ЗАВДАННЯ ДЛЯ ЛАБОРАТОРНИХ РОБІТ

Лабораторна робота №1
ГЕНЕРУВАННЯ ВИПАДКОВИХ ЧИСЕЛ З РІЗНИМИ ЗАКОНАМИ РОЗПОДІЛУ

Мета роботи: навчитися складати програми для генерування випадкових величин з різними законами розподілу.

Теоретичні відомості

Випадковою називається величина, значення якої є невідомим. В процесі проведення експерименту вона може прийняти те, чи інше значення, яке є невідомим наперед. Прикладами випадкових величин є напруга теплового шуму в електричних колах, кількість викликів, які надійшли від абонентів протягом доби, тощо. Значення випадкової величини є невідомим за означенням, проте в теорії ймовірності та статистики запропоновано різні параметри та методи для визначення кількісних характеристик випадкового процесу. За рахунок цього вказати точне значення випадкової величини неможливо, проте можна передбачити певний діапазон її появи. Кількість спостережень (експериментів) в ході яких одержується вибірка випадкового процесу безпосередньо впливає на точність його опису. Чим більшою є вибірка тим точніше можна описати статистичні характеристики випадкового процесу.

Дискретною називається випадкова величина, яка приймає дискретну множину значень (x_1, x_2, ... x_N). Дискретна випадкова величина може описуватися такою таблицею:

$$\begin{matrix} x_1 & x_2 & \dots & x_N \\ p_1 & p_2 & \cdots & p_N \end{matrix} \tag{1.1}$$

де, x_1, x_2, ... x_N — значення випадкової величини;
p_1, p_2, ... p_N - ймовірності появи значень випадкової величини;
N — кількість експериментів.

Таблицю значень випадкової величини та їхніх ймовірностей появи називають законом розподілу випадкової величини. Числа x_1, x_2, ... x_N можуть бути довільними, проте до ймовірності їхньої появи ставляться такі вимоги:
• ймовірність появи значення випадкової величини повинна бути додатною:

$$p_i \geqslant 0 \tag{1.2}$$

• сукупна ймовірність появи значень випадкової величини повинна дорівнювати одиниці, тобто випадкова величина повинна обов'язково прийняти одне із значень:

$$p_1 + p_2 + \ldots + p_N = 1 \tag{1.3}$$

Густина розподілу показує ймовірність появи випадкової величина х під час проведення експериментів і описується функцією f(x).

Функцією розподілу випадкової величини х є функція F(x), яка визначає ймовірність тощо, що випадкова величина х прийме значення менше за х' (буде знаходитися в інтервалі [- ∞, х']):

$$F(x) = P(x < x') \tag{1.4}$$

Функція розподілу та густина розподілу є взаємопов'язаними. Для неперервної випадкової величини функція розподілу визначається формулою:

$$F(x) = \int_{-\infty}^{x} f(x) \cdot dx \tag{1.5}$$

Математичним сподіванням, або середнім значенням дискретної випадкової величини є сума всіх її значень помножена на ймовірність їхньої появи. Дисперсія випадкової величини - це математичне сподівання квадрату її відхилення відносно середнього значення.

Методи генерування випадкових величин

Нормальний закон розподілу

Існують різні закони розподілу випадкових величин. В теорії ймовірності найбільше значення має нормальний закон розподілу випадкових величин, внаслідок кількох причин.

• Згідно центральної граничної теореми, яка стверджує, що закон розподілу суми окремих значень реалізацій деякого випадкового процесу є близьким до нормального. Причому, при збільшенні кількості реалізацій процесу та досліджуваної вибірки випадкової величини цей закон прямуватиме до нормального. Внаслідок цієї властивості нормальний закон розподілу випадкових величин знайшов широке застосування для опису різних випадкових процесів. Він використовується в математичних розрахунках для аналізу зміни розмірів деталей, шумів в радіотехніці, економічних показників, тощо.

• Нормальний закон розподілу піддається аналітичному опису, за рахунок чого можна отримати результати в явному вигляді:

$$f(x)=\frac{1}{\sqrt{2\cdot\pi\cdot\sigma^2}}\cdot\int_{-\infty}^{x} e^{-\frac{(x-M(x))^2}{2\cdot\sigma^2}} \qquad (1.6)$$

Для нормального закону розподілу є характерним “правилом трьох сигм”. Суть цього правила полягає в тому, що якщо випадкова величина розподілена за нормальним законом розподілу, тоді близько 99% всіх значень гарантовано потрапляють в інтервал від -3·σ до +3·σ. На практиці це означає, що ймовірність появи значень величина яких виходить за межі трьох сигм рівна нулеві.

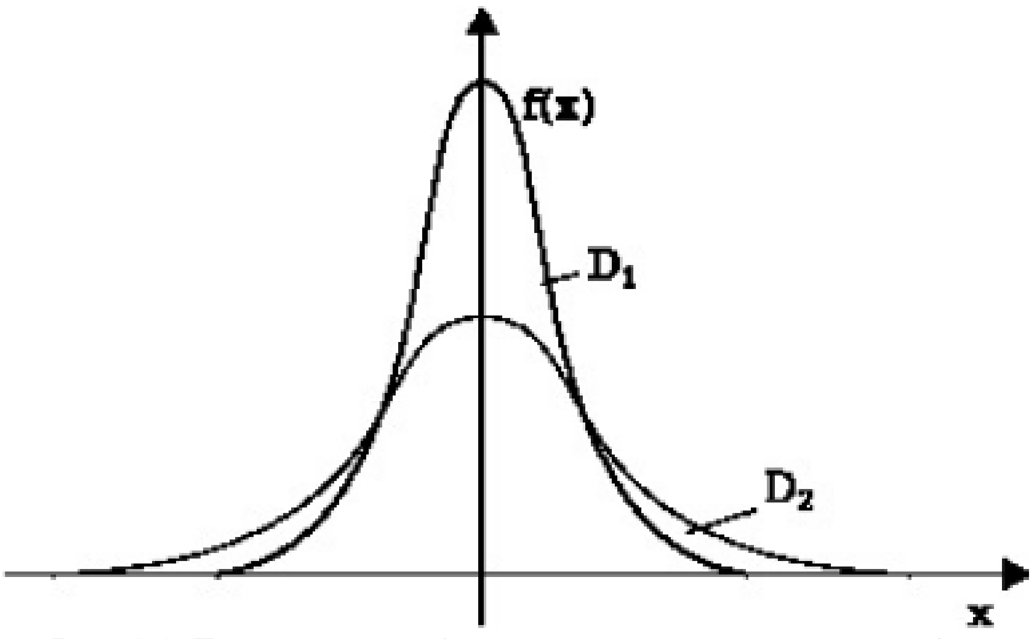

Рис. 1.1. Густина розподілу нормального закону розподілу

Є відомими кілька методів генерування випадкових чисел з нормальним законом розподілу.

• Використання центральної граничної теореми. Для отримання випадкових величин з нормальним законом розподілу необхідно отримати 12 реалізацій випадкового процесу з рівномірним законом розподілу на ділянці (0, 1), просумувати їх і відняти 6 реалізацій цього ж процесу. Отримані випадкові величини будуть мати закон розподілу, який наближається до нормального. Цей метод генерування випадкових величин має малу точність, внаслідок чого необхідно використовувати інші методи.

• Перетворення Бокса-Мюллера - метод моделювання стандартних нормально розподілених випадкових величин запропонований ще у 19 столітті. Цей метод має дві реалізації. Метод є точним, на відміну від методів, що базуються на використанні положень центральної граничної теореми. Перший варіант реалізації методу передбачає використання двох змінних U та V, значення яких повинні бути рівномірно розподілені на ділянці (0, 1). А значення нових змінних X, Y які визначаються на основі формул (7), (8) матимуть нормальний закон розподілу та будуть незалежними між собою.

$$X = \sqrt{-2 \cdot \ln(U)} \cdot \cos(2 \cdot \pi \cdot V) \tag{1.7}$$

$$Y = \sqrt{-2 \cdot \ln(U)} \cdot \sin(2 \cdot \pi \cdot V) \tag{1.8}$$

Існує також удосконалена версія цього методу, запропонована Д. Марсаглі. Згідно методу Д. Марсаглі для генерування випадкових величин з нормальним законом розподілу використовувати тригонометричні функції sin() та cos() непотрібно. Спочатку генеруються дві змінні U та V з рівномірним законом розподілу, але вже на інтервалі (-1, 1) і визначається нова змінна S:

$$S = U^2 + V^2 \tag{1.9}$$

Якщо S ≥ 1, відбувається повторне генерування змінних U і V, а в протилежному випадку згідно формул (10), (11) визначається значення змінних X та Y, які будуть розподілені за нормальним законом розподілу.

$$X = U \cdot \sqrt{\frac{-2 \cdot \ln(S)}{S}} \tag{1.10}$$

$$Y = V \cdot \sqrt{\frac{-2 \cdot \ln(S)}{S}} \tag{1.11}$$

Робота циклу відбувається до моменту настання заданої умови, а далі обчислюється значення X, або Y. Після отримання стандартних нормально розподілених випадкових величин X, або Y, за допомогою перетворення Бокса-Мюллера, остаточні значення випадкової величини визначаються з врахуванням середньоквадратичного відхилення.

З точки зору програмування, процес генерування змінної з випадковим значенням за допомогою перетворення Бокса-Мюллера відбувається в одному циклі.

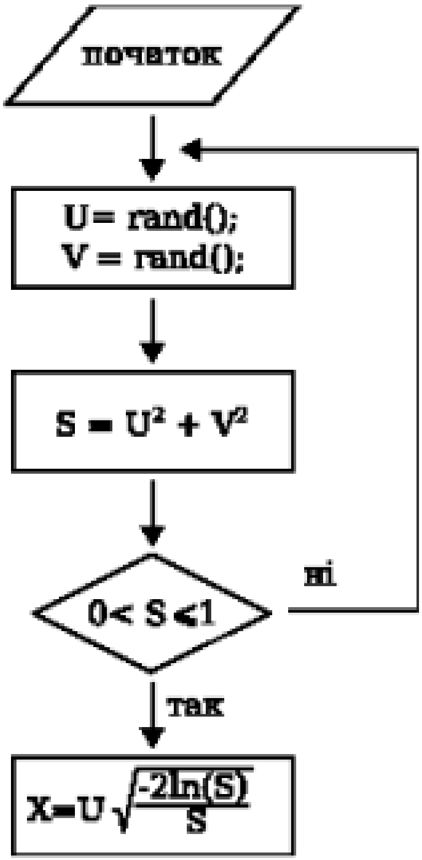

Рис. 1.2. Блок-схема алгоритму генерування випадкових чисел за допомогою удосконаленого перетворення Бокса-Мюллера

Розподіл Пуассона

Іншим прикладом законів розподілу випадкових величин є розподіл Пуассона. Математичну модель розподілу було запропоновано Пуассоном ще в 1830-х роках, для моделювання надходження потоків викликів на телефонну станцію. Розподіл також можна використовувати для моделювання різних випадкових подій, які відбулися за фіксований проміжок часу з деякою середньою інтенсивністю λ та незалежно одна від одної. Цей закон розподілу володіє цікавою властивістю. Дисперсія випадкової величини, розподіленої за законом Пуассона, дорівнює її математичному сподіванню. Цю властивість часто використовують на практиці для перевірки гіпотези про те, чи випадкова величина розподілена саме за законом Пуассона. Для цього на основі проведених експериментів визначають статистичні характеристики випадкової величини - математичне сподівання і дисперсію. Якщо ці значення близькі, це є підставою вважати, що значення випадкової величини розподілені згідно розподілу Пуассона.

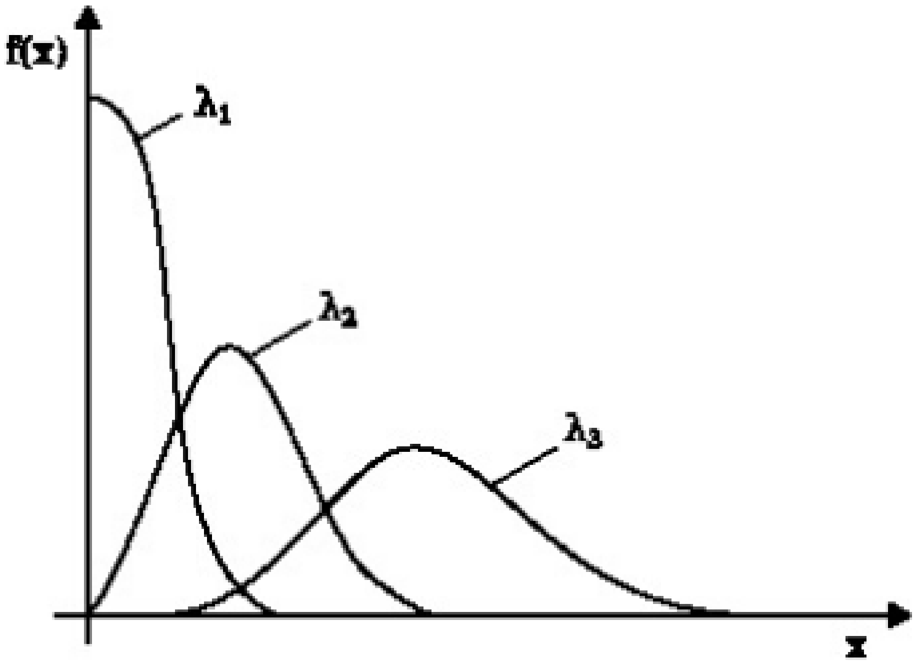

Рис. 1.3. Густина розподілу Пуасона

На практиці цей тип розподілу найчастіше використовується для моделювання надходження викликів від абонентів до автоматичних телефонних станцій із заданою інтенсивністю (математичне сподівання чи дисперсія). Математична модель закону розподілу Пуассона є такою:

$$f(x)=\frac{\lambda^x \cdot e^{-\lambda}}{x!} \tag{1.12}$$

де x — значення випадкової величини x > 0;
λ — параметр, який задає інтенсивність.

Методику (алгоритм) генерування випадкових величин з розподілом Пуассона було запропоновано Дональдом Кнутом (Knuth D.), відомим науковцем в області комп'ютерних наук. Цей алгоритм представлено на рис. 1.4.

У запропонованому алгоритмі для генерування випадкових величин використовується кілька змінних (L, k, p, u), значення яких визначається в циклі. Коли досягнуто виконання умови, робота циклу завершується та повертається значення змінної k.

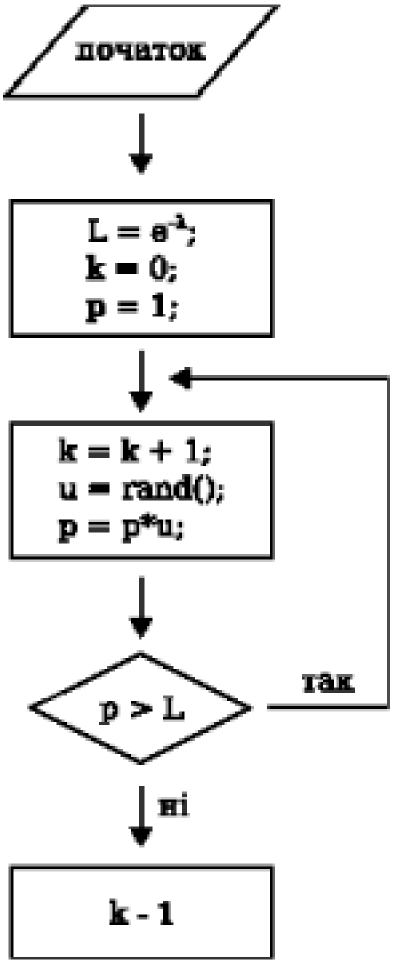

Рис. 1.4. Блок-схема алгоритму генерування випадкових величин з розподілом Пуассона

Розподіл Релея

Одним із найважливіших є закон розподілу випадкових величин Релея. Він був запропонований лордом Релеєм у 1880-х роках для опису різних процесів в акустиці та оптиці. Пізніше розподіл Релея почав використовуватися в теорії електричного зв'язку для визначення співвідношення між сигналом і шумом. З розвитком засобів передавання радіосигналів та появою систем радіозв'язку цей закон розподілу почали застосовувати для дослідження та моделювання поширення, випромінювання і прийому радіохвиль. Закон розподілу Релея також використовується для моделювання роботи радіолокаційних систем.

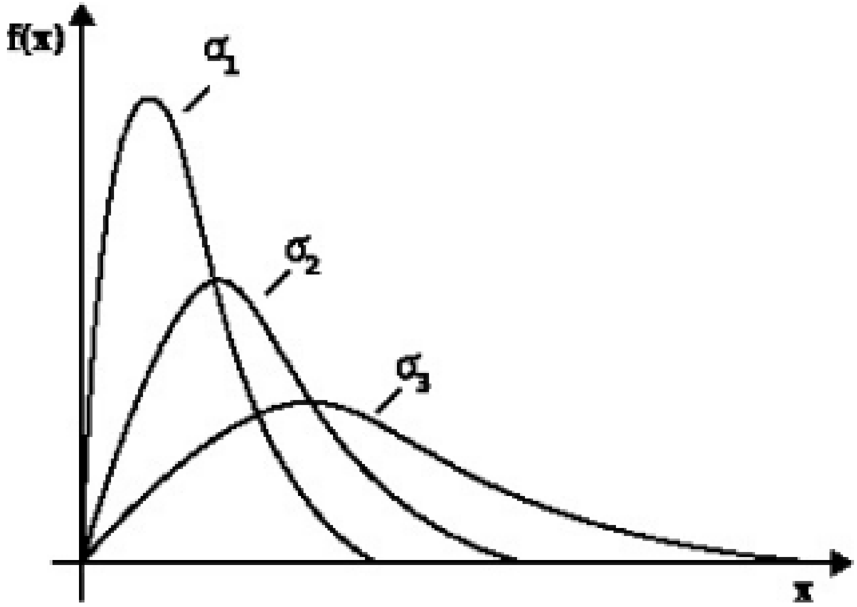

Рис. 1.5. Густина розподілу Релея

Математична модель закону Релея визначаються такою формулою:

$$f(x)=\frac{x}{\sigma^2}\cdot e^{-\frac{x^2}{2\cdot\sigma^2}} \quad (1.13)$$

де x — значення випадкової величини, x ≥ 0;
σ — середньоквадратичне відхилення, σ > 0.

Генерування випадкових величин з розподілом Релея полягає у виконанні кількох операцій. Спочатку створюємо змінну u, значення якої повинне бути рівномірно розподілено на інтервалі (0, 1), а далі визначаємо остаточне значення змінної x.

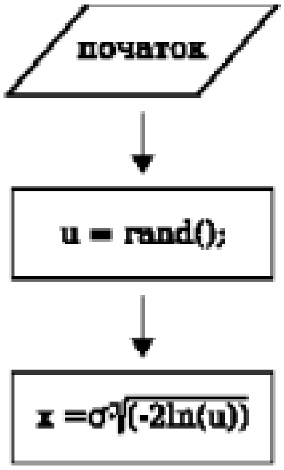

Рис. 1.6. Блок-схема алгоритму генерування випадкових величин з розподілом Релея

Яскравим прикладом застосування цього розподілу є створення математичної моделі каналу Релея, яка використовується для дослідження багатопроменевого поширення радіосигналів.

Порядок роботи

1. Запустити середовище розробки програм на мові C/C++ (BorlandC, GCC, MinGW, Dev-C++, Visual Studio, тощо).
2. Вибрати із залікової книжки дві останні цифри **m** (передостання) та **n** (остання). Якщо будь-яка із цифр рівна нулеві замінити її на 1.
3. Скласти програму генерування випадкових величин з нормальним законом розподілу за допомогою удосконаленого перетворення Бокса-Мюллера.
4. Скласти програму генерування випадкових величин з законом розподілу Пуассона та заданою інтенсивністю:

$$\lambda = \frac{m}{n}$$

Контрольні запитання

1. Що таке випадкова величина?
2. Що містить таблиця значень випадкової величини?
3. Що таке закон розподілу випадкової величини?
4. Чому дорівнює сукупна ймовірність появи значень випадкової величини?
5. Методи генерування випадкових чисел з нормальним законом розподілу?
6. Удосконалене перетворення Бокса-Мюллера?
7. Які властивості нормального закону розподілу?
8. Генерування випадкових чисел з законом розподілу Пуассона?
9. Які властивості розподілу Пуассона?
10. За що відповідає параметр λ?

Зміст звіту

1. Титульний лист
2. Тема та мета роботи
3. Короткі теоретичні відомості
4. **Результати виконаної роботи**
5. **Висновок**

ЛІТЕРАТУРА

1. Березин Б.И., Березин С.Б. Программирование на С и С++ - М.: ДИАЛОГ-МИФИ, 2001. - 288 с.
2. Керниган Б., Ритчи Д.. Язык программирования С, 2-е издание - М.: Вильямс — 2009. - 292 с.
3. Кибзун А.И. Теория вероятности и математическая статистика. Базовый курс с примерами и задачами / Учебн. Пособие. - М.:ФИЗМАТЛИТ, 2002. -224 с.
4. Лафоре Л. Объектно-ориентирование программирование в С++, 4-е издание — М.: Питер, 2004. - 923 с.
5. Минашкин В.Г. Теория статистики: учебно-методический комплекс. - М.: Изд. Центр ЕАОИ. 2008. - 296 с.
6. Папас К., Мюррей У. Программирование на С и С++ - К.: Издательская група BHV, 2000. - 320 с.
7. Пугачев В.С. Теория вероятности и математическая статистика: Учеб. Пособие. - 2-е изд., исправл. и дополн. - М.:ФИЗМАТЛИТ, 2002. - 496 с.
8. Шилдт Г. Справочник программиста С/С++, 3-е изд.: Пер. с англ. - М. Издательский дом “Вильямс”, 2003. - 432 с.