Лекция 5.

5. Метод выборки с отклонением.

5.1. Моделирование вещественнозначной случайной величины с заданной плотностью распределения.

Метод, рассмотренный нами в предыдущем параграфе, позволяет моделировать случайные величины, заданные своей ***функцией распределения.*** На практике с. в. часто заданы своей ***плотностью распределения***, и, поэтому, желательно иметь метод, основанный на использовании функции плотности, а не функции распределения. Опишем один из таких методов. Относительно плотности распределения $\rho(x)$ сделаем следующие предположения.

Гипотезы:

i) Плотность $\rho(x)$ имеет компактный носитель в некотором интервале $[a, b]$: это означает, что для всех $x \in R \backslash [a, b]$ плотность $\rho(x) = 0$.

ii) Плотность $\rho(x)$ непрерывна (или кусочно-непрерывна) на $[a, b]$ и ограничена (известной) константой C: для всех x, принадлежащих носителю плотности, $0 \leq \rho(x) \leq C$.

Метод:

a) Сначала моделируют последовательность *независимых* двумерных случайных векторов $Z_k = \left(V_k, V_k^{'}\right), k = 1,2, \ldots$, имеющих равномерное распределение в прямоугольнике $\mathbf{\Pi} = [a, b] \times [0, C]$.

Такую последовательность можно моделировать следующим образом:

$$V_k = a + (b-a)U_{2k-1}, V_k^{'} = CU_{2k}.$$

(напомним, что $U_1, U_2, \dots$ - исходная последовательность независимых, равномерно распределённых на [0,1] с.в., заданных на некотором вероятностном пространстве $(\Omega, \mathcal{F}, \mathbf{P})$).

Упражнение.

I. Доказать, что построенная таким образом последовательность $V_1, V_1^{'}, V_2, V_2^{'}, V_3, V_3^{'}, \dots$ является последовательностью независимых случайных величин, и для $k \geq 1$ случайные величины V_k имеют равномерное распределение на отрезке $[a,b]$, а $V_k^{'}$ – равномерное распределение на отрезке $[0,C]$.

II. Доказать, что вектор $\left(V_k, V_k^{'}\right)$ имеет равномерное распределение в прямоугольнике $\mathbf{\Pi} = [a,b] \times [0,C]$, то есть, что плотность распределения вектора $\left(V_k, V_k^{'}\right)$ представима в виде

$$\frac{1}{b-a} \times 1_{[a,b]}(x) \cdot \frac{1}{C} \times 1_{[0,C]}(y).$$

III. Доказать, что если рассмотреть «подграфик» $\mathcal{F}_I$ функции плотности над некоторым интервалом I, то есть

$$\mathcal{F}_I \triangleq \{(v, v') \in [a,b] \times [0,C] : v \in I, v' \leq \rho(v)\},$$

то

$$P\left(Z_k = \left(V_k, V_k^{'}\right) \in \mathcal{F}_I\right) = \frac{\text{площадь } (\mathcal{F}_I)}{\text{площадь}(\mathbf{\Pi})} = \frac{\int_I \rho(x)dx}{C(b-a)}.$$

b) Определим теперь случайный индекс $\nu(\omega)$ как ***первый индекс*** k***, для которого вектор*** $\left(V_k(\omega), V_k'(\omega)\right)$ ***принадлежит подграфику*** $\boldsymbol{\mathcal{E}}$ ***плотности*** $\rho(x)$:

$$\nu(\omega) = \inf\left\{k \geq 1 : V_k' \leq \rho\left(V_k(\omega)\right)\right\}.$$

Функция $\nu(\omega): \Omega \to N \cup \{\infty\}$ является случайной величиной, т.е. измерима относительно сигма - алгебры событий $\mathcal{F}$ пространства (Ω, $\mathcal{F}$, **P**). В самом деле, $\nu(\omega)$ принимает значения 1,2,…, {∞} и

$$\{\nu(\omega) = n\} = (Z_1 \in \mathcal{E}^c) \bigcap (Z_2 \in \mathcal{E}^c) \dots \bigcap (Z_{n-1} \in \mathcal{E}^c) \bigcap (Z_n \in \mathcal{E}).$$

Из независимости векторов $\{Z_k\}$ следует, что $\nu(\omega)$ имеет геометрическое распределение с параметром α

$$P\{\nu(\omega) = n\} = P(Z_1 \in \mathcal{E}^c) \times \dots \times P(Z_{n-1} \in \mathcal{E}^c) \times P(Z_n \in \mathcal{E}) =$$

$$(1-\alpha)^{n-1}\alpha,$$

где $\alpha = P(Z_k \in \mathcal{E})$. Поскольку вектора $Z_k = (V_k, V_k')$ равномерно распределёны в **П,** имеем:

$$\alpha = \frac{\int_a^b \rho(x)dx}{C(b-a)} = \frac{1}{C(b-a)}.$$

Имеем также

$$P(\nu(\omega) < \infty) = \sum_{n=1}^{\infty} P(\nu(\omega) = n) = \sum_{n=1}^{\infty} (1-\alpha)^{n-1}\alpha = 1$$

Отсюда, в частности, следует, что $\nu(\omega)$ конечна почти наверное.

c) Рассмотрим ***проекцию*** вектора $\left(V_\nu, V_\nu'\right)$ на ось x. Положим

$X = V_\nu$, то есть

$$X(\omega) = V_{\nu(\omega)}(\omega).$$

Теорема 5.1. *Случайная величина* $X(\omega)$ *имеет плотность распределения ρ(x).*

Доказательство. Найдём распределение X: очевидно, что X принимает значения на отрезке $[a, b]$, и для каждого интервала $I \in [a, b]$, поскольку Ω является объединением счётного числа непересекающихся множеств $\{\omega: \nu(\omega) = n\}$, имеем:

$$P(X \in I) = P((X \in I) \cap \bigcup_{n\geq 1} \{\nu = n\}) = \sum_{n\geq 1} P((X \in I) \cap \{\nu = n\}) =$$

$$\sum_{n\geq 1} P((V_n \in I) \cap \{\nu = n\}) =$$

$$\sum_{n\geq 1} P[(Z_1 \in \mathcal{E}^c) \cap (Z_2 \in \mathcal{E}^c) \dots \cap (Z_{n-1} \in \mathcal{E}^c) \cap (Z_n \in \mathcal{F}_I)].$$

Поскольку векторы Z_k независимы и поскольку (см. пункт iii) упражнения выше)

$$P(Z_n \in \mathcal{F}_I) = \frac{\int_I \rho(x)dx}{C(b-a)}$$

Получим

$$P(X \in I) = \frac{\int_I \rho(x)dx}{C(b-a)} \sum_{n\geq 1} (1-\alpha)^{n-1} = \frac{\int_I \rho(x)dx}{C(b-a)} \cdot \frac{1}{\alpha} = \int_I \rho(x)dx.$$

Из последнего равенства, справедливого для любого интервала I, следует, что X имеет плотность распределения,

и эта плотность равна $\rho(x)$.

Следующая теорема показывает, как построить ***последовательность*** независимых одинаково распределённых случайных величин, каждая из которых имеет плотность распределения $\rho(x)$.

Теорема 5.2. *Положим* $\nu_1 = \nu$ *и, далее, определим* ***рекуррентно***

$$\nu_r = \inf \left\{k \geq \nu_{r-1} + 1: V_k^{'} \leq \rho(V_k)\right\}.$$

Тогда последовательность X_k *, определённая соотношением*

$$X_k(\omega) = V_{\nu_{k(\omega)}}(\omega)$$

независима и каждая случайная величина $X_k(\omega)$ *имеет плотность распределения* $\rho(x)$.

Доказательство. Вычислим для любого r и для любой последовательности интервалов $I_1, \dots, I_r$ вероятность

$$P(X_1 \in I_1, \dots, X_r \in I_r).$$

Для этого обозначим

$$\mathcal{F}_{I_1,\dots,I_r}(k_1, \dots, k_r) = \left[\bigcap_{j \notin \{k_1,\dots,k_r\}} (Z_j \in \mathcal{E}^c)\right] \cap \left[\bigcap_{i=1}^{r} (Z_{k_i} \in \mathcal{F}_{I_i})\right],$$

$1 \leq k_1 < k_2 < \dots < k_r$. (1)

Интересующее нас событие $(X_1 \in I_1, \dots, X_r \in I_r)$ является объединением всех (непересекающихся для разных наборов) событий вида

$$\bigcap_{i=1}^{r} (X_i(\omega) \in I_i) \cap \bigcap_{i=1}^{r} (\nu_i(\omega) = k_i)$$

по всем наборам $1 \leq k_1 < k_2 < \dots < k_r$. Поскольку

$$\bigcap_{i=1}^{r}(X_i(\omega) \in I_i) \cap \bigcap_{i=1}^{r}(\nu_i(\omega) = k_i) = \mathcal{F}_{I_1,\dots,I_r}(k_1, \dots, k_r)$$

имеем

$$P\left(\bigcap_{i=1}^{r}(X_i \in I_i)\right) = \sum_{1\le k_1<k_2<\cdots<k_r} P\left(\mathcal{F}_{I_1,\dots,I_r}(k_1, \dots, k_r)\right)$$

Поскольку все Z_i независимы, из равенства (1) получаем

$$P\left(\bigcap_{i=1}^{r}(X_i \in I_i)\right) =$$

$$\sum_{1\le k_1<k_2<\cdots<k_r} (1-\alpha)^{k_1-1}\frac{\int_{I_1}\rho(x)dx}{C(b-a)}(1-\alpha)^{k_2-k_1-1}\frac{\int_{I_2}\rho(x)dx}{C(b-a)} \times$$

$$\dots \times (1-\alpha)^{k_r-k_{r-1}-1}\frac{\int_{I_r}\rho(x)dx}{C(b-a)}.$$

Сделав замену переменных $p_1 = k_1 - 1, p_2 = k_2 - k_1 - 1, \dots, p_r = k_r - k_{r-1} - 1$ в предыдущей сумме, получим

$$P\left(\bigcap_{i=1}^{r}(X_i \in I_i)\right) =$$

$$\sum_{p_1=0}^{\infty}(1-\alpha)^{p_1} \dots \sum_{p_r=0}^{\infty}(1-\alpha)^{p_r}\frac{\int_{I_1}\rho(x)dx}{C(b-a)} \times \dots \times \frac{\int_{I_r}\rho(x)dx}{C(b-a)} =$$

$$\left(\frac{1}{\alpha}\right)^r \frac{\int_{I_1} \rho(x)dx}{C(b-a)} \times \dots \times \frac{\int_{I_r} \rho(x)dx}{C(b-a)} = \int_{I_1} \rho(x)dx \times \dots \times \int_{I_r} \rho(x)dx, \tag{2}$$

поскольку $\frac{1}{\alpha} = (b-a) \cdot C$. Доказательство теоремы следует теперь непосредственно из равенства (2).

Зададимся теперь вопросом: каково среднее время ожидания до попадания в подграфик плотности $\rho(x)$? Иначе говоря, чему равно $E(\nu_1)$? Имеем:

$$E(\nu_1) = \sum_{k=0}^{\infty} kP(\nu_1 = k) = \sum_{k=1}^{\infty} k(1-\alpha)^{k-1}\alpha =$$

$$\alpha \frac{d}{d\alpha}\left(-\sum_{k=0}^{\infty}(1-\alpha)^k\right) = \alpha \frac{d}{d\alpha}\left(-\frac{1}{\alpha}\right) = \frac{1}{\alpha} = C(b-a).$$

Таким образом, $\boldsymbol{E(\nu_1)}$ **равно *площади прямоугольника*, *содержащего график плотности***, и разумным является выбор наименьшего из всех прямоугольников, содержащих график $\rho(x)$.

Теорему 5.2 можно обобщить на случай ***плотности, определённой на всей прямой R.*** Доказательство близко к доказательству Теоремы 5.2.

Теорема 5.3. *Пусть* $U_1, V_1, U_2, V_2, \dots, U_n, V_n, \dots$ *последовательность независимых случайных величин, причём все* U_i *равномерно распределены на отрезке* $[0,1]$, *а все* V_i *распределены с плотностью* $q(x)$. *Предполагается, что эта плотность известна,* *и мы умеем моделировать с.в. V с плотностью $q(x)$. Пусть существует константа $M \geq 1$ такая, что для всех $x \in R$*

$$\rho(x) \leq Mq(x).$$

Определим $\nu_0 = 0$ и, далее, рекуррентно положим:

$$\nu_r = \inf\{k \geq \nu_{r-1} + 1 : M \cdot U_k \cdot q(V_k) \leq \rho(V_k)\}.$$

Тогда последовательность X_k ,определённая следующим образом:

$$X_k(\omega) = V_{\nu_k(\omega)}(\omega)$$

является последовательностью независимых, одинаково распределённых с плотностью $\rho(x)$ случайных величин.

Упражнение. Доказать сформулированную теорему.

Указание: Ввести случайный вектор $(V_k, V_k^{'})$, где V_k имеет плотность распределения $q(x)$, а $V_k^{'} = M \cdot U_k \cdot q(V_k)$ и доказать, что

$$P\left((V_k, V_k^{'}) \in \mathcal{F}_I\right) = \frac{1}{M}\int_I \rho(x)dx,$$

то есть в этом случае $\alpha = \frac{1}{M}$. Далее следовать приведённому выше доказательству.

5.2. Некоторые задачи моделирования, связанные с методом выборки с отклонением.

Вопрос: как моделировать равномерное распределение в более сложных областях в R^d, например, в открытом диске $D(0,1)$ на плоскости? Первый метод состоит в использовании *полярных координат.*

Упражнение: Пусть U *и* V две независимые с.в., равномерно распределённые на отрезе [0,1]. Положим

$$R = \sqrt{U}, \ \ \Theta = 2\pi V.$$

Пусть

$$\begin{cases} X = R \cdot \cos\Theta \\ Y = R \cdot \sin\Theta \end{cases}$$

Доказать, что случайный вектор (X, Y) имеет равномерное распределение на диске $D(0,1) = \{(x, y): x^2 + y^2 < 1\}$.

Второй метод состоит в использовании метода выборки с отклонением.

Упражнение: Пусть $U_1, \dots, U_n, \dots$ последовательность независимых одинаково распределённых с.в., имеющих равномерное распределение на [0,1].

a) Доказать, что вектор $Z_i = (2U_{2i-1} - 1, 2U_{2i} - 1), i \geq 1$, равномерно распределён в квадрате $\mathbf{\Pi} = [-1,1] \times [-1,1]$.

b) Диск $D(0,1)$ содержится в квадрате **П.** Положим

$$\nu = \inf\{k \geq 1: Z_k \in D(0,1)\}$$

и $W(\omega) = Z_{\nu(\omega)}(\omega)$. Доказать, что с.в. $W(\omega)$ равномерно распределена на диске $D(0,1)$.

c) Положим $\nu_0 = 0$ и, далее, положим

$$\nu_r = \inf\{k \geq \nu_{r-1} + 1: Z_k \in D(0,1)\}.$$

Доказать, что случайные векторы

$$W_r(\omega) = Z_{\nu_r(\omega)}(\omega)$$

независимы и имеют равномерное распределение на диске $D(0,1)$.

d) Каково среднее время ожидания первого попадания в диск?

Второй метод интересен тем, что он «работает» в любой размерности. Например, этим методом можно моделировать случайные вектора, имеющие равномерное распределение на сфере любой размерности. Напомним, что распределение на поверхности *единичной* сферы S^d с центром в начале координат называется равномерным, если для любого борелевского множества A на поверхности этой сферы, вероятность попадания в него равна отношению «площади» этого множества A к объёму сферы, а «площадь» множества A равна, по определению, мере Лебега конуса C_A, опирающегося на A:

$$\text{“площадь” } A = leb(C_A) = leb\{tx : t \in [0,1], x \in A\}.$$

Например, в случае окружности S^1 радиуса r и дуги этой окружности A, вероятность попадания в эту дугу для равномерного распределения на S^1 должна быть равной $\frac{\alpha r^2/2}{\pi r^2} = \alpha/2\pi$, где α – угол сектора, опирающегося на эту дугу.

Упражнение.

a) Предложить вариант метода выборки с отклонением для моделирования последовательности независимых случайных векторов W_k, имеющих равномерное распределение на евклидовом шаре B^{d+1} размерности $d+1$

b) Доказать, что вектор

$$S_k = W_k/\|W_k\|$$

равномерно распределён на единичной сфере S^d размерности d.

5.3. Моделирование нормальных законов распределения. Метод Бокса – Мюллера.

Нормальные законы распределения играют важную роль в теории вероятностей и математической статистике, поэтому важно иметь простой метод их моделирования. Достаточно уметь моделировать с.в. X, имеющую распределение $N(0,1)$, поскольку для моделирования с.в. Y, имеющей распределение $N(\mu, \sigma^2)$ достаточно положить $Y = \mu + \sigma X$. Естественный приближённый метод моделирования связан с центральной предельной теоремой: он основан на суммировании независимых, равномерно распределённых с.в. Однако существует и прямой, достаточно простой метод моделирования – ***метод Бокса – Мюллера***. Этот метод описан в следующей теореме.

Теорема 5.4. *Пусть* U, V *– две независимые с.в., равномерно распределённые на отрезке* $[0,1]$ *. Тогда с.в.* Z *и* T*, определённые следующим образом:*

$$Z = \sqrt{-2 \cdot \ln U} \cdot \cos(2\pi V)\,, T = \sqrt{-2 \cdot \ln U} \cdot \sin\,(2\pi V)$$

имеют стандартное нормальное распределение. Более того, Z *и* T *независимы. В качестве следствия получаем: если* $U_1, V_1, U_2, V_2, \dots, U_n, V_n, \dots$ *последовательность независимых, равномерно распределённых на* $[0,1]$ *с.в., то последовательность*

$$Z_1, T_1, Z_2, T_2, \dots, Z_n, T_n, \dots$$

(полученная вышеописанной процедурой) является последовательностью независимых с.в., имеющих стандартное нормальное распределение.

Доказательство. Пусть $f(x,y): R^2 \to R$, произвольная непрерывная и ограниченная функция. Для доказательства теоремы достаточно доказать, что

$$Ef(Z,T) = \int_{R^2} f(z,t)\frac{1}{2\pi}\exp\left(-\frac{z^2+t^2}{2}\right)dzdt.$$

В одномерном случае это следует из Предложения 1.1 Лекции 1, в многомерном случае доказательство аналогично. Имеем

$$Ef(Z,T) = E\big(f(\sqrt{-2\cdot \ln U}\cdot\cos(2\pi V), \sqrt{-2\cdot \ln U}\cdot\sin(2\pi V))\big).$$

Поскольку с.в. U *и* V независимы и равномерно распределены на $[0,1]$, плотность распределения вектора (U,V) равна $1_{[0,1]}(u)\cdot 1_{[0,1]}(v)$. По формуле перехода получим

$$Ef(Z,T) = \int_{[0,1]^2} f(\sqrt{-2\cdot \ln u}\cdot\cos(2\pi v), \sqrt{-2\cdot \ln u}\cdot\sin(2\pi v))dudv.$$

Обозначим через ϕ: $]0,1[^2 \to R^2\backslash\{0,0\}$ преобразование, заданное формулой

$$\phi(u,v) = \big(\sqrt{-2\cdot \ln u}\cdot\cos(2\pi v), \sqrt{-2\cdot \ln u}\cdot\sin(2\pi v)\big).$$

Это C^∞ диффеоморфизм (т.е. взаимно-однозначное, гладкое в обе стороны C^∞ отображение), обратное отображение задано формулой

$$(u,v) = \phi^{-1}(z,t) = \left(\exp\left(-\frac{z^2+t^2}{2}\right), \frac{1}{2\pi}\arccos\left(\frac{z}{\sqrt{z^2+t^2}}\right)\right).$$

Якобиан обратного отображения $J(\phi^{-1})$ равен $\frac{1}{J(\phi)\circ\phi^{-1}}$, где $J(\phi)$ якобиан отображения ϕ, который равен

$$J(\phi) = \det\begin{pmatrix} -\frac{1}{u}\cdot(-2\ln u)^{-1/2}\cdot\cos(2\pi v) & -2\pi\cdot(-2\ln u)^{1/2}\cdot\sin(2\pi v) \\ -\frac{1}{u}\cdot(-2\ln u)^{-1/2}\cdot\sin(2\pi v) & 2\pi\cdot(-2\ln u)^{1/2}\cdot\cos(2\pi v) \end{pmatrix}$$

$$= -\frac{2\pi}{u}$$

Таким образом,

$$J(\phi^{-1})(z,t) = -\frac{1}{2\pi}\exp\left(-\frac{z^2+t^2}{2}\right)$$

По формуле замены переменных в двойных интегралах, получим

$$Ef(Z,T) = \int_{R^2} f(z,t)|J(\phi^{-1})(z,t)|dzdt =$$

$$\int_{R^2} f(z,t)\frac{1}{2\pi}\exp\left(-\frac{z^2+t^2}{2}\right)dzdt.$$

Последнее равенство, справедливое для любой ограниченной непрерывной функции $f(z,t)$, показывает, что Z и T независимы и имеют стандартное нормальное распределение.