

## 4.5. Моделювання неперервних випадкових величин

Існує кілька методів моделювання значень неперервних випадкових величин з довільним законом розподілу на основі випадкових чисел, рівномірно розподілених у інтервалі  $[0, 1]$ : метод оберненої функції, метод відсіювання, наближені методи тощо.

### 4.5.1. Метод оберненої функції

Розглянемо метод моделювання випадкової величини, яка має функцію щільності ймовірностей  $f(x)$  і монотонно зростаючу функцію розподілу  $F(x)$  (рис. 4.9). Суть методу така. За допомогою генератора випадкових чисел генеруємо значення випадкової величини  $r_i$ , якому відповідає точка на осі ординат. Значення випадкової величини  $x_i$  з функцією розподілу  $F(x)$  можемо одержати з рівняння  $F(x_i) = r_i$ .

Дійсно, якщо на осі ординат відкласти значення  $r_i$  випадкової величини, розподіленої рівномірно в інтервалі  $[0, 1]$ , і на осі абсцис знайти значення  $x_i$  випадкової

величини (рис. 4.9), при якому  $F(x_i) = r_i$ , то випадкова величина  $X = F^{-1}(r)$  буде мати функцію розподілу  $F(x)$ . За визначенням функція розподілу  $F(x)$  випадкової величини  $X$  дорівнює ймовірності  $P(X < x)$ :

$$P(X < x) = P(r < F(X)) = \int_0^{F(x)} f(r) dr = F(x).$$

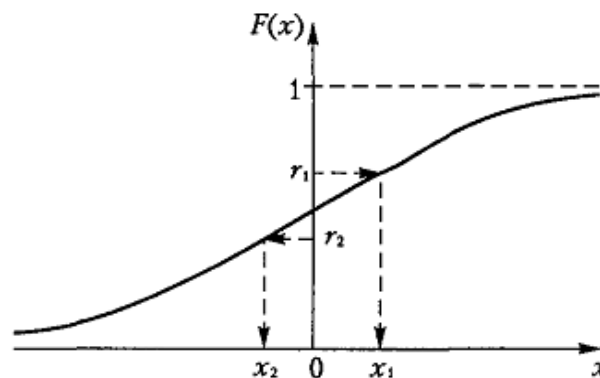


Рис. 4.9. Використання методу оберненої функції для генерування неперервної випадкової величини

Таким чином, послідовність випадкових чисел  $r_1, r_2, r_3, \dots$  перетворюється на послідовність  $x_1, x_2, x_3, \dots$ , яка має задану функцію щільності розподілу  $f(x)$ . Звідси випливає загальний алгоритм моделювання випадкових неперервних величин, що мають задану функцію розподілу ймовірностей:

- ◆ генерується випадкове число  $r_i \in [0, 1]$ ;
- ◆ обчислюється випадкове число  $x_i$ , яке є розв'язком рівняння

$$r_i = \int_{-\infty}^{x_i} f(x) dx.$$

Приклади застосування методу наведені нижче.

#### 4.5.2. Рівномірний розподіл

У загальному випадку випадкова величина  $X$  є рівномірно розподіленою на відрізьку  $[a, b]$ , якщо її щільність розподілу ймовірностей має вигляд

$$f(x) = \begin{cases} 0, & x < a, \\ \frac{1}{b-a}, & a \leq x \leq b, \\ 0, & x > b. \end{cases}$$

Функцію розподілу ймовірностей можна знайти як

$$F(x) = \int_a^x \frac{1}{b-a} dx = \frac{x-a}{b-a},$$

тобто

$$F(x) = \begin{cases} 0, & x < a, \\ \frac{1}{b-a}, & a \leq x \leq b, \\ 1, & x > b. \end{cases}$$

Графіки функцій щільності  $f(x)$  та ймовірності  $F(x)$  зображено на рис. 4.10.

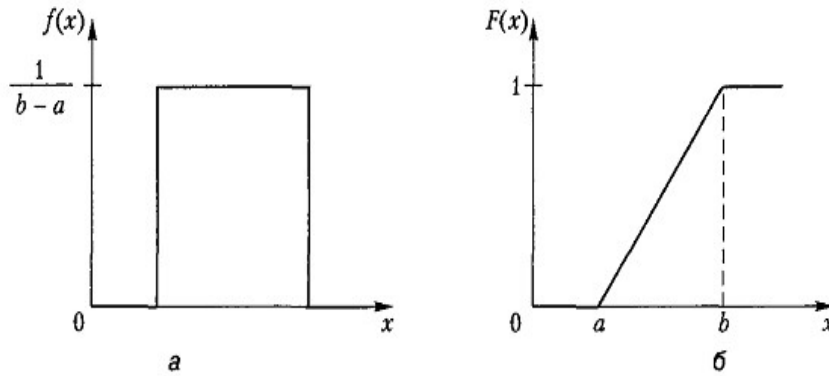


Рис. 4.10. Функції щільності (а) і розподілу (б) рівномірно розподіленої випадкової величини

Математичне сподівання та дисперсія випадкової величини  $X$  визначаються як

$$M(X) = \frac{a+b}{2}, \quad D(X) = \frac{b-a}{12}.$$

Для моделювання випадкової рівномірно розподіленої на відрізку  $[a, b]$  величини можна скористатись методом оберненої функції. Обчислимо функцію розподілу випадкової величини та прирівняємо її до значення  $r_i$ :

$$r_i = \int_a^{x_i} \frac{dx}{b-a} = \frac{x_i - a}{b-a}.$$

Звідси знаходимо значення випадкової величини з функцією розподілу  $f(x)$ :

$$x_i = (b-a)r_i + a.$$

Цю формулу також можна отримати, якщо виконати лінійне перетворення інтервалу  $[0, 1]$  у відрізок  $[a, b]$ . Для цього потрібно змінити масштаб функції рівномірного розподілу, помноживши її на  $(b-a)$ , а потім змістити її на величину  $a$ .

Функція рівномірного розподілу широко застосовується для моделювання випадкових величин, для яких функція розподілу невідома, а відоме лише її середнє значення. У такому випадку припускають, що відомими є середнє значення випадкової величини та деяке розсіювання  $\pm \Delta$  її значень відносно середнього. Це

Прикладами реальних задач, в яких виникає необхідність моделювання рівномірно розподілених випадкових величин, можуть бути аналіз помилок округлення під час проведення числових розрахунків (точність задається як кількість десяткових знаків), час переміщення головок у магнітних накопичувачів (мінімальний та максимальний час), відхилення від розкладу руху транспортних засобів (наприклад, метро).

### 4.5.3. Експоненціальний розподіл

Експоненціальний закон розподілу набув широкого використання в теорії надійності складних систем. Функція щільності експоненціального розподілу випадкової величини має вигляд

$$f(x) = \lambda e^{-\lambda x}.$$

Для її моделювання скористаємося методом оберненої функції. Маємо

$$r_i = \int_0^{x_i} f(x) dx = \int_0^{x_i} \lambda e^{-\lambda x} dx = 1 - e^{-\lambda x_i}.$$

З виразу (4.6) знаходимо значення  $x_i$ :

$$x_i = -\frac{1}{\lambda} \ln(1 - r_i).$$

**Таблиця 4.3.** Вузли та значення апроксимованої функції

$x_i$	$F(x_i)$	$x_i$	$F(x_i)$	$x_i$	$F(x_i)$	$x_i$	$F(x_i)$	$x_i$	$F(x_i)$	$x_i$	$F(x_i)$
0	0	0,4	0,509	0,75	1,38	2,3	0,9	3,2	0,96	5,3	0,995
0,1	0,1	0,5	0,69	0,8	1,6	2,52	0,92	3,5	0,97	6,2	0,998
0,2	0,222	0,6	0,915	0,84	1,83	2,81	0,94	3,9	0,98	7	0,999
0,3	0,355	0,7	1,2	0,88	2,12	2,99	0,95	4,6	0,99	8	0,9998

На рис 4.11 зображено лінійну апроксимацію експоненціальної функції розподілу  $F(x)$  з параметром  $\lambda = 1$ , а на рис. 4.12 – функцію, обернену до неї. Перша функція відображає задані в табл. 4.2 значення, а друга використовується під час моделювання випадкових величин з експоненціальним розподілом.

Можна показати, що випадкові величини  $(1 - r_i)$  мають такий самий розподіл, що і величини  $r_i$ . Тоді, замінивши  $1 - r_i$  на  $r_i$ , отримаємо

$$x_i = -\frac{1}{\lambda} \ln r_i.$$

Випадкові величини з експоненціальним розподілом широко застосовуються в задачах моделювання та аналізу СМО, наприклад під час моделювання процесів виходу з ладу та ремонту обладнання, які виникають у складних системах, у разі визначення інтервалів часу між послідовними викликами абонентів у телефонній мережі або замовлень від незалежних клієнтів у будь-якій мережі обслуговування (швидка допомога, служби ремонту, виклик таксі і т. ін.)

Покажемо ще один підхід до моделювання випадкової величини, розподіленої за експоненціальним законом, з використанням методу оберненої функції, який прийнятий у мові GPSS [68]. Цей підхід передбачає заміну функції розподілу ймовірностей кусково-лінійною апроксимуючою функцією.

Розглянемо найпростіший випадок, коли  $\lambda = 1$ . Апроксимуємо функцію експоненціального розподілу лінійними відрізками таким чином, щоб кожний відрізок можна було використовувати для моделювання за допомогою методу оберненої функції. У мові GPSS функція розподілу апроксимована 23 відрізками. Точки апроксимації  $x_i$  та значення функції  $F(x_i)$  у цих точках наведені в табл. 4.3.

За необхідності моделювання випадкової величини  $X$ , розподіленої за експоненціальним законом з математичним сподіванням  $\lambda_x \neq 1$ , діють таким чином:

- ◆ генерують значення випадкової величини  $r$ , яке використовують як аргумент оберненої функції (рис. 4.12) експоненціального розподілу з параметром  $\lambda = 1$ , і знаходять значення функції  $F(r)$  (для підвищення точності оцінювання параметрів моделювання функцію розподілу  $F(r)$  інтерполюють лінійними відрізками);
- ◆ знаходять добуток отриманого значення  $F(r)$  на  $1/\lambda_x$ .

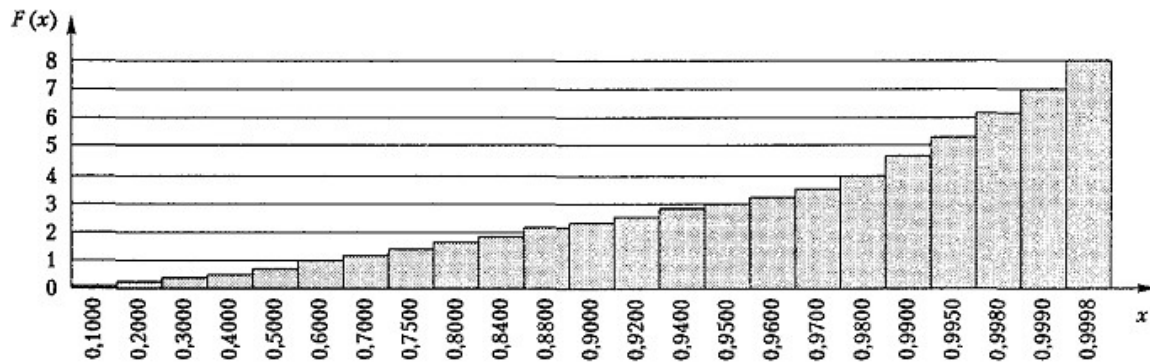


Рис. 4.11. Апроксимація експоненціальної функції розподілу

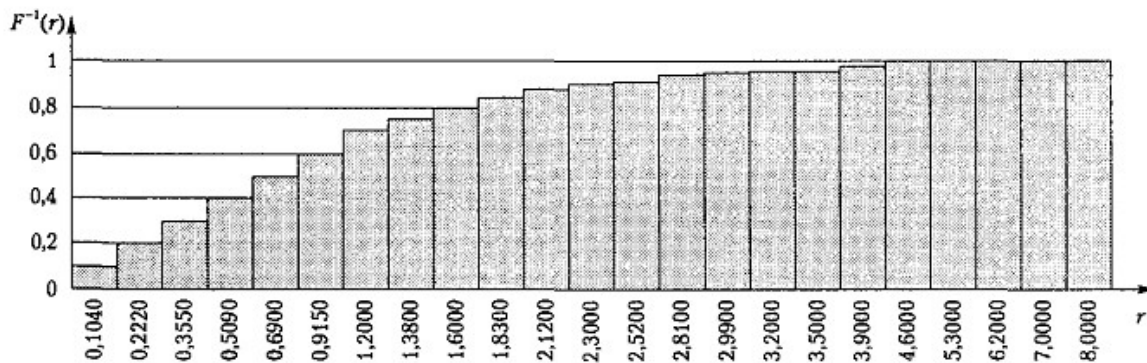


Рис. 4.12. Обернена функція до апроксимованої

### 4.5.5. Нормальний розподіл

Випадкова величина  $X$  має нормальний розподіл (розподіл Гаусса), якщо її щільність розподілу ймовірностей описується виразом

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}},$$

де  $m$  – математичне сподівання, а  $\sigma$  – середньоквадратичне відхилення.

Наприклад, для моделювання часу затримки транзактів, що має експоненціальний закон розподілу з параметром  $\lambda_x = 0,01$ , у мові GPSS використовують блок ADVANCE 100, FN\$XPDIS

У цьому операторі функція XPDIS (див. табл. 4.3) задає експоненціальний розподіл з інтенсивністю  $\lambda = 1$ :

```
XPDIS FUNCTION RN1.C24: exponential distribution function
0.0/.1/.104/.2/.222/.3/.355/.4/.509/.5/.69/
.6/.915/.7.1.2/.75.1.38/.8.1.6/.84.1.83/
.88.2.12/.9.2.3/.92.2.52/.94.2.81/.95.2.99/
.96.3.2.97.3.5/.98.3.9/.99.4.6/.995.5.3/
.998.6.2/.999.7/.9998.8
```

де  $t_j$  –  $j$ -й проміжок часу між надходженнями двох сусідніх вимог;  $\bar{T} = 1/\lambda$  – середнє значення проміжку часу між надходженнями двох сусідніх вимог;  $r_j$  –  $j$ -е число в послідовності випадкових чисел з рівномірним розподілом у інтервалі  $[0, 1]$ .

У мові GPSS для моделювання пуассонівського потоку вимог з  $\bar{T} = 2$  год (одиниця часу в моделі дорівнює 1 хв) використовується блок GENERATE 120, FN\$XPDIS.

#### 4.5.4. Пуассонівський потік

Розглянемо моделювання пуассонівського потоку з інтенсивністю  $\lambda$ , основна властивість якого полягає в тому, що ймовірність надходження  $k$  вимог протягом інтервалу довжиною  $t$  становить

$$p_k(t) = \frac{e^{-\lambda t} (\lambda t)^k}{k!}, \quad k = 1, 2, \dots$$

Для пуассонівського потоку інтервали часу між надходженням двох сусідніх вимог мають експоненціальний закон розподілу (див. розділ 2.1). Тому для його моделювання достатньо отримати ряд чисел з таким розподілом. Це можна реалізувати за допомогою методу оберненої функції, якщо ряд випадкових чисел  $r_j$  рівномірно розподілених у інтервалі  $[0, 1]$ , перетворити згідно з функцією, оберненою до експоненціальної функції розподілу

$$t_j = F^{-1}(x) = -\bar{T} \ln(r_j),$$

Функція розподілу нормально розподіленої величини  $X$  має вигляд

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(x-m)^2}{2\sigma^2}} dx.$$

Графіки функцій щільності ймовірностей  $f(x)$  і розподілу  $F(x)$  зображено на рис. 4.13.

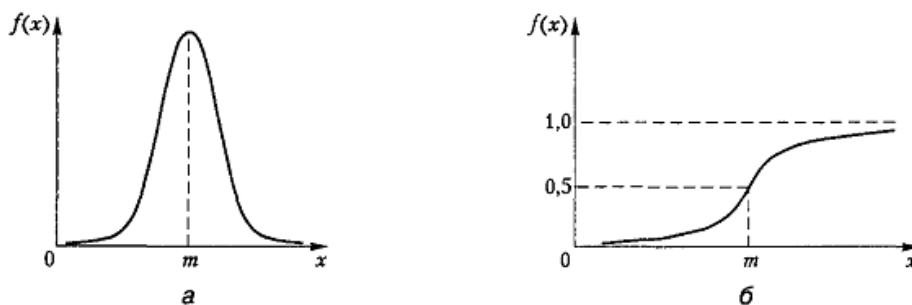


Рис. 4.13. Графіки функції щільності ймовірностей (а) та розподілу (б) випадкової величини

Для моделювання випадкової величини з нормальним законом розподілу безпосередньо скористатися методом оберненої функції не можна, оскільки неможливо аналітично виконати перетворення виду  $X = F^{-1}(r)$ . Тому для моделювання слід скористатися методом згорток.

*Метод згорток* базується на центральній граничній теоремі – одному із найбільш видатних результатів теорії ймовірностей: за широких припущень відносно розподілів суми великої кількості взаємно незалежних малих випадкових величин має місце розподіл, який близький до нормального. Метод згорток передбачає зображення випадкової величини як суми незалежних однаково розподілених випадкових величин зі скінченними математичним сподіванням і дисперсією.

Центральна гранична теорема формулюється таким чином.

Якщо  $X_1, \dots, X_n$  – послідовність незалежних випадкових величин із скінченим математичним сподіванням  $M[X_i] = a, i = \overline{1, n}$  і дисперсією  $D[X_i] = \sigma^2, i = \overline{1, n}$ , то у разі необмеженого збільшення значення  $n$  функція розподілу випадкової величини



$$\bar{X}^*(n) = \frac{1/n (X_1 + \dots + X_n) - a}{\sigma/\sqrt{n}} = \frac{(\bar{X}(n) - a) \sqrt{n}}{\sigma}$$

наближається до функції розподілу стандартного нормального закону  $\Phi(z)$  при всіх значеннях аргументу, тобто

$$F_{\bar{X}^*(n)} \xrightarrow{n \rightarrow \infty} \Phi(z),$$

де

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_0^z e^{-u^2/2} du, \quad u = \frac{x - m}{\sigma}.$$

Функція  $\Phi(z)$  називається функцією Лапласа, для якої є детальні таблиці.

Найпростіший метод отримання значення випадкової величини, що має заданий нормальний розподіл, передбачає виконання таких кроків. Спочатку формують послідовність  $r_i$  ( $i = 1, n$ ) незалежних, рівномірно розподілених у інтервалі  $[0, 1]$

величин і обчислюють суму  $Z = \sum_{i=1}^{12} r_i - 6$ . Величина  $n = 12$  є хорошим наближенням до нормально розподіленої випадкової величини з нульовим математичним сподіванням  $m_z = 0$  і одиничним середньоквадратичним відхиленням  $\sigma_z = 1$ . Нормальний розподіл з параметрами  $m_z = 0$  та  $\sigma_z = 1$  називається стандартним.

Перейти від випадкової величини  $Z$  з нульовим математичним сподіванням і одиничним середньоквадратичним відхиленням до випадкової величини  $X$ , яка має математичне сподівання  $m_x$  і середньоквадратичне відхилення  $\sigma_x$ , дає змогу виконати лінійне перетворення

$$X = \sigma_x Z + m_x. \quad (4.7)$$

У системі моделювання GPSS [68] для моделювання випадкової величини з нормальним розподілом прийнято підхід, який базується на методі оберненої функції. За такого підходу функція нормального розподілу випадкової величини  $Z$  з параметрами  $m_z = 0$  і  $\sigma_z = 1$  наближається кусково-лінійною функцією. Як відзначили розробники інтерпретатора GPSS/PC [68], для цього достатньо 24 відрізки. У табл. 4.4 занесено відповідні значення аргументу  $z_i$  і функції  $\Phi(z)$ .

**Таблиця 4.4.** Вузли апроксимації і значення функції нормального розподілу

$z_i$	$\Phi(z)$	$z_i$	$\Phi(z)$	$z_i$	$\Phi(z)$	$z_i$	$\Phi(z)$	$z_i$	$\Phi(z)$
-5	0	-1,5	0,06681	-0,4	0,34458	0,6	0,72575	2	0,97725
-4	0,00003	-1,2	0,11507	-0,2	0,42074	0,8	0,78814	2,5	0,99379
-3	0,00135	-1	0,15866	0	0,5	1	0,84134	3	0,99865
-2,5	0,00621	-0,8	0,21186	0,2	0,57964	1,2	0,88493	4	0,99997
-2	0,02275	-0,6	0,27425	0,4	0,65542	1,5	0,93319	5	1

Щоб одержати за цим методом 100 пар нормально розподілених чисел, потрібно генерувати 127 пар випадкових чисел. Це простий та швидкий метод, у разі його застосування більша частина часу роботи алгоритму витрачається на обчислення логарифму.

#### 4.5.6. Логарифмічно-нормальний розподіл

Логарифмічно-нормальний розподіл – це такий розподіл випадкової величини, в якій нормальний розподіл має натуральний логарифм її значень. Цей розподіл придатний для моделювання мультиплікативних процесів так само, як нормальний розподіл – для адитивних. Дійсно, використовуючи центральну граничну теорему, можна показати, що добуток незалежних додатних випадкових величин прямує до логарифмічно-нормального розподілу.

Логарифмічно-нормальна величина є результатом взаємодії великої кількості незалежних малих випадкових факторів. Внесок кожного фактора пропорційний

Існує два методи моделювання нормального розподілу, які використовують цю властивість:

**Метод Бокса–Мюллера (Box–Muller).** Генеруємо пару нормально розподілених чисел з  $m_x = 0$  і  $\sigma_x = 1$  за допомогою двох випадкових чисел  $r_1$  і  $r_2$ :

$$x_1 = -2 \ln r_1 \cos(2\pi r_1); \quad x_2 = -2 \ln r_2 \cos(2\pi r_2).$$

Таким чином, отримуємо два числа  $x_1$  і  $x_2$  з нормальним розподілом.

**Метод Марсальї–Брея (Marsaglia–Bray).** Існує більш швидка модифікація цього методу. Генерують два випадкових числа  $r_1$  і  $r_2$ , вважаючи, що  $v_1 = -1 + 2r_1$ ,  $v_2 = -1 + 2r_2$ , обчислюють суму  $s = v_1 + v_2$ . Якщо  $s \geq 1$ , то повторюють процедуру, якщо  $s < 1$ , то одержують два нормально розподілених числа:

$$x_1 = v_1 \sqrt{\frac{-2 \ln s}{s}}, \quad x_2 = v_2 \sqrt{\frac{-2 \ln s}{s}}.$$

Для того щоб одержати нормально розподілену випадкову величину з математичним сподіванням  $m_x \neq 0$  і середньоквадратичним відхиленням  $\sigma_x \neq 1$ , необхідно виконати обчислення за формулою (4.7). На рис. 4.14 зображено графік функції, отриманої в результаті апроксимації функції нормального розподілу  $\Phi(z)$ , а на рис. 4.15 – графік функції  $\Phi^{-1}(r)$ , якою зручніше користуватися під час моделювання (як аргумент використовують значення  $r$  від генератора випадкових чисел і отримують значення функції).

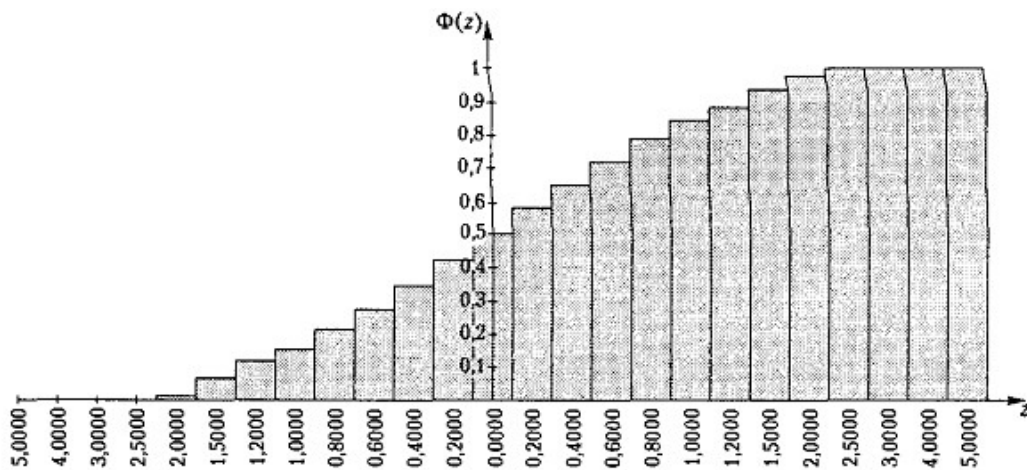


Рис. 4.14. Кусково-лінійна апроксимація функції нормального розподілу

уже досягнутому рівню досліджуваної величини, тобто характер впливу є мультиплікативним. Функція щільності логарифмічно-нормального розподілу має вигляд

$$f_{\eta}(x) = \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{(\ln x - \ln a)^2}{2\sigma^2}}.$$

Щоб перевірити, чи мають емпіричні вибіркові дані логарифмічно-нормальний розподіл, потрібно обчислити логарифм від кожного елемента вибірки і перевірити її на нормальність, наприклад за допомогою критерію  $\chi^2$ . Якщо трансформований набір даних має нормальний розподіл, то вхідні дані є логарифмічно-нормально розподіленими.

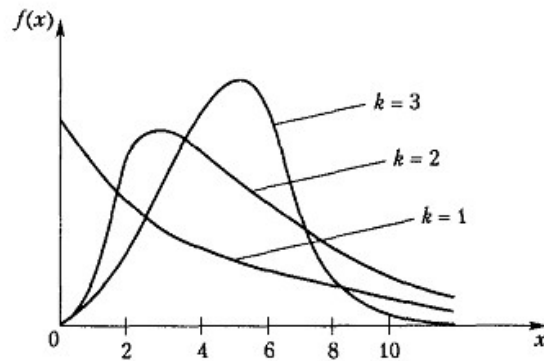
#### 4.5.7. Розподіл і потоки Ерланга

Випадкові величини з експоненціальним розподілом не завжди адекватно описують деякі реальні процеси та події, наприклад час обслуговування і моменти надходження вимог до СМО. Для більш точного моделювання таких процесів доцільніше використовувати гамма-розподілені випадкові величини або ті, що мають розподіл Ерланга. Розподіл Ерланга є результатом підсумовування взаємно незалежних і однаково розподілених експоненціальних випадкових величин і є окремим випадком гамма-розподілу.

Функція щільності розподілу Ерланга  $k$ -го порядку з інтенсивністю  $\lambda$  має такий вигляд:

$$f_{\eta}(x) = \frac{\lambda(\lambda x)^{k-1}}{(k-1)!} e^{-\lambda x}, \quad x \geq 0.$$

Математичне сподівання і дисперсія розподілу Ерланга визначаються як  $M[x] = 1/k\lambda$ ,  $D[x] = 1/k\lambda^2$ . Графік функції щільності розподілу Ерланга зображено на рис. 4.18.



**Рис. 4.18.** Графік функції щільності розподілу Ерланга