17 МОДЕЛИРОВАНИЕ НЕПРЕРЫВНЫХ СЛУЧАЙНЫХ ВЕЛИЧИН-1

17.1 Равномерное распределение

Равномерное распределение случайной величины X (рисунок 1) задается плотностью

$$f(x)=\begin{cases}0 & при\ x<a,\\ 1/(b-a) & при\ a\le x\le b,\\ 0 & при\ x>b.\end{cases} \tag{1}$$

Математическое ожидание и дисперсия X равны соответственно

$$\overline{X}=\frac{a+b}{2},$$

$$D_X=\frac{(b-a)^2}{12}.$$

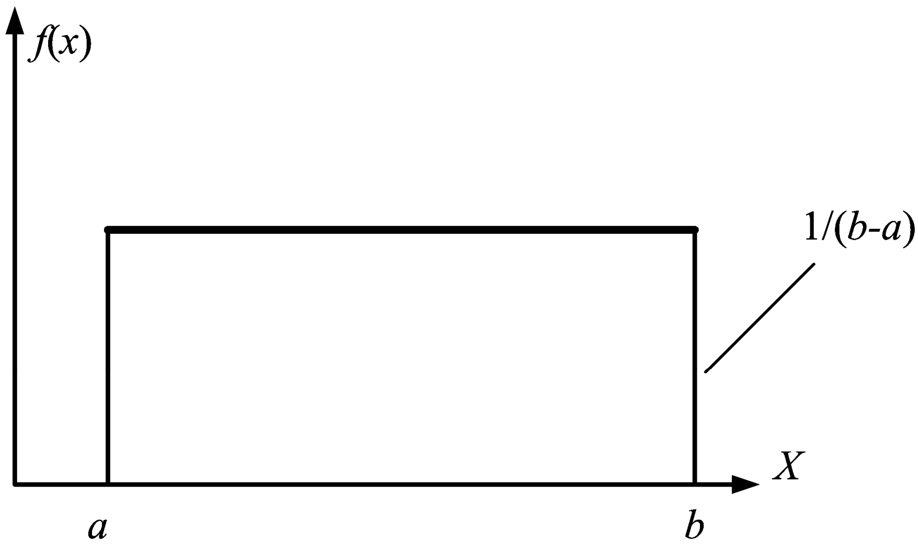

Рисунок 1 – График плотности равномерного распределения

Произвольное равномерное распределение случайной величины X получается из стандартного равномерного распределения в интервале (0...1) величины γ с помощью следующего преобразования:

$$X = a + \gamma \cdot (b - a). \tag{2}$$

17.2 Моделирование случайной величины, заданной гистограммой

На практике часто распределение случайной величины задается гистограммой, построенной по опытным данным. В этом случае распределение некоторой случайной величины X задается таблицей следующего вида:

Границы интервала	x_0, x_1	x_1, x_2	x_2, x_3	$\dots$	x_{n-1}, x_n
Вероятность	P_1	P_2	P_3	$\dots$	P_n

В этом случае в начале моделируется дискретная случайная величина с вероятностями номеров интервалов P_i, $i = 1,...,n$, а затем равномерно распределенная величина в соответствующем интервале. Естественно, что

$$\sum_{i=1}^{n} P_i = 1.$$

На рисунке 2 приведена блок-схема такого алгоритма.

Для генерации одного значения случайной величины X требуется два стандартных случайных числа g_1 и g_2. Первое число используется для случайного выбора интервала, а второе – для генерации равномерного распределения внутри интервала.

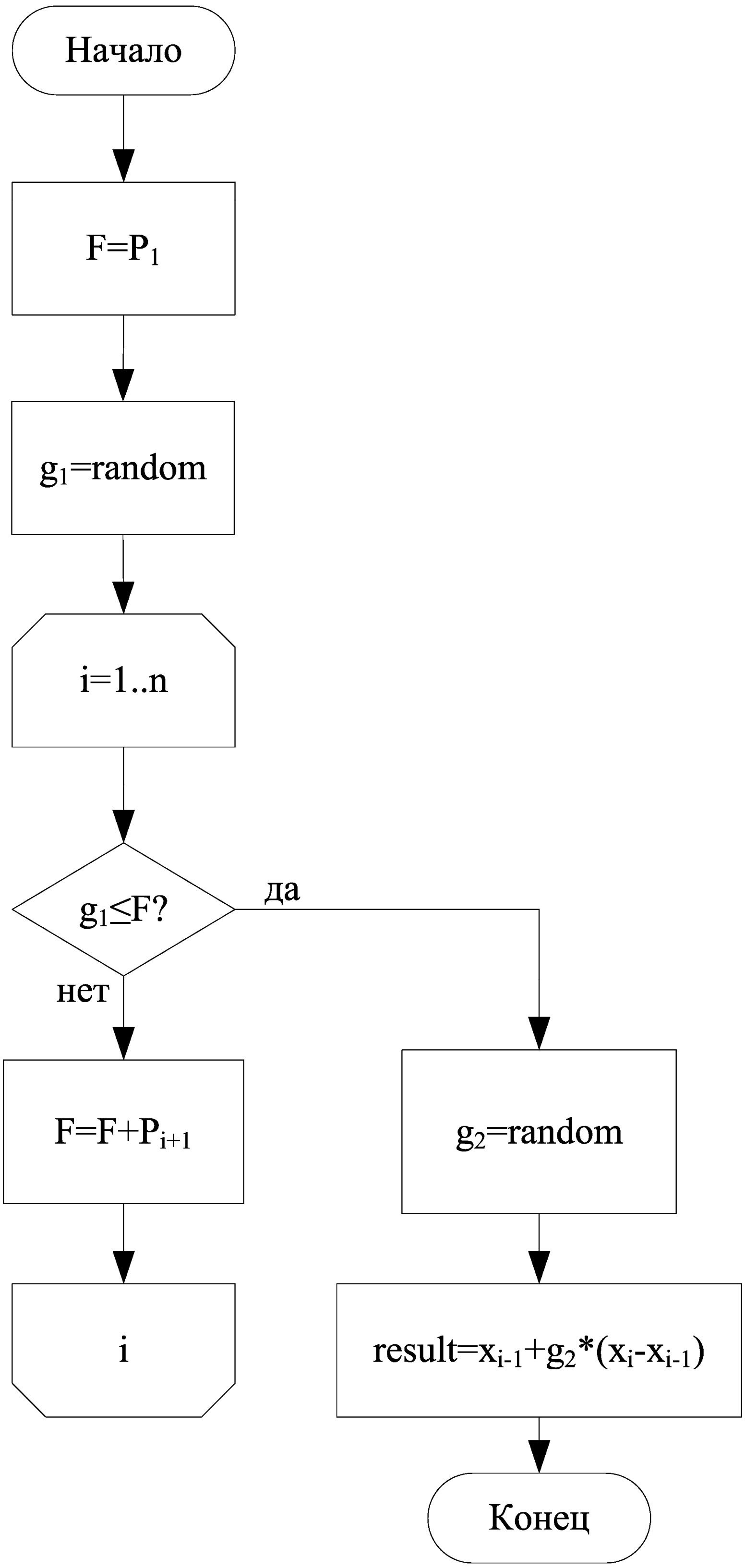

Рис. 5.16. Блок-схема алгоритма моделирования случайной величины, закон распределения которой задан в виде гистограммы:
P, x – входные массивы вероятностей и границ интервалов,
n – число интервалов

17.3 Метод обратной функции

17.3.1 Общая схема метода

Этот метод считается основным для моделирования непрерывных случайных величин. Если X имеет функцию распределения $F(x)$, то обратной считается функция $x = F^{-}(\gamma)$, которая по заданной вероятности γ определяет значение x (рисунок 3).

Если γ с равной вероятностью принимает значения в интервале 0...1, то значения x будут распределены в соответствии с функцией распределения $F(x)$. Таким образом, мы получаем формулу для преобразования равномерного распределения в заданное распределение, то есть, если $\gamma = random$, то

$$X = F^{-}(\gamma), \tag{3}$$

где $F^{-}(\gamma)$ определяется в результате решением уравнения

$$F(x) = \gamma \tag{4}$$

относительно x.

Проиллюстрируем применение метода обратной функции на конкретных примерах.

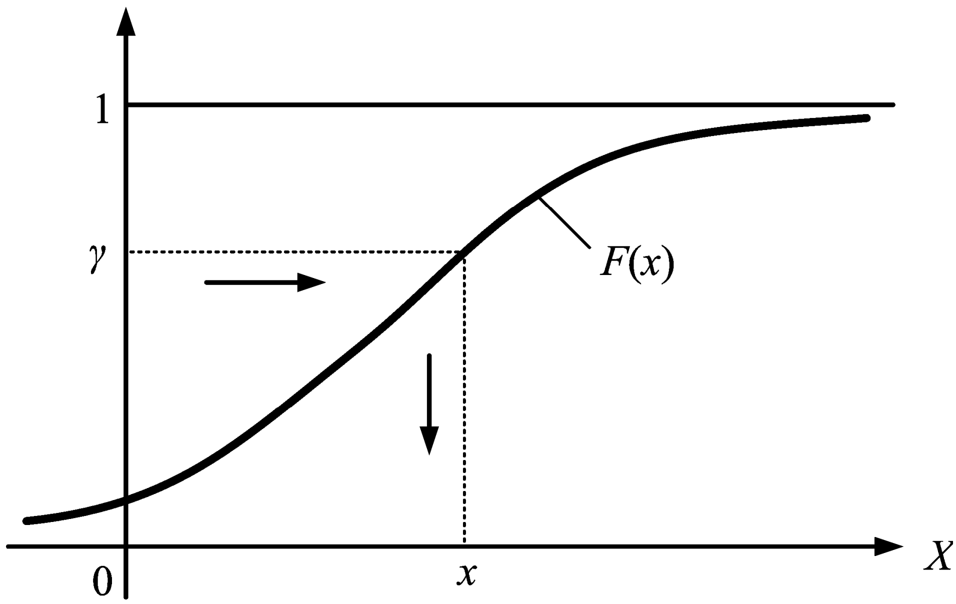

Рисунок 3 – Иллюстрация к определению обратной функции

17.3.2 Показательное распределение

Показательное распределение имеет плотность

$$f(x) = \lambda \cdot \exp(-\lambda \cdot x), \quad x \geq 0. \tag{5}$$

Функция распределения

$$F(x) = 1 - \exp(-\lambda \cdot x), \quad x \geq 0. \tag{6}$$

Подставляя вместо $F(x)$ значение γ, получаем уравнение

$$\gamma = 1 - \exp(-\lambda \cdot x),$$

решая которое относительно x, получаем,

$$x = -\tfrac{1}{\lambda} \cdot \ln(1-\gamma).$$

Но так как значение выражения $1-\gamma$ распределено так же, как и γ,то для генерирования экспоненциально распределенной величины можно пользоваться более простой формулой:

$$X = -\tfrac{1}{\lambda} \cdot \ln\gamma. \tag{7}$$

Если плотность задана в форме

$$f(x) = \frac{1}{a} \cdot \exp(-\frac{x}{a}), \quad x \geq 0 \quad ,$$

то $X = -a \cdot \ln(\gamma)$.

17.3.3 Распределение Вейбулла–Гнеденко

Распределение Вейбулла–Гнеденко выражается следующей формулой:

$$F(x) = 1 - \exp(-(\tfrac{x}{\rho})^{\alpha}), \quad x \geq 0. \tag{8}$$

Подставляя вместо $F(x)$ величину γ, получаем уравнение

$$\gamma = 1 - \exp(-(\tfrac{x}{\rho})^{\alpha}),$$

решение которого относительно x имеет вид

$$x = \rho \cdot (-\ln(1-\gamma))^{\frac{1}{\alpha}}.$$

По тем же соображениям, что и в предыдущем случае, заменяем $1-\gamma$ на γ и окончательно получаем формулу для генерирования распределения Вейбулла-Гнеденко:

$$x = \rho \cdot (-\ln(\gamma))^{\frac{1}{\alpha}}. \tag{9}$$

Если известны математическое ожидание случайной величины X (обозначим его $\overline{X}$) и коэффициент вариации ν_x, то параметры ρ и α определяются из следующих уравнений:

$$\nu_x = \sqrt{\frac{\Gamma(1+\frac{2}{\alpha})}{\Gamma^2(1+\frac{1}{\alpha})} - 1};$$
$$\rho = \frac{\overline{X}}{\Gamma(1+\frac{1}{\alpha})}. \tag{10}$$

Здесь $\Gamma(\ )$ – гамма-функция. При целом аргументе гамма-функция выражается через факториал:

$$\Gamma(n) = (n-1)! = 1 \cdot 2 \cdot 3 \ldots (n-1).$$

При практических расчетах можно пользоваться следующими приближенными зависимостями:

$$\frac{1}{\alpha} = 1.04 \cdot \nu_x - 0.04 \text{ при } \nu_x \geq 0{,}2,$$

$$\frac{1}{\alpha} = 0.7976 \cdot \nu_x + 0.3142 \cdot \nu_x{}^2 \text{ при } \nu_x < 0{,}2.$$

17.3.4 Нормальное распределение

Плотность нормального распределения случайной величины X задается следующей формулой:

$$f(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot \exp[-\frac{(x-a)^2}{2\sigma^2}]. \tag{11}$$

Математическое ожидание $\overline{X} = a$, дисперсия $D_X = \sigma^2$.

Непосредственно применить для этого распределения метод обратной функции сложно, так как нет простого выражения для обратной функции. Но эту трудность удается преодолеть путем перехода к двумерному нормальному распределению в полярных координатах.

Двумерная нормальная плотность вектора (X, Y) при средних значениях $\overline{X} = \overline{Y} = 0$ и квадратичных отклонениях $\sigma_1 = \sigma_2 = 1$ имеет вид

$$f(x, y) = \frac{1}{2\pi} \cdot \exp(-(x^2 + y^2)/2).$$

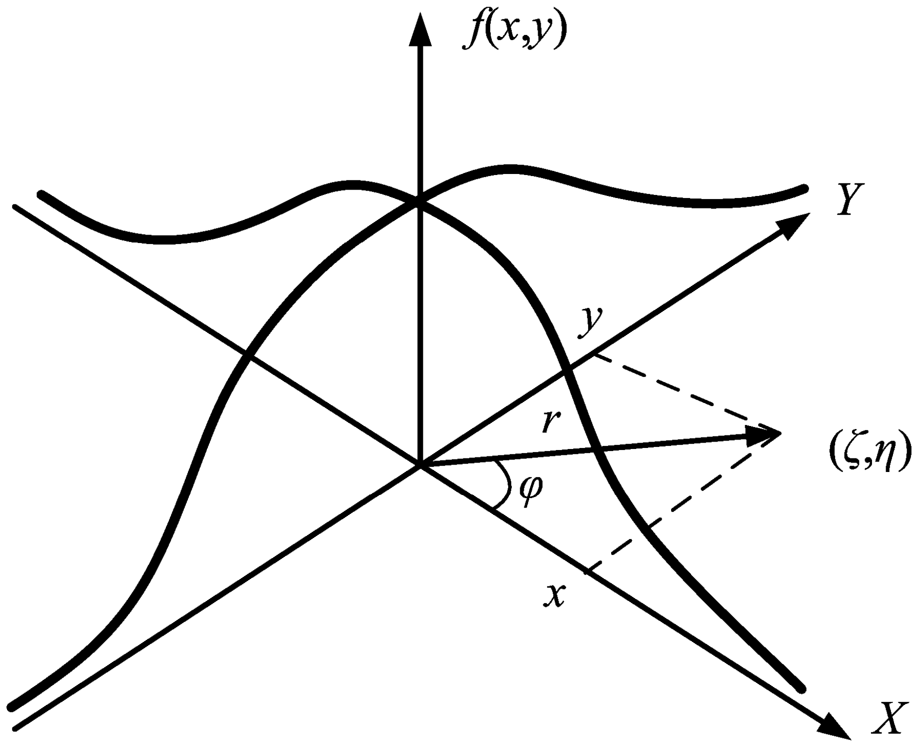

Рисунок 4 – Двумерная нормальная плотность распределения

Перейдем к полярным координатам путем преобразования

$$x = r \cdot \cos(\varphi), \quad y = r \cdot \sin(\varphi).$$

Пусть ρ, θ – случайные полярные координаты точки (X, Y).

Совместная плотность распределения вектора (ρ, θ) в полярных координатах

$$f(r, \varphi) = \frac{r}{2\pi} \cdot \exp(-\frac{r^2}{2}) = f_1(r) \cdot f_2(\varphi).$$

Одномерные плотности

$$f_1(r) = r \cdot \exp(-r^2 / 2), \quad r \ge 0;$$

$$f_2(\varphi) = \frac{1}{2\pi}, \quad 0 \le \varphi \le 2\pi.$$

Плотность $f_1(r)$ в литературе называют законом Релея или распределением эксцентриситета.

Соответствующие интегральные функции распределения имеют вид:

$$F_1(r) = \int_0^r f_1(r)dr = 1 - \exp(-r^2/2), \quad r \ge 0;$$

$$F_2(\varphi) = \int_0^\varphi \frac{1}{2\pi} d\varphi = \frac{\varphi}{2\pi}, \quad 0 \le \varphi \le 2\pi.$$

Обратные функции получаются из следующих уравнений:

$$F_1(\rho) = 1 - \gamma_1;$$

$$F_2(\theta) = \gamma_2,$$

решение которых относительно ρ и θ дает следующие функции для генерации соответствующих распределений:

$$\rho = \sqrt{-2 \cdot \ln(\gamma_1)},$$

$$\theta = 2\pi \cdot \gamma_2,$$

где γ_1, γ_2 – пара стандартных случайных числа, распределенных в интервале (0..1.0).

Перейдем теперь от полярных координат (ρ, θ) к обычным (X, Y):

$$X = \sqrt{-2 \cdot \ln \gamma_1} \cdot \cos(2\pi \cdot \gamma_2),$$

$$Y = \sqrt{-2 \cdot \ln \gamma_1} \cdot \sin(2\pi \cdot \gamma_2),$$

где X, Y – пара независимых нормально распределенных случайных величин с нулевыми средними значениями и квадратичными отклонениями, равными 1.

Для получения произвольного нормального распределения величины Z со средним значением a и дисперсией σ^2 воспользуемся формулами:

$$Z_1 = a + \sigma \cdot X = a + \sigma \cdot \sqrt{-2 \cdot \ln(\gamma_1)} \cdot \cos(2\pi \cdot \gamma_2),$$
$$Z_2 = a + \sigma \cdot Y = a + \sigma \cdot \sqrt{-2 \cdot \ln(\gamma_1)} \cdot \sin(2\pi \cdot \gamma_2).$$

Таким образом, из пары стандартных случайных чисел γ_1, γ_2 получаем два нормально распределенных числа Z_1, Z_2 со средним значением a и квадратичным отклонением σ.

К сожалению, для многих других распределений метод обратных функций оказывается непригодным, так как простого явного выражения для обратной функции не удается получить.

17.3.5 Треугольное распределение

Плотность в этом случае выражается так:

$$f(x) = \begin{cases} 0 \quad \text{при} \quad x > a, \\ \dfrac{2(x-a)}{(b-a)^2} \quad \text{при} \quad a \le x \le b, \\ 0 \quad \text{при} \quad x > b. \end{cases}$$

График плотности представлен на рисунке 5.

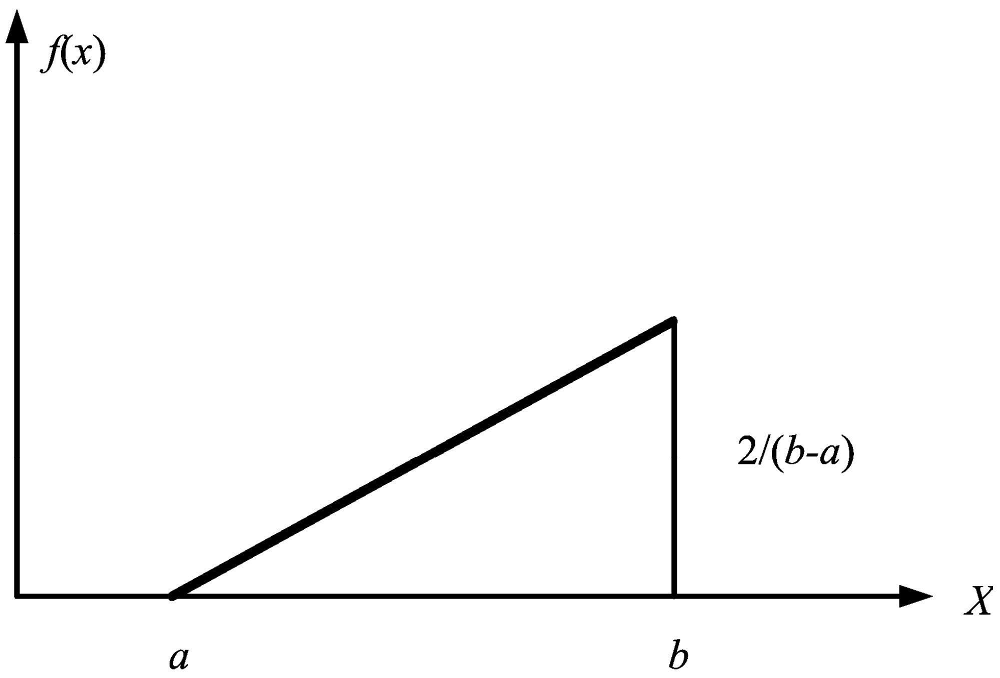

Рисунок 5 – Треугольное распределение

Функция распределения

$$F(x)=\int\limits_{-\infty}^{x} f(x)dx=\begin{cases}0 \quad \text{при} \quad x<a,\\ \dfrac{(x-a)^2}{(b-a)^2} \quad \text{при} \quad a\le x\le b,\\ 1 \quad \text{при} \quad x>b.\end{cases}$$

Уравнение $F(x)=\gamma$ в данном случае имеет вид

$$\frac{(x-a)^2}{(b-a)^2}=\gamma .$$

В результате решения этого уравнения получаем формулу для генерации случайных чисел, распределенных по треугольному распределению:

$$X=a+(b-a)\sqrt{\gamma} ,$$

где γ – случайное число, распределенное равномерно в интервале 0...1.

17.4 Метод Неймана

17.4.1 Общая схема метода

Метод Неймана называют также методом исключения. Метод состоит в следующем. Пусть нужно промоделировать случайную величину X с плотностью $f(x)$, сосредоточенную на интервале (a,b), причем максимум плотности не превышает M (рисунок 6).

Такая случайная величина моделируется следующим образом:

1) генерируем два стандартных случайных числа γ_1 и γ_2 и определяем соответствующую координату точки (x,y) в прямоугольнике $(a\le x\le b; 0\le y\le M)$ по следующим формулам:

$$\begin{aligned} x&=a+\gamma_1\cdot(b-a),\\ y&=\gamma_2\cdot M; \end{aligned} \qquad (12)$$

2) если точка (x,y) лежит под кривой $f(x)$, то искомая случайная величина $X=x$, а если же эта точка лежит выше $f(x)$, то генерируется новая пара

стандартных случайных чисел γ_1, γ_2 и снова выполняется шаг 1.

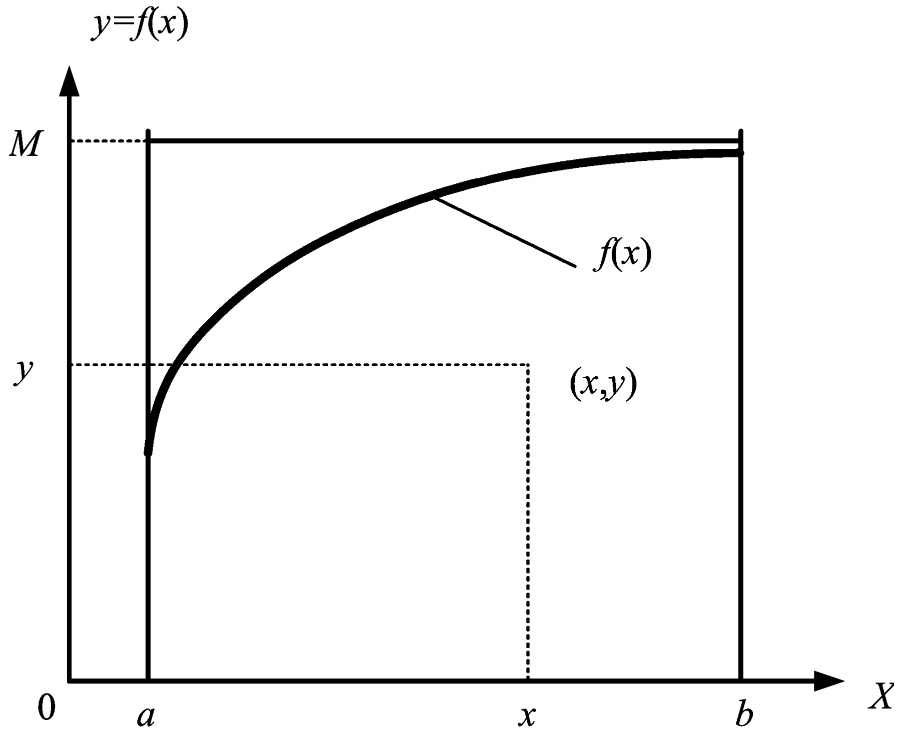

Рисунок 6 – Иллюстрация к методу Неймана

Так как число x равномерно распределено в интервале (a,b), то вероятность того, что точка (x, y) в полосе $(x, x+dx)$, пропорциональна dx. А так как y равномерно распределено в интервале $(0, M)$, то вероятность того, что отобранное значение $X = x$ окажется в интервале $(x, x+dx)$, равна $f(x)dx$.

Чтобы получить одно случайное число X, распределенное с плотностью $f(x)$, требуется в среднем $1/q$ пар стандартных чисел (γ_1, γ_2), где q – вероятность того, что точка (x, y) попадет под график $f(x)$.

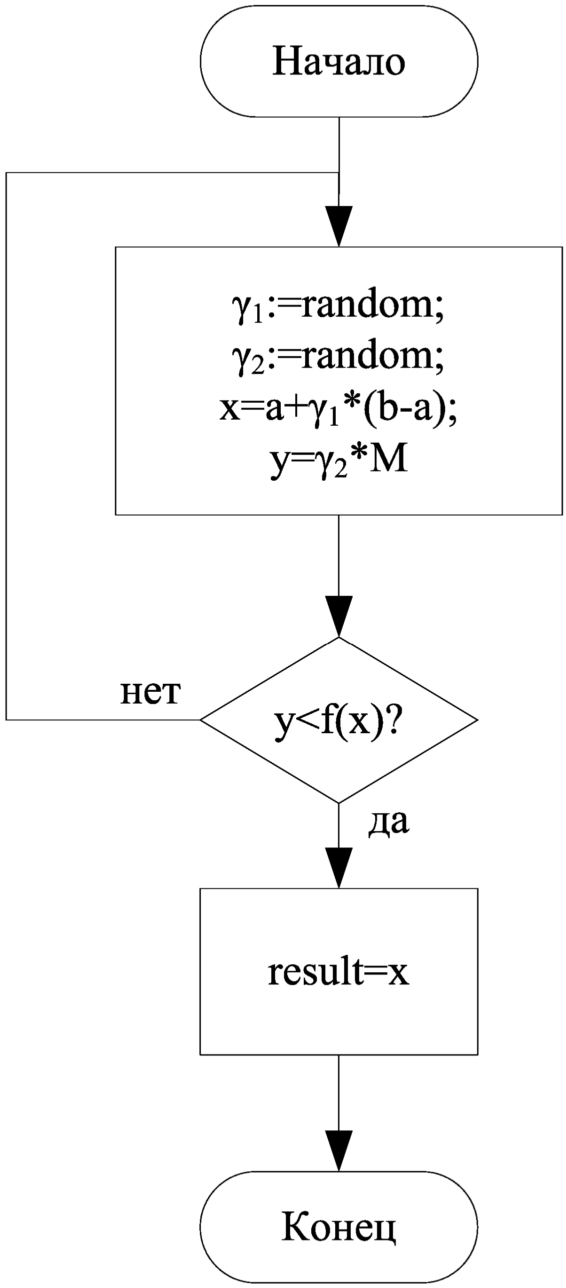

Рисунок 7 – Блок-схема алгоритма Неймана

Эта вероятность рассчитывается по формуле

$$q = \int_a^b f(x)dx/(M\cdot(b-a)) = \frac{1}{M\cdot(b-a)}.$$

Для генерации одного случайного числа с плотностью распределения $f(x)$ требуется в среднем $1/q$ пар стандартных чисел, генерируемых, например, функцией Турбо Паскаля *random*.

17.4.2 Моделирование трапецеидального распределения 1

Плотность трапецеидального распределения выражается формулой

$$f(x) = A \cdot x + B \text{ при } a \le x \le b,$$

где $A = \dfrac{c-d}{a-b}$, $B = \dfrac{a \cdot d - c \cdot b}{a-b}$. При x вне отмеченного интервала $f(x) = 0$.

График плотности приведен на рисунке 8.

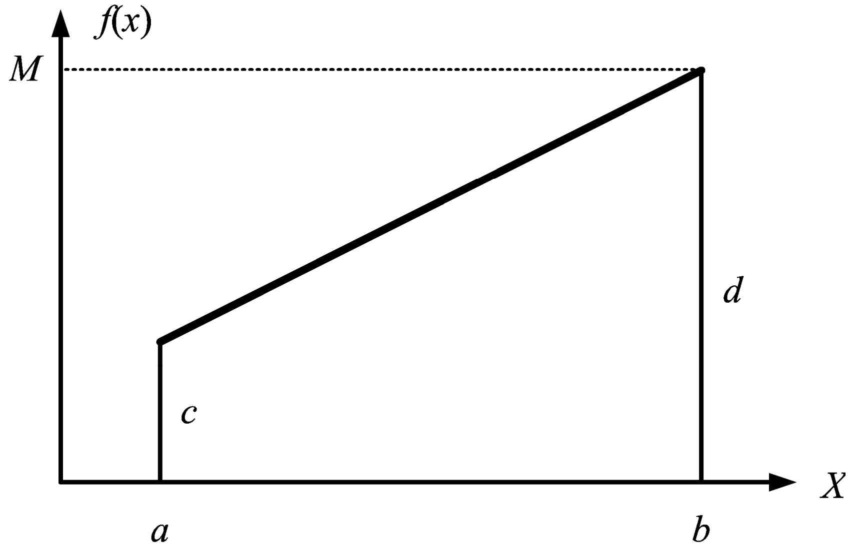

Рисунок 8 – Плотность трапецеидального распределения

Параметр d находится из условия нормировки вероятностей $\int_a^b f(x)dx = 1$,

то есть

$$d = \frac{2}{b-a} - c, \quad c \ge 0, \quad d \ge 0.$$

Формулы аналогичные (12) в данном случае имеют вид:

$$x = a + \gamma_1 \cdot (b-a);$$
$$y = \gamma_2 \cdot M,$$

где $M = \max(c, d)$.

Алгоритм моделирования состоит в следующем. Если

$$y < \frac{c-d}{a-b} \cdot x + \frac{a \cdot d - c \cdot b}{a-b},$$

то $X = x$, иначе разыгрывается новая пара случайных чисел γ_1, γ_2, и процесс расчета и сравнения продолжается.

Вероятность q в данном случае равна $\dfrac{1}{\max(c,d) \cdot (b-a)}$, а среднее число стандартных чисел на одно распределенное по трапецеидальному закону

$$1/q = \max(c,d) \cdot (b-a).$$

17.5 Контрольные вопросы

1. Как можно получить произвольное равномерное распределение случайной величины X из стандартного равномерного распределения в интервале (0...1) величины γ?

2. Как моделируется случайная величина, заданная гистограммой?

3. В чем суть метода обратной функции?

4. Как записывается формула для генерирования распределения Вейбулла-Гнеденко?

5. Как из пары стандартных случайных чисел γ_1, γ_2 получить два нормально распределенных числа Z_1, Z_2 со средним значением a и квадратичным отклонением σ?

6. В чем суть метода Неймана?