МНОГОМЕРНЫЕ СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

1. **Способы задания двумерных СВ**
2. **Критерии независимости двух СВ**
3. **Числовые характеристики системы СВ**
4. **Двумерное нормальное распределение**

1. Способы задания двумерных СВ

Пусть имеется некоторое вероятностное пространство $(\Omega, \mathfrak{F}, P)$. ***Двумерной СВ*** $(\xi; \eta)$ называется совокупность двух числовых функций, заданных на одном и том же пространстве элементарных исходов Ω, если для любых действительных чисел x, y существует $P(\xi < x,\ \eta < y)$.

Двумерная СВ $(\xi; \eta)$ называется ***дискретной***, если обе ее составляющие ξ и η являются дискретными СВ.

Двумерная СВ $(\xi; \eta)$ называется ***непрерывной***, если обе ее составляющие ξ и η являются непрерывными СВ.

Если одна из СВ дискретная, а другая непрерывная, то двумерная СВ относится к смешанному типу.

Универсальным способом задания двумерной СВ является функция распределения.

Функция распределения двумерной СВ $(\xi; \eta)$ – это функция двух действительных переменных x и y, которая определяется с помощью равенства

$$\boxed{F_{\xi;\,\eta}(x; y) = P(\xi < x; \eta < y).} \tag{1}$$

Отметим, что для наглядности значения двумерной СВ $(\xi; \eta)$ могут изображаться точками на плоскости Oxy. Геометрически (1) означает вероятность попадания значения СВ в четверть плоскости левее и ниже точки с координатами $(x; y)$ (рис. 1).

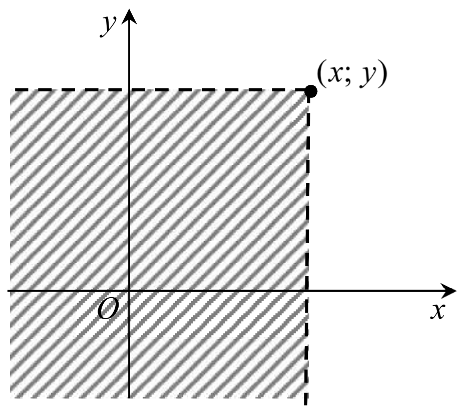

Рис. 1. К понятию функции распределения двумерной СВ

Свойства функции распределения двумерной СВ.

1. $0 \le F(x;\, y) \le 1$ при всех $(x;\, y)$.

2. Функция распределения является неубывающей по каждому из своих аргументов:

$$F(x_1;\, y) \le F(x_2;\, y), \text{ если } x_1 < x_2;$$

$$F(x;\, y_1) \le F(x;\, y_2), \text{ если } y_1 < y_2.$$

3. $F(-\infty;\, y) = F(x;\, -\infty) = F(-\infty;\, -\infty) = 0.$

4. $F(+\infty;\, +\infty) = 1.$

5. $F_{\xi;\,\eta}(x;\, +\infty) = F_{\xi}(x)$ – функция распределения СВ ξ;

$F_{\xi;\,\eta}(+\infty;\, y) = F_{\eta}(y)$ – функция распределения СВ η.

6. Функция распределения непрерывна слева по каждому из своих аргументов.

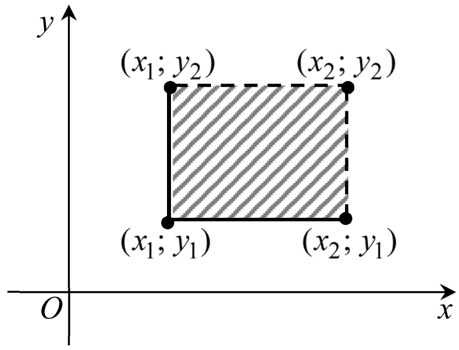

Рис. 2. К вычислению вероятности попадания двумерной СВ в прямоугольник

Можно показать, что вероятность попадания двумерной СВ $(\xi;\, \eta)$ в прямоугольник (рис. 2) равна

$$\boxed{P(x_1 \le \xi < x_2;\, y_1 \le \eta < y_2) = F(x_2;\, y_2) - F(x_1;\, y_2) - F(x_2;\, y_1) + F(x_1;\, y_1).}$$

Распределение *дискретной* двумерной СВ $(\xi;\, \eta)$ проще задать, перечислив все возможные значения этой СВ, т. е. пары чисел $(x_i;\, y_j), 1 \le i \le n, 1 \le j \le m$, и соответствующие им вероятности $P(\xi = x_i;\, \eta = y_j) = p_{ij}$, причем сумма всех вероятностей равна 1: $\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij} = 1.$

Распределение двумерной дискретной СВ $(\xi;\, \eta)$ удобно записывать в виде таблицы.

$\xi \diagdown \eta$	y_1	y_2	$\dots$	y_m	$P(\xi = x_i)$
x_1	p_{11}	p_{12}	$\dots$	p_{1m}	p_1^*
x_2	p_{12}	p_{22}	$\dots$	p_{2m}	p_2^*
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
x_n	p_{n1}	p_{n2}	$\dots$	p_{nm}	p_n^*
$P(\xi = y_j)$	p_1^{**}	p_2^{**}	$\dots$	p_m^{**}	$\sum p_{ij} = 1$

Здесь

$$p_i^* = P(\xi = x_i) = \sum_{j=1}^{m} P(\xi = x_i; \eta = y_j) = \sum_{j=1}^{m} p_{ij};$$

$$p_j^{**} = P(\eta = y_j) = \sum_{i=1}^{n} P(\xi = x_i; \eta = y_j) = \sum_{i=1}^{n} p_{ij}.$$

Таким образом, чтобы по таблице двумерного распределения найти законы распределения составляющих, нужно просуммировать вероятности по строкам – для одной СВ, по столбцам – для другой СВ.

Распределение непрерывной двумерной случайной величины может быть задано с помощью плотности распределения.

Функция $f_{\xi;\eta}(x;y)$ называется ***плотностью распределения*** двумерной СВ (ξ; η), если

$$F_{\xi;\eta}(x;y) = P(\xi < x; \eta < y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{\xi;\eta}(x;y)dxdy.$$

Следовательно, плотность распределения двумерной СВ (ξ; η) может быть найдена по формуле

$$\boxed{f_{\xi;\eta}(x;y) = \frac{\partial^2 F_{\xi;\eta}(x;y)}{\partial x \partial y}.} \qquad (2)$$

Свойства плотности распределения.

1. $f(x;y) \geq 0.$

2. $\boxed{\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x;y)dxdy = 1.}$

3. Вероятность попадания СВ (ξ; η) в область D равна

$$\boxed{P((\xi;\eta) \in D) = \iint_D f(x;y)dxdy.} \qquad (3)$$

4. Плотности распределения составляющих двумерной СВ $(\xi;\eta)$:

$$f_{\xi}(x)=\int\limits_{-\infty}^{+\infty} f_{\xi;\,\eta}(x;y)dy; \quad f_{\eta}(y)=\int\limits_{-\infty}^{+\infty} f_{\xi;\,\eta}(x;y)dx.$$

Говорят, что двумерная СВ (ξ; η) ***распределена равномерно в области*** *D*, если ее плотность распределения постоянна внутри области *D* и равна 0 вне ее:

$$f_{\xi;\,\eta}(x;\,y)=\begin{cases}\dfrac{1}{S_D}, & \text{если } (x;\,y)\in D,\\ 0, & \text{если } (x;\,y)\notin D,\end{cases}$$

где S_D – площадь области *D*.

2. Критерии независимости двух СВ

Две СВ ξ и η называются ***независимыми***, если для любых числовых множеств X и Y события $\{\xi\in X\}$ и $\{\eta\in Y\}$ независимы, т. е. $P(\xi\in X,\eta\in Y)=P(\xi\in X)P(\eta\in Y)$.

***Критерий независимости двух произвольных СВ*: ** СВ ξ и η независимы тогда и только тогда, когда

$$\boxed{F_{\xi;\,\eta}(x;\,y)=F_{\xi}(x)F_{\eta}(y)}$$

для всех действительных *x* и *y*, т. е. их совместная функция распределения представима в виде произведения функций распределения этих СВ.

***Критерий независимости двух дискретных СВ*: ** дискретные СВ ξ и η независимы тогда и только тогда, когда

$$\boxed{P(\xi=x_i;\,\eta=y_j)=P(\xi=x_i)P(\eta=y_j)}$$

для всех возможных пар значений $(x_i;\,y_j),1\le i\le n,1\le j\le m,$ этих СВ.

***Критерий независимости двух независимых СВ*: ** Если двумерная СВ $(\xi;\eta)$ имеет плотность распределения, то СВ ξ и η независимы тогда и только тогда, когда их совместная плотность распределения представима в виде произведения плотностей распределения этих СВ:

$$\boxed{f_{\xi;\,\eta}(x;\,y)=f_{\xi}(x)f_{\eta}(y)}$$

для всех *x* и *y*.

3. Числовые характеристики системы СВ

Основными *числовыми характеристиками* двумерной СВ $(\xi;\eta)$ являются математические ожидания и дисперсии ее составляющих, т. е. СВ ξ и η, а также корреляционный момент и коэффициент корреляции.

Для *дискретной* двумерной СВ $(\xi;\eta)$ с $p_{ij}=P(\xi=x_i;\eta=y_j)$, $1\le i\le n, 1\le j\le m$, математические ожидания СВ ξ и η равны

$$\boxed{M\xi=\sum_{i=1}^{n}\sum_{j=1}^{m}x_i p_{ij};}\quad \boxed{M\eta=\sum_{i=1}^{n}\sum_{j=1}^{m}y_j p_{ij}.}$$

Более того, при некоторых ограничениях на функцию $g(x;y)$ для математического ожидания от функции двух дискретных СВ имеет место формула

$$Mg(\xi;\eta)=\sum_{i=1}^{n}\sum_{j=1}^{m}g(x_i;y_j)p_{ij}.$$

Аналогично, для *непрерывной* двумерной СВ $(\xi;\eta)$ с плотностью распределения $f_{\xi;\eta}(x;y)$ при некоторых ограничениях на функцию $g(x;y)$ для математического ожидания от функции двух непрерывных СВ имеет место формула

$$Mg(\xi;\eta)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}g(x;y)f_{\xi;\eta}(x;y)dxdy.$$

Следовательно,

$$\boxed{M\xi=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xf_{\xi;\eta}(x;y)dxdy;}\qquad \boxed{M\eta=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}yf_{\xi;\eta}(x;y)dxdy.}$$

Дисперсии $D\xi$ и $D\eta$ можно найти по формулам $D\xi=M(\xi-M\xi)^2$ или $D\xi=M(\xi^2)-(M\xi)^2$.

Математические ожидания $M\xi$, $M\eta$ и дисперсии $D\xi$, $D\eta$ характеризуют среднее значение и рассеяние каждой из составляющих двумерной СВ.

Для характеристики степени зависимости двух СВ ξ и η вводится новая числовая характеристика – ***ковариация*** (***корреляционный момент***), которая определяется формулой

$$\boxed{\operatorname{cov}(\xi;\eta)=K_{\xi;\eta}=M(\xi-M\xi)(\eta-M\eta).}$$

Для вычисления ковариации на практике, как правило, удобнее пользоваться следующей формулой:

$$\boxed{\operatorname{cov}(\xi;\eta)=M(\xi\eta)-M\xi M\eta.}$$

Если $\operatorname{cov}(\xi;\eta)=0$, то СВ ξ и η называются ***некоррелированными***.

Если СВ ξ и η независимы, то $\operatorname{cov}(\xi;\eta)=0$. Обратное, вообще говоря, неверно: если $\operatorname{cov}(\xi;\eta)=0$, то это не означает, что СВ ξ и η независимы.

Следует помнить, что:

$$\boxed{\begin{array}{c} \xi \text{ и } \eta \text{ независимы } \Rightarrow \ \xi \text{ и } \eta \text{ некоррелированы } (\mathrm{cov}(\xi;\eta)=0); \\ \xi \text{ и } \eta \text{ зависимы } \Leftarrow \ \xi \text{ и } \eta \text{ коррелированы } (\mathrm{cov}(\xi;\eta)\neq 0). \end{array}}$$

Поскольку ковариация имеет размерность, равную произведению размерностей СВ ξ и η, то для удобства анализа степени зависимости двух СВ вводят безразмерную характеристику – коэффициент корреляции.

Коэффициентом корреляции двух СВ ξ и η – это число, равное

$$\boxed{r_{\xi;\,\eta}=\frac{\mathrm{cov}(\xi;\eta)}{\sqrt{D\xi D\eta}}.}$$

Свойства коэффициента корреляции.

1. $\boxed{-1\le r_{\xi;\,\eta}\le 1.}$

2. Если СВ ξ и η независимы, то $r_{\xi;\,\eta}=0.$

Обратное утверждение неверно: если $r_{\xi;\,\eta}=0,$ то СВ ξ и η *могут быть как зависимыми, так и независимыми*.

3. СВ ξ и η связаны линейной зависимостью в том и только том случае, если $r_{\xi;\,\eta}=\pm 1$:

$$\boxed{\begin{array}{c} \eta=k\xi+b,\ k>0 \quad \Leftrightarrow \quad r_{\xi;\,\eta}=+1; \\ \eta=k\xi+b,\ k<0 \quad \Leftrightarrow \quad r_{\xi;\,\eta}=-1. \end{array}}$$

Итак, *коэффициент корреляции $r_{\xi;\,\eta}$ показывает степень линейной зависимости между СВ ξ и η.*

4. Двумерное нормальное распределение

Двумерная СВ $(\xi;\eta)$ имеет ***нормальный (гауссовский) закон распределения***, если ее плотность распределения

$$f_{\xi;\eta}(x;y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}\exp\left\{-\frac{1}{2(1-r^2)}\left(\frac{(x-a_1)^2}{\sigma_1^2}-2r\frac{(x-a_1)(y-a_2)}{\sigma_1\sigma_2}+\frac{(y-a_2)^2}{\sigma_2^2}\right)\right\}.$$

Параметры двумерного нормального распределения имеют следующий смысл:

$$a_1=M\xi;\ a_2=M\eta;\ \sigma_1^2=D\xi;\ \sigma_2^2=D\eta;\ r=r_{\xi;\,\eta}.$$

Свойства двумерного нормального распределения.

1. Если СВ $(\xi;\eta)$ имеет двумерное нормальное распределение, то $\xi \sim \mathcal{N}(a_1;\sigma_1)$, $\eta \sim \mathcal{N}(a_2;\sigma_2)$.

2. Если СВ $(\xi;\eta)$ имеет двумерное нормальное распределение, то СВ ξ и η независимы тогда и только тогда, когда они некоррелированы, т. е. $r_{\xi;\eta}=0$.

3. Если СВ $(\xi;\eta)$ имеет двумерное нормальное распределение, то условное распределение одной компоненты при фиксированном значении другой также является нормальным:

распределение СВ ξ при условии $\eta = y$ является нормальным с

$$M(\xi\,|\,\eta=y)=a_1+r\frac{\sigma_1}{\sigma_2}(y-a_2), \qquad D(\xi\,|\,\eta=y)=\sigma_1^2(1-r^2);$$

распределение СВ η при условии $\xi = x$ является нормальным с

$$M(\eta\,|\,\xi=x)=a_2+r\frac{\sigma_2}{\sigma_1}(x-a_1), \qquad D(\eta\,|\,\xi=x)=\sigma_2^2(1-r^2).$$

Зависимость между значениями одной СВ и условным математическим ожиданием другой СВ называется ***регрессионной***, а функции, выражающие эту зависимость, называются ***функциями регрессии***:

$M(\xi\,|\,\eta=y)=\psi(y)$ – функция регрессии ξ на η;

$M(\eta\,|\,\xi=x)=\varphi(x)$ – функция регрессии η на ξ.

Таким образом, если СВ $(\xi;\eta)$ имеет двумерное нормальное распределение, то обе функции регрессии (ξ на η и η на ξ) являются линейными.