

ДИСКРЕТНЫЕ СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Рассмотрим вероятностный эксперимент. Пусть (Ω, F, P) – математическая модель данного эксперимента. Часто при решении экономических и инженерных задач каждому элементарному событию $\omega \in \Omega$ сопоставляется некоторое действительное число $\xi(\omega)$, которое зависит от ω . В этом случае результаты эксперимента будут описываться некоторой числовой функцией $\xi(\omega)$, которая определена на пространстве элементарных событий Ω , то есть функция $\xi(\omega)$ отображает множество Ω на множество действительных чисел. Такие функции называются случайными величинами. Условимся случайные величины обозначать греческими буквами, а принимаемые ими значения – строчными латинскими буквами.

Определение 1.1.1 *Дискретной случайной величиной* называется такая величина $\xi = \xi(\omega)$, которая в зависимости от элементарных исходов ω принимает конечное или счётное множество различных значений x .

Определение 1.1.2 *Распределением вероятностей* или *законом распределения* вероятностей дискретной случайной величины называется соответствие между возможными значениями случайной величины и их вероятностями, по которому каждому возможному значению x_i случайной величины ξ поставлено в соответствие неотрицательное число $p_i = P(\xi = x_i)$, где $\sum_i p_i = 1$.

Дискретная случайная величина считается заданной, если определён её закон распределения в различных формах. Закон распределения дискретной случайной величины может быть задан рядом распределения, многоугольником распределения, аналитическим образом и функцией распределения.

Определение 1.1.3 *Рядом распределения* дискретной случайной величины ξ называется таблица, в которой перечислены возможные значения $x_1, x_2, \dots, x_n, \dots$ случайной величины ξ и соответствующие им вероятности $p_1, p_2, \dots, p_n, \dots$, причём $\sum_i p_i = \sum_i P(\xi = x_i) = 1$.

Если дискретная случайная величина ξ принимает конечное множество значений, например, пусть число значений случайной величины равно n , то ряд распределения имеет вид:

$\xi = x_i$	x_1	x_2	$\dots$	x_n
$P(\xi = x_i)$	p_1	p_2	$\dots$	p_n

Если же дискретная случайная величина ξ принимает счётное множество значений, то ряд распределения записывается в виде:

$\xi = x_i$	x_1	x_2	$\dots$	x_n	$\dots$
$P(\xi = x_i)$	p_1	p_2	$\dots$	p_n	$\dots$

Графическое изображение ряда распределения называется *многоугольником распределения*. По оси абсцисс откладываются значения случайной величины, а по оси ординат – соответствующие вероятности. Полученные точки соединяются линиями.

При аналитическом задании закона распределения ДСВ указывается формула нахождения вероятностей для каждого значения случайной величины. Например, для биномиального распределения, значения вероятностей случайной величины определяется формулой Бернулли $P_n(\xi = x_i) = C_n^{x_i} \cdot p^{x_i} \cdot q^{n-x_i}$.

Определение 1.1.4 *Функцией распределения* случайной величины ξ называется функция $F(x)$ действительного переменного x , определяющее вероятность того, что случайная величина ξ в результате эксперимента примет значение меньше некоторого числа x , то есть

$$F(x) = P(\xi < x). \quad (1.1.1)$$

Если случайную величину рассматривать как случайную точку на числовой оси, то функция распределения представляет собой вероятность того, что случайная величина попадёт в интервал расположенный левее некоторого числа x .

Приведём основные свойства функции распределения:

- 1) $\lim_{x \rightarrow -\infty} F(x) = F(-\infty) = 0$;
- 2) $\lim_{x \rightarrow +\infty} F(x) = F(+\infty) = 1$;
- 3) функция распределения является неубывающей функцией, то есть $F(x_1) \leq F(x_2)$, если $x_1 < x_2$;

4) функция распределения непрерывна слева в точке x_0 , то есть $\lim_{x \rightarrow x_0 - 0} F(x) = F(x_0)$;

5) функция распределения $F(x)$ случайной величины ξ однозначно определяет распределение вероятностей этой случайной величины, то есть вероятность произвольного события можно выразить через функцию распределения;

6) если $F(x)$ – функция распределения случайной величины ξ , то вероятность попадания случайной величины ξ в интервал $[\alpha; \beta)$, полузамкнутый справа, определяется формулой

$$P(\alpha \leq \xi < \beta) = F(\beta) - F(\alpha), \quad (1.1.2)$$

для любого значения x из этого интервала;

7) для любого значения $\xi = \alpha$ справедлива формула $P(\xi = \alpha) = \lim_{x \rightarrow \alpha + 0} F(x) - F(\alpha)$.

Функция распределения для ДСВ вычисляется по формуле

$$F(x) = \sum_{x_i < x} P(\xi = x_i), \quad (1.1.3)$$

где суммирование ведётся по всем значениям i , для которых $x_i < x$.

Рассмотрим числовые характеристики дискретных случайных величин.

Определение 1.1.5 Математическим ожиданием дискретной случайной величины ξ называется её среднее значение, которое вычисляется по формуле

$$M(\xi) = \sum_{i=1}^n x_i p_i. \quad (1.1.4)$$

Математическое ожидание также обозначается m_ξ или $E\xi$.

Приведём основные свойства математического ожидания:

- 1) если C – неслучайная величина, то $M(C) = C$;
- 2) если C – постоянный множитель, то его можно выносить за знак математического ожидания, то есть $M(C \cdot \xi) = C \cdot M(\xi)$;
- 3) $M(\xi \pm \eta) = M(\xi) \pm M(\eta)$ – для любых случайных величин ξ и η ;
- 4) $M\left(\sum_{i=1}^n \xi_i\right) = \sum_{i=1}^n M(\xi_i)$ – для любых случайных величин ξ_i , $i = \overline{1; n}$;
- 5) $M(\xi \cdot \eta) = M(\xi) \cdot M(\eta) + M((\xi - m_\xi) \cdot (\eta - m_\eta))$;
- 6) математическое ожидание произведения независимых случайных величин равно произведению их математических ожиданий;
- 7) если $\alpha \leq \xi \leq \beta$, то $\alpha \leq M(\xi) \leq \beta$;
- 8) $M(\xi) \leq M|\xi|$;

$$9) \quad M(\xi - m_\xi) = 0.$$

Определение 1.1.6 *Центрированной случайной величиной* называется разность между случайной величиной ξ и её математическим ожиданием m_ξ :

$$\hat{\xi} = \xi - m_\xi. \quad (1.1.5)$$

Определение 1.1.7 *Модой $Mo(\xi)$ дискретной случайной величины ξ* называется наиболее вероятное значение случайной величины, то есть такое её значение, при котором вероятность максимальна.

Определение 1.1.8 *Медианой $Me(\xi)$ дискретной случайной величины ξ* называется такое значение случайной величины, относительно которого равновероятны получение значений случайной величины, то есть $P(\xi < Me(\xi)) = P(\xi > Me(\xi))$.

Значение медианы можно определить из равенства $2P(\xi < Me(\xi)) = 1$.

Определение 1.1.9 *Дисперсией случайной величины ξ* называется математическое ожидание квадрата отклонения случайной величины от её математического ожидания и обозначается $D(\xi)$

$$D(\xi) = M(\hat{\xi}^2) = M((\xi - m_\xi)^2). \quad (1.1.6)$$

Дисперсия также обозначается σ_ξ^2 , или D_ξ , или $\text{Var}(\xi)$.

Согласно определению математического ожидания, дисперсия дискретной случайной величины ξ вычисляется по формуле

$$D(\xi) = \sum_{i=1}^n (x_i - m_\xi)^2 p_i. \quad (1.1.7)$$

Преобразуем формулу (1.1.7), используя определение и свойства математического ожидания:

$$D(\xi) = M(\xi^2) - m_\xi^2. \quad (1.1.8)$$

На основании формулы (1.1.8) получаем формулу вычисления дисперсии для дискретных случайных величин:

$$D(\xi) = \sum_{i=1}^n x_i^2 p_i - m_\xi^2. \quad (1.1.9)$$

Приведём основные свойства дисперсии случайных величин:

- 1) Если C – постоянная величина, то $D(C) = 0$;
- 2) $D(C\xi) = C^2 D(\xi)$;
- 3) $D(\xi \pm \eta) = D(\xi) + D(\eta) \pm 2M((\xi - m_\xi)(\eta - m_\eta))$;

4) Если события независимы, то дисперсия суммы равна сумме дисперсий.

Определение 1.1.10 Положительное значение квадратного корня из дисперсии называется *средним квадратическим отклонением или стандартом* случайной величины ξ и определяется формулой

$$\sigma_{\xi} = \sigma(\xi) = \sqrt{D(\xi)}. \quad (1.1.10)$$

Кроме дисперсии и среднего квадратического отклонения, рассеивание случайной величины ξ может характеризоваться *размахом* $R = x_{\max} - x_{\min}$, где $x_{\max}$ – наибольшее из возможных значений случайной величины ξ ; $x_{\min}$ – наименьшее из возможных значений случайной величины ξ .

Часто на практике, особенно в экономических расчётах, рассеивание принято характеризовать *коэффициентом вариации* $V = \frac{\sigma_{\xi}}{M(\xi)}$.

Начальным моментом k -го порядка случайной величины ξ называется математическое ожидание случайной величины ξ^k , то есть $\nu_k = M(\xi^k)$.

Следовательно, для дискретной случайной величины начальный момент выражается суммой:

$$\nu_k = \sum_{i=1}^n x_i^k p_i. \quad (1.1.11)$$

Начальный момент первого порядка является математическим ожиданием $\nu_1 = M(\xi)$, а второго порядка равен $\nu_2 = M(\xi^2)$. Следовательно, формулу для вычисления дисперсии можно записать в виде: $D(\xi) = \nu_2 - \nu_1^2$.

Центральным моментом k -го порядка случайной величины ξ называется математическое ожидание случайной величины $(\xi - m_{\xi})^k$, то есть $\mu_k = M((\xi - m_{\xi})^k)$.

Следовательно, для дискретной случайной величины центральный момент выражается суммой:

$$\mu_k = \sum_{i=1}^n (x_i - m_{\xi})^k p_i. \quad (1.1.12)$$

Центральный момент первого порядка всегда равен нулю. Центральный момент второго порядка является дисперсией случайной величины ξ : $\mu_2 = D(\xi) = \nu_2 - \nu_1^2$. Центральный момент третьего порядка определяется по формуле $\mu_3 = \nu_3 - 3\nu_1\nu_2 + 2\nu_1^3$, а четвёртого - $\mu_4 = \nu_4 - 4\nu_3\nu_1 + 6\nu_2\nu_1^2 - 3\nu_1^4$.

НЕПРЕРЫВНЫЕ СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Если множество возможных значений случайной величины ξ несчётно, то возможные значения этой случайной величины целиком заполняют некоторый интервал. Такие случайные величины называются непрерывными.

Определение 2.1.1 Функция $F(x)$, определённая на множестве действительных чисел, называется *абсолютно непрерывной*, если существует такая функция f , что для любого действительного числа x выполняется равенство

$$F(x) = \int_{-\infty}^x f(t)dt. \quad (2.1.1)$$

Определение 2.1.2 Случайная величина ξ называется *абсолютно непрерывной*, если её функция распределения вероятностей $F(x)$ является абсолютно непрерывной функцией.

Функцию $f(x)$ называют *плотностью распределения вероятностей*, а её график – кривой распределения.

Так как функция распределения является неубывающей и кусочно-непрерывной, то плотность распределения обладает следующими свойствами:

- а) из того, что $F(x_2) \geq F(x_1) \quad \forall x_2 > x_1$, следует $f(x) \geq 0$;
- б) функцию распределения можно дифференцировать, за исключением, быть может, конечного или счётного числа точек:

$$F'(x) = \left(\int_{-\infty}^x f(t)dt \right)' = f(x).$$

С другой стороны, если для некоторой функции $f(x)$ выполняются условия:

- 1) $f(x) \geq 0$;
- 2) $f(x)$ - непрерывна или кусочно-непрерывна;
- 3) $\int_{-\infty}^{+\infty} f(x)dx = 1$,

то функция распределения $F(x) = \int_{-\infty}^x f(t)dt$ является абсолютно непрерывной функцией и, следовательно, функция $f(x)$ - плотность распределения.

Первое условие означает, что кривая распределения располагается выше оси Ox , второе - $f(x)$ может быть либо непрерывной, либо состоять из частей непрерывных функций, третье – площадь под кривой распределения численно равна единице.

Определение 2.1.3 *Плотностью распределения вероятностей* непрерывной случайной величины ξ в произвольно фиксированной точке x называется предел отношения вероятности попадания этой случайной величины в интервал $[x; x + \Delta x)$ к величине этого интервала при стремлении Δx к нулю, и если этот предел существует и конечен, то есть

$$f(x) = \lim_{\Delta x \rightarrow 0} \frac{P(x \leq \xi < x + \Delta x)}{\Delta x}.$$

Смысл плотности распределения вероятностей $f(x)$ состоит в том, что она указывает на то, как часто группируются возможные значения случайной величины ξ в некоторой окрестности точки x .

Функция распределения $F(x)$ выражается через плотность распределения вероятностей $f(x)$ формулой

$$F(x) = \int_{-\infty}^x f(x) dx. \quad (2.1.2)$$

Плотность распределения вероятностей $f(x)$ выражается через функцию распределения $F(x)$, если она дифференцируема, формулой

$$f(x) = F'(x). \quad (2.1.3)$$

Для плотности распределения вероятностей справедливо *условие нормировки*:

$$\int_{-\infty}^{+\infty} f(x) dx = 1. \quad (2.1.4)$$

Так как, функция распределения вероятностей является непрерывной функцией, то предел функции в точке равен значению функции в этой точке, а, следовательно, согласно свойствам функции распределения, вероятность принятия случайной величиной конкретного значения равна нулю, то есть $P(\xi = \alpha) = 0$. Следовательно, справедлива *формула нахождения вероятности попадания непрерывной случайной величины в заданный интервал*.

$$\begin{aligned} P(\alpha < \xi < \beta) &= P(\alpha \leq \xi < \beta) = P(\alpha < \xi \leq \beta) = P(\alpha \leq \xi \leq \beta) = \\ &= \int_{\alpha}^{\beta} f(x) dx = F(\beta) - F(\alpha). \end{aligned} \quad (2.1.5)$$

Таким образом, зная функцию распределения непрерывной случайной величины или плотность распределения вероятностей можно вычислить вероятности произвольных событий. Указание функции распределения или плотности распределения вероятностей являются математическими моделями эксперимента. Эти модели называются законами распределения случайных величин.

Рассмотрим числовые характеристики непрерывных случайных величин. Определения и свойства числовых характеристик для произвольных

случайных величин приведены в теме № 1 «Дискретные случайные величины».

Определение 2.1.4 Математическим ожиданием непрерывной случайной величины ξ , для которой функция $f(x)$ является плотностью распределения вероятностей, называется величина несобственного интеграла

$$M(\xi) = m_{\xi} = \int_{-\infty}^{+\infty} x f(x) dx, \quad (2.1.6)$$

если он сходится.

Определение 2.1.5 Модой $Mo(\xi)$ непрерывной случайной величины ξ называется такое её значение, при котором плотность распределения вероятностей имеет максимум.

Определение 2.1.6 Медианой $Me(\xi)$ непрерывной случайной величины ξ называется такое её значение, относительно которого равновероятно получение большего или меньшего значения случайной величины, то есть $P(\xi < Me(\xi)) = P(\xi > Me(\xi))$.

Значение медианы можно определить из равенства $\int_{-\infty}^{Me(\xi)} f(x) dx = \frac{1}{2}$.

Определение 2.1.7 Дисперсией непрерывной случайной величины ξ , для которой функция $f(x)$ является плотностью распределения вероятностей, называется величина несобственного интеграла

$$D(\xi) = \int_{-\infty}^{+\infty} (x - m_{\xi})^2 f(x) dx, \quad (2.1.7)$$

если он сходится.

Из равенства (1.1.8) следует ещё одна формула для вычисления дисперсии:

$$D(\xi) = \int_{-\infty}^{+\infty} x^2 f(x) dx - m_{\xi}^2. \quad (2.1.8)$$

Определение 2.1.8 Положительное значение квадратного корня из дисперсии называется средним квадратическим отклонением или стандартом случайной величины ξ и определяется формулой

$$\sigma_{\xi} = \sigma(\xi) = \sqrt{D(\xi)}. \quad (2.1.9)$$

Размах R и коэффициент вариации для непрерывных случайных величин определяются по формулам, которые указаны на практическом занятии № 11.

Для непрерывной случайной величины *начальный момент k-го порядка* выражается несобственным интегралом:

$$\nu_k = \int_{-\infty}^{+\infty} x^k f(x) dx, \quad (2.1.10)$$

если он сходится.

Для непрерывной случайной величины *центральный момент k-го порядка* находится по формуле

$$\mu_k = \int_{-\infty}^{+\infty} (x - m_\xi)^k f(x) dx, \quad (2.1.11)$$

если он сходится.

Центральный момент первого порядка всегда равен нулю, второго порядка является дисперсией случайной величины ξ : $\mu_2 = D(\xi) = \nu_2 - \nu_1^2$. Как правило, для описания характерных особенностей вероятностной модели используются моменты не выше четвёртого порядка. Центральный момент третьего порядка определяется по формуле $\mu_3 = \nu_3 - 3\nu_2\nu_1 + 2\nu_1^3$, а четвёртого - $\mu_4 = \nu_4 - 4\nu_3\nu_1 + 6\nu_2\nu_1^2 - \nu_1^4$.

Центральный момент третьего порядка μ_3 служит характеристикой асимметрии («скошенности») распределения. Если случайная величина ξ распределена симметрично относительно своего математического ожидания, то третий центральный момент $\mu_3 = M\left((\xi - m_\xi)^3\right) = 0$.

Центральный момент четвёртого порядка μ_4 служит характеристикой островершинности или плосковершинности распределения. Ввиду того, что эти центральные моменты имеют различные размерности, удобнее пользоваться безразмерными коэффициентами: $A = \frac{\mu_3}{\sigma^3}$ - коэффициент асимметрии; $B = \frac{\mu_4}{\sigma^4}$ - коэффициент плосковершинности.

Так как для наиболее распространённого нормального закона распределения $\mu_4/\sigma^4 = 3$, то в качестве характеристики островершинности используют безразмерный коэффициент, называемый эксцессом $\Theta = \frac{\mu_4}{\sigma^4} - 3$.

Кроме рассмотренных начальных и центральных моментов, на практике иногда применяются так называемые *абсолютные начальные моменты* $\alpha_k = M(|\xi|^k)$ и *абсолютные центральные моменты* $\beta_k = M(|\xi - m_\xi|^k)$.

ОСНОВНЫЕ ВЕРОЯТНОСТНЫЕ МОДЕЛИ РАСПРЕДЕЛЕНИЯ ДИСКРЕТНЫХ СЛУЧАЙНЫХ ВЕЛИЧИН

Закон распределения случайной величины представляет собой математическую модель вероятностного эксперимента. Рассмотрим наиболее распространённые модели дискретных распределений.

Одноточечное (вырожденное) распределение

Случайная величина ξ имеет одноточечное (вырожденное) распределение, если она принимает только одно значение x_0 с вероятностью $P(\xi = x_0) = 1$.

$$\text{Функция распределения } F(x) = \begin{cases} 0, & \text{если } x \leq x_0; \\ 1, & \text{если } x > x_0. \end{cases}$$

Математическое ожидание $M(\xi) = x_0$.

Дисперсия $D(\xi) = 0$.

Среднее квадратическое отклонение $\sigma_\xi = 0$.

Распределение Бернулли (двухточечное)

Случайная величина ξ имеет двухточечное распределение, если она принимает только два значения x_1 и x_2 с вероятностями: $P(\xi = x_1) = p_1$; $P(\xi = x_2) = p_2$; $p_1 + p_2 = 1$.

Обычно в теории вероятностей рассматривают случайную величину ξ , принимающую два значения $\xi = 1$ («успех») и $\xi = 0$ («неудача») с вероятностями: $P(\xi = 1) = p$; $P(\xi = 0) = q = 1 - p$; $p + q = 1$. Распределение такой случайной величины называется распределением Бернулли.

Ряд распределения запишем в виде таблицы:

$\xi = x_i$	0	1
$P(\xi = x_i)$	$1 - p$	p

$$\text{Функция распределения } F(x) = \begin{cases} 0, & \text{если } x \leq 0; \\ 1 - p, & \text{если } 0 < x \leq 1; \\ 1, & \text{если } x > 1. \end{cases}$$

Математическое ожидание $M(\xi) = p$.

Дисперсия $D(\xi) = pq$.

Среднее квадратическое отклонение $\sigma_\xi = \sqrt{pq}$.

Биномиальное распределение

Биномиальное распределение является математической моделью независимых экспериментов, которые повторяются при одних и тех же условиях:

- а) число экспериментов должно быть конечным (будем считать, что число экспериментов равно числу n);
- б) каждый эксперимент должен закончиться одним из взаимоисключающих друг друга исходов, то есть «успех» или «неудача»;
- в) эксперименты должны быть независимы друг от друга, то есть наступления события в каждом отдельно взятом эксперименте не зависят от того, какие результаты наступят в других экспериментах;
- г) вероятность p наступления события, то есть «успех», и вероятность $q = 1 - p$ противоположного события, то есть «неудача», в каждом отдельно взятом испытании постоянна.

Данные условия называются схемой Бернулли.

Вероятность того, что случайная величина ξ примет конкретное значение x_0 определяется по формуле

$$P(\xi = x_0) = C_n^{x_0} p^{x_0} (1 - p)^{n - x_0}. \quad (3.1.3.1)$$

$$\text{Функция распределения } F(x) = P(\xi < x) = \sum_{x_i < x} C_n^{x_i} p^{x_i} q^{n - x_i}.$$

$$\text{Математическое ожидание } M(\xi) = np.$$

$$\text{Дисперсия } D(\xi) = npq.$$

$$\text{Среднее квадратическое отклонение } \sigma_\xi = \sqrt{npq}.$$

Биномиальное распределение (этот термин был впервые использован в работе Yule, 1911 г.) возникает в тех случаях, когда ставится вопрос: сколько раз происходит некоторое событие в серии из определенного числа независимых наблюдений (опытов), выполняемых в одинаковых условиях. Биномиальное распределение возникло из наблюдений за простейшей азартной игрой: бросание правильной монеты. Во многих ситуациях эта модель служит хорошим первым приближением для более сложных игр и случайных процессов, возникающих при игре на бирже. Необходимо отметить, что существенные черты многих сложных процессов можно понять, исходя из простой биномиальной модели.

Геометрическое распределение

Пусть случайная величина ξ - число экспериментов по схеме Бернулли (вероятность наступления события в отдельно взятом испытании равна p , а не наступление этого события равна $q = 1 - p$) до наступления первого

успеха. В условиях данной схемы первый успех наступит при выполнении условий:

а) первые $(k-1)$ эксперименты закончились неудачей с вероятностью $(1-p)^{k-1}$;

б) k -ое испытание закончилось успехом с вероятностью p .

Геометрическое распределение, определяющее вероятность того, что в условиях схемы Бернулли, событие наступит ровно в k -ом испытании равно

$$P(\xi = k) = pq^{k-1}. \quad (3.1.4.1)$$

Функция распределения $F(x) = P(\xi < x_i) = \sum_{x_i < x} pq^{x_i-1}$.

Математическое ожидание $M(\xi) = 1/p$.

Дисперсия $D(\xi) = \frac{1-p}{p^2}$.

Среднее квадратическое отклонение $\sigma_\xi = \frac{\sqrt{1-p}}{p}$.

Геометрический закон распределения имеет место в таких науках как микробиология, генетика, физика и экономика. На практике эксперимент или опыт осуществляют до первого появления успешной события A . Число проведенных попыток будет целочисленной случайной величиной $1, 2, \dots$

Гипергеометрическое распределение

Гипергеометрическое распределение используется в теории статистического контроля качества. Используя это распределение, определяется доля качественных деталей в выборке из контролируемой партии.

Предположим, что в контролируемой партии имеется N элементов, среди которых M не удовлетворяют стандарту и $N-M$ удовлетворяют стандарту. Из этой партии выбираем n элементов. Пусть ξ - число элементов, не удовлетворяющих стандарту, среди n взятых элементов. Вероятность того, что в выборке объема n будет ровно x_i элементов, которые не удовлетворяют стандарту, определяется по формуле

$$P(\xi = x_i) = \frac{C_M^{x_i} C_{N-M}^{n-x_i}}{C_N^n}. \quad (3.1.5.1)$$

Функция распределения $F(x) = P(\xi < x) = \sum_{x_i < x} \frac{C_M^{x_i} C_{N-M}^{n-x_i}}{C_N^n}.$

Математическое ожидание $M(\xi) = np$, где $p = M/N$, $q = 1 - p$.

Дисперсия $D(\xi) = npq \cdot \left(1 - \frac{n-1}{N-1}\right).$

Среднее квадратическое отклонение $\sigma_\xi = \sqrt{npq \cdot \left(\frac{N-n}{N-1}\right)}.$

Распределение Пуассона

Рассмотрим схему Бернулли при большом числе испытаний n и достаточно малой вероятности p наступления события в отдельно взятом испытании, причём произведение np принимает постоянное значение $np = \mu$. Тогда вероятность того, что случайная величина ξ принимает значение m , определяется формулой Пуассона

$$P(\xi = m) = \frac{\mu^m}{m!} e^{-\mu}. \quad (3.1.6.1)$$

Значения распределения Пуассона можно определить по специальным таблицам.

Функция распределения $F(x) = P(\xi < x) = \sum_{x_i < x} \frac{\mu^{x_i}}{x_i!} e^{-\mu}.$

Математическое ожидание $M(\xi) = \mu$.

Дисперсия $D(\xi) = \mu$.

Среднее квадратическое отклонение $\sigma_\xi = \sqrt{\mu}.$

Пуассоновское распределение присутствует в теории случайных процессов. Например, при расчёте нагрузки линий связи обычно предполагают, что количества вызовов, поступивших за непересекающиеся интервалы времени, независимые случайные величины, подчиняющиеся пуассоновскому распределению с параметрами, значения которых пропорциональны длинам соответствующих интервалов времени.

ОСНОВНЫЕ ВЕРОЯТНОСТНЫЕ МОДЕЛИ РАСПРЕДЕЛЕНИЯ НЕПРЕРЫВНЫХ СЛУЧАЙНЫХ ВЕЛИЧИН

Рассмотрим основные законы непрерывных случайных величин.

Равномерный закон распределения

На практике встречаются случайные величины, о которых заранее известно, что они могут принять какое-либо значение в строго определенных границах, причем в этих границах все значения случайной величины имеют одинаковую вероятность (обладают одной и той же плотностью вероятностей). Например, при поломке часов остановившаяся минутная стрелка будет с одинаковой вероятностью показывать время, прошедшее от начала данного часа до поломки часов. Это время является случайной величиной, принимающей с одинаковой плотностью вероятности значения, которые не выходят за границы, определенные продолжительностью одного часа. К подобным случайным величинам относится также и погрешность округления. Про такие величины говорят, что они распределены равномерно, т. е. имеют равномерное распределение.

Непрерывная случайная величина ξ имеет равномерное распределение на отрезке $[a; b]$, если на этом отрезке плотность распределения вероятности принимает постоянное значение, то есть

$$f(x) = \begin{cases} \frac{1}{b-a}, & \text{если } a \leq x \leq b; \\ 0, & \text{если } x < a \text{ или } x > b. \end{cases} \quad (4.1.1.1)$$

$$\text{Функция распределения } F(x) = \begin{cases} 0, & \text{если } x < a; \\ \frac{x-a}{b-a}, & \text{если } a \leq x \leq b; \\ 1, & \text{если } x > b. \end{cases}$$

$$\text{Математическое ожидание } M(\xi) = \frac{a+b}{2}.$$

$$\text{Дисперсия } D(\xi) = \frac{(b-a)^2}{12}.$$

$$\text{Среднее квадратическое отклонение } \sigma_{\xi} = \frac{b-a}{2\sqrt{3}}.$$

Вероятность попадания равномерно распределённой случайной величины ξ на интервал $[\alpha; \beta)$, представляющий собой часть отрезка $[a; b]$, вычисляется по формуле

$$P(\alpha \leq \xi < \beta) = \frac{\beta - \alpha}{b - a}. \quad (4.1.1.2)$$

Показательное (экспоненциальное) распределение

Расстояние T между двумя соседними событиями в стационарном пуассоновском потоке с интенсивностью λ представляет собой непрерывную случайную величину, распределённую по показательному закону с плотностью вероятности

$$f(t) = \begin{cases} 0, & \text{если } t < 0; \\ \lambda e^{-\lambda t}, & \text{если } t \geq 0. \end{cases} \quad (4.1.2.1)$$

$$\text{Функция распределения } F(t) = \begin{cases} 0, & \text{если } t < 0; \\ 1 - e^{-\lambda t}, & \text{если } t \geq 0. \end{cases}$$

$$\text{Математическое ожидание } M(T) = \frac{1}{\lambda}.$$

$$\text{Дисперсия } D(T) = \frac{1}{\lambda^2}.$$

$$\text{Среднее квадратическое отклонение } \sigma_T = \frac{1}{\lambda}.$$

Для показательного распределения вероятность того, что случайная величина T примет значение, которое принадлежит интервалу $[\alpha; \beta)$, определяется по формуле

$$P(\alpha \leq T < \beta) = e^{-\lambda \alpha} - e^{-\lambda \beta}. \quad (4.1.2.2)$$

Примером непрерывной случайной величины, распределенной по показательному закону, может служить время между появлениями двух последовательных событий простейшего потока.

Гамма-функция и её свойства

Гамма-функцией или интегралом Эйлера второго рода называется функция вида

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx. \quad (4.1.3.1)$$

Гамма-функция является интегралом, зависящим от параметра α . Она удовлетворяет следующим свойствам:

1) $\Gamma(1) = \Gamma(2) = 1$;

2) $\Gamma(0,5) = \sqrt{\pi}$;

3) $\Gamma(\alpha + 1) = \alpha \cdot \Gamma(\alpha)$;

4) Если α кратно 0,5, то $\Gamma(\alpha)$ может быть легко вычислено на основании третьего свойства:

$$\begin{aligned} \Gamma(1,5) &= \Gamma(0,5 + 1) = 0,5 \cdot \Gamma(0,5) = 0,5 \cdot \sqrt{\pi}, \\ \Gamma(2,5) &= \Gamma(1,5 + 1) = 1,5 \cdot \Gamma(1,5) = 1,5 \cdot 0,5 \cdot \sqrt{\pi}, \\ &\dots \end{aligned}$$

$$\Gamma\left(\frac{n+1}{2}\right) = \Gamma\left(\frac{n-1}{2} + 1\right) = \frac{n-1}{2} \cdot \Gamma\left(\frac{n-1}{2}\right) = \frac{n-1}{2} \cdot \frac{n-3}{2} \cdot \dots \cdot \frac{5}{2} \cdot \frac{3}{2} \cdot \frac{1}{2} \cdot \sqrt{\pi};$$

5) при больших значениях α гамма-функция вычисляется по формуле

$$\Gamma(\alpha) = (\alpha-1) \cdot \Gamma(\alpha-1) = (\alpha-1) \cdot (\alpha-2) \cdot \Gamma(\alpha-2) = \dots$$

Распределение Вейбулла ($a > 0$, $b > 0$, $-\infty < c < +\infty$)

Данное распределение имеет наработка до отказа некоторых невосстанавливаемых изделий. К ним относятся, в частности, некоторые изделия у которых отказ наступает вследствие усталостного разрушения. При оценке надежности механических узлов этим законом описывается надежность подшипников. Распределение Вейбулла используют для описания закономерностей отказов под действием износа и старения, отказов последовательно соединенных и дублированных элементов. Особенно часто распределение Вейбулла применяют тогда, когда поток отказов не стационарен и интенсивность отказов меняется с течением времени.

Непрерывная случайная величина ξ имеет распределение Вейбулла, если плотность распределения вероятности имеет вид

$$f(x) = \begin{cases} \frac{b}{a} \left(\frac{x-c}{a} \right)^{b-1} e^{-\left(\frac{x-c}{a} \right)^b}, & \text{если } x > c; \\ 0, & \text{если } x \leq c, \end{cases} \quad (4.1.4.1)$$

где a - параметр масштаба, b - параметр формы кривой распределения, c - параметр сдвига.

С помощью параметра b удобно подбирать аналитическое описание для различных экспериментальных зависимостей. При значении $b=1$ распределение Вейбулла становится экспоненциальным. При значениях параметра $b < 1$ интенсивность отказов монотонно убывает с течением времени, а при значениях $b > 1$ монотонно возрастает. Подбирая значения a , b и c , можно добиться приближения аналитической функции плотности распределения к опытным данным.

$$\text{Математическое ожидание } M(X) = c + a \cdot \Gamma\left(1 + \frac{1}{b}\right).$$

$$\text{Дисперсия } D(T) = a^2 \cdot \left(\Gamma\left(1 + \frac{2}{b}\right) - \Gamma^2\left(1 + \frac{1}{b}\right) \right).$$

$$\text{Среднее квадратическое отклонение } \sigma_x = a \cdot \sqrt{\Gamma\left(1 + \frac{2}{b}\right) - \Gamma^2\left(1 + \frac{1}{b}\right)}.$$

β -распределение($\alpha > 0, \beta > 0$)

Функция β -распределения обычно используется для изучения вариации в процентах какой-либо величины между выборками - например, части дня, которую студенты проводят для самостоятельного изучения дисциплин.

Непрерывная случайная величина ξ имеет β -распределение, если плотность распределения вероятности имеет вид

$$f(x) = \begin{cases} \frac{x^{\alpha-1} \cdot (1-x)^{\beta-1}}{B(\alpha; \beta)}, & \text{если } 0 < x < 1; \\ 0, & \text{если } x \leq 0 \text{ или } x \geq 1, \end{cases} \quad (4.1.5.1)$$

где $B(\alpha; \beta) = \int_0^1 x^{\alpha-1} (1-x)^{\beta-1} dx$ называется β -функцией.

$$\text{Математическое ожидание } M(X) = \frac{\alpha}{\alpha + \beta}.$$

$$\text{Дисперсия } D(X) = \frac{\alpha \cdot \beta}{(\alpha + \beta)^2 \cdot (\alpha + \beta + 1)}.$$

$$\text{Среднее квадратическое отклонение } \sigma_x = \frac{1}{\alpha + \beta} \cdot \sqrt{\frac{\alpha \cdot \beta}{\alpha + \beta + 1}}.$$

Нормальный закон распределения

Нормальное распределение имеет большое значение в теории вероятностей и математической статистике. Фундаментальная роль, которую играет нормальное распределение, объясняется тем, что при некоторых условиях суммы случайных величин $\sum_{i=1}^n \xi_i$ с ростом числа слагаемых асимптотически нормальны, то есть имеют распределение, близкое к нормальному распределению. Нормальное распределение часто встречается на практике в самых различных областях. Принято считать, что ошибки измерений, вес и размер деталей, вес и рост людей имеют нормальное распределение. Нормальное распределение используется при анализе производственных погрешностей, контроле технологических процессов и режимов, а также при анализе и прогнозировании различных явлений в экономике, медицине и других областях знаний.

Случайная величина ξ имеет нормальное распределение, которое также называется распределением Гаусса, если плотность распределения вероятностей имеет вид

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}, \quad (5.1.1.1)$$

где a - математическое ожидание случайной величины ξ , а σ - среднее квадратическое отклонение.

Нормальное распределение зависит от двух параметров a и σ , и поэтому его часто называют *двухпараметрическим распределением*.

Функция нормального распределения имеет следующий вид

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(x-a)^2}{2\sigma^2}} dx. \quad (5.1.1.2)$$

График плотности распределения вероятностей, задаваемой формулой (5.1.1.1), называется *нормальной кривой или кривой Гаусса*. Нормальная кривая представляет собой колоколообразную кривую, симметричную относительно прямой $x = a$. Если $x \rightarrow \pm\infty$, то график нормальной кривой асимптотически приближается к оси абсцисс.

Приведём основные свойства кривой Гаусса.

1) Функция плотности распределения вероятностей определена на множестве действительных чисел.

2) Нормальная кривая расположена выше оси абсцисс, $f(x) > 0$.

3) Ветви нормальной кривой асимптотически приближаются к оси абсцисс, $\lim_{x \rightarrow \pm\infty} f(x) = 0$.

4) Кривая имеет максимум в точке $x = a$, равный $f(a) = \frac{1}{\sigma\sqrt{2\pi}}$.

При увеличении (уменьшении) параметра σ максимальная ордината уменьшается (увеличивается). Функция плотности распределения вероятностей возрастает на промежутке $(-\infty; a]$ и убывает на промежутке $[a; +\infty)$

5) Кривая Гаусса симметрична относительно прямой $x = a$. Следовательно, для нормально распределённой случайной величины математическое ожидание совпадает с модой и медианой.

6) Нормальная кривая имеет две точки перегиба $M_1\left(a - \sigma; \frac{1}{\sigma\sqrt{2\pi}e}\right)$ и $M_2\left(a + \sigma; \frac{1}{\sigma\sqrt{2\pi}e}\right)$. Кривая Гаусса выпуклая на промежутке $(a - \sigma; a + \sigma)$, а вогнутая на промежутках $(-\infty; a - \sigma)$ и $(a + \sigma; +\infty)$.

7) Площадь фигуры, ограниченной кривой Гаусса и осью абсцисс, равна 1; площадь криволинейной трапеции, ограниченной кривой Гаусса, осью абсцисс и прямыми $x = a \pm \sigma$, приблизительно равна 0,6428; площадь криволинейной трапеции, ограниченной кривой Гаусса, осью абсцисс и прямыми $x = a \pm 2\sigma$, приблизительно равна 0,9545; площадь криволинейной трапеции, ограниченной кривой Гаусса, осью абсцисс и прямыми $x = a \pm 3\sigma$, приблизительно равна 0,9973. Так как площадь под кривой Гаусса всегда

равна 1 ($\int_{-\infty}^{+\infty} f(x)dx=1$), то при увеличении σ кривая Гаусса становится плосковершинной; при уменьшении σ - кривая Гаусса вытягивается вверх. Поэтому параметр σ часто называют *параметром масштаба*.

8) Если изменять математическое ожидание a при неизменном значении σ , то кривая Гаусса будет смещаться вдоль оси Ox , то есть параметр a характеризует положение кривой при неизменной форме. Поэтому часто параметр a называют *параметром сдвига*.

9) Коэффициенты асимметрии и эксцесса для нормального распределения равны нулю: $A = \frac{\mu_3}{\sigma^3} = 0$; $\mathcal{E} = \frac{\mu_4}{\sigma^4} - 3 = 0$.

Вероятность попадания нормально распределённой случайной величины ξ в интервал $(\alpha; \beta)$ вычисляется по следующей формуле

$$P(\alpha < \xi < \beta) = \Phi\left(\frac{\beta - a}{\sigma}\right) - \Phi\left(\frac{\alpha - a}{\sigma}\right), \quad (5.1.1.3)$$

где $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$ - функция Лапласа, значение которой можно определить по таблице (приложение 3). Свойства функции Лапласа приведены на практическом занятии № 10.

Используя определение 1.1.3 (определение функции распределения) и формулу (5.1.1.3), функцию распределения нормально распределённой случайной величины ξ можно записать в виде

$$F(x) = \Phi\left(\frac{x - a}{\sigma}\right) + 0,5. \quad (5.1.1.4)$$

Правило трёх сигм

В практических расчётах часто требуется найти вероятность осуществления неравенства $|\xi - a| < \delta$. Для вычисления вероятности этого события, то есть вероятности того, что нормально распределённая случайная величина ξ отклонится от своего математического ожидания по абсолютной величине на величину, меньшую δ , воспользуемся формулой (5.1.1.3) и свойством нечётности функции Лапласа:

$$P(|\xi - a| < \delta) = 2\Phi\left(\frac{\delta}{\sigma}\right). \quad (5.1.2.1)$$

Преобразуем формулу (5.1.2.1), положив $\delta = t\sigma$. В результате получим $P(|\xi - a| < t\sigma) = 2\Phi(t)$. Полагая последовательно $t = 1, 2, 3$, имеем:

$$P(|\xi - a| < \sigma) = 2\Phi(1) = 0,6428;$$

$$P(|\xi - a| < 2\sigma) = 2\Phi(2) = 0,9545;$$

$$P(|\xi - a| < 3\sigma) = 2\Phi(3) = 0,9973.$$

Из последнего равенства следует, что практическое рассеивание нормально распределённой случайной величины лежит в пределах $(a - 3\sigma; a + 3\sigma)$. Вероятность того, что случайная величина, окажется вне этого промежутка, равна 0,0027, то есть событие может произойти лишь в 0,27 % случаев. В теории вероятностей и математической статистике такие события называются *практически невозможными*. В этом и состоит *правило трёх сигм*: если случайная величина имеет нормальное распределение, то отклонение этой величины от математического ожидания по абсолютной величине не превосходит утроенного среднего квадратического отклонения.

Поэтому трёхсигмовые границы принимаются за границы практически предельных возможных значений нормально распределённой случайной величины. Используя это правило, можно ориентировочно оценить среднее квадратическое отклонение. Для этого вычисляют размах наблюдаемой случайной величины $R = x_{\max} - x_{\min}$. Тогда на основании правила трёх сигм

$$\text{имеем: } \sigma = \frac{1}{6} R.$$

При решении практических задач используются квантили нормального распределения. Квантилем $u_{\alpha/2}$, соответствующим уровню вероятности $P = 1 - \alpha$, называется такое значение нормированной переменной $u = u_{\alpha/2}$,

$$\text{при котором выполняется равенство: } P(|u| < u_{\alpha/2}) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-u_{\alpha/2}}^{u_{\alpha/2}} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = 1 - \alpha.$$

С геометрической точки зрения, нахождение квантиля $u_{\alpha/2}$ заключается в таком выборе значения $u = u_{\alpha/2}$, при котором суммарная площадь криволинейных трапеций, лежащих левее $-u_{\alpha/2}$ и правее $u_{\alpha/2}$ под кривой Гаусса, равна числу α . Квантили нормального распределения, соответствующие вероятности $1 - \alpha$, являются аргументами функции Лапласа: $P(|u| < u_{\alpha/2}) = 2\Phi(u_{\alpha/2}) = 1 - \alpha$. Значения наиболее используемых квантилей нормального распределения приведены в таблице:

$P = 1 - \alpha$	0	0,95	0,99	0,9973	0,9999
$u_{\alpha/2}$	0	1,96	2,58	3,00	3,37

Стандартное нормальное распределение

Нормальное распределение с нулевым математическим ожиданием и единичной дисперсией называется стандартным нормальным распределением.

Непрерывная случайная величина ξ распределена по стандартному нормальному закону с плотностью вероятности

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}. \quad (5.1.3.1)$$

Функция распределения для стандартного нормального распределения имеет вид

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt = \Phi(x) + 0,5. \quad (5.1.3.2)$$

Математическое ожидание $M(\xi) = 0$.

Дисперсия $D(\xi) = 1$.

Среднее квадратическое отклонение $\sigma_{\xi} = 1$.

Вероятность попадания случайной величины ξ , которая имеет стандартное нормальное распределение, в интервал $(\alpha; \beta)$ вычисляется по следующей формуле

$$P(\alpha < \xi < \beta) = \Phi(\beta) - \Phi(\alpha). \quad (5.1.3.3)$$

Логарифмическое нормальное распределение

Случайная величина ξ называется логарифмически нормально распределённой, если её логарифм $\ln \xi$ подчинён нормальному закону распределения. Это означает, в частности, что значения этой случайной величины формируются под воздействием очень большого числа взаимно независимых факторов, причём воздействие каждого отдельного фактора «равномерно незначительно» и равновероятно по знаку. При этом в отличие от схемы формирования механизма нормального закона последовательный характер воздействия случайных факторов таков, что случайный прирост, вызываемый действием каждого следующего фактора, пропорционален уже достигнутому к этому моменту значению исследуемой величины (в этом случае говорят о мультипликативном характере воздействия фактора). Этому распределению с хорошим приближением подчиняется, например, размер частиц при дроблении какого-либо материала (камня и т. п.), содержание многих минералов в породах. Данный закон, чаще всего, используется для описания распределения поступления денежных средств (доходов), банковских вкладов, износа основных средств и т.д. Логнормальное распределение часто используют в математической статистике и эконометрике.

Логарифмическое нормальное распределение является распределением положительной случайной величины, логарифм которой имеет гауссово распределение. Случайная величина ξ распределена по логарифмическому нормальному закону (логнормальному закону) с плотностью вероятности

$$f(x) = \frac{1}{\sigma \cdot x \cdot \sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}. \quad (5.1.4.1)$$

В данном законе параметр сдвига μ и параметр масштаба σ являются средним значением и стандартным отклонением не самой случайной величины ξ , а ее логарифма $\ln \xi$.

Функция распределения для логнормального распределения имеет вид

$$F(x) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \int_0^x \frac{1}{t} e^{-\frac{(\ln t - \mu)^2}{2\sigma^2}} dt = \frac{1}{2} \left(1 + \operatorname{erf} \left(\frac{\ln x - \mu}{\sigma \sqrt{2}} \right) \right), \quad (5.1.4.2)$$

где $\operatorname{erf} t = \frac{2}{\sqrt{\pi}} \int_0^t e^{-z^2} dz$ - функция ошибок.

Плотность распределения, функцию распределения данного закона можно определить по специальным таблицам для нормированного нормального распределения, вычисляя квантиль по формуле $t = \frac{\ln x - \mu}{\sigma}$.

Математическое ожидание $M(\xi) = e^{\mu + \frac{\sigma^2}{2}}$.

Дисперсия $D(\xi) = e^{2\mu + \sigma^2} (e^{\sigma^2} - 1)$.

Среднее квадратическое отклонение $\sigma_\xi = \sqrt{e^{2\mu + \sigma^2} (e^{\sigma^2} - 1)}$.

Вероятность попадания случайной величины ξ , которая имеет логарифмическое нормальное распределение, в интервал $(\alpha; \beta)$ вычисляется по следующей формуле

$$P(\alpha < \xi < \beta) = \operatorname{erf} \left(\frac{\ln \beta - \mu}{\sigma} \right) - \operatorname{erf} \left(\frac{\ln \alpha - \mu}{\sigma} \right). \quad (5.1.4.3)$$

Распределение χ^2 (хи-квадрат)

Распределение хи-квадрат используют при оценивании дисперсии (с помощью доверительного интервала), при проверке гипотез согласия, однородности, независимости, прежде всего для качественных (категоризованных) переменных, принимающих конечное число значений, и во многих других задачах статистического анализа данных.

Рассмотрим случайную величину ξ , распределённую по нормальному закону с математическим ожиданием $M(\xi) = a$ средним квадратическим отклонением σ .

Случайная величина $\chi = \xi - a / \sigma$ называется стандартизованной случайной величиной, распределённой по нормальному закону с параметрами $M(\chi) = 0$ и $\sigma_\chi = 1$. Квадрат стандартизованной случайной

величины $\chi^2 = \left(\frac{\xi - a}{\sigma} \right)^2$ называется случайной величиной χ^2 с одной степенью свободы.

Рассмотрим n независимых величин $\xi_1, \xi_2, \dots, \xi_n$, распределённых по нормальному закону с математическими ожиданиями $a_1, a_2, \dots, a_n$ и средними квадратическими отклонениями $\sigma_1, \sigma_2, \dots, \sigma_n$. Сумма квадратов стандартизованных переменных $\chi^2 = \sum_{i=1}^n \left(\frac{\xi_i - a_i}{\sigma_i} \right)^2$ называется *случайной величиной* χ^2 с числом степеней свободы $\nu = n$.

Плотность распределения случайной величины χ^2 имеет вид

$$f(\chi^2) = \begin{cases} \frac{1}{2^{\nu/2} \Gamma(\nu/2)} \cdot (\chi^2)^{\frac{\nu}{2}-1} \cdot e^{-\frac{\chi^2}{2}}, & \text{если } \chi^2 \geq 0; \\ 0, & \text{если } \chi^2 < 0. \end{cases} \quad (5.1.5.1)$$

Функция распределения случайной величины χ^2 имеет вид

$$F(\chi^2) = P(\chi^2 < \chi_0^2) = \begin{cases} \frac{1}{2^{\nu/2} \Gamma(\nu/2)} \cdot \int_0^{\chi_0^2} (\chi^2)^{\frac{\nu}{2}-1} e^{-\frac{\chi^2}{2}} d(\chi^2), & \text{если } \chi^2 \geq 0; \\ 0, & \text{если } \chi^2 < 0. \end{cases} \quad (5.1.5.2)$$

На практике часто используются квантили χ^2 -распределения - $\chi_{\alpha, \nu}^2$. Квантилем $\chi_{\alpha, \nu}^2$, отвечающим заданному уровню α , называется такое значение $\chi^2 = \chi_{\alpha, \nu}^2$, при котором $P(\chi^2 > \chi_{\alpha, \nu}^2) = \int_{\chi_{\alpha, \nu}^2}^{+\infty} f(\chi^2) d\chi^2 = \alpha$. В приложении 4 приведены значения квантилей $\chi_{\alpha, \nu}^2$ в зависимости от числа ν степеней свободы и вероятностей α .

Распределение Стьюдента

В настоящее время распределение Стьюдента – одно из наиболее известных распределений среди используемых при анализе реальных данных. Его применяют при оценивании математического ожидания, прогнозного значения и других характеристик с помощью доверительных интервалов, по проверке гипотез о значениях математических ожиданий, коэффициентов регрессионной зависимости, гипотез однородности выборок и т.д.

Пусть $X, X_1, X_2, \dots, X_n$ - независимые случайные величины, имеющие нормальное распределение с параметрами $M(X) = M(X_1) = \dots = M(X_n) = 0$ и $\sigma_X = \sigma_{X_1} = \dots = \sigma_{X_n} = 1$.

Случайная величина

$$t = \frac{X}{\sqrt{\frac{1}{n} \sum_{i=1}^n X_i^2}} = \frac{X}{\sqrt{\frac{1}{n} \chi^2}}, \quad (5.1.6.1)$$

являющаяся функцией нормально распределённых случайных величин, называется *безразмерной дробью Стьюдента*.

Плотность распределения вероятностей случайной величины t имеет вид

$$f(t) = S(t, \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu} \cdot \Gamma\left(\frac{\nu}{2}\right)} \cdot \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}. \quad (5.1.6.2)$$

Функция распределения вероятностей случайной величины t имеет вид

$$F(t) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu} \cdot \Gamma\left(\frac{\nu}{2}\right)} \cdot \int_{-\infty}^t \left(1 + \frac{z^2}{\nu}\right)^{-\frac{\nu+1}{2}} dz. \quad (5.1.6.3)$$

Математическое ожидание $M(t) = 0$.

Дисперсия $D(t) = \frac{\nu}{\nu-2} (\nu \geq 3)$.

Среднее квадратическое отклонение $\sigma_t = \sqrt{\frac{\nu}{\nu-2}} (\nu \geq 3)$.

В приложении 5 приведены квантили распределения Стьюдента $t_{\alpha/2; \nu}$ в зависимости от числа ν степеней свободы и заданного уровня вероятности α , найденного из решения уравнения $P(|t| > t_{\alpha/2; \nu}) = 2 \int_{t_{\alpha/2; \nu}}^{+\infty} f(t) dt = \alpha$.

Распределение Фишера

Распределение Фишера (F -распределение) используют при проверке гипотез об адекватности модели в регрессионном анализе, о равенстве дисперсий и в других задачах прикладной статистики.

Пусть случайные величины $\chi_1, \chi_2, \dots, \chi_m$ и $\eta_1, \eta_2, \dots, \eta_n$ независимы и имеют нормальное распределение с параметрами

$$M(\chi_i) = M(\eta_j) = 0; \quad D(\chi_i) = D(\eta_j) = 1 \quad (i = \overline{1; m}; j = \overline{1; n}).$$

Безразмерная случайная величина

$$F = \frac{n \sum_{i=1}^m \chi_i^2}{m \sum_{j=1}^n \eta_j^2}$$

имеет распределение Фишера с плотностью распределения вероятностей

$$f(F) = \begin{cases} \frac{\Gamma\left(\frac{\nu_1 + \nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)} \left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}} \frac{F^{\frac{\nu_1}{2}-1}}{\left(1 + \frac{\nu_1}{\nu_2} F\right)^{\frac{\nu_1 + \nu_2}{2}}}, & \text{если } F > 0; \\ 0, & \text{если } F < 0, \end{cases} \quad (5.1.7.1)$$

где $\nu_1 = m$ - число степеней свободы числителя случайной величины F ;
 $\nu_2 = n$ - число степеней свободы знаменателя случайной величины F .

Функция распределения Фишера имеет вид

$$R(F) = \begin{cases} \frac{\Gamma\left(\frac{\nu_1 + \nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)} \left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}} \int_0^F \frac{Z^{\frac{\nu_1}{2}-1}}{\left(1 + \frac{\nu_1}{\nu_2} Z\right)^{\frac{\nu_1 + \nu_2}{2}}} dZ, & \text{если } Z > 0; \\ 0, & \text{если } Z < 0. \end{cases} \quad (5.1.7.2)$$

Математическое ожидание $M(F) = \frac{\nu_2}{\nu_2 - 2} \quad (\nu_2 \geq 3)$.

Дисперсия $D(F) = \frac{2\nu_2^2}{\nu_2 - 4} \cdot \frac{\nu_1 + \nu_2 - 2}{\nu_1(\nu_2 - 2)^2} \quad (\nu_2 \geq 5)$.

Среднее квадратическое отклонение

$$\sigma_F = \frac{\nu_2}{\nu_2 - 2} \cdot \sqrt{\frac{2 \cdot (\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2 - 4)}}, \quad (\nu_2 \geq 5).$$

В таблицах приведены значения квантилей F - распределения $F_{\alpha; \nu_1; \nu_2}$ в зависимости от числа степеней свободы ν_1 и ν_2 . Значения квантилей

найлены из решения уравнения $P(F > F_{\alpha; \nu_1; \nu_2}) = \int_{F_{\alpha; \nu_1; \nu_2}}^{+\infty} f(F) dF = \alpha$.

ДИСКРЕТНЫЕ МНОГОМЕРНЫЕ СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Во многих экономических приложениях любому элементарному исходу $\omega \in \Omega$ ставится в соответствие не одно, а несколько (в общем случае n) действительных чисел. При этом говорят, что эксперимент описывается n -мерной случайной величиной, или системой n случайных величин, или n -мерным вектором $(\xi_1(\omega); \xi_2(\omega); \dots; \xi_n(\omega))$.

Рассмотрим подробно двумерную случайную величину. Для n -мерной случайной величины указанные понятия определяются аналогично.

Пусть ξ и η - две дискретные случайные величины, которые определены на одном и том же пространстве элементарных событий Ω . Упорядоченную пару $(\xi; \eta)$ дискретных случайных величин ξ и η называют *двумерной случайной величиной*. Множество возможных значений такой случайной величины содержит конечное или счётное число точек $(x_i; y_j)$, $i, j = \overline{1, \infty}$: $\Omega_{\xi\eta} = \{(x_1; y_1); (x_1; y_2); \dots; (x_2; y_1); (x_2; y_2); \dots; (x_3; y_1); (x_3; y_2); \dots\}$.

Определение 6.1.1 *Распределением* дискретной двумерной случайной величины $(\xi; \eta)$ или *законом распределения* называется соответствие, по которому каждой паре точек $(x_i; y_j) \in \Omega_{\xi\eta}$ сопоставляется положительное число p_{ij} , характеризующее вероятность того, что составляющая ξ примет значение x_i и одновременно с этим составляющая η примет значение y_j таким образом, чтобы это число, которое называется вероятностью события $P((\xi = x_i) \cdot (\eta = y_j)) = P(A)$, удовлетворяло аксиомам Колмогорова (определение 5.1.1.1).

Числовую функцию $P_{\xi\eta} = p_{ij}$ называют *функцией вероятностей*. Функцию вероятностей можно задать в виде таблицы.

Функцию вероятностей $P_{\xi\eta} = p_{ij}$ дискретной случайной величины $(\xi; \eta)$ можно задать в виде таблицы:

$\Omega_\eta \backslash \Omega_\xi$	x_1	x_2	...	x_i	...	x_n	Распреде- ление составляю- щей η
y_1	p_{11}	p_{21}	...	p_{i1}	...	p_{n1}	$P(\eta = y_1)$
y_2	p_{12}	p_{22}	...	p_{i2}	...	p_{n2}	$P(\eta = y_2)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
y_j	p_{1j}	p_{2j}	...	p_{ij}	...	p_{nj}	$P(\eta = y_j)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
y_m	p_{1m}	p_{2m}	...	p_{im}	...	p_{nm}	$P(\eta = y_m)$
Распреде- ление составляю- щей ξ	$P(\xi = x_1)$	$P(\xi = x_2)$	...	$P(\xi = x_i)$	...	$P(\xi = x_n)$	Контроль $\sum_i \sum_j p_{ij} = 1$

В клетках, стоящих на пересечении i -го столбца j -ой строки, указана вероятность того, что случайная величина (ξ, η) примет значение $(x_i; y_j)$, то есть $p_{ij} = P((\xi = x_i) \cdot (\eta = y_j))$. Зная функцию вероятностей $P_{\xi\eta} = p_{ij}$, можно вычислить вероятность появления произвольного события C :

$$P(C) = \sum_{(x_i; y_j) \in C} p_{ij}.$$

Зная распределение $P_{\xi\eta}$ дискретной случайной величины $(\xi; \eta)$, можно найти распределение P_ξ и P_η составляющих ξ и η .

Для нахождения вероятности $P(\xi = x_i)$ надо просуммировать вероятности столбца x_i :

$$P(\xi = x_i) = P((\xi = x_i) \cdot (\eta = y_1)) + P((\xi = x_i) \cdot (\eta = y_2)) + \dots \\ \dots + P((\xi = x_i) \cdot (\eta = y_m)) = \sum_j p_{ij} = p_{i\circ}. \quad (6.1.1)$$

Аналогично, сложив вероятность строки j , получаем вероятность

$$P(\eta = y_j) = P((\xi = x_1) \cdot (\eta = y_j)) + P((\xi = x_2) \cdot (\eta = y_j)) + \dots \\ \dots + P((\xi = x_n) \cdot (\eta = y_j)) = \sum_i p_{ij} = p_{\circ j}. \quad (6.1.2)$$

Распределения вероятностей, определяемыми наборами $p_{i\circ}$ и $p_{\circ j}$, называются *маргинальными распределениями* составляющих ξ и η

дискретной двумерной случайной величины $(\xi; \eta)$, соответственно. На краях таблицы указаны маргинальные распределения («margo» (лат.) – граница, край).

По аналогии с распределением двумерной случайной величины определяется распределение $P_{\xi_1, \xi_2, \dots, \xi_n} = p_{i_1 j_1 \dots n}$ n -мерной случайной величины:

$$p_{i_1 j_1 \dots n} = P((\xi_1 = x_{i_1}) \cdot (\xi_2 = x_{j_2}) \cdot \dots \cdot (\xi_n = x_{j_n})).$$

Распределение произвольной n -мерной случайной величины (дискретной или непрерывной) однозначно может быть выражено через функцию распределения этой случайной величины.

Определение 6.1.2 Функцией распределения n -мерной случайной величины $(\xi_1, \xi_2, \dots, \xi_n)$ называется вероятность появления события $\{(\xi_1 < x_1) \cdot (\xi_2 < x_2) \cdot \dots \cdot (\xi_n < x_n)\}$, которая рассматривается как функция переменных $x_1, x_2, \dots, x_n$: $F(x_1, x_2, \dots, x_n) = P((\xi_1 < x_1) \cdot (\xi_2 < x_2) \cdot \dots \cdot (\xi_n < x_n))$.

Функция распределения двумерной случайной величины $(\xi; \eta)$ имеет вид

$$F(x; y) = P((\xi < x) \cdot (\eta < y)). \quad (6.1.3)$$

Приведём основные свойства функции распределения СВ $(\xi; \eta)$:

- 1) $\forall x, y \in \Omega_{\xi\eta} \quad 0 \leq F(x; y) \leq 1$;
- 2) функция $F(x; y)$ является неубывающей по каждому из аргументов, то есть $\forall x_2 > x_1 \rightarrow F(x_2; y) \geq F(x_1; y)$ и $\forall y_2 > y_1 \rightarrow F(x; y_2) \geq F(x; y_1)$;
- 3) функция $F(x; y)$ непрерывна слева по каждому из аргументов;
- 4) справедливы предельные соотношения:
 - а) $\lim_{x \rightarrow -\infty} F(x; y) = F(-\infty; y) = 0$, б) $\lim_{y \rightarrow -\infty} F(x; y) = F(x; -\infty) = 0$,
 - в) $\lim_{\substack{x \rightarrow -\infty \\ y \rightarrow -\infty}} F(x; y) = F(-\infty; -\infty) = 0$, г) $\lim_{\substack{x \rightarrow +\infty \\ y \rightarrow +\infty}} F(x; y) = F(+\infty; +\infty) = 1$;
- 5) при значении $y \rightarrow +\infty$ функция распределения двумерной случайной величины $(\xi; \eta)$ становится маргинальной функцией распределения составляющей ξ : $F(x; +\infty) = P((\xi < x) \cdot (\eta < +\infty)) = P(\xi < x) = F_1(x)$. Функцию $F_1(x)$ называют *граничной* или *разделённой* функцией распределения составляющей ξ ;
- 6) при значении $x \rightarrow +\infty$ функция распределения двумерной случайной величины $(\xi; \eta)$ становится маргинальной функцией

распределения составляющей η :
 $F(+\infty; y) = P((\xi < +\infty) \cdot (\eta < y)) = P(\eta < y) = F_2(y)$. Функцию $F_2(y)$ называют *граничной* или *разделённой* функцией распределения составляющей η .

7)

Все свойства функции распределения, приведённые для двумерной случайной величины, распространяются на многомерные случайные величины.

1. Функция распределения n -мерной случайной величины неотрицательна и не может быть больше единицы.
2. Если хотя бы одна из составляющих n -мерной случайной величины принимает значение $-\infty$, то функция распределения равна нулю.
3. Если $x_1 = x_2 = \dots = x_n = +\infty$, то функция распределения равна 1.
4. Функция распределения является неубывающей функцией.
5. Функция распределения непрерывна слева по каждому из аргументов.
6. Если какие-либо k из составляющих $\xi_1, \xi_2, \dots, \xi_n$ n -мерной случайной величины принимают значения ∞ , то функция распределения n -мерной случайной величины становится равной функции распределения $(n-k)$ -мерной случайной величины.

Условные распределения составляющих многомерных случайных величин, зависимые и независимые случайные величины, а так же числовые характеристики многомерных случайных величин, рассмотрим на примере двумерной случайной величины $(\xi; \eta)$.

Определение 6.1.3 Условным распределением составляющей ξ дискретной случайной величины $(\xi; \eta)$ называется распределение, вычисляемое при условии, что составляющая η приняла определённое значение.

Из определения следует, что условные распределения характеризуют взаимосвязи между составляющими ξ и η случайной величины $(\xi; \eta)$. Условное распределение составляющей ξ дискретной случайной величины $(\xi; \eta)$ может быть задано условной функцией вероятностей, выражающей вероятность того, что составляющая ξ примет значение x_i при условии, что составляющая η приняла значение y_j :

$$p_{i|j} = P(\xi = x_i | \eta = y_j) = \frac{P((\xi = x_i) \cdot (\eta = y_j))}{P(\eta = y_j)}. \quad (6.1.4)$$

Аналогично, условная функция вероятностей составляющей η имеет вид

$$p_{j|i} = P(\eta = y_j | \xi = x_i) = \frac{P((\xi = x_i) \cdot (\eta = y_j))}{P(\xi = x_i)}. \quad (6.1.5)$$

Определение 6.1.4 Пусть $F(x, y)$ - функция распределения дискретной случайной величины $(\xi; \eta)$, а $F_1(x)$ и $F_2(y)$ - маргинальные функции распределения составляющих ξ и η . Говорят, что две случайные величины, составляющие ξ и η , независимы, если для всех $(x, y) \in \Omega_{\xi\eta}$ справедливо равенство

$$F(x, y) = F_1(x) \cdot F_2(y). \quad (6.1.6)$$

Теорема 6.1.1 Если СВ $(\xi; \eta)$ - дискретная, то равенство

$$p_{ij} = p_{i\circ} \cdot p_{\circ j} \quad (6.1.7)$$

является необходимым и достаточным условием независимости СВ ξ и η .

Вероятность попадания случайной величины $(\xi; \eta)$ в прямоугольник $\Pi = \{(\xi; \eta) | x_1 \leq \xi \leq x_2 \wedge y_1 \leq \eta \leq y_2\}$ находится по формуле

$$P((x_1 \leq \xi \leq x_2) \cdot (y_1 \leq \eta \leq y_2)) = F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) + F(x_1, y_1) \quad (6.1.8)$$

Рассмотрим числовые характеристики двумерных случайных величин.

Основными характеристиками, описывающими центр распределения дискретной случайной величины $(\xi; \eta)$, являются математические ожидания составляющих, определяемые по формулам:

$$m_\xi = \sum_i x_i p_{i\circ} = \sum_i \sum_j x_i p_{ij}, \quad (6.1.9)$$

$$m_\eta = \sum_j y_j p_{\circ j} = \sum_i \sum_j y_j p_{ij}. \quad (6.1.10)$$

Математические ожидания m_ξ и m_η определяют координаты центра распределения. Кроме этих основных характеристик, центр распределения характеризуется модами и медианами (практическое занятие № 11).

Рассеивание составляющих дискретной случайной величины $(\xi; \eta)$ определяется дисперсиями, которые определяются по формулам:

$$D(\xi) = \sigma_\xi^2 = M(\xi - m_\xi)^2 = M(\xi^2) - m_\xi^2, \quad (6.1.11)$$

$$D(\eta) = \sigma_\eta^2 = M(\eta - m_\eta)^2 = M(\eta^2) - m_\eta^2. \quad (6.1.12)$$

Используя формулы (6.1.9) и (6.1.10), имеем:

$$D(\xi) = \sum_i \sum_j (x_i - m_\xi)^2 p_{ij} = \sum_i x_i^2 p_{i\circ} - m_\xi^2, \quad (6.1.13)$$

$$D(\eta) = \sum_i \sum_j (y_j - m_\eta)^2 p_{ij} = \sum_j y_j^2 p_{\circ j} - m_\eta^2. \quad (6.1.14)$$

Положительные квадратные корни из дисперсий составляющих ξ и η называются *средними квадратическими отклонениями* составляющих ξ и η или стандартами: $\sigma_\xi = \sqrt{D(\xi)}$, $\sigma_\eta = \sqrt{D(\eta)}$.

Рассеивание составляющих ξ и η также характеризуется коэффициентами вариации составляющих и их размахами (практическое занятие № 11).

Основной характеристикой, которая описывает связь между составляющими ξ и η случайной величины $(\xi; \eta)$, является *ковариация*, которая также называется *корреляционным моментом* или *моментом связи* $K_{\xi\eta}$ и определяется по формуле (ковариация также обозначается $\mu_{\xi\eta}$):

$$K_{\xi\eta} = \text{cov}(\xi; \eta) = M\left((\xi - m_\xi) \cdot (\eta - m_\eta)\right). \quad (6.1.15)$$

Для дискретной случайной величины расчётную формулу для ковариации можно записать в виде

$$K_{\xi\eta} = \sum_i \sum_j (x_i - m_\xi) \cdot (y_j - m_\eta) \cdot p_{ij} = \sum_i \sum_j x_i y_j p_{ij} - m_\xi \cdot m_\eta. \quad (6.1.16)$$

Для получения числовой характеристики, удобной для сравнения составляющих, эти составляющие ξ и η предварительно нормируют, то есть переходят к случайным величинам $\hat{\xi}$ $\hat{\eta}$, математические ожидания которых равны нулю, а их дисперсия – единице. Если положить $\hat{\xi} = \frac{\xi - m_\xi}{\sigma_\xi}$;

$\hat{\eta} = \frac{\eta - m_\eta}{\sigma_\eta}$, то ковариация нормированных случайных величин называется

коэффициентом корреляции и определяется по формуле

$$\rho_{\xi\eta} = M(\hat{\xi} \cdot \hat{\eta}) = M\left(\left(\frac{\xi - m_\xi}{\sigma_\xi}\right) \cdot \left(\frac{\eta - m_\eta}{\sigma_\eta}\right)\right) = \frac{m_{\xi\eta} - m_\xi m_\eta}{\sigma_\xi \sigma_\eta} = \frac{K_{\xi\eta}}{\sigma_\xi \sigma_\eta}. \quad (6.1.17)$$

Коэффициент корреляции не зависит от выбора единиц измерения случайных величин, то есть является безразмерной случайной величиной. Коэффициент корреляции $\rho_{\xi\eta}$ характеризует тесноту линейной связи между составляющими ξ и η случайной величины $(\xi; \eta)$. Приведём основные свойства коэффициента корреляции.

1) Абсолютная величина коэффициента корреляции не превосходит единицы, то есть $|\rho_{\xi\eta}| \leq 1$.

2) Если $\rho_{\xi\eta} = \pm 1$, то между случайными величинами ξ и η существует линейная функциональная зависимость $(|\rho_{\xi\eta}| = 1 \leftrightarrow \eta = k \cdot \xi + b)$.

3) Если $\rho_{\xi\eta} \neq 0$, то случайные величины ξ и η зависимы.

Зависимость между составляющими ξ и η случайной величины $(\xi; \eta)$ может исследоваться с помощью условных распределений. Если окажется, что условные распределения равны безусловным, то говорят, что случайная величина ξ не зависит от случайной величины η . При этом условные распределения составляющих ξ и η случайной величины $(\xi; \eta)$ характеризуются условными числовыми характеристиками, условными математическими ожиданиями и дисперсиями этих составляющих.

Например, *условное математическое ожидание* составляющей ξ при условии, что дискретная случайная величина η приняла определённое значение y , находится по формуле

$$m_{\xi|\eta} = M(\xi|\eta = y) = \sum_i x_i \cdot p_{i|j}. \quad (6.1.18)$$

Аналогично определяется *условное математическое ожидание* составляющей η :

$$m_{\eta|\xi} = M(\eta|\xi = x) = \sum_j y_j \cdot p_{j|i}. \quad (6.1.19)$$

Условная дисперсия составляющей ξ при условии, что дискретная случайная величина η приняла определённое значение y , находится по формуле

$$\sigma_{\xi|\eta}^2 = D(\xi|\eta = y) = \sum_i (x_i - m_{\xi|\eta})^2 \cdot p_{i|j}. \quad (6.1.20)$$

Аналогично определяется *условная дисперсия* составляющей η :

$$\sigma_{\eta|\xi}^2 = D(\eta|\xi = x) = \sum_j (y_j - m_{\eta|\xi})^2 \cdot p_{j|i}. \quad (6.1.21)$$

НЕПРЕРЫВНЫЕ МНОГОМЕРНЫЕ СЛУЧАЙНЫЕ ВЕЛИЧИНЫ. ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ

Непрерывные многомерные случайные величины

Функция распределения, рассмотренная на предыдущем практическом занятии, существует для любых многомерных случайных величин, как для дискретных, так и непрерывных. Определения и свойства функции распределения дискретных случайных величин будут справедливы для непрерывных случайных величин. На практике чаще всего встречаются многомерные непрерывные случайные величины. Дадим определение двумерной непрерывной случайной величины $(\xi; \eta)$.

Определение 7.1.1.1 Случайная величина $(\xi; \eta)$ называется абсолютно непрерывной, если её *функция распределения* $F(x, y)$ абсолютно непрерывна, то есть существует некоторая функция $f(x, y)$, называемая *плотностью распределения вероятностей*, такая, что

$$F(x, y) = \int_{-\infty}^x dx \int_{-\infty}^y f(x, y) dy, \quad \forall (x, y) \in R^2. \quad (7.1.1.1)$$

Так как функция распределения $F(x, y)$ является неубывающей по каждому аргументу и непрерывной, то плотность распределения вероятностей обладает следующими свойствами:

- 1) $f(x, y) \geq 0$;
- 2) во всех точках $(x; y)$, за исключением, быть может, конечного или счётного множества точек, существует

$$\frac{\partial^2 F}{\partial x \partial y} = \frac{\partial^2}{\partial x \partial y} \left(\int_{-\infty}^x dx \int_{-\infty}^y f(x, y) dy \right) = f(x, y). \quad (7.1.1.2)$$

Из формул (7.1.1.1) и (7.1.1.2) следует, что плотность вероятности и функция распределения однозначно выражаются одна через другую. Это означает, что плотность распределения вероятностей $f(x, y)$ так же, как и функция распределения $F(x, y)$, полностью определяет непрерывную случайную величину (ξ, η) . С геометрической точки зрения, плотность распределения вероятностей $f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y}$ можно рассматривать как предел отношения вероятности попадания случайной величины (ξ, η) в

прямоугольник со сторонами Δx и Δy к площади этого прямоугольника, когда обе стороны этого прямоугольника стремятся к нулю:

$$f(x, y) = \frac{\partial^2 F}{\partial x \partial y} = \lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \frac{P((x < \xi < x + \Delta x) \cdot (y < \eta < y + \Delta y))}{\Delta x \Delta y}. \quad (7.1.1.3)$$

Сформулируем характеристические свойства плотности распределения непрерывной двумерной случайной величины (ξ, η) :

- 1) $f(x, y) \geq 0$;
- 2) $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1$.

Дадим определение плотности вероятности n -мерной случайной величины.

Определение 7.1.1.2 Плотностью распределения вероятностей n -мерной случайной величины $(\xi_1, \xi_2, \dots, \xi_n)$ называется предел отношения вероятности попадания этой случайной величины в бесконечно малую область n -мерного пространства к мере этой области при стягивании в точку

$$f(x_1, x_2, \dots, x_n) = \lim_{\substack{\Delta x_i \rightarrow 0 \\ i=1; n}} \frac{P((x_1 < \xi_1 < x_1 + \Delta x_1) \cdot \dots \cdot (x_n < \xi_n < x_n + \Delta x_n))}{\Delta x_1 \cdot \dots \cdot \Delta x_n}. \quad (7.1.1.4)$$

Оба характеристических свойства плотности вероятности, указанных выше для двумерной случайной величины, справедливы и для n -мерной случайной величины:

- 1) $f(x_1, x_2, \dots, x_n) \geq 0$;
- 2) $\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} f(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n = 1$.

Пусть (ξ, η) - непрерывная случайная величина, то есть такая величина, функция распределения которой представима через плотность распределения вероятностей: $F(x, y) = \int_{-\infty}^x dx \int_{-\infty}^y f(x, y) dy$. Если один из аргументов стремится к значению $+\infty$, а другой остаётся переменным, то функция распределения $F(x, y)$ становится функцией распределения одномерной случайной величины, соответствующей другому аргументу:

$$F_1(x) = F(x, +\infty) = \lim_{y \rightarrow +\infty} F(x, y) = \int_{-\infty}^x dx \int_{-\infty}^{+\infty} f(x, y) dy. \quad (7.1.1.5)$$

$$F_2(y) = F(+\infty, y) = \lim_{x \rightarrow +\infty} F(x, y) = \int_{-\infty}^y dy \int_{-\infty}^{+\infty} f(x, y) dx. \quad (7.1.1.6)$$

Дифференцируя функции распределения составляющих ξ и η случайной величины (ξ, η) по соответствующим переменным x и y , согласно правилу дифференцирования определённого интеграла с переменным верхним пределом интегрирования, получаем формулы для вычисления плотностей вероятностей:

$$f_1(x) = F'_1(x) = \int_{-\infty}^{+\infty} f(x, y) dy; \quad (7.1.1.7)$$

$$f_2(y) = F'_2(y) = \int_{-\infty}^{+\infty} f(x, y) dx. \quad (7.1.1.8)$$

Таким образом, чтобы найти плотность распределения одной из составляющей двумерной случайной величины, необходимо проинтегрировать их совместную плотность распределения вероятностей $f(x, y)$ в бесконечных пределах по аргументу, соответствующему другой составляющей.

Пусть $F(x_1, x_2, \dots, x_n)$ - функция распределения n -мерной случайной величины. Выделим из n -мерной случайной величины $(\xi_1, \xi_2, \dots, \xi_n)$ k -мерную случайную величину $(\xi_1, \xi_2, \dots, \xi_k)$. Тогда функция распределения выделенной случайной величины $(\xi_1, \xi_2, \dots, \xi_k)$ определяется по формуле $F(x_1, x_2, \dots, x_k) = F(x_1, x_2, \dots, x_k, +\infty, \dots, +\infty)$. Плотность распределения k -мерную случайной величины $(\xi_1, \xi_2, \dots, \xi_k)$, выделенной из n -мерной случайной величины, получается в результате интегрирования плотности n -мерной случайной величины по оставшимся $(n - k)$ переменным

$$f(x_1, x_2, \dots, x_k) = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} f(x_1, x_2, \dots, x_n) dx_{k+1} dx_{k+2} \dots dx_n.$$

В частности, функция распределения и плотность распределения вероятностей составляющей ξ_i n -мерной случайной величины $(\xi_1, \xi_2, \dots, \xi_i, \dots, \xi_n)$ находятся по формулам:

$$F_i(x_i) = F(+\infty, +\infty, \dots, x_i, \dots, +\infty);$$

$$f_i(x_i) = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} f(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_{i-1} dx_{i+1} \dots dx_n.$$

Условные распределения составляющих многомерных случайных величин, зависимые и независимые случайные величины, а так же числовые характеристики многомерных случайных величин, рассмотрим на примере двумерной случайной величины $(\xi; \eta)$.

Определение 7.1.1.3 Условным распределением составляющей ξ непрерывной случайной величины $(\xi; \eta)$ называется распределение, вычисляемое при условии, что составляющая η приняла определённое значение.

Условные распределения характеризуют взаимосвязи между составляющими ξ и η случайной величины $(\xi; \eta)$.

Условные распределения составляющих ξ и η непрерывной случайной величины $(\xi; \eta)$ могут быть заданы условными функциями распределения $F(x|y)$ и $F(y|x)$, или условными плотностями $f(x|y)$ и $f(y|x)$.

$$F(x|y) = \frac{\int_{-\infty}^x f(x, y) dx}{f_2(y)} = \frac{\int_{-\infty}^x f(x, y) dx}{\int_{-\infty}^{+\infty} f(x, y) dx}. \quad (7.1.1.9)$$

$$F(y|x) = \frac{\int_{-\infty}^y f(x, y) dy}{f_1(x)} = \frac{\int_{-\infty}^y f(x, y) dy}{\int_{-\infty}^{+\infty} f(x, y) dy}. \quad (7.1.1.10)$$

Дифференцируя функции $F(x|y)$ и $F(y|x)$ по переменным x и y , соответственно получаем формулы для вычисления условных плотностей распределения вероятностей составляющих ξ и η случайной величины $(\xi; \eta)$:

$$f(x|y) = \frac{\partial F(x|y)}{\partial x} = \frac{f(x, y)}{f_2(y)}. \quad (7.1.1.11)$$

$$f(y|x) = \frac{\partial F(y|x)}{\partial y} = \frac{f(x, y)}{f_1(x)}. \quad (7.1.1.12)$$

Используя формулы (7.1.1.11) и (7.1.1.12), можно записать, что

$$f(x, y) = f_1(x) \cdot f(y|x) = f_2(y) \cdot f(x|y). \quad (7.1.1.13)$$

Определение 7.1.1.4 Пусть $F(x, y)$ - функция распределения непрерывной случайной величины $(\xi; \eta)$, а $F_1(x)$ и $F_2(y)$ - маргинальные функции распределения её составляющих. Говорят, что две случайные величины, составляющие ξ и η , независимы, если для всех $(x, y) \in \Omega_{\xi\eta}$ справедливо равенство

$$F(x, y) = F_1(x) \cdot F_2(y). \quad (7.1.1.14)$$

Теорема 7.1.1.1 Если СВ $(\xi; \eta)$ - непрерывная, то равенство

$$f(x, y) = f_1(x) \cdot f_2(y) \quad (7.1.1.15)$$

является необходимым и достаточным условием независимости СВ ξ и η .

Рассмотрим числовые характеристики двумерных случайных величин.

Основными характеристиками, описывающими центр распределения непрерывной случайной величины $(\xi; \eta)$, являются математические ожидания составляющих, определяемые по формулам:

$$m_\xi = \int_{-\infty}^{+\infty} x f_1(x) dx = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x f(x, y) dx dy, \quad (7.1.1.16)$$

$$m_\eta = \int_{-\infty}^{+\infty} y f_2(y) dy = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} y f(x, y) dx dy. \quad (7.1.1.17)$$

Математические ожидания m_ξ и m_η определяют координаты центра распределения. Кроме этих основных характеристик, центр распределения характеризуется модами и медианами (практическое занятие № 11).

Рассеивание составляющих случайной величины $(\xi; \eta)$ определяется дисперсиями, которые определяются по формулам (6.1.10) и (6.1.11).

Используя формулы (7.1.1.16) и (7.1.1.17), имеем:

$$D(\xi) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - m_\xi)^2 f(x, y) dx dy = \int_{-\infty}^{+\infty} x^2 f_1(x) dx - m_\xi^2, \quad (7.1.1.18)$$

$$D(\eta) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (y - m_\eta)^2 f(x, y) dx dy = \int_{-\infty}^{+\infty} y^2 f_2(y) dy - m_\eta^2. \quad (7.1.1.19)$$

Положительные квадратные корни из дисперсий составляющих ξ и η называются *средними квадратическими отклонениями* составляющих ξ и η или стандартами: $\sigma_\xi = \sqrt{D(\xi)}$, $\sigma_\eta = \sqrt{D(\eta)}$.

Основной характеристикой, которая описывает связь между составляющими ξ и η случайной величины $(\xi; \eta)$, является *ковариация*,

которая также называется *корреляционным моментом* или *моментом связи* $K_{\xi\eta}$ и определяется по формуле (6.1.4).

Для непрерывной случайной величины расчётную формулу для ковариации можно записать в виде

$$K_{\xi\eta} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - m_{\xi})(y - m_{\eta}) f(x, y) dx dy = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x y f(x, y) dx dy - m_{\xi} \cdot m_{\eta}. \quad (7.1.1.20)$$

Коэффициент корреляции определяется по формуле (6.1.16).

Условные распределения составляющих ξ и η случайной величины $(\xi; \eta)$ характеризуются условными числовыми характеристиками, условными математическими ожиданиями и дисперсиями этих составляющих.

Условное математическое ожидание составляющей ξ при условии, что непрерывная случайная величина η приняла определённое значение y , находится по формуле

$$m_{\xi|\eta} = M(\xi|\eta = y) = \int_{-\infty}^{+\infty} x f(x|y) dx. \quad (7.1.1.21)$$

Аналогично определяется *условное математическое ожидание* составляющей η :

$$m_{\eta|\xi} = M(\eta|\xi = x) = \int_{-\infty}^{+\infty} y f(y|x) dy. \quad (7.1.1.22)$$

Условная дисперсия составляющей ξ при условии, что непрерывная случайная величина η приняла определённое значение y , находится по формуле

$$\sigma_{\xi|\eta}^2 = D(\xi|\eta = y) = \int_{-\infty}^{+\infty} (x - m_{\xi|\eta})^2 f(x|y) dx. \quad (7.1.1.23)$$

Аналогично определяется *условная дисперсия* составляющей η :

$$\sigma_{\eta|\xi}^2 = D(\eta|\xi = x) = \int_{-\infty}^{+\infty} (y - m_{\eta|\xi})^2 f(y|x) dy. \quad (7.1.1.24)$$

Из распределений двумерной случайной величины $(\xi; \eta)$ рассмотрим нормальное распределение, которое наиболее часто применяется на

практике. Случайная величина $(\xi; \eta)$ имеет нормальное распределение, если её плотность вероятности имеет вид

$$f(x, y) = \frac{1}{2\pi\sigma_\xi\sigma_\eta\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left(\frac{(x-m_\xi)^2}{\sigma_\xi^2} - 2\rho\frac{(x-m_\xi)(y-m_\eta)}{\sigma_\xi\sigma_\eta} + \frac{(y-m_\eta)^2}{\sigma_\eta^2}\right)}, \quad (7.1.1.25)$$

где m_ξ, m_η - математические ожидания составляющих ξ и η ; σ_ξ, σ_η - их средние квадратические отклонения; ρ - коэффициент корреляции между составляющими.

Если случайная величина $(\xi; \eta)$ имеет нормальное распределение, то распределение её составляющих ξ и η являются тоже нормальными

распределениями: $f_1(x) = \frac{1}{\sigma_\xi\sqrt{2\pi}} e^{-\frac{(x-m_\xi)^2}{2\sigma_\xi^2}}, f_2(y) = \frac{1}{\sigma_\eta\sqrt{2\pi}} e^{-\frac{(y-m_\eta)^2}{2\sigma_\eta^2}}.$

Вероятность попадания случайной точки, распределённой по нормальному закону, в прямоугольник $G = \{(x, y) | \alpha \leq \xi \leq \beta \wedge \delta \leq \eta \leq \gamma\}$ с осями, параллельными главным осям рассеивания, выражается формулой

$$P((\xi, \eta) \subset G) = \left(\Phi\left(\frac{\beta - m_\xi}{\sigma_\xi}\right) - \Phi\left(\frac{\alpha - m_\xi}{\sigma_\xi}\right) \right) \left(\Phi\left(\frac{\gamma - m_\eta}{\sigma_\eta}\right) - \Phi\left(\frac{\delta - m_\eta}{\sigma_\eta}\right) \right) \quad (7.1.1.26)$$

Вероятность попадания случайной точки, распределённой по нормальному закону, в область R_n , ограниченную эллипсом рассеивания, равна

$$P((\xi, \eta) \subset G) = 1 - e^{-n^2/2},$$

где n - размеры полуосей эллипса в среднеквадратичных отклонениях: $n = a/\sigma_\xi = b/\sigma_\eta.$

Если $\sigma_\xi = \sigma_\eta = \sigma$, то рассеивание по нормальному закону называется круговым. При круговом нормальном рассеивании с центром рассеивания, совпадающим с началом координат, расстояние R от начала координат до случайной точки (ξ, η) распределено по закону Рэлея с плотностью вероятностей

$$f(r) = \begin{cases} \frac{r}{\sigma^2} e^{-r^2/2\sigma^2} & \text{при } r \geq 0; \\ 0 & \text{при } r < 0, \end{cases} \quad (7.1.1.27)$$

и функцией распределения вероятностей

$$F(r) = \begin{cases} 1 - e^{-r^2/2\sigma^2} & \text{при } r \geq 0; \\ 0 & \text{при } r < 0, \end{cases} \quad (7.1.1.28)$$

Предельные теоремы теории вероятностей

Теорема 7.1.2.1 (о неравенстве Чебышева). Вероятность того, что отклонение случайной величины от её математического ожидания будет по абсолютной величине не меньше любого положительного числа ε , ограничена сверху величиной σ^2/ε^2 :

$$P(|\xi - m| \geq \varepsilon) \leq \sigma^2/\varepsilon^2. \quad (7.1.2.1)$$

Так как $\{|\xi - m| \geq \varepsilon\} \cup \{|\xi - m| < \varepsilon\} = \Omega$, то неравенство Чебышева можно записать в виде $P(|\xi - m| < \varepsilon) \geq 1 - \sigma^2/\varepsilon^2$.

Теорема 7.1.2.2 (Чебышева). Пусть $\xi_i, i = \overline{1, \infty}$ - последовательность независимых случайных величин, у которых дисперсии $D(\xi_i) \leq C$, и все математические ожидания $M(\xi_i)$ конечны. Тогда среднее арифметическое этих случайных величин сходится по вероятности к среднему арифметическому их математических ожиданий, то есть

$$P\left(\left|\frac{1}{n} \sum_{i=1}^n \xi_i - \frac{1}{n} \sum_{i=1}^n M(\xi_i)\right| < \varepsilon\right) \geq 1 - \frac{D(\xi)}{n \cdot \varepsilon^2}. \quad (7.1.2.2)$$

Теорема 7.1.2.3 (Маркова). Пусть случайные величины $\xi_i, i = \overline{1, \infty}$ удовлетворяют условию $\frac{1}{n^2} D\left(\sum_{i=1}^n \xi_i\right) \rightarrow 0$ при значении $n \rightarrow +\infty$. Тогда при любом положительном значении ε выполняется предельное соотношение (7.1.2.2).

Следствие. Пусть независимые случайные величины $\xi_n, n = \overline{1, \infty}$, имеют одинаковые математические ожидания $M(\xi_i) = m$ для всех $i = \overline{1, n}$, а дисперсии удовлетворяют условию $D(\xi_i) \leq C$. Тогда среднее арифметическое случайных величин сходится по вероятности к числу m при условии $n \rightarrow +\infty$:

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{1}{n} \sum_{i=1}^n \xi_i - m\right| < \varepsilon\right) = 1.$$

Второе важное следствие из теоремы Чебышева сформулируем в виде теоремы Бернулли, которое называется законом больших чисел в схеме Бернулли.

Теорема 7.1.2.4 (Бернулли). Если в каждом из n независимых испытаний вероятность наступления события A равна p , то относительная частота сходится по вероятности к числу p при значении $n \rightarrow +\infty$, то есть

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{m}{n} - p\right| < \varepsilon\right) = 1. \quad (7.1.2.3)$$

В результате получаем следующую оценку

$$P\left(\left|\frac{m}{n} - p\right| < \varepsilon\right) \geq 1 - \frac{pq}{n\varepsilon^2}. \quad (7.1.2.4)$$

Теорема 7.1.2.5 (центральная предельная теорема или теорема Ляпунова). Пусть $\xi_i, i = \overline{1, \infty}$ - последовательность независимых случайных величин, для каждой из которых существует математическое ожидание $M(\xi_i) = m_i$, дисперсия $D(\xi_i) = \sigma_i^2$ и третий центральный абсолютный момент $M|\xi_i - m_i|^3$. Пусть, кроме этого, выполнено условие Ляпунова

$$\lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n M|\xi_i - m_i|^3}{\left(\sum_{i=1}^n \sigma_i^2\right)^{3/2}} = 0.$$

Тогда при неограниченном увеличении n распределение случайной величины $\eta_n = \sum_{i=1}^n \xi_i$ сходится по распределению к нормальному

распределению с математическим ожиданием $M(\eta_n) = \sum_{i=1}^n M(\xi_i)$ и

дисперсией $\sigma_{\eta_n}^2 = \sum_{i=1}^n \sigma_i^2$, то есть для случайной величины $\eta_n = \sum_{i=1}^n \xi_i$ при

значении $n \rightarrow \infty$ выполняется соотношение

$$P(\eta_n < y) \xrightarrow[n \rightarrow \infty]{p} \frac{1}{\sqrt{2\pi} \cdot \sigma_{\eta_n}} \cdot \int_{-\infty}^y e^{-\frac{(\eta_n - M(\eta_n))^2}{2\sigma_{\eta_n}^2}} dy. \quad (7.1.2.5)$$