

И.М.Соболь

Численные методы Монте-Карло



И. М. СОБОЛЬ

ЧИСЛЕННЫЕ МЕТОДЫ МОНТЕ-КАРЛО



ИЗДАТЕЛЬСТВО «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
Москва 1973

517.8+518

С 54

УДК 519.2

Численные методы Монте-Карло, И. М. Соболев. Главная редакция физико-математической литературы изд-ва «Наука», 1973.

Книга возникла из курса, который автор неоднократно читал в Московском инженерно-физическом институте, где у слушателей предполагалось знакомство с теорией вероятностей в весьма ограниченном объеме (соответствующем программе втузов). На этом уровне удалось рассмотреть важнейшие разделы теории методов Монте-Карло.

В книге эти разделы изложены значительно полнее, имеется много примеров, подобраны упражнения. Многие результаты излагаются впервые.

Книга рассчитана на студентов втузов, инженеров, научных работников. Она будет особенно полезной специалистам по вычислительной и прикладной математике.

Книга содержит 73 рис., библи. 185 назв.

© Издательство «Наука», 1973.

С $\frac{0224-1857}{042(02)-73}$ 57-73

ОГЛАВЛЕНИЕ

Предисловие	5
Введение	7
Глава 1. Получение случайных величин на ЭВМ	10
§ 1. Три способа получения случайных величин	10
§ 2. Псевдослучайные числа	19
§ 3. Статистическая проверка случайных чисел	30
Упражнения к главе 1	42
Глава 2. Преобразования случайных величин	44
§ 1. Метод обратных функций (основной прием моделирования случайных величин)	44
§ 2. Моделирование многомерных случайных величин	53
§ 3. Преобразования вида $\xi = g(\gamma_1, \gamma_2)$	61
§ 4. Преобразования вида $\xi = g(\gamma_1, \dots, \gamma_n)$	70
§ 5. Методы отбора	74
Упражнения к главе 2	83
Глава 3. Вычисление интегралов	86
§ 1. Общий метод оценки математических ожиданий	86
§ 2. Простейший метод Монте-Карло для вычисления интеграла	93
§ 3. Важнейшие способы построения хороших оценок (способы уменьшения дисперсии)	100
§ 4. Интегралы, зависящие от параметра	123
Упражнения к главе 3	131
Глава 4. Вычисление интегралов (сложные оценки)	135
§ 1. Методы Монте-Карло с повышенной скоростью сходимости	135
§ 2. Случайные квадратурные формулы	143
§ 3. Использование смещенных оценок	151
Упражнения к главе 4	159
Глава 5. Решение линейных уравнений	161
§ 1. Интегральные преобразования	161
§ 2. Неоднородные интегральные уравнения	171
§ 3. Пример: рассеяние частиц	182
§ 4. Однородные интегральные уравнения	185
§ 5. Решение линейных алгебраических систем	193
Упражнения к главе 5	207

ОГЛАВЛЕНИЕ

Глава 6. Моделирование естественных процессов . . .	210
§ 1. Моделирование путем имитации	211
§ 2. Моделирование свободного пробега	221
§ 3. Использование статистических весов	231
§ 4. Статистические веса и интегральные уравнения . . .	247
Упражнения к главе 6	251
Глава 7. Неслучайные точки в алгоритмах Монте-Карло	253
§ 1. Конструктивная размерность алгоритмов Монте-Карло	254
§ 2. n -мерные псевдослучайные точки	259
§ 3. Поиски «универсальных» псевдослучайных чисел . .	270
§ 4. Проверка псевдослучайных чисел с детерминистической точки зрения	274
Упражнения к главе 7	277
Глава 8. Некоторые другие задачи	279
§ 1. Интерполирование функций от большого числа переменных	279
§ 2. Простейший случайный поиск	281
§ 3. Решение уравнения Лапласа	284
§ 4. Вычисление винеровских интегралов	287
Приложения	
I. Вспомогательное неравенство	292
II. Таблицы	292
Литература	298
Указатель	308

ПРЕДИСЛОВИЕ

Книга эта возникла из курса, который я в течение ряда лет читал в Московском инженерно-физическом институте, но довольно сильно отличается от него: все вопросы изложены здесь значительно полнее. Основной материал, предназначенный для общего курса методов Монте-Карло, содержится в главах 1, 2, 3, 6, 7, исключая мелкий шрифт. Предполагается, что читатель знаком с теорией вероятностей в сравнительно небольшом объеме, примерно соответствующем программе по высшей математике для вузов *).

За последнее десятилетие сфера приложений методов Монте-Карло необычайно расширилась. Методы Монте-Карло используются для расчета задач физики (перенос излучения и вещества, ядерная физика, статистическая физика и др.), радиотехники, теории массового обслуживания, теории надежности, химии, биологии, экономики (оптимизация, управление, сетевое планирование и др.), теории автоматов, аэродинамики, гидрологии — перечислить все невозможно. В книге рассмотрены почти все наиболее важные вопросы, связанные с применением методов Монте-Карло, и можно надеяться, что она будет полезна специалистам, использующим эти методы, независимо от области приложений.

По мнению автора, современный курс методов Монте-Карло обязательно должен содержать хотя бы краткое изложение детерминистического подхода к методам Монте-Карло, так как использование так называемых детерминированных псевдослучайных чисел позволяет во многих задачах увеличить скорость сходимости ($1/N^{1-\varepsilon}$ вместо $1/\sqrt{N}$), не нарушая структуры вычислительного алгоритма. Этим вопросам посвящена глава 7.

*) В лекции для инженеров-физиков следует включать также ряд вопросов из главы 5.

Обычно преобразования случайных величин излагаются как справочный материал: различные способы моделирования экспоненциальных величин, различные способы моделирования нормальных величин и т. д. Вместо этого в главе 2 построена общая классификация преобразований, используемых для моделирования различных случайных величин. На первый взгляд принцип этой классификации может показаться формальным. Но полностью его роль выясняется в главе 7 в связи с введенным в книге понятием конструктивной размерности алгоритмов Монте-Карло.

Нередко при изложении методов Монте-Карло много места уделяют способам решения задач линейной алгебры. Однако такие способы редко применяют на практике, где, как правило, используют более быстро сходящиеся численные методы линейной алгебры. Поэтому в главе 5 излагаются в первую очередь способы решения линейных и интегральных уравнений; и лишь в § 5 этой главы как частный случай рассмотрены алгебраические системы.

В главе 6 рассмотрены различные способы введения статистических весов. Эти способы позволяют, отходя от естественного процесса, строить модели для расчета, более выгодные, чем имитация процесса (см. мелкий шрифт, стр. 9). Устанавливается связь этих приемов с методами вычисления интегралов и решения интегральных уравнений.

Главы 4 и 8 выходят за рамки общего курса. В главе 4 указаны наиболее интересные и, вероятно, наиболее перспективные направления исследований методов вычисления интегралов, а в главе 8 рассмотрены некоторые задачи других типов. Таким образом, эти главы как бы иллюстрируют возможности развития методов Монте-Карло «вглубь» и «вширь».

Упражнения предназначены в первую очередь для того, чтобы сообщить читателю дополнительные сведения.

Я пользуюсь случаем, чтобы выразить свою признательность Н. Н. Ченцову, многолетний контакт с которым повлиял на мои взгляды на методы Монте-Карло.

И. Соболев

ВВЕДЕНИЕ

0.1. Методы Монте-Карло. Общепринятого определения методов Монте-Карло пока нет. Назовем методами Монте-Карло *численные методы решения математических задач при помощи моделирования случайных величин* *). При таком определении приходится к методам Монте-Карло причислить некоторые другие методы, как, например, стохастические приближения или случайный поиск, которые по традиции рассматриваются отдельно. Однако специалисты, занимающиеся этими вопросами, нередко сами называют свои приемы методами Монте-Карло.

В то же время в определении подчеркивается что:

а) речь идет о численных методах (и конкурировать они могут с классическими численными методами, а не с аналитическими методами решения задач);

б) решать методами Монте-Карло можно любые математические задачи (а не только задачи вероятностного происхождения, связанные со случайными величинами).

Название «Монте-Карло» произошло от города Монте-Карло (княжество Монако), известного своим казино, ибо одним из простейших приборов для генерирования случайных чисел служит рулетка.

Официальной датой рождения методов Монте-Карло считают 1949 год, когда появилась статья под заглавием «Метод Монте-Карло» [159]. Возникновение метода связывают обычно с именами Дж. Неймана, С. Улама, Н. Метрополиса, а также Г. Кана и Э. Ферми; все они в 40-х годах работали в Лос-Аламосе (США). Необходимо сразу же подчеркнуть, что теоретические основы

*) Можно добавить: «и статистической оценки их характеристик».

методов Монте-Карло были известны значительно раньше. Более того, фактически такие методы не раз использовались для расчетов в математической статистике. Однако до появления электронных вычислительных машин (ЭВМ) методы Монте-Карло не могли стать универсальными численными методами, ибо моделирование случайных величин вручную — весьма трудоемкий процесс.

Развитию методов Монте-Карло способствовало бурное развитие ЭВМ. Алгоритмы Монте-Карло (как правило, обладающие небольшой связностью) сравнительно легко программируются и позволяют рассчитывать многие задачи, недоступные для классических численных методов. Так как совершенствование ЭВМ продолжается, есть все основания ожидать дальнейшего развития методов Монте-Карло и дальнейшего расширения области их применения.

0.2. Общий курс методов Монте-Карло. Важнейший прием построения методов Монте-Карло — сведение задачи к расчету математических ожиданий. Более подробно: для того чтобы приближенно вычислить некоторую скалярную величину a , надо придумать такую случайную величину ξ , что $M\xi = a$; тогда, вычислив N независимых значений ξ_1, \dots, ξ_N величины ξ , можно считать, что

$$a \approx (1/N) (\xi_1 + \dots + \xi_N).$$

Пример. Требуется оценить объем V_G некоторой ограниченной пространственной фигуры G .

Выберем параллелепипед Π , содержащий G , объем которого V_Π известен (рис. 1). Выберем N случайных точек, равномерно распределенных в Π , и обозначим через N' количество точек, попавших в G . Если N велико, то, очевидно, $N' : N \approx V_G : V_\Pi$, откуда получаем оценку

$$V_G \approx V_\Pi (N'/N).$$

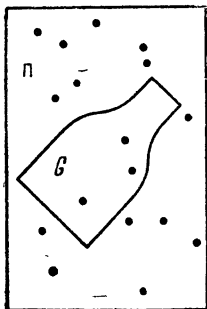


Рис. 1.

В этом примере случайная величина ξ равна V_Π , если случайная точка попадает в G , и ξ равна нулю, если точка попадает в $\Pi - G$. Нетрудно проверить, что математическое ожидание $M\xi = V_G$, а среднее арифметическое

$$(1/N) (\xi_1 + \dots + \xi_N) = V_\Pi (N'/N).$$

Легко видеть, что существует бесконечно много случайных величин ξ таких, что $M\xi = a$. Поэтому теория методов Монте-Карло должна дать ответы на два вопроса:

1) как выбрать удобную величину ξ для расчета той или иной задачи?

2) как находить значения ξ_1, ξ_2, \dots произвольной случайной величины ξ ?

Изучение этих вопросов и должно составить основное содержание практического курса методов Монте-Карло.

Почти все методы, рассмотренные в настоящей книге, основаны на расчете математических ожиданий. За рамками книги остались упомянутые выше методы случайного поиска (кроме простейшего) и стохастических приближений.

Среди методов Монте-Карло можно выделить методы, в которых полностью воспроизводится модель рассчитываемого процесса. Такие методы иногда называют «физическими», хотя автору представляется более удачным другое название этих методов — имитационные. Имитация естественных процессов широко используется в самых различных областях науки, техники, экономики. Однако приемы имитаций в каждой области свои, и подробно излагать их более целесообразно в специальных руководствах, а не в общем курсе методов Монте-Карло.

Нет оснований считать, что имитация естественного процесса — это лучший способ для расчета этого процесса; скорее — наоборот. В самом деле, при имитации вычисляется вся информация о течении процесса; однако в реальных задачах, как правило, не нужна вся информация о всех величинах; поэтому естественно ожидать, что существуют методы, в которых скорость расчета нужных величин повышается благодаря отказу от информации о ненужных величинах.

Слово «численные» включено в заглавие книги, чтобы подчеркнуть, что здесь рассматриваются главным образом не имитационные методы.

ГЛАВА 1

ПОЛУЧЕНИЕ СЛУЧАЙНЫХ ВЕЛИЧИН НА ЭВМ

В алгоритмах Монте-Карло фигурируют значения случайных величин с различными законами распределения. Как будет доказано в гл. 2, для того чтобы вычислять значения любых таких величин, достаточно уметь находить значения какой-нибудь одной случайной величины, ибо всегда можно подобрать такую функцию от этой случайной величины, которая имеет требуемый закон распределения. Поэтому мы сначала рассмотрим вопрос о том, как получать на ЭВМ значения одной, «стандартной», случайной величины.

§ 1. Три способа получения случайных величин

1.1. Случайные числа и случайные цифры. Как правило, в качестве стандартной выбирают непрерывную случайную величину γ , равномерно распределенную в интервале $(0, 1)$. Напомним основные характеристики этой величины: при $0 < x < 1$ плотность $p_\gamma(x) = 1$, а функция распределения $F_\gamma(x) = x$; математическое ожидание $M\gamma = 1/2$, дисперсия $D\gamma = 1/12$.

Заметно реже в качестве стандартной используют дискретную случайную величину ϵ , которая с одинаковой вероятностью может принимать 10 значений 0, 1, 2, ..., 9. Распределение ϵ задается таблицей

$$\begin{pmatrix} 0 & 1 & 2 & \dots & 9 \\ 0,1 & 0,1 & 0,1 & \dots & 0,1 \end{pmatrix}$$

Мы будем называть величину γ *случайным числом*, а величину ϵ *случайной цифрой*. Иногда ϵ называют де-

сятичной случайной цифрой, чтобы отличить ее от двоичной случайной цифры — величины α с распределением

$$\begin{pmatrix} 0 & 1 \\ 0,5 & 0,5 \end{pmatrix}.$$

Чтобы установить связь между γ и ε , разложим число γ в бесконечную десятичную дробь:

$$\gamma = 0, \varepsilon_1 \varepsilon_2 \dots \varepsilon_k \dots \quad (1)$$

Последняя запись означает, что

$$\gamma = \sum_{k=1}^{\infty} \varepsilon_k \cdot 10^{-k}.$$

Теорема 1. Десятичные цифры $\varepsilon_1, \dots, \varepsilon_k, \dots$ случайного числа γ представляют собой независимые случайные цифры. Обратно, если $\varepsilon_1, \dots, \varepsilon_k, \dots$ — независимые случайные цифры, то формула (1) определяет случайное число.

Доказательство. Если величина γ равномерно распределена в $(0, 1)$, то $\varepsilon_k = i$ тогда и только тогда, когда

$$0, \varepsilon_1 \dots \varepsilon_{k-1} i \leq \gamma < 0, \varepsilon_1 \dots \varepsilon_{k-1} i + 0, \underbrace{0 \dots 01}_k, \quad (2)$$

причем $\varepsilon_1, \dots, \varepsilon_{k-1}$ в (2) могут принимать любые значения $0, 1, \dots, 9$. Так как длина каждого из интервалов (2) равна 10^{-k} и интервалы эти попарно не пересекаются, то

$$P\{\varepsilon_k = i\} = \sum_{\varepsilon_1, \dots, \varepsilon_{k-1}=0}^9 10^{-k} = 10^{k-1} \cdot 10^{-k} = 0,1;$$

сумма берется по всем интервалам, так что $\varepsilon_1, \dots, \varepsilon_{k-1}$ независимо пробегает значения $0, 1, \dots, 9$.

Пусть теперь $1 \leq s < k$. Рассмотрим вероятность одновременного выполнения равенств $P\{\varepsilon_k = i, \varepsilon_s = j\}$. Очевидно, оба эти равенства будут выполнены тогда и только тогда, когда выполнено неравенство (2), и в то же время $\varepsilon_s = j$. Следовательно,

$$P\{\varepsilon_k = i, \varepsilon_s = j\} = \sum_{\varepsilon_1, \dots, \varepsilon_{s-1}, \varepsilon_{s+1}, \dots, \varepsilon_{k-1}=0}^9 10^{-k} = 0,01.$$

Точно так же $P\{\epsilon_{k_1}=i_1, \dots, \epsilon_{k_s}=i_s\}=10^{-s}$ и, таким образом,

$$P\{\epsilon_{k_1}=i_1, \dots, \epsilon_{k_s}=i_s\}=P\{\epsilon_{k_1}=i_1\} \dots P\{\epsilon_{k_s}=i_s\},$$

а это означает независимость $\epsilon_{k_1}, \dots, \epsilon_{k_s}$.

Перейдем к доказательству обратного утверждения. Произвольное число x из интервала $(0, 1)$ запишем в форме бесконечной десятичной дроби

$$x=0, a_1 a_2 \dots a_k \dots \quad (3)$$

Если $\gamma < x$, то в разложении (1) либо $\epsilon_1 < a_1$, либо $\epsilon_1 = a_1$ и $\epsilon_2 < a_2$, либо $\epsilon_1 = a_1$, $\epsilon_2 = a_2$ и $\epsilon_3 < a_3$, и т. д. Поэтому

$$P\{\gamma < x\} = \sum_{k=1}^{\infty} P\{\epsilon_1 = a_1, \dots, \epsilon_{k-1} = a_{k-1}, \epsilon_k < a_k\}.$$

Так как по условию теоремы все $\epsilon_1, \dots, \epsilon_k$ независимы, то

$$P\{\gamma < x\} = \sum_{k=1}^{\infty} P\{\epsilon_1 = a_1\} \dots P\{\epsilon_{k-1} = a_{k-1}\} P\{\epsilon_k < a_k\}.$$

Легко видеть, что $P\{\epsilon_k < a_k\} = a_k \cdot 10^{-1}$, ибо значениями ϵ_k в этом случае могут быть $0, 1, \dots, a_k - 1$. Значит,

$$P\{\gamma < x\} = \sum_{k=1}^{\infty} a_k 10^{-1} \cdot 10^{-(k-1)} = \sum_{k=1}^{\infty} a_k 10^{-k} = x,$$

и теорема доказана.

1.2. О приближенных случайных числах. В вычислениях всегда используют числа с конечным количеством десятичных знаков, поэтому вместо случайных чисел γ употребляют конечные десятичные дроби $\gamma=0, \epsilon_1 \dots \epsilon_n$. Никаких специальных исследований по этому поводу мы проводить не будем: мы считаем, что здесь имеет место ошибка округления такая же, как при любых приближенных вычислениях.

Отметим одно простое свойство чисел γ , которое теряется при таком приближении *).

*) В книге использованы следующие обозначения: $\underline{C}(x)$ — целая часть числа x (т. е. наибольшее целое число, не превосходящее x), $\underline{D}(x)$ — дробная часть числа x (т. е. $\underline{D}(x) = x - \underline{C}(x)$).

Теорема 2. Пусть g — произвольное целое положительное число. Случайная величина $\eta = D(g\gamma)$ равномерно распределена в интервале $(0, 1)$.

В самом деле, если x принадлежит интервалу $(0, 1)$, то вероятность

$$\begin{aligned} P\{\eta < x\} &= \sum_{k=0}^{g-1} P\{k \leq g\gamma < k+x\} = \\ &= \sum_{k=0}^{g-1} P\{kg^{-1} \leq \gamma < (k+x)g^{-1}\} = \sum_{k=0}^{g-1} xg^{-1} = x, \end{aligned}$$

что и требовалось доказать

Ясно, что приближенная случайная величина $\tilde{\gamma} = 0, \varepsilon_1 \dots \varepsilon_n$ не обладает таким свойством, так как при $g = p^{2^n}$, где p — целое, $D(g\tilde{\gamma}) = 0$.

Иногда вычислители пытаются использовать цифры $\varepsilon_1, \dots, \varepsilon_k$ числа γ для одной цели, а цифры $\varepsilon_{k+1}, \dots, \varepsilon_n$ — для другой. Теоретически такой прием вполне обоснован: из теоремы 1 вытекает, что одно случайное число γ эквивалентно бесконечной последовательности случайных величин γ :

$$0, \varepsilon_1 \dots \varepsilon_n, \quad 0, \varepsilon_{n+1} \dots \varepsilon_{2n}, \quad 0, \varepsilon_{2n+1} \dots \varepsilon_{3n}, \dots$$

Тем не менее его не всегда можно рекомендовать на практике: как мы увидим ниже, «качество» случайных чисел, используемых в расчетах, проверяется с помощью специальных тестов; и нередко главные разряды — $\varepsilon_1, \varepsilon_2, \varepsilon_3$ чисел γ оказываются гораздо лучше проверенными, чем более далекие разряды.

1.3. Таблицы случайных цифр. Предположим, что мы осуществили N независимых опытов, в результате которых получили N случайных цифр $\varepsilon_1, \dots, \varepsilon_N$. Записав эти цифры (в порядке появления) в таблицу, получим то, что называется *таблицей случайных цифр* (см. стр. 295, где цифры объединены в группы только ради удобства чтения).

Способ употребления такой таблицы весьма прост. Если в ходе расчета некоторой задачи нам потребуется случайная цифра ε , то мы можем взять любую цифру ε_s из этой таблицы. Если нам понадобится случайное число γ , то мы можем взять из таблицы n очередных цифр и считать, что $\gamma = 0, \varepsilon_s \varepsilon_{s+1} \dots \varepsilon_{s+n-1}$. Выбирать цифры из такой таблицы в случайном порядке не обязательно. Их можно выбирать подряд. Но, конечно, можно начинать с любого места, читать в любом направлении, использовать любой заранее заданный алгоритм

выбора, не зависящий от конкретных значений цифр таблицы *).

Все сказанное выше относится к «идеальной» таблице случайных цифр и не вызывает никаких сомнений. Мы не будем подробно останавливаться на трудностях, возникающих при переходе к «реальным» таблицам, но отметим некоторые из них.

Во-первых, изготовление хорошей таблицы — весьма сложное дело, ибо в любом реальном опыте всегда возможны ошибки. Например, для изготовления таблицы [166], содержащей миллион случайных цифр, была построена и тщательно отлажена специальная «рулетка» (с использованием электроники). Тем не менее после некоторого периода хорошей работы она стала выдавать, как показала проверка, не равновероятные цифры. Таким образом, проверка «качества» таблицы абсолютно необходима. Никакие априорные соображения о тщательности постановки опыта не гарантируют нас от ошибок.

Тесты для проверки таблиц случайных цифр рассмотрены ниже в § 3. Здесь только поясним, что проверяется частота каждой из десяти цифр, частота различных соседних пар и т. п.

Вторая трудность связана с «незаконностью» многократного использования одной и той же таблицы. Вычислителей этот вопрос не очень беспокоит, так как интуитивно ясно, что таблицу можно повторно использовать при решении независимых задач (впрочем, что такое «независимые» задачи — не очень ясно). Возможность использования одних и тех же случайных величин для решения целых классов задач доказана ниже в гл. 3, § 4.

Наконец, отметим еще известный парадокс: в большой таблице найдутся плохие участки. (Например, в таблице, содержащей 10^{100} цифр, вполне вероятно найти

*) Таблицы случайных цифр очень удобны для осуществления случайной выборки. Пусть, например, требуется из 55 предметов отобрать случайно 6 предметов. Занумеруем исходное множество предметов и воспользуемся таблицей со стр. 295. Рассмотрим пары цифр из таблицы: 86, 51, 59, 07, 95, 66, 15, 56, 64, ... и отберем предметы с номерами, оказавшимися в этой последовательности: 51, 07, 15, 34, 23, 32.

100 пульс подряд.) Очевидно, самостоятельное использование таких участков недопустимо.

В настоящее время таблицы случайных цифр (или более разнообразные таблицы случайных величин) используют главным образом при расчетах вручную; для расчетов на ЭВМ ими практически не пользуются. Основная причина этого — чисто техническая: внутренний накопитель всех современных ЭВМ сравнительно мал и большая таблица туда не помещается; а обращение к внешним запоминающим устройствам сильно замедляет счет. Вторая причина — ограниченность объема существующих таблиц, — по-видимому, играет второстепенную роль: можно заготовить таблицу любого объема, если это потребуется.

1.4. Датчики случайных чисел. *Генераторами* или *датчиками* случайных величин называют различные технические устройства, вырабатывающие случайные величины. Чаще всего для построения датчика используют «шумящие» радиоэлектронные приборы (диоды, тиратроны, газотроны и др.). Не вдаваясь в технические подробности, рассмотрим один из возможных способов построения датчика, вырабатывающего случайные двоичные цифры α .

Нетрудно представить себе счетчик, который подсчитывает количество ν флуктуаций напряжения шумящего

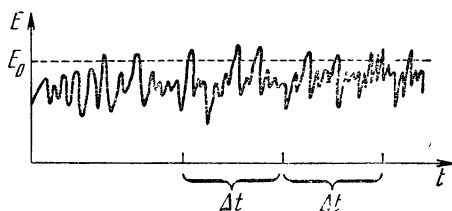


Рис. 2.

прибора, превышающих заданный уровень E_0 за фиксированное время Δt (рис. 2). Еще проще устроить счетчик, который выдавал бы число $\nu \pmod{2}$, т. е. 0 при четном ν и 1 при нечетном ν . Если вероятности появления 0 и 1 в таком процессе равны между собой, то можно считать, что устройство вырабатывает случайную последовательность двоичных цифр.

Если вероятность появления нуля отлична от половины $P\{0\} = p \neq 0,5$, то можно ввести какую-нибудь схему стабилизации вероятности. Например, можно группировать цифры парами и выдавать 0 при получении пары 01 и 1 — при получении пары 10, а пары 00 и 11 просто опускать. Так как $P\{01\} = P\{10\} = p(1-p)$, то в результате получим последовательность нулей и единиц с равными вероятностями *).

Обычно датчики случайных чисел содержат m генераторов описанного типа, работающих независимо, так что датчиком выдается приближенное случайное число $\gamma = 0, \alpha_1 \dots \alpha_m$, записанное в форме m -разрядной двоичной дроби. Для случайных чисел отведена специальная ячейка в накопителе, и скорость генерирования их столь велика, что на каждом такте работы ЭВМ в этой ячейке получается новое случайное число.

Применение датчиков случайных чисел свободно от тех недостатков, которые препятствуют широкому применению таблиц: не требуется места во внутреннем накопителе и запас чисел практически неограничен. Тем не менее подавляющее большинство задач, решенных методом Монте-Карло, сосчитано без применения датчиков. Ибо датчики имеют свои, новые недостатки. Во-первых, числа, выработанные датчиком, нельзя воспроизвести. Это затрудняет контроль расчетов и делает невозможным счет на таких ЭВМ, на которых двойной пересчет является правилом **). Во-вторых, приходится содержать и эксплуатировать дополнительное устройство, которое требует ухода и регулярной проверки «качества» вырабатываемых чисел с помощью специальных тестов.

Основные области применения датчиков — системы автоматического регулирования и аналоговые вычислительные машины, а не методы Монте-Карло.

1.5. Метод псевдослучайных чисел. Мы видели, что пригодность случайных чисел определяется в конечном счете не процессом их получения, а тем, удовлетворяют

*) К сожалению возможен также случай, когда вероятность p меняется во времени или зависит от предшествующих цифр, и тогда простые схемы стабилизации не спасают положения.

**) Можно, конечно, запомнить все выработанные числа, но тогда мы фактически будем использовать таблицу.

ли они некоторым принятым тестам. Но в таком случае совершенно безразлично, как эти числа получены, они могут быть даже сосчитаны по какой-нибудь формуле, лишь бы они удовлетворяли тестам!

Такая точка зрения иногда вызывает возражения, ибо, вычисляя числа по формуле, мы тем самым сразу отказываемся от их «случайности». Любопытно, что такого рода возражения чаще высказывают нематематики, хотя, казалось бы, математики должны более решительно протестовать против нестрогости. Дело в том, что математика лучше понимают, что слово «случайность», взятое выше в кавычки, использовано там не в математическом, а скорее в обыденном смысле. С точки зрения математика равномерно распределенная случайная величина γ — это абстрактное понятие; и только опыт может убедить нас в том, что какая-либо конкретная последовательность чисел $\gamma_1, \gamma_2, \dots, \gamma_N$ обладает интересующими нас свойствами независимых случайных чисел γ .

Числа $\gamma_1, \gamma_2, \dots, \gamma_N$, которые вычисляются по какой-либо заданной формуле и могут быть использованы вместо случайных чисел при решении некоторых задач, называются *псевдослучайными числами*.

Нередко считают, что тесты, приведенные ниже в § 3, обеспечивают возможность использования проверенных случайных или псевдослучайных чисел в любых задачах. Это излишний оптимизм. В действительности каждый тест связан с каким-то достаточно широким, но ограниченным классом задач (см. гл. 7, § 4), и правильнее было бы говорить о таблицах, датчиках или псевдослучайных числах для определенных классов задач *).

Конкретная последовательность псевдослучайных чисел сходна с таблицей случайных чисел: ее можно один раз тщательно проверить и затем многократно применять; все числа легко воспроизводятся; и запас чисел в такой последовательности, как мы увидим позднее, тоже ограничен. Однако метод псевдослучайных чисел свободен от главного недостатка таблиц: существуют простые формулы для расчета псевдослучайных чисел,

*) Любой реальный объект лишь приближенно описывается математическим понятием, и всегда можно указать задачи, в которых данная математическая модель непригодна. Например, как бы аккуратно ни была начерчена прямая, при достаточно большом увеличении она превратится в волнистую или даже прерывную линию.

такие, что на получение каждого числа затрачивается всего 3—5 команд ЭВМ, а программа расчета занимает в накопителе лишь несколько ячеек.

подавляющее большинство расчетов методами Монте-Карло выполнено с помощью псевдослучайных чисел.

1.6. Сравнение трех способов с практической точки зрения. Попутно с изложением методов получения случайных величин мы рассмотрели некоторые их достоинства и недостатки. Остановимся еще на вопросе о быстроте выработки случайных величин.

При решении на ЭВМ большинства задач на получение псевдослучайных чисел затрачивается гораздо меньше времени, чем на расчет самих задач. В самом деле, практика показывает, что в настоящее время количество псевдослучайных чисел, используемых при решении отдельных задач, как правило *), имеет порядок 10^4 — 10^5 и редко достигает 10^6 — 10^7 . Если на выработку каждого числа затрачивать 5 операций, то на получение всех чисел уйдет не более $5 \cdot 10^7$ операций. На современных ЭВМ, скорость которых равна $5 \cdot 10^5$ операций в секунду, это составит не более 100 секунд.

Используя датчики, можно организовать «сверхбыструю» выработку случайных чисел, так как, присоединив к ЭВМ k датчиков, мы будем за один такт работы ЭВМ получать k случайных чисел. Однако выигрыш времени при расчете сложной задачи в указанных выше условиях окажется меньше двух минут. Так что, исходя из интересов методов Монте-Карло, не стоит гнаться за «сверхбыстрой» выработкой случайных величин *). Разумно требовать, чтобы скорость выработки была того же порядка, что скорость счета ЭВМ. Псевдослучайные числа этому требованию удовлетворяют.

В табл. 1 перечислены достоинства и недостатки трех методов по шести признакам. Подчеркнем, что речь идет только об «эксплуатационных» особенностях, а не о «качестве» вырабатываемых чисел, которое должно проверяться во всех трех методах.

*) Исключение составляют такие задачи, в которых огромные массивы случайных чисел подвергаются очень простой обработке. Но такие задачи выгоднее решать на специализированных машинах, а не на универсальных ЭВМ. К тому же, вероятно, для большинства таких задач можно придумать лучшие способы решения...

Из таблицы видно, что метод псевдослучайных чисел — самый удобный с практической точки зрения. Это подтверждается также всей практикой расчетов методами Монте-Карло.

Таблица 1

Метод	Достоинства	Недостатки
таблиц	Проверка однократная. Воспроизводить числа можно	Запас чисел ограничен. Занимает много места в накопителе или медленно вводится. Нужна внешняя память
датчиков	Запас чисел неограничен. Сверхбыстрое получение. Места в накопителе не занимает	Проверка периодическая. Воспроизводить числа нельзя. Требуется специальное устройство
псевдослучайных чисел	Проверка однократная. Воспроизводить числа можно. Быстрое получение. Места в накопителе занимает мало. Внешние устройства не нужны	Запас чисел ограничен

Необходимо, однако, признать, что не все разделяют эту точку зрения. Некоторые авторы отдают предпочтение датчикам, аргументируя тем, что:

а) псевдослучайные числа вырабатываются медленнее и занимают место во внутреннем накопителе;

б) не стоит расходовать рабочее время большой ЭВМ, когда можно использовать время небольшой приставки;

в) а все-таки числа, вырабатываемые датчиком, «случайные», а псевдослучайные числа — нет...

§ 2. Псевдослучайные числа

Общей теории псевдослучайных чисел в настоящее время нет. О возможных путях построения такой теории разговор пойдет в гл. 7. Здесь мы ограничимся изложением некоторых фактов, главным образом эмпирического характера.

2.1. Алгоритмы вида $\gamma_{n+1} = \Phi(\gamma_n)$. Большинство алгоритмов, используемых на практике для получения псевдослучайных чисел, представляют собой рекуррентные формулы первого порядка

$$\gamma_{n+1} = \Phi(\gamma_n), \quad (4)$$

где начальное число γ_0 задано. Легко показать, что функция $y = \Phi(x)$, изображенная на рис. 3, не может

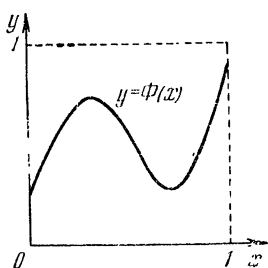


Рис. 3.

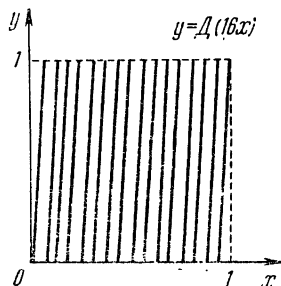


Рис. 4.

породить хорошую последовательность псевдослучайных чисел $\gamma_1, \gamma_2, \dots$. В самом деле, если по настоящим случайным числам построить точки с координатами $(\gamma_1, \gamma_2), (\gamma_3, \gamma_4), \dots$, то они равномерно распределены в единичном квадрате $0 < x < 1, 0 < y < 1$, в то время как соответствующие точки, построенные по числам (4),

$$(\gamma_1, \Phi(\gamma_1)), (\gamma_3, \Phi(\gamma_3)), (\gamma_5, \Phi(\gamma_5)), \dots,$$

располагаются на кривой $y = \Phi(x)$.

Следовательно, «хорошую» последовательность может породить только такая функция $y = \Phi(x)$, график которой весьма плотно заполняет единичный квадрат. Примером такой функции может служить функция $y = D(gx)$ при очень больших g (рис. 4). И действительно, эта функция послужила основой для ряда методов получения псевдослучайных чисел.

Конечно, приведенное условие только необходимо, но далеко не достаточно для того, чтобы формула (4) порождала «хорошие» псевдослучайные числа.

Другая важная черта алгоритмов вида (4) — то, что при реализации их на ЭВМ они всегда порождают пе-

риодические последовательности. В самом деле, так как в коде любой ЭВМ можно записать лишь конечное число N чисел, заключенных между нулем и единицей, то рано или поздно какое-нибудь значение γ_L совпадает с одним из предыдущих значений γ_i . Тогда в силу (4)

$$\gamma_{L+i} = \gamma_{i+l} \quad \text{при } i = 1, 2, \dots \quad (5)$$

Пусть L — наименьшее число, удовлетворяющее (5) при некотором l ($l < L$); множество чисел $\gamma_0, \gamma_1, \dots, \gamma_{L-1}$

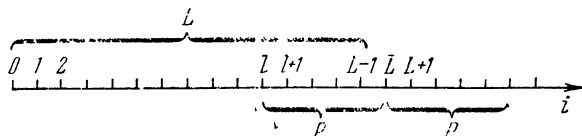


Рис. 5.

называется *отрезком аперидичности* последовательности (4), число L — *длиной отрезка аперидичности*, а $P = L - l$ — *длиной периода* (рис. 5).

Очевидно, в рассматриваемом случае отрезок аперидичности состоит из различных чисел. И обычно для расчета не рекомендуют использовать больше чем L чисел последовательности (4)*). Ясно также, что $L \leq N$.

Для экспериментального определения L и P можно рекомендовать следующий алгоритм. Выбирается целое число T (в несколько раз меньшее, чем предполагаемое значение L). Вычисляются подряд числа $\gamma_0, \gamma_1, \gamma_2, \dots$, причем $\gamma_0, \gamma_T, \gamma_{2T}, \dots$ запоминаются, и все γ_i из отрезка $pT < i \leq (p+1)T$ сравниваются со всеми $\gamma_0, \gamma_T, \dots, \gamma_{pT}$. Как только окажется выполненным равенство $\gamma_{i_0} = \gamma_{mT}$, вычисляется длина периода $P = i_0 - mT$ и число $t = i_0 - pT$.

Чтобы вычислить L , по числу $\gamma_{(p-1)T+t}$ вычисляется $\gamma_{(p-1)T+t}$, после чего считаются и сравниваются между собой две последовательности значений: $\gamma_{(p-1)T+j}$ и $\gamma_{(m-1)T-t+j}$ при $j = t, t+1, t+2, \dots$. Не позднее чем при $j = t+T$ произойдет совпадение этих значений. Если совпадение произойдет при $j = j_0$, то $L = (p-1)T + j_0$.

2.2. Некоторые конкретные алгоритмы.

2.2.1. Первый алгоритм для получения псевдослучайных чисел был предложен Дж. Нейманом. Его называют

*) Впрочем, использование $L+P$ чисел, как правило, ни к каким осложнениям не приводит, особенно при решении сложных задач, когда маловероятно, что повторяющееся число будет вторично использовано для моделирования той же величины.

методом середины квадрата (middle-square method). В этом методе число γ_n предполагается $2k$ -значным:

$$\gamma_n = 0, a_1 a_2 \dots a_{2k}.$$

Чтобы получить число γ_{n+1} , надо γ_n возвести в квадрат

$$\gamma_n^2 = 0, b_1 b_2 \dots b_{4k}$$

и затем отобрать средние $2k$ цифр этого квадрата:

$$\gamma_{n+1} = 0, b_{k+1} b_{k+2} \dots b_{3k}.$$

Нетрудно проверить, что этому методу соответствует функция

$$\Phi(x) = D[10^{-2k} C(10^{3k} x^2)]$$

или, что то же,

$$\Phi(x) = 10^{-2k} C[10^{2k} D(10^k x^2)].$$

Однако от метода середины квадрата вычислители отказались, так как в последовательностях, построенных таким образом, получается больше, чем нужно, малых чисел (см. ниже упражнение 4 гл. 1). Интересно, что при некоторых γ_0 наблюдается вырождение последовательности, т. е. $\gamma_n \equiv 0$ при $n \geq n_0$.

2.2.2. Наибольшее распространение получил алгоритм, предложенный Д. Лемером, который называют методом вычетов (residual method) или методом сравнений (congruential method). В этом методе $\Phi(x) = D(gx)$, т. е.

$$\gamma_{n+1} = D(g\gamma_n), \quad (6)$$

где g — большое целое число.

Докажем, что если задать γ_0 в форме несократимой дроби $\gamma_0 = t_0/M$, где t_0 и M — целые числа, и M взаимно просто с g , то все γ_n будут несократимыми дробями вида $\gamma_n = t_n/M$, где числители t_n определяются формулой

$$t_{n+1} = gt_n \pmod{M}. \quad (7)$$

Запись (7) означает, что t_{n+1} равно остатку, полученному при делении gt_n на M (или, другими словами, t_{n+1} — это наименьший положительный вычет gt_n по модулю M).

Доказательство (по индукции). Пусть $\gamma_n = m_n/M$ — несократимая дробь. Тогда $g\gamma_n = gm_n/M$. Из (7) следует, что $gm_n = rM + m_{n+1}$, где r — целое число. Легко видеть, что m_{n+1} взаимно просто с M . В самом деле, если допустить, что у m_{n+1} и M есть общий множитель — простое число h , то gm_n должно делиться на h ; но g делиться на h не может, так как g взаимно просто с M ; поэтому должно делиться на h число m_n , а это противоречит предположению о том, что m_n/M несократимая дробь. Итак, m_{n+1} взаимно просто с M и

$$\gamma_{n+1} = D(gm_n/M) = D(r + m_{n+1}/M) = m_{n+1}/M.$$

На практике обычно используют формулу (7). Изучению последовательностей m_n посвящено много работ. Методами теории чисел удалось исследовать длину отрезка аперiodичности (см. упражнение 5 гл. 1) и оценить некоторые величины, аналогичные коэффициентам корреляции между γ и $\eta = \Phi(\gamma)$, между γ и $\eta = \Phi(\Phi(\gamma))$ и т. п. Однако вопрос о пригодности таких псевдослучайных чисел в конечном счете решается обычными статистическими тестами § 3, т. е. эмпирически. При некоторых (g, M, m_0) получаются удовлетворительные последовательности, при других — плохие. (См. также гл. 7, п. 3.3.)

Удовлетворительная последовательность псевдослучайных чисел получается, например, при $g=5^{17}$, $M=2^{42}$, $m_0=1$ ($L=P=2^{40}$); результаты статистической проверки этих чисел приведены в [171].

Обширная литература, посвященная методу сравнений, имеется в [69, 139, 140]. Ряд статей с указанием недостатков метода пересчислен в [180].

2.2.3. В качестве примера нелинейного алгоритма, использующего некоторые особенности системы команд ЭВМ, рассмотрим алгоритм, предложенный автором книги для ЭВМ «Стрела». В этом алгоритме число γ_{n+1} получается из числа γ_n тремя командами:

- 1) число γ_n умножается на большую константу g ;
- 2) изображение произведения $g\gamma_n$ в ячейке ЭВМ сдвигается на 7 разрядов влево;
- 3) вычисляется абсолютная величина полученного числа, которая и есть γ_{n+1} (при вычислении абсолютной величины число нормализуется).

Удовлетворительная последовательность псевдослучайных чисел получается, например, при $g=10^{17}$ и $\gamma_0=1$ ($L=87\,835$, $P=53\,535$); результаты статистической проверки этих чисел приведены в [75].

Чтобы пояснить этот алгоритм, необходимо указать, что в ЭВМ «Стрела» числа представляются в нормализованной двоичной форме $x = \text{sgn } x \cdot m \cdot 2^p$, где $p=1+\lfloor \log_2 |x| \rfloor$ — порядок числа, а m — мантисса ($0,5 \leq m < 1$). Ячейка ЭВМ, в которую записывается число x ,

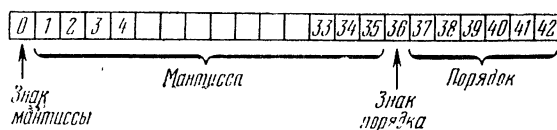


Рис. 6.

состоит из 43 двоичных разрядов (рис. 6). В j -м разряде записана величина e_j , которая может равняться нулю или единице, и

$$m = \sum_{j=1}^{35} e_j 2^{-j}, \quad |p| = \sum_{j=0}^5 e_{42-j} 2^j, \\ \text{sgn } x = (-1)^{e_0}, \quad \text{sgn } p = (-1)^{e_{42}}$$

(в двоичной записи

$$m = 0, e_1 e_2 \dots e_{35}, |p| = e_{37} e_{38} \dots e_{42}).$$

Если произведению $g \gamma_n$ соответствует ячейка $[0, e_1 e_2 \dots e_{42}]$, то после операции сдвига получаем $[e_7 e_8 \dots e_{42} 0 0 0 0 0 0]$. Абсолютная величина этого числа в двоичной записи равна $\gamma_{n+1} = 0, e_8 e_9 \dots e_{42}$.

Нетрудно проверить, что в случае, когда размещение мантиссы m и порядка p в ячейке ЭВМ другое, то же число γ_{n+1} можно получить из $g \gamma_n$ с помощью двух или трех сдвигов и поразрядного сложения результатов этих сдвигов.

2.2.4. Для получения псевдослучайных чисел на различных ЭВМ используют весьма разнообразные алгоритмы, подобные изложенному в п. 2.2.3, которые иногда называют *способами перемешивания*. Сведения о них имеются в [19, 69]. К сожалению, описание алгоритмов далеко не всегда сопровождается указанием количества проверенных чисел и тестов, которым эти числа удовлетворяют, или хотя бы указанием работ, в которых имеются данные о таких проверках.

Например, некоторое распространение получил трехкомандный алгоритм, разработанный Д. И. Голенко для той же ЭВМ «Стрела». Однако позднее сам автор алгоритма обнаружил ([19], стр. 117), что получаемые этим методом числа плохо удовлетворяют одному из важнейших тестов — проверке пар.

2.2.5. Из других алгоритмов вида (4) отметим попытки «улучшения» метода вычетов (6) путем рассмотрения функций вида

$$\Phi(x) = D(\rho + g_1 x)$$

или

$$\Phi(x) = D\left(\rho + \sum_{k=1}^m g_k x^k\right),$$

где ρ — заданное дробное число.

Заметного распространения эти методы не получили.

2.3. Оценка L для алгоритмов вида $\gamma_{n+1} = \Phi(\gamma_n)$.

Для оценки длины отрезка аperiodичности, характерного для последовательностей вида (4), можно воспользоваться следующей элементарной вероятностной моделью. Пусть имеется конечное множество, содержащее N различных чисел. Произведем последовательность независимых опытов, в каждом из которых из этого множества извлекается и записывается одно число. Вероятность извлечения каждого числа в каждом из опытов равна $1/N$ (выборка с возвратом). Обозначим через L случайную величину: номер опыта, в котором впервые будет извлечено уже записанное число.

Теорема 3. В указанной модели для любого $x > 0$

$$\lim_{N \rightarrow \infty} P\{L/\sqrt{N} < x\} = 1 - e^{-x^2/2}. \quad (8)$$

Доказательство. Обозначим записанные числа через $\gamma_0, \gamma_1, \gamma_2, \dots$. Легко вычислить вероятность того, что L окажется равным n , где $1 \leq n \leq N$:

$$P\{L = n\} = \frac{N-1}{N} \cdot \frac{N-2}{N} \cdot \dots \cdot \frac{N-(n-1)}{N} \cdot \frac{n}{N},$$

так как

$$P\{\gamma_1 \neq \gamma_0\} = (N-1)/N,$$

$$P\{\gamma_2 \neq \gamma_0, \gamma_2 \neq \gamma_1 | \gamma_1 \neq \gamma_0\} = (N-2)/N,$$

$$\dots \dots \dots$$

$$P\left\{\gamma_{n-1} \neq \gamma_0, \dots, \gamma_{n-1} \neq \gamma_{n-2} \left| \begin{array}{c} \text{все } \gamma_0, \dots, \gamma_{n-2} \\ \text{различны} \end{array} \right. \right\} = \frac{N-(n-1)}{N},$$

а вероятность того, что γ_n совпадет с одним из различных чисел $\gamma_0, \gamma_1, \dots, \gamma_{n-1}$, равна n/N .

Фиксируем произвольное $x > 0$, и пусть $k = \lfloor x \sqrt{N} \rfloor$. Тогда

$$P\{L < x\sqrt{N}\} = \sum_{n=1}^k P\{L = n\} = \sum_{n=1}^k \left(1 - \frac{1}{N}\right) \left(1 - \frac{2}{N}\right) \cdots \left(1 - \frac{n-1}{N}\right) \frac{n}{N}.$$

Легко видеть, что

$$\begin{aligned} \ln \left[\left(1 - \frac{1}{N}\right) \cdots \left(1 - \frac{n-1}{N}\right) \right] &= \sum_{i=1}^{n-1} \ln \left(1 - \frac{i}{N}\right) = \\ &= - \sum_{i=1}^{n-1} \left(\frac{i}{N} + \frac{i^2}{2N^2} + \cdots \right) = - \frac{n(n-1)}{2N} + O\left(\frac{n^3}{N^2}\right). \end{aligned}$$

Так как $n \leq k = O(\sqrt{N})$, то $n/N = O(1/\sqrt{N})$; поэтому

$$\ln \left[\left(1 - \frac{1}{N}\right) \cdots \left(1 - \frac{n-1}{N}\right) \right] = - \frac{n^2}{2N} + O\left(\frac{1}{\sqrt{N}}\right),$$

и интересующую нас вероятность можно записать в форме

$$P\{L < x\sqrt{N}\} = \sum_{n=1}^k e^{-\frac{n^2}{2N}} (n/N) [1 + O(1/\sqrt{N})]. \quad (9)$$

Рассмотрим теперь разбиение интервала $0 < t < x$ точками $t_n = n/\sqrt{N}$, где $n = 1, 2, \dots, k$ (рис. 7), и составим интегральную сумму для функции $te^{-t^2/2}$ по этому разбиению:

$$\sum_{n=1}^k e^{-t_n^2/2} t_n (t_n - t_{n-1}) + e^{-x^2/2} x (x - t_k).$$

Из неравенства $x\sqrt{N} - 1 < k \leq x\sqrt{N}$ видно, что $t_k = k/\sqrt{N} \rightarrow x$ при

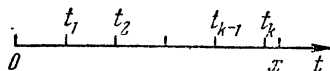


Рис. 7.

$N \rightarrow \infty$. Поэтому последнее слагаемое в этой сумме стремится при $N \rightarrow \infty$ к нулю, и

$$\begin{aligned} \lim_{N \rightarrow \infty} \sum_{n=1}^k e^{-n^2/2N} \frac{n}{N} &= \lim_{N \rightarrow \infty} \sum_{n=1}^k e^{-t_n^2/2} t_n (t_n - t_{n-1}) = \\ &= \int_0^x te^{-t^2/2} dt = 1 - e^{-x^2/2}. \end{aligned}$$

Отсюда и из соотношения (9) следует (8). Теорема доказана.

Следствие. Так как математическое ожидание неотрицательной случайной величины с функцией распределения $1 - e^{-x^{2,2}}$ равно $\sqrt{\pi/2}$, то из (8) следует, что при $N \rightarrow \infty$

$$ML \sim \sqrt{(\pi/2)N}. \quad (10)$$

Рассмотренная в настоящем пункте модель предложена в статье [75], там же получена формула (10), а теорема 3 доказана Л. Н. Большевым и Д. И. Голенко. Несмотря на грубость модели, оценки (8) и (10) оказались полезными во многих опытах, связанных со способами «перемешивания», когда теоретико-числовая оценка периода практически невозможна. В качестве N выбирается количество возможных значений функции $\Phi(x)$ в конкретной ЭВМ; тогда можно ожидать, что длина отрезка аспериодичности окажется порядка $\sqrt{(\pi/2)N}$.

Например, в методе п. 2.2.3 на ЭВМ «Стрела» значение функции $\Phi(x)$ (до нормализации) записывается в форме $0, e_8, e_9 \dots e_{42}$ значит, $N=2^{35}$ и $L \sim 2,3 \cdot 10^5$. На практике для ряда подобных алгоритмов были экспериментально обнаружены отрезки аспериодичности длиной от $0,8 \cdot 10^5$ до $3,5 \cdot 10^5$.

З а м е ч а н и е. В книге [19] параграф, содержащий теорему 3, называется «Критерий проверки периодичности последовательности псевдослучайных чисел». Однако по этому «критерию» судить о качестве псевдослучайных чисел нельзя. У некоторых вполне удовлетворительных последовательностей, полученных по формуле (6), длина L гораздо больше, чем оценка (10) (фактически $L=O(N)$). В то же время некоторые последовательности с $L \approx \sqrt{(\pi/2)N}$ не удовлетворяют ни одному из основных тестов § 3.

2.4. О более сложных алгоритмах.

2.4.1. Вместо формул вида (4), представляющих собой рекуррентные формулы первого порядка, можно пытаться использовать для получения последовательностей псевдослучайных чисел рекуррентные формулы порядка r

$$\gamma_{n+1} = \Phi(\gamma_n, \gamma_{n-1}, \dots, \gamma_{n-r+1}), \quad (11)$$

считая, что начальные значения $\gamma_0, \gamma_1, \dots, \gamma_{r-1}$ заданы. Вероятно, длина отрезка аспериодичности у такой последовательности при $r > 1$ гораздо больше, чем при $r=1$. Однако теорема, аналогичная теореме 3, даже для случая $r=2$ не доказана.

В ряде работ рассматривались методы вида (11), представляющие собой обобщения методов пп. 2.2.1 и 2.2.2:

$$\gamma_{n+1} = D[10^{-2h} C(10^{3h} \gamma_n \gamma_{n-1})]$$

и

$$\gamma_{n+1} = D(\rho + g_1\gamma_n + g_2\gamma_{n-1} + \dots + g_r\gamma_{n-r+1}).$$

Все же применяются такие методы редко. Наверно компенсация за сложность этих методов по сравнению с методами вида (4) недостаточна.

2.4.2. Для получения последовательности $\gamma_0, \gamma_1, \dots$ можно одновременно использовать два алгоритма типа (4) с различными функциями $\Phi(x)$ и $\Psi(x)$:

$$\gamma_{n+1} = \begin{cases} \Phi(\gamma_n), & \text{если } n \not\equiv 0 \pmod{M}, \\ \Psi(\gamma_n), & \text{если } n \equiv 0 \pmod{M}. \end{cases} \quad (12)$$

В счете по методу (12) почти все время используется формула $\gamma_{n+1} = \Phi(\gamma_n)$, и только когда n кратно M , последовательность «возмущается»: $\gamma_{n+1} = \Psi(\gamma_n)$. Поэтому метод этот, предложенный Д. И. Голенко, назван *методом возмущений*, а целое число M — *периодом возмущения*.

Метод возмущений увеличивает длину отрезка апериодичности по сравнению с методом (4). Вместо оценки (10) получается оценка (см. упражнение 6 гл. 1)

$$ML \sim V(\pi/2) NM.$$

Однако необходимо предостеречь вычислителей от некритического применения этого метода: качество всех используемых псевдослучайных чисел должно быть проверено. Наблюдались случаи, когда по каждой из формул $\gamma_{n+1} = \Phi(\gamma_n)$ или $\gamma_{n+1} = \Psi(\gamma_n)$ вычислялись удовлетворительные псевдослучайные числа (с $L \approx V(\pi/2) N$), а в реализованном по этим формулам методе (12) числа оказались плохими: не проходил первый тест п. 3.2 (несмотря на то, что L было близко к $V(\pi/2) NM$).

2.4.3. Рассмотрим рекуррентную формулу r -го порядка

$$\alpha_{n+r} = \sum_{k=0}^{r-1} e_k \alpha_{n+k} \pmod{2}, \quad (13)$$

где $e_0 = 1$, а все остальные коэффициенты e_j и переменные α_j могут принимать лишь два значения — 0 или 1. Если начальные значения $\alpha_0, \alpha_1, \dots, \alpha_{r-1}$, заданы, то (13) позволяет последовательно вычислить $\alpha_r, \alpha_{r+1}, \dots$. Случай $\alpha_0 = \alpha_1 = \dots = \alpha_{r-1} = 0$ из рассмотрения исключается. Уравнение (13) называется *моноклическим*, если период P решения $\alpha_0, \alpha_1, \dots, \alpha_n, \dots$ равен $P = 2^r - 1$.

Нетрудно доказать, что все решения моноциклического уравнения имеют одну и ту же длину периода P , и период их совпадает с отрезком атериодичности.

Пр и м е р. Уравнение $\alpha_{n+4} = \alpha_n + \alpha_{n+1} \pmod{2}$ моноциклическое. Если задать начальные значения $\alpha_0 = 1, \alpha_1 = \alpha_2 = \alpha_3 = 0$, то получим последовательность с периодом 15:

$$100 \ 010 \ 011 \ 010 \ 111 \ 100 \ 010 \ \dots;$$

если начать со значений $\alpha_0 = \alpha_1 = \alpha_2 = \alpha_3 = 1$, то снова получим последовательность с периодом 15:

$$111 \ 100 \ 010 \ 011 \ 010 \ 111 \ 100 \ \dots,$$

которая лишь сдвигом номеров отличается от предыдущей.

Уравнение $\alpha_{n+4} = \alpha_n + \alpha_{n+2} + \alpha_{n+3} \pmod{2}$ не моноциклическое. Если начать со значений $\alpha_0 = 1, \alpha_1 = \alpha_2 = \alpha_3 = 0$, то получим последовательность с периодом 7:

$$1000110 \ 1000110 \ \dots,$$

а если задать начальные значения $\alpha_0 = \alpha_1 = \alpha_2 = \alpha_3 = 1$, то получим последовательность, состоящую из одних единиц.

В статье Н. Ц и р л е р а [185], где исследованы важнейшие свойства уравнений (13), последовательности, порождаемые моноциклическими уравнениями, называются M -последовательностями. Термин «моноциклическое уравнение» введен в [82]. В [176] описаны электронные схемы, реализующие формулы вида (13); по этим схемам построены физические датчики псевдослучайных двоичных цифр. Ибо решения $\alpha_0, \alpha_1, \dots, \alpha_n, \dots$ моноциклических уравнений во многих отношениях ведут себя как двоичные случайные цифры (см. упражнение 7 гл. 1).

Например, формула

$$\alpha_{n+31} = \alpha_n + \alpha_{n+3} \pmod{2}$$

порождает последовательности с $L = P = 2^{31} - 1$. Проверка 600 000 цифр, получающихся при начальных значениях $(\alpha_0, \dots, \alpha_{30}) = (100 \ 101 \ 100 \ 111 \ 110 \ 001 \ 101 \ 110 \ 101 \ 0000)$, показала, что эти цифры удовлетворяют основным тестам п. 3.2, сформулированным применительно к двоичным цифрам (проверка выполнена Ю. Л. Левитаном).

Однако если из цифр $\alpha_0, \alpha_1, \dots, \alpha_n, \dots$ попытаться строить m -разрядные значения случайной величины γ

$$\gamma_0 = 0, \alpha_0 \dots \alpha_{m-1}, \quad \gamma_1 = 0, \alpha_m \dots \alpha_{2m-1},$$

$$\gamma_2 = 0, \alpha_{2m} \dots \alpha_{3m-1}, \dots,$$

то пригодность таких чисел, вероятно, окажется ограниченной: результаты работы Р. Таусворта [172] дают основания ожидать, что эмпирическое распределение групп

$$(\gamma_0, \dots, \gamma_{s-1}), \quad (\gamma_s, \dots, \gamma_{2s-1}), \quad (\gamma_{2s}, \dots, \gamma_{3s-1}), \dots$$

будет близким к теоретическому только при выполнении условия $m \cdot s \leq r_1$

§ 3. Статистическая проверка случайных чисел

Мы видели, что случайные числа, полученные любым из трех рассмотренных в § 1 методов, необходимо проверить. Рассмотрим важнейшие критерии, используемые обычно для такой проверки. Все эти критерии необходимы для случайных чисел, но о достаточности критериев можно говорить, только ограничив класс задач, которые предполагается решать с помощью проверяемых случайных чисел. Такая точка зрения будет развита в гл. 7, § 4.

3.1. Статистические критерии согласия

3.1.1. Теорема К. Пирсона. Рассмотрим произвольную случайную величину ξ , которая может быть одномерной или многомерной, дискретной или непрерывной. Обозначим через X множество возможных значений ξ . Фиксируем какое-нибудь разбиение множества X на r попарно непересекающихся множеств X_1, \dots, X_r таких, что

$$P\{\xi \in X_j\} = p_j > 0 \text{ при } j = 1, 2, \dots, r.$$

Очевидно, $p_1 + \dots + p_r = P\{\xi \in X\} = 1$.

Выберем N независимых значений ξ_1, \dots, ξ_N величины ξ и обозначим через v_j количество значений, принадлежащих X_j . Легко видеть, что математическое ожидание $Mv_j = Np_j$ (*).

В качестве меры отклонения «истинных» значений v_j от «теоретических» Np_j удобно выбрать величину

$$\chi_N^2 = \sum_{j=1}^r \frac{(v_j - Np_j)^2}{Np_j}. \quad (14)$$

Теорема. *Каковы бы ни были исходная величина ξ и разбиение $X = X_1 + \dots + X_r$ (такое, что все $p_j > 0$),*

*) Случайная величина v_j подчиняется биномиальному закону распределения с $Mv_j = Np_j$ и $Dv_j = Np_j(1 - p_j)$. Совокупность величин (v_1, \dots, v_r) подчиняется мультиномиальному закону распределения: при $k_1 + \dots + k_r = N$

$$P\{v_1 = k_1, \dots, v_r = k_r\} = \frac{N!}{k_1! \dots k_r!} p_1^{k_1} \dots p_r^{k_r}.$$

при каждом $x > 0$

$$\lim_{N \rightarrow \infty} \mathbb{P} \{ \chi_N^2 < x \} = \int_0^x k_{r-1}(x) dx, \quad (15)$$

где плотность $k_m(x)$, называемая плотностью распределения χ^2 с m степенями свободы, выражается формулой (рис. 8)

$$k_m(x) = [2^{m/2} \Gamma(m/2)]^{-1} x^{m/2-1} e^{-x/2}.$$

3.1.2. Критерий согласия χ^2 . Теорему предыдущего пункта часто используют в статистике для проверки гипотез о законе распределения случайной величины.

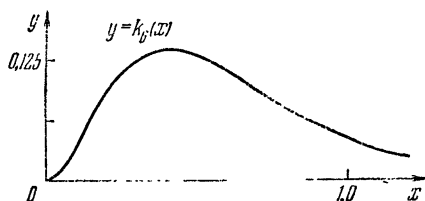


Рис. 8.

Фиксируем достаточно большую вероятность β , которую будем называть *доверительной вероятностью* или *коэффициентом доверия*. Вероятность $1-\beta$ обычно называют *уровнем значимости*. Теоретически в качестве β можно выбрать любое число $0 < \beta < 1$. Практически же выбор β означает, что (в рассматриваемой задаче) мы считаем событие с вероятностью $\geq \beta$ достоверным, а событие с вероятностью $\leq 1-\beta$ невозможным при единичном испытании.

Пусть теперь имеется конкретная гипотеза о законе распределения случайной величины ξ . В результате осуществления N независимых экспериментов были получены N значений ξ_1, \dots, ξ_N этой случайной величины (N достаточно велико). Не противоречат ли эти N значений нашей гипотезе?

Чтобы ответить на этот вопрос, можно рассуждать следующим образом. Выберем какое-нибудь число r и

разбиение множества возможных значений X случайной величины ξ на r попарно непересекающихся множеств $X = X_1 + \dots + X_r$. Исходя из нашей гипотезы, можно вычислить вероятности $p_j = P\{\xi \in X_j\}$. Предположим, что разбиение выбрано так, что все $p_j > 0$. Тогда по значениям ξ_1, \dots, ξ_N нетрудно вычислить величины v_1, \dots, v_r (этот этап в статистике называют *группировкой значений*) и по формуле (14) — величину χ^2_N . Если наша гипотеза справедлива, то (при достаточно большом N) эта величина достаточно хорошо подчиняется закону распределения χ^2 с $(r-1)$ -й степенью свободы. Из уравнения

$$\chi^2(r-1, 1-\beta) \int_0^\infty k_{r-1}(x) dx = 1 - \beta$$

можно найти значение $\chi^2(r-1, 1-\beta)$, отвечающее фиксированному нами уровню значимости $1-\beta$ (рис. 9). Если наше значение $\chi^2_N < \chi^2(r-1, 1-\beta)$, то этот результат не противоречит нашей гипотезе; если же $\chi^2_N \geq$

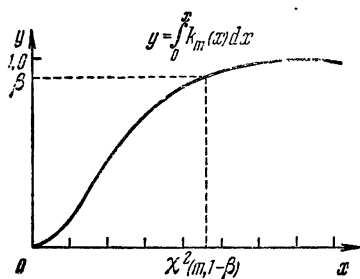


Рис. 9.

$\geq \chi^2(r-1, 1-\beta)$, то, в соответствии со сделанным соглашением, это означает, что наступило невозможное событие, и гипотеза должна быть отброшена, так как она привела к противоречию.

Конечно, вывод этот зависит от выбранной доверительной вероятности и поэтому не носит

абсолютного характера. Чаше других используют доверительные вероятности $\beta = 0,95; 0,99; 0,999$; соответствующие уровни значимости $1-\beta = 0,05; 0,01; 0,001$ называют 5%-ным, 1%-ным и 0,1%-ным уровнями. Если $\chi^2_N \geq \chi^2(r-1, 1-\beta)$ при $\beta = 0,95$, то значение χ^2_N называют *почти значимым*, при $\beta = 0,99$ — *значимым*, а при $\beta = 0,999$ — *высоко значимым* (рис. 10).

На практике широко используют таблицы распределения χ^2 (см. стр. 293), в которых приведены значения

$\chi^2 = \chi^2(m, P)$ — корни уравнения

$$\int_{\chi^2}^{\infty} k_m(x) dx = P, \quad (16)$$

где, очевидно, $P = 1 - \beta$. Если полученному в эксперименте значению χ_N^2 при $m = r - 1$ отвечает в таблице $P < 0,001$, то это значение χ_N^2 высоко значимо. Например, если при $r = 10$ получено значение $\chi_N^2 = 9,4$, то по таблице при девяти степенях свободы находим, что $P \approx 0,40$; следовательно, полученное значение χ_N^2 допустимо.

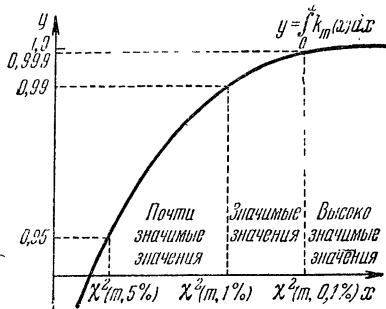


Рис. 10.

3.1.3. Замечания о применении критерия χ^2 .

а) О выборе разбиения. Если N заранее ограничено, то нельзя выбирать r слишком большим, так как тогда будут малы величины Nr_j и неустойчивыми значения v_j . Обычно рекомендуют выбирать X_i так, чтобы минимальные Nr_j были не меньше чем 20.

б) О больших r . В таблицах χ^2 число степеней свободы m обычно не превосходит 30. При больших значениях можно использовать таблицу интеграла вероятностей (стр. 293), так как распределение величины $\xi = \sqrt{2\chi^2} - \sqrt{2m}$ близко к нормальному с параметрами (0; 1):

$$P = \int_{\chi^2}^{\infty} k_m(x) dx \approx \frac{1}{2} [1 - \Phi(z)],$$

где $z = \sqrt{2\chi^2} - \sqrt{2m}$. В [5] имеется таблица поправок к последней приближенной формуле.)

в) О многократном применении критерия χ^2 . Соглашение о невозможности события с вероятностью $\leq 1 - \beta = P$ справедливо для единичного испытания. Если мы будем повторять наш опыт много раз, то (в среднем) для 100 $P\%$ полученных значений χ_N^2 соответствующие вероятности окажутся меньше P (при условии, что гипотеза наша верна).

г) В некоторых случаях гипотетическое распределение величины содержит неизвестные параметры a_1, \dots, a_s и эти параметры оцениваются по результатам тех же N испытаний (т. е. по значениям ξ_1, \dots, ξ_N). Тогда также можно использовать критерий

χ^2 для проверки согласия между значениями ξ_1, \dots, ξ_N и гипотетическим распределением с подобранными параметрами. Однако предельное распределение следует рассматривать с $m=r-s-1$ степенями свободы [44].

д) О «слишком хороших» значениях χ^2 . Допустим, что сосчитанное значение χ_N^2 оказалось столь малым, что соответствующее ему значение $P > 0,99$. В этом случае значения ξ_1, \dots, ξ_N настолько хорошо подчиняются гипотетическому распределению, что... возникает сомнение в их «случайности».

Как мы увидим ниже в гл. 7, для решения некоторых задач важно именно хорошее распределение значений (а не «случайность»), и такие «слишком хорошо» распределенные значения позволяют быстрее решать эти задачи.

3.1.4. К р и т е р и й ω^2 . Рассмотрим одномерную случайную величину ξ с функцией распределения $F(x)$. Выберем N независимых значений ξ_1, \dots, ξ_N этой случайной величины и построим эмпирическую (или выборочную) функцию распределения

$$F_N^*(x) = S_N(x)/N, \quad (17)$$

где $S_N(x)$ равно количеству значений, меньших чем x . (На рис. 11 $N=5$). Так как $F_N^*(x)$ — это частота события $\{\xi < x\}$ при N испытаниях, а вероятность этого события

$P\{\xi < x\} = F(x)$, то $F_N^*(x)$ сходится по вероятности*) к $F(x)$, когда $N \rightarrow \infty$.

В качестве меры отклонения $F_N^*(x)$ от $F(x)$ часто используют величину

$$\omega_N^2 = N \int_{-\infty}^{\infty} [F(x) - F_N^*(x)]^2 dF(x). \quad (18)$$

Теорема (Р. Мизес, Н. В. Смирнов). *Какова бы ни была случайная величина ξ с непрерывной функцией*

*) Последовательность случайных величин $\eta_1, \dots, \eta_N, \dots$ сходится по вероятности к постоянной c , если при любом $h > 0$ вероятность $P\{|\eta_N - c| \geq h\} \rightarrow 0$, когда $N \rightarrow \infty$.

распределения $F(x)$, при каждом $x > 0$

$$\lim_{N \rightarrow \infty} P \{ \omega_N^2 < x \} = a_1(x), \quad (19)$$

где функция $a_1(x)$ от ξ не зависит.

На этой теореме основан критерий согласия ω^2 , позволяющий проверять гипотезы о функции распределения одномерной непрерывной случайной величины ξ . Схема использования этого критерия точно такая же, как схема использования критерия χ^2 : фиксируется доверительная вероятность β ; из уравнения

$$a_1(x_\beta) = \beta$$

находится соответствующее значение x_β (рис. 12; таблица функции $a_1(x)$ на стр. 293); по гипотетической функции распределения $F(x)$ и эмпирической функции распределения $F_N^*(x)$ вычисляется ω_N^2 ; если величина $\omega_N^2 \geq x_\beta$, то это означает, что наступило невозможное событие или, другими словами, наша гипотеза привела

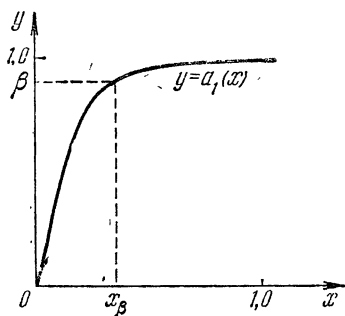


Рис. 12.

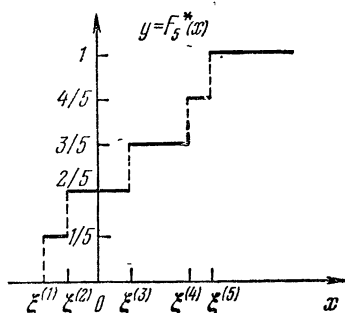


Рис. 13.

к противоречию. Конечно, предполагается, что количество значений N достаточно велико.

По сравнению с критерием χ^2 критерий ω^2 имеет одно преимущество: не нужна группировка значений (и, стало быть, не надо вводить параметр r). Применять критерий можно обычно уже при $N > 50$. Однако, чтобы вычислить ω^2 , нужно расположить значения ξ_1, \dots, ξ_N в порядке возрастания:

$$\xi_{(1)} \leq \xi_{(2)} \leq \dots \leq \xi_{(N)}.$$

(Такой ряд выборочных значений называется в статистике *вариационным рядом*). Если количество значений (т. е. N) превышает объем внутреннего накопителя ЭВМ, то расположение их в порядке возрастания представляет собой весьма трудоемкую процедуру. Из-за этого при очень больших N критерий ω^2 используют редко.

Выведем теперь формулу для расчета ω_N^2 . Во-первых, нетрудно проверить, что

$$F_N^*(x) = k/N \quad \text{при} \quad \xi_{(k)} < x \leq \xi_{(k+1)},$$

где $k=0, 1, \dots, N$, и условно $\xi_{(0)} = -\infty, \xi_{(N+1)} = \infty$ (рис. 13). Далее условимся для краткости писать F_k вместо $F(\xi_{(k)})$. Тогда из (18) следует, что

$$\begin{aligned} \frac{\omega_N^2}{N} &= \sum_{k=0}^N \int_{\xi_{(k)}}^{\xi_{(k+1)}} \left[F(x) - \frac{k}{N} \right]^2 dF(x) = \\ &= \sum_{k=0}^N \left[\frac{F^3}{3} - \frac{2k}{N} \frac{F^2}{2} + \frac{k^2}{N^2} F \right]_{x=\xi_{(k)}}^{x=\xi_{(k+1)}} = \\ &= \sum_{k=0}^N \left[\frac{1}{3} (F_{k+1}^3 - F_k^3) - \frac{k}{N} (F_{k+1}^2 - F_k^2) + \frac{k^2}{N^2} (F_{k+1} - F_k) \right] = \\ &= \sum_{k=0}^N \left\{ \left[\frac{F_{k+1}^3}{3} - \frac{k F_{k+1}^2}{N} + \frac{k^2 F_{k+1}}{N^2} \right] - \right. \\ &\quad \left. - \left[\frac{F_k^3}{3} - \frac{(k-1) F_k^2}{N} + \frac{(k-1)^2 F_k}{N^2} \right] + \frac{1}{N} F_k^2 - \frac{2k-1}{N} F_k \right\}. \end{aligned}$$

Сумма первых двух квадратных скобок легко вычисляется. Последние члены дополним до полного квадрата:

$$\frac{\omega_N^2}{N} = \frac{1}{3} + \frac{1}{N} \sum_{k=0}^N \left[F_k^2 - 2 \frac{k-1/2}{N} F_k + \left(\frac{k-1/2}{N} \right)^2 - \left(\frac{k-1/2}{N} \right)^2 \right].$$

После несложных вычислений получим окончательную формулу, которую удобно записать в следующем виде:

$$\omega_N^2 = \frac{1}{12N} + \sum_{k=1}^N \left[F(\xi_{(k)}) - \frac{k-1/2}{N} \right]^2. \quad (20)$$

3.2. Проверка таблиц случайных цифр. Для проверки таблицы случайных цифр $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_N$ М. Г. Кендалл и Б. Б. Смит предложили использовать четыре теста. В каждом из них цифры классифицируются по некоторому признаку и эмпирические частоты сравниваются с их математическими ожиданиями при помощи критерия χ^2 . Тесты эти: 1) *проверка частот* (frequency test) — проверяется частота различных цифр в таблице; 2) *проверка пар* (serial test) — проверяется частота различных двузначных чисел среди пар $\varepsilon_1 \varepsilon_2, \varepsilon_2 \varepsilon_3, \varepsilon_3 \varepsilon_4, \dots, \varepsilon_{N-1} \varepsilon_N$; 3) *проверка интервалов* (gap test) — проверяется частота различных интервалов между двумя последовательными нулями; 4) *проверка комбинаций* (poker test) — проверяется частота различных типов четверок ($abcd, aabc, aabb, aaab, aaaa$) среди четверок $\varepsilon_1 \varepsilon_2 \varepsilon_3 \varepsilon_4, \varepsilon_2 \varepsilon_3 \varepsilon_4 \varepsilon_5, \dots$

Большинство последующих авторов также использовали эту систему тестов, внося, однако, в нее некоторые изменения, среди которых отметим два. Вместо проверки интервалов обычно используют *проверку серий* (run test): цифры $\varepsilon_{k+1}, \varepsilon_{k+2}, \dots, \varepsilon_{k+l}$ образуют серию длины l , если $\varepsilon_{k+1} = \varepsilon_{k+2} = \dots = \varepsilon_{k+l}$, но $\varepsilon_k \neq \varepsilon_{k+1}$, $\varepsilon_{k+l+1} \neq \varepsilon_{k+l}$. Вместо проверки пар $\varepsilon_1 \varepsilon_2, \varepsilon_2 \varepsilon_3, \varepsilon_3 \varepsilon_4, \dots$ проверяют *независимые пары* $\varepsilon_1 \varepsilon_2, \varepsilon_3 \varepsilon_4, \varepsilon_5 \varepsilon_6, \dots$

С. детерминистической точки зрения (она изложена в гл. 7) проверка частот и проверка независимых пар — важнейшие необходимые тесты. Проверка серий обобщает «критерий случайности», часто используемый в статистике [24]. И только проверка комбинаций носит несколько искусственный характер. Поэтому можно рекомендовать в качестве основных использовать следующие два теста.

Первый тест (*проверка частот и пар*). Предположим, что количество цифр $N = 2N'$ четное. Цифры $\varepsilon_1, \dots, \varepsilon_N$ разобьем на пары $\varepsilon_1 \varepsilon_2, \varepsilon_3 \varepsilon_4, \dots, \varepsilon_{N-1} \varepsilon_N$. Обозначим через v_{ij} количество пар ij . Сосчитав матрицу (v_{ij}) , $0 \leq i, j \leq 9$, легко вычислить также величины

$$v_{i\cdot} = \sum_{j=0}^9 v_{ij}, \quad v_{\cdot j} = \sum_{i=0}^9 v_{ij}, \quad v_i = v_{i\cdot} + v_{\cdot i}.$$

Смысл этих величин очевиден: $v_{i\cdot}$ — количество цифр,

равных i , среди цифр $\varepsilon_1, \varepsilon_3, \varepsilon_5, \dots, \varepsilon_{N-1}$; v_i — количество цифр, равных i , среди цифр $\varepsilon_2, \varepsilon_4, \dots, \varepsilon_N$; v_i — общее количество цифр, равных i .

По значениям v_{ij} , v_i , v_i и v_j можно сосчитать величины χ^2_N для нескольких проверок. Самые важные среди них (индексы N или N' обычно не пишут):

$$\chi^2 = \frac{10}{N} \sum_{i=0}^9 (v_i - 0,1N)^2$$

с 9 степенями свободы, соответствующая проверке частот, и

$$\chi^2 = \frac{100}{N'} \sum_{i,j=0}^9 (v_{ij} - 0,01N')^2$$

с 99 степенями свободы, соответствующая проверке пар. Кроме того, иногда вычисляют величины

$$\chi^2 = \frac{1}{0,1N'} \sum_{i=0}^9 (v_i - 0,1N')^2 \text{ и } \chi^2 = \frac{1}{0,1N'} \sum_{j=0}^9 (v_j - 0,1N')^2,$$

каждая с 9 степенями свободы; эти величины соответствуют проверке частот среди цифр с нечетными и четными номерами.

Используют также критерий независимости строк и столбцов матрицы v_{ij} , которому соответствует величина

$$\chi^2 = N' \left[\sum_{i,j=0}^9 \frac{v_{ij}^2}{v_i \cdot v_j} - 1 \right]$$

с 81 степенью свободы [44]. Впрочем, последний критерий на практике оказывается весьма близким к проверке пар.

Второй тест (проверка серий). Сосчитаем количества n_l серий длины l в последовательности цифр $\varepsilon_1, \dots, \varepsilon_N$ при $l=1, 2, \dots, m$, и пусть n'_{m+1} — количество серий с $l \geq m+1$ (они объединяются в одну группу). Обозначим общее количество серий через $n = n_1 + \dots + n_m + n'_{m+1}$. Величина χ^2 с m степенями свободы вычисляется по формуле

$$\chi^2 = \sum_{l=1}^m \frac{(n_l - np_l)^2}{np_l} + \frac{(n'_{m+1} - np'_{m+1})^2}{np'_{m+1}},$$

где $p_l = 9 \cdot 10^{-l}$, $p'_{m+1} = 10^{-m}$.

Обозначим через λ случайную длину серии, начинающейся с цифры ε_{k+1} (т. е. $\varepsilon_{k+1} \neq \varepsilon_k$). Очевидно, $p_l P\{\lambda = l\} = P\{\varepsilon_{k+1} = \varepsilon_{k+2} = \dots = \varepsilon_{k+l} \neq \varepsilon_{k+l+1}\} = (0,1)^{l-1} 0,9 = 9 \cdot 10^{-l}$ и от номера серии не зависит. Таким образом, каждая серия может с вероятностью p_l иметь длину l . Количество n_l серий длины l среди n серий подчиняется биномиальному распределению и $Mn_l = np_l$.

В качестве примера в табл. 2 приведено распределение по длине серий, полученных при проверке $N=50\,000$ чисел в [75].

Таблица 2

	n_1	n_2	n_3	n_4	n
Теоретические значения	40 500	4050	405	45	45 000
Эмпирические значения	40 568	4010	417	40	45 035
np_l	40 531,5	4053,2	405,3	45,0	—

По данным последних двух строк вычисляем $\chi^2=1,39$. При $m=3$ этому значению соответствует $P=0,71$.

Дополнительные тесты. В качестве дополнительных тестов следует в первую очередь рекомендовать проверку частот независимых троек $\varepsilon_{3h+1}\varepsilon_{3h+2}\varepsilon_{3h+3}$, затем четверок $\varepsilon_{4h+1}\varepsilon_{4h+2}\varepsilon_{4h+3}\varepsilon_{4h+4}$ и т. д.—основания для такой рекомендации приведены в § 4 гл. 7. Так как количество различных троек, четверок и т. д. весьма велико (10^3 , 10^4 , ...), то подсчет всех типов групп имеет смысл проводить лишь для достаточно больших N . При меньших N группы можно объединять по различным признакам. Например, при упомянутой выше проверке комбинаций четверки $\varepsilon_{h+1} \dots \varepsilon_{h+4}$ классифицируются по количеству совпадающих цифр в группе (признак, напоминающий комбинации при игре в покер).

Во многих работах предлагаются различные дополнительные тесты. Имеет место даже чрезмерное увлечение такими тестами. Несмотря на справедливость принципа «каши маслом не испортишь», надо четко представлять себе, что все такие тесты только необходимы. Положительный результат любого теста означает только, что этот результат не противоречит гипотезе

о случайности цифр $\varepsilon_1, \dots, \varepsilon_N$, но, может быть, какой-нибудь другой тест эту гипотезу опровергнет...*)).

3.3. Проверка псевдослучайных чисел. В качестве основных тестов для проверки псевдослучайных чисел $\gamma_1, \dots, \gamma_N$ используют те же тесты, что и для проверки таблиц: проверяются первые десятичные цифры $\varepsilon_k = \Pi(10 \gamma_k)$. Фигурирующая в первом тесте величина v_{ij} равна количеству пар $(\gamma_{2k-1}, \gamma_{2k})$ таких, что

$$0,1 i \leq \gamma_{2k-1} < 0,1(i+1), \quad 0,1 j \leq \gamma_{2k} < 0,1(j+1).$$

Более детальную проверку распределения чисел $\gamma_1, \dots, \gamma_N$ можно осуществить с помощью критерия ω^2 . Если расположить эти числа в вариационный ряд

$$\gamma_{(1)} \leq \gamma_{(2)} \leq \dots \leq \gamma_{(N)},$$

то из (20) вытекает, что

$$\omega_N^2 = \frac{1}{12N} + \sum_{k=1}^N \left(\gamma_{(k)} - \frac{k-1/2}{N} \right)^2.$$

Однако, как уже отмечалось в п. 3.1.4, при очень больших N построение вариационного ряда весьма трудоемко.

Естественно, что при проверке псевдослучайных чисел всегда используют различные дополнительные тесты для проверки последующих десятичных цифр γ_k . Иногда десятичные (или двоичные) цифры чисел γ_k выделяют и проверяют независимо. Конечно, лучше было бы использовать тесты п. 3.2 с более мелким разбиением, но уже при совместной проверке двух десятичных цифр пришлось бы вычислять матрицу (v_{ij}) размером 100×100 . Поэтому на практике ограничиваются более простыми проверками, и далекие цифры чисел γ_k обычно оказываются хуже проверенными.

В пользу указанной системы тестов можно привести аргумент практического характера: в литературе, пожалуй, нет примеров, когда числа, удовлетворяющие всем

*) Более того, имея конечную группу цифр $\varepsilon_1, \dots, \varepsilon_N$, всегда можно придумать такой тест, зависящий от конкретных значений $\varepsilon_1, \dots, \varepsilon_N$, который опровергает нашу гипотезу. Но не надо такие тесты придумывать!...

тестам, оказались бы непригодными для решения конкретной задачи (в которой не предъявлялись повышенные требования к точности решения); есть, однако, примеры неудачных расчетов с помощью чисел, которые не удовлетворяли одному из тестов. В частности, из-за того, что числа, упомянутые в п. 2.2.4, не удовлетворяли тесту проверки пар, плохо моделировалось нормальное распределение.

Таким образом, строго говоря, разумность приведенной системы тестов — факт эмпирический. В действительности эти тесты не гарантируют универсальной пригодности чисел. Поэтому иногда целесообразно вводить дополнительные тесты, связанные с характером решаемых задач. Впрочем, успешное решение нужной задачи — самая лучшая проверка случайных чисел.

3.4. О проверке датчиков случайных чисел. Для проверки чисел, выдаваемых датчиком, можно использовать те же тесты, что и для проверки псевдослучайных чисел. Особенность проверки датчиков в том, что проверяются не те числа, которые используются в счете. Поэтому, кроме проверки «качества» выдаваемых чисел, должна еще как-то гарантироваться устойчивость работы датчика.

Мы не будем останавливаться на технических способах контроля. Отметим только, что обычно расчет тестов, вроде рассмотренных в пп. 3.2 и 3.3, проводят и до, и после расчета нужной задачи, если она не слишком продолжительна во времени.

3.5. О проверке больших массивов случайных чисел. Так как в расчетах часто используется не вся таблица, а только часть ее, то имеет смысл проверить также части таблицы. Например, таблицу, содержащую N цифр, разбивают на s частей по $N_1 = N/s$ цифр и проверяют каждую из этих частей. Если $s > 50$, то вполне вероятно среди значений $\chi^2_{N_1}$, соответствующих одному (какому-нибудь) критерию, встретить значение, отвечающее $P < 0,01$. Это не означает, что вся таблица плоха: просто данную часть не надо использовать как самостоятельную таблицу в расчетах, в которых требуется $\approx N_1$ случайных цифр.

Можно проверить, согласуются ли s полученных значений $\chi^2_{N_1}$ с теоретическим распределением χ^2 для данного критерия. Такая проверка нередко используется в качестве дополнительного теста. И с этой точки зрения ясно, что отдельные «плохие» значения $\chi^2_{N_1}$ обязаны появляться, если гипотеза справедлива и s достаточно велико.

Некоторые авторы предлагают использовать такую проверку таблицы по частям вместо проверки таблицы как целого. К сожалению, это неверная рекомендация. Автору известен пример, когда 500 000 цифр проверялись и как целая таблица и по частям: 20 частей по 25 000 цифр. Все части удовлетворяли всем пяти критериям первого теста (п. 3.2), распределение полученных значений $\chi^2_{N_1}$ для каждого из критериев хорошо согласовалось с теоретическим распределением χ^2 . Однако весь массив из 500 000 цифр не удовлетворял ни одному из критериев первого теста (причем для одного из критериев было $P < 0,00001!$)

Такому несколько неожиданному результату можно дать следующее объяснение. Можно предположить, что рассматриваемая последовательность цифр имеет распределение, несколько отличающееся от равномерного, но для расчета по N_1 цифрам такое отличие допустимо. Отсюда не следует, что это отличие по-прежнему будет допустимым для расчета, в котором используются $N \gg N_1$ цифр. Наоборот, если увеличение количества цифр вызвано желанием увеличить точность расчета, то ясно, что при достаточно большом N отличие это станет недопустимым

Итак, к большим массивам случайных цифр (или чисел) предъявляются большие требования. Этот факт хорошо известен практикам: подобрать хорошую псевдослучайную последовательность, содержащую 10^3 — 10^4 чисел, гораздо проще, чем подобрать такую же последовательность, содержащую 10^5 — 10^6 чисел *).

Это следует иметь в виду также при проверке датчиков случайных чисел: нехорошо проверять датчик тестами, охватывающими 10^3 — 10^4 чисел, а в расчете использовать 10^6 — 10^7 чисел.

Упражнения к главе I

1. Доказать, что коэффициент корреляции случайных величин γ и $\eta = D(g\gamma)$ при целом $g \geq 1$ равен $1/g$.

2. Как с помощью таблицы случайных цифр (стр. 295) моделировать последовательность независимых испытаний, в каждом из которых вероятность наступления некоторого события A равна $P\{A\} = 0,752$?

3. Предположим, что количество ν элементарных частиц, попадающих в детектор за время Δt , подчиняется распределению Пуассона со средним значением $M\nu = a\Delta t$. Пусть $\xi = \nu \pmod{2}$. Доказать, что при $a\Delta t \rightarrow \infty$ вероятности $P\{\xi=0\}$ и $P\{\xi=1\}$ стремятся к $1/2$. На этом принципе построены некоторые датчики случайных цифр.

4. Рассмотреть случайную величину $\eta = D(g\gamma^2)$ с целым g и доказать, что $P\{\eta < x\} = x + \sqrt{x/g} + O(x/\sqrt{g})$ при $x > 0$ и $g \rightarrow \infty$ (К. Д. Точер [174]).

*) Иногда высказывают мнение, что критерии χ^2 или ω^2 слишком жестки при очень больших N . Однако, как показывают формулы (27) и (30) гл. 7, именно постоянство χ^2_N или ω^2_N регулирует упомянутое в тексте повышение требований к большим массивам случайных чисел, если требовать, чтобы погрешности, фигурирующие в этих формулах, убывали как $1/\sqrt{N}$.

Из этой формулы следует, что хотя $P\{\eta < x\} \rightarrow x$ при $g \rightarrow \infty$ тем не менее при $g=10^k$ и $x=10^{-k}$ вероятность $P\{\eta < 10^{-k}\} \approx 2 \cdot 10^{-k}$, что в два раза больше, чем при равномерном распределении. Так как алгоритм $\gamma_{n+1} = D(10^k \gamma_n^2)$ близок к алгоритму середины квадрата: у произведения γ_n^2 отбрасываются k старших цифр слева, то этот результат в какой-то мере объясняет, почему в методе середины квадрата получается больше, чем надо, малых чисел.

5. Доказать, что если M с g и M с m_0 взаимно простые, то из формулы (7) следует, что

а) $m_n = g^n m_0 \pmod{M}$;

б) для последовательности m_0, m_1, \dots, m_n всегда $L = P$;

в) длина периода P равна наименьшему целому корню сравнения $g^P - 1 \equiv 0 \pmod{M}$.

6. Обобщить теорему 3 на случай метода возмущений (12):

$$\lim_{N \rightarrow \infty} P\{L / \sqrt{NM} < x\} = 1 - e^{-x^2/2}.$$

У к а з а н и е. Предположить, что при всех возможных аргументах x функции $\Phi(x)$ и $\Psi(x)$ не равны и использовать вероятностную модель п. 2.3. Роль L играет номер опыта, в котором впервые $\gamma_L = \gamma_l$, $l < L$, и в то же время $L - l \equiv 0 \pmod{M}$ (И. М. Соболев [79]).

7. Доказать, что если последовательность $\alpha_0, \alpha_1, \dots, \alpha_n \dots$ удовлетворяет моноциклическому уравнению (13), то при каждом s таком, что $1 \leq s \leq r$, среди групп вида $(\alpha_n, \dots, \alpha_{n+s-1})$ при $0 \leq n \leq P-1$ встречаются по 2^{r-s} раз все возможные группы, состоящие из нулей и единиц, кроме группы $(0, \dots, 0)$, которая встречается $2^{r-s} - 1$ раз. (Н. Цирлер [185]).

У к а з а н и е. Доказать сначала это для $s=r$.

8. В качестве меры отклонения $F_N^*(x)$ от $F(x)$ (см п. 3.1.4) можно использовать величину

$$D = \sup_{-\infty < x < \infty} |F_N^*(x) - F(x)|.$$

Критерий Колмогорова, близкий к критерию ω^2 и в некоторых случаях более удобный, основан на теореме А. Н. Колмогорова.

Теорема А. Какова бы ни была случайная величина ξ с непрерывной функцией распределения $F(x)$, при каждом $x > 0$

$$\lim_{N \rightarrow \infty} P\{\sqrt{N} D < x\} = K(x),$$

$$\text{где } K(x) = 1 + 2 \sum_{k=1}^{\infty} (-1)^k e^{-2k^2 x^2}.$$

(Таблица функции распределения Колмогорова $K(x)$ имеется на стр. 293)

Доказать, что для расчета D можно использовать формулу

$$D = \max_{1 \leq k \leq N} \left[\left| F(\xi_{(k)}) - \frac{k-1}{N} \right|; \left| \frac{k}{N} - F(\xi_{(k)}) \right| \right].$$

ГЛАВА 2

ПРЕОБРАЗОВАНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

В этой главе рассмотрены преобразования, позволяющие с помощью случайных чисел γ вычислять значения любой случайной величины ξ . Такие вычисления называют *моделированием* случайной величины ξ или *формированием реализаций* случайной величины ξ . Впрочем, вычислители предпочитают их называть *разыгрыванием* величины ξ . Владение этими преобразованиями необходимо каждому, кто желает овладеть «техникой» применения методов Монте-Карло.

Автор не старался изложить здесь максимальное количество рецептов для моделирования различных величин — это сделано в специальной литературе (см. [19, 69, 111]). Упор сделан на изложение методов и их систематизацию. Впервые в основу классификации преобразований положено количество случайных чисел, используемых при расчете одного значения ξ . Поначалу этот принцип может показаться несколько искусственным, но в полной мере его роль выяснится в гл. 7.

Некоторые из общих методов, наиболее важные с точки зрения практики, сформулированы в виде теорем. Замечания практического характера см. в п. 5.7.

§ 1. Метод обратных функций (основной прием моделирования случайных величин)

1.1. Моделирование дискретных случайных величин. Рассмотрим дискретную случайную величину ξ с распределением

$$\begin{pmatrix} x_1 & x_2 & \dots & x_n \\ p_1 & p_2 & \dots & p_n \end{pmatrix}, \quad (1)$$

где $p_i = P\{\xi = x_i\}$. Для того чтобы вычислить значения этой величины разделим интервал $0 \leq y < 1$ на интервалы Δ_i такие (рис. 14), что длина Δ_i равна p_i .

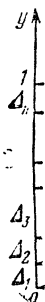
Теорема 1. Случайная величина ξ , определенная формулой

$$\xi = x_i, \text{ когда } \gamma \in \Delta_i, \quad (2)$$

имеет распределение вероятностей (1).

Доказательство занимает одну строку:

$$P\{\xi = x_i\} = P\{\gamma \in \Delta_i\} = \text{длина } \Delta_i = p_i.$$



Для практической реализации формулы (2) Рис. 14. удобно в накопителе ЭВМ расположить подряд значения x_1, x_2, \dots, x_n и $p_1, p_1 + p_2, p_1 + p_2 + p_3, \dots, 1$. Для того чтобы вычислить очередное значение ξ , находим очередное γ . Затем сравниваем γ с p_1 . Если $\gamma < p_1$, то $\xi = x_1$; если $\gamma \geq p_1$, то сравниваем γ с $p_1 + p_2$. Если $\gamma < p_1 + p_2$, то $\xi = x_2$; если $\gamma \geq p_1 + p_2$, то сравниваем γ с $p_1 + p_2 + p_3$, и т. д.

1.1.1. Легко видеть, что в случае, когда $\xi = x_i$ ($1 \leq i \leq n-1$), приходится осуществить i сравнений, и лишь в случае, когда $\xi = x_n$, число сравнений равно $n-1$. Поэтому среднее число сравнений, затрачиваемых при получении одного значения ξ , равно

$$t = \sum_{i=1}^{n-1} i p_i + (n-1) p_n.$$

Так как порядок значений x_1, \dots, x_n в (1) произволен, то выгодно расположить их в порядке убывания вероятностей, т. е. так, чтобы $p_1 \geq p_2 \geq \dots \geq p_n$. Тогда величина t будет минимальной (Ю. Г. Полляк [69]).

1.1.2. Расчет по формуле (2) заметно упрощается в случае, когда все значения x_1, \dots, x_n равновероятны: $p_1 = \dots = p_n = 1/n$. В этом случае многократные сравнения не нужны: так как Δ_i — это интервал $(i-1)/n \leq y < i/n$, то условие $\gamma \in \Delta_i$ равносильно условию $i-1 \leq n\gamma < i$, или $C(n\gamma) = i-1$. Вместо формулы (2) можно записать, что

$$\xi = x_i, \text{ где } i = 1 + C(n\gamma).$$

1.1.3. Теорему 1 легко обобщить на случайную величину, которая может принимать бесконечную последовательность значений $x_1, x_2, \dots, x_n, \dots$ и имеет распределение

$$\begin{pmatrix} x_1 & x_2 & \dots & x_n & \dots \\ p_1 & p_2 & \dots & p_n & \dots \end{pmatrix}.$$

В этом случае числа x_n и p_n задаются формулами, и вычисление их при каждом расчете ξ может оказаться весьма трудоемким. Тогда можно выбрать число n_0 так, чтобы сумма вероятностей $p_1 + \dots + p_{n_0}$ была достаточно близкой к 1, и значения x_1, \dots, x_{n_0} и p_1, \dots, p_{n_0} заготовить заранее. Вычислять x_i и p_i по формулам придется только при $i > n_0$, а это будет достаточно редко.

1.2. Моделирование случайных событий. Моделирование случайных событий сводится к моделированию дискретных случайных величин. Чтобы показать, как это делается, рассмотрим четыре задачи, в каждой из которых требуется моделировать последовательность одинаковых независимых испытаний.

1.2.1. В каждом из испытаний может наступить или не наступить некоторое событие A , вероятность наступления которого $P\{A\} = p$ задана.

Рассмотрим случайную величину ξ , называемую индикатором события A , которая равна 1 при наступлении A и 0 при наступлении противоположного события \bar{A} . Распределение ξ задается таблицей

$$\begin{pmatrix} 1 & 0 \\ p & 1 - p \end{pmatrix}.$$

Согласно теореме 1 для осуществления каждого испытания надо найти случайное число γ и проверить неравенство $\gamma < p$. Если оно выполнено, то событие A в этом испытании произошло, а если $\gamma \geq p$, то нет.

1.2.2. С испытанием связана полная группа попарно несовместных событий*) A_1, \dots, A_n и заданы вероятности $P\{A_i\} = p_i$.

Для моделирования таких испытаний рассмотрим случайную величину ξ — номер наступившего события.

*) Это значит, что сумма $A_1 + \dots + A_n$ есть достоверное событие и $A_i \cdot A_j = 0$ при $i \neq j$.

Очевидно, распределение ξ выражается таблицей

$$\begin{pmatrix} 1 & 2 & \dots & n \\ p_1 & p_2 & \dots & p_n \end{pmatrix}.$$

Для осуществления каждого испытания надо выбрать случайное число γ и по теореме 1 разыграть значение ξ . Если $\xi=i$, то произошло событие A_i .

Пример. Столкнувшись с ядром атома урана, нейтрон может рассеяться, быть захваченным или вызвать деление ядра. Если через Σ_s , Σ_c , Σ_f обозначить соответствующие этим событиям сечения взаимодействия, а через $\Sigma = \Sigma_s + \Sigma_c + \Sigma_f$ — полное сечение взаимодействия нейтрона с ядром, то вероятности трех возможных событий равны соответственно Σ_s/Σ , Σ_c/Σ и Σ_f/Σ . Чтобы разыграть «судьбу» нейтрона при столкновении, выбирают случайное число γ ; если $\gamma < \Sigma_s/\Sigma$, то считают, что нейтрон рассеялся; если $\Sigma_s/\Sigma \leq \gamma < (\Sigma_s/\Sigma) + (\Sigma_c/\Sigma)$, то нейтрон поглощен; если $(\Sigma_s/\Sigma) + (\Sigma_c/\Sigma) \leq \gamma$, то нейтрон вызвал деление ядра.

1.2.3. С испытанием связаны два независимых совместных события A и B , вероятности которых заданы: $P\{A\} = p_A$, $P\{B\} = p_B$.

Ввиду независимости событий A и B можно последовательно моделировать их наступление в каждом испытании: сперва по числу γ_1 методом п. 1.2.1 определить, наступило ли событие A , а затем точно также по числу γ_2 определить, наступило ли событие B .

Однако часто более экономен другой способ. Рассмотрим полную группу попарно несовместных событий, состоящую из четырех событий:

$$A_1 = AB, A_2 = A\bar{B}, A_3 = \bar{A}B, A_4 = \bar{A}\bar{B}. \quad (3)$$

Вероятности этих событий легко вычислить:

$$p_1 = p_A p_B, \quad p_2 = p_A (1 - p_B),$$

$$p_3 = p_B (1 - p_A), \quad p_4 = (1 - p_A) (1 - p_B).$$

Следовательно, метод п. 1.2.2 позволяет, используя одно случайное число γ , определить, какой из этих четырех исходов наступил в моделируемом испытании.

1.2.4. С испытанием связаны два зависимых совместных события A и B , и заданы вероятности $P\{A\} = p_A$, $P\{B\} = p_B$, $P\{AB\} = p_{AB}$.

В этом случае также следует рассмотреть полную группу событий (3), только вероятности этих событий вычисляются иначе:

$$p_1 = p_{AB}, \quad p_2 = p_A - p_{AB}, \quad p_3 = p_B - p_{AB}, \quad p_4 = 1 - p_A - p_B + p_{AB}.$$

Впрочем, и в этом случае можно осуществить последовательное моделирование событий A и B , используя два случайных числа

γ_1, γ_2 . Сперва по числу γ_1 (методом п. 1.2.1) определяем, наступило ли событие A . Если A наступило, то, зная условную вероятность $P\{B/A\} = p_{AB}/p_A$, можно по числу γ_2 определить, наступило ли событие B ; условием наступления B служит выполнение неравенства $\gamma_2 < P\{B/A\}$. Если же событие A не наступило, то наступление B придется разыгрывать с помощью условной вероятности $P\{B/\bar{A}\}$, которая равна $P\{B/\bar{A}\} = (p_B - p_{AB})/(1 - p_A)$.

1.3. Моделирование непрерывных случайных величин. Предположим, что случайная величина ξ определена в интервале $a < x < b$ и имеет плотность $p(x) > 0$ при $a < x < b$. Обозначим через $F(x)$ функцию распределения ξ , которая при $a < x < b$ равна

$$F(x) = \int_a^x p(u) du.$$

Случай $a = -\infty$ и (или) $b = \infty$ не исключается.

Теорема 2. Случайная величина ξ , удовлетворяющая уравнению.

$$F(\xi) = \gamma, \quad (1)$$

имеет плотность распределения $p(x)$.

Доказательство. Так как функция $F(x)$ строго возрастает в интервале (a, b) от $F(a) = 0$ до $F(b) = 1$, то уравнение (4) имеет единственный корень при каждом γ (рис. 15). При этом равны вероятности

$$\begin{aligned} P\{x < \xi < x+dx\} &= \\ &= P\{F(x) < \gamma < F(x+dx)\}. \end{aligned}$$

И так как случайная величина γ равномерно распределена в интервале $(0, 1)$, то

$$\begin{aligned} P\{x < \xi < x+dx\} &= \\ &= F(x+dx) - F(x) = p(x) dx, \end{aligned}$$

что и требовалось доказать.

В тех случаях, когда уравнение (4) аналитически разрешимо относительно ξ , получается явная формула $\xi = G(\gamma)$ для разыгрывания случайной величины ξ , где

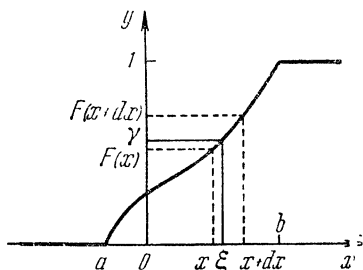


Рис. 15.

$G(y)$ — обратная функция по отношению к $y=F(x)$. В других случаях можно уравнение (4) решать численно. Если объем накопителя позволяет, то удобно составить таблицу функции $G(y)$, $0 < y < 1$, и по ней находить значения ξ . Иногда удобно использовать таблицу функции $F(x)$, $a < x < b$, и находить значения ξ обратной интерполяцией. О некоторых приемах составления таблиц см. [8, 9, 90].

Пример Экспоненциальная случайная величина ξ определена при $x_0 < x < \infty$ с плотностью

$$p(x) = ae^{-a(x-x_0)}.$$

Так как

$$F(x) = \int_{x_0}^x ae^{-a(u-x_0)} du = 1 - e^{-a(x-x_0)},$$

то уравнение (4) примет вид

$$1 - e^{-a(\xi - x_0)} = \gamma.$$

Отсюда получаем явное выражение для расчета ξ

$$\xi = x_0 - (1/a) \ln(1-\gamma). \quad (5)$$

1.4. Метод обратных функций. Теоремы 1 и 2 представляют собой частные случаи общего метода, который естественно назвать *методом обратных функций* (наряду с названием *inverse functions method* встречается также *direct method*).

Рассмотрим произвольную случайную величину ξ с функцией распределения $F(x) = P\{\xi \leq x\}$ (рис. 16, а). Обратную по отношению к $F(x)$ функцию $G(y)$ определим следующим образом. Во-первых, дополним график функции $y=F(x)$ вертикальными отрезками в точках разрыва до непрерывной линии $y=F_0(x)$ (рис. 16, б); функция $y=F_0(x)$, вообще говоря, неоднозначна. Эту же линию можно записать уравнением вида $x=G_0(y)$, где функция $G_0(y)$ опять-таки не обязана быть однозначной: интервалам постоянства $F_0(x)$ соответствуют вертикальные отрезки $G_0(y)$ и наоборот. Положим $G(y)=G_0(y)$ в точках непрерывности и $G(y)=G(y+0)$ в точках разрыва (рис. 16, в).

Построенная таким образом однозначная функция $G(y)$ не убывает при $0 < y < 1$ и непрерывна справа во

всех точках *). Функции $F(x)$ и $G(y)$ связаны следующим свойством: $G(y) < x$ тогда и только тогда, когда $y < F(x)$. Для доказательства этого свойства достаточно проверить, что каждое из двух неравенств $G(y) < x$ и $y < F(x)$ означает, что точка $(G(y), y)$ расположена на линии $y = F_0(x)$ одновременно и левее и ниже точки $(x, F(x))$ (рис. 17).

Теорема 3. Случайная величина

$$\xi = G(\gamma) \quad (6)$$

имеет функцию распределения $F(x)$.

Для доказательства теоремы нужно вычислить функцию распределения случайной величины ξ , определенной формулой (6):

$$\begin{aligned} P\{\xi < x\} &= P\{G(\gamma) < x\} = \\ &= P\{\gamma < F(x)\} = F(x). \end{aligned}$$

То, что теорема 1 предоставляет собой частный случай теоремы 3, видно из сравнения рис. 18, на котором изображена функция $x = G_0(y)$, соответствующая дискретной случайной величине, с рис. 14: если $\gamma \in \Delta_i$ на рис. 14, то $G_0(\gamma) = x_i$ на рис. 18.

В условиях теоремы 2

функция $G(y)$ совпадает с обычной обратной функцией к $F(x)$ и уравнение (6) равносильно (4).

*) Аналитическое определение: функция $G(y)$ равна нижней грани множества чисел x , для которых $y < F(x)$, т. е.

$$G(y) = \inf_{\{x | y < F(x)\}} x.$$

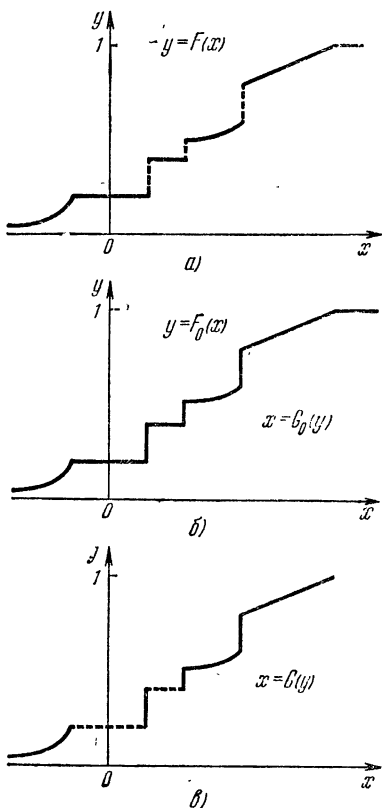


Рис. 16.

Заметим, что так как случайная величина $1-\gamma$ имеет то же распределение, что γ , то в формулах (2), (4), (6) можно вместо γ написать $1-\gamma$. Следовательно, указанные способы моделирования не единственно возможные. Иногда замена γ на $1-\gamma$ несколько упрощает формулы расчета. Например, вместо формулы (5) можно использовать формулу

$$\xi = x_0 - (1/a) \ln \gamma. \quad (7)$$

Итак, метод обратных функций позволяет записать формулы для моделирования любой случайной величины ξ . Но нередко этот метод приводит к сложным или

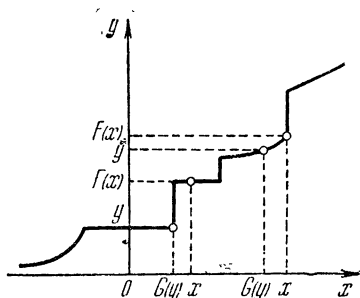


Рис. 17.

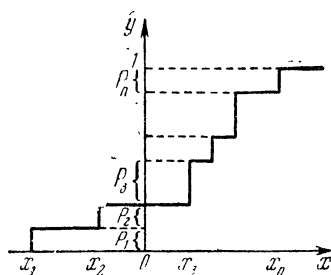


Рис. 18.

просто неудобным алгоритмам. Например, для того чтобы вычислять значения гауссовской (нормальной) случайной величины ξ с параметрами $(0; 1)$, приходится решать уравнение

$$\int_{-\infty}^{\xi} e^{-t^2/2} dt = \sqrt{2\pi} \gamma.$$

В таких случаях обычно прибегают к помощи других методов моделирования, связанных с другими преобразованиями случайных чисел γ .

1.5. Преобразования вида $\xi = g(\gamma)$. Попытаемся найти всевозможные функции $g(y)$ такие, что случайная величина $g(\gamma)$ имеет функцию распределения $F(x)$. Для этого необходимо и достаточно, чтобы $P\{g(\gamma) < x\} = F(x)$.

Введем единичную функцию

$$e(z) = \begin{cases} 0 & \text{при } z \leq 0, \\ 1 & \text{при } z > 0. \end{cases}$$

Так как плотность $p_Y(y) = 1$ при $0 < y < 1$, то

$$P\{g(Y) < x\} = \int_{\{y | g(y) < x\}} p_Y(y) dy = \int_0^1 e(x - g(y)) dy.$$

Таким образом, получаем уравнение, которому должна удовлетворять функция $g(y)$:

$$\int_0^1 e(x - g(y)) dy = F(x). \quad (8)$$

Общее решение уравнения (8), по-видимому, неизвестно*). Однако легко указать частные классы функций $g(y)$, в которых решения существуют. Для простоты ограничимся случаем, когда моделируемая величина ξ принимает значения в интервале $a < x < b$ и имеет плотность вероятностей $p(x) > 0$ при $a < x < b$.

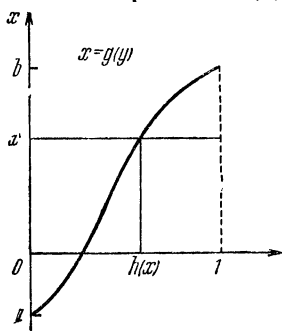


Рис. 19.

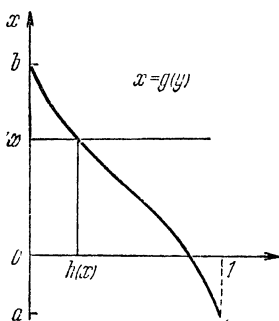


Рис. 20.

Пусть $g(y)$ строго возрастает при $0 < y < 1$ и $g(0) = a$, $g(1) = b$. Тогда из рис. 19 видно, что $e(x - g(y)) = 1$ при $0 < y < h(x)$, где $h(x)$ — функция, обратная по отношению к $g(y)$. Из (8) вытекает, что

$$F(x) = \int_0^{h(x)} dy = h(x). \quad (9)$$

Переходя к обратным функциям, запишем, что $g(y)$ равна $G(y)$ — обратной функции к $F(x)$. Мы пришли, таким образом, к методу обратных функций $\xi = G(Y)$.

*) Если моделируемая случайная величина ξ непрерывна, то плотность ее $p(x) = F'(x)$. Дифференцируя (8) с учетом того, что $e'(z) = \delta(z)$ — дельта-функция Дирака, получим вместо (8) уравнение

$$\int_0^1 \delta(x - g(y)) dy = p(x).$$

Пусть теперь $g(y)$ строго убывает при $0 < y < 1$ и $g(0) = b$, $g(1) = a$. Тогда из рис. 20 видно, что $e(x - g(y)) = 1$ при $h(x) < y < 1$, и из (8) вытекает, что

$$F(x) = \int_{h(x)}^1 dy = 1 - h(x). \quad (10)$$

Сделав замену переменной $y = h(x)$, получим соотношение $F(g(y)) = 1 - y$, откуда $g(y) = G(1 - y)$. Таким образом, в этом случае $\xi = G(1 - \gamma)$, и мы снова пришли к методу обратных функций с заменой y на $1 - y$.

Эти же решения уравнения (8) можно получить для любой случайной величины, если предположить, что $g(y)$ обладает свойствами обратных функций $G(y)$ в смысле п. 1.4. Помимо этих двух монотонных решений, существует бесконечное количество немонотонных решений. Однако используются они сравнительно редко. Пожалуй, единственный общий метод, основанный на использовании немонотонных функций $g(y)$, это модифицированный метод суперпозиции Г. А. Михайлова, изложенный в п. 3.3.3.

Прежде чем перейти к преобразованиям более общего вида, рассмотрим основные методы моделирования многомерных случайных величин.

§ 2. Моделирование многомерных случайных величин

2.1. Моделирование n -мерной случайной точки с независимыми координатами. Если координаты n -мерной случайной величины $Q = (\xi_1, \dots, \xi_n)$ независимы, то функция распределения

$$F_Q(x_1, \dots, x_n) = F_1(x_1) \dots F_n(x_n),$$

где $F_i(x_i)$ — функция распределения величины ξ_i . Естественно ожидать, что в этом случае можно моделировать каждую величину ξ_i независимо:

$$F_i(\xi_i) = \gamma_i, \quad i = 1, 2, \dots, n, \quad (11)$$

где $\gamma_1, \dots, \gamma_n$ — независимые случайные числа.

Действительно, так как γ_i независимы, то и ξ_i , определенные формулами (11), независимы. Поэтому их совместная функция распределения равна произведению

$$\begin{aligned} P\{\xi_1 < x_1, \dots, \xi_n < x_n\} &= \prod_{i=1}^n P\{\xi_i < x_i\} = \\ &= \prod_{i=1}^n F_i(x_i) = F_Q(x_1, \dots, x_n). \end{aligned}$$

Пример. Случайная точка Q с декартовыми координатами (ξ_1, \dots, ξ_n) равномерно распределена в n -мерном параллелепипеде $\Pi = \{a_i < x_i < b_i; i=1, 2, \dots, n\}$ (рис. 21 для $n=2$).

Плотность вероятностей точки Q постоянна в Π :

$$\rho_Q(x_1, \dots, x_n) = \begin{cases} c & \text{при } (x_1, \dots, x_n) \in \Pi, \\ 0 & \text{при } (x_1, \dots, x_n) \notin \Pi, \end{cases}$$

где $1/c = \prod_{i=1}^n (b_i - a_i)$ — объем Π (n -мерный объем). Интегрируя

ρ_Q по всем переменным, кроме x_i , легко получить, что плотность ξ_i равна

$$p_i(x_i) = \begin{cases} 1/(b_i - a_i) & \text{при } x_i \in (a_i, b_i), \\ 0 & \text{при } x_i \notin (a_i, b_i). \end{cases}$$

Следовательно, каждая из координат ξ_i равномерно распределена в интервале $a_i < x_i < b_i$, и координаты эти независимы.

Согласно (11) запишем уравнения

$$F_i(\xi_i) \equiv (\xi_i - a_i)/(b_i - a_i) = \gamma_i^1,$$

откуда вытекают явные формулы для расчета координат

$$\xi_i = a_i + \gamma_i(b_i - a_i), \quad i = 1, 2, \dots, n.$$

2.2. Моделирование n -мерной непрерывной случайной точки (с произвольными координатами). В общем случае, когда ξ_1, \dots, ξ_n

зависимы, их совместную плотность можно представить в виде произведения условных плотностей вероятностей этих величин:

$$\begin{aligned} \rho_Q(x_1, \dots, x_n) &= \\ &= p_1(x_1) p_2(x_2 | x_1) p_3(x_3 | x_1, x_2) \dots \\ &\dots p_n(x_n | x_1, \dots, x_{n-1}). \end{aligned}$$

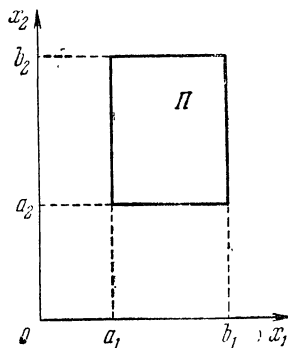


Рис. 21.

Все условные плотности вероятности выражаются через совместную плотность $\rho_Q(x_1, \dots, x_n)$.

Приводим выражения условных плотностей в общем виде; все интегралы берется от $-\infty$ до ∞ :

$$\begin{aligned} p_1(x_1) &= \int \dots \int \rho_Q dx_2 \dots dx_n, \\ p_2(x_2 | x_1) &= \int \dots \int \rho_Q dx_3 \dots dx_n [p_1(x_1)]^{-1}, \end{aligned}$$

Представление плотности $p_Q(x_1, \dots, x_n)$ в форме произведения условных плотностей координат ξ_1, \dots, ξ_n возможно $n!$ способами. В частности, при $n=2$

$$p_Q(x_1, x_2) = p_1(x_1) p_2(x_2 | x_1) = p_2(x_2) p_1(x_1 | x_2).$$

Разным произведениям соответствуют разные порядки разыгрывания величин ξ_1, \dots, ξ_n и, вообще говоря, разные уравнения (12). Нижеследующий пример показывает, что иногда удачный выбор порядка позволяет упростить эти уравнения.

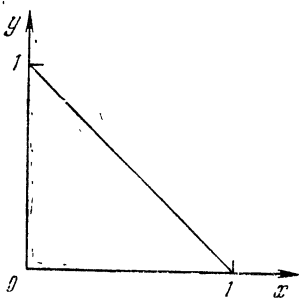


Рис. 22.

Если ξ_1, \dots, ξ_n независимы, то все их условные распределения равны безусловным $p_Q(x_1, \dots, x_n) = p_1(x_1) \dots p_n(x_n)$ и порядок разыгрывания величин роли не играет: уравнения (12) превращаются в (11).

Пример Рассмотрим случайную точку (ξ, η) , которая может принимать значения в треугольнике $x+y \leq 1$, $x > 0$, $y > 0$ (рис. 22) с плотностью $p(x, y) = 6x$.

а) Выберем в качестве первой величины ξ . Тогда

$$p_{\xi}(x) = \int_0^{1-x} p(x, y) dy = 6x(1-x) \quad \text{при } 0 < x < 1;$$

$$p_{\eta}(y | x) = p(x, y) / p_{\xi}(x) = (1-x)^{-1} \quad \text{при } 0 < y < 1-x.$$

Соответствующие этим плотностям функции распределения:

$$F_{\xi}(x) = \int_0^x p_{\xi}(u) du = 3x^2 - 2x^3 \quad \text{при } 0 < x < 1,$$

$$F_{\eta}(y | x) = \int_0^y p_{\eta}(v | x) dv = y(1-x)^{-1} \quad \text{при } 0 < y < 1-x.$$

Из формул (12) получаем уравнения для последовательного вычисления ξ и η

$$3\xi^2 - 2\xi^3 = \gamma_1, \quad \eta = \gamma_2(1-\xi).$$

б) Выберем теперь в качестве первой величины η . Тогда

$$\rho_{\eta}(y) = \int_0^{1-y} \rho(x, y) dx = 3(1-y)^2 \quad \text{при } 0 < y < 1.$$

$$\rho_{\xi}(x|y) = \rho(x, y)/\rho_{\eta}(y) = 2x(1-y)^{-2} \quad \text{при } 0 < x < 1-y.$$

Соответствующие функции распределения:

$$F_{\eta}(y) = \int_0^k \rho_{\eta}(v) dv = 1 - (1-y)^3 \quad \text{при } 0 < y < 1,$$

$$F_{\xi}(x|y) = \int_0^x \rho_{\xi}(u|y) du = x^2(1-y)^{-2} \quad \text{при } 0 < x < 1-y.$$

Из формулы (12), используя $1 - \gamma_1$ вместо γ_1 , получим уравнения для последовательного вычисления η и ξ

$$(1 - \eta)^3 = \gamma_1, \quad \xi^2 = \gamma(1 - \eta)^2.$$

Сравним теперь оба алгоритма для расчета ξ и η : в первом из них для нахождения ξ необходимо решать кубические уравнения, в то время как во втором можно использовать явные формулы

$$\eta = 1 - \sqrt[3]{\gamma_1}, \quad \xi = \sqrt[3]{\gamma_2} \sqrt[3]{\gamma_1}. \quad (13)$$

З а м е ч а н и е. Учащиеся часто допускают ошибку и вместо $\rho_Q(x, y) = \rho_1(x)\rho_2(y|x)$ пишут соотношение $F_Q(x, y) = F_1(x)F_2(y|x)$. Однако последнее тождество неверно! В рассмотренном примере в треугольнике $F_Q(x, y) = 3x^2y$, а

$$F_{\xi}(x)F_{\eta}(y|x) = \frac{3-2x}{1-x}x^2y, \quad F_{\eta}(y)F_{\xi}(x|y) = \frac{1-(1-y)^3}{(1-y)^2}x^2.$$

2.3. Возможные обобщения теоремы 4. Важнейший вывод из теоремы 4 состоит в том, что моделирование многомерной случайной величины может быть сведено к последовательному моделированию ее координат. Формулы (12) используют для этого метод обратных функций. Но это вовсе не обязательно: в некоторых случаях формулы расчета окажутся проще, если использовать для моделирования случайной величины ξ_i с условной плотностью $\rho_i(x|\xi_1, \dots, \xi_{i-1})$ какой-нибудь из методов, рассмотренных в последующих §§ 3, 4, 5.

Например, из результата п. 4.1 следует, что вместо формул (13) можно использовать следующий алгоритм, в котором на расчет каждой точки (ξ , η) затрачиваются пять независимых случайных чисел, но зато не надо извлекать корней:

$$\theta = \max(\gamma_1; \gamma_2; \gamma_3), \quad \eta = 1 - \theta, \quad \xi = \theta \max(\gamma_4; \gamma_5).$$

2.4. Использование замены переменных. Во многих случаях удается упростить формулы моделирования многомерной случайной величины путем удачного выбора координат.

Правило преобразования плотности при преобразовании координат. Пусть $y_i = g_i(x_1, \dots, x_n)$, $i=1, 2, \dots, n$, — взаимно однозначное дифференцируемое отображение области B в пространстве x_1, \dots, x_n на область B' в пространстве y_1, \dots, y_n . Если плотность случайной точки $Q = (\xi_1, \dots, \xi_n)$ в B равна $p_Q(x_1, \dots, x_n)$, то плотность случайной точки $Q' = (\eta_1, \dots, \eta_n)$ в B' , где $\eta_i = g_i(\xi_1, \dots, \xi_n)$, равна

$$p_{Q'}(y_1, \dots, y_n) = p_Q(x_1, \dots, x_n) \left| \frac{\partial(x_1, \dots, x_n)}{\partial(y_1, \dots, y_n)} \right|;$$

в правой части x_i должны быть выражены через y_i .

Доказательство. Пусть D' — произвольная область внутри B' , а D — ее прообраз при рассматриваемом отображении. Очевидно, $P\{Q \in D\} = P\{Q' \in D'\}$. По правилу замены переменных в интеграле

$$P\{Q \in D\} = \int_D p_Q dx_1 \dots dx_n = \int_{D'} p_Q \left| \frac{\partial(x_1, \dots, x_n)}{\partial(y_1, \dots, y_n)} \right| dy_1 \dots dy_n,$$

а по определению плотности $p_{Q'}$

$$P\{Q' \in D'\} = \int_{D'} p_{Q'}(y_1, \dots, y_n) dy_1 \dots dy_n.$$

Приравнявая эти вероятности и принимая во внимание произвольность D' , получим требуемый результат.

2.4.1. Пример. Случайная точка Q равномерно распределена в шаре $x^2 + y^2 + z^2 < R^2$.

Обозначим через ξ, η, ζ декартовы координаты точки Q . Их совместная плотность распределения в шаре постоянна: $p_Q(x, y, z) = [(4/3)\pi R^3]^{-1}$. Однако плотности распределения каждой из координат достаточно громоздки (см. ниже п. 3.2.1). Поэтому перейдем к сферическим координатам (рис. 23):

$$x = r \sin \theta \cos \varphi, \quad y = r \sin \theta \sin \varphi, \quad z = r \cos \theta.$$

В новых координатах шар превращается в параллелепипед $0 \leq r < R$, $0 \leq \theta < \pi$, $0 \leq \varphi < 2\pi$. Так как якобиан преобразования

$$\partial(x, y, z) / \partial(r, \theta, \varphi) = r^2 \sin \theta,$$

то в новых координатах плотность

$$p_Q(r, \theta, \varphi) = [(4/3)\pi R^3]^{-1} r^2 \sin \theta.$$

Легко видеть, что эта плотность представляет собой произведение трех плотностей

$$\rho_Q(r, \theta, \varphi) = (3r^2 R^{-3}) (2^{-1} \sin \theta) (2\pi)^{-1}$$

и, следовательно, сферические координаты r_Q , θ_Q , φ_Q точки Q независимы. Уравнения (11) для их нахождения можно записать так:

$$\int_0^{r_Q} \frac{3r^2 dr}{R^3} = \gamma_1, \quad \int_0^{\theta_Q} \frac{\sin \theta d\theta}{2} = 1 - \gamma_2, \quad \int_0^{\varphi_Q} \frac{d\varphi}{2\pi} = \gamma_3.$$

Отсюда получаем широко распространенные формулы

$$r_Q = R \sqrt[3]{\gamma_1}, \quad \cos \theta_Q = 2\gamma_2 - 1, \quad \varphi_Q = 2\pi\gamma_3. \quad (14)$$

По этим значениям нетрудно вычислить и декартовы координаты точки Q :

$$\xi = r_Q \sin \theta_Q \cos \varphi_Q, \quad \eta = r_Q \sin \theta_Q \sin \varphi_Q, \quad \xi = r_Q \cos \theta_Q.$$

2.4.2. Пример. Выбор случайного направления в пространстве.

Обычно, говоря о «случайном» направлении, подразумевают выбор случайного направления в условиях, когда все направления

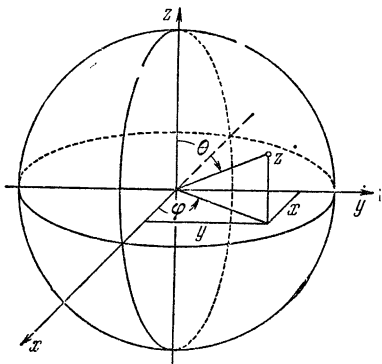


Рис. 23

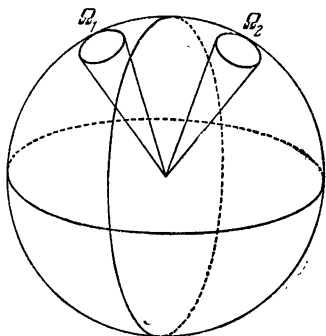


Рис. 24.

равновероятны (в противном случае должно быть задано распределение вероятностей различных направлений). Направление условимся характеризовать единичным вектором

$$\omega = i\omega_1 + j\omega_2 + k\omega_3,$$

где $\omega_1^2 + \omega_2^2 + \omega_3^2 = 1$. Нас интересует такой случайный вектор ω , что для любого телесного угла Ω

$$P\{\omega \in \Omega\} = \Omega/4\pi.$$

Легко видеть, что если Q — случайная точка, равномерно распределенная в шаре $x^2 + y^2 + z^2 < R^2$, то направление ее радиуса-вектора обладает нужным нам свойством. Действительно, если Ω_1 и Ω_2 — два равных телесных угла, то объемы соответствующих им шаровых секторов равны (рис. 24), и вероятность того, что точка Q попадет в каждый из них, одинакова. Поэтому из (14) получаем формулы для выбора «случайного» направления.

$$\cos \theta = 2\gamma_1 - 1, \quad \varphi = 2\pi\gamma_2. \quad (15)$$

Декартовы координаты вектора ω вычисляются по обычным формулам:

$$\omega_1 = \cos \varphi \sqrt{1 - \cos^2 \theta}, \quad \omega_2 = \sin \varphi \sqrt{1 - \cos^2 \theta}, \quad \omega_3 = \cos \theta.$$

2.4.3. Пример [8]. Случайная точка $Q = (\xi_1, \dots, \xi_n)$ подчиняется n -мерному нормальному (гауссовскому) распределению с математическими ожиданиями $M\xi_i = a_i$ и вторыми моментами $M[(\xi_i - a_i)(\xi_j - a_j)] = b_{ij}$. Определитель матрицы $\hat{B} = (b_{ij})$ положителен: $\Delta_B > 0$.

Плотность такой случайной точки выражается формулой

$$p_Q(x_1, \dots, x_n) = \sqrt{\frac{\Delta_C}{(2\pi)^n}} \exp \left\{ -\frac{1}{2} \sum_{i,j=1}^n c_{ij} (x_i - a_i)(x_j - a_j) \right\},$$

где $C = (c_{ij})$ — матрица обратная, по отношению к B , а Δ_C — ее определитель.

Как известно, линейным преобразованием координат можно привести положительно определенную квадратичную форму, стоящую в показателе, к сумме квадратов. Удобно при этом использовать векторные обозначения: если $u = (u_1, \dots, u_n)$ и $v = (v_1, \dots, v_n)$ — векторы, то их скалярное произведение

$$(u, v) = \sum_{i=1}^n u_i v_i;$$

если $C = (c_{ij})$ — квадратная матрица ($1 \leq i, j \leq n$), то $w = Cu$ — это вектор с компонентами $w_i = \sum_{\alpha=1}^n c_{i\alpha} u_\alpha$; квадратичная форма выражается через скалярное произведение

$$\sum_{i,j=1}^n c_{ij} (x_i - a_i)(x_j - a_j) = (C(x - a), x - a).$$

Выберем новые координаты y_1, \dots, y_n , и пусть $x - a = Ty$. Тогда

$$(C(x - a), x - a) = (CTy, Ty) = (T'CTy, y),$$

где T' — транспонированная матрица T . Последнее выражение обратится в (y, y) , если $T'CT = E$ — единичная матрица. Отсюда

$C = (T')^{-1}T^{-1}$, а $B = C^{-1} = TT'$. Таким образом, матрица преобразования T должна удовлетворять уравнению

$$TT' = B. \quad (16)$$

Якобиан преобразования $x = Ty + a$ равен определителю матрицы: $\partial(x)/\partial(y) = \Delta_T$. Из (16) следует, что $(\Delta_T)^2 = \Delta_B$, а так как $B = C^{-1}$, то $(\Delta_T)^2 = (\Delta_C)^{-1}$. Значит,

$$\partial(x)/\partial(y) = (\Delta_C)^{-1,2}.$$

Теперь можно записать плотность точки Q в новых координатах

$$\rho_Q(y_1, \dots, y_n) = (2\pi)^{-\frac{n}{2}} e^{-\frac{1}{2}(y,y)} = \prod_{i=1}^n \left[(2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}y_i^2} \right],$$

откуда видно, что новые координаты точки Q независимы и нормальны с параметрами $(0; 1)$.

Итак, для того чтобы вычислить значения ξ_1, \dots, ξ_n , надо найти n независимых значений ξ_1, \dots, ξ_n нормальной величины с параметрами $(0; 1)$ — как это сделать, см. п. 3.2.2. или п. 4.4, и тогда

$$\xi = T\xi + a.$$

При практической реализации этого метода единственное сложное место — расчет матрицы T . Из теории матриц следует [64], что существует треугольная матрица $T = (t_{ij})$, удовлетворяющая (16). Если при $j > i$ все $t_{ij} = 0$, то (16) превращается в систему, состоящую из $n(n+1)/2$ уравнений

$$\sum_{k=1}^j t_{ik} t_{jk} = b_{ij}, \quad 1 \leq j \leq i \leq n,$$

и все t_{ij} могут быть последовательно вычислены в порядке, схематически указанном в рис. 25. Матрицу C вычислять не надо.

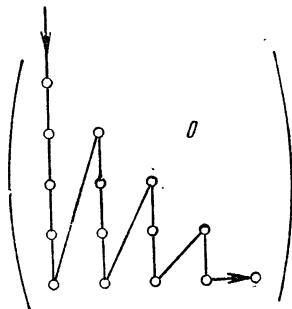


Рис. 25.

§ 3. Преобразования вида $\xi = g(\gamma_1, \gamma_2)$

3.1. Постановка задачи. Пусть γ_1 и γ_2 — два независимых случайных числа. По аналогии с п. 1.5 можно попытаться найти всевозможные функции $g(x, y)$ такие, что случайная величина $g(\gamma_1, \gamma_2)$ имеет функцию распределения $F(x)$. Так как случайная точка (γ_1, γ_2) равномерно распределена в единичном квадрате $0 < x < 1$,

$0 < y < 1$, то, повторяя рассуждения п. 1.5, получим для определения $g(x, y)$ уравнение, аналогичное (8):

$$\int_0^1 \int_0^1 e(z - g(x, y)) dx dy = F(z). \quad (17)$$

Мы не будем заниматься исследованием решений этого уравнения. Оно, по-видимому, более удобно для доказательства известных формул (см. п. 4.1), чем для получения новых. Вместо этого рассмотрим несколько методов построения преобразований вида $\xi = g(\gamma_1, \gamma_2)$. Эти методы имеют много практических приложений.

Во всех методах вместо одномерной величины ξ моделируется двумерная случайная величина Q , по значениям которой нетрудно вычислить ξ . Большой произвол в выборе распределения Q используется для того, чтобы упростить формулы счета. В п. 3.2 плотность $p_Q(x, y) = p_Q(r)$ зависит только от r , и моделируются полярные координаты точки Q . В пп. 3.3 и 3.4 ξ — это декартова координата точки $Q = (\xi, \eta)$, но сперва моделируется η .

3.2. Применение полярных координат. Допустим, что к случайной величине ξ , которую надо моделировать, удалось подобрать случайную величину η так, что плотность точки Q с декартовыми координатами ξ и η зависит только от расстояния до начала координат $r = \sqrt{x^2 + y^2}$:

$$p_Q(x, y) = c(r) \quad \text{при} \quad R_1 \leq r < R_2.$$

Здесь $0 \leq R_1 < R_2 \leq \infty$. Тогда удобно моделировать полярные координаты точки Q , а уже по ним вычислять ξ .

Если $x = r \cos \varphi$, $y = r \sin \varphi$, то якобиан преобразования $\partial(x, y)/\partial(r, \varphi) = r$ и плотность точки Q в полярных координатах равна

$$p_Q(r, \varphi) = rc(r).$$

Область изменения полярных координат точки Q — назовем их ρ и θ — прямоугольник $R_1 \leq r < R_2$, $0 \leq \varphi < 2\pi$. Поэтому легко доказать, что они независимы:

$$p_\rho(r) = 2\pi rc(r), \quad p_\theta(\varphi) = (2\pi)^{-1}.$$

По формуле (11) для вычисления ρ и θ получим уравнения

$$2\pi \int_{R_1}^{\rho} rc(r) dr = \gamma_1, \quad \theta(2\pi)^{-1} = \gamma_2. \quad (18)$$

Вычислив ρ и θ , нетрудно найти и декартовы координаты точки Q :

$$\xi = \rho \cos \theta, \quad \eta = \rho \sin \theta.$$

3.2.1. Пример. Случайная величина ξ определена в интервале $-R < x < R$ с плотностью $p(x) = 2(\pi R^2)^{-1} \sqrt{R^2 - x^2}$.

Нетрудно показать, что ξ представляет собой абсциссу случайной точки Q , равномерно распределенной в круге $x^2 + y^2 < R^2$. В самом деле, если $p_Q(x, y) = (\pi R^2)^{-1}$ в этом круге, то

$$p_{\xi}(x) = \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} p_Q(x, y) dy = \frac{2}{\pi R^2} \sqrt{R^2 - x^2} \quad \text{при } |x| < R.$$

Из формулы (18) при $c(r) = (\pi R^2)^{-1}$, $R_1 = 0$, $R_2 = R$ получаем явные формулы $\rho = R \sqrt{\gamma_1}$, $\theta = 2\pi\gamma_2$. Таким образом,

$$\xi = R \sqrt{\gamma_1} \cos 2\pi\gamma_2.$$

Если в этом примере сразу применить метод обратных функций для моделирования ξ , то уравнение (4) $F(\xi) = \gamma$ для нахождения ξ окажется весьма сложным:

$$F(\xi) \equiv \frac{1}{2} + \frac{1}{\pi} \arcsin \frac{\xi}{R} + \frac{\xi}{\pi R^2} \sqrt{R^2 - \xi^2} = \gamma.$$

3.2.2. Пример. Случайная величина ζ нормальна с параметрами $(0; 1)$

$$p_{\zeta}(x) = (2\pi)^{-1/2} e^{-x^2/2}, \quad 0 < x < \infty.$$

Выберем независимую от ζ случайную величину η , также нормальную с параметрами $(0; 1)$, и рассмотрим на плоскости x, y случайную точку Q с декартовыми координатами ζ и η . Очевидно,

$$p_Q(x, y) = p_{\zeta}(x) p_{\eta}(y) = (2\pi)^{-1} e^{-r^2/2}, \quad 0 \leq r < \infty.$$

По формулам (18), где удобно вместо γ_1 взять $1 - \gamma_1$, получим уравнения

$$\int_0^{\rho} r e^{-r^2/2} dr = 1 - \gamma_1, \quad \theta = 2\pi\gamma_2,$$

так что $\rho = \sqrt{-2 \ln \gamma_1}$. Следовательно,

$$\zeta = \sqrt{-2 \ln \gamma_1} \cos 2\pi\gamma_2, \quad \eta = \sqrt{-2 \ln \gamma_1} \sin 2\pi\gamma_2. \quad (19)$$

Формулы (19), полученные в [107], позволяют по двум случайным числам γ_1 и γ_2 сосчитать сразу два независимых значения случайной величины ξ . Если нужно лишь одно такое значение, то можно ограничиться одной из этих двух формул.

З а м е ч а н и е. Если случайная величина ξ нормальна с параметрами $(0; 1)$, то случайная величина

$$\xi = \sigma \zeta + a$$

нормальна с параметрами $(a; \sigma)$.

3.3 Метод суперпозиции. Допустим, что функция распределения $F(x)$ интересующей нас случайной величины ξ представима в виде

$$F(x) = \sum_{k=1}^m c_k F_k(x), \quad (20)$$

где все $F_k(x)$ — также функции распределения, а $c_k > 0$. Из (20) при $x \rightarrow \infty$ следует, что $c_1 + \dots + c_m = 1$. Следовательно, можно ввести дискретную случайную величину η с распределением

$$\begin{pmatrix} 1 & 2 & \dots & m \\ c_1 & c_2 & \dots & c_m \end{pmatrix},$$

так что $P\{\eta = k\} = c_k$.

Теорема 5. Пусть γ_1 и γ_2 — независимые случайные числа. Если по числу γ_1 разыграть значение $\eta = k$ случайной величины η , а затем из уравнения $F_k(\xi) = \gamma_2$ определить ξ , то функция распределения ξ равна $F(x)$.

Доказательство. Воспользуемся теоремой о полной вероятности и вычислим функцию распределения величины ξ , построенной в теореме:

$$\begin{aligned} P\{\xi < x\} &= \sum_{k=1}^n P\{\xi < x | \eta = k\} P\{\eta = k\} = \\ &= \sum_{k=1}^m F_k(x) c_k = F(x), \end{aligned}$$

что и требовалось доказать.

Функции распределения вида (20) встречаются тогда, когда мы имеем дело со смесью случайных величин. Например, если у нас всего N деталей, среди которых N_k деталей с функцией распределения «времени жизни» $F_k(t)$, $k=1, 2, \dots, m$, то функция распределения

«времени жизни» для случайно выбранной детали равна

$$F(t) = \sum_{k=1}^m (N_k/N) F_k(t).$$

Однако представление (20) часто придумывают искусственно, чтобы облегчить процедуру разыгрывания ξ .

Метод суперпозиции был предложен Дж Батлером [110] и развит в работах [40, 58, 108, 109, 155, 156]. Возможность обобщения его на случай бесконечного числа слагаемых в (20) и на многомерные распределения очевидна.

3.3.1. Пример [110]. Случайная величина ξ определена в интервале $0 < x < 1$ и имеет функцию распределения

$$F(x) = \sum_{k=1}^{\infty} c_k x^k,$$

где все $c_k \geq 0$.

Можно считать, что $F_k(x) = x^k$ при $0 < x < 1$, и воспользоваться методом суперпозиции. Из теоремы 5, используя теоремы 1 и 2, получим формулу

$$\text{если } \sum_{j=1}^{k-1} c_j \leq \gamma_1 < \sum_{j=1}^k c_j, \quad \text{то } \xi = (\gamma_2)^{1/k}$$

(при $k=1$ левую часть неравенства полагать равной нулю).

3.3.2. Пример. Случайная величина ξ определена в интервале $0 < x < 2$ с плотностью

$$p(x) = \frac{5}{12} [1 + (x-1)^4].$$

Если для нахождения значений величины ξ воспользоваться методом обратных функций (4), то получим формулу

$$(\xi-1)^5 + 5\xi = 12\gamma-1,$$

так что придется решать уравнение пятой степени.

Можно, однако, представить $p(x)$ в виде суперпозиции плотностей $p_1(x) = 1/2$ и $p_2(x) = (5/2)(x-1)^4$:

$$p(x) = \frac{5}{6} p_1(x) + \frac{1}{6} p_2(x).$$

На основании теоремы 5 получим следующий явный алгоритм для вычисления значений ξ :

$$\xi = \begin{cases} 2\gamma_2, & \text{если } \gamma_1 < 5/6, \\ 1 + \sqrt[5]{2\gamma_2 - 1}, & \text{если } \gamma_1 \geq 5/6. \end{cases}$$

Следующая модификация метода суперпозиции принадлежит Г. А. Михайлову [58].

3.3.3. Модифицированный метод суперпозиции. Оказывается при реализации метода суперпозиции можно ограничиться одним случайным числом γ .

Теорема 5'. Если в условиях теоремы 5 по числу γ разыграть значение $\eta = k$ случайной величины η , а затем определить ξ из урав-

нения $F_k(\xi) = \theta$, где $\theta = \left(\gamma - \sum_{j=1}^{k-1} c_j \right) c_k^{-1}$, то функция распределения ξ

равна $F(x)$.

Для доказательства этой теоремы достаточно заметить, что $P\{\theta < y | \eta = k\} = y$, т. е. θ равномерно распределена в интервале $(0, 1)$. Поэтому уравнение $F_k(\xi) = \theta$ определяет случайную величину с функцией распределения $F_k(x)$ так же, как в теореме 5*).

В примере п. 3.3.2 величина θ равна $(6/5)\gamma$ при $\eta = 1$ и θ равна $6\gamma - 5$ при $\eta = 2$. Для ξ получаем формулу

$$\xi = \begin{cases} (12/5)\gamma, & \text{если } \gamma < 5/6, \\ 1 + \sqrt[5]{12\gamma - 11}, & \text{если } \gamma \geq 5/6, \end{cases}$$

которая выгоднее формулы п. 3.3.2, ибо не требует вычисления второго случайного числа.

Необходимо отметить, однако, что модифицированный метод более чувствителен к качеству псевдослучайных чисел, используемых в расчете: для успеха обычного метода важно, чтобы частота попа-

дания псевдослучайных чисел в каждый из интервалов $\Delta_k =$

$= \left[\sum_{j=1}^{k-1} c_j \leq x < \sum_{j=1}^k c_j \right]$ равнялась c_k ; для модифицированного метода важно также, чтобы распределение этих чисел внутри каждого Δ_k было достаточно хорошим.

Модифицированному методу суперпозиции соответствует преобразование вида $\xi = g(\gamma)$ с разрывной функцией $g(y)$, которую проще записать, если ввести функции $G_k(y)$, обратные к $F_k(x)$:

$$g(y) = G_k \left(\left(y - \sum_{j=1}^{k-1} c_j \right) c_k^{-1} \right) \quad \text{при } y \in \Delta_k.$$

Проверим непосредственно, что эта функция удовлетворяет уравнению (8). Представим интеграл в виде суммы:

$$\int_0^1 e(x - g(y)) dy = \sum_{k=1}^m \int_{\Delta_k} e \left[x - G_k \left(\left(y - \sum_{j=1}^{k-1} c_j \right) c_k^{-1} \right) \right] dy.$$

^{*} Непрудно доказать, что θ и η независимы [68].

В k -м слагаемом сделаем замену переменной $y = \sum_{j=1}^{k-1} c_j + c_k z$; тогда Δk преобразуется в $[0, 1]$ и

$$\int_0^1 e(x - g(y)) dy = \sum_{k=1}^m c_k \int_0^1 e(x - G_k(z)) dz.$$

Так как $G_k(z) < x$ тогда и только тогда, когда $z < F_k(x)$, то

$$\int_0^1 e(x - G_k(z)) dz = \int_0^1 e(F_k(x) - z) dz = F_k(x).$$

Следовательно,

$$\int_0^1 e(x - g(y)) dy = \sum_{k=1}^m c_k^- F_k(x) = F(x).$$

Тем самым мы получили новое доказательство теоремы 5'.

Функцию $g(y)$ можно сделать непрерывной, если использовать при нечетных k уравнение $F_k(\xi) = 0$, а при четных k — уравнение $F_k(\xi) = 1 - 0$.

Заметим, наконец, что, в отличие от обычного метода суперпозиции, модифицированный метод не может быть так просто обобщен на случай многомерной случайной величины ξ .

3.4. Метод интегральной суперпозиции. Рассмотрим непрерывную случайную точку Q с декартовыми координатами ξ и η на плоскости x, y . Если плотность Q равна $p(x, y)$, то плотность ξ равна

$$p_\xi(x) = \int_{-\infty}^{\infty} p(x, y) dy.$$

Как отмечалось в п. 2.2, моделировать координаты точки Q можно в любом порядке. Воспользуемся представлением $p(x, y) = p_\eta(y)p_\xi(x|y)$ и будем сперва моделировать η , а затем ξ . Иными словами, сперва найдем η из уравнения $F_\eta(\eta) = \gamma_1$, а затем ξ — из уравнения $F_\xi(\xi|\eta) = \gamma_2$. В тех случаях, когда последние уравнения решаются проще, чем уравнение (4) метода обратных функций, такой алгоритм может оказаться выгодным. Вообще же метод интегральной суперпозиции используется сравнительно редко, главным образом тогда, когда плотность $p_\xi(x)$ задана в форме интеграла по параметру.

Пример [110]. Плотность случайной величины ξ при $0 < x < \infty$ пропорциональна интегральной показательной функции n -го порядка ($n > 0$)

$$p_\xi(x) = n \int_1^{\infty} y^{-n} e^{-xy} dy.$$

Так как здесь $p(x, y) = ny^{-n}e^{-xy}$, то

$$p_{\eta}(y) = \int_0^{\infty} p(x, y) dx = ny^{-n-1} \quad \text{при } 1 < y < \infty;$$

$$p_{\xi}(x|y) = p(x, y)/p_{\eta}(y) = ye^{-xy} \quad \text{при } 0 < x < \infty.$$

Соответствующие функции распределения равны

$$F_{\eta}(y) = 1 - y^{-n}, \quad F_{\xi}(x|y) = 1 - e^{-xy}.$$

Из уравнения $F_{\eta}(\eta) = 1 - \gamma_1$ найдем $\eta = (\gamma_1)^{-1/n}$, а из уравнения $F_{\xi}(\xi|\eta) = 1 - \gamma_2$ найдем $\xi = -\eta^{-1} \ln \gamma_2$. Итак,

$$\xi = -(\gamma_1)^{1/n} \ln \gamma_2.$$

З а м е ч а н и е. *Общий метод суперпозиции* может быть описан одной формулой

$$p_{\xi}(x) = \int_{-\infty}^{\infty} p_{\xi}(x|y) dF_{\eta}(y).$$

Именно в таком виде он был сформулирован Дж. Батлером (composition method). Если случайная величина η дискретна, то получаем метод п. 3.3, а если η непрерывна — метод интегральной суперпозиции. Однако в приложениях гораздо большую роль играет дискретный случай.

3.5. Некоторые приложения метода суперпозиции.

3.5.1. Поправки к приближенным распределениям. Предположим, что плотность $p(x)$ случайной величины ξ аппроксимируется снизу достаточно простой линией $y(x)$, как это изображено на рис. 26. Очевидно, в качестве приближения к $p(x)$ можно выбрать плотность

$$p_1(x) = y(x)/c_1,$$

где $c_1 = \int_a^b y(x) dx$, и находить приближенные значения ξ по плотности $p_1(x)$.

Можно, однако, представить $p(x)$ в форме суперпозиции двух плотностей

$$p_1(x) = y(x)/c_1 \text{ и } p_2(x) = [p(x) - y(x)]/c_2,$$

и получить таким образом метод для точного моделирования ξ . Алгоритм расчета ξ по плотности $p_2(x)$ может оказаться весьма сложным; но на времени счета это

почти не скажется, ибо $p_2(x)$ будет использоваться очень редко: $P\{\eta=2\} = c_2 = 1 - c_1 \ll c_1$.

Итак, метод суперпозиции дает возможность учесть «поправку» $p_2(x)$, практически не увеличивая времени счета, а лишь ценою усложнения программы (впрочем, обычно это весьма нежелательно). (Дж. Марсалья [155]).

3.5.2. Дробление области определения случайной величины. Этот прием иногда используют при моделировании случайной величины, плотность которой резко различна в различных областях.

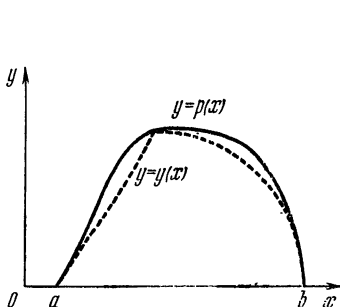


Рис. 26.

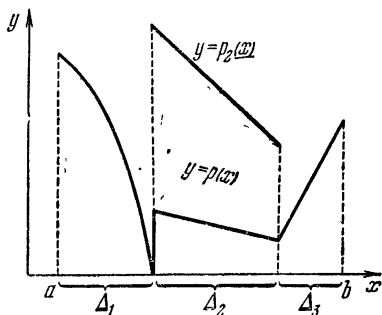


Рис. 27.

Пусть $p(x)$ — плотность случайной величины ξ , определенной в интервале $a < x < b$. Разобьем этот интервал на сумму непересекающихся интервалов Δ_k , так что $(a, b) = \Delta_1 + \dots + \Delta_m$ (рис. 27) и вероятности попадания ξ в Δ_k положительны: $c_k = \int_{\Delta_k} p(x) dx > 0$.

Введем в рассмотрение плотности

$$p_k(x) = \begin{cases} p(x)/c_k & \text{при } x \in \Delta_k, \\ 0 & \text{при } x \notin \Delta_k. \end{cases}$$

Очевидно, $c_1 + \dots + c_m = 1$ и при всех x из (a, b)

$$p(x) = c_1 p_1(x) + \dots + c_m p_m(x).$$

Согласно теореме 5, для того чтобы найти значение ξ , можно сперва по числу γ_1 разыграть номер области $\eta = k$, а затем вычислить ξ из уравнения

$$\int_{a_k}^{\xi} p(x) dx = c_k \gamma_2, \quad (21)$$

где a_k — левый конец Δ_k .

Легко проверить, что с точки зрения количества вычислений этот метод хуже, чем метод обратных функций. В самом деле, уравнение (4) для нахождения ξ

$$\int_a^{\xi} p(x) dx = \gamma$$

можно решать следующим образом: сперва найдем номер k такой, что

$$\sum_{j=1}^{k-1} c_j \leq \gamma < \sum_{j=1}^k c_j; \quad (22)$$

тогда это уравнение превратится в уравнение

$$\int_{a_k}^{\xi} p(x) dx = \gamma - \sum_{j=1}^{k-1} c_j, \quad (23)$$

решая которое и найдем ξ . Уравнение (23) проще, чем (21), и совпадает с уравнением модифицированного метода суперпозиции для рассматриваемой задачи.

Положение может резко измениться в пользу метода дробления области, если вместо (21) использовать для моделирования ξ с плотностью $p_k(x)$, в Δ_k какой-нибудь другой способ. Правда, тогда на получение одного значения ξ будет затрачиваться больше двух случайных чисел.

Метод дробления области применим также для моделирования многомерных случайных величин [19].

§ 4. Преобразования вида $\xi = g(\gamma_1, \dots, \gamma_n)$

Мы ограничимся несколькими весьма разнообразными примерами преобразований указанного вида. Во всех формулах $\gamma_1, \dots, \gamma_n$ — независимые случайные числа.

4.1. Извлечение корней из случайных чисел. Докажем, что значение случайной величины ξ , определенной при $0 < x < 1$ с функцией распределения $F(x) = x^n$, можно вычислять по формуле

$$\xi = \max(\gamma_1; \dots; \gamma_n). \quad (24)$$

Заметим сперва, что функция распределения случайной величины $\xi = g(\gamma_1, \dots, \gamma_n)$ равна $F(x)$ тогда и только тогда, когда $g(y_1, \dots, y_n)$ удовлетворяет уравнению

$$\int_0^1 \dots \int_0^1 e(x - g(y_1, \dots, y_n)) dy_1 \dots dy_n = F(x), \quad (25)$$

вполне аналогичному уравнению (8). Затем рассмотрим величину ξ , определенную уравнением (24). Так как $\max(y_1; \dots; y_n) < x$ в том и только в том случае, когда одновременно $y_1 < x, \dots, y_n < x$, то

$$\int_0^1 \dots \int_0^1 e(x - \max_{1 \leq i \leq n} y_i) dy_1 \dots dy_n = \int_0^x \dots \int_0^x dy_1 \dots dy_n = x^n,$$

что и требовалось доказать.

Если эту же случайную величину ξ моделировать методом обратных функций, то, очевидно,

$$\xi = \sqrt[n]{\gamma}. \quad (26)$$

Сравнивая формулы (26) и (24), приходим к выводу, что в любом алгоритме можно заменить извлечение корня из случайного числа взятием наибольшего из нескольких независимых случайных чисел.

На ранних этапах развития ЭВМ формула (24) часто использовалась даже при $n=2$, так как извлечение корня осуществлялось по весьма громоздкой подпрограмме. Например, на вычисление $\sqrt[3]{\gamma}$ на ЭВМ «Стрела» затрачивалось $3+24=27$ операций (3 операции — на расчет γ), а на вычисление $\max(\gamma_1; \gamma_2)$ — всего $3+3+2=8$ операций.

Обобщение формулы (24) приведено в упражнении 9 гл. 2.

4.2. Моделирование гамма-распределения. Во многих задачах встречаются величины $\xi^{(n)}$, определенные при $0 < x < \infty$ с плотностью вероятностей

$$p_n(x) = [(n-1)!]^{-1} x^{n-1} e^{-x}, \quad (27)$$

где $n \geq 1$ — целое число. Закон (27) называется *гамма-распределением* *), так как

$$\int_0^{\infty} x^{n-1} e^{-x} dx = \Gamma(n) = (n-1)!$$

(Встречаются также распределения (27) с дробными n .)

Метод обратных функций приводит к явной формуле для вычисления $\xi^{(n)}$ только в случае $n=1$

$$\xi^{(1)} = -\ln \gamma. \quad (28)$$

*) Закону (27) подчиняются интервалы между событиями в потоках Эрланга (см. гл. 6, п. 1.3.1).

Докажем, что при любом n значения $\xi^{(n)}$ можно вычислять по формуле

$$\xi^{(n)} = -\ln(\gamma_1 \dots \gamma_n). \quad (29)$$

Доказательство (по индукции). При $n=1$ формула (29) превращается в уже доказанную формулу (28). Допустим, что плотность величины (29) выражается формулой (27), и рассмотрим величину

$$\xi^{(n+1)} = -\ln(\gamma_1 \dots \gamma_{n+1}) = \xi^{(n)} + \xi^{(1)}.$$

По известному правилу композиции плотностей независимых слагаемых

$$p_{\xi^{(n+1)}}(x) = \int_{-\infty}^{\infty} p_n(x-t) p_1(t) dt = \int_0^{\infty} p_n(x-t) e^{-t} dt.$$

Используем теперь индукционное допущение:

$$\begin{aligned} p_{\xi^{(n+1)}}(x) &= [(n-1)!]^{-1} \int_0^x (x-t)^{n-1} e^{-x} dt = \\ &= [n!]^{-1} x^n e^{-x} = p_{n+1}(x). \end{aligned}$$

Пример [56, 118]. Часто при неупругом рассеянии нейтрона ядрами энергия ξ рассеянного нейтрона представляет собой случайную величину с плотностью

$$p(E) = (E/T^2) e^{-(E/T)}, \quad 0 < E < \infty$$

(это так называемый «испарительный спектр»; параметр T зависит от вида ядра и от энергии нейтрона перед столкновением). Используя замену переменной $E=Tx$ и формулу (29) при $n=2$, получим для расчета энергии после рассеяния формулу

$$\xi = -T \ln(\gamma_1 \gamma_2). \quad (30)$$

4.3. Моделирование семейства биномиальных распределений. Рассмотрим случайную величину ξ , которая подчиняется биномиальному распределению с параметром p , т. е. при $k=0, 1, 2, \dots, n$

$$P\{\xi = k\} = C_n^k p^k (1-p)^{n-k}. \quad (31)$$

Это дискретная величина, и моделировать ее можно методом п. 1.1.

Предположим теперь, что для расчета некоторой задачи необходимо многократно моделировать распределение (31) с различными значениями p , получающимися

в ходе расчета. Вместо того чтобы каждый раз вычислять все вероятности (31), можно использовать следующий алгоритм, который представляет собой алгоритм моделирования n независимых испытаний (ср. п. 1.2.1): для каждого из чисел $\gamma_1, \dots, \gamma_n$ проверяется неравенство $\gamma_i < p$. Если это неравенство оказалось выполненным k раз, то $\xi = k$ [174].

Формулу, выражающую ξ через $\gamma_1, \dots, \gamma_n$, можно записать в виде

$$\xi = \sum_{i=1}^n e(p - \gamma_i).$$

4.4. Приближенное моделирование нормального (гауссовского) распределения. Рассмотрим нормированную *) сумму n независимых равномерно распределенных величин:

$$\xi^{(n)} = \sqrt{\frac{3}{n}} \sum_{i=1}^n (2\gamma_i - 1). \quad (32)$$

Согласно центральной предельной теореме при $n \rightarrow \infty$

$$P\{\xi^{(n)} < x\} \rightarrow (2\pi)^{-1/2} \int_{-\infty}^x e^{-t^2/2} dt. \quad (33)$$

Следовательно, по формуле (32) при достаточно больших n можно вычислять приближенные значения нормальной случайной величины ξ с параметрами $(0; 1)$.

Асимптотика в формуле (33) устанавливается весьма быстро (на рис. 28 изображены плотности ξ и $\xi^{(3)}$), и поэтому на практике можно ограничиться значением $n=12$:

$$\xi \approx \xi^{(12)} = \sum_{i=1}^{12} \gamma_i - 6. \quad (34)$$

Иногда ограничиваются в (32) лишь пятью слагаемыми, но зато добавляют поправку, которая ускоряет сходимость распределения к нормальному:

$$\xi \approx 0,01\xi^{(5)}[97 + (\xi^{(5)})^2] \quad (35)$$

*) Случайная величина η_0 называется *нормированной*, если $M\eta_0=0$, $D\eta_0=1$. Любую случайную величину η с конечными $M\eta$ и $D\eta$ можно нормировать: величина $\eta_0=(\eta - M\eta)/\sqrt{D\eta}$ нормирована.

Последние две формулы нередко оказываются удобнее, чем (19), так как расчет по ним возможен без обращения к подпрограммам \ln и \sin .

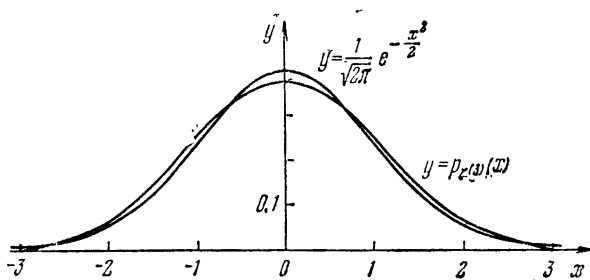


Рис. 28.

О поправках, ускоряющих сходимость, см. [4]. В [47] имеется пример, при расчете которого пришлось использовать $n=30$.

§ 5. Методы отбора

5.1. Общая характеристика методов отбора. Предположим, что в m -мерном пространстве переменных y_1, \dots, y_m заданы случайная точка $Q' = (\eta_1, \dots, \eta_m)$ с функцией распределения $F_Q(y_1, \dots, y_m)$ и некоторая область B' . Рассмотрим одномерную случайную величину ξ , определенную формулой

$$\xi = \varphi(Q') \text{ при } Q' \in B'. \quad (36)$$

Для расчета по этой формуле можно выбрать случайную точку Q' в пространстве; если $Q' \in B'$, то вычисляется $\xi = \varphi(Q')$; если $Q' \notin B'$, то точка Q' отбрасывается и выбирается новая. Таким образом, при расчете по формуле (36) из случайных точек Q' с функцией распределения F_Q отбирают точки, принадлежащие B' , и по ним вычисляют ξ . Мы будем говорить, что формула (36) определяет *метод отбора* для моделирования ξ (в литературе встречается также термин «метод отказов», более соответствующий английскому выражению *rejection technique*).

Эффективностью метода отбора называют вероятность отбора или, более подробно, вероятность того, что

точка Q' будет использована для расчета ξ , а не будет отброшена. Очевидно, эффективность метода (36) равна вероятности

$$\vartheta = P\{Q' \in B\}. \quad (37)$$

Выбрав N точек Q' , мы получим в среднем всего ϑN значений ξ . Следовательно, на расчет каждого значения ξ затрачивается в среднем $1/\vartheta$ точек Q' . Ясно, что при малых ϑ метод (36) становится практически неэффективным.

Если на реализацию каждой точки Q' затрачивается n случайных чисел $\gamma_1, \dots, \gamma_n$, где, очевидно, $n \geq m$, то в среднем на одно значение ξ затрачивается n/ϑ случайных чисел. В вычислительной практике (при моделировании одномерных величин ξ) чаще всего встречается случай $m=2$ и $n=m$.

5.2. Моделирование усеченных распределений. Рассмотрим случайную величину η , определенную в интервале $a < x < b$ с плотностью $p(x)$ (так что $\int_a^b p(x) dx = 1$).

Говорят, что случайная величина ξ имеет *усеченное распределение* $p(x)$, если она определена в интервале $(a', b') \subset (a, b)$ и плотность ее пропорциональна $p(x)$. Очевидно,

$$p_{\xi}(x) = p(x) \left[\int_{a'}^{b'} p(x) dx \right]^{-1} > p(x);$$

см. рис. 29, где $a=a'=0$, $b=\infty$.

Если мы умеем вычислять значения η , то значения ξ можно находить методом отбора:

$$\xi = \eta, \text{ если } \eta \in (a', b'). \quad (38)$$

Эффективность такого метода равна

$$\vartheta = \int_{a'}^{b'} p(x) dx.$$

5.3. Метод Неймана [163]. Этот метод очень часто используется на практике. Иногда все методы отбора называют методом Неймана.

Рассмотрим случайную величину ξ , определенную на конечном интервале $a < x < b$ с ограниченной плотностью $p(x) \leq c$ (рис. 30).

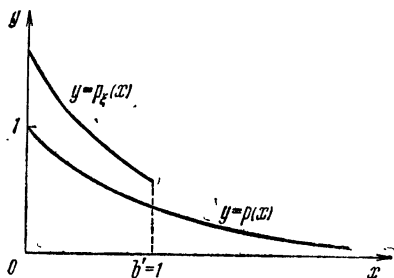


Рис. 29.

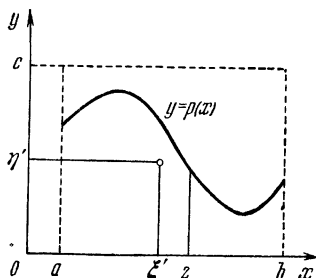


Рис. 30.

Теорема 6. Пусть γ_1 и γ_2 — независимые случайные числа и $\xi' = a + \gamma_1(b-a)$, $\eta' = c\gamma_2$. Случайная величина ξ , определенная условием

$$\xi = \xi', \text{ если } \eta' < p(\xi'), \quad (39')$$

имеет плотность вероятностей, равную $p(x)$.

Доказательство. Во-первых, заметим, что точка (ξ', η') равномерно распределена в прямоугольнике $a < x < b$, $0 < y < c$ (см. п. 2.1). Далее вычислим условную вероятность:

$$P\{\xi < z\} = P\{\xi' < z \mid \eta' < p(\xi')\} = \frac{P\{\xi' < z, \eta' < p(\xi')\}}{P\{\eta' < p(\xi')\}}.$$

Знаменатель последнего выражения — это вероятность попадания точки (ξ', η') под кривую $y = p(x)$. Так как плотность точки (ξ', η') постоянна и равна $[c(b-a)]^{-1}$, то

$$P\{\eta' < p(\xi')\} = \int_a^b dx \int_0^{p(x)} [c(b-a)]^{-1} dy = [c(b-a)]^{-1}.$$

Числитель равен вероятности того, что точка (ξ', η') окажется под кривой и в то же время $\xi' < z$:

$$\begin{aligned} P\{\xi' < z, \eta' < p(\xi')\} &= \\ &= \int_a^z dx \int_0^{p(x)} [c(b-a)]^{-1} dy = [c(b-a)]^{-1} \int_a^z p(x) dx. \end{aligned}$$

Таким образом, получаем, что

$$P\{\xi < z\} = \int_a^z p(x) dx,$$

а это как раз и требовалось доказать.

Эффективность метода (39), согласно (37), равна вероятности попадания точки (ξ', η') под кривую $y=p(x)$, т. е. $\varepsilon = P\{\eta' < p(\xi')\}$. Последняя вероятность уже вычислялась в ходе доказательства теоремы. Значит,

$$\varepsilon = [c(b-a)]^{-1};$$

эффективность метода Неймана будет наибольшей, если выбрать наименьшее возможное c , т. е. положить $c = \sup_{a < x < b} p(x)$. Впрочем, это очевидно также из геометрических соображений.

Пример. Случайная величина ξ определена при $-R < x < R$ с плотностью $p(x) = 2(\pi R^2)^{-1} \sqrt{R^2 - x^2}$.

Согласно теореме 6 нужно выбрать два значения γ_1 и γ_2 и вычислить $\xi' = R(2\gamma_1 - 1)$ и $\eta' = 2(\pi R)^{-1}\gamma_2$; если $\eta' < p(\xi')$, то $\xi = \xi'$. Однако условие $\eta' < p(\xi')$ в данном примере выгодно преобразовать: оно равносильно условию $\gamma_2 < \sqrt{1 - (2\gamma_1 - 1)^2}$. После дальнейших упрощений окончательно запишем:

$$\xi = R(2\gamma_1 - 1), \text{ если } (2\gamma_1 - 1)^2 < 1 - \gamma_2^2.$$

Эффективность этого метода $\varepsilon = \pi/4$. По сравнению с формулой

$\xi = R\sqrt{\gamma_1} \cos 2\pi\gamma_2$, полученной в п. 3.2.1, формула метода отбора проще: не надо извлекать корень и вычислять косинус.

5.4. О некоторых обобщениях метода Неймана. Во многих работах рассмотрены самые разнообразные обобщения метода п. 5.3. (например, [73, 109, 110]). Большинство обобщений, относящихся к моделированию одномерной случайной величины ξ , могут быть получены из нижеследующих формул (40) — (41).

Предположим, что нас интересует случайная величина ξ , определенная в интервале $a < x < b$. Рассмотрим случайную точку $Q' = (\xi', \eta')$ с плотностью $p(x, y)$ в полосе $a < x < b$, $-\infty < y < \infty$, и кривую $y = f(x)$, заданную при $a < x < b$ (рис. 31). Определим метод отбора:

$$\xi = \xi', \quad \text{если} \quad \eta' < f(\xi'). \quad (40)$$

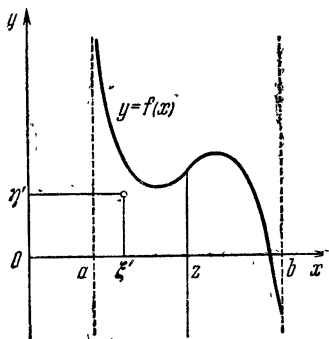


Рис. 31.

Нетрудно вычислить плотность случайной величины ξ , определенной формулой (40): при $a < z < b$

$$F_{\xi}(z) = P\{\xi < z\} = P\{\xi' < z \mid \eta' < f(\xi')\} = \\ = \frac{P\{\xi' < z, \eta' < f(\xi')\}}{P\{\eta' < f(\xi')\}} = \frac{\int_a^z dx \int_{-\infty}^{f(x)} p(x, y) dy}{\int_a^b dx \int_{-\infty}^{f(x)} p(x, y) dy};$$

продифференцировав по z , получим плотность

$$p_{\xi}(z) = \int_{-\infty}^{f(z)} p(z, y) dy \bigg/ \int_a^b dx \int_{-\infty}^{f(x)} p(x, y) dy. \quad (41)$$

Если требуется моделировать случайную величину ξ с заданной плотностью $p_{\xi}(x)$, то существует бесконечное количество способов выбрать $p(x, y)$ и $f(x)$, удовлетворяющие (41).

5.4.1. Сперва рассмотрим частный случай, когда ξ' и η' независимы: $p(x, y) = p_{\xi'}(x) p_{\eta'}(y)$. Тогда из (41) вытекает, что

$$p_{\xi}(z) = p_{\xi'}(z) \int_{-\infty}^{f(z)} p_{\eta'}(y) dy \bigg/ \int_a^b p_{\xi'}(x) dx \int_{-\infty}^{f(x)} p_{\eta'}(y) dy. \quad (42)$$

Последняя формула записывается более коротко, если ввести функцию распределения величины η' , равную $F_{\eta'}(y)$:

$$p_{\xi}(z) = p_{\xi'}(z) F_{\eta'}(f(z)) \bigg/ \int_a^b p_{\xi'}(x) F_{\eta'}(f(x)) dx.$$

Итак, если плотность $p_{\xi}(x)$ моделируемой случайной величины ξ представима в форме произведения

$$p_{\xi}(x) = k p_{\xi'}(x) F_{\eta'}(f(x)),$$

где k — постоянная, то ξ можно моделировать по формуле (40), в которой ξ' имеет плотность $p_{\xi'}(x)$, а η' — функцию распределения $F_{\eta'}(y)$.

5.4.2. Предположим дополнительно, что $0 \leq f(x) \leq c$, и выберем случайную величину η' , равномерно распределенную при $0 < y < c$. Тогда $p_{\eta'}(y) = 1/c$ и из (42) вытекает, что

$$p_{\xi}(z) = p_{\xi'}(z) f(z) \bigg/ \int_a^b p_{\xi'}(x) f(x) dx.$$

Следовательно, если плотность $p_{\xi}(x)$ моделируемой случайной величины ξ представима в форме произведения

$$p_{\xi}(x) = k p_{\xi'}(x) f(x), \quad (43)$$

где $0 \leq f(x) \leq c$, а k — постоянная, то ξ можно моделировать по формуле (40), в которой ξ' имеет плотность $p_{\xi'}(x)$, а η' равномерно распределена в интервале $(0, c)$.

Эффективность метода в этом случае равна

$$\varepsilon = P\{\eta' < f(\xi')\} = \int_a^b dx \int_0^{f(x)} p_{\xi'}(x) c^{-1} dy = (ck)^{-1}.$$

Представляя $p_{\xi}(x)$ в форме различных произведений (43), можно построить различные методы вида (40) для моделирования величины ξ . В частности, таким методом можно моделировать случайные величины, плотности которых неограничены.

Пример. Случайная величина ξ определена при $0 < x < 1$ с плотностью

$$p_{\xi}(x) = v(x)/\sqrt{x},$$

где $0 < v(x) \leq c$.

Представим $p_{\xi}(x)$ в форме произведения (43) с $p_{\xi'}(x) = (2\sqrt{x})^{-1}$. Значения ξ' легко вычислять методом обратных функций, так как из (4) вытекает, что $\sqrt{\xi'} = \gamma_1$. Следовательно, $\xi' = (\gamma_1)^2$, $\eta' = c\gamma_2$ и

$$\xi = \xi' \quad \text{при} \quad \eta' < v(\xi').$$

Эффективность в этом примере равна $\varepsilon = (2c)^{-1}$. Ясно, что для повышения ε следует выбрать наименьшее возможное значение c , равное $c = \sup_{a < x < b} v(x)$.

5.4.3. Снова предположим, что $0 \leq f(x) \leq c$. Если ξ' равномерно распределена при $a < x < b$, а η' равномерно распределена при $0 < y < c$, то $\xi' = a + \gamma_1(b-a)$ и $\eta' = c\gamma_2$. Из (42) вытекает, что при $a < z < b$

$$p_{\xi}(z) = kf(z).$$

Значит, чтобы моделировать случайную величину ξ с ограниченной плотностью $p_{\xi}(x)$, можно выбрать

$$f(x) = p_{\xi}(x).$$

В этом и состоит метод Неймана (39).

5.5. Выбор равномерно распределенных точек в сложных областях. Методы отбора (36) легко обобщить так, чтобы отбирались значения многомерной случайной величины ξ . Мы рассмотрим лишь один прием, который имеет много практических приложений. Его можно считать частным случаем метода моделирования усеченных распределений (п. 5.2), если иметь в виду обобщение этого метода на многомерный случай.

Пусть B — ограниченная область на плоскости x, y , «сложная» с точки зрения вычислительной практики:

например, невыпуклая или несвязная или такая, что границы на отдельных участках трудно записать в явном виде. Предположим лишь, что существует достаточно простой алгоритм, позволяющий определить, принадлежит ли области B любая заданная точка (x, y) или нет.

Выберем прямоугольник $\Pi = \{a_1 < x < b_1, a_2 < y < b_2\}$, содержащий область B (рис. 32). Координаты случайной

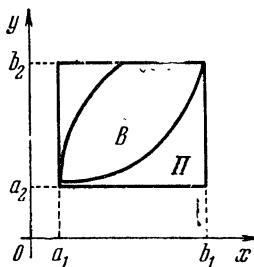


Рис. 32.

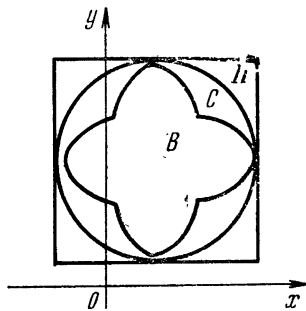


Рис. 33.

точки $Q' = (\xi', \eta')$, равномерно распределенной в Π , легко вычислить (п. 2.1):

$$\xi' = a_1 + \gamma_1(b_1 - a_1), \quad \eta' = a_2 + \gamma_2(b_2 - a_2).$$

Для нахождения точек Q , равномерно распределенных в B , можно вычислять точки Q' , равномерно распределенные в Π , и отбирать среди них те, которые принадлежат B . В самом деле, для любой области $G \subset B$

$$P\{Q \in G\} = P\{Q' \in G \mid Q' \in B\} = P\{Q' \in G \mid Q' \in B\}.$$

Так как Q' равномерно распределена в Π , то вероятность попадания Q' в любую область пропорциональна площади этой области: $P\{Q' \in G\} = S_G/S_\Pi$, $P\{Q' \in B\} = S_B/S_\Pi$. Следовательно,

$$P\{Q \in G\} = S_G/S_B.$$

или, что то же, плотность $p_Q(x, y) = 1/S_B$ в области B .

Эффективность такого метода равна отношению площадей

$$\vartheta = S_B/S_\Pi.$$

Поэтому ε будет наибольшей тогда, когда площадь Π минимальна — результат очевидный геометрически. Ясно также, что в тех случаях, когда область B хорошо вписывается, например, в круг C (рис. 33), лучше не пользоваться прямоугольником Π , а отбирать точки Q из числа точек Q' , равномерно распределенных в C . Эффективность такого метода $\varepsilon = S_B/S_C$ будет выше, ибо $S_C < S_\Pi$.

Пример. Случайные точки $Q = (\xi, \eta, \zeta)$, равномерно распределенные в шаре $x^2 + y^2 + z^2 < R^2$, можно выбирать следующим образом: 1) находим три случайных числа $\gamma_1, \gamma_2, \gamma_3$; 2) вычисляем координаты $\xi' = 2\gamma_1 - 1$, $\eta' = 2\gamma_2 - 1$, $\zeta' = 2\gamma_3 - 1$; 3) если $(\xi')^2 + (\eta')^2 + (\zeta')^2 < 1$, то полагаем $\xi = R\xi'$, $\eta = R\eta'$, $\zeta = R\zeta'$; в противном случае выбираем новую тройку $\gamma_1, \gamma_2, \gamma_3$.

Эффективность метода $\varepsilon = (4/3)\pi R^3 : (2R)^3 = \pi/6$. Несмотря на то, что ε близка к $1/2$, по полученным здесь формулам случайные точки в шаре вычисляются быстрее, чем по формулам п. 2.4.1, в которых приходится вычислять $\sqrt[3]{\gamma_1}$ и $\cos 2\pi\gamma_3$.

5.6. Алгоритмы, соответствующие методам отбора. Формула (36) не определяет полностью алгоритма для расчета значений ξ , так как вычислять значения Q' можно различными методами. Если каждая точка Q' вычисляется по n случайным числам $\gamma_1, \dots, \gamma_n$, где $n \geq m$, то все координаты Q' представляют собой функции $\eta_i = g_i(\gamma_1, \dots, \gamma_n)$, $i = 1, 2, \dots, m$. Подставив эти выражения в (36), получим, что $\xi = \Phi(g_1(\gamma_1, \dots, \gamma_n), \dots, g_m(\gamma_1, \dots, \gamma_n))$. Следовательно, ξ

выражается через $\gamma_1, \dots, \gamma_n$ по формуле вида $\xi = \tilde{g}(\gamma_1, \dots, \gamma_n)$.

Рассмотрим теперь преобразование

$$y_i = g_i(x_1, \dots, x_n), \quad 1 \leq i \leq m, \quad (44)$$

которое отображает единичный n -мерный куб

$$K^n = \{0 < x_1 < 1, \dots, 0 < x_n < 1\}$$

на область возможных значений Q' в m -мерном пространстве переменных y_1, \dots, y_m . Обозначим через B прообраз B' при отображении (44) (иными словами, B — это множество таких точек (x_1, \dots, x_n) из K^n , которым при отображении (44) соответствуют точки (y_1, \dots, y_m) , принадлежащие B'). Тогда условие $Q' \in B'$, фигурирующее в (36), равносильно условию $(\gamma_1, \dots, \gamma_n) \in B$.

Итак, формуле (36) соответствует следующий алгоритм расчета ξ по $\gamma_1, \dots, \gamma_n$:

$$\xi = \tilde{g}(\gamma_1, \dots, \gamma_n) \quad \text{при} \quad (\gamma_1, \dots, \gamma_n) \in B. \quad (45)$$

Формально можно считать, что (45) — это частный случай (36), так

5.7. Заключительные замечания. При выборе алгоритмов для расчета методами Монте-Карло различных задач возникает вопрос, какие же преобразования использовать для моделирования той или иной конкретной случайной величины ξ . Ответить на такой вопрос однозначно нельзя. Целесообразность тех или иных преобразований зависит от многих факторов.

Например, часто стремятся к минимизации времени, затрачиваемого на получение одного значения ξ . Однако за такую минимизацию приходится расплачиваться удлинением программы и (или) расходом места во внутреннем накопителе ЭВМ. Если во внутреннем накопителе места много, то можно использовать таблицу с интерполяцией (см. конец п. 1.3): такой способ будет, как правило, самым быстрым. Поэтому выбор любого другого способа моделирования — это компромисс *).

Далее, выбирая преобразования, редко учитывают качество псевдослучайных чисел. Но однородное распределение чисел γ_i обычно тщательно проверено, а распределение групп $(\gamma_{in+1}, \dots, \gamma_{in+n})$ при $n \geq 3$ проверяется хуже, в результате чего формула вида $\xi = g(\gamma_1, \dots, \gamma_n)$ при $n \geq 3$ может оказаться менее точной, чем формула вида $\xi = g(\gamma)$. Поэтому уменьшение количества случайных чисел, затрачиваемых на получение одного значения ξ , иногда представляется весьма желательным (особенно с точки зрения гл. 7).

К счастью, при расчете многих задач время, затрачиваемое на преобразование случайных величин, не слишком велико, и можно выбирать самые простые и естественные способы моделирования.

Упражнения к главе 2

1. Вывести явную формулу для расчета значений случайной величины ξ с функцией распределения

$$F(x) = 1 - (1/3)(2e^{-x} + e^{-5x}), \quad 0 < x < \infty.$$

*) Еще быстрее — осуществлять случайную выборку из заранее заготовленной таблицы значений ξ . Для некоторых случайных величин такие таблицы имеются [165], однако в расчетах на ЭВМ применяются они редко (ср. конец п. 1.3 гл. 1).

Интересно отметить, что точно так же можно использовать таблицу для моделирования случайной величины, значения которой получают экспериментально и функция распределения которой неизвестна (даже в многомерном случае).

2. Вывести явную формулу для расчета значений случайной величины ξ с плотностью распределения $p(x) = \cos^2 2\pi mx$, $0 < x < 2$; число $m \geq 1$ целое.

3. Вывести явные формулы для расчета случайных точек, равномерно распределенных в плоском кольце $R_1^2 < x^2 + y^2 < R_2^2$.

4. Вывести явные формулы для расчета реализаций случайной точки (ξ, η) с плотностью $p(x, y) = 3y$, определенной в треугольнике, ограниченном прямыми $x=0$, $y=x$ и $y=1$.

5. Вычислить плотность случайной величины $\xi = \eta - \ln \gamma$, если η имеет плотность $p(x)$ при $0 < x < \infty$ (η и γ независимы).

6. Доказать, что случайную величину ξ , определенную в интервале $0 < x < l$, с плотностью

$$p(x) = ae^{-ax}/(1 - e^{-cl}),$$

можно моделировать с помощью любой из четырех формул:

$$\xi = -(1/a) \ln [1 - \gamma (1 - e^{-al})];$$

$$\xi = -(1/a) \ln [\gamma + (1 - \gamma) e^{-al}];$$

$$\xi = l D [-(al)^{-1} \ln \gamma];$$

$$\xi = -(1/a) \ln \gamma, \quad \text{если } \xi < l.$$

7. Если гамма-квант с энергией E рассеивается в результате комптон-эффекта, то его энергия после рассеяния ξ представляет собой случайную величину с плотностью $p(x)$, пропорциональной функции

$$f(E, x) = x/E + E/x + (1/E - 1/x)(2 + 1/E - 1/x)$$

при $E(1+2E)^{-1} < x < E$ (закон Клейна — Нишина). Доказать, что ξ можно вычислять методом отбора:

$$\xi = \xi', \text{ если } \gamma_2[1+2E + (2E)^{-1}] < f(E, \xi'),$$

где $\xi' = E(1+2E\gamma_1)(1+2E)^{-1}$

(И. Г. Дядькин [26]).

8. Энергию нейтрона, испущенного при делении ядра U^{235} , часто считают случайной величиной ξ , определенной при $0 < x < \infty$, с плотностью

$$p(x) = ce^{-x/T} \operatorname{sh} b \sqrt{2x/T},$$

где b и T — параметры, $c = \sqrt{2/\pi} e^{-b^2/2} (bT)^{-1}$ — нормировочная постоянная. Доказать, что ξ можно вычислять по формуле

$$\xi = T[(1/2)(\zeta + b)^2 - \ln \gamma],$$

где ζ — нормальная случайная величина с параметрами $(0; 1)$; ζ и γ независимы.

(Г. А. Михайлов [56])

9. Независимые случайные числа $\gamma_1, \gamma_2, \dots, \gamma_{n+m-1}$ расположены в порядке возрастания $\gamma_{(1)} \leq \gamma_{(2)} \leq \dots \leq \gamma_{(n+m-1)}$. Доказать,

что $\xi = \gamma_{(n)}$ подчиняется бэта-распределению с параметрами n и m :

$$\rho_{\xi}(x) = [B(n, m)]^{-1} x^{n-1} (1-x)^{m-1}.$$

(J. C. Butcher, H. Messel [109]).

(Способы моделирования бэта-распределения с дробными параметрами рассмотрены в статьях [62, 73, 102, 103, 141].)

10. Доказать, что формулы $l_0=1$, $l_i = l_{i-1} - \xi_i$ и

$$\xi_i = l_{i-1} [1 - (\gamma_i)^{1/(n-i+1)}], \quad i = 1, 2, \dots, n,$$

определяют случайную точку (ξ_1, \dots, ξ_n) , равномерно распределенную в n -мерной пирамиде $x_1 + \dots + x_n < 1$, $x_i > 0$.

11. Допустим, что случайная точка Q с плотностью $\rho(\varphi(x, y, z))$ определена в области B . Обозначим через L_c часть поверхности $\varphi(x, y, z) = c$, принадлежащую B , и предположим, что семейство L_c при $c_1 < c < c_2$ заполняет B . Доказать, что если $|\text{grad } \varphi|_{L_c} = h(c)$, то точку Q можно моделировать в два этапа: сперва выбирается случайное значение параметра ρ с плотностью $\rho_{\rho}(c) = [\rho(c)/h(c)] S_{L_c}$, где S_{L_c} — площадь L_c , а затем на поверхности L_{ρ} выбирается случайная равномерно распределенная точка.

(В. А. Герман, И. М. Соболев [17].)

12. Доказать, что случайная точка с декартовыми координатами $(\xi_1/\eta, \dots, \xi_n/\eta)$, где $\xi_i = -\ln \gamma_i$, а $\eta = \xi_1 + \dots + \xi_n$, равномерно распределена в симплексе $x_1 + \dots + x_n = 1$, $0 < x_i < 1$ ($i = 1, 2, \dots, n$).

13. Доказать, что случайная точка с декартовыми координатами $(\zeta_1/\rho, \dots, \zeta_n/\rho)$, где ζ_1, \dots, ζ_n — независимые нормальные случайные величины с параметрами $(0; 1)$, а $\rho = (\zeta_1^2 + \dots + \zeta_n^2)^{1/2}$, равномерно распределена на поверхности единичной многомерной сферы $x_1^2 + \dots + x_n^2 = 1$.

ГЛАВА 3

ВЫЧИСЛЕНИЕ ИНТЕГРАЛОВ

Во введении уже упоминалось, что в подавляющем большинстве задач, решаемых методами Монте-Карло, вычисляют математические ожидания некоторых случайных величин. Так как чаще всего математические ожидания представляют собой обычные интегралы, то центральное положение в теории методов Монте-Карло занимают методы вычисления интегралов.

В настоящей главе изложены наиболее важные и вместе с тем наиболее простые методы, знакомство с которыми необходимо каждому специалисту в области Монте-Карло. Более сложные методы, не имеющие пока значительных практических применений, но намечающие, как нам думается, основные направления развития методов вычисления интегралов, рассмотрены в гл. 4.

Впрочем, п. 1.1 гл. 4 мог бы быть отнесен к гл. 3: выборка по группам не раз использовалась в практических вычислениях. Численный пример, приведенный в конце гл. 4 (п. 3.3), относится к оценкам, рассмотренным в обеих главах.

Заметим, что почти все результаты этих глав легко обобщаются на интегралы Римана — Стильтьеса (и даже Лебега — Стильтьеса [33]). Мы, однако, всюду рассматриваем только интегрирование в смысле Римана.

§ 1. Общий метод оценки математических ожиданий

1.1. Сходимость метода. Рассмотрим произвольную случайную величину ξ , у которой существует математическое ожидание $M\xi = a$. (Напомним, что по определению математическое ожидание $M\xi$ существует тогда и

только тогда, когда существует $M[\xi]$. Чтобы оценить величину a , выберем N независимых реализаций $\xi_1, \dots, \dots, \xi_N$ случайной величины ξ и вычислим среднее арифметическое

$$\bar{\xi}_N = (1/N) \sum_{i=1}^N \xi_i. \quad (1)$$

Так как *последовательность одинаково распределенных независимых случайных величин, у которых существуют математические ожидания, подчиняется закону больших чисел* (теорема А. Я. Хинчина [44]), то среднее арифметическое этих величин сходится по вероятности*) к математическому ожиданию: при $N \rightarrow \infty$

$$\bar{\xi}_N \xrightarrow{P} a.$$

Таким образом, при больших N величина $\bar{\xi}_N \approx a$; и оценку (1) можно использовать во всех случаях, когда существует $M\xi = a$.

1.2. Погрешность метода. Предположим дополнительно, что случайная величина ξ имеет конечную дисперсию

$$D\xi = M(\xi^2) - (M\xi)^2. \quad (2)$$

Из курса теории вероятностей известно, что *последовательность одинаково распределенных независимых случайных величин с конечными дисперсиями подчиняется центральной предельной теореме*. Последнее означает, что для любых $x_1 < x_2$

$$\begin{aligned} \lim_{N \rightarrow \infty} P \left\{ x_1 < (1/\sqrt{ND\xi}) \sum_{i=1}^N (\xi_i - a) < x_2 \right\} = \\ = (1/\sqrt{2\pi}) \int_{x_1}^{x_2} e^{-t^2/2} dt. \end{aligned}$$

Выберем $x_2 = -x_1 = x$. Тогда из последнего соотношения получим, что

$$\lim_{N \rightarrow \infty} P \left\{ \left| (1/N) \sum_{i=1}^N (\xi_i - a) \right| < x \sqrt{D\xi/N} \right\} = \Phi(x),$$

где $\Phi(x)$ — *интеграл вероятностей*, таблица которого

*) Определение сходимости по вероятности приведено на стр. 34.

приведена на стр. 293:

$$\Phi(x) = (2/\sqrt{2\pi}) \int_0^x e^{-t^2/2} dt.$$

Следовательно, при достаточно больших значениях N

$$P\{|\bar{\xi}_N - a| < x\sqrt{D\xi/N}\} \approx \Phi(x). \quad (3)$$

Формула (3) содержит целое семейство оценок, зависящее от параметра x . Если задать любой коэффициент доверия β (см. стр. 31), то можно найти (по таблице) корень $x = x_\beta$ уравнения $\Phi(x) = \beta$. Тогда из (3) вытекает, что *вероятность неравенства*

$$|\bar{\xi}_N - a| < x_\beta \sqrt{D\xi/N} \quad (4)$$

приблизительно равна β .

Чаше других используют коэффициент доверия $\beta = 0,997$, которому отвечает $x_\beta = 3$, или $\beta = 0,95$, которому отвечает $x_\beta = 1,96$. (Значение $x_\beta = 3$ соответствует так называемому «правилу трех сигм», ибо случайная величина $\bar{\xi}_N$ приближенно нормальна и ее среднее квадратичное отклонение $\sigma = \sqrt{D\xi/N}$.)

1.3. Вероятная ошибка метода. Несколько иной подход к оценке ошибки связан с понятием *вероятной ошибки*

$$r_N = 0,6745 \sqrt{D\xi/N}. \quad (5)$$

Численный множитель 0,6745, фигурирующий в (5), — это значение x_β , отвечающее $\beta = 0,50$. Название «вероятная ошибка» вызвано тем, что

$$P\{|\bar{\xi}_N - a| < r_N\} \approx 1/2 \approx P\{|\bar{\xi}_N - a| > r_N\},$$

т. е. одинаково вероятны ошибки, большие чем r_N , и ошибки, меньшие чем r_N .

Величина r_N часто используется на практике для характеристики порядка ошибки: действительная ошибка $|\bar{\xi}_N - a|$ зависит от использованных в расчете случайных чисел и может оказаться в 2—3 раза больше, чем r_N , но может быть и меньше. Таким образом, используя r_N , мы оцениваем порядок ошибки, а используя (4) — верхнюю границу ошибки (с коэффициентом доверия β).

1.4. Эмпирическая оценка дисперсии. Как правило, когда мы приступаем к расчету $M\xi$, значение дисперсии $D\xi$ неизвестно. Хорошую теоретическую оценку для $D\xi$ удастся получить редко. Однако в большинстве задач величину $D\xi$ нетрудно оценить эмпирически, в ходе расчетов a . В самом деле, достаточно одновременно с вычислением $\Sigma \xi_i$ вычислять также $\Sigma (\xi_i)^2$; так как при больших N

$$(1/N) \sum_{i=1}^N (\xi_i)^2 \approx M(\xi^2),$$

то из (2) видно, что

$$D\xi \approx (1/N) \sum_{i=1}^N (\xi_i)^2 - \left[(1/N) \sum_{i=1}^N \xi_i \right]^2. \quad (6)$$

Формулу (6) постоянно используют на практике.

Правда, ниже в п. 1.7 доказано, что при небольших N более точна формула

$$D\xi \approx \frac{1}{N-1} \sum_{i=1}^N (\xi_i)^2 - \frac{1}{N(N-1)} \left(\sum_{i=1}^N \xi_i \right)^2, \quad (7)$$

отличающаяся от предыдущей множителем $(1-1/N)$. Но в расчетах, выполняемых методами Монте-Карло, всегда $N \gg 10$, и разница между (7) и (6) невелика. К тому же надо иметь в виду, что $D\xi$ используется только для оценки ошибки, так что погрешность порядка 10% в значении $D\xi$ роли не играет.

1.5. Оценка ошибки без расчета дисперсии. Будем по-прежнему считать, что дисперсия $D\xi$ конечна. Допустим, что по каким-либо причинам (иногда из-за отсутствия места во внутреннем накопителе ЭВМ) мы не можем (или не хотим) одновременно с $\Sigma \xi_i$ вычислять также $\Sigma (\xi_i)^2$. Оценку погрешности $\bar{\xi}_N - a$ тем не менее можно получить.

Предположим, что $N = mN_1$, где m — небольшое натуральное число $m \geq 3$, а N_1 настолько велико, что распределение случайной величины

$$\zeta = (1/N_1) \sum_{i=1}^{N_1} \xi_i \quad (8)$$

(где ξ_1, \dots, ξ_{N_1} — независимые случайные величины, распределение которых совпадает с распределением ξ) можно считать близким к нормальному (по той же центральной предельной теореме). Очевидно, $M\xi = a$.

Вместо того чтобы сразу вычислять $\bar{\xi}_N$, разделим задачу на m «вариантов» и вычислим m величин, которые можно считать независимыми реализациями ξ :

$$\xi_1 = \frac{1}{N_1} \sum_{i=1}^{N_1} \xi_i, \quad \xi_2 = \frac{1}{N_1} \sum_{i=1}^{N_1} \xi_{i+N_1}, \dots,$$

$$\xi_m = \frac{1}{N_1} \sum_{i=1}^{N_1} \xi_{i+N_1(m-1)}.$$

Воспользуемся теперь следующей теоремой Р. Фишера [24, 44].

Если ξ_1, \dots, ξ_m — независимые одинаково распределенные нормальные (гауссовские) случайные величины с математическим ожиданием a , то случайная величина

$$t = \sqrt{m-1} \frac{\bar{x} - a}{s},$$

где

$$\bar{x} = \frac{1}{m} \sum_{k=1}^m \xi_k, \quad s^2 = \frac{1}{m} \sum_{k=1}^m (\xi_k - \bar{x})^2,$$

подчиняется закону распределения Стьюдента с $(m-1)$ -й степенью свободы*). Это означает, что для любых $x_1 < x_2$

$$P\{x_1 < t < x_2\} = \int_{x_1}^{x_2} s_{m-1}(x) dx.$$

*) Плотность распределения Стьюдента с r степенями свободы выражается формулой

$$s_r(x) = C_r \left(1 + \frac{x^2}{r}\right)^{-\frac{r+1}{2}}, \quad -\infty < x < \infty,$$

где нормировочная постоянная

$$C_r = \frac{1}{\sqrt{\pi r}} \left[\Gamma\left(\frac{r+1}{2}\right) \Gamma\left(\frac{r}{2}\right) \right].$$

В нашем случае величины ξ_1, \dots, ξ_m приближенно нормальные; $M\xi_k = a$; $\bar{x} = \bar{\xi}_N$;

$$s^2 = (1/m) \sum_{k=1}^m (\xi_k - \bar{\xi}_N)^2. \quad (9)$$

Выбрав $x_2 = -x_1 = x$, получим приближенное равенство

$$P\{|\bar{\xi}_N - a| < x \sqrt{s^2/(m-1)}\} \approx 2 \int_0^x s_{m-1}(y) dy,$$

которое будет тем точнее, чем ближе распределение ξ к нормальному.

На стр. 294 приведена таблица корней $x = t_{m,\beta}$ уравнения

$$2 \int_0^x s_m(y) dy = \beta.$$

Если задан коэффициент доверия β , то по этой таблице нетрудно найти соответствующее значение $x = t_{m-1,\beta}$ и записать окончательную оценку: вероятность неравенства

$$|\bar{\xi}_N - a| < t_{m-1,\beta} \sqrt{s^2/(m-1)} \quad (10)$$

приблизительно равна β .

Оценка (10) подобна оценке (4), но вместо неизвестной дисперсии $D\xi$ сюда входит эмпирическая величина s^2 , которую легко вычислить по формуле (9). Однако, вообще говоря, оценка (4) применима при меньших N .

Неравенство (10) позволяет определить также вероятную ошибку метода (5):

$$r_N = t_{m-1, 0.5} \sqrt{s^2/(m-1)}. \quad (11)$$

Оценка (10) использовалась в некоторых работах автора и, независимо, в работе Г. Герцеля и М. Калоса [127].

1.6. Случай $D\xi = \infty$. Из п. 1.1 вытекает, что бесконечность дисперсии не препятствует приближению $\bar{\xi}_N$ к a . Однако последующие оценки погрешности теряют силу. Из (4) видно, что в случае конечной дисперсии ошибка убывает как $N^{-1/2}$. При $D\xi = \infty$ порядок убывания ошибки оказывается хуже. Поэтому обычно рекомендуется избегать методов расчета, в которых дисперсия осредняемой величины бесконечна.

Тем не менее в некоторых задачах такие методы используются. Оценки погрешности при $D\xi = \infty$ имеются в монографии [33] (см. также статьи [144, 176].)

1.7. Замечание. О некоторых терминах, употребляемых в математической статистике [24, 44].

Независимые реализации ξ_1, \dots, ξ_N случайной величины ξ называются *выборкой*. Предположим, что закон распределения величины ξ зависит от некоторого параметра a . Любая функция от выборочных значений $\varphi(\xi_1, \dots, \xi_N)$, используемая в качестве приближения к a , называется *оценкой* a . Если $M\varphi = a$, то оценка φ называется *несмещенной*. Если $\varphi \xrightarrow{P} a$ при $N \rightarrow \infty$, то оценка φ называется *состоятельной*. При небольших N более существенно отсутствие смещения, при больших — важнее состоятельность.

Очевидно, оценка (1) представляет собой несмещенную и состоятельную оценку математического ожидания a .

Легко доказать, что оценка дисперсии

$$(1/N) \sum_{i=1}^N \xi_i^2 - \bar{\xi}_N^2,$$

фигурирующая в (6), смещенная. В самом деле,

$$M \left[(1/N) \sum_{i=1}^N \xi_i^2 - \bar{\xi}_N^2 \right] = M(\xi^2) - M(\bar{\xi}_N^2) = D\xi - D\bar{\xi}_N,$$

ибо $M\xi = M\bar{\xi}_N$. Из (1) следует, что $D\bar{\xi}_N = D\xi/N$. Поэтому

$$M \left[(1/N) \sum_{i=1}^N \xi_i^2 - \bar{\xi}_N^2 \right] = (1 - 1/N) D\xi.$$

Несмещенную оценку дисперсии дает формула (7).

Рассмотрим две функции $\varphi'(\xi_1, \dots, \xi_N; \beta)$ и $\varphi''(\xi_1, \dots, \xi_N; \beta)$. Интервал (φ', φ'') называется *доверительным интервалом для параметра a с коэффициентом доверия β* , если вероятность неравенства $\varphi' < a < \varphi''$ равна β :

$$P\{\varphi' < a < \varphi''\} = \beta.$$

В приведенных выше оценках (4) и (10) использованы доверительные интервалы для среднего значения a нормальной случайной величины: в первом случае — при известной дисперсии, во втором — при неизвестной. Так как величины $\bar{\xi}_N$ и ξ_k лишь приближенно (асимптотически) нормальны, то вероятности неравенств

$$\bar{\xi}_N - x_\beta \sqrt{D\xi/N} < a < \bar{\xi}_N + x_\beta \sqrt{D\xi/N}$$

и соответственно

$$\bar{\xi}_N - t_{m-1, \beta s} \sqrt{m-1} < a < \bar{\xi}_N + t_{m-1, \beta s} \sqrt{m-1}$$

лишь приближенно равны β

§ 2. Простейший метод Монте-Карло для вычисления интеграла

2.1. Простейший метод Монте-Карло. Обозначим через G произвольную область (ограниченную или неограниченную, связную или несвязную) плоскости x, y . Точки плоскости будем обозначать одной буквой $P = (x, y)$, а элемент площади $dP = dx dy$. Рассмотрим задачу о приближенном вычислении интеграла

$$I = \int_G f(P) p(P) dP, \quad (12)$$

где $p(P)$ — некоторая заданная плотность вероятностей, определенная в G , так что $\int_G p(P) dP = 1$.

Заметим сразу, что любой интеграл

$$\int_G f(P) dP$$

по ограниченной области G можно считать интегралом вида (12). Действительно, если площадь G обозначить через S_G , то $p_1(P) \equiv 1/S_G$ при $P \in G$ представляет собой плотность вероятностей случайной точки, равномерно распределенной в G . Если ввести функцию $f_1(P) \equiv S_G \cdot f(P)$ то, очевидно,

$$\int_G f(P) dP = \int_G f_1(P) p_1(P) dP.$$

Чтобы построить метод Монте-Карло для расчета интеграла (12), рассмотрим случайную точку Q с плотностью $p(P)$ и введем скалярную случайную величину $Z = f(Q)$, математическое ожидание которой равно искомому значению интеграла

$$MZ = \int_G f(P) p(P) dP = I. \quad (13)$$

Для расчета MZ можно использовать оценку (1).

Итак, если Q_1, \dots, Q_N — независимые реализации случайной точки Q и $Z_1 = f(Q_1), \dots, Z_N = f(Q_N)$, то оценкой интеграла (12) служит величина

$$\theta_N = (1/N) \sum_{i=1}^N Z_i. \quad (14)$$

Очевидно, $M\theta_N = I$, и если существует $M|Z|$, то, согласно п. 1.1, оценка $\theta_N \xrightarrow{P} I$. В рассматриваемом случае

$$M|Z| = \int_G |f(P)| p(P) dP.$$

Следовательно, если интеграл (12) сходится абсолютно, то θ_N сходится по вероятности к I .

Пример. Требуется вычислить интеграл

$$I = \int_0^{\infty} f(x) e^{-kx} dx, \quad \text{где } k > 0.$$

Выберем плотность $p(x) = k e^{-kx}$ и функцию $f_1 = k^{-1} f(x)$. Если ξ_i — значение случайной величины ξ с плотностью $p(x)$, то оценка интеграла

$$\theta_N = N^{-1} \sum_{i=1}^N f_1(\xi_i) = (kN)^{-1} \sum_{i=1}^N f(\xi_i).$$

Находить значения ξ можно по формуле (7) гл. 2. Поэтому формулу для вычисления I можно записать в виде

$$I \approx (kN)^{-1} \sum_{i=1}^N f(-k^{-1} \ln \gamma_i),$$

где $\gamma_1, \dots, \gamma_N$ — независимые случайные числа.

В дальнейшем вычисляются только абсолютно сходящиеся интегралы. Однако метод Монте-Карло позволяет вычислять и условно сходящиеся интегралы, если преобразовать их надлежащим образом (см. упражнение 5 гл. 3).

2.2. Геометрический метод Монте-Карло. Предположим, что в области G

$$0 \leq f(P) \leq c. \quad (15)$$

В трехмерном пространстве x, y, z рассмотрим цилиндрическую область $G = G \times (0, c)$ (рис. 34), а в G рассмотрим случайную точку \tilde{Q} с плотностью

$\tilde{p}(x, y, z) = (1/c)p(x, y)$. Очевидно, проекция точки \tilde{Q} на плоскость x, y представляет собой случайную

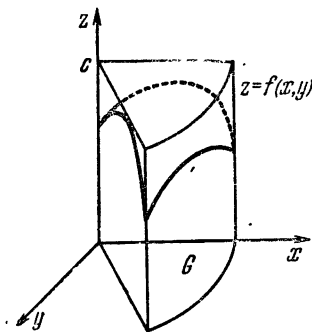


Рис. 34.

точку $Q = (\xi, \eta)$ из п. 2.1 (с плотностью $p(x, y)$), а третья координата \tilde{Q} , назовем ее ξ , не зависит от ξ и η и равномерно распределена в интервале $0 < z < c$, так что ее плотность $p_{\xi}(z) = 1/c$.

Выберем N независимых реализаций $\tilde{Q}_1, \dots, \tilde{Q}_N$ случайной точки \tilde{Q} ; обозначим через ν количество точек, оказавшихся ниже поверхности $z = f(P)$, и составим оценку

$$\tilde{\theta}_N = c\nu/N. \quad (16)$$

Дискретная случайная величина ν подчиняется распределению Бернулли $P\{\nu = m\} = C_N^m p^m (1-p)^{N-m}$ ($m = 0, 1, \dots, N$), где p — вероятность того, что точка \tilde{Q} окажется ниже поверхности $z = f(P)$. Вычислить эту вероятность нетрудно:

$$p = P\{\xi < f(\xi, \eta)\} = \int_G dx dy \int_0^{f(x,y)} \tilde{p}(x, y, z) dz = (1/c)I.$$

Так как $M\nu = Np = (1/c)NI$, то из (16) вытекает, что $M\tilde{\theta}_N = I$. Сходимость $\tilde{\theta}_N \xrightarrow{P} I$ следует из известной теоремы Бернулли о сходимости частот к вероятностям.

Впрочем, оценку (16) также можно представить в форме (1). Введем случайную величину \tilde{Z} , зависящую от точки $\tilde{Q} = (\xi, \eta, \xi)$:

$$\tilde{Z} = \begin{cases} c, & \text{если } \xi < f(\xi, \eta), \\ 0, & \text{если } \xi \geq f(\xi, \eta). \end{cases}$$

Если точкам $\tilde{Q}_1, \dots, \tilde{Q}_N$ соответствуют значения $\tilde{Z}_1, \dots, \tilde{Z}_N$, то

$$\tilde{\theta}_N = (1/N) \sum_{i=1}^N \tilde{Z}_i. \quad (17)$$

И поэтому утверждения о том, что $M\tilde{\theta}_N = I$ и $\tilde{\theta}_N \xrightarrow{P} I$ вытекают также из результатов § 1. Абсолютная сходимость интеграла (12) следует из ограничения (15).

Геометрический метод представляет собой обобщение метода вычисления объема, рассмотренного во введении. В самом деле, если

область G ограничена и $p(P) \equiv 1/S_G$ при $P \in G$, то при больших N

$$v/N \approx I/c = V/V_{\tilde{G}},$$

где $V = \int_G f(P) dP$ — объем части \tilde{G} , ограниченной сверху поверхностью $z=f(x, y)$, а $V_{\tilde{G}} = cS_G$ — объем всей цилиндрической области \tilde{G} .

2.3. Сравнение точности методов Монте-Карло.

В оценке (14) фигурируют значения случайной величины $Z=f(Q)$, где Q — случайная точка с плотностью $p(P)$. Так как для существования дисперсии DZ необходимо и достаточно, чтобы существовал второй момент

$$M(Z^2) = \int_G f^2(P) p(P) dP \quad (18)$$

(ср. (2)), то условием применимости оценок погрешности § 1 в рассматриваемом случае служит существование интеграла (18)*).

Если через L_2 обозначить множество функций $f(P)$, для которых интеграл (18) сходится, то требование сходимости этого интеграла можно записать в форме $f(P) \in L_2$. В тех случаях, когда важно указать область G и плотность $p(P)$, будем писать, что $f(P) \in L_2(G; p)$.

Итак, если $f(P) \in L_2(G; p)$, то дисперсия осредняемой величины Z в простейшем методе п. 2.1 конечна:

$$DZ = \int_G f^2(P) p(P) dP - I^2. \quad (19)$$

Рассмотрим теперь геометрический метод п. 2.2. В оценке (17) осредняются значения случайной величины \tilde{Z} , для которой

$$M(\tilde{Z}^2) = c^2 P\{\xi < f(\xi, \eta)\} = cI,$$

так что дисперсия

$$D\tilde{Z} = cI - I^2. \quad (20)$$

*) Из сходимости интеграла (18) следует абсолютная сходимость интеграла (12), ибо

$$\left[\int_G |f(P)| p(P) dP \right]^2 \leq \int_G f^2(P) p(P) dP.$$

Сравним величины (19) и (20): если $0 \leq f(P) \leq c$, то

$$\int_G f^2(P) p(P) dP \leq c \int_G f(P) p(P) dP = cI,$$

следовательно,

$$DZ \leq D\tilde{Z}. \quad (21)$$

Неравенство (21) показывает, что в каком-то отношении простейший метод Монте-Карло лучше геометрического метода: при одинаковом количестве N осредняемых величин вероятная ошибка оценки (14) будет не больше, чем вероятная ошибка оценки (17). Можно сказать, что *точность метода Монте-Карло зависит от дисперсии осредняемой случайной величины. И простейший метод всегда точнее геометрического.*

Пример. Требуется вычислить интеграл

$$I = \int_0^1 e^x dx.$$

Оценки (14) и (16) в этом случае равны

$$\theta_N = (1/N) \sum_{i=1}^N e^{\gamma_i}, \quad \tilde{\theta}_N = e\nu/N,$$

где ν — количество пар $(\gamma_1, \gamma'_1), \dots, (\gamma_N, \gamma'_N)$ таких, что $e\gamma'_i < e^{\gamma_i}$ (как всегда $\gamma_1, \dots, \gamma_N, \gamma'_1, \dots, \gamma'_N$ — независимые случайные числа). По формуле (19) вычислим дисперсию DZ :

$$DZ = \int_0^1 e^{2x} dx - I^2 = (1/2)(e^2 - 1) - (e - 1)^2 = 0,2420.$$

По формуле (20) вычислим дисперсию $D\tilde{Z}$:

$$D\tilde{Z} = eI - I^2 = e - 1 = 1,7183.$$

2.4. Сравнение трудоемкости алгоритмов Монте-Карло. Из результатов предыдущего пункта напрашивается вывод, что простейший метод Монте-Карло всегда выгоднее геометрического. Однако такой вывод был бы слишком поспешным: хотя вероятная ошибка оценки θ_N и меньше вероятной ошибки $\tilde{\theta}_N$ при одинаковых N , но время, затрачиваемое на расчет θ_N , может оказаться гораздо больше времени, затрачиваемого на расчет $\tilde{\theta}_N$.

И тогда оценка $\tilde{\theta}_N$ окажется практически более эффективной, чем θ_N .

Обычно говорят, что «построен метод Монте-Карло» для расчета интеграла I тогда, когда определена оценка вида (1), приближающая I :

$$I \approx (1/N) \sum_{i=1}^N \xi_i;$$

при этом обычно предполагается, что $M\xi = I$. Условимся говорить, что «построен алгоритм Монте-Карло», соответствующий данному методу, если указаны также формулы для моделирования используемых в этом методе случайных величин. Например, оценка (14) и формула $Z=f(Q)$ определяют метод Монте-Карло; чтобы получить алгоритм Монте-Карло, надо указать еще формулы для расчета точки Q по стандартным случайным числам. Как мы видели в гл. 2, это можно делать различными способами, так что один и тот же метод Монте-Карло может быть реализован в виде различных алгоритмов.

Рассмотрим два алгоритма Монте-Карло (безразлично соответствуют ли они разным методам или одному) для вычисления интеграла I . Пусть в одном алгоритме осредняются значения случайной величины $\xi' = \xi'(\gamma_1, \dots, \gamma_n)$, а во втором $\xi'' = \xi''(\gamma_1, \dots, \gamma_n)$, так что соответствующие оценки интеграла равны

$$\theta'_N = (1/N) \sum_{i=1}^N \xi'_i \quad \text{и} \quad \theta''_N = (1/N) \sum_{i=1}^N \xi''_i. \quad (22)$$

Обозначим дисперсии осредняемых величин через $D' = D\xi'$ и $D'' = D\xi''$, и пусть t' и t'' — время, затрачиваемое на расчет одного значения ξ' и соответственно ξ'' . Естественно считать более эффективным тот алгоритм, который требует меньшего времени счета для достижения заданной вероятной ошибки.

Для того чтобы вероятная ошибка для первого из сравниваемых алгоритмов равнялась $r_{N'} = \varepsilon$, количество N' осредняемых значений ξ' , согласно (5), должно равняться

$$N' = D' (0,6745/\varepsilon)^2.$$

Время счета при этом окажется равным

$$t'N' = t'D'(0,6745/\epsilon)^2.$$

Для второго алгоритма время счета, необходимое для достижения той же вероятной ошибки ϵ , равно

$$t''N'' = t''D''(0,6745/\epsilon)^2.$$

Назовем *трудоемкостью* алгоритма Монте-Карло произведение $tD\xi$ дисперсии осредняемой случайной величин ξ на время t расчета одного значения ξ . Тогда время счета, необходимое для достижения заданной вероятной ошибки, окажется пропорциональным трудоемкости алгоритма. Ясно, что из двух алгоритмов более эффективен менее трудоемкий.

Так как трудоемкость используется обычно для сравнения алгоритмов, то вычислять ее можно в любых единицах, в частности, t можно выражать в секундах, а можно — в количествах элементарных операций. Конечно, надо предполагать, что сравниваемые алгоритмы реализуются с помощью одних и тех же вычислительных средств. Кстати, не исключена возможность того, что при переходе к другим средствам вычисления менее трудоемкий алгоритм станет более трудоемким.

Пример. Вычислить интеграл

$$I = \int_0^1 \sqrt[5]{x} dx = \frac{5}{6}.$$

Можно считать, что под интегралом стоит также плотность $p(x) \equiv 1$ случайной величины γ . Тогда простейший метод приводит к оценке

$$I \approx (1/N) \sum_{i=1}^N \sqrt[5]{\gamma_i}.$$

Дисперсия осредняемой случайной величины $Z = (\gamma)^{1/5}$ равна

$$DZ = \int_0^1 x^{2/5} dx - \left(\frac{5}{6}\right)^2 = \frac{5}{252}.$$

Так как $0 \leq x^{1/5} \leq 1$, то можно воспользоваться также геометрическим методом:

$$I \approx (1/N) \sum_{i=1}^N \tilde{Z}_i,$$

где $\tilde{Z}_i = 1$, если $\gamma'_i < (\gamma_i)^{1/5}$ и $\tilde{Z}_i = 0$ в противном случае. Дисперсию

осредняемой величины легко вычислить по формуле (20):

$$D\tilde{Z} = \frac{5}{6} - \left(\frac{5}{6}\right)^2 = \frac{5}{36}.$$

Таким образом, в полном согласии с (21), $DZ < D\tilde{Z}$.

Оценим количество операций, затрачиваемых на расчет одного значения Z_i или \tilde{Z}_i . Для расчета Z_i надо найти одно случайное число γ_i — три операции — и вычислить $(\gamma_i)^{1/5}$. На ЭВМ вычисление $x^{1/5}$ осуществляется путем логарифмирования с помощью стандартных программ $\ln x$ и $\exp x$. Общее количество используемых в таком расчете элементарных операций $t \approx 70$.

Для расчета \tilde{Z}_i надо найти два случайных числа γ_i и γ'_i — шесть операций — и вместо неравенства $\gamma'_i < (\gamma_i)^{1/5}$ проверить равносильное неравенство $(\gamma'_i)^5 < \gamma_i$. Общее количество элементарных операций в таком расчете $t \approx 10$.

Значит, $t \cdot DZ \approx 1,4$, $tD\tilde{Z} \approx 1,4$ и второй алгоритм оказывается не более трудоемким, чем первый, несмотря на то, что дисперсия $D\tilde{Z}$ в семь раз больше, чем DZ .

Однако тот же простейший метод можно реализовать с помощью другого алгоритма (см. п. 4.1 гл. 2).

$$I \approx (1/N) \sum_{i=1}^N \max(\gamma_i; \gamma'_i; \gamma''_i; \gamma'''_i; \gamma^{IV}_i);$$

тогда $t \approx 20$ и $t \cdot DZ \approx 0,40$, так что трудоемкость этого алгоритма оказывается заметно меньшей.

§ 3. Важнейшие способы построения хороших оценок (способы уменьшения дисперсии)

Из формулы (5) видно, что вероятная ошибка оценки (1) пропорциональна $\sqrt{D\xi/N}$. Скорость убывания этой ошибки с ростом N невелика. Поэтому очень важно научиться выбирать для расчета интегралов такие вычислительные схемы или, другими словами, такие случайные величины ξ , для которых дисперсия $D\xi$ по возможности мала. Часто способы построения таких схем называют способами уменьшения дисперсии, имея в виду, что для этих способов дисперсия должна быть меньше, чем дисперсия простейшего метода Монте-Карло (п. 2.1).

3.1. Частичное аналитическое интегрирование. Если часть задачи можно решить аналитически, то, используя это частичное решение, обычно удается построить метод

Монте-Карло для решения всей задачи с дисперсией, меньшей, чем (19). Правда, вообще говоря, построенные таким путем методы могут оказаться более трудоемкими и в конечном счете невыгодными.

Мы будем говорить об аналитическом интегрировании, хотя в некоторых случаях такую же роль может сыграть численное интегрирование, если точность этого интегрирования значительно выше, чем точность метода Монте-Карло.

3.1.1. Выделение главной части. Это весьма очевидный и весьма общий принцип, относящийся ко всем методам Монте-Карло: если главную часть задачи можно вычислить аналитически, то, как правило, выгодно считать методом Монте-Карло не всю задачу, а только «поправку» — разницу между всей задачей и главной частью. Уменьшение дисперсии при этом может оказаться очень значительным.

Пусть, например, требуется вычислить интеграл (12)

$$I = \int_G f(P) p(P) dP,$$

где $f(P) \in L_2(G; p)$, и имеется функция $h(P) \in L_2(G; p)$, «близкая» к $f(P)$, такая, что значение интеграла

$$\int_G h(P) p(P) dP = C$$

известно. Тогда вместо оценки (14) можно использовать оценку

$$\theta'_N = C + (1/N) \sum_{i=1}^N [f(Q_i) - h(Q_i)], \quad (23)$$

ибо математическое ожидание осредняемой величины

$$Z' = C + f(Q) - h(Q)$$

равно $MZ' = C + I - C = I$. Дисперсия Z' в этом случае равна

$$DZ' = \int_G [f(P) - h(P)]^2 p(P) dP - (I - C)^2.$$

Если $h(P)$ настолько близка к $f(P)$, что

$$\int_G [f(P) - h(P)]^2 p(P) dP \leq \varepsilon,$$

то, очевидно, и $DZ' \leq \varepsilon$.

Пример. Вычислить интеграл

$$I = \int_0^1 e^x dx,$$

рассмотренный в п. 2.3. Так как $e^x = 1 + x + \dots$, то выберем $h(x) = x$ (постоянное слагаемое на дисперсию не влияет). Согласно (23) получим расчетную формулу

$$\theta'_N = (1/2) + (1/N) \sum_{i=1}^N (e^{\gamma_i} - \gamma_i).$$

Дисперсия осредняемой величины $Z' = 1/2 + e^{\gamma} - \gamma$ равна

$$DZ' = \int_0^1 (e^x - x)^2 dx - (I - 1/2)^2 = (e - 1)(5 - e)/2 - \\ - 23/12 = 0,0437$$

и значительно меньше, чем DZ и $D\tilde{Z}$ в п. 2.3.

Общие правила выделения главной части указать трудно: в различных задачах это делается по-разному. Например, при расчете больших сцинтилляционных детекторов необходимо учитывать многократно рассеянные нейтроны (их вклад в световыход составляет до 10%). В. Г. Золотухину с сотрудниками [37] удалось добиться значительного увеличения точности метода Монте-Карло путем аналитического учета вклада нерассеянных и однократно рассеянных нейтронов.

В ряде областей физики используют теорию возмущений, которая позволяет оценить решение сложной задачи по известному решению «близкой» простой задачи. Если рассчитывать методом Монте-Карло не всю сложную задачу, а только интегралы теории возмущений, то можно значительно повысить точность результатов [57].

3.1.2. Интегрирование по части области. Допустим, что мы умеем (аналитически) вычислить интегралы по некоторой части B области G :

$$\int_B f(P) p(P) dP = C, \quad \int_B p(P) dP = c,$$

где $0 < c < 1$. Докажем, что, как правило, всегда выгодно представить интеграл (12) в виде суммы

$$I = \int_{G_1} f(P) p(P) dP + C \quad (24)$$

и вычислять простейшим методом только интеграл по области $G_1 = G - B$. (Если C близко к I , то можно считать, что мы выделяем главную часть; по данный прием выгоден и тогда, когда область B заметно меньше, чем G (рис. 35); правда, и понижение дисперсии в этом случае будет заметно меньше.)

В области G_1 определим плотность $p_1(P) = p(P)/(1-c)$ и рассмотрим случайную величину

$$Z' = C + (1-c)f(Q'),$$

где Q' — случайная точка с плотностью $p_1(P)$ в G_1 . Легко видеть, что $MZ' = I$. Поэтому для расчета I можно использовать оценку

$$\theta'_N = C + (1/N)(1-c) \sum_{i=1}^N f(Q'_i), \quad (25)$$

где Q'_1, \dots, Q'_N — независимые реализации точки Q' . Чтобы сравнить эту оценку с оценкой (14), сравним дисперсии DZ' и DZ , где $Z = f(Q)$.

Теорема 1. Если существует дисперсия DZ , то

$$DZ' \leq (1-c)DZ. \quad (26)$$

Доказательство. Согласно определению дисперсии

$$DZ = \int_G f^2 p dP - I^2 = \int_{G_1} f^2 p dP + \int_B f^2 p dP - I^2,$$

$$\begin{aligned} DZ' &= (1-c)^2 \int_{G_1} f^2 p_1 dP - \left[(1-c) \int_{G_1} f p_1 dP \right]^2 = \\ &= (1-c) \int_{G_1} f^2 p dP - \left[\int_{G_1} f p dP \right]^2. \end{aligned}$$

Умножив DZ на $1-c$ и вычтя DZ' , получим, что

$$(1-c)DZ - DZ' = (1-c) \int_B f^2 p dP - (1-c)I^2 + (I-C)^2.$$

Оставшийся интеграл по области B выразим через неотрицательную величину

$$b^2 \equiv \int_B (f - C/c)^2 p dP = \int_B f^2 p dP - C^2/c.$$

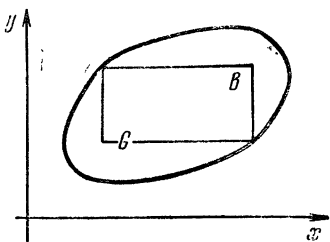


Рис. 35.

Тогда окажется, что

$$(1 - c)DZ - DZ' = (1 - c)b^2 + (\sqrt{c}I - C/\sqrt{c})^2 \geq 0,$$

что и требовалось доказать.

Предположим, что задан какой-либо алгоритм расчета точек Q в G . Тогда трудоемкость алгоритма (14) равна $(t_q + t_f)DZ$, где t_q — время расчета одной точки Q , а t_f — время расчета одного значения $f(Q)$.

Рассмотрим оценку (25). Если никакого более удобного способа моделирования точек Q' в G_1 нет, то можно отбирать точки Q' среди точек Q (п. 5.2 гл. 2). Эффективность такого отбора $\varepsilon = P\{Q \in G_1\} = 1 - c$. Поэтому время расчета одной точки Q' в среднем равно $t_{q'} = (t_q + t_o)/(1 - c)$, где t_o — время, затрачиваемое на отбор (т. е. на проверку условия $Q \in G_1$). Трудоемкость алгоритма (25) в этих условиях равна $(t_{q'} + t_f)DZ'$.

Легко доказать, что если

$$t_o \leq ct_f, \quad (27)$$

то трудоемкость алгоритма (25) не больше, чем трудоемкость алгоритма (14). В самом деле,

$$(t_{q'} + t_f)DZ' \leq [(t_q + ct_f)/(1 - c) + t_f]DZ' \leq (t_q + t_f)DZ.$$

Условие (27) легко проверяется на практике. Если область B «простая», а функция $f(P)$ «сложная», то, очевидно, $t_f \gg t_o$.

Пример. Требуется вычислить объем фигуры V , ограниченной поверхностью (в сферических координатах)

$$r = a + hu(\varphi, \theta),$$

где $-1 \leq u(\varphi, \theta) \leq 1$ (рис. 36).

Обозначим через V_0 и V_1 вписанный в V и описанный около V шары, радиусы которых равны $a - h$ и $a + h$. Объемы V_0 , V , V_1 будем обозначать теми же буквами. Воспользуемся геометрическим методом: вы-

берем случайные точки Q_1, \dots, Q_N , равномерно распределенные в V_1 , и если v из этих точек попадут внутрь V , то будем считать, что объем V приближенно равен

$$\bar{V}_N = V_1 (v/N).$$

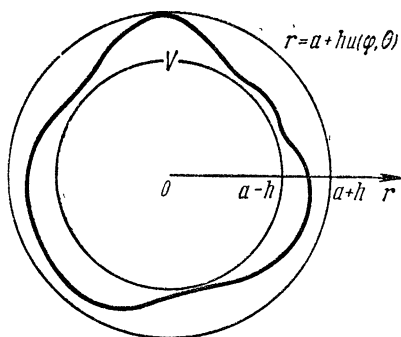


Рис. 36.

Выделим теперь объем шара V_0 . Для этого достаточно выбрать случайные точки Q'_1, \dots, Q'_N , равномерно **распределенные** в шаровом слое $V_1 - V_0$; если v' из этих точек принадлежит V , то объем V приближенно равен

$$\tilde{V}'_N = V_0 + (V_1 - V_0) (v'/N).$$

Вместо сравнения дисперсий осредняемых величин сравним в этом примере дисперсии самих оценок. Так как v и v' подчиняются биномиальным распределениям с параметрами $p = V/V_1$ и соответственно $p' = (V - V_0)/(V_1 - V_0)$, то $Dv = Np(1-p)$, $Dv' = Np'(1-p')$. Поэтому $D\tilde{V}_N = (1/N)V(V_1 - V)$, $D\tilde{V}'_N = (1/N)(V - V_0)(V_1 - V)$. Очевидно, всегда $D\tilde{V}'_N < D\tilde{V}_N$.

Если отношение $\varepsilon = h/a \ll 1$, то в оценке \tilde{V}'_N мы фактически выделяем главную часть задачи, и уменьшение дисперсии $D\tilde{V}'_N$ по сравнению с $D\tilde{V}_N$ должно быть особенно заметным. Действительно, так как $V_0 = (4/3)\pi(a-h)^3$, $V_1 = (4/3)\pi(a+h)^3$, то можно записать, что $V = (4/3)\pi(a+u_0h)^3$, где $|u_0| < 1$. Тогда нетрудно сосчитать, что величина

$$V(V_1 - V) = [(4/3)\pi a^3]^2 3(1 - u_0)\varepsilon + O(\varepsilon^2)$$

пропорциональна ε , а величина

$$(V - V_0)(V_1 - V) = [(4/3)\pi a^3]^2 9(1 - u_0^2)\varepsilon^2 + O(\varepsilon^3)$$

— второго порядка малости.

Наконец, покажем, что если моделировать точки Q и Q' в сферических координатах (п. 2.4.1 гл. 2), то алгоритмы, соответствующие обоим рассмотренным методам, примерно одинаково сложны: $t \approx t'$. В самом деле, в обоих случаях для i -го испытания нужны три случайных числа γ_i , γ'_i , γ''_i . Координаты точек Q_i и Q'_i вычисляются соответственно по формулам

$$r_i = (a+h)\sqrt[3]{\gamma_i}, \quad \varphi_i = 2\pi\gamma'_i, \quad \cos \theta_i = 2\gamma''_i - 1$$

или

$$r_i = (a+h)\sqrt[3]{\gamma_i + b(1-\gamma_i)}, \quad \varphi_i = 2\pi\gamma'_i, \quad \cos \theta_i = 2\gamma''_i - 1,$$

где $b = (a-h)^3/(a+h)^3$. Условие принадлежности точки Q_i или Q'_i объему V проверяется одинаково: $r_i < a+hu$ (φ_i, θ_i). Если это условие выполнено, то к счетчику v (соответственно v') добавляется единица.

3.1.3. Интегрирование по части переменных (понижение порядка интеграла). Докажем, что если аналитически взять интеграл по некоторым из переменных, а по остальным переменным использовать тот же метод Монте-Карло, то дисперсия уменьшится [143]. Правда, в отличие от случая, рассмотрен-

ного в п. 3.1.2, нередко бывает, что после интегрирования по некоторым из переменных получаются более сложные формулы счета и, несмотря на уменьшение дисперсии, трудоемкость возрастает.

Перейдем к точной формулировке задачи. Пусть требуется вычислить интеграл

$$I = \int_G dP \int_{G'} f(P, P') p(P, P') dP',$$

где $p(P, P')$ — совместная плотность вероятностей случайных точек $Q \in G$ и $Q' \in G'$; $\int_G \int_{G'} p(P, P') dP dP' = 1$.

Если интеграл этот вычислять простейшим методом, то осредняемая случайная величина $Z = f(Q, Q')$.

Предположим, что по переменному P' мы умеем выполнить интегрирование и можем вычислить плотность точки Q :

$$p_1(P) = \int_{G'} p(P, P') dP',$$

а также функцию

$$f_1(P) = \int_{G'} f(P, P') p(P, P') dP' [p_1(P)]^{-1}.$$

Тогда, очевидно,

$$I = \int_G f_1(P) p_1(P) dP,$$

и если вычислять простейшим методом этот интеграл, то придется осреднять величину $Z' = f_1(Q)$.

Теорема 2. Если дисперсия DZ конечна, то $DZ' \leq \leq DZ$.

Доказательство. Воспользуемся неравенством (1) на стр. 292.

$$[p_1(P) f_1(P)]^2 = \left[\int_{G'} f \sqrt{Vp} \cdot \sqrt{Vp} dP' \right]^2 \leq \int_{G'} f^2 p dP' \int_{G'} p dP'.$$

Отсюда следует, что

$$f_1^2(P) p_1(P) \leq \int_{G'} f^2(P, P') p(P, P') dP'.$$

Пронтегрировав это неравенство по P , получим, что

$$\int_G f_1^2(P) p_1(P) dP \leq \int_G dP \int_{G'} f^2(P, P') p(P, P') dP',$$

или, что то же, $M(Z'^2) \leq M(Z^2)$. Так как $MZ' = MZ = I$, то тем самым доказано, что $DZ' \leq DZ$.

Пример. Рассмотрим интеграл

$$I = \iint_B (1/y) dx dy,$$

где область интегрирования B представляет собой треугольник, ограниченный прямыми $y=1$, $x=2$ и $y=x$ (рис. 37). Интеграл этот легко вычисляется:

$$I = \int_1^2 dx \int_1^x y^{-1} dy = \int_1^2 \ln x dx =$$

$$= 2 \ln 2 - 1 = 0,38630.$$

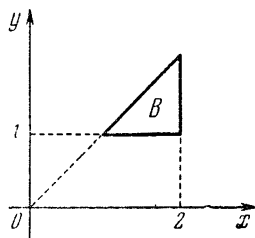


Рис. 37.

А. Введем случайную точку (ξ, η) , плотность которой $p(x, y) \equiv 2$ в B . Тогда $I = MZ$, где $Z = (2\eta)^{-1}$, а дисперсия

$$DZ = \int_1^2 dx \int_1^x 2^{-1} y^{-2} dy - I^2 = (1/2)(7 \ln 2 - 1) - 4 \ln^2 2 = 0,0043.$$

Так как плотность $p_\eta(y) = 2(2-y)$, то из уравнения $F_\eta(\eta) = 1 - \gamma$ получим, что $\eta = 2 - \sqrt{\gamma}$. Следовательно, оценка интеграла I равна

$$0_N = (2N)^{-1} \sum_{i=1}^N (2 - \sqrt{\gamma_i})^{-1}.$$

Б. Пронтегрируем аналитически плотность и подынтегральную функцию по y :

$$p_1(x) = \int_1^x 2 dy = 2(x-1), \quad p_1(x) f(x) = \int_1^x (1/y) dy = \ln x.$$

В этом случае функция распределения ξ равна $F_1(x) = (x-1)^2$. Из уравнения $F_1(\xi) = \gamma$ следует, что $\xi = 1 + \sqrt{\gamma}$. Так как осредняемая величина $Z' = f_1(\xi) = 2^{-1}(\xi-1)^{-1} \ln \xi$, то оценка интеграла

$$0'_N = (2N)^{-1} \sum_{i=1}^N \gamma_i^{-1/2} \ln(1 + \gamma_i^{1/2}).$$

Дисперсия в этом случае равна

$$DZ' = \int_1^2 2^{-1} (x-1)^{-1} \ln^2 x dx - I^2 = 0,15025 - 0,14923 = 0,0010.$$

В. Из первых цифр табл. 4 (стр. 295) образуем десять случайных чисел и вычислим по ним значения θ_{10} и θ'_{10} . Случайные числа: $\gamma_1=0,86515$, $\gamma_2=0,90795$, $\gamma_3=0,66155$, $\gamma_4=0,66434$, $\gamma_5=0,56558$, $\gamma_6=0,12332$, $\gamma_7=0,94337$, $\gamma_8=0,57802$, $\gamma_9=0,69186$, $\gamma_{10}=0,03393$. Сосчитанные по ним значения:

$$\theta_{10}=0,408, \theta'_{10}=0,378.$$

Ошибки этих значений $\theta_{10} - I = 0,022$ и $\theta'_{10} - I = 0,008$ близки к соответствующим вероятным ошибкам $r_{10}=0,014$ и $r'_{10}=0,007$.

3.2. Метод существенной выборки. До сих пор мы рассматривали интегралы вида (12) и использовали при вычислении их случайные точки с плотностью $p(P)$. Предположим теперь, что требуется вычислить абсолютно сходящийся интеграл

$$I_0 = \int_G f(P) dP, \quad (28)$$

где область G может быть как ограниченной, так и неограниченной, и квадрат функции $f(P)$ не обязательно интегрируем. Предполагается только, что $\int_G |f(P)| dP > 0$.

3.2.1. Плотность $p(P)$, определенную в G , назовем допустимой по отношению к $f(P)$, если $p(P) > 0$ в тех точках, в которых $f(P) \neq 0$.

Если $p(P) > 0$ всюду в G , то эта плотность допустима по отношению к любым $f(P)$. Вообще же допустимая плотность может обращаться в нуль, но только там, где $f(P) = 0$. Множество точек, в которых $f(P) = 0$, назовем G_0 и пусть $G^+ = G - G_0$.

Выберем произвольную допустимую плотность $p(P)$ и рассмотрим функцию

$$Z_0(P) = \begin{cases} f(P)/p(P) & \text{при } P \in G^+, \\ 0 & \text{при } P \in G_0. \end{cases}$$

Если Q — случайная точка, определенная в G с плотностью $p(P)$, то

$$MZ_0(Q) = \int_G Z_0(P) p(P) dP = \int_{G^+} f(P) dP = I_0,$$

ибо вне множества G^+ функция $f(P) \equiv 0$.

Согласно п. 2.1 для приближенного расчета I_0 можно использовать независимые реализации Q_1, \dots, Q_N случайной точки Q и оценку

$$\theta_N = (1/N) \sum_{i=1}^N Z_0(Q_i).$$

Вероятная ошибка этой оценки зависит от дисперсии $DZ_0(Q)$, которую нетрудно вычислить: так как

$$DZ_0 = \int_G Z_0^2(P) p(P) dP - I_0^2,$$

то

$$DZ_0 = \int_{G^+} [f^2(P)/p(P)] dP - I_0^2. \quad (29)$$

Величина эта зависит от выбора плотности $p(P)$ и даже не обязательно конечна. Естественнo поставить вопрос о выборе $p(P)$ так, чтобы минимизировать DZ_0 .

Теорема 3. *Минимальная дисперсия DZ_0 реализуется в случае, когда плотность $p(P)$ пропорциональна $|f(P)|$ и равна*

$$D\hat{Z}_0 = \left[\int_G |f(P)| dP \right]^2 - I^2. \quad (30)$$

Доказательство. Если плотность $p(P)$ пропорциональна $|f(P)|$, то она равна

$$\hat{p}(P) = |f(P)| \left[\int_G |f(P)| dP \right]^{-1}. \quad (31)$$

Подставив (31) в (29), получим, что $DZ_0 = D\hat{Z}_0$.

Осталось доказать, что, какова бы ни была допустимая плотность $p(P)$, дисперсия $DZ_0 \geq D\hat{Z}_0$. А это легко доказывается с помощью неравенства (1), стр. 292:

$$\begin{aligned} \left[\int_G |f| dP \right]^2 &= \left[\int_{G^+} |f| dP \right]^2 = \left[\int_{G^+} |f| p^{-1/2} p^{1/2} dP \right]^2 \leq \\ &\leq \int_{G^+} f^2 p^{-1} dP \int_{G^+} p dP \leq \int_{G^+} f^2 p^{-1} dP. \end{aligned}$$

Следствие. Если подынтегральная функция $f(P)$ не меняет знака в G , то, $D\hat{Z}_0 = 0$.

Отметим, что плотность $\hat{p}(P)$ тождественно равна нулю в области G_0 , в которой $f(P) \equiv 0$. Этот результат согласуется с выводом из теоремы 1, что область, в которой $f(P) \equiv 0$, выгодно при интегрировании исключить.

В действительности использовать плотность (31) для расчета интеграла (28) нельзя, ибо в (31) входит значение $\int_G |f(P)| dP$, вычисление которого представляет собой задачу, эквивалентную по трудности исходной задаче (в случае знакопостоянной функции $f(P)$ — в точности эквивалентную). Однако из теоремы 3 можно сделать вывод, что *желательно выбирать плотность $p(P)$ по возможности пропорциональной $|f(P)|$* . Такой метод выбора $p(P)$ часто приводит к величинам Z_0 с небольшими дисперсиями. Он был предложен Г. Каном [142] и называется методом существенной выборки (importance sampling), ибо если $p(P)$ пропорциональна $|f(P)|$, то в тех частях области G , в которых $|f(P)|$ больше и вклад которых в I_0 более существен, будет выбираться больше случайных точек.

Возможна и другая интерпретация целесообразности выбора $p(P)$ пропорционально $f(P)$ (в случае знакопостоянной $f(P)$): чем ближе $Z_0 = f(P)/p(P)$ к постоянной, тем меньше дисперсия $DZ_0(Q)$.

Очень сложные плотности $p(P)$ использовать не рекомендуется, так как тогда процесс реализации случайных точек Q с плотностью $p(P)$ станет очень трудоемким. Заметим, что для вычисления интеграла

$$I_0 = \int_G Z_0(P) p(P) dP$$

не обязательно использовать оценку θ_N : можно применить какой-нибудь из методов п. 3.1.

Пример. Интеграл

$$I = \int_0^1 e^x dx,$$

рассмотренный в пп. 2.3 и 3.1.1, можно представить в виде

$$I = (3/2) \int_0^1 e^x (1+x)^{-1} p(x) dx,$$

где $p(x) = (2/3)(1+x)$. Значения случайной величины ξ с плотностью $p(x)$ вычисляются по формуле $\xi = \sqrt{1+3\gamma} - 1$ (метод обратных функций), а оценка равна

$$\theta_N = (3/2) N^{-1} \sum_{i=1}^N e^{\xi_i} (1 + \xi_i)^{-1}.$$

В этом примере дисперсия осредняемой величины $Z(\xi) = (3/2) e^{\xi} (1 + \xi)^{-1}$ равна

$$DZ = (3/2) \int_0^1 e^{2x} (1+x)^{-1} dx - I^2 = 0,0269.$$

Это гораздо меньше, чем в п. 2.3, и меньше, чем в п. 3.1.1.

3.2.2. Метод существенной выборки позволяет строить хорошие оценки для несобственных интегралов (28).

Предположим, что область G ограничена, но

$$\int_G f^2(P) dP = \infty. \quad (32)$$

В этом случае простейший метод Монте-Карло позволяет записать оценку интеграла (28)

$$\theta_N = (V_G/N) \sum_{i=1}^N f(Q_i),$$

где точки Q_i равномерно распределены в G , а V_G — объем G . Однако оценка эта плохая, ибо $D\theta_N = \infty$. В то же время оценка, полученная методом существенной выборки будет иметь конечную дисперсию, если выбрать допустимую плотность $p(P)$ так, чтобы интеграл, фигурирующий в выражении (29), сходился. Такие плотности всегда существуют. В частности, этому требованию удовлетворяет плотность (31).

На практике, если подынтегральная функция $f(P)$ имеет особенность, то стараются выбрать плотность $p(P)$ с той же особенностью, так, чтобы отношение $f(P)/p(P)$ было ограниченным. Прием этот часто называют *включением особенности в плотность*.

Предположим теперь, что функция $f(P)$ в (28) особенностей не имеет, но область интегрирования G неограничена. Чтобы оценить такой интеграл, можно

выбрать ограниченную область $G_e \subset G$ так, что

$$\left| \int_{G-G_e} f(P) dP \right| < \varepsilon,$$

и строить оценку для интеграла по области G_e . Вместо этого можно попытаться заменой переменных преобразовать G в конечную область. Однако наиболее естественным, по-видимому, надо считать использование метода существенной выборки. И в этом случае также рекомендуется включать особенность в плотность, т. е. выбирать $p(P)$ так, чтобы отношение $f(P)/p(P)$ стремилось к постоянной при $|P| \rightarrow \infty$, $P \in G$.

3.2.3. Один тип интегралов с особенностью. Обозначим через G шар $x^2 + y^2 + z^2 < R^2$. Во многих разделах физики и механики встречаются интегралы вида

$$\iint_{GG} h(P, P') \rho^{-\alpha} dP dP',$$

где обе точки P и P' принадлежат G , а ρ — расстояние между этими точками: $\rho = |P - P'|$.

Рассмотрим интеграл такого типа с особенностью

$$I = \iint_{GG} h(\rho) \rho^{-\alpha} dP dP', \quad (33)$$

где функция $h(\rho)$ ограничена и $h(0) \neq 0$. Интеграл (33) абсолютно сходится при $\alpha < 3$.

А. Для того чтобы вычислить интеграл (33) простейшим методом Монте-Карло, выберем две независимые случайные точки Q и Q' , равномерно распределенные в G . Так как плотности их $p_Q(P) = p_{Q'}(P') = 1/V_G$, то положим $Z = V_G^2 h(\rho) \rho^{-\alpha}$, где $\rho = |Q - Q'|$, $V_G = (4/3)\pi R^3$. Тогда $MZ = I$, а дисперсия Z равна

$$DZ = V_G^2 \iint_{GG} h^2(\rho) \rho^{-2\alpha} dP dP' - I^2.$$

При $1,5 \leq \alpha < 3$ последний интеграл расходится и $DZ = \infty$.

Б. Воспользуемся методом существенной выборки и выберем совместную плотность $p(P, P')$ случайных точек Q и Q' так, чтобы она содержала такую же особенность, как подынтегральная функция. Так как $p(P, P') = p_Q(P) p_{Q'}(P' | P)$, то точку Q будем по-прежнему считать равномерно распределенной в G : $p_Q(P) \equiv 1/V_G$.

Для определения $p_{Q'}(P' | P)$ перенесем начало координат в точку P и выберем сферические координаты r', θ', φ' с центром в P (рис. 38). Направление (из точки P) условимся задавать единичным вектором ω , а расстояние по этому направлению от точки P до границы шара G назовем $l(\omega)$. Пусть

$$p_{Q'}(P' | P) = (4\pi)^{-1} (3 - \alpha) (r')^{-\alpha} l^{-(3-\alpha)}, \quad (34)$$

где $r' = |P' - P|$, а $\omega = (P' - P)/r'$.

Нетрудно проверить, что формула (34) действительно определяет условную плотность вероятностей: при любой фиксированной точке P

$$\int_G p_{Q'}(P' | P) dP' = \oint (4\pi)^{-1} d\omega \int_0^{l(\omega)} (3 - \alpha) (r')^{2-\alpha} l^{\alpha-3} dr' = 1.$$

Рассмотрим скалярную случайную величину

$$Z(Q, Q') = 4\pi V_G (3 - \alpha)^{-1} h(\rho) l^{3-\alpha}(\omega),$$

где $\rho = |Q' - Q|$, $\omega = (Q' - Q)/\rho$. Нетрудно вычислить, что

$$MZ = \iint_{GG} Z p_Q(P) p_{Q'}(P' | P) dP dP' = \iint_{GG} h(r') (r')^{-\alpha} dP dP' = I,$$

а дисперсия этой величины равна

$$DZ = (4\pi V_G)^2 (3 - \alpha)^{-2} \iint_{GG} h^{2l^{6-2\alpha}} dP dP' - I^2.$$

Так как функция $h(\rho)$ ограничена, а l не превосходит диаметра шара G , то последний интеграл сходится, и дисперсия DZ конечна при всех $\alpha < 3$.

В. Из-за симметрии задачи интеграл (33) может быть сведен к трехкратному. В расчетной схеме методов Монте-Карло этот факт учитывается «автоматически», если при реализации случайных точек принимать во внимание симметрию. Чтобы показать это выведем расчетные формулы для обоих способов расчета I .

Сферические координаты точки Q в обоих случаях можно вычислять по формулам (14) гл. 2. Однако из соображений симметрии ясно, что точку Q можно выбрать на оси Oz .

Поэтому положим $x_Q = y_Q = 0$,

$$z_Q = r_Q = R \sqrt[3]{\gamma}.$$

В способе А точка Q' также равномерно распределена в G . Из соображений симметрии ясно, что можно ее выбирать в плоскости $q' = 0$ и считать, что

$$x_{Q'} = r_{Q'} \sin \theta_{Q'}, \quad y_{Q'} = 0, \quad z_{Q'} = r_{Q'} \cos \theta_{Q'},$$

$$\text{где } \cos \theta_{Q'} = 2\gamma' - 1, \quad r_{Q'} = R \sqrt[3]{\gamma'}.$$

Получаем следующий алгоритм для расчета I методом А:

1) Формулы для i -го испытания

$$r_Q = R \sqrt[3]{\gamma}, \quad \cos \theta_{Q'} = 2\gamma' - 1, \quad r_{Q'} = R \sqrt[3]{\gamma'},$$

$$\rho = \sqrt{r_{Q'}^2 - 2r_Q r_{Q'} \cos \theta_{Q'} + r_Q^2},$$

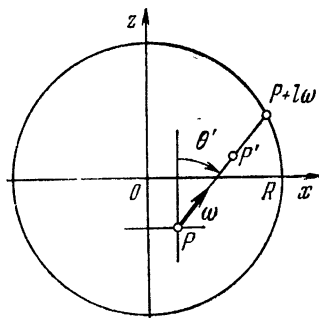
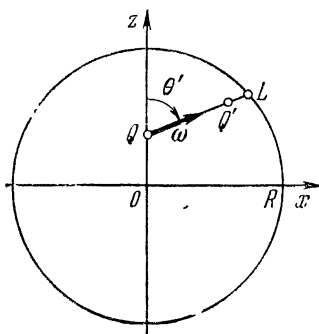


Рис. 38.

2) Если $\rho = \rho_i$ — значение ρ , полученное в i -м испытании, то

$$I_N^A = V_G^2 N^{-1} \sum_{i=1}^N h(\rho_i) \rho_i^{-\alpha}. \quad (35)$$

В способе Б случайную точку Q' с плотностью (34) можно строить следующим образом: сперва из точки Q надо выбрать случайное направление ω (п. 2.4.2 гл. 2); затем на луче $P' = Q + r'\omega$ выбрать случайное расстояние ρ с функцией распределения $F_\rho(r') = (r'/l)^{3-\alpha}$,



$0 < r' \leq l$ (где $l = l(\omega)$); тогда $Q' = Q + \rho\omega$.

В самом деле, условная плотность точки Q' (при условии $Q = P$) в сферических координатах r', θ', φ' с центром в P равна

$$\rho_{Q'}(P'/P) r'^2 \sin \theta' = \\ = [(4\pi)^{-1} \sin \theta'] [(3-\alpha) (r')^{2-\alpha} l'^{\alpha-2}].$$

Первый множитель равен плотности случайного направления ω (в сферических координатах), а второй — это условная плотность случайного расстояния на луче $P' = P + r'\omega$ при условии, что ω уже выбрано.

Те же соображения симметричны,

что в случае А, позволяют выбирать направление ω в плоскости $\varphi' = 0$. Оно определяется одним параметром $\mu = \cos \theta' = 2\gamma' - 1$ (рис. 39). Вычислив*) расстояние $l = |L - Q|$, получим следующий алгоритм для расчета I методом Б:

1) Формулы для i -го испытания

$$r_Q = R \sqrt[3]{\gamma}, \quad \mu = 2\gamma' - 1, \\ l = -\mu r_Q + \sqrt{R^2 - r_Q^2 (1 - \mu^2)}, \quad \rho = l (\gamma'')^{1/(3-\alpha)};$$

2) Если $l = l_i$ и $\rho = \rho_i$ — значения l и ρ , полученные в i -м испытании, то

$$I_N^B = 4\pi V_G (3 - \alpha)^{-1} N^{-1} \sum_{i=1}^N h(\rho_i) l_i^{3-\alpha}. \quad (36)$$

В обоих алгоритмах — (35) и (36) — на каждое испытание затрачивается всего три случайных числа, так что фактически вычисля-

*) В векторных обозначениях $L = Q + l\omega$. Так как $(L, L) = R^2$, то получаем квадратное уравнение $l^2(\omega, \omega) + 2l(Q, \omega) + (Q, Q) = R^2$, в котором $(\omega, \omega) = 1$. Решение: $l = -(Q, \omega) + \sqrt{(Q, \omega)^2 + R^2 - (Q, Q)}$. Чтобы перейти к декартовым координатам, заметим, что $Q = \{0, 0, r_Q\}$, а $\omega = \{\sin \theta', 0, \cos \theta'\}$. Следовательно, $(Q, Q) = r_Q^2$, $(Q, \omega) = r_Q \cos \theta'$.

сся трехкратный интеграл. В общем случае, когда функция $h(P, P')$ зависит не только от расстояния $\rho = |P - P'|$, а от самих точек P и P' , пришлось бы моделировать все шесть координат точек Q и Q' .

Г. Рассмотрим численный пример: интеграл (33) в случае $h(\rho) = \pi^{-2}$, $R=1$:

$$I_{\alpha} = (1/\pi^2) \iint_{GG} \rho^{-\alpha} dP dP',$$

где G — единичный шар $x^2 + y^2 + z^2 < 1$. Некоторые значения этого интеграла: $I_2 = 4$, $I_1 = 32/15$, $I_{-1} = 64/35$.

Запишем оценки (35) и (36) для I_1 и I_2 :

$$I_{1,N}^A = (16/9) N^{-1} \sum_{i=1}^N \rho_i^{-1}; \quad I_{2,N}^A = (16/9) N^{-1} \sum_{i=1}^N \rho_i^{-2};$$

$$I_{1,N}^B = (8/3) N^{-1} \sum_{i=1}^N l_i^2; \quad I_{2,N}^B = (16/3) N^{-1} \sum_{i=1}^N l_i.$$

Нетрудно заметить, что в этом случае оценка I^B зависит лишь от двух случайных чисел.

Вычислим дисперсии соответствующих этим методам величин $Z_{(\alpha)}$ (формулы приведены в пунктах А и Б):

$$DZ_{(\alpha)}^A = (4/3)^2 \pi^{-2} \iint_{GG} \rho^{-2\alpha} dP dP' - I_{\alpha}^2 = (4/3)^2 I_{2\alpha} - I_{\alpha}^2,$$

откуда следует, что $DZ_{(1)}^A = 2,560$, $DZ_{(2)}^A = \infty$. Для оценок типа Б

$$DZ_{(\alpha)}^B = (16/3)^2 (3 - \alpha)^{-2} \pi^{-2} \iint_{GG} l^{6-2\alpha} \rho(P, P') dP dP' - I_{\alpha}^2.$$

Пусть $\alpha' = 2\alpha - 3$, так что $6 - 2\alpha = 3 - \alpha'$. Тогда

$$DZ_{(\alpha)}^B = (32/3) (3 - \alpha)^{-1} I_{\alpha'} - I_{\alpha}^2,$$

откуда следует, что $DZ_{(1)}^B = 5,201$, $DZ_{(2)}^B = 6,756$. Интересно отме-

тить, что $DZ_{(1)}^B > DZ_{(1)}^A$. Этот пример показывает, что неудачное использование существенной выборки приводит к увеличению дисперсии.

Т а б л и ц а 1

Точные значения	$I_1=2,13$		$I_2=4,00$	
Результаты расчета	$I_1^A=1,78$	$I_1^B=1,48$	$I_2^A=1,94$	$I_2^B=3,37$
Ошибки	-0,35	-0,65	-2,06	-0,63
Вероятные ошибки r_{10}	0,34	0,49	∞	0,55

В табл. I приведены результаты расчета всех четырех оценок при $N=10$ с использованием случайных чисел, указанных на стр. 108. Здесь же указаны ошибки расчета и вероятные ошибки r_{10} .

3.3. Симметризация подынтегральной функции.

3.3.1. Простая симметризация. Пусть требуется вычислить интеграл

$$I_0 = \int_a^b f(x) dx$$

по конечному интервалу $a < x < b$. Рассмотрим случайную величину ξ , равномерно распределенную в этом интервале, и величину $Z = (b-a)f(\xi)$. Так как $MZ = I_0$, то простейший метод Монте-Карло приводит к оценке интеграла

$$\theta_N = \frac{b-a}{N} \sum_{i=1}^N f(\xi_i),$$

где ξ_1, \dots, ξ_N — независимые значения ξ .

Рассмотрим теперь *симметризованную функцию*

$$f^{(1)}(x) = \frac{1}{2} [f(x) + f(a+b-x)],$$

интеграл которой по-прежнему равен I_0 , и пусть $Z^{(1)} = (b-a)f^{(1)}(\xi)$. Ввиду того, что $MZ^{(1)} = I_0$, можно записать симметризованную оценку интеграла

$$\theta_N^{(1)} = \frac{b-a}{2N} \sum_{i=1}^N [f(\xi_i) + f(a+b-\xi_i)].$$

Так как математическое ожидание квадрата $Z^{(1)}$ равно

$$M[Z^{(1)}]^2 = \frac{b-a}{4} \int_a^b [f^2(x) + 2f(x)f(a+b-x) + f^2(a+b-x)] dx = \frac{b-a}{2} \left[\int_a^b f^2(x) dx + \int_a^b f(x)f(a+b-x) dx \right]$$

а математическое ожидание квадрата Z равно

$$MZ^2 = (b-a) \int_a^b f^2(x) dx,$$

то легко доказать, что всегда $M[Z^{(1)}]^2 \leq MZ^2$ и, следовательно, $DZ^{(1)} \leq DZ$. Однако для расчета одного значения $Z^{(1)}$ надо вычислить два значения $f(x)$. Поэтому трудоемкость оценки $\theta_N^{(1)}$ будет меньше трудоемкости θ_N только тогда, когда $DZ^{(1)}$ по крайней мере вдвое меньше, чем DZ . Оказывается, для монотонных функций это всегда выполнено.

Теорема 4. Если кусочно непрерывная функция $f(x)$ монотонна при $a \leq x \leq b$, то

$$DZ^{(1)} \leq (1/2) DZ.$$

Доказательство*). Из выражений для дисперсий

$$2DZ^{(1)} = (b-a) \int_a^b f^2(x) dx + (b-a) \int_a^b f(x) = f(a+b-x) dx - 2I_0^2$$

и

$$DZ = (b-a) \int_a^b f^2(x) dx - I_0^2$$

вытекает, что утверждение теоремы равносильно неравенству

$$(b-a) \int_a^b f(x) f(a+b-x) dx \leq I_0^2. \quad (37)$$

Предположим для определенности, что $f(x)$ не убывает и $f(b) > f(a)$. Введем вспомогательную функцию

$$v(x) = (b-a) \int_a^x f(a+b-t) dt - (x-a) I_0,$$

которая обращается в нуль на концах отрезка $a \leq x \leq b$. Производная этой функции

$$v'(x) = (b-a) f(a+b-x) - I_0$$

монотонна, $v'(a) > 0$, $v'(b) < 0$; следовательно, $v(x) \geq 0$

*) Мы рассматриваем случай, когда $f(x)$ непрерывна и дифференцируема. Однако нетрудно видоизменить доказательство (вводя интегралы Стильеса) так, что ни дифференцируемость, ни непрерывность $f(x)$ не понадобятся.

при $a \leq x \leq b$. И можно записать очевидное неравенство

$$\int_a^b v(x) f'(x) dx \geq 0.$$

Проинтегрировав по частям, получим неравенство

$$\int_a^b f(x) v'(x) dx \leq 0.$$

Подставив сюда выражение для $v'(x)$, получим (37).

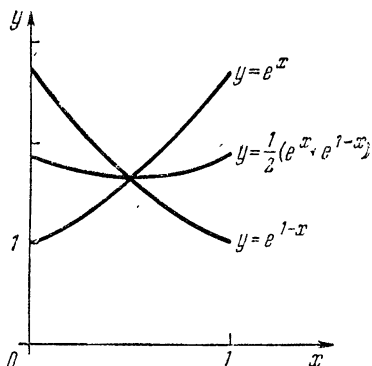
Случай невозрастания $f(x)$ рассматривается точно так же, так как тогда $v(x) \leq 0$, $f'(x) \leq 0$.

Пример. Рассмотрим интеграл

$$I = \int_0^1 e^x dx.$$

Симметризованная функция в этом примере равна $f^{(1)}(x) = (1/2)(e^x + e^{1-x})$ (рис. 40), а соответствующая оценка

$$\theta_N^{(1)} = (2N)^{-1} \sum_{i=1}^N (e^{\gamma_i} + e^{1-\gamma_i}).$$



Дисперсия осредняемой величины равна

$$DZ^{(1)} = (1/4) [2e - (e - 1)(3e - 5)] = 0,00392.$$

Это во много раз меньше, чем $DZ = 0,2420$.

3.3.2. О сложной симметризации. Интервал $a < x < b$ можно разбить на конечное число частей и на каждой из них использовать простую симметризацию.

Рассмотрим случай разбиения (a, b) на две равные части. Пусть $c = (1/2)(a+b)$. Тогда

$$\begin{aligned} I_0 &= \int_a^c f(y) dy + \int_c^b f(y) dy = \\ &= (1/2) \int_a^c [f(y) + f(a+c-y)] dy + (1/2) \int_c^b [f(y) + f(c+b-y)] dy. \end{aligned}$$

В первом из этих интегралов сделаем замену переменной $x = 2y - a$, которая преобразует интервал (a, c) в (a, b) , а во втором — замену

$x = 2y - b$, которая преобразует интервал (c, b) в (a, b) . Получим выражение

$$I_0 = \int_a^b f^{(2)}(x) dx,$$

где

$$f^{(2)}(x) = (1/4) \left[f\left(\frac{a+x}{2}\right) + f\left(\frac{2a+b-x}{2}\right) + f\left(\frac{b+x}{2}\right) + f\left(\frac{2'+a-x}{2}\right) \right].$$

Следовательно, для оценки интеграла I_0 можно использовать величину

$$0_N^{(2)} = (1/N) \sum_{i=1}^N Z_i^{(2)},$$

где $Z_i^{(2)}$ — значения случайной величины $Z^{(2)} = (b-a) f^{(2)}(\xi)$.

Пример. Рассмотрим интеграл

$$I = \int_0^1 (2 - 3x - x^2) dx = 1/6.$$

Так как $f(x) = 2 - 3x - x^2$, то нетрудно вычислить, что

$$f^{(1)}(x) = x - x^2, \quad f^{(2)}(x) = (1/8)(1 + 2x - 2x^2).$$

Сравним дисперсии величин

$$Z = f(\gamma), \quad Z^{(1)} = f^{(1)}(\gamma) \text{ и } Z^{(2)} = f^{(2)}(\gamma):$$

$$DZ = \int_0^1 (2 - 3x - x^2)^2 dx - 1/36 = 41/30 - 1/36 = 241/180;$$

$$DZ^{(1)} = \int_0^1 (x - x^2)^2 dx - 1/36 = 1/30 - 1/36 = 1/180;$$

$$DZ^{(2)} = (1/64) \int_0^1 (1 + 2x - 2x^2)^2 dx - 1/36 = 9/320 - 1/36 = 1/2880.$$

Если обозначить время расчета $f(x)$ через t , то времена расчета $f^{(1)}(x)$ и $f^{(2)}(x)$ приблизительно равны $2t$ и $4t$. Следовательно, трудоемкости трех способов расчета равны соответственно

$$241t/180 > 2t/180 > 0,25t/180.$$

3.3.3. К сожалению, различные методы симметризации, весьма наглядные и эффективные в одномерном случае, становятся громоздкими и трудно оцениваемыми при переходе к функциям многих переменных. Даже простая симметризация функции $f(x, y, z)$ по всем пере-

менным в единичном кубе содержит уже 8 слагаемых:

$$f^{(1)} = (1/8) [f(x, y, z) + f(1-x, y, z) + f(x, 1-y, z) + \\ + f(x, y, 1-z) + f(1-x, 1-y, z) + f(1-x, y, 1-z) + \\ + f(x, 1-y, 1-z) + f(1-x, 1-y, 1-z)].$$

Некоторые способы одномерной симметризации рассмотрены в работах [133, 136, 149]. См. также упражнение 6.

При расчете методом Монте-Карло ряда физических задач [29] используют приемы, которые по существу представляют собой частичную симметризацию (по отдельным переменным). Например, вместо того, чтобы моделировать случайное направление каждой частицы,

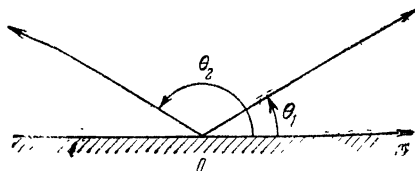


Рис. 41.

вылетающей из заданной точки O (рис. 41), по формуле $\cos \theta = 2\gamma - 1$ рассматривают пару частиц, которые вылетают по направлениям $\theta_1 = \arccos(2\gamma - 1)$ и $\theta_2 = \pi - \theta_1$. Выбор направления θ_2 равенсильен использованию $1 - \gamma$ вместо значения γ . Такой способ моделирования обеспечивает более равномерное расположение траекторий частиц в пространстве.

3.4. Двухэтапные схемы расчета. В некоторых задачах можно указать не одну «хорошую» оценку, а целое семейство, зависящее от параметров. Возникает вопрос о наилучшем выборе параметров. Обычно условием выбора служит требование минимума дисперсии оценки (при этом молчаливо предполагается, что время расчета одного испытания слабо зависит от значений параметров).

Аналитическое решение в этой ситуации, как правило, невозможно. Однако можно рекомендовать численный подход: на первом этапе весьма грубо вычисляются дисперсии оценки при различных значениях параметров (это нетрудно сделать методом п. 1.4 по небольшому количеству N испытаний); на втором этапе решается основная задача, при помощи оценки с наилучшей системой параметров. Известные по первому этапу время счета и дисперсия позволяют довольно точно оценить объем работы, необходимый для достижения заданной вероятной ошибки.

Мы рассмотрим два примера однопараметрических оценок.

3.4.1. В п. 3.1.1 для приближенного вычисления интеграла

$$I = \int_G f(P) p(P) dP$$

использовалась оценка

$$\theta_N = C + (1/N) \sum_{i=1}^N [f(Q_i) - h(Q_i)],$$

где Q_i — случайные точки с плотностью $p(P)$, а $h(P)$ — некоторая функция, «близкая» к $f(P)$, интеграл которой известен: $C = \int_G h(P) p(P) dP$. Обе функции $f(P)$ и $h(P)$ принадлежат $L_2(G; p)$.

Можно выбрать произвольный параметр α и рассмотреть более общую оценку интеграла I

$$\theta_N^{(\alpha)} = \alpha C + (1/N) \sum_{i=1}^N [f(Q_i) - \alpha h(Q_i)]. \quad (38)$$

Действительно, в (38) осредняется величина $Z^{(\alpha)} = \alpha C + f(Q) - \alpha h(Q)$, математическое ожидание которой $MZ^{(\alpha)} = I$. По известной формуле о дисперсии суммы

$$DZ^{(\alpha)} = Df(Q) - 2\alpha r \sqrt{Df(Q) Dh(Q)} + \alpha^2 Dh(Q), \quad (39)$$

где r — коэффициент корреляции величин $f(Q)$ и $h(Q)$:

$$r = M \{ [f(Q) - I] [h(Q) - C] \} [Df(Q) Dh(Q)]^{-1/2}.$$

Простые вычисления показывают, что минимум $DZ^{(\alpha)}$ реализуется при $\alpha = \alpha_0 = r \sqrt{Df(Q)/Dh(Q)}$ и равен

$$\min_{\alpha} DZ^{(\alpha)} = DZ^{(\alpha_0)} = (1 - r^2) Df(Q).$$

(Теоретически возможен даже случай $DZ^{(\alpha_0)} = 0$, который реализуется при $|r| = 1$; однако на практике, если $|r| = 1$, то случайные величины $f(Q)$ и $h(Q)$ связаны линейной зависимостью $f(P) = Ah(P) + B$, и искомый интеграл равен $I = AC + B$; никакие приближенные оценки не нужны).

Пример. Требуется вычислить интеграл

$$I = \int_0^1 e^x dx = e - 1.$$

Так же, как в п. 3.1.1, положим $h(x) = x$, но вместо оценки (23) воспользуемся оценкой (38)

$$\theta_N^{(\alpha)} = (\alpha/2) + (1/N) \sum_{i=1}^N (e^{y_i} - \alpha y_i).$$

Так как

$$Df = \int_0^1 e^{2x} dx - I^2 = 0,2420, \quad Dh = \int_0^1 x^2 dx - (1/4) = 1/12,$$

$$r \sqrt{Df/Dh} = \int_0^1 (e^x - 1)(x - 1/2) dx = (1/2)(3 - e),$$

то наилучшее значение $\alpha_0 = 6(3 - e) = 1,6903$, и соответствующее ему значение дисперсии равно

$$DZ(\alpha_0) = (1/2)(e^2 - 1) - I^2 - 3(3 - e)^2 = 0,00393.$$

Это примерно в 10 раз меньше, чем DZ' в п. 3.1.1. Ничего удивительного в таком результате нет: из рис. 42 видно, что прямая $y = I + \alpha_0(x - 1/2)$ гораздо лучше приближает функцию $y = e^x$, чем прямая $y = I + x - 1/2$.

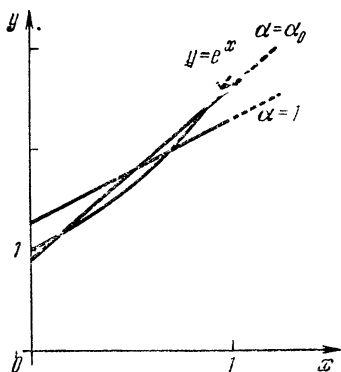


Рис. 42.

3.4.2. Чаше других встречается метод существенной выборки, зависящий от параметра. Если при всех α (из некоторого множества) плотности $p(P; \alpha)$ допустимы по отношению к $f(P)$, то интеграл $I_0 = \int_G (P) dP$ можно вычислять

методом существенной выборки (п. 3.2) с любым α . Наилучшее значение $\alpha = \alpha_0$ — это значение, при котором достигает минимума интеграл

$$\int_{G^+} [f^2(P)/p(P; \alpha)] dP,$$

фигурирующий в выражении (29) для дисперсии. Аналитическое исследование способов выбора α имеется в статье А. Маршалла [157]. Однако на практике обычно выбор α осуществляется экспериментальным способом, указанным в начале п. 3.4.

Пример. Интеграл

$$I = \int_0^1 e^x dx$$

в п. 3.2.1 вычислялся методом существенной выборки с плотностью $p(x) = (2/3)(1+x)$. Рассмотрим теперь семейство допустимых плотностей

$$p(x; \alpha) = (1 + \alpha x)/(1 + \alpha/2), \quad \alpha > 0.$$

Для разыгрывания значений случайной величины ξ с плотностью $p(x; \alpha)$ методом обратных функций получаем уравнение

$$(\xi + \alpha \xi^2/2)/(1 + \alpha/2) = \gamma,$$

решение которого трудно записать в явном виде. Окончательная оценка интеграла I :

$$\theta_N^{(\alpha)} = (1 + \alpha/2) N^{-1} \sum_{i=1}^N e^{\xi_i} (1 + \alpha \xi_i)^{-1},$$

где $\xi_i = (1/\alpha) [\sqrt{1 + \alpha(\alpha + 2)\gamma_i} - 1]$.

В этом примере дисперсия осредняемой величины $Z^{(\alpha)} = (1 + \alpha/2) e^{\xi} (1 + \alpha \xi)^{-1}$ равна

$$DZ^{(\alpha)} = \int_0^1 (1 + \alpha/2) (1 + \alpha x)^{-1} e^{2x} dx - I^2.$$

После замены переменной $\alpha y = 2(1 + \alpha x)$ это выражение превратится в

$$DZ^{(\alpha)} = (2^{-1} + \alpha^{-1}) e^{-2/\alpha} \alpha^{\frac{2+\alpha}{2}} \int_{2/\alpha}^{\frac{2+\alpha}{2}} e^y y^{-1} dy - I^2,$$

где интеграл выражается через интегральную показательную функцию $\bar{E}i(x)$:

$$DZ^{(\alpha)} = (2^{-1} + \alpha^{-1}) e^{-2/\alpha} [\bar{E}i(2 + 2/\alpha) - \bar{E}i(2/\alpha)] - I^2.$$

При $\alpha=1$ получаем для $DZ^{(1)}$ выражение из п. 3.2.1, которое равно 0,0269. Однако расчеты показывают, что дисперсия $DZ^{(\alpha)}$ будет минимальной при $\alpha = \alpha_0 \approx 1,81$, когда $DZ^{(\alpha_0)} = 0,0040$.

§ 4. Интегралы, зависящие от параметра

4.1. Использование зависимых испытаний. Предположим, что требуется приближенно вычислить значение интеграла

$$I(\lambda) = \int_G f(P, \lambda) p(P) dP \quad (40)$$

при нескольких значениях действительного параметра λ , например $\lambda = \lambda_1, \dots, \lambda_s$. Если при каждом из этих λ условие применимости простейшего метода Монте-Карло выполнено

$$DZ(\lambda) = \int_G f^2(P, \lambda) p(P) dP - I^2(\lambda) < \infty,$$

то можно записать оценку для $I(\lambda)$:

$$\theta_N(\lambda) = (1/N) \sum_{i=1}^N f(Q_i, \lambda), \quad (41)$$

где Q_1, \dots, Q_N — случайные точки с плотностью $p(P)$.

Оценка (41) справедлива при каждом из интересующих нас значений $\lambda_1, \dots, \lambda_s$. Если функция $f(P, \lambda)$ непрерывно зависит от λ , то и $\theta_N(\lambda)$ будет непрерывной функцией от λ . Вычислив $\theta_N(\lambda_1), \dots, \theta_N(\lambda_s)$ мы, конечно, получим значения, отличные от $I(\lambda_1), \dots, I(\lambda_s)$, но эти значения будут расположены на гладкой кривой. Если N достаточно велико, то кривая $\theta_N(\lambda)$ окажется достаточно близкой к $I(\lambda)$ и по значениям $\theta_N(\lambda_k)$ можно будет даже оценить производную $I'(\lambda)$ (если она существует).

Несколько неожиданным может показаться утверждение, что результаты п. 1.2 не дают достаточного основания для применения оценки (41). В самом деле, из (4) следует, что вероятность неравенства

$$|\theta_N(\lambda) - I(\lambda)| < x_\beta \sqrt{DZ(\lambda)/N} \quad (42)$$

приблизительно равна β (если N достаточно велико). Это справедливо при каждом $\lambda = \lambda_k$. Но нельзя утверждать, что вероятность одновременного выполнения нескольких неравенств (42) при $\lambda = \lambda_1, \dots, \lambda_s$ тоже равна β .

Формула (4) позволяет оценить все ошибки $|\theta_N(\lambda_k) - I(\lambda_k)|$, если для каждого $\lambda = \lambda_k$ вычислять $\theta_N(\lambda_k)$ по независимым испытаниям, т. е. считать, что

$$I(\lambda_k) \approx (1/N) \sum_{i=1}^N f(Q_{i,k}, \lambda_k),$$

где $Q_{i,k}$ при $i=1, 2, \dots, N$; $k=1, 2, \dots, s$ — независимые случайные точки с плотностью $p(P)$. Однако в этом случае приближенные значения интеграла даже при очень близких λ могут различаться между собой на величину порядка $\sqrt{DZ/N}$ и результаты расчета не дают верного представления о разностях $I(\lambda_{k+1}) - I(\lambda_k)$. (Кроме того, при таком расчете приходится вычислять в s раз больше случайных точек.)

Пример. Вычислить интеграл

$$I(\lambda) = \int_0^2 \cos \lambda x \, dx = \lambda^{-1} \sin 2\lambda$$

для $\lambda = 0,2, 0,4; \dots; 2,6$.

Т а б л и ц а 2

λ	$I(\lambda)$	θ^0	θ^+	r_{10}	λ	$I(\lambda)$	θ^0	θ^+	r_{10}
0,2	1,947	1,929	1,929	0,01	1,6	-0,036	-0,628	0,196	0,30
0,4	1,793	1,722	1,732	0,04	1,8	-0,246	-0,851	-1,040	0,31
0,6	1,553	1,401	1,474	0,08	2,0	-0,378	-0,955	-0,299	0,31
0,8	1,249	0,997	0,915	0,14	2,2	-0,433	-0,939	-0,175	0,30
1,0	0,909	0,551	0,377	0,19	2,4	-0,415	-0,814	-0,056	0,29
1,2	0,563	0,105	0,627	0,24	2,6	-0,340	-0,600	-0,163	0,28
1,4	0,239	-0,300	0,348	0,28					

По формуле

$$\theta_{10}(\lambda) = 0,2 \sum_{i=1}^{10} \cos 2\lambda \gamma_i$$

проведены два расчета, результаты которых даны в табл. 2 и на рис. 43. В первом случае для расчета всех $\theta_{10}(\lambda)$ использовались

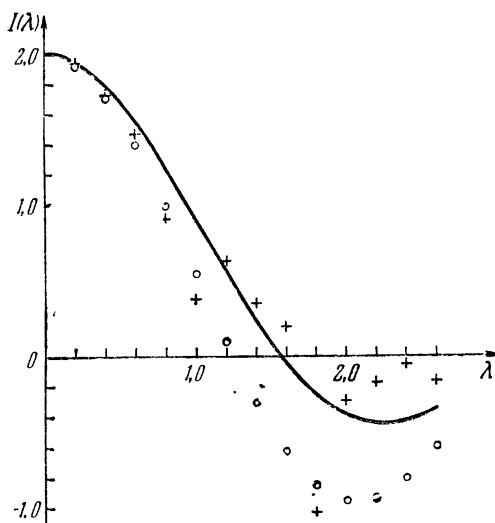


Рис. 43.

одни и те же $\gamma_1, \dots, \gamma_{10}$ приведенные на стр. 108; соответствующие результаты обозначены кружками: θ^0 . Во втором случае для расчета каждого $\theta_{10}(\lambda)$ использовались новые случайные числа (образованные из пятерок цифр таблицы на стр. 295, которые выбирались подряд); соответствующие результаты обозначены крестиками: θ^+ .

На рис. 43 отчетливо видно, что значения θ^0 расположены на гладкой кривой, сходной с $I(\lambda)$ (непрерывная линия), а значения θ^+ разбросаны около $I(\lambda)$. В табл. 2 приведены также вероятные ошибки

$$r_{10}(\lambda) = 0,675 \sqrt{D\theta_{10}} = \\ = 0,675 \cdot 10^{-1/2} [2 - \lambda^{-1} \sin 2\lambda (\lambda^{-1} \sin 2\lambda - \cos 2\lambda)]^{1/2}.$$

Значение $N=10$ невелико, и нельзя быть уверенным в том, что распределение θ_{10} достаточно близко к нормальному. Поэтому интересно отметить, что все ошибки $\theta_{10}(\lambda) - I(\lambda)$ по порядку равны $r_{10}(\lambda)$.

На протяжении многих лет метод (41) неоднократно и с успехом использовался в самых различных расчетах, но строгое обоснование его было впервые опубликовано лишь в 1962 году А. С. Фроловым и Н. Н. Ченцовым [92]. Этому же вопросу посвящены работы [67, 78, 84]. В дальнейшем изложении мы следуем статье [78].

4.2. Вспомогательная теорема о погрешности простейшей квадратурной формулы. Допустим, что для приближенного вычисления интеграла

$$I = \int_0^1 f(x) dx$$

используются N фиксированных точек x_1, \dots, x_N , принадлежащих интервалу $(0, 1)$, и простейшая формула

$$I \approx (1/N) \sum_{i=1}^N f(x_i).$$

Погрешность такого приближения зависит от функции $f(x)$:

$$\delta(f) = (1/N) \sum_{i=1}^N f(x_i) - \int_0^1 f(x) dx. \quad (43)$$

Обозначим через $W'_2(L)$ множество непрерывных функций $f(x)$ с кусочно непрерывными производными $f'(x)$, такими, что

$$\int_0^1 [f'(x)]^2 dx \leq L^2. \quad (44)$$

По сравнению с § 2 здесь мы сужаем класс подынтегральных функций: там требовалось только, чтобы $f(x) \in L_2$, а здесь — $f'(x) \in L_2$.

Обозначим через $S_N(x)$ количество точек x_i с номерами $1 \leq i \leq N$, удовлетворяющих неравенству $x_i \leq x$. Нетрудно проверить, что

$$S_N(x) = \sum_{i=1}^N e(x - x_i),$$

и представляет собой ступенчатую функцию (рис. 44).

Теорема 5. *Каковы бы ни были точки x_1, \dots, x_N , верхняя грань погрешности (43) равна*

$$\sup_{f \in W_2(L)} |\delta(f)| = \frac{L}{N} \left\{ \int_0^1 |S_N(x) - Nx|^2 dx \right\}^{1/2}. \quad (45)$$

Доказательство. Функция $f(x)$ может быть выражена через свою производную

$$\begin{aligned} f(x) &= f(1) - \int_x^1 f'(t) dt = \\ &= f(1) - \int_0^1 f'(t) e(t - x) dt. \end{aligned}$$

Полагая здесь $x = x_i$ и суммируя по i , получим, что

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N f(x_i) &= \\ &= f(1) - \frac{1}{N} \int_0^1 f'(t) S_N(t) dt. \end{aligned}$$

С другой стороны, интегрируя по частям, легко получить, что

$$\int_0^1 f(t) dt = f(1) - \int_0^1 f'(t) t dt.$$

Вычитая последнее равенство из предпоследнего, получим формулу для погрешности

$$\delta(f) = \frac{1}{N} \int_0^1 f'(t) [Nt - S_N(t)] dt. \quad (46)$$

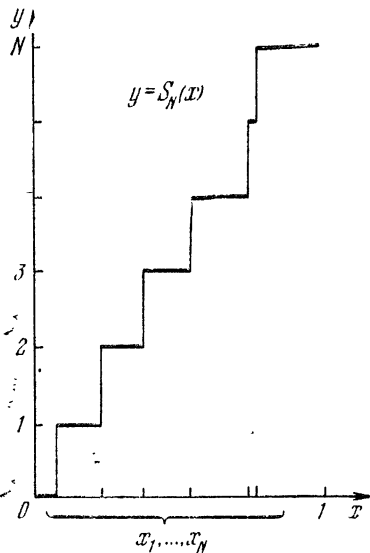


Рис. 44.

Из формулы (46), используя неравенство (1) (стр. 292) и условие (44), выводим, что если $f(x) \in W'_2(L)$, то

$$|\delta(f)| \leq \frac{1}{N} \int_0^1 |f'| |S_N - Nt| dt \leq \frac{L}{N} \left\{ \int_0^1 |S_N - Nt|^2 dt \right\}^{1/2};$$

эта величина как раз фигурирует в (45).

Осталось доказать неувлучшаемость последнего неравенства. Для этого рассмотрим функцию

$$g(x) = L \int_0^x [Nt - S_N(t)] dt \left\{ \int_0^1 |Nt - S_N(t)|^2 dt \right\}^{-1/2}.$$

Так как $\int_0^1 [g'(x)]^2 dx = L^2$, то $g(x) \in W'_2(L)$. Подставив

$f = g(x)$ в (46), получим, что

$$\delta(g) = \frac{L}{N} \left\{ \int_0^1 |Nt - S_N(t)|^2 dt \right\}^{1/2}.$$

Таким образом, теорема доказана.

4.3. Оценка погрешности метода Монте-Карло с помощью распределения ω^2 . Выберем N независимых случайных чисел $\gamma_1, \dots, \gamma_N$ и рассмотрим погрешность простейшего метода Монте-Карло

$$\delta(f) = \frac{1}{N} \sum_{i=1}^N f(\gamma_i) - \int_0^1 f(x) dx. \quad (47)$$

В этом случае величины, входящие в (45), имеют весьма простой вероятностный смысл: $S_N(x)/N$ — это эмпирическая функция распределения $F_N^*(x)$ выборки $\gamma_1, \dots, \gamma_N$ (ср. (17) гл. 1), а x — это функция распределения $F(x)$ случайной величины γ (при $0 < x < 1$). Поэтому, согласно (18) гл. 1,

$$\frac{1}{N} \left\{ \int_0^1 [S_N(x) - Nx]^2 dx \right\}^{1/2} = \left\{ \int_0^1 [F_N^* - F]^2 dF(x) \right\}^{1/2} = \left(\frac{\omega_V^2}{N} \right)^{1/2}.$$

Из теоремы предыдущего пункта вытекает, что для простейшего метода Монте-Карло (47)

$$\sup_{f \in W_2(L)} |\delta(f)| = L \sqrt{\omega_N^2/N}. \quad (48)$$

Воспользуемся теоремой Мизеса — Смирнова, приведенной на стр. 34: *если N достаточно велико, то $P\{\omega_N^2 < \epsilon\} \approx a_1(\epsilon)$* . Следовательно, можно выбрать любую доверительную вероятность β и найти соответствующий ей корень $\epsilon = \epsilon_\beta$ уравнения $a_1(\epsilon_\beta) = \beta$ и утверждать, что с вероятностью, приблизительно равной β ,

$$\sup_{f \in W_2(L)} |\delta(f)| \leq L \sqrt{\epsilon_\beta/N}. \quad (49)$$

Из оценки (49) следует, что *при достаточно большом N с вероятностью, приблизительно равной β , неравенство*

$$|\delta(f)| \leq L \sqrt{\epsilon_\beta/N}$$

справедливо одновременно для всех функций $f(x) \in W_2'(L)$.

Обратимся к интересующему нас случаю, когда вычисляется интеграл вида (40)

$$I(\lambda) = \int_0^1 f(x, \lambda) dx \quad (50)$$

с помощью оценки вида (41)

$$\theta_N(\lambda) = \frac{1}{N} \sum_{i=1}^N f(\gamma_i, \lambda). \quad (51)$$

Из предыдущего результата вытекает, что если при всех $a \leq \lambda \leq b$ производная $f'_x(x, \lambda) \in L_2$, то с вероятностью, не меньшей чем β , при всех этих λ одновременно

$$|\theta_N(\lambda) - I(\lambda)| \leq L(\lambda) \sqrt{\epsilon_\beta/N}, \quad (52)$$

где, очевидно,

$$L(\lambda) = \left\{ \int_0^1 [f'_x(x, \lambda)]^2 dx \right\}^{1/2}.$$

4.4. Численное дифференцирование оценки (51). Рассмотрим случай, когда существует производная $I'(\lambda)$ от интеграла (50), и докажем, что по значениям $\theta_N(\lambda)$ можно обычным способом эту производную оценить:

$$I'(\lambda) \approx \frac{\theta_N(\lambda + h) - \theta_N(\lambda - h)}{2h}.$$

Допустим, что при $a \leq \lambda \leq b$ существуют вторые производные $f''_{\lambda x}(x, \lambda) \in L_2$ и $|f''_{\lambda \lambda}(x, \lambda)| \leq C$. Тогда

$$I'(\lambda) = \int_0^1 f'_\lambda(x, \lambda) dx,$$

а оценка (41) для этого интеграла равна производной от оценки (51):

$$\theta'_N(\lambda) = \frac{1}{N} \sum_{i=1}^N f'_\lambda(\gamma_i, \lambda).$$

На основании (49) можно утверждать (мы по-прежнему считаем, что N достаточно велико), что с вероятностью, не меньшей чем β , одновременно справедливы и неравенство (52), и неравенство

$$|\theta'_N(\lambda) - I'(\lambda)| < L_1(\lambda) \sqrt{x_\beta/N}, \quad (53)$$

где

$$L_1(\lambda) = \left\{ \int_0^1 [f''_{\lambda x}(x, \lambda)]^2 dx \right\}^{1/2}.$$

Далее, в силу сделанных предположений, при любых фиксированных значениях $\gamma_1, \dots, \gamma_N$ функция (51) дважды дифференцируема по λ . Из разложения

$$\theta_N(\lambda \pm h) = \theta_N(\lambda) \pm h\theta'_N(\lambda) \pm (1/2)h^2\theta''_N(\lambda \pm)$$

вытекает, что

$$\left| \frac{\theta_N(\lambda + h) - \theta_N(\lambda - h)}{2h} - \theta'_N(\lambda) \right| < \frac{1}{2} Ch. \quad (54)$$

Наконец, из (53) и (54) следует, что

$$\left| \frac{\theta_N(\lambda + h) - \theta_N(\lambda - h)}{2h} - I'(\lambda) \right| \leq L_1(\lambda) \sqrt{\frac{x_\beta}{N}} + \frac{1}{2} Ch,$$

и это неравенство, одновременно, с (52) и (53), имеет место с вероятностью, не меньшей чем β .

В этих условиях, по-видимому, целесообразно выбирать h порядка $1/\sqrt{N}$.

4.5. О таблицах случайных чисел. Оценка (49) в какой-то степени оправдывает многократное использование таблиц случайных чисел, ибо по одним и тем же числам («по таблице») можно решать много задач: вычислять интегралы от любых функций $f(x)$, принадлежащих $W_2(L)$. Оценку, аналогичную (49), можно получить для гораздо более широкого класса задач, включающих вычисление интегралов от функций многих переменных. Однако для всех функций $f(x)$ из L_2 оценка погрешности такого типа невозможна.

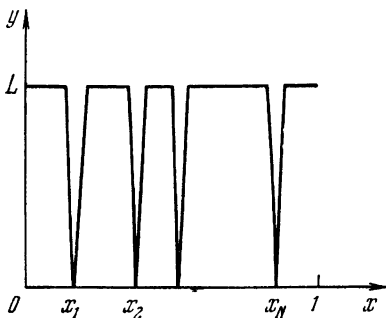


Рис. 45.

В самом деле, каковы бы ни были точки x_1, \dots, x_N , найдется функция $f(x)$ (рис. 45) такая, что

а) $f(x_1) = f(x_2) = \dots = f(x_N) = 0$;

б) $\int_0^1 f(x) dx = L - \varepsilon$, где $L > \varepsilon > 0$ — любые числа;

в) $\int_0^1 f^2(x) dx < L^2$.

Для такой функции, очевидно, $\delta(f) = L - \varepsilon$. Следовательно, при любых x_1, \dots, x_N

$$\sup_{f \in L_2} |\delta(f)| \geq L,$$

и величина эта не стремится к нулю, когда $N \rightarrow \infty$.

Упражнения к главе 3

1. Записать формулы для расчета методом Монте-Карло интеграла

$$I = \iiint_G f(x, y, z) dx dy dz \quad (55)$$

от произвольной ограниченной функции $f(x, y, z)$. Область интегрирования G определена неравенствами $x^2 + y^2 < z < 2$.

2. Требуется вычислить интеграл (55) от произвольной ограниченной функции $f(x, y, z)$ по области G , расположенной между

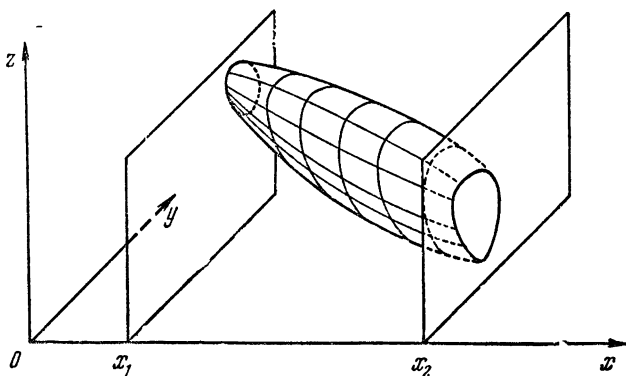


Рис. 46.

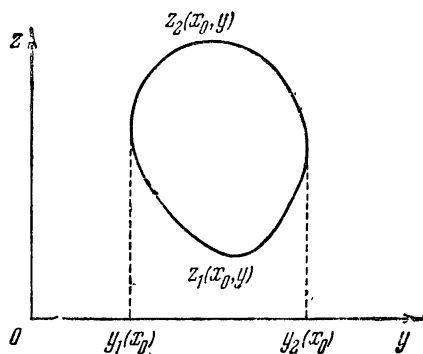


Рис. 47.

плоскостями $x=x_1$ и $x=x_2$ (рис. 46); сечение G плоскостью $x=x_0$ изображено на рис. 47.

Случайные точки $Q_i = (\xi_i, \eta_i, \zeta_i)$ в области G будем выбирать по формулам

$$\xi_i = x_1 + \gamma_i(x_2 - x_1),$$

$$\eta_i = y_1(\xi_i) + \gamma'_i[y_2(\xi_i) - y_1(\xi_i)],$$

$$\zeta_i = z_1(\xi_i, \eta_i) + \gamma''_i[z_2(\xi_i, \eta_i) - z_1(\xi_i, \eta_i)].$$

Можно ли в качестве оценки для I использовать среднее арифметическое

$$\frac{1}{N} \sum_{i=1}^N f(Q_i)?$$

Если нет, то написать верную оценку, содержащую значения $f(Q_1), \dots, f(Q_N)$.

3. Построить оценку с конечной дисперсией для вычисления интеграла

$$\int_0^{\infty} x^{-5/2} f(x) dx$$

в случае, когда $f(x) \sim x$ при $x \rightarrow \infty$ и $f(x) \sim x^2$ при $x \rightarrow 0$.

4. Записать формулы для расчета интеграла

$$I = \int_0^{\infty} f(x) e^{-kx} dx, \quad k > 0,$$

с помощью значений случайной величины ξ , плотность которой равна $p(x) = \alpha e^{-\alpha x}$. Доказать, что если $f(x) \approx Ax^n$, то дисперсия будет наименьшей при $\alpha \approx \alpha_0 = k/(n+1)$.

5. Условно сходящийся интеграл

$$I = \int_1^{\infty} x^{-1} \sin 2\pi x dx$$

можно вычислить методом Монте-Карло с помощью оценки

$$\theta_N = \frac{1}{N} \sum_{i=1}^N h(\gamma_i) \sin \pi \gamma_i,$$

где

$$h(x) = \sum_{k=1}^{\infty} [(2k+x)(2k+1+x)]^{-1}.$$

Доказать, что $M[h(\gamma) \sin \pi \gamma] = I$, $D[h(\gamma) \sin \pi \gamma] \leq (1/2)h^2(0) - I^2$.

6. Рассмотреть симметризации функции

$$f(x) = a_0 + \sum_{k=1}^{\infty} (a_k \cos 2\pi kx + b_k \sin 2\pi kx)$$

на интервале $0 < x < 1$ и выразить дисперсии $Df(\gamma)$, $Df^{(1)}(\gamma)$ и $Df^{(2)}(\gamma)$ через коэффициенты Фурье a_k и b_k .

7. Рассмотреть функцию $f^{(3)}(x) = (1/2)[f(x/2) + f(1-x/2)]$ и доказать, что хотя, вообще говоря, $f^{(3)}(x) \neq f^{(1)}(x)$ (где $f^{(1)}(x) = (1/2)[f(x) + f(1-x)]$), но по отношению к простейшему методу Монте-Карло эти функции равносильны:

$$Mf^{(3)}(\gamma) = Mf^{(1)}(\gamma) = Mf(\gamma), \quad Df^{(3)}(\gamma) = Df^{(1)}(\gamma).$$

8. Требуется вычислить интеграл (12) от функции $f(P) \in L_2(G; \rho)$. Пусть $\varphi_1(P), \dots, \varphi_s(P)$ — ортонормированные функции со средними значениями, равными нулю, т. е.

$$\int_G \varphi_j(P) \varphi_k(P) \rho(P) dP = \delta_{jk}, \quad \int_G \varphi_j(P) \rho(P) dP = 0.$$

Рассмотрим семейство оценок для I

$$\theta_N = \frac{1}{N} \sum_{i=1}^N \left[f(Q_i) - \sum_{k=1}^s \alpha_k \varphi_k(Q_i) \right],$$

где Q_i — независимые случайные точки с плотностью $\rho(P)$, а $\alpha_1, \dots, \alpha_s$ параметры.

Доказать, что дисперсия этой оценки минимальна тогда, когда

$$\alpha_k = \int_G f(P) \varphi_k(P) \rho(P) dP.$$

9. Случайная величина v_k подчиняется биномиальному распределению: $P\{v_k = j\} = C_{n_k}^j p_k^j (1 - p_k)^{n_k - j}$, $j = 0, 1, \dots, n_k$. Случайные величины v_1, \dots, v_s независимы. Требуется вычислить математическое ожидание ограниченной величины $\eta = f(v_1, \dots, v_s)$, которое равно

$$M\eta = \sum_{j_1, \dots, j_s=0}^{n_1, \dots, n_s} f(j_1, \dots, j_s) P\{v_1 = j_1\} \dots P\{v_s = j_s\}.$$

Если все $n_k \approx 10$ и $s \approx 20$, то количество слагаемых в этой сумме $\approx 10^{20}$. Ясно, что на современных ЭВМ такая сумма вычислена быть не может.

Построить простейший метод Монте-Карло для расчета $M\eta$ и какой-нибудь алгоритм, соответствующий этому методу.

ГЛАВА 4

ВЫЧИСЛЕНИЕ ИНТЕГРАЛОВ (СЛОЖНЫЕ ОЦЕНКИ)

§ 1. Методы Монте-Карло с повышенной скоростью сходимости

1.1. Выборка по группам. Этот хорошо известный в статистике прием ([24], стр. 218) может быть с успехом использован для уменьшения дисперсии. По идее он весьма близок к методу существенной выборки: здесь также предлагается выбирать больше точек в более «существенных» областях, однако выбор регулируется не специальной плотностью, а указанием количества точек в различных областях.

Итак, пусть требуется вычислить интеграл

$$I = \int_G f(P) p(P) dP. \quad (1)$$

Разобьем область интегрирования G на m частей

$$G = G_1 \cup \dots \cup G_m$$

и введем обозначения

$$p_j = \int_{G_j} p(P) dP, \quad I_j = \int_{G_j} f(P) p(P) dP.$$

Очевидно,

$$p_1 \cup \dots \cup p_m = 1, \quad I_1 \cup \dots \cup I_m = I.$$

В области G_j рассмотрим случайную точку $Q^{(n)}$ с плотностью $p(P)/p_j$ и для оценки I_j воспользуемся

простейшим методом Монте-Карло: так как

$$I_j = M[p_j f(Q^{(j)})],$$

то, выбрав N_j независимых реализаций $Q_1^{(j)}, \dots, Q_{N_j}^{(j)}$ точки $Q^{(j)}$ можем записать оценку для I_j :

$$\theta_{N_j} = \frac{p_j}{N_j} \sum_{s=1}^{N_j} f(Q_s^{(j)}).$$

Складывая такие оценки для всех I_j , получим новую несмещенную оценку

$$\theta_N^* = \sum_{j=1}^m \frac{p_j}{N_j} \sum_{s=1}^{N_j} f(Q_s^{(j)}). \quad (2)$$

Общее количество случайных точек в формуле (2) обозначим по-прежнему через N :

$$N = N_1 + \dots + N_m. \quad (3)$$

Заметим сразу, что оценку (2) можно считать квадратурной суммой со случайными узлами. В самом деле, обычная квадратурная формула записывается в форме

$$I \approx \sum_{k=1}^N C_k f(P_k),$$

где точки P_1, \dots, P_N , принадлежащие G , называются узлами, а числа C_1, \dots, C_N — весами. Если квадратурная формула точна для функций $f(P) = \text{const}$, то $C_1 + \dots + C_N = 1$.

Оценка (2) дает нам такую же приближенную формулу

$$I \approx \sum_{j=1}^m \sum_{s=1}^{N_j} (p_j/N_j) f(Q_s^{(j)}),$$

причем и здесь

$$\sum_{j=1}^m \sum_{s=1}^{N_j} (p_j/N_j) = \sum_{j=1}^m p_j = 1.$$

Предположим, что $f(P) \in L_2(G; p)$. Тогда точность простейшего метода Монте-Карло для расчета I характеризуется дисперсией оценки (14) гл. 3, которая равна

$$D\theta_N = DZ/N,$$

где

$$DZ = \int_G f^2(P) p(P) dP - I^2.$$

Найдем теперь дисперсию оценки (2). Очевидно,

$$D\theta_N^* = \sum_{j=1}^m (p_j / N_j)^2 \sum_{s=1}^{N_j} Df(Q_s^{(j)}),$$

и так как здесь точки $Q_s^{(j)}$ при $s=1, \dots, N_j$ независимы, то

$$D\theta_N^* = \sum_{j=1}^m (p_j^2 / N_j) Df(Q^{(j)}). \quad (4)$$

Легко вычислить, что

$$Df(Q^{(j)}) = \frac{1}{p_j} \int_{G_j} f^2(P) p(P) dP - \left(\frac{I_j}{p_j} \right)^2. \quad (5)$$

Теорема 1. Если разбиение $G = G_1 + \dots + G_m$ и число N фиксированы, то минимум выражения (4) при дополнительном условии (3) равен

$$b_N^2 = \frac{1}{N} \left[\sum_{j=1}^m p_j \sqrt{Df(Q^{(j)})} \right]^2 \quad (6)$$

и реализуется при

$$N_j = N p_j \sqrt{Df(Q^{(j)})} / \sum_{j=1}^m p_j \sqrt{Df(Q^{(j)})}. \quad (7)$$

Доказательство. Согласно (6) величина b_N^2 равна

$$b_N^2 = \left[\sum_{j=1}^m p_j \sqrt{Df(Q^{(j)})/N_j} \cdot \sqrt{N_j/N} \right]^2.$$

Используя неравенство Коши — Буняковского, получим

$$b_N^2 \leq \sum_{j=1}^m \frac{p_j^2 Df(Q^{(j)})}{N_j} \sum_{j=1}^m \frac{N_j}{N} = \sum_{j=1}^m \frac{p_j^2}{N_j} Df(Q^{(j)}),$$

где справа стоит (4). Остается проверить, что при подстановке (7) в (4) получается (6).

В реальных задачах дисперсии $Df(Q^{(j)})$, как правило, заранее неизвестны. Известны лишь вероятности p_j . Оказывается, этого уже достаточно, чтобы выбрать N_j , обеспечивающие уменьшение дисперсии (т. е. неравенство $D\theta_N^* \leq D\theta_N$).

Теорема 2. Если разбиение $G = G_1 + \dots + G_m$ фиксировано, то при

$$N_j = N p_j \quad (8)$$

величина (4) не превосходит $D\theta_N$.

Доказательство. Умножив (5) на p_j и просуммировав по j от 1 до m , получим

$$\sum_{j=1}^m p_j Df(Q^{(j)}) = \int_G f^2(P) p(P) dP - \sum_{j=1}^m (I_j^2 / p_j).$$

Далее

$$\begin{aligned} I^2 &= \left(\sum_{j=1}^m I_j \right)^2 = \left[\sum_{j=1}^m (I_j / \sqrt{p_j}) \sqrt{p_j} \right]^2 \leq \\ &\leq \sum_{j=1}^m (I_j^2 / p_j) \sum_{j=1}^m p_j = \sum_{j=1}^m (I_j^2 / p_j). \end{aligned}$$

Следовательно,

$$\sum_{j=1}^m p_j Df(Q^{(j)}) \leq \int_G f^2(P) p(P) dP - I^2 = DZ.$$

Подставив (8) в (4), получим выражение

$$D\theta_N^*|_{N_j = N p_j} = (1/N) \sum_{j=1}^m p_j Df(Q^{(j)}),$$

которое в силу предыдущего неравенства не превосходит $(1/N) DZ = D\theta_N$. Таким образом теорема доказана.

Теоремы 1 и 2 доказаны в [24] для случая выборки из конечного множества. Если выполнено условие (8), то выборка называется *типической*.

В действительности выбирать N_j по формуле (7) или (8) нельзя, так как N_j обязаны быть целыми. Обычно выбирают любые из ближайших к (7) или (8) целых чисел, лишь бы удовлетворялось условие (3).

Пример. Чтобы вычислить интеграл

$$I = \int_0^1 e^x dx$$

с помощью $N=10$ точек, разобьем $(0, 1)$ на две равные части и выберем в $(0, 1/2)$ всего 4 равномерно распределенные точки $\xi^{(1)}$, а в $(1/2, 1)$ — 6 равномерно распределенных точек $\xi^{(2)}$. Тогда $p_1=p_2=1/2$, $N_1=4$, $N_2=6$ и

$$\theta_{10}^* = \frac{1}{8} \sum_{s=1}^4 \exp \xi_s^{(1)} + \frac{1}{12} \sum_{s=1}^6 \exp \xi_s^{(2)}.$$

Дисперсия этой оценки $D\theta_{10}^* = (1/16) De^{\xi^{(1)}} + (1/24) De^{\xi^{(2)}}$, где

$$De^{\xi^{(1)}} = 2 \int_0^{1/2} e^{2x} dx - \left(2 \int_0^{1/2} e^x dx \right)^2 = e - 1 - 4(\sqrt{e} - 1)^2 = 0,03492,$$

$$De^{\xi^{(2)}} = 2 \int_{1/2}^1 e^{2x} dx - \left(2 \int_{1/2}^1 e^x dx \right)^2 = e^2 - e - 4(e - \sqrt{e})^2 = 0,09493.$$

Следовательно, $D\theta_{10}^* = 0,00614$, в то время как дисперсия простейшего метода $D\theta_{10} = 0,0242$. Значения $\xi^{(1)}$ и $\xi^{(2)}$ легко вычисляются по формулам $\xi^{(1)} = 0,5\gamma$ и $\xi^{(2)} = 0,5(1+\gamma')$.

1.2. Оценки с повышенной скоростью сходимости. Пусть G — конечная область в n -мерном пространстве переменных x_1, \dots, x_n , так что интеграл (1) n -кратный: $P = (x_1, \dots, x_n)$, $dP = dx_1, \dots, dx_n$. Рассмотрим оценку θ_N^* в случае, когда $m=N$ и все $N_j=1$:

$$\theta_N^* = \sum_{j=1}^N p_j f(Q^{(j)}). \quad (9)$$

Обозначим через d_j диаметр*) области G_j и предположим, что существуют положительные постоянные c_1 и c_2 такие, что при всех $j=1, 2, \dots, N$

$$p_j \leq c_1/N, \quad (10)$$

$$d_j \leq c_2/N^{1/n}. \quad (11)$$

Условия (10) — (11) означают, что все области G_j равномерно малы как по вероятности, так и по геометрическим размерам.

*) Диаметром области G называется верхняя грань расстояний между двумя точками этой области:

$$d = \sup_{P_1, P_2 \in G} |P_1 - P_2|.$$

Например, если разбить единичный куб

$$K^n = \{0 < x_1 < 1, \dots, 0 < x_n < 1\},$$

в котором $p(P) \equiv 1$, на $N = M^n$ равновеликих кубов с ребром $1/M$ (рис. 48), то, очевидно, все $p_j = 1/N$, а все

$$d_j = \sqrt[n]{n}/M.$$

Теорема 3. *Предположим, что функция $f(P)$ и все ее частные производные первого порядка $\partial f / \partial x_k$ непрерывны в G и при всех $1 \leq k \leq n$*

$$|\partial f / \partial x_k| \leq L.$$

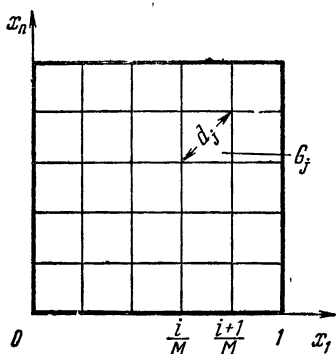


Рис. 48.

Если выполнены условия (10) и (11), то для дисперсии оценки (9) справедливо неравенство

$$D\theta_N^* \leq c^2 L^2 N^{-1-2/n}, \quad (12)$$

где $c = nc_1 c_2$.

Доказательство. Выберем внутри G_j произвольную точку $S_j = (s_{j,1}, \dots, s_{j,n})$ и воспользуемся формулой Тейлора:

$$f(P) = f(S_j) + \sum_{k=1}^n a_{j,k} (x_k - s_{j,k}), \quad (13)$$

где $a_{j,k}$ — значения производных $\partial f / \partial x_k$ в некоторых точках (зависящих и от S_j , и от P), так что все $|a_{j,k}| \leq L$. Из (13) следует, что

$$Df(Q^{(j)}) = D \sum_{k=1}^n a_{j,k} (\xi_k^{(j)} - s_{j,k}),$$

где $\xi_1^{(j)}, \dots, \xi_n^{(j)}$ — координаты случайной точки $Q^{(j)}$.

Воспользуемся известным в теории вероятностей соотношением

$$D \sum_{k=1}^n \eta_k \leq \left[\sum_{k=1}^n \sqrt{D\eta_k} \right]^2$$

и запишем цепочку неравенств:

$$\begin{aligned} \sqrt{Df(Q^{(j)})} &\leq \sum_{k=1}^n \sqrt{D[a_{j,k}(\xi_k^{(j)} - s_{j,k})]} \leq \\ &\leq \sum_{k=1}^n \sqrt{M[a_{j,k}(\xi_k^{(j)} - s_{j,k})]^2} \leq L \sum_{k=1}^n \sqrt{M(\xi_k^{(j)} - s_{j,k})^2}. \end{aligned}$$

Так как $Q^{(j)} \in G_j$, то $|\xi_k^{(j)} - s_{j,k}| \leq d_j$ и, используя (11), получим

$$\sqrt{Df(Q^{(j)})} \leq L \sum_{k=1}^n d_j \leq Lnc_2 N^{-1/n}.$$

Последнее неравенство вместе с (10) позволяет оценить дисперсию $D\theta_N^*$ из (9)

$$D\theta_N^* = \sum_{j=1}^N p_j^2 Df(Q^{(j)}) \leq \sum_{j=1}^N (Lnc_1 c_2 N^{-1-1/n})^2,$$

откуда следует (12).

З а м е ч а н и е. В математической статистике оценки, дисперсии которых убывают быстрее чем $1/N$, иногда называют *сверхэффективными*.

С помощью теоремы 3 оценим погрешность $\theta_N^* - I$. Для этого в известном неравенстве Чебышева

$$P\{|\eta - M\eta| < h\} \geq 1 - (D\eta/h^2), \quad (14)$$

справедливом при любом $h > 0$, положим $\eta = \theta_N^*$, а $h = (1/\varepsilon) \sqrt{D\theta_N^*}$, где $\varepsilon > 0$ — малое число. Получим неравенство $P\{|\theta_N^* - I| < (1/\varepsilon) \sqrt{D\theta_N^*}\} \geq 1 - \varepsilon^2$. В силу (12) тем более имеет место неравенство

$$P\{|\theta_N^* - I| < (cL/\varepsilon) N^{-(1/2)-(1/n)}\} \geq 1 - \varepsilon^2, \quad (15)$$

которое показывает, что со сколь угодно большой вероятностью ошибка $\theta_N^* - I$ убывает как $N^{-(1/2)-(1/n)}$, т. е. быстрее чем $N^{-1/2}$.

Конечно, при больших n это ускорение порядка сходимости незначительно. И поэтому на практике оценка (9) используется редко.

Теорема 3 впервые была доказана В. Дупачем [120] для случая $G = K^n$, $p(P) \equiv 1$ и разбиения, изображенного на рис. 48. См. также [1, 128].

1.3. Сравнение с квадратурными формулами. Фиксируем произвольные точки $S_j \in G_j$ и рассмотрим квадратурную формулу

$$I \approx \sum_{j=1}^N p_j f(S_j). \quad (16)$$

Для того чтобы оценить ее погрешность, умножим (13) на $p(P)$ и проинтегрируем по G_j

$$I_j = p_j f(S_j) + \sum_{k=1}^n \int_{G_j} a_{j,k}(x_k - s_{j,k}) p(P) dP. \quad (17)$$

Интеграл, стоящий справа, легко оценить с помощью (10) и (11):

$$\left| \int_{G_j} a_{j,k}(x_k - s_{j,k}) p dP \right| \leq L d_j p_j \leq L c_1 c_2 N^{-1-1/n}.$$

Суммируя равенства (17) по j от 1 до N и используя последнюю оценку, получим, что в условиях теоремы 3

$$\left| I - \sum_{j=1}^N p_j f(S_j) \right| \leq c L N^{-1/n}. \quad (18)$$

Сравнение (15) и (18) показывает, что порядок вероятностной оценки погрешности (15) на $N^{-1/2}$ лучше, чем порядок оценки (18). Это не случайность. Следует подчеркнуть различный характер обеих оценок: оценка (18) справедлива одновременно для всех функций рассматриваемого класса $D_1(L)$ — с непрерывными частными производными $|\partial f / \partial x_k| \leq L$; можно переписать ее в виде

$$\sup_{f \in D_1(L)} \left| \int_G f(P) p(P) dP - \sum_{j=1}^N p_j f(S_j) \right| \leq c L N^{-1/n}.$$

В то же время неравенство (15) справедливо для одной функции $f(P)$ (хотя и любой) из $D_1(L)$.

Вопрос о наилучших порядках сходимости квадратурных формул, в которых используются значения $f(P)$ в N заданных точках, и недетерминированных методов вычисления интегралов, в которых используются значения $f(P)$ в N случайных точках, исследовался Н. С. Вах-

валовым [2]. Для некоторых классов функций, заданных в K^n , оказалось, что если наилучший порядок сходимости квадратурных формул (на рассматриваемом классе) равен $N^{-\alpha}$, то можно построить недетерминированный метод интегрирования, порядок сходимости которого (для каждой функции класса) будет с большой вероятностью равен $N^{-\alpha-1/2}$. И этот порядок сходимости тоже является наилучшим. (См. упражнение 1 на стр. 159).

Практического применения такие недетерминированные методы интегрирования пока не имеют, вероятно, из-за того, что сами многомерные квадратурные формулы, на базе которых должны строиться эти методы, достаточно сложны.

Другой подход к оценкам с повышенной скоростью сходимости использован в [85], где для некоторых классов функций доказано, что если имеется семейство квадратурных формул с порядком сходимости, равным $N^{-\alpha}$, $0 < \alpha \leq 1$ (на рассматриваемом классе), то, выбрав подынтегральную функцию из этого класса случайно, можно с большой вероятностью гарантировать для нее порядок сходимости $N^{-\alpha-1/2}$, если $\alpha + 1/2 \leq 1$ и N^{-1} , если $\alpha + 1/2 \geq 1$.

§ 2. Случайные квадратурные формулы

2.1. Случайные квадратурные формулы. Обычно *квадратурной формулой* для вычисления интегралов вида

$$I = \int_G f(P) p(P) dP$$

при фиксированной функции $p(P)$ называют выражение

$$I \approx \sum_{j=1}^N C_j f(P_j), \quad (19)$$

где и узлы $P_1, \dots, P_N \in G$ и веса C_1, \dots, C_N заданы. Естественно назвать *случайной квадратурной формулой* выражение

$$I \approx \sum_{j=1}^N \kappa_j f(Q^{(j)}), \quad (20)$$

где $Q^{(1)}, \dots, Q^{(N)}$ — случайные точки, определенные в G (случайные узлы), а $\kappa_1, \dots, \kappa_N$ — случайные скалярные величины (случайные веса).

Рассмотренные в § 1 оценки (2) и (9) представляют собой частные случаи суммы (20) с фиксированными (не случайными) весами.

Другой класс случайных квадратурных формул, в которых веса κ_j являются функциями от узлов $\kappa_j = \kappa_j(Q^{(1)}, \dots, Q^{(N)})$, был построен С. М. Ермаковым и В. Г. Золотухиным [34], использовавшими структуру интерполяционных квадратурных формул, точных для заданной системы функций $\varphi_0(P), \varphi_1(P), \dots, \varphi_m(P)$.

2.2. Некоторые свойства интерполяционных квадратурных формул. Рассмотрим произвольную n -мерную область G . Все функции будем считать кусочно непрерывными и принадлежащими $L_2(G; 1)$. Выберем систему ортонормированных функций $\varphi_0(P), \varphi_1(P), \dots, \varphi_m(P)$ так, что

$$\int_G \varphi_k(P) \varphi_j(P) dP = \delta_{kj}.$$

Для приближенной оценки интеграла

$$I = \int_G f(P) \varphi_0(P) dP \quad (21)$$

рассмотрим квадратурную формулу

$$I \approx \sum_{j=0}^m C_j f(P_j). \quad (22)$$

Фиксируем произвольные различные узлы P_0, P_1, \dots, P_m и выберем веса C_0, C_1, \dots, C_m так, чтобы эта формула была точной для функций $f = \varphi_0(P), f = \varphi_1(P), \dots, f = \varphi_m(P)$.

Полученную при таком выборе формулу (22) удобно записать в форме отношения двух определителей

$$I \approx W_f(P_0, \dots, P_m) / W_{\varphi_0}(P_0, \dots, P_m), \quad (23)$$

где по определению

$$W_g(P_0, \dots, P_m) = \begin{vmatrix} g(P_0) & \varphi_1(P_0) & \dots & \varphi_m(P_0) \\ g(P_1) & \varphi_1(P_1) & \dots & \varphi_m(P_1) \\ \dots & \dots & \dots & \dots \\ g(P_m) & \varphi_1(P_m) & \dots & \varphi_m(P_m) \end{vmatrix}. \quad (24)$$

В самом деле, если в (23) подставить $f = \varphi_j$ при $1 \leq j \leq m$, то в определителе $W_{\varphi_j}(P_0, \dots, P_m)$ окажутся два совпадающих столбца и он обратится в нуль; в этом случае $I = \int_G \varphi_j(P) \varphi_0(P) dP = 0$. Если же $f = \varphi_0$, то числитель и знаменатель в (23) совпадут, но в этом случае $I = \int_G \varphi_0^2(P) dP = 1$. Единственное ограничение применимости формулы (23) — это требование, чтобы $W_{\varphi_0}(P_0, \dots, P_m) \neq 0$. Если оно выполнено, то формула точна для любых линейных комбинаций вида $f = a_0 \varphi_0(P) + \dots + a_m \varphi_m(P)$, и поэтому представляет собой в каком-то смысле «разумную» квадратурную формулу.

Лемма 1. Пусть $\varphi_0(P), \varphi_1(P), \dots, \varphi_m(P)$ — любая совокупность ортонормированных функций в G . Тогда

$$\int_G \dots \int_G W_{\varphi_0}^2 dP_0 \dots dP_m = (m+1)! \quad (25)$$

Доказательство проведем индукцией по m . При $m=2$

$$W_{\varphi_0}(P_0, P_1) = \begin{vmatrix} \varphi_0(P_0) & \varphi_1(P_0) \\ \varphi_0(P_1) & \varphi_1(P_1) \end{vmatrix}.$$

Отсюда

$$W_{\varphi_0}^2 = \varphi_0^2(P_0) \varphi_1^2(P_1) - 2\varphi_0(P_0) \varphi_1(P_0) \varphi_0(P_1) \varphi_1(P_1) + \varphi_0^2(P_1) \varphi_1^2(P_0),$$

и

$$\int_G \int_G W_{\varphi_0}^2 dP_0 dP_1 = 1 + 1 = 2.$$

Допустим теперь, что для определителей порядка m формула (25) справедлива и интеграл равен $m!$. Рассмотрим определитель $(m+1)$ -го порядка $W_{\varphi_0}(P_0, \dots, P_m)$ и разложим его по элементам первой строки:

$$W_{\varphi_0} = \sum_{j=1}^m (-1)^j \varphi_j(P_0) \begin{vmatrix} \varphi_0(P_1) \dots \varphi_{j-1}(P_1) & \varphi_{j+1}(P_1) & \dots \varphi_m(P_1) \\ \dots & \dots & \dots \\ \varphi_0(P_m) \dots \varphi_{j-1}(P_m) & \varphi_{j+1}(P_m) & \dots \varphi_m(P_m) \end{vmatrix} \quad (26)$$

Нетрудно заметить, что каждый из определителей в (26) есть определитель типа W_{Φ_0} порядка m : его столбцы образованы значениями m ортонормированных функций в m независимых точках. Согласно индукционному допущению

$$\int_G \dots \int_G \left| \begin{array}{cccc} \Phi_0(P_1) & \dots & \Phi_{j-1}(P_1) & \Phi_{j+1}(P_1) & \dots & \Phi_m(P_1) \\ \vdots & & \vdots & \vdots & & \vdots \\ \Phi_0(P_m) & \dots & \Phi_{j-1}(P_m) & \Phi_{j+1}(P_m) & \dots & \Phi_m(P_m) \end{array} \right|^2 dP_1 \dots dP_m = m!$$

Поэтому, возведя (26) в квадрат и проинтегрировав (сперва по P_0 , затем по P_1, \dots, P_m), получим

$$\int_G \dots \int_G W_{\Phi_0}^2 dP_0 \dots dP_m = \sum_{j=0}^m \int_G \Phi_j^2(P_0) dP_0 m! = (m+1)!$$

Лемма 2. Пусть $\Phi_0(P), \Phi_1(P), \dots, \Phi_m(P), \Psi(P)$ — любая совокупность ортонормированных функций в G . Тогда

$$\int_G \dots \int_G W_{\Phi_0} W_{\Psi} dP_0 \dots dP_m = 0.$$

Доказательство. Выберем новую функцию $\Phi = (\sqrt{2}/2)[\Phi_0(P) + \Psi(P)]$, которая также ортогональна ко всем $\Phi_1(P), \dots, \Phi_m(P)$ и нормирована:

$$\int_G \Phi^2(P) dP = \frac{1}{2} \left[\int_G \Phi_0^2 dP + 2 \int_G \Phi_0 \Psi dP + \int_G \Psi^2 dP \right] = 1.$$

По известным правилам действий с определителями

$$W_{\Phi} = (\sqrt{2}/2) (W_{\Phi_0} + W_{\Psi}),$$

так, что

$$W_{\Phi}^2 = (1/2) (W_{\Phi_0}^2 + 2W_{\Phi_0}W_{\Psi} + W_{\Psi}^2).$$

Проинтегрировав это равенство по всем переменным и используя лемму 1, получим, что

$$\begin{aligned} (m+1)! &= \\ &= \frac{1}{2} \left[(m+1)! + 2 \int_G \dots \int_G W_{\Phi_0} W_{\Psi} dP_0 \dots dP_m + (m+1)! \right], \end{aligned}$$

откуда вытекает утверждение леммы 2.

2.3. Случайные интерполяционные квадратурные формулы. Обозначим для краткости $n(m+1)$ -мерную область, точки которой $T = (P_0, P_1, \dots, P_m)$, через $B = G \times \dots \times G$ и пусть $dT = dP_0 dP_1 \dots dP_m$. Пусть далее B_0 — множество точек T , в которых $W_{\varphi_0} = 0$, а $B^+ = B - B_0$.

Определим случайные точки $Q^{(0)}, \dots, Q^{(m)}$ в G и в качестве оценки интеграла (21) рассмотрим случайную величину

$$\check{\theta}[f] = \begin{cases} \frac{W_f(Q^{(0)}, \dots, Q^{(m)})}{W_{\varphi_0}(Q^{(0)}, \dots, Q^{(m)})}, & \text{если } (Q^{(0)}, \dots, Q^{(m)}) \in B^+, \\ 0, & \text{если } (Q^{(0)}, \dots, Q^{(m)}) \in B_0. \end{cases} \quad (27)$$

Теорема 4. Если совместная плотность распределения случайных точек $Q^{(0)}, \dots, Q^{(m)}$ в B равна

$$p(P_0, \dots, P_m) = \frac{1}{(m+1)!} [W_{\varphi_0}(P_0, \dots, P_m)]^2, \quad (28)$$

то для любой функции $f(P)$ из $L_2(G; 1)$

$$M\check{\theta}[f] = \int_G f(P) \varphi_0(P) dP, \quad (29)$$

$$D\check{\theta}[f] \leq \int_G f^2(P) dP - \sum_{j=0}^m \left[\int_G f(P) \varphi_j(P) dP \right]^2. \quad (30)$$

Доказательство. Выберем произвольную конечную функцию $f(P)$ из $L_2(G; 1)$ и обозначим через c_j ее коэффициенты Фурье $c_j = \int_G f(P) \varphi_j(P) dP$. Пусть

$$a^2 = \int_G \left[f(P) - \sum_{j=0}^m c_j \varphi_j(P) \right]^2 dP = \int_G f^2(P) dP - \sum_{j=0}^m c_j^2.$$

Если $a^2 \neq 0$, то, введя функцию $\psi = (1/a) \left[f(P) - \sum_{j=0}^m c_j \varphi_j(P) \right]$ получим представление

$$f(P) = \sum_{j=0}^m c_j \varphi_j(P) + a\psi(P). \quad (31)$$

Легко проверить, что $\psi(P)$ ортогональна ко всем $\varphi_j(P)$ и нормирована:

$$\int_G \psi^2(P) dP = 1; \quad \int_G \psi(P) \varphi_j(P) dP = 0, \quad 0 \leq j \leq m.$$

Если $a=0$, то представление (31) также справедливо: в качестве $\psi(P)$ можно выбрать любую нормированную функцию, ортогональную ко всем $\varphi_j(P)$, $0 \leq j \leq m$.

Подставив (31) в (24), получим, что

$$W_f = \sum_{j=0}^m c_j W_{\varphi_j} + a W_{\psi} = c_0 W_{\varphi_0} + a W_{\psi}.$$

С помощью последней формулы нетрудно вычислить математическое ожидание оценки (27):

$$\begin{aligned} M\check{\theta}[f] &= \int_{B^+} (W_f/W_{\varphi_0}) p dT = \frac{1}{(m+1)!} \int_{B^+} W_f W_{\varphi_0} dT = \\ &= \frac{1}{(m+1)!} \int_B W_f W_{\varphi_0} dT = \frac{c_0}{(m+1)!} \int_B W_{\varphi_0}^2 dT + \\ &\quad + \frac{a}{(m+1)!} \int_B W_{\varphi_0} W_{\psi} dT. \end{aligned}$$

Из последних двух интегралов первый равен $(m+1)!$ по лемме 1, а второй равен нулю по лемме 2. Таким образом, $M(\check{\theta}[f]) = c_0$, что равносильно (29).

Перейдем к вычислению дисперсии:

$$\begin{aligned} M(\check{\theta}[f])^2 &= \int_{B^+} (W_f/W_{\varphi_0})^2 p dT = \frac{1}{(m+1)!} \int_{B^+} W_f^2 dT = \\ &= \frac{1}{(m+1)!} \int_{B^+} [c_0^2 W_{\varphi_0}^2 + 2ac_0 W_{\varphi_0} W_{\psi} + a^2 W_{\psi}^2] dT = \\ &= \frac{c_0^2}{(m+1)!} \int_B W_{\varphi_0}^2 dT + \frac{2ac_0}{(m+1)!} \int_B W_{\varphi_0} W_{\psi} dT + \\ &\quad + \frac{a^2}{(m+1)!} \int_{B^+} W_{\psi}^2 dT. \end{aligned}$$

Так как здесь $\int_{B^+} W_{\psi}^2 dT \leq \int_B W_{\psi}^2 dT = (m+1)!$ — по той же лемме 1, то $M(\check{\theta}[f])^2 \leq c_0^2 + a^2$, откуда вытекает

неравенство, равносильное (30):

$$D\check{\theta}[f] = M(\check{\theta}[f])^2 - c_0^2 \leq a^2.$$

Впервые эта теорема была доказана в [34], а уточнение к ней — в [32]. Из доказательства видно, что знак равенства в (30) реализуется тогда и только тогда, когда

$$\int_{B_0} W_{\varphi}^2 dT = 0. \quad (32)$$

Это условие будет выполнено для любых функций $f(P)$, если объем B_0 ($n(m+1)$ -мерный) равен нулю. Иными словами, если $W_{\varphi_0} = 0$ только на многообразных меньшего числа измерений, чем $n(m+1)$, как, например, $P_0 = P_1$ и т. п.

Равенство (32) не будет выполнено для некоторых функций $f(P)$, если функции $\varphi_0(P), \dots, \varphi_m(P)$ линейно зависимы в какой-то области $G' \subset G$ с положительным n -мерным объемом: в этом случае $W_{\varphi_0} = 0$ в области $G' \times \dots \times G'$ с положительным $n(m+1)$ -мерным объемом, и объем B_0 положителен.

Легко показать, что оба эти случая возможны. В самом деле, пусть G — интервал $0 < x < 1$, $m=1$, так что B — квадрат $\{0 < x_0 < 1, 0 < x_1 < 1\}$. Если выбрать $\varphi_0(x) = 1$, $\varphi_1(x) = \sqrt{3}(1-2x)$, то

$$W_{\varphi_0}(x_0, x_1) = \begin{vmatrix} 1 & \sqrt{3}(1-2x_0) \\ 1 & \sqrt{3}(1-2x_1) \end{vmatrix} = 2\sqrt{3}(x_0 - x_1)$$

и множество B_0 состоит из точек, расположенных на диагонали $x_0 = x_1$ квадрата B . Если выбрать $\varphi_0(x) = 1$, $\varphi_1(x) = \operatorname{sgn}(1/2 - x)$, то легко вычислить, что

$$W_{\varphi_0}(x_0, x_1) = \begin{vmatrix} 1 & \varphi_1(x_0) \\ 1 & \varphi_1(x_1) \end{vmatrix} = \varphi_1(x_1) - \varphi_1(x_0)$$

и $W_{\varphi_0}(x_0, x_1) \equiv 0$ в двух четвертях $\{0 < x_0 < 1/2, 0 < x_1 < 1/2\}$ и $\{1/2 \leq x_0 < 1, 1/2 \leq x_1 < 1\}$ квадрата B^* .

Верхняя граница (30) для дисперсии $D\check{\theta}[f]$ имеет простой геометрический смысл: она равна квадрату расстояния (в метрике пространства L_2) от функции $f(P)$ до линейного подпространства, определяемого функциями $\varphi_0(P), \dots, \varphi_m(P)$.

*) В качестве $\varphi_1(x)$ в этом примере использована вторая функция системы Хаара [82].

Пример. Рассмотрим интеграл $I = \int_0^1 f(x) dx$. Пусть $m=1$ и заданы две ортонормированные функции: $\varphi_0(x)=1$, $\varphi_1(x) = \sqrt{3}(1-2x)$.

Обозначим через ξ и η две случайные точки (вместо $Q^{(0)}$ и $Q^{(1)}$) и вычислим определители, входящие в (27):

$$W_f = f(\xi) \varphi_1(\eta) - f(\eta) \varphi_1(\xi) = \sqrt{3} [f(\xi)(1-2\eta) - f(\eta)(1-2\xi)],$$

$$W_{\varphi_0} = \varphi_0(\xi) \varphi_1(\eta) - \varphi_0(\eta) \varphi_1(\xi) = 2\sqrt{3}(\xi - \eta).$$

Из (27) и (29) вытекает оценка интеграла I :

$$\check{\theta}[f] = 2^{-1}(\xi - \eta)^{-1} [(1-2\eta)f(\xi) - (1-2\xi)f(\eta)],$$

где ξ и η имеют совместную плотность распределения

$$p_{\xi, \eta}(x, y) = 6(x-y)^2, \quad 0 < x, y < 1. \quad (33)$$

Для расчета интеграла $I = \int_0^1 e^x dx$ по десяти значениям подынтегральной функции запишем оценку

$$\check{\theta}_{10} = \frac{1}{5} \sum_{i=1}^5 \frac{(1-2\eta_i) e^{\xi_i} - (1-2\xi_i) e^{\eta_i}}{2(\xi_i - \eta_i)}. \quad (34)$$

Дисперсия этой оценки, согласно (30), где имеет место знак равенства, есть

$$\begin{aligned} D\check{\theta}_{10} &= \frac{1}{5} D\check{\theta}[e^x] = \frac{1}{5} \left\{ \int_0^1 e^{2x} dx - I^2 - \left[3 \int_0^1 (1-2x) dx \right]^2 \right\} = \\ &= 0,2 \{ 0,5(e^2 - 1) - (e-1)^2 - 3(3-e)^2 \} = 0,000788. \end{aligned}$$

Уменьшение дисперсии по сравнению с простейшим методом весьма значительное.

Так как $p_{\xi, \eta}(x, y) \leq 6$, то удобно находить значения ξ и η методом Неймана (п. 5.3 гл. 2): выбираем три случайных числа γ_i , γ'_i , γ''_i и проверяем условие $\gamma''_i < (\gamma_i - \gamma'_i)^2$. Если оно выполнено, то $\xi_i = \gamma_i$, $\eta_i = \gamma'_i$. Эффективность отбора $\vartheta = 1/6$. Пример расчета по формуле (34) приведен в табл. 1. Результат этого расчета $\check{\theta}_{10} = 1,731$.

2.4. Замечания. Методу п. 2.3 посвящено несколько работ [20, 31, 137]. Исследователей привлекает большая общность метода и значительное уменьшение дисперсии. Однако он имеет и свои недостатки. Во-первых, пока нет удобных и достаточно общих приемов для разыгрывания точек $Q^{(0)}, \dots, Q^{(m)}$ с плотностью (28)*. Во-вто-

*) Если известна верхняя граница плотности $p(Q^{(0)}, \dots, Q^{(m)}) \leq c$, то для разыгрывания точек $Q^{(0)}, \dots, Q^{(m)}$ можно использовать метод Неймана, эффективность которого $\vartheta = 1/c$.

Таблица 1

i	ξ_i	η_i	Числитель	Знаменатель	$\vartheta[f_i]$
1	0,69186	0,03393	2,2588	0,6579	1,716
2	0,91682	0,12904	2,8043	0,7878	1,780
3	0,22561	0,70177	-1,6128	-0,4762	1,694
4	0,95114	0,16021	2,8182	0,7909	1,782
5	0,40837	0,93030	-1,7569	-0,5219	1,683

рых, в формуле (27), вообще говоря, присутствуют отрицательные случайные веса (ср. пример п. 2.3); а формула со знакопеременными весами практически плоха, если число слагаемых велико. Чтобы устранить эти недостатки, надо использовать другие плотности $p(P_0, \dots, P_m)$, отличные от (28) (см. [33]). Наиболее подробно изучен случай, когда все точки $Q^{(1)}, \dots, Q^{(m)}$ представляют собой функции от одной случайной точки $Q^{(0)}$ (Б. Л. Грановский [20, 21]).

§ 3. Использование смещенных оценок

Все указанные в гл. 3 и 4 оценки θ интеграла

$$I = \int_G f(P) p(P) dP \quad (35)$$

представляют собой несмещенные оценки: $M\theta = I$. Большинство авторов ([33, 130]) включают условие несмещенности даже в определение случайной квадратурной формулы (20), требуя, чтобы $\sum_{i=1}^N \kappa_i f(Q^{(i)}) = I$ для всех $f(P) \in L_2(G; p)$. И мы отдали дань этой традиции в гл. 3, п. 2.4.

Однако, как указано в гл. 3, п. 1.7, при больших N (когда количество используемых значений $f(Q^{(i)})$ велико) для практических целей достаточно требовать только, чтобы оценка была состоятельной: $\theta \xrightarrow{P} I$. В самом деле, обычно порядок дисперсии $D\theta = O(N^{-1})$ и вероятная ошибка r_N оказывается $r_N = O(N^{-1/2})$. В тех случаях, когда смещение $\theta - I = O(N^{-1})$, ясно, что при больших N порядок ошибки определяется величиной r_N , а не смещением.

3.1. Взвешенная равномерная выборка. Предположим, что нам удалось найти допустимую по отношению к $f(P)$ плотность $p(P)$, приблизительно пропорциональную $|f(P)|$. Д. Хэндсcombe [137] предложил в качестве оценки интеграла по конечной области G

$$I_0 = \int_G f(P) dP$$

использовать величину

$$\bar{\theta}_N = \frac{\sum_{i=1}^N f(Q'_i)}{\sum_{i=1}^N p(Q'_i)}, \quad (36)$$

где Q'_1, \dots, Q'_N — независимые случайные точки, равномерно распределенные в G .

Естественно сравнить оценку (36) с оценкой метода существенной выборки, где рекомендуется использовать такую же плотность $p(P)$. Согласно п. 3.2.1 гл. 3 получаем оценку

$$\theta_N = \frac{1}{N} \sum_{i=1}^N \frac{f(Q_i)}{p(Q_i)},$$

где Q_1, \dots, Q_N — независимые реализации случайной точки Q с плотностью $p(P)$. Сразу видно, что при сложных $p(P)$ расчет по формуле (36) гораздо проще, так как не требуется разыгрывать точки Q_i с этой плотностью.

Предположим, что $\int_G |f(P)| dP < \infty$. Тогда по теореме Хинчина (стр. 87) величина $(V_G/N) \sum_{i=1}^N f(Q'_i) \xrightarrow{P} I_0$,

а величина $(V_G/N) \sum_{i=1}^N p(Q'_i) \xrightarrow{P} 1$, когда $N \rightarrow \infty$ (здесь V_G — объем области G). Следовательно, по известному свойству сходимости по вероятности, величина $\bar{\theta}_N$ также сходится по вероятности к I_0 , так что оценка $\bar{\theta}_N$ состоятельна.

Оценка (36) исследовалась в работе М. Поуэлла и Дж. Свэнна [164]. При некоторых предположениях относительно $f(P)$ и $p(P)$ удалось доказать, что, когда

$N \rightarrow \infty$, дисперсия

$$D\bar{\theta}_N = \frac{V_G}{N} \int [f(P) - I_0 p(P)]^2 dP + O\left(\frac{1}{N^2}\right), \quad (37)$$

а смещение

$$M\bar{\theta}_N - I_0 = O\left(\sqrt{D\bar{\theta}_N/N}\right).$$

Легко видеть, что если $f(P) \geq 0$ и $p = \hat{p}(P) = f(P)/I_0$, то главные члены в $D\bar{\theta}_N$ и $M\bar{\theta}_N - I_0$ обращаются в нули (ср. с методом существенной выборки).

Пример. Вычислить интеграл $I = \int_0^1 e^x dx = e - 1$. Пусть $p(x) = = (2/3)(1+x)$. Оценка (36) при $N=10$ равна

$$\bar{\theta}_{10} = 1,5 \sum_{i=1}^{10} e^{\gamma_i} \left/ \left(10 + \sum_{i=1}^{10} \gamma_i \right) \right.$$

Главный член дисперсии в этом примере легко вычисляется:

$$\frac{1}{10} \int_0^1 [e^x - I p(x)]^2 dx = \frac{1}{10} (e - 1) \left(\frac{1}{2} - \frac{5}{6} e + \frac{28}{27} I \right) = 0,00287,$$

в то время как для простейшего метода $D\theta_{10} = 0,0242$.

3.2. Простейший метод Монте-Карло с поправочным множителем. Пусть требуется вычислить интеграл вида (35), где $p(P)$ — заданная плотность вероятностей, определенная в G и Q_1, \dots, Q_N — независимые реализации случайной точки Q с плотностью $p(P)$. В качестве оценки интеграла I рассмотрим величину

$$\bar{\theta}_N = \left[\frac{1}{N} \sum_{i=1}^N f(Q_i) \right] \left[\frac{1}{N} \sum_{i=1}^N v(Q_i) \right], \quad (38)$$

где функция $v(P)$ пока не определена. Если $Mv(Q) = 1$, то оценка (38) будет состоятельной (доказывается так же, как в п. 3.1).

Оценка (38) сходна с оценкой (36), так как при больших N с большой вероятностью

$$\begin{aligned} \left[\frac{1}{N} \sum_{i=1}^N p_i \right]^{-1} &= \left[1 - \frac{1}{N} \sum_{i=1}^N (1 - p_i) \right]^{-1} \approx \\ &\approx 1 + \frac{1}{N} \sum_{i=1}^N (1 - p_i) = \frac{1}{N} \sum_{i=1}^N (2 - p_i). \end{aligned}$$

В то же время для $\bar{\theta}_N$ гораздо легче вычислить математическое ожидание и дисперсию.

Теорема 5. Если $f(P) \in L_2(G; p)$, $v(P) \in L_2(G; p)$ и $Mv(Q) = 1$, то

$$M\bar{\theta}_N = I - \frac{1}{N} \int_G f(1-v) p dP, \quad (39)$$

$$D\bar{\theta}_N = \frac{1}{N} \int_G [f - I(2-v)]^2 p dP + O\left(\frac{1}{N^2}\right). \quad (40)$$

Доказательство. Для краткости обозначим $f_i = f(Q_i)$, $v_i = v(Q_i)$. Заметив, что при $i \neq k$ величины эти независимы, запишем

$$\begin{aligned} M\bar{\theta}_N &= \frac{1}{N^2} M \sum_{i,k=1}^N f_i v_k = \frac{1}{N^2} \left[\sum_{i \neq k} M f_i M v_k + \sum_{i=1}^N M (f_i v_i) \right] = \\ &= N^{-2} [N(N-1) M f M v + N M (f v)]. \end{aligned} \quad (41)$$

Так как $Mf = I$, $Mv = 1$, то

$$M\bar{\theta}_N = I - N^{-1} M (f - f v),$$

что равносильно (39).

Перейдем к вычислению дисперсии $D\bar{\theta}_N$. Так как

$$M\bar{\theta}_N^2 = \frac{1}{N^4} M \sum_{i,j,k,l=1}^N f_i f_j v_k v_l,$$

то в сумме необходимо выделить слагаемые с четырьмя различными индексами (i, j, k, l) , с тремя различными индексами, например вида (i, i, k, l) с двумя различными индексами вида (i, i, k, k) или вида (i, i, i, l) и со всеми совпадающими индексами (i, i, i, i) . Тогда нетрудно получить, что

$$\begin{aligned} M\bar{\theta}_N^2 &= N^{-4} \{ N(N-1)(N-2)(N-3) (Mf)^2 (Mv)^2 + \\ &+ N(N-1)(N-2) [M(f^2) (Mv)^2 + 4M(fv) Mf Mv + \\ &+ (Mf)^2 M(v^2)] + N(N-1) [M(f^2) M(v^2) + 2\{M(fv)\}^2 + \\ &+ 2Mf M(fv^2) + 2Mv M(f^2 v)] + NM(f^2 v^2) \}. \end{aligned}$$

Вычитая из $M\bar{\theta}_N^2$ квадрат выражения (41) и принимая

во внимание равенства $Mf=I$, $Mv=1$, найдем выражение для дисперсии

$$\begin{aligned} D\bar{\bar{\theta}}_N = N^{-4} \{ & N(N-1)(-4N+6)I^2 + \\ & + N(N-1)(N-2)[M(f^2) + I^2M(v^2)] + \\ & + 2N(N-1)(N-4)IM(fv) + N(N-2)[M(fv)]^2 + \\ & + N(N-1)[M(f^2)M(v^2) + 2IM(fv^2) + 2M(f^2v)] + \\ & + NM(f^2v^2) \}. \end{aligned}$$

В этом выражении легко выделить главные члены

$$\begin{aligned} D\bar{\bar{\theta}}_N = N^{-1} \{ & -4I^2 + M(f^2) + I^2M(v^2) + 2IM(fv) \} + \\ & + O(N^{-2}) = N^{-1}M[(f-2I+Iv)^2] + O(N^{-2}). \end{aligned}$$

Последняя формула совпадает с (40), и таким образом теорема доказана.

Из (39) видно, что оценка $\bar{\bar{\theta}}_N$ будет несмещенной для любой функции $f(P)$ тогда и только тогда, когда $v(P) \equiv 1$. В этом случае (38) обращается в оценку простейшего метода Монте-Карло. Однако для каждой конкретной $f(P)$ существует бесконечно много таких $v(P)$, что $\int_0^1 f(1-v)p dP = 0$.

Из (40) видно, что если $v(P) = 2 - f(P)/I$, то главный член выражения $D\bar{\bar{\theta}}_N$ обращается в нуль и $D\bar{\bar{\theta}}_N = O(N^{-2})$. При этом смещение в нуль не обращается и, согласно (39), равно

$$M\bar{\bar{\theta}}_N - I = -(IN)^{-1} Df(Q).$$

В этом смысле оценка $\bar{\bar{\theta}}_N$ хуже, чем оценка $\bar{\theta}_N$. Неясно, однако, играет ли указанное свойство какую-либо роль на практике, когда выбор $p = f/I$ или $v = 2 - f/I$ в точности невозможен. Заметим, что функция $v(P)$ в теореме 5 не обязана быть знакопостоянной, так что если $h(P) \approx f(P)$ и значение интеграла $Mh(Q) = \int_0^1 h(P)p(P)dP = C \neq 0$ известно, то целесообразно положить $v = 2 - h(P)/C$.

Пример. Вычислить интеграл $I = \int_0^1 e^x dx = 1,7183$. Так как

$$h(x) = 1+x \approx e^x, \quad \text{то положим} \quad v = 2 - (2/3)(1+x) = (2/3)(2-x).$$

Согласно (38) при $N=10$ получим оценку

$$\bar{\theta}_{10} = \frac{2}{300} \left(\sum_{i=1}^{10} e^{\gamma_i} \right) \left(20 - \sum_{i=1}^{10} \gamma_i \right).$$

Главный член дисперсии тот же, что в примере п. 3.1, так что $D\theta_{10} \approx 0,00287$, и вероятная ошибка $r_{10} = 0,675 \sqrt{D\bar{\theta}_{10}} \approx 0,036$. Смещение же равно $M\bar{\theta}_{10} - I = -0,1(1 - e/3) = -0,0094$, т. е. в несколько раз меньше. (С увеличением N различие это еще увеличивается: например, при $N=100$ вероятная ошибка $r_{100} \approx 0,011$, а смещение равно $-0,00094$).

3.3. Численный пример. Большинство оценок в гл. 3 и 4 иллюстрировались одним и тем же примером: формулой для расчета интеграла

$$I = \int_0^1 e^x dx = 1,718.$$

Выпишем все эти формулы при $N=10$ (точнее, с использованием десяти значений подынтегральной функции).

Во-первых, две грубые оценки:

1. Простейший метод

$$\theta_{10} = \frac{1}{10} \sum_{i=1}^{10} e^{\gamma_i}.$$

2. Геометрический метод

$$\theta_{10} = \frac{1}{10} \sum_{i=1}^{10} \tilde{z}_i, \quad \text{где} \quad \tilde{z}_i = \begin{cases} e, & \text{если } e^{\gamma'_i} < e^{\gamma_i}, \\ 0, & \text{если } e^{\gamma'_i} \geq e^{\gamma_i}. \end{cases}$$

Во-вторых, четыре оценки, соответствующие основным методам уменьшения дисперсии:

3. Выделение главной части $h=1+x$

$$\theta_{10} = \frac{1}{2} + \frac{1}{10} \sum_{i=1}^{10} (e^{\gamma_i} - \gamma_i).$$

4. Существенная выборка с плотностью $p = (2/3)(1+x)$

$$\theta_{10} = \frac{3}{20} \sum_{i=1}^{10} \frac{e^{\xi_i}}{1 + \xi_i}, \quad \text{где} \quad \xi_i = \sqrt{1 + 3\gamma_i} - 1.$$

5. Симметризованная оценка

$$= \frac{1}{10} \sum_{i=1}^5 (e^{\gamma_i} + e^{1-\gamma_i}).$$

6. Выборка по двум группам

$$\theta_{10} = \frac{1}{8} \sum_{i=1}^4 e^{\xi_i^{(1)}} + \frac{1}{12} \sum_{i=1}^6 e^{\xi_i^{(2)}}$$

где

$$\xi_i^{(1)} = \frac{\gamma_i}{2}, \quad \xi_i^{(2)} = \frac{1 + \gamma_{i+4}}{2}.$$

Затем две оценки, соответствующие двухэтапным схемам расчета:

7. Выделение главной части $h = \alpha_0(1+x)$ при $\alpha_0 = 1,690$

$$\theta_{10} = 0,8451 + \frac{1}{10} \sum_{i=1}^{10} (e^{\gamma_i} - 1,690\gamma_i).$$

8. Существенная выборка с $p = (1 + \alpha_0 x) / (1 + 0,5\alpha_0)$ при $\alpha_0 = 1,81$

$$\theta_{10} = 0,1905 \sum_{i=1}^{10} \frac{e^{\gamma_i}}{1 + 1,81\xi_i},$$

где

$$\xi_i = 0,5525 (\sqrt{1 + 6,896\gamma_i} - 1).$$

Далее случайная квадратурная формула интерполяционного типа:

$$9. \quad \theta_{10} = \frac{1}{10} \sum_{i=1}^5 \frac{(1 - 2\eta_i) e^{\xi_i} - (1 - 2\xi_i) e^{\eta_i}}{\xi_i - \eta_i},$$

где для расчета пары ξ_i, η_i надо выбрать три случайных числа $\gamma_i, \gamma'_i, \gamma''_i$ и проверить условие $\gamma''_i < (\gamma_i - \gamma'_i)^2$. Если это условие выполнено, то $\xi_i = \gamma_i, \eta_i = \gamma'_i$, в противном случае надо выбрать новую тройку случайных чисел. В среднем на получение каждой пары ξ_i, η_i придется затратить 6 проб.

И, наконец, две смещенные оценки:

10. Взвешенная равномерная выборка

$$\theta_{10} = \frac{3}{2} \left(\sum_{i=1}^{10} e^{\gamma_i} \right) / \left(10 + \sum_{i=1}^{10} \gamma_i \right).$$

11. Простейшая оценка с поправочным множителем

$$\theta_{10} = \frac{2}{300} \left(\sum_{i=1}^{10} e^{\gamma_i} \right) \left(20 - \sum_{i=1}^{10} \gamma_i \right).$$

В табл. 2 сравниваются трудоемкости $t_{10} \cdot D\theta_{10}$ этих оценок по отношению к вычислительной машине БЭСМ-4: t_{10} — это время расчета θ_{10} в миллисекундах (Для смещенных оценок в качестве $D\theta_{10}$ приведено значение главного члена дисперсии.)

Таблица 2

	Грубые оценки		Основные методы уменьшения дисперсии					Двухэтапные схемы		Сл. кв. ф-ла	Смещенные оценки	
	1	2	3	4	5	6	7	8	9	10	11	
$D\theta_{10}$	0,0242	0,172	0,0044	0,0027	0,00078	0,0061	0,00039	0,00040	0,00079	0,0029	0,0029	
t_{10}	20,3	23,3	20,8	26,1	19,4	22,1	24,8	28,0	47,1	20,9	20,8	
$t_{10} \cdot D\theta_{10}$	0,49	4,0	0,091	0,070	0,015	0,13	0,009	0,011	0,037	0,061	0,069	

Таблица 3

	Грубые оценки		Основные методы уменьшения дисперсии						Двухэтапные схемы		Сл. кв. ф-ла	Смещенные оценки	
	1	2	3	4	5	6	7	8	9	10	11		
θ_{10}	1,901	1,359	1,798	1,782	1,708	1,804	1,726	1,745	1,731	1,778	1,770		
$\theta_{10}-I$	0,183	-0,359	0,080	0,064	-0,010	0,086	0,008	0,027	0,013	0,060	0,052		
r_{10}	0,10	0,28	0,044	0,035	0,019	0,053	0,013	0,014	0,019	0,036	0,036		

В табл. 3 приведены значения θ_{10} и ошибки $\theta_{10} - I$, полученные при расчете всех этих оценок с помощью случайных чисел γ_i , выписанных на стр. 108. (В качестве γ'_i и γ''_i , необходимых для расчета оценок 2 и 9, выбирались дальнейшие группы цифр из табл. 4 (стр. 295), умноженные на 10^{-5} .) Здесь же указаны вероятные ошибки r_{10} . Легко видеть, что фактические ошибки по порядку хорошо согласуются с вероятными ошибками, хотя количество слагаемых во всех оценках слишком мало для того, чтобы можно было гарантировать применимость центральной предельной теоремы.

Упражнения к главе 4

1. Пусть G — конечная область в n -мерном пространстве. Предположим, что дано разбиение $G = G_1 + \dots + G_N$ и каждая из областей G_j центрально симметрична с центром S_j . Пусть $Q^{(j)}$ — случайная точка, равномерно распределенная в G_j , а $Q^{(j) '}$ симметрична $Q^{(j)}$ (относительно центра S_j): $Q^{(j)} + Q^{(j) '}$ — симметрична $Q^{(j)}$. В качестве оценки для интеграла $I_0 = \int_G f(P) dP$ рассмотрим симметризованную формулу

$$\hat{\theta}_N^{(1)} = \sum_{j=1}^N \frac{V_j}{2} [f(Q^{(j)}) + f(Q^{(j) '})],$$

где V_j — объем области G_j .

Доказать, что если $f(P)$ и все ее частные производные первого и второго порядков непрерывны в G , причем для всех $k, l = 1, 2, \dots, n$

$$|\partial^2 f / \partial x_k \partial x_l| \leq L,$$

и выполнено условие (11), то для дисперсии $\hat{\theta}_N^{(1)}$ справедливо неравенство

$$D\hat{\theta}_n^{(1)} \leq c^2 L^2 N^{-1-4/n},$$

где $c = 0,5n^2 c_2^{n+2}$.

При этих же условиях оценить погрешность квадратурной формулы

$$\left| I_0 - \sum_{j=1}^N V_j f(S_j) \right| \leq c L N^{-2/n}.$$

У к а з а н и е. Доказательства аналогичны пп. 1.2 и 1.3.

(Для случая $G = K^n$ и разбиения, изображенного на рис. 48, это предложение доказано Н. С. Бахваловым [1]; см. также С. Хабер [128]).

2. Доказать, что математическое ожидание оценки

$$\check{\theta}_N = \frac{1}{M^n} \sum_{i_1, \dots, i_n=0}^{M-1} f\left(\frac{i_1 + \gamma_1}{M}, \dots, \frac{i_n + \gamma_n}{M}\right),$$

все узлы которой зависят от одной случайной точки $(\gamma_1, \dots, \gamma_n)$, равномерно распределенной в единичном кубе K^n , равно

$$M\check{\theta}_N = \int_{K^n} f(P) dP.$$

(Б. Л. Грановский)

3. Доказать, что величина $\sqrt{N}(\bar{\theta}_N - I_0)$ из п. 3.1 асимптотически нормальна, и вычислить главный член (37) дисперсии $D\bar{\theta}_N$.

ГЛАВА 5

РЕШЕНИЕ ЛИНЕЙНЫХ УРАВНЕНИЙ

§ 1. Интегральные преобразования

1.1. Итерированные функции. Рассмотрим функцию $\varphi(P)$, определенную в некоторой области G на плоскости $P = (x, y)$, и функцию $K(P, P')$, определенную при $P \in G, P' \in G$. Интегральное преобразование

$$K\varphi(P) = \int_G K(P, P') \varphi(P') dP' \quad (1)$$

преобразует функцию $\varphi(P)$ в функцию $K\varphi(P)$, которую называют *итерацией* функции $\varphi(P)$ с помощью ядра $K(P, P')$. *Второй итерацией* функции $\varphi(P)$ (с помощью ядра $K(P, P')$) называется функция $KK\varphi(P)$, которая обозначается $K^2\varphi(P)$.

Очевидно,

$$K^2\varphi(P) = \int_G \int_G K(P, P') K(P', P'') \varphi(P'') dP' dP''.$$

Точно так же определяются $K^3\varphi(P), \dots, K^i\varphi(P), \dots$

Вычислять такие интегралы можно методами, указанными в гл. 3 и 4. Однако задачи, в которых приходится вычислять итерации функций, имеют свою специфику: обычно требуется не одна какая-нибудь итерация, а несколько или даже все итерации до некоторого порядка. Поэтому вычислительные схемы Монте-Карло выбирают так, чтобы все эти итерации вычислялись одновременно по одним и тем же случайным испытаниям.

Далее, многие методы приближенного решения интегральных уравнений используют не сами значения $K^i\varphi(P)$,

а некоторые функционалы от $K^i\varphi(P)$, чаще всего — линейные, представимые в форме скалярных произведений.

Условимся записывать скалярное произведение функций $\varphi(P)$ и $\psi(P)$ в виде

$$(\varphi, \psi) = \int_G \varphi(P) \psi(P) dP. \quad (2)$$

В следующем пункте мы рассмотрим задачу о вычислении интегралов вида $(\psi, K^i\varphi)$. Заметим, что если область G n -мерная, то интеграл $(\psi, K^i\varphi)$ представляет собой $n(i+1)$ -кратный интеграл.

В дальнейшем мы будем предполагать, что

$$\varphi(P) \in L_2(G), \quad \psi(P) \in L_2(G), \quad K(P, P') \in L_2(G \times G).$$

Эта запись означает, что

$$\int_G \varphi^2 dP < \infty, \quad \int_G \psi^2 dP < \infty, \quad \int_G \int_G K^2 dP dP' < \infty.$$

Легко доказать, что если $\varphi(P) \in L_2(G)$ и $\psi(P) \in L_2(G)$, то скалярное произведение (2) конечно. Это вытекает из неравенства (1) на стр. 292:

$$\left| \int_G \varphi \psi dP \right| \leq \int_G |\varphi \psi| dP \leq \left\{ \int_G \varphi^2 dP \cdot \int_G \psi^2 dP \right\}^{1/2} < \infty.$$

Так же легко доказать, что если $\varphi(P) \in L_2(G)$ и $K(P, P') \in L_2(G \times G)$, то $K\varphi(P) \in L_2(G)$. В самом деле, из (1) следует, что

$$|K\varphi(P)|^2 \leq \left\{ \int_G |K \cdot \varphi| dP' \right\}^2 \leq \int_G K^2(P, P') dP' \int_G \varphi^2(P') dP'.$$

Интегрируя это неравенство по P , получим

$$\int_G |K\varphi|^2 dP \leq \int_G \int_G K^2(P, P') dP' dP \int_G \varphi^2(P') dP' < \infty.$$

Отсюда вытекает, что и $K^2\varphi(P), \dots, K^i\varphi(P), \dots$ принадлежат $L_2(G)$.

Пример. Область G представляет собой треугольник $x > 0, y > 0, x + y < 1$ (рис. 49), ядро $K(x, y, x', y') = xx' + yy'$. Если выбрать

$\varphi(x, y) = 1$, то

$$K\varphi(x, y) = \int_0^1 dx' \int_0^{1-x'} dy' (xx' + yy') = \frac{1}{6}(x + y),$$

$$K^2\varphi(x, y) = \int_0^1 dx' \int_0^{1-x'} dy' (xx' + yy') \frac{1}{6}(x' + y') = \frac{1}{6 \cdot 8}(x + y);$$

нетрудно проверить, что $K^i\varphi = 6^{-1} 8^{1-i}(x + y)$ при $i=1, 2, \dots$

1.2. Вычисление линейных функционалов от итерированных функций. Предположим, что в области G заданы функции $\varphi(P) \in L_2(G)$, $\psi(P) \in L_2(G)$, $K(P, P') \in L_2(G \times G)$ и требуется вычислить интегралы $(\psi, K^j\varphi)$ для $j=0, 1, \dots, i$.

Выберем в G произвольную плотность вероятностей $p(P)$, допустимую по отношению к $\psi(P)$ (см. гл. 3, п. 3.2), и произвольную условную плотность вероятностей $p(P, P') = p(P'|P)$, допустимую по отношению к ядру $\Lambda(P, P')$. Условия нормировки

$$\int_G p(P) dP = 1, \quad \int_G p(P, P') dP' = 1.$$

Определим в G случайную траекторию T_i (рис. 50)

$$T_i = (Q_0 \rightarrow Q_1 \rightarrow \dots \rightarrow Q_i),$$

где точка Q_0 имеет плотность $p(P)$, а плотность точки

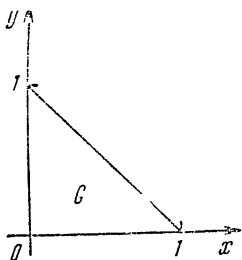


Рис. 49.

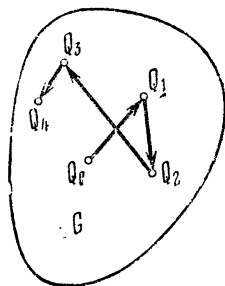


Рис. 50.

Q_i при известном значении Q_{i-1} равна $p(Q_{i-1}, P)$. Функцию $p(P, P')$ часто называют *плотностью вероятностей перехода* из точки P в точку P' и обозначают $p(P \rightarrow P')$,

а функцию $p(P)$ называют *начальной плотностью*. Траекторию можно интерпретировать как точку в $G \times \dots \times G = G^{i+1}$. Плотность вероятностей этой точки равна

$$p_i(Q_0, \dots, Q_i) = p(Q_0)p(Q_0, Q_1) \dots p(Q_{i-1}, Q_i). \quad (3)$$

Функцию (3) называют *плотностью траектории* T_i .

Введем в рассмотрение функции от траектории, называемые обычно *весами*:

$$W_i = \frac{K(Q_0, Q_1) K(Q_1, Q_2) \dots K(Q_{i-1}, Q_i)}{p(Q_0, Q_1) p(Q_1, Q_2) \dots p(Q_{i-1}, Q_i)}, \quad (4)$$

которые определены при $j=1, 2, \dots, i$. Для $j=0$ положим $W_0=1$; тогда справедлива рекуррентная формула

$$W_j = W_{j-1} [K(Q_{j-1}, Q_j) / p(Q_{j-1}, Q_j)], \quad (5)$$

которая позволяет последовательно вычислять все W_j по мере расчета траектории.

Обозначим через $\theta_i[\psi]$ случайную величину

$$\theta_i[\psi] = [\psi(Q_0) / p(Q_0)] W_i \varphi(Q_i). \quad (6)$$

Теорема 1. Если условия, перечисленные в начале пункта, выполнены, то математическое ожидание величины $\theta_i[\psi]$ равно $(\psi, K^i \varphi)$

$$M\theta_i[\psi] = (\psi, K^i \varphi). \quad (7)$$

Для доказательства теоремы вычислим это математическое ожидание:

$$M\theta_i[\psi] = \int_G \dots \int_G \theta_i[\psi] p_i dQ_0 \dots dQ_i.$$

Приняв во внимание соотношения (3), (4) и (6), получим *)

$$\begin{aligned} M\theta_i[\psi] &= \int_G \frac{\psi(Q_0)}{p(Q_0)} dQ_0 \int_G \dots \int_G W_i p_i \varphi(Q_i) dQ_1 \dots dQ_i = \\ &= \int_G \psi(Q_0) dQ_0 \int_G \dots \int_G K(Q_0, Q_1) \dots K(Q_{i-1}, Q_i) \varphi(Q_i) \times \\ &\quad \times dQ_1 \dots dQ_i = \int_G \psi(Q_0) dQ_0 K^i \varphi(Q_0) = (\psi, K^i \varphi). \end{aligned}$$

*) В этой главе переменные интегрирования иногда обозначаются той же буквой Q_i , что и случайная точка.

Доказательство теоремы нуждается в одном уточнении. Плотности $p(P)$ и (или) $p(P, P')$ могут обращаться в нуль в тех точках, в которых $\psi(P)$ и (или) $K(P, P')$ обращаются в нуль. В таких точках формулы (4), (5) и (6) не определены. Можно доопределить в этих точках W_j и $\theta_i[\psi]$, полагая их равными нулю. На значение интеграла это не повлияет, ибо если произведение $p(Q_0)p(Q_0, Q_1) \dots p(Q_{i-1}, Q_i)$ равно нулю в некоторой области $B \subset G \times \dots \times G$, то произведение $\psi(Q_0)K(Q_0, Q_1) \dots K(Q_{i-1}, Q_i)$ также равно нулю в B и интеграл по области B равен нулю. В дальнейшем мы оговаривать доопределение соответствующих величин не будем, так как это всегда может быть осуществлено таким же образом.

Теорема 1 позволяет построить метод Монте-Карло для расчета величин $(\psi, K'\psi)$. В самом деле, если по указанным правилам построить N траекторий вида T_i и для каждой из них вычислить $\theta_i[\psi]$, то при достаточно большом N

$$(\psi, K'\psi) \approx \frac{1}{N} \sum_{s=1}^N \theta_s[\psi], \quad (8)$$

где $\theta_i[\psi]$, — значение $\theta_i[\psi]$, сосчитанное для s -й траектории. Так как часть $(Q_0 \rightarrow Q_1 \rightarrow \dots \rightarrow Q_j)$ траектории T_i при $j < i$ представляет собой траекторию того же типа T_j , то по тем же точкам можно вычислить $\theta_j[\psi]$, и оценить $(\psi, K'\psi)$ для $j < i$.

Если требуется вычислить несколько скалярных произведений $(\psi, K'\psi)$ с различными функциями ψ и ψ , то это можно осуществить с помощью одних и тех же траекторий T_i , так как закон построения траекторий зависит только от $p(P)$ и $p(P, P')$, но не от ψ и ψ ; и если выбрать положительную плотность $p(P)$, то она будет допустимой по отношению к любой $\psi(P)$.

1.3. Вычисление значений итерированных функций. Мы укажем три способа оценки $K'\psi(P)$.

1.3.1. Траектории с фиксированной начальной точкой. Формально, для того чтобы вычислить значение $K'\psi(P)$ в точке $P = P_0$, можно выбрать $p(P) = \psi(P) = \delta(P - P_0)$, где $\delta(P)$ — дельта-функция Дирака*). Тогда из (7) вытекает, что

$$M\theta_i[\delta(P - P_0)] = (\delta(P - P_0), K'\psi) = K'\psi(P_0).$$

*) Хотя дельта-функция и не принадлежит $L_2(G)$, но скалярное произведение $(\delta, K^i\psi)$ конечно.

Выбор в качестве начальной плотности дельта-функции означает, что начальная точка фиксирована: $Q_0 = P_0$. Со-
кратив в (6) $\psi(Q_0)$ с $\rho(Q_0)$, можно переписать (7) в виде

$$M\{W_i\varphi(Q_i) | Q_0 = P_0\} = K^i\varphi(P_0). \quad (9)$$

Итак, если все N траекторий T_i начинать с фиксированной точки $Q_0 = P_0$, то

$$K^i\varphi(P_0) \approx \frac{1}{N} \sum_{s=1}^N \{W_i\varphi(Q_i) | Q_0 = P_0\}_s, \quad (10)$$

где индекс s имеет тот же смысл, что в (8). Конечно, формулу (9) можно вывести не прибегая к помощи дельта-функции, а рассматривая лишь траектории, начинающиеся из точки P_0 и их плотность.

Изложенный метод, очевидно, неудобен, если нужны значения $K^i\varphi(P)$ во многих точках, ибо из каждой такой точки пришлось бы строить «свои» N траекторий. Следующие два способа позволяют использовать одни и те же траектории для оценки $K^i\varphi(P)$ во многих точках.

1.3.2. Средние значения по области. Обозначим через $\chi_B(P)$ индикатор произвольной области $B \subset G$:

$$\chi_B(P) = \begin{cases} 1, & \text{если } P \in B, \\ 0, & \text{если } P \notin B. \end{cases}$$

Пусть $\rho(P)$ — произвольная «весовая» функция. Если в формуле (7) положить $\psi = \rho\chi_B$, то

$$M0_i[\rho\chi_B] = (\rho\chi_B, K^i\varphi) = \int_B \rho(P) K^i\varphi(P) dP. \quad (11)$$

Формула (11) позволяет по N траекториям вида T_i оценить среднее значение $K^i\varphi(P)$ с весом $\rho(P)$ по любой области B :

$$\overline{(K^i\varphi)_B} = \frac{\int_B \rho(P) K^i\varphi(P) dP}{\int_B \rho(P) dP}.$$

так как стоящий в числителе интеграл

$$\int_B \rho(P) K^i\varphi(P) dP \approx \frac{1}{N} \sum_{s=1}^N \theta_i[\rho\chi_B], \quad (12)$$

Наличие справа величины χ_B имеет простой смысл: суммируются $\theta_i[\rho]$, но только по тем траекториям, для которых начальная точка $Q_0 \in B$.

Конечно, надо иметь в виду, что если область B очень мала, то в (12) окажется очень мало слагаемых, отличных от нуля, и точность такого метода расчета будет невысокой.

Обозначим интересующую нас функцию через $u = K^i \varphi(P)$. Метод настоящего пункта позволяет вместо значений $u(P_1), \dots, u(P_k)$ в заданных точках P_1, \dots, P_k вычислить средние значения $\bar{u}_{B_1}, \dots, \bar{u}_{B_k}$ по некоторым областям B_1, \dots, B_k , содержащим соответствующие точки P_1, \dots, P_k . Оценить погрешность такого приближения нетрудно, если по полученным значениям $\bar{u}_{B_1}, \dots, \bar{u}_{B_k}$ можно составить себе представление о поверхности $u(P)$. Например, пусть вместо значения $u(x_0)$ мы вычислили среднее значение

$$\bar{u}_0 = \frac{\int_{x_0-h}^{x_0+h} u(x) x dx}{\int_{x_0-h}^{x_0+h} x dx}.$$

Если $u(x) \approx v(x) = a + bx^2$, то ошибка $\bar{u}_0 - u(x_0)$ приближенно равна

$$\bar{u}_0 - v(x_0) = \frac{b}{2} \frac{(x_0 + h)^4 - (x_0 - h)^4}{(x_0 + h)^2 - (x_0 - h)^2} - bx_0^2 = bh^2.$$

1.3.3. Оценка коэффициентов Фурье. Выберем в качестве $\psi(P)$ в (7) несколько ортонормированных функций $\psi_1(P), \dots, \psi_m(P)$, так что $(\psi_k, \psi_j) = \delta_{kj}$. Здесь δ_{kj} — символ Кронекера: $\delta_{kj} = 0$ при $k \neq j$ и $\delta_{kk} = 1$. Формула (7) позволяет оценить коэффициенты Фурье $c_{i,k}$ функции $K^i \varphi(P)$ относительно $\psi_k(P)$:

$$c_{i,k} = (K^i \varphi, \psi_k) = M \theta_i[\psi_k].$$

Если система функций ψ_1, \dots, ψ_m выбрана достаточно разумным образом, то

$$K^i \varphi(P) \approx \sum_{k=1}^m c_{i,k} \psi_k(P), \quad (13)$$

и, вычислив приближенные значения

$$c_{i,k} \approx \frac{1}{N} \sum_{s=1}^N \theta_i[\psi_k]_s, \quad (14)$$

получим приближение (13) для $K^i \varphi(P)$ во всей области.

1.4. Случайные траектории с поглощением. Траектории типа T_i — далеко не единственный способ вычисления итераций методом Монте-Карло. Зададим произвольную

функцию $a(P)$, удовлетворяющую в G условию

$$0 < a(P) < 1.$$

Пусть $s(P) = 1 - a(P)$, а плотности $p(P)$ и $p(P, P')$ — такие же, как в п. 1.2.

Определим в G траектории T_v случайной длины v

$$T_v = (Q_0 \rightarrow Q_1 \rightarrow \dots \rightarrow Q_v)$$

по следующим правилам:

а) точка Q_0 выбирается в соответствии с плотностью $p(P)$.

б) в точке Q_j (где $j=0, 1, \dots$) траектория с вероятностью $a(Q_j)$ заканчивается и с вероятностью $s(Q_j)$ продолжается;

в) если траектория не заканчивается, то точка Q_{j+1} выбирается в соответствии с плотностью $p(Q_j, P)$.

Естественно назвать $a(P)$ *вероятностью поглощения* случайной точки в точке P и $s(P)$ — *вероятностью рас-сеяния* этой точки.

Вероятность получить i -звенную траекторию T_v с соответствующими вершинами, расположенными в окрестностях точек P_0, P_1, \dots, P_i , равна произведению

$$p(P_0) dP_0 s(P_0) \prod_{j=1}^{i-1} p(P_{j-1}, P_j) dP_j s(P_j) p(P_{i-1}, P_i) dP_i a(P_i).$$

Введем обозначение

$$\tilde{p}_i(P_0, \dots, P_i) = p(P_0) \prod_{j=1}^i s(P_{j-1}) p(P_{j-1}, P_j) a(P_i); \quad (15)$$

тогда вероятность получить i -звенную траекторию T_v равна

$$P\{v=i\} = \int_G \dots \int_G \tilde{p}_i(P_0, \dots, P_i) dP_0 \dots dP_i.$$

Запишем также *условную плотность вероятностей траектории* T_v при условии, что $v=i$:

$$\tilde{p}_v(P_0, \dots, P_i | v=i) = \tilde{p}_i(P_0, \dots, P_i) / P\{v=i\}.$$

Вдоль траектории T_v рассмотрим функции \tilde{W}_i (веса), определенные (для $j \leq v$) рекуррентной формулой

$$\tilde{W}_0 = 1. \quad \tilde{W}_j = \tilde{W}_{j-1} [K(Q_{j-1}, Q_j) / s(Q_{j-1}) p(Q_{j-1}, Q_j)]. \quad (16)$$

Сравнивая (16) с (5), легко заметить, что

$$W_j(P_0, \dots, P_j) = \tilde{W}_j(P_0, \dots, P_j) s(P_0) \dots s(P_{j-1}). \quad (17)$$

Обозначим через $\tilde{\theta}_i[\psi]$ случайную величину

$$\tilde{\theta}_i[\psi] = [\psi(Q_0)/p(Q_0)] \tilde{W}_i[\varphi(Q_i)/a(Q_i)]. \quad (18)$$

Теорема 2. Если условия, перечисленные в начале п. 1.2, выполнены, то условное математическое ожидание величины $\tilde{\theta}_i[\psi]$ равно

$$M\{\tilde{\theta}_i[\psi] | v = i\} = (\psi, K^i \varphi) / P\{v = i\}. \quad (19)$$

Доказательство. По определению

$$\begin{aligned} M\{\tilde{\theta}_i[\psi] | v = i\} &= \int_G \dots \int_G \tilde{\theta}_i[\psi] \tilde{p}_v(Q_0, \dots, Q_i | v = i) dQ_0 \dots \\ &\dots dQ_i = \int_G \dots \int_G \psi(Q_0) \prod_{j=1}^i K(Q_{j-1}, Q_j) \varphi(Q_i) \frac{dQ_0 \dots dQ_i}{P\{v = i\}} = \\ &= \frac{(\psi, K^i \varphi)}{P\{v = i\}}. \end{aligned}$$

Теорема 2 позволяет сформулировать метод Монте-Карло для расчета $(\psi, K^i \varphi)$ с помощью траекторий \tilde{T}_v . В самом деле, реализовав N таких траекторий, получим среди них N_i траекторий, состоящих из i звеньев. Очевидно,

$$N_i/N \approx P\{v = i\}.$$

А из (19) вытекает, что

$$(\psi, K^i \varphi) \approx \frac{P\{v = i\}}{N_i} \sum_{s=1}^{N_i} \tilde{\theta}_i[\psi]_s, \quad (20)$$

где $\tilde{\theta}_i[\psi]_s$ — значение $\tilde{\theta}_i[\psi]$ на s -й траектории (из числа i -звенных траекторий). Заметим, что последнюю формулу можно заменить формулой

$$(\psi, K^i \varphi) \approx \frac{1}{N} \sum_{s=1}^{N_i} \tilde{\theta}_i[\psi]_s. \quad (21)$$

1.5. Сравнение точности оценок (8) и (20). В оценке (8) осредняются значения случайной величины $\theta_i[\psi]$, математическое ожидание которой равно $(\psi, K^i \varphi)$. В (20) осредняются значения случайной

величины $\xi_i = P\{v=i\} \tilde{\theta}_i[\psi]$, условное математическое ожидание которой равно также $(\psi, K^i \psi)$. Для того чтобы сравнить дисперсии этих величин, достаточно сравнить математические ожидания их квадратов. В первом случае

$$M \theta_i^2[\psi] = \int_G \dots \int_G \theta_i^2[\psi] p_i(Q_0, \dots, Q_i) dQ_0 \dots dQ_i, \quad (22)$$

а во втором

$$\begin{aligned} M\{\xi_i^2 | v=i\} &= \int_G \dots \int_G \xi_i^2 \tilde{p}_v(Q_0, \dots, Q_i | v=i) dQ_0 \dots dQ_i = \\ &= P\{v=i\} \int_G \dots \int_G \tilde{\theta}_i^2[\psi] \tilde{p}_i(Q_0, \dots, Q_i) dQ_0 \dots dQ_i. \end{aligned}$$

С помощью формул (3), (6) и (15), (18), приняв во внимание соотношение (17), нетрудно проверить, что имеет место тождество

$$\theta_i^2[\psi] p_i = \tilde{\theta}_i^2[\psi] \tilde{p}_i s(Q_0) \dots s(Q_{i-1}) a(Q_i). \quad (23)$$

Используем (23), чтобы преобразовать последнее выражение для $M\{\xi_i^2 | v=i\}$:

$$M\{\xi_i^2 | v=i\} = \int_G \dots \int_G \theta_i^2[\psi] p_i \frac{P\{v=i\} dQ_0 \dots dQ_i}{s(Q_0) \dots s(Q_{i-1}) a(Q_i)}. \quad (24)$$

Вообще говоря, выражение (24) может быть как меньше, так и больше, чем (22). Однако в весьма важном частном случае, когда поглощение постоянно во всей области G или, другими словами, $a(P) \equiv a$, эти выражения равны. В самом деле, в этом случае

$$P\{v=i\} = a s^i \int_G \dots \int_G p_i(P_0, \dots, P_i) dP_0 \dots dP_i = a s^i,$$

и дробь, стоящая в (24), равна $dQ_0 \dots dQ_i$.

Следовательно, в случае $a(P) \equiv a$ дисперсии равны

$$D \theta_i[\psi] = D\{\xi_i | v=i\}.$$

Для того чтобы количество слагаемых в формулах (8) и (20) было одинаковым и равнялось N^* , в первом случае надо построить $N = N^*$ траекторий типа T_i , а во втором случае, чтобы $N_i = N^*$, надо построить в среднем $N^*/P\{v=i\} > N^*$ траекторий типа \tilde{T}_v . Таким образом, оценка (20) для $(\psi, K^i \psi)$ оказывается менее выгодной.

На различных других оценках величины $(\psi, K^i \psi)$ мы останавливаться не будем. Кроме случайных траекторий типа T_i и \tilde{T}_v , возможны и другие типы траекторий. Например, «поглощение» может происходить при пересечении некоторых заранее заданных линий или зависеть только от направления звена траектории и т. п.

§ 2. Неоднородные интегральные уравнения

2.1. Постановка задачи. Ряд Неймана. Рассмотрим интегральное уравнение

$$z(P) = \int_G K(P, P') z(P') dP' + f(P),$$

которое с учетом (1) можно записать в виде

$$z = Kz + f. \quad (25)$$

Здесь $z(P)$ — искомая функция, $K(P, P')$ — заданная функция, называемая *ядром уравнения*, $f(P)$ — заданная функция (*свободный член*).

Предположим, что $f(P) \in L_2(G)$, $K(P, P') \in L_2(G \times G)$. Можно попытаться искать решение методом последовательных приближений. Пусть $z^{(0)} = \varphi(P)$ — произвольная функция из $L_2(G)$. Пусть далее при $i = 1, 2, \dots$

$$z^{(i)} = Kz^{(i-1)} + f. \quad (26)$$

Нетрудно вычислить, что

$$z^{(i)} = f + Kf + \dots + K^{(i-1)}f + K^i\varphi. \quad (27)$$

Последовательные приближения $z^{(i)}$ сходятся при $i \rightarrow \infty$ к решению z уравнения (25) тогда и только тогда, когда это решение представимо в виде ряда Неймана

$$z = \sum_{i=0}^{\infty} K^i f. \quad (28)$$

Условия сходимости ряда Неймана имеются, например, в [63].

В частности, если

$$\iint_G K^2(P, P') dP dP' < 1,$$

то ряд (28) сходится в среднем:

$$\lim_{m \rightarrow \infty} \left(z - \sum_{i=0}^m K^i f, z - \sum_{i=0}^m K^i f \right) = 0.$$

Если, кроме того, для всех $P \in B \subseteq G$

$$\int_G K^2(P, P') dP' \leq C,$$

то ряд (28) сходится абсолютно и равномерно в B .

Лемма. Если ряд (28) сходится в среднем, то для любой функции $\psi(P)$ из $L_2(G)$

$$\lim_{m \rightarrow \infty} \left(\psi, \sum_{i=0}^m K^i f \right) = (\psi, z).$$

Доказательство. Рассмотрим разность

$$(\psi, z) - \left(\psi, \sum_{i=0}^m K^i f \right) = \left(\psi, z - \sum_{i=0}^m K^i f \right).$$

Из неравенства (1) на стр. 292 следует, что

$$\left(\psi, z - \sum_{i=0}^m K^i f \right)^2 \leq (\psi, \psi) \left(z - \sum_{i=0}^m K^i f, z - \sum_{i=0}^m K^i f \right) \rightarrow 0,$$

когда $m \rightarrow \infty$.

Формулы (27) и (28) показывают, что вычисление $z^{(i)}$ или z сводится к вычислению суммы итерированных функций.

2.2. Оценка линейных функционалов от $Z^{(i)}$. Результаты § 1 позволяют указать методы Монте-Карло для вычисления скалярных произведений $(\psi, z^{(i)})$, где $\psi(P)$ — любая заданная функция из $L_2(G)$.

В самом деле, рассмотрим случайные траектории T_i , определение которых дано в п. 1.2, и введем случайные величины $\xi_i[\psi]$, зависящие от таких траекторий:

$$\xi_i[\psi] = \frac{\psi(Q_0)}{p(Q_0)} \left[\sum_{j=0}^{i-1} W_{jf}(Q_j) + W_{i\varphi}(Q_i) \right]. \quad (29)$$

(При $i=0$ полагать квадратную скобку в (29) равной $W_{0\varphi}(Q_0)$.)

Легко доказать, что математическое ожидание этих величин равно

$$M \xi_i[\psi] = (\varphi, z^{(i)}), \quad (30)$$

ибо каждое слагаемое в (29) — это величина типа $\theta_i[\psi]$:

$$\begin{aligned} M \xi_i[\psi] &= \sum_{j=0}^{i-1} M \left\{ \frac{\psi(Q_0)}{p(Q_0)} W_{jf}(Q_j) \right\} + M \left\{ \frac{\psi(Q_0)}{p(Q_0)} W_{i\varphi}(Q_i) \right\} = \\ &= \sum_{j=0}^{i-1} (\psi, K^j f) + (\psi, K^i \varphi) = (\psi, z^{(i)}). \end{aligned}$$

Соответствующий (30) метод Монте-Карло:

$$(\psi, z^{(i)}) \approx \frac{1}{N} \sum_{s=1}^N \xi_i[\psi]_s, \quad (31)$$

где $\xi_i[\psi]_s$ — значение $\xi_i[\psi]$ на s -й траектории.

Очевидно, по одним и тем же траекториям типа T_i можно вычислять значения $(\psi, z^{(i)})$ для нескольких функций $\psi(P)$ и даже для нескольких уравнений (25) — с одним и тем же ядром $K(P, P')$, но с разными $f(P)$. Сами значения $z^{(i)}(P)$ можно вычислять любым из трех методов, указанных в п. 1.3. Если ряд Неймана (28) сходится, то при достаточно больших i значения $(\psi, z^{(i)})$ могут служить приближениями к (ψ, z) .

2.3. Метод существенной выборки для траекторий T_i . В предыдущем пункте рассматривались траектории T_i , построенные по произвольным допустимым $p(P)$ и $p(P')$, но не было никаких указаний на то, как лучше выбирать эти плотности. Однако от их выбора зависят дисперсии $D\xi_i[\psi]$, определяющие точность метода Монте-Карло (31). Особенно нежелательна большая дисперсия тогда, когда начальное приближение $z^{(0)} = \varphi(P)$ близко к точному решению $z(P)$, так как можно «испортить» это приближение.

Если функция $z^{(i)}(P)$ известна, и интеграл $(\psi, z^{(i)})$ вычисляется с помощью существенной выборки (гл. 3, п. 3.2), то минимальная дисперсия такой оценки, согласно теореме 3 гл. 3, равна

$$\hat{D}_i = (|\psi|, |z^{(i)}|)^2 - (\psi, z^{(i)})^2. \quad (32)$$

Теорема 3. *Предположим, что ядро уравнения (25) и его решение неотрицательны*

$$K(P, P') \geq 0, \quad z(P) \geq 0,$$

а траектории T_i строятся с начальной плотностью

$$p(P) = |\psi(P)| \varphi(P) / (|\psi|, \varphi) \quad (33)$$

и с плотностью вероятностей перехода

$$p(P, P') = K(P, P') \varphi(P') / K \varphi(P). \quad (34)$$

Если начальное приближение $\varphi(P)$ равно решению $z(P)$, то при любом $j=0, 1, 2, \dots, i$

$$D\xi_j[\psi] = \hat{D}_j. \quad (35)$$

Доказательство. Если $\varphi = K\varphi + f$ и выполнено равенство (34), то

$$\begin{aligned}\xi_1[\psi] &= [\psi(Q_0)/p(Q_0)] [f(Q_0) + K(Q_0, Q_1)\varphi(Q_1)/p(Q_0, Q_1)] = \\ &= [\psi(Q_0)/p(Q_0)] [f(Q_0) + K\varphi(Q_0)] = \\ &= \varphi(Q_0)\psi(Q_0)/p(Q_0) = \xi_0[\psi].\end{aligned}$$

В то же время при $j \geq 2$ из (29) следует, что

$$\begin{aligned}\xi_{j-1}[\psi] - \xi_j[\psi] &= \frac{\psi(Q_0)}{p(Q_0)} [W_{j-1}\varphi(Q_{j-1}) - W_{j-1}f(Q_{j-1}) - \\ &- W_j\varphi(Q_j)] = [\psi(Q_0)/p(Q_0)] W_{j-1} [\varphi(Q_{j-1}) - f(Q_{j-1}) - \\ &- K\varphi(Q_{j-1})] = 0.\end{aligned}$$

Следовательно, при любом j значение $\xi_j[\psi]$ зависит лишь от Q_0 и, с учетом (33), равно $\xi_j[\psi] = \operatorname{sgn} \psi(Q_0) \times \times (|\psi|, \varphi)$. Поэтому дисперсия $\xi_j[\psi]$ равна

$$D\xi_j[\psi] = M\xi_j^2[\psi] - (\psi, z^{(j)})^2 = (|\psi|, \varphi)^2 - (\psi, z^{(j)})^2.$$

А так как из $\varphi(P) = z(P)$ следует, что все $z^{(j)}(P) = z(P)$, то последнее выражение для $D\xi_j[\psi]$ совпадает с (32). Теорема доказана.

Как и в гл. 3, п. 3.2.1, плотности (33) и (34) с $\varphi(P) = z(P)$ практически использовать нельзя, так как решение $z(P)$ неизвестно. Однако, имея «хорошее» приближение $z_0(P) \approx z(P)$, мы можем (в принципе) выбрать $\varphi(P) = z_0(P)$, и дисперсии $D\xi_j[\psi]$ будут близки к минимальным.

Если функция $\psi(P)$ знакопостоянна, то очевидно, $\hat{D}_j = 0$.

2.4. Оценка линейных функционалов от z . Рассмотрим бесконечные случайные траектории

$$T_\infty = (Q_0 \rightarrow Q_1 \rightarrow \dots \rightarrow Q_i \rightarrow \dots),$$

которые строятся по тем же правилам, что траектории T_i . Тогда любой начальный участок $(Q_0 \rightarrow \dots \rightarrow Q_i)$ этой траектории представляет собой траекторию типа T_i и плотность его $p_i(Q_0, \dots, Q_i)$ выражается формулой (3). Предположим, что начальная плотность $p(P)$ допустима по отношению к $\psi(P)$, а плотность вероятностей перехода $p(P, P')$ допустима по отношению к $K(P, P')$. Тогда из теоремы 1 следует, что при каждом i

$$M\{[\psi(Q_0)/p(Q_0)] W_i f(Q_i)\} = (\psi, K^i f).$$

Рассмотрим случайную величину $\xi[\psi]$, зависящую от траектории T_∞ :

$$\xi[\psi] = \frac{\psi(Q_0)}{p(Q_0)} \sum_{i=0}^{\infty} W_i f(Q_i). \quad (36)$$

При некоторых условиях справедливы равенства

$$\begin{aligned} M\xi[\psi] &= \sum_{i=0}^{\infty} M\left\{ \frac{\psi(Q_0)}{p(Q_0)} W_i f(Q_i) \right\} = \\ &= \sum_{i=0}^{\infty} (\psi, K^i f) = \left(\psi, \sum_{i=0}^{\infty} K^i f \right) = (\psi, z), \end{aligned}$$

и математическое ожидание величины $\xi[\psi]$ оказывается равным вычисляемому функционалу:

$$M\xi[\psi] = (\psi, z). \quad (37)$$

Например, достаточно потребовать, чтобы было $K(P, P') \geq 0$, $f(P) \geq 0$ и ряд Неймана (28) сходиллся в среднем. Более общие условия приведены ниже в теореме 4.

Если (37) справедливо, то для оценки (ψ, z) можно использовать N траекторий типа T_∞ , по каждой из них вычислить значение $\xi[\psi]$, (s — номер траектории) и осреднить результат

$$(\psi, z) \approx \frac{1}{N} \sum_{s=1}^N \xi[\psi]_s. \quad (33)$$

Конечно, построить бесконечную траекторию численно невозможно. На практике траекторию строят до тех пор, пока слагаемые в (36) не становятся пренебрежимо малыми по сравнению со всей суммой, и тогда траекторию «обрывают». Часто в качестве условия обрыва используют неравенство

$$W_i < \varepsilon,$$

где $\varepsilon > 0$ — некоторое наперед заданное малое число, а i — номер последней точки траектории.

Можно обойтись без искусственного обрыва траекторий, если вместо траекторий T_∞ использовать траектории с поглощением \tilde{T}_v (п. 1.4). В самом деле, рассмотрим случайную величину

$$\tilde{\xi}_v[\psi] = [\psi(Q_0)/p(Q_0)] \tilde{W}_v[f(Q_v)/u(Q_v)] \quad (39)$$

со случайным номером v , равным количеству звеньев траектории \tilde{T}_v .

При некоторых условиях математическое ожидание случайной величины $\tilde{\xi}_v[\psi]$, зависящей от случайного индекса v , можно вычислить по формуле

$$M\tilde{\xi}_v[\psi] = \sum_{i=0}^{\infty} M\{\tilde{\xi}_v[\psi] | v = i\} P\{v = i\},$$

аналогичной формуле полной вероятности. Так как при $v=i$ и $f=\varphi$ величина $\tilde{\xi}_v[\psi] = \theta_i[\psi]$, то из теоремы 2 вытекает, что

$$M\{\tilde{\xi}_v[\psi] | v = i\} = (\psi, K^i f) / P\{v = i\}.$$

Следовательно,

$$M\tilde{\xi}_v[\psi] = \sum_{i=0}^{\infty} (\psi, K^i f) = \left(\psi, \sum_{i=0}^{\infty} K^i f \right) = (\psi, z)$$

и окончательно

$$M\tilde{\xi}_v[\psi] = (\psi, z). \quad (40)$$

Формула (40) также будет справедлива в случае, когда $K(P, P') \geq 0$, $f(P) \geq 0$ и ряд Неймана (28) сходится в среднем. Более общие условия см. ниже в теореме 5.

Соответствующий формуле (40) метод Монте-Карло

$$(\psi, z) \approx \frac{1}{N} \sum_{s=1}^N \tilde{\xi}_v[\psi]_s, \quad (41)$$

где $\tilde{\xi}_v[\psi]_s$ — значение $\tilde{\xi}_v[\psi]$ на s -й траектории типа \tilde{T}_v , а v — номер последней точки этой траектории.

2.4.1. Метод Монте-Карло (31) позволяет оценить функционал (ψ, z) , если только ряд Неймана (28) сходится в среднем, так как можно фиксировать столь большое i , чтобы разность $|(\psi, z^{(i)}) - (\psi, z)|$ была меньше любого заданного $\varepsilon > 0$. Казалось бы, методы (38) и (41) также должны сходиться при этом единственном условии. К сожалению, переход в бесконечномерное пространство $G \times G \times \dots$ заставляет наложить несколько более жесткие ограничения, связанные с тем, что сходимость ряда $\sum_{i=0}^{\infty} M\eta_i$ не обеспечивает вообще

говоря, существования $M \sum_{i=0}^{\infty} \eta_i$ и приходится требовать, чтобы схо-

дился ряд $\sum_{i=0}^{\infty} M|\eta_i|$.

Интегральное уравнение

$$\bar{z}(P) = \int_G |K(P, P')| \bar{z}(P') dP' + |f(P)| \quad (\bar{25})$$

назовем *мажорантным* по отношению к уравнению (25). Так как $|K^i f| \leq |K|^i |f|$, то из сходимости (в среднем) ряда Неймана для мажорантного уравнения следует абсолютная сходимость (в среднем) ряда (28).

Теорема 4. Если ряд Неймана для мажорантного уравнения (25) сходится в среднем, то имеет место равенство (37).

Теорема 5. Если ряд Неймана для мажорантного уравнения (25) сходится в среднем, то имеет место равенство (40).

Легко видеть, что в случае, когда $K(P, P') \geq 0$ и $f(P) \geq 0$, мажорантное уравнение (25) совпадает с уравнением (25) и из теорем 4 и 5 вытекают достаточные условия, приведенные в п. 2.4.

Перейдем к доказательству теоремы 4. Из теоремы 1 следует, что

$$M\{|\psi(Q_0)/\rho(Q_0)| |W_i f(Q_i)|\} = (|\psi|, |K|^i |f|).$$

Сумма таких математических ожиданий равна

$$\sum_{i=0}^m M\left|\frac{\psi(Q_0)}{\rho(Q_0)} W_i f(Q_i)\right| = \sum_{i=0}^m (|\psi|, |K|^i |f|) = \left(|\psi|, \sum_{i=0}^m |K|^i |f|\right)$$

и (по лемме п. 2.1) стремится к конечному пределу, равному $(|\psi|, \bar{z})$. Сходимость этого ряда обеспечивает существование $M\zeta[\psi]$ и справедливость равенства

$$M\zeta[\psi] = \sum_{i=0}^{\infty} M\left\{\frac{\psi(Q_0)}{\rho(Q_0)} W_i f(Q_i)\right\} = \sum_{i=0}^{\infty} (\psi, K^i f).$$

Еще раз применяя лемму п. 2.1, получим, что

$$\sum_{i=0}^{\infty} (\psi, K^i f) = \lim_{m \rightarrow \infty} \sum_{i=0}^m (\psi, K^i f) = \lim_{m \rightarrow \infty} \left(\psi, \sum_{i=0}^m K^i f\right) = (\psi, z).$$

Аналогично доказывается и теорема 5. Из теоремы 2 вытекает, что

$$M\{|\tilde{\zeta}_v[\psi]| |v = i\} = M\left|\frac{\psi(Q_0)}{\rho(Q_0)} \tilde{W}_i \frac{f(Q_i)}{a(Q_i)}\right| = \frac{(|\psi|, |K|^i |f|)}{P\{v = i\}}.$$

Затем с помощью леммы п. 2.1 доказывается сходимость ряда

$$\sum_{i=0}^{\infty} M\{|\tilde{\zeta}_v[\psi]| |v = i\} P\{v = i\} = \sum_{i=0}^{\infty} (|\psi|, |K|^i |f|) = (|\psi|, \bar{z}),$$

которая обеспечивает существование $M\tilde{\zeta}_v[\psi]$ и справедливость

равенства

$$M\tilde{\zeta}_v[\psi] = \sum_{i=0}^{\infty} M\{\tilde{\zeta}_v[\psi] | v=i\} P\{v=i\} = \sum_{i=0}^{\infty} (\psi, K^i f).$$

То, что $\sum_{i=0}^{\infty} (\psi, K^i f) = (\psi, z)$, мы уже доказали.

Метод Монте-Карло для расчета (ψ, z) , справедливый и тогда, когда ряд Неймана для мажорантного уравнения расходится, рассмотрен в статье [148].

2.5. Сравнение точности оценок (38) и (41). Для того чтобы сравнить дисперсии $D\tilde{\zeta}[\psi]$ и $D\tilde{\zeta}_v[\psi]$, рассмотрим математические ожидания квадратов этих величин. Во-первых,

$$\begin{aligned} M\tilde{\zeta}_v^2[\psi] &= \sum_{i=0}^{\infty} M\{\tilde{\zeta}_v^2[\psi] | v=i\} P\{v=i\} = \\ &= \sum_{i=0}^{\infty} P\{v=i\} \int_G \dots \int_G \left[\frac{\psi(Q_0)}{p(Q_0)} \tilde{w}_i \frac{f(Q_i)}{a(Q_i)} \right]^2 \times \\ &\quad \times \tilde{p}_v(Q_0, \dots, Q_i | v=i) dQ_0 \dots dQ_i. \end{aligned}$$

Используя (17) и (15) для перехода от \tilde{w}_i и \tilde{p}_v к w_i и p_i и введя (для краткости) обозначение

$$\theta_i = [\psi(Q_0)/p(Q_0)] w_i f(Q_i), \quad (42)$$

запишем результат в виде

$$M\tilde{\zeta}_v^2[\psi] = \sum_{i=0}^{\infty} \int_G \dots \int_G \theta_i^2 \frac{p_i(Q_0, \dots, Q_i) dQ_0 \dots dQ_i}{s(Q_0) \dots s(Q_{i-1}) a(Q_i)}. \quad (43)$$

Во-вторых, из (36) и (42) видно, что

$$\zeta^2[\psi] = \left(\sum_{j=0}^{\infty} \theta_j \right)^2 \leq \sum_{i,j=0}^{\infty} |\theta_i| |\theta_j|.$$

Для того чтобы оценить это выражение, запишем его в виде

$$\zeta^2[\psi] \leq \sum_{i,j=0}^{\infty} t^{i+j} (|\theta_i| t^{-i}) (|\theta_j| t^{-j}),$$

где t — любое число из интервала $0 < t < 1$, и воспользуемся очевидным неравенством $2uv \leq u^2 + v^2$

$$\zeta^2[\psi] \leq \frac{1}{2} \sum_{i,j=0}^{\infty} t^{i+j} \left(\frac{\theta_i^2}{t^{2i}} + \frac{\theta_j^2}{t^{2j}} \right) = \sum_{i=0}^{\infty} \theta_i^2 t^{-i} \sum_{j=0}^{\infty} t^j.$$

Следовательно,

$$\xi^2[\psi] \leq \frac{1}{1-t} \sum_{i=0}^{\infty} 6_i^2 t^{-i},$$

и математическое ожидание этой величины не превосходит

$$M\xi^2[\psi] \leq \sum_{i=0}^{\infty} \frac{t^{-i}}{1-t} \int_G \dots \int_G 6_i^2 p_i(Q_0, \dots, Q_i) dQ_0 \dots dQ_i. \quad (44)$$

Теорема 6. Если траектории \tilde{T}_v строятся с постоянным поглощением (т. е. $a(P) \equiv a$ во всей области G), то

$$D\xi_v[\psi] \leq D\tilde{\xi}_v[\psi].$$

Для доказательства этой теоремы воспользуемся формулами (43) и (44). Так как $s(P) \equiv 1-a$, то из (43) вытекает, что

$$M\tilde{\xi}_v^2[\psi] = \sum_{i=0}^{\infty} \frac{(1-a)^{-i}}{a} \int_G \dots \int_G 6_i^2 p_i(Q_0, \dots, Q_i) dQ_0 \dots dQ_i.$$

Положим в (44) $t=1-a$. Тогда (44) превратится в неравенство $M\xi^2[\psi] \leq M\tilde{\xi}_v^2[\psi]$, а отсюда сразу вытекает, что $D\xi[\psi] \leq D\tilde{\xi}_v[\psi]$.

З а м е ч а н и е. Вообще говоря, ряды (43) и (44) не обязательно сходятся, и дисперсии могут быть бесконечными. Условия конечности дисперсий $D\xi[\psi]$ и $D\tilde{\xi}_v[\psi]$ приведены в упражнениях 2 и 3.

Теорема 6 показывает, что в конкретном случае постоянного поглощения оценка (38) для (ψ, z) обеспечивает лучшую точность, чем (41) (при одном и том же количестве траекторий). Впрочем, оценка (41) все-таки может оказаться менее трудосемкой, чем (38). Так будет, например, в случае, когда функция $f(P)$ очень сложна (по сравнению с формулами расчета траекторий), так как для расчета $\xi[\psi]$ необходимо вычислять $\tilde{f}(Q_j)$ во всех точках траектории T_{∞} , а для расчета $\tilde{\xi}_v[\psi]$ нужно лишь значение $f(Q_v)$ в последней точке траектории \tilde{T}_v .

2.6. Использование сопряженного уравнения. Нетрудно убедиться, что вместо того, чтобы вычислять функционал (ψ, z) от решения z уравнения

$$z = Kz + f, \quad (45)$$

можно решать сопряженную задачу: вычислять функционал (f, u) от решения u уравнения

$$u = K^*u + \psi, \quad (46)$$

где $K^*(P, P') = K(P', P)$. В самом деле, умножив скалярно (45) на u , а (46) на z , получим

$$(u, z) = (u, Kz) + (u, f), \quad (u, z) = (K^*u, z) + (\psi, z).$$

Так как *) $(K^*u, z) = (u, Kz)$, то из этих соотношений вытекает, что $(u, f) = (\psi, z)$.

Уравнения (45) и (46) мы будем называть *сопряженными*.

Если плотности $p(P)$ и $p(P, P')$, по которым строятся траектории T_∞ (п. 2.4), допустимы по отношению к $f(P)$ и $K^*(P, P')$ (соответственно), то, согласно п. 2.4, можно рассмотреть случайную величину

$$\xi^*[f] = \frac{f(Q_0)}{p(Q_0)} \sum_{j=0}^{\infty} W_j^* \psi(Q_j), \quad (47)$$

где веса $W_0^* = 1$, $W_j^* = W_{j-1}^* [K^*(Q_{j-1}, Q_j)/p(Q_{j-1}, Q_j)]$. Нетрудно доказать, что если условия теоремы 4 выполнены, то математическое ожидание $\xi^*[f]$ равно

$$M\xi^*[f] = (f, u) = (\psi, z).$$

Следовательно, при больших N

$$(\psi, z) \approx \frac{1}{N} \sum_{s=1}^N \xi^*[f]_s, \quad (48)$$

где $\xi^*[f]_s$ — это значение $\xi^*[f]$ на s -й траектории.

Интересно, что даже в случае уравнения (45) с симметричным ядром, когда $K^*(P, P') = K(P, P')$ и $W_j^* = W_j$ оценки (36) и (47) представляют собой различные оценки для функционала (ψ, z) :

$$\xi[\psi] = \frac{\psi(Q_0)}{p(Q_0)} \sum_{j=0}^{\infty} W_j f(Q_j), \quad \xi^*[f] = \frac{f(Q_0)}{p(Q_0)} \sum_{j=0}^{\infty} W_j \psi(Q_j).$$

Очевидно, различие в объеме работы, затрачиваемой на расчет этих оценок: для расчета $\xi[\psi]$ надо один раз вычислить $\psi(P)$ и много раз $f(P)$, а для расчета

*) $(K^*u, z) = \int K^*u(P) z(P) dP = \iint K^*(P, P') u(P') z(P) dP dP' = \iint K(P', P) z(P) u(P') dP dP' = \int Kz(P') u(P') dP' = (u, Kz)$.

Обычно ядро $K(P', P)$ называют *транспонированным*, а комплексно сопряженное с ним ядро $\overline{K(P', P)}$ — *сопряженным* (по отношению к $K(P, P')$). Однако для действительных ядер $\overline{K(P', P)} = K(P', P)$,

$\xi^*[f]$ — наоборот, один раз вычислить $f(P)$ и много раз $\psi(P)^*$.

Некоторые задачи сводятся к уравнению (45) с дельта-функцией: $f(P) = c_0 \delta(P - P_0)$ («источник» расположен в точке P_0 и имеет «интенсивность» c_0). В таких задачах величина

$$\xi[\psi] = c_0 [\psi(Q_0)/p(Q_0)] \sum_{j=0}^{\infty} W_j \delta(Q_j - P_0)$$

для расчета практически бесполезна. В то же время величину $\xi^*[c_0 \delta(P - P_0)]$ можно с успехом использовать, если положить $p(P) = \delta(P - P_0)$, т. е. (ср. п. 1.3.1) строить траектории T_∞ с фиксированной начальной точкой $Q_0 = P_0$. Получим величину

$$\xi^* = c_0 \sum_{j=0}^{\infty} W_j^* \psi(Q_j),$$

для которой

$$M\{\xi^* | Q_0 = P_0\} = (\psi, z).$$

2.7. Усложненные оценки линейных функционалов от z . Используя те же траектории T_∞ или \tilde{T}_v , можно построить различные более сложные оценки для функционала (ψ, z) . Рассмотрим одну из них, которая фактически означает оценку (ψ, Kz) вместо (ψ, z) .

В самом деле, из (25) следует, что $(\psi, z) = (\psi, Kz) + (\psi, f)$. Предположим, что интеграл (ψ, f) мы умеем вычислить аналитически (или численно). Тогда расчет (ψ, Kz) позволит нам найти (ψ, z) . В таких задачах, в которых ядро $K(P, P')$ мало по абсолютной величине и $z \approx f$, выделение (ψ, f) из (ψ, z) может сыграть роль выделения главной части и существенно увеличить точность оценки (ψ, z) .

Легко указать два способа расчета (ψ, Kz) . Во-первых, так как $(\psi, Kz) = (K^*\psi, z)$, то можно использовать случайные величины $\xi[K^*\psi]$ или $\tilde{\xi}_v[K^*\psi]$. Во-вторых, так как функция $v = Kz$ удовлетворяет уравнению

$$v = Kv + Kf$$

с тем же ядром, что у исходного уравнения (25), то для расчета (ψ, v) можно использовать величины $\xi[\psi]$ и $\tilde{\xi}_v[\psi]$ с функцией Kf вместо f .

В обоих случаях расчетные формулы по сравнению с п. 2.4 усложняются, так как либо вместо $\psi(Q_0)$ придется вычислять $K^*\psi(Q_0)$, либо вместо $f(Q_i)$ придется вычислять $Kf(Q_i)$

***)** Встречаются уравнения вида (25), в которых P и P' принадлежат различным пространствам. Тогда сопряженные уравнения могут обладать весьма различными свойствами.

Методы Монте-Карло, основанные на использовании случайных траекторий, были первоначально предназначены для решения линейных алгебраических систем (см. ниже § 5). Траектории с поглощением были построены Дж. Форсайтом и Р. Лейблером [121] (по идее Дж. Неймана и С. Улама), а бесконечные траектории — В. Вазовым [181]. Начальные и переходные вероятности, аналогичные (33) и (34), были указаны Дж. Кэртиссом [114]. Обобщения на случай линейных интегральных уравнений имеются у многих авторов, начиная с работы Р. Каткоски [116]. О более сложных оценках для (ψ, z) см. [33, 94, 170].

§ 3. Пример: рассеяние частиц

3.1. Основное уравнение теории рассеяния. Методы Монте-Карло часто используются для расчета различных задач, связанных с прохождением частиц (нейтронов, гамма-квантов и др.) через вещество. Мы не будем касаться здесь специальных вопросов, а рассмотрим лишь общую схему рассеяния, предполагая, что частицы при столкновении с атомами (точнее, с ядрами атомов) среды могут либо рассеиваться, либо поглощаться. Обозначим через P точку шестимерного фазового пространства координат r и скоростей v частицы. Обозначим элемент объема $dr dv$ этого пространства через dP .

Пусть $f(P)dP$ — количество первых столкновений, а $z(P)dP$ — количество всех столкновений в элементе объема dP около точки P (за единицу времени). Функцию $f(P)$ можно явно вычислить, если задан источник частиц. Функцию $z(P)$, называемую *плотностью столкновений*, требуется найти.

Введем *ядро столкновений* $K_{ст}(P', P)$, которое определяется следующим условием: $K_{ст}(P', P)dP$ — это вероятность того, что частица, испытавшая столкновение в точке P' , испытает следующее столкновение в элементе объема dP около точки P (за единицу времени). Конкретный вид ядра столкновений в одnogрупповой теории переноса нейтронов имеется на стр. 223.

Нетрудно составить интегральное уравнение, которому подчиняется плотность столкновений:

$$z(P) = \int K_{ст}(P', P) z(P') dP' + f(P), \quad (49)$$

ибо столкновение в окрестности точки P может быть либо первым столкновением, либо следует за столкновением в окрестности некоторой точки P' , а количество таких столкновений (за единицу времени) равно $z(P')dP'$. Область интегрирования в (49) — все пространство. Введем сопряженное ядро $K(P, P') = K_{ст}(P' P)$. Тогда уравнение (49) совпадает с уравнением (25), а сопряженное уравнение (46) можно записать в виде

$$u(P) = \int K_{ст}(P, P') u(P') dP' + \psi(P). \quad (50)$$

В рассматриваемом случае итерации $f(P)$ — функции $Kf(P), K^2f(P), \dots$ имеют простой физический смысл: это плотности вторых, третьих и т. д. столкновений. И ряд Неймана

$$z(P) = f(P) + Kf(P) + K^2f(P) + \dots$$

означает, что плотность столкновений есть сумма плотностей первых, вторых и т. д. столкновений.

Обычно требуется вычислить какие-нибудь функционалы вида (ψ, z) . Это можно сделать любым из методов, указанных в § 2.

3.2. Использование истинных траекторий. Нетрудно вычислить вероятность $s(P')$ того, что частица после столкновения в точке P' снова испытает столкновение

$$s(P') = \int K_{ct}(P', P) dP.$$

Следовательно, $s(P')$ — это вероятность рассеяния частицы при столкновении в точке P' , а $a(P') = 1 - s(P')$ — вероятность поглощения.

Истинными траекториями частиц будут, очевидно, траектории с поглощением (типа \tilde{T}_v), которые строятся по плотности вероятностей перехода

$$p(P, P') = K_{ct}(P, P')/s(P) \quad (51)$$

с истинной вероятностью поглощения $a(P) = 1 - s(P)$. Плотность (51) всегда допустима по отношению к ядру сопряженного уравнения (50); поэтому естественно вместо функционала (ψ, z) вычислять функционал (f, u) (см. п. 2.6).

Запишем случайную величину (39) применительно к уравнению (50) и функционалу (f, u) :

$$\tilde{\xi}_v^*[f] = [f(Q_0)/p(Q_0)] \tilde{W}_v^*[\psi(Q_v)/a(Q_v)].$$

Нетрудно, однако, убедиться, что при каждом i

$$\tilde{W}_i^* = \prod_{j=1}^i \frac{K_{ct}(Q_{j-1}, Q_j)}{s(Q_{j-1})p(Q_{j-1}, Q_j)} = 1;$$

это — важнейшее следствие закона (51). Таким образом, в рассматриваемом случае

$$\tilde{\xi}_v^*[f] = [f(Q_0)/p(Q_0)] [\psi(Q_v)/a(Q_v)],$$

и $\tilde{\xi}_v^*[f] = (\psi, z)$. Напомним, что v — это случайный номер последней точки траектории.

3.3. Использование искусственных траекторий. Из теоремы 6 вытекает, что в некоторых случаях оценка для (ψ, z) будет иметь меньшую дисперсию, если в расчете использовать траекторию без поглощения (типа $(T_\infty)^*$).

Пусть плотность вероятностей перехода снова определена формулой (51). Согласно (17) и п. 3.2 в этом случае

$$W_j^* = s(Q_0) s(Q_1) \dots s(Q_{j-1}),$$

*) Ограничение $a(P) \equiv a$, фигурирующее в теореме 6, строго говоря, справедливо только в бесконечной однородной среде в одногрупповом приближении.

а величина (36) применительно к уравнению (50) и функционалу (f, u) равна

$$\zeta^*[f] = \frac{f(Q_0)}{p(Q_0)} \sum_{j=0}^{\infty} w_j^* \psi(Q_j).$$

И в этом случае $M\zeta^*[f] = (\psi, z)$.

Вес W_j^* имеет простой физический смысл: это вероятность того, что частица уцелеет (т. е. не поглотится) после столкновений в точках Q_0, Q_1, \dots, Q_{j-1} .

3.4. Существенная выборка. Рекомендации (33) и (34), обеспечивающие при $f(P) = u(P) \geq 0$ минимальные дисперсии оценок \hat{D}_j , в случае уравнения (50) принимают вид

$$p(P) = \frac{f(P) u(P)}{(f, u)}, \quad p(P, P') = \frac{K_{\text{ст}}(P, P') u(P')}{K_{\text{ст}} u(P)}.$$

Так как $f(P) \geq 0$ (по физическому смыслу), то здесь $\hat{D}_j = 0$. Стоит подчеркнуть, что в эти формулы входит не решение $z(P)$ уравнения рассеяния, а решение $u(P)$ сопряженного уравнения.

Конечно, можно обеспечить малые дисперсии, также используя правила (33) и (34) применительно к уравнению (49) с ядром $K(P, P') = K_{\text{ст}}(P', P)$. Однако построение траекторий с плотностью переходов (51) делает метод расчета физически более наглядным и легче контролируемым.

3.5. Рассеяние в области G . Часто встречаются задачи, в которых существенно только рассеяние внутри заданной области G : частица, вылетевшая из G , исключается из рассмотрения. Все изложенные выше методы применимы и к таким задачам, при расчете которых можно считать, что все пространство вне области G заполнено поглощающим веществом: если $K_{\text{ст}}(P', P) \equiv 0$ при $P' \notin G$, то $s(P') \equiv 0$ и $a(P') \equiv 1$ при $P' \notin G$.

Однако, так как нас интересуют значения $z(P)$ только внутри G , то вместо уравнения (49) можно решать уравнение

$$z(P) = \int_G K_{\text{ст}}(P', P) z(P') dP' + f(P), \quad (52)$$

где $P \in G$. Естественнo ожидать, что, применяя методы § 2 к последнему уравнению, можно получить лучшую точность, чем при решении уравнения (49) во всем пространстве (ср. гл. 3, п. 3.1.2). Любые траектории в этом случае не будут истинными траекториями.

В практических расчетах часто используют оценки $\tilde{\zeta}_v^*[f]$ и $\zeta^*[f]$, а плотность $p(P, P')$ выбирают пропорциональной $K_{\text{ст}}(P, P')$. Некоторые методы решения уравнений (49) и (52) для нейтронов в одноструйном приближении имеются в гл. 6, § 4).

Литература к § 3: [8, 25, 33, 36, 50, 51, 53, 59, 93, 105, 127, 144, 145, 153, 168].

§ 4. Однородные интегральные уравнения

4.1. Расчет первого собственного значения и первой собственной функции. Рассмотрим интегральное уравнение

$$z(p) = \lambda \int_G K(P, P') z(P') dP',$$

где λ — действительный параметр. Это же уравнение с учетом (1) можно записать в виде

$$z = \lambda Kz. \quad (53)$$

Если при некотором λ уравнение (53) имеет решение $z(P) \neq 0$, то это значение λ называется *собственным значением уравнения* *) (53), а $z(P)$ — *собственной функцией*, соответствующей этому собственному значению λ . Собственные функции будем считать нормированными так, что $(z, z) = 1$.

Предположим, что наименьшее по абсолютной величине собственное значение уравнения (53) положительно $\lambda_1 > 0$, и соответствующая собственная функция $z_1(P) > 0$ внутри G .

При весьма широких предположениях относительно ядра $K(P, P')$ для приближенного расчета λ_1 и $z_1(P)$ можно воспользоваться методом Келлога [10, 63]: каковы бы ни были положительные внутри G функции $\varphi(P)$ и $\psi(P)$, отношения $(\psi, K^i \varphi) : (\psi, K^{i+1} \varphi)$ стремятся к λ_1

$$\lim_{i \rightarrow \infty} [(\psi, K^i \varphi) / (\psi, K^{i+1} \varphi)] = \lambda_1,$$

а нормированные итерации φ в каждой точке стремятся к собственной функции

$$\lim_{i \rightarrow \infty} K^i \varphi(P) (K^i \varphi, K^i \varphi)^{-1/2} = z_1(P).$$

Расчет величин $(\psi, K^i \varphi)$ и $K^i \varphi(P)$ можно осуществлять методами Монте-Карло, рассмотренными в § 1. Заметим, что траектории T_i позволяют одновременно вычислить все $(\psi, K^j \varphi)$ и $K^j \varphi(P)$ при $j=0, 1, \dots, i$.

*) Собственным значением оператора K называют обратную величину: $1/\lambda$.

Скалярным произведением $(K'\varphi, K^i\varphi)$ на практике пользуются редко: по значениям $K'\varphi(P)$ можно получить приближенный рельеф функции $z_1(P)$, а нормировку осуществить отдельно.

Все вышеупомянутые предположения выполнены, например, в случае, когда ядро $K(P, P')$ симметрично $K(P', P) = K(P, P')$ и положительно определено: $(K\varphi, \varphi) > 0$ при $\varphi \neq 0$ (конечно, предполагается также, что $\varphi(P) \in L_2(G)$, $K(P, P') \in L_2(G \times G)$ и $\psi(P) \in L_2(G)$).

Для того чтобы пояснить сущность метода Келлога, рассмотрим случай (весьма важный), когда все собственные значения уравнения (53) суть

$$0 < \lambda_1 < \lambda_2 \leq \lambda_3 \leq \dots,$$

а соответствующие им собственные функции $z_1(P) > 0$, $z_2(P)$, и т. д. образуют полную ортонормированную систему $(z_i, z_m) = \delta_{im}$ (символ Кронекера). Пусть неотрицательная функция $\varphi(P)$ равна

$$\varphi(P) = \sum_{m=1}^{\infty} b_m z_m(P),$$

где, очевидно, $b_1 = (\varphi, z_1) > 0$. Из этого разложения следует, что

$$K^i\varphi = \sum_{m=1}^{\infty} b_m K^i z_m = \sum_{m=1}^{\infty} b_m \lambda_m^{-i} z_m.$$

Запишем скалярные произведения, фигурирующие в методе Келлога:

$$(\psi, K^i\varphi) = \sum_{m=1}^{\infty} b_m \lambda_m^{-i} (\psi, z_m), \quad (K^i\varphi, K^i\varphi) = \sum_{m=1}^{\infty} b_m^2 \lambda_m^{-2i},$$

и выделим главные члены в интересующих нас величинах:

$$\frac{(\psi, K^i\varphi)}{(\psi, K^{i+1}\varphi)} = \frac{b_1 \lambda_1^{-i} (\psi, z_1) + b_2 \lambda_2^{-i} (\psi, z_2) + \dots}{b_1 \lambda_1^{-i-1} (\psi, z_1) + b_2 \lambda_2^{-i-1} (\psi, z_2) + \dots} = \lambda_1 + O\left(\left|\frac{\lambda_1}{\lambda_2}\right|^i\right),$$

$$\frac{K^i\varphi(P)}{\sqrt{(K^i\varphi, K^i\varphi)}} = \frac{b_1 \lambda_1^{-i} z_1(P) + b_2 \lambda_2^{-i} z_2(P) + \dots}{\sqrt{b_1^2 \lambda_1^{-2i} + b_2^2 \lambda_2^{-2i} + \dots}} = z_1(P) + O\left(\left|\frac{\lambda_1}{\lambda_2}\right|^i\right).$$

Отсюда видно, что сходимость будет хорошей, если $\lambda_1 \ll \lambda_2$, и плохой, если λ_2 близко к λ_1 . Положительность $\psi(P)$, вообще говоря, не нужна: нужно лишь, чтобы $(\psi, z_1) \neq 0$.

Заметим, что если выбрать $\varphi(P)$ так, что $b_1 = (\varphi, z_1) = 0$, но $b_2 = (\varphi, z_2) \neq 0$, то

$$(\psi, K^i\varphi) / (\psi, K^{i+1}\varphi) \rightarrow \lambda_2, \quad K^i\varphi(P) (K^i\varphi, K^i\varphi)^{-1/2} \rightarrow z_2(P).$$

4.2. Метод существенной выборки для траекторий T_i . Предположим, что $K(P, P') \geq 0$ и $\varphi(P) \geq 0$. Интеграл

$$(\psi, K^i \varphi) = \int_G \dots \int_G \psi(P_0) K(P_0, P_1) \dots K(P_{i-1}, P_i) \varphi(P_i) dP_0 \dots dP_i$$

можно вычислить методом п. 3.2 гл. 3 с помощью любой допустимой плотности $p(P_0, \dots, P_i)$ в $G \times \dots \times G$. Из теоремы 3 гл. 3 следует, что если выбрать плотность

$$\hat{p} = |\psi(P_0)| K(P_0, P_1) \dots K(P_{i-1}, P_i) \varphi(P_i) / (|\psi|, K^i \varphi),$$

то дисперсия будет минимальной, и равна она

$$\hat{D}_i = (|\psi|, K^i \varphi)^2 - (\psi, K^i \varphi)^2;$$

при знакопостоянной функции $\psi(P)$ минимум этот $\hat{D}_i = 0$.

Легко убедиться в том, что плотности \hat{p} отвечают траектории $(Q_0 \rightarrow Q_1 \rightarrow \dots \rightarrow Q_i)$ с вероятностями перехода, зависящими от номера точки. Это следует из представления \hat{p} в виде произведения условных плотностей

$$\hat{p} = \frac{|\psi(P_0)| K^i \varphi(P_0)}{(|\psi|, K^i \varphi)} \cdot \frac{K(P_0, P_1) K^{i-1} \varphi(P_1)}{K^i \varphi(P_0)} \dots \frac{K(P_{i-1}, P_i) \varphi(P_i)}{K \varphi(P_{i-1})}.$$

Однако здесь, как и в п. 2.3, законы построения траекторий (33) и (34) обеспечивают минимальность дисперсий, если $\varphi(P)$ пропорциональна $z_1(P)$.

Теорема 7. Предположим, что ядро уравнения (53) и первая собственная функция $z_1(P)$ неотрицательны

$$K(P, P') \geq 0, \quad z_1(P) \geq 0,$$

а траектории T_i строятся по законам (33) и (34):

$$p(P) = \frac{|\psi(P)| \varphi(P)}{(|\psi|, \varphi)},$$

$$p(P, P') = \frac{K(P, P') \varphi(P')}{K \varphi(P)}.$$

Если $\varphi(P) = c z_1(P)$, $c \neq 0$, то при каждом i дисперсия

$$D \theta_i[\psi] = \hat{D}_i.$$

Доказательство. Если $\varphi = c z_1$, то $K \varphi = \lambda_1^{-1} c z_1 = \lambda_1^{-1} \varphi$.

Тогда плотность (3) траекторий T_i равна

$$\begin{aligned} p_i &= \frac{|\psi(P_0)| \varphi(P_0)}{(|\psi|, \varphi)} \cdot \frac{K(P_0, P_1) \varphi(P_1)}{\lambda_1^{-1} \varphi(P_0)} \cdot \dots \cdot \frac{K(P_{i-1}, P_i) \varphi(P_i)}{\lambda_1^{-1} \varphi(P_{i-1})} = \\ &= \frac{|\psi(P_0)| K(P_0, P_1) \dots K(P_{i-1}, P_i) \varphi(P_i)}{\lambda_1^{-i} (|\psi|, \varphi)}. \end{aligned}$$

А так как $K^i \varphi = \lambda_1^{-i} \varphi$, то нетрудно проверить, что плотность \hat{p} также равна этому выражению.

4.3. О других методах расчета λ_1 и $z_1(P)$. Большинство методов для приближенного расчета λ_1 (метод Рунге, метод моментов и др. [63, 60]) так или иначе требует вычисления некоторых интегралов. Если эти интегралы достаточно сложные, то может оказаться целесообразным применение методов Монте-Карло. Рассмотрим схему одного из таких методов, который использовался в [14] для оценки эффективного коэффициента размножения нейтронов в реакторе.

Допустим, что мы можем указать несколько ортонормированных функций $\psi_1(P), \dots, \psi_m(P)$ так, что первая собственная функция $z_1(P)$ уравнения (53) достаточно хорошо аппроксимируется их линейной комбинацией

$$z_1(P) \approx \sum_{i=1}^m c_i \psi_i(P).$$

Вычислим (методом Монте-Карло) интегралы

$$b_{ij} = (K\psi_i, \psi_j), \quad i, j = 1, 2, \dots, m,$$

и составим линейную алгебраическую систему уравнений с неизвестными c_1, \dots, c_m :

$$c_i = \lambda \sum_{j=1}^m b_{ji} c_j, \quad i = 1, 2, \dots, m.$$

Решать эту систему можно методами высшей алгебры. Если $\mu = \tilde{\mu}_1$ — наибольший корень характеристического уравнения

$$\begin{vmatrix} b_{11} - \mu & b_{21} & \dots & b_{m1} \\ b_{12} & b_{22} - \mu & \dots & b_{m2} \\ \dots & \dots & \dots & \dots \\ b_{1m} & b_{2m} & \dots & b_{mm} - \mu \end{vmatrix} = 0,$$

а $\tilde{c}_1, \dots, \tilde{c}_m$ — соответствующее ему решение системы, нормированное так, что $\tilde{c}_1^2 + \dots + \tilde{c}_m^2 = 1$, то

$$\lambda_1 \approx 1/\mu_1, \quad z_1(P) \approx \sum_{i=1}^m \tilde{c}_i \psi_i(P).$$

Идею метода легко объяснить следующим образом. Пусть $\psi_1(P), \dots, \psi^m(P), \dots$ — полная система ортонормированных функций из $L_2(G)$. Разложим в ряды по системе $\{\psi_i\}$ искомую функцию $z(P)$ и итерированные функции $K\psi_i(P)$:

$$z(P) = \sum_{j=1}^{\infty} c_j \psi_j(P), \quad \text{где} \quad c_j = (z, \psi_j),$$

$$K\psi_i(P) = \sum_{j=1}^{\infty} b_{ij} \psi_j(P), \quad \text{где} \quad b_{ij} = (K\psi_i, \psi_j)^*.$$

Подставим эти выражения в уравнение (53): с одной стороны,

$$\lambda Kz(P) = \lambda \sum_{j=1}^{\infty} c_j K\psi_j(P) = \sum_{i=1}^{\infty} \psi_i(P) \lambda \sum_{j=1}^{\infty} b_{ji} c_j,$$

а с другой

$$z(P) = \sum_{i=1}^{\infty} c_i \psi_i(P).$$

Приравнявая коэффициенты при $\psi_i(P)$, получим бесконечную линейную систему уравнений для нахождения величин c_1, \dots, c_m, \dots , эквивалентную уравнению (53),

$$c_i = \lambda \sum_{j=1}^{\infty} b_{ji} c_j, \quad i = 1, 2, \dots$$

Если в этой системе пренебречь величинами c_{m+1}, c_{m+2}, \dots , то придем к конечной системе, выписанной выше.

4.4. Пример: интегральное уравнение Пайерлса. Нахождение критических параметров ядерного реактора в простейших случаях (в однорупповом приближении) сводится к вычислению первого собственного значения λ_1 интегрального уравнения

$$n(P) = \lambda \int_{G_0} \frac{\beta(P') e^{-\left| \int_P^{P'} \alpha ds \right|}}{4\pi |P - P'|^2} n(P') dP', \quad (54)$$

которое называется *уравнением Пайерлса* [25, 51]. Здесь G_0 — трехмерная область (объем реактора), в которой происходит диффузия нейтронов, $\alpha(P)$ и $\beta(P)$ — заданные положительные функции*), ин-

*) Они выражаются через нейтронные сечения (см. стр. 47) и среднее число $\bar{\nu}$ нейтронов деления: $\alpha(P) = \Sigma$, $\beta(P) = \Sigma_s + \bar{\nu} \Sigma_f$. См. также упражнение 5 гл. 6.

теграл в показателе берется по отрезку прямой, соединяющей точки P и P' и равен

$$\left| \int_P^{P'} \alpha ds \right| = \int_0^{|P-P'|} \alpha(P + \omega s) ds,$$

где $\omega = (P' - P)/|P' - P|$ — единичный вектор.

Если $\lambda_1 > 1$, то область G_0 подкритическая, если $\lambda_1 < 1$, то область G_0 надкритическая. Область будет критической при $\lambda_1 = 1$; в этом случае собственная функция $n_1(P)$ равна плотности нейтронов.

Ядро уравнения (54) не симметрично, но симметризуемо и имеет слабую особенность. В. С. Владимиров [10] исследовал сходимость метода Келлога для уравнений такого типа и методы Монте-Карло для расчета приближений. Оказалось, что метод п. 4.1 полностью применим для решения уравнения (54), а если выбрать $\psi = \beta(P)\varphi(P)$, то приближения $\lambda^{(i)} = (K\varphi, \beta\varphi)/(K^{i+1}\varphi, \beta\varphi)$ монотонно убывают:

$$\lambda^{(0)} \geq \lambda^{(1)} \geq \dots \geq \lambda^{(i)} \geq \dots \rightarrow \lambda_1.$$

Расчеты λ_1 и $n_1(P)$ этим методом были осуществлены в [13, 76].

4.4.1. Численный пример [13, 82]. Пусть G_0 — однородный шар радиуса $R=1$ и $\alpha \equiv 1,279$, $\beta \equiv h\alpha$, где $h=1,724$.

Выберем $\varphi(P) \equiv 1$ и $\psi(P) \equiv \beta$. Плотность начальной точки Q_0 зададим формулой $\rho = (2\pi R r_0)^{-1}$, где r_0 — расстояние Q_0 от центра шара. Плотность вероятностей перехода $\rho(P, P')$ из точки P в точку P' выберем так, чтобы в знаменателе стояла величина

$|P - P'|^2$ (ср. гл. 3, п. 3.2.3). Для этого воспользуемся сферическими координатами (ρ, φ, θ) с центром в точке P (рис. 51). Пусть

$$\rho(P, P') = \alpha e^{-\alpha l} [4\pi \rho^2 F(P, \omega)]^{-1},$$

где $F(P, \omega) = 1 - \exp(-\alpha l)$, а l — расстояние PL по направлению ω от точки P до границы шара. Нетрудно проверить, что

$$\oint_{G_0} \rho(P, P') dP' = \oint \frac{d\omega}{4\pi F(P, \omega)} \int_0^1 \frac{\alpha e^{-\alpha \rho}}{\rho^2} \rho^2 d\rho = \oint \frac{d\omega}{4\pi} = 1,$$

так что $\rho(P, P')$ действительно есть условная плотность вероятностей.

Выведем формулы для расчета траектории T_i . Расстояние от начальной точки Q_0 до центра шара вычисляется по формуле $r_0 = R \sqrt{y}$. Две другие координаты можно не разыгрывать из-за симметрии задачи. Направление случайного луча $P' = Q_i + \rho \omega_i$ из точки Q_i определяется величинами $\mu_i = \cos \vartheta_i$ и φ_i . Первую из них можно вы-

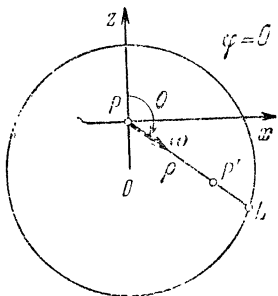


Рис. 51.

числять по формуле $\mu_i = 2\gamma - 1$, а вторую (из соображений симметрии) можно полагать равной нулю $\varphi_i = 0$. Расстояние от точки Q_i по направлению ω_i до границы шара равно

$$l_i = -\mu_i r_i + \sqrt{R^2 - r_i^2 (1 - \mu_i^2)},$$

где r_i — расстояние от Q_i до центра шара. Если обозначить $F_i = F(Q_i, \omega_i)$, то, очевидно, $F_i = 1 - \exp(-\alpha l_i)$. Для определения случайного расстояния $\rho_i = |Q_{i+1} - \zeta_i|$, получаем уравнение

$$\int_0^{\rho_i} \alpha e^{-\alpha s} ds = \gamma F_i,$$

откуда следует, что $\rho_i = -(1/\alpha) \ln(1 - \gamma F_i)$. Вычислив точку $Q_{i+1} = Q_i + \rho_i \omega_i$, можем найти расстояние от нее до центра шара

$$r_{i+1} = \sqrt{r_i^2 + \rho_i^2 + 2\mu_i \rho_i r_i}.$$

Так как $\varphi(P) \equiv 1$, $\psi(P) \equiv \beta$, то (по формуле (6)) нужная нам величина равна $\theta_i \equiv \theta_i[\beta] = 2\pi\beta R r_0 W_i$. Так как из (5) следует, что $W_i = W_{i-1} h F_{i-1}$, а $W_0 = 1$, то легко получить рекуррентную формулу для расчета непосредственно θ_i :

$$\theta_0 = 2\pi\beta R r_0, \quad \theta_i = \theta_{i-1} h F_{i-1}.$$

Полный набор расчетных формул.

1) Начальная точка траектории:

$$r_0 = R \sqrt{\gamma_1}; \quad \theta_0 = 2\pi\beta R r_0.$$

2) Звено номер $j+1$ ($j=0, 1, \dots, i-1$):

$$\mu_j = 2\gamma_{2j+2} - 1; \quad l_j = -\mu_j r_j + \sqrt{R^2 - r_j^2 (1 - \mu_j^2)};$$

$$F_j = 1 - e^{-\alpha l_j}; \quad \rho_j = -(1/\alpha) \ln(1 - \gamma_{2j+3} F_j);$$

$$\theta_{j+1} = \theta_j h F_j; \quad r_{j+1} = \sqrt{r_j^2 + \rho_j^2 + 2\mu_j \rho_j r_j}.$$

3) Если $\theta_{j,s}$ — значение θ_j , полученное при расчете траектории номер s , то нужные нам скалярные произведения $(K^j 1, \beta)$ приближенно равны

$$(K^j 1, \beta) \approx (K^j 1, \beta)_N = \frac{1}{N} \sum_{s=1}^N \theta_{j,s}.$$

Таблица 1
Значения $(K^j, 1, \beta)_N$

$j \backslash N$	2°	2°	2 ¹⁰	2 ¹¹	2·2 ¹¹	3·2 ¹¹	4·2 ¹¹	5·2 ¹¹	6·2 ¹¹	7·2 ¹¹	8·2 ¹¹
0	9,2085	9,2225	9,2294	9,2329	9,2346	9,2346	9,2354	9,2351	9,2354	9,2354	9,2358
1	8,6580	8,6485	8,6528	8,6536	8,6535	8,6536	8,6535	8,6534	8,6530	8,6533	8,6534
2	8,4102	8,4239	8,4086	8,4030	8,3993	8,4002	8,3983	8,3972	8,3961	8,3970	8,3968
3	8,3510	8,3381	8,2746	8,2840	8,2837	8,2823	8,2770	8,2761	8,2721	8,2732	8,2723
4	8,2895	8,2426	8,2596	8,2263	8,2109	8,2045	8,2080	8,2079	8,1951	8,1963	8,2003
5	8,1543	8,1212	8,2042	8,2236	8,2467	8,2154	8,1858	8,1679	8,1421	8,1579	8,1651
6	7,9722	8,0546	8,1441	8,1865	8,2467	8,1977	8,1658	8,1536	8,1297	8,1548	8,1526

Таблица 2
Значения $\lambda_N^{(j)}$

$j \backslash N$	2°	2°	2 ¹⁰	2 ¹¹	2·2 ¹¹	3·2 ¹¹	4·2 ¹¹	5·2 ¹¹	6·2 ¹¹	7·2 ¹¹	8·2 ¹¹
0	1,0636	1,0664	1,0666	1,0669	1,0671	1,0671	1,0672	1,0672	1,0673	1,0673	1,0673
1	1,0285	1,0267	1,0290	1,0298	1,0303	1,0302	1,0304	1,0305	1,0306	1,0305	1,0306
2	1,0080	1,0105	1,0162	1,0144	1,0139	1,0142	1,0147	1,0146	1,0150	1,0149	1,0150
3	1,0074	1,0113	1,0018	1,0070	1,0089	1,0095	1,0084	1,0083	1,0094	1,0094	1,0088
4	1,0166	1,0150	1,0067	1,0003	0,9957	0,9987	1,0027	1,0049	1,0065	1,0047	1,0043
5	1,0222	1,0083	1,0074	1,0045	1,0000	1,0022	1,0024	1,0015	1,0015	1,0004	1,0015

В табл. 1 приведены значения $(K^j 1, \beta)_N$, полученные при расчете по этим формулам, а в табл. 2 — соответствующие отношения

$$\lambda_N^{(j)} = (K^j 1, \beta)_N / (K^{j+1} 1, \beta)_N.$$

При расчете этого же примера методом характеристик *) получено значение $\lambda_1 = 1,000$.

В ходе расчета были сосчитаны дисперсии $D\theta_j$, которые при $j=0, 1, \dots, 6$ равны соответственно 10,7; 21,9; 37,5; 57,3; 81,1; 110,9; 147,0. Если по этим дисперсиям вычислить вероятные ошибки величин $(K^j 1, \beta)_N$, например, при $N=2^{10}$, то получим значения $r_{1024} = 0,07; 0,10; 0,13; 0,17; 0,19; 0,22; 0,26$. Нетрудно заметить, что погрешности соответствующих величин $(K^j 1, \beta)_N$ при $N=2^{10}$ в табл. 1 на порядок меньше, чем r_{1024} (если судить по изменениям этих величин с ростом N). Это вызвано тем, что пример считался с помощью детерминированных псевдослучайных точек, о которых речь пойдет в гл. 7.

§ 5. Решение линейных алгебраических систем

5.1. Алгебраическая система как частный случай интегрального уравнения. Рассмотрим линейную алгебраическую систему, состоящую из m уравнений с m неизвестными z_1, \dots, z_m :

$$z_\alpha = \sum_{\beta=1}^m a_{\alpha\beta} z_\beta + f_\alpha, \quad 1 \leq \alpha \leq m, \quad (55)$$

которую сокращенно будем записывать как

$$z = Az + f. \quad (56)$$

Запись (56) можно считать одним векторным уравнением, где $z = (z_1, \dots, z_m)$ и $f = (f_1, \dots, f_m)$ — m -мерные векторы, $A = (a_{\alpha\beta})$ — квадратная матрица размера $m \times m$ (которая определяет линейное преобразование в пространстве векторов):

$$(Az)_\alpha = \sum_{\beta=1}^m a_{\alpha\beta} z_\beta.$$

Скалярное произведение векторов будем по-прежнему

*) Метод характеристик [11] позволяет находить решения уравнения (54) с большой точностью, но практически применим только в случае достаточно простой «геометрии» области G_0 .

обозначать скобкой

$$(f, z) = \sum_{\alpha=1}^m f_{\alpha} z_{\alpha}.$$

Рассмотрим интегральное уравнение вида (25)

$$z(x) = \int_G K(x, x') z(x') dx' + f(x), \quad (57)$$

где в качестве G выберем отрезок оси абсцисс $0 \leq x < m$. Обозначим через G_{α} отрезки $\alpha-1 \leq x \leq \alpha$, так что

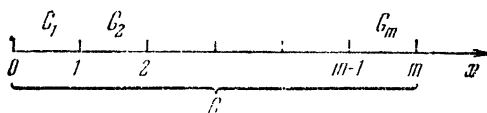


Рис. 52.

$G = G_1 + \dots + G_m$ (рис. 52). В качестве свободного члена и ядра уравнения (57) выберем кусочно постоянные функции

$$\begin{aligned} f(x) &= f_{\alpha} \quad \text{при } x \in G_{\alpha}, \\ K(x, x') &= a_{\alpha\beta} \quad \text{при } x \in G_{\alpha}, \quad x' \in G_{\beta}. \end{aligned}$$

Для такого уравнения при $x \in G_{\alpha}$

$$z(x) = \sum_{\beta=1}^m \int_{G_{\beta}} K(x, x') z(x') dx' + f_{\alpha} = \sum_{\beta=1}^m a_{\alpha\beta} \int_{G_{\beta}} z(x') dx' + f_{\alpha}.$$

И так как последнее выражение от x не зависит, то решение $z(x)$ постоянно в G_{α} .

Обозначив

$$z(x)|_{G_{\alpha}} = z_{\alpha}, \quad (58)$$

сможем переписать последнее уравнение как

$$z_{\alpha} = \sum_{\beta=1}^m a_{\alpha\beta} z_{\beta} + f_{\alpha}.$$

Следовательно, решение $z(x)$ уравнения (57) представляет собой функцию, постоянную в каждом из G_{α} , значения z_{α} которой на этих отрезках удовлетворяют

алгебраической системе (55). Обратно, если (z_1, \dots, z_m) — решение системы (55), то формула (58) определяет решение уравнения (57).

5.2. Случайная цепь для решения алгебраической системы. Очевидно, все методы Монте-Карло, приведенные в § 2, применимы для оценки решений уравнения (57) и дают нам возможность оценить решение системы (55). Однако специфика уравнения (57) позволяет упростить эти методы и дать им иную интерпретацию.

Так как $f(x)$, $K(x, x')$ и $z(x)$ постоянны при $x \in G_\alpha$, $x' \in G_\beta$, то естественно выбирать плотности $p(x)$ и $p(x, x')$ также постоянными

$$\begin{aligned} p(x) &= p_\alpha & \text{при } x \in G_\alpha, \\ p(x, x') &= p_{\alpha\beta} & \text{при } x \in G_\alpha, x' \in G_\beta. \end{aligned} \quad (59)$$

Величины p_α и $p_{\alpha\beta}$, очевидно, должны быть неотрицательными и удовлетворять условиям нормировки

$$\sum_{\alpha=1}^m p_\alpha = 1, \quad \sum_{\beta=1}^m p_{\alpha\beta} = 1. \quad (60)$$

Правила (59) можно интерпретировать как равномерное распределение случайной точки внутри соответственно G_α или G_β . Можно, однако, совсем отказаться от фиксации положения этой точки и говорить только об отрезке, в котором эта точка расположена. Тогда p_α — это вероятность того, что начальная точка траектории попадет в G_α , а $p_{\alpha\beta}$ — вероятность того, что случайная точка из G_α перейдет в G_β . При такой интерпретации вместо траектории T_i случайной точки достаточно рассмотреть последовательность случайных номеров

$$k_0 \rightarrow k_1 \rightarrow \dots \rightarrow k_i, \quad (61)$$

тех отрезков G_α , в которые эта точка попадет.

Итак, для решения системы (55) мы будем строить *цепь случайных номеров* (61), каждый из которых может принимать значения $1, 2, \dots, m$. Правила построения цепи (61):

$$P\{k_0 = \alpha\} = p_\alpha, \quad P\{k_j = \beta | k_{j-1} = \alpha\} = p_{\alpha\beta}, \quad (62)$$

где начальные вероятности p_α и вероятности переходов $p_{\alpha\beta}$ должны удовлетворять условиям нормировки (60).

В теории вероятностей такие цепи называются *цепями Маркова* с конечным числом m состояний [71]. Веса W_j вдоль цепи (61) вычисляются по формуле

$$W_j = \frac{a_{k_0 k_1} a_{k_1 k_2} \cdots a_{k_{j-1} k_j}}{p_{k_0 k_1} p_{k_1 k_2} \cdots p_{k_{j-1} k_j}} \quad (63)$$

или по рекуррентной формуле

$$W_j = W_{j-1} (a_{k_{j-1} k_j} / p_{k_{j-1} k_j}), \quad W_0 = 1.$$

Формулы эти — следствие (4) и (5).

Условимся говорить, что распределение вероятностей (p_1, \dots, p_m) *допустимо* по отношению к вектору $\psi = (\psi_1, \dots, \psi_m)$, если для тех α , для которых $\psi_\alpha \neq 0$, значение $p_\alpha > 0$. Аналогично распределение $(p_{\alpha\beta})$ *допустимо* по отношению к матрице $A = (a_{\alpha\beta})$, если $p_{\alpha\beta} > 0$ для тех пар (α, β) , для которых $a_{\alpha\beta} \neq 0$.

Как известно, из теории матриц, для того чтобы ряд $\sum_{j=0}^{\infty} A^j f$ сходиллся для любого вектора f , необходимо и достаточно, чтобы все собственные значения $\mu = \mu_\alpha$ матрицы A , представляющие собой корни уравнения $\det |a_{\alpha\beta} - \mu \delta_{\alpha\beta}| = 0$, лежали внутри единичного круга (на комплексной плоскости): $|\mu_\alpha| < 1$ при $\alpha = 1, 2, \dots, m$. Достаточным условием может служить неравенство

$$\sum_{i,j=0}^m a_{ij}^2 < 1 \quad \text{или} \quad \text{неравенство} \quad \max_{1 \leq i \leq m} \sum_{j=1}^m |a_{ij}| < 1.$$

Сформулируем для системы (55) теорему, вытекающую из теоремы 4 (§ 2). Пусть бесконечная цепь

$$k_0 \rightarrow k_1 \rightarrow \dots \rightarrow k_i \rightarrow \dots$$

строится по правилам (62), где p_α и $p_{\alpha\beta}$ допустимы по отношению к ψ и $(a_{\alpha\beta})$ соответственно, и рассмотрим случайную величину

$$\xi[\psi] = (\psi_{k_0} / p_{k_0}) \sum_{j=0}^{\infty} W_j f_{k_j}. \quad (64)$$

Теорема 8. Если все собственные значения $\bar{\mu}_\alpha$ матрицы $(|a_{\alpha\beta}|)$ по абсолютной величине меньше единицы, то математическое ожидание случайной величины $\xi[\psi]$ равно

$$M\xi[\psi] = (\psi, z). \quad (65)$$

Повторим доказательство со стр. 177 применительно к рассматриваемому случаю. Так как

$$P\{k_0 = t_0, \dots, k_j = t_j\} = p_{t_0} p_{t_0 t_1} \dots p_{t_{j-1} t_j},$$

то, принимая во внимание (63), получим, что

$$M\left\{\frac{\psi_{k_0}}{p_{k_0}} W_{j|f_{k_j}}\right\} = \sum_{t_0, \dots, t_j=1}^m \psi_{t_0} a_{t_0 t_1} \dots a_{t_{j-1} t_j} f_{t_j} = (\psi, A^j f).$$

Затем

$$M_z^{\infty}[\psi] = \sum_{j=0}^{\infty} M[(\psi_{k_0}/p_{k_0}) W_{j|f_{k_j}}] = \sum_{j=0}^{\infty} (\psi, A^j f) = (\psi, z),$$

причем существование этого математического ожидания обеспечивается условиями теоремы

Если какие-нибудь из элементов матрицы $(a_{\alpha\beta})$ равны нулю, то обычно целесообразно выбрать соответствующие вероятности перехода $p_{\alpha\beta} = 0$ (ср. п. 3.2.1 гл. 3). В противном случае, если в цепи окажется переход $k_{j-1} \rightarrow k_j$, которому отвечает $p_{k_{j-1} k_j} > 0$ и $a_{k_{j-1} k_j} = 0$, то из (63) видно, что все $W_j = W_{j+1} = \dots = 0$; можно считать, что цепь, попав в k_j , останавливается: любые дальнейшие переходы $k_j \rightarrow k_{j+1} \rightarrow \dots$ значения величины $\xi[\psi]$ не изменяют.

Если матрица A содержит целую строку из нулей: $a_{\alpha 1} = a_{\alpha 2} = \dots = a_{\alpha m} = 0$, то, конечно, нельзя все соответствующие вероятности выбрать нулевыми: это противоречило бы условию (60). Тогда можно положить $p_{\alpha\alpha} = 1$, $p_{\alpha\beta} = 0$ при $\beta \neq \alpha$, так что $p_{\alpha\beta} = \delta_{\alpha\beta}$.

5.3. Вычисление одной компоненты решения. Иногда встречаются задачи, сводящиеся к системе (55), где, однако, нас интересует не все решение $z = (z_1, \dots, z_m)$, а только одна из неизвестных, например z_r . Методы Монте-Карло позволяют приближенно оценить одну эту компоненту. Для этого достаточно в качестве ψ выбрать единичный вектор $e^{(r)} = (0, \dots, 0, 1, 0, \dots, 0)$, в котором лишь на r -м месте стоит 1. Тогда скалярное произведение равно

$$(\psi, z) = (e^{(r)}, z) = \sum_{\alpha=1}^m e_{\alpha}^{(r)} z_{\alpha} = z_r.$$

В качестве начальных вероятностей можно выбрать $p_{\alpha} = e_{\alpha}^{(r)}$ или, другими словами, начинать цепь с $k_0 = r$.

Итак, строим цепи $r \rightarrow k_1 \rightarrow k_2 \rightarrow \dots \rightarrow k_i \rightarrow \dots$, вычисляем вдоль цепей веса

$$W_i = \frac{a_{rk_1} a_{k_1 k_2} \dots a_{k_{i-1} k_i}}{p_{rk_1} p_{k_1 k_2} \dots p_{k_{i-1} k_i}} \quad (66)$$

и случайные величины

$$\xi[e^{(r)}] = \sum_{i=0}^{\infty} W_i f_{k_i} \quad (67)$$

Если количество цепей N достаточно велико, то

$$z_r \approx \frac{1}{N} \sum_{s=1}^N \xi[e^{(r)}]_s,$$

где $\xi[e^{(r)}]_s$ — значение $\xi[e^{(r)}]$, полученной на s -й цепи.

5.4. Обращение матрицы. Пусть задана квадратная матрица $B = (b_{\alpha\beta})$, $1 \leq \alpha, \beta \leq m$. Требуется вычислить элементы обратной матрицы $C = (c_{\alpha\beta})$, удовлетворяющей соотношению $BC = E$, где E — единичная матрица:

$$E = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Если все собственные значения μ_α матрицы B удовлетворяют условию

$$|\mu_\alpha - 1| < 1,$$

то задача эта может быть решена методом последовательных приближений*). В самом деле, обозначим через A матрицу

$$A = E - B. \quad (68)$$

Легко проверить, что собственные значения матрицы A равны $1 - \mu_\alpha$ и, следовательно, по абсолютной величине меньше единицы. В силу этого сходится ряд

$$C = \sum_{i=0}^{\infty} A^i$$

и

$$BC = (E - A)C = C - AC = \sum_{i=0}^{\infty} A^i - \sum_{i=1}^{\infty} A^i = E^{**}).$$

*) Более общий случай см. упражнение 6 стр. 209.

**) Для запоминания можно заметить, что ряд $(E - A)^{-1} = \sum_{i=0}^{\infty} A^i$ аналогичен геометрической прогрессии $(1 - q)^{-1} = \sum_{i=0}^{\infty} q^i$.

Докажем, что если матрица A удовлетворяет условиям теоремы 8, то для того, чтобы вычислить элемент $c_{rr'}$ матрицы C , можно использовать цепи п. 5.3; надо лишь, кроме $\psi = e^{(r')}$, выбрать еще $f = e^{(r)}$.

В самом деле, мы видели, что $M\xi[e^{(r)}] = z_r$, где z_r — r -я компонента решения уравнения (56). Однако это решение равно

$$z = \sum_{i=0}^{\infty} A^i f = Cf,$$

так что его r -я компонента есть

$$z_r = \sum_{\alpha=1}^m c_{r\alpha} f_{\alpha}.$$

В частном случае, когда $f = e^{(r')}$, получим, что $z_r = c_{rr'}$, что и требовалось доказать.

Подставив в (67) вместо f вектор $e^{(r')}$, получим соответствующий этому случаю метод Монте-Карло: при достаточно большом N

$$c_{rr'} \approx \frac{1}{N} \sum_{s=1}^N \left[\sum_{(i|k_i=r')} W_i \right]_s; \quad (69)$$

здесь снова s —номер цепи, а сумма с условием $(i|k_i=r')$ означает, что суммируются только те W_i , для которых номер $k_i=r'$.

5.4.1. Численный пример: найти матрицу C , обратную к матрице

$$B = \begin{pmatrix} 1/2 & 0 & 0 \\ -2 & 1/2 & 0 \\ -1 & 0 & 1 \end{pmatrix}.$$

собственные значения которой, очевидно, равны $\mu_1 = \mu_2 = 1/2$, $\mu_3 = 1$.

Сперва по формуле (68) вычислим матрицу

$$A = \begin{pmatrix} 1/2 & 0 & 0 \\ 2 & 1/2 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Затем выберем матрицу вероятностей перехода, допустимую по отношению к матрице A , например:

$$(p_{\alpha\beta}) = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 1 & 0 & 0 \end{pmatrix}. \quad (70)$$

1) Чтобы оценить элементы c_{11} , c_{12} и c_{13} , надо строить цепи, начинающиеся с $k_0=r=1$. Так как в первой строке матрицы (70) лишь $p_{11} \neq 0$, то все цепи окажутся одинаковыми:

$$k_0=1, \quad k_1=1, \quad k_2=1, \dots$$

Веса вдоль такой цепи меняются на множитель $a_{11}/p_{11}=1/2$:

$$W_0 = 1, \quad W_1 = \frac{1}{2}, \quad W_2 = \frac{1}{4}, \dots$$

Осреднение в (69) не нужно, достаточно одной цепи для получения точного результата

$$c_{11} = \sum_{(i|k_i=1)} W_i = 1 + \frac{1}{2} + \frac{1}{4} + \dots = 2,$$

$$c_{12} = \sum_{(i|k_i=2)} W_i = 0; \quad c_{13} = \sum_{(i|k_i=3)} W_i = 0.$$

2) Столь же просто оцениваются элементы c_{31} , c_{32} и c_{33} : цепь начинается с $k_0=r=3$; так как в третьей строке матрицы (70) лишь один ненулевой элемент $p_{31}=1$, то $k_1=1$, далее $k_2=1$, $k_3=1, \dots$. Итак, все цепи будут одинаковыми

$$k_0=3, \quad k_1=1, \quad k_2=1, \quad k_3=1, \dots,$$

а соответствующие веса (с учетом того, что $a_{31}/p_{31}=1$) равны

$$W_0 = 1, \quad W_1 = 1, \quad W_2 = \frac{1}{2}, \quad W_3 = \frac{1}{4}, \dots$$

Поэтому

$$c_{31} = \sum_{(i|k_i=1)} W_i = 1 + \frac{1}{2} + \frac{1}{4} + \dots = 2,$$

$$c_{32} = \sum_{(i|k_i=2)} W_i = 0, \quad c_{33} = \sum_{(i|k_i=3)} W_i = 1.$$

3) И только для вычисления c_{21} , c_{22} и c_{23} нужно действительно строить случайные цепи, начинающиеся с $k_0=r=2$, и производить расчеты. Впрочем, в этом случае различные типы цепей легко классифицируются, ибо во второй строке матрицы (70) лишь p_{21} и p_{22} отличны от нуля.

Цепь, начало которой $k_0=2$, может несколько раз оставаться в $k_1=k_2=\dots=k_t=2$. Но если окажется, что $k_{t+1}=1$, то дальше уже все $k_{t+1}=k_{t+2}=\dots=1$. Следовательно, общий вид такой цепи ($0 \leq t < \infty$)

$$k_0 = k_1 = \dots = k_t = 2, \quad k_{t+1} = k_{t+2} = \dots = 1. \quad (71)$$

Соответствующие такой цепи веса легко вычислить: так как $a_{22}/p_{22}=1$, $a_{21}/p_{21}=4$, $a_{11}/p_{11}=1/2$, то

$$W_0 = W_1 = \dots = W_t = 1; \quad W_{t+1} = 4; \quad W_{t+2} = 2, \quad W_{t+3} = 1, \dots$$

Интересующие нас суммы весов вдоль такой цепи равны

$$\sum_{(i/k_i=1)} W_i = 4 + 2 + 1 + 1/2 + \dots = 8,$$

$$\sum_{(i/k_i=2)} W_i = t + 1, \quad \sum_{(i/k_i=3)} W_i = 0.$$

Из этих соотношений сразу следует, что $c_{21}=8$, $c_{23}=0$.

Оставшийся элемент c_{22} можно вычислить аналитически, если записать, что вероятность получить цепь (71) равна (см. начало стр. 39)

$$(p_{22})^t p_{21} = (1/2)^{t+1},$$

и вместо формулы (69) воспользоваться точным выражением для математического ожидания

$$c_{22} = M \sum_{(i/k_i=2)} W_i = \sum_{t=0}^{\infty} (t+1) 2^{-(t+1)} = 2.$$

Итак, обратная матрица C равна

$$C = \begin{pmatrix} 2 & 0 & 0 \\ 8 & 2 & 0 \\ 2 & 0 & 1 \end{pmatrix}.$$

Конечно, в общем случае (когда в матрице A много ненулевых элементов) разнообразие цепей столь велико, что заменить формулу (69) аналитическим расчетом невозможно. Заметим также, что по цепям, начинающимся с $k_0=r$, можно одновременно вычислять все $c_{rr'}$ с $r'=1, 2, \dots, m$, ибо каждое W_i входит в сумму $\sum_{(i/k_i=r')} W_i$ при одном и только одном r' .

5.5. Решение дифференциальных уравнений Лапласа и Пуассона. В ограниченной связной области G плоскости x, y с простой границей G^0 рассмотрим

дифференциальное уравнение с частными производными

$$\partial^2 u / \partial x^2 + \partial^2 u / \partial y^2 = F(x, y), \quad (72)$$

где $u = u(x, y)$ — искомая функция. Уравнение (72) при $F(x, y) \equiv 0$ называется *уравнением Лапласа*, а при $F(x, y) \neq 0$ — *уравнением Пуассона*.

Предположим, что на границе G^0 задана некоторая функция $g(x, y)$ (часто пишут $g(S)$, где S — длина дуги границы, отсчитываемая от какой-нибудь фиксирован-

ной точки). Требуется найти такое решение $u(x, y)$ уравнения (72), которое на границе совпадает с $g(x, y)$:

$$u(x, y)_{G^0} = g. \quad (73)$$

Задачу об отыскании решения уравнения (72), удовлетворяющего граничному условию (73), называют *задачей Дирихле*.

Для приближенного решения этой задачи [63, 88] выбирают на плоскости достаточно мелкую квадратную сетку с шагом h (рис. 53). Координаты узлов этой сетки

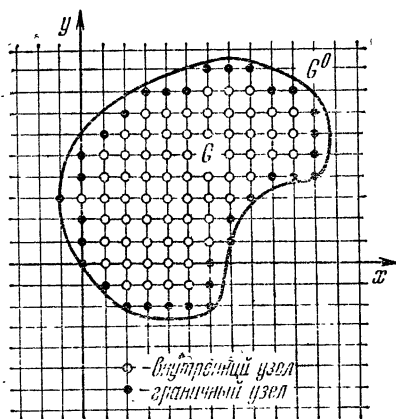


Рис. 53.

пусть будут $x_j = jh$, $y_l = lh$, а значения $u(x_j, y_l)$ и $F(x_j, y_l)$ для краткости обозначим $u_{j,l}$ и $F_{j,l}$. Узел (j, l) называют *внутренним*, если и он, и все четыре соседних с ним узла $(j-1, l)$, $(j+1, l)$, $(j, l-1)$, $(j, l+1)$ принадлежат $G \cup G^0$, в противном случае узел (j, l) , принадлежащей $G \cup G^0$, называют *граничным*.

Во внутреннем узле (x_j, y_l) уравнение (72) заменяется разностным уравнением

$$\frac{u_{j+1,l} - 2u_{j,l} + u_{j-1,l}}{h^2} + \frac{u_{j,l+1} - 2u_{j,l} + u_{j,l-1}}{h^2} = F_{j,l},$$

которое можно переписать в виде

$$u_{j,l} = (1/4)(u_{j-1,l} + u_{j+1,l} + u_{j,l-1} + u_{j,l+1} - h^2 F_{j,l}). \quad (74)$$

В граничных узлах полагают

$$u_{j,i} = g_{j,i}. \quad (75)$$

(Значения $g_{j,i}$ «сносят» с ближайших точек границы G^0 .) Решение алгебраической системы (74)—(75) при $h \rightarrow 0$ приближается к решению задачи Дирихле для уравнения (72).

Если перенумеровать все узлы, принадлежащие $G+G^0$ (в произвольном порядке), и переписать в том же порядке уравнения (74) и (75), то получим систему вида (55)

$$u_\alpha = \sum_{\beta=1}^m a_{\alpha\beta} u_\beta + f_\alpha, \quad \alpha = 1, 2, \dots, m$$

(m — количество узлов), с весьма своеобразной матрицей A : внутреннему узлу с номером α отвечает строка $a_{\alpha 1}, \dots, a_{\alpha m}$, в которой 4 элемента равны $1/4$, а остальные — нули; граничному узлу с номером α отвечает строка $a_{\alpha 1} = a_{\alpha 2} = \dots = a_{\alpha m} = 0$; все диагональные элементы $a_{\alpha\alpha} = 0$. (Можно доказать, что все собственные значения такой матрицы по абсолютной величине меньше единицы.) Свободные члены этой системы равны $f_\alpha = -(1/4)h^2 F_\alpha$, если узел номер α внутренний, и $f_\alpha = g_\alpha$, если узел номер α граничный.

Воспользуемся методом п. 5.3. и построим метод Монте-Карло для расчета u_r — значения решения в одном (заранее заданном) узле. Выберем матрицу переходов

$$p_{\alpha\beta} = \begin{cases} a_{\alpha\beta}, & \text{если узел номер } \alpha \text{ внутренний,} \\ \delta_{\alpha\beta}, & \text{если узел номер } \alpha \text{ граничный} \end{cases}$$

($\delta_{\alpha\beta}$ — символ Кронекера: $\delta_{\alpha\alpha} = 1$, $\delta_{\alpha\beta} = 0$ при $\alpha \neq \beta$).

Процесс построения цепи по такому закону оказывается очень наглядным: 1) начинаем с $k_0 = r$; 2) если узел k_i внутренний, то с одинаковой вероятностью $1/4$ выбираем в качестве k_{i+1} номер одного из соседних с ним узлов; 3) если узел k_i граничный, то цепь останавливается: $k_i = k_{i+1} = k_{i+2} = \dots$

Расчет весов W_i вдоль такой цепи также чрезвычайно прост: пока цепь не попала на границу $W_0 = W_1 = \dots, \dots = W_i = 1$; далее $W_{i+1} = W_{i+2} = \dots = 0$.

Поэтому случайная величина (67) оказывается равной

$$\zeta = f_{k_0} + f_{k_1} + \dots + f_{k_i}, \quad (76)$$

где k_i — номер первого выхода цепи на границу. В (76) все $f_{k_0}, \dots, f_{k_{i-1}}$ вычисляются по формуле $f_\alpha = -(1/4)h^2 F_\alpha$ и лишь последнее f_{k_i} равно значению g_{k_i} .

З а м е ч а н и е. Если вместо граничных условий (73), заданы более сложные условия, например:

$$[c_1 u + c_2 (\partial u / \partial x) + c_3 (\partial u / \partial y)]_{\partial \Omega} = g,$$

то уравнения (75) наряду с $u_{j,i}$ будут содержать также значения u_α в некоторых соседних узлах. И случайная цепь, попав на границу, останавливаться не будет.

Пр и м е р. Пусть $u(x, y)$ — решение уравнения Лапласа $(\partial^2 u / \partial x^2) + (\partial^2 u / \partial y^2) = 0$ в единичном квадрате $0 \leq x \leq 1, 0 \leq y \leq 1$, удовлетворяющее граничным условиям $u(x, 0) = 0, u(0, y) = 0, u(x, 1) = x, u(1, y) = y$. Вычислить значение $u(1/2, 1/2)$.

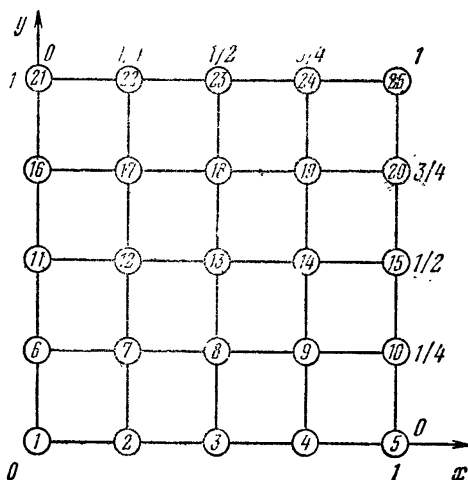


Рис. 54.

Выберем в квадрате сетку с шагом $h=1/4$ и перенумеруем узлы так, как это указано на рис. 54. Для уравнения Лапласа формула (76) еще более упрощается: $\zeta = g_{k_i}$, так что ζ равно значению g в том узле, в котором цепь попадает на границу. Возле каждого граничного узла на рис. 54 проставлено значение g для нашего примера.

Для построения цепей воспользуемся таблицей случайных цифр, приведенной на стр. 295. Если случайная цифра ε окажется 0 или 4, то будем перемещаться в соседний узел справа; если ε равно 1 или 5, то будем перемещаться влево; если ε равно 2 или 6, то перемещаемся вверх; если ε равно 3 или 7, то перемещаемся вниз; значения ε , равные 8 или 9, опускаем.

Таблица 3

$\begin{array}{ccc} 6 & 5 & 1 \\ \uparrow & \leftarrow & \leftarrow \\ 13-18-17-16 \end{array}$	$\begin{array}{ccccccc} 5 & 0 & 7 & 5 & 6 & 6 & 1 \\ \leftarrow & \rightarrow & \downarrow & \leftarrow & \uparrow & \uparrow & \leftarrow \\ 13-12-13-8-7-12-17-16 \end{array}$	$\begin{array}{cc} 5 & 5 \\ \leftarrow & \leftarrow \\ 13-12-11 \end{array}$	
$\begin{array}{cc} 6 & 6 \\ \uparrow & \uparrow \\ 13-18-23 \end{array}$	$\begin{array}{ccc} 4 & 3 & 4 \\ \rightarrow & \downarrow & \rightarrow \\ 13-14-9-10 \end{array}$	$\begin{array}{ccc} 5 & 6 & 5 \\ \leftarrow & \uparrow & \leftarrow \\ 13-12-17-16 \end{array}$	$\begin{array}{cc} 5 & 1 \\ \leftarrow & \leftarrow \\ 13-12-11 \end{array}$
$\begin{array}{ccccccc} 2 & 3 & 3 & 2 & 4 & 3 & 7 \\ \uparrow & \downarrow & \downarrow & \uparrow & \rightarrow & \downarrow & \downarrow \\ 13-18-13-8-13-14-9-4 \end{array}$	$\begin{array}{ccc} 7 & 5 & 7 \\ \downarrow & \leftarrow & \downarrow \\ 13-8-7-2 \end{array}$	$\begin{array}{ccc} 0 & 2 & 6 \\ \rightarrow & \uparrow & \uparrow \\ 13-14-19-24 \end{array}$	
$\begin{array}{ccccccc} 1 & 6 & 0 & 3 & 3 & 3 & \\ \leftarrow & \uparrow & \rightarrow & \downarrow & \downarrow & \downarrow & \\ 13-12-17-18-13-8-3 \end{array}$	$\begin{array}{ccccc} 4 & 2 & 5 & 0 & 2 \\ \rightarrow & \uparrow & \leftarrow & \rightarrow & \uparrow \\ 13-14-19-18-19-24 \end{array}$	$\begin{array}{cc} 2 & 2 \\ \uparrow & \uparrow \\ 13-18-23 \end{array}$	
$\begin{array}{cccc} 4 & 5 & 5 & 5 \\ \rightarrow & \leftarrow & \leftarrow & \leftarrow \\ 13-14-13-12-11 \end{array}$	$\begin{array}{cc} 3 & 7 \\ \downarrow & \downarrow \\ 13-18-3 \end{array}$	$\begin{array}{cc} 5 & 1 \\ \leftarrow & \leftarrow \\ 13-12-11 \end{array}$	

В табл. 3 приведены 16 случайных цепей. В первой строке записаны использованные случайные цифры, во второй — схематически указано направление перемещения, а в третьей — сама цепь (номера k). Соответствующие этим цепям значения ζ равны 0, 0, 0, 1/2, 1/4, 0, 0, 0, 0, 3/4, 0, 3/4, 1/2, 0, 0, 0. Среднее арифметическое этих величин дает нам приближенное значение решения в точке (1/2, 1/2):

$$u\left(\frac{1}{2}, \frac{1}{2}\right) \approx \frac{1}{16} \sum_{s=1}^{16} \zeta_s = 0,17.$$

Из эмпирической оценки дисперсии

$$D\zeta \approx \frac{1}{15} \sum_{s=1}^{16} \zeta_s^2 - \frac{1}{15 \cdot 16} \left(\sum_{s=1}^{16} \zeta_s \right)^2 = 0,081$$

следует, что вероятная ошибка $r_{16} = 0,675 \sqrt{D\zeta/16} \approx 0,05$.

Точное решение рассмотренной задачи $u=xy$, так что $u(1/2, 1/2) = 0,25$, и фактическая ошибка расчета равна 0,08.

Метод настоящего пункта позволяет вычислять решения разностных уравнений, аппроксимирующих дифференциальные уравнения. Связь таких задач с блужданиями по сетке была впервые установлена в [113]. В § 3 гл. 8 указан метод Монте-Карло для решения задачи Дирихле, не связанный с разностными уравнениями.

5.6. Случай очень большого числа переменных. Методы Монте-Карло, изложенные в §§ 2, 3, 4, часто используются в вычислительной практике, так как классические численные методы решения интегральных уравнений в многомерном случае приводят к весьма сложным расчетным схемам. Методы настоящего параграфа, напротив, сравнительно редко используются, так как эти задачи хорошо решаются численными средствами линейной алгебры. Пожалуй, только в задачах с большим числом переменных методы Монте-Карло могут успешно конкурировать с методами линейной алгебры [115].

Предположим, что требуется оценить одну компоненту z_r решения z системы (55), у которой все $|a_{\alpha\beta}| < q/m$, $|f_\alpha| < c$, и $q < 1$. В этом случае ряд $z_r = f_r + (A^1 f)_r + (A^2 f)_r + \dots$ сходится быстрее, чем геометрическая прогрессия $c + cq + cq^2 + \dots$, и при заданной точности можно ограничиться каким-то определенным количеством t слагаемых:

$$z_r = \sum_{i=0}^t (A^i f)_r. \quad (77)$$

Будем считать, что $t \geq 3$. Для того чтобы по компонентам вектора $A^i f$ вычислить все компоненты вектора $A^{i+1} f$, необходимо проделать m^2 умножений (операциями сложения, как более простыми, мы пренебрегаем). У вектора $A^i f$ нам нужна лишь одна компонента $(A^i f)_r$. Следовательно, общее количество умножений, затрачиваемых при расчете по формуле (77), равно $(t-1)m^2 + m$.

Рассмотрим теперь метод п. 5.3. Пусть все $\rho_{\alpha\beta} = 1/m$; это значит, что каждый номер k_i с равной вероятностью может принимать значения $1, 2, \dots, m$ независимо от k_{i-1} . Такой выбор можно (см. стр. 46) осуществлять по формуле $k_i = 1 + U(\gamma \nu)$, где ν — очередное случайное число.

Для реализации цепи с такими вероятностями перехода надо: 1) выбрать $k_0 = r$; 2) выбрать k_1 и найти элемент $a_{k_0 k_1}$; 3) выбрать k_2 и найти $a_{k_1 k_2}$ и т. д. до $a_{k_{t-1} k_t}$. Значение случайной величины (67), соответствующей (77), будет вычисляться по формуле

$$\zeta = \sum_{i=0}^t W_i f_{k_i}, \quad \text{где } W_i = W_{i-1} m a_{k_{i-1} k_i}, \quad W_0 = 1.$$

Если матрицу $a_{\alpha\beta}$ заранее умножить на m , то на расчет одного значения ζ будет затрачиваться всего $3t$ умножений (m на γ , $ma_{k_t-k k_t}$ на W_{i-1} и f_{k_t} на W_i). И если для достижения требуемой точности придется вычислить N цепей, то общее количество умножений при таком способе расчета равно $3tN + m^2$.

Очевидно, расчет методом Монте-Карло будет экономичнее, чем расчет по формуле (77), если

$$(t-2)m^2 + m > 3tN. \quad (78)$$

Легко показать, что величина N — количество цепей, необходимое для достижения заданной вероятной ошибки — ограничена при $m \rightarrow \infty$ и $t \rightarrow \infty$. В самом деле, как известно, N пропорционально дис-

персии $D\zeta$. Но $D\zeta \leq M\zeta^2$ и так как $|\zeta| \leq c \sum_{i=0}^t q^i < c(1-q)^{-1}$, то

$D\zeta < c^2(1-q)^{-2}$. Последняя граница ни от m , ни от t не зависит.

Квадратному неравенству (78) удовлетворяют все m такие, что

$$m > [-1 + \sqrt{1 + 12t(t-2)N}] / [2(t-2)].$$

Так как $N \gg 1$, то это неравенство можно заменить более простым

$$m > \sqrt{12t(t-2)N} / [2(t-2)] = \sqrt{3Nt/(t-2)}.$$

Наконец, так как при $t \geq 3$ отношение $t/(t-2) \leq 3$, то последнее неравенство, а вместе с ним и (78), будет выполнено при любом $t \geq 3$, если только

$$m > 3\sqrt{N}.$$

Например, если $N=900$, то метод Монте-Карло выгоднее, чем расчет по формуле (77), для систем порядка $m > 90$.

Упражнения к главе 5

1. Рассмотрим произвольную функцию $R(P, P')$ из $L_2(G \times G)$ и решение $z(P)$ уравнения (25), представимое в виде сходящегося ряда (28). Чтобы вычислить квадратичный функционал (Rz, z) методом Монте-Карло, можно использовать пары независимых случайных траекторий (Л. В. Майоров, А. М. Суховой [49]).

а) Пусть две траектории типа $T_\infty (Q_0 \rightarrow \dots \rightarrow Q_i \rightarrow \dots)$ и $(Q'_0 \rightarrow \dots \rightarrow Q'_j \rightarrow \dots)$ строятся по плотностям

$$p(Q_0), p(Q_{i-1}, Q_i) \text{ и } p'(Q'_0), p'(Q'_{j-1}, Q'_j)$$

соответственно. Определим случайную величину η , зависящую от пары таких траекторий:

$$\eta = \frac{R(Q_0, Q'_0)}{p(Q_0)p'(Q'_0)} \sum_{i=0}^{\infty} W_i f(Q_i) \sum_{j=0}^{\infty} W'_j f(Q'_j).$$

б) Пусть две траектории типа $\tilde{T}_v(Q_0 \rightarrow \dots \rightarrow Q_v)$ и $(Q'_0 \rightarrow \dots \rightarrow Q'_v)$ строятся по законам $p(Q_0)$, $p(Q_{i-1}, Q_i)$, $a(Q_i)$ и $p'(Q'_0)$, $p'(Q'_{i-1}, Q'_i)$, $a'(Q'_i)$ соответственно. Определим случайную величину $\tilde{\eta}_{vv'}$, зависящую от пары таких траекторий:

$$\tilde{\eta}_{vv'} = \frac{R(Q_0, Q'_0)}{p(Q_0) p'(Q'_0)} \tilde{W}_v \frac{f(Q_v)}{a(Q_v)} \tilde{W}'_{v'} \frac{f(Q'_v)}{a(Q'_v)}.$$

Доказать, что если выполнены условия теоремы 4, то

$$M\eta = M\tilde{\eta}_{vv'} = (Rz, z).$$

2. Доказать, что если плотности $p(P)$ и $p(P, P')$, по которым строятся траектории T_∞ , удовлетворяют условиям

$$\int_G [\psi(P) f(P) p^{-1}(P)]^2 p(P) dP < \infty$$

и

$$\int_G [K(P, P') f(P') p^{-1}(P, P') f^{-1}(P)]^2 p(P, P') dP' \leq q < 1$$

(при любом $P \in G$), то дисперсия $D\xi[\psi] < \infty$.

У к а з а н и е. Воспользоваться неравенством (44).

3. Предположим, что функции

$$\psi_D(P) = \frac{\psi^2(P)}{p(P)}, \quad f_D(P) = \frac{f^2(P)}{a(P)}, \quad K_D(P, P') = \frac{K^2(P, P')}{s(P)p(P, P')}$$

принадлежат соответственно $L_2(G)$, $L_2(G)$, $L_2(G \times G)$, и ряд Неймана для решения z_D уравнения $z_D = K_D z_D + f_D$ сходится. Доказать, что дисперсия оценки (39) конечна и равна

$$D\tilde{\xi}_v[\psi] = (\psi_D, z_D) - (\psi, z)^2$$

(В. Г. Золотухин, С. М. Ермаков [36]).

4. Интегральное уравнение

$$z(x) = \frac{3}{20} \int_0^1 \frac{x+x'}{\sqrt{x'}} z(x') dx' + 45x - 3$$

имеет решение $z=50x$.

а) Доказать, что последовательные приближения (26) сходятся, если только интеграл $K\phi(x)$ конечен.

б) Пусть $\Delta_1, \dots, \Delta_l$ — заданные интервалы, расположенные внутри $[0, 1]$, и требуется вычислить методом Монте-Карло п. 2.2

числа

$$\bar{z}_{i,k} = \int_{\Delta_k} z^{(i)}(x) dx, \quad k = 1, 2, \dots, l,$$

в случае, когда начальное приближение $z^{(0)}(x) \equiv 1$. Выписать все расчетные формулы, если начальная плотность $p(x) \equiv 1$, а плотность вероятностей перехода из x в x' пропорциональна $(x+x')/\sqrt{x'}$.

5. Доказать, что в условиях п. 3.2 (расчет по истинным траекториям) можно для оценки (ψ, z) вместо случайной величины

$$\tilde{\kappa}_v^* [f] = [f(a_0)/p(Q_0)] [\psi(Q_v)/a(Q_v)],$$

называемой *оценкой по поглощениям*, использовать случайную величину

$$\tilde{\kappa}_v^* [f] = [f(Q_0)/p(Q_0)] \sum_{j=0}^v \psi(Q_j), \quad \text{называемую оценкой по}$$

столкновениям.

Указание. Получив выражение

$$M\tilde{\kappa}_v^* = \sum_{i=0}^{\infty} \int \dots \int f(P_0) \sum_{j=0}^i \psi(P_j) \prod_{s=1}^i K_{\text{ст}}(P_{s-1}, P_s) a(P_i) dP_0 \dots dP_i,$$

использовать тождество $a(P_i) = 1 - \int K_{\text{ст}}(P_i, P_{i+1}) dP_{i+1}$.

6. Рассмотрим квадратную матрицу $B = (b_{\alpha\beta})$, $1 \leq \alpha, \beta \leq m$, все собственные значения μ_α которой удовлетворяют условию

$$|\mu_\alpha - h| < h, \quad h > 0. \quad (79)$$

Построить метод Монте-Карло для обращения матрицы B при помощи итераций матрицы $A = E - h^{-1}B$. (Если все собственные значения μ_1, \dots, μ_m матрицы B действительны и положительны, то условие (79) всегда выполнено при $h = \max(\mu_1; \dots; \mu_m)$.)

7. В области $G = \{0 \leq x \leq 1, t \geq 0\}$ рассмотрим дифференциальное уравнение теплопроводности

$$\partial u / \partial t = \partial^2 u / \partial x^2 + F(x, t).$$

Требуется найти решение $u = u(x, t)$ этого уравнения, удовлетворяющее граничным условиям

$$u(x, 0) = g(x), \quad u(0, t) = g_0(t), \quad u(1, t) = g_1(t).$$

Выберем прямоугольную сетку с шагом h по x и Δt по t . Координаты узлов этой сетки $x_j = jh$, $t_l = l\Delta t$, а значения $u(x_j, t_l)$ и $F(x_j, t_l)$ обозначим $u_{j,l}$ и $F_{j,l}$. Во всех внутренних узлах заменим дифференциальное уравнение разностным

$$\frac{u_{j,l} - u_{j,l-1}}{\Delta t} = \frac{u_{j+1,l-1} - 2u_{j,l-1} + u_{j-1,l-1}}{h^2} + F_{j,l-1},$$

которое, введя параметр $\kappa = \Delta t/h^2$, можно переписать в форме

$$u_{j,l} = \kappa u_{j+1,l-1} + (1 - 2\kappa) u_{j,l-1} + \kappa u_{j-1,l-1} + \Delta t F_{j,l-1}.$$

Условие устойчивости этого уравнения $\kappa \leq 1/2$.

Определить случайную цепь, аналогичную цепи из п. 5.5, для расчета $u_{j,l}$ методом Монте-Карло.

ГЛАВА 6

МОДЕЛИРОВАНИЕ ЕСТЕСТВЕННЫХ ПРОЦЕССОВ

Рассмотрим какой-нибудь естественный процесс, на протекание которого влияют различные случайные факторы. Законы распределения этих факторов будем считать известными: иногда эти законы можно вычислить теоретически, чаще — экспериментально. Так как мы умеем находить значения случайных величин с любыми законами распределения (гл. 1 и 2), то, разыграв конкретные значения случайных факторов, мы сможем рассчитать конкретную случайную реализацию процесса. Такой расчет называют *имитацией* (simulation).

Большинство приложений метода Монте-Карло связано именно с имитацией. Технические приемы имитации в различных областях науки, техники, экономики подробно рассматриваются в специальных книгах. Поэтому мы ограничимся лишь очень кратким изложением нескольких простых примеров (§ 1).

С точки зрения математики гораздо интереснее то, что для многих задач можно построить более выгодные алгоритмы расчета, если отказаться от имитации естественного процесса и воспользоваться некоторыми искусственными моделями.

Именно таким способам расчета — моделированию с использованием статистических весов — посвящена настоящая глава. Основная область применения этих методов (пока) — нейтронная физика. Однако материалы конференции [55] показывают, что такие методы начинают проникать и в другие области.

§ 1. Моделирование путем имитации

1.1. Задача о поглощении нейтронов. Рассмотрим ограниченную выпуклую (для простоты) область G_0 в пространстве, не содержащую делящихся ядер. Считаем известными сечения рассеяния и поглощения $\Sigma(r, E) = \Sigma_s(r, E) + \Sigma_a(r, E)$ и индикатрису рассеяния $p_s(r', \Omega'; \Omega)$, которая представляет собой условную плотность распределения направлений рассеяния Ω . В точке $r_0 \in G_0$ расположен источник нейтронов с энергией E_0 и с равновероятными направлениями Ω_0 начальной скорости. Требуется вычислить вероятность p_A того, что нейтрон, вылетевший из источника, поглотился в области G_0 .

Задача рассматривается в одnogрупповом приближении, так что энергия нейтрона E_0 при рассеянии не меняется.

Возможностью возвращения вылетевшего нейтрона обратно в G_0 пренебрегаем. Последнее условие означает, что нейтрон, вылетевший из G_0 , перестает нас интересовать. Поэтому можно внешнюю среду заменить любой поглощающей средой с сечением $\Sigma = \Sigma_a \geq \varepsilon_0 > 0$.

Рассмотрим какой-нибудь нейтрон, порожденный источником. Выберем случайное направление Ω_0 его скорости (методом гл. 2, п. 2.4.2). Затем разыграем для него случайную длину ξ_0 свободного пробега — как это делать указано ниже в § 2. Получим точку столкновения нейтрона $r_1 = r_0 + \xi_0 \Omega_0$.

Если $r_1 \notin G_0$, то мы считаем, что история нейтрона закончилась вылетом из области G_0 , и полагаем случайную величину $\eta_A = 0$.

При $r_1 \in G_0$ разыгрываем судьбу нейтрона (методом гл. 2, п. 1.2.2). Если нейтрон поглотился, то история его заканчивается и полагаем $\eta_A = 1$. Если же нейтрон рассеялся, то в соответствии с индикатрисой $p_s(r_1, \Omega_0; \Omega)$ разыгрываем новое направление скорости $\Omega = \Omega_1$, затем новую длину свободного пробега ξ_1 , и вычисляем следующую точку столкновения $r_2 = r_1 + \xi_1 \Omega_1$.

Расчет траектории продолжается до вылета нейтрона из области G_0 или до его поглощения. (Можно доказать, что при весьма общих условиях вероятность бесконечной последовательности рассеяний внутри G_0 равна нулю.) Если история нейтрона заканчивается погло-

щением, то полагаем $\eta_A=1$, в противном случае полагаем $\eta_A=0$ (рис. 55).

По определению искомая вероятность $p_A=P\{\eta_A=1\}$. Так как распределение случайной величины η_A задается таблицей

$$\begin{pmatrix} 1 & 0 \\ p_A & 1-p_A \end{pmatrix},$$

то $M\eta_A=p_A$, $D\eta_A=p_A-p_A^2$. Оценкой величины p_A

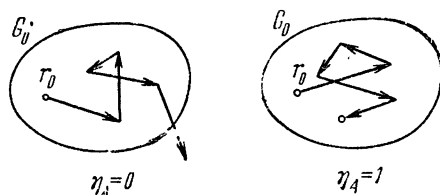


Рис. 55.

служит среднее арифметическое

$$p_A \approx \frac{1}{N} \sum_{s=1}^N (\eta_A)_s, \quad (1)$$

где $(\eta_A)_s$ — значение η_A , полученное на s -й траектории, а N — общее количество реализованных траекторий. Последнюю формулу можно записать также в виде

$$p_A \approx N_A/N,$$

где N_A — количество траекторий, закончившихся поглощением нейтрона.

Изложенный метод Монте-Карло для расчета p_A основан на имитации поведения нейтронов в среде. В каком-то смысле это самый естественный метод решения задачи. Однако это не лучший метод. Ниже, в § 3 указаны несколько видоизменений этого метода, позволяющие строить случайные величины η такие, что $M\eta=p_A$ и $D\eta \leq D\eta_A$. При этом будет использована следующая лемма.

Лемма. Если случайная величина η , для которой $M\eta=p_A$, удовлетворяет неравенствам $0 \leq \eta \leq 1$, то $D\eta \leq D\eta_A$.

Доказательство. Так как $M\eta^2 \leq M\eta=p_A$, то дисперсия $D\eta = M\eta^2 - p_A^2 \leq p_A - p_A^2 = D\eta_A$.

1.2. Задача о размножении нейтронов. Предположим, что область G_0 , рассмотренная в п. 1.1, содержит расщепляющиеся вещества, так что полное сечение $\Sigma = \Sigma_s + \Sigma_c + \Sigma_f$, где Σ_c — сечение захвата, а Σ_f — сечение деления. При захвате поглощенный нейтрон «исчезает», а при делении вместо поглощенного нейтрона появляются ν новых нейтронов, называемых нейтронами деления. Распределение случайной величины ν задано *).

Снова ограничимся одногрупповым приближением и будем считать, что энергии всех нейтронов равны E_0 ; рассеяние изотропно; распределение скоростей каждого из нейтронов деления также изотропно.

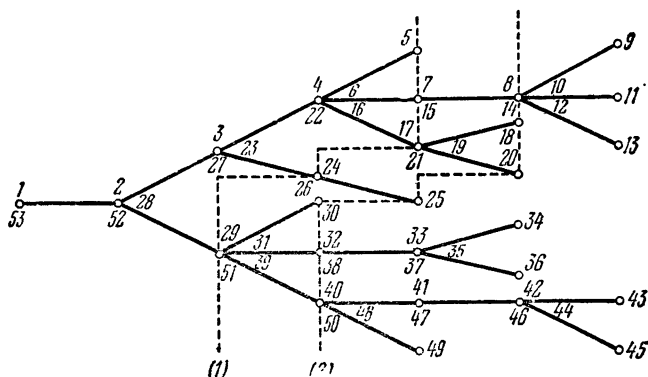


Рис 56.

Рассмотрим историю одного нейтрона, оказавшегося в области G_0 . Каждое звено его траектории моделируется так же, как в п. 1.1: разыгрываются случайное направление Ω_i , случайный пробег ξ_i и «судьба» при столкновении в точке $r_{i+1} = r_i + \xi_i \Omega_i$ (гл. 2, п. 1.2.2). Если произошло деление, то разыгрывается величина ν , и траектория в точке r_{i+1} разветвляется на ν ветвей, каждая из которых моделируется независимо по тем же формулам. Полученную ветвящуюся траекторию часто называют *деревом*. Схема такого дерева изображена на рис. 56.

*) Обычно, Если $M\nu = \bar{\nu}$, заключено между m и $m+1$, то считают, что $P\{\nu = m+1\} = m+1 - \bar{\nu}$, $P\{\nu = m\} = \bar{\nu} - m$.

Область G_0 называется *критической*, если количество нейтронов в этой области постоянно во времени. Наиболее распространенный метод расчета критичности — *счет по поколениям* [25, 51, 53] — состоит в следующем. Поместим в G_0 некоторое количество n_1 нейтронов, которые будем считать нейтронами 1-го поколения. Затем разыграем историю каждого из этих нейтронов до его «исчезновения» (т. е. до вылета из G_0 , до захвата или до поглощения с делением). При этом мы получим некоторое количество n_2 нейтронов деления, которые будем считать нейтронами второго поколения. После расчета достаточно большого числа i поколений, отношения n_{i+1}/n_i устанавливаются, так как $n_{i+1}/n_i \rightarrow k_{эфф}$. Величина $k_{эфф}$ называется *эффективным коэффициентом размножения нейтронов* в G_0 . Если $k_{эфф} = 1$, то область G_0 критична. (При $k_{эфф} < 1$ цепная реакция в G_0 затухает, а при $k_{эфф} > 1$ — приводит к взрыву.).

Очевидно, для расчета по поколениям строить ветвящиеся траектории не нужно.

1.2.1. О технике расчета ветвящихся траекторий. Ветвящиеся траектории встречаются во многих задачах, связанных с прохождением элементарных частиц, когда, например, фотон может привести к образованию пары электрон — позитрон или к появлению свободного электрона (фотоэффект). При расчете деревьев вовсе не обязательно сперва строить все дерево, а потом производить его обработку (т. е. отбор нужных данных и осуществление нужных вычислений). Укажем два основных алгоритма постепенной обработки дерева.

а) Обработка дерева по поколениям. Этот способ удобен в тех задачах, в которых деревья длинные, но не слишком сильно ветвятся. Схема расчета достаточно очевидна: по частицам одного поколения вычисляем все частицы следующего поколения и одновременно производим обработку построенной части дерева. Затем информацию о старом поколении можно уничтожить. Так что в памяти приходится хранить не более двух поколений.

Вообще говоря, к поколению можно отнести любые две частицы, не являющиеся «предками» друг друга (например, нейтроны, соединенные пунктиром (1) на рис. 56). Тогда следующее поколение определяется однозначно (пунктир (2)).

б) Лексикографическая обработка дерева. Этот способ удобен в тех задачах, в которых деревья не очень длинные, но сильно ветвятся. Схема расчета такова: движемся по одной (какой-нибудь) ветви, производя обработку и записывая все ответвления. Достигнув конца ветви, возвращаемся на одно колено назад и начинаем двигаться по какому-нибудь из записанных ответвлений. Если записанных ответвлений нет, то возвращаемся еще на колено

назад. Счет закончится тогда, когда мы возвратимся к основанию дерева.

Номера на рис. 56 указывают порядок одного из возможных лексикографических обходов изображенного дерева. Ясно, что при каждом возвращении на одно колено можно уничтожить информацию об обработанной части. Поэтому в любой момент в памяти ЭВМ сохраняется лишь одна полная или неполная ветвь с записанными ответвлениями.

1.3. Системы массового обслуживания. Рассмотрим модель одной из простейших (однофазных) систем массового обслуживания. Такая система состоит из n линий (или каналов, или пунктов обслуживания), каждая из которых может независимо «обслуживать» заявки. В любой момент времени линия может быть либо свободной, либо занятой.

Примерами таких систем могут служить автозаправочная станция (линии — бензоколонки, заявки — приезжающие машины), телефонная станция (линии — каналы связи, заявки — вызовы абонентов), парикмахерская (линии — мастера, заявки — клиенты) и т. д.

В систему поступает случайный поток заявок, вероятностные характеристики которого предполагаются известными. Если в момент поступления k -й заявки, назовем его t_k , имеются свободные линии, то одна из них (правило выбора должно быть задано) принимает на себя обслуживание этой заявки. Продолжительность обслуживания заявки i -й линией представляет собой случайную величину τ , плотность распределения которой $p_i(t)$ также должна быть задана.

Предположим, что наша система представляет собой *систему с отказами*: если в момент t_k все линии заняты, то система выдает отказ. (Такая модель годится для телефонной станции, но не подходит для автозаправочной станции или парикмахерской, которые ближе к так называемым системам с ожиданием). Важнейшая характеристика системы с отказами — среднее число отказов в заданном интервале времени. Конечно, число отказов зависит от поступающего потока заявок.

Аналитические методы позволяют рассчитывать системы массового обслуживания только в простейших случаях. Во всех более сложных задачах приходится прибегать к статистическому моделированию [7, 38, 43].

1.3.1. Описание потока заявок. Чаще других рассматривается *стационарный пуассоновский поток*, называемый также *простейшим потоком*, в котором интервалы между моментами поступления двух последовательных заявок независимы и подчиняются экспоненциальному распределению: $t_k = t_{k-1} + \theta$, и функция распределения θ равна

$$F(x) = 1 - e^{-ax}, \quad 0 < x < \infty. \quad (2)$$

Величина $a = 1/M\theta$ называется *интенсивностью* потока заявок.

Из более сложных стационарных потоков отметим только *потoki Эрланга*, в которых величины $\theta = t_k - t_{k-1}$ также независимы, но подчиняются более общему гамма-распределению (см. гл. 2, п. 4.2) с плотностью

$$p(x) = \frac{a^m}{(m-1)!} x^{m-1} e^{-ax}, \quad 0 < x < \infty.$$

Интенсивность такого потока равна $1/M\theta = a/m$.

Сравнительно редко используют *нестационарные пуассоновские* потоки, в которых функция распределения $\theta = t_{k+1} - t_k$ зависит от t_k :

$$F(x) = 1 - \exp \left\{ - \int_0^x a(t_k + s) ds \right\}, \quad 0 < x < \infty.$$

Нетрудно проверить, что $F(x)$ в последнем случае имеет тот же вид, что функция распределения свободных пробегов (см. ниже формулу (6)). Изложенный в п. 2.2.3 метод постоянного сечения позволяет легко моделировать такие нестационарные потоки, вводя фиктивные заявки.

1.3.2. Алгоритм расчета системы с отказами. Рассмотрим конкретную систему с отказами, состоящую из n линий. Плотности $p_i(t)$ считаем известными. Правило выбора свободной линии сформулируем так: заявка поступает на линию, которая освободилась раньше всех, а если таких несколько, то на ту из них, номер которой меньше.

Пусть в рассматриваемую систему поступает простейший поток заявок с интенсивностью a . Требуется оценить среднее число отказов за время $t_0 \leq t < t_{\text{кон.}}$

Поставим в соответствие линии с номером i ячейку ЭВМ, в которую будем записывать момент освобождения этой линии T_i . В начальный момент, очевидно, $T_1 = T_2 = \dots = T_n = t_0$. В отдельную ячейку будем записывать величину $\hat{T} = \min_{1 \leq i \leq n} T_i$.

Момент поступления первой заявки разыграем по формуле

$$t_1 = t_0 - (1/a) \ln \gamma_1.$$

Очевидно, первую заявку начнет обслуживать первая линия. Время обслуживания τ найдем из уравнения

$$\int_0^{\tau} p_1(t) dt = \gamma_2.$$

Линия номер 1 будет занята до момента $(T_1)_{\text{новое}} = t_1 + \tau$. Следовательно, в соответствующей первой линии ячейке надо заменить T_1 на $(T_1)_{\text{новое}}$ и вычислить новое \hat{T} .

Предположим теперь, что «судьбу» $(k-1)$ -й заявки, поступившей в момент t_{k-1} , мы уже выяснили. Тогда разыгрываем момент поступления k -й заявки

$$t_k = t_{k-1} - (1/a) \ln \gamma_{2k-1}$$

и проверяем, если ли свободные линии. Для этого достаточно проверить неравенство

$$\hat{T} \leq t_k. \quad (3)$$

а) Если (3) выполнено, то линия, для которой $T_i = \hat{T}$, приступает к обслуживанию k -й заявки (если линий, для которых $T_i = \hat{T}$, несколько, то выбирается та из них, номер которой меньше). Продолжительность обслуживания τ разыгрывается по формуле

$$\int_0^{\tau} p_i(t) dt = \gamma_{2k};$$

затем вычисляется $(T_i)_{\text{новое}} = t_k + \tau$ и новое значение \hat{T} . После этого можно перейти к рассмотрению следующей, $(k+1)$ -й заявки.

б) Если неравенство (3) невыполнено, то свободных линий в момент t_k нет и система должна выдать отказ,

В этом случае необходимо прибавить единицу к счетчику отказов, после чего можно перейти к рассмотрению $(k+1)$ -й заявки.

Расчет продолжается до тех пор, пока момент поступления очередной заявки не окажется больше, чем

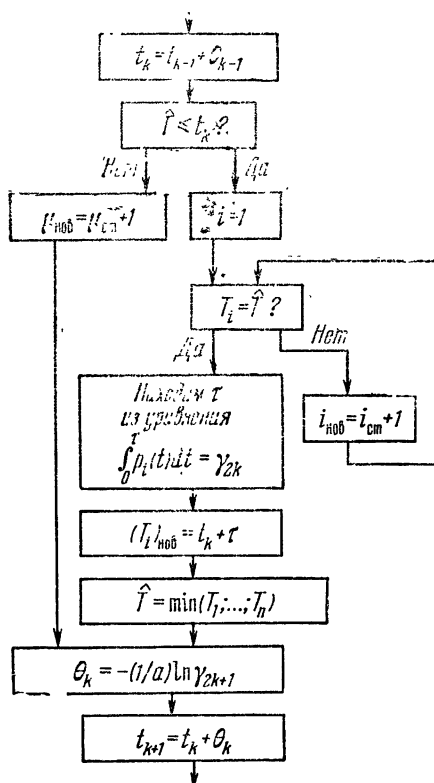


Рис 57.

$t_{кон}$. К этому времени в счетчике отказов будет стоять количество отказов μ за время $t_0 \leq t < t_{кон}$.

Повторив такой опыт N раз (с различными случайными числами), получим N значений μ_1, \dots, μ_N , по которым можно оценить среднее число отказов за время $t_0 \leq t < t_{кон}$: $M\mu \approx (1/N)(\mu_1 + \dots + \mu_N)$. Схема обработки одной заявки изображена на рис. 57.

1.4. Расчет вероятностных характеристик сложной случайной величины. Рассмотрим n -мерную случайную точку $Q = (\xi^1, \dots, \xi^n)$, закон распределения которой $F_Q(x_1, \dots, x_n)$ известен. Требуется вычислить какие-нибудь вероятностные характеристики скалярной случайной величины

$$\eta = f(Q),$$

например $M\eta$ или $P\{\eta \in \Delta\}$, где Δ — заданный интервал.

Если функция $f(x_1, \dots, x_n)$ хоть сколько-нибудь сложная (или задана алгоритмом, а не аналитическим выражением), то аналитически вычислить функцию распределения $F_\eta(y)$ не удастся. Весьма естественно в этой ситуации использовать метод Монте-Карло.

В самом деле, приемы гл. 2 позволяют находить независимые реализации Q_i точки Q . По ним можно вычислить соответствующие значения $\eta_i = f(Q_i)$ величины η . Достаточно большая выборка η_1, \dots, η_N позволяет оценить любые характеристики величины η , например,

$$M\eta \approx \frac{1}{N} \sum_{i=1}^N \eta_i, \quad P\{\eta \in \Delta\} \approx \frac{1}{N} \sum_{i=1}^N \chi_\Delta(\eta_i),$$

где $\chi_\Delta(y)$ — индикатор интервала Δ (см. стр. 166). В п. 3.1.4 гл. 1 указано, как по выборке η_1, \dots, η_N построить эмпирическую функцию распределения $F_N(y)$, которая при больших N близка к $F_\eta(y)$ *).

*) Для того, чтобы оценить плотность $\rho_\eta(y)$ случайной величины η , обычно строят *гистограмму* $h_N(y)$. Для этого выбирают интервал $[a, b)$, содержащий все выборочные значения η_1, \dots, η_N , разбивают его на r интервалов:

$$[a, b) = \sum_{j=1}^r [y_{j-1}, y_j), \quad \text{где } y_0 = a, \quad y_r = b,$$

и вычисляют количества выборочных значений, принадлежащих этим интервалам: v_1, \dots, v_r (ср. гл. 1, п. 3.1.2). Гистограммой называется ступенчатая линия $h_N(\cdot) = (v_j/N)/(y_j - y_{j-1})$, при $y \in [y_{j-1}, y_j)$; ($j = 1, 2, \dots, r$).

О гистограммах в n -мерном случае см. [98].

В указанную схему укладываются способы расчета многих практических задач. Интересно, что в большинстве случаев приходится выполнять серии расчетов, отличающихся значениями некоторого параметра λ , так что $\eta(\lambda) = f(Q; \lambda)$, и метод гл. 3, п. 4.1 оказывается весьма полезным.

Мы ограничимся только одним примером.

1.4.1. Оценка качества приборов. Рассмотрим электрический прибор, который состоит из сопротивлений $R^{(1)}, R^{(2)}, \dots$ емкостей $C^{(1)}, C^{(2)}, \dots$ и др. Предположим, что качество прибора определяется величиной U , которая может быть вычислена по параметрам деталей:

$$U = f(R^{(1)}, R^{(2)}, \dots, C^{(1)}, C^{(2)}, \dots, \dots) \quad (4)$$

В действительности параметры всех деталей несколько отличаются от номинальных, и поэтому значения U для разных экземпляров прибора будут отличаться от (4). Можно пытаться оценить пределы изменения U , выбирая самые неблагоприятные значения параметров. Однако далеко не всегда ясно, какой набор параметров будет наихудшим. К тому же, как правило, такая оценка большого практического значения не имеет, так как в действительности мало вероятно, чтобы все параметры одновременно были наихудшими (особенно если их много). Поэтому более рационально считать параметры всех деталей независимыми случайными величинами и методом п. 1.4 определить закон распределения случайной величины U .

Для такого расчета нужно знать функции распределения параметров всех деталей. К сожалению, вероятностные характеристики деталей заводами-изготовителями пока не выдаются. Конечно, можно получить такие данные путем просмотра больших партий однотипных деталей. Но это довольно большая работа. И чтобы избежать ее, некоторые исследователи поступают проще: если указано, что величина сопротивления может отклоняться от номинала R на $\pm 5\%$, то считают, что величина эта нормальна с параметрами $(a; \sigma)$, которые определяются из условий

$$a = R, \quad 3\sigma = 0,05R$$

(ср. с «правилом трех сигм», стр. 88).

Заметим, что экспериментальное определение вероятностных характеристик прибора возможно лишь тогда, когда имеется достаточно большая партия готовых приборов. Метод настоящего пункта позволяет оценить эти характеристики еще на стадии проектирования.

§ 2. Моделирование свободного пробега

2.1. Закон распределения длины свободного пробега.

Предположим, что вдоль оси Ox движется некоторая элементарная частица (нейтрон, фотон, протон и т. п.), которая может сталкиваться с частицами (обычно с ядрами атомов) среды. Во всех исследованиях, явно или неявно, фигурирует следующее предположение: вероятность того, что частица, долетевшая до точки x , испытает столкновение в интервале $(x, x+\Delta x)$, равна

$$\Sigma \Delta x + o(\Delta x). \quad (5)$$

Множитель пропорциональности Σ называется *полным сечением* или подробнее: *полным макроскопическим эффективным сечением взаимодействия частицы со средой*. Значение Σ зависит как от состояния среды в точке x , так и от типа и энергии частицы.

Формулу (5) можно считать определением полного сечения*). Вполне аналогично определяются сечения различных типов взаимодействия (различных реакций), которые могут произойти при столкновении частицы. Например, Σ_s — сечение рассеяния, Σ_f — сечение деления и др. Такие сечения иногда называют *парциальными сечениями*.

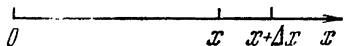


Рис. 58.

Рассмотрим теперь частицу, вылетевшую вдоль оси Ox из точки $x=0$ (рис. 58). Обозначим через ξ случайную длину свободного пробега этой частицы,

*) Обычно рассматривают пучок, содержащий N одинаковых частиц, долетевших до точки x . Если в среднем в интервале $(x, x+\Delta x)$ испытывают столкновение $N \cdot \Delta x$ частиц такого пучка, то $\Sigma(x) = N'/N$.

а через $F(x)$ — функцию распределения ξ , так что $F(x) = P\{\xi \leq x\}$.

Вероятность того, что частица испытает первое столкновение в интервале $(x, x+\Delta x)$, равна

$$F(x+\Delta x) - F(x) = [1 - F(x)] [\Sigma(x)\Delta x + o(\Delta x)].$$

(Здесь $1 - F(x)$ — вероятность того, что частица долетит до точки x .) Разделив это соотношение на Δx и перейдя к пределу при $\Delta x \rightarrow 0$, получим дифференциальное уравнение

$$dF/dx = [1 - F(x)] \Sigma(x).$$

Решение этого уравнения, удовлетворяющее начальному условию $F(0) = 0$, легко записать:

$$F(x) = 1 - \exp \left\{ - \int_0^x \Sigma(s) ds \right\}. \quad (6)$$

Это и есть искомая функция распределения ξ .

Интеграл $I(x) = \int_0^x \Sigma(s) ds$ часто называют *оптической длиной* интервала $(0, x)$. Из условия нормировки $F(\infty) = 1$ следует необходимое требование $I(\infty) = \infty$.

2.1.1. Ядро столкновений в одногрупповой теории переноса *). В этой теории скорость частиц предполагается постоянной по абсолютной величине (но не по направлению) и не меняется во время свободного пробега. Поэтому фазовое пространство (см § 3 гл. 5) можно считать пятимерным $P = (r, \Omega)$, где r — координаты частицы, а Ω — единичный вектор направления скорости.

По определению $K_{st}(P', P) dP$ — это вероятность того, что после столкновения в точке P' фазового пространства частица испытает следующее столкновение в элементе dP около точки P . Вероятность того, что частица при столкновении в P' рассеется, равна $\Sigma_s(r')/\Sigma(r')$.

Обозначим $l = |r - r'|$, $\omega = (r - r')/l$, и в качестве координат точки r выберем сферические координаты с центром в r' (рис. 59), так что $dr = l^2 dl d\omega$ — элемент «физического» объема. Для того чтобы рассеянная частица могла попасть в элемент dr около точки r , направление ее движения после рассеяния должно оказаться в конусе $d\omega$ около направления ω . Вероятность такого события выражается через индикатрису рассеяния $p(r', \Omega'; \Omega)$ и равна $p_s(r', \Omega'; \omega) d\omega$. Если это условие выполнено, и части-

*) Одногрупповая (или односкоростная) теория часто используется для расчета диффузии быстрых нейтронов, особенно в среде, состоящей из тяжелых атомов [25, 51].

на рассеялась по направлению ω , то вероятность столкновения в интервале $(l, l+dl)$, согласно (6), равна

$$\Sigma(r) \exp \left\{ - \int_0^l \Sigma(r' + \omega s) ds \right\} dl.$$

Наконец, направление скорости Ω (как независимая величина в

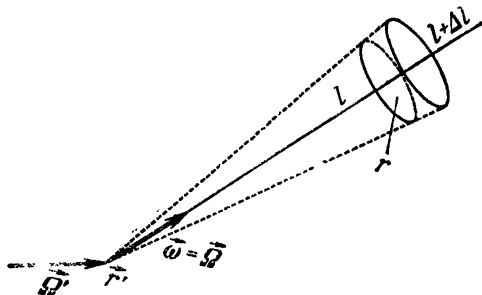


Рис. 59.

фазовом пространстве) обязана совпадать с ω и вероятность этого равна $\delta(\Omega - \omega) d\Omega$. Следовательно,

$$K_{\text{ст}}(P', P) dP = \frac{\Sigma_s(r')}{\Sigma(r')} p_s(r', \Omega'; \omega) d\omega \Sigma(r) e^{-\int_0^l \Sigma ds} dl \delta(\Omega - \omega) d\Omega.$$

Введя сюда $dP = dr d\Omega = l^2 dl d\omega d\Omega$ и сократив dP , получим формулу

$$K_{\text{ст}}(P', P) = \frac{\Sigma_s(r')}{\Sigma(r')} p_s(r', \Omega'; \omega) \delta(\Omega - \omega) \frac{\Sigma(r) e^{-\int_0^l \Sigma(r' + \omega s) ds}}{l^2}. \quad (7)$$

2.2. Моделирование свободного пробега нейтрона.

2.2.1. В однородной среде. Так как во время пробега энергия нейтрона не меняется, то полное сечение постоянно: $\Sigma(E, x) = \Sigma(E) \equiv \Sigma$. Из уравнения (6) следует

$$F(x) = 1 - e^{-\Sigma x},$$

и метод обратных функций позволяет записать явную формулу для расчета ξ

$$\xi = -(1/\Sigma) \ln \gamma. \quad (8)$$

Легко вычислить, что *средняя длина свободного пробега* $M\xi = 1/\Sigma$. Если в качестве единицы длины выбрать $1/\Sigma$, то $\xi = -\ln \gamma$ средних длин свободного пробега.

2.2.2. В кусочно однородной среде (метод обратных функций). Мы рассмотрим случай, когда интересующая нас область пространства G состоит из конечного числа однородных областей. Именно такова геометрия большинства задач, встречающихся на практике. В то же время очевидно, что всякую неоднородную среду можно с любой точностью аппроксимировать такой кусочно однородной средой.

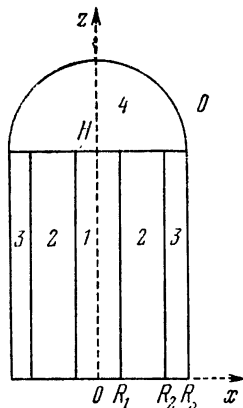


Рис. 60.

Поверхности, являющиеся границами областей, как правило, не сложнее, чем поверхности второго порядка: плоские, сферические, цилиндрические, конические. Поэтому не представляет труда найти аналитически пересечение любого луча с любой из этих поверхностей.

Рассмотрим простейший (по своей логике) алгоритм для расчета свободного пробега в такой среде, разработанный автором в 1955 г. В качестве иллюстрации выберем область G , представляющую собой трехслойный конечный цилиндр с «крышкой» в форме полусферы.

Сечение G плоскостью $y=0$ изображено на рис. 60, где цифры означают номера различных областей. В этом примере шесть граничных поверхностей:

- 1) $x^2 + y^2 = R_1^2$; 4) $z = 0$;
- 2) $x^2 + y^2 = R_2^2$; 5) $z = H$;
- 3) $x^2 + y^2 = R_3^2$; 6) $x^2 + y^2 + (z - H)^2 = R_3^2$.

Пусть нейтрон вылетает из точки r_0 по направлению единичного вектора ω . Тогда уравнение $F(\xi) = 1 - \gamma$ запишется в виде

$$I(\xi) \equiv \int_0^\xi \Sigma(r_0 + \omega s) ds = -\ln \gamma. \quad (9)$$

Чтобы решить это уравнение, находим пересечение луча $r=r_0+\omega s$ со всеми граничными поверхностями, и все положительные значения s , соответствующие точкам пересечения, располагаем в порядке возрастания:

$$0=s_{(0)}<s_{(1)}<s_{(2)}<\dots<s_{(m)}. \quad (10)$$

Вычисляем длину отрезков луча $l_{(i)}=s_{(i)}-s_{(i-1)}$, каждый из которых принадлежит одной области, и находим соответствующие этим областям значения Σ , которые обозначим через $\alpha_{(i)}$, $i=1, 2, \dots, m^*$.

Обозначим через I_k интегралы $I_k=I(s_{(k)})$. Они легко вычисляются по формуле

$$I_k = \sum_{i=1}^k \alpha_{(i)} l_{(i)}.$$

Нетрудно доказать, что если $I_k \leq -\ln \gamma < I_{k+1}$, то

$$\xi = s_{(k)} + (1/\alpha_{(k+1)}) (-\ln \gamma - I_k); \quad (11)$$

если же $-\ln \gamma \geq I_m$, то $\xi \geq s_{(m)}$ и нейтрон вылетает из области G .

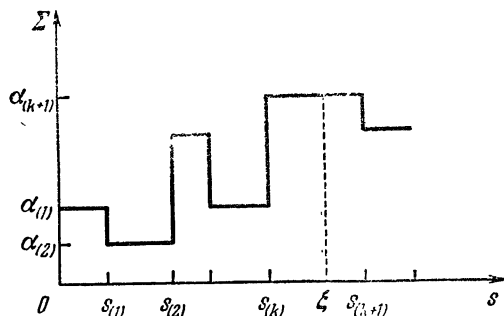


Рис. 61.

В самом деле, легко видеть, что $s_{(k)} \leq \xi < s_{(k+1)}$ тогда и только тогда, когда $I_k \leq I(\xi) < I_{k+1}$. Уравнение (9)

*) Программа расчета должна содержать блок, позволяющий определить, какой области принадлежит любая заданная точка. Выбрав середину отрезка $l_{(i)}$, можно найти номер области, которой принадлежит $l_{(i)}$.

в этом случае можно переписать в виде (рис. 61)

$$I_k + \alpha_{(k+1)} (\xi - s_{(k)}) = -\ln \gamma,$$

откуда сразу вытекает (11).

На рис. 62 изображен пример такого луча с указанием всех $s_{(i)}$. Легко заметить, однако, что пересечения $s_{(2)}$, $s_{(6)}$ и $s_{(8)}$ здесь «лишние». Чтобы исключить «лишние» пересечения, можно, например, изменить определение граничных поверхностей, добавив к ним неравенства (одно или несколько), выделяющие действительный участок границы. В примере, изображенном на рис. 62, граничные поверхности запишутся так:

- 1) $x^2 + y^2 = R_1^2$, $0 < z < H$; 4) $z = 0$, $x^2 + y^2 < R_3^2$;
- 2) $x^2 + y^2 = R_2^2$, $0 < z < H$; 5) $z = H$, $x^2 + y^2 < R_3^2$;
- 3) $x^2 + y^2 = R_3^2$, $0 < z < H$; 6) $x^2 + y^2 + (z - H)^2 = R_3^2$, $z > H$.

Определив точку пересечения луча $r = r_0 + \omega s$ с какой-нибудь из граничных поверхностей и получив положительное значение s , следует подставить координаты точки пересечения в соответствующие этой поверхности неравенства. Если координаты точки пересечения хотя бы одному из этих неравенств не удовлетворяют, то данная точка пересечения исключается (рис. 63).

Если количество граничных поверхностей в задаче велико, а фактические пробеги нейтронов малы, то такой алгоритм невыгоден:

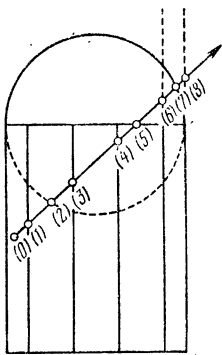


Рис. 62.

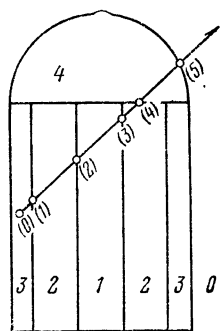


Рис. 63.

каждый раз пришлось бы вычислять очень много пересечений, хотя в действительности каждый нейтрон пересекает лишь одну-две поверхности. Для построения более экономного алгоритма в [46] предлагается разбить граничные поверхности на элементарные граничные поверхности, каждая из которых разделяет две области. В нашем примере (рис. 60) окажется 10 таких поверхностей (в скоб-

ках указаны *индексы* поверхности, т. е. номера областей, которые эта поверхность разделяет):

$$1) \ x^2 + y^2 = R_1^2, \quad 0 < z < H, \quad (1; 2);$$

$$2) \ x^2 + y^2 = R_2^2, \quad 0 < z < H, \quad (2; 3);$$

$$3) \ x^2 + y^2 = R_3^2, \quad 0 < z < H, \quad (0; 3);$$

$$4) \ z = 0, \quad x^2 + y^2 < R_1^2, \quad (0; 1);$$

$$5) \ z = 0, \quad R_1^2 < x^2 + y^2 < R_2^2, \quad (0; 2);$$

$$6) \ z = 0, \quad R_2^2 < x^2 + y^2 < R_3^2, \quad (0; 3);$$

$$7) \ z = H, \quad x^2 + y^2 < R_1^2, \quad (1; 4);$$

$$8) \ z = H, \quad R_1^2 < x^2 + y^2 < R_2^2, \quad (2; 4);$$

$$9) \ z = H, \quad R_2^2 < x^2 + y^2 < R_3^2, \quad (3; 4);$$

$$10) \ x^2 + y^2 + (z - H)^2 = R_3^2, \quad z > H, \quad (0; 4).$$

Если, например, исходная точка r_0 принадлежит области 3, то надо найти пересечения луча $r = r_0 + \omega s$ только с теми элементарными граничными поверхностями, в индексах которых фигурирует 3. Если область 3 выпуклая, то такая точка пересечения будет единственной; в общем случае выбираем точку пересечения с наименьшим s , превосходящим $s_{(0)}$, которое и назовем $s_{(1)}$. Второй индекс пересекаемой поверхности показывает, в какую область переходит нейтрон. И т. д.

В примере, изображенном на рис. 63, $s_{(1)}$ соответствует точке пересечения луча с поверхностью 2, а нейтрон переходит в область 2. Затем находим пересечения луча со всеми элементарными поверхностями, в индексе которых имеется 2, и отбираем пересечение, которому отвечает наименьшее s , превосходящее $s_{(1)}$. Таким образом, получим значение $s_{(2)}$, соответствующее первому пересечению луча с поверхностью 1, и номер следующей области — 1.

Конечно, условия $I_k \leq -\ln \gamma < I_{k+1}$, входящие в (11), следует проверять постепенно, по мере нахождения

$$s_{(1)}, I_{(1)}, \alpha_{(1)}, I_1; s_{(2)}, I_{(2)}, \alpha_{(2)}, I_2; \dots$$

2.2.3. В произвольной среде (метод постоянного сечения). В последние годы широкое распространение получил совсем другой метод моделирования пробегов в сложной среде, предложенный, по-видимому, Е. Р. Вудкоком. Кусочная однородность среды при этом не предполагается.

Выберем произвольную постоянную $\alpha \geq \sup \Sigma$, и обозначим через Σ_Φ разность $\Sigma_\Phi = \alpha - \Sigma \geq 0$. Условимся счи-

тять, что при столкновении нейтрона с ядрами, кроме реакций, входящих в Σ , возможна еще одна — *фиктивное столкновение*, при котором ни энергия, ни направление движения нейтрона не меняются. Сечение фиктивного столкновения будем считать равным Σ_{Φ} . Если (ср. пример гл. 2, п. 1.2.2) $\Sigma = \Sigma_s + \Sigma_c + \Sigma_f$, то вероятности соответствующих реакций в условной задаче равны Σ_s/α , Σ_c/α , Σ_f/α , а вероятность фиктивного столкновения равна Σ_{Φ}/α . Пробеги в этой задаче легко вычисляются по формуле (8)

$$\xi = -(1/\alpha) \ln \gamma, \quad (12)'$$

а тип столкновения разыгрывается с учетом всех четырех возможностей. Ниже доказано, что сумма таких пробегов до первого нефиктивного столкновения подчиняется тому же закону распределения (6), что истинный случайный пробег.

Теорема 1 ([112]). *Рассмотрим нейтрон, вылетающий из точки $x=0$ по направлению оси Ox и подчиняющийся законам условной задачи. Обозначим через ξ координату первого нефиктивного столкновения. Тогда функция распределения ξ выражается формулой (6)*

Доказательство. Обозначим через v случайное количество фиктивных столкновений в интервале $(0, \xi)$ и рассмотрим случай $v=i$ ($i=0, 1, 2, \dots$). Выберем произвольные числа x_1, x_2, \dots, x_{i+1} , удовлетворяющие неравенствам

$$0 < x_1 < x_2 < \dots < x_i < x_{i+1}.$$

Вероятность того, что фиктивные столкновения окажутся в окрестностях точек x_1, \dots, x_i , а первое нефиктивное столкновение — в окрестности точки x_{i+1} равна (ср. (15) гл. 5)

$$\tilde{p}(x_1, \dots, x_{i+1}) dx_1 \dots dx_{i+1} =$$

$$\begin{aligned} &= \prod_{j=1}^i \left\{ \alpha e^{-\alpha(x_j - x_{j-1})} dx_j \frac{\Sigma_{\Phi}(x_j)}{\alpha} \right\} \times \\ &\times \alpha e^{-\alpha(x_{i+1} - x_i)} \left[1 - \frac{\Sigma_{\Phi}(x_{i+1})}{\alpha} \right] dx_{i+1} = \\ &= e^{-\alpha x_{i+1}} \Sigma_{\Phi}(x_1) \dots \Sigma_{\Phi}(x_i) \left[\alpha - \Sigma_{\Phi}(x_{i+1}) \right] dx_1 \dots dx_{i+1}. \end{aligned}$$

Проинтегрировав эту вероятность по всем возможным x_1, \dots, x_{i+1}

таким, что $x_{i+1} < x$, получим вероятность

$$\begin{aligned} P\{\xi < x, \nu = i\} &= \int_0^x dx_{i+1} \int_0^{x_{i+1}} dx_i \dots \int_0^{x_i} \tilde{p}(x_1, \dots, x_{i+1}) dx_1 = \\ &= \int_0^x e^{-\alpha x_{i+1}} [\alpha - \sum_{\Phi}(x_{i+1})] dx_{i+1} \int_0^{x_{i+1}} \sum_{\Phi}(x_i) dx_i \dots \\ &\dots \int_0^{x_i} \sum_{\Phi}(x_1) dx_1. \end{aligned}$$

Все внутренние интегралы легко вычисляются. Если ввести обозначение $I_{\Phi}(x) = \int_0^x \sum_{\Phi}(s) ds$, то (заменив x_{i+1} на y) можно будет записать результат в виде

$$P\{\xi < x, \nu = i\} = \int_0^x e^{-\alpha y} [\alpha - \sum_{\Phi}(y)] \frac{[I_{\Phi}(y)]^i}{i!} dy.$$

Отсюда следует, что

$$\begin{aligned} P\{\xi < x\} &= \sum_{i=0}^{\infty} P\{\xi < x, \nu = i\} = \int_0^x e^{-\alpha y + I_{\Phi}(y)} [\alpha - \sum_{\Phi}(y)] dy = \\ &= 1 - e^{-\alpha x + I_{\Phi}(x)} = 1 - \exp\left\{-\int_0^x \sum(s) ds\right\}, \end{aligned}$$

что и требовалось доказать.

Так как при увеличении α количество фиктивных столкновений возрастает, то обычно стараются выбрать минимально возможное значение, т. е. $\alpha = \sup \Sigma$. Некоторые другие применения метода постоянного сечения указаны в работах [152, 169].

2.3. Моделирование свободного пробега заряженной частицы. Рассмотрим пробег быстрой заряженной частицы в плазме, состав и температуры которой (как функции времени t и координаты r) заданы. Физические ограничения на плазму: а) не слишком большое разрежение; б) ларморовский радиус значительно больше длины свободного пробега, так что траекторию частицы можно считать прямолинейной: $r = r_0 + \omega s$.

В результате взаимодействия с электронами и ионами среды частица теряет свою энергию (тормозится). Этот процесс можно считать непрерывным, так что вдоль траектории

$$dE/dt = -q(E, r, t).$$

Формулы для расчета q предполагаются заданными.

Полное сечение Σ заряженной частицы зависит от ее энергии и положения, так что в конечном счете $\Sigma = \Sigma(E, r, t)$. Скорость частицы $v = ds/dt$ связана с ее энергией соотношением $E = (1/2)Mv^2$, где M — масса частицы.

Для того чтобы вычислить значения функции

$$I(t) = \int_0^{s(t)} \sum ds,$$

можно численно интегрировать систему обыкновенных дифференциальных уравнений

$$\frac{dI}{dt} = v \sum, \quad \frac{dE}{dt} = -q, \quad \frac{ds}{dt} = v \quad (13)$$

с начальными условиями $I(t_0)=0$, $E(t_0)=E_0$, $s(t_0)=0$. Если при $t=t_k$ значение $I_k = I(t_k) \leq -\ln \gamma$, а при $t=t_{k+1}$ значение $I_{k+1} = I(t_{k+1}) > -\ln \gamma$, то значения всех величин в момент столкновения $-t=t_\xi$, $E=E_\xi$ и $s=\xi$ — можно проинтерполировать:

$$t_\xi = t_k + z(t_{k+1} - t_k), \quad E_\xi = E_k - z(E_k - E_{k+1}),$$

$$\xi = s_k + z(s_{k+1} - s_k), \quad \text{где } z = (-\ln \gamma - I_k)/(I_{k+1} - I_k).$$

Возможен также случай, когда при $t=t_k$ значение I_k все еще не превосходит $-\ln \gamma$, а частица вылетает из области G или ее энергия оказывается ниже интересующего нас уровня $E_k < E_{\min}$. Тогда частица из дальнейшего рассмотрения исключается. При расчете некоторых задач оказалось удобным выбрать в качестве независимой переменной в (13) энергию E и численно интегрировать уравнения

$$\frac{dt}{dE} = -\frac{1}{q}, \quad \frac{ds}{dE} = -\frac{v}{q}, \quad \frac{dI}{dE} = -\frac{v\Sigma}{q}$$

при $E_0 \geq E \geq E_{\min}$ с начальными условиями $t(E_0)=t_0$, $s(E_0)=0$, $I(E_0)=0$.

Для моделирования пробегов заряженных частиц также можно использовать метод постоянного сечения (п. 2.2.3), если только в интересующем нас диапазоне величин $\sup \Sigma = \alpha < \infty$. Тогда можно длину свободного пробега разыгрывать по той же формуле (12); затем придется интегрировать два уравнения

$$\frac{dE}{ds} = -\frac{q}{v}, \quad \frac{dt}{ds} = \frac{1}{v}$$

от $s=0$ до $s=\xi$ (начальные значения $E(0)=E_0$, $t(0)=t_0$) и, только вычислив E_ξ , t_ξ и $\sum(E_\xi, r_\xi, t_\xi)$, можно будет определить (разыграть), фиктивное ли это столкновение или нет.

З а м е ч а н и е. Изложенный метод разработан в [81]. Иногда взаимодействие заряженной частицы с электронами и ионами среды целесообразно учитывать двояко: взаимодействия, приводящие к небольшим изменениям энергии частицы, осредняются и включаются в непрерывное торможение q , а сравнительно редкие столкновения, влекущие за собой значительное изменение энергии («катастрофические столкновения»), включаются в Σ и разыгрываются [39, 104].

§ 3. Использование статистических весов

В гл. 5 мы встречались с величинами W_j или \tilde{W}_j , которые называли весами. Однако при решении многих физических задач веса можно вводить, руководствуясь чисто физическими соображениями, отправляясь при этом от естественного процесса и не пользуясь макроскопическими уравнениями (например, уравнением переноса). Нередко использование весов заметно повышает эффективность расчета.

Мы рассмотрим несколько способов введения весов на примере задачи о поглощении нейтронов (п. 1.1).

3.1. Веса, заменяющие розыгрыш поглощения. Предположим, что из источника r_0 в направлении Ω_0 вылетел не один нейтрон, а «пакет», состоящий из большого числа w_0 идентичных нейтронов. Разыграв длину пробега ξ_0 , определим точку столкновения для всего пакета $r_1 = r_0 + \xi_0 \Omega_0$. В среднем при таком столкновении $[\Sigma_a(r_1)/\Sigma(r_1)] w_0$ нейтронов поглощаются, а $[\Sigma_s(r_1)/\Sigma(r_1)] w_0$ нейтронов рассеиваются. Поэтому, разыграв (в соответствии с индикатрисой рассеяния) новое направление движения пакета Ω_1 будем считать, что в этом направлении движется пакет, состоящий из

$$w_1 = [\Sigma_s(r_1)/\Sigma(r_1)] w_0$$

нейтронов.

Правила построения траектории оказываются во многом такими же, как в п. 1.1: так же разыгрываются пробеги ξ_i , функции распределения которых в соответствии с (6) равны

$$F_i(x) = 1 - \exp \left\{ - \int_0^x \Sigma(r_i + \Omega_i s) ds \right\}, \quad (14)$$

так же разыгрываются направления Ω_i . Однако при столкновении в точке $r_{i+1} = r_i + \xi_i \Omega_i$ «судьба» нейтрона не разыгрывается: вместо этого предполагается, что $w_i [\Sigma_a(r_{i+1})/\Sigma(r_{i+1})]$ нейтронов из пакета поглотились, а в рассеянном пакете остаются лишь

$$w_{i+1} = [\Sigma_s(r_{i+1})/\Sigma(r_{i+1})] w_i \quad (15)$$

нейтронов. История пакета заканчивается тогда, когда

он вылетает из области G_0 . Количество поглощенных за всю историю нейтронов равно

$$\eta_A^{(1)} = \sum_{i=0}^{v-1} [\sum_a (r_{i+1}) / \sum (r_{i+1})] \omega_i, \quad (16)$$

где v — номер последней точки траектории внутри G_0 (другими словами, $r_{v+1} \notin G_0$).

Наконец, нетрудно заметить, что величины (15) и (16) пропорциональны ω_0 . Поэтому, несмотря на рассуждения о «большом количестве» ω_0 можем считать, что

$$\omega_0 = 1. \quad (17)$$

Тогда величина $\eta_A^{(1)}$ — количество поглощений в расчете на один испущенный нейтрон — окажется оценкой искомой вероятности: $M\eta_A^{(1)} = p_A$. (Формальное доказательство имеется в § 4).

Для приближенного расчета p_A нужно реализовать достаточно большое число N траекторий указанного вида (с $\omega_0 = 1$) и положить

$$p_A \approx \frac{1}{N} \sum_{s=1}^N (\eta_A^{(1)})_s,$$

где $(\eta_A^{(1)})_s$ — значение $\eta_A^{(1)}$, полученное на траектории номер s .

Докажем теперь, что оценка $\eta_A^{(1)}$ всегда не хуже оценки η_A , используемой при имитации поведения нейтронов (п. 1.1):

$$D\eta_A^{(1)} \leq D\eta_A. \quad (18)$$

Согласно лемме п. 1.1. для этого достаточно доказать, что $0 \leq \eta_A^{(1)} \leq 1$.

Введем для краткости обозначение

$$s_i = \sum_s (r_i) / \sum (r_i). \quad (19)$$

Из формул (15), (16) и (17) вытекает, что при $i \geq 1$

$$\begin{aligned} \omega_i &= s_1 s_2 \dots s_i, \\ \eta_A^{(1)} &= \sum_{i=0}^{v-1} s_1 s_2 \dots s_i (1 - s_{i+1}). \end{aligned} \quad (20)$$

Последнее выражение легко преобразовать к виду

$$\eta_A^{(1)} = 1 - s_1 s_2 \dots s_v, \quad (21)$$

откуда сразу следует требуемое неравенство.

3.1.1. В большинстве реальных задач дисперсия $D\eta_A^{(1)}$ заметно меньше, чем $D\eta_A$, однако общих оценок на этот счет мало. Обратимся к частному случаю — однородной области G_0 , когда

$$\sum_s (r) / \sum (r) \equiv s > 0.$$

Обозначим через p_i вероятность того, что траектория пакета закончится в точке r_{i+1} с номером $i+1$ или, другими словами, $p_i = P\{v=i\}$. Вычислять эти вероятности нам не потребуется. Заметим только, что p_0 равно вероятности того, что нейтрон, испущенный источником, вылетит из области G_0 , не испытав ни одного столкновения.

Теорема 2. *Рассмотрим задачу о поглощении нейтронов в однородной области G_0 (п. 1.1). Если*

$$p_0 \leq (1 - p_A)^2, \quad (22)$$

то имеет место неравенство

$$D\eta_A^{(1)} \leq s D\eta_A. \quad (23)$$

Доказательство. Из формулы (21) видно, что в однородной области случайная величина $\eta_A^{(1)}$ может принимать только значения $1 - s^l$, где $l=0, 1, 2, \dots$. Следовательно, ее распределение задается таблицей

$$\begin{pmatrix} 0 & 1-s & 1-s^2 & \dots & 1-s^l & \dots \\ p_0 & p_1 & p_2 & \dots & p_l & \dots \end{pmatrix}.$$

Математическое ожидание этой величины равно

$$M\eta_A^{(1)} = \sum_{l=1}^{\infty} (1-s^l) p_l = \sum_{l=1}^{\infty} p_l - \sum_{l=1}^{\infty} p_l s^l = 1 - p_0 - \sum_{l=1}^{\infty} p_l s^l.$$

Так как $M\eta_A^{(1)} = p_A$, то получаем выражение

$$p_A = 1 - p_0 - \sum_{l=1}^{\infty} p_l s^l. \quad (24)$$

Дисперсию $\eta_A^{(1)}$ запишем в форме

$$D\eta_A^{(1)} = \sum_{l=1}^{\infty} (1-s^l)^2 p_l - p_A^2 = \sum_{l=1}^{\infty} p_l - 2 \sum_{l=1}^{\infty} p_l s^l + \sum_{l=1}^{\infty} p_l s^{2l} - p_A^2.$$

Так как $s^i \leq s$ при всех $i \geq 1$, то отсюда следует, что

$$D\eta_A^{(1)} \leq 1 - p_0 - 2 \sum_{i=1}^{\infty} p_i s^i + s \sum_{i=1}^{\infty} p_i s^i - p_A^2.$$

Входящую сюда сумму исключим с помощью формулы (24). Получим неравенство

$$D\eta_A^{(1)} \leq 1 - p_A^2 - (2 - s)(1 - p_A) + (1 - s)p_0.$$

Наконец, воспользуемся условием (22) теоремы, после чего последнее неравенство превратится в (23):

$$D\eta_A^{(1)} < 1 - p_A^2 - (2 - s)(1 - p_A) + (1 - s)(1 - p_A)^2 = s(p_A - p_A^2).$$

3.2. Веса, учитывающие вылет из области G_0 . Снова рассмотрим задачу о поглощении нейтронов из п. 1.1. Пусть из точки r_i в направлении Ω_i вылетает пакет,

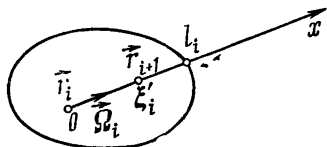


Рис. 64.

состоящий из большого числа w_i идентичных нейтронов. Обозначим через l_i расстояние от точки r_i до границы области G_0 (по направлению полета, рис. 64). Обозначим через $F_i(x)$ функцию распределения (14) длины свободного пробега ξ для

одного нейтрона из пакета. Вероятность того, что нейтрон этот вылетит из области G_0 , равна

$$P\{\xi > l_i\} = 1 - F_i(l_i).$$

В среднем из области G_0 вылетят $w_i[1 - F_i(l_i)]$ нейтронов пакета, а $w_i F_i(l_i)$ нейтронов испытают столкновение внутри G_0 . Будем считать, что в следующую точку столкновения r_{i+1} прилетит пакет, содержащий

$$v_{i+1} = w_i F_i(l_i) \quad (25)$$

нейтронов. Тогда свободный пробег ξ' для того пакета надо разыгрывать *внутри* G_0 . Это значит, что величина ξ' подчиняется усеченному распределению (14) на интервале $0 < x < l_i$ (см. п. 5.2 гл. 2). Функция распределения ξ' равна

$$F(x) = F_i(x) / F_i(l_i).$$

Метод обратных функций позволяет записать уравнение для расчета $\xi' = \xi'_i$ в виде

$$F_i(\xi') = (1-\gamma)F_i(l_i) \quad (26)'$$

или, по аналогии с (9), в виде

$$I(\xi') \equiv \int_0^{\xi'} \sum (r_i + \Omega_i s) ds = -\ln [\gamma + (1-\gamma)e^{-I(\bar{l}_i)}].$$

Определив точку столкновения пакета $r_{i+1} = r_i + \Omega_i \xi'_i$, разыгрываем (обычным способом), рассеялся ли пакет или поглотился? Если он рассеялся, то количество нейтронов в пакете после рассеяния равно

$$w_{i+1} = v_{i+1}. \quad (27)'$$

Если он поглотился, то количество поглощенных нейтронов $\eta_A^{(2)} = v_{i+1}$. Очевидно, история пакета не может закончиться вылетом и продолжается до его поглощения в некоторой точке r_{v+1} .

Если положить $w_0 = 1$, то из (25) и (27) следует, что

$$w_i = F_0(l_0)F_1(l_1) \dots F_{i-1}(l_{i-1}), \quad (28),$$

а случайная величина

$$\eta_A^{(2)} = w_{v+1}. \quad (29)$$

В этом случае $\eta_A^{(2)}$ равно количеству поглощенных нейтронов, приходящихся на один нейтрон источника, и служит оценкой для искомой вероятности p_A . Формальное доказательство равенства $M\eta_A^{(2)} = p_A$ имеется в § 4. Расчетная формула такого метода Монте-Карло:

$$p_A \approx \frac{1}{N} \sum_{s=1}^N (\eta_A^{(2)})_s,$$

где $(\eta_A^{(2)})_s$ — значение $\eta_A^{(2)}$, полученное по траектории номер s .

Из формул (28) и (29) видно, что $\eta_A^{(2)} \leq 1$ и (по лемме п. 1.1) дисперсия

$$D\eta_A^{(2)} \leq D\eta_A.$$

Таким образом, точность оценки $\eta_A^{(2)}$ всегда не хуже, чем точность оценки η_A , получаемой при имитации поведения

нейтронов. Правда, расчет величины $\eta_A^{(2)}$ может оказаться более трудоемким, чем расчет η_A , из-за необходимости вычислять l_i .

3.3. Веса, заменяющие розыгрыш поглощения и учитывающие вылет из области G_0 . Следующий способ расчета p_A объединяет особенности двух предыдущих методов. Пусть из точки r_i в направлении Ω_i вылетает пакет, состоящий из большого числа w_i идентичных нейтронов. Если (как в п. 3.2) $w_i[1-F_i(l_i)]$ нейтронов пакета вылетают из области G_0 , то в следующую точку столкновения $r_{i+1}=r_i+\xi'_i \Omega_i$ прилетит всего $v_{i+1}=w_i F_i(l_i)$ нейтронов (длина пробега ξ'_i определяется из уравнения (26)).

А теперь, как в п. 3.1, будем считать, что

$$v_{i+1} [\sum_a (r_{i+1}) / \sum (r_{i+1})]$$

нейтронов при столкновении поглотились, а

$$w_{i+1} = v_{i+1} [\sum_s (r_{i+1}) / \sum (r_{i+1})]$$

нейтронов рассеялись и продолжают полет по направлению Ω_{i+1} . При таком способе расчета все траектории оказываются бесконечными и не могут закончиться ни вылетом, ни поглощением. На практике расчет продолжают до тех пор, пока вес пакета w_i не станет пренебрежимо малым, например меньше некоторого заданного числа ε .

Если использовать обозначение (19), то, выбрав $w_0=1$, получим, что

$$w_i = F_0(l_0) s_1 F_1(l_1) s_2 \dots F_{i-1}(l_{i-1}) s_i, \quad (30)$$

а количество поглощенных за всю историю пакета нейтронов равно

$$\eta_A^{(3)} = \sum_{i=0}^{\infty} v_{i+1} (1 - s_{i+1}). \quad (31)$$

Формальное доказательство того, что $M\eta_A^{(3)} = p_A$, имеется в § 4. Формула для оценки p_A , очевидно, запишется так:

$$p_A \approx \frac{1}{N} \sum_{s=1}^N (\eta_A^{(3)})_s,$$

где $(\eta_A^{(3)})_s$ — значение $\eta_A^{(3)}$ на траектории номер s .

Из формулы (31) видно, что $\eta_A^{(3)} \geq 0$. Чтобы доказать, что $\eta_A^{(3)} \leq 1$, представим эту величину в виде

$$\eta_A^{(3)} = \\ = F_0(l_0) - \sum_{i=0}^{\infty} F_0(l_0) \dots F_i(l_i) [1 - F_{i+1}(l_{i+1})] s_1 \dots s_{i+1};$$

тогда, очевидно, $\eta_A^{(3)} \leq F_0(l_0) \leq 1$ и по лемме п. 1.1 дисперсия $D\eta_A^{(3)} \leq D\eta_A$.

Доказательство. Формулу (31) запишем в виде

$$\eta_A^{(3)} = \sum_{i=0}^{\infty} F_0 \dots F_i s_1 \dots s_i - \sum_{i=0}^{\infty} F_0 \dots F_i s_1 \dots s_i s_{i+1}.$$

В первой сумме выделим первое слагаемое, а в остальных слагаемых заменим индекс i на $i+1$:

$$\begin{aligned} \sum_{i=0}^{\infty} F_0 \dots F_i s_1 \dots s_i &= F_0 + \sum_{i=1}^{\infty} F_0 \dots F_i s_1 \dots s_i = \\ &= F_0 + \sum_{i=0}^{\infty} F_0 \dots F_{i+1} s_1 \dots s_{i+1}. \end{aligned}$$

Подставив это выражение в предыдущее равенство, получим

$$\begin{aligned} \eta_A^{(3)} &= F_0 + \sum_{i=0}^{\infty} F_0 \dots F_{i+1} s_1 \dots s_{i+1} - \sum_{i=0}^{\infty} F_0 \dots F_i s_1 \dots s_{i+1} = \\ &= F_0 - \sum_{i=0}^{\infty} F_0 \dots F_i (1 - F_{i+1}) s_1 \dots s_{i+1}, \end{aligned}$$

что и требовалось доказать.

3.3.1. На рис. 65 приведены схемы расчета одного звена траектории во всех четырех методах оценки ρ_A . Номер (0) соответствует методу п. 1.1, номера (1), (2) и (3) — методам пп. 3.1, 3.2 и 3.3. Нетрудно заметить, что основное различие в количестве вычислений связано с тем, что в методах (0) и (1) нужно проверять условие « $r_{i+1} \in G_0$ », а в методах (2) и (3) приходится вычислять l_i и $F_i(l_i)$. Кроме того, в методах (0) и (2) на расчет одного звена траектории затрачиваются 4 случайных числа γ , а в методах (1) и (3) — только 3. Остальные различия сводятся к нескольким элементарным арифметическим операциям.

Некоторые пояснения к схемам на рис. 65. По условию задачи точка r_0 задана; в методах (1), (2) и (3) заданы также $w_0=1$, а в методах (1) и (3), кроме того, полагаем $h_0=0$. Величина h_i — это суммарное количество поглощенных в точках r_1, \dots, r_i нейтронов.

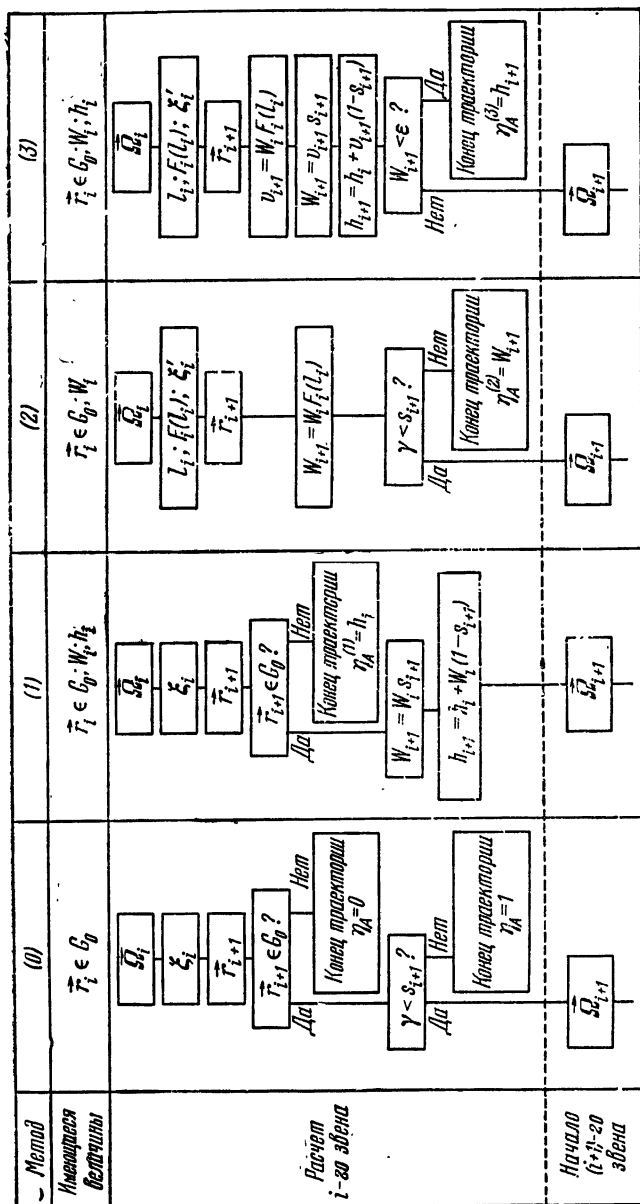


Рис. 65.

Во всех методах Ω_0 выбирается случайно (формула (15) гл. 2); Ω_i разыгрывается по плотности $p_s(r_i, \Omega_{i-1}; \Omega)$; величина $s_i = \sum_s (r_i) / \sum (r_i)$.

В методах (0) и (1) пробег ξ_i определяется из уравнения $F_i(\xi_i) = \gamma$, а $r_{i+1} = r_i + \xi_i \Omega_i$. В методах (2) и (3) пробег ξ'_i определяется из уравнения $F_i(\xi'_i) = \gamma F_i(l_i)$, а $r_{i+1} = r_i + \xi'_i \Omega_i$ (l_i — расстояние от точки r_i до границы G_0 по направлению Ω_i).

З а м е ч а н и е. Из рис. 65 видно, что количество вычислений в методе (1) не больше, чем в методе (0). Поэтому из (18) вытекает, что для рассматриваемой задачи алгоритм, использующий оценку $\eta_A^{(1)}$, всегда более экономичен, чем алгоритм, основанный на имитации процесса, использующий оценку η_A .

3.4. Различные статистические веса. Величины w_i , использованные в пп. 3.1, 3.2 и 3.3, называют *весами* (или *статистическими весами*). И вместо того чтобы говорить о пакете, содержащем w_i нейтронов (число w_i обычно дробное!), говорят о нейтроне с весом w_i . При движении нейтрона по траектории вес его меняется.

3.4.1. Расщепление траекторий. Используемые выше рассуждения (с пакетом идентичных нейтронов) показывают, что после очередного столкновения в точке r_i вместо одного нейтрона с весом w_i можно рассматривать m идентичных нейтронов, вес каждого из которых равен w_i/m . Таким образом, траектория нейтрона после точки r_i как бы расщепляется на m траекторий, каждая из которых дальше строится независимо.

Этот прием используют для лучшего учета прохождения нейтронов в некоторых «важных» областях.

3.4.2. Случайный обрыв траекторий. Этот прием в каком-то смысле противоположен предыдущему: после очередного столкновения в точке r_i для нейтрона с весом w_i можно разыграть дополнительную альтернативу: либо с вероятностью p вес нейтрона увеличивается и становится равным w_i/p , либо с вероятностью $1-p$ нейтрон «гибнет», и траектория заканчивается. Действительно, в среднем при такой процедуре вес нейтрона остается прежним:

$$(w_i/p) \cdot p + 0 \cdot (1-p) = w_i.$$

Случайный обрыв траекторий используют в тех случаях, когда вес нейтрона становится весьма малым и продолжать рассчитывать его траекторию «невыгодно»:

вклад от всех последующих столкновений сравнительно невелик, а вычислений много; и в то же время пренебречь таким весом нельзя.

3.4.3. Систематическая выборка. Рассмотрим источник нейтронов с энергией E_0 и равновероятными направлениями скоростей, распределенный в области G_0 с плотностью $p(r)$. Требуется промоделировать N нейтронов из этого источника.

Вместо того чтобы независимо разыгрывать начальные положения всех этих нейтронов в соответствии с плотностью $p(r)$, можно разбить G_0 на m непересекающихся областей

$$G_0 = G_{01} + G_{02} + \dots + G_{0m}$$

и выбрать в них по N_1, N_2, \dots, N_m начальных положений (сумма $N_1 + \dots + N_m$ не обязана равняться N). При этом нейтронам, у которых $r_0 \in G_{0k}$, придется приписать начальные веса

$$w_{0k} = (N/N_k) \int_{G_{0k}} p(r) dr, \quad (32)$$

так, чтобы сумма начальных весов всех $N_1 + \dots + N_m$ нейтронов равнялась

$$N_1 w_{01} + \dots + N_m w_{0m} = N.$$

Заметим, что начальные направления для каждого из этих нейтронов можно выбирать независимо, а плотность распределения r_0 внутри G_{0k} равна $p(r) / \int_{G_{0k}} p(r) dr$.

3.5. Статистические веса и существенная выборка. Обоснование статистических весов в простейших задачах может быть получено при помощи метода существенной выборки (гл. 3, п. 3.2). В самом деле, пусть требуется вычислить математическое ожидание $Mf(Q)$, где случайная точка $Q = (\xi_1, \dots, \xi_n)$ имеет плотность $p_Q(x_1, \dots, x_n)$, определенную во всем n -мерном пространстве. Можно ввести любую другую случайную точку $H = (\eta_1, \dots, \eta_n)$ с плотностью $p_H(x_1, \dots, x_n)$, допустимой по отношению к произведению $f(x, \dots, x_n) p_Q(x_1, \dots, x_n)$, и

заменить расчет $Mf(Q)$ расчетом $M[f(H)w(H)]$, где

$$w(H) = p_Q(H)/p_H(H), \quad (33)$$

ибо

$$\begin{aligned} M[f(H)w(H)] &= \int f(P)w(P)p_H(P)dP = \\ &= \int f(P)p_Q(P)dP = Mf(Q); \end{aligned}$$

здесь $P = (x_1, \dots, x_n)$. Если координаты η_1, \dots, η_n моделируются последовательно, с использованием условных плотностей (п. 2.2 гл. 2), то из тождеств

$$p_Q(x_1, \dots, x_n) = p_{\xi_1}(x_1)p_{\xi_2}(x_2|x_1)\dots p_{\xi_n}(x_n|x_1, \dots, x_{n-1})$$

и $p_H(x_1, \dots, x_n) = p_{\eta_1}(x_1)p_{\eta_2}(x_2|x_1)\dots p_{\eta_n}(x_n|x_1, \dots, x_{n-1})$ вытекает, что «вес» равен

$$w = \frac{p_Q(H)}{p_H(H)} = \frac{p_{\xi_1}(\eta_1)}{p_{\eta_1}(\eta_1)} \cdot \frac{p_{\xi_2}(\eta_2|\eta_1)}{p_{\eta_2}(\eta_2|\eta_1)} \dots \frac{p_{\xi_n}(\eta_n|\eta_1, \dots, \eta_{n-1})}{p_{\eta_n}(\eta_n|\eta_1, \dots, \eta_{n-1})}.$$

Вычислять вес можно рекуррентно, используя формулы

$$\begin{aligned} w_0 &= 1; \quad w_i = w_{i-1} \frac{p_{\xi_i}(\eta_i|\eta_1, \dots, \eta_{i-1})}{p_{\eta_i}(\eta_i|\eta_1, \dots, \eta_{i-1})}, \quad 1 \leq i \leq n; \\ w &= w_n. \end{aligned}$$

Именно так вводились в гл. 5 величины W_i , которые там назывались весами.

Обычно на практике считают, что *введение весов приводит к уменьшению дисперсии, если $w < 1$* . Эта рекомендация требует пояснения, так как во всем пространстве неравенство $w < 1$ невозможно: оно равносильно неравенству $p_Q < p_H$, которое противоречит требованию нормированности плотностей. В действительности желательно, чтобы неравенство $w < 1$ выполнялось в той части B пространства, которая вносит основной вклад в интегралы $Mf^2(Q)$ и $M[f^2(H)w^2(H)]$; тогда

$$\begin{aligned} Mf^2(Q) &= \int f^2(P)p_Q(P)dP \approx \int_B f^2(P)p_Q(P)dP, \\ M[f^2(H)w^2(H)] &= \int f^2(P)w^2(P)p_H(P)dP = \\ &= \int f^2(P)w(P)p_Q(P)dP \approx \int_B f^2(P)w(P)p_Q(P)dP < \\ &< \int_B f^2(P)p_Q(P)dP, \end{aligned}$$

и, следовательно, $D[f(H)w(H)] < Df(Q)$.

Во многих практических задачах выполнение этого условия гарантируется тем, что $f(P) \neq 0$ только в некоторой части фазового пространства, и как раз в этой части $\omega(P) < 1$. Например, в задаче п. 1.1 распределение нейтронов, вылетевших из области G_0 , на величину p_A не влияет. Введя веса, учитывающие вылет из области (п. 3.2), мы исключаем моделирование нейтронов вне G_0 и получаем метод расчета, в котором всегда $\omega < 1$.

3.6. Метод подобных траекторий. Нередко приходится рассчитывать методами Монте-Карло серии геометрически подобных задач. В такой ситуации можно ограничиться моделированием случайных траекторий лишь для одной из этих задач, а траектории для всех других задач получать преобразованием подобия. Этот метод был предложен К. Мортонем [161], который, моделируя прохождение нейтронов через однородную пластинку толщины $h=1$, вычислял вероятности прохождения $p(h)$ для целой серии однородных пластинок и даже оценил производную dp/dh при $h=0$. Дальнейшее развитие метода имеется в [8, 95]. Обоснование его следует из § 4 гл. 3.

В качестве примера рассмотрим задачу о поглощении нейтронов из п. 1.1. Предположим, что требуется вычислить вероятности поглощения $p_A(\lambda)$ для серии однородных областей $G_0(\lambda)$ с центром подобия в источнике r_0 (рис. 66), λ — коэффициент подобия. Расчет траекторий T в $G_0(1)$ будем осуществлять методом п. 1.1 или 3.1., а траектории T' в $G_0(\lambda)$ будем считать подобными T .

Нетрудно заметить, что направления звеньев траекторий T' и судьбы нейтронов при столкновениях в точках этих траекторий окажутся разыгранными правильно, так как законы рассеяния и поглощения во всех точках одинаковы*). Однако длины звеньев траекторий T' будут разыграны неверно.

В самом деле, длина свободного пробега ξ' для траектории T' должна выбираться в соответствии с плотностью

$$p_{\xi}(x) = \Sigma e^{-\Sigma x}.$$

*) Метод подобных траекторий применим также в задачах, в которых учитываются потери энергии нейтронов при столкновениях: потери в точках траекторий T' будут такими же, как в соответствующих точках траекторий T .

Вместо этого мы полагаем ξ' равным $\lambda\xi$, где пробег ξ для траектории T имеет ту же плотность $p_\xi(x)$, так что плотность $\lambda\xi$ равна

$$p_{\lambda\xi}(x) = (1/\lambda) p_\xi(x/\lambda) = (\Sigma/\lambda) e^{-(\Sigma/\lambda)x}.$$

Согласно (33) этот произвол надо компенсировать весовым множителем

$$p_\xi(\lambda\xi)/p_{\lambda\xi}(\lambda\xi) = \lambda e^{(1-\lambda)\Sigma\xi}.$$

Поэтому, если на траектории T' из точки r'_i по направлению Ω_i вылетел нейтрон с весом w_i , то надо считать,

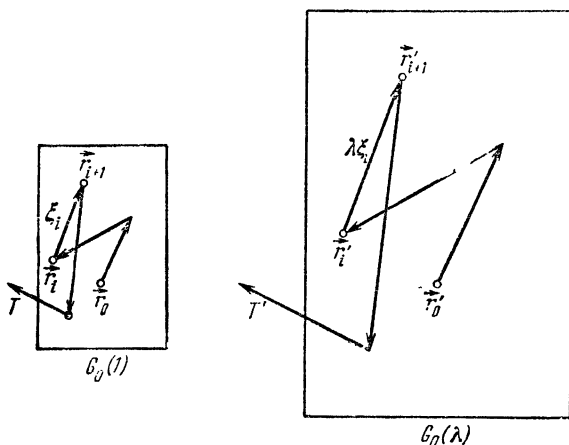


Рис 66.

что в точку $r'_{i+1} = r'_i + \Omega_i \lambda \xi_i$ прилетит нейтрон с весом

$$w'_{i+1} = w'_i \lambda e^{(1-\lambda)\Sigma\xi_i}.$$

Если расчет ведется методом п. 3.1, то после учета поглощения в точке r'_{i+1} останется нейтрон с весом

$$w'_{i+1} = w'_i (\Sigma_s/\Sigma) \lambda e^{(1-\lambda)\Sigma\xi_i}. \quad (34)$$

Для того чтобы выразить веса w'_i через веса w_i в области $G_0(1)$, разделим (34) на (15):

$$w'_{i+1}/w_{i+1} = (w'_i/w_i) \lambda e^{(1-\lambda)\Sigma\xi_i}.$$

Выбрав $\omega'_0 = \omega_0 = 1$, получим формулу

$$\omega'_i = \omega_i \lambda^i e^{(1-\lambda)\Sigma(\xi_0 + \dots + \xi_{i-1})}.$$

(Легко доказать, что последняя формула справедлива также в случае метода п. 1.1.)

Согласно (16) количество поглощенных на траектории T' в $G_0(\lambda)$ нейтронов равно

$$\eta_A^{(1)}(\lambda) = \frac{\Sigma_a}{\Sigma} \sum_{i=0}^{v-1} \omega_i \lambda^i e^{(1-\lambda)\Sigma(\xi_0 + \dots + \xi_{i-1})}.$$

3.7. Векторные веса. При расчете многих задач нейтронной физики (в частности, связанных с ядерными реакторами) вместо одногруппового приближения используют более точное многогрупповое приближение [25, 51, 53]: предполагается, что энергия каждого нейтрона может принимать конечное число значений $E_1 > E_2 > \dots > E_m$. Таким образом, в каждый момент нейтроны оказываются распределенными на m групп. При столкновении нейтрона с ядром атома среды возможно как «собственно» рассеяние, при котором нейтрон остается в той же группе, так и замедление, когда нейтрон из группы j переходит в группу k , причем $k > j$.

Рассмотрим задачу о поглощении нейтронов (п. 1.1) в многогрупповом приближении. Пусть заданы числа p_i — вероятности того, что нейтрон, испущенный источником, принадлежит группе номер j . Задана матрица сечений рассеяния

$$S = \begin{pmatrix} \Sigma_1^1 & 0 & \dots & 0 \\ \Sigma_2^1 & \Sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_m^1 & \Sigma_m^2 & \dots & \Sigma_m^m \end{pmatrix}$$

и полные сечения для нейтронов всех групп

$$\Sigma^j = \Sigma_j^j + \Sigma_{j+1}^j + \dots + \Sigma_m^j + \Sigma_a^j,$$

так что вероятность того, что нейтрон группы j при столкновении перейдет в группу k , равна Σ_k^j / Σ^j , а вероятность того, что он поглотится, равна Σ_a^j / Σ^j . Для простоты будем считать, что направления рассеянных и замедленных нейтронов (так же, как и нейтронов источника) распределены равномерно по пространству. Требуется вычислить вероятность поглощения p_A в области G_0 для одного нейтрона источника.

Легко видеть, что метод п. 1.1 без труда переносится на такую задачу: сперва разыгрывается энергия испущенного нейтрона, а затем прослеживается его траектория (до поглощения или до вылета из области G_0); конечно, при каждом столкновении нейтрон может

перейти в другую энергетическую группу. Методы пп. 3.1—3.3 также применимы *).

В работе [80] использованы статистические веса более сложного вида, и строится одна траектория для нейтронов всех групп. Предположим, что вдоль траектории движется «большой пакет», содержащий нейтроны всех групп. Пусть вектор $w(i) = \{w_1(i), \dots, w_m(i)\}$ описывает состав пакета после столкновения в точке r_i , так что $w_j(i)$ — количество нейтронов, принадлежащих группе j .

Воспользуемся весами, заменяющими розыгрыш поглощения и учитывающими вылет из области G_0 (п. 3.3). Фиксируем номер одной из групп $j=j_0$ и условимся свободный пробег ξ' пакета разыгрывать по закону (26) для этой группы, так что плотность ξ' равна

$$p_{\xi'}(x) = \sum^{j_0} e^{-\Sigma^{j_0} x} \left(1 - e^{-\Sigma^j l_i} \right), \quad 0 < x < l_i,$$

где l_i — расстояние от точки r_i до границы G_0 по направлению полета Ω_i **). Так как истинная длина свободного пробега для нейтрона группы j внутри G_0 имеет плотность

$$p_j(x) = \sum^j e^{-\Sigma^j x} \left(1 - e^{-\Sigma^j l_i} \right), \quad 0 < x < l_i,$$

то, согласно (33), необходимо умножить количество таких нейтронов на весовой множитель

$$\frac{p_j(\xi')}{p_{\xi'}(\xi')} = \frac{\sum^j \left(1 - e^{-\Sigma^{j_0} l_i} \right)}{\sum^{j_0} \left(1 - e^{-\Sigma^j l_i} \right)} e^{-(\Sigma^j - \Sigma^{j_0}) \xi'}. \quad (35)$$

Количество нейтронов группы j в пакете после столкновения мы обозначили $w_j(i)$. Из них в области G_0 останутся лишь

$w_j(i) \left(1 - e^{-\Sigma^j l_i} \right)$ нейтронов. С учетом (35) надо считать, что в точку $r_{i+1} = r_i + \xi' \Omega_i$ прилетят всего

$$v_j(i+1) = w_j(i) \sum^j L^j(i)$$

нейтронов группы j , где множитель

$$L^j(i) = (1/\Sigma^{j_0}) \left(1 - e^{-\Sigma^{j_0} l_i} \right) e^{-(\Sigma^j - \Sigma^{j_0}) \xi'}.$$

*) Можно также вычислить все вероятности поглощения — назовем их p_A^j — при условии, что испущен нейтрон группы j , и затем сосчитать величину $p_A = p_A^1 p_1 + \dots + p_A^m p_m$.

**) Будем писать ξ' без индекса i , чтобы не загромождать изложение индексами.

В результате столкновения в точке r_{i+1} из этих $v_j(i+1)$ нейтронов часть, а именно $v_j(i+1) (\sum_a^j / \sum^j)$ нейтронов, поглотятся, а $v_j(i+1) (\sum_k^j / \sum^j)$ нейтронов перейдут в группу номер k . Из всего пакета в группу номер k попадут

$$w_k(i+1) = \sum_{j=1}^k (\sum_k^j / \sum^j) v_j(i+1)$$

нейтронов, а количество поглощенных нейтронов равно

$$\sum_{j=1}^m (\sum_a^j / \sum^j) v_j(i+1).$$

Нетрудно проверить, что если ввести m -мерные векторы

$$\rho = \{\rho_1, \dots, \rho_m\}, \quad \hat{v}(i) = \{v_1(i)/\sum^1, \dots, v_m(i)/\sum^m\},$$

$a = \{\sum_a^1, \dots, \sum_a^m\}$ и диагональную матрицу

$$L(i) = \begin{pmatrix} L^1(i) & & & 0 \\ & \cdot & & \\ & & \cdot & \\ 0 & & & L^m(i) \end{pmatrix},$$

то схему расчета весов можно записать в векторной форме:

$$\begin{aligned} w(0) &= \rho, \quad \hat{v}(i+1) = L(i) w(i), \\ w(i+1) &= \hat{S} \hat{v}(i+1) \end{aligned} \quad (36)$$

при $i=0, 1, 2, \dots$. Количество поглощенных за всю историю пакета нейтронов выражается через скалярные произведения

$$\eta_A^{(3)} = \sum_{i=0}^{\infty} (a, \hat{v}(i+1)). \quad (37)$$

В [80] в качестве j_0 выбирался номер самой быстрой группы: $j_0=1$. Вероятно, в некоторых случаях выгоднее в качестве j_0 выбирать номер самой важной (или самой многочисленной) группы.

Формулы для расчета траектории пакета от количества групп m не зависят. При увеличении m меняется лишь размерность векторов и матриц в формулах (36).

Совсем другой метод введения векторных весов для решения интегральных уравнений предложен в [61].

§ 4. Статистические веса и интегральные уравнения

4.1. Вероятность p_A — линейный функционал от плотности столкновений. Предположим, что источник нейтронов, описанный в п. 1.1, излучает 1 нейтрон в единицу времени. Обозначим через $z(P)$ плотность столкновений за единицу времени в 5-мерном фазовом пространстве (см. п. 2.1.1) и рассмотрим уравнение (49) гл. 5

$$z(P) = \int K_{\text{ст}}(P', P) z(P') dP' + f(P). \quad (38)$$

Область интегрирования по $P' = (r', \Omega')$ в этом уравнении: по координате r' — все пространство, а по Ω' — все направления. Ядро столкновений выписано на стр. 223. Свободный член $f(P)$ — это плотность первых столкновений. Явное выражение для $f(P)$ приведено ниже в п. 4.2.

Искомая вероятность поглощения p_A в области G_0 может быть записана в форме скалярного произведения $p_A = (z, \psi)$, где

$$\psi(P) = [\sum_a (r)_i \sum_j (r)] \chi_{G_0}(r), \quad (39)$$

а $\chi_{G_0}(r)$ — индикатор области G_0 (т. е. $\chi_{G_0} = 1$ при $r \in G_0$, $\chi_{G_0} = 0$ при $r \notin G_0$). В самом деле,

$$\int z(P) \psi(P) dP = \int_{G_0} (\sum_a / \sum) dr \oint z(r, \Omega) d\Omega;$$

внутренний интеграл представляет собой плотность полного количества столкновений в точке r за единицу времени, или, что в нашем случае одно и то же, в расчете на один нейтрон источника; умножив это количество на $(\sum_a / \sum) dr$, получим количество поглощений в элементе dr около точки r , а проинтегрировав по G_0 , получим количество поглощений в области G_0 — также в расчете на один нейтрон источника.

4.2. Плотность первых столкновений. Обозначим

$$\omega = (r - r_0)/l, \quad l = |r - r_0|.$$

Вероятность того, что направление скорости нейтрона, испущенного источником, окажется в конусе $d\omega$ около направления ω , равна $d\omega/(4\pi)$. Далее, рассуждая в точности так же, как в п. 2.1.1, получим, что вероятность первого столкновения в элементе dP около точки P равна

$$f(P) dP = \frac{d\omega}{4\pi} \sum_i (r) e^{-\int_0^l \sum ds} d l \delta(\Omega - \omega) d\Omega,$$

откуда следует формула

$$f(P) = \frac{1}{4\pi l^2} \sum_i (r) \exp \left\{ - \int_0^l \sum (r_0 + \omega s) ds \right\} \delta(\Omega - \omega). \quad (40)$$

Важно подчеркнуть, что приведенный в п. 1.1 алгоритм расчета Ω_0 и первой точки столкновения r_1 представляет собой способ моделирования случайной точки $Q_0 = (r_1, \Omega_0)$ фазового пространства с плотностью (40). Плотность оказывается нормированной из-за того, что источник наш единичной мощности:

$$\begin{aligned} \int i(P) dP &= \oint \frac{d\omega}{4\pi} \int_0^\infty \sum (r) e^{-\int_0^l \Sigma ds} dl \oint \delta(\Omega - \omega) d\Omega = \\ &= \oint \frac{d\omega}{4\pi} \int_0^\infty \sum (r) e^{-\int_0^l \Sigma ds} dl = \oint \frac{d\omega}{4\pi} = 1. \end{aligned}$$

Рассмотрим теперь случайные точки столкновения $Q_i = (r_{i+1}, \Omega_i)$ в фазовом пространстве и выясним, какие траектории $Q_0 \rightarrow Q_1 \rightarrow Q_2 \rightarrow \dots$ соответствуют рассмотренным выше алгоритмам расчета p_A .

4.3. Имитация прохождения нейтронов как метод решения интегрального уравнения. Из п. 4.2 следует, что при реализации алгоритма п. 1.1 плотность начальной точки $(p(P))$ из гл. 5) равна $f(P)$. Плотность вероятностей перехода $(p(P, P'))$ из гл. 5) при $r_i \in G_0$ равна

$$p(Q_{i-1}, P) = p_s(r_i, \Omega_{i-1}; \omega) l^{-2} F'_{i-1}(l) \delta(\Omega - \omega),$$

где $l = |r - r_i|$, $\omega = (r - r_i)/l$. Если $r_i \notin G_0$, то можно считать, что траектория останавливается в точке Q_{i-1} , т. е. $p(Q_{i-1}, P) = \delta(P - Q_{i-1})$ и $Q_i = Q_{i-1}$. Таким образом,

$$p(P, P') = \begin{cases} K_{ст}(P, P')/s(P) & \text{при } r \in G_0, \\ \delta(P' - P) & \text{при } r \notin G_0, \end{cases} \quad (41)$$

причем $s(P) = \sum_s(r)/\sum(r)$.

Очевидно, траектории $Q_0 \rightarrow Q_1 \rightarrow \dots \rightarrow Q_v$ при таком методе расчета представляют собой траектории с поглощением типа \hat{T}_v во всем пространстве. Поглощение $a(P)$ из гл. 5) равно

$$a(P) = \sum_a(r)/\sum(r). \quad (42)$$

При $r \in G_0$ можно считать, что $\sum_a = \sum$, так что $a(P) = 1$.

Согласно гл. 5, п. 3.2 оценкой функционала (ψ, z) служит величина $\tilde{\zeta}_v^*[f] = [f(Q_0)/p(Q_0)] [\psi(Q_v)/a(Q_v)]$, которая в нашем случае,

когда $p(P) = f(P)$, а $\psi(P)$ определена формулой (39), равна

$$\tilde{\zeta}_v^*[f] = \begin{cases} 1, & \text{если } r_{v+1} \in G_0, \\ 0, & \text{если } r_{v+1} \notin G_0. \end{cases}$$

Таким образом, случайная величина η_A из п. 1.1 равна $\zeta_v^*[f]$ и $M\eta_A = (\psi, z) = p_A$.

4.4. Веса, заменяющие розыгрыш поглощения, как метод решения интегрального уравнения. Легко видеть, что алгоритму п. 3.1 соответствуют траектории без поглощения (типа T_∞) во всем пространстве, которые строятся по той же начальной плотности $p(P) = f(P)$ и плотности вероятностей перехода (41). Согласно п. 3.3 (гл. 5), в качестве оценки (ψ, z) можно выбрать величину $\zeta^*[f]$, которая в рассматриваемом случае запишется как

$$\zeta^*[f] = \sum_{j=0}^{\infty} W_j^* \psi(Q_j),$$

причем $W_j^* = s(Q_0)s(Q_1) \dots s(Q_{j-1})$. Если r_{v+1} — первая точка траектории, оказавшаяся вне G_0 , то $\psi(Q_j) = 0$ при всех $j = v, v+1, \dots$. Поэтому сумму можно закончить членом с $j=v-1$:

$$\zeta^*[f] = \sum_{j=0}^{v-1} W_j^* [\sum_a (r_{j+1}) / \sum (r_{j+1})].$$

Сопоставив формулы для W_j^* и $\zeta^*[f]$ с формулами (20) и (16), нетрудно заметить, что в этом случае $w_i = W_i^*$, $\eta_A^{(1)} = \varepsilon^*[f]$, следовательно, $M\eta_A^{(1)} = (\psi, z) = p_A$.

4.5. Веса, учитывающие вылет, как метод решения интегрального уравнения. Обозначим через G область фазового пространства, состоящую из таких точек $P = (r, \Omega)$, что $r \in G_0$. Так как при $P' \in G$ ядро столкновений $K_{ст}(P', P) = 0$, то (п. 3.5 гл. 5) уравнение (38) можно заменить уравнением

$$z(P) = \int G K_{ст}(P', P) z(P') dP' + f(P). \quad (43)$$

Алгоритму п. 3.2 отвечает начальная плотность $p(P)$ точки $Q_0 = (r_1, \Omega_0)$ в G такая, что

$$p(P) = (4\pi l^2)^{-1} [F'_0(l)/F_0(l_0)] \delta(\Omega - \omega),$$

где l_0 — расстояние от источника r_0 до границы области G_0 по направлению Ω_0 , $l = |r - r_0|$, $\omega = (r - r_0)/l$. Сравнивая это выражение с (40), видим, что

$$p(P) = f(P)/F_0(l_0). \quad (44)$$

Нетрудно также записать плотность вероятностей перехода, соответствующую алгоритму п. 3.2:

$$p(Q_{i-1}, P) = p_s(r_i, \Omega_{i-1}; \omega) \frac{F'_{i-1}(l)}{F_{i-1}(l_{i-1})} \frac{\delta(\Omega - \omega)}{l^2}.$$

Сравнивая эту формулу с (7), получим, что

$$p(P, P') = K_{ст}(P, P')/h(P, P'), \quad (45)$$

где, как нетрудно вычислить,

$$h(P, P') = \frac{\sum s(r)}{\sum (r)} \left[1 - \exp \left\{ - \int_0^{l(r, \omega')} \sum (r + \omega' s) ds \right\} \right],$$

$l(r, \omega')$ — расстояние от точки r до границы G_0 по направлению вектора $\omega' = (r' - r)/|r' - r|$

В качестве поглощения в п. 3.2 выбирается истинное поглощение: $a(P) = \sum a(r)/\sum (r)$. Таким образом, рассматриваемому алгоритму соответствуют траектории с поглощением типа \tilde{T}_v внутри G .

Согласно п. 3.5 гл. 5 в качестве оценки для (ψ, z) можно выбрать величину

$$\tilde{\xi}_v^* [f] = [f(Q_0)/p(Q_0)] \tilde{W}_v^* [\psi(Q_v)/a(Q_v)],$$

где (см. (16) гл. 5)

$$\tilde{W}_i^* = \prod_{j=1}^i \frac{K_{s_i}(Q_{j-1}, Q_j)}{s(Q_{j-1}) p(Q_{j-1}, Q_j)} = \prod_{j=1}^i \frac{h(Q_{j-1}, Q_j)}{s(Q_{j-1})},$$

а $s(P) = 1 - a(P)$. Здесь v — номер точки $Q_v = (r_{v+1}, \Omega_v)$, в которой траектория закончилась поглощением

С учетом выражения для $h(P, P')$ можем записать \tilde{W}_i^* в виде

$$\tilde{W}_i^* = F_1(l_1) \dots F_i(l_i).$$

Приняв во внимание формулы (39) и (44), получим, что

$$\tilde{\xi}_v^* [f] = F_0(l_0) F_1(l_1) \dots F_v(l_v).$$

Наконец, сравнив формулы для \tilde{W}_i^* и $\tilde{\xi}_v^* [f]$ с формулами (28) и (29), увидим, что $\omega_i = F_0(l_0) \tilde{W}_{i-1}^*$ и $\eta_A^{(2)} = \tilde{\xi}_v^* [f]$. Следовательно,

$$M_{1A}^{(2)}(\psi, z) = p_A.$$

4.6. Веса, заменяющие розыгрыш поглощения и учитывающие вылет, как метод решения интегрального уравнения. Снова рассмотрим уравнение (43) внутри G . Алгоритму п. 3.3 отвечает та же начальная плотность (44) и та же плотность вероятностей перехода (45). Однако соответствующие траектории в фазовом пространстве представляют собой траектории типа T_∞ без поглощения. Для оценки скалярного произведения (ψ, z) используем оценку $\xi^* [f]$ из гл. 5:

$$\xi^* [f] = F_0(l_0) \sum_{i=0}^{\infty} W_i^* \psi(Q_i) = \sum_{i=0}^{\infty} F_0(l_0) W_i^* \frac{\sum a(r_{i+1})}{\sum (r_{i+1})}.$$

Однако, в отличие от предыдущего пункта, здесь

$$W_i^* = \prod_{j=1}^i \frac{K_{\text{ст}}(Q_{j-1}, Q_j)}{\rho(Q_{j-1}, Q_j)} = \prod_{j=1}^i h(Q_{j-1}, Q_j) = \prod_{j=1}^i \frac{\sum_s (r_j)}{\sum (r_j)} F_j(l_j).$$

Сравнивая последние две формулы с (31) и (30), нетрудно заметить, что $v_{i+1} = W_i^* F_0(l_0)$ и $\eta_A^{(3)} = \xi^*[f]$. Следовательно, $M\eta_A^{(3)} = (\psi, z) = \rho_A$.

4.7. Веса п. 3.3 при решении уравнения Пайерлса. В п. 1.2 изложена схема расчета критичности по поколениям. Однако можно рассчитывать критичность также по времени, сравнивая количества нейтронов в G_0 через заданные промежутки времени [118] или по пробегам, сравнивая количества нейтронов в G_0 после каждого пробега всех нейтронов.

Нетрудно показать, что численный метод решения уравнения Пайерлса, использованный в п. 4.4 гл. 5, можно интерпретировать как расчет по пробегам с использованием весов п. 3.3. В самом деле, пусть на j -м шаге в G_0 оказалось N нейтронов, один из которых — нейтрон номер s — расположен в точке r_j и имеет вес θ_j (индекс s мы опускаем). Разыграем случайное направление скорости этого нейтрона μ_j и найдем расстояние до границы шара по этому направлению l_j . Тогда, согласно п. 3.3, $\theta_j e^{-\alpha l_j}$ нейтронов вылетят из шара, и столкновение в точке r_{j+1} испытают всего $\theta_j F_j$ нейтронов, где $F_j = 1 - e^{-\alpha l_j}$. Из них $\theta_j F_j (\sum_s / \sum_i)$ нейтронов испытают рассеяние, а $\theta_j F_j (\sum_f / \sum)$ вызовут деления, в результате которых получится всего $v \theta_j F_j (\sum_f / \sum)$ нейтронов. Всего в точке r_{j+1} после столкновения окажется

$$\theta_{j+1} = \theta_j F_j (\sum_s / \sum) + \theta_j F_j (\bar{v} \sum_f / \sum) = \theta_j F_j (\beta / \alpha)$$

нейтронов. Эта формула совпадает с соответствующей формулой п. 4.4.1 гл. 5. Легко также проверить, что отношение количеств нейтронов в G_0 , равное $\left(\sum_{s=1}^N \theta_{j,s} \right) / \left(\sum_{s=1}^N \theta_{j+1,s} \right)$ совпадает с $\lambda_N^{(j)}$.

Упражнения к главе 6

1. Записать формулы для расчета методом п. 2.2.2 свободного пробега нейтрона в многослойном шаре $r^2 < R_n^2$, в котором

$$\sum \equiv \alpha_k \quad \text{при} \quad R_{k-1}^2 < r^2 < R_k^2, \quad R_0 = 0, \quad R_1 < R_2 < \dots < R_n.$$

2. Записать формулы для расчета $k_{\text{эфф}}$ однородного шара $r^2 < R^2$ методом п. 1.2. Плотность нейтронов первого поколения считать

пропорциональной $1 - (r/R)^2$; предполагать, что $\bar{v} = 2,5$, а $k_{эфф}$ близок к 1*).

3. Записать формулу для расчета $k_{эфф}$ однородного шара из упражнения 2 методом поколений с помощью весов п 3.3.

4 Доказать, что плотности (44) и (45) нормированы.

5 Из уравнения (43) вывести уравнение Пайерлса для плотности нейтронов $n(r)$ в критической области G_0

$$n(r) = \int_{G_0} \frac{\beta(r')}{4\pi l^2} \exp \left\{ - \int_0^l \sum (r + \omega s) ds \right\} n(r') dr',$$

где $l = |r' - r|$.

У к а з а н и е. В критическом случае $f(P)$ выражается через $z(P)$ при помощи интегрального преобразования с ядром $[\bar{v}(r') \sum_f(r') / \sum_s(r')] K_{сг}(P', P)$. Плотность $n(r)$ также выражается через $z(P)$, так как $\sum(r) n(r) = \oint z(P) d\Omega$.

*) Если $k_{эфф}$ сильно отличается от единицы, то n_i либо быстро обращается в нуль, либо переполняет накопитель ЭВМ. Приходится искусственно пополють или разреживать некоторые поколения, или использовать нормировку числа нейтронов.

ГЛАВА 7

НЕСЛУЧАЙНЫЕ ТОЧКИ В АЛГОРИТМАХ МОНТЕ-КАРЛО

В этой главе рассматриваются только такие методы Монте-Карло, в которых выборочные средние используются для оценки математических ожиданий. Сюда относятся как методы вычисления интегралов, так и методы расчета величин, вырабатываемых в ходе моделирования естественных процессов: например, формула (8) гл. 5

$$(\psi, K^j \varphi) \approx \frac{1}{N} \sum_{s=1}^N \theta_j[\psi]_s \quad (1)$$

или формула (1) гл. 6

$$p_A \approx \frac{1}{N} \sum_{s=1}^N (\eta_A)_s \quad (2)$$

Для этих методов можно указать бесконечные множества заведомо не случайных чисел, которые гарантируют сходимость соответствующих алгоритмов Монте-Карло не по вероятности, а в обычном смысле, — если использовать эти числа вместо случайных. Такие числа часто называют *квазислучайными*. Очевидно, достаточно большие группы квазислучайных чисел в каком-то смысле удовлетворяют определению псевдослучайных чисел, приведенному на стр. 17. (Если проверять такие числа с помощью тестов § 3 гл. 1, то часто результаты оказываются слишком хорошими, например $P > 0,9999$).

Использование таких чисел в расчетах имеет свои минусы и свои плюсы. С одной стороны, вероятная ошибка перестает быть характеристикой порядка ошибки, числа пригодны лишь для определенных классов

задач и должны использоваться в определенном порядке.

С другой — гарантирована сходимость вычислений, не нужны статистические тесты и, самое главное, — во многих задачах удается добиться более быстрой сходимости: вместо $1/\sqrt{N}$ порядок ошибки оказывается равным $1/N^{1-\varepsilon}$ (где $\varepsilon > 0$ как угодно мало). Пренебрегать таким выигрышем нельзя!

Отношение специалистов к квазислучайным числам различное: одни считают, что только на таком пути можно строго обосновать методы Монте-Карло; другие считают, что такой подход уничтожает методы Монте-Карло... ([70, 130, 183]). Однако все алгоритмы, изложенные в гл. 2—6 книги, сохраняют свою силу независимо от того, какие псевдослучайные числа предполагается использовать. Так что говорить о гибели методов Монте-Карло вряд ли стоит.

Можно не сомневаться, что и для других методов Монте-Карло (например, связанных с расчетом средних по одной эргодичной траектории) будут найдены достаточно удобные квазислучайные числа.

§ 1. Конструктивная размерность алгоритмов Монте-Карло

Так же как в гл. 3, п. 2.4, условимся говорить, что задан метод Монте-Карло для расчета некоторой скалярной величины a , если указана такая случайная величина η , что ее математическое ожидание равно a :

$$M\eta = a, \quad (3)$$

и оценкой для a служит среднее арифметическое

$$a \approx \frac{1}{N} \sum_{i=1}^N \eta_i; \quad (4)$$

здесь η_1, \dots, η_N — независимые значения η .

Однако формула (4) не определяет алгоритма расчета, так как значения одной и той же случайной величины η можно вычислить различными способами (гл. 2). Так же как в п. 2.4 гл. 3, условимся говорить, что задан алгоритм Монте-Карло для расчета величины a , если, кроме формулы (4), задана формула

$$\eta = \Phi(\gamma^{(1)}, \dots, \gamma^{(n)}, \dots), \quad (5)$$

выражающая нужную нам величину η через независимые случайные числа.

Таким образом, формулы (1) и (2) определяют лишь методы Монте-Карло для расчета величин $(\psi, K^1\psi)$ и ρ_A . А соответствующие им алгоритмы будут заданы только тогда, когда будут записаны формулы для моделирования всех входящих в определение $\theta_j[\psi]$ и η_A величин при помощи случайных чисел γ .

1.1. Алгоритмы с конечной конструктивной размерностью.

Определение. Если функция Φ зависит от n аргументов

$$\Phi = \Phi(\gamma^{(1)}, \dots, \gamma^{(n)}),$$

то мы скажем, что *конструктивная размерность* (к. р.) алгоритма (4)—(5) равна n .

В этом случае для реализации i -го «испытания» достаточно выбрать n случайных чисел $\gamma_i^{(1)}, \dots, \gamma_i^{(n)}$ и вычислить по ним случайное значение

$$\eta_i = \Phi(\gamma_i^{(1)}, \dots, \gamma_i^{(n)}).$$

Конечно, может случиться, что при каких-то конкретных значениях аргументов функция Φ зависит не от всех $\gamma^{(1)}, \dots, \gamma^{(n)}$. Так что конструктивная размерность n — это максимальное количество случайных чисел, которое может понадобиться для реализации одного испытания.

Так как каждая из независимых величин $\gamma^{(1)}, \dots, \gamma^{(n)}$ равномерно распределена в интервале $(0, 1)$, то функция $\Phi(y_1, \dots, y_n)$ определена в единичном n -мерном кубе

$$K^n = \{0 < y_1 < 1, \dots, 0 < y_n < 1\},$$

и случайная точка $\Gamma = (\gamma^{(1)}, \dots, \gamma^{(n)})$ равномерно распределена в K^n : плотность ее $p_\Gamma(y_1, \dots, y_n) \equiv 1$ при $(y_1, \dots, y_n) \in K^n$. Следовательно, искомая величина α может быть записана в форме n -мерного интеграла по K^n :

$$\alpha = M\eta = M\Phi(\Gamma) = \int_0^1 \dots \int_0^1 \Phi(y_1, \dots, y_n) dy_1 \dots dy_n. \quad (6)$$

Мы приходим к следующей общей интерпретации алгоритмов Монте-Карло: если конструктивная размерность алгоритма равна n (к. р. $= n$), то этот алгоритм представляет собой приближенный метод вычисления

n -мерного интеграла (6) по случайным точкам $\Gamma_i = (\gamma_i^{(1)}, \dots, \gamma_i^{(n)})$, равномерно распределенным в K^n

$$\int_{K^n} \Phi(P) dP \approx \frac{1}{N} \sum_{i=1}^N \Phi(\Gamma_i). \quad (7)$$

Здесь и ниже для краткости используется запись

$$\int_{K^n} \Phi(P) dP \equiv \int_0^1 \dots \int_0^1 \Phi(y_1, \dots, y_n) dy_1 \dots dy_n.$$

Формула (7) равносильна формулам (4) и (5).

Пример. Вычисляется однократный интеграл

$$I = \int_a^b f(x) p(x) dx, \quad (8)$$

где $p(x)$ — плотность вероятностей некоторой случайной величины ξ , определенной в интервале $a < x < b$, с помощью простейшего метода Монте-Карло

$$I \approx \frac{1}{N} \sum_{i=1}^N f(\xi_i).$$

Если случайную величину ξ моделировать методом обратных функций (п. 1.4 гл. 2), то $\xi = G(y)$. Функция $G(y)$ это обратная функция

по отношению к $y = F(x)$, где $F(x) = \int_a^x p(t) dt$. Замена $y = F(x)$

преобразует интеграл (8) к виду

$$I = \int_0^1 f(G(y)) dy.$$

Очевидно, в этом случае $\Phi = f(G(y))$ и к.р. равна 1.

Однако, вообще говоря, для моделирования ξ можно использовать какую-нибудь формулу вида $\xi = g(\gamma_1, \dots, \gamma_n)$ (§ 4 гл. 2). Согласно (25) гл. 2 (ср. сноску на стр. 52) в этом случае

$$p(x) = \int_0^1 \dots \int_0^1 \delta(x - g(y_1, \dots, y_n)) dy_1 \dots dy_n.$$

Подставив это равенство в (8) и поменяв порядок интегрирований — сперва по x , а потом по y_1, \dots, y_n — получим, что

$$I = \int_0^1 \dots \int_0^1 f(g(y_1, \dots, y_n)) dy_1 \dots dy_n.$$

Таким образом, в этом случае $\Phi = f(g(y_1, \dots, y_n))$ и к.р. равна 1.

1.2. Алгоритм с бесконечной конструктивной размерностью. Такие алгоритмы встречаются довольно часто при моделировании физических задач. Например, в задаче о поглощении нейтронов (гл. 6, п. 1.1) траектория нейтрона может (теоретически) состоять из сколь угодно большого числа звеньев, так что нельзя указать заранее, сколько значений $\gamma^{(1)}, \dots, \gamma^{(n)}, \dots$ понадобится нам для реализации такой траектории.

Алгоритмы с к. р., равной ∞ , получаются также при моделировании случайных величин методами отбора (гл. 2, § 5), когда количество значений γ , используемых для реализации одного значения η , случайно и (теоретически) может оказаться сколь угодно большим.

Пример. Рассмотрим снова вычисление интеграла (8). Предположим, что $p(x) \leq c$ и значения ξ вычисляются методом Неймана (гл. 2, п. 5.3):

$$\xi = a + (b-a)\gamma, \text{ если } c\gamma' < p(a + (b-a)\gamma);$$

если последнее неравенство не выполнено, то пара случайных чисел (γ, γ') отбрасывается и выбирается новая пара.

Нетрудно проверить, что в этом случае

$$\xi = g(\gamma^{(1)}, \gamma^{(1)'}, \dots, \gamma^{(k)}, \gamma^{(k)'}, \dots),$$

где функция $g(y_1, y_1', \dots, y_k, y_k', \dots)$ определяется следующими условиями:

$$\text{если } cy_1' \geq p(g_1), \dots, cy_{k-1}' \geq p(g_{k-1}),$$

$$\text{но } cy_k' < p(g_k), \text{ то } g = g_k,$$

где (для краткости) мы обозначили g_k величину

$$a + (b-a)y_k;$$

$k=1, 2, \dots$ (Конечно, при желании можно ввести единую нумерацию переменных). И в этом случае $\Phi = f(g)$, так что к. р. $= \infty$.

Запишем (формально; строгое определение имеется в [82]) интеграл по бесконечномерному единичному кубу, в котором все $0 < y_k < 1$, $0 < y_k' < 1$.

$$I_\infty = \int_0^1 \dots \int_0^1 \dots f(g) dy_1 dy_1' \dots dy_k dy_k' \dots$$

и докажем, что $I_\infty = I$. Для этого представим I_∞ в виде бесконечной суммы интегралов по областям, в которых $g = g_k$:

$$\begin{aligned}
 I_\infty = & \int_0^1 \int_0^{\frac{p(g_1)}{c}} f(g_1) dy_1 dy'_1 + \int_0^1 \int_{\frac{p(g_1)}{c}}^1 dy_1 dy'_1 \int_0^1 \int_0^{\frac{p(g_2)}{c}} f(g_2) dy_2 dy'_2 + \dots \\
 & \dots + \int_0^1 \int_{\frac{p(g_1)}{c}}^1 dy_1 dy'_1 \int_0^1 \int_{\frac{p(g_2)}{c}}^1 dy_2 dy'_2 \dots \\
 & \dots \int_0^1 \int_{\frac{p(g_{k-1})}{c}}^1 dy_{k-1} dy'_{k-1} \int_0^1 \int_0^{\frac{p(g_k)}{c}} f(g_k) dy_k dy'_k + \dots
 \end{aligned}$$

Здесь каждый из интегралов вида

$$\int_0^1 \int_0^{\frac{p(g_k)}{c}} f(g_k) dy_k dy'_k = \int_0^1 f(a + (b-a)y_k) \frac{1}{c} p(a + (b-a)y_k) dy_k$$

легко вычисляется с помощью замены $x = a + (b-a)y_k$ и равен

$$\frac{1}{c(b-a)} \int_a^b f(x) p(\bar{x}) dx = \varepsilon I,$$

где ε — эффективность метода Пеймана. А каждый из «наружных» интегралов равен

$$\int_0^1 \int_{\frac{p(g_k)}{c}}^1 dy_k dy'_k = \int_0^1 \left[1 - \frac{1}{c} p(a + (b-a)y_k) \right] dy_k = 1 - \varepsilon.$$

Следовательно,

$$I_\infty = \sum_{k=1}^{\infty} (1 - \varepsilon)^{k-1} \varepsilon I = I.$$

Реализация алгоритмов Монте-Карло с $k, p = \infty$ на практике затруднений не вызывает, если предполагать, что в расчете используются «настоящие» случайные числа γ . Иногда каждое испытание доводят до конца, и количество использованных случайных чисел оказывается конечным (хотя и случайным). Иногда расчет испытания прекращают после выполнения некоторого условия. Например, при решении интегрального уравнения можно

ограничиться конечным (фиксированным) числом членов ряда Неймана, а при использовании алгоритма п. 3.3 гл. 6 учесть условие обрыва $\omega_k < \varepsilon$. В этой ситуации мы по существу аппроксимируем алгоритм с к. р. $= \infty$ алгоритмом с к. р. $= n_0$. Величина n_0 зависит от допустимой погрешности расчета и может зависеть от общего количества испытаний N или даже от конкретных значений γ , использованных во время расчета. Поэтому оценка n_0 иногда весьма затруднительна.

Значение к. р. играет важную роль, когда мы в качестве случайных значений γ хотим использовать псевдослучайные (детерминированные) числа. В этом случае алгоритм с различными к. р. приходится рассматривать отдельно.

§ 2. n-мерные псевдослучайные точки

2.1. Равномерно распределенные последовательности. Рассмотрим произвольный алгоритм Монте-Карло с к. р. $= n$ и соответствующую ему функцию $\Phi(y_1, \dots, y_n)$. Как мы видели в п. 1.1, этот алгоритм сводит решаемую задачу к вычислению интеграла (7).

Естественно поставить вопрос: нельзя ли указать неслучайную последовательность точек P_1, \dots, P_i, \dots из K^n такую, что

$$\int_{K^n} \Phi(P) dP = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \Phi(P_i) \quad (9)$$

для всех функций Φ из достаточно широкого класса?

Определение. Последовательность точек P_1, \dots, P_i, \dots называется *равномерно распределенной* в K^n , если соотношение (9) справедливо для любой функции $\Phi(y_1, \dots, y_n)$, интегрируемой в K^n по Риману *).

Понятие это было введено в 1916 г. Г. Вейлем [182], который построил также примеры равномерно распределенных последовательностей.

*) Напомним, что интеграл Римана определяется только для ограниченных функций.

Однако автор книги недавно доказал, что если выбрать $P_i = Q_i^*$ (см. п. 2.4.2), то формула (9) справедлива также для функций $\Phi(P)$ с любыми степенными особенностями вида $y_1^{-\beta_1} \dots y_n^{-\beta_n}$, где $\beta_1 < 1, \dots, \beta_n < 1$ (см. Докл. АН СССР 210, № 2, 1973 г., 278–281).

Сопоставление формул (9) и (7) показывает, что для реализации алгоритмов Монте-Карло с к.р. = n можно попытаться вместо случайных точек Γ_i использовать точки равномерно распределенной последовательности P_i . Для этого надо при реализации i -го «испытания» вместо случайных чисел $\gamma_i^{(1)}, \dots, \gamma_i^{(n)}$ использовать декартовы координаты $y_{i,1}, \dots, y_{i,n}$ точки P_i . Соотношение (9) гарантирует сходимость такого способа вычислений для большинства встречающихся на практике алгоритмов.

Легко заметить, что равенство (9) не нарушается, если изменить в последовательности P_1, \dots, P_i, \dots любое конечное число точек. Однако сходимость средних к пределу может при этом очень замедлиться. Поэтому далеко не каждую равномерно распределенную последовательность разумно использовать на практике в качестве псевдослучайных точек. Среди всех равномерно распределенных последовательностей следует отобрать в некотором смысле (см. ниже п. 2.2) «хорошие». Отыскание таких последовательностей обычно наталкивается на серьезные трудности.

Например, еще Г. Вейль доказал, что любые последовательности точек с декартовыми координатами

$$P_i = (D(i\theta_1), \dots, D(i\theta_n)), \quad i = 1, 2, \dots,$$

где $\theta_1, \dots, \theta_n$ — алгебраически независимые иррациональные чис-

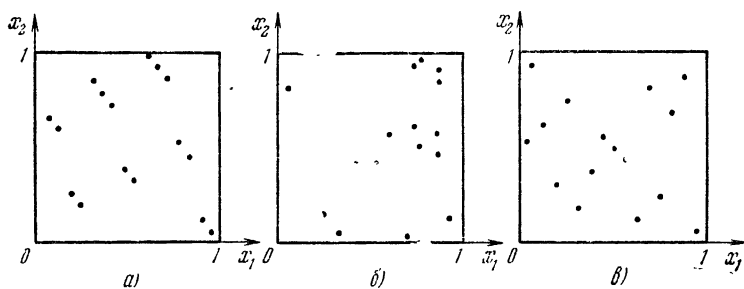


Рис. 67.

ла, равномерно распределены в K^n . Но ни одного «хорошего» набора $\theta_1, \dots, \theta_n$ при $n > 2$ до сих пор не известно. (На рис. 67, а изображены точки P_1, \dots, P_{16} в квадрате, полученные при $\theta_1 = \sqrt{2}/2, \theta_2 = \sqrt{3}/2$.)

2.2. Геометрическая характеристика равномерно распределенных последовательностей. Обозначим через G произвольную n -мерную область, принадлежащую K^n , а через V_G — ее объем (n -мерный). Обозначим через $S_N(G)$ — количество точек с номерами $1 \leq i \leq N$ принадлежащих G .

Теорема (Г. Вейль). Для того чтобы последовательность точек P_1, \dots, P_i, \dots была равномерно распределенной в K^n , необходимо и достаточно, чтобы для любой области G

$$\lim_{N \rightarrow \infty} [S_N(G)/N] = V_G. \quad (10)$$

Отсюда видно, что при больших N количество точек, принадлежащих G , среди точек P_1, \dots, P_N , приблизительно пропорционально объему V_G .

Рассмотрим случайную точку Γ , равномерно распределенную в K^n , и N ее независимых реализаций $\Gamma_1, \dots, \Gamma_N$. Так как вероятность $P\{\Gamma \in G\} = V_G$, то сходимость частоты попадания этих реализаций в G к вероятности попадания означает, что

$$[S_N(G)/N] \xrightarrow{P} V_G.$$

Сравнение этой формулы с (10) снова показывает, что точки равномерно распределенной последовательности являются аналогами независимых реализаций случайной точки Γ .

Мы уже отмечали, что не все равномерно распределенные последовательности одинаково хорошо распределены. Оценить «равномерность» распределения можно при помощи величины, называемой отклонением. Чтобы определить ее, выберем в K^n произвольную точку P и обозначим через Π_P параллелепипед с диагональю OP и со сторонами, параллельными координатным осям (рис. 68).

Отклонением группы точек P_1, \dots, P_N называется величина

$$D_N = \sup_{P \in K^n} |S_N(\Pi_P) - NV_{\Pi_P}|. \quad (11)$$

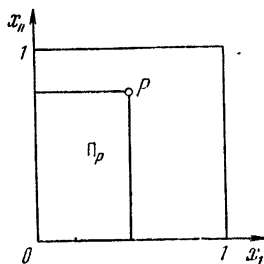


Рис. 68.

Следующая теорема легко вытекает из работ Г. Вейля:

Теорема. *Для того чтобы последовательность точек P_1, \dots, P_N, \dots была равномерно распределенной в K^n , необходимо и достаточно, чтобы*

$$\lim_{N \rightarrow \infty} (D_N/N) = 0.$$

Очевидно, чем быстрее убывает отношение D_N/N , тем более равномерно распределена последовательность *).

Можно доказать, что $1/2 \leq D_N \leq N$, но неясно, каков наилучший порядок роста D_N при $N \rightarrow \infty$. В настоящее время известны лишь два класса последовательностей точек в K^n , для которых при всех N

$$D_N = O(\ln^n N). \quad (12)$$

Это последовательности Холтона и ЛП $_{\tau}$ -последовательности. Примеры таких последовательностей — P_i^* и Q_i^* — приведены ниже в п. 2.4. Существуют ли последовательности, для которых $D_N = o(\ln^n N)$ при всех N неизвестно. Однако для точек Q_1^*, \dots, Q_N^* при $N = 2^m$ отклонение равно $D_N = O(\ln^{n-1} N)$.

В книге [82] изучается другая количественная характеристика расположения группы точек P_1, \dots, P_N , называемая *неравномерностью* $\varphi_{\infty} = \varphi_{\infty}(P_1, \dots, P_N)$. Для нее справедлива теорема, аналогичная предыдущей: для того чтобы последовательность точек P_1, \dots, P_N была равномерно распределенной в K^n , необходимо и достаточно, чтобы $\lim_{N \rightarrow \infty} (\varphi_{\infty}/N) = 0$.

Можно доказать, что $1 \leq \varphi_{\infty} \leq N$. Поэтому наиболее равномерно распределенными следует считать такие последовательности, неравномерности которых при всех N ограничены. Среди известных в настоящее время последовательностей точек в K^n лишь ЛП $_{\tau}$ -последовательности обладают этим свойством:

$$\varphi_{\infty}(Q_1^*, \dots, Q_N^*) \leq O(n, \tau). \quad (13)$$

Для последовательностей Холтона получена только более слабая оценка: $\varphi_{\infty} = O(\ln^n N)$.

*) В литературе часто отклонением называют отношение D_N/N , так как это есть верхняя грань отклонений эмпирической функции распределения $S_N(\Pi_P)/N$ точек P_1, \dots, P_N от теоретической функции распределения случайной точки Γ , которая в точке P равна V_{Π_P} . Ср. также упреждение 8 гл. 1.

2.3. Ускорение сходимости. Формула (9) справедлива для всех интегрируемых по Риману функций $\Phi(y_1, \dots, y_n)$. Если рассмотреть более узкие классы функций, то возможны оценки погрешности этой формулы. Например, неравенство

$$\left| \int_{K^n} \Phi(P) dP - \frac{1}{N} \sum_{i=1}^N \Phi(P_i) \right| \leq c(\Phi) \frac{D_N}{N}, \quad (14)$$

где $c(\Phi)$ ни от N , ни от точек P_i не зависит, справедливо для любых P_1, \dots, P_N и для всех функций $\Phi(y_1, \dots, y_n)$, которые непрерывны и ограничены в K_n вместе со своими частными производными, содержащими не более одного дифференцирования по каждой переменной. (Все эти производные можно записать формулой $\partial^k \Phi / \partial y_{j_1} \dots \partial y_{j_k}$, где $1 \leq j_1 < j_2 < \dots < j_k \leq n$ и k может принимать значения $1, 2, \dots, n$. Старшая среди этих производных, $\partial^n \Phi / \partial y_1 \dots \partial y_n$).

В случае $n=1$ доказательство неравенства (14) аналогично доказательству теоремы 5 гл. 3. В самом деле, пусть $c(\Phi) = \int_0^1 |\Phi'(x)| dx$. Тогда, полагая в (46) стр. 128 $f = \Phi$, получим, что каковы бы ни были точки x_1, \dots, x_N из интервала $(0,1)$

$$\left| \int_0^1 \Phi dx - \frac{1}{N} \sum_{i=1}^N \Phi(x_i) \right| \leq \frac{c(\Phi)}{N} \sup_{0 < x < 1} |S_N(x) - Nx| = \frac{c(\Phi) D_N}{N}.$$

Однако алгоритмы Монте-Карло весьма часто приводят к разрывным функциям Φ . Поэтому важно отметить, что оценка (14) справедлива для гораздо более широких классов функций (Э. Хлавка [138], И. М. Соболев [82]). Например, если все разрывы функции $\Phi(y_1, \dots, y_n)$ расположены на конечном числе гиперплоскостей вида $y_k = \text{const}$ (т. е. параллельных координатным гиперплоскостям), а в остальных точках куба K^n функция Φ и все вышеупомянутые производные непрерывны и ограничены, то неравенство (14) выполнено.

Интересно, что разрывы, возникающие при моделировании дискретных случайных величин и при использовании метода суперпозиции (гл. 2, п. 3.3) как раз такого типа. Напротив, при использовании метода Неймана

(гл. 2, п. 5.3) возникают разрывы, которые не обязаны располагаться в гиперплоскостях, параллельных координатным.

Пример. Снова предположим, что вычисляется интеграл (8), а случайная величина ξ моделируется методом суперпозиции:

$$\xi = G_k(\gamma^{(2)}), \text{ если } c_1 + \dots + c_{k-1} < \gamma^{(1)} < c_1 + \dots + c_k;$$

($G_k(y)$ — обратная функция по отношению к $F_k(x)$). В этом случае

$$\Phi = f(G_k(y_2)) \quad \text{при} \quad \sum_{j=1}^{k-1} c_j \leq y_1 < \sum_{j=1}^k c_j,$$

так что $\Phi(y_1, y_2)$, вообще говоря, имеет разрывы при $y = c_1, y = c_1 + c_2, \dots$.

В этом же случае метод Неймана (см. пример п. 1.2) приводит к функции $\Phi = j(g(y_1, y'_1, \dots))$, которая разрывна при $cy'_k = p(a + (b-a)y_k)$, $k=1, 2, \dots$.

Если в (14) подставить $P_i = P_i^*$ или $P_i = Q_i^*$, то в соответствии с (12) правая часть окажется порядка $O(N^{-1} \ln^2 N)$. Так как при всех достаточно больших N справедливо неравенство $\ln^2 N < N^\epsilon$ (при любых фиксированных $n \geq 1$ и $\epsilon > 0$), то можно сказать, что *погрешность (14) убывает быстрее, чем $N^{-(1-\epsilon)}$ с любым $\epsilon > 0$* . Напомним, что порядок погрешности формулы (7) с «настоящими» случайными точками равен $N^{-1/2}$, т. е. заметно хуже.

Численный пример, сосчитанный с помощью точек Q_i^* , имеется в гл. 5, п. 4.4.1. См. также [13а].

Заметим, что порядок сходимости формулы (9) не может быть $o(N^{-1})$ даже на весьма узких классах функций (см. упражнение 5 на стр. 278).

2.4. «Хорошие» псевдослучайные точки.

2.4.1 Последовательность Холтона P_i^* . Для построения этой последовательности необходимо определить числовые последовательности $p_r(i)$. Фиксируем натуральное число $r \geq 2$.

Определение. Если в r -ичной системе счисления $i = a_m a_{m-1} \dots a_2 a_1$, то снова в r -ичной системе

$$p_r(i) = 0, a_1 a_2 \dots a_{m-1} a_m.$$

Здесь все a_s — целые r -ичные цифры, т. е. равны одному из значений $0, 1, 2, \dots, r-1$. В десятичной системе последние две

формулы выглядят так:

$$i = \sum_{s=1}^m a_s r^{s-1}, \quad p_r(i) = \sum_{s=1}^m a_s r^{-s}.$$

Первые 10 значений $p_3(i)$ приведены в табл. 1.

Таблица 1

i	1	2	3	4	5	6	7	8	9	10
i троичн.	1	2	10	11	12	20	21	22	100	101
$p_3(i)$ троичн.	0,1	0,2	0,01	0,11	0,21	0,02	0,12	0,22	0,001	0,101
$p_3(i)$	1/3	2/3	1/9	4/9	7/9	2/9	5/9	8/9	1/27	10/27

Пусть r_1, \dots, r_n — попарно взаимно простые числа. Последовательностью Холтона называется последовательность точек в K^n с декартовыми координатами

$$(p_{r_1}(i), \dots, p_{r_n}(i)), \quad i=1, 2, \dots$$

Эти последовательности были построены Дж. Холтоном [131], получившим для них оценку (12). Все такие последовательности равномерно распределены в K^n .

На практике обычно в качестве r_1, \dots, r_n выбирают первые n простых чисел: $r_1=2, r_2=3, r_3=5, \dots$ и используют n -мерные точки

$$P_i^* = (p_2(i), p_3(i), \dots, p_{r_n}(i)), \quad i=1, 2, \dots$$

2.4.2. ЛП_τ-последовательность Q_i^* . Свойства ЛП_τ-последовательностей подробно изучаются в [82]. Здесь мы укажем лишь алгоритм для расчета точек

$$Q_i^* = (q_{i,1}, \dots, q_{i,n}), \quad i=1, 2, \dots,$$

образующих ЛП_τ-последовательность. Программа расчета на ЭВМ БЭСМ-4 имеется в [86].

Определение. Если в двоичной системе счисления

$$i = e_m e_{m-1} \dots e_2 e_1, \quad (15)$$

то для всех $j=1, 2, \dots, n$

$$q_{i,j} = e_1 V_j^{(1)} * e_2 V_j^{(2)} * \dots * e_m V_j^{(m)}. \quad (16)$$

Здесь e_1, \dots, e_m — двоичные цифры, каждая из которых равна 0 или 1. В десятичной системе

$$i = 2^{m-1} e_m + 2^{m-2} e_{m-1} + \dots + 2 e_2 + e_1.$$

Числа $V_j^{(s)}$ можно найти по табл. 6 (стр. 297), которая позволяет вычислить более $2 \cdot 10^6$ точек Q_i в кубе K^n размерности $n \leq 13$. Звездочкой (*) обозначена операция поразрядного сложения по модулю два в двоичной системе.

Более подробно, чтобы вычислить «сумму» $a * b$, надо оба слагаемых записать в двоичной системе

$$a = 0, a_1 a_2 \dots a_M, \quad b = 0, b_1 b_2 \dots b_M;$$

тогда в двоичной системе

$$a * b = 0, c_1 c_2 \dots c_M,$$

где $c_k = (a_k + b_k) \pmod{2}$ или, другими словами, $c_k = 1$, если $a_k \neq b_k$ и $c_k = 0$, если $a_k = b_k$.

В системе команд любой ЭВМ имеется специальная команда, осуществляющая операцию *. Обычно ее называют командой сравнения. Она относится к числу логических команд и выполняется быстрее, чем арифметические команды. Вообще, для расчета по формуле (16) нужны лишь логические команды (произведение $e_k V_j^{(k)}$ либо равно $V_j^{(k)}$, если $e_k = 1$, либо равно нулю, если $e_k = 0$).

Пример. Вычислить первые 10 точек Q_i^* в трехмерном кубе. По табл. 6 (стр. 297) находим нужные значения $V_j^{(s)}$. Они написаны в табл. 2.

Вычисления по формуле (16) сведены в табл. 3.

Результаты в десятичной системе:

$$Q_1^* = (1/2, 1/2, 1/2), \quad Q_6^* = (3/8, 3/8, 5/8),$$

$$Q_2^* = (1/4, 3/4, 1/4), \quad Q_7^* = (7/8, 7/8, 1/8),$$

$$Q_3^* = (3/4, 1/4, 3/4), \quad Q_8^* = (1/16, 15/16, 11/16),$$

$$Q_4^* = (1/8, 5/8, 7/8), \quad Q_9^* = (9/16, 7/16, 3/16),$$

$$Q_5^* = (5/8, 1/8, 3/8), \quad Q_{10}^* = (5/16, 3/16, 15/16).$$

На рис. 67, а изображены точки Q_1^*, \dots, Q_{16}^* в квадрате, а на рис. 67, б — 16 «настоящих» случайных точек.

Замечание. Хотя табл. 6 на стр. 297 рассчитана на к.р. ≤ 13 , ее можно иногда использовать при любых к.р., даже к.р. $= \infty$. В качестве значений недостающих координат псевдослучайных точек можно выбирать обычные псевдослучайные числа γ , так что, например,

$$P_i = (q_{i,1}, \dots, q_{i,13}, \gamma_i^{(14)}, \dots, \gamma_i^{(r)}, \dots).$$

Целесообразно вычислять по $q_{i,j}$ наиболее существенные переменные в $\Phi(y_1, \dots, y_n, \dots)$, а по $\gamma_i^{(j)}$ — все остальные. Если в действительности существенных координат немного, то такой способ расчета может ускорить сходимость (по сравнению с расчетом по формуле (7)). (Двумерные точки, у которых одна координата случайная, а вторая детерминированная использовались в [18].)

Таблица 2

В десятичной системе					В двоичной системе				
$i \backslash s$	1	2	3	4	$j \backslash s$	1	2	3	4
1	1/2	1/4	1/8	1/16	1	0,1	0,01	0,001	0,0001
2	1/2	3/4	5/8	15/16	2	0,1	0,11	0,101	0,1111
3	1/2	1/4	7/8	11/16	3	0,1	0,01	0,111	0,1011

Таблица 3

i	$i_{ДВ}$	$q_{i,j}$	$q_{i,1}$	$q_{i,2}$	$q_{i,3}$
1	1	$V_j^{(1)}$	0,1	0,1	0,1
2	10	$V_j^{(2)}$	0,01	0,11	0,01
3	11	$V_j^{(1)} * V_j^{(2)}$	$0,1 * 0,01 =$ $= 0,11$	$0,1 * 0,11 =$ $= 0,01$	$0,1 * 0,01 =$ $= 0,11$
4	100	$V_j^{(3)}$	0,001	0,101	0,111
5	101	$V_j^{(1)} * V_j^{(3)}$	$0,1 * 0,001 =$ $= 0,101$	$0,1 * 0,101 =$ $= 0,001$	$0,1 * 0,111 =$ $= 0,011$
6	110	$V_j^{(2)} * V_j^{(3)}$	$0,01 * 0,001 =$ $= 0,011$	$0,11 * 0,101 =$ $= 0,011$	$0,01 * 0,111 =$ $= 0,101$
7	111	$V_j^{(1)} * V_j^{(2)} * V_j^{(3)}$	$0,1 * 0,01 *$ $* 0,001 = 0,111$	$0,1 * 0,11 *$ $* 0,101 = 0,111$	$0,1 * 0,01 *$ $* 0,111 = 0,001$
8	1000	$V_j^{(4)}$	0,0001	0,1111	0,1011
9	1001	$V_j^{(1)} * V_j^{(4)}$	$0,1 * 0,0001 =$ $= 0,1001$	$0,1 * 0,1111 =$ $= 0,0111$	$0,1 * 0,1011 =$ $= 0,0011$
10	1010	$V_j^{(2)} * V_j^{(4)}$	$0,01 * 0,0001 =$ $= 0,0101$	$0,11 * 0,1111 =$ $= 0,0011$	$0,01 * 0,1011 =$ $= 0,1111$

2.4.3. Общие требования. Если подходить к вопросу более строго, то от хороших псевдослучайных точек естественно потребовать, чтобы:

1° асимптотика D_N или φ_∞ была наилучшей (или хотя бы близкой к наилучшей);

2° константы в (12) или в (13) были наилучшими (или хотя бы достаточно малыми);

3° значения D_N/N или φ_∞/N были небольшими уже при больших N ;

4° алгоритм расчета этих точек на ЭВМ был достаточно простым.

К сожалению, проверить все эти требования в настоящее время невозможно, так как наилучшие значения констант (13) неизвестны (а для D_N неизвестен даже наилучший порядок роста).

Однако первому требованию в какой-то мере удовлетворяют точки P_i^* и вполне — точки Q_i^* . При небольших n второму и третьему требованию удовлетворяют точки Q_i^* . Наконец, время расчета точек Q_i^* того же порядка, что время расчета стандартных псевдослучайных точек Γ_i (если только имеется готовая таблица $V_j^{(s)}$). Для расчета точек P_i^* нужны лишь n простых чисел, но по сравнению со временем расчета Q_i^* время расчета P_i^* примерно в n раз больше.

2.5. Замечание о роли дисперсии. В расчетах, выполненных по точкам Q_i^* , фактическая ошибка часто оказывается на порядок меньше вероятной.

Пример. Рассмотрим значения $(K^31, \beta)_N$ при $N=2^m$, приведенные в табл. 1 гл. 5 (стр. 192). Условимся считать последнее значение, соответствующее $N=2^{14}$, точным. Фактические ошибки $\delta_N = |(K^31, \beta)_N - K^31, \beta_{2^{14}}|$ выписаны в табл. 4. Рядом приведены вероятные ошибки $r_N = 0,675 \sqrt{D\theta_3/N}$, где дисперсия $D\theta_3 = 57,3$. Очевидно, δ_N заметно меньше, чем r_N .

Таблица 4

v	$(K^31, 1)_N$	δ_N	r_N
2^8	8,3510	0,079	0,32
2^9	8,3381	0,066	0,22
2^{10}	8,2746	0,002	0,16
2^{11}	8,2840	0,012	0,11
2^{12}	8,2837	0,011	0,08
2^{13}	8,2770	0,005	0,06
2^{14}	8,2723	—	0,04

Не следует думать, что если дисперсия не определяет ошибку, то все присмы гл. 3, 4, направленные на уменьшение дисперсии, теряют смысл. Во-первых, алгоритмам с меньшей дисперсией отвечают функции Φ с меньшим изменением, кото-

рые, вообще говоря, интегрируются лучше.

Простейший результат в этом направлении для случая $n=1$ можно получить с помощью формулы (48) гл. 3, стр. 129. В самом деле, для любой функции $\Phi(y)$ класса $W_2'(L)$,

$$\left| \int_0^1 \Phi(y) dy - \frac{1}{N} \sum_{i=1}^N \Phi(\gamma_i) \right| \leq L \sqrt{\omega_{N'}^2/N}, \quad (17)$$

где можно считать, что

$$L^2 = \int_0^1 [\Phi'(y)]^2 dy.$$

Как доказано в [78], если $\Phi(y) \in W_2'(L)$, то дисперсия случайной величины $\Phi(\gamma)$ удовлетворяет неравенству

$$D\Phi(\gamma) \leq L^2/\pi^2. \quad (18)$$

Фиксируем числа $\gamma_1, \dots, \gamma_N$. Из (17) видно, что для тех функций, для которых L меньше, погрешность интегрирования будет меньше. Но таким функциям, согласно (18), соответствуют также меньшие значения дисперсии.

Во-вторых, алгоритмам с меньшей дисперсией часто отвечают более гладкие функции Φ , удовлетворяющие оценке (14). Например, при расчете задачи о поглощении нейтронов методом п. 1.1 гл. 6, осредняемая функция может принимать лишь два значения: 0 и 1. Нетрудно убедиться в том, что ее разрывы не параллельны координатным гиперплоскостям*). А при расчете той же задачи методом п. 3.3 (гл. 6), осредняемая функция непрерывна и даже дифференцируема.

Вообще говоря, можно ожидать большего ускорения сходимости за счет использования детерминированных псевдослучайных чисел тогда, когда используются более совершенные алгоритмы метода Монте-Карло.

*) Если область G_0 — шар $r^2 < R^2$, то из формулы $r_{i+1} = r_i + \xi_i \Omega_i$ следует условие вылета

$$r_{i+1}^2 = r_i^2 + 2\xi_i(r_i, \Omega_i) + \xi_i^2 > R^2.$$

Поэтому одна из поверхностей разрыва осредняемой функции определяется уравнением

$$\xi_i^2 + 2\xi_i(r_i, \Omega_i) + r_i^2 - R^2 = 0.$$

В случае изотропного рассеяния $(r_i, \Omega_i) = |r_i|/(2\gamma - 1)$, длина пробега $\xi_i = -(1/\Sigma) \ln \gamma'$, так что уравнение это связывает γ с γ' .

§ 3. Поиски «универсальных» псевдослучайных чисел

В предыдущем параграфе были указаны детерминированные точки, координаты которых можно использовать в качестве псевдослучайных чисел при реализации алгоритмов Монте-Карло с конечными к. р. Однако весьма часто встречаются также алгоритмы с к. р. $= \infty$. Рассмотрим некоторые попытки построить детерминированные псевдослучайные числа, пригодные для расчета задач с любыми к. р. Мы говорим о попытках, так как нельзя считать, что эти поиски уже закончены: до сих пор нет последовательностей, удовлетворяющих требованиям практики (типа описанных в п. 2.4.3), да и сами эти требования не вполне четко сформулированы.

3.1. Использование бесконечномерных точек. Такой подход к проблеме был предложен Н. Н. Ченцовым в 1961 г. [97]. На этом пути удалось указать весьма широкие классы функций $\Phi(y_1, \dots, y_n, \dots)$, зависящих от бесконечного числа переменных y_1, \dots, y_n, \dots ($0 < y_n < 1$) и такие классы последовательностей P_1, \dots, P_i, \dots , состоящих из точек бесконечномерного единичного куба

$$P_i = (y_{i,1}, \dots, y_{i,n}, \dots),$$

что для каждой функции класса

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \Phi(P_i) = \int_0^1 \dots \int_0^1 \dots \Phi(y_1, \dots, y_n, \dots) \prod_{k=1}^{\infty} dy_k,$$

и порядок сходимости лучше, чем $1/N^{1-\varepsilon}$ с любым $\varepsilon > 0$. В частности, этим классам функций принадлежат любые достаточно гладкие функции $\Phi(y_1, \dots, y_n)$, зависящие от любого конечного числа переменных n . Поэтому при реализации всех упомянутых в § 1 алгоритмов с к. р. $= \infty$ можно пытаться вместо псевдослучайных чисел использовать координаты точек P_i .

Пока известны лишь две конкретные бесконечномерные последовательности с более или менее «хорошими» свойствами. Одна из них — *обобщенная ЛП-последовательность* — вычисляется по формулам (15), (16), где $1 \leq j < \infty$. К сожалению, достаточно простых формул для расчета всех элементов $V_j^{(s)}$ нет.

Вторая последовательность, впервые рассмотренная в [77], называется *обобщенной последовательностью Холтона*. Точки этой последовательности P_i^* , $i=1, 2, \dots$, имеют координаты

$$P_i^* = (p_{r_1}(i), p_{r_2}(i), \dots, p_{r_n}(i), \dots),$$

где $r_1 < r_2 < \dots < r_n < \dots$ — последовательность всех простых чисел.

Для расчета точек P_i^* на практике можно задать достаточно большую таблицу простых чисел и (или) запрограммировать какой-нибудь алгоритм их нахождения (например, метод решета). Первые две точки этой последовательности:

$$P_1^* = \left(\frac{1}{2}, \frac{1}{3}, \frac{1}{5}, \frac{1}{7}, \dots, \frac{1}{r_n}, \dots \right),$$

$$P_2^* = \left(\frac{1}{4}, \frac{2}{3}, \frac{2}{5}, \frac{2}{7}, \dots, \frac{2}{r_n}, \dots \right).$$

Если мы хотим использовать точки P_i^* для расчета задачи о поглощении нейтронов (гл. 6, п. 1.1), то для построения первой траектории надо вместо случайных чисел использовать значения $1/2, 1/3, 1/5, \dots$, для второй — $1/4, 2/3, 2/5, \dots$, и т. д. Пример такого расчета имеется в [80].

3.2. Вполне равномерно распределенные последовательности чисел.

Определение. Последовательность чисел x_1, \dots, x_i, \dots , принадлежащих интервалу $(0, 1)$, называется *вполне равномерно распределенной*, если при каждом натуральном n последовательность точек

$$(x_1, \dots, x_n), (x_{n+1}, \dots, x_{2n}), (x_{2n+1}, \dots, x_{3n}), \dots \quad (19)$$

равномерно распределена в K^n . Понятие это было введено Н. М. Коробовым в 1949 г. [41].

Из п. 2.1 вытекает, что если функция $\Phi(y_1, \dots, y_n)$ интегрируема по Риману в K^n , то

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=0}^{N-1} \Phi(x_{in+1}, \dots, x_{in+n}) = \int_{K^n} \Phi(P) dP, \quad (20)$$

и соотношение это справедливо для любого натурального n или, другими словами, для функций Φ от любого ко-

печного числа аргументов. Следовательно, вполне равномерно распределенные последовательности чисел можно (в принципе) использовать для практической реализации алгоритмов Монте-Карло с любыми конечными к.р., а в некоторых случаях и с к.р. $= \infty$ [99]. К сожалению, до сих пор неизвестно ни одной вполне равномерно распределенной последовательности, которая в какой-то степени удовлетворяла бы требованиям п. 2.4.3 при различных n .

Пример [126] Известно, что существуют бесконечно много таких чисел $\alpha > 1$, что последовательность дробных долей $x_i = D(\alpha^i)$, $i = 1, 2, \dots$, вполне равномерно распределена*). Легко доказать, что такие α обязаны быть трансцендентными числами. Но до сих пор ни одного конкретного значения α не найдено.

Можно доказать, что вышеприведенное определение вполне равномерно распределенных последовательностей эквивалентно следующему.

Определение. Последовательность чисел $x_1, \dots, x_i, \dots, x_n, \dots$, принадлежащих интервалу $(0, 1)$, называется *вполне равномерно распределенной*, если при каждом натуральном n последовательность точек

$$(x_1, \dots, x_n), (x_2, \dots, x_{n+1}), (x_3, \dots, x_{n+2}), \dots \quad (21)$$

равномерно распределена в K^n .

Из этого определения и п. 2.1 следует, что если $\Phi(y_1, \dots, y_n)$ интегрируема по Риману в K^n , то, каково бы ни было натуральное n ,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=0}^{N-1} \Phi(x_{i+1}, \dots, x_{i+n}) = \int_{K^n} \Phi(P) dP. \quad (22)$$

Может показаться, что расчет по формуле (22) выгоднее, чем расчет по формуле (20), так как при этом затрачивается примерно в n раз меньше чисел x_i . Несомненно однако, что вычислители, инстинктивно стремясь к «независимости» псевдослучайных точек, отдадут предпочтение точкам (19) и формуле (20). Вполне вероятно, что они правы: может оказаться, что для точек (21) отклонение D_N (при любом n ?) растет быстрее, чем для точек (19). Вопрос этот не исследован.

*) Таким свойством обладают почти все $\alpha > 1$,

3.3. Асимптотически вполне равномерно распределенные последовательности чисел. Рассмотрим семейство последовательностей x_1, \dots, x_i, \dots , зависящее от натурального параметра $g (g=1, 2, \dots)$:

$$x_1(g), x_2(g), \dots, x_i(g), \dots \quad (23)$$

Семейство (23) естественно назвать *асимптотически вполне равномерно распределенным*, если для любой функции $\Phi(y_1, \dots, y_n)$, интегрируемой по Риману в K^n , при любом натуральном n справедливо равенство

$$\lim_{g \rightarrow \infty} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=0}^{N-1} \Phi(x_{in+1}(g), \dots, x_{in+n}(g)) = \int_{K^n} \Phi(P) dP, \quad (24)$$

или равенство

$$\lim_{g \rightarrow \infty} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=0}^{N-1} \Phi(x_{i+1}(g), \dots, x_{i+n}(g)) = \int_{K^n} \Phi(P) dP. \quad (25)$$

Термин этот введен И. Франклином в 1963 г. [126]. Ему же принадлежит второе утверждение следующей теоремы, первое утверждение которой доказано в [83].

Теорема. *Существует бесконечно много чисел α , $0 < \alpha < 1$, таких, что семейство последовательностей*

$$x_i(g) = D(\alpha g^i), \quad i=1, 2, \dots$$

асимптотически вполне равномерно распределено как в смысле (24), так и в смысле (25).*

Интересно отметить, что при каждом фиксированном g последовательность $x_i = D(\alpha g^i)$ может быть равномерно распределенной в K^1 , но точки (x_{2i-1}, x_{2i}) не могут быть равномерно распределенными в K^2 . В самом деле, точки $(x_1, x_2), (x_3, x_4), \dots$ можно записать в виде $(x_1, D(gx_1)), (x_3, D(gx_3)), \dots$. Отсюда видно, что все они расположены на линии $y = D(gx)$, которая состоит из g отрезков прямых $y = gx - k$ при

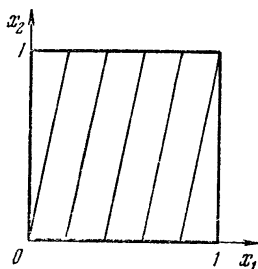


Рис. 69.

* Таким свойством обладают почти все α из интервала $(0, 1)$.

$k=0, \dots, g-1$ (рис. 69). Тем более последовательность x_i не может быть вполне равномерно распределенной.

Сформулированная теорема в какой-то мере объясняет, почему метод сравнений (п. 2.2.2 гл. 1) позволяет строить «хорошие» псевдослучайные числа и почему приходится выбирать большие g : из (7) гл. 1 вытекает, что $\gamma_i = D(\gamma_0 g^i)$. Однако никаких выводов о том, как выбирать «хорошие» g , M , m_0 отсюда сделать нельзя.

Как доказал С. М. Ермаков [33], скорость сходимости в формуле (24) (но не в (25)!) связана с $\sqrt{D\Phi(\Gamma)/N}$.

§ 4. Проверка псевдослучайных чисел с детерминистической точки зрения

4.1. Интерпретация простейших критериев. В гл. 1, § 3 мы рассмотрели некоторые тесты, используемые для проверки псевдослучайных чисел, и отметили, что все такие тесты только необходимы, но не достаточны: они могут опровергнуть гипотезу о том, что $\gamma_1, \dots, \gamma_N$ — независимые значения случайной величины γ , но не могут доказать ее. Сузим теперь постановку вопроса: вместо «можно ли числа $\gamma_1, \dots, \gamma_N$ считать независимыми значениями γ ?» зададим вопрос: можно ли числа $\gamma_1, \dots, \gamma_N$ использовать вместо независимых значений γ в формуле

$$\int_0^1 \Phi(x) dx \approx \frac{1}{N} \sum_{i=1}^N \Phi(\gamma_i) \quad (26)$$

при любых $\Phi(x)$ из достаточно широкого класса?

При такой постановке вопроса рассмотренные в гл. 1 критерии оказываются и необходимыми и достаточными — каждый на своем классе функций $\Phi(x)$. Впервые такая интерпретация была изложена автором в 1968 г.

4.1.1. Критерий ω^2 . Из формулы (48) стр. 129 вытекает, что

$$\sup_{\Phi \in W_2(L)} \left| \int_0^1 \Phi(x) dx - \frac{1}{N} \sum_{i=1}^N \Phi(\gamma_i) \right| = L \sqrt{\frac{\omega_N^2}{N}}. \quad (27)$$

Следовательно, малость величины ω_N^2 необходима и достаточна для того, чтобы ошибка приближения (26) была мала для всех функций класса $W_2'(L)$.

4.1.2. Критерий χ^2 . Разобьем интервал $(0, 1)$ на r интервалов $X_1 + \dots + X_r = (0, 1)$, длины которых равны p_1, \dots, p_r так, что $p_1 + \dots + p_r = 1$ и все $p_j > 0$. Обозначим через v_j количе-

ство значений γ_i при $1 \leq i \leq N$, принадлежащих X_j . Критерий χ^2 основан на асимптотическом распределении величины

$$\chi_N^2 = \sum_{j=1}^r \frac{N}{p_j} \left(p_j - \frac{v_j}{N} \right)^2, \quad (28)$$

которое не зависит от X_1, \dots, X_r , а только от числа степеней свободы $r - 1$.

Введем множество функций $A_1(L)$, постоянных на каждом из X_j , т. е. функций вида

$$\Phi(x) \equiv c_j \text{ при } x \in X_j, \quad 1 \leq j \leq r,$$

и таких, что

$$\sum_{j=1}^r p_j c_j^2 \leq L^2. \quad (29)$$

Нетрудно вычислить, что если $\Phi(x) \in A_1(L)$, то

$$\int_0^1 \Phi dx - \frac{1}{N} \sum_{i=1}^N \Phi(\gamma_i) = \sum_{j=1}^r p_j c_j - \frac{1}{N} \sum_{j=1}^r c_j v_j = \sum_{j=1}^r c_j \left(p_j - \frac{v_j}{N} \right).$$

Используя известное неравенство $(\sum u_j v_j)^2 \leq \sum u_j^2 \sum v_j^2$, получим оценку

$$\left| \int_0^1 \Phi dx - \frac{1}{N} \sum_{i=1}^N \Phi(\gamma_i) \right| \leq \left\{ \sum_{j=1}^r p_j c_j^2 \right\}^{1/2} \left\{ \sum_{j=1}^r \frac{1}{p_j} \left(p_j - \frac{v_j}{N} \right)^2 \right\}^{1/2} \leq \\ \leq L \sqrt{\frac{\chi_N^2}{N}},$$

причем оценка эта точная, так как при

$$c_j = \frac{L}{p_j} \left(p_j - \frac{v_j}{N} \right) \left\{ \sum_{j=1}^r \frac{1}{p_j} \left(p_j - \frac{v_j}{N} \right)^2 \right\}^{-1/2}$$

последнее неравенство обращается в равенство. Итак,

$$\sup_{\Phi \in A_1(L)} \left| \int_0^1 \Phi(x) dx - \frac{1}{N} \sum_{i=1}^N \Phi(\gamma_i) \right| = L \sqrt{\frac{\chi_N^2}{N}}. \quad (30)$$

Отсюда видно, что малость величины χ_N^2 необходима и достаточна для того, чтобы ошибка приближения (26) была мала для всех функций класса $A_1(L)$.

Мы знаем, однако, что к формуле (26) сводятся лишь такие алгоритмы Монте-Карло, у которых к.р.=1. Стало быть, рассмотренные критерии (как и критерий Колмогорова, см. выражение 8

гл. 1 и упражнение 6 гл. 7) гарантируют применимость проверенных псевдослучайных чисел $\gamma_1, \dots, \gamma_N$ только для расчетов по алгоритмам с к.р. = 1.

4.2. Многомерные критерии. Если мы хотим использовать эти же числа для расчета по алгоритмам с к.р. = n , то необходимо проверить их с помощью какого-либо n -мерного теста. Например, разобьем куб K^n на r областей $X_1 + \dots + X_r = K^n$, объемы которых равны p_1, \dots, p_r , так что $p_1 + \dots + p_r = 1$ и все $p_j > 0$. Из чисел

$\gamma_1, \dots, \gamma_N$ образуем $N_n = \Pi(N/n)$ точек с координатами

$$(\gamma_1, \dots, \gamma_n), (\gamma_{n+1}, \dots, \gamma_{2n}), \dots, (\gamma_{n(N_n-1)+1}, \dots, \gamma_{nN_n}),$$

и пусть v_j — количество таких точек, принадлежащих X_j .

Тот же критерий χ^2 с $(r-1)$ -й степенью свободы позволяет проверить распределение этих точек в K^n с помощью величины

$$\chi_{N_n}^2 = \sum_{j=1}^r \frac{N_n}{p_j} \left(p_j - \frac{v_j}{N_n} \right)^2.$$

Обозначим через $A_n(L)$ множество функций $\Phi(x_1, \dots, x_n)$, постоянных на каждом X_j , т. е. функций вида

$$\Phi(x_1, \dots, x_n) \equiv c_j \text{ при } (x_1, \dots, x_n) \in X_j, \quad 1 \leq j \leq r,$$

и таких, что выполнено условие (29). Повторяя вычисления п. 4.1.2, получим, что

$$\sup_{\Phi \in A_n(L)} \left| \int_{K_n} \Phi(P) dP - \frac{1}{N_n} \sum_{i=0}^{N_n-1} \Phi(\gamma_{in+1}, \dots, \gamma_{in+n}) \right| = L \sqrt{\frac{\chi_{N_n}^2}{N_n}}. \quad (31)$$

Таким образом, малость $\chi_{N_n}^2$ действительно гарантирует применимость чисел $\gamma_1, \dots, \gamma_N$ для расчета по некоторым алгоритмам с к.р. = n .

Конечно, класс алгоритмов, которым отвечают функции $\Phi(x_1, \dots, x_n)$ из $A_n(L)$ весьма узок. Однако если r велико и все X_j малы, то многие функции допускают хорошее приближение кусочно постоянными функциями из $A_n(L)$. И тогда псевдослучайные числа, пригодные для интегрирования функций из $A_n(L)$ окажутся пригодными для интегрирования гораздо более разнообразных функций.

4.3. О системах тестов. С точки зрения изложенного в пп. 4.1 и 4.2, система тестов, рекомендованная в гл. 1, представляется логичной, но довольно слабой. В самом деле, она включает простейшие проверки одномерного и двумерного распределения γ_i и заодно распределения некоторых попутно вычисляемых второстепенных величин. Лишь проверка серий может рассматриваться как тест,

рассчитанный на алгоритмы с к.р. = ∞ , правда, алгоритмы весьма специального вида

Если бы речь шла только об использовании алгоритмов с любыми к.р. $\leq n$, то можно было бы предложить очень хороший тест — расчет отклонения D_N или, другими словами, n -мерный критерий Колмогорова (см. упражнения 8 гл. 1 и упражнение 6 гл. 7). Правда, вычисление D_N при больших N очень трудоемко.

Обеспечить же эмпирической проверкой «универсальность» псевдослучайных чисел, по-видимому, крайне трудно. Например, без теоретических результатов, подобных результатам § 3, пришлось бы проверять n -мерное распределение не один раз, а n раз, так как из того, что точки $(\gamma_1, \dots, \gamma_n), (\gamma_{n+1}, \dots, \gamma_{2n}) \dots$ хорошо распределены, не следует, что точки $(\gamma_j, \dots, \gamma_{j+n}), (\gamma_{j+n+1}, \dots, \gamma_{j+2n}) \dots$ при $1 < j < n$ также хороши. Это видно из следующего простого примера: в последовательности 1111000011110000 ... пары 11, 11, 00, 00, 11, 11, 00, 00 распределены неравномерно (так как комбинаций 01 и 10 совсем нет); отбросив первую единицу последовательности, получим пары 11, 10, 00, 01, 11, 10, 00, 01, ... , где каждая из возможных комбинаций встречается одинаково часто.

Результаты этого параграфа содержатся в статье [84].

Упражнения к главе 7

1. Рассмотрим пример п 1.2, где значения ξ вычисляются методом Неймана. Ограничим число проб величиной l и условимся полагать $\Phi=0$, если $cy'_1 \geq p(g_1), \dots, cy'_l \geq p(g_l)$. Для такого алгоритма к.р. = $2l$.

Доказать, что для того, чтобы погрешность такого приближения не превосходила вероятной ошибки метода $r_N = 0,675 \sqrt{Df(\xi)/N}$, должно выполняться неравенство

$$l \geq \ln(r_N/l) \ln^{-1}(1 - \varepsilon).$$

2. Выберем в кубе K^n равномерную кубическую сетку, состоящую из $N = M^n$ точек с координатами

$$\left(\frac{i_1 - 1/2}{M}, \dots, \frac{i_n - 1/2}{M} \right),$$

где $i_1, \dots, i_n = 1, 2, \dots, M$. Доказать, что отклонение этой сетки равно $D_N = 0,5N^{1-1/n}$ (при $n=1$ величина $D_N = 0,5$ минимальна, но при больших n отношение D_N/N убывает крайне медленно)

3. Вычислить точку Q_{13}^* в 4-мерном кубе. Ответ: $Q_{13}^* = (11/16, 13/16, 13/16, 15/16)$.

4. Используя равенства $V_1^{(s)} = 2^{-s}$ при $s=1, 2, \dots$, доказать, что первые координаты $q_{i,1}$ точек Q_i совпадают с числами $p_2(i)$.

5. Какова бы ни была последовательность чисел x_1, \dots, x_i, \dots из интервала $(0, 1)$, равенство

$$\int_0^1 f(x) dx - \frac{1}{N} \sum_{i=1}^N f(x_i) = o\left(\frac{1}{N}\right)$$

для $f=x$ и $f=x^2$ одновременно невозможно. (Н. Н. Ченцов).

6. Обозначим $W'_1(L)$ множество непрерывных функций $f(x)$, определенных при $0 < x < 1$ и имеющих кусочно непрерывные производные $f'(x)$ такие, что

$$\int_0^1 |f'(x)| dx \leq L,$$

Доказать, что для любых чисел $\gamma_1, \dots, \gamma_N$

$$\sup_{\Phi \in W'_1(L)} \left| \int_0^1 \Phi(x) dx - \frac{1}{N} \sum_{i=1}^N \Phi(\gamma_i) \right| = L \frac{K_N}{\sqrt{N}},$$

где $K_N = \sqrt{N}D$ — величина, фигурирующая в критерии Колмогорова (упражнение 8 гл. 1) (И. М. Соболев [84]).

ГЛАВА 8

НЕКОТОРЫЕ ДРУГИЕ ЗАДАЧИ

§ 1. Интерполирование функций от большого числа переменных

В книге уже встречались задачи, в которых методы Монте-Карло оказываются эффективнее классических методов при большом числе переменных (см. упражнение 9 гл. 3 и п. 5.6 гл. 5). Здесь изложена одна задача такого типа, рассмотренная впервые Дж. Хэммерсли [134].

1.1. Постановка задачи.

Рассмотрим функцию $f(x_1, \dots, x_n)$, значения которой известны лишь в вершинах единичного n -мерного куба K^n . Требуется проинтерполировать значение этой функции в точке (x_1, \dots, x_n) , расположенной внутри K^n (рис. 70). Интерполяция линейная по каждому из переменных.

В случае $n=1$ интерполяционная формула всем хорошо знакома:

$$f(x_1) = (1-x_1)f(0) + x_1f(1).$$

Нетрудно проверить, что при $n=2$ получается аналогичная формула:

$$f(x_1, x_2) = (1-x_1)(1-x_2)f(0,0) + x_1(1-x_2)f(1,0) + \\ + (1-x_1)x_2f(0,1) + x_1x_2f(1,1). \quad (1)$$

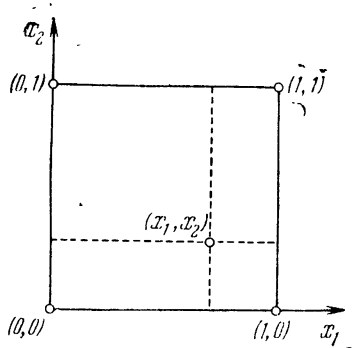


Рис. 70.

Чтобы записать эту формулу в компактной форме, введем индексы вершин k_1 и k_2 , каждый из которых может принимать два значения -0 и 1 . Пусть

$$c(k_i) = \begin{cases} 1 - x_i & \text{при } k_i = 0, \\ x_i & \text{при } k_i = 1. \end{cases} \quad (2)$$

Тогда (1) можно переписать в виде

$$f(x_1, x_2) = \sum_{k_1, k_2=0}^1 c(k_1) c(k_2) f(k_1, k_2).$$

Точно так же выглядит интерполяционная формула в случае n переменных:

$$f(x_1, \dots, x_n) = \sum_{k_1, \dots, k_n=0}^1 c(k_1) \dots c(k_n) f(k_1, \dots, k_n). \quad (3)$$

Расчет по формулам (2) и (3) в принципе весьма прост. Однако количество слагаемых в (3) быстро растет с ростом n и уже при $n=30$ превышает 10^9 .

Может показаться, что для больших n такая постановка задачи нереальна: невозможно хранить 2^{30} значений функции f . В действительности это и не нужно: достаточно иметь алгоритм, позволяющий вычислить значения f в вершинах K^n .

1.2. Метод Монте-Карло. Введем n независимых случайных величин $\xi^{(1)}, \dots, \xi^{(n)}$, каждая из которых может принимать два значения -0 и 1 , и пусть

$$P\{\xi^{(i)}=0\}=1-x_i, P\{\xi^{(i)}=1\}=x_i. \quad (4)$$

Совокупность этих величин $(\xi^{(1)}, \dots, \xi^{(n)})$ определяет случайную вершину куба K^n .

Теорема 1. Математическое ожидание $f(\xi^{(1)}, \dots, \xi^{(n)})$ равно

$$Mf(\xi^{(1)}, \dots, \xi^{(n)}) = f(x_1, \dots, x_n). \quad (5)$$

Доказательство. Из (2) и (4) нетрудно заметить, что и при $k_i=0$, и при $k_i=1$ справедливо равенство $P\{\xi^{(i)}=k_i\}=c(k_i)$. Поэтому правая часть обычной

формулы для математического ожидания

$$\begin{aligned} Mf(\xi^{(1)}, \dots, \xi^{(n)}) &= \\ &= \sum_{k_1, \dots, k_n=0}^1 f(k_1, \dots, k_n) P\{\xi^{(1)} = k_1, \dots, \xi^{(n)} = k_n\} \end{aligned}$$

превращается в правую часть формулы (3), откуда сразу вытекает (5).

Соответствующий формуле (5) метод Монте-Карло: при больших N

$$f(x_1, \dots, x_n) \approx \frac{1}{N} \sum_{s=1}^N f(\xi_s^{(1)}, \dots, \xi_s^{(n)}),$$

где $(\xi_1^{(1)}, \dots, \xi_1^{(n)}), \dots, (\xi_N^{(1)}, \dots, \xi_N^{(n)})$ — независимые реализации случайной величины $(\xi^{(1)}, \dots, \xi^{(n)})$ (или, другими словами, набор случайных вершин куба).

В последней формуле легко выразить все $\xi^{(i)}$ через случайные числа γ , так как $\xi^{(i)} = e(x_i - \gamma)$. Получим формулу

$$f(x_1, \dots, x_n) \approx \frac{1}{N} \sum_{s=1}^N f(e(x_1 - \gamma_{1,s}), \dots, e(x_n - \gamma_{n,s})), \quad (6)$$

где все $\gamma_{i,s}$ — независимые случайные числа. Формула (6) позволяет интерполировать значения $f(x_1, \dots, x_n)$ в нескольких точках, используя одни и те же случайные числа $\gamma_{i,s}$ (ср. гл. 3, § 4).

В статьях [179] и [177] построены методы Монте-Карло для экстраполирования и нелинейной интерполяции функции $f(x_1, \dots, x_n)$

§ 2. Простейший случайный поиск

2.1. Случайный поиск. Рассмотрим ограниченную кусочно непрерывную функцию $\Phi(P)$, определенную в замкнутом единичном n -мерном кубе $\overline{K^n}$, так что $\overline{K^n}$ — это куб K^n вместе с его границей. Обозначим через \hat{P} точку абсолютного максимума $\Phi(P)$ в $\overline{K^n}$, т. е. такую точку, что $\Phi(\hat{P}) \geq \Phi(P)$ для всех $P \in \overline{K^n}$.

Простейший поиск точки \hat{P} состоит в том, что в кубе задается произвольная последовательность точек P_1, \dots

\dots, P_N, \dots , называемых пробными точками; в каждой из этих точек вычисляется значение $\Phi(P_i)$; и находится точка P_{i_0} , в которой

$$\Phi(P_{i_0}) = \max_{1 \leq i \leq N} \Phi(P_i).$$

Точка P_{i_0} служит приближением к точке \hat{P} .

Рассмотрим произвольную область $B \subset K^n$ с положительным n -мерным объемом $V_B > 0$, содержащую точку \hat{P} . Если при $N \rightarrow \infty$ хотя бы одна пробная точка попадает в B , то говорят, что *процесс поиска сходится*.

Докажем, что если в качестве пробных точек выбирать независимые случайные точки с плотностью $p(P) > 0$ внутри K^n , то процесс поиска сходится. Для доказательства фиксируем произвольную область B с $V_B > 0$, содержащую точку \hat{P} и обозначим через p_B вероятность того, что одна пробная точка попадет в B .

Очевидно,

$$p_B = \int_B p(P) dP > 0.$$

Вероятность того, что хотя бы одна из N пробных точек окажется в B , равна $1 - (1 - p_B)^N$ и стремится к 1, когда $N \rightarrow \infty$.

Следовательно, какой бы коэффициент доверия β мы ни выбрали, при достаточно большом N с вероятностью, большей чем β , хотя бы одна пробная точка попадет в B .

Если никакой предварительной информации о расположении \hat{P} нет, то естественно выбрать $p(P) \equiv 1$, т. е. использовать пробные точки Γ_i , равномерно распределенные в K^n . Такой поиск называют *простейшим* случайным поиском (или слепым поиском).

Ясно, что процесс поиска можно улучшить, если в ходе поиска менять плотность $p(P)$ с учетом уже полученных значений. Например, точку P_i выбирать по плотности $p_i(P)$, которая строится с учетом значений $\Phi(P_1), \dots, \Phi(P_{i-1})$. На таких более совершенных (но и более сложных) алгоритмах поиска мы здесь останавливаться не будем [22, 66, 72, 184]. Отметим только некоторые ситуации, в которых простейший случайный поиск весьма полезен.

а) Функция $\Phi(P)$ достаточно гладкая, но многоэкстремальная. Простейший случайный поиск целесообразно использовать для отбора начальных точек, из которых можно локальными методами (метод градиентов, наискорейший спуск и др.) попасть в ближайший максимум. Чем тщательнее выбор начальных точек, тем меньше шансов пропустить абсолютный максимум.

б) Требуется найти максимумы нескольких функций $\Phi_1(P), \dots, \Phi_m(P)$. При простейшем случайном поиске можно одновременно (по одним и тем же пробным точкам G_i) искать максимумы (и минимумы) всех этих функций.

Такая ситуация довольно типична для задач оптимального конструирования: обычно качество конструкции можно оценивать по ряду весьма различных критериев, и выбор решающего (или компромиссного) критерия удастся сделать лишь тогда, когда известно, чего можно добиться, оптимизируя тот или иной критерий в отдельности.

2.2. ЛП-поиск. В полном согласии с идеями гл. 7, можно попытаться использовать в качестве пробных точек для простейшего поиска любые точки, образующие равномерно распределенную последовательность в K^n (гл. 7, п. 2.1): чем более равномерно распределены точки, тем лучших результатов можно ожидать от поиска.

Сходимость такого поиска легко доказать: так как для любой области B при $N \rightarrow \infty$ отношение $S_N(B)/N \rightarrow V_B$ (см. (10) гл. 7), то $S_N(B) \sim NV_B$; следовательно, количество пробных точек, попавших в B , неограниченно возрастает с ростом N (если только $V_B > 0$).

ЛП-поиском называется простейший поиск, в котором пробными точками служат точки Q_1, \dots, Q_N, \dots , образующие ЛП-последовательность (п. 2.4. гл. 7).

Сравнение ЛП-поиска с простейшим случайным поиском проводилось на многих задачах. И неизменно ЛП-поиск оказывался более эффективным.

Пример В табл. 1 приведены результаты поиска максимума некоторой (достаточно сложной) функции $\Phi(P)$ от 9 переменных [86]. Обозначения: $\Gamma_{(1)}, \Gamma_{(2)}, \Gamma_{(3)}$ — три наилучшие точки, полученные при случайном поиске, а $Q_{(1)}^*, Q_{(2)}^*, Q_{(3)}^*$ — три наилучшие точки полученные при ЛП-поиске с тем же количеством N пробных точек

2.3. Поиск в произвольной конечной области. Если функция $\Phi(P)$ определена в конечной замкнутой области \bar{G} , то для реализации простейшего случайного поиска надо выбирать случайные точки (ξ_1, \dots, ξ_n) , равномерно

Таблица 1

N	$\Phi(Q_{(1)}^*)$	$\Phi(Q_{(2)}^*)$	$\Phi(Q_{(3)}^*)$	$\Phi(\Gamma_{(1)})$	$\Phi(\Gamma_{(2)})$	$\Phi(\Gamma_{(3)})$
512	48,12	47,00	37,84	42,29	39,52	38,51
1024	48,61	48,12	47,83	42,29	40,21	39,52
2048	48,61	48,12	47,83	43,52	42,29	40,22

распределенные в G . Делается это с помощью преобразований гл. 2, которые можно записать в форме $\xi_k = g_k(\gamma_1, \dots, \gamma_m)$, $1 \leq k \leq n$, где, вообще говоря, $m \geq n$.

Обозначим через $\Gamma = (\gamma_1, \dots, \gamma_m)$ случайную точку, равномерно распределенную в K^m . Пусть $\Gamma_1, \dots, \Gamma_N, \dots$ — независимые значения Γ . Тогда i -я пробная точка в G имеет координаты

$$(g_1(\Gamma_i), \dots, g_n(\Gamma_i)), \quad i=1, 2, \dots, N, \dots$$

Эти же преобразования позволяют осуществить в области G и ЛП-поиск. При ЛП-поиске координаты i -й пробной точки равны

$$(g_1(Q_i^*), \dots, g_n(Q_i^*)), \quad i=1, 2, \dots, N, \dots,$$

где функции g_k — те же, а $Q_1^*, \dots, Q_N^*, \dots$ — точки m -мерной ЛП $_{\tau}$ -последовательности.

§ 3. Решение уравнения Лапласа

3.1. Построение случайных траекторий. Пусть задана ограниченная связная область G и точка $P_0 \in G$. Определим случайную траекторию $Q_0 \rightarrow Q_1 \rightarrow \dots \rightarrow Q_n \rightarrow \dots$ следующим образом: положим $Q_0 = P_0$; далее, если точка Q_n известна, то построим окружность произвольного радиуса l_n , расположенную внутри G , и на этой окружности выберем случайную точку Q_{n+1} (рис. 71). Таким образом,

$$Q_{n+1} = Q_n + l_n \omega_n, \quad n=0, 1, 2, \dots,$$

где $\omega_n = \{\cos \varphi_n, \sin \varphi_n\}$, и угол φ_n равномерно распределен в интервале $(0, 2\pi)$.

Теорема 2. Если функция $u(P) \equiv u(x, y)$ удовлетворяет в области G уравнению Лапласа

$$\partial^2 u / \partial x^2 + \partial^2 u / \partial y^2 = 0, \quad (7)$$

то при каждом n и при любых l_0, \dots, l_n математическое ожидание $Mu(Q_{n+1})$ равно значению $u(P_0)$ в начале траектории.

Доказательство. Придадим более точный смысл утверждению о произвольности радиуса l_n . Будем считать, что задана некоторая плотность $q_n(l)$, которая тождественно равна нулю при всех l , превосходящих минимальное расстояние от Q_n до границы G^0 , а также при $l < 0$; случай $q_n(l) = \delta(l - l_n)$ также допускается; и выбор l_n осуществляется в соответствии с плотностью $q_n(l)$.

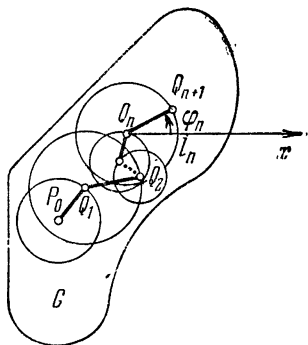


Рис. 71.

Пусть $p_n(P)$ — плотность распределения точки Q_n в G . Тогда математическое ожидание величины $u(Q_{n+1}) \equiv u(Q_n + l_n \omega_n)$ равно

$$Mu(Q_{n+1}) = \int_G p_n(P) dP \int_{-\infty}^{\infty} q_n(l) dl \int_0^{2\pi} u(P + l\omega) \frac{d\varphi}{2\pi}.$$

По известной теореме о среднем значении гармонической функции [88]

$$\frac{1}{2\pi} \int_0^{2\pi} u(P + l\omega) d\varphi = u(P).$$

Поэтому

$$Mu(Q_{n+1}) = \int_G u(P) p_n(P) dP = Mu(Q_n).$$

При $n=0$ точка $Q_0 \equiv P_0$, $u(Q_0) \equiv u(P_0)$ и $Mu(Q_0) \equiv u(P_0)$. Применяя индукцию, получим утверждение теоремы.

Построение траекторий рассмотренного типа в трехмерном случае иногда называют *блужданиями по сферам*.

3.2. Решение задачи Дирихле для уравнения Лапласа. Траекторию предыдущего пункта можно использовать для приближенного решения задачи Дирихле (гл. 5, п. 5.5). Пусть на границе G^0 области G задана ограниченная функция $g(P)$. Обозначим через $u(P)$ искомое решение, удовлетворяющее внутри G уравнению (7) и обращающееся в $g(P)$ при $P \in G^0$.

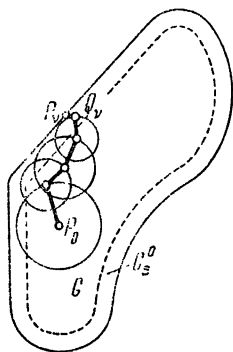


Рис. 72.

Фиксируем достаточно малую окрестность G_ϵ^0 границы G^0 (рис. 72). Чтобы вычислить $u(P_0)$, будем строить траектории вида $P_0 \rightarrow Q_1 \rightarrow \dots \rightarrow Q_v$ до тех пор, пока случайная точка Q_v не попадает в G_ϵ^0 . Пусть P_v — ближайшая к Q_v точка границы G^0 . Можем считать, что значение случайной величины $u(Q_v)$ приближенно равно $u(Q_v) \approx g(P_v)$. Построив N траекторий такого типа, получим значения $g(P_{v_1}), \dots, g(P_{v_N})$, по которым оценивается искомое решение

$$u(P_0) \approx \frac{1}{N} \sum_{s=1}^N g(P_{v_s}).$$

Заметим, что сходимость по вероятности

$$\frac{1}{N} \sum_{s=1}^N u(Q_{v_s}) \xrightarrow{P} u(P_0), \quad (8)$$

когда $N \rightarrow \infty$ не вытекает из теоремы Хинчина (стр. 87), ибо в сумме (8) фигурируют N различных случайных величин, различающихся правилами выбора l_0, l_1, \dots . Можно, однако воспользоваться другой формой закона больших чисел — теоремой Чебышева:

Если величины $\eta_1, \dots, \eta_s, \dots$ независимы и существуют $M\eta_s = a_s$ и $D\eta_s \leq C$, то при $N \rightarrow \infty$

$$\frac{1}{N} \sum_{s=1}^N \eta_s - \frac{1}{N} \sum_{s=1}^N a_s \xrightarrow{P} 0.$$

(Доказательство этой теоремы легко получить, применяя к величине $(1/N)(\eta_1 + \dots + \eta_N)$ неравенство Чебышева, стр. 141).

В нашем случае все $Mu(Q_{v_s}) = u(P_0)$, а дисперсии $Du(Q_{v_s}) \leq Mu^2(Q_{v_s}) \leq C^2$, где $C = \sup_{P \in G_0} |g(P)|$. В самом деле, как известно [88], максимум и минимум гармонической функции достигаются на границе области, так что $|u(P)| \leq C$ при всех $P \in G$.

Такой метод расчета $u(P_0)$ считается более быстрым, чем метод п. 5.5 гл. 5, так как вдали от границы G^0 позволяет делать большие шаги (l_n). Обычно рекомендуют выбирать максимально возможные радиусы l_n . Впрочем, аккуратного численного сравнения этих методов автор не видел.

Изложенный метод был предложен Дж. Брауном [6] и обоснован М. Мюллером [162], который доказал, в частности, что вероятность того, что траектория $P_0 \rightarrow Q_1 \rightarrow \dots \rightarrow Q_n \rightarrow \dots$ никогда не попадет в G^0 , равна нулю. Дальнейшее развитие метода — организация зависимых испытаний, решение уравнений более общего вида, использование вместо кругов других фигур (для которых известны функции Грина) — имеются в работах [28, 30, 65].

§ 4. Вычисление винеровских интегралов

Решения многих задач теории вероятностей, статистической и квантовой физики, теории дифференциальных уравнений могут быть выражены через так называемые континуальные интегралы [16]. Вычисление континуальных интегралов классическими методами весьма сложно, хотя некоторые «квадратурные» формулы для этого имеются [12, 123]. Методы Монте-Карло для расчета таких интегралов были, по-видимому, впервые использованы в работе И. М. Гельфанда и Н. Н. Ченцова [15].

4.1. Винеровские интегралы. Чаше других встречаются континуальные интегралы по мере Винера, называемые обычно *винеровскими интегралами*:

$$\int_a^b F[x] d_W x, \quad (9)$$

где C — пространство всех непрерывных на отрезке $0 \leq t \leq T$ функций $x(t)$, удовлетворяющих начальному условию $x(0) = 0$, а $F[x(t)]$ — произвольный непрерывный и ограниченный функционал, заданный на C .

При определении меры Винера здесь, как и в [16], коэффициент диффузии D полагается равным $1/4$: этого всегда можно добиться изменением масштаба времени. Пусть $x = \xi(t)$ — координата

частицы, совершающей броуновское движение вдоль оси Ox , начинающееся из точки $\xi(0)=0$. Если $D=1/4$, то плотность вероятностей $\xi(t)$ равна

$$p_t(x) = (\pi t)^{-1/2} e^{-(x^2/t)}.$$

Рассмотрим случайную траекторию $x=\xi(t)$ частицы, совершающей броуновское движение, с начальным условием $\xi(0)=0$ (такую траекторию называют также винеровским процессом [71]). Тогда математическое ожидание случайной величины $F[\xi(t)]$ равно интегралу

$$MF[\xi(t)] = \int_G F[x] d_W x.$$

Отсюда вытекает простейший метод Монте-Карло для расчета интеграла (9)

$$\int_G F[x] d_W x \approx \frac{1}{N} \sum_{s=1}^N F[\xi_{(s)}], \quad (10)$$

где $\xi_{(1)}(t), \dots, \xi_{(N)}(t)$ — независимые реализации броуновской траектории $x=\xi(t)$.

4.2. Приближенное построение броуновских траекторий. Укажем два способа приближенной реализации таких траекторий. В обоих способах отрезок $[0, T]$ делится на n равных частей абсциссами

$$0=t_0 < t_1 < t_2 < \dots < t_n=T, \quad (11)$$

разыгрываются случайные значения траектории $\xi(t_i)$, и

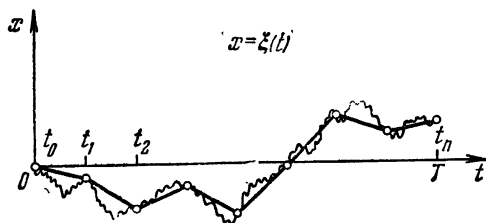


Рис. 73.

полученные точки $(t_i, \xi(t_i))$ на плоскости t, x соединяются отрезками прямых (рис. 73). Построенная ломаная и есть приближенная броуновская траектория.

В приведенных ниже расчетных формулах ξ_1, ξ_2, \dots — это независимые нормальные случайные величины с параметрами $a=0$ и $\sigma^2=1$.

Первый способ основан непосредственно на определении броуновского движения: так как условное распределение $\xi(t_i)$ при известном значении $\xi(t_{i-1})$ нормально с параметрами $a=\xi(t_{i-1})$ и $\sigma^2=(1/2)|t_i-t_{i-1}|$, то при $i=1, 2, \dots, n$

$$\xi(t_i) = \xi(t_{i-1}) + \sqrt{T/(2n)} \xi_i. \quad (12)$$

Формула (12) с начальным условием $\xi(0)=0$ позволяет разыграть все значения $\xi(t_i)$.

Второй способ [14, 48] основан на том, что если значения $\xi(t')$ и $\xi(t'')$ известны, то условное распределение значения $\xi\left[\frac{1}{2}(t' + t'')\right]$ также нормально с параметрами

$$a=(1/2)[\xi(t') + \xi(t'')] \quad \text{и} \quad \sigma^2=(1/8)|t''-t'|.$$

Пусть $n=2^m$. Используя условие $\xi(0)=0$, можно разыграть значение $\xi(T) = \xi(0) + \sqrt{T/2} \xi_1$ (это фактически формула (12) при $n=1$), а затем разыгрывать остальные значения $\xi(t_i)$ в серединах отрезков:

$$\xi(T/2) = \frac{1}{2} [\xi(0) + \xi(T)] + \sqrt{T/8} \xi_2;$$

$$\xi(T/4) = \frac{1}{2} [\xi(0) + \xi(T/2)] + \sqrt{T/16} \xi_3;$$

$$\xi(3T/4) = \frac{1}{2} [\xi(T/2) + \xi(T)] + \sqrt{T/16} \xi_4;$$

$$\xi(T/8) = \frac{1}{2} [\xi(0) + \xi(T/4)] + \sqrt{T/32} \xi_5,$$

.....

Численный пример [8]. Вычислить винеровский интеграл, точное значение которого известно:

$$\int_G \|x\|^2 d_W x = \frac{1}{4},$$

где

$$\|x\|^2 = \int_0^1 x^2(t) dt.$$

Случайные траектории будем строить вторым способом при $n=4$. Расчетные формулы:

$$\begin{aligned}\xi(1) &= 0,70711\zeta_1, \\ \xi(1/2) &= 0,5\xi(1) + 0,35355\zeta_2, \\ \xi(1/4) &= 0,5\xi(1/2) + 0,25\zeta_3, \\ \xi(3/4) &= 0,5[\xi(1/2) + \xi(1)] + 0,25\zeta_4.\end{aligned}$$

Значение функционала $\|x\|^2$ на ломаной вычисляется точно: на каждом отрезке можно воспользоваться формулой

$$\int_a^b f^2(t) dt = \frac{b-a}{3} [f^2(a) + f(a)f(b) + f^2(b)],$$

которая точна для функций $f(t)$, линейных на $[a, b]$. Получим расчетную формулу

$$\begin{aligned}\|\xi\|^2 &= \frac{1}{12} \{ \xi^2(0) + 2\xi^2(1/4) + 2\xi^2(1/2) + 2\xi^2(3/4) + \xi^2(1) + \\ &+ \xi(0)\xi(1/4) + \xi(1/4)\xi(1/2) + \xi(1/2)\xi(3/4) + \xi(3/4)\xi(1) \}.\end{aligned}$$

Результат расчета, выполненного в [8] при $N=10$, случайно оказался исключительно хорошим:

$$\frac{1}{10} \sum_{s=1}^{10} \|\xi_{(s)}\|^2 = 0,2587.$$

В самом деле, так как

$$D(\|\xi\|^2) = \int_C \|x\|^4 d_W x - \left(\frac{1}{4}\right)^2 = \frac{7}{48} - \frac{1}{16} = \frac{1}{12},$$

то при $N=10$ вероятная ошибка $r_{10} = 0,675/\sqrt{120} = 0,062$ (именно такой порядок имеет погрешность, если взять всего девять из сотенных десяти траекторий). Кроме статистической ошибки из-за малого значения N , возможна еще ошибка от замены траекторий ломаными, т. е. из-за малости n .

4.3. Замена континуального интеграла многомерным. Фиксируем разбиение (11) отрезка $[0, T]$ и условимся каждую непрерывную кривую $x(t)$ заменять ломаной $\tilde{x}(t)$, совпадающей с $x(t)$ во всех точках деления:

$$\tilde{x}(t_i) = x(t_i) = x_i, \quad i = 0, 1, \dots, n.$$

Значение функционала $F[x]$ на таких ломаных можно

рассматривать как функцию от n переменных: $F[x] = F(x_1, \dots, x_n)$. Известно [16, 89], что

$$\int_{\mathcal{G}} F[x] d_W x = \lim_{n \rightarrow \infty} \left(\frac{n}{\pi T} \right)^{n/2} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} F(x_1, \dots, x_n) \exp \left\{ -\frac{n}{T} \sum_{i=1}^n (x_i - x_{i-1})^2 \right\} dx_1 \dots dx_n.$$

Поэтому для приближенного расчета интеграла (9) можно вычислять многомерные интегралы, стоящие справа, при достаточно больших n . Можно, в частности, использовать методы Монте-Карло (гл. 3 и 4).

В статьях [124, 125, 173, 178] методами Монте-Карло вычисляются континуальные интегралы, к которым сводятся некоторые задачи теории обыкновенных дифференциальных уравнений и, в частности, уравнения Шрёдингера (другие методы Монте-Карло для расчета уравнения Шрёдингера рассмотрены в работах [27, 119, 122, 146, 150]).

ПРИЛОЖЕНИЯ

Вспомогательное неравенство

Пусть в области G задана неотрицательная функция $\rho(P) \geq 0$. Рассмотрим две произвольные функции $u(P)$ и $v(P)$, принадлежащие $L_2(G; \rho)$. Тогда справедливо неравенство

$$\left[\int_G uv \rho \, dP \right]^2 \leq \int_G u^2 \rho \, dP \int_G v^2 \rho \, dP. \quad (1)$$

Доказательство. Запишем очевидное неравенство

$$\int_G [tu(P) + v(P)]^2 \rho(P) \, dP \geq 0,$$

где t — любое действительное число. Это неравенство можно переписать в виде

$$t^2 \int_G u^2 \rho \, dP + 2t \int_G uv \rho \, dP + \int_G v^2 \rho \, dP \geq 0.$$

Как известно, квадратный трехчлен $At^2 + 2Bt + C$, где $A \geq 0$, неотрицателен при всех $-\infty < t < \infty$ тогда и только тогда, когда $B^2 - AC \leq 0$. А в нашем случае неравенство $B^2 \leq AC$ совпадает с (1).

Неравенство (1) представляет собой одну из известных форм неравенства Коши — Буняковского, называемого также неравенством Шварца.

Таблицы

Таблица 1 содержит некоторые значения интеграла вероятностей $\Phi(x)$, функции распределения $a_1(x)$ и функции распределения Колмогорова $K(x)$. Определения этих функций см. на стр. 88, 35 и 43.

Таблица 2 содержит некоторые значения распределения χ^2 . Определение величины $\chi^2(m, P)$ см. на стр. 33.

Таблица 3 содержит некоторые значения распределения Стюдента. Определение величины $t_{m,\beta}$ см. на стр. 90.

Таблица 4 содержит 1000 случайных цифр, которые имитируют независимые значения случайной величины ε (гл. 1, § 1).

Таблица 5 содержит 200 нормальных величин, которые имитируют независимые значения нормальной (гауссовской) случайной величины ζ с параметрами

$$\alpha = M\zeta = 0, \quad \sigma^2 = D\zeta = 1.$$

Таблица 6 содержит числители величин $V^{(s)}$, необходимых для расчета точек Q_i^* (стр. 265). Знаменатели всех $V_j^{(s)}$, равны 2^s , так что, например, $V_4^{(9)} = 433/512$.

Таблица 1

Некоторые значения интеграла вероятностей $\Phi(x)$, функции распределения $a_1(x)$ и функции распределения Колмогорова $K(x)$.

x	$\Phi(x)$	x	$a_1(x)$	x	$K(x)$
0,675	0,50	0,12	0,50	0,82	0,50
0,842	0,60	0,15	0,60	0,89	0,60
1,036	0,70	0,18	0,70	0,97	0,70
1,282	0,80	0,24	0,80	1,07	0,80
1,645	0,90	0,35	0,90	1,22	0,90
1,96	0,95	0,46	0,95	1,36	0,95
2,58	0,99	0,74	0,99	1,63	0,99
2,81	0,995	0,87	0,995	1,73	0,995
3,29	0,999	1,17	0,999	1,95	0,999
3,72	0,9998	1,49	0,9998	2,15	0,9998

Таблица 2

Значения $\chi^2(m, P)$

m	P	0,99	0,90	0,70	0,50	0,30	0,10	0,05	0,01	0,005	0,001
	β	0,01	0,10	0,30	0,50	0,70	0,90	0,95	0,99	0,995	0,999
2	0,02	0,2	0,7	1,4	2,4	4,6	6,0	9,2	10,6	13,8	
3	0,12	0,6	1,4	2,4	3,7	6,3	7,8	11,3	12,8	16,3	
4	0,30	1,1	2,2	3,4	4,9	7,8	9,5	13,3	14,9	18,5	
5	0,55	1,6	3,0	4,4	6,1	9,2	11,1	15,1	16,7	20,5	
6	0,87	2,2	3,8	5,3	7,2	10,6	12,6	16,8	18,5	22,5	
7	1,24	2,8	4,7	6,3	8,4	12,0	14,1	18,5	20,3	24,3	
8	1,65	3,5	5,5	7,3	9,5	13,4	15,5	20,1	22,0	26,1	
9	2,09	4,2	6,4	8,3	10,7	14,7	16,9	21,7	23,6	27,9	
10	2,56	4,9	7,3	9,3	11,8	16,0	18,3	23,2	25,2	29,6	

Таблица 2 (продолжение)

m	P	0,99	0,90	0,70	0,50	0,30	0,10	0,05	0,01	0,005	0,001
	β	0,01	0,10	0,30	0,50	0,70	0,90	0,95	0,99	0,995	0,999
12	3,6	6,3	9,0	11,3	14,0	18,5	21,0	26,2	28,3	32,9	
14	4,7	7,8	10,8	13,3	16,2	21,1	23,7	29,1	31,3	36,1	
16	5,8	9,3	12,6	15,3	18,4	23,5	26,3	32,0	34,3	39,2	
18	7,0	10,9	14,4	17,3	20,6	26,0	28,9	34,8	37,2	42,3	
20	8,3	12,4	16,3	19,3	22,8	28,4	31,4	37,6	40,0	45,3	
25	11,5	16,5	20,9	24,3	28,2	34,4	37,7	44,3	46,9	52,6	
30	15,0	20,6	25,5	29,3	33,5	40,3	43,8	50,9	53,7	59,7	
35	18,5	24,8	30,2	34,3	38,9	46,1	49,8	57,3	60,3	66,6	
40	22,2	29,1	34,9	39,3	44,2	51,8	55,8	63,7	66,8	73,4	
45	25,9	33,3	39,6	44,3	49,5	57,5	61,7	70,0	73,2	80,1	
50	29,7	37,7	44,3	49,3	54,7	63,2	67,5	76,2	79,5	86,7	
55	33,6	42,1	49,1	54,3	60,0	68,8	73,3	82,3	85,7	93,2	
60	37,5	46,5	53,8	59,3	65,2	74,4	79,1	88,4	92,0	99,6	

Таблица 3

Значения t_m, β

$\beta \backslash m$	0,50	0,70	0,90	0,95	0,99	0,995	0,999
2	0,817	1,34	2,92	4,30	9,92	14,09	31,60
3	0,765	1,25	2,35	3,18	5,84	7,45	12,92
4	0,741	1,19	2,13	2,78	4,60	5,60	8,61
5	0,727	1,16	2,02	2,57	4,03	4,77	6,87
6	0,718	1,13	1,94	2,45	3,71	4,32	5,96
7	0,711	1,12	1,90	2,36	3,50	4,03	5,41
8	0,706	1,11	1,86	2,31	3,36	3,83	5,04
9	0,703	1,10	1,83	2,26	3,25	3,69	4,78
10	0,700	1,09	1,81	2,23	3,17	3,58	4,59
12	0,696	1,08	1,78	2,18	3,06	3,43	4,32
14	0,692	1,08	1,76	2,14	2,98	3,33	4,14
16	0,690	1,07	1,75	2,12	2,92	3,25	4,02
18	0,688	1,07	1,73	2,10	2,88	3,20	3,92
20	0,687	1,06	1,72	2,09	2,85	3,15	3,85
25	0,684	1,06	1,71	2,06	2,79	3,08	3,73
30	0,683	1,05	1,70	2,04	2,75	3,03	3,65
50	0,679	1,05	1,68	2,01	2,68	2,94	3,50
100	0,677	1,04	1,66	1,98	2,63	2,87	3,39
∞	0,675	1,04	1,65	1,96	2,58	2,81	3,29

Т а б л и ц а 4

1000 случайных цифр

86515	90795	66155	66434	56558	12332
69186	03393	42502	99224	88955	53758
41686	42163	85181	38967	33181	72664
86522	47171	88059	89342	67248	09082
72587	93000	89688	78416	27589	99528
94377	57802	52452	42499	33346	83935
91641	18867	76773	97526	27256	66447
53807	00607	04825	82134	80317	75120
12311	90316	87113	84778	45863	24520
14480	50961	84754	57616	38132	64294
79130	90410	45420	77757	75593	51435
25731	37525	16287	66181	73244	61870
45904	75601	70492	10274	23974	14783
19976	04925	07824	76044	32373	05312
15218	49286	89571	42903	59598	56774
73189	64448	31276	70795	33071	96929
28709	38238	76208	76575	53163	58481
17932	66686	64254	57598	26623	91730
94590	22561	70177	03569	21302	17381
08749	43448	28484	16325	62766	31466
91682	12904	29142	65877	64517	31466
87653	98088	75162	97496	59297	79636
79429	66186	59157	95114	16021	30890
85444	39453	67981	49687	36801	38666
85739	44326	91641	40837	93030	03675
02555	52905	84637	76154	14150	07876
74364	16796	59575	32764	91090	66515
21656	93662	81305	58846	69558	41675
50055	11244	29835	35801	23472	22700
18788	91332	32795	54313	39072	16809
41899	69207	66785	87225		
05498	51512	16107	52141		
88898	23775	30649	86545		
39976	21279	36694	85970		
22148	60102	18465	87650		

Т а б л и ц а 5

200 нормальных величин

0,201	1,192	—0,008	0,035	1,042	—1,814
1,161	—0,669	—1,589	0,582	1,882	0,739
0,586	—0,925	0,090	1,507	—1,115	0,278
0,143	—0,286	1,281	0,404	0,638	—0,443
0,952	—1,771	2,885	0,469	1,466	1,685
1,180	0,003	—0,586	0,857	—0,556	0,812
—0,274	1,083	1,157	0,999	—0,103	0,541
0,101	—1,357	—0,443	—0,556	—0,510	—1,193
—2,301	—0,645	—0,392	1,798	0,614	—1,360
—0,969	—0,083	0,832	0,427	—0,889	0,417
—0,268	—1,250	—1,213	1,388	0,978	—0,768
—0,602	0,009	0,212	—1,465	0,408	0,117
—0,057	—0,506	—0,156	—1,238	0,251	0,145
1,494	—0,441	—0,203	—0,132	—0,616	—1,660
—0,851	1,105	1,224	—0,700	2,196	0,837
0,896	0,515	—0,717	0,856	—1,163	1,889
0,438	0,215	0,247	1,222	—0,533	0,677
—0,107	1,516	—0,115	1,717	—0,425	—0,433
0,650	—1,139	—0,079	0,072	—2,140	—0,511
0,833	0,084	0,557	—0,847	0,006	—0,227
0,595	—0,957	—0,453	1,714	0,117	—0,304
0,888	1,440	—0,167	1,989	1,391	1,141
1,360	0,626	0,777	—0,471	—1,877	—0,332
0,396	0,317	1,306	1,775	—0,783	1,141
1,400	—0,541	—0,351	0,096	—1,298	—0,209
—0,865	—0,276	0,610	0,214	—1,765	0,887
0,141	1,110	1,057	0,086	—0,520	0,119
—0,455	0,541	—0,793	1,405	1,573	0,061
—1,232	—1,550	0,709	—0,878	—0,447	—1,697
1,499	—0,843	0,419	—1,102	—0,397	0,276
1,225	1,224	1,399	0,218		
0,418	—1,338	—1,170	—1,514		
2,855	0,188	—0,084	0,766		
1,702	—0,054	0,750	0,889		
—0,821	—2,101	0,376	0,580		

Таблица 6

j	Числители V^s																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	3	5	15	17	51	85	255	257	771	1285	3855	4369	13107	21845	65535	65537	196611	327685	983055
3	1	1	7	11	13	61	67	79	465	721	823	4091	4125	4141	28723	45311	53505	250113	276231	326411
4	1	3	1	5	31	29	81	147	433	149	719	3693	3841	11523	16641	49925	16671	83229	515921	482707
5	1	1	5	3	15	51	125	141	177	759	267	1839	6929	16241	16565	17139	82207	50979	252717	851901
6	1	3	7	7	9	11	79	63	193	707	1351	1479	6857	1227	32527	36863	102401	241667	258055	323591
7	1	1	3	11	7	31	13	161	309	901	2007	1311	6235	2915	27745	35885	34225	112177	378387	246791
8	1	3	3	13	23	57	11	219	101	687	449	579	3779	717	17751	20473	23627	261147	140453	658927
9	1	1	1	15	11	29	13	85	475	63	1233	337	6577	16383	24859	24845	110365	175013	97387	453871
10	1	3	5	1	29	61	79	169	293	177	689	2259	2197	6225	22861	64589	15455	107641	103861	247585
11	1	1	7	5	5	3	127	97	47	819	765	3341	3379	15217	21513	19463	15387	50237	130435	790151
12	1	3	1	11	25	15	9	191	417	631	789	1327	2885	4295	29021	37603	62509	248899	87493	716363
13	1	1	5	7	19	51	51	151	157	141	379	501	5889	16121	12039	11055	90495	123227	49489	574417

ЛИТЕРАТУРА

1. Бахвалов Н. С., О приближенном вычислении кратных интегралов. Вестник МГУ, сер. матем., механ., astron., физ., 1959, № 4, 3—18.
2. Бахвалов Н. С., Об оптимальных оценках сходимости квадратурных процессов и методов интегрирования типа Монте-Карло на классах функций. Сб. «Числ. методы решения дифф. и интегр. уравнений», «Наука», 1964, 5—63.
3. Беляев Б. Н., Царицына И. В., Расчет взаимодействия быстрых нуклонов с ядрами методом случайных испытаний. Сб «Методы вычислений», ЛГУ, 1963, № 1, 76—107.
4. Большев Л. Н., О преобразованиях случайных величин. Теория вероятн. и ее применение, 1959, 4, № 2, 136—149.
5. Большев Л. Н., Смирнов Н. В., Таблицы математической статистики, «Наука», 1965.
6. Браун Дж. В., Метод Монте-Карло. «Современная математика для инженеров» под ред. Э. Беккенбаха, ИЛ, 1958.
7. Бусленко Н. П., Моделирование сложных систем, «Наука», 1968.
8. Бусленко Н. П., Голенко Д. И., Соболев И. М., Срагович В. Г., Шрейдер Ю. А., Метод статистических испытаний (метод Монте-Карло). СМБ, Физматгиз, М., 1962.
9. Бусленко Н. П., Шрейдер Ю. А., Метод статистических испытаний (Монте-Карло) и его реализация на ЦВМ, Физматгиз, М., 1961.
10. Владимиров В. С., О применении метода Монте-Карло для отыскания наименьшего характеристического числа и соответствующей собственной функции линейного интегрального уравнения. Теория вероятн. и ее применение, 1956, 1, № 1, 113—130.
11. Владимиров В. С., Численное решение кинетического уравнения для сферы. Вычислит. математика, 1958, № 3, 3—33.
12. Владимиров В. С., О приближенном вычислении вине-ровских интегралов. Успехи матем. наук, 1960, 15, № 4, 129—135.
13. Владимиров В. С., Соболев И. М., Расчет наименьшего характеристического числа уравнения Пайерлса методом Монте-Карло. Вычислит. математика, 1958, № 3, 130—137.
- 13а. Воронцов Ю. В., Полляк Ю. Г., Об использовании ква-зислучайных последовательностей при прямом вероятностном моделировании систем. Автоматика и вычислительная техника, 1971, № 6, 23—27.

14. Гельфанд И. М., Фролов А. С., Ченцов Н. Н., Вычисление континуальных интегралов методом Монте-Карло. Изв. вузов, сер. матем., 1958, № 5, 32—45.
15. Гельфанд И. М., Ченцов Н. Н., Численное вычисление континуальных интегралов. ЖЭТФ, 1956, 31, № 6, 1106—1107.
16. Гельфанд И. М., Яглом А. М., Интегрирование в функциональных пространствах и его применение в квантовой физике. Успехи матем. наук, 1966, 11, №1, 77—114.
17. Герман В. А., Соболев И. М., Об использовании поверхностей постоянной плотности при моделировании многомерных случайных величин. Ж. вычисл. матем. и матем. физ., 1971, 11, № 3, 789—791.
18. Голенко Д. И., Образование случайных величин с произвольным законом распределения. Вычислит. математика. 1959, № 5, 83—92.
19. Голенко Д. И., Моделирование и статистический анализ псевдослучайных чисел на ЭВМ, «Наука», 1965.
20. Грановский Б. Л., О случайных квадратурах гауссовского типа. Ж. вычисл. матем. и матем. физ., 1968, 8, № 4, 879—884.
21. Грановский Б. Л., О допустимости случайных квадратур гауссовского типа. Ж. вычисл. матем. и матем. физ., 1971, 11, № 1, 44—50.
22. Гурин Л. С., Опыт применения метода Монте-Карло для нахождения экстремальных значений функций. Сб. «Вопросы вычисл. матем. и вычисл. техники» под ред. Л. А. Люстерника, Машгиз, М., 1963, 118—123.
23. Гутер Р. С., Муратова Т. А., Метод Монте-Карло и парадокс Бертрана. Заводская лаборатория 1970, № 5, 582—584.
24. Дунин-Барковский И. В., Смирнов Н. В., Теория вероятностей и математическая статистика в технике (общая часть), Гостехиздат, М., 1955.
25. Дэвисон Б., Теория переноса нейтронов, Атомиздат, М., 1960.
26. Дядькин И. Г., Моделирование случайной энергии гамма-кванта, рассеянного в результате комптон-эффекта. Ж. вычисл. матем. и матем. физ., 1966, 6, № 2, 384—385.
27. Дядькин И. Г., Жукова С. А., О постановке задачи и алгоритме решения уравнения Шредингера методом Монте-Карло. Ж. вычисл. матем. и матем. физ., 1968, 8, № 1, 222—229.
28. Дядькин И. Г., Стариков В. Н., Об одной возможности экономии машинного времени при решении уравнений Лапласа методом Монте-Карло. Ж. вычисл. матем. и матем. физ., 1965, 5, № 5, 936—938.
29. Дядькин И. Г., Стариков В. Н., Об использовании симметрии и других особенностей нейтронных траекторий для ускорения расчетов методом Монте-Карло. Ж. вычисл. матем. и матем. физ., 1968, 8, № 5, 1001—1012.
30. Елепов Б. С., Михайлов Г. А., О решении задачи Дирихле для уравнения $\Delta u - cu = -g$ моделированием «блуждающий по сферам». Ж. вычисл. матем. и матем. физ., 1969, 9, № 3, 647—654.
31. Ермаков С. М., Интерполирование по случайным точкам. Ж. вычисл. матем. и матем. физ., 1963, 3, № 1, 186—190.

32. Ермаков С. М., О допустимости процедур метода Монте-Карло. Докл. АН СССР, 1967, 172, № 2, 262—264.
33. Ермаков С. М., Метод Монте-Карло и смежные вопросы, «Наука», 1971.
34. Ермаков С. М., Золотухин В. Г., Полиномиальные приближения и метод Монте-Карло. Теория вероятн. и ее применение, 1960, 4, № 4, 473—476.
35. Золотухин В. Г., Деев Г. Е., Модификация метода Монте-Карло для расчета дифференциального потока нейтронов и γ -квантов. [53], 53—66.
36. Золотухин В. Г., Ермаков С. М., Применение метода Монте-Карло для расчета защиты от ядерных излучений. Сб. «Вопросы физики защиты реакторов», Госатомиздат, М., 1963, 171—182.
37. Золотухин В. Г., Ефименко Б. А., Дорошенко Г. Г., Расчет амплитудных распределений импульсов в однокристалльном сцинтилляционном спектрометре быстрых нейтронов. [53], 149—179.
38. Климов Г. П., Стохастические системы обслуживания, «Наука», 1966.
39. Кольчужкин А. М., Шевцова И. Н., К расчету характеристик потока электронов методом Монте-Карло. Изв. вузов, сер. физика, 1969, № 11, 134—136.
40. Кондюрин Ю. М., О моделировании случайных величин методом суперпозиции. Ж. вычисл. матем. и матем. физ., 1970, 10, № 1, 262—266.
41. Коробов Н. М., Некоторые проблемы распределения дробных долей. Успехи матем. наук, 1949, 4, № 1, 189—190.
42. Коробов Н. М., Теоретико-числовые методы в приближенном анализе. Физматгиз, М., 1963.
43. Кофман А., Крюон Р., Массовое обслуживание, теория и приложения, «Мир», 1965.
44. Крамер Г., Математические методы статистики. ИЛ, 1948.
45. Кубилюс Й. П., Линник Ю. В., Арифметическое моделирование броуновского движения. Изв. вузов, сер. матем., 1959, № 6, 88—95.
46. Куропатенко Э. С., Огибин В. Н., Об одной схеме моделирования траекторий частиц в системах сложной геометрии. Ж. вычисл. матем. и матем. физ., 1968, 8, № 1, 211—216.
47. Кутай А. К., Применение метода Монте-Карло в исследованиях точности приборостроения. [54], 98—101.
48. Леви П. Стохастические процессы и броуновское движение, «Наука», 1972.
49. Майоров Л. В., Суховой А. М., О вычислении некоторых нелинейных функционалов от решения уравнения переноса. Сб. тезисов докл. совещания по методам Монте-Карло, Минск, 1970.
50. Майоров Л. В., Франк-Каменецкий А. Д., Сравнение различных нелокальных оценок потока. [54], 113—115.
51. Марчук Г. И., Методы расчета ядерных реакторов, Госатомиздат, М., 1961.
52. Марчук Г. И., Михайлов Г. А., Назаралиев М. А., Дарбинян Р. А., Решение прямых и некоторых обратных

- задач атмосферной оптики методом Монте-Карло, «Наука», Новосибирск, 1968.
- 53 Метод Монте-Карло в проблеме переноса излучений. Сб. под ред. Г. И. Марчука, Атомиздат, М., 1967.
 - 54 Методы Монте-Карло и их применения. Сб. тезисов докл. на 2-м Всесоюзн. совещ. по методам Монте-Карло. Тбилиси, 1969.
 - 55 Методы Монте-Карло и их применения. Сб. тезисов докл. на 3-й Всесоюзн. конф. по методам Монте-Карло, Новосибирск, 1971.
 - 56 Михайлов Г. А., О моделировании случайных величин для одного класса законов распределения. Теория вероятн. и ее примен., 1965, 10, № 4, 749—751.
 - 57 Михайлов Г. А., О расчетах возмущений ядерных реакторов методом Монте-Карло. Ж. вычисл. матем. и матем. физ., 1966, 6, № 2, 380—384.
 - 58 Михайлов Г. А., К вопросу о построении экономичных алгоритмов моделирования случайных величин. Ж. вычисл. матем. и матем. физ., 1966, 6, № 6, 1134—1136.
 - 59 Михайлов Г. А., Использование приближенных решений сопряженной задачи для улучшения алгоритмов метода Монте-Карло. Ж. вычисл. матем. и матем. физ., 1969, 9, № 5, 1145—1152.
 - 60 Михайлов Г. А., Новый алгоритм метода Монте-Карло для оценки максимального собственного значения интегрального оператора. Докл. АН СССР, 1970, 191, № 5, 993—996.
 - 61 Михайлов Г. А., О сочетании метода конечных сумм с методом Монте-Карло для решения интегральных уравнений 2-го рода. Мат. заметки. 1971, 9, № 4, 425—434.
 - 62 Михайлов Г. А., Моделирования бета-распределения, один из параметров которого целый. [55], 107—108.
 - 63 Михлин С. Г., Смолицкий Х. Л., Приближенные методы решения дифференциальных и интегральных уравнений. СМБ, «Наука». 1965.
 - 64 Мишина А. П., Проскуряков И. В., Высшая алгебра. СМБ, Физматгиз, М., 1962.
 - 65 Мишустин Б. А., О решении задачи Дирихле для уравнения Лапласа методом статистических испытаний. Ж. вычисл. матем. и матем. физ., 1967, 7, № 5, 1179—1185.
 - 66 Моцкус И. Б., Многоэкстремальные задачи в проектировании, «Наука», 1967. ✓
 - 67 Полляк Ю. Г., О сравнении двух вероятностей. Теория вероятн. и ее применение, 1963, 8, № 2, 195.
 - 68 Полляк Ю. Г., Об экономичных алгоритмах моделирования случайных величин. Ж. вычисл. матем. и матем. физ., 1968, 8, № 3, 710.
 - 69 Полляк Ю. Г., Вероятностное моделирование на ЭВМ, «Сов. радио», М., 1971.
 - 70 Постников А. Г., Арифметическое моделирование случайных процессов. Труды матем. ин-та АН СССР, 1960, т. 57.
 - 71 Прохоров Ю. В., Розанов Ю. А., Теория вероятностей. Основные понятия, предельные теоремы, случайные процессы СМБ, «Наука», 1967.
 - 72 Растринин Л. А., Статистические методы поиска, «Наука», 1968. ✓

73. Романовский И. В., О методах моделирования непрерывных случайных величин из величин с равномерным распределением. Сб. «Методы вычислений», ЛГУ, 1966, № 3, 113—121.
74. Сизова А. Ф., Оценка плотности делений в сферическом реакторе. [53], 228—231.
75. Соболев И. М., Псевдослучайные числа для машины «Стрела». Теория вероятн. и ее применение, 1958, 3, № 2, 205—211.
76. Соболев И. М., О решении интегрального уравнения Пайерлса методом Монте-Карло. Теория вероятн. и ее применение, 1960, 5, № 3, 361—366.
77. Соболев И. М., О вычислении бесконечномерных интегралов Ж. вычисл. матем. и матем. физ., 1961, 1, № 5, 917—922.
78. Соболев И. М., Применение распределения ω^2 для оценки погрешности при вычислении интегралов методом Монте-Карло. Ж. вычисл. матем. и матем. физ., 1962, 2, № 4, 717—723.
79. Соболев И. М., О периодах псевдослучайных последовательностей Теория вероятн. и ее применение, 1964, 9, № 2, 367—373.
80. Соболев И. М., Метод Монте-Карло для расчета критичности в многогрупповом приближении. [53], 232—254.
81. Соболев И. М., О расчете движения заряженных частиц в плазме методом Монте-Карло. [54], 180—182.
82. Соболев И. М., Многомерные квадратурные формулы и функции Хаара, «Наука», 1969.
83. Соболев И. М., Об одном подходе к вычислению многомерных интегралов. Вопросы вычислит. и приклад. математики, Ташкент, 1970, № 38, 100—111.
84. Соболев И. М., Детерминистическая интерпретация критериев согласия и проверка псевдослучайных чисел. Сб. «Исследование операций и статистическое моделирование», Ленинград, 1972, 162—169.
85. Соболев И. М., Вероятностная оценка погрешности для неслучайных сеток интегрирования. Вопросы вычислит. и приклад. матем., Ташкент, 1972, № 14, 5—11.
86. Соболев И. М., Статников Р. Б., ЛП-поиск и задачи оптимального конструирования. Сб. «Проблемы случайного поиска». «Зинатне», Рига, 1972, № 1, 117—135.
87. Срагович В. Г., Моделирование некоторых классов случайных процессов. Ж. вычисл. матем. и матем. физ., 1963, 3, № 3, 586—593.
88. Тихонов А. Н., Самарский А. А., Уравнения математической физики, изд. 3, «Наука», 1966.
89. Фомин С. В., О включении интеграла по мере Винера в общую теорию интеграла Лебега. Научн. докл. высш. школы, физ.-матем. науки, 1958, № 2, 83—85.
90. Франк-Каменецкий А. Д., О решении кинетического уравнения методом Монте-Карло. Ж. вычисл. матем. и матем. физ., 1963, 3, № 4, 766—769.
91. Франк-Каменецкий А. Д., Применение метода Монте-Карло для многогруппового расчета реакторов. Атомная энергия, 1964, 16, № 2, 119—122.

- 92 Фролов А. С., Ченцов Н. Н., О вычислении методом Монте-Карло определенных интегралов, зависящих от параметра. Ж. вычисл. матем. и матем. физ., 1962, 2, № 4, 714—717.
- 93 Хисамутдинов А. И., Оценка функционалов от решения сопряженного уравнения переноса излучения методом Монте-Карло. Ж. вычисл. матем. и матем. физ., 1968, 8, № 2, 467—471.
- 94 Хисамутдинов А. И., Единичный класс оценок для вычисления по методу Монте-Карло функционалов от решения интегрального уравнения 2-го рода. Ж. вычисл. матем. и матем. физ., 1970, 10, № 5, 1269—1280.
- 95 Хисамутдинов А. И., К уменьшению дисперсии оценки вероятности методом Монте-Карло. Ж. вычисл. матем. и матем. физ., 1970, 10, № 6, 1547—1549.
- 96 Чавчанидзе В. В., Джапаридзе К. Г., Кумсиашвили В. А., Статистико-вероятностное моделирование цепных химических реакций. Тр. Ин-та кибернет. АН Груз. ССР, 1963, 1, 85—91.
- 97 Ченцов Н. Н., О квадратурных формулах для функций бесконечно большого числа переменных. Ж. вычисл. матем. и матем. физ., 1961, 1, № 3, 418—424.
- 98 Ченцов Н. Н., Оценка неизвестной плотности распределения по наблюдениям. Докл. АН СССР, 1962, 147, № 1, 45—48.
- 99 Ченцов Н. Н., Псевдослучайные числа для моделирования марковских цепей. Ж. вычисл. матем. и матем. физ., 1967, 7, № 3, 632—643.
- 100 Шилов Г. Е., Интегрирование в бесконечномерных пространствах и интеграл Винера. Успехи матем. наук, 1963, 18, № 2, 99—120.
- 101 Шрейдер Ю. А., Решение систем линейных алгебраических уравнений по методу Монте-Карло. Вопросы теории матем. машин, 1958, № 1, 167—171.
- 102 B á n k ö v i G., A note on the generation of beta distributed and gamma distributed random variables. Magyar Tud. Akad. Mat. Kut. Int. Közleményei, 1964, 9A, № 3, 555—563.
- 103 B é k é s s y A., Remarks on beta distributed random numbers. Magyar Tud. Akad. Mat. Kut. Int. Közleményei, 1964, 9A, № 3, 565—571.
- 104 Berger M. J., Monte Carlo calculation of the penetration and diffusion of fast charged particles. [158], 135—215.
- 105 Berger M. J., Reflection and transmission of gamma radiation by barriers, Monte Carlo calculation by a collision-density method. J. Research NBS, 1955, 55, № 6, 343—350.
- 106 Bouquet G., Feix M., Nicourd P., Sajaloli C., Emploi de la méthode de Monte-Carlo pour la détermination du volume critique d'un cylindre. C. R. acad. sci., 1958, 246, № 9, 1382—1384.
- 107 Box G. E. P., Muller M. E., A note on the generation of random normal deviates. Ann. Math. Statistics, 1958, 29, № 2, 610—611.
- 108 Butcher J. C., Random sampling from the normal distribution. Comput. J., 1961, 3, № 4, 251—253.

109. Butcher J. C., Messel H., Electron number distribution in electron-photon showers. *Phys. Rev.*, 1958, **112**, № 6, 2096—2106.
110. Butler J. W., Machine sampling from given probability distributions. [170], 249—264.
111. Cashwell E. D., Everett C. J., A practical manual on the Monte Carlo method for random walk problems. Pergamon Press, 1959.
112. Coleman W. A., Mathematical verification of a certain Monte Carlo sampling technique and applications of the technique to radiation transport problems. *Nucl. Sci. Engng*, 1968, **32**, № 1, 76—81.
113. Courant R., Friedrichs F., Lewy H., Über die partiellen Differenzengleichungen der mathematischen Physik. *Math. Annalen*, 1928, **100**, 32—74.
114. Curtiss J. H., «Monte Carlo» methods for the iteration of linear operators. *J. Math. Phys.*, 1954, **32**, № 4, 209—232.
115. Curtiss J. H., A theoretical comparison of the efficiencies of two classical methods and a Monte Carlo method for computing one component of the solution of a set of linear algebraic equations. [170], 191—233.
116. Cutkosky R. E., A Monte Carlo method for solving a class of integral equations. *J. Research NBS*, 1951, **47**, № 2, 113—115.
117. Dahlquist G., Monte-Carlo-metoden. *Nordisk Mat. Tidskr.*, 1954, **2**, 27—43.
118. Davis D. H., Critical-size calculations for neutron systems by the Monte Carlo method. [158], 67—88.
119. Donsker M. D., Kac M., A sampling method for determining the lowest eigenvalue and the principal eigenfunction of Schrödinger's equation. *J. Research NBS*, 1950, **44**, 551—557.
120. Dupač V., Stochastické početni metody. *Cas. pro pešt. mat.*, 1956, **81**, № 1, 55—68.
121. Forsyth G. E., Leibler R. A., Matrix inversion by a Monte Carlo method. *Math. Tables Aids Comput.*, 1950, **4**, № 31, 127—129.
122. Fortet R., On the estimation of an eigenvalue by an additive functional of a stochastic process, with special reference to the Kac—Donsker method. *J. Research NBS*, 1952, **48**, № 1, 68—75.
123. Fosdick L. D., Approximation of a class of Wiener integrals. *Math. of Comput.*, 1965, **19**, № 90, 225—233.
124. Fosdick L. D., The Monte Carlo method in quantum statistics. *SIAM Rev.*, 1968, **10**, № 3, 315—328.
125. Fosdick L. D., Jordan H. F., Path-integral calculation of the two-particle Slater sum for He^4 . *Phys. Rev.*, 1966, **143**, № 1, 58—66.
126. Franklin J. N., Deterministic simulation of random processes. *Math. of Comput.*, 1963, **17**, № 81, 28—59.
127. Goertzel G., Kalos M., Monte Carlo methods in transport problems. *Progress in nuclear energy, ser. 1: Phys. and Math.*, 1958, **2**, 315—369.
128. Haber S., A modified Monte Carlo quadrature. *Math. of Comput.*, 1966, **20**, № 95, 361—368; 1967, **21**, № 99, 388—397.

129. Haber S., Stochastic quadrature formulas. *Math. of Comput.*, 1969, 23, 108, 751—764.
130. Haber S., Numerical evaluation of multiple integrals. *SIAM Rev.*, 1970, 12, № 4, 481—526.
131. Malton J. H., On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. *Numer. Math.*, 1960, 2, № 2, 84—90.
132. Halton J. H., A retrospective and prospective survey of the Monte Carlo method. *SIAM Rev.*, 1970, 12, № 1, 1—63.
133. Halton J. H., Handscomb D. C., A method for increasing the efficiency of Monte Carlo integrations. *J. Assoc. Comput. Machinery*, 1957, 4, № 3, 329—340.
134. Hammersley J. M., Monte Carlo methods for solving multivariable problems. *Ann. New York Acad. Sci.*, 1960, 86, № 3, 844—874.
135. Hammersley J. M., Handscomb D. C., Monte Carlo methods. Methuen, London, 1964.
136. Hammersley J. M., Morton K. W., A new Monte Carlo technique: antithetic variates. *Proc. Cambridge Philos. Soc.*, 1956, 52, № 3, 449—475.
137. Handscomb D. C., Remarks on a Monte Carlo integration method. *Numer. Math.*, 1964, 6, № 4, 261—268.
138. Hlawka E. Uniform distribution mod. 1 and numerical analysis *Compositio Math.* 16, № 1—2, 92—105.
139. Hull T. E., Dobell A. R., Random number generators. *SIAM Rev.*, 1962, 4, № 3, 230—254.
140. Jansson B., Random numbers generators. Stockholm, 1966.
141. Jönnk M. D., Erzeugung von betaverteilten und gamma-verteilten Zufallszahlen, *Metrika*. 1964, 8, № 1, 5—15.
142. Kahn H., Random sampling (Monte Carlo) techniques in neutron attenuation problems. *Nucleonics*, 1950, 6, № 5, 27—33; 1950, 6, № 6, 60—65.
143. Kahn H., Use of different Monte Carlo sampling techniques. [170], 146—190.
144. Kalos M. H., Monte Carlo method for estimation of flux at a point. *Nuclear Sci. Engng.* 1963, 16, № 1, 111—117.
145. Kalos M. H., Importance sampling in Monte Carlo shielding calculations. *Nuclear Sci. Engng.* 1963, 16, № 2, 227—234.
146. Kalos M. H., Stochastic wave function for atomic Helium. *J. Comput. Phys.*, 1966, 1, № 2, 257—275.
147. Kendall M. G., Smith B. B., Randomness and random sampling numbers. *J. Royal Statist. Soc.*, 1938, 101, № 1, 147—166.
148. Kschwendt H., Convergence limits in the Monte Carlo theory of integral equations. *Numer. Math.*, 1968, 11, № 4, 307—314.
149. Laurent P.-J., Remarque sur l'évaluation d'intégrales par la méthode de Monte-Carlo. *C. R. acad. sci.*, 1961, 253, № 4, 610—611.
150. Lawande S. V., Jensen C. A., Sahlin H. L., Monte Carlo integration of the Feynman propagator in imaginary time. *J. Comput. Phys.*, 1969, 3, № 3, 416—443.

151. Lehmer D. H., Mathematical methods in large-scale computing units. Proc. second symposium on large-scale digital calculating machinery (1949). Harvard Univ. Press, 1951, 141—146.
152. Leimdörfer M., On the use of Monte Carlo methods for calculating the deep penetration of neutrons in shields. Chalmers Tekn. Högsk. Handl., Gothenburg (Sweden), 1964, № 287.
153. MacMillan D. B., Comparison of statistical estimators for neutron Monte Carlo calculations. Nuclear Sci. Engng, 1966, 26, № 3, 366—372.
154. Mann N. B., Wald A., On the choice of the number of class intervals in the application of the chi square test. Ann. Math. Stat., 1942, 13, № 3, 306—317.
155. Marsaglia G., Expressing a random variable in terms of uniform random variables. Ann. Math. Stat. 1961, 32, № 3, 894—898.
156. Marsaglia G., MacLaren M. D., Bray T. A., A fast procedure for generating normal random variables. Comm. Assoc. Comput. Machinery, 1964, 7, № 1, 4—10.
157. Marshall A. W., The use of multi-stage sampling schemes in Monte Carlo computations. [170], 123—140.
158. Methods in computational physics, ed. B. Alder, S. Fernbach, M. Rotenberg. Vol. 1: Statistical Physics. Acad. Press, New York, 1963.
159. Metropolis N., Ulam S. M., The Monte Carlo method. J. Amer. Statist. Assoc., 1949, 44, № 247, 335—341.
160. Morton K. W., A generalization of the antithetic variate technique for evaluating integrals. J. Math. Phys., 1957, 36, № 3, 289—293.
161. Morton K. W., Scaling neutron tracks in Monte Carlo shielding calculations. J. Nuclear Energy, 1957, 5, № 3/4, 320—324.
162. Muller M. E., Some continuous Monte Carlo methods for the Dirichlet problem. Ann. Math. Stat., 1956, 27, № 3, 569—589.
163. Neumann J., Various techniques used in connection with random digits. NBS Appl. Math. series, 1951, № 12, 36—38.
164. Powell M. J. D., Swann J., Weighted uniform sampling — a Monte Carlo technique for reducing variance. J. Inst. Math. and its Appl., 1966, 2, № 3, 228—236.
165. Quenouille M. H., Tables of random observations from standard distributions. Biometrika, 1959, 46, № 1/2, 178—202.
166. RAND Corporation. A million random digits with 100 000 normal deviates. The Free Press, 1955.
167. Satoshi I., A new pseudo-random generator using binary shift-register. Bull. Tokyo Inst. Technology, 1965, № 68, 1—7.
168. Spanier J., Gelbard E. M., Monte Carlo principles and neutron transport problems. Addison-Wesley, Reading, 1969.
169. Steen N. M., A simple method to improve the efficiency of the Σ_a/Σ_t estimator in Monte Carlo programs. Trans. Amer. Nuclear Soc., 1966, 9, № 1, 204.
170. Symposium on Monte Carlo methods, ed. H. A. Meyer. Wiley, New York, 1956.
171. Tausky O., Todd J., Generation and testing of pseudo-random numbers. [170], 15—28.

172. Tausworthe R. C., Random numbers generated by linear recurrence modulo two. *Math. of Comput.*, 1965, 19, № 90, 201—209.
173. Thouless D. J., Integration of the Schrödinger equation by a Monte Carlo method. *Proc. Phys. Soc.*, 1965, 86, № 553, 905—913.
174. Tocher K. D., The application of automatic computers to sampling experiments. *J. Royal Statist. Soc., ser. B*, 1954, 16, № 1, 39—61.
175. Todd J., Experiments in the solution of differential equations by Monte Carlo methods. *J. Washington Acad. Sci.*, 1954, 44, № 12, 377—381.
176. Törn A., Crude Monte Carlo-quadrature in infinite variance case and the central limit theorem. *Nordisk. Tidskr. Inform. Behandl.*, 1966, 6, № 4, 339—346.
177. Tsuda T., Ichida K., Nonlinear interpolation of multivariable functions by the Monte Carlo method. *J. Assoc. Comput. Machinery*, 1970, 17, № 3, 420—425.
178. Tsuda T., Ichida K., Kiyono T., Monte Carlo path-integral calculations for two-point boundary-value problems. *Numer. Math.*, 1967, 10, № 2, 110—116.
179. Tsuda T., Matsumoto H., A note on linear extrapolation of multivariable functions by the Monte Carlo method. *J. Assoc. Comput. Machinery*, 1966, 13, № 1, 143—150.
180. Verdier P. H., Relations within sequences of congruential pseudo-random numbers. *J. Research NBS. ser. B*, 1969, 73B, № 1, 41—44.
181. Wasow W., A note on the inversion of matrices by random walks. *Math. Tables Aids Comput.*, 1952, 6, 78—81.
182. Weyl H., Über die Gleichverteilung von Zahlen mod Eins. *Math. Annalen*, 1916, 77, № 3, 313—352.
183. Zaremba S. K., The mathematical basis of Monte Carlo and quasi-Monte Carlo methods. *SIAM Rev.*, 1968, 10, № 3, 303—314.
184. Zieliński R., *Metody Monte Carlo*. Warszawa, 1970.
185. Zierler N., Linear recurring sequences. *J. SIAM*, 1959, 7, № 1, 31—48.

УКАЗАТЕЛЬ

- Алгебраические системы 193
 Алгоритм Монте-Карло 98, 254
 — — —, трудоемкость 99
 — — —, конструктивная размерность 255
 Асимптотически вполне равномерное распределение 273
- Бесконечномерная последовательность точек 270
 Блуждания по сферам 285—286
 Броуновская траектория 288
- Вариационный ряд 36
 Вероятная ошибка 88
 Вероятность доверительная 31
 — перехода 163
 — поглощения 168
 — рассеяния 47, 168
 Веса 164, 168
 — векторные 244
 — статистические 239, 240
 Ветвящаяся траектория 213
 Взвешенная равномерная выборка 152
 Винеровский интеграл 287
 Включение особенности в плотность 111
 Вполне равномерное распределение 271
 Выборка 92
 — взвешенная равномерная 152
 — по группам 135
 — систематическая 240
 — существенная 108
 Выборка типическая 138
- Выделение главной части 101
 Вычисление обратной матрицы 198
 — одной компоненты решения 197
- Геометрический метод Монте-Карло 94
 Гистограмма 219
 Группировка 32
- Датчик случайных чисел 15
 Двухэтапная схема расчета 120
 Дерево 213
 —, лексикографический обход 214
 —, обход по поколениям 214
 Длина отрезка апериодичности 21
 — свободного пробега 221
 Доверительная вероятность 31
 Доверительный интервал 92
 Допустимая плотность 108
 Допустимое распределение 196
 Дробная часть $D(x)$ числа x 12
- Зависимые испытания 123
 Задача Дирихле 202, 284
- Имитация 210
 Интеграл вероятностей 88
 — Винера 287
 — континуальный 287
 Интегральное уравнение неоднородное 171

Интегральное уравнение Пай-
ерлса 189
— теории рассеяния 182
Интегральной суперпозиции ме-
тод 67
Интегрирование по части обла-
сти 102
— — — переменных 105
Интенсивность потока заявок 216
Интерполирование 279
Итерации 161
Итерированные функции 161

Квадратурная формула 143
— — — интерполяционная 144
— — — случайная 143
— — — интерполяционная 146
Квазислучайные числа 253
Класс функций $L_2(G)$ 162
— $L_2(G; p)$ 96
Конструктивная размерность ал-
горитма 255
Континуальный интеграл 287
Коэффициент доверия 31, 92
— Фурье 167
К. р. 255
Критерий согласия Колмогорова
43
— — χ^2 31
— — ω^2 34
Критическая область 214

ЛП-поиск 283
ЛПт-последовательность 265
— обобщенная 270

Матрицы обращение 198
Метод возмущений 28
— вычетов 22
— Неймана 75
— обратных функций 49
— подобных траекторий 242
— постоянного сечения 227
— псевдослучайных чисел 16
— середины квадрата 22
— сравнений 22
— суперпозиции 64
— — интегральной 67
— — модифицированный 66
— существенной выборки 108
— — для траекторий 173, 184,
187

Методы отбора 74
— перемешивания 24
— уменьшения дисперсии 100
Многогрупповая теория переноса
244
Моделирование биномиального
распределения 72
— бэта-распределения 85
— в полярных координатах 62
— гамма-распределения 71
— дискретной случайной величин-
ны 44
— комптон-эффекта 84
— многомерного нормального
распределения 60
— многомерной случайной точки
54
— — — с независимыми коор-
динатами 53
— непрерывной случайной вели-
чины 48
— неупругого рассеяния нейтро-
нов 72
— нормального распределения 63,
73
— равномерного распределения в
параллелепипеде 54
— — — в пирамиде 85
— — — в симплексе 85
— — — в сложных областях 79
— — — в шаре 58, 81
— — — на поверхности сферы 59,
85
— свободного пробега заряжен-
ной частицы 229
— — — нейтрона в кусочно одно-
родной среде 224
— — — в однородной среде 223
— — — в произвольной среде
227
— случайного направления 59
— — события 46
— «судьбы» нейтрона при столк-
новении 47
— усеченных распределений 75
— экспоненциального распреде-
ления 49, 84
— энергии нейтронов деления 84

Начальная плотность 164
Нейтроны деления 84, 213

- Неравенство Коши-Буняковского 292
 — Чебышева 141
 Неравномерность 262
 Одногрупповая теория переноса 222
 Оптическая длина 222
 Отклонение 261
 Отрезок аперидичности 21
 Оценка качества 220
 — квадратичного функционала 207
 — несмещенная 92
 — по поглощениям 209
 — по столкновениям 209
 — погрешности 87, 89, 128
 — сверхэффективная 141
 — смещенная 92, 151
 — трудоемкости 99
 — усложненная 181
- Плотность бэта-распределения 85
 — вероятностей перехода 163
 — гамма-распределения 71
 — допустимая 108
 — начальная 164
 — первых столкновений 247
 — распределения Стюдента 90
 — — χ^2 31
 — столкновений 182
 Погрешности оценка 87, 89, 128
 Подобных траекторий метод 242
 Последовательность точек бесконечномерная 270
 — — равномерно распределенная 259
 — Холтона 264
 — — обобщенная 271
 — чисел асимптотически вполне равномерно распределенная 273
 — — вполне равномерно распределенная 271
 Постоянного сечения метод 227
 Поток Пуассона 216
 — — нестационарный 216
 — Эрланга 216
 Правило «трех сигм» 88
 Преобразование плотности 58
 Простейший метод Монте-Карло 93
- — — — с поправочным множителем 153
 Проверка интервалов 37
 — комбинаций 37
 — пар 37
 — псевдослучайных чисел 40, 274
 — серий 37
 — частот 37
 Псевдослучайные числа 17
- Равномерно распределенная последовательность 259
 Разыгрывание случайной величины 44
 Распределение допустимое 196
 — длины свободного пробега 221
 — Колмогорова 43
 — равномерное в многомерной пирамиде 85
 — — в многомерном параллелепипеде 54
 — — — симплексе 85
 — — в шаре 58, 81
 — — на поверхности многомерной сферы 85
 Рассеяние частиц 182
 Расщепление траектории 239
 Ряд Неймана 171
- Сечение взаимодействия 221
 Символ Кронекера δ_{ij} 167
 Симметризация 116, 159
 Система массового обслуживания 215
 Систематическая выборка 240
 Случайная интерполяционная квадратурная формула 146
 — квадратурная формула 143
 — траектория T_i 163
 — — T_∞ 174
 — — ветвящаяся 213
 — — —, лексикографический обход 214
 — — —, обход по поколениям 214
 — — с поглощением 168
 Случайная цепь 195
 — цифра 10

- Случайное направление 59
 — событие 46
 — число 10
 Случайный обрыв траектории 239
 — поиск 281
 — поток заявок 216
 Случайных цифр таблица 13, 131, 295
 — чисел датчик 15
 Сопряженное уравнение 179
 Суперпозиции метод 64
 Существенная выборка 108
 Сходимость по вероятности 34
 Счет по поколениям 214
- Таблица случайных цифр 13, 131, 295
 Теорема Колмогорова 43
 — Мизеса — Смирнова 34
 — Пирсона 30
 — Фишера 90
 — Хинчина 87
 — Чебышева 286
 Точность метода Монте-Карло 96
 Траектории броуновские 288
 — ветвящиеся 213
 — с поглощением \tilde{T}_v 168
 — T_i 163
 — T_∞ 174
 Трудоемкость алгоритма Монте-Карло 99
- Уравнение дифференциальное Лапласа 201, 284
- — Пуассона 201
 — — теплопроводности 209
 — интегральное неоднородное 171
 — — однородное 185
 — — Пайерлса 189, 252
 — мажорантное 177
 — моноциклическое 28
 — сопряженное 179
 — теории рассеяния 182
 — Шредингера 291
 Уровень значимости 31
- Функции от большого числа переменных 134, 206, 279
- Целая часть $C(x)$ числа x 12
- Эмпирическая оценка дисперсии 89
 — функция распределения 34
 Эффективное сечение взаимодействия 221
 Эффективность метода Неймана 77
 — отбора 74
 Эффективный коэффициент размножения нейтронов 214
- Ядро столкновений 182, 222
 — интегрального уравнения 171

Илья Меерович Соболев
ЧИСЛЕННЫЕ МЕТОДЫ
МОНТЕ-КАРЛО

М., 1973 г., 312 стр. с илл.

Редактор Г. Я Пирогова
Техн. редактор Е. Н. Земская
Корректор В. П. Сорокина

Сдано в набор 13/VII 1973 г. Подписано к печати
17/XII 1973 г. Бумага 84×108/32. Физ. печ. л. 9,75.
Условн. печ. л. 16,38. Уч.-изд. л. 16,08 Тираж
22 000 экз. Т-19945. Цена книги 1 р. 22 к.
Заказ № 132

Издательство «Наука»
Главная редакция
физико-математической литературы
117071, Москва, В-71, Ленинский проспект, 15

4-я типография изд. ва «Наука».
Новосибирск, 77, Станиславского, 25,