

Практикум по специальности.

Весенний семестр 2019/2020 учебного года.

Теория.

1.1 Генерация случайных величин, метод Монте-Карло.

1.1.1 Метод обратной функции

Начнем с известного метода, который известен многим. 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, но необходимо правильно определить обратную функцию:

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

Таким образом, 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 $Y \leq 1 - p$, мы получим бернуллиевскую величину.

1.1.2 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:

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

Тогда мы можем:

- 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 $x = 1, \dots, 7$, $g(x) = 1/8, x = 1, \dots, 8$.

Следовательно,

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

при любом x . Таким образом, мы должны сгенерировать величину Y с помощью трех бросков монеты, если полученное значение y будет из набора $\{1, \dots, 7\}$, 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 $1, \dots, 7$ мы заканчиваем эксперимент с таким результатом, а при выпадении 8 перебрасываем три броска монеты.

1.1.3 Многомерный 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, \dots, 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 $G(x) = F_{X_2|X_1}(x|x_1) = P(X_2 \leq x|X_1 = x_1)$, $X_2 = G^{-1}(R_2)$ 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 $H(x) = F_{X_3|X_1, X_2}(x|x_1, x_2)$ 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 также 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.

1.1.4 Метод Монте-Карло. Вычисление интегралов.

В широком смысле под методом Монте-Карло понимается численный метод решения математической задачи при помощи псевдослучайных чисел. Метод этот получил достаточное распространение в силу того, что в ряде практических задач, не являющихся непосредственно случайными, бывает удобно получать приближенный ответ с помощью вероятностных методов.

Название метода происходит от города Монте-Карло в княжестве Монако, знаменитого своими игорными домами. Одним из классических примеров задач такого рода служит приближенный подсчет интеграла.

Нужно вычислить приближенно интеграл $I = \int_a^b f(x)dx$, где f — заданная функция.

Наиболее простым методом является метод прямоугольников, предлагающий заменить интеграл на сумму

$$\sum_{k=1}^n f(\tilde{x}_k) (x_{k+1} - x_k) = \frac{(b-a)}{n} \sum_{k=1}^n f(\tilde{x}_k),$$

где x_k — разбиение отрезка $[a, b]$ на равные части, а $\tilde{x}_k = (x_{k+1} + x_k)/2$ — отмеченные точки. Погрешность такого метода $O(1/n)$ для любой гладкой f и $O(1/n^2)$ для дважды гладкой f .

Метод Монте-Карло предлагает взамен взять

$$\hat{I} = \frac{b-a}{n} \sum_{k=1}^n f(X_k),$$

где X_k — н.о.р. $R[a, b]$. То есть, в качестве "узлов" используются псевдослучайные числа.

В силу ЗБЧ, примененной к $f(X_k)$, видим, что \hat{I} сходится к

$$(b-a)Ef(X) = (b-a) \int_a^b f(x) \frac{1}{b-a} dx = I,$$

причем в силу ЦПТ \hat{I} будет асимптотически нормальной оценкой для I , т.е.

$$\sqrt{n} \frac{\hat{I} - I}{\sigma} \xrightarrow{d} Z \sim \mathcal{N}(0, 1),$$

где $\sigma^2 = (b-a)^2 \mathbf{D}f(X)$. Таким образом, погрешность метода Монте-Карло будет порядка $O(n^{-1/2})$.

На данной задаче мы не видим особенных преимуществ замены детерминированных точек \tilde{x}_k на случайные X_k — точки попадают более разреженно и вместо возможной погрешности $O(n^{-2})$ мы получаем погрешность $O(n^{-1/2})$, да еще и случайную.

Но, во-первых, не налагается никаких условий на гладкость или даже непрерывность f , нужна лишь ЦПТ, то есть интегрируемость квадрата функции по Лебегу.

Во-вторых, выгода становится заметной, когда рассматривается многомерную задачу

$$\int_A f(x_1, \dots, x_k) dx_1 \dots dx_k.$$

Тогда метод прямоугольников для той же погрешности $O(n^{-2})$ заставляет брать разбиение на n^k кубиков, т.е. сделать порядка $O(n^k)$ операций суммирования.

В то же время при оценивании погрешности метода Монте-Карло мы не опираемся на размерность и для той же погрешности $O(n^{-2})$ понадобится $O(n^4)$ операций суммирования.

1.1.5 Метод Монте-Карло. Общая концепция и вычисление погрешности.

Допустим, что нам требуется вычислить какую-то неизвестную величину m . По сути, при реализации метода Монте-Карло, мы пытаемся придумать такую случайную величину ξ , чтобы $\mathbf{E}\xi = m$. Пусть при этом $\mathbf{D}\xi = b^2$.

Рассмотрим N независимых случайных величин $\xi_1, \xi_2, \dots, \xi_N$, распределения которых совпадают с распределением ξ .

Если N достаточно велико, то согласно центральной предельной теореме распределение суммы $S_N = \sum_i \xi_i$ будет стремиться к нормальному с параметрами $\mathbf{E}S_N = N \cdot m$, $\mathbf{D}S_N = N \cdot b^2$.

Из ЦПТ также не трудно получить соотношение:

$$P\left(\left|\frac{S_N}{N} - m\right| \leq k \frac{b}{\sqrt{N}}\right) = P\left(\left|\frac{1}{N} \sum_i \xi_i - m\right| \leq k \frac{b}{\sqrt{N}}\right) \rightarrow 2\Phi(k) - 1,$$

где $\Phi(x)$ — функция распределения стандартного нормального распределения.

Это соотношение позволяет нам получить оценку погрешности метода Монте-Карло.

В самом деле, найдем N значений случайной величины ξ . Из указанного соотношения видно, что среднее арифметическое этих значений будет приближенно равно m . С вероятностью близкой к $(2\Phi(k) - 1)$ ошибка такого приближения не превосходит величины kb/\sqrt{N} . Очевидно, эта ошибка стремится к нулю с ростом N .

В зависимости от целей исследования последнее соотношение можно использовать по-разному:

1) если взять $k = 3$, то получим так известное "правило 3σ ":

$$P\left(\left|\frac{S_N}{N} - m\right| \leq 3 \frac{b}{\sqrt{N}}\right) \approx 0.9973.$$

2) если же требуется найти определенный уровень погрешности вычислений α , то

$$P\left(\left|\frac{S_N}{N} - m\right| \leq (\Phi^{-1}((1 + \alpha)/2)) \frac{b}{\sqrt{N}}\right) \approx \alpha.$$