Моделирование статистических распределений с использованием метода Монте-Карло

А. А. Маляр, email: alexei002@yandex.ru
А. И. Драбо, email: alexei002@yandex.ru
А. Е. Пигарев, email: alexei002@yandex.ru
А. С. Животворев, email: alexei002@yandex.ru

ВУНЦ ВВС «ВВА им. проф. Н.Е. Жуковского и Ю.А. Гагарина»
(г. Воронеж)

Аннотация. *Представляется алгоритм, позволяющий моделировать экспоненциальное и нормальное распределение случайной величины с использованием метода Монте-Карло. Рассмотрено его применение на модельном примере.*

Ключевые слова: *равномерное распределение, метод Монте-Карло, экспоненциальное распределение, нормальное распределение.*

Введение

Статистическое моделирование является мощным инструментом при проведении различного рода аналитических исследований. Оно позволяет перейти от натурных экспериментов по изучению процессов и явлений к их описанию различными статистическими распределениями. В основе этого моделирования лежит метод, получивший название метода Монте-Карло. Результаты процесса моделирования позволяют сделать выводы о поведении рассматриваемого процесса или явления в случае если нет возможности применять другие методы его исследования. Возможности современного пакета Microsoft Excel позволяют проводить большое число статистических испытаний, что вполне может быть использовано, для решения некоторых конкретных задач, например для расчета определенных интегралов.

1. Моделирование распределений

За основу для моделирования принято равномерное распределение случайной величины ($X_{\text{равн}}$) в прямоугольнике с основанием $(b-a)$ и шириной c . Функция плотности этого распределения имеет вид [1, 2]:

© Маляр А. А., Драбо А. И., Пигарев А. Е., Животворев А. С., 2021

$$f(X_{\text{равн}}) = \begin{cases} 0, & \text{при } X \le a, \\ 1/(b-a)c, & \text{при } a < X \le b, \\ 0, & \text{при } X > b. \end{cases} \quad (1)$$

Для задания равномерного распределения в Microsoft Excel с использованием редактора кода VBA создан макрос, содержание которого приведено в листинге 1.

Листинг 1

Задание равномерного распределения

```
Range("F4").Select
ActiveCell.FormulaR1C1 = "=RAND()"
Range("F4").Select
Selection.AutoFill Destination:=Range("F4:F1003"),
Type:=xlFillDefault
Range("F4:F1003").Select
ActiveWindow.SmallScroll Down:=-111
Range("G4").Select
ActiveCell.FormulaR1C1 = "=RAND()"
Range("G4").Select
Selection.AutoFill Destination:=Range("G4:G1003"),
Type:=xlFillDefault
Range("G4:G1003").Select
ActiveWindow.SmallScroll Down:=-111
End Sub
```

Использование макросов удобно при многократном повторении одной и той же процедуры. В тесте редактора макроса имеется возможность легко установить необходимое количество испытаний. Результатом работы данного макроса являются два ряда равномерно распределенной случайной величины r_i и R_i длиной по 1000 значений.

Для моделирования экспоненциально распределенной случайной величины ($X_{\text{эксп}}$) использована функция плотности [1, 2]

$$f(X_{\text{эксп}}) = \begin{cases} 0, & \text{при } X \le 0, \\ \lambda e^{-\lambda X}, & \text{при } X > 0. \end{cases} \quad (2)$$

где λ – постоянная положительная величина.

Тогда

$$F(X_{\text{эксп}}) = \int_0^X f(x)dx = 1 - e^{-\lambda X} \quad (3)$$

Если

$$R_{\text{эксп}} = 1 - e^{-\lambda X}, \tag{4}$$

то можно получить

$$X_{\text{эксп}} = -\frac{\ln\left|1 - R_{\text{равн}}\right|}{\lambda}, \tag{5}$$

где $R_{\text{равн}}$ – равномерно распределенные случайные числа, полученные в результате работы макроса, приведенного в листинге 1.

На рис. 1 показана гистограмма смоделированного по формуле (5) экспоненциального распределения для n = 1000 испытаний.

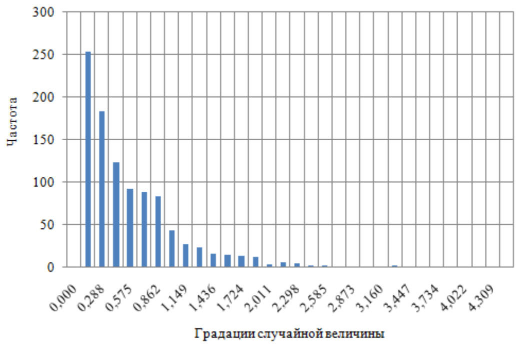

Рис. 1. Гистограмма экспоненциального распределения

Для получения распределения случайной величины, распределенной по нормальному закону $X_{\text{норм}}$, использована функция плотности распределения в виде [1, 2]

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \tag{6}$$

где $z = \dfrac{X_{\text{равн}} - M(X_{\text{равн}})}{\sigma(X_{\text{равн}})}$ – центрированная и нормированная случайная величина $X_{\text{равн}}$, $M(X_{\text{равн}})$ и $\sigma(X_{\text{равн}})$ – оценки математического ожидания и СКО равномерно распределенной случайной величины $X_{\text{равн}}$.

Интегрирование (6) позволяет перейти к функции распределения:

$$F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{t^2}{2}} dt = \Phi(z) + 0{,}5 \quad (7)$$

,

где $\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{t^2}{2}} dt$ – функция Лапласа, для определения которой использована встроенная функция Microsoft Excel.

Как и в случае с экспоненциально распределенной случайной величиной, примем соответствие

$$R_{\text{норм}} = F(z) = \Phi(z) + 0{,}5 \quad (8)$$

,

на основании которого можно записать:

$$X_{\text{норм}} = \sigma(X_{\text{равн}})\Phi^{-1}(R_{\text{норм}}) + M(X_{\text{равн}}) - 0{,}5 \quad (9)$$

.

На рис. 2 показана гистограмма смоделированного по формуле (9) нормального распределения для того же числа испытаний.

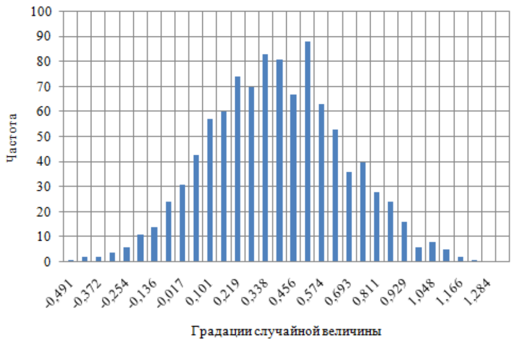

Рис. 2. Гистограмма нормального распределения

Использование выражений (5) и (9) позволяет получить случайные величины $X_{\text{эксп}}$ и $X_{\text{норм}}$, распределенные по экспоненциальному и нормальному законам распределения соответственно.

2. Оценка погрешности метода Монте-Карло при расчете определенных интегралов на модельном примере

Применение на практике результатов работы макроса и формул (5) и (9) делает возможным расчет определенных интегралов, аналитическое вычисление которых вызывает существенные затруднения, то есть первообразная у которых не является элементарной функцией. В качестве модельного интеграла выбран интеграл

$$I_1 = \int_0^2 (x^2 - x + 4)dx \quad (10)$$

,

истинное значение которого легко можно получить путем непосредственного интегрирования:

$$I_1 = \int_0^2 (x^2 - x + 4)dx = \frac{1}{3}x^3\Big|_0^2 - \frac{1}{2}x^2\Big|_0^2 + 4x\Big|_0^2 \approx 8,(6) \quad (11)$$

,

Формула для расчета определенного интеграла с использованием одной из разновидностей метода Монте-Карло имеет вид:

$$I^* = \frac{(b-a)}{n}\sum_{i=1}^{n} f\left[a + (b-a)r_i\right] \quad (12)$$

,

где a и b – пределы интегрирования, r_i – наборы случайных величин, распределенных по равномерному, экспоненциальному и нормальному законам распределения соответственно, полученные с использованием макроса и смоделированные по формулам (5) и (9).

В таблице приведены результаты расчета интеграла (10) и значения погрешностей как модулей разности между истинным и рассчитанными значениями.

Таблица

Рассчитанные значения интеграла (10) и погрешности расчета

Вид распределения	Число испытаний n			
	1000		5000	
	Значение интеграла	Погрешность	Значение интеграла	Погрешность
Равномерное	8,399	0,201	8,671	0,071
Экспоненциальное	8,656	0,056	9,946	1,346
Нормальное	8,080	0,520	8,262	0,338
Истинное значение	8,6(6)		8,6(6)	

Заключение

Проведенное исследование показало, что разработанный алгоритм позволяет моделировать экспоненциальное и нормальное распределение случайной величины с использованием метода Монте-Карло. Оценка погрешности метода Монте-Карло, полученная на модельном примере, говорит о возможности расчета определенных интегралов, используя не только равномерное, но и другие виды распределений. Значительное увеличение числа испытаний показывает, что получаемые значения интегралов имеют удовлетворительную погрешность, однако погрешность увеличивается, следовательно, увеличение числа испытаний теряет смысл. Полученный алгоритм можно использовать для решения более сложных интегралов, например от быстро осциллирующих функций.

Список литературы

1. Теория вероятностей и математическая статистика / под ред. В. И. Ермакова. – М. : ИНФРА-М, 2010. – 288 с.
2. Гмурман, В. Е. Теория вероятностей и математическая статистика : учебное пособие для бакалавров / В. Е. Гмурман; пер. с англ. П. А. Волковой. – 12-е изд-е. М. : Юрайт, 2013. – 480 с.