5 Теория. Генерация случайных величин, метод Монте-Карло.

5.1 Метод обратной функции

Начнем с "прямого"метода. Для генерации величины X с функцией распределения $F(x)$ нужно взять величину Y, имеющую равномерное распределение на отрезке $[0,1]$, и положить $X = F^{-1}(Y)$. Этот метод называется *методом обратной функции.*

Этот принцип работает для любой функции $F(x)$, включая разрывные, при этом обратная функция задается следующим образом:

$$F^{-1}(y) = \inf\{x : F(x) \geqslant y\}.$$

Итак, чтобы сгенерировать случайную величину с распределением $F(x)$ требуются:

1) датчик равномерных $R[0,1]$ чисел,

2) функция G, обратная к функции распределения F.

Пример. Функция распределения бернуллиевской случайной величины с параметром p имеет вид

$$F(x) = \begin{cases} 0, & x < 0, \\ 1-p, & x \in [0,1), \\ 1, & x \geq 1. \end{cases}$$

Тогда определим обратную функцию

$$F^{-1}(y) = \begin{cases} 0, & y \in [0, 1-p], \\ 1, & y \in (1-p, 1] \end{cases}$$

Следовательно, рассматривая случайную величину $Y \sim R[0,1]$ и полагая $X = 1$ при $Y > 1-p$ и $X = 0$ при $Y \leqslant 1-p$, мы получим бернуллиевскую величину.

5.2 Смеси распределений.

Если заданную функцию распределения можно представить в виде $F(x) = \sum p_i F_i(x)$, где $F_i(x)$ – функции распределения каких-то распределений, $p_i > 0$, $\sum p_i = 1$, то такое распределение называют *смесью* соответствующих распредлений с весами $p_1, p_2, \ldots$ (аналлогичное представление в виде суммы верно и для плотностей/ дискретных распределений).

Такое распределение можно генерировать следующим образом. Пусть X_i – случайная величина с i-м распределением, с.в. Y принимает значения $1, 2, \ldots$ с вероятностями $p_1, p_2, \ldots$ (если количество распредлений в смеси конечно, то $Y \sim Polinomial(k, p_1, \ldots, p_k)$. Тогда искомую с.в. X можно преставить в виде $X = \sum X_i \, \mathrm{I}(Y = i)$.

Отсюда получаем алгоритм:
1) гененируем значение дискретной с.в. Y,
2) генерируем значение с.в. X_Y.

Пример. Пусть нам надо сгенерировать с.в. X с распределением

$$\mathbf{P}(X = k) = \frac{\lambda_1^k e^{-\lambda_1} + 2\lambda_2^k e^{-\lambda_2}}{3k!}.$$

В этом случае $X_i \sim Poisson(\lambda_i)$, $p_1 = 1/3$, $p_2 = 2/3$.

Алгоритм: бросаем монетку $Y \sim Bern(1/3)$. Если $Y = 1$, генерируем $X \sim Poisson(\lambda_1)$, если $Y = 0$, то генерируем $X \sim Poisson(\lambda_2)$.

5.3 Метод выбора с отклонением

На практике метод обратной функции может оказаться весьма трудозатратным (часто более трудоемкой операцией является подсчет обратной функции). Существуют и другие методы, например, так называемый метод выборки с отклонением (в англоязычной литературе Rejection Sampling или Acceptance-Rejection algorithm). Опишем вкратце идею метода.

Предположим, что X имеет плотность $f(x)$ (или дискретное распределение $f(x) = \mathbf{P}(X = x)$, а мы умеем генерировать величину Y с плотностью $g(x)$ (вероятностью $g(x) = \mathbf{P}(X = x)$), причем выполнено соотношение

$$f(x)/g(x) \leqslant c.$$

Тогда мы можем воспользоваться следущим алгоритмом.

1) Сгенерировать величину Y с плотностью (распределением) $g(x)$.

2) При $Y = y$ с вероятностью $f(y)/(cg(y))$ положить $X = y$, иначе повторить 1), 2).

Пример. Для генерирования случайной величины со значением от 1 до 7 мы можем трижды бросить монетку и сопоставить полученным исходам числа от 1 до 8. Тогда $f(x) = 1/7$ при $x = 1, \ldots, 7$, $g(x) = 1/8, x = 1, \ldots, 8$.

Следовательно,

$$f(x) \leqslant 8g(x)/7$$

при любом x. Таким образом, мы должны сгенерировать величину Y с помощью трех бросков монеты, если полученное значение y будет из набора $\{1, \ldots, 7\}$, то мы оставляем y с вероятностью $f(y)/(cg(y) = 1$, а если $y = 8$, то мы оставляем y с вероятностью $f(y)/(cg(y) = 0$.

Иначе говоря, при выпадении $1, \ldots, 7$ мы заканчиваем эксперимент с таким результатом, а при выпадении 8 перебрасываем три броска монеты.

5.4 Многомерномый случай. Метод условных распределений.

Зная способы генерации случайных величин, возможно сгенерировать вектор с независимыми компонентами, имеющими заданные функции распределения. Для этого достаточно иметь датчик независимых $R[0,1]$ величин. Но моделирование вектора, чьи компоненты зависимы, требует, очевидно, другого подхода.

В основе метода лежит условное распределение. Для получения вектора $(X_1, \ldots, X_n)$ мы:

1) Производим генерацию X_1 одним из известных нам методов, на данный момент, используем $F^{-1}(R_1)$, где $R_1 \sim R[0,1]$ (метод обратной функции);

2) Рассматриваем X_1 как фиксированный параметр x_1 и производим генерацию X_2 на основе условной функции распределения $G(x) = F_{X_2|X_1}(x|x_1) = P\left(X_2 \leq x | X_1 = x_1\right), X_2 = G^{-1}\left(R_2\right)$ где $R_2 \sim R[0,1]$;

3) Рассматриваем X_1, X_2 как на фиксированные параметры x_1, x_2 и производим генерацию X_3 как $H^{-1}(R_3)$, где $H(x) = F_{X_1|X_1,X_2}(x|x_1, x_2)$ и так далее.

Пример. Хотим смоделировать вектор с совместной плотностью

$$f_{X,Y}(x,y) = \begin{cases} e^{-y} & 0 < x < y \\ 0, & \text{иначе.} \end{cases}$$

Найдем сперва плотность $f_X(x)$ и ф.р. $F_X(x)$. Плотность компоненты X двумерного вектора (X,Y) находится путем интегрирования по второй переменной: при $x > 0$ имеем

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy = \int_{x}^{\infty} e^{-y}dy = e^{-x}, F_X(x) = 1 - e^{-x}.$$

Следовательно, X генерируется как $-\ln(1-R_1)$, где $R_1 \sim R[0,1]$.

Можем брать $-\ln R_1$, поскольку $1-R_1$ распределена так же, как R_1.

Условная плотность $Y|X$ при $0 < x < y$ равна

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = e^{-y}/e^{-x} = e^{x-y},$$

откуда

$$F_{Y|X}(y|x) = \int_x^y e^{x-z} dz = 1 - e^{-(y-x)},$$

при $0 < x < y$.

Следовательно, получив значение x величины $-\ln R_1$, мы можем произвести генерацию Y как $x - \ln R_2$, где $R_2 \sim R[0,1]$ и не зависит от R_1.

Таким образом, вектор $(-\ln R_1, -\ln(R_1 R_2))$ будет иметь требуемое распределение.