6. Двумерные случайные величины

Функция распределения случайного вектора, ее свойства. Плотность распределения непрерывного случайного вектора, свойства. Плотности распределения компонент случайного вектора. Вероятность попадания в область. Независимые случайные величины. Корреляционный момент, коэффициент корреляции и его свойства. Прямые среднеквадратической регрессии.

6.1. Определение n-мерной случайные величины

Во многих реальных задачах мы имеем не одну, а несколько случайных величин в одном и том же эксперименте. Иногда их удобно рассматривать как единый объект. Это приводит нас к следующему определению.

Определение 6.1. *n -**мерным случайным вектором** или n-**мерной случайной величиной** называется набор $\xi = (\xi_1, \xi_2 \ldots, \xi_n)$ случайных величин, заданных на одном и том же вероятностном пространстве (Ω, A, P).*

Фактически случайный вектор ξ есть отображение $\xi : \Omega \to R^n$.

Приведём примеры многомерных случайных величин.

1. Результаты экзаменационной сессии студенческих групп характеризуется системой n случайных величин $\xi_1, \xi_2, \ldots, \xi_n$ — оценками по различным предметам.

2. Отклонение пули от центра мишени в виде квадрата можно задавать как четырёхмерный случайный вектор: $\mathbf{X} = (\xi; \eta; \zeta; \tau)$, где случайные величины: ξ, η, ζ, τ — отклонение пули вправо, вверх, влево, вниз, соответственно.

Случайные величины $\xi_1, \xi_2, \ldots, \xi_n$, входящие в систему, могут быть дискретными и непрерывными.

6.2. Двумерные случайные векторы

Для простоты и большей наглядности, рассмотрим двумерные случайные векторы. Будем рассматривать точку на плоскости со случайными координатами $(\xi; \eta)$. Сначала рассмотрим случай, когда обе составляющие — дискретные случайные величины.

Определение 6.2. ***Законом распределения** дискретной двумерного случайного вектора называют перечень возможных значений этой величины, т.е. пар чисел $(x_i; y_j)$, $i = 1, \ldots, n$, $j = 1, \ldots, m$, и их вероятностей $p_{ij} = P(\xi = x_i; \eta = y_j)$.*

Закон распределения задают в виде таблицы с двойным входом, в которой указывают все значения x_i, y_i и вероятности p_{ij}.

$\xi \backslash \eta$	y_1	$\ldots$	y_j	$\ldots$	y_m
x_1	p_{11}	$\ldots$	p_{1j}	$\ldots$	p_{1m}
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
x_i	p_{i1}	$\ldots$	p_{ij}	$\ldots$	p_{im}
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
x_n	p_{n1}	$\ldots$	p_{nj}	$\ldots$	p_{nm}

Добавим к этой таблице ещё справа один столбец и снизу одну строку, в которые запишем суммы элементов.

Таблица 6.1

Распределение двумерной дискретной случайной величины						
$\xi \backslash \eta$	y_1	$\ldots$	y_j	$\ldots$	y_m	$P(\xi = x_i)$
x_1	p_{11}	$\ldots$	p_{1j}	$\ldots$	p_{1m}	$p_{1\,*}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
x_i	p_{i1}	$\ldots$	p_{ij}	$\ldots$	p_{im}	$p_{i\,*}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
x_n	p_{n1}	$\ldots$	p_{nj}	$\ldots$	p_{nm}	$p_{n\,*}$
$P(\eta = y_j)$	p_{*1}	$\ldots$	p_{*j}	$\ldots$	p_{*m}	$\boxed{1}$

Так как события $\{\xi = x_i, \eta = y_i\}$, $i = 1, \ldots, n$, $j = 1, \ldots, m$ попарно несовместны и в сумме дают достоверное событие, сумма всех вероятностей равна 1.

Зная двумерный закон распределения, можно найти закон распределения каждой составляющей (но не наоборот). Действительно:

$$\begin{gathered} P(\xi = x_i) = P(\xi = x_i, \eta = y_1) + P(\xi = x_i, \eta = y_2) + \ldots \\ \ldots + P(\xi = x_i, \eta = y_m) = \sum_{j=1}^{m} p_{ij} = p_{i\,*}. \end{gathered} \tag{6.1}$$

Аналогично

$$P(\eta = y_i) = \sum_{i=1}^{n} p_{ij} = p_{*j}. \tag{6.2}$$

Итак, сложив вероятности «по строкам» и записав их в последний столбец, мы получим распределение составляющей ξ (первый и последний столбец таблицы 6.1). Сложив вероятности по столбцам и записав их в последнюю строчку, мы получим распределение составляющей η (первая и последняя строки таблицы 6.1).

Зная распределение составляющих, можем найти числовые характеристики каждой из них:

$$M(\xi) = \sum_{i=1}^{n} x_i p_{i\,*}, \qquad M(\eta) = \sum_{j=1}^{m} y_j p_{*\,j}\,. \tag{6.3}$$

Определение 6.3. *Точка с координатами $\big(M(\xi); M(\eta)\big)$ называется центром распределения.*

Отметим, что таблица 6.1, кроме информации о распределении каждой составляющей, содержит также информацию об их взаимном влиянии.

Найдём, например, условные вероятности $P(\eta = y_j/\xi = x_i)$ и $P(\xi = x_i/\eta = y_j)$. Из формулы (2.3) следует, что

$$P(B/A) = \frac{P(A \cdot B)}{P(A)}. \tag{6.4}$$

Поэтому

$$P(\eta = y_j/\xi = x_i) = \frac{P(\xi = x_i, \eta = y_j)}{P(\xi = x_i)} = \frac{p_{ij}}{p_{i\,*}}. \tag{6.5}$$

Аналогично:

$$P(\xi = x_i/\eta = y_j) = \frac{p_{ij}}{p_{*\,j}}. \tag{6.6}$$

Очевидно, что $\sum\limits_{j=1}^{m} P(\eta = y_j/\xi = x_i) = 1$ для $i = 1, \ \ldots, n$, так же, как и $\sum\limits_{i=1}^{n} P(\xi = x_i/\eta = y_j) = 1$ для $j = 1, \ \ldots, m$ (докажите самостоятельно).

Вероятности $P(\eta = y_j/\xi = x_i)$ для $j = 1, \ \ldots, m$ образуют условное распределение случайной величины η при фиксированном значении ξ. В частности, можно найти условное математическое ожидание η при фиксированном значении ξ:

$$M(\eta/\xi = x_i) = \sum_{j=1}^{m} y_j P(\eta = y_j/\xi = x_i) \quad \text{для} \ \ i = 1, \ \ldots, n \tag{6.7}$$

и условное математическое ожидание ξ при фиксированном значении η:

$$M(\xi/\eta = y_j) = \sum_{i=1}^{n} x_i P(\xi = x_i/\eta = y_j) \quad \text{для} \quad j = 1, \ \ldots, m. \tag{6.8}$$

Замечание 6.1. *Легко показать, что для независимых дискретных случайных величин ξ и η*

$$P(\eta = y_j/\xi = x_i) = P\{\eta = y_j\} \quad и \quad P(\xi = x_i/\eta = y_j) = P(\xi = x_i).$$

Другими словами, закон распределения каждой из них не зависит от значений, принимаемых другой.

Действительно, по определению для независимых дискретных случайных величин ξ и η вероятность $p_{ij} = p_{i\,*} \cdot p_{*j}$, поэтому:

$$P(\eta = y_j/\xi = x_i) = \frac{p_{ij}}{p_{i\,*}} = \frac{p_{i\,*} \cdot p_{*j}}{p_{i\,*}} = p_{*j} \ .$$

Аналогично получаем:

$$P(\xi = x_i/\eta = y_j) = \frac{p_{ij}}{p_{*j}} = \frac{p_{i\,*} \cdot p_{*j}}{p_{*\,j}} = p_{i\,*} .$$

Пример 6.1. *Дискретная двумерная случайная величина задана таблицей 6.2.*

Таблица 6.2

Условие примера 6.1			
$\xi \backslash \eta$	1	3	5
1	0,1	0,2	0,3
2	0,0	0,3	0,1

Найти безусловное и условное математическое ожидание η при условии $\xi = 2$, а также безусловное и условное математическое ожидание ξ при условии $\eta = 1$. Найти математическое ожидание и дисперсию произведения $\xi\eta$.

◀Сначала найдём безусловные распределения ξ и η, суммируя вероятности по строкам и столбцам таблицы 6.2, и допишем их в таблицу распределения (в последний столбец и строку) (см. табл. 6.3).

Таблица 6.3

Решение примера 6.1				
$\xi\backslash\eta$	1	3	5	$P(\xi = x_i)$
1	0,1	0,2	0,3	0,6
2	0,0	0,3	0,1	0,4
$P(\eta = y_i)$	0,1	0,5	0,4	1

Искомые безусловные математические ожидания получатся как обычно для дискретных распределений:

$$M(\xi) = 1 \cdot 0{,}6 + 2 \cdot 0{,}4 = 1{,}4,$$
$$M(\eta) = 1 \cdot 0{,}1 + 3 \cdot 0{,}5 + 5 \cdot 0{,}4 = 3{,}6.$$

Далее, по формулам (6.5) и (6.6) найдём условные распределения $P(\eta = y_j/\xi = 2)$ и $P(\xi = x_i/\eta = 1)$:

$$P(\eta = y_j/\xi = 2) = \frac{(\eta = y_j,\ \xi = 2)}{P(\xi = 2)} = \frac{P(\eta = y_j,\ \xi = 2)}{0{,}4},$$
$$P(\xi = x_i/\eta = 1) = \frac{P(\xi = x_i,\ \eta = 1)}{P(\eta = 1)} = \frac{P(\xi = x_i,\ \eta = 1)}{0{,}1}.$$

Результаты представлены в таблицах 6.4, 6.5.

Условные распределения

Таблица 6.4

η	1	3	5
$P(\eta = y_j/\xi = 2)$	0	3/4	1/4

Таблица 6.5

ξ	1	2
$P(\xi = x_i/\eta = 1)$	1	0

Найдём теперь условные математические ожидания по формулам (6.7), (6.8) для данных из таблиц 6.4, 6.5.

$$M(\xi/\eta = 1) = 1 \cdot 1 + 2 \cdot 0 = 1,$$
$$M(\eta/\xi = 2) = 1 \cdot 0 + 3 \cdot \frac{3}{4} + 5 \cdot \frac{1}{4} = 3{,}5.$$

Как видим, условные и соответствующие безусловные математические ожидания различаются.

Найдём теперь математическое ожидание произведения $\xi\eta$. Для этого напишем статистический ряд этой случайной величины.

$M(\xi\eta) = 1 \cdot 0{,}1 + 2 \cdot 0 + 3 \cdot 0{,}2 + 5 \cdot 0,3 + 6 \cdot 0{,}3 + 10 \cdot 0{,}1 = 5{,}1$

$D(\xi\eta) = M((\xi\eta)^2) - M^2(\xi\eta) = 1 \cdot 0{,}1 + 4 \cdot 0 + 9 \cdot 0{,}2 + 25 \cdot 0,3 + 36 \cdot 0{,}3 + 100 \cdot 0{,}1 - 5{,}1^2 = 30{,}2 - 26{,}01 = 4{,}19.$

Таблица 6.6

Распределение произведения случайных величин						
$\xi\eta$	1	2	3	5	6	10
P	0,1	0,0	0,2	0,3	0,3	0,1

Ответ: $M(\xi) = 1{,}4$; $M(\eta) = 3{,}6$; $M(\xi/\eta = 1) = 1$; $M(\eta/\xi = 2) = 3{,}5$; $M(\xi\eta) = 5{,}1$; $D(\xi\eta) = 4{,}19$.

6.3. Двумерная функция распределения и плотность

Приводимое ниже определение 6.4 функции распределения справедливо для любой двумерной случайной величины. Заметим, однако, что дискретная случайная величина полностью определяется таблицей 6.1, работать с которой удобнее, чем с функцией распределения двумерной дискретной случайной величины.

Определение 6.4. ***Функцией распределения*** *двумерной случайной величины* $(\xi;\eta)$ *называют*

$$F(x;y) = P(\xi < x; \eta < y). \tag{6.9}$$

Двумерная функция распределения обладает следующими свойствами:

(1) $0 \leqslant F(x;y) \leqslant 1$;

(2) $F(-\infty;y) = F(x;-\infty) = F(-\infty;-\infty) = 0 \quad F(+\infty;+\infty) = 1$;

(3) $F(x;y)$ есть неубывающая функция по каждому аргументу;

(4) Функции распределения каждой составляющей получаются предельным переходом:

$$F_\xi(x) = P(\xi < x) = F(x;+\infty),$$
$$F_\eta(y) = P(\eta < y) = F(+\infty;y);$$

(5) Вероятность попадания в прямоугольник выражается через функцию распределения по формуле:

$$\begin{gathered} P(x_1 \leqslant \xi < x_2; y_1 \leqslant \eta < y_2) = \\ = \big(F(x_2;y_2) - F(x_2;y_1)\big) - \big(F(x_1;y_2) - F(x_1;y_1)\big). \end{gathered} \tag{6.10}$$

Доказательства свойств 1, 2 непосредственно следуют из определения 6.4 (проведите их самостоятельно). Доказательство свойства 3 аналогично доказательству свойства 3 функции распределения $F(x)$ в п. 3.2 лекции 3.

Свойство 4 очевидно:

$F(x; +\infty) = P(\xi < x; \eta < +\infty) = P(\xi < x) = F_\xi(x)$.

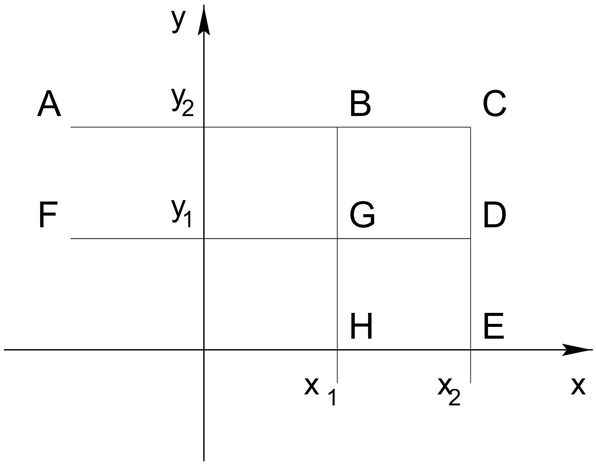

Рисунок 25. *Вероятность попадания в прямоугольник*

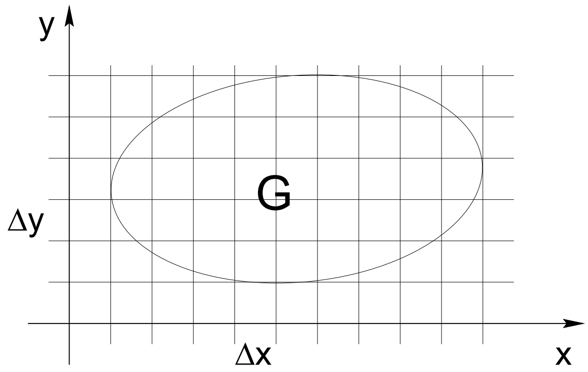

Рисунок 26. *Вероятность попадания в область*

Для доказательства свойства 5 заметим, что согласно определению 6.4 $F(x_2; y_2)$ есть вероятность попадания двумерной случайной величины в угол ACE, $F(x_2; y_1)$ — в угол FDE; следовательно $\big(F(x_2; y_2) - F(x_2; y_1)\big)$ есть вероятность попадания в полуполосу $ACDF$ (рис. 25).

Аналогично $\big(F(x_1; y_2) - F(x_1; y_1)\big)$ есть вероятность попадания в полуполосу $ABGF$. Следовательно, разность этих вероятностей есть вероятность попадания в прямоугольник $BCDG$.

6.4. Непрерывные двумерные случайные величины

ОПРЕДЕЛЕНИЕ 6.5. *Двумерная случайная величина $(\xi; \eta)$ называется непрерывной, если её функция распределения $F(x; y)$ непрерывна и имеет непрерывные частные производные второго порядка всюду (за исключением быть может, конечного числа кривых).*

ОПРЕДЕЛЕНИЕ 6.6. ***Плотностью распределения*** *двумерной непрерывной случайной величины $(\xi; \eta)$ называется вторая смешанная частная производная функции распределения:*

$$f(x; y) = \frac{\partial^2 F(x; y)}{\partial x \partial y}. \tag{6.11}$$

Двумерная плотность распределения обладает следующими свойствами:

(1) $f(x;y) \geqslant 0$;

(2) $f(-\infty;y) = f(x;-\infty) = f(\pm\infty;\pm\infty) = 0$;

(3) $F(x;y) = \int\limits_{-\infty}^{x} \int\limits_{-\infty}^{y} f(s;t)dsdt$;

(4) Вероятность попадания двумерной случайной величины $(\xi;\eta)$ в область G равна:

$$P((\xi;\eta) \in G) = \iint\limits_{G} f(x;y)dxdy;$$

(5) $\iint\limits_{-\infty}^{+\infty} f(x;y)dxdy = 1$.

Свойство 1 есть следствие свойства 3 $F(x;y)$: производная от неубывающей функции неотрицательна. Свойство 2 вытекает из свойства 2 $F(x;y)$, т.к. производная константы равна нулю.

Свойство 3 следует из определения 6.6, поскольку $F(x;y)$ является первообразной для $f(x;y)$.

Для доказательства свойства 4 область G следует разбить на множество прямоугольников со сторонами Δx и Δy (рис. 26). Вероятность попадания в i-й из них определяется с помощью свойства 5 функции распределения $F(x;y)$. Применим к правой части этого равенства формулу Лагранжа:

$$\begin{aligned} P(x_{1i} \leqslant \xi < x_{2i}; y_{1i} \leqslant \eta < y_{2i}) = \big(F(x_{2i};y_{2i}) - F(x_{2i}y_{1i})\big) - \\ -\big(F(x_{1i};y_{2i}) - F(x_{1i}y_{1i})\big) = F''_{xy}(s_i;t_i)\Delta x\Delta y = f(s_i;t_i)\Delta x\Delta y, \end{aligned} \tag{6.12}$$

где точка $(s_i;t_i)$ находится внутри i-го прямоугольника.

Очевидно, что вероятность попадания в область G приближённо равна сумме вероятностей попадания в эти прямоугольники:

$$P((\xi;\eta) \in G) \approx \sum_{i=1}^{n} f(s_i;t_i)\Delta x\Delta y.$$

Переходя к пределу при $\Delta x \to 0,\ \Delta y \to 0$ $(n \to \infty)$, получим свойство 4 плотности $f(x;y)$.

Теперь свойство 5 очевидно, т.к. вероятность попасть во всю плоскость с одной стороны равна $\iint\limits_{-\infty}^{+\infty} f(x;y)dxdy$, а с другой стороны — есть достоверное событие.

Плотности распределения составляющих двумерной непрерывной случайной величины получаются из её плотности $f(x;y)$ по формулам (6.13):

$$f_\xi(x) = \int\limits_{-\infty}^{+\infty} f(x;y)dy; \qquad f_\eta(y) = \int\limits_{-\infty}^{+\infty} f(x;y)dx. \tag{6.13}$$

Действительно, поскольку $F(x;y) = \int\limits_{-\infty}^{x} \int\limits_{-\infty}^{y} f(s;t)dsdt$, получаем $F_\xi(x) = F(x;+\infty) = \int\limits_{-\infty}^{x} \int\limits_{-\infty}^{+\infty} f(s;t)dsdt$.

Продифференцировав обе части этого равенства, получим:

$$f_\xi(x) = \frac{dF_\xi(x)}{dx} = \frac{d}{dx}\left(\int\limits_{-\infty}^{x} \int\limits_{-\infty}^{+\infty} f(s;t)dsdt\right) = \int\limits_{-\infty}^{+\infty} f(x;t)dt.$$

Из равенства (6.12) следует, что вероятностный смысл двумерной плотности состоит в том, что $f(x;y)$ равна вероятности попадания случайной точки в прямоугольник с вершиной $(x;y)$, с малыми сторонами $\Delta x, \Delta y$, отнесённой к площади этого прямоугольника.

Аналогично тому, как это было сделано для дискретной случайной величины, найдём условную плотность составляющей η при фиксированной величине ξ и наоборот.

Определение 6.7. ***Условной плотностью*** $f(y/\xi = x)$ *распределения* η *при условии, что* $\xi = x$*, называется:*

$$f(y/\xi = x) = \left[\begin{array}{ll} 0, & f_\xi(x) = 0, \\ \dfrac{f(x;y)}{f_\xi(x)}, & f_\xi(x) \neq 0. \end{array}\right. \tag{6.14}$$

Условной плотностью $f(x/\eta = y)$ *распределения* ξ *при условии, что* $\eta = y$*, называется:*

$$f(x/\eta = y) = \left[\begin{array}{ll} 0, & f_\eta(y) = 0, \\ \dfrac{f(x;y)}{f_\eta(y)}, & f_\eta(y) \neq 0. \end{array}\right. \tag{6.15}$$

Заметим, что формулы (6.14), (6.15) соответствуют формуле (6.4), если учесть вероятностный смысл плотности. Так, например:

$$\frac{f(x;y)\Delta x\Delta y}{f_\xi(x)\Delta x}=\frac{f(x;y)\Delta y}{f_\xi(x)}=f(y/\xi=x)\Delta y.$$

Определение 6.8. ***Условным математическим ожиданием*** *η при условии, что $\xi=x$, называется:*

$$M(\eta/\xi=x)=\int\limits_{-\infty}^{+\infty} yf(y/\xi=x)dy. \tag{6.16}$$

Условным математическим ожиданием ξ при условии, что $\eta=y$, называется:

$$M(\xi/\eta=y)=\int\limits_{-\infty}^{+\infty} x\cdot f(x/\eta=y)dx. \tag{6.17}$$

Заметим, что $M(\eta/\xi=x)$ есть функция от x: $M(\eta/\xi=x)=f_{\eta/\xi}(x)$. Аналогично $M(\xi/\eta=y)$ является функцией от y:
$M(\xi/\eta=y)=\psi_{\xi/\eta}(y)$.

Определение 6.9. *Функцию $f_{\eta/\xi}(x)$ называют* ***регрессией*** *η на ξ. Другими словами, регрессией η на ξ называется условное математическое ожидание η при фиксированном $\xi=x$. Аналогично $\psi_{\xi/\eta}(y)$ называется регрессией ξ на η.*

6.5. Коэффициент корреляции

Напомним, что в соответствии с определением 3.22, две случайные величины называются независимыми, если

$$F(x;y)=F_\xi(x)\cdot F_\eta(y).$$

Теорема 6.1. *Для независимости непрерывных случайных величин ξ и η необходимо и достаточно, чтобы $f(x;y)=f_\xi(x)\cdot f_\eta(y)$.*

Доказательство. Если ξ и η независимы, то по определению 3.22

$$F(x;y) = F_\xi(x)\cdot F_\eta(y) \implies f(x;y) = \frac{\partial^2 F(x;y)}{\partial x \partial y} =$$
$$= \frac{\partial}{\partial y}\Big(F'_\xi(x)\cdot F'_\eta(y)\Big) = F'_\xi(x)\cdot F'_\eta(y) = f_\xi(x)\cdot f_\eta(y).$$

Если $f(x;y) = f_\xi(x)\cdot f_\eta(y)$, то

$$F(x;y) = \int\limits_{-\infty}^{x}\int\limits_{-\infty}^{y} f(s;t)dsdt = \int\limits_{-\infty}^{x}\int\limits_{-\infty}^{y} f_\xi(s)\cdot f_\eta(t)dsdt =$$
$$= \int\limits_{-\infty}^{x} f_\xi(s)ds \int\limits_{-\infty}^{y} f_\eta(t)dt = F_\xi(x)\cdot F_\eta(y).$$

Следовательно, ξ и η независимы по определению 3.22.

Замечание 6.2. *Можно показать, что для независимых непрерывных случайных величин ξ и η*

$$f(y/\xi = x) = f_\eta(y) \text{ и } f(x/\eta = y) = f_\xi(x) \text{ при } f_\xi(x)\neq 0,\ f_\eta(y)\neq 0.$$

Т.е. закон распределения каждой из них не зависит от значений, принимаемых другой. Действительно, по теореме 6.1 для независимых непрерывных ξ и η выполняется $f(x;y) = f_\xi(x)\cdot f_\eta(y)$, поэтому при $f_\xi(x)\neq 0$ получаем:

$$f(y/\xi = x) = \frac{f(x;y)}{f_\xi(x)} = \frac{f_\xi(x)\cdot f_\eta(y)}{f_\xi(x)} = f_\eta(y).$$

Аналогично доказывается второе равенство.

Пример 6.2. *Плотность $f(x;y)$ определяется формулой:*

$$f(x;y) = \begin{cases} C & \text{при } x^2+y^2 \leqslant R^2, \\ 0 & \text{при } x^2+y^2 > R^2. \end{cases}$$

Определить константу C и функции регрессии η на ξ и ξ на η.

◀Для определения константы C воспользуемся свойством 5 плотности:

$$\iint\limits_{-\infty}^{+\infty} f(x;y)dxdy = 1 \implies \iint\limits_{x^2+y^2<R} C\,dxdy = 1 \implies C\iint\limits_{x^2+y^2<R} dxdy = 1.$$

Воспользуемся тем, что $\iint\limits_{x^2+y^2<R} dxdy$ равен объёму цилиндра с основанием, площадь которого πR^2, и высотой равной 1.

$$C\cdot\pi R^2 = 1 \implies C = \frac{1}{\pi R^2}.$$

Определим теперь плотности составляющих по формулам (6.13): при $|x|>R \quad f_\xi(x)=0;\quad$ при $|x|<R$

$$f_\xi(x) = \int\limits_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} \frac{1}{\pi R^2}dy = \frac{2\sqrt{R^2-x^2}}{\pi R^2}.$$

Окончательно:

$$f_\xi(x) = \begin{cases} \dfrac{2\sqrt{R^2-x^2}}{\pi R^2} & \text{при} \quad |x|<R, \\ 0 & \text{при} \quad |x|>R. \end{cases} \tag{6.18}$$

Аналогично:

$$f_\eta(y) = \begin{cases} \dfrac{2\sqrt{R^2-y^2}}{\pi R^2} & \text{при} \quad |y|<R, \\ 0 & \text{при} \quad |y|>R. \end{cases} \tag{6.19}$$

Теперь по формулам (6.14), (6.15) определяем:

$$f(y/\xi=x) = \begin{cases} \dfrac{1}{2\sqrt{R^2-x^2}} & \text{при} \quad x^2+y^2<R^2, \\ 0 & \text{при} \quad x^2+y^2>R^2; \end{cases},$$

$$f(x/\eta=y) = \begin{cases} \dfrac{1}{2\sqrt{R^2-y^2}} & \text{при} \quad x^2+y^2<R^2, \\ 0 & \text{при} \quad x^2+y^2>R^2. \end{cases}$$

Наконец, по формулам (6.16), (6.17) найдём уравнения регрессии:

$$M(\eta/\xi = x) = \int\limits_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} y\, \frac{1}{2\sqrt{R^2-x^2}}dy = 0,$$

$$M(\xi/\eta = y) = \int\limits_{-\sqrt{R^2-y^2}}^{\sqrt{R^2-y^2}} x\, \frac{1}{2\sqrt{R^2-y^2}}dx = 0.$$

Пример 6.3. *Установить, будут ли зависимы составляющие ξ и η примера 6.2.*

◀Как было установлено в примере 6.2, плотности равны:

$$f(x;y) = \begin{cases} \dfrac{1}{\pi R^2} & \text{при} \quad x^2+y^2 < R^2, \\ 0 & \text{при} \quad x^2+y^2 > R^2; \end{cases}$$

$$f_\xi(x) = \begin{cases} \dfrac{2\sqrt{R^2-x^2}}{\pi R^2} & \text{при} \quad |x| < R, \\ 0 & \text{при} \quad |x| > R; \end{cases}$$

$$f_\eta(y) = \begin{cases} \frac{2\sqrt{R^2-y^2}}{\pi R^2} & \text{при} \quad |y| < R, \\ 0 & \text{при} \quad |y| > R. \end{cases}$$

Поскольку $f(x;y) \neq f_\xi(x) \cdot f_\eta(y)$, случайные величины ξ и η зависимы. Этот факт следует также из того, что $f(x/\eta = y) \neq f_\xi(x)$ и $f(y/\xi = x) \neq f_\eta(y)$.

Ответ: ξ и η зависимы.

Для описания зависимости между двумя случайными величинами ξ и η введённые ранее числовые характеристики $M(\xi), D(\xi), M(\eta), D(\eta)$ неприменимы. Введём понятие корреляционного момента и коэффициента корреляции.

Определение 6.10. ***Корреляционным моментом*** *$K_{\xi\eta}$ случайных величин ξ и η называют:*

$$K_{\xi\eta} = M\left(\big(\xi - M(\xi)\big)\big(\eta - M(\eta)\big)\right).$$

Легко убедиться, что корреляционный момент можно также вычислять по формуле:

$$K_{\xi\eta} = M(\xi \cdot \eta) - M(\xi) \cdot M(\eta). \tag{6.20}$$

Действительно, пользуясь свойствами математического ожидания, получаем:

$$K_{\xi\eta} = M\left[\big(\xi - M(\xi)\big)\big(\eta - M(\eta)\big)\right] = M\big(\xi\eta - \xi M(\eta) - \eta M(\xi) +$$
$$+ M(\xi)\cdot M(\eta)\big) = M(\xi\eta) - M(\xi)M(\eta) - M(\eta)M(\xi) + M(\xi)M(\eta) =$$
$$= M(\xi\eta) - M(\xi)M(\eta).$$

Вычисление корреляционного момента по формуле (6.20) для дискретных случайных величин сводится к вычислению суммы:

$$K_{\xi\eta} = \sum_{j=1}^{m}\sum_{i=1}^{n} p_{ij}x_i y_j - M(\xi)\cdot M(\eta),$$

а для непрерывных — интеграла:

$$K_{\xi\eta} = \iint\limits_{-\infty}^{+\infty} xyf(xy)dxdy - M(\xi)\cdot M(\eta).$$

Теорема 6.2. *Для независимых случайных величин корреляционный момент равен нулю.*

Действительно, пользуясь свойствами математического ожидания из формулы (6.20), получаем для независимых ξ и η:

$$K_{\xi\eta} = M(\xi\cdot\eta) - M(\xi)\cdot M(\eta) = M(\xi)\cdot M(\eta) - M(\xi)\cdot M(\eta) = 0.$$

Теорема 6.3. *Модуль корреляционного момента не превышает произведения среднеквадратических отклонений:* $|K_{\xi\eta}| \leqslant \sigma_\xi \sigma_\eta$.

◀ Рассмотрим $D(\sigma_\eta\cdot\xi - \sigma_\xi\cdot\eta) \geqslant 0$.
Учитывая (6.20), а также:
$\sigma_\xi^2 = M(\xi^2) - M^2(\xi)$, $\sigma_\eta^2 = M(\eta^2) - M^2(\eta)$,

получаем:

$$\begin{gathered}
D(\sigma_\eta \cdot \xi - \sigma_\xi \cdot \eta) = M(\sigma_\eta \cdot \xi - \sigma_\xi \cdot \eta)^2 - \big(M(\sigma_\eta \cdot \xi - \sigma_\xi \cdot \eta)\big)^2 = \\
= M(\sigma_\eta^2 \cdot \xi^2 - 2\sigma_\xi\sigma_\eta \cdot \xi \cdot \eta + \sigma_\xi^2 \cdot \eta^2) - \big(\sigma_\eta M(\xi) - \sigma_\xi M(\eta)\big)^2 = \\
= \sigma_\eta^2 M(\xi^2) - 2\sigma_\xi\sigma_\eta M(\xi\eta) + \sigma_\xi^2 M(\eta^2) - \sigma_\eta^2 M^2(\xi) + 2\sigma_\xi\sigma_\eta M(\xi)M(\eta) - \\
-\sigma_\xi^2 M^2(\eta) = \sigma_\eta^2\big(M(\xi^2) - M^2(\xi)\big) + \sigma_\xi^2\big(M(\eta^2) - M^2(\eta)\big) - \\
-2\sigma_\xi\sigma_\eta\big(M(\xi\eta) - M(\xi)M(\eta)\big) = \\
= \sigma_\eta^2\sigma_\xi^2 + \sigma_\xi^2\sigma_\eta^2 - 2\sigma_\xi\sigma_\eta K_{\xi\eta} = 2\sigma_\xi^2\sigma_\eta^2 - 2\sigma_\xi\sigma_\eta K_{\xi\eta}.
\end{gathered}$$

Из неравенства $2\sigma_\xi^2\sigma_\eta^2 - 2\sigma_\xi\sigma_\eta K_{\xi\eta} \geqslant 0$ получаем: $K_{\xi\eta} \leqslant \sigma_\xi\sigma_\eta$. Аналогично, рассмотрев $D(\sigma_\eta\xi + \sigma_\xi\eta) \geqslant 0$, получим: $K_{\xi\eta} \geqslant -\sigma_\xi\sigma_\eta$. Объединяя два неравенства, получим: $-\sigma_\xi\sigma_\eta \leqslant K_{\xi\eta} \leqslant \sigma_\xi\sigma_\eta$. ▶

На практике пользуются безразмерной характеристикой — коэффициентом корреляции.

Определение 6.11. ***Коэффициентом корреляции*** $r_{\xi\eta}$ *случайных величин* ξ *и* η *называется*

$$r_{\xi\eta} = \frac{K_{\xi\eta}}{\sigma_\xi\sigma_\eta} = \frac{M(\xi\eta) - M(\xi)M(\eta)}{\sigma_\xi\sigma_\eta}. \tag{6.21}$$

Пример 6.4. *Дано распределение двумерного случайного вектора* $(\xi;\eta)$ *с дискретными компонентами.*

$\xi\backslash\eta$	*0*	*2*	*4*
-1	*0,05*	*0,15*	*0,15*
0	*0,2*	*0,1*	*0,1*
1	*0,15*	*0,1*	*0*

Найти коэффициент корреляции $r_{\xi\eta}$.

◀Выпишем отдельно ряды распределения случайных величин ξ, η и $\xi\eta$.

ξ	-1	0	1
P	0,35	0,4	0,25

η	0	2	4
P	0,4	0,35	0,25

$\xi\eta$	-4	-2	0	2	4
P	0,15	0,15	0,6	0,1	0

Найдём математические ожидания $M(\xi)$, $M(\eta)$ и $M(\xi\eta)$.

$M(\xi) = -1 \cdot 0{,}35 + 1 \cdot 0{,}25 = -0{,}1$.

$M(\eta) = 2 \cdot 0{,}35 + 4 \cdot 0{,}25 = 1{,}7$.

$M(\xi\eta) = -4 \cdot 0{,}15 - 2 \cdot 0{,}15 + 2 \cdot 0{,}1 = -0{,}7$.

Применяя формулу (6.20) найдём корреляционный момент

$K_{\xi\eta} = M(\xi \cdot \eta) - M(\xi) \cdot M(\eta) = -0{,}7 + 0{,}1 \cdot 1.7 = -0{,}53$.

$D(\xi) = M(\xi^2) - M^2(\xi) = 1 \cdot 0{,}35 + 1 \cdot 0{,}25 - 0{,}01 = 0{,}59.$
$D(\eta) = 4 \cdot 0{,}35 + 16 \cdot 0{,}25 - 1{,}7^2 = 5{,}4 - 2{,}89 = 2{,}51.$
Используя формулу 6.21 найдём коэффициент корреляции
$r_{\xi\eta} = \dfrac{-0{,}53}{\sqrt{0{,}59 \cdot 2{,}51}} \approx 0,436.$ ►

Пример 6.5. *Определить коэффициент корреляции случайных величин из примера 6.2.*

◄Поскольку плотности составляющих ξ и η, определяемые по формулам (6.18), (6.19), являются чётными функциями, математические ожидания составляющих равны нулю:

$$M(\xi) = \int\limits_{-R}^{R} x \cdot \frac{2\sqrt{R^2 - x^2}}{\pi R^2} dx = 0$$

как интеграл от нечётной функции по симметричному относительно нулю интервалу. Аналогично:

$$M(\eta) = \int\limits_{-R}^{R} y \frac{2\sqrt{R^2 - y^2}}{\pi R^2} dy = 0.$$

Найдём $M(\xi \cdot \eta)$:

$$M(\xi \cdot \eta) = \iint\limits_{-\infty}^{+\infty} x \cdot y f(xy)\ dxdy = \frac{1}{\pi R^2} \int\limits_{-R}^{+R} x dx \int\limits_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} y\ dy = 0.$$

Итак:

$$K_{\xi\eta} = M(\xi \cdot \eta) - M(\xi) \cdot M(\eta) = 0 \implies r_{\xi\eta} = \frac{K_{\xi\eta}}{\sigma_\xi \sigma_\eta} = 0.$$

Ответ: $r_{\xi\eta} = 0$.
Перечислим свойства коэффициента корреляции.

(1) Для независимых ξ и η коэффициент корреляции равен нулю: $r_{\xi\eta} = 0$,

(2) $|r_{\xi\eta}| \leqslant 1$,

(3) $|r_{\xi\eta}| = 1 \iff \eta = k\xi + b$ или $\xi = k\eta + b$.

Свойство 1 является следствием определения 6.11 и теоремы 6.2.
Свойство 2 немедленно следует из теоремы 6.3:

$$r_{\xi\eta} = \frac{K_{\xi\eta}}{\sigma_\xi \sigma_\eta} \implies -1 \leqslant r_{\xi\eta} \leqslant 1.$$

Свойство 3 будет доказано в следующем пункте.

ЗАМЕЧАНИЕ 6.3. *Из равенства нулю коэффициента корреляции не следует независимость случайных величин.*

Действительно, в примере 6.5 определено, что коэффициент корреляции случайных величин из примера 6.2 равен нулю, а в примере 6.3 установлено, что эти случайные величины зависимы.

ОПРЕДЕЛЕНИЕ 6.12. *Случайные величины ξ и η называются некоррелированными, если их коэффициент корреляции равен нулю: $r_{\xi\eta} = 0$.*

Из свойства 1 и замечания 6.3 следует связь между независимостью и некоррелированностью:

$$\begin{aligned}
\text{независимость} &\implies \text{некоррелированность;}\\
\text{некоррелированность} &\;\not\!\!\implies \text{независимость;}\\
\text{коррелированность} &\implies \text{зависимость;}\\
\text{зависимость} &\;\not\!\!\implies \text{коррелированность.}
\end{aligned}$$

6.6. Прямые среднеквадратической регрессии

Рассмотрим двумерную случайную величину $(\xi; \eta)$. Поставим задачу: «наилучшим образом» приблизить случайную величину η функцией $g(\xi)$. «Наилучшим образом» будет пониматься в смысле минимизации среднеквадратического отклонения, т.е. $M\left[\left(\eta - g(\xi)\right)^2\right]$ должно принимать наименьшее возможное для данного класса функций $g(\xi)$ значение.

ОПРЕДЕЛЕНИЕ 6.13. *Функция $y = g(x)$ такая, что $M\left[\left(\eta - g(\xi)\right)^2\right]$ принимает наименьшее возможное для данного класса функций $g(\xi)$ значение, называется среднеквадратической регрессией η на ξ. Если наименьшее значение ищется в классе линейных функций $g(x) = kx + b$, то регрессия называется линейной среднеквадратической регрессией η на ξ; её графиком является, очевидно, прямая.*

Теорема 6.4. *Линейная среднеквадратическая регрессия η на ξ имеет вид:*

$$y = M(\eta) + r_{\xi\eta}\frac{\sigma_\eta}{\sigma_\xi}\big(x - M(\xi)\big). \tag{6.22}$$

Аналогично

$$x = M(\xi) + r_{\xi\eta}\frac{\sigma_\xi}{\sigma_\eta}\big(y - M(\eta)\big). \tag{6.23}$$

Обе прямые регрессии (6.22) и (6.23) проходят через точку $\big(M(\xi); M(\eta)\big)$ — центр распределения.

Обе прямые совпадают, если $r_{\xi\eta} = \pm 1$.

Коэффициент $r_{\xi\eta}\dfrac{\sigma_\eta}{\sigma_\xi}$ называется коэффициентом регрессии η на ξ, а $r_{\xi\eta}\dfrac{\sigma_\eta}{\sigma_\xi}$ — коэффициент регрессии ξ на η. Знак коэффициента регрессии совпадает со знаком коэффициента корреляции $r_{\xi\eta}$. Так, например, при $r_{\xi\eta} > 0$ линейная среднеквадратическая регрессия η на ξ возрастает, при $r_{\xi\eta} < 0$ — убывает.

Пример 6.6. *Случайный вектор (ξ, η) распределен равномерно в области G, рис.27.*

1) Найти плотность распределения вероятностей компонент случайного вектора и проверить, являются ли они зависимыми.

2) Выяснить, коррелированы ли компоненты случайного вектора (ξ, η).

3) Найти $P((\xi, \eta) \in D)$, где $D = \{(x, y) | x^2 + y^2 \leqslant 1\}$.

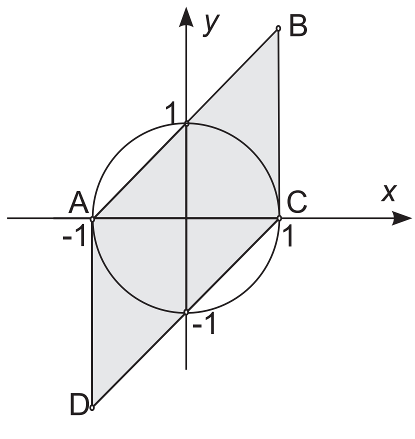

Рисунок 27. *Пример 6.6*

◄1) На рис. 27 представлена область равномерного распределения случайного вектора G, представляющая параллелограмм и область D. Из свойств плотности распределения следует, что функция плотности постоянна и равна $1/S$ (S площадь параллелограмма) на области G и равна нулю вне её.

$S = AC \cdot AD = 4$. Следовательно, функция плотности двумерного распределения равна

$$f(x,y) = \begin{cases} 0, & (x,y) \notin G, \\ \frac{1}{4}, & (x,y) \in G. \end{cases}$$

Плотности распределения составляющих двумерной непрерывной случайной величины получаются из её плотности $f(x;y)$ по формулам (6.13). При $x \notin [-1;1]$ $f_\xi(x) = 0$, т.к. $f(x,y) = 0$. При закрашивании области G вертикальными линиями необходимо двигаться от нижней линии $y = x - 1$ до верхней линии $y = x + 1$, поэтому

$$f_\xi(x) = \int\limits_{-\infty}^{+\infty} f(x;y)dy = \int\limits_{x-1}^{x+1} \frac{1}{4}dy = \frac{1}{4}y\Big|_{x-1}^{x+1} = \frac{1}{4}(x+1-x+1) = \frac{1}{2}.$$

Следовательно, случайная величина ξ распределена равномерно на отрезке $[-1;1]$ и её функция плотности равна

$$f_\xi(x) = \begin{cases} \frac{1}{2}, & x \in [-1;1], \\ 0, & x \notin [-1;1]. \end{cases}$$

Аналогично получим плотность распределения компоненты η.
При $y \notin [-2;2]$ $f_\eta(y) = 0$, т.к. $f(x,y) = 0$. При закрашивании области G горизонтальными линиями необходимо разбить область на две подобласти: DAC, которая слева ограничивается прямой $x = -1$, а справа прямой $x = y + 1$ и ABC, ограниченную прямыми $x = y - 1$ (слева) и $x = 1$ (справа).

При $y \in [-2;0]$

$$f_\eta(y) = \int\limits_{-\infty}^{+\infty} f(x;y)dx = \int\limits_{-1}^{y+1} \frac{1}{4}dx = \frac{1}{4}x\Big|_{-1}^{y+1} = \frac{1}{4}(y+1+1) = \frac{1}{4}y + \frac{1}{2}.$$

При $y \in [0;2]$

$$f_\eta(y) = \int\limits_{-\infty}^{+\infty} f(x;y)dx = \int\limits_{y-1}^{1} \frac{1}{4}dx = \frac{1}{4}x\Big|_{y-1}^{1} = \frac{1}{4}(1-y+1) = -\frac{1}{4}y + \frac{1}{2}.$$

Следовательно, плотность распределения компоненты η имеет равна:

$$f_\eta(y) = \begin{cases} 0, & y \notin [-2;2], \\ \dfrac{1}{4}y + \dfrac{1}{2}, & y \in [-2;0), \\ \dfrac{1}{2} - \dfrac{1}{4}y & y \in [0;2]). \end{cases}$$

На рис. 28, представлены графики плотности распределения вероятностей компонент случайного вектора $(\xi;\, \eta)$. Отметим, что все свойства функции плотности выполняются.

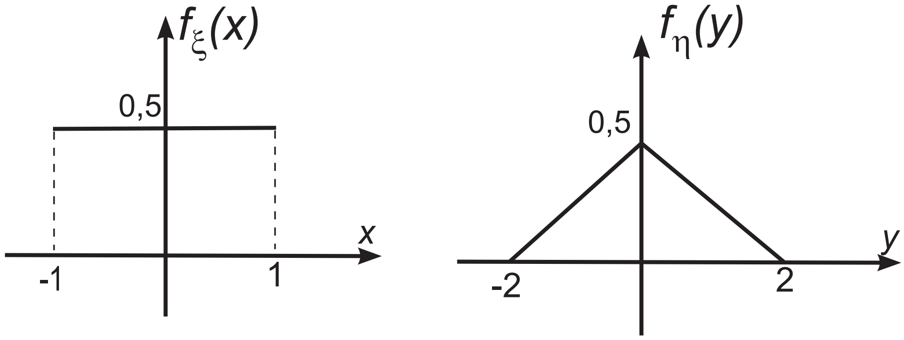

Рисунок 28. *Компоненты плотности вектора* $(\xi;\, \eta)$

Согласно теореме 6.1, которая утверждает, что для независимости непрерывных случайных величин ξ и η необходимо и достаточно, чтобы $f(x;y) = f_\xi(x) \cdot f_\eta(y)$, делаем вывод, что случайные величины ξ и η зависимы.

Докажем это ещё вторым методом. Для этого найдём условные плотности компонент по формулам (6.14) и (6.15).

$$f(y/\xi = x) = \begin{cases} 0, & f_\xi(x) = 0, \\ \dfrac{f(x;y)}{f_\xi(x)}, & f_\xi(x) \neq 0. \end{cases}$$

$$f(y/\xi = x) = \begin{cases} \dfrac{1}{2}, & x \in [-1;1], \\ 0, & x \notin [-1;1]. \end{cases}$$

$$f(x/\eta = y) = \begin{cases} 0, & y \leqslant -2, \\ \dfrac{\frac{1}{4}}{\frac{1}{4}y + \frac{1}{2}}, & -2 < y \leqslant 0, \\ \dfrac{\frac{1}{4}}{\frac{1}{2} - \frac{1}{4}y}, & 0 \leqslant y < 2, \\ 0, & y \geqslant 2, \end{cases} = \begin{cases} 0, & y \leqslant -2, \\ \dfrac{1}{y+2}, & -2 < y \leqslant 0, \\ \dfrac{1}{2-y}, & 0 \leqslant y < 2, \\ 0, & y \geqslant 2. \end{cases}$$

В рассмотренном примере условные плотности распределения $f(x/\eta = y)$ и $f(y/\xi = x)$ не совпадают с безусловными плотностями $f_\eta(y)$ и $f_\xi(x)$. Это имеет место тогда и только тогда, когда случайные величины ξ и η *зависимы*.

2) Выяснить, коррелированы ли компоненты случайного вектора (ξ, η). По вычисленным плотностям распределения компонент случайного вектора найдём их математические ожидания.

$$M(\xi) = \int\limits_{-\infty}^{+\infty} x f_\xi(x)dx = \int\limits_{-1}^{1} \frac{1}{2}xdx = \frac{1}{2} \cdot \frac{x^2}{2}\Big|_{-1}^{1} = 0.$$

$$M(\eta) = \int\limits_{-\infty}^{+\infty} y f_\eta(y)dy = \int\limits_{-2}^{0} y\left(\frac{1}{4}y + \frac{1}{2}\right)dy + \int\limits_{0}^{2} y\left(\frac{1}{2} - \frac{1}{4}y\right)dy =$$
$$= \left(\frac{y^3}{12} + \frac{y^2}{4}\right)\Big|_{-2}^{0} + \left(\frac{y^2}{4} - \frac{y^3}{12}\right)\Big|_{0}^{2} = \frac{8}{12} - 1 + 1 - \frac{8}{12} = 0.$$

Следует отметить, что можно было без вычислений интегралов найти значения математических ожиданий, используя тот факт, что обе функции симметричны относительно оси ординат.

Следовательно, математическое ожидание случайного вектора (ξ, η) равно нуль-вектору $(0; 0)$.

Корреляционный момент (ковариация) $K_{\xi\eta}$ вычисляется по формуле (6.20)

$$K_{\xi\eta} = M(\xi\eta) - M(\xi)M(\eta).$$

Вычислим $M(\xi\eta)$

$$M(\xi\eta) = \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} xy\, f(x,y)\,dx\,dy = \frac{1}{4}\int\limits_{G} xy\,dxdy = \frac{1}{4}\int\limits_{-1}^{1} x\,dx \int\limits_{x-1}^{x+1} y\,dy =$$

$$= \frac{1}{4}\int\limits_{-1}^{1} x\,dx \cdot \frac{y^2}{2}\Bigg|_{x-1}^{x+1} = \frac{1}{8}\int\limits_{-1}^{1} x(x^2+2x+1-x^2+2x-1)\,dx =$$

$$= \frac{1}{2}\int\limits_{-1}^{1} x^2\,dx = \frac{1}{2}\frac{x^3}{3}\Bigg|_{-1}^{1} = \frac{1}{3}.$$

Теперь найдем корреляционный момент

$$K_{\xi\eta} = M(\xi\eta) - M(\xi)M(\eta) = \frac{1}{3}.$$

Следовательно случайные величины ξ и η находятся в корреляционной зависимости.

3) Найти $P((\xi,\eta) \in D)$, где $D = \{(x,y) | x^2 + y^2 \leqslant 1\}$.

Найдём площадь круга радиуса 1, за вычетом двух сегментов круга выходящих за пределы параллелограмма. Эта площадь равна

$S_1 = 2 \cdot (\frac{\pi}{4} + 0{,}5 \cdot 1 \cdot 1) = \frac{\pi}{2} + 1.$

$$P((\xi,\eta) \in D) = \iint\limits_{D \bigcap G} \frac{1}{4} dxdy = \frac{S_1}{4} = \frac{\frac{\pi}{2}+1}{4} = \frac{\pi}{8} + \frac{1}{2}. \blacktriangleleft$$

6.7. Двумерное нормальное распределение

Определение 6.14. *Двумерным нормальным распределением (нормальным законом распределения на плоскости) называют распределение непрерывной двумерной случайной величины $(\xi; \eta)$ с плотностью:*

$$f(x;y) = \frac{1}{2\pi\sigma_\xi\sigma_\eta\sqrt{1-r_{\xi\eta}^2}}\times$$
$$\times exp\left(-\frac{1}{2(1-r_{\xi\eta}^2)}\left(\frac{(x-a_\xi)^2}{\sigma_\xi^2}+\frac{(y-a_\eta)^2}{\sigma_\eta^2}-2r_{\xi\eta}\frac{(x-a_\xi)}{\sigma_\xi}\frac{(y-a_\eta)}{\sigma_\eta}\right)\right). \tag{6.24}$$

Можно доказать, что его параметры имеют следующий вероятностный смысл: $a_\xi = M(\xi)$, $a_\eta = M(\eta)$, $\sigma_\xi^2 = D(\xi)$, $\sigma_\eta^2 = D(\eta)$, $r_{\xi\eta}$ — коэффициент корреляции ξ и η.

Замечание 6.4. *Используя формулы (6.13), можно доказать, что составляющие ξ и η имеют нормальное распределение с параметрами $\xi \sim N(a_\xi;\, \sigma_\xi)$ и $\eta \sim N(a_\eta; \sigma_\eta)$ соответственно.*

Теорема 6.5. *Если составляющие двумерной нормальной случайной величины некоррелированы, то они независимы.*

Доказательство. Если $r_{\xi\eta} = 0$, то из (6.24) следует, что

$$f(x;y) = \frac{1}{\sigma_\xi\sqrt{2\pi}}\cdot e^{-\frac{(x-a_\xi)^2}{2\sigma_\xi^2}}\cdot\frac{1}{\sigma_\eta\sqrt{2\pi}}\cdot e^{-\frac{(y-a_\eta)^2}{2\sigma_\eta^2}} = f_\xi(x)\cdot f_\eta(y).$$

Т.е. двумерная плотность равна произведению плотностей составляющих, что в соответствии со следствием 6.3 означает их независимость.

Итак, для нормального распределения двумерной случайной величины понятие некоррелированности и независимости составляющих равносильны.

Определение 6.15. *Если обе функции регрессии η на ξ (т.е. $y = M(\eta/\xi = x)$) и ξ на η (т.е. $x = M(\xi/\eta = y)$) линейны, то говорят, что ξ и η связаны линейной корреляционной зависимостью.*

Теорема 6.6. *Составляющие двумерной нормальной случайной величины связаны линейной корреляционной зависимостью.*

Доказательство. Обозначив $u = \dfrac{x - a_\xi}{\sigma_\xi}$, $v = \dfrac{y - a_\eta}{\sigma_\eta}$, запишем плотность (6.24) в виде:

$$f(x;y) = \frac{1}{2\pi\sigma_\xi\sigma_\eta\sqrt{1 - r_{\xi\eta}^2}} \cdot e^{-\frac{1}{2(1 - r_{\xi\eta}^2)}\left(u^2 + v^2 - 2r_{\xi\eta}u \cdot v\right)}.$$

Плотность распределения составляющей ξ в соответствии с замечанием 6.4 имеет вид:

$$f_\xi(x) = \frac{1}{\sigma_\xi\sqrt{2\pi}}\, e^{-\frac{u^2}{2}}.$$

Найдём условную плотность распределения η при фиксированной ξ:

$$f(y/\xi = x) = \frac{f(x;y)}{f_\xi(x)} = \frac{1}{\sqrt{2\pi}\sigma_\eta\sqrt{1 - r_{\xi\eta}^2}} \cdot exp\left(-\frac{1}{2(1 - r_{\xi\eta}^2)}(v - r_{\xi\eta}u)^2\right) =$$

$$= \frac{1}{\sqrt{2\pi}\left(\sigma_\eta\sqrt{1 - r_{\xi\eta}^2}\right)} \cdot exp\left(-\frac{\left(\frac{y - a_\eta}{\sigma\eta} - r_{\xi\eta}\frac{x - a_\xi}{\sigma_\xi}\right)^2}{2\left(\sqrt{1 - r_{\xi\eta}^2}\right)^2}\right) =$$

$$= \frac{1}{\sqrt{2\pi}\left(\sigma_\eta\sqrt{1 - r_{\xi\eta}^2}\right)} \cdot exp\left(-\frac{\left(y - \left(a_\eta + r_{\xi\eta}\frac{\sigma_\eta}{\sigma_\xi}(x - a_\xi)\right)\right)^2}{2\left(\sigma_\eta\sqrt{1 - r_{\xi\eta}^2}\right)^2}\right).$$

Как видим, полученное условное распределение нормально с математическим ожиданием (функцией регрессии η на ξ):

$$M(\eta/\xi = x) = a_\eta + r_{\xi\eta}\frac{\sigma_\eta}{\sigma_\xi}(x - a_\xi)$$

и дисперсией $\sigma_\eta^2(1 - r_{\xi\eta}^2)$.

Аналогично можно получить функцию регрессии ξ на η:

$$M(\xi/\eta = y) = a_\xi + r_{\xi\eta}\frac{\sigma_\xi}{\sigma_\eta}(y - a_\eta).$$

Так как обе функции регрессии линейны, утверждение теоремы доказано.