**Лекция 4. Моделирование непрерывных случайных величин.**

*Классификация методов моделирования непрерывных случайных величин*

Для моделирования случайных величин с заданным законом распределения, как это следует из основного принципа моделирования случайных объектов, необходимо преобразовать базовую случайную величину $\xi$. Можно выделить четыре направления такого преобразования: аналитическое, избирательное, вероятностное и комбинированное.

При аналитическом преобразовании над реализацией $z$ случайной величины $\xi$ осуществляется некоторая операция, формирующая число $x$, которое можно рассматривать как реализацию случайной величины $\eta$ с заданным законом распределения. Наибольшую известность здесь получил метод обратной функции.

Однако для ряда важных распределений, у которых закон распределения не выражается через элементарные функции, этот метод практически невозможно реализовать.

Суть следующего направления состоит в выборе чисел из базовой последовательности так, чтобы они образовали новую последовательность с требуемым распределением.

Широкое распространение среди "избирательных" методов получил метод исключения Неймана. К сожалению, этот метод характеризуется большим числом "холостых ходов" и применяется для моделирования случайных величин, которые определены на конечном отрезке $[a, b]$.

Третье направление связано с моделированием условий соответствующих предельных теорем теории вероятностей, обеспечивающих практически приемлемое приближение к заданным законам распределения. Очевидно, область применения этого направления ограничена числом предельных теорем.

Название четвертого направления предопределяет использование для моделирования случайных величин одновременно нескольких методов. Такой подход оправдан при моделировании непрерывных случайных величин с достаточно сложными законами распределения.

Ниже будут рассмотрены по одному методу из каждого направления. На наш взгляд, эти методы, дополняя друг друга, обеспечат моделирование непрерывных случайных величин с любым законом распределения, заданных как аналитически, так и в виде графиков или таблиц.

*Метод обратной функции*

Пусть непрерывная случайная величина $\eta$ определена в интервале $(a, b)$ и имеет плотность распределения $f(x) > 0$ при $a < x < b$ (случай $a = -\infty$ , $b = \infty$ не исключается).

Функция распределения $F(x) = \int\limits_a^x f(x)dx.$

Принцип работы метода обратной функции сформулируем в виде теоремы 3. Случайная величина $\eta$, реализации $x$ которой определяются из выражения

$$F(x) = z \text{ или } x = F^{-1}(z) \tag{14}$$

где $z$ – реализация базовой случайной величины $\xi$, имеет плотность распределения $f(x)$.

Для практического применения метода обратной функции необходимо разрешить относительно xj уравнение

$$\int\limits_a^{x_j} f(x)dx = z_j \,. \tag{15}$$

Алгоритм, реализующий метод обратной функции, состоит из следующих процедур:

Шаг 1. Положить $j = 1$.

Шаг 2. Получить реализацию $z_j$ случайной величины $\xi$.

Шаг 3. Вычислить реализацию $x_j$ случайной величины $\eta$ $x_j = F^{-1}\left(z_j\right)$.

Шаг 4. Положить $j = j + 1$.

Шаг 5. Проверить выполнение условия $j > n$, где $n$ – требуемое число реализаций случайной величины $\eta$. При нарушении этого условия переход на шаг 2.

Шаг 6. Вывод значений $\{x_j\}$.

*Метод исключения Неймана*

Суть метода исключения, предложенного Джон фон Нейманом, заключается в том, что из равномерно распределенной базовой последовательности исключается часть чисел таким образом, чтобы оставшиеся числа подчинялись заданному закону.

Пусть случайная величина $\eta$ определена на отрезке $[a, b]$ и имеет ограниченную сверху функцию плотности $f(x) < M$ (рисунок 4).

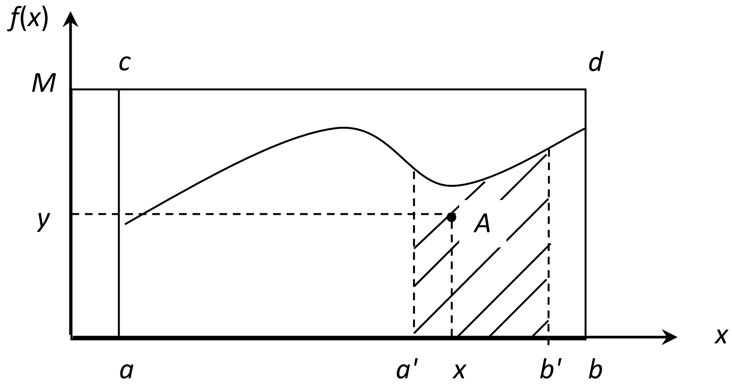


Рисунок 4

Сформулируем теорему 4. Пусть $z_1$ и $z_2$ – независимые реализации базовой случайной величины $\xi$ и

$$x = a + z_1(b - a),\ \ y = Mz_2. \tag{16}$$

Тогда случайная величина $\eta$, определенная из условия

$$\eta = x \ \text{ при } \ y < f(x) \tag{17}$$

имеет плотность распределения $f(x)$.

Для практической реализации метода исключения можно предложить следующий алгоритм:

Шаг 1. Положить $i = 1, j = 1$.

Шаг 2. Получить две независимые реализации $z_{2j-1}$, $z_{2j}$ случайной величины $\xi$.

Шаг 3. Вычислить $x_j = a + z_{2j-1}(b - a)$ и $y_j = Mz_{2j}$.

Шаг 4. Проверить условие $y_j < f\left(x_j\right)$. При его нарушении перейти на шаг 6.

Шаг 5. Положить $x_i = x_j$ и $i = i + 1$.

Шаг 6. Положить $j = j + 1$.

Шаг 7. Проверить выполнение условия окончания счета $i > n$. При нарушении этого условия переход на шаг 2.

Шаг 8. Вывод $\{x_i\}$.

Эффективность метода исключения Неймана прямо пропорциональна вероятности попадания точки $A(x, y)$ под кривую $y = f(x)$, т.е.

$$P\{y < f(x)\} = \frac{1}{M(b - a)}.$$

Следовательно, эффективность метода будет наибольшей, если выбрать наименьшее возможное значение $M$ , т.е. принять

$$M = \sup_{a \le x \le b} f(x).$$

Важным достоинством метода исключения Неймана является возможность задания закона распределения случайной величины как аналитически, так и графически.

К сожалению, этот метод характеризуется большим числом "холостых ходов" и применяется для моделирования случайных величин, которые определены на конечном отрезке $[a,b]$.

Таким образом, применяя выше предложенные методы можно смоделировать наиболее часто встречающиеся непрерывные распределения.

*Метод предельных теорем*

Этот метод моделирования случайных величин основывается на приближенном воспроизведении условий, при которых оказываются справедливыми соответствующие предельные теоремы. Так, центральная предельная теорема теории вероятностей позволяет получить случайную величину с нормальным законом распределения. Эта теорема впервые была сформулирована Лапласом. Ее обобщением занимались многие выдающиеся математики, в том числе Чебышев П.Л., Марков А.А. и Ляпунов А.М.

Приведем центральную предельную теорему в следующей формулировке:

Теорема 5. Пусть $\tau_1, \tau_2, ..., \tau_n$ – взаимно независимые нормированные случайные величины с одним и тем же распределением. Тогда при $n \to \infty$ распределение нормированной суммы

$$\eta^H = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \tau_i \tag{18}$$

стремится к нормальному распределению с плотностью вероятностей

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

Если в качестве независимых случайных величин $\tau_i$ использовать базовые случайные величины $\xi_i$ с математическим ожиданием $m_z$ = 1/2 и дисперсией $\sigma_z^2$ = 1/12, то нормированная сумма (18) примет вид

$$\eta^H = \sqrt{\frac{3}{n}} \sum_{i=1}^{n} (2\xi - 1). \tag{19}$$

Следовательно, по формуле (19) при достаточно больших $n$ можно вычислить приближенные значения нормального распределения с параметрами $m$ = 0 и $\sigma^2$ = 1. Проведенные исследования показали, что уже при $n$ = 12 погрешность суммы (19) не превышает $9*10^{-3}$. Поэтому на практике для моделирования нормального закона распределения с заданными параметрами $m_x$ и $\sigma_x^2$ используется формула

$$x = m_x + \sigma_x \left( \sum_{i=1}^{12} z_i - 6 \right). \tag{20}$$

Здесь $z$ и $x$ – реализации соответственно базовой ($\xi$) и моделируемой ($\eta$) случайных величин.

Алгоритм, реализующий формулу (20), заключается в следующем:

Шаг 1. Положить $j$ = 1.

Шаг 2. Принять $S$ = 0 и $i$ = 1.

Шаг 3. Получить реализацию $z$ случайной величины $\xi$.

Шаг 4. Положить $S$ = $S$ + $z$ и $i$ = $i$ + 1.

Шаг 5. Проверить условие $i \le 12$. При выполнении условия перейти на шаг 3.

Шаг 6. Вычисление очередной реализации $x_j$ случайной величины $\eta$ $x = m_x + \sigma_x(S - 6)$

.

Шаг 7. Положить $j = j + 1$.

Шаг 8. Проверить условие завершения счета $j > n$, где $n$ – число реализаций нормального распределения. При нарушении этого условия возврат на шаг 2.

Шаг 9. Вывод реализации $\{x_j\}$.

*Метод композиции*

Если функция распределения $F(x)$ некоторой случайной величины $\eta$ имеет достаточно сложный вид, то во многих случаях ее можно представить как композицию более простых распределений

$$F(x) = \sum_{k=1}^{m} C_k * F_k(x) \tag{21}$$

где $C_k > 0$. Из (21) при $x \to \infty$ также следует, что $\sum_{k=1}^{m} C_k = 1$. Следовательно, можно ввести полную группу событий $\{A_k\}$ заданной таблицей

$$\begin{pmatrix} A_1 & A_2 \ldots & A_m \\ C_1 & C_2 \ldots & C_m \end{pmatrix},$$

где $C_k = P(A_k)$.

Теорема 6. Пусть $\xi_1$ и $\xi_2$ – независимые базовые случайные величины. Если с помощью $\xi_1$ определить номер наступившего события $A_k$, а затем из уравнения $F_k(\eta) = \xi_2$ получить $\eta$, то функция распределения $\eta$ равна $F(x)$.

Для практической реализации метода композиции удобнее пользоваться функцией плотности распределения $f(x)$ моделируемой случайной величины $\eta$

$$f(x) = \sum_{k=1}^{m} C_k * f_k(x) \tag{22}$$

Алгоритм, реализующий условия теоремы 6, состоит из семи шагов:

Шаг 1. Положить $j = 1$.

Шаг 2. Получить реализации $z_{2j-1}$, $z_{2j}$ случайной величины $\xi$.

Шаг 3. С помощью $z_{2j-1}$ разыграть событие $A_k$.

Шаг 4. Получить реализацию $x_j$, соответствующую функции плотности $f_k(x)$.

Шаг 5. Положить $j = j + 1$.

Шаг 6. Проверить выполнение условия $j > n$, где $n$ – требуемый объем реализаций случайной величины $\eta$. При нарушении этого условия возврат на шаг 2.

Шаг 7. Вывести полученные реализации.

**Контрольные вопросы**:

1. В каком случае применяется метод композиции?
2. Сформулируйте теорему, применяемую при моделировании непрерывной случайной величины методом исключения Неймана.
3. Какой метод моделирования применяется в случае задания функции плотности непрерывной величины графически?