Глава 2

Теория вероятностей и математическая статистика

2.1. Введение

В общем случае исследуемая система содержит ряд элементов, обладающих некоторой неопределенностью. Подобные системы обычно называют стохастическими, так как их поведение не может быть полностью предсказано заранее. При имитации стохастических систем требуется описывать изменчивость элементов в терминах теории вероятностей. Поскольку результаты, полученные с помощью имитационной модели, также носят вероятностный характер, они требуют статистической интерпретации. Мы предполагаем, что читатель уже знаком с теорией вероятностей и статистикой, и в данной главе приводим лишь некоторые основные вероятностные и статистические понятия, относящиеся непосредственно к имитационному моделированию[1]. Кроме того, рекомендуем читателю предварительно ознакомиться с одной из книг, указанных в библиографии в конце данной главы [7, 12, 14, 19].

2.2. Эксперимент, пространство выборки и результат

Эксперимент является хорошо отработанной процедурой или процессом, результат которого можно наблюдать, но нельзя точно предсказать. Множество всех возможных результатов называется пространством выборки. Если пространство выборки является конечным или конечно-счетным, оно называется дискретным, в противном случае — непрерывным.

Комбинация результатов при помощи операций теории множеств, таких, например, как объединение ($\cup$) и пересечение ($\cap$), может давать новые результаты[2]. Если результат C определяется как объединение множества результатов A и множества результатов B, т. е. $C = A \cup B$, это означает, что C состоит из множества всех результатов, принадлежащих A или B. Если результат O является пересечением A и B, т. е. $D = A \cap B$, это означает, что D является множеством результатов, принадлежащих и A и B.

В качестве примера использования приведенных выше понятий рассмотрим модель банка с одним кассиром. Посетители приходят в банк, ожидают обслуживания и обслуживаются кассиром. Будем полагать, что интервалы времени между приходами посетителей и продолжительностью их

---

1 Читатели, интересующиеся проблемами статистики, относящимися только к имитации, могут сразу переходить к чтению разд. 2.15.

2 Обычно для описания комбинаций результатов используется термин «событие». В имитационной терминологии понятие события имеет иной смысл, поэтому здесь мы его употреблять не будем.

обслуживания кассиром имеют разброс. Пусть наш первый эксперимент заключается в наблюдении интервалов времени между приходами посетителей в банк. Пространство выборки данного эксперимента состоит, таким образом, из всевозможных наблюдений за интервалами между приходами посетителей. Поскольку интервал времени может быть любым неотрицательным вещественным числом, пространство выборки является непрерывным. Результат определяется как подмножество пространства выборки. Таким результатом может быть, например, совокупность интервалов времени продолжительностью от 8 до 9 мин.

В качестве второго примера рассмотрим эксперимент, состоящий в наблюдении числа посетителей, обслуженных в течение первого часа работы банка. Это число может быть любым неотрицательным целым числом, следовательно, пространство выборки дискретно. Результат же, например, может быть определен как обслуживание пяти посетителей в течение первого часа работы банка.

2.3. Вероятность

Вероятность является мерой возможности осуществления результата. Формально мера вероятности является функцией P( ), которая ставит в соответствие результатам некоторые вещественные числа и удовлетворяет следующим аксиомам:

1. $0 \leq P(E) \leq 1$ для любого результата E.

2. $P(S) = 1$, где S - пространство выборки, или «достоверный результат».

3. Если E_1, E_2, E_3,... - взаимно исключающие результаты, то $P(E_1 \cup E_2 \cup E_3 \cup \ldots) = P(E_1) + P(E_2) + P(E_3) + \ldots$.

Из этих трех аксиом и правил теории множеств могут быть выведены основные законы теории вероятностей. Однако использования этих аксиом недостаточно для вычисления вероятности результата. Обычно числовые значения вероятностей получить довольно сложно, хотя весьма полезно постулировать их существование.

В ряде простых случаев точную вероятность результата можно вычислить, используя комбинаторный анализ. Примером может служить вычисление вероятности выпадения h «орлов» при n бросаниях монеты или вычисление вероятности нахождения трех тузов в раздаче из пяти карт. Однако в большинстве случаев точная вероятность результатов не может быть вычислена. Иногда в таких случаях может быть получено приблизительное значение вероятности результата с помощью ее частотной интерпретации. Если при повторении эксперимента n раз результат E произойдет k раз, то отношение k/n будет частотой появления E. Вероятность результата E может быть выражена так:

$$P(E) = \lim_{n \to \infty} k/n$$

при условии существования данного предела. При выборе достаточно большого значения n частота k/n будет асимптотически стремиться к вероятности результата E. Можно показать, что полученные подобным образом приблизительные значения вероятностей удовлетворяют аксиомам вероятностей, определенным выше. На практике применение данного подхода ограничивается имеющимися возможностями или целесообразностью проведения требуемого эксперимента.

2.4. Случайные величины и вероятностные распределения

Функция, которая ставит в соответствие каждому результату из пространства выборки некоторое вещественное число, называется случайной величиной. Дискретными называются те случайные величины, которые принадлежат конечному или счетному множеству значений. Непрерывные случайные величины могут принадлежать континууму значений. В нашей модели банка с одним кассиром интервал времени между приходами посетителей является непрерывной случайной величиной, а число посетителей, обслуженных за первый час работы банка, — дискретной.

Вероятностное распределение представляет собой некоторое правило задания вероятности для каждого из всех возможных значений случайной переменной. Правило задания вероятности имеет две различные формы в зависимости от того, является случайная величина дискретной или непрерывной.

Для дискретной случайной величины вероятность каждого ее значения задается функцией вероятности $p(x)$, определяемой как[1]

$$p(x_1) = P(X = x_1,).$$

Для каждого возможного значения x_1 функция устанавливает конкретную вероятность того, что случайная переменная X принимает значение x_1. Аксиомы вероятностей накладывают следующие ограничения на $p(x_1)$:

$$0 \leq p(x_1) \leq 1 \text{ для всех } i,$$

$$\sum_i p(x_1) = 1.$$

Альтернативой функции вероятности является функция распределения или кумулятивная функция распределения $F(x)$, определяемая следующим образом:

---

[1] При определении понятий теории вероятностей и математической статистики мы будем стараться использовать прописные буквы для обозначения случайных величин, а строчные для обозначения их возможных значений. При отступлении от этого правила смысл того или иного обозначения будет обязательно поясняться.

$$F(x) = P(X \le x).$$

Здесь функция $P(x)$ определяет вероятность того, что случайная величина X примет значение не большее, чем x. Из аксиом вероятностей вытекают следующие свойства $P(x)$:

$$O \le F(x) \le 1 \quad \text{для всех } x,$$

$$F(-\infty) = 0,$$

$$F(+\infty) = 1.$$

Функция распределения связана с функцией вероятности следующим образом:

$$F(x) = \sum_{x_i<x} p(x_i)$$

В качестве примера дискретного распределения вероятности рассмотрим эксперимент, заключающийся в трех бросаниях монеты. Пусть случайная величина X обозначает число «орлов», выпавших в результате трех бросаний. Случайная величина X может, таким образом, принимать целые значения от 0 и до 3. При этом возможны восемь результатов, из которых в одном выпадет 0 «орлов», в трех — 1 «орел», в трех — 2 «орла» и в одном — 3 «орла». На рис.2.1 показана функция вероятности для переменной X, а на рис. 2.2 — функция распределения.

Для непрерывных случайных величин требуется иная форма представления вероятностного распределения. Поскольку случайная величина может принимать любое из бесконечного-несчетного множества значений, вероятность конкретного значения равна нулю.

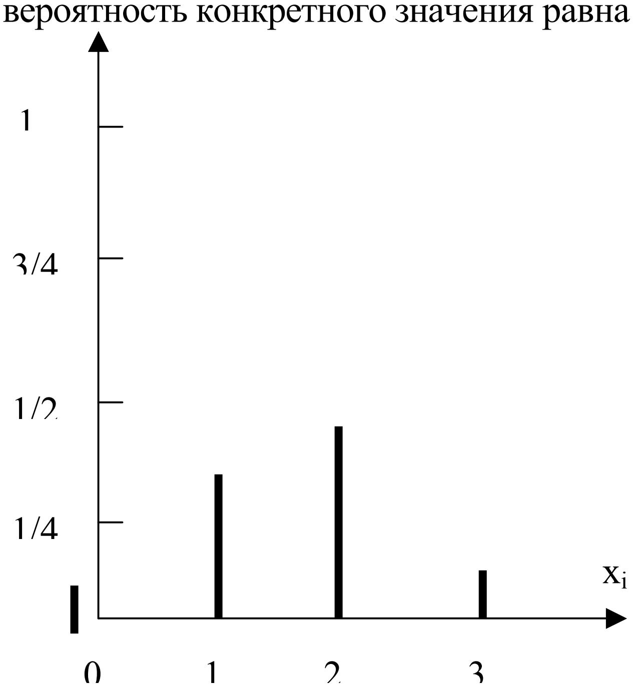


Рис. 2.1 Пример функции вероятности (дискретный случай)

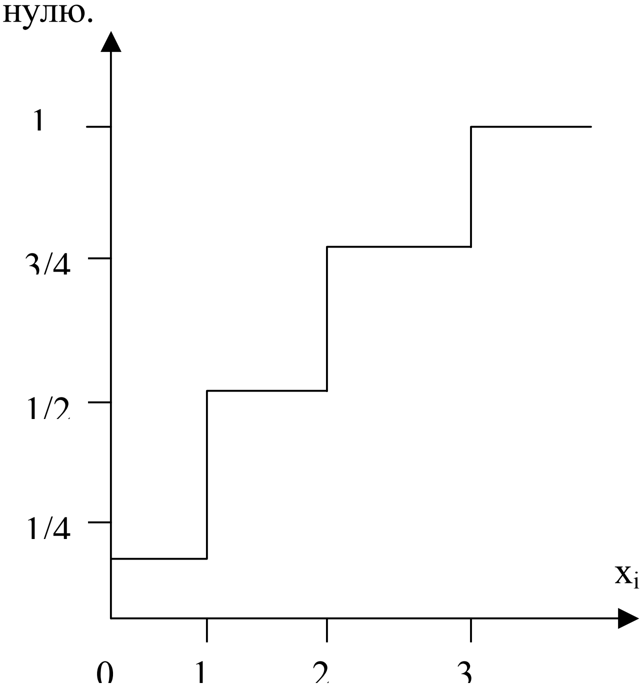


Рис. 2.2 Пример распределения (дискретный случай)

Это говорит не о том, что данное значение невозможно, а о том, что оно крайне невероятно вследствие бесконечного числа альтернативных значений. При этом, конечно, вероятность того, что переменная примет значение в интервале между точками a и b, в большинстве случаев не будет равна нулю. Следовательно, функция вероятности для дискретного случая заменяется на непрерывную функцию плотности вероятности $f(x)$, определяемую следующим выражением:

$$P(a \le X \le b) = \int_a^b f(x)dx$$

Таким образом, функция плотности вероятности при интегрировании на интервале от a до b дает вероятность того, что случайная величина примет значение из этого интервала. В соответствии с аксиомами вероятностей функция плотности должна удовлетворять следующим условиям:

$$f(x) \ge 0, \int_a^b f(x)dx = 1$$

Функция распределения $F(x)$ определяется для непрерывных случайных величин следующим образом:

$$F(x) = \int_{-\infty}^{x} f(y)dy = P(X \le x)$$

Функция $P(x)$ определяет вероятность того, что непрерывная случайная величина X примет значение, не большее их.

В качестве примера непрерывного вероятностного распределения рассмотрим случайную величину X, которая может принимать значения на интервале от 0 до 1. На рис. 2.3 и 2.4 показаны функция плотности вероятности и функция распределения соответственно для случая, когда любое из несчетного бесконечного множества значений равновероятно. Вероятность того, что случайная величина примет значение на интервале от 0,5 до 0,75, равна площади под кривой функции плотности вероятности на этом интервале. Для случайной величины на рис. 2.3 эта вероятность равна 0,25.

Случайная величина может быть одновременно и дискретной, и непрерывной. Обычно говорят, что такая случайная величина имеет «смешанное» распределение. Случайная переменная со смешанным распределением может принимать как дискретные значения с конечной вероятностью, так и континуум значений, описанных функцией плотности вероятности. На рис. 2.5 изображено смешанное распределение с дискретными значениями 1 и 2, вероятность каждого из которых равна 1/3, что показано вертикальными прямыми со стрелками. Значения между 1 и 2 описываются функцией плотности f(x) = 1/3.

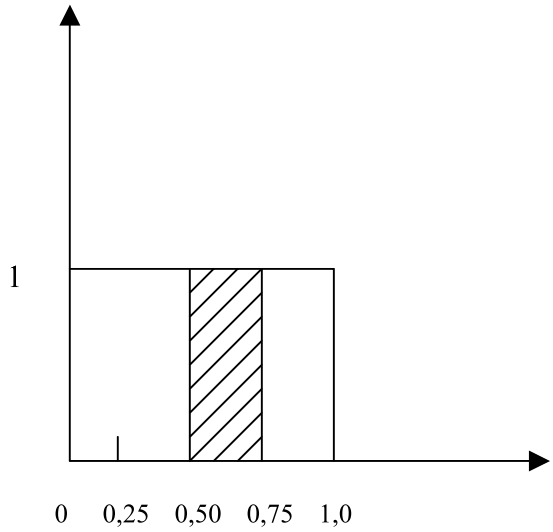


Рис. 2.3. Пример функции плотности вероятности (непрерывный случай)

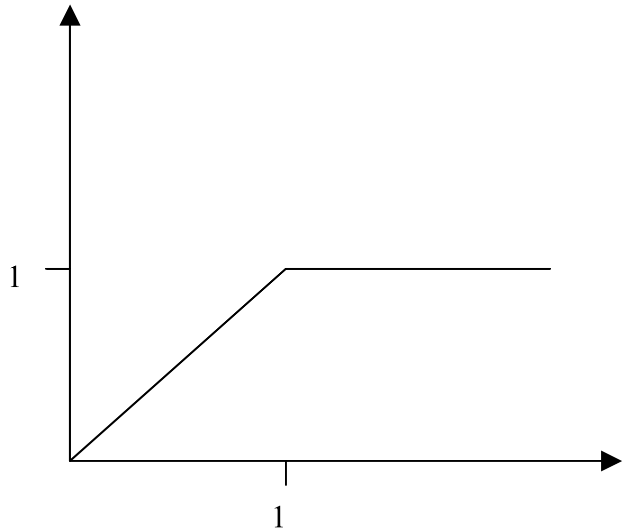


Рис. 2.4. Пример функции распределения вероятности (непрерывный случай)

Подобное распределение можно получить из равномерного непрерывного распределения на интервале от 0 до 3, присваивая всем значениям, большим 2, значение, равное 2, а значениям, меньшим 1, — значение, равное 1. Уравнение для функции распределения этой случайной величины имеет следующий вид:

$$F(x) = \begin{cases} 0, x < 1 \\ 1/3 + (x-1)/3, 1 \le x < 2, \\ 1, x \ge 2 \end{cases}$$

Из этого уравнения и рис. 2.5 видно, что F(x) имеет разрывы в точках x=1 и x=2. В этих точках P(X=x) равна величине скачка, который делает F(x) в точке x. Например, P(X=1)=1/3. В то же время при 1<x<2 функция F(x) непрерывна по x, а P(X=x)=0.

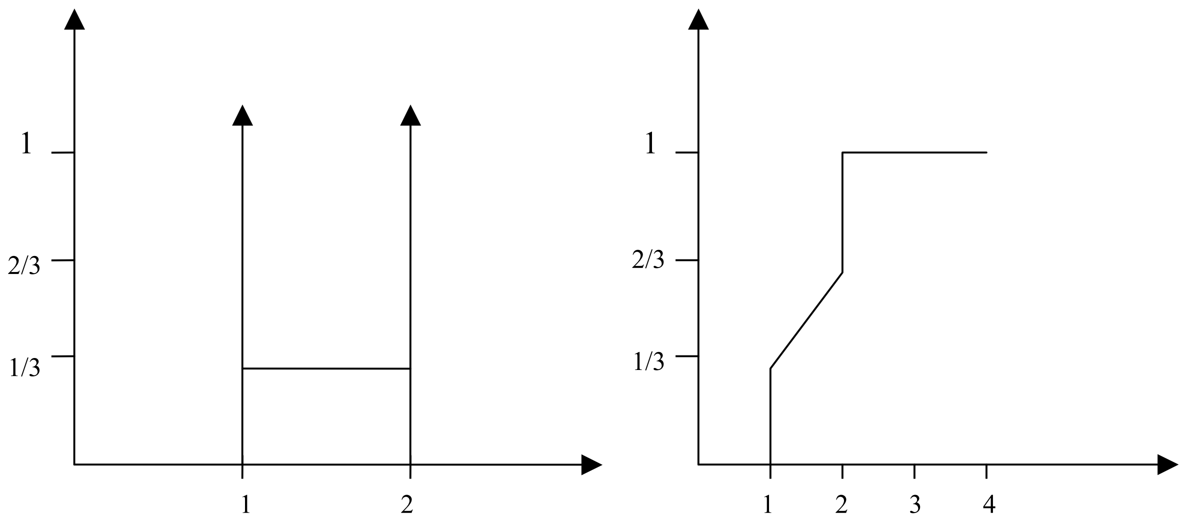


Рис. 2.5. Пример смешенного распределения

2.5. Математическое ожидание и моменты

Часто необходимо охарактеризовать случайную переменную одним или несколькими значениями, которые суммируют информацию, содержащуюся в функции распределения вероятности. Математическим ожиданием случайной величины X, обозначаемым E[X], является значение, определяемое следующим образом:

$E[X] = \sum x_i\, p(x_i)$, если величина X дискретна;

$3[X] = {}_x\!\int xF(x)dx$, если величина X непрерывна.

Математическим ожиданием является, таким образом, взвешенная по вероятности средняя величина всех возможных значений X, определяющая меру центральности распределения. Поэтому эта величина часто называется средним значением.

Математическое ожидание можно находить также для функций случайных величин. В частности, математическое ожидание-X, называется n-м моментом случайной переменной и определяется следующим образом:

$E[X^n] = \sum_i x_i^n\, p(x_i)$, если величина X дискретна;

$E[X^n] = \int x^n f(x)dx$, если величина X непрерывна.

Математическое ожидание является частным случаем данного выражения при n = 1 и называется первым моментом.

Вариацией n-го момента называется n-й момент среднего, который определяется выражением

$$E[(X-E[X])^n].$$

Следовательно, перед вычислением n-го момента математическое ожидание X вычитается из X.

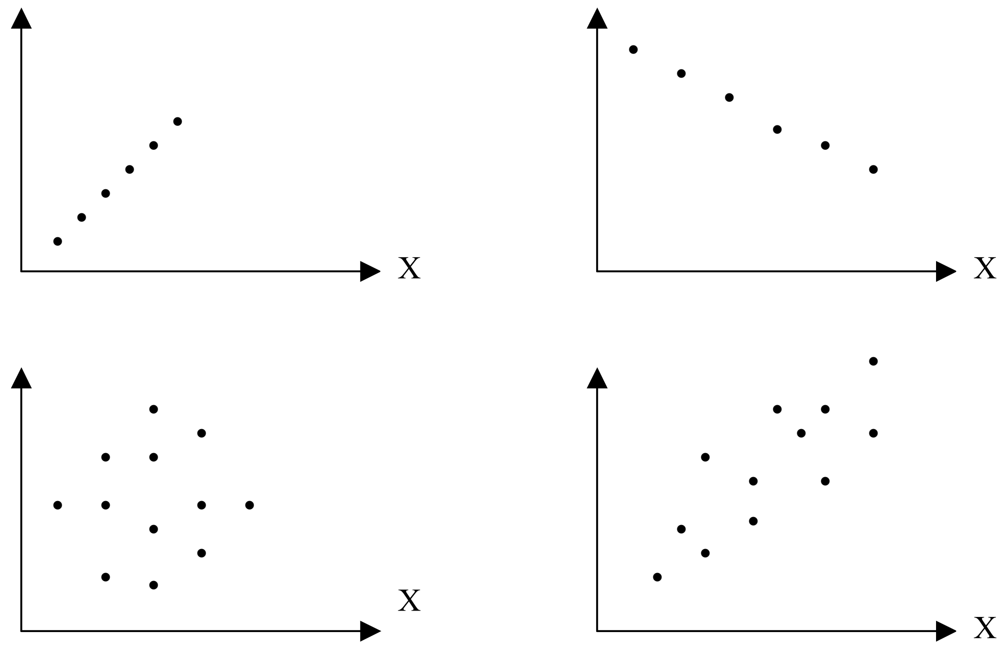


Рис. 2.6. Зависимость Y и X для различных значений p

Особое значение в теории вероятностей имеет второй момент среднего, называемый обычно дисперсией X и обозначаемый как σ^2. Дисперсия случайной переменной является мерой разброса вероятностного распределения. Если дисперсия случайной величины мала, вся выборка лежит вблизи математического ожидания. Квадратный корень из дисперсии называется среднеквадратичным отклонением случайной величины.

Если X и Y — случайные величины, то ковариацией X и Y - называется величина Cov[X,Y], определяемая следующим образом:

$$\mathrm{Cov}[X, Y] = E\,[(X-E\,[X])\,(Y-E\,[Y])].$$

Ковариация измеряет линейную связь между X и Y. Если результат X не влияет на результат Y, говорят, что X и Y независимы, а Cov[X,Y]=0. В общем случае X и Y независимы тогда и только тогда, когда

$p(y \mid x) = p(y)$ (в дискретном случае $p(y \mid x)$ является вероятностью того, что Y = y при X = x);

$f(y \mid x) = f(y)$ (в непрерывном случае $f(y \mid x)$ является условной функцией плотности Y при X = x).

Данные выражения утверждают, что вероятностное распределение Y при наличии информации об X аналогично вероятностному распределению Y при отсутствии информации об X.

Мерой зависимости, связанной с ковариацией, является коэффициент корреляции ρ, определяемый выражением

$$\rho = \frac{Cov[X,Y]}{\sqrt{Var[X]*Var[Y]}}$$

Коэффициент корреляции может лежать в интервале от –1 до + 1, причем нулевое значение свидетельствует об отсутствии корреляции между X и Y. Положительное значение ρ показывает, что Y увеличивается с увеличением X, а отрицательное значение — что Y уменьшается с увеличением X. Величина р отражает степень линейности зависимости Y от X. Если Y линейно зависит от X, то ρ = ±1. Если X и Y независимы, то график зависимости Y от X представляет собой набор случайных точек, а ρ = 0. На рис. 2.6 приведены типичные графики зависимости Y от X при различных значениях ρ

2.6. Функции случайных величин

Функция случайной величины также является случайной величиной. В данном разделе показан ряд важных свойств функций случайных величин.

Если X и Y — случайные величины, а k — некоторая константа, то для математического ожидания могут быть выведены следующие свойства:

$$E[X+Y] = E[X]+E[Y],$$
$$E[kX] = kE[X],$$
$$E[X + k] = E[X]+k.$$

Для дисперсии же аналогичные свойства являются менее очевидными:

$$Var[X+Y]=Var[X]+Var[Y]+2\ Cov[X,Y],$$
$$Var[кX]=k^2Var[X],$$
$$Var[X+k]=Var[X],$$
$$Var[kX+nY]=k^2Var[X]+n^2Var[Y]+2knCov[X,Y].$$

Отметим, что в случае независимости случайных величин X и Y ковариация Cov[X, Y] равна нулю, и, следовательно,

$$Var[X+Y]=Var[X]+Var[Y].$$

В математической статистике определенную роль играет случайная величина, называемая выборочным средним (средним по выборке) $\overline{X}_i$ где I — размер выборки из вероятностного распределения. Выборочное среднее определяется отношением суммы всех значений выборки к ее размеру, т. е.

$$\overline{X}_I = \frac{1}{i}\sum_{1}^{i} X_I$$

Предположив, что все X_i независимы и одинаково распределены (НОР), можно получить следующие свойства для математического ожидания и дисперсии $\overline{X}_i$

$$E[\overline{X}_I]=E[X]$$

$$Var[\overline{X}_I]=\frac{Var[X]}{I}$$

Дисперсия среднего по выборке размером I в I раз меньше, чем дисперсия случайной величины, по которой взята выборка. Следовательно, выбрав I достаточно большим, можно уменьшить дисперсию среднего до любой малой величины.

Отметим, что зависимости, приведенные выше для дисперсии $\overline{X}_I$ справедливы только для случая независимых наблюдений X_i. Если наблюдения не являются независимыми, вычисление $Var[X_i]$ требует принятия во внимание ковариации между ними. Например, в модели работы кассира в банке времена ожидания приходящих друг за другом посетителей будут коррелированны из-за наличия вероятности того, что (i+1) -й посетитель будет ждать дольше, если 1-й посетитель по прибытии в банк помещается в очередь, чем если 1-й посетитель сразу же обслуживается. Следовательно, дисперсия среднего времени ожидания не может быть оценена простым делением дисперсии времени ожидания на число наблюдений. Подобная последовательность коррелированных наблюдений называется автокоррелированной выборкой. В разд. 2.15 и гл. 13 мы уделим внимание проблеме оценки дисперсии выборочного среднего для автокоррелированных выборок.

2.7. Генераторы

Функции случайных величин часто называют функциями-генераторами. Существует целый ряд генераторов, однако мы рассмотрим только производящие функции вероятностей и производящие функции моментов. Производящая функция вероятностей для дискретной случайной величины имеет следующий вид:

$$A(s) = \sum p(x_i)s^i$$

Если вид A(s) известен, то значения $p(x_i)$ могут быть получены путем дифференцирования i раз по s при s = 0. Математическое ожидание X можно получить из A(s) взятием первой производной по s при s=1. Моменты более высокого порядка вычисляются аналогичным способом, но требуют комбинирования производных. Примером генератора случайной функции является Z-преобразование.

Производящая функция моментов (ПФМ) случайной величины определяется следующим образом:

$$M(s)=E[e^{sX}]$$

n-й момент вычисляется путем дифференцирования данного выражения по s:

$$\frac{d^n M(s)}{ds^n} = E[X^n e^{sX}]$$

При s = 0 имеем $E[X^n]$. Производящая функция моментов называется характеристической функцией.

Кроме вычисления моментов случайной величины по производящим функциям ПФМ полезны для вычисления моментов сумм независимых случайных величин. Например, если W = X+Y, причем X и Y независимы, то

$$E[e^{sW}] = E[e^{s(X+Y)}] = E[e^{sX}]E[e^{sY}]$$

Таким образом, ПФМ W является произведением ПФМ X и ПФМ Y. Моменты W могут быть получены затем с помощью соответствующей ПФМ. В ряде публикаций дается перечень производящих функций вероятностей и моментов [4, 10].

2.8. Закон больших чисел и центральная предельная теорема

Поведение $\overline{X}_I$ при бесконечно большом увеличении выборки определяется двумя весьма важными теоремами. Первой теоремой формулируется сильный закон больших чисел, действие которого интуитивно понятно и заключается в том, что по мере увеличения размера выборки I величина $\overline{X}_I$ стремится к E[X] с вероятностью, равной 1. С данной теоремой связано и действие слабого закона больших чисел, заключающегося в том, что

$$\lim_{I\to\infty} P\{|\overline{X}_I — E[X]|>\varepsilon\}=0 \text{ для любого положительного } \varepsilon.$$

Другими словами, для любого положительного, сколь угодно малого е вероятность того, что модуль разности $\overline{X}_I$ и E[X] превысит е, стремится к нулю при $I\to\infty$.

Второй важной теоремой, определяющей поведение $\overline{X}_I$, является центральная предельная теорема. Она утверждает, что при определенных благоприятных условиях распределение суммы I независимых наблюдений X стремится к нормальному (см. разд. 2.9), когда $I\to\infty$, независимо от характера распределения X. Отсюда следует также, что выборочные средние асимптотически нормально распределены при достаточно больших I. При этом, однако, трудно сказать, какой размер выборки является достаточным для того, чтобы считать величину X_I нормально распределенной. Очень часто достаточным бывает относительно небольшой размер выборки (примерно 10—15 наблюдений). Существует целый ряд модификаций центральной предельной теоремы. В частности, при определенных условиях она применима для последовательностей зависимых случайных величин. Описание этих условий приводится в приложении к гл. 13.

2.9. Распределения

В предыдущих разделах мы описывали свойства случайных величин и их распределений в общих чертах. Ниже описаны несколько конкретных распределений, весьма важных при моделировании случайных процессов. Описаны также характеристики этих распределений для того, чтобы сориентировать разработчика модели в выборе конкретного типа случайной величины при моделировании некоторого случайного процесса. Более формализованное обсуждение этой проблемы с подробными графическими иллюстрациями приводится в работах [11, 13].

Ниже будут использованы следующие обозначения:

X - случайная величина;

$f(x)$ - функция плотности вероятности X;

$p(x)$ - функция вероятности X;

a - минимальное значение;

b - максимальное значение;

m - мода;

μ - математическое ожидание $E[X]$;

σ^2 - дисперсия $E[(X—\mu)^2]$;

σ - среднеквадратичное отклонение;

α - параметр функции плотности вероятности;

β - параметр функции плотности вероятности.

Для тех функций плотности вероятности, которые нельзя записать с помощью μ и σ, приводятся выражения для μ и σ.

2.9.1. Равномерное распределение

Функция плотности вероятности равномерного распределения задает одинаковую вероятность для всех значений, лежащих между минимальным и максимальным значениями переменной. Другими словами, вероятность того, что значение попадает в указанный интервал, пропорциональна длине этого интервала. Применение равномерного распределения часто вызвано полным отсутствием информации о случайной величине, кроме ее предельных значений. Равномерное распределение называют также прямоугольным.

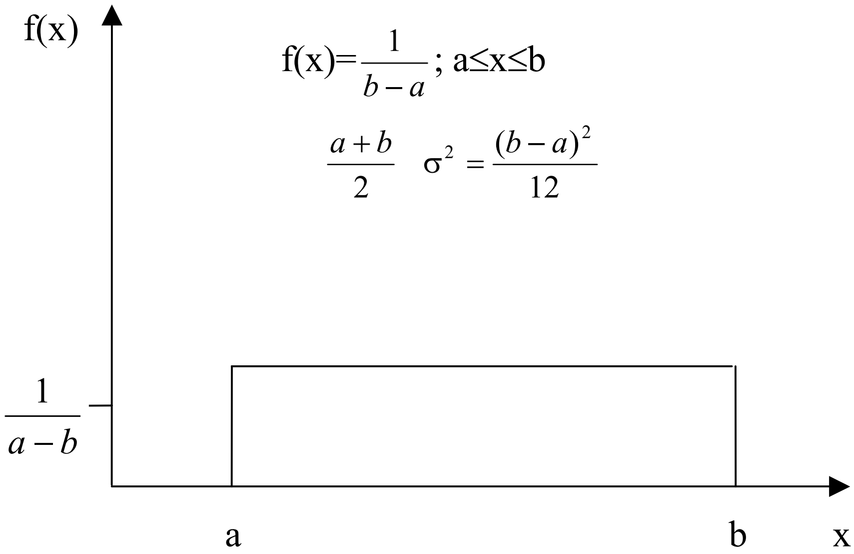


Рис. 2.7. Функция плотности вероятности и характеристики равномерного распределения.

На рис. 2.7 приводятся характеристики равномерного распределения и график его функции плотности вероятности.

2.9.2. Треугольное распределение

Треугольное распределение является более информативным, чем равномерное. Для этого распределения определяются три величины — минимум, максимум и мода. График функции плотности состоит из двух отрезков прямых, одна из которых возрастает при изменении X от минимального значения до моды, а другая убывает при изменении X от значения моды до максимума. Значение математического ожидания треугольного распределения равно одной трети суммы минимума, моды и максимума. Треугольное распределение используется тогда, когда известно наиболее вероятное значение на некотором интервале и предполагается кусочно-линейный характер функции плотности. На рис. 2,8 приведены характеристики треугольного распределения и график его функции плотности вероятности.

Треугольное распределение легко применять и интерпретировать, однако для его выбора необходимы веские основания.

$$f(x) = \begin{cases} \dfrac{2(x-a)}{(m-a)(b-a)}; a \le x \le m \\ \dfrac{2(b-x)}{(b-m)(b-a)}; m \le x \le b \end{cases}$$

$$\mu = \frac{a+m+b}{3}$$

$$\sigma^2 = \frac{a(a-m)+b(b-a)+m(m-b)}{18}$$

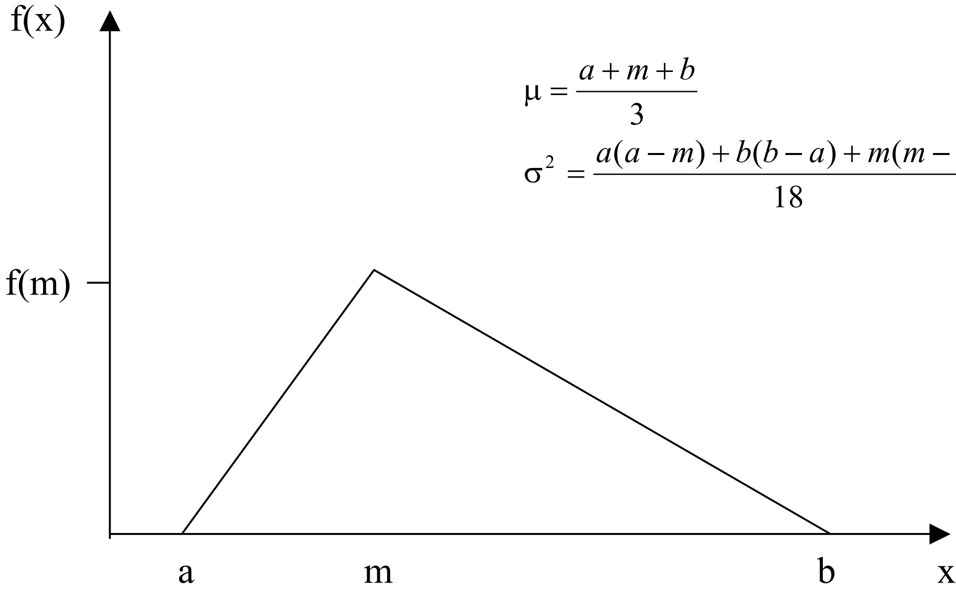


Рис. 2.8. Функция плотности вероятности и характеристики треугольного распределения.

2.9.3. Экспоненциальное распределение

Если вероятность того, что один и только один результат наступит на интервале Δt, пропорциональна Δt и если наступление результата не зависит от наступления других результатов, величины интервалов между результатами распределены экспоненциально. Другими словами, работа, продолжительность которой экспоненциально распределена, имеет одинаковую вероятность завершения в течение любого последующего периода времени Δt. Таким образом, работа, выполняемая за Δt единиц времени, имеет ту же вероятность окончания в последующий период Δt, что и только что начатая работа. Подобное отсутствие временной обусловленности называется марковским свойством или свойством отсутствия последействия. Существует прямая связь между предположением об экспоненциальности распределения продолжительности работы и марковским свойством. Экспоненциальное распределение предполагает значительную вариабельность переменной. Если математическое ожидание продолжительности работы равно μ, то дисперсия равна μ^2. По сравнению с большинством остальных распределений экспоненциальное обладает большей дисперсией. С экспоненциальным распределением легко осуществлять математические преобразование, благодаря чему оно применяется в целом ряде исследований. На рис. 2.9. представлены характеристики экспоненциального распределения и его функции плотности вероятности.

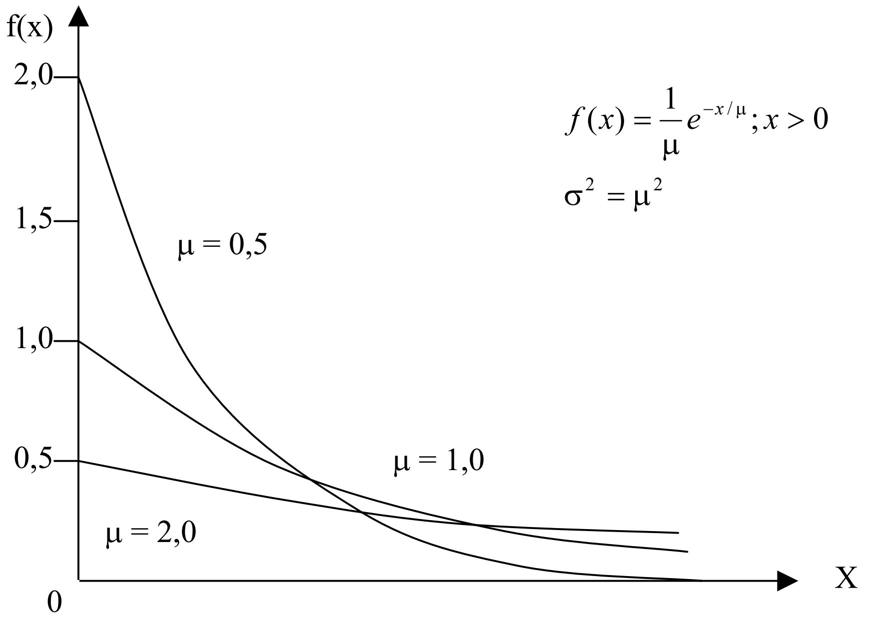


Рис. 2.9. Функция плотности вероятности и характеристики экспоненциального распределения

2.9.4. Распределение Пуассона

Распределение Пуассона является дискретным и обычно связано с числом результатов за определенный период времени. Если продолжительность интервалов времени между результатами распределена экспоненциально и в каждый момент времени может произойти только один результат, то можно доказать, что число результатов на фиксированном интервале времени распределено по закону Пуассона.

Другими словами, если интервалы между прибытиями распределены экспоненциально, то распределение числа прибытий будет пуассоновским. Пуасоновское распределение используется часто как аппроксимация биномиального распределения в том случае, когда оно моделирует последовательности независимых испытаний Бернулли (результаты таких испытаний могут быть типа «да - нет», «стоять - идти», «успех - неудача» и т.п.). При больших значениях математического ожидания пуассоновское распределение аппроксимируется нормальным. На рис. 2.10 приведена функция плотности вероятности распределения Пуассона при различных значениях μ.

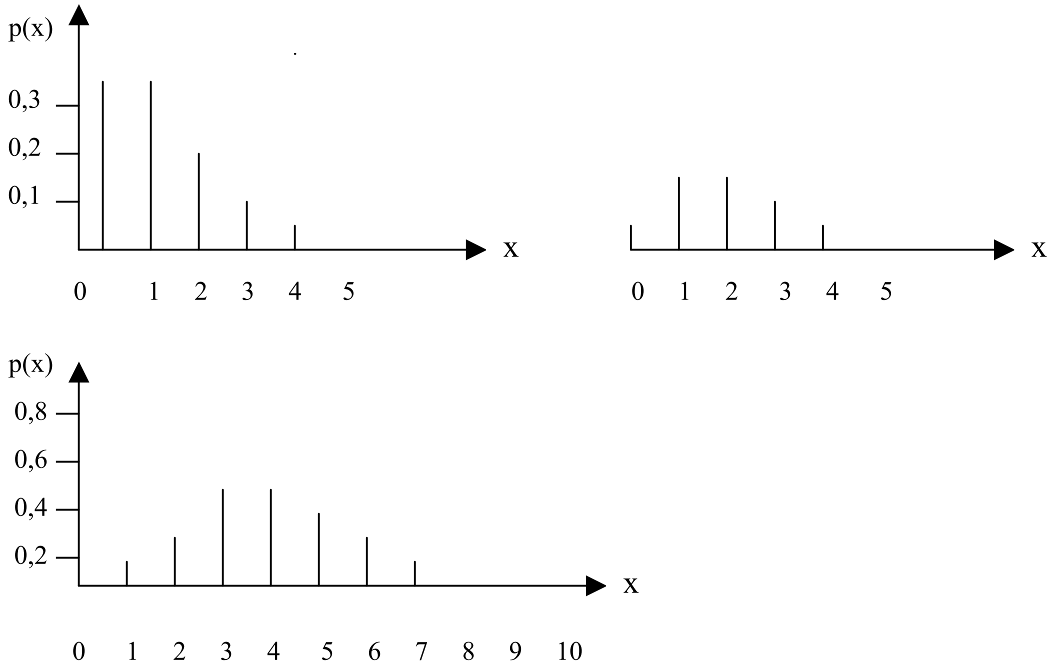


Рис. 2.10. Функция плотности вероятности и характеристики распределения Пуассона

2.9.5. Нормальное распределение

Нормальное или гауссово, распределение является наиболее важным в теории вероятностей и математической статистике. Эту роль нормальное распределение приобрело в связи с центральной предельной теоремой, которая, как это говорилось выше, утверждает, что при весьма нестрогих условиях распределение средней величины или суммы I независимых наблюдений из любого распределения стремится к нормальному по мере увеличения I. Таким образом, сумму случайных величин часто можно считать нормально распределенной.

Именно благодаря центральной предельной теореме нормальное распределение так часто применяется в исследованиях по теории вероятностей и математической статистике. Существует и другая причина частого применения нормального распределения. Его преимуществом является легкость математического трактование, в связи с чем многие методы доказательств в таких областях, как, например, регрессионный или вариационный анализ, основаны на предположении о нормальном характере функции плотности.

Как уже говорилось выше, при больших значениях среднего нормальное распределение является хорошей аппроксимацией для распределения Пуассона, которое в свою очередь является аппроксимацией биномиального распределения. На рис. 2.11 приведена функция плотности нормального распределения при различных значениях математического ожидания и среднеквадратичного отклонения.

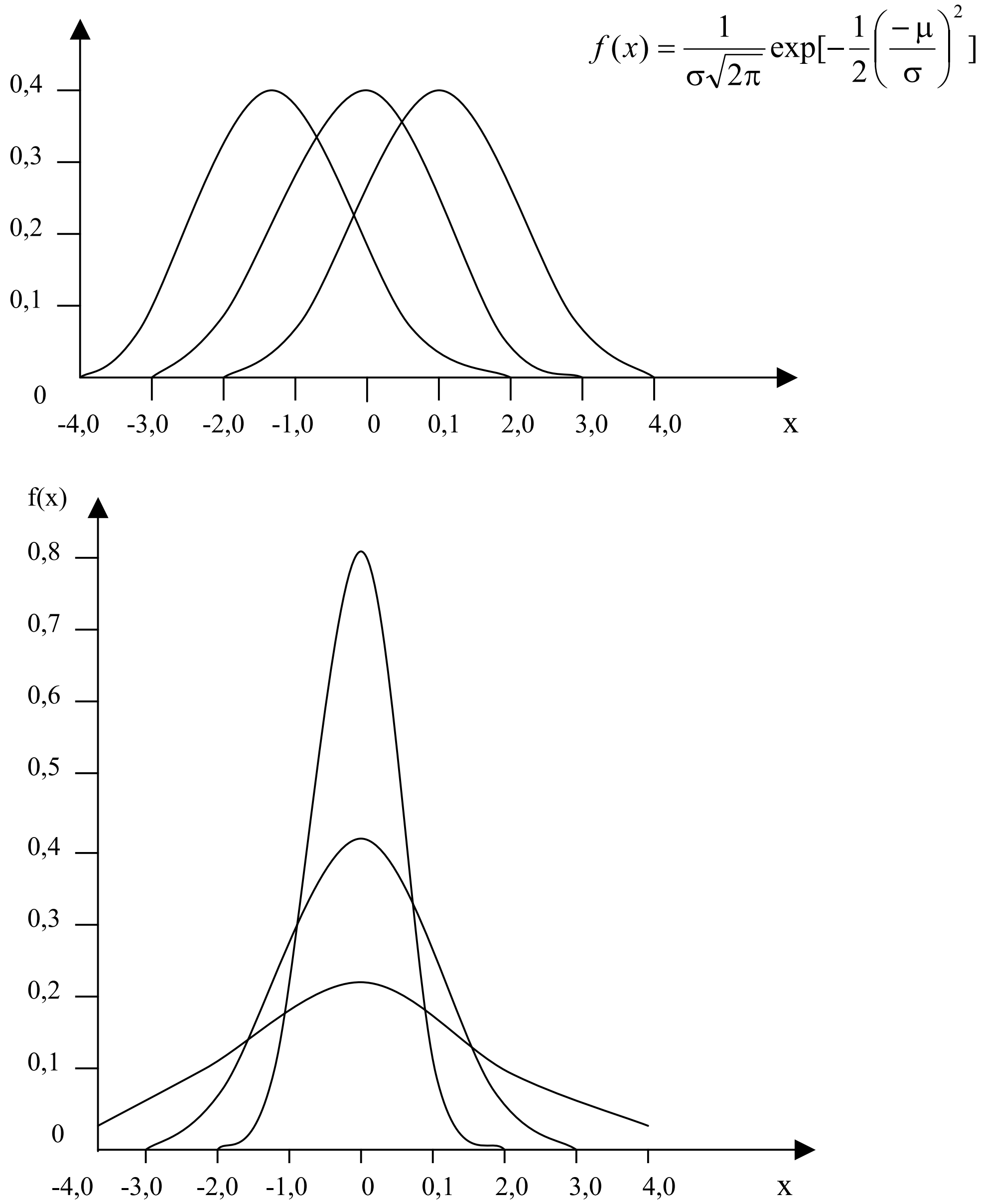


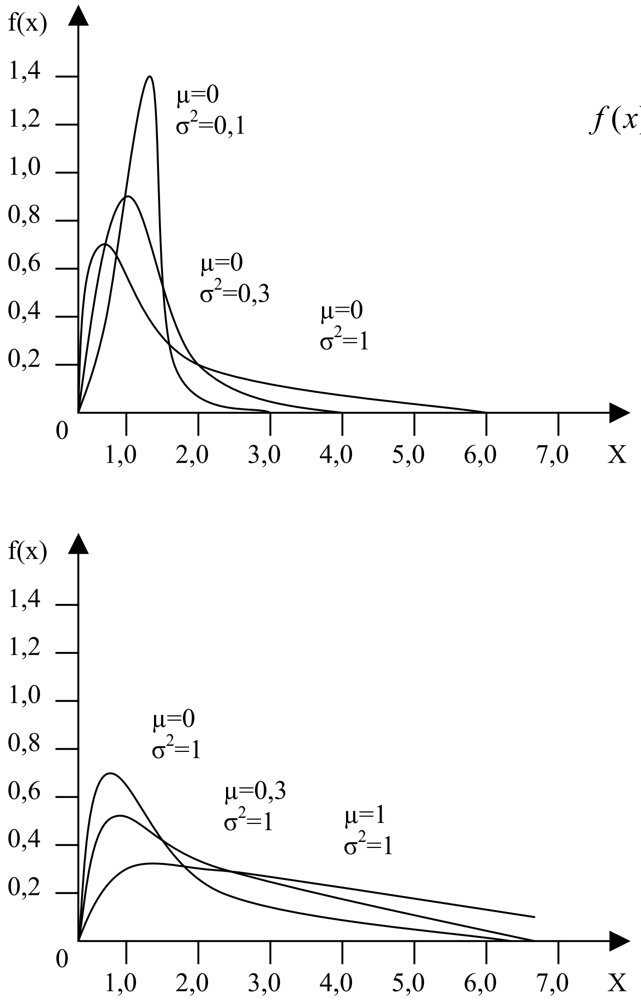
Рис 2.11. Функция плотности вероятности и характеристики нормального распределения.

2.9.6. Логарифмическое нормальное (логнормальное) распределение

Логнормальное распределение является таким распределением случайной величины, натуральный логарифм которой нормально распределен [11]. Это распределение пригодно для моделирования мультипликативных процессов так же, как нормальное - для аддитивных. С помощью центральной предельной теоремы можно показать, что распределение произведения независимых положительных случайных величин стремится к логнормальному.

Если после логарифмирования каждого элемента некоторого набора данных этот трансформированный набор данных нормально распределен, можно сказать, что исходные данные распределены по логнормальному закону. Логнормальное распределение широко используется для моделирования биологических и экономических систем. Оно хорошо моделирует процессы, в которых значение наблюдаемой переменной является случайной долей от значения предыдущего наблюдения. Примерами подобных процессов являются распределение личных доходов, наследства или банковских вкладов, распределение длины слов и т. п.

На рис. 2.12 приведены характеристики логнормального распределения, а также график функции его плотности для различных значений среднего и дисперсии.



$$f(x)=\frac{1}{\sigma x\sqrt{2\pi}}\quad \exp[-\frac{1}{2}(\frac{1nx-\mu}{\sigma})^2]$$

Рис.2.12. Функция плотности вероятности и характеристики логнормального распределения.
μ и σ приведены для соответствующего нормального распределения

2.9.7. Распределение Эрланга

Распределение Эрланга является результатом суммирования независимых и одинаково распределенных экспоненциальных случайных величин. Оно представляет собой частный случай гамма-распределения, поэтому все, что касается функции плотности, интерпретации и замечаний относительно гамма-распределения, справедливо также и для распределения

Эрланга. Это распределение широко используется в теории массового обслуживания, когда исследуется выполнение работ в течение экспоненциально распределенных промежутков времени.

2.9.8. Гамма-распределение

Гамма-распределение является обобщением распределения Эрланга для случая, когда число суммируемых экспоненциальных величин не является целым. Гамма-распределенная величина может принимать значения от 0 до бесконечности. Функция плотности гамма-распределения принимает различные формы при различных значениях параметров, что позволяет моделировать различные физические процессы.

Гамма-распределение можно интерпретировать также как сумму квадратов нормально распределенных случайных переменных, т. е. χ^2 как распределение. Таким образом, χ^2-распределение, распределение Эрланга и, следовательно, экспоненциальное распределение являются частными случаями гамма - распределения.

На рис. 2.13 приведены характеристики гамма-распределения, а также график его функции плотности для различных значений этих характеристик.

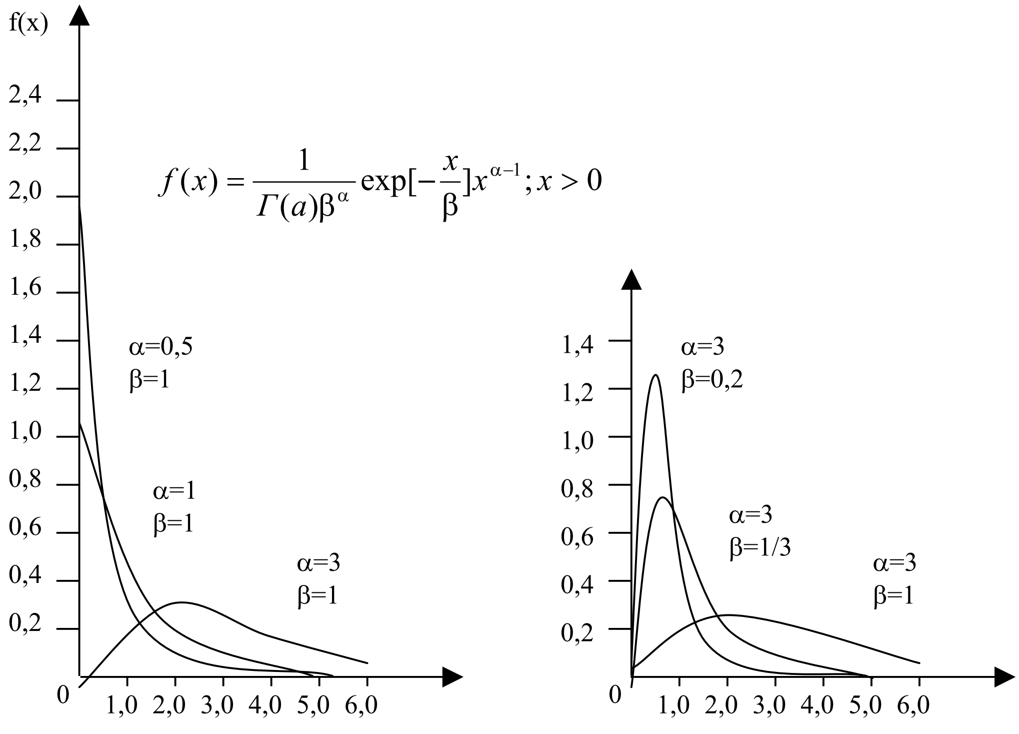
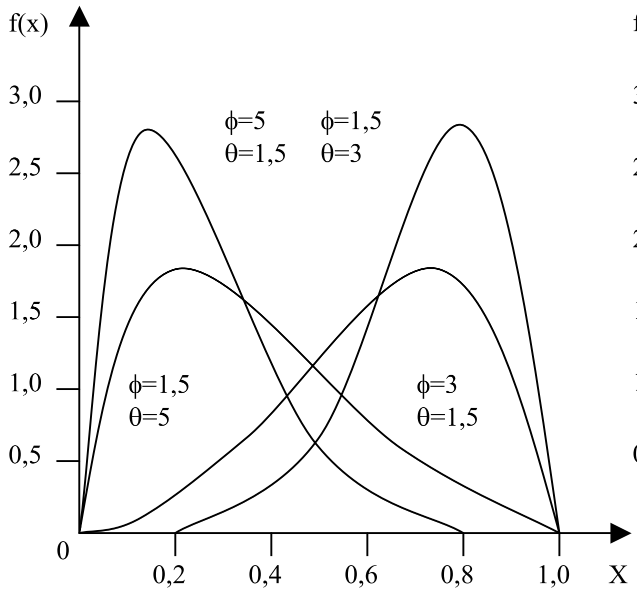
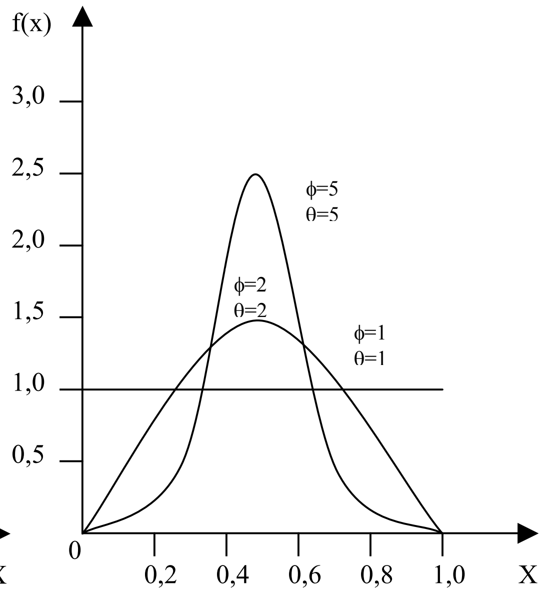


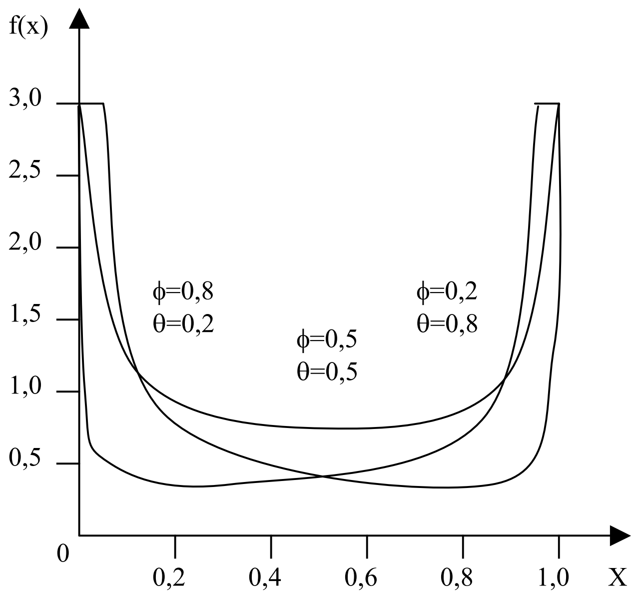
Рис. 2.13. Функция плотности вероятности и характеристики гамма-распределения

2.9.9. Бета-распределение

Бета-распределение определено на конечном интервале и при различных значениях параметров описывается разными кривыми. Эти кривые могут иметь либо форму «колокола» (симметричного или асимметричного), либо U-образную форму. Для U-образных бета-функций значение функции плотности стремится к бесконечности на концах интервала. Один из простейших случаев бета - распределения называется распределением Парето и используется обычно для моделирования распределения доходов. Благодаря тому, что бета-функция описывает множество кривых различного вида, она используется для моделирования множества различных данных[1].







[1] Здесь необходимо напомнить о предупреждении Феллера [8] относительно закона логистического роста: «единственной теоретической проблемой является то, что не только логистическое распределение, но и нормальное распре деление, распределение Коши и другие могут описывать один и тот же статистический материал, давая ничуть не худшие значения критерия согласия распределения. Весьма противоречивые теоретические модели могут опираться на одни и те же наблюдения».

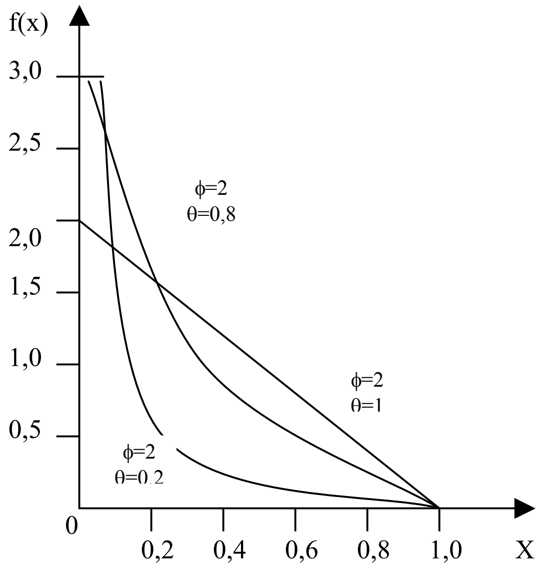


Рис. 2.14. Функция плотности вероятности в характеристики бета-распределения

Поскольку бета-распределение определено на конечном интервале, это обусловливает выбор объекта моделирования. Примерами могут служить функции плотности оценок вероятностей или долей чего-либо. Часто бета-распределение применяется в качестве априорного распределения параметра биномиального процесса в байесовских исследованиях. Кроме того, бета-распределение часто применяют в качестве описательной функции плотности продолжительности работы в сетевом моделировании. Субъективные оценки продолжительности работы, основанные на оптимистическом (a), пессимистическом (b) и наиболее вероятном (m) значениях, определяют значение оценки среднего в виде (a+4m+b)/6 и оценки дисперсии в виде

$$(b-a)^2/36.$$

На рис. 2.14 приведены функции плотности бета-распределения для различных значений его параметров.

2.10. Генерация псевдослучайных чисел

В имитационных моделях иногда необходимо получать случайные выборки из одного или нескольких распределений, описанных в предыдущем разделе. Наиболее часто применимым на практике методом получения выборок случайных чисел из заданного распределения на цифровом компьютере является генерация одного или нескольких случайных чисел, равномерно распределенных на интервале между 0 и 1, и последующее преобразование этого числа или чисел в новое случайное число, распределенное по желаемому закону. Независимые случайные числа, равномерно распределенные на интервале от 0 до 1, являются, таким образом, основой для генерации выборок всевозможных распределений. Мы остановимся ниже на проблеме получения случайных чисел, а затем обсудим процедуры их преобразования в случайные величины, распределенные по различным законам.

В цифровой имитации существует по крайней мере три способа получения случайных чисел. Первым методом является хранение в компьютере таблицы случайных чисел [23] и получение затем из нее данных для имитационного моделирования. Недостаток этого метода заключается в относительно медленной скорости считывания компьютером данных с внешнего устройства ввода и в необходимости хранения большого объема табличных данных. Вторым методом является использование некоторого физического устройства, например электронной лампы, для генерации случайного шума. Недостаток этого метода заключается в невозможности повторного воспроизведения результатов имитации, а следовательно, невозможности осуществления верификации модели и направленного эксперимента с ее параметрами. Третьим методом, которому отдается предпочтение, является применение рекурсивных формул, по которым на основании 1-го случайного числа вычисляется (i+1)-е случайное число. Поскольку последовательность чисел вычисляется в уравнении детерминировано, они, естественно, не являются случайными, и их обычно называют псевдослучайными числами.

В дальнейшем для краткости будем называть эти числа случайными, имея в виду, что на самом деле они псевдослучайны. Генераторы псевдослучайных[1] чисел должны удовлетворять следующим требованиям:

1. Числа равномерно распределены на интервале (0, 1) и независимы, т. е. корреляция между случайными числами последовательности отсутствует.

2. Генерируется достаточное количество неповторяющихся Чисел, т. е. период (цикл) генератора довольно длинный.

3. Последовательность случайных чисел воспроизводима. Это предполагает, что различные начальные значения (корни) дают различные последовательности.

4. Генератор должен быть быстродействующим, поскольку для моделирования может потребоваться большое число чисел.

5. Желательно использование малого объема памяти.

Наилучшим образом удовлетворяет данным требованиям широко распространенный в настоящее время конгруэнтный метод.

[1] Из-за того, что псевдослучайные числа генерируются на цифровых компьютерах детерминированными методами, вокруг их определения ведется большая дискуссия. С точки зрения нашего подхода к имитационному моделированию наиболее приемлемым является определение Левиса [16]: «некоторый хитроумный способ записи, скрывающий принцип формирования последовательности, при котором значение каждого последующего члена непредсказуемо, а сами члены удовлетворяют некоторым тестам в зависимости от того, для таких целей формируется данная последовательность».

Конгруэнтный метод использует следующее рекурсивное уравнение:

$$Z_{i+1}=(aZ_i+b)\ (\mathrm{mod}\ c),\ i=0,\ 1,\ 2,...,$$
$$R_{i+1} = Z_{i+1}/c$$

где z_0— значение корня, а r_i является 1-м псевдослучайным числом. Это уравнение определяет, что ненормализованное случайное число 21+1 равно остатку (az_i+b), деленному на с, где z_i— предыдущее ненормализованное случайное число, z_0 — начальное значение (корень), а a, b и с — константы. Выбор значений констант a, b и с является предметом постоянных исследований. В работе [9] представлен исчерпывающий обзор методов определения констант и процедур тестирования генераторов случайных чисел. В приложении (стр. 66) кратко излагаются правила определения констант a, b и с для конгруэнтных генераторов. Эти правила дают лишь общие рекомендации для выбора a, b и с, выбор же наилучших значений непосредственно зависит от используемого компьютера. Мы рекомендуем разработчикам моделей использовать генераторы случайных чисел, разработанные специально для компьютера, на котором затем реализуется имитационная модель.

В процессе имитации часто необходимо в одной модели работать с несколькими потоками случайных чисел. Например, раздельные потоки случайных чисел могут быть использованы в системе массового обслуживания для моделирования процессов прибытия и обслуживания заявок. При этом можно генерировать одни и те же последовательности моментов прибытия заявок независимо от порядка их обслуживания и, таким образом, оценивать различные процедуры обслуживания для одной и той же последовательности заявок. Разработчику модели предоставляется возможность выбора различных значений корней генератора случайных чисел для параллельных случайных потоков.

2.10.1. Метод обратной функции

Случайные числа являются основой для получения величин, распределенных по заданным законам. Простейшим и наиболее фундаментальным методом, на основе которого генерируются подобные величины, является метод обратной функции [1, 9]. В основе этого метода лежит тот факт, что случайная величина R= F(X) равномерно распределена на интервале [0, 1]. Другими словами, для генерации случайной величины из распределения X генерируется случайное число r и решается уравнение r = F(x) относительно значения $x = P^{-1}(r)$. Доказательство состоятельности метода очевидно [10, 22] и основано на следующих соображениях. Пусть R = F(X) имеет функцию распределения G( ). Тогда для 0≤r≤1 имеем

$$G(r) = P[F(X)\le r] = P[X\le F^{-1}(r)] = F[F^{-1}(r)] = r$$

Следовательно, R равномерно распределена на интервале [0,1].

Для иллюстрации метода на примере непрерывного распределения рассмотрим генерацию экспоненциально распределенной случайной величины. Функция экспоненциального распределения имеет вид $F(x) = 1 — e^{\lambda x}$, где $1/\lambda$ — математическое ожидание. Приравнивая $P(x)=г$ и решая уравнение относительно x, получаем

$$x = -(1/\lambda)\ln(r)$$

Если г равномерно распределена на интервале [0, 1], из данного уравнения вытекает, что x распределена экспоненциально с ожиданием, равным $1/\lambda$.

Данный метод применим также и для дискретных распределений. Рассмотрим, например, следующую функцию вероятности:

$$p(0) = 0{,}25;\ p(1) = 0{,}50;\ p(2) = 0{,}25.$$

Кумулятивная функция распределения $P(x)$ показана на рис. 2.15. Для получения случайной величины из этого распределения необходимо на интервале от 0 до 1 сгенерировать случайное число и нанести его на ось ординат на графике функции распределения.

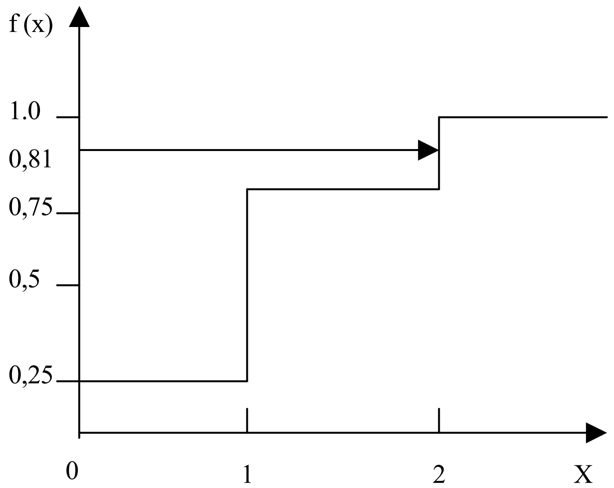


Рис. 2.15. Иллюстрация метода обратной функции для получения выборки из заданного распределения.

Проведя горизонтальную прямую из этой точки до линии графика функции P(x), а затем проведя вертикальную прямую из этой точки до оси абсцисс, получаем результирующую случайную величину. Например, случайное число 0,81 даст случайную величину, равную 2. Очевидно, что при такой процедуре 25% случайных чисел попадет на интервал (0; 0,25), 50%— на интервал (0,25; 0,75) и 25%— на интервал (0,75; 1,00) в соответствии с заданной функцией распределения. При этом необходимо принять некоторое правило определения результирующей величины для случайных чисел, попадающих в точки разрыва функции.

Затруднения при использовании метода обратной функций происходят обычно при поиске обратного преобразования $F^{-1}(r)$. В ряде случаев метод приводит к простым преобразованиям, подобным тем, что были проделаны для экспоненциального распределения. Тем не менее для ряда непрерывных распределений представление обратной функции в явном виде отсутствует. К счастью, для всех основных распределений, не имеющих явного представления обратной функции, разработаны специальные методы генерации. Применяемые в СЛАМ методы получения случайных величин описаны в приложении 7.

Статистическая состоятельность результатов имитационного моделирования зависит от степени «случайности» применяемого генератора случайных чисел. В связи с этим разработан целый ряд статистических процедур тестирования генераторов случайных чисел. Однако, как отмечено в работе [15], не существует некоторого набора тестов, гарантирующего пригодность конечной случайной последовательности вообще. Для данного набора тестов всегда существует удовлетворяющая ему последовательность случайных чисел, но полностью не пригодная для целого ряда частных случаев. Эта оговорка не порождает серьезных проблем, поскольку исследователь, как правило, не нуждается в дополнительных свойствах случайности, выходящих за рамки описанных выше.

Для исследования свойств случайности генераторов применяются как аналитические, так и эмпирические тесты. К ним относятся тесты: частотный, сериальный, интервальный, на сумму цифр, циклический и др. Эмпирические результаты по использованию тестов приведены в работах [9, 16]. В работе [6] описаны спектральные, а в работах [5, 17]—решеточные процедуры оценки качества конгруэнтных генераторов случайных чисел относительно их отклонения от идеальных свойств случайности.

2.12. Сбор и анализ данных

Важной функцией имитационного моделирования является сбор и анализ данных. Выполнение этой функции необходимо как при определении входных данных для модели, так и при получении результатов эксперимента. Ниже дается обзор ряда важных понятий математической статистики, применяемых при сборе и анализе данных.

2.12.1. Подготовка данных

Под подготовкой данных подразумевается процесс получения (сбора) данных об изучаемом явлении. Существует ряд методов получения исходных данных. В некоторых случаях исходные данные содержатся в существующей документации, и тогда задачей исследователя является выявление требуемых данных и организация доступа к ним. В других случаях подготовка данных может включать анкетирование, обзор проблематики или же физическое экспериментирование.

В крупномасштабных моделях, таких, как модели городских шли экономических систем, требуемые данные обычно можно получить из существующей документации. Источниками данных для таких моделей служат, например, официальные отчеты, статистические сборники, а также материалы правительственных и международных организаций. Все чаще подобные материалы не только появляются в виде бумажных документов, но записываются и на машинные носители информации (например, на магнитную ленту или гибкий диск).

Для моделей производственных систем важным источником данных может служить финансовая и техническая документация. Хотя эта документация часто недостаточна при формировании целостной основы для оценки спроса, стоимости продукции и других важных факторов, она тем не менее служит отправной точкой при моделировании. Анкетирование и обзор предметной области также являются одними из возможных методов получения данных при моделировании производственной деятельности.

Физическое экспериментирование, как правило, является наиболее дорогостоящим и трудоемким методом получения исходных данных. Оно включает в себя измерение, запись и обработку данных. Особое внимание здесь следует уделять планированию эксперимента, которое позволяет убедиться в представительности условий эксперимента и правильности записываемых данных. Проблема планирования эксперимента при сборе данных подробно рассмотрена в работе [3].

В ряде случаев исходных данных может не существовать, при этом сама природа моделируемой системы часто исключает возможность экспериментирования. Примером такой ситуации может служить имитационное моделирование различных вариантов размещения оборудования сборочной линии. Возможным подходом к подготовке данных в таких случаях может быть предварительный синтез данных [2, 18], который предполагает вычисление оценок продолжительности работ с использованием таблиц стандартных исходных данных. Таким образом, этот метод позволяет оценивать продолжительность выполнения работ еще до того, как они будут выполняться в действительности.

2.12.2. Описание статистических данных

Как при сборе реальных данных для определения входов модели, так и при сборе данных о функционировании системы на основе моделирования мы сталкиваемся с проблемой преобразования «сырых» данных к удобному для анализа виду. Поэтому нас интересуют способы нахождения и описания наиболее важных свойств набора данных. Эти способы обычно позволяют агрегировать данные за счет потери некоторой содержащейся в них информации.

Группировка данных**.** Одним из методов преобразования данных к удобному виду является группировка данных по классам (интервалам). Данные затем сводятся в таблицу, содержащую количество попаданий элементов данных в каждый класс. Такая таблица называется частотной и обычно дает хорошее «общее представление о данных. Примером частотной таблицы могут служить представленные ниже данные о временах ожидания обслуживания посетителями:

Время ожидания	Число посетителей
0→20	21
20→40	35
40→60	42
60→80	35
80→100	19
100→120	10
>120	10

Числа в правом столбце таблицы указывают количество, попавших в каждый класс посетителей, и называются частотами классов. Числа в левом столбце для каждого класса определяют интервалы значений наблюдаемой величины и называются границами класса. Разница между верхней и нижней границами класса называется размером класса. Классы, не ограниченные сверху или снизу, называются открытыми. Классы, ограниченные с двух сторон, называются закрытыми. Довольно часто первый и (или) последний классы частотной таблицы бывают открытыми.

Существует несколько видов частотных таблиц, полезных для отображения сгруппированных данных. Одним из видов является таблица накопленных частот, получаемая последовательным сложением значений частот. Ниже приводится таблица накопленных частот для данных о времени ожидания обслуживания посетителями:

Время ожидания (с) меньше, чем	Число посетителей (с накоплением)
20	21
40	56
60	98
80	133
100	152

120	162
∞	172

Числа в правом столбце указывают общее число посетителей, время ожидания у которых оказалось меньше, чем указанная в левом столбце верхняя граница класса. Еще один вид таблиц можно получить, преобразовав частотную таблицу (или таблицу накопленных частот) в таблицу распределения частот путем деления частоты каждого класса (накопленной частоты) на общее число имеющихся в нем элементов данных.

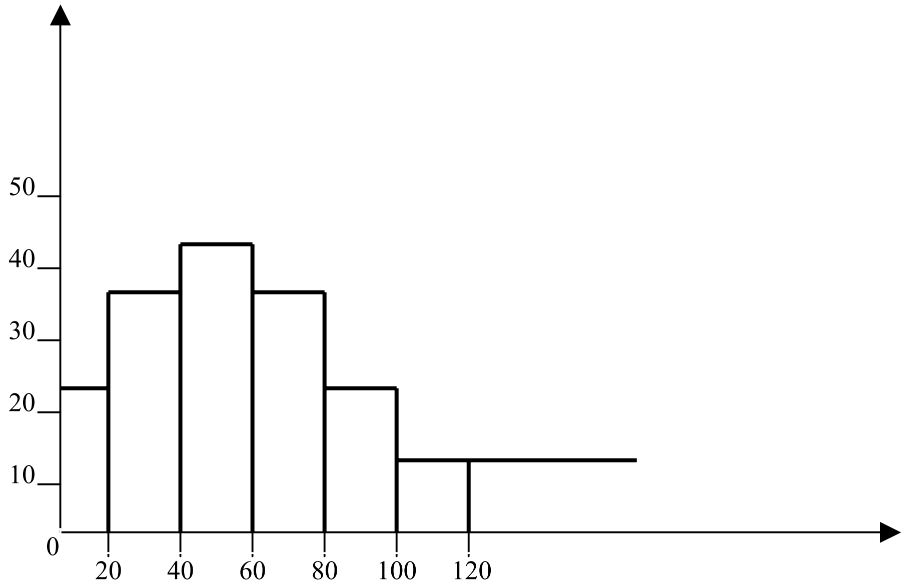


Рис. 2.16 Гистограмма распределение времени ожидания обслуживания посетителями.

Распределение деления частот полезны, в частности, при сравнении нескольких законов распределений.

С целью улучшения отображения данных частоты и накопленные частоты иногда представляются графически. Наиболее общим графическим представлением является гистограмма, которая отображает в виде прямоугольников частоты классов, причем высота прямоугольников пропорциональна частотам. На рис. 2.16 приведена гистограмма времени ожидания обслуживания посетителями.

Особое внимание при построении распределений частот необходимо уделять выбору числа классов и границ интервалов. При этом необходимо, учитывая, конечно, природу данных и цель их использования, следовать некоторым общим рекомендациям:

1. По возможности классы должны иметь одинаковый размер, хотя первый и последний классы при этом могут быть открытыми.

2. Интервалы классов не должны перекрываться. Другими словами, каждый элемент данных должен принадлежать одному и только одному классу.

3. Не следует обычно делать меньше 5 и больше 20 классов.

Оценка параметров. Если множество элементов данных содержит множество всех возможных наблюдений, оно называется популяцией. Если же содержит только часть из них, оно называется выборкой. Одним из методов агрегации множества данных является рассмотрение данных как выборки, используемой для оценки параметров исходной популяции. Наиболее интересными параметрами популяции являются среднее, оценивающее меру центральности, и дисперсия, оценивающая меру рассеивания.

Для примера рассмотрим снова данные о времени ожидания обслуживания посетителями. Эти данные можно рассматривать как выборку из популяции, состоящей из всех возможных значений времени ожидания. Мы можем далее по данным: этой выборки оценить среднее время ожидания посетителя и дисперсию времени ожидания для всей популяции.

Для отличия параметров популяции от оценок этих параметров на основе выборки они обозначаются по-разному. Для обозначения среднего и дисперсии популяции используются обычно греческие символы μ и σ^2 соответственно. Оценки же этих параметров на основе выборки x_1, x_2, …, $\bar{x}_i$ обозначаются символами $\bar{x}_i$ и s^2x соответственно. Для разграничения в дальнейшем этих понятий характеристики популяции будем называть параметрами, а характеристики выборки – статистиками.

Прежде чем продолжить обсуждение статистических характеристик, сделаем ряд уточнений относительно обозначений, принятых для случайных величин, экспериментальных оценок случайной величины и случайной последовательности. Случайная величина до осуществления наблюдения обозначается через $\bar{X}_I$,а после его осуществления – символом x_i. Выборочным средним $\bar{X}_I$ будем называть случайную величину, являющуюся суммой I случайных величин до осуществления наблюдений, деленную на I. После осуществления наблюдений x_i, среднее будем обозначать $\bar{x}_i$. Аналогично случайную величину, являющуюся оценкой дисперсии выборки до осуществления наблюдения, будем обозначать символом S^2x, а после осуществления – символом s^2x. Таким образом, как было условленно выше, случайные величины обозначаются прописными буквами, а их количественные оценки – строчными.

При построении оценок параметров популяции по данным выборки необходимо рассматривать два различных случая. В первом случае выборка содержит только значения самих наблюдений без учета моментов времени осуществления этих наблюдений. Примером такой выборки могут служить данные о времени ожидания обслуживания посетителями. Статистики по независимой от времени выборке называются статистиками по наблюдениям или точечными статистиками.

Таблица 2.1. Формулы для вычисления значений среднего и дисперсии по выборке

Статистика	Формула	
	Точечные статистические оценки	Интервальные статистические оценки
Выборочное среднее	$\overline{x_i} = \frac{\sum_{i=1}^{I} x_i}{I}$	$\overline{x_T} = \frac{\int_{i=1}^{I} x(t)dt}{T}$
Дисперсия выборки	$s^2x = \frac{\sum_{i=1}^{I} x_i^2 - I\overline{x^2}I}{I-1}$	$s^2x = \frac{\int_0^T x^2(t)dt}{T} - \overline{x^2}T$

Во втором случае значения случайных величин определены во времени. Например, число занятых кассиров в банке является во времени. Например, число занятых кассиров в банке является случайной величиной, значение которой меняется во времени. При этом нас интересует информация о том, какие значения принимала наблюдаемая случайная величина и на каких интервалах времени. Статистики по зависимой от времени выборке называются временными или интервальными статистиками.

В табл. 2.1 приведены формулы для вычисления как точечных, так и интервальных статистик $\overline{x}_i$ и s^2x. Для интервального случая выборочное среднее обозначается x_T, где T равно общей продолжительности интервала времени наблюдения. Для вычисления s^2x существует несколько формул, однако приводимая здесь формула наиболее удобна с вычислительной точки зрения. Отметим, что для вычисления точечных статистик необходимо знать значения и размер выборки I. Аналогично для вычисления интервальных оценок необходимо знать $\int_0^T xdt$, $\int_0^T x^2dt$ и T.

Еще одной часто используемой при обработке данных оценкой является коэффициент вариации, равный $S_x/\overline{X}_I$. Он определяет отношение среднеквадратичного отклонения выборки к выборочному среднему. Коэффициент вариации применяется обычно для сравнения дисперсией нескольких наборов данных.

2.12.3. Подбор распределения

Предыдущий раздел был посвящен проблеме оценки параметров популяции по выборке. Аналогичной, но более сложной проблемой является идентификация распределения популяции по данным выборки. Она часто возникает в моделировании, так как вероятностные элементы необходимо представлять в модели в виде конкретных распределений. Хотя понимание свойств теоретических распределений, описанных в разд. 2.9, помогает разработчику модели выдвинуть гипотезу о выборе подходящего распределения, все же необходимо проверить ее с помощью одного из статистических тестов. Наиболее пригодными из известных тестов являются тест х-квадрат и тест Колмогорова – Смирнова, описание и примеры использования которых можно найти в большинстве книг по математической статистике [21]. Существует удобное для использования программное обеспечение (AID), реализующее как графические, так и статистические тесты согласия для подбора теоретического распределения по данным выборке [23].

2.13. Статистический вывод

В имитационных исследованиях часто необходимо на основе результатов имитационных экспериментов делать некоторые выводы или прогнозы относительно поведения моделируемой системы. Поскольку имитационная модель содержит вероятностные элементы, результаты имитации являются наблюдениями случайных величин. Как следствие, любая интерпретация поведения системы на основе анализа полученных результатов имитации должна учитывать их разброс. Подобный учет осуществляется с помощью вычисления доверительных интервалов или проверки гипотез.

2.13.1 Доверительные интервалы

В разд. 2.12.2 обсуждались методы оценки параметров среднего и дисперсии популяции на основе данных выборотки. Оценками параметров были некоторые числа, называемые точечными оценками. В общем случае вследствие случайного разброса оценка отличается от действительного, но известного значения параметра. Недостатком точечной оценки является то, что она не позволяет лицу, принимающему решение, судить о ее точности. Мерой точности является доверительный интервал, определение которого носит вероятностный характер. Доверительный интервал характеризует вероятность попадания значения оцениваемого параметра в заданный интервал.

Наибольший интерес для имитационного анализа представляет среднее значение популяции. Классическое определение доверительного интервала для среднего значения подразумевает независимость и одинаковую распредленность наблюдений. Следовательно, в соответствии с центральной теоремой выборочное среднее $\overline{X}_{\mathrm{I}}$ распределено приблизительно нормально при достаточно больших I. Как уже говорилось выше, предположение о независимости не является необходимым условием центральной предельной теоремы.

Если предположить, что $\overline{X}_I$ нормально распределено, то статистика

$$Z = \frac{\overline{X_i - \mu}}{\sigma_{\bar{x}}}$$

является нормально распределенной случайной величиной со средним значением, равным нулю, и средне квадратическому отклонению, равным единице. Кроме того,

$$P[-Z_{\alpha/2} < Z < Z_{\alpha/2}] = 1-\alpha,$$

где $Z_{\alpha/2}$ – такое значение Z при котором площадь под кривой функции плотности вероятности нормального распределения равна $\alpha/2$. Следовательно, можно с вероятностью $1-\alpha$ утверждать, что

$$\overline{X}_I - Z_{\alpha/2\sigma x} < \mu < \overline{X}_I + Z_{\alpha/2\sigma x}$$

В данной формуле предполагается, что величина средне квадратического отклонения среднего σ_x известна, хотя это верно далеко не всегда. Если в качестве оценки σ_x принимается выборочное среднеквадратическое отклонение среднего S_x, то, поскольку статистика

$$t = \frac{\overline{X_I - \mu}}{S_{\overline{X}}}$$

является случайной величиной с t-распределением Стьюдента с I-1 степенями свободы, можно вывести следующие аналогичное соотношение, позволяющее определить доверительный интервал $1-\alpha$ для μ с оценкой S_x :

$$\overline{X}_I - t_{\alpha/2,\, I-1}\ S\,\overline{X} < \mu < \overline{X}_I + t_{\alpha/2,\, I-1}\ S\,\overline{X}\,,$$

где $t_{\alpha/2,\, I-1}$ - критическое значение t – статистика с I-1 степенями свободы.

Если наблюдения $\overline{X}_I$ независимых и одинаково распределены, то выражения (2.1) и (2.2) для доверительных интервалов преобразуются путем замены

$$\sigma\overline{X} = \sigma\overline{X}/\sqrt{I}$$

$$S\overline{X} = S\overline{X}/\sqrt{I}$$

Соответственно. Подобная замена позволяет вычислить доверительный интервал по данным выборки. Однако это простое соотношение между дисперсией выборки и дисперсией среднего выборки справедливо только в случае независимости наблюдений.

Методы определения 5х для выражения (2.2) в случае автокоррелированных наблюдений описаны в разд. 13.3. Наиболее очевидным подходом является организация эксперимента для получения независимых наблюдений, что достигается путем; повтора имитационных прогонов или группировки данных.

2.14. Проверка гипотез

В ряде имитационных исследований необходимо принять решение о том, справедливо или ложно определенное утверждение относительно некоторого параметра. Например, надо решить, снижает ли изменение правила распределения работ в, цехе среднее время запаздывания обрабатываемых заданий, В связи с экспериментальной природой имитации мы должны учитывать случайную вариацию оценок сравниваемых параметров. Это осуществляется с помощью проверки гипотез.

В общем виде процедура проверки гипотез требует определения . нулевой гипотезы (обозначается H_0) и альтернативной гипотезы (обозначается H_1). Нулевая гипотеза задается обычно с целью определения того, может ли она быть отвергнута или нет. Например, если мы хотим установить, что правило распределения работ А снижает среднее время запаздывания по отношению к правилу распределения В, нам необходимо определить нулевую и альтернативную гипотезу следующим образом:

H_0: среднее время ожидания при правиле А равно среднему времени ожидания при правиле В.

H_1 среднее время ожидания при правиле А меньше среднего времени ожидания при правиле В.

Затем можно использовать результаты имитационного эксперимента при правилах А и В и попытаться отвергнуть гипотезу H_0 в пользу гипотезы H_1

Проверка нулевой гипотезы относительно альтернативной подразумевает выбор правила решения, основанного на данных выборки и приводящего к принятию или отказу от нулевой гипотезы. Принятие нулевой гипотезы означает не то, что она справедлива, а то, что на основании данных выборки нельзя сделать уверенного заключения об отказе от нее.

При использовании данного правила решения можно сделать ошибки двух типов. Ошибка первого типа заключается в отказе от нулевой гипотезы, в то время как она верна. Ошибка второго типа состоит в принятии нулевой гипотезы, в то время как она неверна. Принятие решения определяется значением (вероятностей, связанных с ошибками первого и второго типа. Эти вероятности обозначаются обычно как αи β-вероятности соответственно. Вероятность ошибки типа а называют уровнем значимости теста.

Критерий принятия решения формируется с помощью построения тестовой статистики, имеющей известное распределение. Тестовая статистика вычисляется по данным выборки и проверяется по правилу исключения. Если значение тестовой статистики попадает в тестовую область, нулевая гипотеза отвергается.

Тестовая статистика и правило исключения при проверке гипотез «работают» с характеристиками, приведенными в табл. 2.2. Тесты 1 и 2 проверяют, равно ли значение среднего данному значению μ_0. Тесты 3 и 4 сравнивают значения двух средних. Поскольку нельзя сформулировать предположение о независимости наблюдений, уравнения для тестовых статистик даны в терминах σ_x и s_x.

2.15. Статистические проблемы имитационного моделирования

Анализ решений, принимаемых на основе анализа результатов имитационного моделирования, требует обычно получения оценок усредненного отклика имитационной модели и его дисперсии. Обе эти оценки зависят от условий эксперимента. К условиям эксперимента, которые должен определить разработчик модели, относятся начальное, или исходное, состояние имитируемой системы, момент начала сбора статистических данных, продолжительность прогона модели и число повторных прогонов. Ниже кратко описаны проблемы, связанные с заданием этих условий. Они подробно обсуждаются в гл. 13.

2.15.1. Начальное состояние имитируемой системы

Любая имитационная модель в неявном виде подразумевает наличие начальных условий или исходного состояния имитации. Простейшим и наиболее общепринятым начальным состоянием является состояние «пуст и свободен», при котором имитация начинается в условиях отсутствия в модели компонентов и свободного состояния всех обслуживающих устройств. Приемлемость подобного начального условия зависит от природы моделируемой системы, а также от того, переходный или установившийся режим функционирования) системы нас интересует.

Если целью исследования является анализ установившегося режима работы системы, обычно всегда есть возможность улучшить качество статистических оценок с помощью выбора начального состояния, отличного от состояния «пуст и свободен». Начальные условия могут быть заданы на основе определения наиболее характерного для установившегося режима состояния системы, полученного в результате пробного имитационного прогона. Если же целью является анализ переходного режима, начальные условия должны отражать исходное моделируемое состояние системы.

Таблица 2.2 Тексты гипотез для средних значений

Нулевая гипотеза	Условие	Статистика текста	Распределение статистики текста	Число степеней свободы	Альтернативная гипотеза	Правило отказа от нулевой гипотезы
1. $\mu = \mu_0$	$\sigma_{\overline{X}}$ - неизвестно	$Z=\dfrac{\overline{X}-\mu}{\sigma_{\overline{X}}}$	Стандартное нормальное	----	$\mu>\mu_0$ $\mu<\mu_0$ $\mu\neq\mu_0$	$Z>Z\alpha$ $Z<-Z\alpha$ $\lvert Z\rvert>Z\alpha/2$
2. $\mu = \mu_0$	$\sigma_{\overline{X}}$ - неизвестно	$t=\dfrac{\overline{X}-\mu}{S_{\overline{X}}}$	Т – стьюдента	I-1	$\mu>\mu_0$ $\mu<\mu_0$ $\mu\neq\mu_0$	$t>t\alpha$ $t<-t\alpha$ $\lvert t\rvert>t\alpha/2$
3. $\mu_x = \mu_y$	$\sigma_{\overline{X}}$ и $\sigma_{\overline{Y}}$ неизвестно	$Z=\dfrac{X-Y}{\sqrt{\sigma^2_X+\sigma^2_Y}}$	Стандартное нормальное	----	$\mu_X>\mu_Y$ $\mu_X<\mu_Y$ $\mu_X\neq\mu_Y$	$Z>Z\alpha$ $Z<-Z\alpha$ $\lvert Z\rvert>Z\alpha/2$
$\mu_x = \mu_y$	$\sigma_{\overline{X}}$ и $\sigma_{\overline{Y}}$ неизвестно	$Z=\dfrac{X-Y}{\sqrt{S^2_{\overline{X}}+S^3_{\overline{Y}}}}$	t - стьюдента	Ближайшее целое к $\dfrac{(S^2_{\overline{X}}+S^2_{\overline{Y}})^2}{\dfrac{S^4_{\overline{X}}}{I_X+1}+\dfrac{S^4_{\overline{Y}}}{I_Y+1}}$	$\mu_X>\mu_Y$ $\mu_X<\mu_Y$ $\mu_X\neq\mu_Y$	$t>t\alpha$ $t<-t\alpha$ $\lvert t\rvert>t\alpha/2$

Индекс I в X и Y опущен для упрощения. I – рамер выборки; I_X – размер выборотки для X; I_Y - размер выборки для Y; μ_0 - гипотетическое среднее; α - уровень значимости.

2.15.2. Момент начала сбора статистических данных

Уменьшение смещения оценок характеристик в установившемся режиме, вызванного воздействием начальных условий, часто осуществляется с помощью метода, при котором начало сбора статистических данных задерживается до момента завершения периода «разогрева». Обычно для этой цели определяется момент отсечения, указывающий что собранные до него данные не учитываются при вычислении статистических оценок. Желаемое снижение влияния начальных условий достигается, следовательно, путем уменьшения числа собранных наблюдений в течение переходного периода имитации. Однако отсечение части данных может привести к увеличению значения оценки дисперсии среднего. Таким образом, улучшение качества оценки среднего достигается ценой увеличения разброса результатов имитационного моделирования.

Наиболее часто точка отсечения определяется по графику отклика, полученному в результате пробного прогона имитационной модели. Момент времени отсечения выбирается так, что значение отклика по графику можно считать установившимся. Существует ряд методов формализации данной процедуры в виде правила, которое включается в программу имитационной модели и автоматически определяет точку отсечения в ходе имитации. Эти правила описываются в гл. 13.

2.15.3. Продолжительность имитационного прогона и число повторных прогонов

Важным моментом планирования имитационного эксперимента является определение соотношения между продолжительностью прогона модели и числом повторных прогонов. Использование нескольких продолжительных прогонов предпочтительнее, чем использование множества коротких, так как в общем случае это дает лучшую оценку для среднего в условиях, установившегося режима, поскольку меньшее число раз вносятся искажения переходного режима и отсекается меньше данных. Однако уменьшение числа наблюдений в связи с уменьшением числа повторных прогонов, с одной стороны, может увеличить оценку дисперсии среднего. Большое число коротких прогонов, с другой стороны, может внести искажения, вызываемые начальными условиями. Чем больше переходный период, тем важнее использовать более продолжительные прогоны с целью снижения влияния начальных условий.

Существует ряд методов задания продолжительности имитационного прогона. Наиболее часто, по-видимому, задается момент времени завершения моделирования. Недостатком этого метода является то, что число наблюдений, будучи случайным, может быть различным в каждом из повторных прогонов. Метод, который позволяет управлять размером выборки, заключается в задании определенного числа компонентов, поступающих на вход модели. В этом случае имитация продолжается до тех пор, пока заданное число компонентов не будет полностью обработано в модели. Таким образом, после окончания имитационного прогона система будет находиться в состоянии «пуст и свободен». Другим подходом, аналогичным данному, является задание числа компонентов, обрабатываемых в системе. При этом имитационный прогон может завершиться, когда система находится в любом, отличном от пустого состоянии. Используя такой подход, необходимо всегда обеспечивать, чтобы компоненты, оставшиеся необработанными, были типичными представителями выборки. Например, этот метод непригоден, когда в модели системы используется правило распределения работ по минимуму времени их выполнения, и, следовательно, к концу имитации в очереди могут накопиться работы, для выполнения каждой из которых требуется много времени.

Еще одним подходом к управлению продолжительностью имитационного прогона является применение правил автоматической остановки, которые позволяют автоматически отслеживать результаты моделирования через заданные интервалы времени в процессе имитации. Имитация прекращается, когда оценка дисперсии среднего становится меньше заданной величины. Более подробно правила автоматической остановки рассматриваются в гл. 13.

Если мы оцениваем дисперсию выходной переменной X с помощью повторных прогонов и если предполагаем, что X нормально распределена (если X является выборочным средним, это вполне справедливо), то число независимых повторных прогонов, которое необходимо осуществлять для достижения заданного доверительного интервала для X, будет равно

$$I = \left(\frac{t_{\alpha/2, I-1} S_x}{g} \right)^2$$

$t_{\alpha/2,I-1}$ - величина, взятая из таблицы критических значений t-статистика с I-1 степенями свободы; g – половина длины заданного доверительного интервала.

К сожалению, применение этой формулы требует информации о t-статистике с I-1 степенями свободы и S_X. Обычно устанавливают значение I, проводят I повторных прогонов имитационной модели, на основе проведенных прогонов вычисляют значения t и s_x, а затем применяют приведенную выше формулу для проверки достаточности начальных предположений или для определения необходимого числа дополнительных прогонов.

2.16. Заключение

В главе изложены основы теории вероятностей и математической статистики, знание которых необходимо при проведении имитационного анализа. При этом дается широкий обзор вероятностных и статистических положений, имеющих отношение к имитационному моделированию, без подробного изложения каждого из них. Представленный материал вполне достаточен для понимания цели имитационного моделирования и экспериментальной природы имитационного анализа. Он также дает возможность читателю перейти далее к изучению более тонких аспектов статистического анализа в имитационном моделировании, изложенных в гл.13.

2.17. Упражнения

2.1. Используя приведенные ниже данные 20 имитационных прогонов о времени пребывания посетителя в системе, вычислить оценки для выборочного среднего, дисперсии и коэффициента вариации. Построить гистограмму, содержащую 5 интервалов (длина каждого равна 1), причем нижняя граница первого интервала равна 0.

1,1; 2,8; 3,7; 1,9; 4,9; 1,6; 0,4; 3,8; 1,5; 3,4; 1,9; 2,1; 3,8; 1,6; 3,2; 2,9; 3,7; 2,0; 4,2; 3,3.

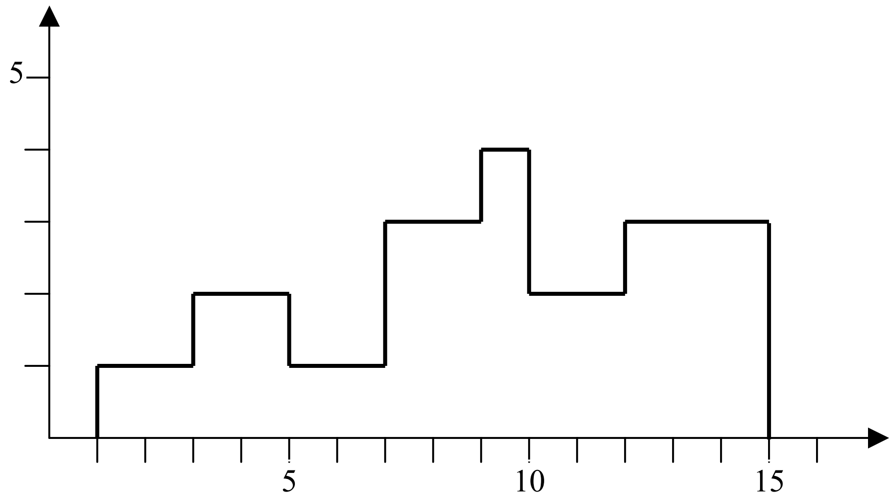
2.2. На рис. 2.17 показано изменение числа посетителей в очереди в течение 15-мин интервала. Вычислить среднее и среднеквадратичное отклонение числа ожидающих в очереди посетителей.

2.3. При проектировании метро применялась имитационная модель. Ниже приведены средние времена ожидания пассажиров в минутах, полученные в результате 20 независимых повторных прогонов этой модели. Построить 99%-ный доверительный интервал для среднего времени ожидания пассажира.

15, 17, 14, 15, 16, 14, 15, 18, 15, 14,

15,20, 17, 14, 16, 16, 15, 18, 14, 15.

2.4. Ниже приведены средненедельные величины стоимости эксплуатации склада (в тысячах долларов) для двух различных стратегий управления запасами. Эти значения были получены по результатам независимых повторных прогонов имитационной модели системы. Проверьте гипотезу о том, что средненедельная стоимость эксплуатации при стратегии А меньше, чем при стратегии В с 5%-ным уровнем значимости.



СТРАТЕГИЯ А:

1,2; 1,3; 1,1; 1,4; 0,9; 1,1; 1,5; 1,2;

1,3; 1,2; 1,3; 1,4; 1,1; 1,2; 1,0.

СТРАТЕГИЯ В: 1,1; 1,5; 1,4; 1,3; 1,6; 1,5; 1,4; 1,5;

1,7; 1,3; 1,6; 1,5; 1,8; 1,7; 1,6.

2.5. При имитационном исследовании проекта сборочной линии было проведено 10 независимых повторных прогонов для оценки средней часовой производительности линии. На основе этих данных вычислить оценку числа прогонов, требуемых для определения средней производительности за 2 единицы времени с уровнем значимости, равным 1%.

157, 162, 151, 170, 162, 157, 166, 165, 152, 160.

2.6. Показать, что сумма двух независимых пуассоновских случайных величин с математическими ожиданиями μ_1 и μ_2 является также пуассоновской величиной с ожиданием, равным $\mu_1+\mu_2$ (напомним, что генератор моментов распределения Пуассона с ожиданием μ имеет вид $e^{\mu(e-1)}$)). Как распределена сумма двух независимых, нормально распределенных случайных величин?

2.7. Багдадский вор заключен в подземелье с тремя дверьми. Одна дверь ведет на свободу, другая — в длинный туннель, а третья — в короткий. Попав в один из туннелей, вор снова оказывается в темнице. Каждый раз после этого он опять пытается выйти на свободу, но при этом не помнит, в какую дверь он входил в прошлый раз, т. е. мы предполагаем марковский процесс. Вероятность того, что вор выберет дверь, ведущую на свободу, равна 0,3; вероятность выбора двери в короткий туннель равна 0,2; вероятность выбора двери в длинный туннель равна 0,5. Пусть время пребывания вора в коротком и длинном туннелях равно соответственно 6 и 3 единицам времени. Определить среднее время, которое затратит вор на поиск пути на свободу.

2.8. Применить мультипликативный конгруэнтный метод для генерации последовательности из 10 случайных чисел с $c = 256$, $a = 13$, $b = 0$, $z_0 = 51$.

2.9. При $z_{i+1}=(az_1+b) \pmod c$ показать, что z_{i+1} является функцией только от z_0, a, b и c, т. е. $z_{i+1} = (a^{i+1}z_0+b(a^{i+1} - 1)/(a - 1)) \pmod c$. Вычислить по этой формуле z_9 для значений, данных в упр. 2.8.

2.10. Использовать метод обратной функции для преобразования случайных чисел, полученных в упр. 2.8, в случайную выборку из непрерывного распределения со следующей функцией плотности вероятности:

$$f(x) = \begin{cases} 3x^2/8, если 0 \le x \le 2 \\ 0, в \cdot противном \cdot случае \end{cases}$$

2.11. Использовать метод обратной функции для преобразования случайных чисел из упр. 2.8 в выборку из дискретного распределения со следующей функцией вероятности:

$$P(0)= 1/5;\ P(1)=1/5;\ P(2) = 2/5,\ P(3) = 1/5$$

Конгруэнтные генераторы [9]

Смешанные конгруэнтные генераторы

Полный период цикла, равный 2^B, будет получен на B-битовом компьютере для генератора

$$z_{i+1}=(az_i+b) \pmod c$$

в том случае, если $c=2^B$, b — простое число относительно c (т. е. наибольший общий делитель b и c равен 1) и $a \equiv 1 \pmod 4$ или $a=1+4k$, где k целое.

Мультипликативные конгруэнтные генераторы

Максимальный период цикла, равный 2^{B-2}, будет получен на В-битовом компьютере для генератора

$z_{i+1} = az_1 (\mathrm{mod}\ c)$

в том случае, если $c=2^B$, $a=\pm 3+8k$ или $a=l+4k$ для целых k и z_0 нечетно.

Такие генераторы называют мультипликативными с максимальным периодом [9].

Для мультипликативных конгруэнтных генераторов период, равный c - 1, может быть получен, если $c=2^в$—1, а а — простой корень с (а является простым корнем с, если $a^{c-1}= l + ck$, где k — целое, и для любого целого q<c-1 выражение $(a^q—l)/c$ не является целым). Эти генераторы называются мультипликативными конгруэнтными с простым модулем.

ЛИТЕРАТУРА

1. Abramowitz, M. and I. A. Stegun, Eds., *Handbook of Mathematical Functions,* Applied Mathematics Series 55, Washington, D.C.: National Bureau of Standards, 1964.
2. Barnes, R. M., *Motion and Time Study: Design and Measurement of Work,* Sixth Edition, John Wiley, 1968,
3. Bartee, E. M., *Engineering Experimental Design Fundamentals,* Prentice-Hall 1968.
4. Beightler, C. S., L. G. Mitten, and G. L. Nemhauser, "A Short Table of Z-Transforms and Generating Functions", *Operations Research,* Vol. 9, 1961, pp. 576—577.
5. Beyer, W. A., R. B. Roof and D. Williamson, "The Lattice Structure of Mul-tiplicotive Congruential Pseudo-Random Vectors", *Math. Comp.,* Vol. 25, 1971, pp. 345—363.
6. Conveyou, R. R. and R. D. MacPherson, "Fourier Analysis of Uniform Random Number Generators", /. *ACM,* Vol. 14, 1967, pp. 100—119.
7. Feller, W.; *An Introduction to Probability Theory and Its Applications,* John Wiley, 1950.
8. Feller, W., *An Introduction to Probability Theory and Its Applications,* John Wiley, Vol. II, 1972.
9. Fishman, G. S., *Principles of Discrete Event Simulation,* John Wiley, 1978.
10. Giffin, W., *Transform Techniques for Probability Modeling,* Academic Press, 1975.
11. Hahn, G. J. and S. S. Shapiro, *Statistical Methods in Engineering,* John Wiley, 1967.
12. Hald, A., *Statistical Theory with Engineering Applications,* John Wiley, 1952.
13 Hastings, N. A. J. and J. B. Peacock, *Statistical Distributions,* Butterworth, 1975.
14. Hogg, R. V. and A. T. Craig, *Introduction to Mathematical Statistics,* Mac-' millan, 1970.
15 Hull, T. E. and A. R. Dobell, "Random Number Generators", *SIAM Review,* Vol. 4, 1962, pp. 230—254.
16 Lewis, T. G., *Distribution Sampling for Computer Simulation,* Lexington Books, 1975.
17. Marsaglia, G., "The Structure of Linear Congruential Sequences", in *Applications of Number Theory to Numerical Analysis, S.* K. Zaremba, ed., Academic Press, 1972.
18. Niebel, B. W., *Motion and Time Study,* Fourth Edition, Richard D. Irwin, 1967.
19. Papoulis, A., *Probability, Random Variables, and Stochastic Processes,* McGraw-Hill, 1965.
20. Parzen, E., *Modern Probability Theory and Its Applications,* John Wiley, 1960.

21. Phillips, D. T., Applied Goodness of Fit Testing, AIIE Monograph Series, AIIE-OR-72-1, Atlanta, Georgia, 1972.
22. Pritsker, A. A. B., *The GASP IV Simulation Language,* John Wiley, 1974.
23. Pritsker & Associates, AID: *Fitting Distributions to Observations* by K. J. Musselman, W. R. Penick and M. E. Grant, P&A, West Lafayette, IN 1981.
24. RAND Corporation, *A Million Random Digits with* 1,000,000 *Normal Deviates,* Free Press, 1955.