Раздел: Теория вероятностей и математическая статистика

# Тема: ***Двумерные случайные величины***

Лектор Пахомова Е.Г.

2015 г.

# §10. Двумерные случайные величины

## 1. Многомерные СВ

ОПРЕДЕЛЕНИЕ. ***n-мерной случайной величиной*** (***n-мерным случайным вектором***) *называется упорядоченная совокупность n одномерных СВ* $(X_1, X_2, \dots, X_n)$.

При этом каждая СВ $X_i$ называется ***составляющей*** или ***компонентой*** *n*-мерной СВ.

Геометрическая интерпретация:

$(X_1, X_2, \dots, X_n)$ – случайная точка в *n*-мерном пространстве с координатами $(X_1, X_2, \dots, X_n)$.

Различают два вида $n$-мерных СВ – дискретные и непрерывные.

ОПРЕДЕЛЕНИЕ. *$n$-мерная СВ называется* ***дискретной****, если все ее компоненты – дискретные.*

*$n$-мерная СВ называется* ***непрерывной****, если все ее компоненты – непрерывные.*

$n$-мерная СВ задается с помощью ***закона распределения*** – *соотношения, устанавливающего связь между значениями $n$-мерной СВ и соответствующими им вероятностями.*

Виды законов распределения:

для $n$-мерных ДСВ: 1) $n$-мерная таблица распределения;

2) функция распределения;

для $n$-мерных НСВ: 1) функция распределения;

2) плотность распределения вероятностей.

Для простоты изложения в дальнейшем будем рассматривать двумерные СВ. На общий случай все результаты легко переносятся.

# 2. Способы задания двумерных СВ
## а) Таблица распределения

Пусть X,Y – одномерные ДСВ

X может принимать значения $x_1, x_2, \ldots, x_n, \ldots$

Y может принимать значения $y_1, y_2, \ldots, y_n, \ldots$

$\Rightarrow$ СВ (X,Y) – двумерная СВ с множеством возможных значений $(x_i; y_j)$, $i = 1,2,\ldots$, $j = 1,2,\ldots$ (множество значений СВ – конечно или счётно).

Пусть $P(x_i; y_j) = P(X = x_i; Y = y_j) = p_{ij}$

***Таблицей распределения*** двумерной ДСВ (X,Y) называется двумерная таблица, в которой перечислены все возможные значения $(x_i; y_j)$ и соответствующие им вероятности:

| | $x_1$ | $x_2$ | … | $x_i$ | … |
|---|---|---|---|---|---|
| $y_1$ | $p_{11}$ | $p_{21}$ | … | $p_{i1}$ | … |
| $y_2$ | $p_{12}$ | $p_{22}$ | … | $p_{i2}$ | … |
| … | … | … | … | … | … |
| $y_j$ | $p_{1j}$ | $p_{2j}$ | … | $p_{ij}$ | … |
| … | … | … | …. | … | …. |

***Замечание***. Поскольку в результате опыта СВ обязательно примет одно из своих возможных значений ($x_i$; $y_j$), и эти события всегда несовместны, то

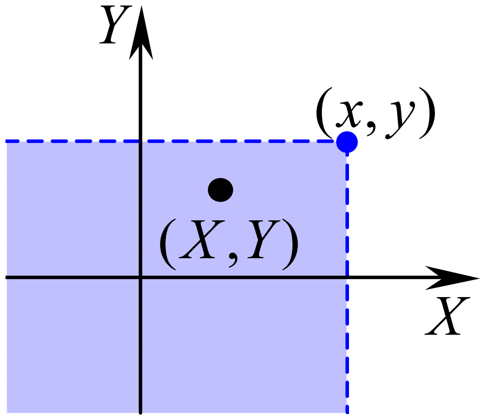
$$\sum_i \sum_j p_{ij} = 1.$$

# б) Функция распределения двумерной ДСВ и НСВ

Пусть (X,Y) – двумерная СВ (дискретная или непрерывная)

ОПРЕДЕЛЕНИЕ. ***Функций распределения*** (*или* ***интегральным законом распределения***) двумерной СВ (X,Y) (*дискретной или непрерывной*) *называется функция* $F(x,y)$, *определяемая равенством* $F(x,y) = P\{X < x,\ Y < y\}$.

Геометрически $F(x,y)$ – вероятность попадания случайной точки (X,Y) в квадрант, расположенный левее и ниже точки $(x;y)$ .




Так как при нахождении $F(x,y)$ СВ X и Y рассматриваются совместно, то $F(x,y)$ называют также ***совместной функцией распределения СВ*** X и Y.

# СВОЙСТВА ФУНКЦИИ РАСПРЕДЕЛЕНИЯ

1) Значения $F(x,y)$ принадлежат [0; 1].

2) Функция $F(x,y)$ не убывает по каждому аргументу в отдельности, т.е.

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

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

3) Справедливы предельные соотношения:

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

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

4) Зная функцию $F(x,y)$ можно найти функции распределения $F(x)$ и $F(y)$ ее составляющих X и Y:

$$F(x) = F(x;+\infty) = \lim_{y\to+\infty} F(x,y),$$

$$F(y) = F(+\infty;y) = \lim_{x\to+\infty} F(x,y).$$

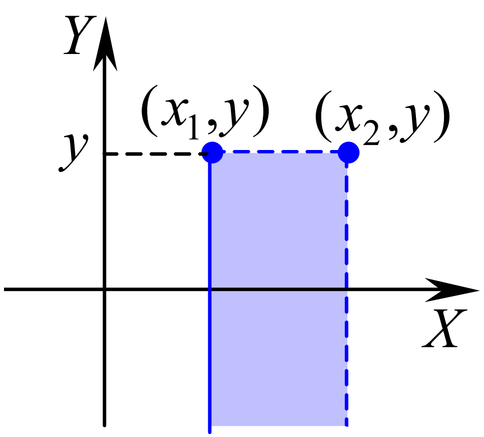
5) $F(x,y)$ – непрерывная слева по каждому аргументу.

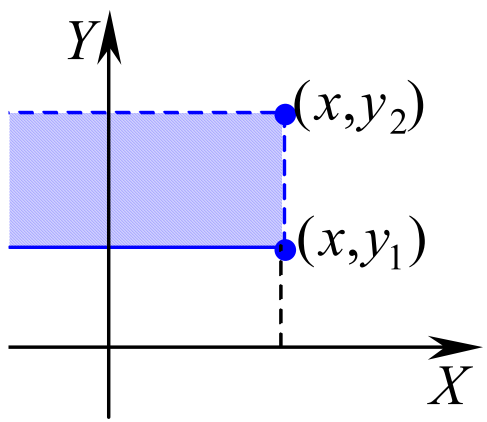
6) Вероятность попадания случайной точки в полуполосу:

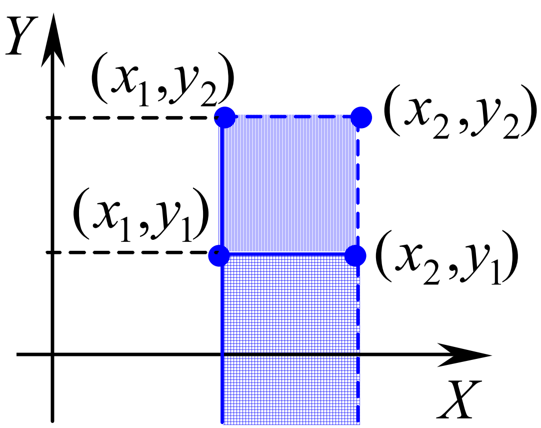
$$P(x_1 \le X < x_2,\ Y < y) = F(x_2,y) - F(x_1,y);$$
$$P(X < x,\ y_1 \le Y < y_2) = F(x,y_2) - F(x,y_1).$$

7) Вероятность попадания случайной точки в прямоугольник:

$$P(x_1 \le X < x_2,\ y_1 \le Y < y_2) =$$
$$=[F(x_2,y_2) - F(x_1,y_2)] - [F(x_2,y_1) - F(x_1,y_1)]$$



## в) Плотность распределения вероятностей

Пусть (X,Y) – непрерывная двумерная СВ,

$F(x,y)$ – функция распределения СВ (X,Y),

$F(x,y)$ – имеет непрерывную $\dfrac{\partial^2 F(x,y)}{\partial x \partial y}$ (всюду, за исключением возможно конечного числа кривых).

ОПРЕДЕЛЕНИЕ. ***Плотностью распределения вероятностей*** (***дифференциальной функцией распределения***) двумерной СВ (X,Y) называется функция

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y}$$

График функции $f(x,y)$ называют ***поверхностью распределения*** двумерной СВ.

# СВОЙСТВА ПЛОТНОСТИ ВЕРОЯТНОСТЕЙ

Аналогичны свойствам плотности распределения вероятностей для одномерной СВ. А именно:

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

2) Для функции $f(x,y)$ выполняется условие нормирования:

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

3) Для функций $f(x,y)$ и $F(x,y)$ справедливо равенство:

$$F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(x,y)dxdy.$$

5) Вероятность попадания двумерной СВ в заданную область.

Пусть (σ) – произвольная квадрируемая плоская область.

Тогда

$$P[(X,Y)\in(\sigma)]=\iint\limits_{(\sigma)} f(x,y)dxdy.$$

$\Rightarrow$ вероятность попадания случайной точки (X,Y) в плоскую область (σ) численно равна объему цилиндрического тела с основанием (σ) и ограниченного функцией $f(x,y)$ (вероятностный смысл $f(x,y)$).

## 3. Двумерные СВ с независимыми компонентами

Пусть X,Y – одномерные СВ.

ОПРЕДЕЛЕНИЕ 1. *СВ X и Y называются* ***независимыми****, если* $\forall x,y\in\mathbb{R}$ *независимы события* «$X < x$» и «$Y < y$». *В противном случае СВ X и Y называются зависимыми.*

Из определения независимых событий и теоремы умножения вероятностей получаем:

События «$X < x$» и «$Y < y$» – независимы $\Leftrightarrow$ вероятность их совместного наступления равна произведению их вероятностей, т.е.

$$P(X < x\ ,\ \ Y < y) = P(X < x)\cdot P(Y < y)\ .$$

$\Rightarrow$ можно дать другое определение независимых СВ

ОПРЕДЕЛЕНИЕ 2. *СВ X и Y называются* ***независимыми****, если функция распределения двумерной СВ (X,Y) равна произведению функций распределения составляющих ее одномерных СВ, т.е.*

$$F(x,y) = F(x)\cdot F(y).$$

Пусть X,Y – независимые СВ,

$f(x)$, $f(y)$ – плотность распределения вероятностей СВ X и Y соответственно,

$f(x,y)$ – плотность вероятностей двумерной СВ (X,Y).

Справедлива следующая теорема.

ТЕОРЕМА.

*СВ* X *и* Y *независимы* $\Leftrightarrow$ *плотность вероятностей их совместного распределения* (т.е. двумерной СВ (X,Y)) *равна произведению их плотностей распределения вероятностей*, т.е. $f(x,y) = f(x) \cdot f(y)$.

ДОКАЗАТЕЛЬСТВО

## 4. Числовые характеристики двумерной СВ

Основные числовые характеристики двумерной СВ (X,Y):

а) математические ожидания M[X] и M[Y],

б) дисперсии D[X] и D[Y],

в) ковариация

г) коэффициент корреляции.

## а) Математические ожидания M[X] и M[Y]

Математические ожидания M[X] и M[Y] определяются по следующим формулам:

1) если (X,Y) – дискретная СВ, то

$$M[X] = \sum_{i=1}^{n(\infty)} x_i \cdot P(X = x_i) = \sum_{i=1}^{n(\infty)} x_i \cdot \left( \sum_{j=1}^{m(\infty)} p_{ij} \right),$$

$$M[Y] = \sum_{j=1}^{m(\infty)} y_j \cdot P(Y = y_j) = \sum_{j=1}^{m(\infty)} y_j \cdot \left( \sum_{i=1}^{n(\infty)} p_{ij} \right).$$

***Замечание***. Принято записывать:

$$M[X]=\sum_{i=1}^{n(\infty)}\sum_{j=1}^{m(\infty)}x_i p_{ij}, \qquad M[Y]=\sum_{i=1}^{n(\infty)}\sum_{j=1}^{m(\infty)}y_j p_{ij}.$$

2) если (X,Y) – непрерывная СВ с плотностью вероятностей $f(x,y)$, то

$$M[X]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}x\cdot f(x,y)dxdy,$$

$$M[y]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}y\cdot f(x,y)dxdy.$$

***Замечание***. M[X] и M[Y] существуют, если ряды (интегралы) в формулах их определяющих, сходятся абсолютно.

В противном случае говорят, что СВ ***не имеет математического ожидания***.

Математические ожидания M[X] и M[Y] – характеристики положения значений СВ на плоскости.

Они являются координатами центра рассеивания значений двумерной СВ.

Математические ожидания M[X] и M[Y] обладают теми же свойствами, что и математическое ожидание одной СВ.

## б) Дисперсии D[X] и D[Y]

Дисперсии D[X] и D[Y] определяются по формулам:

1) если (X,Y) – дискретная СВ, то

$$D[X] = \sum_{i=1}^{n(\infty)} \sum_{j=1}^{m(\infty)} (x_i - M[X])^2 p_{ij},$$

$$D[Y] = \sum_{i=1}^{n(\infty)} \sum_{j=1}^{m(\infty)} (y_j - M[Y])^2 p_{ij}.$$

2) если (X,Y) – непрерывная СВ с плотностью вероятностей $f(x,y)$, то

$$D[X] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - M[X])^2 \cdot f(x,y)\,dxdy,$$

$$D[Y] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (y - M[Y])^2 \cdot f(x,y)\,dxdy.$$

Дисперсии D[X] и D[Y] характеризуют рассеивание значений двумерной СВ вдоль осей $Ox$ и $Oy$ соответственно.

Дисперсии D[X] и D[Y] обладают теми же свойствами, что и дисперсия одной СВ.

В частности, остается справедливой формула

$$\mathrm{D}[X] = \mathrm{M}[X^2] - \mathrm{M}[X]^2 .$$

# в) ковариация

Пусть $X,Y$ – СВ, имеющие математические ожидания $M[X]$ и $M[Y]$ соответственно.

ОПРЕДЕЛЕНИЕ. ***Ковариацией*** (***корреляционным моментом***) *СВ* $X$ *и* $Y$ *называется математическое ожидание произведения их отклонений*:

$$\mathrm{cov}(X,Y) = M[(X - M[X]) \cdot (Y - M[Y])].$$

Используя свойства математического ожидания, получаем:

$$\mathrm{cov}(X,Y) = M[X \cdot Y] - M[X] \cdot M[Y] \qquad (1)$$

В частности, при $X = Y$ из (1) получаем:

$$\mathrm{cov}(X,Y) = M[X^2] - M[X]^2 = D[X]$$

Кроме того, из (1) получаем: $\mathrm{cov}(X,Y) = \mathrm{cov}(Y,X)$

Понятие ковариации позволяет получить формулу для дисперсии суммы 2-х произвольных СВ. А именно:

$$D[X+Y] = D[X] + D[Y] + \mathrm{cov}(X,Y).$$

# ВЕРОЯТНОСТНЫЙ СМЫСЛ КОВАРИАЦИИ

Пусть X и Y – независимые СВ.

Тогда X – M[X] и Y – M[Y] – тоже независимы.

$$\Rightarrow \ \mathrm{cov}(X,Y) = M[(X - M[X]) \cdot (Y - M[Y])] =$$
$$= M[X - M[X]] \cdot M[Y - M[Y]] = 0$$

ВЫВОД: cov(X,Y) = 0 – *необходимое условие независимости СВ* X *и* Y.

Однако это условие не является достаточным.

Существуют зависимые X и Y, для которых cov(X,Y) = 0.

В общем случае справедливы утверждения:

1) $\mathrm{cov}(X,Y) \neq 0 \ \Rightarrow$ X и Y – зависимые СВ.

2) $\mathrm{cov}(X,Y) = 0 \ \Rightarrow$ о характере связи X и Y ничего сказать нельзя

Таким образом, *ковариация – числовая характеристика взаимосвязи СВ*.

## г) коэффициент корреляции

Недостаток ковариации – ее размерность равна произведению размерностей отклонений СВ от своих математических ожиданий (т.е. произведению размерностей СВ).

Для устранения этой величины вводят безразмерную величину – коэффициент корреляции.

ОПРЕДЕЛЕНИЕ. *Величина* $r(X,Y)=\dfrac{\operatorname{cov}(X,Y)}{\sigma_X\cdot\sigma_Y}$ *называется* ***коэффициентом корреляции*** (линейным коэффициентом корреляции, коэффициентом корреляции Пирсона)

$\Rightarrow$ коэффициент корреляции также является числовой характеристикой взаимосвязи СВ.

ОПРЕДЕЛЕНИЕ. СВ называются ***коррелированными***, если

$$r(\mathrm{X},\mathrm{Y})\neq 0.$$

Если $r(\mathrm{X},\mathrm{Y})=0$, то СВ называют ***некоррелированными***.

$\Rightarrow$ независимые СВ – некоррелированные.

Зависимые СВ – могут быть как коррелированными, так и некоррелированными.

## СВОЙСТВА КОЭФФИЦИЕНТА КОРРЕЛЯЦИИ

1) $|\, r(X,Y) \,| \le 1$;

2) если $|\, r(X,Y) \,| = 1$, то СВ X и Y связаны линейной зависимостью, т.е.

$$Y = aX + b, \text{ где } a,b \text{ – числа.}$$