных числа. Одно число соответствует моменту поступления, другое — времени обслуживания. Следовательно, в нашем примере $m = 2m_1$. Мы еще вернемся к условию о известной константе m позднее.

Из (16) в III.2.1 вытекает, что хорошая стратифицирующая переменная y должна коррелировать с интересующей нас переменной x . Как замечает Мой [Мой, 1965, р.90; 1966, р.5, 17—18], такая корреляция возникает, если имитационную модель конструировать так, чтобы плотность очереди возрастала с ростом значения случайного числа r_i . Тогда в приведенном примере простой системы массового обслуживания мы можем генерировать экспоненциальное время прибытия, назовем его v , и экспоненциальное время обслуживания, назовем его w , с помощью (43) и (44) соответственно (эти уравнения основываются на выражениях (18) и (20) из I.6.4):

$$v_i = -\lambda_v^{-1} \ln(r_{2i-1}) = \lambda_v^{-1} |\ln(r_{2i-1})| \quad (43)$$

и

$$w_i = -\lambda_w^{-1} \ln(1 - r_{2i}) = \lambda_w^{-1} |\ln(1 - r_{2i})| \quad (i = 1, 2, \dots, m_1), \quad (44)$$

где λ_v и λ_w — параметры экспоненциального распределения v и w соответственно. Это распределение было определено в (15) из I.6.4. Заметим далее, что случайное число r удовлетворяет условию $0 \leq r \leq 1$ и, следовательно, $\ln(r)$ и $\ln(1 - r)$ отрицательны. Итак, большое значение случайной величины означает, что время между двумя последовательными поступлениями v мало или что время обслуживания w велико. Это ведет к большому скоплению очереди, т. е. к долгому времени ожидания, которое и есть интересующая нас переменная x . Окончательно, из определения стратифицирующей переменной в (41) и (42) следует, что большое значение r приводит к большому значению стратифицирующей переменной y . Итак, x и y — положительно коррелированные переменные. Легко видеть, что наблюдается положительная корреляция между стратифицирующей переменной в (41) и откликом системы эксплуатации, описанной в приложении I.3. Мы считаем, что в моделях систем других типов такая корреляция тоже может быть, так что переменная y в выражении (41) может служить стратифицирующей переменной в моделировании многих систем.

Другое условие, которое должно выполняться для стратифицирующей переменной — известные веса слов r_k , определенные в (4).

Эти веса можно вычислить следующим образом. Из определения \mathbf{y} в (41) и (42) мы видим, что \mathbf{y} — биномиально распределенная переменная с «вероятностью успеха», равной $(1 - c)$. Для

$$P(t_i = 1) = P(r_i \geq c) = 1 - c \quad (45)$$

на основании [Fisz, 1967, p.136] получаем, что функция вероятности равна:

$$P(\mathbf{y} = j) = \binom{m}{j} (1 - c)^j c^{m-j} \quad (j = 0, 1, \dots, m). \quad (46)$$

Мы можем образовать K слоев, разделив область значений \mathbf{y} на K непересекающихся и исчерпывающих эту область интервалов. Например, если $m = 500$ и $k = 5$, то мы образуем следующие слои ${}^{14}S_k$:

$$[0, 100), [100, 200), [200, 300), [300, 400), [400, 500]. \quad (47)$$

Вес слоя (например, p_2) задается таким образом:

$$p_2 = P(\mathbf{y} \in S_2) = P(100 \leq \mathbf{y} < 200) = \sum_{k=100}^{199} P_k(\mathbf{y} = h), \quad (48)$$

где $P(\mathbf{y} = h)$ дается выражением (46).

Мой применяет принципы *пропорциональной* стратифицированной выборки. (Позже мы рассмотрим стратификацию *после* выборки со стратифицирующей переменной *Моя*). Из (14) следует, что при пропорциональной стратификации число наблюдений в слое определяется формулой

$$n_k = p_k n, \quad (49)$$

где n — общее число наблюдений, т. е. общее число имитационных прогонов опытов (каждый длиной m). Заметим, что n — известная константа (в III.8.5 и главе V мы вернемся к этому предположению, так как число опытов может зависеть от имеющегося машинного времени и желаемой надежности). Зная n и вычисляя p_k из (48), можно получить значения n_k с помощью (49).

Вычислив n_k — число наблюдений в k -м слое, мы далее должны выбрать n_k наблюдений из слоя *случайным образом*. Наблюдение принадлежит слою k , если в результате имитации получаем величину стратифицирующей переменной \mathbf{y} из этого слоя. Мой предлагает следующую процедуру.

1. *Выбираем* случайным образом \mathbf{y} из слоя k способом, описанным далее. Нам понадобится определение (50) условной вероятности, данное в [Fisz, 1967, p.20].

$$P(A|B) = \frac{P(AB)}{P(B)} \quad \text{для } P(B) > 0, \quad (50)$$

т. е. вероятность события A при условии, что произошло событие B , равна вероятности произведения событий A и B , разделенной на вероятность события B . Если мы хотим выбрать \mathbf{y} случайным образом

из слоя 2, то мы должны выбрать y на основе функции распределения вероятностей:

$$P(y = g | y \in S_2) = \frac{P(y = g \text{ и } y \in S_2)}{P(y \in S_2)} = \frac{P(y = g)}{P(y \in S_2)} \quad (g = 100, 101, \dots, 199). \quad (51)$$

Подставляя (46) и (48) в (51), получаем

$$P(y = g | y \in S_2) = \frac{\binom{m}{g} (1-c)^g c^{m-g}}{\sum_{h=100}^{199} \binom{m}{h} (1-c)^h c^{m-h}}. \quad (52)$$

Значения из (52) y генерируются с помощью метода обращения, определенного выражением (33) в I.6.4. На рис. 9 изображена блок-схема машинной программы Моя [Мой, 1965, р.91, fig. 6—1].

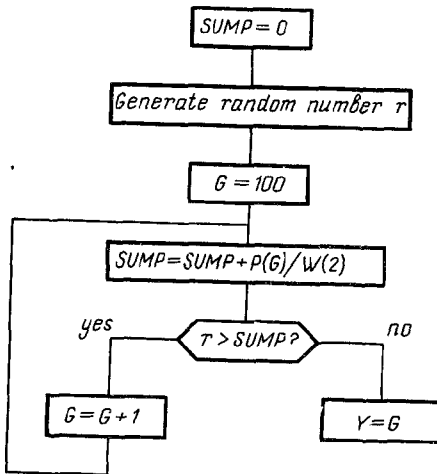


Рис 9. Генерирование значения стратифицирующей переменной y для слоя 2

как это определено в (46), а $W(2)$ — вес слоя p_2 , данный в (48).

2. После того как получено выборочное значение стратифицирующей переменной y , мы должны генерировать наблюдение, т. е. имитационный опыт, который соответствует этому значению y . Предположим, что с помощью (52) или ее эквивалента на рис. 9 мы получили $y = 175$ из слоя 2. Из определения y в (41) следует, что 175 из 500 (= m) случайных чисел должны быть не меньше чем c . Есть много, а именно $\binom{500}{175}$, возможных векторов, содержащих 500 случайных чисел, 175 из кото-

рых не меньше c . Мы должны выбрать один из таких векторов случайным способом. На рис. 10 приведена блок-схема такой выборки¹⁵ [Мой, 1965, р.93, fig. 6—2]. Заметим, что мы обращаемся к рис. 9 всякий раз, когда требуется случайное число, т. е. 500 раз в течение опыта. В начале каждого опыта $M = 500$ и Y равно значению, которое было выбрано с помощью процедуры, изображенной на рис. 9. (Мы производим выборку n значений для стратифицирующей переменной по блок-схеме рис. 9, повторяя процедуру n раз.) Заметим, что на рис. 10 изображена процедура *выборки без возвращения*, так как вероятность выбора некоторого события меняется в ходе имитационного эксперимента. Для процесса, изображенного на рис. 10, вероятность того, что случайное число в блоке 1 меньше, чем Y/M , меняется в течение опыта,

так как меняются Y и M . В результате мы получаем, что количество случайных чисел, которое не меньше c , в точности равно значению, отбираемому для стратифицирующей переменной y . В приложении III.6 мы дадим простой пример, демонстрирующий рациональность процедуры, изображенной на рис. 10.

Мой [Моу, 1965, р.97,103, 104; 1966, р.29, 32] применял пропорционально стратифицированную выборку к одноканальным системам массового обслуживания и к сложной системе товарообмена в порту (он

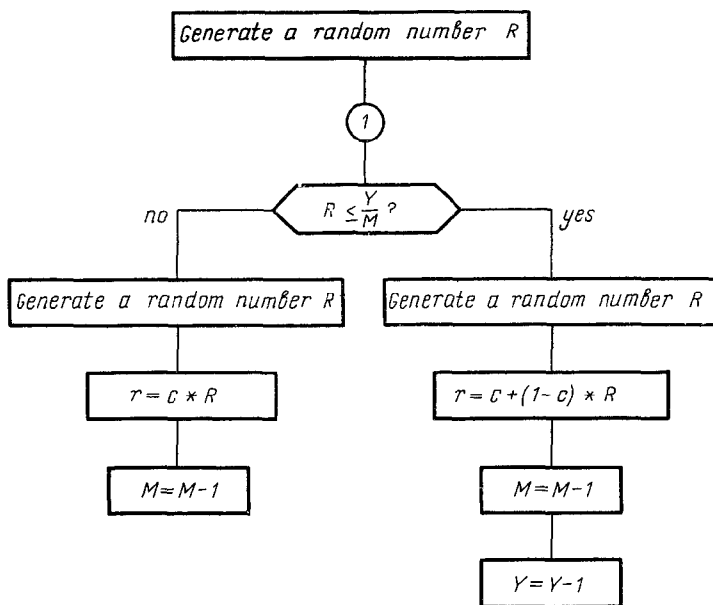


Рис. 10. Генерирование случайного числа r , удовлетворяющего желаемому значению стратифицирующей переменной

положил константу c в (42) равной $1/2$). Для системы с одним пунктом обслуживания получилось среднее снижение дисперсии на 20%. Однако, принимая во внимание дополнительное машинное время, необходимое для стратификации, Мой [Моу, 1965, р.101] приходит к выводу, что для этого типа систем стратификация неэффективна. Для сложной системы товарообмена в порту дисперсия даже повысилась! Мой [Моу, 1965, р.103—104] также применял стратификацию, «зависящую от системы», т. е. в определении стратификационной переменной y случайные числа используются только для генерирования специального типа событий (например, времени между поступлениями). «Зависящий от системы» вариант дает лучший результат для системы товарообмена в порту, но пока худшие результаты по сравнению с «независящим от системы» вариантом при моделировании простой системы массового обслуживания. Мы хотели бы сделать два замечания о резуль-

татах Моя, прежде чем решить, эффективна ли стратификация в имитационных исследованиях ¹⁶.

1. В системе товарообмена в порту порядка m число случайных чисел при имитации может меняться от опыта к опыту. Заменяем m в определении стратифицирующей переменной, данном в (41), на m , где m обозначает случайное число случайных чисел, необходимых для опыта. В [Моу, 1965, р.101—102; 1966, р.18] содержится попытка решить эту проблему, предположив, что длина вектора случайных чисел равна длине наибольшего из всех векторов, нужных для имитации. Затем для генерации наблюдений используется ровно столько случайных чисел, сколько надо, остальное отбрасывается. Однако, по нашему мнению, эта процедура не корректна и мы считаем, что стратификация невозможна, если число случайных чисел меняется от опыта к опыту. Поскольку вывод формулы весов для слоев (48) основан на биномиальном распределении, где m — константа, как видно из (46), выбор значений стратифицирующей переменной основывается также на постоянном значении m , как видно из (52). Мы продемонстрируем наши возражения далее на примере. Допустим, что наибольшей разумной величиной m является 750. Разделим область стратифицирующей переменной на следующие пять слоев S_k :

$$\{0, 150\}, \{150, 300\}, \{300, 450\}, \{450, 600\}, \{600, 750\}. \quad (53)$$

Допустим, что по схеме, изображенной на рис. 9, мы выбрали значение стратифицирующей переменной из последнего слоя, равное 723. Далее, согласно рис.10 нам нужно получить вектор случайных чисел, содержащий 750 случайных чисел, 723 из которых не меньше s . Однако может случиться, что после получения, скажем, 500-го случайного числа имитационный процесс остановится. Величина y будет равна приблизительно:

$$723 \times (500/750) = 482. \quad (54)$$

В этом случае имитационный опыт относится к 4-му слою, а не к 5-му! Здесь выглядит разумным поместить отклик x в 4-й слой и воспользоваться весом

$$p_4 = P(y \in S_4 | m = 500). \quad (55)$$

Но фактически x порождается при величине y , равной 723, которая выбрана согласно (56), что представляет собой аналог выражения (52):

$$P(y = g | y \in S_5 \text{ и } m = 750) = \frac{P(y = g)}{P(y \in S_5)} \quad (g = 600, 601, \dots, 750) \quad (56)$$

Мой [Моу, 1965 и 1966] не показывает, как он выходит из этого положения, т. е. как определить, к какому слою отнести отклик и вес какого слоя надо брать. Нужны дополнительные исследования, чтобы выявить, вызваны ли плохие результаты Моя при моделировании системы товарообмена в порту *некорректной процедурой стратификации* или *сложностью системы* (которая может накладывать возмущения на желаемую высокую корреляцию стратифицирующей переменной и отклика). Можно предположить, что в сложной системе с *постоянным*

m мы получим понижение дисперсии, как и в случае простых систем массового обслуживания с постоянным m . (Заметим, что в сложных системах постоянное число m можно легко реализовать в моделировании *установившихся состояний*. Согласно П.8 длину прогона экспериментатор может зафиксировать и, следовательно, мы можем остановить процесс, когда число случайных чисел станет равным m . С другой стороны, установившиеся системы могут меньше подходить для стратификации, так как вероятно увеличение объема выборки из-за продолжения прогона, а не из-за его повторения, и это может привести к нарушению требования, что слой не должен быть пуст.)

2. Мы видели, что из-за дополнительного машинного времени Мой отказывается от пропорционально стратифицированной выборки даже для простых систем массового обслуживания с постоянным m . В качестве альтернативы мы предлагаем обсудить процедуру стратификации, а именно *стратификацию после выборки*. В соответствии с разъяснениями в III.2.2 здесь мы не фиксируем число наблюдений в каждом слое, а традиционным способом генерируем имитационные опыты с использованием в каждом m случайных чисел. После этого согласно определению (41) стратифицирующей переменной y выясняем, к какому слою принадлежит каждый из опытов, и вычисляем стратифицированную оценку (17) или, более точно, одну из оценок (22) или (29). В III.2.2 мы видели, что дисперсия или средняя квадратичная ошибка этой оценки больше, чем в случае пропорционально стратифицированной оценки. Однако для получения стратифицирующей переменной по схеме, изображенной на рис. 9, и для получения необходимого вектора случайных чисел по схеме на рис. 10 не требуется *дополнительного машинного времени*. В III.2.2 мы показали, что при стратификации после выборки в отличие от случая с фиксированным числом наблюдений в слоях определение *доверительных пределов* для среднего совокупности—более трудное дело. Обсуждение уравнения (12) показывает, что в последнем случае точные значения доверительных пределов установить невозможно. Определение же приближенных доверительных пределов при стратификации требует дополнительных исследований.

III.3. СЕЛЕКТИВНАЯ ВЫБОРКА ИЛИ МЕТОД ФИКСИРОВАННОЙ ПОСЛЕДОВАТЕЛЬНОСТИ

III.3.1. Описание метода

Мы рассмотрим метод понижения дисперсии (МПД), предложенный Бреннером [Brenner, 1963]. Мы уже видели, что *стратификация* (кроме стратификации после выборки) изменяет выборочный процесс таким образом, что число имитационных опытов со значением стратификационной переменной y в определенном слое k становится постоянным, а именно равным n_k . Следовательно, число имитационных опытов определенного типа событий тоже постоянно, что реализуется выборкой без возвращения согласно рис. 10. Бреннер фиксирует тип событий *внутри* имитационного опыта, так что у него все n опытов одного и того

же типа; при этом число событий определенного вида точно равно теоретически ожидаемому числу. Это будет ясно из последующего описания процедуры.

Бреннер неявно предполагает, что существует только *один вид случайной входной* переменной. В рассмотренных Бреннером [Brenner, 1963, p.293—294] системах входная переменная — это требование на товар со склада. Эта входная переменная, скажем y , *дискретна*, и мы можем, следовательно, различать M возможных значений, которые она принимает ¹⁷:

$$P(y = y_i) = q_i \quad (i = 1, 2, \dots, M). \quad (57)$$

Пусть нам нужно m значений y в имитационном опыте. Обозначим число раз, когда y принимает значение y_i , через v_i . Очевидно, что математическое ожидание v_i определяется формулой

$$E(v_i) = q_i m. \quad (58)$$

Бреннер *фиксирует* число раз, когда некоторое значение y_i попадает в выборку при «проигрывании» опыта, т. е. выполняется равенство

$$v_i = v_i. \quad (59)$$

Он хочет сделать v_i равным *математическому ожиданию* величины v_i , т. е.

$$v_i = q_i m, \quad (60)$$

где

$$\sum_1^M v_i = m \sum_1^M q_i = m. \quad (61)$$

В общем виде, однако, $q_i m$ не будет *целым числом*, в то время как v_i должно быть целым. Следовательно, Бреннер [Brenner, 1963, p.292, 925] присваивает $q_i m$ целое значение, такое, что сумма квадратов отклонений между «теоретическими» величинами $q_i m$ и целыми v_i минимизируется, т. е. он минимизирует z , определенное в (62):

$$z = \sum_{i=1}^M (v_i - q_i m)^2 \quad (62)$$

при условии (61). Бреннер приводит алгоритм для нахождения v_i в выражении (62). Этот алгоритм дает значение v_i , равное целому, которое наиболее близко к $q_i m$. Если $\sum v_i > m$, то v_i , для которых наименьшие дробные части превышают 0,5, уменьшаются на единицу до тех пор, пока не будет выполняться равенство $\sum v_i = m$. Аналогично, если $\sum v_i < m$, то те v_i , для которых наибольшие дробные части меньше чем 0,5, увеличиваются на единицу до выполнения равенства $\sum v_i = m$.

Обозначим v_i , которое минимизирует z в равенстве (62) при условии (61), через v_i^* . Итак, число появлений y_i теперь фиксировано, однако *порядок* появлений не задан. Выборка производится так, что появление каждого порядка, удовлетворяющего $v_i = v_i^*$, имеет одну и ту же вероятность. Технически это реализуется с помощью *выборки без возвра-*

щения. Этот метод мы упоминали при обсуждении рис. 10 в III.2.4, где была описана процедура стратификации Мойя. Метод предполагает, что вероятность выбора значения y_i при t ($t = 1, 2, \dots, m$) равна

$$(v_i^* - g_i)/(m - t + 1), \quad (63)$$

где g_i обозначает число раз, которое y_i было отобрано в имитационном опыте. Бреннер [Brenner, 1963, p.295—296] дает программу на Фортране для вычисления v_i^* и генерирования последовательности m значений y , удовлетворяющих равенству $v_i = v_i^*$.

III.3.2. Комментарий

После описания метода селективной выборки Бреннера дадим некоторый комментарий. Прежде всего, по нашему мнению, процедура дает *смещенные* результаты: 1) в процедуре селективного выборочного метода используется не теоретическое значение математического ожидания v_i , равное $q_i m$, а значение $v_i = v_i^*$, которое представляет собой целое число, близкое к $q_i m$, но не обязательно точно равное $q_i m$; 2) процедура Бреннера оценивает математическое ожидание отклика по формуле $v_i = v_i^*$ ($i = 1, \dots, M$). Согласно [Fisz, 1967, p.84] можно написать, что математическое ожидание отклика x определяется следующим выражением:

$$\begin{aligned} E(x) &= \sum_v [E(x | v_1 = v_1, v_2 = v_2, \dots, v_M = v_M)] = \\ &= \sum \sum \dots \sum [E(x | v_1 = v_1, v_2 = v_2, \dots, v_M = v_M) \times P(v_1 = \\ &= v_1, v_2 = v_2, \dots, v_M = v_M)], \end{aligned} \quad (64)$$

где сумма по всем возможным значениям v_1, v_2, \dots, v_M , т. е. по всем комбинациям целых v_1, v_2, \dots, v_M , удовлетворяет условиям

$$v_i \geq 0 \text{ и } \sum_1^M v_i = m.$$

Бреннер, однако, оценивает¹⁸

$$E(x | v_1 = v_1^*, v_2 = v_2^* \dots, v_M = v_M^*). \quad (65)$$

В общем случае условное математическое ожидание (65) не совпадает со значением, вычисленным по формуле (64), даже если $v_i^* = E(v_i)$. Далее мы покажем, что метод селективной выборки дает те же результаты, что и *метод фиксированной последовательности*, предложенный Эренфельдом и Бен-Тувия [Ehrenfeld and Ben-Tuvia, 1962, p.258, 263, 274—275]. Эти авторы соглашались с тем, что их метод дает смещенную оценку, и они подтверждают это примером, приведенным в табл. 1.

Конечно, возможно, что понижение дисперсии компенсирует смещение, так что *средняя квадратичная ошибка* (25) уменьшится. В простых системах массового обслуживания, изучаемых в [Ehrenfeld and Ben-Tuvia, 1963, p.259, 263—264, 274—275], смещение было небольшим, и авторы вычисляли дисперсию, игнорируя смещение. Оцененное снижение дисперсии при этом способе оказалось равным 56 % в системе с одним каналом и 71 % в системе с двумя каналами. В простой систе-

ме запасов, исследуемой в [Brenner, 1963, p.293], селективная выборка также дала хорошие результаты. (Бреннер не измерял среднюю квадратичную ошибку, но он измерял «потери», т. е. разность между истинной ценой запасов при оптимальном сочетании факторов, оцененных с помощью имитационного моделирования, и ценой при точном оптимуме, определенном с помощью аналитического решения.) Ни Бреннер, ни Эренфельд и Бен-Тувия не рассчитывали дополнительное *время*, необходимое для применения МПД.

В начале этого параграфа мы указывали, что Бреннер ограничивается только системами с одним типом входных переменных, которые к тому же предполагаются дискретными, а число входных величин m представляет собой известную константу. Очевидно, процедуру Бреннера можно обобщить на системы с *большим числом типов входных переменных*. В этом случае мы к каждому из типов входов можем применить (62) и (63). В [Ehrenfeld and Ben-Tuvia, 1962, p.267] мы можем найти примеры с двумя типами входных переменных. Далее, если взять систему с *непрерывным входом*, то мы можем образовать M классов, как в (57), и подставить в (57) средние классов вместо y_i . Такая процедура была применена Эренфельдом и Бен-Тувия в системе массового обслуживания, где время между заявками и время обслуживания были непрерывными входными переменными. Они взяли M классов, так что в каждом из них оказалось по $1/M$ наблюдений, а в уравнении (57) они воспользовались для y_i медианой класса. Очевидно, что замена непрерывной входной переменной дискретной создает некоторое смещение помимо того смещения, которое, как указано в [Ehrenfeld and Ben-Tuvia, 1962, p.259, 264], возникает при замене (64) на (65). То, что m представляет собой случайное число значений входа, — одна из наиболее серьезных проблем МПД. Эта проблема не упоминается ни у Бреннера, ни у Эренфельда и Бен-Тувия. Мы можем заменить m некоторой оценкой m из (62) и (63). К сожалению, это вносит дополнительные возмущения и v_i^* — желаемое число появлений y_i — может не плучиться.

В заключение укажем, что в [Orcutt et al., 1961, p.33—34, 215, 294—299, 373—374] также исследуется выборка типа селективной (не названная так) и число событий определенного типа, генерируемых в имитационном эксперименте, также зависит от ожидаемого числа событий (либо, если моделируются прошлые события, это число должно совпадать с числом событий в прошлом). Этот метод был применен как *сервомеханизм*, но без выборки. Здесь регулируются первоначальные вероятности с помощью мультипликативного множителя, который измеряет, насколько действительное число событий, генерируемых для определенного имитируемого периода времени, отклоняется от ожидаемого (либо прошлого) числа событий. Для детального ознакомления с этим вопросом мы отсылаем к [Orcutt et al., 1961, p.296—299]. Заметим, что авторы указанной работы считают, что этот механизм не дает смещения, но мы с этим не согласны и уравнения (64) и (65) в предыдущем обсуждении подтверждают это. В [Meier et al, 1969, p.275—277] кратко обсуждаются некоторые другие исследования, использующие «следящие устройства».

III.4. КОНТРОЛЬНЫЕ ВЕЛИЧИНЫ ИЛИ МЕТОД РЕГРЕССИОННОЙ ВЫБОРКИ

III.4.1. Основные понятия метода контрольных величин

При применении МПД, метода селективной выборки, рассматривается v_i — число раз, когда входная переменная y принимает значение y_i . Метод контрольных величин — более общий инструмент, так как он предполагает рассмотрение только *среднего* значения входной случайной величины (в III.4.2 мы обсудим более общее определение этого метода). Среднее значение входной переменной y в имитационном опыте, где мы используем m значений входной переменной, определяется выражением

$$\bar{y} = \sum_{i=1}^m y_i / m \quad (66)$$

(заметим, что наше определение (66) с дискретной случайной величиной m включает также случай $P(m = m) = 1$, т. е. случай, когда m — константа). В отличие от метода селективной выборки метод контрольной величины не *закрепляет* \bar{y} на значении, равном теоретическому, скажем η , где η определяется выражением

$$\eta = E(\bar{y}). \quad (67)$$

Итак, выборочный процесс сам по себе не управляется. Вместе этого в конце каждого опыта полученное значение \bar{y} сравнивается с *истинной* или ожидаемой величиной η . Эта величина η предполагается известной. Поскольку y_i из (66) порождается известным распределением входной переменной, которой может быть, например, случайное число заявок или время обслуживания, мы можем вычислить математическое ожидание y_i , которое равно математическому ожиданию \bar{y} . Если \bar{y} отличается от своего теоретического значения η , то МПД позволяет подправить отклонение с помощью формулы

$$xc = x - a(\bar{y} - \eta), \quad (68)$$

где x — отклик (например, общий доход или среднее время ожидания в очереди) одного имитационного опыта, xc — оценка в случае контрольной величины y , a — константа в течение этого опыта. Очевидно, что оценка xc будет *несмещенной*. Из (67) и (68) следует, что

$$E(xc) = E(x) - aE(\bar{y} - \eta) = E(x) - a\{E(\bar{y}) - \eta\} = E(x). \quad (69)$$

Ее дисперсия определяется по формуле

$$\begin{aligned} \text{var}(xc) &= \text{var}(x) + a^2 \text{var}(\bar{y}) - 2a \text{cov}(x, \bar{y}) = \\ &= \text{var}(x) + a^2 \text{var}(\bar{y}) - 2ap(x, \bar{y}) \sqrt{\text{var}(x) \text{var}(\bar{y})}. \end{aligned} \quad (70)$$

Из выражения (70) видно, что дисперсия новой оценки меньше, чем $\text{var}(x)$, дисперсии грубой оценки, если:

а) при положительной константе a (71а) удовлетворяется неравенство

$$a < 2\rho(\bar{x}, \bar{y})\sigma(x)/\sigma(\bar{y}), \quad (71a)$$

б) при отрицательном a (71б) удовлетворяется неравенство

$$a > 2\rho(\bar{x}, \bar{y})\sigma(x)/\sigma(\bar{y}). \quad (71б)$$

Следовательно, в случае, когда x и \bar{y} положительно коррелируют, дисперсия понижается, если константа a положительна и меньше правой части неравенства (71а); в случае отрицательной корреляции константа a должна быть отрицательной и по абсолютной величине меньше абсолютной величины правой части неравенства (71б). Заметим, что знак коэффициента понятен из интуитивных соображений. Если \bar{y} и x положительно коррелируют, тогда значение величины \bar{y} , большее чем ее математическое ожидание η , ведет к тому, что и величина x будет больше своего математического ожидания, скажем μ . Далее, если значение $(\bar{y} - \eta)$ будет положительным, то и $a(\bar{y} - \eta)$ будет положительным, когда положительно a . Итак, $a(\bar{y} - \eta)$ должно быть положительным, если x больше своего математического ожидания μ , поэтому в данном случае мы вычитаем $a(\bar{y} - \eta)$ из x в (68). *Оптимальный* выбор константы получается из (70) после дифференцирования и решения относительно a : оптимальное значение a , скажем a_0 , определяется следующим выражением:

$$a_0 = \frac{\text{cov}(x, \bar{y})}{\text{var}(\bar{y})} = \rho(\bar{x}, \bar{y})\sigma(x)/\sigma(\bar{y}). \quad (72)$$

Это значение оптимальной константы, подставленное в (70), дает минимальную дисперсию (т. е. дисперсию x_{c_0} — оптимальную оценку контрольной величины):

$$\text{var}(x_{c_0}) = \text{var}(x)\{1 - \rho^2(\bar{x}, \bar{y})\}. \quad (73)$$

Укажем, что предельные значения a , задаваемые неравенствами (71а) и (71б), и оптимальное значение a , данное в (72), зависят от корреляции между выходом x и входом \bar{y} . Мы не знаем этой корреляции и, возможно, нам понадобится ее оценить. В III.4.3 будет показано, как можно оценить оптимальное значение коэффициента a и что следует из такого оценивания. Сначала, однако, обсудим более общие определения метода контрольной величины.

III.4.2. Типы контрольных величин

Здесь мы опишем несколько типов контрольных величин. Варианты *оптимальных* коэффициентов этих новых типов и их *оценивание* представлены в следующем параграфе.

1. Если имеется *несколько разных входных переменных*, например время обслуживания и время между двумя поступлениями, то оценка

множественной контрольной величины, очевидно, выразится в виде

$$xct = x - \sum_{k=1}^K a_k (\bar{y}_k - \eta_k), \quad (74)$$

где

$$\bar{y}_k = \sum_{t=1}^{m_k} y_{kt} / m_k, \quad (75)$$

$$\eta_k = E(\bar{y}_k) \quad (76)$$

и где y_{kt} есть t -наблюдение k -типа входной переменной в имитационном прогоне. Легко видеть, что оценка не является смещенной. Ее дисперсия будет зависеть от выбора коэффициента a_k , который мы обсудим в III.4.3.

2. Мы обобщили метод контрольной величины на случай нескольких типов входной переменной. Следующее обобщение вытекает из определения контрольной величины y в (68) не как *среднего* (66), а как *некоторой функции* от значений входных переменных y_i , как в (77):

$$z = g(y_1, y_2, \dots, y_m). \quad (77)$$

Следовательно, (77) определяет контрольную величину z как некоторую функцию индивидуальных наблюдений случайного входа y_i ; мы изучим общий случай, когда m , число индивидуальных наблюдений, может быть случайным. Необходимое условие, накладываемое на функцию g , состоит в том, что мы должны иметь возможность вычислить *математическое ожидание* новой контрольной величины. Обозначим его через ξ , т. е.

$$E(z) = E[g(y_1, y_2, \dots, y_m)] = \xi. \quad (78)$$

Аналогично (68), (69) и (70) оценка xsg обобщенной контрольной величины, ее средняя и дисперсия даны выражениями (79), (80) и (81) соответственно:

$$xsg = x - z + E(z) = x - g(y_1, y_2, \dots, y_m) + \xi \quad (79)$$

при

$$E(xsg) = E(x) - E(z) + E(\xi) = E(x) - \xi + \xi = E(x) \quad (80)$$

и

$$\text{var}(xsg) = \text{var}(x) + \text{var}(z) - 2\text{cov}(x, z). \quad (81)$$

Из (81) следует второе условие, накладываемое на функцию g . Функция g выбирается так, чтобы контрольная величина z имела «сильную» положительную *корреляцию* с откликом x . Это условие заменяет условия (71a) и (71b) для более частного случая контрольной величины (68). Мой [Мой, 1965, р.60—61; 1966, р.9—10] с успехом экспериментировал с функцией

$$z_1 = g_1(y_1, y_2, \dots, y_m) = b_1 + b_2 \sum_{t=1}^m y_t, \quad (82)$$

которая после подстановки в (79) дает следующую оценку контрольной величины:

$$\begin{aligned} xc1 &= x - \left(b_1 + b_2 \sum_1^m y_t \right) - E \left(b_1 + b_2 \sum_1^m y_t \right) = \\ &= x - b_2 \left\{ \sum_1^m y_t - E \left(\sum_1^m y_t \right) \right\}. \end{aligned} \quad (83)$$

Эта оценка напоминает исходную оценку простой контрольной величины, данную в (68). Мой пользуется также выражением более высокой степени:

$$z_2 = g_2(y_1, y_2, \dots, y_m) = b_1 + b_2 \sum_1^m y_t + b_3 \left(\sum_1^m y_t \right)^2. \quad (84)$$

Чтобы применять выражение (84), мы должны иметь возможность вычислить его *математическое ожидание*. Обсудим более общее определение контрольной величины, а именно контрольной величины со случайным m (не обсуждаемой Моем). Оценка контрольной величины (84) со случайным m дается выражением

$$\begin{aligned} xc2 &= x - \left\{ b_1 + b_2 \sum_1^m y_t + b_3 \left(\sum_1^m y_t \right)^2 \right\} + E \left[b_1 + b_2 \sum_1^m y_t + \right. \\ &+ \left. b_3 \left(\sum_1^m y_t \right)^2 \right] = x - b_2 \left\{ \sum_1^m y_t - E \left(\sum_1^m y_t \right) \right\} - b_3 \left\{ \left(\sum_1^m y_t \right)^2 - \right. \\ &\quad \left. - E \left[\left(\sum_1^m y_t \right)^2 \right] \right\}. \end{aligned} \quad (85)$$

В приложении III.7 мы получим

$$E \left[\sum_1^m y_t \right] = \eta_1 E(m) \quad (86)$$

и

$$E \left[\left(\sum_1^m y_t \right)^2 \right] = \eta_1^2 E(m^2) + \sigma_1^2 E(m), \quad (87)$$

где моменты $\eta_1 = E(y_t)$ и $\sigma_1^2 = \text{var}(y_t)$ можно вычислить, исходя из распределения входа y_t . Если можно вычислить значения $E(m)$ и $E(m^2)$, подставив их в (86) и (87), получим формулу (85). С другой стороны, мы можем применить следующую оценку контрольной величины:

$$xc3 = x - b_2 \left\{ \sum_1^m y_t - \eta_1 m \right\} - b_3 \left\{ \left(\sum_1^m y_t \right)^2 - (\eta_1^2 m^2 + \sigma_1^2 m) \right\}. \quad (88)$$

Легко проверить с помощью (86) и (87), что (88) определяет *несмещенную* оценку. Итак, мы можем генерировать имитационный опыт. Следя за генерированием входных переменных y_t в конце опыта, мы знаем значения m и $\sum y_t$, которые и надо подставить в выражение (88).

3. Можно комбинировать эти два обобщения, а именно множественную контрольную величину и обобщенную функцию g . Мой [Моу, 1965, р.60], например, экспериментировал с функцией, у которой m_1 и m_2 обозначают число значений входов типа 1 и 2 соответственно.

$$g_3(y_{11}, y_{12}, \dots, y_{1m_1}, y_{21}, y_{22}, \dots, y_{2m_2}) = \\ = b_1 + b_2 \sum_1^{m_1} y_{1t} + b_3 \left(\sum_1^{m_1} y_{1t} \right)^2 + b_4 \sum_1^{m_2} y_{2t} + b_5 \left(\sum_1^{m_2} y_{2t} \right)^2. \quad (89)$$

4. Поскольку статистическая входная переменная генерируется с помощью случайных чисел r , мы можем определить контрольную величину как функцию случайных чисел. Мой [Моу, 1965, р.60; 1966, р. 10—11] экспериментировал с функцией

$$h_1(r_1, r_2, \dots, r_m) = b_1 + b_2 \sum_1^{m_1} r_t + b_3 \left(\sum_1^{m_1} r_t \right)^2 + \\ + b_4 \sum_{m_1+1}^m r_t + b_5 \left(\sum_{m_1+1}^m r_t \right)^2, \quad (90)$$

где первые m_1 случайных чисел генерируют вход типа 1 (например, время между поступлениями), а остальные случайные числа — вход типа 2 (например, время обслуживания).

5. Все упомянутые контрольные величины названы Моем зависящими от системы, так как случайный вход в (68), (74), (79) и (89) есть преобразование случайных чисел, а это преобразование зависит от моделируемой системы. Даже в выражении (90) мы также имеем величину, зависящую от системы, так как все случайные числа разделены на две части в зависимости от имеющихся в системе двух типов входных переменных. Пример не зависящей от системы контрольной величины дает выражение

$$h_2(r_1, r_2, \dots, r_m) = b_1 + b_2 \sum_1^m r_t + b_3 \left(\sum_1^m r_t \right)^2. \quad (91)$$

При обсуждении выражения (81) мы указывали на то, что контрольную величину стоит определить так, чтобы она сильно (положительно) коррелировала с откликом. Если мы объединяем все случайные числа, как в [91] и как предлагает Мой [Моу, 1965, р.36; 1966, р.5, 9], то мы должны писать машинную программу так, чтобы большие случайные числа приводили к большим значениям переменной. Способ, которым можно реализовать это отношение, дается в III.2.4 при обсуждении (43) и (44). В случае контрольной величины, зависящей от системы, такое объединение невозможно, так как каждый тип входной переменной берется отдельно. С помощью знака коэффициентов b корреляцию между контрольной величиной и откликом можно сделать положительной (см. также III.4.3).

Мой [Моу, 1965, р.35—36] предложил еще несколько типов контрольных величин, но оказалось, что их вычисление занимает слишком много дополнительного времени или они слишком слабо коррелируют

с откликом. Экспериментальные результаты, полученные в исследовании Моя, а также в [Radema, 1969, p.22—24], показывают, что лучше всего работать с контрольными величинами, *зависящими от системы*. Интуитивно это кажется возможным, так как контрольная величина, зависящая от системы, содержит больше априорных знаний о системе. Итак, кажется, что лучше определить контрольную величину в терминах входных переменных, чем функцию случайных чисел/саих по себе, и, далее, лучше сохранить раздельными несколько типов входных переменных. Экспериментальные результаты в [Моу, 1965, p.62] и [Radema, 1969, p.23, 47] показывают, что корреляция между контрольной величиной и откликом не усиливается при добавлении членов *более высокого порядка*, как в формулах (84), (89), (90) и (91), поэтому лучше коэффициенты второго порядка положить равными 0. Но даже в этом случае в системе может быть много коэффициентов b , если есть много типов входных переменных. Как замечает Мой [Моу, 1965, p.63], можно положить коэффициент b_k равным 0, если наши *априорные знания* о системе позволяют предположить, что входные переменные k -го типа не сильно коррелируют с выходом. Поскольку нам необходимо оценить оптимальную величину каждого коэффициента, то важно, чтобы количество коэффициентов было небольшим. (Мы еще вернемся к этой проблеме в III.4.3.)

6. Контрольные переменные Моя, как зависящие от системы, так и не зависящие от нее, применимы во всех видах имитационных исследований. Конечно, мы можем сформировать *конкретную* контрольную переменную для конкретного типа системы, которую мы хотим моделировать. Радема [Radema, 1969] использовал переменную

$$xcr = x - a \{ \bar{w}/\bar{v} - E(\bar{w}/\bar{v}) \}, \quad (92)$$

где \bar{v} и \bar{w} — среднее время между поступлениями и среднее время обслуживания соответственно в одноканальной системе массового обслуживания и отношение \bar{w}/\bar{v} оценивает *загрузку* системы. Для m независимых экспоненциально распределенных времени поступления \bar{v} и времени обслуживания \bar{w} Радема [Radema, 1969, p.23] доказал, что

$$E\left(\frac{\bar{w}}{\bar{v}}\right) = \frac{m}{m-1} \cdot \frac{E(\bar{w})}{E(\bar{v})}. \quad (93)$$

Бейа [Beja, 1969, p.5—7] экспериментировал с некоторыми другими контрольными величинами в системе массового обслуживания с *приоритетом*, т. е.

$$z = \sum_{j=1}^m s_j/m \quad (94)$$

при

$$s_j = \max(w_{j-1} - v_j, 0), \quad (95)$$

где v_j — время между поступлениями требований $(j-1)$ и j , w_{j-1} — время обслуживания $(j-1)$ требования. Переменные типа z могут определяться для класса с приоритетом. Для вывода $E(z)$ мы отсылаем

к Бейя [Beja, 1969, p.5—6]. Этот автор упоминает несколько других контрольных величин, специально применяющихся в системах массового обслуживания с приоритетом. В [Gaver, 1969] обсуждаются другие контрольные величины для систем массового обслуживания. В [Raj, 1968, p.85—106, 140—142, 150—151] исследуются контрольные величины не только в задачах Монте-Карло, но и в выборочных исследованиях.

7. К контрольным величинам можно отнести так называемую *оценку отношений* $(x/y)E(y)$. Эта оценка смещена в том случае, если $E(x/y) \neq E(x)/E(y)$. Вообще оценка полезна, если x и y имеют коэффициент корреляции, превышающий $1/2$ [Gray and Schucany, 1972, p.83]. Для уменьшения смещения может быть применен метод «складного ножа» (см. обсуждение этого метода в приложении I). Для более подробного рассмотрения оценок отношения и их связи с контрольными величинами и методом «складного ножа» отсылаем к [Rao, 1969], [Gray and Schucany, 1972, p.80—102], [Mosteller and Tukey, 1968, p.144—147] и [Raj, 1968, p.85—106].

Мы увидели, что контрольные величины всех типов оставляют процесс выборки неизменным. Следовательно, для изучения других характеристик, кроме среднего отклика, особенно динамического поведения систем, можно применять генерацию траектории во времени.

III.4.3. Оценивание контрольных коэффициентов

На основании выражений (71a) и (71б) мы видели, что оценка с контрольной величиной дает меньшую дисперсию, чем грубая оценка, в том случае, когда контрольный коэффициент лежит в определенных пределах. Эти пределы зависят от ковариации (или корреляции) между откликом и контрольной величиной, и эта ковариация неизвестна. Как показано в (72), *оптимальный* контрольный коэффициент будет также функцией от неизвестной ковариации. Обсудим оценку оптимальных контрольных коэффициентов для нескольких вариантов контрольных величин, которые были уже представлены.

1. Оптимальный коэффициент оценки *простой* контрольной величины (68) выводится в виде

$$a_0 = \text{cov}(x, \bar{y}) / \text{var}(\bar{y}) = \rho(x, \bar{y})\sigma(x) / \sigma(\bar{y}). \quad (96)$$

Теперь можно вычислить

$$\text{var}(\bar{y}) = \text{var}(y_i)E(m^{-1}). \quad (97)$$

Выполнение равенства (97) доказывается просто, если \bar{y} есть среднее m независимых наблюдений y_i , где все y_i распределены одинаково. Первый множитель в (97) — $\text{var}(y_i)$ — можно вычислить из известного распределения y_i . Второй множитель равен $1/m$, если m — известная константа. В противном случае можно вычислить $\text{var}(\bar{y})$ приближенно, применяя выражение (98), где нужно еще определить $E(m)$.

$$E(m^{-1}) \approx 1/E(m). \quad (98)$$

Более серьезная проблема состоит в том, что мы не знаем $\text{cov}(\bar{x}, \bar{y})$ в (96). Мой [Моу, 1965, р.32—33] замечает, что можно оценить либо только $\text{cov}(\bar{x}, \bar{y})$, либо обе величины $\text{cov}(\bar{x}, \bar{y})$ и $\text{var}(\bar{y})$. В одном эксперименте с одноканальной системой массового обслуживания Радема [Radema, 1969, table 3.4.1.2] применил обе процедуры и нашел одинаковое понижение дисперсии для обеих результирующих оценок контрольных величин. Оценивание $\text{cov}(\bar{x}, \bar{y})$ и $\text{var}(\bar{y})$ вместе желательно, если ошибка в одной оценке компенсируется ошибкой в другой¹⁹. Мой, Радема и Точер оценивают значения $\text{cov}(\bar{x}, \bar{y})$ и $\text{var}(\bar{y})$, и мы будем следовать за ними, так как этот путь может привести к компенсации ошибок, как мы только что заметили. Более того, мы увидим, что эта процедура приводит к уже знакомому нам алгоритму МНК. Пусть a_0 оценивается из (99):

$$\hat{a}_0 = \frac{c(\bar{x}, \bar{y})}{s^2(\bar{y})} = r(\bar{x}, \bar{y}) \frac{s(\bar{x})}{s(\bar{y})}, \quad (99)$$

где

$$c(\bar{x}, \bar{y}) = \sum_{i=1}^n (x_i - \bar{x})(\bar{y}_i - \bar{y}) / (n-1), \quad (100)$$

$$s^2(\bar{y}) = \sum_{i=1}^n (\bar{y}_i - \bar{y})^2 / (n-1), \quad (101)$$

$$s^2(\bar{x}) = \sum_{i=1}^n (x_i - \bar{x})^2 / (n-1), \quad (102)$$

$$r(\bar{x}, \bar{y}) = \sum_{i=1}^n (\bar{x}_i - \bar{x})(\bar{y}_i - \bar{y}) / \{(n-1) s(\bar{x}) s(\bar{y})\}, \quad (103)$$

$$\bar{y} = \sum_{i=1}^n \bar{y}_i / n, \quad (104)$$

$$\bar{x} = \sum_{i=1}^n x_i / n, \quad (105)$$

т. е. мы используем хорошо известные оценки дисперсии, ковариации и коэффициента корреляции. (Заметим, что в литературе индивидуальное наблюдение обычно обозначается x_i и y_i , в нашем случае таким «индивидуальным» наблюдением служат отклики x_i имитационного опыта i и среднее входной переменной \bar{y}_i опыта i при общем числе опытов, равном n .) Интересен тот факт, что \hat{a}_0 равно оценке МНК регрессионного коэффициента α в простом линейном регрессионном уравнении (106) [Johnston, 1963 р.12]:

$$x = \alpha \bar{y} + \beta + u, \quad (106)$$

где u обозначает случайный шум относительно линии регрессии. По определению оценка α с помощью МНК минимизирует

$$\sum_1^n (x_i - \tilde{x}_i)^2 = \sum_1^n (x_i - \alpha \bar{y}_i - \beta)^2, \quad (107)$$

где x обозначает величину, предсказываемую с помощью регрессионного уравнения.

Оценивание контрольного коэффициента a_0 ведет и приводит к осложнениям, так как \hat{a}_0 есть случайная величина и это вызывает *смещение* оценки контрольной величины x_{ce} . Оценка контрольной величины вместе с оценкой a_0 определяется следующим выражением:

$$x_{ce_i} = x_i - \hat{a}_0 (\bar{y}_i - \eta) \quad (i = 1, \dots, n). \quad (108)$$

Следовательно,

$$E(x_{ce_i}) = E(x_i) - E(\hat{a}_0 \bar{y}_i) + \eta E(\hat{a}_0), \quad (109)$$

где

$$E(\hat{a}_0 \bar{y}_i) = E(\hat{a}_0) E(\bar{y}_i) = E(\hat{a}_0) \eta \quad (110)$$

будет выполняться, если только \hat{a}_0 и \bar{y}_i независимы [Fisz, 1967, p.82]. Однако (99) показывает, что \hat{a}_0 зависит от \bar{y}_i . Следовательно, оценка x_{ce_i} имеет смещение, так же как и среднее n оценок x_{ce_i} ($i = 1, \dots, n$). Покажем, как можно это исправить.

Для того чтобы получить *несмещенную* оценку контрольной величины, Точер [Tocher, 1963, p.115] предлагает *разбить* n пар (x_i, \bar{y}_i) на две равные группы. Из первой группы α_1 определяется так, чтобы минимизировать

$$\sum_{h=1}^{n_1} (x_h - \alpha_1 \bar{y}_h - \beta_1)^2. \quad (111)$$

Обозначим МНК-оценку α_1 через \hat{a}_{01} . Из данных второй группы α_2 подбирается так, чтобы минимизировать

$$\sum_{g=n_1+1}^n (x_g - \alpha_2 \bar{y}_g - \beta_2)^2. \quad (112)$$

Пусть \hat{a}_{02} обозначает МНК-оценку величины α_2 . Комбинируем оценку \hat{a}_{01} , полученную из наблюдений над первой группой, с наблюдениями второй группы. Это дает нам «оценку контрольной величины с коэффициентами, оцененными в результате разбиения», определенную в (113):

$$x_{ces_g} = x_g - \hat{a}_{01} (\bar{y}_g - \eta) \quad (g = n_1 + 1, \dots, n). \quad (113)$$

Так как \hat{a}_{01} не зависит от \bar{y}_g , легко доказать с помощью (110), что x_{ces_g} — несмещенная оценка. Следовательно, из n пар (x_i, \bar{y}_i) можно получить несмещенную оценку

$$\bar{x}_{ces} = \frac{1}{2} \bar{x}_{ces_I} + \frac{1}{2} \bar{x}_{ces_{II}}, \quad (114)$$

где

$$\bar{x}_{ces_I} = \sum_{g=n_1+1}^n x_{ces_g} / n_1 \quad (115)$$

и \bar{x}_{cesII} определена аналогично комбинированием \hat{a}_{02} с наблюдениями из первой группы.

Точер [Tosher, 1963, p. 116] показал, что дисперсия x_{ces} в два раза больше дисперсии оценки с неслучайной a_0 , которая была определена в (73). Это увеличение дисперсии можно уменьшить с помощью процедуры *обобщенного разбиения*. Все n пар наблюдений делятся на J групп ($2 < J \leq n$) и оценка \bar{a}_{0j} ($j = 1, 2, \dots, J$) основана на всех парах, исключая пары группы j ; аналогично (113) мы образуем

$$x_{cesj_i} = x_{ji} - \hat{a}_{0j} (\bar{y}_{ji} - \eta) \quad (i = 1, \dots, n/J), \quad (116)$$

где x_{ji} и \bar{y}_{ji} образуются из пар i внутри группы j . Точер показал, что дисперсия оценки ниже, когда разбиение производится на число групп, большее чем 2. Недостаток — потребность в дополнительном машинном времени, так как вычисляется большее число коэффициентов a_{0j} , особенно если J велико. Технику разбиения можно встретить также в статистической литературе по оценкам регрессии «складного ножа» (и отношения) для понижения смещения (см., например, [Rao, 1969, p.216] или, в более общем виде, [Gray and Schucany, 1972, p.80—107] и [Raj, 1968, p.236]). В литературе по методу «складного ножа» рекомендуется брать возможно большее число групп или $J = n$ [Rao, 1969, p.217]. В исследованиях по имитационному моделированию [Gaver and Shedler, 1971, p.449] и [Radema, 1969, p.25] берется ${}^{20}J = n$.

2. Далее мы обсудим более общий случай, когда контрольная величина \bar{y}_k ($k = 1, \dots, K$) многомерна, как мы ее определили в (74). Удобно взять контрольную величину с *нулевым математическим ожиданием*. Определим новую контрольную величину z_k :

$$z_k = \bar{y}_k - E(\bar{y}_k) = \bar{y}_k - \eta_k, \quad (117)$$

так что

$$E(z_k) = 0. \quad (118)$$

Тогда оценка контрольной величины (74) эквивалентна

$$x_{ст} = x - \sum_1^K a_k z_k. \quad (119)$$

Реалистично предположить, что контрольная величина z_k *независима* от других контрольных величин $z_{k'}$ ($k \neq k'$). В большинстве имитационных моделей входная переменная y_k независима от входной переменной $y_{k'}$. Например, время обслуживания одной очереди предполагается обычно независимым от времени обслуживания другой очереди и от времени поступления. (Далее здесь будут рассмотрены и зависимые контрольные величины.) Для независимых контрольных величин (119) справедливо выражение

$$\text{var}(x_{ст}) = \text{var}(x) + \sum_1^K a_k^2 \text{var}(z_k) - 2 \sum_1^K a_k \text{cov}(x, z_k). \quad (120)$$

Следовательно, оптимальные значения коэффициентов можно определить, беря частные производные (120) и решая относительно a_k . Реше-

ние дает оптимальные значения \hat{a}_{k0} :

$$a_{k0} = \frac{\text{cov}(x, z_k)}{\text{var}(z_k)} \quad (k = 1, 2, \dots, K). \quad (121)$$

Так как $E(z_k)$ равно нулю, (121) эквивалентно (122).

$$a_{k0} = \frac{E(xz_k)}{E(z_k^2)} \quad (k = 1, 2, \dots, K). \quad (122)$$

Следовательно, простая²¹ оценка оптимального контрольного коэффициента равна:

$$\hat{a}_{k0} = \frac{\sum_{i=1}^n x_i z_{ki}}{\sum_{i=1}^n z_{ki}^2} \quad (k = 1, 2, \dots, K). \quad (123)$$

Теперь рассмотрим *линейное регрессионное уравнение*

$$x_i = \sum_{k=1}^K a_k z_{ki} + u_i \quad (i = 1, 2, \dots, n) \quad (124)$$

или его матричный эквивалент

$$\vec{X} = \vec{Z}\vec{A} + \vec{U} \quad (125)$$

при

$$\vec{X}' = (x_1, x_2, \dots, x_n), \quad (126)$$

$$\vec{Z} = \begin{bmatrix} z_{11}, & z_{21}, & \dots, & z_{K1} \\ z_{12}, & z_{22}, & \dots, & z_{K2} \\ \vdots & \vdots & & \vdots \\ z_{1n}, & z_{2n}, & \dots, & z_{Kn} \end{bmatrix}, \quad (127)$$

$$\vec{A}' = (a_1, a_2, \dots, a_K), \quad (128)$$

$$\vec{U}' = (u_1, u_2, \dots, u_n). \quad (129)$$

Согласно [Johnston, 1963, p.109]²² МНК-оценки определяются следующим выражением:

$$\vec{\hat{A}} = (\vec{Z}'\vec{Z})^{-1}\vec{Z}'\vec{X}. \quad (130)$$

Элемент k -го столбца и k' -й строки матрицы $(\vec{Z}'\vec{Z})$ — это

$$\sum_{i=1}^n z_{ki} z_{k'i} \quad (k, k' = 1, 2, \dots, K). \quad (131)$$

Если эту сумму разделить на $(n - 1)$, то получим несмещенную оценку ковариации между z_k и $z_{k'}$. При выводе значений контрольных коэффициентов мы предполагали независимость контрольных величин z_k . Следовательно, мы можем воспользоваться этими априорными зна-

ниями при оценке коэффициентов регрессии (124), т. е. приравнять нулю внедиагональные элементы матрицы $(\vec{Z}'\vec{Z})$. На главной диагонали согласно (131) останутся элементы $\Sigma_i z_{ki}^2$. Следовательно, $(\vec{Z}'\vec{Z})^{-1}$ сводится к диагональной матрице с элементами на диагонали $(\Sigma_i z_{ki}^2)^{-1}$. Итак, оценки регрессионных коэффициентов в (124), вычисленные по методу наименьших квадратов с учетом независимости между переменными z_k , идентичны оптимальным оценкам контрольных коэффициентов (123).

3. Рассмотрим обобщенную контрольную величину

$$z = a_1 + a_2 g_2(r_1, r_2, \dots, r_m) + a_3 g_3(r_1, r_2, \dots, r_m) + \dots + a_K g_K(r_1, r_2, \dots, r_m), \quad (132)$$

где g_k — функции от m случайных чисел r_i ($i = 1, 2, \dots, m$). Поэтому (132) включает различные типы контрольных величин, представленных в III.4.2. Как там же было отмечено, от g_k требуется, чтобы была возможность вычисления его математического ожидания. Значит, мы можем рассмотреть также и контрольные величины z_k с нулевым математическим ожиданием:

$$z_k = g_k(r_1, r_2, \dots, r_m) - E[g_k(r_1, \dots, r_m)] \quad (k = 1, 2, \dots, K). \quad (133)$$

Заметим, что из (132) и (133) следует, что

$$z_1 = 1 - 1 = 0. \quad (134)$$

Отсюда оценка контрольной величины дается выражением

$$x_{cgz} = x - \sum_1^K a_k z_k. \quad (135)$$

Правая часть (135) идентична правой части (119). В (135), однако, возможно, что контрольные величины z_k зависимы. Эта зависимость видна в формулах (88)—(91), где члены второго порядка зависят от соответствующих членов первого порядка. В [Вежа, 1969, р.2] приводятся оптимальные значения коэффициентов a_k в формуле (135)²³:

$$\vec{A}_0 = \vec{\Omega}_z^{-1} \vec{\Psi}, \quad (136)$$

где $\vec{\Omega}_z$ — ковариационная матрица контрольных величин z_k , т. е.

$$\vec{\Omega}_z = \{E[(z_k - E(z_k))(z_{k'} - E(z_{k'}))]\} = \{E(z_k z_{k'})\} \quad (k, k' = 1, 2, \dots, K). \quad (137)$$

Последнее равенство следует из (133), и $\vec{\Psi}$ в (136) означает вектор ковариаций между x и z_k , т. е.

$$\vec{\Psi} = \{E(xz_1), E(xz_2), \dots, E(xz_k)\}, \quad (138)$$

где мы имеем

$$\text{cov}(x, z_k) = E(xz_k) - E(x)E(z_k) = E(xz_k). \quad (139)$$

Итак, мы можем *оценить* оптимальные коэффициенты a_{0k} по простой аналогии с (136):

$$\vec{\hat{A}}_0 = (\vec{Z}' \vec{Z})^{-1} \vec{Z}' \vec{X}, \quad (140)$$

где \vec{X} и \vec{Z} определены в (126) и (127). Сравнение (140) с (130) показывает, что *оптимальные* коэффициенты контрольных величин в (135) можно вычислить, применяя алгоритм *метода наименьших квадратов* для вычисления коэффициентов в линейном регрессионном уравнении (125)²⁴. Поэтому метод контрольных величин с оценкой коэффициентов часто называется «регрессионной выборкой». Аналогично (108)—(110) оценка (135) имеет *смещение*, так как коэффициенты становятся зависимыми от контрольных величин. Но это можно исправить, как в (116), *разбиваем* n наблюдений x, z_1, \dots, z_K на J групп ($2 \leq J \leq n$) с n/J наблюдениями в каждой группе и оценкой a_{kj} , коэффициента контрольной величины k в группе j ($j = 1, \dots, J$), на основании всех оставшихся наблюдений ($n - n/J$). Это дает «обобщенную оценку контрольной величины с оптимально оцененными коэффициентами»:

$$x_{cze_{ji}} = x_{ji} - \sum_{k=1}^K \hat{a}_{kj} z_{kji} \\ (j=1, \dots, J) \quad (i=1, \dots, n/J). \quad (141)$$

Даже если мы не знаем $E(\hat{a}_{kj})$, то нам известно, что $E(\hat{a}_{kj} z_{kji}) = 0$, так как \hat{a}_{kj} независимы от z_{kji} и $E(z_{kji}) = 0$.

Отметим, что для процедуры оценки коэффициентов необходимо, чтобы было «*достаточно*» наблюдений, т.е. повторных имитационных опытов. В [Kleijnen, 1969, p.291—292] указывается, что это число может быть мало в случае моделирования установившегося поведения. Слишком малое число опытов может стать препятствием для оценки большого числа контрольных коэффициентов. Но даже если эта оценка окажется возможной, то дисперсии оценок возрастают с уменьшением числа наблюдений, что приведет к меньшему понижению дисперсии. Эти соображения приводят к тому, что оценкой контрольной величины можно пользоваться в случае малого количества коэффициентов, если число наблюдений не слишком велико. Так, контрольная величина первой степени из (82) более привлекательна с этой точки зрения, чем величина второй степени из (88)—(91). Укажем также на численные результаты Моя [Моя, 1965, p.49—54]. Радема [Radema, 1969, p.23, 47] обнаружил даже повышение дисперсии при применении квадратичной контрольной величины (89) по сравнению с контрольной величиной первой степени.

Проблема, которая не затрагивается в литературе по имитационному моделированию, — это *оценивание дисперсии* оценки контрольной величины в (141). В приложении III.7 показано, что оценки контрольной величины $x_{cze_{ji}}$ не коррелируют *внутри* одной группы, но коррелируют *между* группами. Последний результат очевиден, так как оценка $x_{cze_{ji}}$ зависит от \hat{a}_{kj} , а \hat{a}_{kj} зависят от значений переменных

вне группы j . Но если оценки \hat{x}_{cze} внутри группы j зависят от оценок \hat{x}_{cze} вне группы j от a_{hj} , то (142) выполняется только *приближенно*:

$$\text{var}(\bar{x}_{cze}) = \sum_{j=1}^J \hat{v}_j \hat{a}_j / J, \quad (142)$$

так как \bar{x}_{cze_j} — среднее группы j — не независимо от средних других групп. Далее заметим, что Бейя [Beja, 1969, p.7] дает некорректную формулу, так как он использует формулу, подобную (120), заменяя $\text{var}(\mathbf{x})$, $\text{var}(\mathbf{z}_h)$ и $\text{cov}(\mathbf{x}, \mathbf{z}_h)$ их оценками. Формула, подобная (120), однако, пригодна только для неслучайных контрольных коэффициентов a_h . Радема [Radema, 1969, p.39—40] выводит формулу для дисперсии оценки отдельной контрольной величины, принимая во внимание стохастический характер \hat{a}_h . Однако дисперсия среднего индивидуальных оценок включает неизвестные ковариации между оценками. Очевидно, что из-за указанной выше зависимости *доверительные интервалы*, подобные (12), будут лишь приближенными. Изложенные далее подходы также основаны на известной литературе по методам выборки.

Рао [Rao, 1969, p.222] нашел, что в примерах, которые он изучал, смещенная оценка (108) (или t_s в обозначениях Рао), где a оценивается на основании всех наблюдений, имеет меньшую СКО, чем несмещенная оценка (116) (или t_s у Рао), основанная на общем разбиении на $J = n$ групп. Следовательно, можно пользоваться смещенной оценкой и оценить ее дисперсию на основании асимптотического отношения, выведенного в [Raj, 1968, p.100]:

$$\text{var}(\bar{x}_{ce}) = \sigma_x^2 (1 - \rho^2) / n. \quad (143)$$

Очевидно, что полученный в результате интервал основан на грубом приближении (большая дисперсия выборки, неучтенное смещение; Рао [Rao, 1969, p. 224] нашел, что его оценка дисперсии v_s имеет значительное смещение). Необходимы дополнительные исследования для того, чтобы найти хорошую оценку дисперсии регрессионной оценки. Эта оценка, возможно, будет основана на применении метода разбиения либо метода «складного ножа» [Miskey, 1959, p.600—601], [Rao, 1969, p.216—217]. В качестве еще одной альтернативы мы предлагаем регрессионную оценку, основанную на методе «складного ножа».

1. Вычисляем оценку (смещенную) $\theta = E(\mathbf{x})$.

$$\hat{\theta} = \bar{\mathbf{x}} - \sum_{k=1}^K \hat{a}_k \bar{\mathbf{z}}_k, \quad (144)$$

где $\bar{\mathbf{x}}$ — средний отклик из n повторных опытов, $\bar{\mathbf{z}}_k$ — среднее значений контрольной величины \mathbf{z}_k и \hat{a}_k — оценка оптимального коэффициента на основе *тех же* опытов; сравните со (108).

2. Образует N групп, в каждой по $M = n/N$ наблюдений. Вычисляем $\hat{\theta}_j$ аналогично (144) после исключения группы:

$$\hat{\theta}_j = \bar{\mathbf{x}}_j - \sum_{k=1}^K \hat{a}_{kj} \bar{\mathbf{z}}_{kj} \quad (j = 1, \dots, N), \quad (145)$$

где \bar{x}_j — среднее $(n - M)$ наблюдений после исключения группы j , \hat{a}_{kj} оценивается по тем же $(n - M)$ наблюдениям и т. д. Заметим, что в противоположность (141) оценка $\hat{\theta}_j$ остается смещенной. «Псевдозначения» статистики по методу «складного ножа» есть

$$J_j = N\hat{\theta} - (N - 1)\hat{\theta}_j. \quad (146)$$

3. Вычисляем оценку по методу «складного ножа»:

$$\bar{J} = N\hat{\theta} - (N - 1) \sum_{j=1}^N \hat{\theta}_j / N = \bar{x} + (N - 1) \sum_j \sum_k \alpha_{kj} \bar{z}_{kj} / N. \quad (147)$$

При хороших условиях эта оценка имеет смещение меньше, чем $\hat{\theta}$ в (144). Общее описание метода «складного ножа» приведено в приложении III.1.

4. Вычисляем доверительный интервал для \bar{J} , считая J_j независимыми наблюдениями и используя t -распределение. Можно построить консервативные доверительные интервалы с помощью оценок возможных положительных корреляций между J_j (см. приложение III.1).

III.4.4. Приложения контрольных величин в моделировании и исследованиях по методу Монте-Карло

Мой [Моу, 1965, 1966] применял различные контрольные величины в одноканальных системах массового обслуживания и к уже упоминавшейся ранее системе товарообмена в порту. Радема [Radema, 1969] также применял контрольные величины к одноканальной системе массового обслуживания. В III.4.2 (пункт 5) указывалось на то, что их эксперименты имеют тем лучшие результаты, чем больше контрольные величины *зависят от системы*. В III.4.3 мы также указывали на то, что оценка должна содержать по возможности *меньше коэффициентов*. Очевидно, что мы получим лучшие оценки этих коэффициентов, если будем иметь *большее число наблюдений*, т. е. большее число повторений опытов; экспериментальные результаты можно найти в [Radema, 1969, table 3.4.1.2].

Экспериментальные результаты Моя [Моу, 1965, р.55, 60—62; 1966, р.29] и Радема [Radema, 1969, tables 3.4.1.1 и 3.4.1.2] показывают, что с помощью метода контрольных величин в *простых* системах массового обслуживания можно добиться *достаточно понижения дисперсии*. Понижение дисперсии становится более важным по мере роста коэффициента использования системы. Понижение дисперсии у Моя равно 20—50%, у Радема — от 15% (при коэффициенте использования 50%, 10 повторных опытах и оценке контрольной величины с двумя коэффициентами, как в (74), для $K = 2$) до 68% (при коэффициенте использования 90%, 40 опытах, одном коэффициенте, как в (92)). Мой далее применил контрольные величины для сложной *системы товарообмена в порту*. Для независимой от системы величины, определенной как в (91), понижение дисперсии оказалось равным только 7% для 90 параллельных опытов [Моу, 1965, р.56—57]. Согласно Мою [Моу, 1966,

р.33] для системы товарообмена в порту и зависимой от системы контрольной величины понижение дисперсии было между 14 и 47%. Бейа [Beja, 1969, p.6] применяет контрольную величину, основанную на (94), которая даст понижение дисперсии, равное 60%; при применении комбинации контрольных величин из среднего времени обслуживания и среднего времени прибытия понижение дисперсии оказалось равным 92%. Бейа [Beja, 1969, p.8] сообщает также, что для системы массового обслуживания с *приоритетом* «в ряду опытов не более чем с 4 контрольными величинами отношение дисперсии было около 2 для неблагоприятных условий и около 10 для благоприятных». Андреассон [Andreasson, 1971] приводит подробные результаты применения контрольных величин в моделировании простого телефонного коммутатора, для которого оценивалась некая вероятность p . Для ряда вариантов системы ему удалось уменьшить дисперсию вдвое, но понижение дисперсии уменьшается по мере того, как p приближается к 0 или 1. Очевидно, что при оценке метода понижения дисперсии нужно также учитывать *дополнительное машинное время*, потребное для вычисления контрольных величин и их оптимальных коэффициентов. Мой [Moy, 1966, p.33] также вычисляет это скорректированное понижение дисперсии для системы товарообмена в порту. Понижение изменяется от 1 до 41%. Как замечает Радема [Radema, 1969, p.32], дополнительное время становится менее важным при увеличении времени одного имитационного опыта. Можно попытаться сберечь машинное время, не оценивая контрольные оптимальные коэффициенты, а беря их значения из аналогичных предыдущих имитационных задач. К сожалению, Мой [Moy, 1965, p.67—70] показал, что оценки оптимальных контрольных коэффициентов меняются в зависимости от вариантов системы, следовательно, кажется необходимым оценивать коэффициенты для каждого варианта отдельно. Радема [Radema, 1969] исследовал, как связано понижение дисперсии с коэффициентом использования одноканальной системы массового обслуживания, длиной каждого имитационного опыта, начальными условиями и числом повторных опытов. Кроме того, в [Gaver, 1969] и [Lombaers, 1968, p.253] описаны эксперименты с контрольной величиной в простой системе массового обслуживания, в [Toscher, 1963, p.115—117, 172—173] также кратко обсуждаются контрольные величины в моделировании.

Применение метода контрольной величины в исследованиях по методу *Монте-Карло* довольно просто. Так, в примере из приложения I.1 мы оценивали *интеграл*

$$\xi = \int_0^{\infty} \frac{1}{y} \lambda e^{-\lambda y} dy \quad (\lambda, v > 0) \quad (148)$$

с помощью выборки y из плотности распределения с экспоненциальным законом

$$f(y) = \lambda e^{-\lambda y} \quad (y > 0). \quad (149)$$

Очевидной контрольной величиной будет y с известным средним $1/\lambda$. В приложении I.2 мы рассмотрели пример *выборочного распределения*,

где мы оценивали вероятность p , определенную следующим выражением:

$$p = P(\mathbf{x} < a), \quad (150)$$

где

$$\mathbf{x} = \min(\mathbf{y}_1, \mathbf{y}_2) \quad (151)$$

имеет нормальное распределение $N(\eta_i, \sigma_i^2)$ ($i = 1, 2$). Здесь также очевидными контрольными величинами будут \mathbf{y}_i с известными средними η_i . (В действительности метод контрольных величин не применен нами к этим двум примерам, поэтому мы не можем привести численные расчеты эффективности.)

Другие типы контрольных величин описаны в литературе²⁵. К ним относится величина, которая не является случайным *входом* моделируемой системы. Вместе с нашей моделируемой системой моделируется и ее *упрощенный вариант*. Относительно этой последней системы предполагается, что она имеет *известную среднюю*, скажем μ_z . При ее моделировании пользуются той же самой последовательностью случайных чисел, что и при моделировании основной системы, следовательно, \mathbf{z} — отклик упрощенной системы — коррелирован с \mathbf{x} — откликом интересующей нас системы (см. III.7). Можно оценить μ_x — среднее интересующей нас системы — на основании²⁶

$$\hat{\mu}_x = \mathbf{x} - a(\mathbf{z} - \mu_z), \quad (152)$$

где оптимальное значение a можно оценить, если есть повторные опыты с обеими системами. Этот подход относится к тому типу стратификации, который описан в начале III.2.4, и (152) обладает всеми его недостатками. Во-первых, должна существовать подходящая система с известной средней. Во-вторых, программирование и счет могут занимать слишком много времени (однако, возможно, уже существует запрограммированная упрощенная система для просчетов обоснованности, верификации и проверки чувствительности, см. главу II). При *выборочном распределении* можно пренебречь тем дополнительным рабочим временем, которое требуется для генерирования соответствующего отклика \mathbf{z} . Рассмотрим оценку математического ожидания для выборки из m наблюдений из нормально распределенной совокупности. На основании m наблюдений $\hat{\mu}_x$ мы можем оценить не только математическое ожидание, но также и дисперсию. Теперь в формуле (152) \mathbf{x} подставляется вместо размаха, а \mathbf{z} — дисперсия. Этот пример рассматривается в [Tocher, 1963, p.102—103]. Другой пример выборочного распределения обсуждается в [Hillier and Lieberman, 1968, p.460—461], где в выражении (152) \mathbf{x} — это некоторая сложная функция от случайного числа r , а \mathbf{z} — простая функция r с известным средним. В этих примерах время, затрачиваемое на программирование и счет для определения \mathbf{z} в (152), можно не учитывать. В [Burt et al., 1970] используется контрольная величина, аналогичная величине из (152) для моделирования вероятностных сетей. Авторы не сообщают время, необходимое для моделирования простой сети с известным решением. В работе указано, что использовались несколько систем управления. В [Maxwell, 1965, p.7—9] обсуждается применение формулы (152) при \mathbf{z} , которое есть

либо случайный вход нашей системы, либо случайный выход упрощенной системы. Кроме того, там рассмотрен случай, когда z — отклик упрощенной системы с *неизвестным средним*, когда μ_z в (152) заменяется на $\hat{\mu}_z$ — оценку отклика этой системы, полученную в имитационных опытах *ранее*. Очевидно, в последнем случае трудно оценить пониженные дисперсии, так как $\hat{\mu}_z$ повышает дисперсию оценки контрольных величин²⁷. В [Lewis, 1972, p.11] предлагается оценку переменной одного варианта системы на основании большого числа опытов брать в качестве контрольной величины при оценивании переменных в другом варианте с малым числом опытов (при том же потоке случайных чисел). Или можно воспользоваться

$$\hat{\mu}_x = \bar{x}_m - a(\bar{z}_m - \bar{z}_n), \quad (153)$$

где \bar{x}_m и \bar{z}_m — средние откликов 1-го и 2-го вариантов системы, рассчитанных для одних и тех же случайных чисел, а \bar{z}_n — среднее системы 2-го варианта, основанное на n опытах, которые включают m ($\ll n$) опытов для \bar{z}_m .

Наконец, в литературе можно найти метод, который также называется методом контрольных величин. Однако этот метод значительно отличается от описанного выше. Вместо оценки *только* среднего отклика он предполагает оценку *распределения* отклика. Этот метод основан на предположении, что помимо интересующего нас отклика мы рассматриваем еще дополнительную переменную с *известным* распределением. Метод был впервые предложен Файллером и Хартли [Fieller and Hartley, 1954], он описан также Фармером [Farmer, 1968, p. 7 — 9] и Точером [Tocher, 1963, p. 90 — 93]. Два последних автора иллюстрируют метод примером из *Монте-Карло*, в котором оценивается распределение медианы и определяется среднее, имеющее известное распределение. Мейтерн [Matern, 1962] описывает несколько вариантов метода контрольной величины с оценкой распределения.

III.5. ЗНАЧИМАЯ ВЫБОРКА

III.5.1. Основания значимой выборки

Элементарное обсуждение значимой выборки можно найти в [Clark, 1961]. Ее основная идея состоит в полной замене исходного выборочного процесса совершенно *другой* выборкой. В новом выборочном процессе наблюдения корректируются своего рода взвешиванием так, чтобы среднее подкорректированных наблюдений было несмещенной оценкой среднего исходного процесса. Таким образом, в методе значимой выборки в отличие от метода контрольных величин происходит изменение исходной выборочной процедуры, и, следовательно, выход нового выборочного процесса неприменим для исследования динамического поведения интересующей нас системы. Значимая выборка больше похожа на *стратифицированную* выборку, потому что: 1) исходный выборочный процесс изменяется (хотя при значимой выборке это изменение больше, чем при стратификации), 2) наблю-

дения имеют разные веса. Мы отсылаем к [Clark, 1961, p. 608 — 610], [Handscorn, 1968, p. 6 — 7] и [Hillier and Lieberman, 1968, p. 458] для дальнейших сравнений методов значимой и стратифицированной выборок. Далее мы дадим более точное представление основной идеи значимой выборки, как это описано также в [Kahn and Marshall, 1953, p. 269 — 270].

Допустим, что мы хотим оценить ξ -интеграл:

$$\xi = \int_{-\infty}^{\infty} g(\mathbf{x}) f(\mathbf{x}) dx, \quad (154)$$

где $f(\mathbf{x})$ — функция плотности вероятности. Следовательно, ξ — математическое ожидание $g(\mathbf{x})$, функцией плотности вероятности которой будет $f(\mathbf{x})$:

$$\xi = E[g(\mathbf{x})]. \quad (155)$$

Грубой оценочной процедурой будет выборка n величин \mathbf{x} из $f(\mathbf{x})$ и вычисление

$$\hat{\xi} = \frac{1}{n} \sum_{i=1}^n g(\mathbf{x}_i). \quad (156)$$

Этот пример рассматривался в приложении I.1, где $f(\mathbf{x})$ и $g(\mathbf{x})$ были определены в (1.3) и (1.4) соответственно. Однако можно записать (154) и в виде

$$\xi = \int_{-\infty}^{\infty} \frac{g(\mathbf{x}) f(\mathbf{x})}{h(\mathbf{x})} h(\mathbf{x}) dx. \quad (157)$$

Теперь если $h(\mathbf{x})$ есть другая плотность вероятности, то можно оценить ξ , выбирая \mathbf{x} из $h(\mathbf{x})$ и подставляя \mathbf{x} в (158):

$$g^*(\mathbf{x}) = \frac{g(\mathbf{x}) f(\mathbf{x})}{h(\mathbf{x})} = g(\mathbf{x}) \left[\frac{f(\mathbf{x})}{h(\mathbf{x})} \right], \quad (158)$$

где $f(\mathbf{x})/h(\mathbf{x})$ можно интерпретировать как взвешивающий множитель. Математическое ожидание $g^*(\mathbf{x})$ определяется из выражения

$$E[g^*(\mathbf{x})] = \int_{-\infty}^{\infty} g^*(\mathbf{x}) h(\mathbf{x}) dx, \quad (159)$$

которое сводится к ξ . Можно выбрать \mathbf{x} из $h(\mathbf{x})$ n раз и оценить ξ :

$$\hat{\xi}^* = \frac{1}{n} \sum_{i=1}^n g^*(\mathbf{x}_i). \quad (160)$$

В [Kahn and Marshall, 1953, p. 270] показано, что дисперсия оценки значимой выборки минимизируется, если в качестве новой плотности вероятности взять

$$h_0(\mathbf{x}) = \frac{|g(\mathbf{x})| f(\mathbf{x})}{\int_{-\infty}^{\infty} |g(\mathbf{x})| f(\mathbf{x}) dx}. \quad (161)$$

Если $g(x) \geq 0$ для всех x , то (161) сводится к

$$h_0(x) = \frac{g(x)f(x)}{\int_{-\infty}^{\infty} g(x)f(x) dx} = \frac{g(x)f(x)}{\xi}. \quad (162)$$

Итак, *оптимальность* новой плотности вероятностей означает, что выборка в основном производится из «значимых» областей x , т. е. из тех областей x , которые дают высокие значения отклика $g(x)$ (кроме тех случаев, когда вероятность такого значения x мала, т. е. $f(x)$ мала). В [Kahn and Marshall, 1953, p. 271] показано также, что если $g(x) \geq 0$ для всех x , то дисперсия оценки значимой выборки становится равной нулю. К сожалению, мы не сможем вычислить оптимальную плотность для (162), так как (162) содержит ξ и ее надо еще оценить. Тем не менее (162) можно воспользоваться при аппроксимации оптимальной плотности вероятности $h_0(x)$. Сначала мы продемонстрируем, как можно осуществить аппроксимацию с помощью методов Монте-Карло. Далее мы представим аппроксимацию Моей для более сложного случая моделирования.

III.5.2. Пример значимой выборки в методе Монте-Карло²⁸

Предположим, что нам нужно оценить интеграл из приложения I.1:

$$\xi(\lambda, v) = \int_v^{\infty} \frac{1}{x} \lambda e^{-\lambda x} dx \quad (\lambda, v > 0). \quad (163)$$

Сравнение с общим выражением для ξ в (154) показывает, что мы имеем

$$\begin{aligned} g(x) &= 0, & \text{если } x < v, \\ &= \frac{1}{x}, & \text{если } x \geq v \end{aligned} \quad (164)$$

и

$$\begin{aligned} f(x) &= \lambda e^{-\lambda x}, & \text{если } x \geq 0, \\ &= 0, & \text{если } x < 0, \end{aligned} \quad (165)$$

где (164) и (165) идентичны (1.3) и (1.4) в приложении I.1. Из (162) следует, что оптимальная плотность распределения для значимой выборки в этой задаче определяется выражением

$$\begin{aligned} h_0(x) &= 0, & \text{если } x < v, \\ &= \frac{\lambda}{\xi} \frac{1}{x} e^{-\lambda x}, & \text{если } x \geq v. \end{aligned} \quad (166)$$

Эта оптимальная плотность вероятности представлена на рис. 11.

Рассмотрим три *аппроксимации* этого оптимального распределения. (Пусть звездочка обозначает оценку значимой выборки.)

1. Сдвинутое экспоненциальное распределение. Сначала делаем выборку x' из экспоненциального распределения с параметром c

(вместо параметра λ в (165); значение c определяется позже): Далее мы прибавляем к x' константу v (см. рис. 11). Плотность распределения $h_1(x)$ сдвинутого экспоненциального распределения $x = x' + v$ определяется выражением (167) [Fisz, 1967, p. 39 — 40]:

$$h_1(x) = ce^{-c(x-v)} = ce^{cv} e^{-cx}, \text{ если } x \geq v, \\ = 0, \text{ если } x < v. \quad (167)$$

Если мы выбираем $h_1(x)$ в качестве новой функции плотности, по которой будет производиться значимая выборка x , то новой функцией $g(x)$ будет $g_1^*(x)$, определенная из (158), (164), (165) и (167):

$$g_1^*(x) = \left(\frac{1}{x}\right) (\lambda e^{-\lambda x}) (ce^{cv} e^{-cx})^{-1} = \frac{\lambda}{c} e^{-cv} \frac{1}{x} e^{-(\lambda-c)x}, \quad (168)$$

где $x \geq v$, так как согласно (167) мы не сможем генерировать $x < v$.

Из выражения (160) видно, что после получения n значений x из $h_1(x)$ можно оценить ξ :

$$\xi_1^* = \frac{1}{n} \sum_{i=1}^n g_1^*(x_i). \quad (169)$$

Далее нам нужно найти константу c в (167). Величину c выбираем так, чтобы дисперсия ξ_1^* была минимальной для данного n , т. е. минимизируем дисперсию $g_1^*(x)$. В [Kleijnen, 1968 b, p. 195 — 197] указано, что оптимальное значение c невозможно вычислить, так как выражение для определения c включает неизвестную ξ и поэтому целесообразен следующий подход [Kleijnen, 1968 b, p. 183]. В III.5.1 мы видели, что при оптимальной плотности распределения $h_0(x)$ в (162) дисперсия оценки значимой выборки равна 0. Следовательно, размах R также равен нулю.

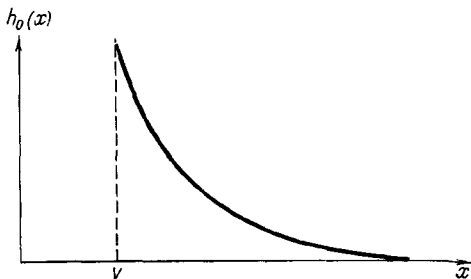


Рис. 11. Оптимальная плотность вероятности значимой выборки при оценивании $\xi(\lambda, v)$

$$R[g_1^*(x)] = \max_x [g_1^*(x)] - \min_x [g_1^*(x)]. \quad (170)$$

Поскольку мы не можем вычислить c из условия минимума дисперсии $g_1^*(x)$, выберем c так, чтобы минимизировать размах $g_1^*(x)$. Если $c > \lambda$, то, как видно из (168), $g_1^*(x) = \infty$ для $x = \infty$ или $R = \infty$. Поэтому мы должны взять $c \leq \lambda$, чтобы область была конечной. Для любого значения c из интервала $(0, \lambda]$ минимальной величиной $g_1^*(x)$ будет 0, если $x = \infty$. Чтобы минимизировать R , мы должны минимизировать максимальное значение $g_1^*(x)$, выбирая подходящее c в $(0, \lambda]$. Этот максимум $\max [g_1^*(x)]$ достигается при $x=v$, так как

из (168) видно, что $g_1^*(x)$ есть убывающая функция x (если $0 < c \leq \lambda$), т. е. выполняется

$$\max_x [g_1^*(x)] = g_1^*(v) = \frac{\lambda}{c} e^{-cv} \frac{1}{v} e^{-(\lambda-c)v} = \frac{\lambda}{c} \frac{1}{v} e^{-\lambda v}. \quad (171)$$

Следовательно, $\max [g_1^*(x)]$ минимизируется, если мы выберем c наибольшим в интервале $(0, \lambda]$, или область R минимальна при $c = \lambda$. Подстановка $c = \lambda$ в (167) дает

$$h_1(x) = \lambda e^{\lambda v} e^{-\lambda x}, \text{ если } x \geq v, \quad (172)$$

$$= 0, \text{ если } x < v,$$

т. е. мы делаем выборку x' из исходной плотности распределения, данной в (165), затем прибавляем к x' константу v и новое x подставляем в (173), которое получается из (168), если положить $c = \lambda$.

$$g_1^*(x) = e^{-\lambda v} \frac{1}{x} \quad (x > v). \quad (173)$$

Заметим, что $g(x)$ — грубая оценка по методу Монте-Карло в (164) — имеет больший размах, чем $g_1^*(x)$. Действительно,

$$R[g_1^*(x)] = e^{-\lambda v} \frac{1}{v} - e^{-\lambda v} \frac{1}{\infty} = e^{-\lambda v} \frac{1}{v}, \quad (174)$$

а

$$R[g(x)] = \frac{1}{v} - 0 = \frac{1}{v}. \quad (175)$$

Дисперсию $g_1^*(x)$ мы рассмотрим далее.

2. Гамма-распределение. На рис. 11 видно, что $h_0(x)$ можно аппроксимировать некоторым *непрерывным асимметрическим* распределением. В приложении III.10 показано, что подходящим будет гамма-распределение, описанное в [Fisz, 1967, p. 151—154], с параметрами $p = 2$ и $b = 2/v$. Такая оценка обозначена $g_2^*(x)$.

3. Экспоненциальное распределение с модифицированным параметром. В [Kahn and Marshall, 1953, p. 272] отмечено, что во имя простоты $h_0(x)$ можно аппроксимировать той же функцией плотности, что и $f(x)$ в (154), но с новыми значениями параметров³⁰. Легко показать, что вместо λ в исходном распределении нужно взять параметр $1/v$. Обозначим полученную оценку $g_3^*(x)$.

В заключение рассмотрим случай, когда $\xi(\lambda, v)$ неизвестна и нам нужно оценить ξ , т. е. при выводе трех аппроксимаций $h_0(x)$ мы будем предполагать $\xi(\lambda, v)$ неизвестной. Следовательно, мы не сможем выбрать параметры приближений так, чтобы дисперсия была минимальной. Вместо этого мы выберем параметры так, чтобы минимизировать размах. Если мы зафиксируем эти параметры, то не сможем сказать, какая из трех аппроксимаций значимых выборок дает наименьшую дисперсию. В [Kleijnen, 1968b, p. 186, 187, 198] показано, что такой выбор должен основываться на знании³¹ $\xi(\lambda, v)$. Этого достигнет экспериментатор, который с помощью метода Монте-Карло попытается оценить $\xi(\lambda, v)$. Фактически мы знаем *аналитическое решение* $\xi(\lambda, v)$;

оно дано в (1.2) приложения I. 1. С помощью этого решения можно вычислить дисперсию каждой из четырех оценок, $g(x)$ и $g_i^*(x)$ ($i = 1, 2, 3$). В [Kleijnen, 1968 b, p. 195 — 198] вычислена дисперсия ³² для $g(x)$, $g_1^*(x)$ и $g_2^*(x)$, в приложении III.9 мы вычисляем дисперсию $g_3^*(x)$; все эти выражения включают $\xi(\lambda, v)$. Это дает нам возможность составить табл. 1³³.

Таблица 1

Значение дисперсий трех оценок значимых выборок $g_i^*(x)$ ($i=1, 2, 3$)
в процентах от дисперсии $\text{var}[g(x)]$

λ	v	$\text{var}[g_i^*(x)]/\text{var}[g(x)] \times 100\%$		
		$i=1$	$i=2$	$i=3$
0,5	4	0,9%	31,4%	50,4%
0,5	3,5	1,4%	37,5%	59,9%
0,5	2	6,5%	76,6%	100,0%
0,25	8,5	0,7%	28,8%	46,1%

Табл. 1 показывает, что использование $g_1^*(x)$, основанное на смещенном экспоненциальном распределении, дает значительное понижение дисперсии. Две другие аппроксимации тоже дают существенное снижение дисперсии³⁴. Эти методы понижения дисперсии требуют дополнительного машинного времени, так как в случае значимой выборки мы конструируем более сложную функцию $g^*(x)$ вместо $g(x)$. Более того, для функции $g_2^*(x)$, основанной на гамма-распределении, мы должны генерировать две экспоненциальные величины, чтобы получить одну величину с гамма-распределением (см. приложение III.8). С помощью грубых оценок дополнительного машинного времени Клейнен [Kleijnen, 1968 b, p. 191, table 2] показал, что для $g_2^*(x)$ и $g_3^*(x)$ это время сильно снижает эффективность их работы. Применение функции $g_1^*(x)$ дает уменьшение дисперсии по сравнению с $g(x)$ от 3 до 17%.

Мы закончим этот раздел некоторыми ссылками, касающимися применения значимой выборки в исследованиях по методу Монте-Карло. В [Clark, 1961, p. 610—614] дано элементарное описание значимой выборки для оценки интеграла

$$\int_{-\infty}^1 \lambda e^{-\lambda x} dx. \quad (176)$$

Подобные примеры можно найти в [Haitsma and Oosterhoff, 1964, p. 29 — 30], [Hillier and Lieberman, 1968, p. 458 — 460], [Kahn and

Marshall, 1953, p. 270], [Molenaar, 1968, p. 121]. Далее в [King 1953, p. 50 — 51] кратко обсуждается метод значимой выборки (без этого названия) в «случайных блужданиях». Сравните это с примером из [Kahn and Marshall, 1953, p. 272], в котором оценивается вероятность прохождения ядерной частицы через экран. В [Evans, 1963] и [Шапиро, 1970] рассмотрено применение значимой выборки как части одного метода понижения дисперсии, предложенного авторами этих работ. В [Pugh, 1966] обсуждается последовательная процедура для оценки новой оптимальной плотности распределения $h_0(x)$. В [Clark, 1961, p. 618] и [Kahn and Marshall, 1953, p. 271] приведены примеры, показывающие, как неправильный выбор новой плотности распределения может привести к весьма большому *повышению дисперсии!* В [Kahn and Marshall, 1953, p. 273] доказано, что метод значимой выборки может понизить дисперсию оцениваемой дисперсии «отклика» $g(x)$ в (154). В [Kahn, 1955, p. 11 — 18] и [Kahn and Marshall, 1953, p. 271 — 272] предложена новая оптимальная плотность распределения $h(x, y)$ для случая двух *входных случайных переменных*, т. е. приводится (154) к (177):

$$\xi = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y) dx dy. \quad (177)$$

Пример (177) рассмотрен также в [Clark, 1961, p. 614 — 616]. Там же отмечено, что если несколько входных переменных и процесс, естественно, становится сложнее, то труднее найти новую подходящую плотность распределения и оценить эффективность значимой выборки. Это также проиллюстрировано на примере в [Clark, 1961, p. 616 — 619], касающемся автомашины со случайным временем работы и некоторой вероятностью поломки. Этот пример вводит нас в область моделирования, которую мы обсудим далее.

III.5.3. Значимая выборка в моделировании

В [Handscornb, 1968, p. 7] кратко обсуждаются различия между задачами метода Монте-Карло и имитационного моделирования, а также их связь со значимой выборкой. Рассматриваемая в указанной работе процедура значимой выборки, адаптированная к моделированию, не выглядит достаточно общей в применении к моделированию. Автор показывает трудности применения значимой выборки в моделировании. В [Clark, 1961, p. 616 — 619] обсуждается задача поломки автомобиля; задача массового обслуживания рассмотрена в [Ehrenfeld and Ben-Tuvia, 1962, p. 270 — 271]. В [Gürtler, 1969, p. 84 — 93] вновь демонстрируется эта трудность, так как для выбора новой плотности распределения требуется глубокое знание системы. Мой [Моу, 1965, 1966] предлагает *стандартный* тип новой плотности распределения до значимой выборки в моделировании. В этом случае нет необходимости в глубоком анализе системы при выборе плотности распределения (далее мы рассмотрим такой метод).

При моделировании мы получаем значение отклика, скажем \mathbf{x} , в одном имитационном опыте на основании последовательности случайных входных переменных, которые, в свою очередь, генерируются на основании последовательности случайных чисел r_t ($t=1, 2, \dots, m$).

Следовательно, \mathbf{x} есть некоторая функция, скажем g (эта функция определена машинной программой), случайных чисел. Следуя Мюю, сначала предположим, что m фиксировано, тогда выполняется (178).

$$\mathbf{x} = g(r_1, r_2, \dots, r_m) = g(\vec{R}), \quad (178)$$

где \vec{R} — вектор случайных чисел в имитационном опыте. Математическое ожидание \mathbf{x} определяется следующим выражением:

$$\begin{aligned} E(\mathbf{x}) &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} g(r_1, r_2, \dots, r_m) f(r_1, r_2, \dots, r_m) dr_1 dr_2 \dots dr_m = \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} g(\vec{R}) f(\vec{R}) d\vec{R}. \end{aligned} \quad (179)$$

Итак, (179) представляет собой многомерный аналог (154) и (155). Поскольку случайные числа r_t независимы, можно записать совместную плотность распределения в виде произведения частных плотностей, скажем f_t ($t=1, 2, \dots, m$) (см. I. 5, уравнение (4)). Далее, пусть эти плотности распределения все равны, допустим, $f_t(r_t) = s(r_t)$. Тогда

$$\begin{aligned} f(\vec{R}) &= f(r_1, r_2, \dots, r_m) = f_1(r_1) f_2(r_2) \dots f_m(r_m) = \\ &= s(r_1) s(r_2) \dots s(r_m). \end{aligned} \quad (180)$$

В I.5, уравнение (2), мы видели, что $s(r_t)$ удовлетворяет условиям

$$\begin{aligned} s(r_t) &= 1 \text{ для } 0 \leq r_t \leq 1, \\ &= 0 \text{ для остальных значений.} \end{aligned} \quad (181)$$

Сочетание (180) и (181) дает

$$\begin{aligned} f(\vec{R}) &= f(r_1, r_2, \dots, r_m) = 1, \text{ если все } r_t \text{ в } [0, 1], \\ &= 0, \text{ если любой из } r_t \text{ лежит вне } [0, 1]. \end{aligned} \quad (182)$$

После подстановки (182) в (179) мы получим

$$E(\mathbf{x}) = \int_0^1 \dots \int_0^1 g(\vec{R}) d\vec{R}. \quad (183)$$

Аналогично (157) можно выразить в (179) так:

$$E(\mathbf{x}) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \frac{g(\vec{R}) h(\vec{R})}{h(\vec{R})} h(\vec{R}) d\vec{R}. \quad (184)$$

Или, пользуясь менее общим выражением (183), можно записать:

$$E(\mathbf{x}) = \int_0^1 \dots \int_0^1 \frac{g(\vec{R})}{h(\vec{R})} h(\vec{R}) d\vec{R}. \quad (185)$$

Следовательно, если $h(\vec{R})$ есть совместная плотность распределения, то можно сделать выборку r_t из $h(\vec{R})$ и воспользоваться оценкой значимой выборки, например z , в (186) аналогично (158):

$$z = g^*(\vec{R}) = g(\vec{R})f(\vec{R})/h(\vec{R}) = \mathbf{x}f(\vec{R})/h(\vec{R}). \quad (186)$$

С помощью (182) можно (186) свести к (187):

$$z = g^*(\vec{R}) = g(\vec{R})/h(\vec{R}) = \mathbf{x}/h(\vec{R}). \quad (187)$$

Так, машинная программа генерирует отклик \mathbf{x} не из случайных чисел, т. е. случайных величин \vec{R} , имеющих плотность распределения $f(\vec{R})$, определенную в (182), а используя случайную величину \vec{R} с плотностью $h(\vec{R})$. Полученный отклик корректируется весовым множителем $1/h(\vec{R})$. Мы будем употреблять термин *значимые числа*, чтобы отличить случайные величины с плотностью $h(\vec{R})$ от псевдослучайных чисел³⁵. Дисперсия оценки значимой выборки дается выражением

$$\begin{aligned} \text{var}(z) &= E(z^2) - [E(z)]^2 = \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \frac{g^2(\vec{R}) f^2(\vec{R})}{h^2(\vec{R})} h(\vec{R}) d\vec{R} - [E(\mathbf{x})]^2, \end{aligned} \quad (188)$$

которое можно дальше упростить. Легко проверить, что если взять $h(\vec{R})$ в (188), равным

$$h_0(\vec{R}) = \frac{g(\vec{R})f(\vec{R})}{E(\mathbf{x})}, \quad (189)$$

то дисперсия $\text{var}(z)$ понизится до нуля. Значит, (189) дает *оптимальную* новую плотность распределения при условии, что $g(\vec{R}) \geq 0$ для всех \vec{R} , иначе (189) дает отрицательные плотности. Это не выглядит ограничением, так как $x = g(\vec{R})$ не может быть отрицательным. Заметим, что (189) — многомерный аналог (162).

Прежде чем продолжить обсуждение новой плотности распределения $h(\vec{R})$, обобщим определение отклика \mathbf{x} для *случайного* числа случайных чисел в одном опыте³⁶. Выражение (178) обобщается в этом общем случае как

$$\mathbf{x} = g(r_1, r_2, \dots, r_m). \quad (190)$$

Следовательно,

$$E(\mathbf{x}) = \sum_m E(\mathbf{x} | m = m) P(m = m), \quad (191)$$

где Σ пробегает все возможные значения m . Обозначим совместную плотность распределения m случайных величин $f_m(\vec{R})$ и функцию g в (190) для $m = m$ через g_m . Тогда

$$E(\mathbf{x} | m = m) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} g_m(\vec{R}) f_m(\vec{R}) d\vec{R}. \quad (192)$$

Аналогично (186) можно определить z_m , оценку для данного m и для значимой выборки, как

$$z_m = g_m(\vec{R}) f_m(\vec{R}) / h_m(\vec{R}), \quad (193)$$

где $h_m(\vec{R})$ — совместная плотность распределения, из которой мы выбираем \vec{R} . Из (192) и (193) следует, что

$$E(z_m) = E(\mathbf{x} | m = m). \quad (194)$$

Перед началом опыта мы не знаем m -реализации случайной величины m . Значит, математическое ожидание отклика в опыте надо определить по всем m , т. е. на основании (194) и (191) можно записать:

$$E_m[E(z_m)] = E_m[E(\mathbf{x} | m = m)] = E(\mathbf{x}). \quad (195)$$

Следовательно, имитационный опыт при значимой выборке приводит к несмещенной оценке. Мы заметим, что z_m может генерироваться, даже если мы не знаем m до конца опыта, потому что, как мы увидим, значимые числа выбираются независимо, т. е. их совместная функция плотности может быть записана в виде произведения частных функций плотности, скажем $v_t(r_t)$, т. е.

$$h_m(\vec{R}) = \prod_{i=1}^m v_i(r_i), \quad (196)$$

где каждый сомножитель $v_t(r_t)$ не содержит m . Итак r_t выбирается из $v_t(r_t)$ независимо от $r_{t'}$ ($t' \neq t$). В конце опыта можно вычислить $h_m(\vec{R})$ на основании (196) и подставить результат в (193), чтобы взвесить отклик $g_m(\vec{R})$. Дисперсия z_m следует из (193) и записывается аналогично (188):

$$\text{var}(z_m) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \frac{g_m^2(\vec{R}) f_m^2(\vec{R})}{h_m^2(\vec{R})} h_m(\vec{R}) d\vec{R} - [E(z_m)]^2. \quad (197)$$

Математическое ожидание дисперсии по m равно

$$E_m \text{var}(z_m) P(m = m). \quad (198)$$

Следовательно, аналогично (189) мы берем $h_m(\vec{R})$ равным

$$h_{m_0}(\vec{R}) = \frac{g_m(\vec{R}) f_m(\vec{R})}{E(\mathbf{x} | m = m)}, \quad (199)$$

это снижает дисперсию $\text{var}(z_m)$ до нуля и, следовательно, понижает (198) также до нуля. Итак, аппроксимацию оптимальной новой плотности распределения можно обосновать результатами для *нестационарного* m , данного в (199), или его эквивалента (189). Далее мы рассмотрим аппроксимацию, предложенную Моём.

Из выражений (189) и (199) мы видим, что для того, чтобы определить $h_0(\vec{R})$, необходимо знать математическое ожидание отклика и явный вид функции g , которая фактически заложена в сложную машинную программу. Следовательно, нам нужно найти *приближение* для $h_0(\vec{R})$. Мой [Моу, 1965, р. 107; 1966, р. 21] предлагает конструировать *машинную программу* так, чтобы отклик увеличивался с ростом случайных чисел. Мы показали (43), (44), как это можно сделать. Такое свойство приводит к тому, что по крайней мере приближенно выполняется

$$g(\vec{R}_a) < g(\vec{R}_b), \quad (200)$$

где \vec{R}_a и \vec{R}_b представляют собой m - a -й и m - b -й векторы соответственно; $0 \leq a < b \leq 1$. Из (182) вытекает, что

$$f(\vec{R}_a) = f(\vec{R}_b) = 1. \quad (201)$$

Следовательно, подстановка (200) и (201) в (189) при условии $E(\mathbf{x}) > 0$ приводит к тому, что справедливо неравенство

$$h_0(\vec{R}_a) < h_0(\vec{R}_b). \quad (202)$$

Предположим, что мы выбираем значимые числа независимо, так что выполняется (196); тогда (202) выполняется, если *индивидуальные* или безусловные плотности распределений v_t удовлетворяют условию

$$v_t(a) < v_t(b) \text{ для } a < b. \quad (203)$$

Поэтому Мой предлагает формировать новые функции плотности $v_t(r_t)$, *возрастающие* с ростом r_t . Эти возрастающие плотности распределения заменяют старые одинаковые функции $s(r_t)$ в (181).

Очевидно, что существует много возрастающих функций плотности. Мы хотим выбрать ту, которая дает наибольшее понижение дисперсии и не требует много машинного времени для выборки из нее и вычисления весов $f(\vec{R})/h(\vec{R})$. Далее мы увидим, что функция может иметь такую форму, что ее оптимальные параметры вычисляемы без большой работы. Мой [Моу, 1965, р. 107] исследует несколько таких функций. На основании своих экспериментов Мой [Моу, 1965, р. 119] нашел подходящей только одну новую плотность распределения ³⁷, которая приводится также в [Моу, 1966, р. 22]:

$$\begin{aligned} v_t(r_t) &= \alpha_t^t (\ln \alpha_t) / (\alpha_t - 1) \text{ для } 0 \leq r_t \leq 1 (\alpha_t > 1), \\ &= 0 \text{ для всех остальных значений.} \end{aligned} \quad (204)$$

Определив конкретный *вид* аппроксимации, надо вычислить *параметры*, т. е. α_t . В задаче Монте-Карло из III.5.2 были определены

параметры, минимизирующие размах (так как невозможно было минимизировать дисперсию), но в сложных имитационных исследованиях мы не можем минимизировать размах. Мой [Мой, 1965, р. 109 — 113] предлагает следующий подход.

Уменьшим число параметров в новой плотности распределения, фиксируя α_t , т. е. $\alpha_t = \alpha$ в (204). Значение α выберем так, чтобы дисперсия z , оценки значимой выборки, стала минимальной. Дисперсия z дается в (188), где $h(\vec{R})$ зависит только от α . Нужно выбрать α так, чтобы выполнялось³⁸

$$\begin{aligned} \frac{\partial \text{var}(z)}{\partial \alpha} &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} g^2(\vec{R}) f^2(\vec{R}) \frac{\partial [1/h(\vec{R}, \alpha)]}{\partial \alpha} d\vec{R} = \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} g^2(\vec{R}) f^2(\vec{R}) \frac{1}{h^2(\vec{R}, \alpha)} \left[-\frac{\partial h(\vec{R}, \alpha)}{\partial \alpha} \right] d\vec{R} = \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \frac{g^2(\vec{R}) f(\vec{R}) \left[-\frac{\partial h(\vec{R}, \alpha)}{\partial \alpha} \right]}{h^2(\vec{R}, \alpha)} f(\vec{R}) d\vec{R} = 0. \end{aligned} \quad (205)$$

Интеграл в (205) можно интерпретировать как среднее от

$$W(\vec{R}, \alpha) = \frac{g^2(\vec{R}) f(\vec{R}) \left[-\frac{\partial h(\vec{R}, \alpha)}{\partial \alpha} \right]}{h^2(\vec{R}, \alpha)}, \quad (206)$$

где \vec{R} имеет плотность распределения $f(\vec{R})$. Но (205) означает, что среднее $W(\vec{R}, \alpha)$ равно 0. Поскольку мы не можем решить (205), Мой предлагает *выборочный* эквивалент, т. е.

$$\bar{W}(\vec{R}, \alpha) = \sum_{i=1}^{n_1} W(\vec{R}_i, \alpha) / n_1 = 0, \quad (207)$$

где \vec{R}_i обозначает вектор случайных чисел, генерируемый из $f(\vec{R})$ в i -м имитационном опыте ($i = 1, \dots, n_1$). Мы можем положить $f(\vec{R}) = 1$ и в (204) подставить $\alpha_t = \alpha$, т. е.

$$h(\vec{R}, \alpha) = \prod_{t=1}^m \alpha^{r_t} (\ln \alpha) / (\alpha - 1) = \alpha^{\sum_{t=1}^m r_t} [(\ln \alpha) / (\alpha - 1)]^m. \quad (208)$$

Затем Мой [Мой, 1965, р. 112] показал, что (207) сводится к следующему выражению:

$$\sum_{i=1}^{n_1} \frac{(\sum_j r_{ij}) (\ln \alpha) (\alpha - 1) + m_i (\alpha - 1 - \alpha \ln \alpha)}{\alpha^{\sum_j r_{ij}}} x_i^2 = 0, \quad (209)$$

где Σ_j обозначает $\sum_{i=1}^{m_j}$ и $x_i = g(\vec{R}_i)$ — отклик в i -м опыте. Так можно получить α из численного решения (209). Для оставшихся $n_2 (= n - n_1)$ опытов можно взять значимую выборку с новой плотностью распределения (208) с оптимальным параметром $\hat{\alpha}$, вычисленным из (209).

Мой [Моу, 1965, р. 117, 122 — 123] предложил важный вариант усовершенствования описанной выше процедуры, в котором α_i в (204) — это константа для *каждого типа входной переменной*. Поэтому если мы имеем K типов входных переменных, то (205) примет вид

$$\frac{\partial \text{var}(z)}{\partial \alpha_k} = 0 \quad (k = 1, 2, \dots, K). \quad (210)$$

Вместо выражения (207) мы должны в данном случае решить систему K уравнений, что потребует дополнительного машинного времени. Поэтому Мой [Моу, 1965, р. 125] предлагает новую плотность только для того типа переменных, которые оказывают наибольшее влияние на отклик, или

$$h(\vec{R}, \alpha) = f(\vec{R}_2, \vec{R}_3, \dots, \vec{R}_K) h_1(\vec{R}_1, \alpha), \quad (211)$$

где полный вектор случайных чисел \vec{R} разложен на K частей в соответствии с K типами входных переменных и только переменная первого типа выбирается из новой плотности распределения. Оптимальное значение α можно вычислить из (205). Интересной темой дальнейших исследований может быть применение «градиентного метода» Пью в сложных имитационных исследованиях. Пью [Pugh, 1966] применил метод *наискорейшего спуска* для *последовательного* оценивания новой оптимальной плотности распределения.

Мой применял различные типы плотностей распределения к некоторым системам. Применяя независимый от системы вариант (208) к одноканальной системе массового обслуживания, Мой [Моу, 1965, р. 118; 1966, р. 29] получил снижение дисперсии от 33 до 54%. Он оценил оптимальный параметр α_0 на основании 175 наблюдений, т. е. n_1 в (209) было равно 175. Дополнительные эксперименты, осуществленные Моем [Моу, 1965, р. 126 — 128], показали, что снижение дисперсии не очень чувствительно к флуктуациям в оценке α_0 , поэтому для оценки α_0 достаточно небольшого числа опытов. Применение того же варианта, не зависящего от системы, к сложной системе *товарообмена в порту* не дало понижения дисперсии, наоборот, дисперсия увеличилась. Использование отдельной функции плотности для каждого из типов входных переменных дало понижение дисперсии на 43% и 57% для двух вариантов этой сложной системы. Новая плотность распределения только для одной входной переменной, как в (211), дает понижение дисперсии на 30 и 39% для двух вариантов этой же системы. Параметры были оценены по большому числу опытов (180 или 360).

Эксперименты Моя [Моу, 1965, р. 129] показывают, что *оптимальное значение параметра $\hat{\alpha}_0$* остается относительно *постоянным* для различных типов систем и равно примерно 1,12. Поэтому он предлагает

начинать моделирование не со случайных чисел, т. е. $f(\vec{R}) = 1$, а со значимых, имеющих плотности распределения $h(\vec{R})$ в виде (208) при $\alpha = 1, 12$. (Если есть несколько типов входных переменных, то на основании $h(\vec{R})$ генерируется наиболее значимый тип.) Можно добавить, что если мы ранее провели моделирование упрощенной системы, то можно вычислить α_0 на основе экспериментов с этой системой с помощью уравнения, аналогичного (209). Итак, на первом этапе мы ставим опыты, применяя «искаженную» функцию плотности $h(\vec{R}, \alpha)$, где α либо равно 1, 12, либо α оценивается из предыдущих опытов с упрощенной системой. Далее мы можем оценить α_0 на основании опытов первого этапа и воспользоваться этой оценкой на втором этапе. (Укажем, что при этом подходе $f(\vec{R})$ в (206) не равно единице, а обозначает плотность, «искаженную» на первом этапе.)

Обе процедуры — рассматриваемая в предыдущем параграфе и исходная — являются *двухступенчатыми*. На первом этапе эксперименты генерируются для $f(\vec{R}) = 1$ или $h(\vec{R}, \alpha)$, где α равно 1, 12 или оценено на основании предыдущих опытов с упрощенной системой; α_0 оценивается из опытов первого этапа. Далее, n_2 ($= n - n_1$) дополнительных опытов генерируется с помощью значимой выборки с последующей оценкой α_0 . Сначала n_1 опытов дают средний отклик \bar{x}_1 , далее, n_2 опытов дают среднее \bar{x}_2 ; несмещенная оценка среднего отклика получается так:

$$\bar{x} = \omega \bar{x}_1 + (1 - \omega) \bar{x}_2, \quad (212)$$

где ω — вес среднего опытов первой группы. Дисперсия взвешенного среднего \bar{x} определяется выражением

$$\text{var}(\bar{x}) = \omega^2 \text{var}(\bar{x}_1) + (1 - \omega)^2 \text{var}(\bar{x}_2) + 2\omega(1 - \omega) \text{cov}(\bar{x}_1, \bar{x}_2), \quad (213)$$

где ковариационный член объясняется тем, что \bar{x}_2 есть оценка значимой выборки, зависящая от оценки оптимального параметра α_0 , который, в свою очередь, оценен на основании первых экспериментов, дающих \bar{x}_1 . Мы можем оценить дисперсии $\text{var}(\bar{x}_1)$ и $\text{var}(\bar{x}_2)$ на основании n_1 и n_2 повторов соответственно, но мы не можем оценить из этих повторов $\text{cov}(\bar{x}_1, \bar{x}_2)$. Мой [Моу, 1965, р. 114] предположил, что эта ковариация очень мала и ею можно пренебречь; он провел эксперимент, чтобы проверить это предположение, и оценил корреляцию между \bar{x}_1 и \bar{x}_2 на основании 67 наблюдений, где \bar{x}_1 и \bar{x}_2 — средние для простой системы массового обслуживания, вычисленные на основании 25 опытов. Полученный коэффициент, равный 0,101, оказался незначимым [Моу, 1965, р. 132]. Если не принимать во внимание ковариацию (213), то очень просто определить оптимальный вес, равный

$$\omega_0 = \frac{\text{var}(\bar{x}_2)}{\text{var}(\bar{x}_1) + \text{var}(\bar{x}_2)} = \frac{n_1}{n_1 + n_2 \text{var}(x_{1i}) / \text{var}(x_{2i})}, \quad (214)$$

где x_{1i} и x_{2j} — i -й и j -й опыты в группах 1 и 2 соответственно. Мой [Мой, 1965, р. 115; 1966, р. 27] указывает, что для получения этих оптимальных оценок можно брать оценки дисперсий. Мы наблюдали в противоположность Мюю, что таким образом оцененный вес приводит к несмещенной взвешенной оценке среднего:

$$E(\bar{x}) = E[E(\bar{x} | w_0)] = E[\mu_x] = \mu_x. \quad (215)$$

Дисперсия, данная в (216), вычисляется на основе (4.1) в приложении III.4.

$$\begin{aligned} \text{var}(\bar{x}) &= E[\text{var}(\bar{x} | w_0)] + \text{var}[E(\bar{x} | w_0)] = E[w_0^2 \text{var}(\bar{x}_1) + \\ &+ (1 - w_0)^2 \text{var}(\bar{x}_2) + 2w_0(1 - w_0) \text{cov}(\bar{x}_1, \bar{x}_2)] + \text{var}[\mu_x] = \\ &= E(w_0^2) \text{var}(\bar{x}_1) + E[(1 - w_0)^2] \text{var}(\bar{x}_2) + 2E[w_0(1 - w_0) \text{cov}(\bar{x}_1, \bar{x}_2)]. \end{aligned} \quad (216)$$

Поэтому (216) есть аналог (213); последнее уравнение действительно для неслучайного веса w_0 . Даже если мы отбросим ковариационный член в (216), оценка (217) будет смещенной.

$$\hat{\text{var}}(\bar{x}) = w_0^2 \hat{\text{var}}(\bar{x}_1) + (1 - w_0)^2 \hat{\text{var}}(\bar{x}_2). \quad (217)$$

Так как w_0 зависит от дисперсии $\hat{\text{var}}(\bar{x}_1)$ и $\hat{\text{var}}(\bar{x}_2)$, то w_0 определяется с помощью простого аналога (214).

Следовательно,

$$E[w_0^2 \hat{\text{var}}(\bar{x}_1)] \neq E(w_0^2) E[\hat{\text{var}}(\bar{x}_1)]. \quad (218)$$

Аналогичное соотношение выполняется для второго члена в (217). Поэтому оценка веса w_0 — это еще один источник смещения в оценке $\text{var}(\bar{x})$ (кроме смещения за счет ковариационного члена, которым мы пренебрегаем). Поэтому можно положить $\text{var}(x_{1i}) = \text{var}(x_{2j})$ в (214) и w_0 оценить на основании либо априорной информации, либо части из n опытов, а $\text{var}(\bar{x}_1)$ и $\text{var}(\bar{x}_2)$ в (217) — из остальных опытов. Определение доверительного интервала для среднего — трудное дело. (Нужно учитывать, что дисперсии x_1 и x_2 не равны.)

Мой ограничил свои исследования значимой выборки системами с остановом, когда каждый имитационный опыт дает единственный отклик. Мы кратко рассмотрели последствия этих ограничений, следуя [Kleijnen, 1969, р. 291 — 293]. Если мы моделируем систему без остановов согласно сказанному в главе II, можно использовать несколько удлиненных опытов вместо многих повторных. В методе значимой выборки первые опыты применяются для оценки оптимального параметра α_0 , и, следовательно, если их мало, то оценка α_0 будет плохой и повысится уровень дисперсии (хотя очевидно, что отклик останется несмещенным). Если нас интересует несколько откликов, то каждый опыт должен давать несколько значений откликов. Много-

мерный отклик означает, что на основании уравнения, подобного (207), вычисляется параметр α_0 для каждой из переменных выхода. Поэтому каждая такая переменная дает свою величину для $\hat{\alpha}_0$. Тем не менее мы берем одно значение α в $h(\vec{R}, \alpha)$. Остается неизвестным, дает ли снижение дисперсии компромиссно выбранное α . Если различные выходные переменные реагируют в разных направлениях на случайные числа \vec{R} , то становится трудно найти подходящую *параметрическую* форму $h(\vec{R})$ при аппроксимации оптимальной плотности $h_0(\vec{R})$. Мы можем составить программу таким образом, чтобы время ожидания требований возрастало с ростом случайных чисел и выполнялось выражение (200). Однако это вызывает уменьшение времени простоя машин при росте случайных чисел, и в итоге для этой выходной переменной условие (200) не выполняется. Выбор параметрической формы для $h(\vec{R})$ весьма непросто, так как мы даже не знаем, какую функцию — возрастающую или убывающую — мы выбираем. Случай многомерных откликов не вызывает проблем при применении других, уже обсуждавшихся методов понижения дисперсии, но в методе значимой выборки приводит к упомянутым *конфликтным требованиям* к новой плотности распределения $h(\vec{R}, \alpha)$.

Мы видели, что метод значимой выборки довольно сложен. Он требует дополнительного *машинного времени* для оценки оптимального параметра α в уравнении типа (209). Кроме того, дополнительное время нужно для генерирования значимых чисел из чисто случайных (эти значимые числа используются для генерирования входных переменных). Наконец, надо вычислять *вес* $f(\vec{R})/h(\vec{R})$ в (186). Дополнительное машинное время для системы товарообмена в порту в случае значимой выборки можно найти в [Моу, 1966, р. 33]. Если значимая выборка применяется ко всем четырем типам входов, то это дает понижение дисперсии около 65%, а с учетом дополнительного времени — около 43%. Если же значимая выборка применяется только к одному типу событий в этой системе, то понижение уменьшается с 69 до 67%. Далее будут рассмотрены два метода понижения дисперсии, которые почти не требуют дополнительного времени.

III.6. ДОПОЛНЯЮЩИЕ ВЕЛИЧИНЫ

Мы увидим, что в методе с дополняющими величинами мы стремимся создать отрицательную *корреляцию* между наблюдениями, генерируя одно наблюдение с помощью случайного числа r , а другое — с помощью его «дополняющего» партнера — числа $(1 - r)$. Этот метод был предложен в [Hammersley and Morton, 1956] для оценивания интеграла с помощью метода Монте-Карло. Авторы оценили интеграл, применяя стратифицированную выборку; выборка из разных слоев производилась так, чтобы наблюдения не были независимыми, а имели между собой отрицательную корреляцию (см. примечание 13 в III.2.4). При случайных числах r и $(1 - r)$ отрицательная корреляция полу-

чалась только для двух слоев. Интересующихся проблемами дополняющих величин в методе Монте-Карло мы отсылаем к [Hammersley and Morton, 1956], дополнительные ссылки на старые работы можно найти в [Andreasson, 1972b, p. 2]. Кратко применение метода Монте-Карло обсуждается в [Haitsma and Costerhoff, 1964, p. 34—35], [Handscornb, 1968, p. 9 — 10], [Newman and Odell, 1971, p. 64 — 65] и [Tocher, 1963, p. 117—118]. Очень простой вариант метода дополняющих величин был применен к одной задаче Монте-Карло в [Hil-lier and Lieberman, 1968, p. 461]. Применение дополняющих величин к оценке не только средней, а всего распределения описано в [Burt et al., 1970, p. 442].

Уже в 1958 г. Харлинг [Harling, 1958, p. 12 — 13] предложил этот метод понижения дисперсии для *моделирования*. Однако он не объяснил, как это можно сделать. Эренфельд и Бен-Тувия (Ehrenfeld and Ben-Tuvia, 1962, p. 273 — 274) предложили применять метод дополняющих величин к простым системам массового обслуживания. Наконец, Точер [Tocher, 1963, p. 173 — 175] рекомендует простой способ получения отрицательной корреляции между двумя «наблюдениями», т. е. между двумя повторными опытами при моделировании сложных систем. По его мнению, следует генерировать один опыт, используя случайные числа r_1, r_2, \dots , и другой дополняющий, используя случайные числа $(1 - r_1), (1 - r_2), \dots$ Вспомним, что в I.6.4 (19) мы доказали, что $(1 - r)$ есть тоже случайное число. Точер предположил, что с помощью этой процедуры можно получить отрицательную корреляцию между двумя откликами. Мы вернемся к этому предположению чуть позже. Отрицательная корреляция между двумя откликами желательна потому, что она уменьшает дисперсию оцениваемого отклика. Для μ , среднего отклика системы, оценку можно написать в виде

$$\bar{x} = \frac{1}{2} (x_1 + x_2), \quad (219)$$

где x_1 и x_2 — отклики в первом и втором опытах соответственно. Дисперсия оценки \bar{x} равна:

$$\text{var}(\bar{x}) = \frac{1}{4} \{ \text{var}(x_1) + \text{var}(x_2) + 2 \text{cov}(x_1, x_2) \}. \quad (220)$$

Следовательно, дисперсия \bar{x} уменьшается, если x_1 и x_2 имеют отрицательную корреляцию или отрицательную величину $\text{cov}(x_1, x_2)$, что эквивалентно. Теперь мы выясним, почему последовательность случайных чисел $(1 - r_1), (1 - r_2), \dots$ создает отрицательную корреляцию.

1. Предположим, что отклик x зависит от единственного значения одной случайной входной переменной y (вместо последовательности различных случайных входных переменных, как это действительно бывает при моделировании). Далее, предположим, что x есть *монотонная* функция от y , например монотонно возрастающая функция $g_1(y)$, т. е.

$$g_1(y_1) > g_1(y_2), \text{ если } y_1 > y_2. \quad (221)$$

Мы знаем, что если y генерируется случайным числом r с помощью метода обращения, определенного (13) в 1.6.4, то y будет монотонно возрастающей функцией, скажем g_2 от r^{39} . Даже если y генерируется не методом обращения, то все равно обычно получается такой же тип функции ⁴⁰. Итак,

$$g_2(r_1) > g_2(r_2), \text{ если } r_1 > r_2. \quad (222)$$

Следовательно, отклик x есть монотонно возрастающая функция, например g_3 , случайного числа r , т. е.

$$g_3(r_1) > g_3(r_2), \text{ если } r_1 > r_2, \quad (223)$$

или большие значения x получаются из больших значений r . Но «большая» величина r приводит к «малой» дополняющей случайной величине

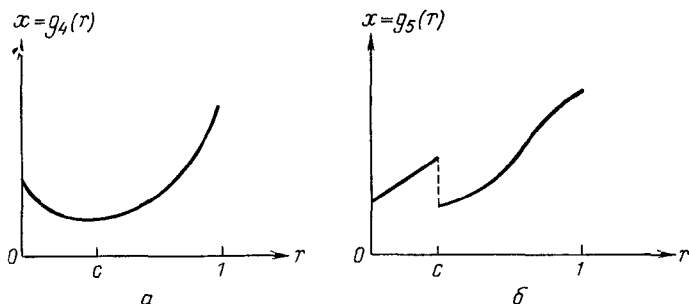


Рис. 12. Немонотонная зависимость отклика x от случайного числа r

$(1 - r)$, что, в свою очередь, приводит к малым значениям x . Итак, «большим» значениям (большим, чем математическое ожидание μ) $x = g_3(r)$ соответствуют «малые» значения (т. е. меньшие, чем величина μ) дополняющего отклика, например $x_a = g_3(1 - r)$, а это значит, что x и x_a отрицательно коррелируют между собой (индекс a обозначает дополняющую величину). Формальное доказательство факта отрицательной корреляции можно найти в [Andreasson, 1972 b, p. 4]. Оба отклика имеют ту же среднюю величину, так как генерируются с помощью случайной величины с тем же самым распределением. Или, более формально,

$$F(x) = \int_{-\infty}^{\infty} g_3(r) f(r) dr = \int_0^1 g_3(r) dr. \quad (224)$$

Обозначая дополняющее случайное число $(1 - r)$ через z , мы получим

$$E(x_a) = \int_{-\infty}^{\infty} g_3(z) f(z) dz = \int_0^1 g_3(z) dz, \quad (225)$$

где z имеет ту же самую плотность распределения f , что и r , это было показано в 1.6.4 (19).

Примеры немонотонных зависимостей между откликом x и случайным числом r приведены на рис. 12, а и 12, б. Если мы знаем значение c ,

то можно определить дополняющие случайные числа, скажем r_2 . Так,

$$\begin{aligned} r_2 &= c - r, \text{ если } 0 \leq r \leq c, \\ &= 1 + c - r, \text{ если } c < r \leq 1. \end{aligned} \quad (226)$$

Легко показать, что r_2 — случайное число и дает отклик, который отрицательно коррелирует с $x = g_a(r)$ и $x = g_b(r)$ (см. также [Kleijnen, 1968 а, р. 28]). К сожалению, при моделировании сложных систем мы не знаем c и не можем применить (226). Поэтому мы предполагаем, что существует *монотонная* зависимость между откликом и случайным числом⁴¹. Это предположение выполняется во многих приложениях, например в системах массового обслуживания, где время ожидания требования растет с уменьшением времени между поступлениями и с ростом времени обслуживания. В этих системах, однако, мы имеем дело не с одним случайным входом, а с последовательностью входов; кроме того, типы входов могут быть разными. Мы рассмотрим эти более сложные случаи.

2. При моделировании отклик зависит от последовательности значений случайной входной переменной, т. е. вместо $x = g_3(r)$ мы имеем⁴²

$$x = g_6(r_1, r_2, \dots, r_m). \quad (227)$$

Можно рассуждать следующим образом. Пусть x — средний отклик (т. е. среднее время ожидания) за имитируемый период, т. е.

$$x = \sum_{j=1}^m x_j / m. \quad (228)$$

Ковариация между x и откликом x_a определяется следующим выражением:

$$\text{cov}(x, x_a) = \sum_{j=1}^m \sum_{h=1}^m \text{cov}(x_j, x_{ah}) / m^2, \quad (229)$$

где x_{ah} означает h -й отклик в дополняющем опыте. Ни одну из m^2 ковариаций нельзя предположить априорно равной нулю, так как для каждой j и h мы знаем, что x_j и x_{ah} частично используют одну и ту же последовательность случайных чисел. Мы можем предположить, что отдельные наблюдения x_j и x_{ah} сильно зависят от наиболее часто встречающихся значений случайных чисел, т. е.

$$x_j = g_7(r_j). \quad (230)$$

Однако это означает, что можно считать отдельные последовательные наблюдения независимыми и, следовательно, можно руководствоваться рассуждениями из пункта 1. Мы считаем, что при моделировании сложных систем невозможно показать аналитически, что применение дополняющих величин приводит к отрицательной корреляции между опытами. Неудивительно, что Точер [Tocher, 1963, р. 174] допускает, что «в некотором смысле предположение об отрицательной корреляции неоправданно», и Пэйдж [Page, 1965, р. 303] замечает: «Насколько большей будет отрицательная корреляция между новыми

наблюдениями $[x_{an}]$ и исходными $[x_j]$, можно определить только опытным путем». Андреассон [Andreasson, 1972 b, p. 4 — 6, 15 — 30] показывает аналитически, что в одноканальной системе массового обслуживания дополняющие величины дают отрицательную корреляцию. Можно также сослаться на [Radema, 1969, p. 18 — 20]. Мы опишем результаты применения метода дополняющих величин к некоторым системам, чтобы продемонстрировать возникающую отрицательную корреляцию. Сначала, однако, введем еще одно усложнение, а именно допустим несколько типов входных переменных.

3. Пусть мы имеем K типов случайных входных переменных, например время обслуживания, время между поступлениями требований и т. д. Точер [Tocher, 1963, p. 174, 175] считает, что в дополняющем опыте только одна переменная генерируется как дополняющая, «генерирование остальных переменных из других распределений начинается с тех же самых начальных значений... Лучше не менять более одной переменной, так как такое изменение может ослабить корреляцию, вызванную изменением первой переменной». Далее он предлагает подход: 2^K -факторный эксперимент, т. е. если K равно, скажем, 3 то число опытов равно: $2^K = 2^3 = 8$, каждый из 8 опытов проводится в соответствии с табл. 2.

Т а б л и ц а 2

2^K -факторный эксперимент для генерирования дополняющих опытов при $K=3^*$

Входная переменная	Число опытов в серии							
	1	2	3	4	5	6	7	8
1	+	-	+	-	+	-	+	-
2	+	+	-	-	+	+	-	-
3	+	+	+	+	-	-	-	-

* «-» означает дополняющие случайные числа, «+» означает общие случайные числа.

Плюс в строке табл. 2 означает, что эта переменная генерируется с теми же случайными числами, которые применялись для этой переменной в других экспериментах; минус означает, что генерируется дополняющая величина. Например, сравнение опытов 1 и 2 в серии из 8 опытов показывает, что переменная 1 генерируется как дополняющая, в то время как переменные 2 и 3 генерируются на основе тех же случайных чисел в обоих опытах. Опыты в разных сериях независимы, так как для каждой серии берутся новые случайные числа.

Для того чтобы увидеть результаты подхода Точера, возьмем систему с большим числом входных случайных переменных, т. е. когда K велико. Мы рассмотрим только опыты 1 и 2 в серии из 2 опытов. При подходе Точера одна переменная генерируется как дополняющая, в то время как все остальные $(K - 1)$ входных переменных генерируются с теми же векторами случайных чисел \vec{R}_k ($k' = 2, 3, \dots, K$), которые брались в первом опыте. Поэтому если входная переменная 1 не оказывает доминирующего влияния на отклик, то можно ожидать, что отклик опыта 2 близок к отклику опыта 1. Следовательно, между

опытами. Это возможно положительная, а не отрицательная корреляция. В качестве примера с *небольшим* значением K мы рассмотрим однокликовую систему массового обслуживания. Если в опыте 1 получится много малых значений входной переменной 1 (время между двумя требованиями), а для входной переменной 2 (время обслуживания) — много больших значений, то в результате получим большое значение среднего времени ожидания. Если же в отличие от предположений Точера взять *оба* входа в опыте 2 из дополняющих величин, то мы получим много больших значений времени между требованиями и много малых значений для времени обслуживания, что в результате даст низкое значение среднего времени ожидания. Этот пример вновь демонстрирует, что лучше не следовать подходу Точера, а вместо этого генерировать входные переменные, дополняющие в парных опытах. Другой *простейший* пример «системы» с двумя типами входных переменных — система с откликом, зависящим только от одного значения каждой входной переменной, т. е. $x = g(r_1, r_2)$. Для некоторых видов функций g мы вычисляли корреляцию для случаев, когда оба входа либо один из них — дополняющие. Эти примеры показали, что наиболее высокую отрицательную корреляцию мы получаем, если делаем оба входа дополняющими. Для дальнейшего изучения 2^K -факторного подхода отсылаем к [Andreasson, 1972 b, p. 9 — 10], [Kleijnen, 1968 a, p. 4 — 7] и [Radema, 1969, p. 16 — 18]. Радема аналитически и экспериментально показал, что подход Точера в случае одноканальной системы массового обслуживания дает плохие результаты. Андреассон [Andreasson, 1971b, p. 8] воспользовался идентичным временем поступления и дополняющим временем обслуживания в моделировании простой телефонной системы обслуживания; результирующая корреляция между двумя парными опытами оказалась положительной. (В частном сообщении Точер согласился с выводом, что 2^K -факторный подход «не работает», хотя остаются разногласия по поводу причин, приводящих к этой неудаче.)

Обобщая сказанное, мы предлагаем вместо 2^K -коррелированных опытов генерировать пары опытов. Один опыт генерируется на основании K векторов случайных чисел \vec{R}_k ($k = 1, 2, \dots, K$), где \vec{R}_k служит для генерирования k -й входной переменной и имеет m_k элементов; другой опыт генерируется на основании векторов дополняющих случайных чисел ($\vec{I}_k - \vec{R}_k$), где \vec{I}_k — вектор, состоящий из m_k элементов. Чтобы каждый тип входной переменной имел свой собственный вектор случайных чисел, каждая из K входных переменных должна иметь *свой собственный генератор* случайных чисел. (Мы вернемся к таким генераторам в этом параграфе.)

Для простоты рассмотрим только пример с неслучайным m_k . В некоторых системах m_k бывает случайным, так что в дополняющем опыте понадобится больше (либо меньше) случайных чисел, чем в исходном опыте. Соглашаясь с Точером [Tocher, 1963, p. 174], мы предполагаем, что это не оказывает серьезного воздействия, хотя и уменьшает корреляцию. Стохастический характер m_k приводит нас к более важной проблеме *синхронизации*. Если i -е случайное число r_i генери-

рует определенное событие (например, поступление требования j), то в дополняющем опыте $(1 - r_i)$ должно генерировать то же самое событие (т. е. не прибытие требования j' , где $j' \neq j$, и без времени обслуживания). Такая синхронизация повышает желаемую отрицательную корреляцию. Синхронизация упрощается, если каждая входная переменная k имеет свой собственный генератор (сравните с обсуждением синхронизации [Andreasson, 1971b, p. 7 — 8] и [Emshoff and Sisson, 1971, p. 197]). При рассмотрении общих случайных чисел мы вернемся к проблеме синхронизации и обсудим пример.

Далее мы рассмотрим более подробно несколько *методов* реализации *дополняющих* случайных чисел и случайных величин.

(а) Генерируем случайное число r и *вычитаем* его из 1. Вычитание — единственное действие, которое надо выполнить для каждого случайного числа, что требует лишь одной дополнительной операции. По сравнению с общим временем, затрачиваемым на опыт, это дополнительное время незначительно. Если машинная память достаточна, то можно сохранить исходную последовательность случайных чисел. Тогда в дополняющем опыте мы не генерируем r вновь, а только вычисляем $(1 - r)$; эта процедура применялась в [Andreasson, 1971b, p. 7] и [Garman, 1971, p. 9 — 10].

(б) Если есть *мультипликативный* генератор случайных чисел, т. е.

$$z_i = az_{i-1} \pmod{m} \quad (i = 1, 2, \dots), \quad (231)$$

$$r_i = z_i/m, \quad (232)$$

то дополняющие случайные числа $(1 - r_i)$ можно получить, просто взяв дополняющее начальное значение z_0 , т. е. если в качестве начального значения выбрать число

$$z_0^* = m - z_0, \quad (233)$$

то в (231) результирующие случайные числа r^* будут дополняющими, т. е. $r_i^* = 1 - r_i$. Это доказано в приложении III.10⁴³. Тогда не надо дополнительного машинного времени (см. также [Andreasson, 1972b, p. 12 — 13]).

(в) Если наша переменная *дискретна*, то можно применить процедуру⁴⁴, представленную на рис. 13, а, б. Здесь мы меняем порядок значений переменной на абсциссе. Большое значение случайного числа r генерирует большое значение y_2 на рис. 13, а и малое значение y_1 на рис. 13, б. Поэтому для генерирования дополняющих значений входной переменной нам нужна *дополнительная таблица* в машинной памяти, где входная переменная записывается в *обратном* порядке. Дополнительное время, требуемое на обращение программы к таблице, незначительно. Это приводит к другой процедуре для генерирования дополняющих величин дискретных входных переменных⁴⁵.

(г) Точер [Tocher, 1963, p. 16] упоминает возможность аппроксимации обратной кумулятивной функции распределения F^{-1} в $y = F^{-1}(r)$ так называемой функцией Рассела, т. е. $y = F^{-1}(r)$ аппроксимируется выражением

$$y = A + Br + Cr^2 + ar^2 \log(1 - r) + b(1 - r)^2 \log r. \quad (234)$$

Фармер [Farmer, 1966, p. 11] утверждает, что дополняющие значения y получаются, если в выражении (234) заменить коэффициенты по следующим формулам:

$$A^* = A + B + C, B^* = -(B + 2C), C^* = C, a^* = b, b^* = a. \quad (235)$$

Однако, поскольку аппроксимация Рассела мало применяется в моделировании, процедура Фармера также не выглядит очень полезной.

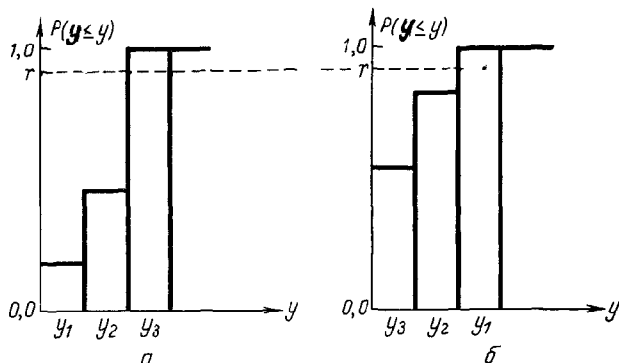


Рис. 13. Генерирование дополняющей случайной величины y из случайного числа r

(д) Андреассон [Andreasson, 1970, p. 6—7; 1971 а, p. 20] предлагает генерировать дополняющие значения y_a для входной переменной y (при симметричном непрерывном распределении и среднем η) с помощью соотношения

$$y_a = 2\eta - y. \quad (236)$$

Легко показать, что при этом методе генерируются те же самые значения, что и при методе обращения, использующем дополняющие случайные числа в соответствии с (237):

$$y_a = F^{-1}(1 - r). \quad (237)$$

Если старые значения y сохранены, то в общем процедура Андреассона занимает меньше времени, так как требует только вычитаний, а не вычисления значений функции F^{-1} . Его техника может применяться и тогда, когда старые значения y генерируются не с помощью обращения, т. е. y не есть $F^{-1}(r)$ (например, $y = \sum_{i=1}^2 r_i$, когда требуются величины с нормальным распределением).

Мы указывали ранее, что, по нашему мнению, лучшие результаты достигаются, когда каждая из входных переменных генерируется собственной последовательностью случайных чисел. Если, например, мы имеем мультипликативный генератор случайных чисел, определенный в (231) и (232), то каждый тип входной переменной должен генерироваться своим множителем в (231). В I.5 мы видели, что выбор множителя надо производить весьма тщательно, и поэтому можно

взять в качестве множителя только одно определенное значение. Есть еще другая возможность — разные *начальные* значения z_0 в (231) для разных типов входных переменных. Тогда мы избегаем пересечений в разных последовательностях случайных чисел. Пересечения можно избежать, если при применении генератора случайных чисел мы записываем начальное и конечные значения случайных последовательностей. В [Mize, 1973, p. 11 — 13, 17] предлагается другой способ порождения многомерных случайных последовательностей. Однако подход автора указанной работы, по нашему мнению, сомнителен. В [Mihgam, 1972, p. 246 — 249] также обсуждаются многомерные последовательности, но в другом контексте, не для понижения дисперсии, а в целях верификации.

Другая ситуация возникает в одноканальной системе массового обслуживания, когда после выборки каждого значения времени между поступлениями происходит выборка времени обслуживания, т. е. r_{2i-1} служит для генерирования времени между поступлениями, а r_{2i} — для генерирования времени обслуживания ($i = 1, 2, \dots$). В таком случае нам не нужны два генератора случайных чисел, можно работать с одним для двух типов входных переменных. В такой системе отрицательную корреляцию между двумя опытами можно вызвать следующим образом. В 1-м опыте время поступления генерируется на основе случайных чисел \vec{R}_1 , а время обслуживания — \vec{R}_2 ; во 2-м опыте \vec{R}_1 и \vec{R}_2 *меняются местами*, т. е. время поступления генерируется из \vec{R}_2 , а время обслуживания — на \vec{R}_1 . Этот метод был впервые предложен⁴⁶ Маршаллом и применен Пэйджем [Page, 1965], а позднее — Радемой [Radema, 1969]. Его нельзя применить в сложных системах, и поэтому мы не будем обсуждать его в дальнейшем.

Теперь мы назовем ряд *применений* метода дополняющих величин в моделировании. Большинство из них касается моделирования уже хорошо известной *одноканальной* системы массового обслуживания. Дополняющие величины в такой системе рассматриваются в [Lombaers, 1968, p. 252 — 253], [Mou, 1965, p. 72 — 74; 1966, p. 11 — 14, 27 — 35], [Page, 1965], [Radema, 1969], а также в [Gaver, 1969]. В первых трех работах получено следующее среднее понижение дисперсии (в круглых скобках коэффициент использования системы): 71% (80%), 41% (от 60 до 100%) и 46% (от 75 до 90%) соответственно. Радема исследовал влияние *коэффициента использования* системы на понижение дисперсии. Было обнаружено сильное влияние на понижение дисперсии: если коэффициент использования низок, то время ожидания будет равно нулю даже для малых промежутков между поступлениями и высоких значений времени обслуживания, т. е. монотонно возрастающая функция в (223) будет заменена отношением типа

$$g_3(r_1) = g_3(r_2) = 0, \text{ если } r_1 > r_2. \quad (238)$$

Таблица в [Radema, 1969, p. 13] показывает, что понижение дисперсии меняется от 5% для коэффициента использования, равного 20%, до 42% для коэффициента использования, равного 100%. Андреас-

сон [Andreasson, 1972b, p. 15 — 59] провел солидное исследование эффекта дополняющих величин (включая варианты, упомянутые в примечании 41) для различных коэффициентов использования, распределений входа, значений отклика. Он также вывел аналитические результаты для экстремальных значений коэффициента использования (для 0 или 1) в одноканальной системе массового обслуживания. Он получил, что в общем рост коэффициента использования приводит к уменьшению дисперсии; однако для некоторых переменных отклика уменьшение дисперсии понижалось при экстремальном значении коэффициента использования. Андреассон [Andreasson, 1971b, п. 5] изучал простую телефонную систему обслуживания, где оценивалась некоторая вероятность p . Он получил аналитическое выражение для понижения дисперсии, которое происходит при условии минимизации (отрицательной) корреляции между выходами двух парных опытов. Понижение дисперсии уменьшается, если p приближается к 1 или 0. На практике экстремальные значения корреляции между двумя выходами не были реализованы, тем не менее оценка понижения дисперсии уменьшалась при стремлении p к 0 или 1 (см. диаграммы в [Andreasson, 1971b, p. 21 — 22]). Радема также изучал влияние, которое оказывают длина опыта и начальные условия на понижение дисперсии. Применение метода дополняющих величин в системе массового обслуживания с четырьмя последовательными станциями обслуживания, где оценивался эффект изменения параметров распределения обслуживания этих станций, обсуждалось в [Kleijnen, 1969, p. 291]; оценка понижения дисперсии оказалась 41%. В [Shedler, Yang, 1971, p. 122—123] рассмотрено моделирование системы массового обслуживания на машине с мультипрограммированием. Авторы этой работы обнаружили, что если для одного отклика имелось значительное понижение дисперсии, то для другого дисперсия могла и повыситься. В [Gürtler, 1969, p. 83] получено понижение дисперсии от 5 до 10% в системе массового обслуживания с параллельными станциями. Андреассон [Andreasson, 1970, p. 8, 17] применил метод дополняющих величин для двух различных систем массового обслуживания с параллельным и последовательным порядком обслуживания. Понижение дисперсии было порядка 40%; применение дополняющих величин потребовало 11% дополнительного машинного времени. Мой применил дополняющие величины к своей сложной системе товарообмена в порту и получил понижение дисперсии⁴⁷ от 18 до 27%. В [Burt et al., 1970] дополняющие величины применяются к различным простым вероятностным сетям, и по грубой оценке авторы понизили дисперсию вдвое (см. также [Burt and Garman, 1971 a, p. 251 — 254]). В [Maune, 1966] обсуждается метод дополняющих величин для решения различных задач (оценка градиента функции стоимости в нелинейных системах управления с дискретным временем). Автор указанной работы нашел, что оценка дисперсии понижалась в 50 — 500 раз. В [Garman, 1971, p. 19 — 20] освещен метод дополняющих величин для искусственной системы (т. е. алгоритм не воспроизводит реальную систему), обеспечивающий снижение дисперсии на 22% при условии, что последовательности случайных чисел синхронизированы. Многие из упомянутых примене-

ний собраны и обобщены в [Andreasson, 1972b, p. 60 — 64]. Заметим, что ни в одном из этих применений нет 2^K -факторного подхода Точера.

Статистический анализ имитационных опытов не осложняется методом дополняющих величин. Поэтому можно взять среднее каждой дополняющей пары, т. е.

$$z_j = (x_{2j} + x_{2j-1})/2 \quad (j = 1, 2, \dots, n/2) \quad (n \text{ четное}), \quad (239)$$

где x_{2j} есть дополняющий опыт-партнер для x_{2j-1} ; все x_{2j-1} берут новые последовательности случайных чисел, т. е. x_{2j-1} независимы и, следовательно, все z_j *независимы*. Поэтому легко оценить дисперсию среднего отклика. Для

$$\bar{x} = \frac{\sum_{h=1}^n x_h}{n} = \frac{\sum_{j=1}^{n/2} z_j}{n/2} = \bar{z} \quad (240)$$

имеем

$$\hat{\text{var}}(\bar{x}) = \hat{\text{var}}(\bar{z}) = \sum_{j=1}^{n/2} (z_j - \bar{z})^2 / (n/2 - 1). \quad (241)$$

Доверительные интервалы аналогично (12) можно найти с помощью (240), (241) и t -статистики с $(n/2 - 1)$ степенями свободы. Следовательно, мощность t -критерия уменьшается или, что эквивалентно, длина доверительного интервала увеличивается из-за потери степеней свободы, но это компенсируется уменьшением дисперсии z по сравнению⁴⁸ с x .

Очевидно, что метод дополняющих величин привлекателен, так как он удобен в применении и требует мало дополнительного времени как на программирование, так и на счет. При этом достигается значительное понижение дисперсии и не усложняется статистический анализ результатов. Метод пригоден также для систем с остановом и без него, для одномерного и многомерного отклика. Динамическое поведение системы не искажается при работе с дополняющими величинами, следовательно, оно также может быть изучено с помощью этого метода.

III.7. ОБЩИЕ СЛУЧАЙНЫЕ ЧИСЛА

Все рассмотренные методы понижения дисперсии (МПД) предназначены для повышения надежности оценок отклика некоторой системы. Обычно мы моделируем несколько систем, чтобы сравнить их. В этом случае интересуются не *абсолютными* величинами откликов, а *различиями* между системами. Интуитивно ясно, что мы должны сравнивать системы, находящиеся в *равных условиях*. Поэтому желательно моделировать системы без останова с одинаковыми начальными условиями. Другое требование — по возможности пользоваться теми же самыми значениями входных переменных, т. е. если у систем 1 и 2 случайное время между поступлениями, то обе системы моделируются при одной и той же последовательности случайных чисел. Статистически применение общих случайных чисел приводит к *корреляции* между откликами. Рассмотрим дисперсию разности между x , оценкой отклика системы 1, и y , оценкой отклика системы 2. Эта дисперсия

выражается так:

$$\text{var} (x - y) = \text{var} (x) + \text{var} (y) - 2 \text{cov} (x, y). \quad (242)$$

Следовательно, дисперсия оценки уменьшается, если ковариация в (242) будет положительной. Такая положительная ковариация получается при использовании общих случайных чисел в предположении, что обе системы реагируют на случайные входы в одном направлении. Это иллюстрируется рис. 14. На рис. 14, а изображена ситуация, когда x и y отрицательно коррелируют при тех же самых случайных числах. Рис. 14, б показывает слабую корреляцию между x и y , когда

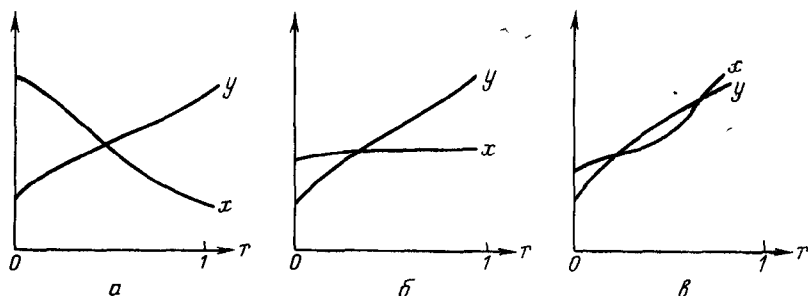


Рис. 14. Моделирование двух систем одинаковыми случайными числами

почти все случайные величины x равны между собой, в то время как значения y меняются с изменением r . На рис. 14, в, где x и y сильно изменяются в одинаковом направлении с изменением случайной величины r , имеет место положительная корреляция. Конечно, в моделировании могут встретиться и более сложные соотношения между откликом и случайным числом, однако разумно предположить, что обе системы реагируют в одном и том же направлении. Рассмотрим моделирование систем массового обслуживания. Обычно системы отличаются друг от друга тем, что они имеют различные числа обслуживающих станций, средние скорости обслуживания, средние скорости поступления, правила очередности и т. д. Тем не менее варианты этой системы реагируют в одном направлении, так как частые требования и длительное время обслуживания приводят к более длительному ожиданию независимо от числа обслуживающих станций и т. д.

Далее мы рассмотрим более детально, что означает «использование тех же случайных чисел». Очевидно, что метод предполагает, что каждая случайная входная переменная имеет свою собственную последовательность случайных чисел. При моделировании следующего варианта системы мы будем для генерирования входов брать те же начальные значения для генераторов случайных чисел, что и в предыдущем варианте системы. В простой системе массового обслуживания этого будет достаточно. Простая система массового обслуживания с эффектами различных случайных потоков для входной переменной изучена в [Mize, 1973]. Однако в [Kosten, 1968, р. 8] отмечено, что для других систем потребуются дополнительные затраты, чтобы сохранять синхронизированность последовательности случайных чисел. В [Kahn and

Marshall, 1953, p. 268] описана *синхронизация* как «использование идентичных случайных чисел при идентичных обстоятельствах в реализации двух процессов». Более формализованный анализ синхронизации представлен в [Garman, 1971, p. 15 — 23], где показано, что понижение дисперсии уменьшается при возрастании сложности системы, кроме

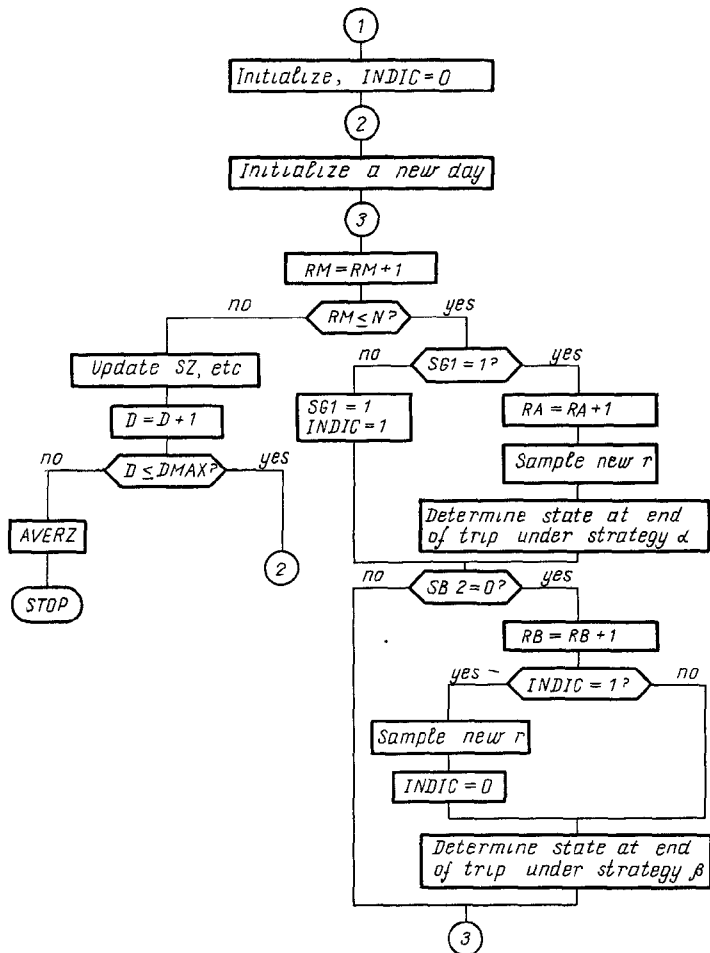


Рис. 15 Моделирование двух стратегий эксплуатации общими случайными числами

тех случаев, когда осуществляется тщательная синхронизация. Автор указанной работы предлагает вместо *нескольких* потоков случайных чисел брать один поток, в котором некоторые случайные числа пропускаются («фиктивные»), для того чтобы синхронизировать события (см. конец примера далее).

Проблема синхронизации проиллюстрирована на рис. 15. На этом рисунке мы даем блок-схему для задачи *эксплуатации* автобу-

са, сформулированной в приложении I.3 (блок-схему на рис. 15 можно сравнить с исходной диаграммой на рис.7 в приложении I.3). Состояние автобуса в конце маршрута задается выборочным случайным числом r и представлено в табл. 3. Положительную корреляцию между числом маршрутов в день для двух стратегий α и β , описанных в приложении I.3, можно сделать так. Выбираем состояние автобуса в конце маршрута j ($j = 1, 2, \dots, N$) в день i ($i = 1, 2, \dots, M$) для обеих стратегий α и β с использованием одних и тех же случайных чисел. Следовательно, если r «мало» (допустим, r меньше a и b в табл. 3), то при варианте α автобус заканчивает маршрут в состоянии A и должен ремонтироваться. Один маршрут отменяется, и это уменьшает общее число маршрутов в этот день. При стратегии β для этого же случайного числа автобус находится либо в состоянии A , либо B , в зависимости от того, в каком из состояний — соответственно «хорошем» или A — он был вначале. Оба события приводят к понижению общего числа маршрутов за день. Подводя итоги, можно сказать, что обе стратегии α и β уменьшают число маршрутов в день для малых случайных чисел; следовательно, наблюдается положительная корреляция.

Таблица 3

Выбор состояния автобуса в конце маршрута

Автобус начинает маршрут в состоянии	Автобус кончает маршрут в состоянии		
	хорошее	A	B
Хорошее A	$a \leq r < 1$...	$0 \leq r < a$ $b \leq r < 1$	$0 \leq r < b$

Чтобы *синхронизировать* случайные числа в стратегиях α и β , нам не надо моделировать эксплуатацию сначала M дней для стратегии α , потом M дней для стратегии β . Вместо этого моделируется сначала работа автобуса в стратегии α , потом в стратегии β ; далее идет моделирование следующего маршрута в стратегии α , затем в стратегии β и т. д. Более того, мы должны избегать *одного и того же* случайного числа в двух последовательных маршрутах *одной и той же* стратегии, так как в модели требуется, чтобы состояние в конце каждого маршрута определялось новым случайным числом. Поэтому рис. 15 содержит индикатор «INDIC». Индикатор принимает значение 1, если при стратегии α не выбирается новое случайное число (автобус будет ремонтироваться). При стратегии β должно генерироваться новое случайное число, которого не было в стратегии α . Альтернативный подход состоит в исключении этого индикатора. При этом новое случайное число генерируется для стратегии α , даже если автобус ремонтируется, но это случайное число более не применяется. Такой подход с «фиктивным» случайным числом предложен в [Garman, 1971; 1973].

Далее мы опишем конкретный пример *понижения дисперсии* с помощью МПД. Мы моделировали представленную систему эксплуатации автобуса в течение 1500 дней при значениях параметров $a =$

$= \frac{1}{4}$, $b = \frac{4}{5}$, $N = 25$. Моделирование проводилось с независимыми и общими случайными числами соответственно. Применение метода общих последовательностей дало понижение дисперсии на 38%. В [Kleijnen, 1969, p. 291] сообщается, что в системе массового обслуживания с четырьмя станциями при последовательном обслуживании требований понижение дисперсии было равным 95%. Андреассон [Andreasson, 1970, p. 19] получил понижение дисперсии, равное 90%, в его многоканальной системе массового обслуживания. Картер и Игнолл [Carter and Ignall, 1970, p. 31] получили оценку понижения дисперсии, равную 73%, в моделировании пожарного депо. Бреннер [Brenner, 1963, p. 293 — 294] оценил, как метод общих случайных чисел, названный им методом «коррелированной выборки», понижает «потери». Вспомним, что он определяет «потери» как разность между ценой при оптимальной комбинации факторов, найденной в имитационном эксперименте, и ценой точного оптимума, определенного с помощью аналитического решения. Для системы запасов Бреннер отметил значительное уменьшение потерь для общих случайных чисел. Игнолл [Ignall, 1972] работал с общими случайными числами при моделировании системы запасов с помощью методологии планирования экстремального эксперимента, но он не подсчитывал получаемые сбережения. Майз [Mize, 1973, p. 13 — 16] моделировал работу цеха общими случайными числами с разделением и без разделения потоков для входной переменной; улучшенная синхронизация дала более гладкие кривые стоимости. Мы не смогли найти других результатов применения метода общих случайных чисел для понижения дисперсии. Тем не менее применение общих случайных чисел кратко упоминается во многих работах, например [Emshoff and Sisson, 1971, p. 197], [Farmer, 1966, p. 11], [Fishman and Kiviat, 1967, p. 26], [Haitsma and Oosterhoff, 1964, p. 34], [Maxwell, 1965, p. 10], [Molenaar, 1968, p. 118], [Wurl, 1971, p. 94] и др. В [Mihram, 1972, p. 275] неявно отклоняется метод общих случайных чисел (только на с. 401 отмечено, что частичное применение общих случайных чисел создает «блоки» в терминологии планирования экспериментов).

Очевидно, что общие случайные числа применимы и при сравнении *более* чем двух систем. Далее, как мы уже упоминали в параграфе о контрольных величинах, специальная форма общих случайных чисел предполагает моделирование одновременно системы, представляющей интерес, и упрощенной системы с *известным* аналитическим решением так, чтобы можно было брать выражение (152). Для критической оценки этого метода мы отсылаем к III.4.4. Кроме того, с помощью общих случайных чисел в машинной программе одного опыта можно оценить несколько систем. Примеры можно найти в [Handscorn, 1968, p. 8 — 9] и [Lombaers, 1968, p. 254], см. также рис. 15. Как указывает Радема [Radema, 1969, p. 16], если при моделировании нам нужны экспоненциальные входные переменные, можно запоминать $\ln(r_i)$ (вместо r_i ($i = 1, 2, \dots$)), так как вычисление логарифмов требует много машинного времени. Наконец, как замечено в [Conway, 1963, p. 53], [Conway, Johnson and Maxwell, 1959, p. 106] и в [Hillier and Lieberman,

1968, р. 453], общие случайные числа приводят к тому, что нарушается независимость наблюдений, а это усложняет *статистический анализ*. Очевидно, сравнивая две системы, можно вычислить разности

$$d_i = x_i - y_i \quad (i = 1, 2, \dots, n), \quad (243)$$

где x_i и y_i есть отклики систем 1 и 2 соответственно и каждая система моделируется одно и то же число раз. Хотя значения x_i зависимы от y_i , новые наблюдения d_i уже взаимно независимы, и, следовательно, можно просто вычислить доверительный интервал для средней разности \bar{d} , где

$$\bar{d} = \sum_{i=1}^n d_i/n = \bar{x} - \bar{y}. \quad (244)$$

Однако если мы сравниваем более чем две системы, то, как будет объяснено в главе V, вычисление доверительных интервалов, как для двух систем, становится некорректным; вместо этого мы будем применять методы *множественных сравнений* и *множественного ранжирования*. Все эти методы предполагают независимость наблюдений. Мы вернемся к проблеме анализа коррелированных наблюдений в III.8.7.

Подводя итоги этого параграфа, можно сказать, что общие случайные числа дают значительное понижение дисперсии без особых затрат на машинное время. Применение этой техники очень просто, и потому она наиболее распространена среди всех процедур МПД. Фактически это единственная процедура, которая, как правило, применяется в практике моделирования.

III.8. СОВМЕСТНОЕ ПРИМЕНЕНИЕ ДОПОЛНЯЮЩИХ ВЕЛИЧИН И ОБЩИХ СЛУЧАЙНЫХ ЧИСЕЛ⁴⁹

III.8.1. Введение

Мой исследовал четыре метода понижения дисперсии (МПД), которые можно применить ко многим типам систем. К ним относятся методы стратифицированной, регрессионной, значимой и дополняющей выборки. Он сравнил эти четыре процедуры, применяя их к нескольким вариантам одноканальной системы массового обслуживания и к системе товарообмена в порту. Некоторые из его результатов описаны в предыдущих параграфах. Тех, кто интересуется деталями выводов, мы отсылаем к [Моу, 1965, р. 135 — 139; 1966, р. 35 — 36]. Если мы рассмотрим только результаты Мой для *сложной* системы товарообмена в порту, то мы заключим, что метод *дополняющих величин* — самый лучший из не зависящих от системы. Метод значимой выборки с зависимыми от системы переменными и метод контрольных величин дают большее понижение дисперсии, чем метод дополняющих величин. Однако эффективность методов понижения дисперсии надо скорректировать на дополнительное машинное время, которое необходимо для написания программ и счета. Мы бы рекомендовали метод

дополняющих величин как наиболее простой и эффективный среди этих четырех МПД, который легко применять на практике. Мой не исследовал эффективность метода общих случайных чисел. В III.7 мы видели, что этот МПД дает значительное понижение дисперсии. Кроме того, метод настолько прост, что его можно считать единственным действительно широко распространенным в практике. Другой не исследованный Моем метод — это рассмотренный нами метод селективной выборки. В III.3.2 мы показали, что метод дает смещение (хотя СКО может быть и малым) и не годится для моделирования систем со случайным числом случайных чисел в опыте. Кроме того, он требует дополнительного машинного времени для вычисления и реализации желаемых частот v_i^* с помощью выборки без возвращения (непрерывные случайные входные переменные надо сгруппировать перед проведением расчета). Подводя итоги, скажем, что имеются два метода, наиболее пригодных для обычных применений, а именно метод *дополняющих* величин и метод *общих* случайных чисел.

В некоторых работах упоминается о возможности одновременного применения более чем одного МПД к одной задаче. Точер [Tocher, 1963, p. 103] предлагает комбинацию методов контрольных величин и стратификации в его примере Монте-Карло. В [Burt et al., 1970, p. 452] методы дополняющих и контрольных величин объединены (это дало дополнительное понижение дисперсии на 50%), так же поступил и Андреассон [Andreasson, 1971b, p. 12], но он не получил дополнительного понижения дисперсии при добавлении дополняющих величин к контрольным. В [Gaver, 1969] получено понижение дисперсии при комбинации этих двух методов, зато в [Gaver, Shedler, 1971, p. 448] это понижение оказалось очень незначительным. В [Shedler, Yang, 1971] обсуждается объединение метода дополняющих величин со стратификацией, что дало большее понижение дисперсии (см. также [Fishman, 1972] и [Lewis, 1972, p. 12]). В [Handscorn, 1968, p. 4] утверждается, что «различные методы не являются ни в коем случае взаимно исключаящими». В [Tocher, 1963, p. 177] и особенно в [Emshoff and Sisson, 1971, p. 198] и [Fishman, 1967, p. 23 — 24] рекомендуется *комбинация методов дополняющих величин и общих случайных чисел*. Но, по нашему мнению, вовсе не очевидно, что объединение именно этих двух методов даст нам лучшие результаты.

III.8.2. Конфликт между дополняющими величинами и общими случайными числами

На первый взгляд кажется разумным применять оба метода — дополняющих величин и общих случайных чисел — для оценки разности откликов двух систем. Пусть

$$\bar{x} = \sum_{i=1}^M x_i / M, \quad (245)$$

$$\bar{y} = \sum_{j=1}^N y_j / N, \quad (246)$$

x_j и y_j — отклики в опыте i системы 1 и опыте j системы 2 соответственно. Тогда разность откликов двух систем оценивается выражением

$$\bar{d} = \bar{x} - \bar{y} \quad (247)$$

при

$$\text{var}(\bar{d}) = \text{var}(\bar{x}) + \text{var}(\bar{y}) - 2 \text{cov}(\bar{x}, \bar{y}). \quad (248)$$

Дополняющие системы понижают дисперсию оценки среднего откликов каждой системы, т. е. дисперсии $\text{var}(\bar{x})$ и $\text{var}(\bar{y})$ понижаются. Предполагается, что общая последовательность случайных чисел создает положительную ковариацию между \bar{x} и \bar{y} . Рассмотрим более подробно совместное применение обоих методов. В этом случае опыты с системами 1 и 2 генерируются в соответствии с табл. 4. Столбец 2 табл. 4

Таблица 4

Совместное применение дополняющих величин и общих случайных чисел

Опыты	Система 1		Система 2	
	случайные числа	отклик	случайные числа	отклик
1	2	3	4	5
1	\vec{R}_1	x_1	\vec{R}_1	y_1
2	$\vec{I} - \vec{R}_1$	x_2	$\vec{I} - \vec{R}_1$	y_2
3	\vec{R}_2	x_3	\vec{R}_2	y_3
4	$\vec{I} - \vec{R}_2$	x_4	$\vec{I} - \vec{R}_2$	y_4
⋮	⋮	⋮	⋮	⋮

показывает, что при моделировании системы 1 применяется метод дополняющих величин (\vec{I} означает вектор из единиц). Мы предполагаем, что метод создает желаемую отрицательную корреляцию между x_1 и x_2 , x_3 и x_4 и т. д. Аналогично столбец 4 означает, что к системе 2 применяются дополняющие величины и это создает отрицательную корреляцию между y_1 и y_2 , y_3 и y_4 и т. д. Далее, табл. 4 показывает, что те же случайные числа используются для обеих систем. Следовательно, можно предположить положительную корреляцию между y_1 и x_1 , y_2 и x_2 и т. д. Из табл. 4 также вытекает, что существует отрицательная корреляция между x_1 и y_2 , x_2 и y_1 , x_3 и y_4 и т. д. Эти отрицательные *перекрестные корреляции нежелательны* (в III.8 мы употребляем термин *перекрестная корреляция* для корреляции между откликом x системы 1 и откликом y системы 2; поэтому корреляции между опытами одной и той же системы — это не перекрестные корреляции). Отрицательные перекрестные корреляции нежелательны, так как, например, большое значение x_1 соответствует малому значе-

нию y_2 и это мешает сравнению двух систем. Эффект различных корреляций или ковариаций показан более строго в (249) (это уравнение выведено в приложении III.11).

$$\begin{aligned} \text{var}(\bar{a}) = & M^{-2} \sum_{i \neq g}^M \sum^M \text{cov}(\mathbf{x}_i, \mathbf{x}_g) + M^{-1} \sigma_1^2 + N^{-2} \sum_{i \neq h}^N \sum^N \text{cov}(\mathbf{y}_i, \mathbf{y}_h) + \\ & + N^{-1} \sigma_2^2 - 2M^{-1} N^{-1} \sum_{i=1}^M \sum_{j=1}^N \text{cov}(\mathbf{x}_i, \mathbf{y}_j), \end{aligned} \quad (249)$$

где

$$\sigma_1^2 = \text{var}(\mathbf{x}_i) = \text{var}(\mathbf{x}_g), \quad i, g = 1, \dots, M, \quad (250)$$

$$\sigma_2^2 = \text{var}(\mathbf{y}_j) = \text{var}(\mathbf{y}_h), \quad j, h = 1, \dots, N. \quad (251)$$

Сравнивая (249) с (248), мы видим, что первые два члена в (249) соответствуют первому члену в (248), следующие два члена в (249) — второму члену в (248) и последний член в (249) — последнему члену в (248). Каждый опыт системы имеет одну и ту же дисперсию, т. е. выполняются (250) и (251).

Величины σ_1^2 и σ_2^2 определяются моделируемой системой, но ковариации задаются выбором случайных чисел. На основании (249) можно заключить, что случайные числа нужно выбирать так, чтобы выполнялись следующие неравенства:

$$\text{cov}(\mathbf{x}_i, \mathbf{x}_g) < 0, \quad i \neq g; \quad i, g = 1, \dots, M, \quad (252)$$

$$\text{cov}(\mathbf{y}_j, \mathbf{y}_h) < 0, \quad j \neq h; \quad j, h = 1, \dots, N, \quad (253)$$

$$\text{cov}(\mathbf{x}_i, \mathbf{y}_j) > 0, \quad i = 1, \dots, M; \quad j = 1, \dots, N. \quad (254)$$

При обсуждении табл. 4 мы видели, что совместное применение обоих МПД приводит в некотором случае к отрицательным перекрестным корреляциям, т. е. $\text{cov}(\mathbf{x}_i, \mathbf{y}_j)$ отрицательны, и, следовательно, для некоторых значений i и j (254) нарушается. Далее мы рассмотрим три очевидные альтернативы.

III.8.3. Три альтернативы

В этом параграфе мы обсудим три альтернативных метода генерирования коррелированных опытов и выведем соответствующие дисперсии оценок разностей между откликами систем.

А. Только дополняющие величины. Для метода А табл. 4 заменяется табл. 5, где \vec{P}_j обозначает вектор случайных чисел, отличный от \vec{R}_i ($j = 1, 2, \dots, N/2$ и $i = 1, 2, \dots, M/2$). (Символ P выбран из мнемонических соображений для обозначения псевдослучайного числа.)

В. Только общие случайные числа (в первых N опытах). Оставляем открытым случай, когда число опытов для обеих систем неодинаково, т. е. $N \neq M$. Назовем систему с меньшим числом опытов системой 2, тогда $N \leq M$. Если $N < M$, тогда $\text{var}(\bar{a})$, очевидно, уменьшится, если мы применим дополняющие переменные в последних $(M - N)$ опытах системы 1, потому что в этом случае последние $(M - N)$ опы-

Таблица 5

Только дополняющие величины ($\vec{P} \neq \vec{R}$)

Опыты	Система 1		Система 2	
	случайные числа	отклик	случайные числа	отклик
1	2	3	4	5
1	\vec{R}_1	x_1	\vec{P}_1	y_1
2	$\vec{I} - \vec{R}_1$	x_2	$\vec{I} - \vec{P}_1$	y_2
3	\vec{R}_2	x_3	\vec{P}_2	y_3
4	$\vec{I} - \vec{R}_2$	x_4	$\vec{I} - \vec{P}_2$	y_4
⋮	⋮	⋮	⋮	⋮

тов не создают нежелательную отрицательную перекрестную корреляцию. Следовательно, можно использовать общие случайные числа в первых N опытах, а оставшиеся $(M - N)$ опытов системы 1 генерировать с дополняющими величинами. Отсюда табл. 6 представляет собой альтернативу В.

Таблица 6

Общие случайные числа в первых N опытах;
дополняющие переменные в последних $(M - N)$

опытах системы 1 ($\vec{R} \neq \vec{P}$, $z = (M - N)/2$)

Опыты	Система 1		Система 2	
	случайные числа	отклик	случайные числа	отклик
1	2	3	4	5
1	\vec{R}_1	x_1	\vec{R}_1	y_1
2	\vec{R}_2	x_2	\vec{R}_2	y_2
⋮	⋮	⋮	⋮	⋮
N	\vec{R}_N	x_N	\vec{R}_N	y_N
$N+1$	\vec{P}_1	x_{N+1}		
$N+2$	$\vec{I} - \vec{P}_1$	x_{N+2}		
⋮	⋮	⋮		
$M-1$	\vec{P}_z	x_{M-1}		
M	$\vec{I} - \vec{P}_z$	x_M		

С. Совместное применение дополняющих величин и общих случайных чисел. Этот случай был уже показан в табл. 4. В табл. 4, 5 и 6 понадобится внести исправления при нечетных значениях M и N , как будет видно из табл. 7.

Далее мы выведем дисперсию оцененной разности \bar{d} для каждой из трех альтернатив. Вывод дисперсии основан на формуле (249). Введем следующие символы:

$$c_1 = \text{cov}(\mathbf{x}_i, \mathbf{x}_g), \text{ если } \mathbf{x}_i \text{ и } \mathbf{x}_g \text{ генерируются с отрицательной корреляцией}; \quad (255)$$

$$c_2 = \text{cov}(\mathbf{y}_j, \mathbf{y}_h), \text{ если } \mathbf{y}_j \text{ и } \mathbf{y}_h \text{ генерируются с отрицательной корреляцией}; \quad (256)$$

$$c_3 = \text{cov}(\mathbf{x}_i, \mathbf{y}_j), \text{ если } \mathbf{x}_i \text{ и } \mathbf{y}_j \text{ генерируются с положительной (перекрестной) корреляцией}; \quad (257)$$

$$c_4 = \text{cov}(\mathbf{x}_i, \mathbf{y}_j), \text{ если } \mathbf{x}_i \text{ и } \mathbf{y}_j \text{ генерируются с отрицательной (перекрестной) корреляцией}. \quad (258)$$

Заметим, что если \mathbf{x}_1 и \mathbf{x}_2 генерируются с помощью дополняющих величин, как \mathbf{x}_3 и \mathbf{x}_4 и т. д., то их ковариации равны, т. е.

$$\text{cov}(\mathbf{x}_1, \mathbf{x}_2) = \text{cov}(\mathbf{x}_3, \mathbf{x}_4) = \dots = \text{cov}(\mathbf{x}_{M-1}, \mathbf{x}_M) = c_1, \quad (259)$$

где M — четное число. Легко видеть, что (259) выполняется, так как

$$\mathbf{x}_1 = g(\vec{R}_1) \text{ и } \mathbf{x}_2 = g(\vec{I} - \vec{R}_1), \quad (260)$$

$$\mathbf{x}_3 = g(\vec{R}_2) \text{ и } \mathbf{x}_4 = g(\vec{I} - \vec{R}_2), \quad (261)$$

где функция g обозначает машинную программу, при работе которой отклик становится функцией случайных чисел; \vec{R}_1 и \vec{R}_2 имеют одно и то же распределение. Тем же путем находим, что если \mathbf{y}_1 и \mathbf{y}_2 , \mathbf{y}_3 и \mathbf{y}_4 и т. д. генерируются на основании дополняющих величин, то выполняется

$$\text{cov}(\mathbf{y}_1, \mathbf{y}_2) = \text{cov}(\mathbf{y}_3, \mathbf{y}_4) = \dots = \text{cov}(\mathbf{y}_{N-1}, \mathbf{y}_N) = c_2, \quad (262)$$

где N — четное. Если опыт j ($j = 1, 2, \dots, N$) для обеих систем генерируется на основе общих случайных чисел, то выполняется

$$\text{cov}(\mathbf{x}_1, \mathbf{y}_1) = \text{cov}(\mathbf{x}_2, \mathbf{y}_2) = \dots = \text{cov}(\mathbf{x}_N, \mathbf{y}_N) = c_3. \quad (263)$$

Окончательно, если применяются оба метода — дополняющих величин и общих случайных чисел, — возникает нежелательная отрицательная перекрестная корреляция:

$$\text{cov}(\mathbf{x}_1, \mathbf{y}_2) = \text{cov}(\mathbf{x}_2, \mathbf{y}_1) = \dots = \text{cov}(\mathbf{x}_N, \mathbf{y}_{N-1}) = c_4. \quad (264)$$

Заметим, что вероятность генерирования коррелированных опытов зависит от значений M и N и эта зависимость двоякая.

1. Если число опытов в системе *нечетное*, то один опыт нельзя генерировать с помощью дополняющих величин.

2. Как мы видели при обсуждении альтернативы B , если число опытов в разных системах *различно* ($M > N$), то мы можем генери-

Дисперсия оценок разности между откликками систем I и 2
(к каждому значению нужно добавить общие члены $(M^{-1}\sigma_1^2 + N^{-1}\sigma_2^2)$)

Случай	А: только дополняющие величины	В: общие случайные числа в первых N опытах (дополняющие величины в последних $(M-N)$ опытах системы I)
I. $M = N$ — четное II. $M = N$ — нечетное III. M (четное) $> N$ (четное) IV. M (четное) $> N$ (нечетное) V. M (нечетное) $> N$ (четное) VI. M (нечетное) $> N$ (нечетное)	$M^{-1}c_1 + N^{-1}c_1$ $(M^{-1} - M^{-2})c_1^2 + (N^{-1} - N^{-2})c_2$ $M^{-1}c_1 + N^{-1}c_2$ $M^{-1}c_1 + (N^{-1} - N^{-2})c_2$ $(M^{-1} - M^{-2})c_1 + N^{-1}c_2$ $(M^{-1} - M^{-2})c_1 + (N^{-1} - N^{-2})c_2$	$-2M^{-1}c_3$ $-2M^{-1}c_3$ $-2M^{-1}c_3 + M^{-2}(M-N)c_1$ $-2M^{-1}c_3 + M^{-2}(M-N-1)c_1$ $-2M^{-1}c_3 + M^{-2}(M-N-1)c_1$ $-2M^{-1}c_3 + M^{-2}(M-N)c_1$
Случай	С: совместное применение дополняющих величин и общих случайных чисел во всех опытах	
I. $M = N$ — четное II. $M = N$ — нечетное III. M (четное) $> N$ (четное) IV. M (четное) $> N$ (нечетное) V. M (нечетное) $> N$ (четное) VI. M (нечетное) $> N$ (нечетное)	$M^{-1}c_1 + N^{-1}c_2 - 2M^{-1}(c_3 + c_4)$ $(M^{-1} - M^{-2})c_1 + (N^{-1} - N^{-2})c_2 - 2M^{-1}(c_3 + (1 - M^{-1})c_4)$ $M^{-1}c_1 + N^{-1}c_2 - 2M^{-1}(c_3 + c_4)$ $M^{-1}c_1 + (N^{-1} - N^{-2})c_2 - 2M^{-1}(c_3 + c_4)$ $(M^{-1} - M^{-2})c_1 + N^{-1}c_2 - 2M^{-1}(c_3 + c_4)$ $(M^{-1} - M^{-2})c_1 + (N^{-1} - N^{-2})c_2 - 2M^{-1}(c_3 + c_4)$	

ровать только первые N опытов системы 1 совместно с опытами системы 2 (беря те же самые случайные числа для опыта с системами 1 и 2).

Проиллюстрируем процедуру вывода $\text{var}(\bar{d})$ для ситуации в табл. 7, где применяются только дополняющие переменные (аналогия с табл. 5) и где $M = N$ — четное число. Первый член в выражении $\text{var}(\bar{d})$ в (249) сводится к

$$\begin{aligned} M^{-2} \{ \text{cov}(x_1, x_2) + \text{cov}(x_2, x_1) + \text{cov}(x_3, x_4) + \text{cov}(x_4, x_3) + \\ + \dots + \text{cov}(x_{M-1}, x_M) + \text{cov}(x_M, x_{M-1}) \} = \\ = M^{-2} \{ 2 \text{cov}(x_1, x_2) + 2 \text{cov}(x_3, x_4) + \dots + 2 \text{cov}(x_{M-1}, x_M) \} = \\ = M^{-2} \left\{ 2 \frac{M}{2} c_1 \right\} = M^{-1} c_1. \end{aligned} \quad (265)$$

Тем же способом находим, что третий член (249) равен $N^{-1} c_2$. Так как не существует корреляции между опытами в различных системах, последний член в (249) исчезает и, следовательно,

$$\text{var}(\bar{d}) = M^{-1} c_1 + M^{-1} \sigma_1^2 + N^{-1} c_2 + N^{-1} \sigma_2^2. \quad (266)$$

Вывод $\text{var}(\bar{d})$ для других ситуаций аналогичен. Поскольку (249) всегда содержит два члена ($M^{-1} \sigma_1^2 + N^{-1} \sigma_2^2$), они не показаны в табл. 7⁵⁰.

III.8.4. Сравнение альтернатив

Из табл. 7 следует, что для того, чтобы определить, какая из трех альтернатив A, B, C дает наименьшую дисперсию, мы должны знать *относительные величины ковариации* $c_1 \div c_4$. Например, если $(c_3 + c_4) > 0$, или, что эквивалентно, $c_3 > |c_4|$, то метод C лучше, чем метод A ; метод A лучше, чем метод B , если $c_3 < |c_1 + c_2|/2$ ($M = N$ — четное); метод B лучше, чем C , если $|c_4| > |c_1 + c_2|/2$ (и $M = N$ — четное) и т. д. Мы хотим знать, выполняются ли такие соотношения между ковариациями в *общем* случае в моделируемых системах. Рассмотрим сначала очень *простые* системы и попытаемся проверить *аналитически*, можно ли предположить, что такие соотношения выполняются для всех систем. Пусть в рассматриваемых системах отклик x в опыте зависит от одного случайного числа r и является монотонной возрастающей функцией от r . Например,

$$x = r^2. \quad (267)$$

Следовательно, в дополняющем опыте

$$x_a = (1 - r)^2 \quad (268)$$

и

$$\begin{aligned} c_1 = \text{cov}(x, x_a) &= E(x x_a) - E(x) E(x_a) = \\ &= E[r^2 (1 - r)^2] - E(r^2) E[(1 - r)^2] = E(r^2 - 2r^3 + \\ &+ r^4) - E(r^2) E(1 - 2r + r^2) = \int_0^1 (r^2 - 2r^3 + r^4) dr - \\ &- \int_0^1 r^2 dr \int_0^1 (1 - 2r + r^2) dr = 1/30 - (1/3)(1/3) = -7/90. \end{aligned} \quad (269)$$

Другие простые системы определены в столбцах 1 и 2 табл. 8. Результирующие ковариации от c_1 до c_4 вычисляются аналогично (269) и даются в столбцах 3 — 6 табл. 8. В столбцах 7 — 9 проводятся некоторые сравнения между ковариациями, что было описано выше. Сравнения в табл. 8 доказывают, что не существует постоянных соотношений между ковариациями для всех систем, т. е. ни один из методов — А, В или С — нельзя назвать наилучшим при всех обстоятельствах.

Таблица 8

Относительные величины ковариации в некоторых простых системах

Системы		Ковариации				Сравнения		
$x =$ $x(r)$	$y =$ $y(r)$	c_1	c_2	c_3	c_4	$ c_1 + c_2 $ $2c_3$	$ c_1 + c_2 $ $2 c_4 $	$\frac{c_3}{ c_4 }$
1	2	3	4	5	6	7	8	9
r^2	$r^2/2$	-7/90	-7/360	8/810	-7/180	>1	>1	>1
$2r+5$	r^2+r	-1/3	-59/180	1/3	-1/3	<1	<1	.
$2r^2$	\sqrt{r}	.	.	8/63	-132/905	.	.	<1

Кроме аналитических результатов для простых систем, представленных в табл. 8, мы получили экспериментальные результаты для некоторых *более сложных систем* (см. табл. 9).

Таблица 9

Оценки дисперсий разностей между откликами двух систем

Система	Метод			
	А (дополняющих величин)	В (общих случайных чисел)	С (совместный)	Д (независимый)
1. Один канал:				
а) экспоненциальная	0,057	0,11	0,053	0,12
б) экспонента/постоянная	0,090	0,18	0,081	0,33
2. Четыре канала	0,022	0,0019	0,0032	0,037

1. Одноканальная система массового обслуживания моделировалась в двух вариантах:

а) система 1 имеет экспоненциально распределенное время между поступлениями требований и время обслуживания с параметрами 1,5 и 2 соответственно; система 2 также имеет экспоненциальное время, но с параметрами 0,5 и 2,5 соответственно;

б) система 1 имеет экспоненциальное время поступления и время обслуживания с параметрами 1,5 и 2; система 2 имеет экспоненциальное время поступления с параметром 1,5 и постоянное время обслуживания, равное 1/2.

2. Система массового обслуживания с *четырьмя последовательными* станциями моделировалась в двух вариантах; варианты разли-

чаются параметрами четырех экспоненциально распределенных времен. Система описана⁵¹ подробно в [Naylor et al., 1967b, p. 703 — 705].

Все системы с одним каналом обслуживания были промоделированы 50 раз, т. е. $M = N = 50$; системы с четырьмя станциями обслуживания — 20 раз, $M = N = 20$. Эти эксперименты были осуществлены по методам A, B, C , кроме того, по методу D , который является грубым методом, где все опыты независимы. Оценки дисперсий разностей между откликами двух систем с применением методов A, B, C, D даются в таблице 9⁵². Из таблицы видно, что наименьшая дисперсия не обязательно достигается, когда мы применяем оба метода — дополняющих величин и общих случайных чисел (метод C). В действительности для системы с одной станцией обслуживания получаем порядок C, A, B, D , а для системы с четырьмя станциями обслуживания порядок будет B, C, A, D (очевидно, что мы пренебрегаем возможными ошибками выборки)⁵³.

III.8.5. Оптимальная альтернатива и распределение машинного времени

Как аналитические результаты для совсем простых систем, так и экспериментальные результаты для более сложных систем показывают, что не существует наилучшей альтернативы для *всех* систем. Допустим, что мы генерируем некоторые *предварительные* опыты для двух систем, чтобы *оценить* σ_1^2 и σ_2^2 и ковариации $c_1 \div c_4$, а также выбрать метод, который мы применим в оставшихся опытах (в III.8.6 мы рассмотрим оценочную процедуру для дисперсий и ковариаций в деталях). Если мы хотим воспользоваться дисперсиями и ковариациями из табл. 7, то мы должны знать, какой из 6 случаев, перечисленных в таблице, будет обоснованным. Нам нужно знать, являются ли M и N четными или нечетными числами и равны ли M и N (или $M < N$ так, как в табл. 7; мы назвали эту систему с меньшим числом опытов системой 2, но если мы не знаем заранее, что система 2 со своими σ_2^2 и c_2 будет «проиграна» меньшее число раз, чем система 1, то мы должны также не отбрасывать и возможность $M < N$). Поэтому мы фиксируем M и N после предварительного этапа так, чтобы *затрачиваемое машинное время*, т. е. (270), удовлетворяло дополнительным ограничениям (271) и (272).

$$t_1 M + t_2 N = T, \quad (270)$$

$$M \geq N_p, \quad (271)$$

$$N \geq N_p, \quad (272)$$

где N_p — число предварительных опытов для систем 1 и 2; t_1 и t_2 — машинное время одного опыта с системами 1 и 2 соответственно; T — общее машинное время, данное для моделирования двух систем; $(M - N_p)$ и $(N - N_p)$ — числа опытов, генерируемых после предварительной фазы. Единственное решение возможно, если мы введем ограничение, состоящее в том, что мы будем моделировать

обе системы *одинаковое* число раз (позже мы вернемся к этому ограничению). Итак,

$$M = N. \quad (273)$$

На основе данных M и N , вычисляемых в (270) — (273), данных дисперсий и ковариаций мы можем с помощью табл. 7 найти метод, который дает наименьшую дисперсию. Заметим, что выражение (270), ограничивающее машинное время, очень важно, так как для неограниченного времени M и N можно взять неограниченно большими. Неограниченно большие значения M и N могут дать нулевую дисперсию, $\text{var}(\bar{d}) = 0$, для всех трех методов A , B и C , и, следовательно, выбор метода становится безразличным.

Конечно, нет необходимости брать M и N равными, как в (273). Вместо этого можно выбрать M и N так, чтобы *минимизировать* $\text{var}(\bar{d})$ и удовлетворить условиям (270), (271) и (272). Тот факт, что формула для $\text{var}(\bar{d})$ меняется в зависимости от M и N и методов A , B , C , усложняет оптимальный выбор M и N . Поэтому помимо 6 случаев для M и N , приведенных в табл. 7, рассмотрим еще 4 случая, когда $M < N$ (как мы уже отмечали, систему 2 с ее σ_2^2 и c_2 можно промоделировать большее число раз, чем систему 1). Для простоты мы решили воспользоваться приближенной формулой для $\text{var}(\bar{d})$,

Таблица 10

Аппроксимация для дисперсии $\text{var}(\bar{d})$ при отбрасывании членов с M^{-2} и N^{-2}

Случай	Метод	
$M \geq N$	A	$(\sigma_1^2 + c_1) M^{-1} + (\sigma_2^2 + c_2) N^{-1}$
	B	$(\sigma_1^2 - 2c_3) M^{-1} + \sigma_2^2 N^{-1}$
	C	$(\sigma_1^2 + c_1 - 2c_3 - 2c_4) M^{-1} + (\sigma_2^2 + c_2) N^{-1}$
$M < N$	A	$(\sigma_1^2 + c_1) M^{-1} + (\sigma_2^2 + c_2) N^{-1}$
	B	$(\sigma_1^2 M^{-1} + (\sigma_2^2 - 2c_3) N^{-1})$
	C	$(\sigma_1^2 + c_1) M^{-1} + (\sigma_2^2 + c_2 - 2c_3 - 2c_4) N^{-1}$

в которой отброшены члены с M^{-2} и N^{-2} из табл. 7. Результаты сведены в таблицу 10⁵⁴. Из табл. 10 видно, что $\text{var}(\bar{d})$ можно аппроксимировать выражением

$$\text{var}(\bar{d}) = a_1 M^{-1} + a_2 N^{-1} \quad (274)$$

с коэффициентами a_1 и a_2 , разными в методах A , B , C , и в случаях $M \geq N$ и $M < N$. Итак, мы хотим *минимизировать дисперсию* оценки разности между откликами двух систем, определенную в (274) при заданном ограничении на *объем машинного времени* (270) и заданных условиях (271) и (272) на число *предварительных опытов* N_p с системами 1 и 2.

Оптимальные значения M и N в (274) зависят от знаков коэффициентов a_1 и a_2 . В приложении III.2 мы докажем, что существуют следующие две возможности:

- 1) оба коэффициента a_1 и a_2 положительны;
- 2) один коэффициент положителен, а другой отрицателен.

В приложении III.13 мы выводим, что если оба коэффициента *положительны*, то $\text{var}(\bar{a})$ в (274) минимизируется при ограничении на машинное время (270) при условии, что мы выберем M и N равными (275) и (276) соответственно:

$$M^* = T / \{t_1 + (a_1^{-1} a_2 t_1 t_2)^{1/2}\}, \quad (275)$$

$$N^* = T / \{t_2 + (a_1 a_2^{-1} t_1 t_2)^{1/2}\} \quad (a_1, a_2 > 0). \quad (276)$$

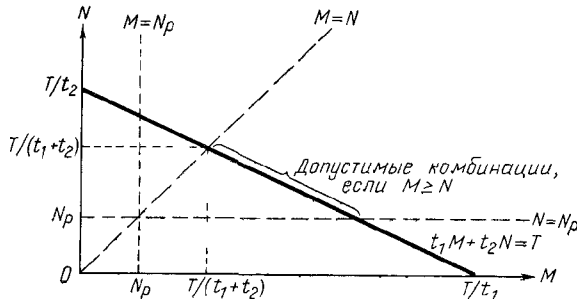


Рис 16 Допустимые комбинации M и N

На рис. 16 мы видим, что решения (275) и (276) могут нарушать условия (271) и (272), требующие, чтобы M и N в крайнем случае были равны N_p . Более того, если мы возьмем определенное значение для a_1 и a_2 в (274), это значит, что либо $M \leq N$, либо $M \geq N$, что видно из табл. 10. При таких дополнительных ограничениях оптимальные величины M и N надо брать из табл. 11, что доказывается в приложении III.13. Из табл. 11 следует, что если (276) дает $N^* < N_p$, то N_0 — оптимальная величина N — становится равной N_p ; если (276) даёт $N^* > T / (t_1 + t_2)$, то нарушается условие $M \geq N$ и мы берем равное число опытов, т. е. $M_0 = N_0 = T / (t_1 + t_2)$.

Если один из коэффициентов a_1 или a_2 *отрицателен*, то, как показано в приложении III.13, мы должны взять M и N из

$$M_0 = N_0 = T / (t_1 + t_2) \quad (a_1 \text{ или } a_2 < 0). \quad (277)$$

Есть ограничение, которое не оговаривалось ранее, а именно M и N должны быть *целыми*. Следовательно, M_0 и N_0 , определенные из табл. 11 или выражения (277), должны быть целыми. Можно взять целые числа в окрестности M_0 и N_0 , как видно из рис. 17. Для простоты можно ограничить свое внимание следующими тремя парами (квадратные скобки означают целую часть числа):

$$M = [M_0], N = [N_0], \quad (278)$$

$$M = [M_0] + 1, N = [N_0], \quad (279)$$

$$M = [\dot{M}_0], N = [N_0] + 1. \quad (280)$$

Оптимальные значения M и N , когда a_1 и a_2 положительны
(M^* и N^* определены в (275) и (276) соответственно)

Случай	Граничные условия	M_0	N_0
$M \geq N$	$N_p < N^* \leq T/(t_1 + t_2)$	M^*	N^*
	$N^* < N_p$	$(T - t_2 N_p)/t_1$	N_p
$M \leq N$	$N^* > T/(t_1 + t_2)$	$T/(t_1 + t_2)$	$T/(t_1 + t_2)$
	$N_p < M^* \leq T/(t_1 + t_2)$	M^*	N^*
	$M^* < N_p$	N_p	$(T - t_1 N_p)/2$
	$M^* > T/(t_1 + t_2)$	$T/(t_1 + t_2)$	$T/(t_1 + t_2)$

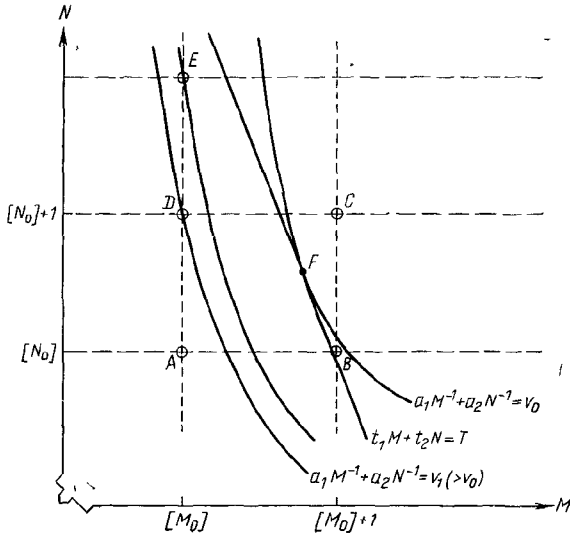


Рис 17. Оптимальная комбинация целых чисел M и N : F — оптимум (M_0, N_0) ; A, B, C, D, E — целые комбинации; C, B — неприемлемые (см. ограничения на машинное время); E лучше, чем D ; E не удовлетворяет (278) — (280)

Мы должны проверить, удовлетворяют ли целые числа условиям машинного времени, условиям предварительных опытов (270) — (272) и являются ли $M \geq N$ или $M \leq N$ (на рис. 17 показано лишь ограничение, накладываемое на машинное время). Если существует более одной пары M и N , удовлетворяющих всем этим условиям, то мы выбираем ту из них, которая дает минимальную дисперсию. С помощью (278) — (280) мы находим пару целых чисел M и N , которая дает наименьшую дисперсию. Однако эта минимальная дисперсия определена

для конкретной пары значений коэффициентов a_1 и a_2 . Из табл. 10 следует, что различные пары коэффициентов годятся для каждого метода и для каждого случая ($M \geq N$ или $M \leq N$), поэтому в результате получаем шесть пар (a_1, a_2) . В действительности метод A дает одни и те же коэффициенты для случаев $M \geq N$ и $M \leq N$, поэтому получаем пять вместо шести пар значений a_1 и a_2 . Каждая из пяти пар (a_1, a_2) дает соответствующую ей пару оптимальных M и N . Подстановка оптимальных M и N в выражение для $\text{var}(\bar{d})$ в (274) дает минимальную дисперсию. Таким образом найдем пять минимальных дисперсий. Наконец, выберем *минимум* среди этих пяти минимальных дисперсий и определим соответствующий ему метод (A , B или C) и число опытов (M и N).

Заметим, что в нашей процедуре выбора мы принимаем во внимание *ковариации*, которые создаются различными методами. Для этих ковариаций определяем величины коэффициентов a_1 и a_2 в (274). Более того, можно учесть возможные различия между *машинным временем на опыт* для разных методов, так как величины t_1 и t_2 в (270) зависят от метода (A , B или C). Заметим, что дисперсия минимизируется не только выбором подходящего *метода, понижающего дисперсию*, но также и выбором *оптимальной* комбинации числа опытов на систему. При оптимальном числе опытов для каждой системы можно получить более точную оценку $\text{var}(\bar{d})$, применяя точную формулу для $\text{var}(\bar{d})$ из табл. 7.

III.8.6. Оценивание коэффициентов в методе оптимизации

Здесь мы обсудим, как можно использовать предварительный эксперимент для получения *оценок коэффициентов* a_1 и a_2 в (274) и t_1 и t_2 в (270). Число N_p предварительных опытов генерируется для систем 1 и 2 методом C при *совместном* применении дополняющих величин и общих случайных чисел. Для метода C вычисляются все четыре значения ковариаций, $c_1 \div c_4$. Мы можем оценить c_1 , ковариацию между дополняющими опытами системы 1, с помощью выражения

$$\hat{c}_1 = \sum_{i=1}^{n_p} (x_{2i-1} - \bar{x})(x_{2i} - \bar{x}_a) / (n_p - 1), \quad (281)$$

где

$$\bar{x} = \sum_{i=1}^{n_p} x_{2i-1} / n_p, \quad (282)$$

$$\bar{x}_a = \sum_{i=1}^{n_p} x_{2i} / n_p, \quad (283)$$

$$n_p = N_p / 2 \quad (N_p \text{ — четное}). \quad (284)$$

Оценка c_2 полностью аналогична (281). Оценивая c_3 , положительную перекрестную корреляцию между x_i и y_i ($i = 1, \dots, N_p$), мы должны принимать во внимание хорошо известные формулы оценок, подобные (281), основанные на предположении, что каждая пара на

блюдений *независимы* от любой другой пары. Однако пара (x_1, y_1) не независима от пары (x_2, y_2) , поскольку, как видно из табл. 4, обе пары используют случайные числа \vec{R}_1 . Мы можем решить это затруднение, рандомизировав предварительные опыты на две группы N_p , как видно из табл. 12. Тогда положительная перекрестная ковариация может быть оценена на основании:

$$\hat{c}_3(1) = \sum_{i=1}^{n_p} (x_{2i-1} - \bar{x})(y_{2i-1} - \bar{y}) / (n_p - 1), \quad (285)$$

$$\hat{c}_3(2) = \sum_{i=1}^{n_p} (x_{2i} - \bar{x}_a)(y_{2i} - \bar{y}_a) / (n_p - 1), \quad (286)$$

$$\hat{c}_3 = \{\hat{c}_3(1) + \hat{c}_3(2)\} / 2. \quad (287)$$

Таблица 12

Группировка N_p предварительного эксперимента для оценки ковариации c_3

Группа 1			Группа 2		
случайные числа	отклик 1	система 2	случайные числа	отклик 1	система 2
\vec{R}_1	x_1	y_1	$\vec{I} - \vec{R}_1$	x_2	y_2
\vec{R}_2	x_3	y_3	$\vec{I} - \vec{R}_2$	x_4	y_4
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
\vec{R}_{n_p}	x_{N_p-1}	y_{N_p-1}	$\vec{I} - \vec{R}_{n_p}$	x_{N_p}	y_{N_p}

Тем же способом можно оценить c_4 , группируя наблюдения, как показано в табл. 13⁵⁵. Хорошо известная формула для оценки дисперсии также основывается на независимых наблюдениях.

Таблица 13

Группировка N_p предварительного эксперимента для оценки c_4

Группа 1				Группа 2			
система 1		система 2		система 1		система 2	
случайные числа	отклик	случайные числа	отклик	случайные числа	отклик	случайные числа	отклик
\vec{R}_1	x_1	$\vec{I} - \vec{R}_1$	y_2	$\vec{I} - \vec{R}_1$	x_2	\vec{R}_1	y_1
\vec{R}_2	x_3	$\vec{I} - \vec{R}_2$	y_4	$\vec{I} - \vec{R}_2$	x_4	\vec{R}_2	y_3
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
\vec{R}_{n_p}	x_{N_p-1}	$\vec{I} - \vec{R}_{n_p}$	y_{N_p}	$\vec{I} - \vec{R}_{n_p}$	x_{N_p}	\vec{R}_{n_p}	y_{N_p-1}

Следовательно, можно воспользоваться группировками из табл. 12 или 13 для оценки дисперсии системы 1 на основе:

$$\hat{\sigma}_1^2(1) = \sum_{i=1}^{n_p} (x_{2i-1} - \bar{x})^2 / (n_p - 1), \quad (288)$$

$$\hat{\sigma}_1^2(2) = \sum_{i=1}^{n_p} (x_{2i} - \bar{x}_a)^2 / (n_p - 1), \quad (289)$$

$$\hat{\sigma}_1^2 = \{\hat{\sigma}_1^2(1) + \hat{\sigma}_1^2(2)\} / 2. \quad (290)$$

Ту же самую процедуру можно применить для оценки σ_2^2 . Оценки для дисперсий и ковариаций можно подставить в табл. 10, чтобы получить оценки коэффициентов a_1 и a_2 в выражении для $\text{var}(\bar{d})$ в (274). Заметим, что оценки дисперсий и ковариаций получаются при применении только метода *C*. Однако эти оценки можно употребить в табл. 10 для получения оценок a_1 и a_2 для *A*, *B*, и *C* (и для $M \gg N$, $M \leq N$). Если нам нужно T_p единиц времени для одного опыта с системой 1, то, применяя метод *C* N_p раз, можно оценить t_1 в ограничении (270) с помощью выражения

$$\hat{t}_1 = T_p / N_p. \quad (291)$$

Тем же способом можно оценить время опыта по методу C^{56} для системы 2. Как указывалось в III.6 и III.7, дополнительное время для применения методов дополняющих величин и общих случайных чисел обычно незначительно. Следовательно, можно пользоваться теми же самыми значениями \hat{t}_1 и \hat{t}_2 для всех трех методов *A*, *B*, *C*. Однако если мы генерируем дополняющие опыты, вычитая r из единицы, то мы можем оценить сэкономленное время одного опыта системы 1 без дополняющих величин, разделив время, необходимое для генерирования $x_1(\vec{R}_1)$, $x_3(\vec{R}_2)$, ..., $x_{N_p-1}(\vec{R}_{n_p})$, на $N_p/2$.

Мы отмечали, что оптимальные значения M , N и $\text{var}(\bar{d})$ — *нелинейные* функции дисперсий, ковариаций и времени на опыт. Следовательно, даже для несмещенных оценок этих дисперсий и т. д. соответствующие оптимальные оценки M , N и $\text{var}(\bar{d})$ все-таки будут смещенными⁵⁷. Этот тип смещения обсуждается в [Fishman, 1967, p. 19—22]. Мы не пытались определить смещение, считая его незначительным⁵⁸.

Итак, мы получили следующую процедуру для выбора одного из методов *A*, *B*, *C* и определения оптимальной комбинации опытов с системами 1 и 2.

1. Проводится N_p предварительных опытов для системы 1 и N_p таких же опытов для системы 2 с применением метода *C* (совместное использование дополняющих величин и общих случайных чисел согласно табл. 4).

2. Оцениваются дисперсии σ_1^2 и σ_2^2 , определенные в (250) и (251), ковариации $c_1 \div c_4$, определенные в (255) — (258), и время на опыт для систем 1 и 2, т. е. t_1 и t_2 в (270). Оценочные формулы — это (281) — (291).

3. Найденные дисперсии и ковариации подставляются в табл. 10, чтобы получить оценки коэффициентов a_1 и a_2 для $\text{var}(\bar{d})$ в (274). Существует 6 пар оценок коэффициентов a_1 и a_2 , так как мы имеем 3 метода (A, B, C) и два случая ($M \geq N$, $M \leq N$); 5 из шести пар различны.

4. Для каждой из 5 пар оцененных a_1 и a_2 и соответствующих t_1 и t_2 определяется оптимальная комбинация M и N (число опытов систем 1 и 2) с помощью табл. 11 и формул (277) — (280).

5. Для оптимальных величин M и N каждой из пяти пар определяется соответствующая дисперсия с помощью аппроксимации для $\text{var}(\bar{d})$ в (274) или точной формулы из табл. 7.

6. Определяется минимум из пяти дисперсий шага 5 и выбираются соответствующий метод и число опытов.

III.8.7. Комментарии к методу оптимизации

Вначале рассмотрим некоторые применения метода оптимизации. Метод применялся к нескольким простым системам массового обслуживания. Результаты приведены в табл. 14⁵⁹. Для применения к одноканальным системам массового обслуживания 1а и 1б из табл. 9 получается тот же порядок рангов, что и в табл. 9, т. е. в порядке возрастания $\text{var}(\bar{d})$ следуют C, A, B, D. Применение к четырем другим ситуациям, где сравнивались системы с одной или двумя станциями обслуживания, дало в результате следующее ранжирование: C, B, A, D.

Таблица 14

Результаты применений метода оптимизации

Сравниваемые системы	Минимальная $\text{var}(\bar{d}) (\times 10^5)$				Оптимум M и N для метода C**		Примечания относительно M и N
	A	B	C	D			
1. Система с одной станцией обслуживания:							
а) Экспоненциальные распределения*	162	76	52	257	100	100	$N^* > T/(t_1 + t_2)$
б) Экспоненциальные распределения*	83	99	77	104	139	50	$N^* < N_p$
в) Постоянное время обслуживания	9	3	2	14	150	150	a_1 — отрицательно
г) одно экспоненциальное время и одно постоянное время обслуживания	114	145	113	180	164	54	
2. Одна система с одной станцией обслуживания и одна система с двумя станциями обслуживания	33	26	24	51	128	68	
3. Две системы с двумя станциями обслуживания	61	45	39	79	95	94	a_1 — отрицательно

* Для 1а и 1б разные машинные программы.

** Для методов A, B, D найдены примерно те же значения.

Оптимальные значения M и N очень сильно различались между собой (кроме случая с отрицательным коэффициентом a_1 или при $N^* > T/(t_1 + t_2)$; тогда $M_0 = N_0$).

Заметим, что порядок при методе оптимизации может не совпадать с порядком, возникающим при применении каждого из четырех методов к двум системам равное число раз, скажем N_p (последняя процедура отражена в табл. 9). Это проиллюстрировано на рис. 18 где A_p и B_p означают дисперсию \bar{d} при применении методов A и B , каждого — N_p раз; A_0 и B_0 — минимальные дисперсии для методов A и B ,

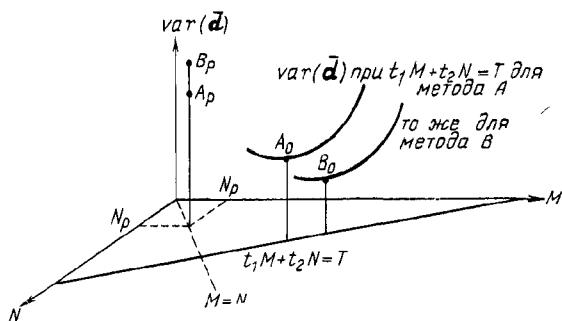


Рис 18 Дисперсия \bar{d} для равного числа опытов ($M = N = N_p$) или для оптимального числа опытов с учетом ограничения на машинное время

удовлетворяющих ограничению на машинное время. На рис. 18 мы имеем $A_p < B_p$, но $A_0 > B_0$. (Различный порядок можно объяснить тем фактом, что оптимизация дает только оценку дисперсии и основана на аппроксимации для $\text{var}(\bar{d})$ в (274).)

Мы считаем, что наш метод оптимизации имеет ряд *недостатков*.

1. Требуется *время*, чтобы оценить коэффициенты и произвести необходимые вычисления с ними. Тем не менее в сложных исследованиях по моделированию это дополнительное время несущественно и оптимизация может быть полезной.

2. Метод основан на *оценках* дисперсий и ковариаций, вычисленных из N_p предварительных опытов с каждой из систем. Следовательно, если N_p возрастает, то эти оценки становятся более надежными и более надежным оказывается выбор методов A , B или C . К сожалению, с ростом N_p может случиться, что C будет не лучшим методом (этот метод применяется на предварительной фазе), а в лучшие выйдут A или B ; однако большинство опытов уже генерировано при применении худшего метода C . Чтобы решить эту дилемму, мы можем ограничить N_p , скажем 10% ожидаемых значений M_0 и N_0 . (В действительности более точная процедура для минимального $\text{var}(\bar{d})$ сравнивает не $\text{var}(\bar{d})$ при приложении метода A , B или C ко всем M_0 и N_0 опытам, она должна сравнивать $\text{var}(\bar{d})$ при применении метода A , B или C к $(M_0 - M_p)$ и $(N_0 - N_p)$ опытам и метода C — к N_p опытам. Если N_p мало по

сравнению с M_0 и N_0 , как и должно быть, то для простоты можно воспользоваться процедурой, описанной в конце III.8.6.)

3. Если мы решим после предварительной фазы перейти от метода C к методу A или B , то получим *неоднородный* выход. Переход от C к A означает, что перекрестные ковариации c_3 и c_4 обращаются в нуль; переход от C к B означает, что мы не ставим дополняющих опытов и c_1 и c_2 обращаются в нуль. Это осложняет *статистический анализ*. Потому что если бы мы применяли метод C ко всем M и N опытам, то можно было бы оценить дисперсию \bar{d} , пользуясь формулой для $\text{var}(\bar{d})$ из табл. 7 и оценками по формулам (281) — (290). Если же мы переходим от метода C к методу A после предварительной фазы, то можно воспользоваться выражением

$$\text{var}(\bar{d}) = \text{var}(\bar{x}) + \text{var}(\bar{y}) - 2 \text{cov}(\bar{x}, \bar{y}). \quad (292)$$

Когда мы переходим от метода C к методу A , мы имеем $M/2$ дополняющих пар⁶⁰. Следовательно, $\text{var}(\bar{x})$ можно оценить с помощью среднего каждой дополняющей пары, как в (239) — (241). Тем же способом мы оцениваем $\text{var}(\bar{y})$ из $N/2$ дополняющих пар. Последний член в (292) можно оценить из выражения

$$\text{cov}(\bar{x}, \bar{y}) = \sum_{i=1}^M \sum_{j=1}^N \text{cov}(x_i, y_j) / (MN) = N_p (c_3 + c_4) / (MN). \quad (293)$$

Последнее равенство в (293) выполняется, так как x_i и y_j коррелируют в первых N_p предварительных опытах каждой системы, но они независимы после предварительной фазы. Если мы после этой фазы переходим от метода C к методу B , то последний член в (292) можно вычислить на основе

$$\text{cov}(\bar{x}, \bar{y}) = \{\min(N, M) c_3 + N_p c_4\} / (MN). \quad (294)$$

Когда переходим от метода C к методу B , дисперсию можно оценить делением M наблюдений \bar{x}_i на две части: в первой части содержатся все дополняющие пары, во второй — все независимые наблюдения x . Тогда \bar{x} есть взвешенная средняя средних этих двух частей:

$$\bar{x} = a\bar{x}_1 + (1 - a)\bar{x}_2. \quad (295)$$

Следовательно,

$$\text{var}(\bar{x}) = a^2 \text{var}(\bar{x}_1) + (1 - a)^2 \text{var}(\bar{x}_2), \quad (296)$$

где $\text{var}(\bar{x}_1)$ можно оценить аналогично (239) — (241), а $\text{var}(\bar{x}_2)$ оценивается традиционно. Как только $\text{var}(\bar{d})$ оценена, можно определить *доверительные пределы* для разности $E(\bar{d})$ с использованием нормального распределения, по крайней мере если есть много наблюдений. В таком случае мы имеем

$$\bar{d} = \bar{x} - \bar{y}, \quad (297)$$

где \bar{x} и \bar{y} асимптотически нормальны, что следует из центральной предельной теоремы⁶¹. Нормальность \bar{x} и \bar{y} может быть также основана на

асимптотической нормальности индивидуальных наблюдений x_i и y_j , если для индивидуальных наблюдений выполняется центральная предельная теорема для стационарных r -зависимых реализаций; эта теорема упоминалась в III.2.1 при обсуждении (12) и (13). Мы предполагаем⁶², что \bar{x} и \bar{y} распределены как двумерные асимптотически нормальные величины. Следовательно, линейное преобразование \bar{d} этой двумерной нормальной величины также будет асимптотически нормально распределенной величиной согласно [Fisz, 1967, p. 162]. Поэтому если мы имеем *много* наблюдений, то можно пользоваться $(1 - \alpha)$ доверительным интервалом

$$\bar{d} \pm z^{\alpha/2} s, \quad (298)$$

где $z^{\alpha/2}$ — соответствующая $\alpha/2$ точка стандартной нормальной переменной и s — оценка стандартного отклонения \bar{d} , вычисленная из выражения (292). Если же число наблюдений ограничено, то (298) нужно заменить интервалом, основанным на t -распределении:

$$\bar{d} \pm t_{n-1}^{\alpha/2} s_d / \sqrt{n}, \quad (299)$$

где

$$s_d = \left\{ \sum_{i=1}^n (d_i - \bar{d})^2 / (n-1) \right\}^{1/2}, \quad (300)$$

$$\bar{d} = \sum_{i=1}^n d_i / n \quad (301)$$

и каждое d_i выбирается из того же нормального распределения. Проблема состоит в том, что d_i при нашем подходе не выбирается из одного и того же нормального распределения, кроме того случая, когда мы применяем метод C в обеих фазах (предварительной и окончательной) и когда $M = N$.

Анализ выхода при моделировании также усложнен в следующих вариантах исходной задачи.

1. Вместо минимизации дисперсии, удовлетворяющей ограничению машинного времени, мы могли бы решить *обратную* задачу, т. е. минимизировать *машинное время*, необходимое для моделирования обеих систем:

$$T = t_1 M + t_2 N. \quad (302)$$

Эта минимизация ограничена условием *фиксированной надежности* оценки разности. Эту надежность можно измерить дисперсией. Следовательно, минимизация (302) подчинена условию

$$\text{var}(\bar{d}) = a_1 M^{-1} + a_2 N^{-1} = V, \quad (303)$$

где V — выбранная подходящая константа. Если мы знаем, что наблюдений достаточно много и можно воспользоваться (298), то берем константу равной

$$V = (\delta^* / z^{\alpha/2})^2, \quad (304)$$

предполагая, что мы хотим быть уверены на $(1 - \alpha)\%$ в том, что выберем систему с наивысшим средним, если лучшее среднее по крайней

мере на δ^* лучше, чем среднее следующей системы. Мы отсылаем к главе V для дальнейших рассуждений об определении V . Если мы не уверены, что получим очень много наблюдений, то возникает проблема, обсуждавшаяся нами в связи с (299), и определение V с помощью (303) становится трудным. Если V определено так, как это сделано в [Fishman, 1967, p. 6], то решение (302) и (303) можно получить в виде (305) и (306); это решение легко видоизменить, следуя приложению III.13⁶³.

$$M' = \{a_1 + (t_1^{-1} t_2 a_1 a_2)^{1/2}\} / V, \quad (305)$$

$$N' = \{a_2 + (t_1 t_2^{-1} a_1 a_2)^{1/2}\} / V. \quad (306)$$

2. Можно *обобщить* нашу задачу на случай $K (\geq 2)$ систем. Следуя [Lombaers, 1968, p. 254], мы хотим предложить такую процедуру. Предположим, что среди систем имеется одна контрольная система, т. е. есть $(K - 1)$ новых вариантов системы и один существующий вариант. Моделируются много опытов (например, M) контрольной системы и меньшее число (например, N) $(K - 1)$ -й экспериментальной системы. В [Scheffé, 1964, p. 88] указано, что при определенных предположениях оптимальное отношение есть $M/N = \sqrt{K} - 1$. Контрольная система генерируется с помощью дополняющих величин $(K - 1)$ экспериментальных систем — с использованием случайных чисел, общих с контрольной системой. Однако в этой процедуре опять появляется нежелательная отрицательная перекрестная ковариация c_4 . Помимо положительной перекрестной ковариации c_3 создается отрицательная перекрестная ковариация c_4 . Поэтому вместо этого возможен оптимизационный подход, предложенный ранее. Вместо минимизации $\text{var}(\bar{x} - \bar{y})$ можно минимизировать сумму дисперсий

$$\sum_{k=2}^K \text{var}(\bar{x}_1 - \bar{x}_k), \quad (307)$$

где \bar{x}_1 — среднее опытов с контрольной системой; \bar{x}_k — среднее экспериментальной системы. Если нет контрольной системы, то можно сравнить все системы друг с другом. Мы получим $K(K - 1)/2$ сравнений. Тогда (307) можно заменить выражением

$$\sum_{i=1}^{K-1} \sum_{k=i+1}^K \text{var}(\bar{x}_i - \bar{x}_k). \quad (308)$$

Оба выражения (307) и (308) ограничены условием на машинное время:

$$\sum_1^K t_k N_k = T. \quad (309)$$

Каждая из дисперсий (307) или (308) после знака суммирования выражается аналогично выражению в табл. 10. Ограничение (309) можно ввести в минимизацию (307) и (308) с помощью метода множителей Лагранжа. Мы должны исследовать больше случаев, чем в табл. 10, так как помимо сравнения N_1 с N_2 надо еще сравнить N_1 с N_3 и т. д. Оптимизацию можно применять, но она становится более трудоемкой.

Статистический анализ также превращается в серьезную проблему, потому что при анализе имитационного эксперимента с K системами мы должны применять метод *множественных сравнений*. Как мы увидим в главе V, этот метод дает совместные доверительные интервалы для сравнения K систем. К сожалению, вывод основан на независимости наблюдений. Следовательно, методы B и C будут противоречить этому предположению. Метод A не противоречит ему, так как можно взять средние дополняющих пар в качестве новых наблюдений аналогично (239). Далее рассмотрим *обратную* задачу. Если мы не ограничиваем возможное машинное время и, следовательно, число наблюдений, то можно продолжать моделирование до достижения желаемого уровня надежности. Для случая двух систем это было сформулировано в (302) и (303). Если имеется K систем, можно применить методы⁶⁴ *многомерного ранжирования*. В главе V будет показано, что эти методы гарантируют вероятность выбора наилучшей системы не ниже $(1 - \alpha)\%$. Одним из предположений при применении этих методов снова служит *независимость* наблюдений. Этому предположению удовлетворяет только метод A .

Мы видели, что возникают серьезные проблемы в статистическом анализе откликов двух и большего числа систем либо при минимизации дисперсии, либо при минимизации машинного времени. Следовательно, мы можем не применять описанную выше оптимизацию для выбора одного из трех методов A , B или C . Вместо этого можно априори выбрать *метод* A только с дополняющими величинами, тогда описанная процедура полезна для оптимального распределения машинного времени между системами. Из приложения III.12 следует, что коэффициенты a_1 и a_2 в методе A всегда положительны. Табл. 10 показывает, что если мы возьмем только метод A , то мы не сделаем различия между $M \geq N$ и $M \leq N$. Следовательно, оптимальные значения M и N можно определить без табл. 11 и (277). Вместо этого мы берем ограничения (278) — (280) и табл. 15.

Таблица 15

Оптимальные значения M и N при применении только дополняющих величин

Граничные условия	M_0	N_0
$M^* \geq N_p, N^* \geq N_p$	M^*	N^*
$M^* < N_p$	N_p	$(T - t_1 N_p) / t_2$
$N^* < N_p$	$(T - t_2 N_p) / t_1$	N_p

III.9. ЗАКЛЮЧЕНИЕ И ВЫВОДЫ

В этой главе мы познакомились с *несколькими* МПД, которые применяются при моделировании сложных систем, и это дает значительное понижение дисперсии. Два метода стоят в стороне из-за *чрезвычайной простоты*, а именно метод *дополняющих величин* и ме-

тод *общих случайных чисел*. Они легки в применении, так как требуют мало времени на программирование и экспериментирование и в то же время дают значительное понижение дисперсии. Более того, они не искажают выборочный процесс, значит помимо среднего значения отклика можно также изучать динамическое поведение системы (динамическое поведение системы интересно само по себе, оно может оказаться полезным для анализа обоснованности модели). Из этих двух методов наиболее широко применяется метод общих случайных чисел. Методом дополняющих величин, несмотря на его простоту, редко пользуются в моделировании сложных систем.

Было показано, что *совместное* применение этих двух методов не обязательно дает наибольшее понижение дисперсии. Мы предложили процедуру для выбора метода, дающего наименьшую дисперсию (метода *A* — дополняющих величин, метода *B* — общих случайных чисел и метода *C* — их совместное применение). Эта процедура дает оптимальное распределение ограниченного машинного времени для сравнения двух систем. Помимо повышения надежности моделируемого эксперимента мы обычно хотим еще и *измерить* эту надежность. Для измерения надежности мы применяли статистические методы, основанные на независимых наблюдениях. Метод общих случайных чисел нарушает это основное допущение; метод дополняющих величин дает независимые наблюдения, если брать среднее дополняющих пар.

До сих пор имитационные эксперименты проводились без формального статистического анализа их результатов. Мы думаем, что в будущем такой анализ будет проводиться чаще для получения достаточно надежных выводов. Поэтому мы считаем, что метод общих случайных чисел здесь менее полезен, так как он противоречит основному предположению большинства процедур статистического анализа; более полезным будет метод дополняющих величин.

III.10. ЛИТЕРАТУРА

Учебником по МПД, на который ссылаются многие авторы, в том числе и Нейлор [Naylor, 1971, p. 25], служит книга Хаммерсли и Хендскомба [Hammersley and Handscomb, 1964]. Эта книга может быть очень полезна для применения МПД к задачам Монте-Карло, но для моделирования, как мы считаем, она полностью непригодна. Диссертация Моя [Moу, 1965] дает хороший обзор применения методов МПД в моделировании (см. также [Moу, 1966]).

ПРИЛОЖЕНИЯ К ГЛАВЕ III

ПРИЛОЖЕНИЕ III.1. СТАТИСТИКА «СКЛАДНОГО НОЖА»

Цель этого метода — уменьшить *смещение* оценки и получить ее *доверительный интервал*. Процедура состоит в следующем. Допустим, что мы имеем оценку $\hat{\theta}$, основанную на n независимых наблюдениях x_1, \dots, x_n . Разделим выборку на N групп равного объема $M = n/N$. Образует подвыборку, исключая из исходной выборки одну

из N групп, и вычислим ту же оценку по оставшимся $(N - 1) M$ наблюдениям. Обозначим такую оценку для i -й группы символом $\hat{\theta}_i$ ($i = 1, \dots, N$). Тогда псевдозначения J_i статистики «складного ножа» определяются так:

$$J_i = n \hat{\theta} - (N - 1) \hat{\theta}_i \quad (i = 1, \dots, N), \quad (1.1)$$

а оценка параметра θ будет

$$\bar{J} = \sum_{i=1}^N J_i / N = N \hat{\theta} - (N - 1) \bar{\hat{\theta}}_i. \quad (1.2)$$

Эта статистика «складного ножа» была введена Кенуем [Queouille, 1956]; чтобы уменьшить возможное смещение оценки $\hat{\theta}$. Так, если θ имеет смещение порядка n^{-1} , то во многих случаях \hat{J} имеет смещение порядка n^{-2} [Gray and Schucany, 1972, p. 7]. Позднее Тьюки [Tukey, 1958] предложил рассматривать псевдозначения J_i как независимые одинаково распределенные переменные. Тогда приближенные доверительные интервалы для \bar{J} можно получить из t -распределения с $(N - 1)$ степенями свободы:

$$t_{N-1} \approx \frac{(\bar{J} - \theta) N^{1/2}}{\{\sum_i (J_i - \bar{J})^2 / (N - 1)\}^{1/2}}. \quad (1.3)$$

Мы отсылаем читателя к [Gray and Schucany, 1972, p. 139, 149, 158] для знакомства с доказательством асимптотической нормальности правой части выражения (1.3) для различных условий. Если в частном случае J_i может принимать только $N_1 < N$ определенных значений, то число степеней свободы будет равно $N_1 - 1$ [Gray and Schucany, 1972, p. 164]. Если значения J_i независимы, но при этом любые пары имеют положительную корреляцию ρ (например, $\rho = 1/n$), то правую часть выражения (1.3) надо умножить на корректирующий множитель

$$f = \left[\frac{1 - \rho}{1 + (N - 1)\rho} \right]^{1/2}, \quad (1.4)$$

который обоснован в [Gray and Schucany, 1972, p. 165, 173 — 174]. В качестве эвристической процедуры, которая требует дальнейшего исследования и осторожного применения, мы предлагаем оценить ρ (> 0) и подставить $\hat{\rho}$ в (1.4). Затем желательно уменьшить возможное отклонение $\hat{\theta}_i$ с помощью подходящего преобразования, например логарифмирования для оценки $\hat{\theta} = \hat{\sigma}^2$ (см. [Schucany and Gray, 1972, p. 169]).

Много приложений метода «складного ножа» можно найти в [Schucany and Gray, 1972], [Arvesen and Salsburg, 1975] и [Goodman et al., 1973]. Статистику «складного ножа» (1.2) можно обобщить;

$$G = \frac{\hat{\theta}_1 - R \hat{\theta}_2}{1 - R} \quad (R \neq 1), \quad (1.5)$$

// $\hat{\theta}_1$ и $\hat{\theta}_2$ — оценки θ . Очевидно, что (1.5) превращается в (1.2),
 и положим,

$$R = \frac{N-1}{N}, \quad \hat{\theta}_1 = \hat{\theta}, \quad \hat{\theta}_2 = \hat{\theta}_i. \quad (1.6)$$

Мы отсылаем читателя к [Arvesen and Salsburg, 1972] и [Gray Schucany, 1972] для более детального анализа статистики «складного ножа» (как исходного, так и обобщенного) и за дополнительными ссылками. Эта статистика может также применяться, если $\hat{\theta}$ основана не только на наблюдениях \mathbf{x} , но и на сопутствующих переменных $\bar{y}_1, \dots, \bar{y}_n$ (см. [Quepouille, 1956, p. 357 — 358], а также наш параграф, где изложен материал о контрольных величинах).

ПРИЛОЖЕНИЕ III.2. СТРАТИФИКАЦИЯ ПОСЛЕ ВЫБОРКИ С ОБЪЕДИНЕНИЕМ СЛОЕВ

В этом приложении мы покажем, как ограничить *смещение* при использовании оценки, основанной на стратификации после выборки с объединением слоев, если какой-нибудь слой останется пустым. Эта оценка \bar{x}_{SAC} была дана в (22) — (24) из III.2.2. Рассмотрим математические ожидания величины \bar{w}_k , определенной в (23а) и в (23б):

$$E(\bar{w}_k) = \sum_{i=0}^n E(\bar{w}_k | n_k = i) P(n_k = i) = \sum_{j=1}^n E(\bar{w}_k | n_k = j) P(n_k = j) + E(\bar{w}_k | n_k = 0) P(n_k = 0). \quad (2.1)$$

Из (23а) вытекает:

$$E(\bar{w}_k | n_k = j) = E\left(\sum_{h=1}^j x_{kh}/j\right) = \sum_{h=1}^j E(x_{kh})/j = \sum_{h=1}^j \mu_k/j = \mu_k, \quad (2.2)$$

а из (23б) вытекает:

$$E(\bar{w}_k | n_k = 0) = E\left(\sum_{h=1}^{n_g} x_{gh}/n_g\right), \quad (2.3)$$

где

$$n_g > 0. \quad (2.4)$$

Следовательно,

$$\begin{aligned} E(\bar{w}_k | n_k = 0) &= \sum_{j=1}^n E\left(\sum_{h=1}^{n_g} x_{gh}/n_g | n_g = j\right) P(n_g = j) = \\ &= \sum_{j=1}^n E\left(\sum_{h=1}^j x_{gh}/j\right) P(n_g = j) = \sum_{j=1}^n \mu_g P(n_g = j) = \\ &= \mu_g \sum_{j=1}^n P(n_g = j) = \mu_g, \end{aligned} \quad (2.5)$$

где последнее равенство включает

$$\sum_{j=1}^n P(\mathbf{n}_g = j) = 1, \quad (2.6)$$

что следует из (2.4). Подстановка (2.2) и (2.5) в (2.1) дает

$$E(\bar{w}_k) = \sum_{j=1}^n \mu_k P(\mathbf{n}_k = j) + \mu_g P(\mathbf{n}_k = 0) = \mu_k \sum_{j=1}^n P(\mathbf{n}_k = j) + \mu_g P(\mathbf{n}_k = 0). \quad (2.7)$$

Если

$$\mu_g = \mu_k \quad (2.8)$$

или имеет место равенство

$$P(\mathbf{n}_k = 0) = 0, \quad (2.9)$$

то можно заменить (2.7) выражением

$$E(\bar{w}_k) = \mu_k \sum_{j=0}^n P(\mathbf{n}_k = j) = \mu_k^1 = \mu_k, \quad (2.10)$$

которое, следуя (2.2), преобразуется так:

$$E(\bar{x}_{SAC}) = \sum_{k=1}^K p_k \mu_k = \mu. \quad (2.11)$$

Если же равенства (2.8) и (2.9) не имеют место, то рассмотренная оценка \bar{x}_{SAC} смещена.

ПРИЛОЖЕНИЕ III.3. СРЕДНЯЯ КВАДРАТИЧНАЯ ОШИБКА

Пусть мы хотим оценить параметр генеральной совокупности μ с помощью статистики \mathbf{y} . Так, например, μ может быть математическим ожиданием, \mathbf{y} — выборочным средним или медианой и т. д. Обозначим математическое ожидание \mathbf{y} через η , т. е.

$$E(\mathbf{y}) = \eta. \quad (3.1)$$

Тогда смещение оценки \mathbf{y} определяется как

$$\text{смещение} = E(\mathbf{y}) - \mu = \eta - \mu. \quad (3.2)$$

Дисперсия оценки \mathbf{y} определяется как

$$\sigma^2 = \text{var}(\mathbf{y}) = E\{[\mathbf{y} - E(\mathbf{y})]^2\} = E[(\mathbf{y} - \eta)^2]. \quad (3.3)$$

Средняя квадратичная ошибка, или, коротко, СКО, — это математическое ожидание квадрата отклонения между \mathbf{y} и тем параметром, который величина \mathbf{y} должна оценивать. Так,

$$\begin{aligned} \text{СКО}(\mathbf{y}) &= E[(\mathbf{y} - \mu)^2] = E\{[(\mathbf{y} - \eta) + (\eta - \mu)]^2\} = \\ &= E[(\mathbf{y} - \eta)^2 + (\eta - \mu)^2 + 2(\mathbf{y} - \eta)(\eta - \mu)] = \\ &= \text{var}(\mathbf{y}) + (\text{смещение})^2. \end{aligned} \quad (3.4)$$

Отметим, что СКО можно *оценить* по N выборкам (равного объема), определить для каждой выборки статистику y_i ($i = 1, \dots, N$) и вычислить оценку, которая дается в (3.5) при предположении, что μ известна:

$$\begin{aligned} \widehat{\text{СКО}} &= \sum_{i=1}^N (y_i - \mu)^2 / N = \sum_{i=1}^N \{(y_i - \bar{y}) + (\bar{y} - \mu)\}^2 / N = \\ &= \sum_{i=1}^N (y_i - \bar{y})^2 / N + \sum_{i=1}^N (\bar{y} - \mu)^2 / N + \sum_{i=1}^N 2(y_i - \bar{y})(\bar{y} - \mu) / N = \\ &= \hat{\sigma}^2 + (\bar{y} - \mu)^2, \end{aligned} \quad (3.5)$$

где \bar{y} — среднее значение оценки в N повторных выборках, т. е.

$$\bar{y} = \sum_{i=1}^N y_i / N \quad (3.6)$$

и $\hat{\sigma}^2$ — оценка дисперсии y , которая дается формулой

$$\hat{\sigma}^2 = \sum_{i=1}^N (y_i - \bar{y})^2 / N. \quad (3.7)$$

Математическое ожидание величины $\hat{\sigma}^2$ можно найти, например, в [Fisz, 1967, p. 347], оно равно:

$$E(\hat{\sigma}^2) = \frac{N-1}{N} \sigma^2. \quad (3.8)$$

Математическое ожидание квадрата оцениваемого смещения (это можно легко показать) есть

$$E(\bar{y} - \mu)^2 = \frac{\sigma^2}{N} + (\text{смещение})^2. \quad (3.9)$$

Формулы (3.8) и (3.9) показывают, что выражение (3.5) — несмещенная оценка СКО:

$$\begin{aligned} E(\widehat{\text{СКО}}) &= E(\hat{\sigma}^2) + E[(\bar{y} - \mu)^2] = \frac{N-1}{N} \sigma^2 + \frac{1}{N} \sigma^2 + (\text{смещение})^2 = \\ &= \sigma^2 + (\text{смещение})^2. \end{aligned} \quad (3.10)$$

ПРИЛОЖЕНИЕ III.4. ОЦЕНКА ДИСПЕРСИИ И СТРАТИФИКАЦИИ ПОСЛЕ ВЫБОРКИ

В стратификации после выборки число наблюдений на один слой — случайная величина. Чтобы получить дисперсию стратифицированной оценки, воспользуемся (4.1). Это соотношение доказано в [Keeping, 1962, p. 398 — 399].

$$\text{var}(\mathbf{x}) = E_y [\text{var}(\mathbf{x} | y)] + \text{var}_y [E(\mathbf{x} | y)]. \quad (4.1)$$

Рассмотрим дисперсию $\bar{\vartheta}_h$ как среднюю величину для слоя, когда мы продолжаем сбор наблюдений, пока не заполнится каждый слой, это определено в (29) — (32) из III.2.2:

$$\begin{aligned} \text{var}(\bar{\vartheta}_h) &= E \left[\text{var} \left(\sum_{j=1}^{n_h} \mathbf{x}_{hj} / n_h \mid n_h = n_h \right) \right] + \\ &+ \text{var} \left[E \left(\sum_{j=1}^{n_h} \mathbf{x}_{hj} / n_h \mid n_h = n_h \right) \right] = E \left[\text{var}(\mathbf{x}_{hj}) / n_h \right] + \text{var}[\mu_h] = \\ &= \text{var}(\mathbf{x}_{hj}) E(1/n_h) + 0, \end{aligned} \quad (4.2)$$

из этого мы сделали вывод, что $\text{var}(\mathbf{x}_{hj})$ и μ_h — неслучайные величины.

Рассмотрим оценку параметра $\text{var}(\bar{\vartheta}_h)$ в (4.3):

$$\hat{\text{var}}(\bar{\vartheta}_h) = \frac{\sum_{j=1}^{n_h} (\mathbf{x}_{hj} - \bar{\vartheta}_h)^2}{(n_h - 1)} \cdot \frac{1}{n_h}. \quad (4.3)$$

Как видно из (4.3), мы заменяем исходное условие (31) модифицированным условием

$$n_h > 1 \text{ для всех } k. \quad (4.4)$$

Тогда (4.3) будет несмещенной оценкой (4.2), поскольку

$$\begin{aligned} E[\hat{\text{var}}(\bar{\vartheta}_h)] &= \sum_{n_h} E[\hat{\text{var}}(\bar{\vartheta}_h) \mid n_h = n_h] P(n_h = n_h) = \\ &= \sum_{n_h} E \left(\frac{\sum_{j=1}^{n_h} (\mathbf{x}_{hj} - \bar{\vartheta}_h)^2}{(n_h - 1)} \cdot \frac{1}{n_h} \right) P(n_h = n_h) = \\ &= \sum_{n_h} \frac{\text{var}(\mathbf{x}_{hj})}{n_h} P(n_h = n_h) = \text{var}(\mathbf{x}_{hj}) E(1/n_h). \end{aligned} \quad (4.5)$$

Из (29) следует, что стратифицированная оценка $\bar{\mathbf{x}}_{SAP}$ — это взвешенное среднее независимых средних по слоям $\bar{\vartheta}_h$. Поэтому ее дисперсия имеет вид

$$\text{var}(\bar{\mathbf{x}}_{SAP}) = \sum_{k=1}^K p_k^2 \text{var}(\bar{\vartheta}_k). \quad (4.6)$$

Подставляя (4.3), т. е. несмещенную оценку дисперсии $\text{var}(\bar{\vartheta}_h)$, в (4.6), получим несмещенную оценку дисперсии $(\bar{\mathbf{x}}_{SAP})$.

ПРИЛОЖЕНИЕ III. ПРИМЕР ГЕНЕРИРОВАНИЯ СЛУЧАЙНЫХ ЧИСЕЛ
В СТРАТИФИЦИРОВАННОЙ ВЫБОРКЕ

В этом приложении мы продемонстрируем рациональную процедуру Моя для генерирования вектора m случайных чисел, когда g случайных чисел не должны быть меньше, чем константа c . Процедура Моя была дана на рис. 10 в III.2.4. Предположим, что $m=4$ и $g=2$. Тогда мы имеем $C_4^2 = 6$ различных векторов, как показано в таблице III.5.1. Из этой таблицы видно, что:

$$\begin{aligned}
 P(r_1 \geq c) &= 3/6 = g/m = 2/4; \\
 &\quad (\text{сравнение векторов 1, 2 и 3}), \\
 P(r_2 \geq c | r_1 \geq c) &= 1/3 = (g-1)/(m-1) \\
 &\quad (\text{сравнение векторов 1, 2 и 3}), \\
 P(r_2 \geq c | r_1 < c) &= 2/3 = g/(m-1) \\
 &\quad (\text{сравнение векторов 4, 5 и 6}), \\
 P(r_3 \geq c | r_1 \geq c \text{ и } r_2 \geq c) &= 0 = (g-2)/(m-2) \\
 &\quad (\text{сравнение вектора 1}), \\
 P(r_3 \geq c | r_1 \geq c \text{ и } r_2 < c) &= 1/2 = (g-1)/(m-2) \\
 &\quad (\text{сравнение векторов 2 и 3}).
 \end{aligned}$$

Таблица III.5.1

Шесть различных векторов при $m=4$ и $g=2^1$

Вектор	r_1	r_2	r_3	r_4
1	*	*		
2	*		*	
3	*			*
4		*	*	
5		*		*
6			*	*

¹ Знак * означает, что $r \geq c$.

ПРИЛОЖЕНИЕ III.6. МАТЕМАТИЧЕСКОЕ ОЖИДАНИЕ КОНТРОЛЬНОЙ
ВЕЛИЧИНЫ ПРИ СЛУЧАЙНОМ m

Контрольная величина — это

$$b_1 + b_2 \sum_1^m y_t + b_3 \left(\sum_1^m y_t \right)^2. \quad (6.1)$$

Взяв обозначения $\eta_1 = E(y_t)$, $\eta_2 = E(y_t^2)$ и $\sigma_1^2 = \eta_2 - \eta_1^2 =$
 $= \text{var}(y_t)$, получим

$$E \left(\sum_1^m y_t \right) = E \left[E \sum_1^m y_t | m = m \right] = E [m\eta_1] = \eta_1 E(m) \quad (6.2)$$

и

$$E \left[\left(\sum_1^m y_t \right)^2 \right] = E \left\{ E \left[\left(\sum_1^m y_t \right)^2 \middle| m = m \right] \right\}, \quad (6.3)$$

где

$$\begin{aligned} E \left[\left(\sum_1^m y_t \right)^2 \right] &= E \left[\left(\sum_{t=1}^m y_t \right) \left(\sum_{s=1}^m y_s \right) \right] = E \left[\sum_{t=1}^m \sum_{s=1}^m y_t y_s \right] = \\ &= \sum_{t \neq s} E(y_t y_s) + \sum_{t=1}^m E(y_t^2) = \sum_{t \neq s} E(y_t) E(y_s) + m\eta_2 = m(m-1)\eta_1^2 + \\ &+ m\eta_2 = m^2\eta_1^2 + m(\eta_2 - \eta_1^2) = m^2\eta_1^2 + m\sigma_1^2. \end{aligned} \quad (6.4)$$

В (6.4) η_1 и σ_1 можно вычислить из распределения входной переменной y_t . Подстановка (6.4) в (6.3) дает

$$E \left[\left(\sum_1^m y_t \right)^2 \right] = E \left[\eta_1^2 m^2 + \sigma_1^2 m \right] = \eta_1^2 E(m^2) + \sigma_1^2 E(m). \quad (6.5)$$

ПРИЛОЖЕНИЕ III.7. КОВАРИАЦИИ МЕЖДУ ОЦЕНКАМИ КОНТРОЛЬНЫХ ВЕЛИЧИН

Определим оценки контрольной величины, как в (141), т. е.

$$\boldsymbol{x}cze_{j_i} = \boldsymbol{x}_{j_i} = \sum_{k=1}^K \hat{\boldsymbol{a}}_{kj} \boldsymbol{z}_{kji}. \quad (7.1)$$

Рассмотрим ковариацию между двумя оценками внутри j -й группы:

$$\begin{aligned} \text{cov}(\boldsymbol{x}cze_{j_i}, \boldsymbol{x}cze_{j_h}) &= E \left[\left\{ \boldsymbol{x}_{j_h} - \sum_k \hat{\boldsymbol{a}}_{kj} \boldsymbol{z}_{kji} - \mu \right\} \left\{ \boldsymbol{x}_{j_h} - \right. \right. \\ &\quad \left. \left. - \sum_k \hat{\boldsymbol{a}}_{kj} \boldsymbol{z}_{kjh} - \mu \right\} \right], \end{aligned} \quad (7.2)$$

где μ — математическое ожидание случайной величины, $\boldsymbol{x}cze_{j_i}$, оно равно математическому ожиданию \boldsymbol{x}_{j_i} , поскольку обе случайные величины несмещенные. Так,

$$\begin{aligned} \text{cov}(\boldsymbol{x}cze_{j_i}, \boldsymbol{x}cze_{j_h}) &= E[(\boldsymbol{x}_{j_i} - \mu)(\boldsymbol{x}_{j_h} - \mu)] - \\ &- E \left[(\boldsymbol{x}_{j_i} - \mu) \left(\sum_k \hat{\boldsymbol{a}}_{kj} \boldsymbol{z}_{kjh} \right) \right] - E \left[\left(\sum_k \hat{\boldsymbol{a}}_{kj} \boldsymbol{z}_{kji} \right) (\boldsymbol{x}_{j_h} - \mu) \right] + \\ &+ E \left[\left(\sum_k \hat{\boldsymbol{a}}_{kj} \boldsymbol{z}_{kji} \right) \left(\sum_k \hat{\boldsymbol{a}}_{kj} \boldsymbol{z}_{kjh} \right) \right]. \end{aligned} \quad (7.3)$$

Первый член в формуле (7.3) — это ковариация между двумя независимыми r раз повторенными имитационными опытами. Поэтому он

равен нулю. Во втором члене первый элемент $(\mathbf{x}_{ij} - \mu)$ независим от второго элемента, поскольку $\hat{\mathbf{a}}_{kj}$ оценивается по наблюдениям вне j -й группы. Величина \mathbf{z}_{hjh} относится к реализации h , которая не зависит от i -й реализации. Обе эти реализации принадлежат j -й группе. Следовательно,

$$E \left[(\mathbf{x}_{ji} - \mu) \left(\sum_k \hat{\mathbf{a}}_{kj} \mathbf{z}_{hjh} \right) \right] = E [(\mathbf{x}_{ji} - \mu)] E \left[\left(\sum_k \hat{\mathbf{a}}_{kj} \mathbf{z}_{hjh} \right) \right] = 0. \quad (7.4)$$

Таким образом, мы получили, что и третий член в формуле (7.3) равен нулю. Четвертый член можно переписать так:

$$E \left[\sum_k \sum_{k'} \hat{\mathbf{a}}_{kj} \mathbf{z}_{hji} \hat{\mathbf{a}}_{k'j} \mathbf{z}_{k'jh} \right] = \sum_k \sum_{k'} E (\hat{\mathbf{a}}_{kj} \hat{\mathbf{a}}_{k'j}) E (\mathbf{z}_{hji}) E (\mathbf{z}_{k'jh}). \quad (7.5)$$

Поскольку $\hat{\mathbf{a}}_{kj}$ и $\hat{\mathbf{a}}_{k'j}$ оцениваются по реализации вне j -й группы, \mathbf{z}_{hji} и $\mathbf{z}_{k'jh}$ независимы, так как реализации с номерами i и h независимы. Поскольку контрольные величины имеют нулевое математическое ожидание, четвертый член в (7.3) равен нулю. Таким образом, ковариация между двумя оценками внутри одной и той же группы равна нулю.

Теперь рассмотрим ковариацию между двумя наблюдениями в различных группах, например i -м наблюдением в j -й группе и h -м наблюдением в p -й группе. Поступим аналогично тому, как в формуле (7.3):

$$\begin{aligned} \text{cov}(\mathbf{x}cze_{ji}, \mathbf{x}cze_{ph}) &= E [(\mathbf{x}_{ji} - \mu) (\mathbf{x}_{ph} - \mu)] - E \left((\mathbf{x}_{ji} - \mu) \left(\sum_k \hat{\mathbf{a}}_{kp} \times \right. \right. \\ &\quad \left. \left. \times \mathbf{z}_{kph} \right) \right) - E \left[\left(\sum_k \hat{\mathbf{a}}_{kj} \mathbf{z}_{kji} \right) (\mathbf{x}_{ph} - \mu) \right] + E \left[\left(\sum_k \hat{\mathbf{a}}_{kj} \mathbf{z}_{kji} \right) \times \right. \\ &\quad \left. \times \left(\sum_k \hat{\mathbf{a}}_{kp} \mathbf{z}_{kph} \right) \right]. \end{aligned} \quad (7.6)$$

Первый член равен нулю. Второй член

$$E \left[\sum_k \hat{\mathbf{a}}_{kp} (\mathbf{x}_{ji} - \mu) \mathbf{z}_{kph} \right] = \sum_k E [\hat{\mathbf{a}}_{kp} (\mathbf{x}_{ji} - \mu)] E (\mathbf{z}_{kph}) = 0. \quad (7.7)$$

Таким образом, видно, что третий член также равен нулю. Последний член равен

$$\sum_k \sum_{k'} E (\hat{\mathbf{a}}_{kj} \mathbf{z}_{hji} \hat{\mathbf{a}}_{kp} \mathbf{z}_{kph}), \quad (7.8)$$

где \mathbf{z}_{hji} нельзя рассматривать отдельно, поскольку эта величина не независима от $\hat{\mathbf{a}}_{kp}$. Следовательно, ковариация между двумя оценками, принадлежащими двум различным группам, не равна нулю.

ПРИЛОЖЕНИЕ III.8. АППРОКСИМАЦИЯ ОПТИМАЛЬНОЙ ПЛОТНОСТИ РАСПРЕДЕЛЕНИЯ ГАММА-РАСПРЕДЕЛЕНИЕМ В ПРИМЕРЕ ЗНАЧИМОЙ ВЫБОРКИ

Рис. 11 из III.5.2 предполагает, что мы можем аппроксимировать $h_0(x)$ непрерывным распределением, имеющим моду, меньшую, чем математическое ожидание (мода — это такое значение случайной ве-

личины, которой соответствует наибольшая вероятность^{*}). Большинство стандартных распределений, однако, либо симметричны, либо дискретны. Гамма-распределение обладает нужными свойствами. Функция плотности этого распределения имеет вид

$$G(x) = \frac{b^p}{\Gamma(p)} x^{p-1} e^{-bx} \quad (x, b, p > 0). \quad (8.1)$$

После выбора типа распределения нужно определить значения параметров. Генерирование значения x , подчиняющегося гамма-распределению, осуществить просто, если параметр p в (8.1) *целочисленный*. В [Naylor et al., 1967a, p. 88] показано, что такую переменную гамма-распределения можно генерировать с помощью суммирования p переменных, имеющих экспоненциальное распределение. Далее известно (см. [Keeping, 1962, p. 81]), что гамма-распределение становится более симметричным при уменьшении p . Наименьшее целое $p = 1$, однако, приводит к экспоненциальному распределению, которое непрерывно уменьшается с ростом x . Сравнение с рис. 11 из III.5.2 показывает, что нам нужна такая функция плотности, которая не сразу спадает с ростом x . Наоборот, эта функция должна иметь наибольшее значение для $x > 0$. Поэтому выберем следующее по порядку наименьшее значение, т. е. $p = 2$. Тогда новая функция плотности $h_2(x)$ получается при подстановке $p = 2$ в (8.1) и определяется таким образом:

$$h_2(x) = b^2 x e^{-bx} \quad (x, b > 0). \quad (8.2)$$

Из (158), (164), (165) и (8.2) следует:

$$g_2^*(x) = 0, \quad \text{если } 0 < x < v \\ = \frac{x^{-1} \lambda e^{-\lambda x}}{b^2 x e^{-bx}} = \frac{\lambda}{b^2} x^{-2} e^{-(\lambda-b)x}, \quad \text{если } x \geq v. \quad (8.3)$$

В [Kleijnen, 1968, p. 197 — 198] доказано, что нельзя вычислить значение b , которое минимизирует дисперсию $g_2^*(x)$, поскольку нужно знать ξ , которую еще саму надо оценить. Поэтому, как объяснялось в III.5.2, мы минимизируем размах функции $g_2^*(x)$. Соответственно, чтобы этот размах был конечным, положим $b \leq \lambda$.

Тогда

$$R[g_2^*(x)] = \max_x [g_2^*(x)] - \min_x [g_2^*(x)] = g_2^*(v) - 0 = \\ = \frac{\lambda}{b^2} v^{-2} e^{-(\lambda-b)v}. \quad (8.4)$$

Следующее значение b , которое минимизирует R , находится дифференцированием R по b , приравниванием производной нулю и решением полученного уравнения относительно b . Эта процедура приводит к значению $b = 2/v$. Полученное значение можно подставить в (8.2) и (8.3), чтобы найти новую плотность распределения $h_2(x)$ и оценку $g_2^*(x)$ в аппроксимации гамма-распределением. Заметим, что новая

^{*} На самом деле ей соответствует максимум функции плотности распределения. — *Прим. пер.*

плотность распределения имеет моду $v/2$ и математическое ожидание v . При выводе формулы (8.4) предполагалось выполнение условия $b \leq \leq \lambda$, которое может противоречить условию $b = 2/v$. Мы отсылаем читателя к [Kleijnen, 1968b, p. 185 — 186], где обсуждается этот момент. Там показано также, что эта проблема не является серьезной. Следуя (174) и (175), мы можем доказать, что размах $g_2^*(x)$ меньше, чем размах $g(x)$, грубой оценки. Дисперсия $g_2^*(x)$ обсуждалась в III.5.2.

ПРИЛОЖЕНИЕ III.9. ДИСПЕРСИЯ ОЦЕНКИ ЗНАЧИМОЙ ВЫБОРКИ $g_3^*(x)$

Оценка значимой выборки $g_3^*(x)$ определяется как

$$g_3^*(x) = \frac{1}{x} \lambda v \exp \left[- \left(\lambda - \frac{1}{v} \right) x \right], \quad \text{если } x \geq v, \\ = 0, \quad \text{если } 0 \leq x < v. \quad (9.1)$$

Дисперсия $g_3^*(x)$ вычисляется из

$$\text{var } [g_3^*(x)] = E\{[g_3^*(x)]^2\} - E[g_3^*(x)] E[g_3^*(x)]. \quad (9.2)$$

Поскольку $g_3^*(x)$ — несмещенная оценка функции $\xi(\lambda, v)$, имеем

$$E[g_3^*(x)] = \xi(\lambda, v). \quad (9.3)$$

Далее,

$$E\{[g_3^*(x)]^2\} = \int_{-\infty}^{\infty} \{g_3^*(x)\}^2 h_3(x) dx, \quad (9.4)$$

где $h_3(x)$ — новая плотность распределения x , т. е.

$$h_3(x) = \frac{1}{v} \exp \left(- \frac{1}{v} x \right), \quad \text{если } x \geq 0 \\ = 0, \quad \text{если } x < 0. \quad (9.5)$$

Следовательно,

$$E\{[g_3^*(x)]^2\} = \int_v^{\infty} \frac{1}{x^2} (\lambda v)^2 \exp \left[- 2 \left(\lambda - \frac{1}{v} \right) x \right] \frac{1}{v} \exp \left(- \frac{1}{v} x \right) dx = \\ = \lambda^2 v \int_v^{\infty} \frac{1}{x^2} \exp \left[- \left(2\lambda - \frac{1}{v} \right) x \right] dx. \quad (9.6)$$

Интегрирование по частям даст

$$E\{[g_3^*(x)]^2\} = \lambda^2 [e^{(1-2\lambda v)}] + (1-2\lambda v) \int_v^{\infty} \frac{1}{x} \exp \times \\ \times \left[- \left(2\lambda - \frac{1}{v} \right) x \right] dx, \quad (9.7)$$

что доказывает необходимость $\lambda > 2/v$, иначе интеграл расходится. Взяв символ ψ ,

$$\psi = -2\lambda + \frac{1}{v}, \quad (9.8)$$

и определяя $\xi(\lambda, v)$ в (163), получим

$$E \{ [g_3^*(\mathbf{x})]^2 \} = \lambda^2 [e^{\psi v} - v\xi(-\psi, v)]. \quad (9.9)$$

Следовательно,

$$\text{var} [g_3^*(\mathbf{x})] = \lambda^2 [e^{\psi v} - v\xi(-\psi, v)] - \{ \xi(\lambda, v) \}^2. \quad (9.10)$$

ПРИЛОЖЕНИЕ III.10. ГЕНЕРИРОВАНИЕ ДОПОЛНЯЮЩИХ СЛУЧАЙНЫХ ЧИСЕЛ ИЗ МУЛЬТИПЛИКАТИВНО-КОНГРУЭНТНОГО ОТНОШЕНИЯ

Мультипликативно-конгруэнтное отношение для генерирования случайных чисел задается следующим образом:

$$x_i = ax_{i-1} \pmod{m} \quad (i = 1, 2, \dots), \quad (10.1)$$

где a — целое число, а x_0 — константа, меньшая, чем m . Дополняющие случайные числа r_i получают из соотношения

$$r_i = 1 - \frac{x_i}{m}. \quad (10.2)$$

Можно генерировать такие же случайные числа с помощью (10.3) — (10.5):

$$y_i = ay_{i-1} \pmod{m}, \quad (10.3)$$

$$y_0 = m - x_0, \quad (10.4)$$

$$r'_i = \frac{y_i}{m}. \quad (10.5)$$

Покажем, что r_i и r'_i совпадают. Для этого сначала докажем, что

$$x_i + y_i = m \text{ для любого } i. \quad (10.6)$$

Из (10.1) следует, что

$$x_i = ax_{i-1} \pmod{m} = ax_{i-1} - pm, \quad (10.7)$$

где p — целое число, но не нуль. Таким образом, из (10.3) имеем

$$y_i = ay_{i-1} \pmod{m} = ay_{i-1} - qt. \quad (10.8)$$

Объединение (10.7) и (10.8) дает

$$x_i + y_i = a(x_{i-1} + y_{i-1}) - (p + q)m. \quad (10.9)$$

Если

$$x_{i-1} + y_{i-1} = m, \quad (10.10)$$

то (10.9) преобразуется к виду

$$x_i + y_i = (a - p - q)m, \quad (10.11)$$

что означает кратность суммы $x_i + y_i$ значению m . Поскольку, однако, $0 < x_i < m$ и $0 < y_i < m$, то должно быть

$$x_i + y_i = m. \quad (10.12)$$

Итак, если

$$x_{i-1} + y_{i-1} = m, \text{ то } x_i + y_i = m. \quad (10.13)$$

Из (10.4) следует, что

$$x_0 + y_0 = m. \quad (10.14)$$

Воспользуемся (10.13) для $i = 1$, получим

$$x_1 + y_1 = m. \quad (10.15)$$

При $i = 2$

$$x_2 + y_2 = m \quad (10.16)$$

и т. д. В общем случае

$$x_i + y_i = m \quad (i = 1, 2, \dots). \quad (10.17)$$

Далее покажем, что

$$r_i = r'_i \quad (10.18)$$

или

$$r_i - r'_i = 0. \quad (10.19)$$

На основании (10.2) и (10.5) имеем

$$r_i - r'_i = 1 - \frac{x_i}{m} - \frac{y_i}{m}. \quad (10.20)$$

Подстановка (10.7) и (10.8) в (10.20) дает

$$r_i - r'_i = 1 - \frac{ax_{i-1} - pm}{m} - \frac{ay_{i-1} - qm}{m} = 1 + p + q - \frac{a(x_{i-1} + y_{i-1})}{m}. \quad (10.21)$$

С помощью (10.17) получим

$$r_i - r'_i = 1 + p + q - a. \quad (10.22)$$

Поскольку p , q и a — целые числа, в то время как $0 < r_i < 1$ и $0 < r'_i < 1$, имеем

$$r_i - r'_i = 0. \quad (10.23)$$

ПРИЛОЖЕНИЕ III.11. ДИСПЕРСИЯ ОЦЕНКИ РАЗНОСТИ МЕЖДУ ОТКЛИКАМИ ДВУХ СИСТЕМ

Различие между откликами двух систем оценивается следующим образом:

$$\bar{a} = \sum_1^M x_i/M - \sum_1^N y_j/N. \quad (11.1)$$

Поэтому

$$\begin{aligned} \text{var}(\bar{a}) &= \text{var}(\sum x_i/M) + \text{var}(\sum y_j/N) - \\ &\quad - 2 \text{cov}(\sum x_i/M, \sum y_j/N) = \\ &= M^{-2} \text{var}(\sum x_i) + N^{-2} \text{var}(\sum y_j) - \\ &\quad - 2 M^{-1} N^{-1} \text{cov}(\sum x_i, \sum y_j). \end{aligned} \quad (11.2)$$

Вспомним, что

$$\begin{aligned} \text{var} \left(\sum_t \mathbf{x}_t \right) &= E \left[\left\{ \sum_t \mathbf{x}_t - E \left(\sum_t \mathbf{x}_t \right) \right\}^2 \right] = E \left[\left\{ \sum_t (\mathbf{x}_t - E(\mathbf{x}_t)) \right\}^2 \right] = \\ &= E \left[\sum_t \sum_g \{ \mathbf{x}_t - E(\mathbf{x}_t) \} \{ \mathbf{x}_g - E(\mathbf{x}_g) \} \right] = \sum_t \sum_g E \{ \{ \mathbf{x}_t - E(\mathbf{x}_t) \} \{ \mathbf{x}_g - \\ &\quad - E(\mathbf{x}_g) \} \} = \sum_t \sum_g \text{cov}(\mathbf{x}_t, \mathbf{x}_g). \end{aligned} \quad (11.3)$$

Следовательно,

$$\begin{aligned} \text{var}(\bar{\mathbf{d}}) &= M^{-2} \sum_{i=1}^M \sum_{g=1}^M \text{cov}(\mathbf{x}_i, \mathbf{x}_g) + N^{-2} \sum_{j=1}^N \sum_{h=1}^N \text{cov}(\mathbf{y}_j, \mathbf{y}_h) - \\ &\quad - 2M^{-1} N^{-1} \sum_{i=1}^M \sum_{j=1}^N \text{cov}(\mathbf{x}_i, \mathbf{y}_j) = M^{-2} \sum_{i \neq g}^M \sum_{j \neq h}^N \text{cov}(\mathbf{x}_i, \mathbf{x}_g) + \\ &\quad + M^{-2} \sum_i^M \text{var}(\mathbf{x}_i) + N^{-2} \sum_{j \neq h}^N \sum_{h}^N \text{cov}(\mathbf{y}_j, \mathbf{y}_h) + N^{-2} \sum_i^N \text{var}(\mathbf{y}_j) + \\ &\quad + 2M^{-1} N^{-1} \sum_i^M \sum_j^N \text{cov}(\mathbf{x}_i, \mathbf{y}_j), \end{aligned} \quad (11.4)$$

где можно заменить $\text{var}(\mathbf{x}_i)$ на σ_1^2 , поскольку все \mathbf{x}_i имеют одинаковые дисперсии. Соответственно $\text{var}(\mathbf{y}_j) = \sigma_2^2$. Это приводит к

$$\begin{aligned} \text{var}(\bar{\mathbf{d}}) &= M^{-2} \sum_{i \neq g}^M \sum_{j \neq h}^N \text{cov}(\mathbf{x}_i, \mathbf{x}_g) + M^{-1} \sigma_1^2 + \\ &\quad + N^{-2} \sum_{j \neq h}^N \sum_{h}^N \text{cov}(\mathbf{y}_j, \mathbf{y}_h) + N^{-1} \sigma_2^2 - \\ &\quad - 2M^{-1} N^{-1} \sum_i^M \sum_j^N \text{cov}(\mathbf{x}_i, \mathbf{y}_j). \end{aligned} \quad (11.5)$$

ПРИЛОЖЕНИЕ III.12. ЗНАКИ КОЭФФИЦИЕНТОВ a_1 И a_2

В этом приложении мы изучим знаки коэффициентов a_1 и a_2 в выражении (274) для дисперсии $\text{var}(\bar{\mathbf{d}})$. Из табл. 10 следует, что нужно рассмотреть три случая.

1. $\sigma_1^2 + c_1$. Из определения σ_1^2 и c_1 в (250) и в (255) соответственно вытекает, что $(\sigma_1^2 + c_1)$ отрицательно, если

$$\text{var}(\mathbf{x}_i) < -\text{cov}(\mathbf{x}_{2i}, \mathbf{x}_{2i-1}) \quad (i = 1, 2, \dots), \quad (12.1)$$

или

$$\rho_1 < -1, \quad (12.2)$$

где ρ_1 обозначает отрицательный коэффициент корреляции между \mathbf{x}_{2i} и \mathbf{x}_{2i-1} . Поскольку коэффициент корреляции должен находиться

в пределах $[-1, 1]$, можно заключить, что $(\sigma_1^2 + c_1)$ всегда положительная величина. Так же можно показать, что и $(\sigma_2^2 + c_2)$ всегда положительно.

2. $\sigma_1^2 - 2c_3$. Из определения c_3 в (257) видно, что $(\sigma_1^2 - 2c_3)$ отрицательно, если

$$\text{var}(\mathbf{x}_i) < 2 \text{cov}(\mathbf{x}_i, \mathbf{y}_i) \quad (i = 1, 2, \dots) \quad (12.3)$$

или

$$\rho_3 > \frac{1}{2} \frac{\sigma(\mathbf{x}_i)}{\sigma(\mathbf{y}_i)}, \quad (12.4)$$

где ρ_3 означает положительный коэффициент корреляции между \mathbf{x}_i и \mathbf{y}_i . Следовательно, если существует сильная (положительная) корреляция между \mathbf{x}_i и \mathbf{y}_i , то $(\sigma_1^2 - 2c_3)$ может быть отрицательной величиной. Таким же образом можно сделать выводы, что $(\sigma_2^2 - 2c_3)$ будет отрицательной или положительной величиной в зависимости от силы корреляционной связи.

3. $\sigma_1^2 + c_1 - 2c_3 - 2c_4$. Как было показано, $(\sigma_1^2 + c_1 - 2c_3 - 2c_4)$ отрицательно, если

$$\frac{\sigma(\mathbf{x}_i)}{\sigma(\mathbf{y}_{2i-1})} < 2 \frac{(\rho_3 + \rho_4)}{(1 + \rho_1)}, \quad (12.5)$$

где ρ обозначает нежелательную отрицательную корреляцию между \mathbf{x}_{2i} и \mathbf{y}_{2i-1} . Следовательно, заключаем, что коэффициент $(\sigma_1^2 + c_1 - 2c_3 - 2c_4)$ может быть положительным или отрицательным. Подобный вывод правомерен и по поводу величины $(\sigma_2^2 + c_2 - 2c_3 - 2c_4)$.

Далее, известно, что σ_1^2 и σ_2^2 положительны. Тогда, как видно из табл. 10, существуют две возможности:

1) оба коэффициента a_1 и a_2 положительны (это всегда верно для метода А);

2) один коэффициент положителен, а другой отрицателен; не может быть такой ситуации, когда одновременно оба коэффициента отрицательны (это неудивительно, поскольку иначе дисперсия должна быть отрицательной).

ПРИЛОЖЕНИЕ III.13. ОПТИМАЛЬНОЕ РАСПРЕДЕЛЕНИЕ МАШИННОГО ВРЕМЕНИ

В этом приложении мы выведем значения M и N , т. е. число опытов с системами 1 и 2 соответственно, которые минимизируют дисперсию оценки разности между откликами систем 1 и 2. Как было показано в табл. 10, эту дисперсию можно аппроксимировать

$$\text{var}(\bar{d}) = aM^{-1} + a_2N^{-1}, \quad (13.1)$$

где частные значения a_1 и a_2 обоснованы, например, для случая

$$M \geq N. \quad (13.2)$$

Числа M и N не могут быть меньше, чем N_p — число предварительных опытов с каждой системой, т. е.

$$M \geq N_p, N \geq N_p, \quad (13.3)$$

где постоянная N_p , очевидно, удовлетворяет неравенству

$$0 < N_p < T/(t_1 + t_2), \quad (13.4)$$

поскольку имеется ограниченное машинное время, как показано в (13.5):

$$t_1 M + t_2 N = T \quad (t_1, t_2, T > 0). \quad (13.5)$$

Следовательно, нужно минимизировать (13.1) по отношению к (13.2), (13.3) и (13.5). Условия (13.2) и (13.3) совместно эквивалентны неравенству

$$M \geq N \geq N_p. \quad (13.6)$$

Условие (13.5) эквивалентно равенству

$$M = t_1^{-1} T - t_1^{-1} t_2 N, \quad (13.7)$$

поскольку $M \geq N$, если $N \leq T(t_1 + t_2)^{-1}$. Следовательно, (13.6) и (13.7) совместно эквивалентны неравенству

$$N_p \leq N \leq T(t_1 + t_2)^{-1}. \quad (13.8)$$

Таким образом, мы сократили число ограничений до одного ограничения (13.8). Подстановка (13.7) в (13.1) дает

$$\text{var}(\bar{d}) = a_1 t_1 (T - t_2 N)^{-1} + a_2 N^{-1}. \quad (13.9)$$

Чтобы найти значение N , минимизирующее (13.9), сначала рассмотрим первую возможность, когда оба коэффициента a_1 и a_2 *положительны*. Определим первую и вторую производные:

$$\frac{d[\text{var}(\bar{d})]}{dN} = a_1 t_1 t_2 (T - t_2 N)^{-2} - a_2 N^{-2} \quad (13.10)$$

и
$$\frac{d^2[\text{var}(\bar{d})]}{dN^2} = 2a_1 t_1 t_2^2 (T - t_2 N)^{-3} + 2a_2 N^{-3}. \quad (13.11)$$

Следовательно, первая производная равна нулю, если N удовлетворяет равенству

$$N = T/\{t_2 \pm (a_1 a_2^{-1} t_1 t_2)^{1/2}\}. \quad (13.12)$$

Поскольку

$$T/\{t_2 - (a_1 a_2^{-1} t_1 t_2)^{1/2}\} > T/(t_2 + t_1), \quad (13.13)$$

одно решение в (13.12) не удовлетворяет ограничению (13.8). Поэтому остается решение

$$N = N^* = T/\{t_2 + (a_1 a_2^{-1} t_1 t_2)^{1/2}\}, \quad (13.14)$$

которое действительно обеспечивает минимум, поскольку подстановка $N = N^*$ во вторую производную (13.11) дает

$$\left[\frac{d^2[\text{var}(\bar{d})]}{dN^2} \right]_{N=N^*} > 0. \quad (13.15)$$

Подставив N^* в (13.7), получим соответствующее значение для M , т. е.

$$M = M^* = T / \{t_1 + (a_1^{-1} a_2 t_1 t_2)^{1/2}\}, \quad (13.16)$$

M^* и N^* — это оптимальные значения, если не нарушаются ограничения (13.8). Чтобы найти решение в случае таких отклонений, выведем вид функции $\text{var}(\bar{a})$, определенной в (13.9). Из формулы для второй производной в (13.11) следует, что

$$\frac{d^2 [\text{var}(\bar{a})]}{dN^2} \geq 0, \text{ если } 0 \leq N \leq T/t_2. \quad (13.17)$$

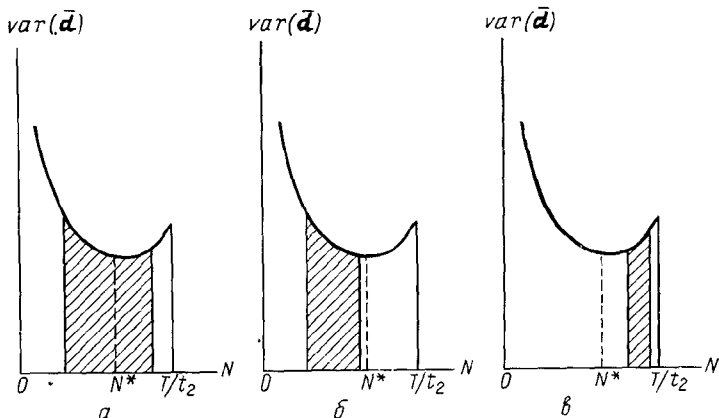


Рис. 19. Минимизированное значение N^* и (заштриховано) допустимая область для N

Из (13.14) вытекает, что N^* лежит внутри интервала $[0, T/t_2]$, а из (13.14) и (13.8) — что допустимая область — это часть интервала $[0, T/t_2]$. Следовательно, в этой области вторая производная не отрицательна, поскольку $\text{var}(\bar{a})$ монотонно убывает в интервале $0 < N < N^*$ и монотонно возрастает для $N^* < N < T/t_2$. Поэтому можно выделить три ситуации, представленные на рис. 19, где заштрихованная область соответствует допустимой области $[N_p, T/(t_1 + t_2)]$. Из рис. 19 N_0 и M_0 , т. е. оптимальное значение N и соответственно через (13.7) оптимальное значение M , следующие:

- а) если $N_p \leq N^* \leq T/(t_1 + t_2)$, то $N_0 = N^*$ и $M_0 = M^*$;
- б) если $N^* > T/(t_1 + t_2)$, то $N_0 = T/(t_1 + t_2)$ и $M_0 = T/(t_1 + t_2)$;
- в) если $N^* < N_p$, то $N_0 = N_p$ и $M_0 = (T - t_2 N_p)/t_1$.

Полностью аналогичное решение можно получить, если в (13.2) заменить $M \leq N$. Теперь рассмотрим вторую возможность, т. е. когда один коэффициент в (13.9) *положителен*, а другой — *отрицателен*. Пусть, например, a_1 отрицателен. Из (13.10) следует, что $\text{var}(\bar{a})$ уменьшается с ростом N . Следовательно, нужно взять N как можно большим. Табл. 10 и результаты приложения III.12 показывают, что *отрицательное* значение a_1 приводит к неравенству $M \geq N$ (если

$M \leq N$, α_1 всегда положительно). Условие $M \geq N$ совместно с машинным временем и условиями предварительных опытов, приведено в (13.8). Так, если мы захотим взять N самым большим и удовлетворим условию (13.8), то мы должны взять

$$N_0 = T/(t_1 + t_2) \quad (13.18)$$

и, следовательно,

$$M_0 = T/(t_1 + t_2). \quad (13.19)$$

Таким образом, можно показать, что если коэффициент a_2 отрицателен, то N нужно взять как можно меньшим. Отрицательное значение коэффициента a_2 ведет к неравенству $M \leq N$, и это, в свою очередь, снова приводит к (13.18) и (13.19).

У П Р А Ж Н Е Н И Я

1. Докажите, что в (3.9) $E[(\bar{y} - \mu)^2] = \sigma^2/N + (\text{смещение})^2$.
2. Докажите, что \bar{x}_{SAP} , определенное в (29), — несмещенная оценка.
3. Докажите, что $\text{var}(\sum_{i=1}^m y_i/m) = \text{var}(y)E(1/m)$.
4. Докажите, что оцениваемые коэффициенты $\vec{A} = (\vec{Z}' \vec{Z})^{-1} \vec{Z}' \vec{X}$ в (140) были бы несмещенными, если бы \vec{Z} была случайной.
5. Пусть мы имеем n независимых пар (x_i, y_i) ($i = 1, 2, \dots, n$). Обычная оценка отношения μ_x/μ_y — это $\hat{\theta} = \bar{x}/\bar{y} = \sum x_i/\sum y_i$. Дайте оценку методом «складного ножа», основанную на образовании двух групп (см. [Rao, 1969, ур. (4)]).
6. Докажите, что дисперсия функции $g(x)$ из (164) дается выражением

$$\text{var}[g(x)] = \lambda \exp(-\lambda v)/v - \lambda \xi - \xi^2.$$

7. Пусть $\xi(\lambda, v)$ в (163) оценивается методом значимой выборки с такой новой функцией плотности, как экспоненциальная $h_3(x)$ с параметром λ' . Докажите, что размах результирующей оценки $g_3^*(x)$ минимизируется при $\lambda' = 1/v$ (доказав, что $1/v \leq \lambda$). Докажите, что результирующий размах $g_3^*(x)$ больше, чем размах $g_1^*(x)$ и $g_2^*(x)$, но не больше, чем размах $g(x)$.

8. Функция $h_0(\vec{R}) = g(\vec{R})f(\vec{R})/E(x)$ в формуле (189) может быть плотностью распределения, если $g(\vec{R}) \geq 0$. Каким будет второе условие для этого? Докажите, что и это условие также выполняется.

9. Докажите, что при осуществлении выборки из симметричного распределения с математическим ожиданием η процедура $y^* = 2\eta - y$ идентична процедуре $y^* = F^{-1}(1 - r)$.

10. Докажите, что для симметричного распределения переменная y и дополняющая ее величина y^* имеют коэффициент корреляции, равный минус единице. Каков коэффициент корреляции между дополняющими переменными, которые генерируются из экспоненциального распределения с помощью процедуры $y^* = F^{-1}(1 - r)$? Каков коэффициент корреляции между переменной x и ее линейной функцией $y = a + bx$?

11. Докажите, что r_2 — случайное число, если оно определяется по формуле (226), т. е. $r_2 = c - r_1$ для $0 \leq r_1 \leq c$ и $r_2 = 1 + c - r_1$ для $c < r_1 < 1$, где r_1 — случайное число.

12. Докажите, что на рис. 14, б x и y станут скоррелированными, если положить $x = c$ (т. е. сделать x константой).

13. Рассмотрим моделирование системы с одним обслуживающим устройством для двух дисциплин обслуживания: 1) обслуживание в порядке поступления и 2) обслуживание в обратном порядке. Продемонстрируйте, как отдельные гене-

рафоры случайных чисел для моделирования времени обслуживания и времени между поступлениями требований не обеспечивают точную синхронизацию. Как можно было бы обеспечить точную синхронизацию в задачах такого типа (сравните также моделирование работы цеха)?

14. Смоделируйте систему с одним обслуживающим устройством, применяя каждый из шести способов уменьшения дисперсии, и оцените это уменьшение дисперсии.

15. Примените дополняющие величины, общие случайные числа и их комбинацию в простом случае моделирования, например в моделировании управления запасами.

П Р И М Е Ч А Н И Я

¹ Эта глава основана на работе [Kleijnen, 1971].

² МПД в задачах оценивания дисперсий обсуждаются в [Beja, 1969, p.7—8], [Fishman, 1972, p.182], [Kahn and Marshall, 1953, p.273], [Lewis, 1972, p.11] и [Matérn, 1962]. Общее обсуждение оценивания дисперсии см. в [Raj, 1968, p.189—199]. Методы понижения дисперсии в оценивании самой плотности вероятности рассматривались в [Fieller and Hartley, 1954] (сравните 111.4.4), в [Burt and Garman, 1971b], [Burt et al., 1970, p.442], [Garman, 1972] и [Ringer, 1965].

³ Исключение составляет работа [Lombaers, 1968, p.253], в которой изучалось уменьшение среднеквадратичной ошибки. Упомянем также концепцию «дефицита», или $1 - \sigma_0^2/\sigma_1^2$ [Andrews et al., 1972, p.121], где σ_0^2 и σ_1^2 обозначают дисперсии обычной и новой оценки соответственно.

⁴ Некоторые из МПД, применяемые в методе Монте-Карло, отличаются от методов выборочного обследования, например от стратификации или от контрольных величин.

⁵ Обычный метод Монте-Карло, применяемый к вероятностным сетям, представлен в [Burt and Garman, 1971b] и [Garman, 1972]. Использование математических ожиданий обсуждается в [Handscorn, 1968, p.4—5], [Kahn and Mann, 1957, p.15—17], [Maxwell, 1965, p.4] и [Moy, 1965, p.21]. Возможные меры работы введены в [Carter and Ignall, 1970, p.13; 1972]. Метод статистического оценивания представлен в [Kahn and Marshall, 1953, p.274—275]. Русская рулетка и разбросные упомянуты в [Kahn and Mann, 1957], [Moy, 1965, p.20] и [Tocher, 1963, p.109—110]. В [Relles, 1970] предложены МПД для выборки из распределения Стюдента. В [Andrews et al., 1972, p.55—61, 349—368] обсуждены МПД в методе Монте-Карло и эксперименты с «загрязненными» распределениями Гаусса. В [Jessor, 1965, p.161]. [Bakes et al., 1970, p.58—64] и [Van Slyke, 1963, p.847] показано, что можно сделать попытку решить часть всей задачи аналитическими методами. Поэтому на статистические методы падает уже оставшаяся часть задачи (см. также методы, использующие математическое ожидание). Многократная выборка была предложена в [Evans, 1963]. Метод, использующий цепи Маркова, обсуждается в [Hastings, 1970]. В [Easterfield, 1961] предложен сокращенный метод для систем массового обслуживания с требованиями, прибывающими партиями. Метод Костена для независимых поступлений требований и времени обслуживания в системах массового обслуживания представлен в [De Boer, 1969] и [Kosten, 1948]. Много МПД и ссылки дано в [Dutton and Starbuck, 1971, p.592], [Halton, 1970, p.7—13], [Hammersley and Handscorn, 1964], [Kahn, 1955], [Moy, 1965, p.12—23], [Nelson, 1966, p.17], [Wurl, 1971, p.60—63]. В [Raj, 1968] обсуждаются выборочные методы, отличающиеся от метода Монте-Карло (см. также [Matérn, 1960]). В [Bauknecht and Neß, 1971, p.35—153] улучшена эффективность методов с помощью развитого языка моделирования (BERSIM), который в системах массового обслуживания (особенно транспортных системах) служит специфическим путем уменьшения необходимого числа случайных чисел и машинного времени.

⁶ Символ \in обозначает, что u принадлежит слою S_k , т. е. значение u попадает внутрь слоя S_k .

⁷ Выражение после вертикальной черты означает условие. Так, формула (9) представляет дисперсию, т. е. математическое ожидание величины $(x - \mu_k)^2$ при условии, что x принадлежит k -му слою. Вспомним, что x_i попадает в S_k , если y_i принадлежит S_k .

⁸ Мы воспользовались уравнением (5.15) из [Cochran, 1966, p. 94], но в предположении бесконечной совокупности (см. также [Мой, 1965, p. 81]). Допущение о бесконечной совокупности вполне реалистично для моделирования и метода Монте-Карло, где можно иметь бесконечно много наблюдений, прибегая к различным случайным числам для каждого наблюдения.

⁹ Сравнение обозначения в (6) и в (18) показывает преимущества разделения между случайными и неслучайными величинами (случайные величины набраны жирным шрифтом). Это соглашение не используется ни Кокреном, ни Моем. Мы увидим, что Моем вообще не замечен стохастический характер некоторых переменных.

¹⁰ Мы отметили, что в [Johnston, 1963, p. 276] также определена СКО, но автор указанной работы смешал среднее значение выборочной статистики по N повторным выборкам и математическое ожидание, т. е. среднее по бесконечно большому числу повторных выборок. Наше определение согласуется с определением в [Togo-Vizcargondo and Wallace, 1968, p. 560].

¹¹ После некоторых обычных алгебраических преобразований получим

$$\begin{aligned} \text{var}(\bar{w}_k) = & \text{var}(\bar{w}_g) P(n_k \leq 1) + \text{var}(\bar{x}_{hh}) E(n_k | n_k \geq 2) + \\ & + \{E(\bar{w}_g) - E(\bar{w}_k)\}^2 P(n_k \leq 1) + \{E(\bar{x}_{hh}) - E(\bar{w}_k)\}^2 P(n_k \geq 2). \end{aligned}$$

¹² Мы возвратимся к подходу такого рода в III.4.4 (особенно при обсуждении формулы (152)).

¹³ Метод получения квазислучайных чисел, коротко описанный в [Handscornb, 1968, p. 10], несколько напоминает процедуру Моея. Метод Моея относится к варианту стратификации для оценивания интеграла. Указанный вариант, как мы упоминали в III.2.3, описывается в [Hammersley and Handscornb, 1964, p. 55]. Недавно этот метод применялся в моделировании в [Burt et al., 1970] (см. также [Andreasson, 1972b, p. 6], [Burt and Garman, 1971a, p. 254—256], [Fishman, 1972], [Garman, 1971, p. 11], [Shedler and Yang, 1971, p. 121]). Их метод подобен методу дополняющих величин (который обсуждался в III.6), поскольку генерируются k парных опытов (вместо двух) и полагается, что эти опыты отрицательно коррелированы. Мы, однако, думаем, что этот метод требует больше дополнительного программирования и времени, чем метод дополняющих величин. Для более подробного знакомства с деталями отсылаем читателя к [Burt et al., 1970].

¹⁴ $[a, b)$ обозначает интервал от a до b , включающий a и не включающий b .

¹⁵ Кроме блока 1 на рис. 10, где мы используем знак \leq , а Мой — знак $>$, что, вероятно, является опечаткой.

¹⁶ Наше первое замечание основано на предыдущих публикациях (см. [Kleijpen, 1969, p. 294—295]).

¹⁷ Переменные u и т. д. в этом параграфе определяются иначе по сравнению с III.2. Мы употребляем символы M , q_i и m в тех случаях, где у Бреннера соответственно m , p и n .

¹⁸ Как мы видели ранее, в селективной выборке каждый опыт удовлетворяет условию $v_i = v_i^*$, хотя опыты все-таки отличаются друг от друга, поскольку отличается порядок событий.

¹⁹ Мы видели, что две оценки должны иметь коэффициент корреляции, превосходящий по модулю $1/2$ при том условии, что обе оценки имеют одинаковые абсолютные значения коэффициентов вариации (этот коэффициент определяется как среднеквадратичная ошибка, деленная на математическое ожидание). Этот вывод здесь не приводится, поскольку данный результат, видимо, должен иметь очень ограниченное применение.

²⁰ В [Мой, 1965, p. 47] применена другая обобщенная процедура расщепления. Видимо, эта процедура менее эффективна, поскольку в ней оценивается коэффициент только по n/J парам, вместо $(n - n/J)$ пар. Поэтому полученный Моем вывод о том, что расщепление на две группы наиболее эффективно, вызывает сомнение.

²¹ Оценка согласно (123) в общем случае не будет несмещенной. В общем случае $E(x_1/x_2) \neq E(x_1)/E(x_2)$ или, следуя (123), $E(\hat{a}_{k0}) \neq E(xz_k)/E(z^2) = a_{k0}$.

²² В [Johnston, 1963, p. 107] допущено, что зависимые переменные — это фиксированные числа, тогда как в (124) это случайные величины z_h . Метод наименьших квадратов совсем не зависит от случайного характера z_h .

²³ Результат Бейна можно легко проверить с помощью вывода, данного в [Scheffé, 1964, p.8]. Если $\vec{W} = \vec{A}\vec{V}$, то $\vec{\Omega}_W = \vec{A}\vec{\Omega}_V\vec{A}'$, где $\vec{\Omega}_V$ обозначает ковариационную матрицу переменных v , а $\vec{\Omega}_W$ — ковариационную матрицу переменных w .

²⁴ Другой вывод о характере оптимально оцениваемых коэффициентов по методу наименьших квадратов можно найти в [Моу, 1965, p. 31—33]. Мой не использовал условие $E(z_h) = 0$, но вместо этого потребовал, чтобы

$$\bar{x} = \sum_{i=1}^n \sum_{k=1}^K a_{hi} z_{hi} / n.$$

²⁵ См. [Burt et al., 1970], [Burt and Garman, 1971a, p. 256—259], [Ehrenfeld and Ben-Tuvia, 1962, p.268—270], [Garman, 1971, p.12—14], [Gaver, 1969], [Gaver and Shedler, 1971], [Haitsma and Oosterhoff, 1964, p. 31—34], [Handscorn, 1968, p.8], [Hillier and Lieberman, 1968, p.460—461], [Kahn and Marshall, 1953, p.269], [Molenaar, 1968, p.121—122], [Newman and Odell, 1971, p. 62—64].

Отметим, что представление в [Ehrenfeld and Ben-Tuvia, 1962] оценки контрольных величин недостаточно корректно, поскольку \bar{W}_1 в формуле (3.2) из указанной работы должна быть постоянной, следуя выводу этих авторов о дисперсии оценки контрольной величины. На самом деле \bar{W}_1 в их примере оценивается.

²⁶ Большинство авторов априори полагают, что множитель a в оценке контрольной величины (152) равен единице, так что оценка (152) сводится к $\mu_x = \mu_z(x - z)$.

²⁷ Условие, которое должно быть наложено на эту оценку контрольной величины, чтобы достигнуть эффективного выигрыша, приведено в [Maxwell, 1965, p.8]. Заметим, что знак минус нужно поместить перед правой частью этого условия.

²⁸ Этот раздел основан на ранее опубликованном исследовании на голландском языке [Kleijneп, 1968в].

²⁹ Константа c положительна, поскольку это параметр экспоненциального распределения.

³⁰ Фактически эта процедура применялась в 2, где исходное экспоненциальное распределение (т. е. гамма-распределение с параметрами $p = 1$ и $b = \lambda$) заменено гамма-распределением с $p = 2$ и $b = 2/\nu$.

³¹ В [Kleijneп, 1968в, p. 197] доказано, что, даже когда $\xi(\lambda, \nu)$ неизвестно, можно показать, что дисперсия функции $g^*(x)$ меньше, чем дисперсия $g(x)$. Но, не зная ничего о $\xi(\lambda, \nu)$, мы не можем доказать, что $g_2^*(x)$ и $g_3^*(x)$ имеют меньшую дисперсию, чем $g(x)$ или $g_1^*(x)$. Поэтому невозможно выбрать такую аппроксимацию, которая имеет наименьшую дисперсию.

³² В [Kleijneп, 1968b, p.198, 3-я строка] опечатка. Чтобы ее устранить, надо умножить на $1/2 \text{ var}[g_2^*(x)]$.

³³ В [Kleijneп, 1968b, p.189] оцениваются дисперсии. Эти оценки очень хорошо согласуются с точными значениями, приведенными в табл. 1.

³⁴ Если $\lambda = 0,5$ и $\nu = 2$, то $h_3^*(x)$ идентична старой плотности распределения $f(x)$, поскольку $g_3^*(x) = g(x)$.

³⁵ Термины *псевдослучайный* и *квазислучайный* уже употреблялись для других случайных величин.

³⁶ Мой не рассматривает случайную величину m при выводе значимой выборки. Тем не менее один из примеров, где он использует значимую выборку, связан со случайным числом случайных чисел.

³⁷ Хотя Мой не определяет $v_t(r_t)$ точно, как это делаем мы в (204), легко получить, что α_t должно быть больше чем единица, для того чтобы сделать $v_t(r_t)$ возрастающей функцией от r_t . Если $\alpha_t > 1$, то можно проверить, что $v_t(r_t)$ не-

отрицательно для любого r_t в интервале $[-\infty, \infty]$. Наконец, положив $v_t(r_t)$ равным нулю для r_t вне интервала $[0, 1]$, получим интеграл

$$\int_{-\infty}^{\infty} v_t(r_t) dr_t,$$

равный единице. Это свидетельствует о том, что функция $v_t(r_t)$ есть функция плотности распределения.

³⁸ Сравните с формулой для дифференцирования под знаком интеграла, приведенной в [Keeping, 1962, p. 392—393]. Если

$$I(\alpha) = \int_a^b f(x, \alpha) dx,$$

то

$$\frac{dI(\alpha)}{d\alpha} = \int_a^b \frac{\partial f}{\partial \alpha} dx.$$

³⁹ Более точно, это монотонная неубывающая функция. Тем не менее если взять r_1 и r_2 достаточно далекими, то соотношение (222) все еще будет выполняться.

⁴⁰ Например, в приложении I.2 переменная y , имеющая нормальное распределение, генерировалась с помощью суммы $\sum r_j$. Поэтому увеличение любого r_j приводит к увеличению y .

⁴¹ Из-за условия монотонности соотношения мы берем $(1 - r)$ вместо (226). В [Andreasson, 1971a, p. 5] доказано, что использование r и $(1 - r)$ приводит к экстремуму коэффициента корреляции (отрицательного) между входной переменной y и дополняющей ее величиной y_a . Отметим, что эта теорема не относится к случаю, когда корреляция между выходными переменными x и x_a минимизируется с помощью r и $1 - r$ (ср. [Andreasson, 1971a, p.25]). В [Andreasson, 1972a и 1972b] приведен интересный анализ обобщенных схем, подобных $r_2 = a \pm r \pmod{1}$ (и перестановка случайных чисел, см. примечание 46). Там применялась оптимальная комбинация нескольких дополняющих методов, принадлежащих к общей схеме. Однако такая оптимальная комбинация (т. е. комбинация, обеспечивающая минимальную дисперсию) может оказаться лишь близкой к оптимальной из-за дополнительного программирования, увеличения времени опыта и усложнения анализа (все опыты внутри комбинации зависимы).

⁴² Для простоты мы берем m как константу, а не случайную величину. Это упрощение не влияет на аргумент в 2.

⁴³ Эта процедура генерирования дополняющих случайных чисел была экспериментально опробована Х. Тильборгхсом (H. Tilborghs, Katholieke Hogeschool Tilburg).

⁴⁴ Этот метод был предложен Х. Ломбаерсом (Technische Hogeschool Delft).

⁴⁵ Заметим, что язык моделирования GPSS позволяет генерировать как дискретные, так и непрерывные случайные входные переменные с помощью таблиц вместо функций (см. [Naylor et al., 1967a, p. 250]).

⁴⁶ Кроме того, переключение \vec{R}_1 и \vec{R}_2 , результирующих входных переменных, регулируется множителем α (см. [Моу, 1965, p.71]).

⁴⁷ В [Моу, 1965, p.76] установлено, что числа в таблице — это отношение дисперсий σ_1^2/σ_0^2 . А поскольку эти числа сравниваются там непосредственно по уменьшению дисперсии $\{1 - (\sigma_1^2/\sigma_0^2)\}$ в сравнении с другими МПД, мы предполагаем, что эта таблица действительно дает уменьшение дисперсии.

⁴⁸ Читателя, который недостаточно хорошо представляет себе мощность критерия, мы отсылаем к [Keeping, 1962, p.133]. В главе V (часть Б) мы упомянем пример, где доверительные интервалы (совместные), основанные на F -статистике, на самом деле уже, чем при использовании дополняющих величин, поскольку результирующая корреляция компенсируется потерей степеней свободы.

⁴⁹ III.8 основан на [Kleijnen, 1968a].

⁵⁰ В [Kleijnen, 1968a, p.15, 30—31] получен следующий результат. Если применить подход Точера, использующий факторные планы 2^K , и допустить, что этот подход вполне приемлем, то $\text{var}(\bar{d})$ будет минимальной при методе А (только с дополняющими величинами). Однако в нашем обсуждении табл. 2 из III.6 мы пришли к выводу, что подход Точера на самом деле неприемлем.

⁵¹ Два варианта, которые мы смоделировали, — это планы I и II в таблице I из [Naylor et al., 1967b].

⁵² Мы благодарны Х. Тильборгхсу (H. Tilborghs, Katholieke Hogeschool Tilburg), который запрограммировал системы с одним обслуживающим устройством, и профессорам Т. Воннакотту (T. Wonnacott, Western Ontario University) и Д. Грэхему (D. Graham, Duke University), которые помогли при моделировании систем с четырьмя станциями.

⁵³ В [Andreasson, 1970, p. 19] смоделирована многоканальная система и найден порядок С, В, А, D. Точер (частное сообщение) сказал, что метод С дает результаты хуже, чем метод В.

⁵⁴ Для метода В более точную формулу для $\text{var}(\bar{d})$ можно получить, добавляя в табл. 10 выражение $\{M^{-2}(M-N)c_1\}$, если $M > N$, или $\{N^{-2}(N-M)c_2\}$, если $M < N$. Эти выражения, однако, не упрощают отыскание оптимальных значений M и N . Для этого в методе В надо решить уравнение

$$\lambda a_1 M^2 + (2c_1 \sigma_2) (-c_1 + \lambda a_2 M^2)^{-1/2} - (\sigma_1^2 + c_1 - 3c_3) = 0,$$

где λ — множитель Лагранжа. Выражение для $\text{var}(\bar{d})$ в табл. 10 на самом деле представляет в чистом виде применение совместных случайных чисел вместо смеси, где используются дополняющие опыты $(M-N)$ или $(N-M)$. Как видно из III.8.7, такое применение в чистом виде может оказаться более предпочтительным, поскольку при этом упрощается анализ.

⁵⁵ Мы допустили, что будет *одинаковое* число (V_p) наблюдений для каждой системы на предварительной стадии. Если взять разное число их, то для некоторых наблюдений в одной системе не найдется парных наблюдений в другой. В [Boas, 1967, p. 291] дана следующая формула для оценки ковариаций, использующая все наблюдения — как парные, так и непарные:

$$\text{cov}(\bar{x}, \bar{y}) = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) (n_1 n_2) / n (n_1 n_2 - n_1 - n_2 + n),$$

где (x_i, y_i) — n парных наблюдений, \bar{x} — среднее по n_1 ($\geq n$) наблюдениям x , \bar{y} — среднее по n_2 ($\geq n$) наблюдениям y . Все x и y взаимно независимы.

⁵⁶ В системах со случайным числом случайных чисел в опыте величина T_p тоже будет случайной. В системах с одним обслуживающим устройством, где мы решили моделировать конечное число требований, T_p — неслучайная величина.

⁵⁷ Если $x=g(y)$, то в общем случае $E[x] \neq g(E[y])$, кроме случая, когда $g(y)$ — линейная функция.

⁵⁸ В [Ghurye and Robbins, 1954] берется оценка разности двух средних, когда выборочные наблюдения независимы. В формуле для оптимального объема выборки авторы указанной работы заменили неизвестные дисперсии оценками, основанными на предварительной выборке. Они показали, что в результирующей двухступенчатой процедуре дисперсия \bar{d} асимптотически стремится к дисперсии этой же величины при известных σ_1^2 и σ_2^2 .

⁵⁹ Эти результаты были получены Х. Тильборгхсом.

⁶⁰ Ради удобства предполагаем, что M равны. Далее через M будем обозначать оптимальное число опытов, полученное из табл. 11 и формулы (277) при применении формулы (280). Отметим, что в формуле (292) мы хотим оценить дисперсию среднего \bar{x} , в то время как в формуле (288) с помощью (290) мы оценивали дисперсию единичного наблюдения x .

⁶¹ В формуле (295) мы сначала применили центральную предельную теорему, чтобы усреднить пары дополняющих величин. Это объясняет причину асимптотической нормальности \bar{x}_1 . Из центральной предельной теоремы также следует, что и \bar{x}_2 асимптотически нормально. Отсюда и \bar{x} асимптотически нормально.

⁶² Отметим, что x и y независимы даже при применении метода A после предварительной стадии, тогда как на этой стадии x_i и y_i ($i = 1, \dots, N_p$) зависимы. Отметим далее, что нормальные условные распределения не обязательно приводят к двумерному нормальному распределению (см. [Ruymgaart, 1973]).

⁶³ В [Fishman, 1967, p. 6] установлено, что обратная задача минимизации V для данного T имеет несколько решений. Сравнение (305) и (306) с (275) и (276) показывает, что оба решения выглядят одинаково, но тем не менее они не идентичны.

⁶⁴ Другой подход (не в плане моделирования) рассмотрен в [Sedransk, 1971]. Мы имеем это доказательство благодаря профессору М. Эйве (M. Euwe, Katholieke Hogeschool Tilburg).

БИБЛИОГРАФИЯ

1. Andreasson I. J. (1970). The Application of Two Variance-Reducing Techniques on a Queueing Simulation in Simula—67, Report NA 70.21, Department of Information Processing, The Royal Institute of Technology, Stockholm.
2. Andreasson I. J. (1971a). On the Generation of Negatively Correlated Random Numbers, Report NA 71.32, Department of Information Processing The Royal Institute of Technology, Stockholm.
3. Andreasson I. J. (1971b). Antithetic and Control Variate Methods for the Estimation of Probabilities in Simulations, Report NA 71.41, Department of Information Processing, The Royal Institute of Technology, Stockholm.
4. Andreasson I. J. (1972a). Combinations of Antithetic Methods in Simulation. In: Working Papers, vol. 1, Symposium Computer simulation versus analytical solutions for business and economic models, Graduate School of Business Administration, Gothenburg, Sweden.
5. Andreasson I. J. (1972b). Antithetic Methods in Queuing Simulations, Report NA 72.58, Department of Computer Science, Royal Institute of Technology, Stockholm.
6. Andrews D. F., Bickel P. J., Hampel F. R., Huber P. J., Rogers W. H. and Tukey J. W. (1972). Robust Estimation of Location. Princeton University Press, Princeton, N. J.
7. Arvesen J. N. and Salsburg D. S. (1975). Approximate tests and confidence intervals using the jackknife. In: Perspectives in Biometrics (R. Elashoff, ed.), Academic, New York.
8. Bakes M. D., Bramson M. J., Freckleton S., Roberts P. C. and Ryan D. (1970). Stochastic—network reduction and sensitivity techniques in a cost effectiveness study of a military communications system. Operational Res. Quart., 21, Spec. Conf. Iss., 45—67.
9. Bauknecht K. and Nef W. (ed.) (1971). Digitale Simulation (Digital Simulation). Springer, Berlin.
10. Beja A. (ca. 1969). Multiple Control Variates in Monte Carlo Simulation (with Application to Queueing Systems with Priorities). W. P. No. 80/70, Gordon and Breach, New York.
11. Boas J. (1967). A note on the estimation of the covariance between two random variables using extra information on the separate variables. Stat. Neerl., 21, 291—292.
12. Brenner M. E. (1963). Selective sampling — a technique for reducing sample size in simulation of decision making problems. J. Ind. Eng., 14, 291—296. (Also published in: Computer Simulation of Human Behavior, by J. M. Dutton and W. H. Starbuck, Wiley, New York.)
13. Burt J. M. and Garman M. B. (1971a). Monte Carlo techniques for stochastic PERT network analysis. Infor, 9, 248—262.
14. Burt J. M. and Garman M. B. (1971b). Conditional Monte Carlo: a simulation technique for stochastic network analysis. Management Sci., 18, 207—217.
15. Burt J. M., Gaver D. P. and Perlas M. (1970). Simple stochastic networks: some problems and procedures. Naval Res. Logistics Quart., 17, 439—459.

16. Carter G. and Ignall E. (1970). A Simulation Model of Fire Department Operations: Design and Preliminary Results, R-632-NYC, The New York City Rand Institute, New York. (Also published in: IEEE Trans. System Sci. Cybernetics, Oct. 1970.)
17. Carter G. and Ignall E. (1972). Virtual Measures for Computer Simulation Experiments, P-4817. The Rand Corp., Santa Monica, Calif.
- 17a. Clark C. E. (1969). The Utility of Statistics of Random Numbers. System Development Corp., Santa Monica, Calif.
18. Clark C. E. (1961). Importance sampling in Monte Carlo analyses. Operations Res., 9, 603—620. (Also published in: Computer Simulation of Human Behavior, by J. M. Dutton and W. H. Starbuck, Wiley, New York.)
19. Cochran W. G. (1966). Sampling Techniques. 2nd ed., Wiley, New York. Русский перевод: Кокрен У. Методы выборочного исследования. М., «Статистика», 1976.
20. Conway R. W. (1963). Some tactical problems in digital simulation. Management Sci., 10, 47—61.
21. Conway R. W., Johnson B. M. and Maxwell W. L. (1959). Some problems of digital systems simulation. Management Sci., 6, 92—110.
22. De Boer J. (1969). Toepassing van Kosten's lotingsas in simulaties (Applying Kosten's sampling axis to simulation). Stat. Neerl., 23, 243—248.
23. Dutton J. M. and Starbuck W. H. (1971). Computer Simulation of Human Behavior. Wiley, New York.
24. Easterfield T. E. (1961). A short cut in a class of simulation problems. Operational Res. Quart., 12, 221—225.
25. Ehrenfeld S. and Ben-Tuvia S. (1962). The efficiency of statistical simulation procedures. Technometrics, 4, 257—275.
26. Emshoff J. R. and Sisson R. L. (1971). Design and Use of Computer Simulation Models, Macmillan, New York.
27. Evans D. H. (1963). Applied multiplex sampling. Technometrics, 5, 341—359.
28. Farmer J. A. (1966). Statistical Aspects of Digital Simulation. The Thomson Organisation, paper read at the symposium on "Simulation techniques and languages", Brunel College, Department of Mathematics, London.
29. Fieller E. C. and Hartley H. O. (1954). Sampling with control variables. Biometrika, 41, 494—501.
30. Fishman G. S. (1967). Digital Computer Simulation: The Allocation of Computer Time in Comparing Simulation Experiments, RM-5288-1-PR. The Rand Corp., Santa Monica, Calif. (Also published in Operations Res., 16, 1968, 280—295. Erratum in Operations Res., 16, 1968, 1087.)
31. Fishman G. S. (1972). Variance reduction in simulation studies. Stat. Computation Simulation, 1, 173—182.
32. Fishman G. S. and Kiviat P. J. (1967). Digital Computer Simulation: Statistical Considerations, RM-5387-PR. The Rand Corp., Santa Monica, Calif.
33. Fisz M. (1967). Probability Theory and Mathematical Statistics. 3rd ed., Wiley, New York.
34. Fraser D. A. S. (1957). Nonparametric Methods in Statistics. Wiley, New York.
35. Garman M. B. (1971). Variance Reduction in Large-Scale Computer Simulations. Working paper No. CP-338, Center for Research in Management Science, University of California, Berkeley, Calif.
36. Garman M. B. (1972). More on conditioned sampling in the simulation of stochastic networks. Management Sci., 19, 1972, 90—95.
37. Garman M. B. (1973). On the "Variate-Matching" Problem of Digital Simulation. Working paper no. 350, Center for Research in Management Science, University of California, Berkeley, Calif.
38. Gaver D. P. (1969). Statistical Methods for Improving Simulation Efficiency. Carnegie-Mellon University, Pittsburgh, Pennsylvania, 1969. (Also in: Proc. 3rd Ann. Conf. Applications Simulation, Los Angeles, Dec. 1969.)
39. Gaver D. P. and Shedler G. S. (1971). Control variable methods in the simulation of a model of a multiprogrammed computer system. Naval Res. Logistics Quart., 18, 435—450.

40. Ghurye S. G. and Robbins H. (1954). Two-stage procedures for estimating the difference between means. *Biometrika*, 41, parts 1 and 2, 146—152.
41. Goldberger A. S. (1964). *Econometric Theory*. Wiley, New York.
- 41a. Goodman A. S., Lewis P. A. W. and Robbins H. E. (1973). Simultaneous estimation of large numbers of extreme quantiles in simulation experiments. *Commun. Stat.*
42. Gray H. L. and Schucany W. R. (1972). *The Generalized Jackknife Statistic*. Marsel Dekker, New York.
43. Gürtler H. (1969). *Kwantitative Modelle zur Optimierung des Schalterverkehrs in einem Postamt (Quantitative models for optimizing traffic at counters in a post office)*. Doctoral dissertation, Wilhelms-Universität, Münster (Germany).
44. Haitsma A. H. and Oosterhoff J. (1964). *Monte Carlo Methoden (Monte Carlo methods)*. Rapport S 265 (C13), Leergang Besliskunde, Hoofdstuk XVI, Stichting Mathematisch Centrum, Amsterdam.
45. Hailton J. H. (1970). A retrospective and prospective survey of the Monte Carlo method. *Siam Rev.*, 12, 1—63.
46. Hammersley J. M. and Handscomb D. C. (1964). *Monte Carlo Methods*. Wiley, New York; Methuen, London.
47. Hammersley J. M. and Mauldon J. (1956). General principles of antithetic variates. *Proc. Cambridge Phil. Soc.*, 52, 476—481.
48. Hammersley J. M. and Morton K. W. (1956). A new Monte Carlo technique: antithetic variates. *Proc. Cambridge Phil. Soc.*, 52, 449—475.
49. Handscomb D. C. (1968). *Variance Reduction Techniques*. Presented at the Symposium on the design of computer simulation experiments, Duke University, Durham, N. C., 14—16 Oct. 1968. (Also published in: *The Design of Computer Simulation Experiments*, T. H. Naylor (ed.), Duke University Press, Durham, N. C., 1969).
50. Harling J. (1958). Simulations techniques in operational research. *Operational Res. Quart.*, 9, 9—21.
51. Hastings W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57, 97—109.
52. Henzler H. (1970). *Die Optimierung der Schichtenbildung beim Zufallsgesteuerten Geschichteten Stichprobenverfahren (The optimisation of strata formation in random stratified sampling)*. Doctoral dissertation, Ludwig-Maximilian University, Munich.
53. Hillier F. S. and Lieberman G. J. (1968). *Introduction to Operations Research*. Holden-Day, San Francisco, Calif., Chap. 14.
54. Ignall E. J. (1972). On experimental designs for computer simulation experiments. *Management Sci.*, 18, 384—388.
55. Jessop W. N. (1956). Monte Carlo methods and industrial problems. *Appl. Stat.*, 5, 158—165.
56. Johnston J. (1963). *Econometric Methods*. McGraw-Hill, New York.
57. Kahn H. and Mann I. (1957). *Monte Carlo*. Report P-1165, The Rand Corp., Santa Monica, Calif.
58. Kahn H. (1955). Use of Different Monte Carlo Sampling Techniques. Report P-766. The Rand Corp., Santa Monica, Calif.
59. Kahn H. and Marshall A. W. (1953). Methods of reducing sample size in Monte Carlo computations. *J. Operations Res., Soc. Amer.*, 1, 263—278.
60. Keeping E. S. (1962). *Introduction to Statistical Inference*. Van Nostrand, Princeton, N. J.
61. King G. W. (1953). The Monte Carlo method as a natural mode of expression in operations research. *J. Operations Res. Soc. Amer.*, 1, 46—51.
62. Kleijnen J. P. (1968a). Increasing the Reliability of Estimates in the Simulation of Systems: Negative and Positive Correlation Between Runs, Preliminary report, Los Angeles, April 1968(a). (Also presented at the European meeting on statistics, econometrics and management science of IMS, TIMS, ES and IASPS, Amsterdam, 2—7 Sept. 1968. Obtainable from the author at the Katholieke Hogeschool, Tilburg, Netherlands.)
63. Kleijnen J. P. C. (1968b). Een toepassing van "importance sampling" (An application of importance sampling). *Stat. Neerl.*, 22, 179—198. (Also pub-

- lished as report EIT, no. 2, of the Economisch Instituut Tilburg, Econometrische Afdeling, Tilburg, Netherlands.)
64. Kleijnen J. P. (1969). Monte Carlo techniques: a comment. In: *The Design of Computer Simulation Experiments* (T. A. Naylor, ed.), Duke University Press, Durham, N. C.
 65. Kleijnen J. P. C. (1971). *Variance Reduction Techniques in Simulation*. Doctoral dissertation, Katholieke Hogeschool, Tilburg (Netherlands).
 66. Kosten L. (1948). On the measurement of congestion quantities by means of fictitious traffic. *Het PTT Bedrijf*, 2, 15—25.
 67. Kosten L. (1968). *Statistische Aspecten van Simulatie* (Statistical aspects of simulation). Mimeographed notes, Afdeling Algemene Wetenschappen, Technische Hogeschool, Delft (Netherlands).
 68. Lewis P. A. W. (1972). Large-Scale Computer-Aided Statistical Mathematics, Naval Postgraduate School, Monterey, Calif. (In: *Proc. Computer Sci. Stat.*, 6th Ann. Symp. Interface, Western Periodical Co., Hollywood, Calif., 1973.)
 69. Lombaers H. J. M. (1968). Enige statistische aspecten van simulatie (Some statistical aspects of simulation). *Stat. Neerl.*, 22, 249—255.
 70. Marshall A. W. (1958). Experimentation by simulation and Monte Carlo. Report P-1174, The Rand Corp., Santa Monica, Calif.
 71. Matérn B. (1962). On the use of information on supplementary variables in estimating a distribution. *Rev. Int. Stat. Inst.* 30, 121—135.
 72. Matérn B. (1970). Spatial Variation. *Meddelanden från statens skogsforskningsinstitut*, band 49, nr. 5, 1960. (Reproduced by offset, Department of Forest Biometry, Royal College of Forestry, Stockholm, August 1970.)
 73. Maxwell W. L. (1965). *Variance Reduction Techniques*. The Rand Corp., Santa Monica, Calif.
 74. Mayne D. Q. (1966). A gradient method for determining optimal control of nonlinear stochastic systems. *Proc. 2nd IFAC Symp. Theory Self-Adaptive Control Systems*, P. H. Hammon (ed.), Plenum Press, New York.
 75. Meier R. C., Newell W. T. and Pazer H. L. (1969). *Simulation in Business and Economics*. Prentice-Hall, Englewood Cliffs, N. Y.
 76. Mickey M. R. (1959). Some finite population unbiased ratio and regression estimators. *J. Amer. Stat. Assoc.*, 54, 594—612.
 77. Mirham G. A. (1972). *Simulation: Statistical Foundations and Methodology*. Academic, New York.
 78. Mize J. H. (1973). *Multiple Sequence Random Number Generators*. School of Industrial Engineering and Management, Oklahoma State University, Stillwater, Oklahoma. Paper prepared for 1973 Winter Simulation Conference, San Francisco, Calif.
 79. Molenaar W. (1968). *Simulatie* (Simulation). In: *Minimaxmethode, Network Planning, Simulatie* (J. Kriens, F. Gobel and W. Molenaar, eds.). *Leerang Besliskunde*, 1.8, Mathematisch Centrum, Amsterdam.
 80. Mosteller F. and Tukey J. W. (1968). *Data analysis, including statistics*. In: *The Handbook of Social Psychology*. 2nd ed., vol. 2 (G. Lindzey and E. Aronson, eds.), Addison-Wesley, Reading, Pa.
 81. Moy W. A. (1965). *Sampling Techniques for Increasing the Efficiency of Simulations of Queuing Systems*. Ph. D. dissertation, Industrial engineering and management science, Northwestern University, August 1965. (Obtainable from University Microfilms, Inc., Ann Arbor, Michigan, order 66—2730).
 82. Moy W. A. (1966). *Practical Variance-Reducing Procedures for Monte Carlo Simulations*. University of Wisconsin, 7 Oct. 1966. (Also published in: *The Design of Computer Simulation Experiments*, T. H. Naylor (ed.), Duke University Press, Durham, N. C., 1969 and in: *Computer Simulation Experiments with Models of Economic Systems*, by T. H. Naylor, Wiley, New York, 1971.)
 83. Naylor T. H. (1971). *Computer Simulation Experiments with Models of Economic Systems*, Wiley, New York. Русский перевод: Нейлор Т. *Машинные имитационные эксперименты с моделями экономических систем*. М., «Мир», 1975.
 84. Naylor T. H., Balintfy J. L., Burdick D. S. and Chu K. (1967a). *Computer Simulation Techniques*, Wiley, New York.

85. Naylor T. H., Wertz K. and Wonnacott T. H. (1967b). Methods for analyzing data from computer simulation experiments. *Commun. ACM*, 10, 703—710.
86. Nelson R. T. (ca. 1966). *Systems, Models and Simulation*, Mimeographed notes, Graduate School of Business Administration, University of California, Los Angeles, Calif.
87. Newman T. G. and Odell P. L. (1971). *The Generation of Random Variables*. Griffin, London.
88. Orcutt G. H., Greenberger M., Korbel J. and Rivlin A. M. (1961). *Microanalysis of Socioeconomic Systems: A Simulation Study*. Harper and Row, New York.
89. Page E. S. (1965). On Monte Carlo methods in congestion problems: II. Simulation of queuing systems. *Operations Res.*, 13, 300—305.
90. Price C. M. (1972). The Design and Implementation of Generalised Simulation Models for the Mining Industry. In: Working papers, vol. 2, Symposium Computer simulation versus analytical solutions for business and economic models, Graduate School of Business Administration, Gothenburg (Sweden).
91. Pugh E. L. (1966). A gradient technique of adaptive Monte Carlo. *SIAM Rev.*, 8, 346—355.
- 91a. Quenouille M. (1956). Notes on bias-in estimation. *Biometrika*, 43, 353—360.
92. Radema F. W. (1969). *Versnellings technieken bij Systemsimulatie (Variance reduction techniques in systems simulation)*. Groep Bedrijfseconometrie, Technische Hogeschool, Eindhoven (Netherlands).
93. Raj D. (1968). *Sampling Theory*. McGraw-Hill, New York.
94. Rao J. N. K. (1969). Ratio and regression estimators. In: *New Developments in Survey Sampling* (N. L. Johnson and H. Smith, eds.), Wiley-Interscience, New York.
95. Relles D. A. (1970). Variance reduction techniques for Monte Carlo sampling from Student distribution. *Technometrics*, 12, 499—515.
100. Ringer L. J. (ca. 1965). Simulation of PERT Project Completion Times by Stratified Sampling Methods, Technical Report no. 3, Texas A and M Research Foundation, processed for Defense Documentation Center, Defense Supply Agency. (Obtainable from Clearinghouse for Federal Scientific and Technical Information.)
101. Ruymgaart F. H. (1973). Non-normal bivariate densities with normal marginals and linear regression functions. *Stat. Neerl.*, 27, 11—17.
102. Satterthwaite F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics*, 2, 110—114.
103. Scheffé H. (1964). *The Analysis of Variance*. Wiley, New York. Русский перевод: Шеффе Г. Дисперсионный анализ. М., Физматгиз, 1963.
104. Sedransk J. (1971). Precision specifications for the estimation of treatment differences. *Biometrics*, 27, 673—680.
105. Shedler G. S. and Yang S. C. (1971). Simulation of a model of paging system performance. *IBM Systems J.*, 10, 113—128.
106. Tocher K. D. (1963). *The Art of Simulation*. The English Universities Press, London.
107. Toro-Vizcarrondo C. and Wallace T. D. (1968). A test of the mean square error criterion for restrictions in linear regression. *J. Amer. Stat. Assoc.*, 63, 558—572.
108. Tukey J. W. (1958). Bias and confidence in not quite large samples (abstract). *Ann. Math. Stat.*, 29, 614.
109. Van Slyke R. M. (1963). Monte Carlo methods and the PERT problem. *Operations Res.*, 11, 839—860.
110. Williams W. H. (1964). Sample selection and the choice of estimator in two-way stratified populations. *J. Amer. Stat. Assoc.*, 59, 1054—1062.
111. Wurl H. (1971). Die Anwendung der Simulationstechnik zur Betriebswirtschaftlichen Beurteilung Industrieller Projekte in Entwicklungsländern (The application of the simulation technique for the economic evaluation of industrial projects in underdeveloped countries). Duncker and Humblot, Berlin.

ОТВЕТЫ К УПРАЖНЕНИЯМ

Глава I

1. После операции взятия модуля остается m целых между 0 и $(m - 1)$; в результате получаем не более m различных чисел. У Фибоначчи те же самые два числа должны быть воспроизведены.

2. Производится выборка x из нормального распределения с $E(x) = b$ и $\text{var}(x) = c^2$. Подставляется x в $g(x) = c\sqrt{2\pi/\ln x}$ для $x \geq a$ и $g(x) = 0$ для $x \leq a$.

3. Генерируются AT и ST; вычисляется $\text{DIF} = \text{WT} + \text{ST} - \text{AT}$; если $\text{DT} > 0$ затем $\text{WT} = \text{DIF}$, или $\text{IT} = -\text{DIF}$.

4. $P(x < a) = 1 - P(x \geq a) = 1 - P(x_1 \geq a) P(x_2 \geq a)$. Для $a = 83$, $p = 0,39$.

5. Передвигаем «SG2 = 1» от нижней связи 2 к верхней 2 и добавляем «SG2 = 1» в правую часть.

6. Быстрее со средними.

7. Из $N(0, \hat{\sigma}^2)$ с $\hat{\sigma}^2 = \sum_1^n (y_1 - \hat{y}_1)^2 / (n - 2)$. Не производится выборка $\hat{\alpha}$ и $\hat{\beta}$ (анализ чувствительности может быть полезным).

Глава II

1. $\text{var}[(\bar{x}_1 + \bar{x}_2)/2] = [\text{var}(\bar{x}_1) + \text{var}(\bar{x}_2) + 2 \text{cov}(\bar{x}_1, \bar{x}_2)]/4$. Последний член есть нуль в двух опытах и положительная величина в продленном опыте.

2. Дополняющие переменные.

3. Общие случайные числа меньше $\sigma(\bar{x}_1 - \bar{x}_2)$. Применим t -критерий для разности $d_i = x_{1i} - x_{2i}$ ($i = 1, 2, 3 \dots$).

4. $(1 - \alpha)^2$. Нет, потому что $\hat{\mu}_1 - \hat{\mu}_2$ и $\hat{\mu}_1 - \hat{\mu}_3$ зависимы.

5. $\alpha = \rho(\hat{\mu}, \bar{y})\sigma(\hat{\mu})/\sigma(\bar{y})$, где ρ обозначает корреляционный коэффициент.

6. Да, формула применяется только для оценки величин в конце периода планирования (сравните с [Naylor et al., 1967c]).

7-10. Сравните с текстом главы II.

Глава III

1. $E\{[(\bar{y} - \eta) + (\eta - \mu)]^2\}$ и т. д.

2. $E(\bar{v}_k) = \sum \Sigma E(\bar{v}_k | n_k = i, n = j)P(n_k = i, n = j)$. Доказывается, как в приложении III.2.

3. Доказывается, как в приложении III.4.

4. $E(\vec{A}) = (\vec{Z}'\vec{Z})^{-1}\vec{Z}'E(\vec{X}) = (\vec{Z}'\vec{Z})^{-1}\vec{Z}'[\vec{Z}\vec{A} + E(\vec{U})] = \vec{A}$.

5. $J_1 = 2\hat{\theta} - \sum_1^M x_j/\sum_1^M y_j$, $J_2 = 2\hat{\theta} - \sum_{M+1}^n x_g/\sum_{M+1}^n y_g$ ($M = n/2$),

$J = 2\hat{\theta} - (\sum x_j/\sum y_j + \sum x_g/\sum y_g)/2$.

6. Аналогично приложению III.9.

7. Аналогично приложению III.8.

8. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h_0(\vec{R})d\vec{R} = 1$.

9. Симметрично: $F(y) = 1 - F(2\eta - y)$. Поэтому равенство $F(y^*) = 1 - r$ дает те же результаты, что и равенство $1 - F(2\eta - y^*) = 1 - r$ или $y^* = -F^{-1}(r) + 2\eta$.

10. $E[(y - \eta)(2\eta - y - \eta)] = \eta^2 - E(y^2) = \sigma_y^2$; $\rho = \sigma_y^2 / \sigma_y \sigma_y = 1$.

Экспоненциальное распределение: $\rho = \int_0^1 \ln(r) \ln(1-r) dr - 1 = 1 - \pi^2/6 \approx -0,645$.

Корреляция между x и $y = a + bx$ равна единице, если $b > 0$, и -1 , если $b < 0$.

11. При $0 \leq a \leq c$, $P(r_2 \leq a) = P(c - a \leq r_1 \leq c) = a$. При $c \leq b \leq 1$, $P(r_2 \leq b) = b$.

12. $\text{Cov}(x, y) = E(cy) - E(c)E(y) = cE(y) - cE(y) = 0$.

13. Время обслуживания j можно использовать для j -го клиента в системе массового обслуживания с дисциплиной очереди «первый прибыл — первый обслужен» и для клиента j' при дисциплине очереди «последний прибыл — последний обслужен» ($j \neq j'$).

Время обслуживания и прибытия каждого клиента далее используется при дисциплине очереди «последний прибыл — последний обслужен».

14—15. См. главу III.

● УКАЗАТЕЛЬ ПЕРЕВОДА ТЕРМИНОВ

- Abstract model математическая модель; абстрактная модель
Abstract system абстрактная система
Accuracy правильность; точность
Adaptive system адаптивная система
Additive generator аддитивный генератор
Analogous model аналоговая модель
Analysis of simulation experiments анализ имитационных экспериментов
Analysis of variance, ANOVA дисперсионный анализ
Antithetic random numbers дополняющие случайные числа
Antithetic variates дополняющие переменные величины
Attribute признаки (обычно дискретные)
- Balance of payments deficit дефицит платежного баланса
Batch партия; группа
Bayesian approach байесовский подход
Binary computer двоичная ЭВМ
Binomially distributed variable биномиально распределенная переменная
Binary number двоичное число; бинарное число; число, записанное в двоичной системе счисления
Block size размер блока
Blocks of observations блоки наблюдений
Building block design блочная конструкция
Business game деловые игры
- Canonical analysis канонический анализ
Capital budgeting статьи движения капитала в бюджете
Capital investment simulation моделирование капиталовложений
Closed-loop feedback control system замкнутая система управления с обратной связью
Closed system закрытая система
Cochran's variance test критерий Кокрена для дисперсий
Collection and processing of real-world data сбор и обработка данных с реального объекта
Combinational problems комбинаторные задачи
Combination of the levels of the factors комбинация уровней факторов
Common random numbers общие случайные числа
Complement дополнение
Complete rankings полные ранжирования; полная ранжировка
Computer system сеть ЭВМ
Concomitant variable сопутствующая переменная
Conditional probability условная вероятность
Confidence interval доверительный интервал
Congestion скопление; перегруженность
«Contaminated» Gaussian distribution «загрязненное» распределение Гаусса
Continuous, increasing function непрерывная возрастающая функция
Control variates контрольные величины
Controllable variables управляемые переменные
Converse problem обратная задача
Correlated sampling коррелированная выборка
Cost curve кривая стоимости
Cost function функция стоимости
Critical path in network критический путь в сети
Cross-correlation перекрестная корреляция
Cross-section data данные по аналогии (сравнительные)
Cumulative distribution кумулятивное распределение; накопленное распределение

Customer требование; заказчик
 Cyclical behavior циклическое поведение
 Debugging отладка (программы)
 Decimal computer десятичная ЭВМ
 Decreasing function убывающая функция
 Deficiency недостаток; дефицит; дефект
 Definition characteristic определяющая характеристика
 Degree of orthogonality of the columns in the matrix of independent variables
 степень ортогональности столбцов матрицы независимых переменных
 Demand спрос; требование; запрос; потребность
 Density function плотность распределения; функция плотности
 Dependence of events зависимые события
 Design columns столбцы матрицы плана
 Deterministic quantities детерминированные величины
 Deterministic system детерминированная система
 Difference equations with side condition дифференциальные уравнения с граничными условиями
 Differential calculus дифференциальное исчисление
 Digital simulation цифровое моделирование
 Discrete measurement дискретное измерение
 Discrete system дискретная система
 Distribution sampling выборочное распределение
 Duel model модель дуэли
 Duopoly монополия двух конкурирующих компаний
 Dynamic behavior динамическое поведение
 Dynamic system динамическая система
 Effect of a factor эффект фактора
 Eigenvalue собственное значение
 Empirical distribution эмпирическое распределение
 Empirical weights эмпирические веса
 Empty state «пустое» состояние
 Endogenous variables эндогенные переменные; внутренние переменные; внутрисистемные переменные
 Entropy criterion энтропийный критерий
 Equilibrium state состояние равновесия
 Equipment failure model модель отказа оборудования
 Ergodic time series эргодический временной ряд
 Error rate доля ошибки
 Euler constant постоянная Эйлера
 Evolutionary model building эволюционное построение модели; адаптация модели
 Exogenous variables экзогенные переменные; внешние переменные
 Expected range ожидаемый размах; математическое ожидание размаха
 Expected value математическое ожидание; ожидаемое значение
 Experimental design экспериментальный план
 Exponential density function экспоненциальная функция; экспоненциальная плотность распределения
 External relationship внешние отношения
 Factor loadings факторные нагрузки
 Factorial experiments факторные эксперименты
 F-distribution *F*-распределение; распределение Финнера
 Feedback system система с обратной связью
 Feedforward system система с предварением
 Financial simulation моделирование операций
 First degree polynomial полином первой степени; полином первого порядка
 First-in-first-out (FIFO) queuing discipline обслуживание в порядке поступления (первым пришел — первым обслужен)
 First-order serial correlation serial correlation первого порядка
 First-order term член первого порядка

Fitting a regression equation подбор уравнения регрессии
Fixed sequence method метод фиксированной последовательности
Flowchart блок-схема программы
Foreign demand внешнеторговый спрос
Fortran subroutine Фортранная подпрограмма
Fractional 2^{k-p} design дробный план типа 2^{k-p}
Fromm's utility function функция полезности Фромма
F-test F-критерий
Full factorial design полный факторный план

Gaming имитационная игра
Gamma distribution гамма-распределение
Gaussian distribution распределение Гаусса; нормальное распределение
General factorial experiments полный факторный эксперимент
Generalized splitting обобщенное разбиение
Goodness-of-fit test критерий согласия
Gradient technique градиентный метод
Group-screening designs планы группового отсеивания
Heteroscedasticity гетероскедастичность; неоднородность дисперсии
Hybrid system гибридная система: смешанная система
Hypothesis testing проверка гипотез

Iconic model модель-изображение
Identified system идентифицированная система
Identify characteristic идентифицирующая система
Importance numbers значимые числа
Importance sampling значимая выборка
Incomplete design неполный план
Incomplete rankings неполные ранжирования
Independent variables независимые переменные; факторы
Indifference zone зона безразличия
Individual distribution индивидуальное распределение; частное распределение
Industrial dynamics промышленная динамика
Inequality coefficient коэффициент неравенства (Тейл)
Inequation неравенство
Information-theoretic measure теоретико-информационная мера
Input data analysis анализ входных данных
Integer целое число
Integral подынтегральное выражение; подынтегральная функция
Intercept свободный член
Interdisciplinary approach междисциплинарный подход
Intermediate variables промежуточные переменные
Internal relationship внутреннее отношение
Invariant distribution стационарное распределение; инвариантное распределение
Inverse cumulative distribution function обратная кумулятивная функция распределения
Inverse function обратная функция
Inversion technique метод обращения
Iterative итеративность; итеративный

Jackknife statistic статистика «складного ножа»
Jackknife technique метод «складного ножа»
Jackknifed regression estimator регрессионная оценка, основанная на методе «складного ножа»
Joint cumulative distribution совместное кумулятивное распределение
Joint density function совместная плотность распределения

Lag лаг; сдвиг
Large-scale simulation крупномасштабное моделирование
Least squares метод наименьших квадратов; МНК
Level уровень (фактора)

Linear combinations of exponential variates линейные комбинации экспоненциальных случайных переменных
 Linear contrasts линейные контрасты
 Link-trainer тренажер
 Location estimator оценка положения
 Logical value логическое значение
 Lognormal distribution логарифмически-нормальное распределение; логнормальное распределение
 Loss function функция потерь
 Low-order interaction взаимодействие низкого порядка
 Lumber industry лесоразработка
 Lump скачок
 Main diagonal главная диагональ
 Management scientist специалист по управлению
 Management system административная система; управленческая система
 Man-machine simulation человеко-машинное моделирование; диалоговое моделирование
 Marginal density functions безусловные плотности частных распределений
 Mathematical derivation математические отклонения
 Mathematical game theory математическая теория игр
 Mean square error (MSE) средняя квадратичная ошибка (СКО)
 Method of Lagrangian multipliers метод множителей Лагранжа
 Midsquare generator среднеквадратичный генератор
 Mixed multiplicative generator смешанный мультипликативный генератор
 Model building построение модели
 Model sampling модель выборки
 Monte Carlo experiments монтекарловские эксперименты; эксперименты по методу Монте-Карло (т. е. полностью рандомизированные)
 Moving average скользящее среднее
 Multichannel queuing system многоканальная система массового обслуживания
 Multiple comparison procedures методы множественных сравнений
 Multiple integral многомерный интеграл
 Multiple ranking procedures методы множественного ранжирования
 Multiple response многомерный отклик
 Multiplex sampling многократная выборка
 Multiplicative congruential method мультипликативно-конгруэнтный метод
 Multiprogramming computer машина с мультипрограммированием; многопрограммная машина
 Multistage approach (много)шаговый подход
 Multivariate normal distribution многомерное нормальное распределение
 Newton's method for approximating the roots of an equation метод Ньютона для приближенного отыскания корней уравнений
 Nonadaptive system неадаптивная система
 Noncontrolled system разомкнутая система; неуправляемая система
 Nonlinear discrete-time control system нелинейная система управления с дискретным временем
 Nonorthogonality неортогональность
 Nonstationary нестационарный
 Nonstationary system нестационарная система
 Nonterminating system система без останова; неограниченная система
 Numerical accuracy численная точность
 Numerical solution численное решение
 Numerical value числовое значение
 Oligopoly монополия многих конкурирующих компаний
 Open system открытая система
 Operational gaming операционная игра с риском; азартная операционная игра
 Optimal allocation оптимальное распределение
 Organization theory теория организации
 Output variables выходные переменные; отклики
 Overestimation переоценивание

Parallel service stations параллельные станции обслуживания
Parametric statistical techniques методы параметрической статистики
Pairwise comparisons попарные сравнения
Partial differentiation частное дифференцирование
Partial integration интегрирование по частям
Penalty потери
Percentile процентиль
Pilot-runs предварительные опыты
Planning of simulation experiment планирование имитационного эксперимента
Population совокупность; популяция
Population mean генеральное среднее; математическое ожидание
Positive serial correlation положительная сериальная корреляция
Posterior uses of data апостериорное использование данных
Power мощность (критерия)
Power residual method метод степенных вычетов
Power of t -test мощность t -критерия
Preference предпочтение; вкус
Priority приоритет
Prior knowledge априорные сведения
Probability distribution распределение вероятностей
Probability of erroneously accepting (nonrejecting) a false hypothesis вероятность ошибочного принятия (неотбрасывания) ложной гипотезы
Probability of erroneously rejecting a true hypothesis вероятность ошибочного отбрасывания правильной гипотезы
Probability success вероятность успеха
Product произведение
Production/inventory model модель «производство-запасы»
Profit прибыль; доход
Programming problems проблемы программирования
Proportionally stratified estimator пропорционально стратифицированная оценка
Proportionally stratified sampling пропорциональная выборка; стратифицированная выборка
Pseudorandom numbers псевдослучайные числа
Purely random numbers чисто случайные числа

Quantile квантиль

Queuing массовое обслуживание; очереди
Queuing system система массового обслуживания

Random designs рандомизированные планы; случайные планы

Random number случайные числа
Random number generation получение случайных чисел; генерирование случайных чисел
Random walks случайные блуждания
Range размах; диапазон значений
Ratio estimator оценка отношения
Real-life object настоящий объект; действительный объект
Real-world system реальная система
Recursive рекурсивный; рекуррентный
Regression equations уравнение регрессии
Regression sampling регрессионная выборка
Reliability надежность
The reliability of statements based on a sample надежность выводов, основанных на некоторой выборке
Remainder остаток
Replication повторение; повтор; дублирование
Response отклик
Response surface methodology концепция поверхности отклика
Reticulation method метод декомпозиции (сеточный метод)
Ridge гребень; стационарное возвышение
Robustness робастность; устойчивость (к нарушениям предпосылок)

Rule of thumb эмпирическое правило; эмпирический прием
 Runlength время счета; длина серии опытов; длительность эксперимента
 Saddlepoint седловая точка; седло
 Sample size объем выборки; размер выборки
 Sample survey выборочное обследование
 Sampling data выборочные данные
 Sampling experiment выборочный эксперимент
 Sampling of stochastic variables выборка случайных величин
 Sampling without replacement выборка без возвращения
 Scraping probability вероятность поломки
 Screening отсеивание факторов
 Screening design отсеивающий план; план отсеивания; план отсеивающего эксперимента
 Second degree polinom полином второй степени; полином второго порядка
 Second-order term член второго порядка
 Selective sampling селективная выборка
 Sensitive чувствительность
 Sensitivity analysis анализ чувствительности
 Sequential design последовательный план; последовательное планирование
 Serial correlation of higher order сериальная корреляция высших порядков
 Serial expansion разложение в ряд
 Service counter канал обслуживания
 Service stations in sequence последовательные станции обслуживания
 Set of simultaneous linear difference equations of n -th order система совместных линейных дифференциальных уравнений n -го порядка
 Shifted exponential distribution сдвинутое экспоненциальное распределение
 Shuffling тасование (например, карт)
 Significance значимость
 Significant digits значащие цифры
 Simulation имитационное моделирование; имитация; моделирование
 Simulation language язык моделирования
 Simulation program программа имитационного моделирования
 Simultaneous confidence intervals совместные доверительные интервалы
 Single population единичная совокупность (популяция)
 Single-server одноканальная система массового обслуживания
 Slope наклон
 Splitting разбиение; расщепление
 Stability устойчивость
 Start state начальное состояние
 Starting condition начальные условия
 Stationary r -dependent central limit theorem центральная предельная теорема для стационарных r -зависимых реализаций
 Statistical design of simulation experiments статистическое планирование имитационных экспериментов
 Statistical performance статистические характеристики
 Statistical reliability статистическая надежность
 Statistical reliability measure мера статистической надежности
 Status variables переменные состояния
 Steady-state установившееся состояние; установившийся режим
 Steepest ascent крутое восхождение
 Steepest descent method метод крутого спуска; метод наискорейшего спуска
 Step шаг; ступень
 Stepwise approximation ступенчатая аппроксимация
 Stochastic disturbance стохастические возмущения
 Stochastic network вероятностная сеть
 Stochastic process случайный процесс
 Stochastic system стохастическая система; вероятностная система, индетерминированная система
 Stochastic variable случайная величина
 Strata слой; страта
 Stratification стратификация; расслоение

Stratification variable стратифицирующая переменная
 Stratum weight вес слоя
 Subjective estimate субъективная оценка
 Subjective probability субъективная вероятность
 Sufficient conditions достаточные условия
 Sum of squared deviations сумма квадратичных отклонений
 Supersaturated designs сверхнасыщенные планы
 Surface of an irregularly shaped figure площадь нерегулярной поверхности
 Symbolic model символьная модель
 Systems approach системный подход
 Systems theory теория систем

T
 Taylor series ряд Тейлора
t-distribution *t*-распределение; распределение Стьюдента
t-test критерий Стьюдента, *t*-критерий
 Technical organizational system технико-организационная система
 Ten-sided die десятигранная игральная кость
 Terminating system система с остановом; система с конечным числом элементов; ограниченная система; периодическая система
 Theoretical weights теоретические веса
 Theory on the statistical design of experiments теория статистического планирования экспериментов; статистическая теория планирования экспериментов
 Time series distribution распределение временного ряда
 Timepath траектория (во времени)
 Tour тур; цикл
 Transformations of random numbers преобразование случайных чисел
 Transient state неустановившееся состояние; переходный режим
 Transition probabilities вероятности перехода
 Trend дрейф; тренд
 Turning point локальный экстремум упорядоченного ряда; точка поворота
 Turing's test критерий Тьюринга
 Two-stage least squares двухступенчатый метод наименьших квадратов; двухступенчатый МНК

U
 Unbiased estimator несмещенная оценка
 Uncontrollable variables управляемые переменные
 Uniformly distributed on the interval 0,1 равномерно распределены на интервале 0,1
 Utilization factor коэффициент использования; КПД
 Utilization коэффициент использования; загрузка (системы массового обслуживания)

V
 Validation обоснованность
 Variable переменная; фактор
 Variance reduction techniques методы понижения дисперсии
 Variate случайная переменная
 Variation coefficient коэффициент вариации
 Vector of covariances вектор ковариаций
 Virtual виртуальный; возможный; дающий тот же эффект в данных условиях
 Verification всрификация; проверка

W
 Waiting line очередь
 Wasted computer time холостое время счета
 Weighted sum of independent normally distributed variables взвешенная сумма независимых нормально распределенных переменных
 Wishart distribution распределение Уншарта

2^k factorial experiment факторный эксперимент типа 2^k
 2^{k-p} designs планы типа 2^{k-p}
 α-error α-ошибка; ошибка I рода; уровень значимости
 β-error β-ошибка; ошибка II рода, мощность критерия

● ИМЕННОЙ УКАЗАТЕЛЬ *

- Ackoff R. L. 51
 Ahrens J. H. 34
 Aigner D. J. 66
 Anderson T. W. 77, 78
 Andreasson I. J. 50, 87, 124, 142,
 145, 146, 147, 148, 150, 151, 155, 157,
 197, 199, 200
 Andrews D. F. 21, 31, 34, 196
 Anker C. J. 73, 74, 75
 Arvesen J. N. 179, 180
Bakes M. D. 78, 196
 Balderston F. E. 26
 Bauknecht K. 29, 196
 Beja A. 114, 115, 120, 122, 124, 196
 Bechhofer R. E. 13
 Benecke R. W. 25, 34
 Ben-Tuvia S. 99, 107, 108, 132, 142,
 198
 Biles W. E. 26
 Blomqvist N. 63, 81
 Blumenthal S. 13
 Boas J. 200
 Bonini C. P. 26
 Boot J. C. G. 17
 Bovet D. P. 25
 Bray T. A. 31
 Brenner M. E. 106, 107, 108, 155
 Bruzelius L. H. 75, 81
 Buijs B. G. F. 17
 Burt J. M. 25, 125, 142, 150, 157,
 196, 197, 198
 Carter E. E. 25
 Carter G. 59, 155, 196
 Chambers J. M. 31
 Chambers J. C. 61, 81
 Chambers R. P. 31, 34, 60, 81
 Chan M. M. W. 60
 Chiu J. S. Y. 22
 Churchman C. W. 16, 17, 27, 37
 Clark C. E. 99, 126, 127, 131, 132
 Clarkson G. P. E. 81
 Clough D. J. 24, 61
 Cochran W. G. 91, 92, 93, 94, 197
 Cohen K. J. 23, 26, 28, 61
 Conway R. W. 15, 27, 29, 63, 64, 75,
 78, 79, 155
 Coplin W. D. 26
 Coveyou R. R. 31
 Cox D. R. 70
 Crane M. A. 63, 64, 77
De Boer J. 196
 Dear R. E. 81
 Dewan P. B. 25
 Dieter U. 34
 Donaldson T. S. 21
 Driehuis W. 23, 26, 81
 Dutton J. M. 26, 30, 34, 68, 79, 81,
 196
 Easterfield T. E. 196
 Ehrenfeld S. 99, 107, 108, 132, 142,
 198
 Elton M. 25, 81
 Emery F. E. 15, 16, 51
 Emshoff J. R. 17, 23, 24, 26, 29, 34,
 51, 60, 61, 63, 65, 67, 73, 81, 147, 155,
 157
 Evans D. H. 132, 196
 Eykhoff P. 51
 Fagerstedt L. 81
 Fain J. B. 81
 Fain W. W. 81
 Farmer J. A. 126, 148, 155
 Fetter R. B. 64
 Fieller E. C. 126, 196
 Finger J. M. 60, 66, 67

* Именной указатель не включает фамилии, встречающиеся только в библиографии. —
 Прим. ред.

- Fishman G. S. 59, 60, 64, 66, 67,
 74, 77, 155, 157, 171, 176, 196, 197,
 201
 Fisz M. 19, 33, 37, 90, 91, 101, 107,
 129, 130, 175, 182
 Flagle C. D. 32, 34
 Forrester J. 16, 23, 28
 Fraser D. A. S. 92
- Gafarian A. V.** 73, 74, 75, 81
Garman M. B. 88, 147, 150, 153,
 154, 196, 197, 198
Gaver D. P. 98, 115, 118, 124, 149,
 157, 198
Geisler M. A. 23
Ghurye S. G. 200
Gilbert R. F. 22
Gillespie 26
Gnugnoli G. 25, 29, 81
Goodman A. S. 32, 179
Goldberger A. S. 95
Granger C. W. J. 21
Graver C. A. 51
Gray H. L. 87, 115, 118, 179, 180
Gregg L. W. 66
Guetzkow H. 26
Gürtler H. 63, 68, 132, 150
- Haitsma A. H.** 19, 51, 131, 155, 198
Halton J. H. 19, 21, 32, 34, 196
Hammersley J. M. 17, 22, 88, 98,
 141, 142, 178, 196, 197
Handscorn D. C. 17, 22, 87, 88, 97,
 98, 127, 132, 142, 155, 157, 178, 196,
 197, 198
Hanken A. F. G. 17
Hanna J. F. 66, 81
Harling J. 81, 142
Hartley H. O. 126, 196
Hartley M. G. 81
Hastings W. K. 132, 196
Hauser N. 29, 75, 77, 79
Henzler H. 93
Herzberg A. M. 70
Heuts R. M. J. 21
Hill W. J. 70
Hillier F. S. 30, 34, 64, 81, 98, 125,
 127, 131, 142, 155, 198
Hoggatt A. C. 26, 77
Holland 26
- Holt C. C.** 81
Horn W. 26, 81
Howrey E. P. 23, 66, 68
Howrey P. 28
Huisman F. 64
Hunter W. G. 70
Hutter R. 25
- Ignall E. J.** 59, 155, 196
Iglehart D. L. 63, 64, 77
- Jansson B.** 31
Jessop W. N. 27, 196
Johnson B. M. 155
Johnston J. 22, 61, 116, 197, 198
- Kahn H.** 127, 128, 130, 131, 132,
 196, 198
Kay I. M. 29
Keeping E. S. 182, 187, 199
Kelejian H. H. 28
King G. W. 132
Kiviat P. J. 15, 16, 23, 30, 59, 60,
 66, 67, 74, 77, 81, 155
Kleijnen J. P. C. 23, 28, 51, 81, 121,
 129, 130, 131, 140, 146, 150, 155, 187,
 188, 197, 198, 199, 200
Klingel A. R. 25
Kmenta J. 22
Knuth D. E. 31, 34
Kohlas J. 34
Koopman B. O. 26
Kosten L. 196
Kotler P. 61
Kriens J. 40
- Lewis P. A. W.** 21, 22, 31, 32, 87,
 126, 157, 196
Lieberman G. J. 30, 34, 64, 81, 98,
 125, 127, 131, 142, 155, 198
Lombaers H. J. M. 51, 79, 124, 155,
 176, 196, 199
Lovell C. C. 81
Lucas H. C. 25
Lyesen D. P. 14, 16, 29, 51
- McCracken D. D.** 20
McKenney J. L. 81
MacPherson R. D. 31

- Maguire J. N. 25, 29
Maisei H. 25, 29
Mann I. 196
Mann N. R. 25
March J. G. 26
Marks B. L. 61, 81
Marsaglia G. 31, 32, 34
Marshall A. W. 88, 127, 128, 130,
132, 153, 196, 198
Matérn B. 126, 196
Mayne D. Q. 150
Maxwell W. L. 88, 125, 155, 196,
198
McKay W. 76, 77
Mechanic H. 76, 77
Meier R. C. 22, 23, 26, 28, 29, 30,
34, 51, 64, 73, 74, 79, 81, 88, 108
Mickey M. R. 122
Mihram G. A. 14, 16, 21, 25, 27, 29,
30, 31, 34, 51, 66, 67, 73, 75, 77, 78,
81, 149, 155
Mikhail W. M. 22
Mitrofi I. I. 81
Mize J. H. 149, 155
Molenaar W. 19, 24, 98, 132, 155,
198
Moore J. M. 25
Morgenthaler G. W. 20, 21, 23, 26,
29, 30, 34
Morse P. M. 40
Morton K. W. 141, 142
Mosteller F. 115
Moy W. A. 87, 88, 93, 99, 100, 102,
103, 104, 111, 113, 114, 116, 121, 123,
124, 132, 136, 137, 138, 139, 140, 141,
149, 156, 178, 196, 197, 198, 199
Mullick S. K. 61, 81
Nance R. E. 32, 73
Naylor T. H. 6, 14, 16, 17, 22, 23,
25, 26, 27, 28, 29, 30, 31, 34, 49, 51,
59, 60, 61, 62, 66, 67, 71, 73, 77, 78,
79, 81, 165, 178, 187, 199, 200
Neave H. R. 21
Neeleman D. 22
Nef 196
Nelson R. T. 14, 15, 34, 63, 64, 65,
79, 81, 196
Neumann (von) J. 20
Newman T. G. 21, 31, 34, 98, 142,
198
Niedereichholtz J. 28, 73
Nisenfeld A. E. 29
Odell P. L. 21, 31, 34, 98, 142, 198
Oosterhoff J. 19, 51, 131, 142, 155,
198
Orcutt G. M. 26, 28, 65, 67, 79,
108
Overholt J. L. 60
Overstreet C. 32
Page E. S. 144, 149
Parzen E. 81
Pegels C. C. 60
Petterson S. 81
Pomerantz A. G. 25
Price C. M. 99
Pugh E. L. 132, 138
Quenouille M. 180
Radema F. W. 87, 114, 116, 118,
121, 123, 124, 145, 146, 149, 150, 155
Raj D. 87, 91, 93, 95, 115, 118, 196
Rao J. N. K. 115, 118, 122
Rechtschaffen R. N. 25
Reith P. F. 28
Rens P. J. 21
Ringer L. J. 99, 196
Robbins H. 200
Rose J. 16
Rosenhead J. 25, 81
Rytz R. 29
Salazar R. C. 25
Saleeb S. 81
Salsburg D. S. 179, 180
Sasser W. E. 22
Satterthwaite F. E. 91
Scheffé H. 91, 176, 198
Schlaifer R. 60
Schmidt J. W. 25, 31, 34, 51, 60, 66,
67, 81
Schrank W. E. 81
Schnaucany W. R. 87, 115, 118, 179,
180
Sen S. K. 25

Shah J. C. 81
 Shannon R. E. 26
 Schink W. A. 22
 Shedler G. S. 118, 150, 157, 197,
 198
 Shreider Y. A. 19, 21
 Shubic M. 23, 34
 Simon H. A. 66
 Sisson R. L. 17, 23, 24, 26, 29, 34,
 51, 60, 61, 63, 65, 67, 73, 81, 147,
 155, 157
 Smith W. P. 81
 Smith J. U. M. 25, 26, 28, 29, 73
 Starbuck 26, 30, 34, 66, 68, 79, 81,
 196
 Stash S. F. 67
 Syert 26

 Taylor R. E. 25, 31, 34, 51, 60, 66,
 67, 81
 Teichroew D. 19, 22, 28
 Ten Broeke A. M. 34
 Theil H. 17
 Tocher K. D. 19, 22, 25, 27, 31, 34,
 63, 75, 88, 92, 96, 97, 98, 99, 117, 118,
 124, 125, 126, 142, 144, 145, 147,
 157, 196
 Thomson J. D. 64
 Tukey J. W. 115, 179
 Turban E. 26

Ulam 20

 Van Daal J. 34, 63
 Van Dixhoorn J. J. 14, 16, 29, 51
 Van Doeland F. 34, 63
 Van Horn R. L. 68, 81
 Van Slyke R. M. 25, 196

 Wagner H. M. 59
 Wallace T. D. 197
 Walsh J. E. 81
 Wetherill G. B. 60, 71
 Westgaard T. P. 25, 34
 Williams W. H. 96
 Wilson R. C. 25
 Wolf G. 23, 66
 Wurl H. 155, 196

 Yang S. C. 150, 157, 197

О Г Л А В Л Е Н И Е

Планирование имитационных экспериментов	5
Предисловие	12
Г л а в а I. Основы моделирования	14
I.1. Введение и резюме	14
I.2. Системы	14
I.3. Модели	16
I.4. Методы решения	17
I.5. Метод Монте-Карло	17
I.6. Имитационное моделирование	22
I.6.1. Описание метода моделирования	22
I.6.2. Моделирование и системы	27
I.6.3. Проблемы моделирования	29
I.6.4. Генерирование случайных чисел и величин	30
I.7. Литература	34
Глоссарий (словарь специальных терминов)	35
Приложения к главе I	35
Приложение I.1. Оценивание интеграла методом Монте-Карло	35
Приложение I.2. Пример выборочного распределения	37
Приложение I.3. Моделирование стратегии обслуживания	39
Приложение I.4. Генератор случайных чисел Льюиса—Лермонта	43
Приложение I.5. Выборка двух коррелированных величин	49
Упражнения	50
Примечания	51
Библиография	52
Г л а в а II. Статистические аспекты моделирования	58
II.1 Введение и резюме	58
II.2. Формулировка задачи	59
II.3. Анализ входных данных	60
II.4. Модель и программа для вычислительной машины	62
II.5. Обоснованность модели	66
II.6. План эксперимента	68
II.7. Анализ нескольких вариантов системы	71
II.8. Объем выборки	72
II.9. Анализ выхода в установившемся состоянии и длина реализации	76
II.10. Прочие статистические аспекты	78
II.11. Литература	79

Упражнения	80
Примечания	81
Библиография	82
Глава III. Методы понижения дисперсии	86
III.1. Введение и резюме (краткое изложение)	86
III.2. Стратифицированная выборка	89
III.2.1. Основы стратифицированной выборки	89
III.2.2. Стратификация после выборки	92
III.2.3. Стратифицированная выборка в методе Монте-Карло	97
III.2.4. Стратифицированная выборка в имитационном моделировании	99
III.3. Селективная выборка или метод фиксированной последовательности	105
III.3.1. Описание метода	105
III.3.2. Комментарий	107
III.4. Контрольные величины или метод регрессионной выборки	109
III.4.1. Основные понятия метода контрольных величин	109
III.4.2. Типы контрольных величин	110
III.4.3. Оценивание контрольных коэффициентов	115
III.4.4. Приложения контрольных величин в моделировании и исследованиях по методу Монте-Карло	123
III.5. Значимая выборка	126
III.5.1. Основания значимой выборки.	126
III.5.2. Пример значимой выборки в методе Монте-Карло	128
III.5.3. Значимая выборка в моделировании	132
III.6. Дополняющие величины	141
III.7. Общие случайные числа	151
III.8. Совместное применение дополняющих величин и общих случайных чисел.	156
III.8.1. Введение	156
III.8.2. Конфликт между дополняющими величинами и общими случайными числами	157
III.8.3. Три альтернативы.	159
III.8.4. Сравнение альтернатив	163
III.8.5. Оптимальная альтернатива и распределение машинного времени	165
III.8.6. Оценивание коэффициентов в методе оптимизации	169
III.8.7. Комментарии к методу оптимизации	172
III.9. Заключение и выводы	177
III.10. Литература	178
Приложения к главе III.	178
Приложение III.1. Статистика «складного ножа».	178
Приложение III.2. Стратификации после выборки с объединением слов	180
Приложение III.3. Средняя квадратичная ошибка	181
Приложение III.4. Оценка дисперсии и стратификации после выборки	182
Приложение III.5. Пример генерирования случайных чисел в стратифицированной выборке	184

Приложение III.6. Математическое ожидание контрольной величины при случайном m	184
Приложение III.7. Ковариации между оценками контрольных величин	185
Приложение III.8. Аппроксимация оптимальной плотности распределения гамма-распределением в примере значимой выборки	186
Приложение III.9. Дисперсия оценки значимой выборки $g_3^*(x)$	188
Приложение III.10. Генерирование дополняющих случайных чисел из мультипликативно-конгруэнтного отношения	189
Приложение III.11. Дисперсия оценки разности между откликами двух систем	190
Приложение III.12. Знаки коэффициентов a_1 и a_2	191
Приложение III.13. Оптимальное распределение машинного времени	192
Упражнения	195
Примечания	196
Библиография	201
Ответы к упражнениям	206
Указатель перевода терминов	208
Именной указатель	215

ОГЛАВЛЕНИЕ ВТОРОГО ВЫПУСКА

Планирование имитационных экспериментов

Глава IV. Планирование и анализ экспериментов

- IV.1. Введение и резюме
- IV.2. Общие факторные планы и их анализ
- IV.3. Факторные планы с факторами на двух уровнях
- IV.4. Основы двухуровневых дробных факторных планов
- IV.5. Планы разрешающей способности III, IV и V
- IV.6. Дисперсия ошибки опыта, неадекватность и последовательное планирование
- IV.7. Планы отсеивания
- IV.8. Прочие вопросы планирования эксперимента
- IV.9. Литература

Приложения к главе IV

- Приложение IV.1. Дисперсионный анализ однофакторного эксперимента
- Приложение IV.2. Оценивание взаимодействий в эксперименте 2^k
- Приложение IV.3. Анализ планов Уэбба
- Приложение IV.4. Пример применения формулы Бокса для расширения планов
- Приложение IV.5. Выбор уровней незначимых факторов в групповом отсеивании
- Приложение IV.6. Роль взаимодействий в групповом отсеивании

Глава V. Объем выборки и надежность

V.A. Надежность для одной совокупности

V.A.1. Аннотация

V.A.2. Оценивание дисперсии в имитационном моделировании

V.A.3. Фиксированный объем выборки и одна совокупность

V.A.4. Определение объема выборки для одной совокупности

Приложения к части V.A.

Приложение V.A.1. Корреляция между последовательными опытами

Приложение V.A.2. Метод подопытов Механика и Маккея

V.B. Методы множественных сравнений

V.B.1. Вступление и краткое резюме

V.B.2. Уровни ошибок

V.B.3. Методы множественных сравнений

V.B.4. Эффективность и робастность ММС в имитационном моделировании

V.B.5. Другие методы и экспериментальная ситуация

Приложения к части V.B.

Приложение V.B.1. Неравные дисперсии в методе Даннета

V.B. Методы множественного ранжирования

V.B.1. Введение и резюме

V.B.2. Подход, использующий «зоны безразличия»

V.B.3. Существующие методы

V.B.4. Эффективность, робастность и приближенные методы

V.B.5. Другие распределения, решения и постановки задач

Приложения к части V.B.

Приложение V.B.1. Выбор одной из лучших совокупностей

Приложение V.B.2. Одновыборочный метод для известных дисперсий

Приложение V.B.3. Критические константы для метода В Сриваставы

Глава VI. Эксперименты Монте-Карло с методом множественного ранжирования Бехгофера-Блюментала: пример

VI.1. Введение и резюме

VI.2. Цель эксперимента

VI.3. Факторы в эксперименте

VI.4. План эксперимента

VI.5. Число повторных реализаций

VI.6. Методы понижения дисперсии

VI.7. Результаты эксперимента

VI.8. Проверка робастности ММР

VI.9. Определение важных факторов

VI.10. Заключение и направление дальнейших исследований

Приложения к главе VI

Приложение VI.1. Общие случайные числа и оценивание эффектов факторов

Приложение VI.2. Простой и обобщенный метод наименьших квадратов МНК

Ответы к упражнениям

Дополнительная библиография по имитационному моделированию и некоторым смежным вопросам

Указатель перевода терминов

Именной указатель

Клейнен Дж.

К48 Статистические методы в имитационном моделировании/Пер. с англ. Ю. П. Адлера, К. Д. Аргуновой, В. Н. Варыгина, А. М. Талалая; Под ред. и с предисл. Ю. П. Адлера и В. Н. Варыгина. — Вып. 1. — М.: Статистика, 1978. — 221 с., ил. — (Математико-статистические методы за рубежом).

В пер.: 1 р. 10 к.

Работа представляет собой обзор статистических аспектов построения имитационных моделей. Выпуск первый (выпуск второй выйдет в свет в 1978 г.) включает главы I—III. В гл. I излагаются основные понятия и дается краткий обзор применения методов; гл. II посвящена обзору статистических аспектов имитации; в гл. III описываются методы понижения дисперсии. Для чтения книги требуется знание основ теории вероятностей и математической статистики.

Книга предназначена для научных работников, занимающихся построением моделей в экономике и других областях. Она может быть полезна также преподавателям статистических дисциплин, аспирантам и студентам старших курсов вузов.

ББК 22.172

К $\frac{10805^1-096}{008(01)-78}$ 36-78

517.8

¹ Второй индекс 20204.

Дж. Клейнен

**СТАТИСТИЧЕСКИЕ МЕТОДЫ
В ИМИТАЦИОННОМ МОДЕЛИРОВАНИИ**

Редакторы *Е. В. Крестьянинова, К. М. Чижевская*
Мл. редактор *О. Б. Степанченко*
Техн. редактор *В. А. Чуракова*
Корректоры *Г. А. Башарина, Я. Б. Островский,*
О. Г. Тарасевич
Худ. редактор *Т. В. Стихно*
Обложка художника *Т. Н. Погореловой*
ИБ № 664

Сдано в набор 23/XII 1977 г. Подписано к печати
16/VI 1978 г. Формат бумаги 60×90¹/₁₆ Бумага № 2
Объем 14 печ. л. Уч.-изд. л. 15,75 Усл. печ. л. 14
Тираж 9 000 экз. (Тематич. план 1978 г. № 36)
Заказ 2248 Цена 1 р. 10 к.

Издательство «Статистика», Москва, ул. Кирова, 39.
Московская типография № 4 Союзполиграфпрома
при Государственном комитете Совета Министров СССР
по делам издательств, полиграфии и книжной торговли,
Москва, И-41, Б. Переяславская ул., дом № 46