Лекция 6

# Система двух случайных величин

Системой случайных величин или случайным вектором называют две или более СВ, в совокупности представляющие некоторый объект той или иной природы.

**Совместная функция распределения системы СВ и ее свойства**

Хотя каждую СВ системы можно рассматривать в отдельности и описывать их индивидуальными функциями распределения $F_X(x)$, $F_Y(y)$, суть системы СВ в том, что *совместная функция распределения*

$$F(x, y) = P(X<x \cap Y<y) \equiv P(X<x,\ Y<y) \tag{6.1}$$

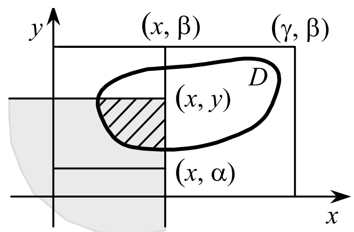
содержит информацию, не выводимую из частных законов распределения. Только в случае независимых $X$ и $Y$ совместную функцию распределения можно получить перемножением частных законов:

$$F(x, y) = P(X<x)\ P(Y<y) = F_X(x)\ F_Y(y). \tag{6.2}$$

Независимо от природы $X$, $Y$ их можно рассматривать как координаты случайного вектора на плоскости. Функцию распределения $F(x, y)$ можно наглядно представить как вероятность попадания случайной точки $(X, Y)$ в бесконечный по отрицательным направлениям квадрант с вершиной $(x, y)$. Значения $F(x, y)$ в области возможных значений $D$ заключены между 0 и 1, как вероятности попадания случайной точки в часть этой области (рис. 6.1). Если хотя бы один из аргументов меньше минимального возможного значения (точка $(x, \alpha)$), функция распределения равна нулю как вероятность невозможного события; если один аргумент превосходит возможные значения (точка $(x, \beta)$), соответствующее событие $(Y<\beta)$ достоверно и потому функция распределения принимает значение частной функции распределения другой СВ: $F(x, \beta) = F_X(x)$. Если же в квадрант $(x< \gamma, y<\beta)$ попадает вся область возможных значений, $F(\gamma, \beta) = 1$. В общем случае:

$$F(\infty, y) = F_Y(y), \qquad F(x, \infty) = F_X(x),$$

$$F(-\infty, y) = F(x, -\infty) = 0,\ \ F(\infty,\infty)=1.$$




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

$F(x,\alpha) = P(X < x,\ Y < \alpha) = 0$,

$F(x,\beta) = P(X < x,\ Y < \beta) = F_X(x)$,

$F(\gamma,\beta) = P(X < \gamma,\ Y < \beta) = 1$.

Рис. 6.1. Свойства функции распределения двух СВ

**Вероятность попадания в прямоугольник**

Подобно тому, как разность значений функции распределения одной СВ на концах интервала дает вероятность попадания СВ в этот интервал, по значениям совместной функции распределения двух СВ в вершинах прямоугольника (рис. 6.2) можно вычислить вероятность попадания случайной точки $(X, Y)$ в эту область:

$$P(\alpha \le X < \beta, \gamma \le Y < \delta) = F(\beta, \delta\ ) - F(\alpha, \delta) - F(\beta,\gamma) + F(\alpha,\gamma).$$

Если система СВ непрерывна, можно говорить о вероятности попадания в бесконечно малый прямоугольник и о плотности вероятности.

### *Система двух непрерывных случайных величин*

**Плотность совместного распределения системы СВ**

Система ($X$, $Y$) называется непрерывной, если $F(x, y)$ – непрерывная функция, дифференцируемая по каждому аргументу и имеющая вторую смешанную производную. Отношение вероятности попадания точки ($X$, $Y$) в элементарную площадку $\Delta R$ (рис. 6.2) к ее площади – *плотность распределения вероятности* (плотность вероятности) – совпадает с определением второй смешанной производной от функции распределения:

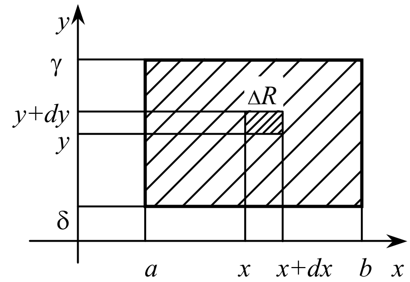


Рис. 6.2. Вероятность попадания в прямоугольник

$$f(x, y) = \lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \frac{P((X, Y) \in \Delta R)}{\Delta x \Delta y} =$$

$$= \lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \frac{F(x + \Delta x, y + \Delta y) - F(x + \Delta x, y) - F(x, y + \Delta y) + F(x, y)}{\Delta x \Delta y} = \frac{\partial^2 F(x, y)}{\partial x \partial y}.$$

**Элемент вероятности и вероятность попадания в произвольную область**

Элемент вероятности двумерного распределения $f(x, y)dxdy$ геометрически его можно представить как объем столбика на элементарном прямоугольнике под поверхностью $f(x, y)$. Вероятность попадания в конечную область $D$ можно вычислить интегрированием по этой области:

$$P((X, Y) \in D) = \int_D f(x, y)dxdy, \quad (6.3)$$

а если $D$ – прямоугольник со сторонами, параллельными осям, интегрирование легко осуществить, по крайней мере, численными методами:

$$P(\alpha \le X < \beta, \delta \le Y < \gamma) = \int_\alpha^\beta \int_\delta^\gamma f(x, y)dxdy. \quad (6.4)$$

**Основное свойство плотности распределения**

В частности, функцию распределения можно определить по известной плотности вероятности

$$F(x, y) = (-\infty \le X < x, -\infty \le Y < y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(x, y)dxdy,$$

а так как $F(x, y) = 1$ при $x \to \infty$, $y \to \infty$, плотность распределения должна обладать свойством

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)dxdy = 1. \quad (6.6)$$

**Интегральная формула полной вероятности**

По условным вероятностям $P(A|x, y)$ события $A$, рассматривая элементы вероятностей как вероятности гипотез для событий $x < X \le x + dx$, $y < Y \le y + dy$, вероятность события найдем по интегральной формуле полной вероятности:

$$P(A) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(A \mid x, y) f(x, y)dxdy. \quad (6.7)$$

**Связь частных распределений с совместной плотностью**

Как и функция распределения, плотность вероятности содержит всю информацию о системе непрерывных СВ. Интегрируя $f(x, y)$ определенным образом, можно получить не только функцию распределения или вероятность попадания в заданную область, но и частные плотности распределения отдельных СВ системы:

$$f_X(x) = F_X'(x) = \frac{d}{dx}\int_{-\infty}^{x}\int_{-\infty}^{\infty} f(x,y)dxdy = \int_{-\infty}^{\infty} f(x,y)dy, \qquad (6.8)$$

$$f_Y(y) = F_Y'(y) = \frac{d}{dy}\int_{-\infty}^{\infty}\int_{-\infty}^{y} f(x,y)dxdy = \int_{-\infty}^{\infty} f(x,y)dx. \qquad (6.9)$$

Итак, *чтобы получить плотность распределения одной из СВ непрерывной системы, надо проинтегрировать совместную плотность распределения по всем возможным значениям другой СВ.* Можно ли, наоборот, получить совместную плотность распределения, зная частные распределения $f_X(x)$, $f_Y(y)$? Ответ утвердителен только в том случае, когда $X$ и $Y$ независимы, так как при этом $F(x, y) = F_X(x)F_Y(y)$ и, следовательно,

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y} = F_X'(x)F_Y'(y) = f_X(x)f_Y(y)\,. \qquad (6.10)$$

**Условная плотность распределения**

Умножив обе части уравнения (6.10) на элемент площади $dxdy$, убедимся, что вероятность попадания случайной точки в элементарную площадку равна произведению вероятностей попадания в горизонтальную и вертикальную полоски, как и полагается независимым событиям:

$$f(x, y)dxdy = f_X(x)dx\, f_Y(y)dy.$$

Это следствие независимости СВ, но в общем случае события попадания в проекции нельзя считать независимыми. Ясно, например, что длина фрагмента естественного дробления может оказаться в некотором интервале с вероятностью, зависящей от массы фрагмента (большая длина вероятнее для тяжелых фрагментов, чем для легких). Вместо элемента безусловной вероятности $f_Y(y)dy = P(y < Y < y + dy)$ нужно подставить условную вероятность $P(y < Y < y + dy \mid X \in [x, x + dx])$, причем условие можно записать короче как $X = x$, не меняя его смысла (условием является то, что значение $X$ близко к $x$). Предел отношения условной вероятности попадания одной из СВ системы в бесконечно малый интервал при условии, что другая СВ приняла некоторое фиксированное значение, называется *условной плотностью распределения*:

$$f_{y|x}(y \mid x) = \lim_{\Delta y \to 0}\frac{P(y < Y < y + \Delta y \mid X = x)}{\Delta y} \qquad (6.11)$$

Аналогичный смысл имеет и функция $f_{x|y}(x|y)$ – условная плотность распределения СВ $X$ при $Y = y$.

**Теорема умножения плотностей**

Из соотношения между элементами вероятностей

$$f(x, y)dxdy = f_X(x)dx\, f_{y|x}(y|x)dy, \qquad (6.12)$$

следует правило *умножения плотностей*:

$$f(x, y) = f_X(x)\, f_{y|x}(y|x) = f_Y(y)\, f_{x|y}(x|y). \qquad (6.13)$$

Для независимых $X$ и $Y$ условные плотности равны соответствующим безусловным $f_{x|y}(x|y) = f(x)$ и $f_{y|x}(y|x) = f(y)$, и, наоборот, выполнение этих равенств доказывает независимость СВ.

**Связь между совместной и условной плотностью распределения**

Условную плотность нельзя получить из одномерных распределений, это та дополнительная информация, которая содержится в плотности совместного распределения и характеризует стохастическую связь между СВ:

$$f_{x|y}(x \mid y)=\frac{f(x,y)}{f_Y(y)}=\frac{f(x,y)}{\int_{-\infty}^{\infty} f(x,y)dx}, \qquad (6.14)$$

$$f_{y|x}(y \mid x)=\frac{f(x,y)}{f_X(x)}=\frac{f(x,y)}{\int_{-\infty}^{\infty} f(x,y)dy}. \qquad (6.15)$$

Условная плотность распределения обладает необходимыми свойствами плотности распределения: она положительна и нормирована единицей:

$$\int_{-\infty}^{\infty} f_X(x / y)dx=\int_{-\infty}^{\infty}\frac{f(x,y)}{f_Y(y)}dx=\frac{f_Y(y)}{f_Y(y)}=1.$$

**Произведение нормальных законов независимых СВ**

Создадим два объекта класса Norm_1, реализующие СВ $X \in N(0, 2)$, $Y \in N(0, 3)$, с помощью функции Net этого класса разделим полные интервалы $m \pm 4\sigma$ по обеим переменным на малые части:

```
>> X=Norm_1(0,2);Y=Norm_1(0,3);x=Net(X,30,4,1);y=Net(Y,30,4,1);
```

Считая $X$ и $Y$ независимыми, вычислим согласно формуле (6.9) совместную плотность распределения $f(x, y)$ на двумерной сетке [xx, yy], построим поверхность функции $f(x, y)$ (рис. 6.1) и проверим выполнение основного свойства (6.6):

```
>> [xx,yy]=meshgrid(x,y); Z=f(X,xx).*f(Y,yy); surf(xx,yy,Z), Trap2(Z,x,y)
ans =    0.9999
```

Объем под выделенной частью поверхности практически равен 1. Численное интегрирование на прямоугольной сетке выполнено с помощью программы Trap2 (Листинг 6.1). Теперь построим сечение поверхности $f(x, y)|_{x=-2}$ и вычислим площадь под этой кривой:

```
>> x1=find(xx==-2);  f1=Trap(Z(x1),y), hold on,plot3(xx(x1),y,Z(x1),'r')
f1 =    0.1210
```

Согласно формуле (6.8) интегрирование практически по всем $y$ (в пределах интервала $m \pm 4\sigma$) – это значение $f_X(-2)$, действительно:

```
>> f1=f(X,-2)
f1 =    0.1210
```

График условной плотности распределения $f_{y|x}(y|\text{-}2)$ построим, поделив аппликаты графика $f(x, y)|_{x=-2}$ на $f_X(-2)$ согласно формуле (6.15):

```
>> plot3(xx(x1),y,Z(x1)/f1*0.2,'k')
```

Умножение аппликат на 0,2 понадобилось для выравнивания масштабов на общем рисунке. Так же вычислим $f(x, y)|_{y=-4}$ и построим график условной плотности $f_{x|y}(x|\text{-}4)$:

```
>> y1=find(yy==-4);  plot3(x,ones(size(x))*(-4),Z(y1)/f(Y,-4)*0.1,'k')
```

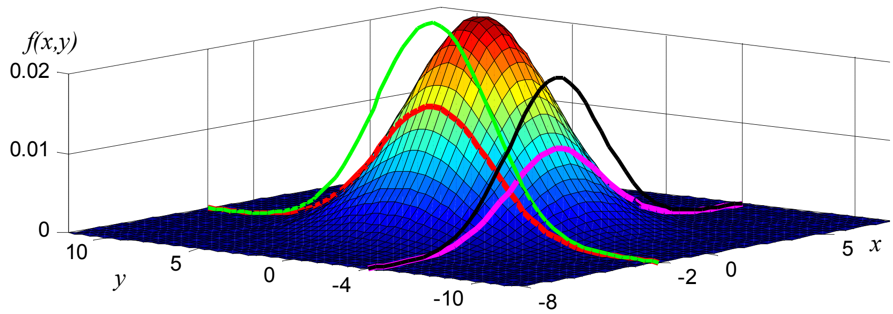



Рис. 6.3. График плотности распределения двух СВ и условных распределений

### *Система двух дискретных случайных величин*

Каждая СВ системы ($X$, $Y$) имеет свой ряд распределения $P(X = x_i) = p_i$, $i = 1, \ldots, n$, $P(Y = y_j) = g_j$, $j = 1, \ldots, m$, (возможно, $n, m \to \infty$). Сочетания пар значений ($x_i$, $y_j$) в системе имеют вероятности $P(X = x_i, Y = y_j) = p_{ij}$, $i = 1, \ldots, n$; $j = 1, \ldots, m$. Прямоугольная ($n \times m$) таблица, в которой всем парам ($x_i$, $y_j$) поставлены в соответствие вероятности $p_{ij}$, называют *матрицей распределения*.

$X$:

| $x_i$ | $\mathbf{x_1}$ | $\mathbf{x_2}$ | … | $\mathbf{x_n}$ |
|---|---|---|---|---|
| $p_i$ | $p_1$ | $p_2$ | … | $p_n$ |

$Y$:

| $y_i$ | $\mathbf{y_1}$ | $\mathbf{y_2}$ | … | $\mathbf{y_m}$ |
|---|---|---|---|---|
| $g_i$ | $g_1$ | $g_2$ | … | $g_m$ |

($X$, $Y$):

| $x_i \backslash y_i$ | $\mathbf{y_1}$ | $\mathbf{y_2}$ | … | $\mathbf{y_m}$ |
|---|---|---|---|---|
| $\mathbf{x_1}$ | $p_{11}$ | $p_{12}$ | … | $p_{1m}$ |
| $\mathbf{x_2}$ | $p_{21}$ | $p_{22}$ | … | $p_{2m}$ |
| … | … | … | … | … |
| $\mathbf{x_n}$ | $p_{n1}$ | $p_{n2}$ | … | $p_{nm}$ |

**Свойства матрицы распределения**

События ($X = x_i$), ($Y = y_j$), $i = 1, \ldots, n$; $j = 1, \ldots, m$ образуют полную группу, поэтому сумма всех элементов матрицы распределения равна единице:

$$\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij} = 1.$$

Матрица распределения содержит всю информацию о системе СВ. Из нее, в частности, можно получить распределения отдельных СВ:

$$p_i = P(X = x_i) = P\left(\sum_{j=1}^{m}(X = x_i, Y = y_j)\right) = \sum_{j=1}^{m} P(X = x_i, Y = y_j) = \sum_{j=1}^{m} p_{ij}, \qquad (6.16)$$

$$g_j = P(Y = y_j) = \sum_{i=1}^{n} p_{ij}, \qquad (6.17)$$

то есть, для того, чтобы получить вероятность возможного значения одной из СВ, надо просуммировать соответствующий этому значению столбец или строку матрицы распределения.

Если $X$ и $Y$ независимы, можно выполнить и обратную процедуру – построить матрицу распределения по одномерным распределениям:

$$p_{ij} = P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j) = p_i g_j. \qquad (6.18)$$

**Условный ряд распределения**

В общем случае приходится учитывать условные вероятности:

$$p_{ij} = P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j | X = x_i) = p_i g_{j|i}, \qquad (6.19)$$

где $g_{j|i}$, $j = 1, \ldots, m$ – условный ряд распределения СВ $Y$ при условии, что СВ $X$ приняла фиксированное значение $X = x_i$.

Учитывая, что $g_{j|i} = \dfrac{p_{ij}}{p_i}$, получить условное распределение можно из $i$-й строки матрицы распределения, разделив все ее элементы на число $p_i$ – сумму элементов этой же строки. Очевидно, что условное распределение обладает обязательным свойством дискретных распределений – сумма вероятностей всех возможных значение равна единице:

$$\sum_{j=1}^{m} g_{j|i} = \sum_{j=1}^{m}\frac{p_{ij}}{p_i} = \frac{p_i}{p_i} = 1.$$

Аналогично, из столбцов матрицы распределения можно получить условный ряд распределения СВ $X$ при фиксированном $Y = y_j$.

Функцию распределения системы дискретных СВ $F(x, y)$ можно получить из матрицы распределения суммированием всех элементов матрицы, находящихся выше $x$ и левее $y$[1].

**Полиномиальное распределение**

В испытаниях Бернулли может быть более двух исходов. Например, попадание в цель, содержащую резервированные функциональные элементы, может привести к трем возможным исходам: поражение цели при попадании в нерезервированный элемент, вывод из строя резервированного элемента без поражения цели, отсутствие эффекта при попадании в непоражаемую зону. Результатом $n$ попаданий будут различные комбинации чисел попадания в каждую из трех зон на проекции цели $m_1, m_2, m_3$. Каждую реализацию можно рассматривать как случайное событие $A_{m_1, m_2, m_3}$, вероятность которого следует вычислять по формуле (2.3). Вероятность поражения цели зависит от двух случайных чисел $X_1$, $X_2$, которые можно рассматривать как функцию случайных событий $A_{m_1, m_2, n - m_1 - m_2}$.

$$P(X_1 = k_1, X_2 = k_2) = \frac{n!}{k_1!k_2!k_3!} p_1^{k_1} p_2^{k_2} p_3^{k_3};\ k_1 + k_2 = 0, \dots, n;\ k_3 = n - k_1 - k_2. \quad (6.20)$$

**Проекции полиномиального распределения**

Для вычисления полиномиального распределения можно использовать электронную формулу p_Polynom (Листинг 6.2). В отличие от функции p_Binom она принимает вектор вероятностей каждого исхода и вектор интересующих чисел наступления каждого из событий. Многоугольник распределения на рис. 6.4 получен при $n = 12$, $p_1 = 0{,}3$, $p_2 = 0{,}4$:

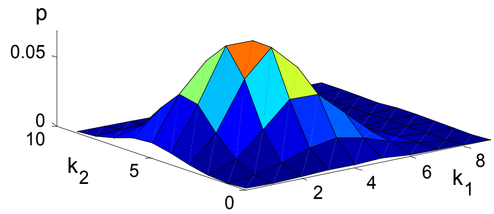


Рис. 6.4. Поверхность полиномиального распределения

```
>> n=9;p=[0.3 0.4]; [I J]=meshgrid([0:n],[0:n]); P=p_Polynom(p,n,[I(:) J(:)]);
>> P= reshape(P,size(I)); surf(I,J,P)
```

Сечения многогранника напоминают многоугольники биномиального распределения, но они таковыми не являются. Только сумма всех аппликат по направлению одной из СВ является распределением другой. Сравним эти суммы с соответствующими биномиальными распределениями, имеющих параметры $p_1$ и $p_2$ :

```
>> sum(P),Ber(p(1),n)
ans =  0.0404   0.1556   0.2668   0.2668   0.1715   0.0735   0.0210   0.0039   0.0004   0.0000
ans =  0.0404   0.1556   0.2668   0.2668   0.1715   0.0735   0.0210   0.0039   0.0004   0.0000

>> sum(P,2)'-p_Binom(p(2),n)
ans =  1.0e-015 *
        0.0087   0.0555   0.0833   0.1665   0.1110   0.0278   0.0139   0.0069    0    0
```

**Операции с дискретными распределениями**

Полиномиальный закон не имеет такого широкого применения, как биномиальный, но выполненные на этом примере операции с матрицей распределения характерны для решения задач с системами дискретных СВ. Часто бывает известен ряд распределения одной из СВ $P(Y = y_i) = g_j$ и условные вероятности другой $P(X = x_i / Y = y_j) = p_{i|j}$, а нужно построить безусловный ряд распределения СВ $X$. Построив матрицу распределения $p_{ij} = g_j p_{i|j}$, $\forall i, j$, искомый закон распределения можно вычислить суммированием соответствующих ее строк (или столбцов): $P(X = x_i) = \sum_j p_{ij}$ .

[1] Элементы матрицы распределения, так же как и ряда распределения, упорядочены по возрастанию возможных значений.

## *Числовые характеристики системы двух случайных величин*

**Моментные характеристики**

Начальными моментами порядка $k+s$ называются МО произведения степеней $X^k$, $Y^s$, а центральными – МО таких же произведений центрированных СВ $\dot{X}^k = (X - m_x)^k$, $\dot{Y}^s = (Y - m_y)^s$:

$$\alpha_{k,s}[X, Y] = M[X^k Y^s],$$
$$\mu_{k,s}[X, Y] = M[\dot{X}^k \dot{Y}^s].$$

Через плотность непрерывной системы $f(x, y)$ и матрицу распределения дискретной системы $p_{ij}$ начальные и центральные моменты вычисляются по формулам:

$$\alpha_{k,s}[X|y] = \begin{cases} \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} x^k y^s f(x,y)dxdy, \\ \sum\limits_i \sum\limits_j x_i^k y_j^s p_{ij}. \end{cases} \qquad \mu_{k,s}[X|y] = \begin{cases} \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} (x-m_x)^k (y-m_x)^s f(x,y)dxdy, \\ \sum\limits_i \sum\limits_j (x_i - m_x)^k (y_j^s - m_y)^s p_{ij}. \end{cases} \tag{6.21}$$

**Числовые характеристики условных распределений**

С системой $(X, Y)$ связаны распределения $F_X(x)$, $F_Y(y)$ и условные распределения $F_{X|y}(x|y)$, $F_{Y|x}(y|x)$, а значит и все числовые характеристики, определенные для одномерных СВ. Моменты одномерных распределений можно вычислять по соответствующим формулам через частные законы или через совместный закон по формулам (6.21) как $\alpha_{k,0}$, $\alpha_{0,s}$ и $\mu_{k,0}$ и $\mu_{0,s}$.

Числовые характеристики условных распределений не отличаются особыми свойствами, но в определенной мере они отражают связи между СВ системы. Условное МО распределения $f_{X|y}(x|y)$

$$m_{x|y} = M[X|y] = \begin{cases} \int\limits_{-\infty}^{\infty} x f_{X|y}(x|y)dx, \\ \sum\limits_i x_i p_{i|j}. \end{cases} \tag{6.22}$$

называется *регрессией* $X$ на $y$ (среднее значение СВ $X$ при условии, что $Y$ приняла значение, близкое к $y$). Аналогично определяется регрессия $Y$ на $x$. График функции $x = M[X|y]$ называется линией регрессии $X$ на $y$, а график функции $y = M[Y|x]$ – линией регрессии $Y$ на $x$.

Если $X$ и $Y$ независимы, $f_{x|y}(x|y) = f_x(x)$ и $m_{x|y} = m_x$ (а также $m_{y|x} = m_y$), линии регрессии параллельны осям координат. С другой стороны, постоянство условных МО означает лишь независимость случайных величин «в среднем», хотя зависимость условных распределений от значений другой СВ может проявляться в условных моментах старших порядков.

Условный центральный момент второго порядка называется *условной дисперсией*:

$$D[X|y) = M[\overset{\circ}{X}{}^2 |y] = \int\limits_{-\infty}^{\infty} (x - m_{x|y})^2 f_x(x|y)dx.$$

**Произведение нормальных законов зависимых СВ**

Плотность распределения двух зависимых СВ получим как произведение одномерного закона СВ $X \in N(0, 2)$ и условной плотности $f_{y|x}(y|x)$ нормального закона с постоянным СКО $\sigma_{y|x} = 3$ и линейной регрессией $m_{y|x} = -1{,}5x$. СВ $X$ представим объектом X=Norm_1(0,2), условное распределение – объектом Yx=Norm_1(0,3). Плотность $f_{y|x}(y|x)$ вычислим с помощью функции f класса Norm_1, передавая ей центрированные аргументы $y - m_{y|x} = y + 1{,}5x$. С помощью объектов X, Yx построим сетку и поверхность $f(x, y)$ на ней (рис. 6.5):

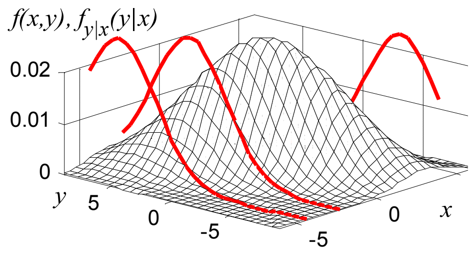



Рис. 6.5. Поверхность распределения и кривые условной плотности

```
>> X=Norm_1(0,2); Yx=Norm_1(0,3); x=Net(X,30,3);y=Net(Yx,30,3);
>> [xx,yy]=meshgrid(x,y); Z=f(X,xx).* f(Yx,yy+1.5*xx); surf(xx,yy,Z), hold on
```

Построим графики условных распределений $f_{y|x}(y|x)$ на 5-м, 10-м и 25-м узлах сетки:

```
>> for i=[5,10,25] x1=find(xx==x(i));f_1=Z(x1)/f(X,x(i))*0.2; plot3(xx(x1), yy, f_1,'r'),end
```

**Центральные моменты второго порядка**

Центральные моменты второго порядка $\mu_{2,0}[X,Y] = M[(X - m_x)^2] = D_x$ и $\mu_{0,2}[X, Y] = M[(Y - m_y)^2] = D_y$ характеризуют разброс СВ $X$, $Y$ в отдельности, а второй смешанный момент $\mu_{1,1}[X,Y] = M[(X - m_x)(Y - m_y)] = K_{xy}$ – еще и степень их взаимного разброса (*ковариацию*) или корреляционный момент. Эти величины объединяют в матрицу корреляционных моментов

$$K^{XY} = \begin{pmatrix} D_x & K_{xy} \\ K_{xy} & D_y \end{pmatrix},$$

поскольку в таком виде удобно представлять все попарные вторые центральные моменты между компонентами многомерных случайных векторов.

В случае независимых СВ матрица корреляционных моментов имеет диагональный вид, так как согласно свойствам МО для независимых СВ

$$K_{xy} = M[(X - m_x)(Y - m_y)] = M[X - m_x]M[Y - m_y] = 0.$$

Обращение в ноль коэффициента корреляции (некоррелированность) означает отсутствие линейной связи, но не исключает зависимость между СВ по основному признаку $f_{Y|x}(y|x) \neq f_y(y)$. Независимые СВ некоррелированы, но обратное утверждение, вообще говоря, неверно. В некоторых случаях корреляционный момент *зависимых* СВ может обращаться в ноль. Преобразуем выражение для $K_{xy}$ непрерывной системы следующим образом:

$$K_{xy} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x - m_x)(y - m_y)f(x, y)dxdy = \int_{-\infty}^{\infty}(x - m_x)(m_{y|x} - m_y)f_X(x)dx. \quad (6.23)$$

Действительно, выразив $f(y, x)$ по формуле умножения плотностей, интеграл по $y$ с учетом определения (6.22) можно представить в виде:

$$\int_{-\infty}^{\infty}(y - m_y)f_X(x)f_{Y|x}(y \mid x)dy = (m_{y|x} - m_y)f_X(x).$$

Из выражения для $K_{xy}$ в форме (6.23) следует, что $K_{xy} = 0$ при $m_{y|x} = m_y$, то есть *независимые в среднем СВ некоррелированы*.

**Коэффициент корреляции**

По величине $K_{xy}$ – можно судить только о наличии, но не о степени зависимости, так как ковариация учитывает и совместное, и индивидуальное рассеивание каждой СВ Не учитывает индивидуальное рассеивание второй смешанный центральный момент нормированных СВ $\dfrac{X}{\sigma_x}, \dfrac{Y}{\sigma_Y}$:

$$r_{xy} = \mu_{1,1}\left[\frac{X}{\sigma_x}, \frac{Y}{\sigma_y}\right] = M\left[\frac{X - m_X}{\sigma_x} \cdot \frac{Y - m_Y}{\sigma_y}\right] = \frac{K_{xy}}{\sigma_x \sigma_y}.$$

Эта безразмерная числовая характеристика называется коэффициентом корреляции между $X$ и $Y$, по ее величине можно судить о степени статистической связи между $X$ и $Y$.

**Границы значений коэффициента корреляции**

Наиболее сильная корреляция имеет место при линейной функциональной зависимости между СВ:

$$Y = aX + b,\ D_y = a^2 D_x,\ \sigma_y = |a|\sigma_x,$$

$$K_{xy} = M[\dot{X}(a\dot{X} + b)] = M[a\dot{X}^2 + b\dot{X}] = aD_x,$$

$$r_{xy} = \frac{K_{xy}}{\sigma_x \sigma_y} = \frac{aD_x}{\sigma_x |a| \sigma_x} = \frac{a}{|a|} = \begin{cases} 1, & a > 0, \\ -1, & a < 0. \end{cases}$$

**Линейная регрессия**

Если от $X$ линейно зависит не сама СВ $Y$, а ее условное МО

$$m_{y|x} = a(x - m_x) + m_y,$$

$Y$ уже не определяется значением аргумента, а имеет случайные отклонения относительно линии регрессии. Подстановка $a(x - m_x)$ вместо $(m_{y|x} - m_y)$ в выражение (6.23) даст такой же корреляционный момент, что и при линейной функциональной зависимости:

$$K_{xy} = \int_{-\infty}^{\infty} (x - m_x) a (x - m_x) f_X(x) dx = aD_x,$$

однако чистая мера зависимости – коэффициент корреляции – по модулю меньше единицы, так как теперь $\sigma_y > |a|\sigma_x$. В самом деле, можно доказать, что

$$D_y = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (y - m_y)^2 f(x, y) dx dy = \int_{-\infty}^{\infty} \sigma_{Y|x}^2 f_X(x) dx + a^2 D_x, \quad (6.24)$$

и если $\sigma_{Y|x} \neq 0$ ($Y$ имеет случайные отклонения от линии регрессии), $D_y > a^2 D_x$. Между коэффициентом пропорциональности, коэффициентом корреляции и дисперсиями существует связь:

$$a^2 D_x = (aD_x)^2 / D_x = K_{xy}^2 / D_x = r^2 D_y D_x / D_x = r^2 D_y. \quad (6.25)$$

Следовательно, коэффициент пропорциональности в линейном уравнении регрессии $m_{Y|x} = a(x - m_x) + m_y$ равен $a = r\sigma_y / \sigma_x$:

$$m_{Y|x} = r\frac{\sigma_y}{\sigma_x}(x - m_x) + m_y. \quad (6.26)$$

Так как $X$ и $Y$ равноправны в системе, существует такая же зависимость и для $m_{X|y}$

$$m_{X|y} = r\frac{\sigma_x}{\sigma_y}(y - m_y) + m_x, \quad (6.27)$$

причем $r$ является коэффициентом пропорциональности между центрированными, нормированными СВ и условным МО:

$$\frac{m_{Y|x} - m_y}{\sigma_y} = r\frac{x - m_x}{\sigma_x}, \quad \frac{m_{X|y} - m_x}{\sigma_x} = r\frac{y - m_y}{\sigma_y}. \quad (6.28)$$

**Нормальная регрессия**

Линейная регрессия с постоянным условным СКО $\sigma_{Y|x} = \text{const}$ называется *нормальной регрессией*. Из соотношения (6.24) при $\sigma_{Y|x} = \text{const}$ с учетом соотношения (6.25) следует

$$D_y = \sigma_{Y|x}{}^2 + a^2 D_x = \sigma_{Y|x}{}^2 + r^2 D_y,$$

откуда получим связь между безусловными и условными СКО:

$$\sigma_{Y|x} = \sqrt{1-r^2}\,\sigma_y\,,\ \sigma_{X|y} = \sqrt{1-r^2}\,\sigma_x\,. \qquad (6.29)$$

Погрешность оценки СВ $Y$ по линии регрессии не зависит от $x$ и тем меньше, чем больше коэффициент корреляции. Из тех же соотношений выразим коэффициент корреляции через коэффициент пропорциональности в уравнении линейной регрессии:

$$r = \frac{a\sigma_x}{\sigma_y} = \frac{a\sigma_x}{\sqrt{\sigma_{Y|x}^2 + a^2\sigma_x^2}} = \operatorname{sign}(a)\left(1 + \frac{\sigma_{Y|x}^2}{a^2\sigma_x^2}\right)^{-2}. \qquad (6.30)$$

Теперь ясно, что $|r_{xy}| < 1$ и корреляция тем слабее, чем меньше $\sigma_{Y|x}$.

**Построение условных распределений**

Для системы СВ $X \in N(0, 2)$, $Y \in N(0, 3)$ с условным МО $m_{y|x} = -1{,}5x$ (из предыдущего примера) построим условные распределения. Из исходных данных известны параметры: $m_x = m_y = 0$, $\sigma_x = 2$, $\sigma_{Y|x} = 3$, $a = -1{,}5$. Вычислим коэффициент корреляции по формуле (6.30):

$$r = -\left(1 + \frac{3^2}{1{,}5^2 2^2}\right)^{-2} = -\frac{\sqrt{2}}{2}.$$

По формуле (6.29) определим неизвестные параметры $\sigma_y$ и $\sigma_{X|y}$:

$$\sigma_y = \frac{\sigma_{Y|x}}{\sqrt{1-r^2}} = \frac{3}{\sqrt{1-0.5}} = 3\sqrt{2}\,,\quad \sigma_{X|y} = \sigma_x\sqrt{1-r^2} = 2\sqrt{0.5} = \sqrt{2}\,.$$

В дополнение к созданным ранее объектам создадим объект Xy – нормальное распределение с параметрами $m = 0$, $\sigma_{X|y} = \sqrt{2}$ , чтобы с его помощью построить графики условных распределений $f_{x|y}(x|y)$ для нескольких $y$:

```
>> X=Norm_1(0,2); Yx=Norm_1(0,3); x=Net(X,30,3);y=Net(Yx,30,3);
>> [xx,yy]=meshgrid(x,y); Z=f(X,xx).* f(Yx,yy+1.5*xx);
>> r=-1/sqrt(2), sX=2;sY=3/sqrt(1-r^2), sXy=sqrt(2), Xy=Norm_1(0, sXy);
r =   -0.7071   sY =    4.2426    sXy =    1.4142
```

Так как объект Xy представляет центрированную СВ, для вычисления плотности в его функцию f нужно подставлять аргументы, уменьшенные на $m_{X|y} = r\sigma_x/\sigma_y\, y$:

```
>> y_1=-9;f_1= f(Xy,x-y_1*r*sX/sY);plot3(x,y_1*ones(size(x)),f_1, 'k'), hold on
>> for i=1:length(x) plot3([x(i) x(i)],[y_1 y_1],[0 f_1(i)],'k'), end
```

Двумя последними командами построен график $f_{x|y}(x|9)$ на рис. 6.6. Повторение команд с y_1=9 дает еще один график $f_{x|y}(x|\,-9)$. Так же построим графики $f_{y|x}(y|6)$, $f_{y|x}(y|\,0)$, $f_{y|x}(y|\,-6)$:

```
>> x_1=-6;f_1= f(Yx,y-x_1*r*sY/sX);plot3(x_1*ones(size(y)),y,f_1, 'r')
>> for i=1:length(y) plot3([x_1 x_1],[y(i) y(i)],[0 f_1(i)],'r'), end
```

Построим линии регрессии по формулам (6.27), а также линии уровня $f(x.\ y)$:

```
>> plot3(x,x*r*sY/sX,zeros(size(x)),'r', y*r*sX/sY,y,zeros(size(y)),'k'),contour(xx,yy,Z), grid
```

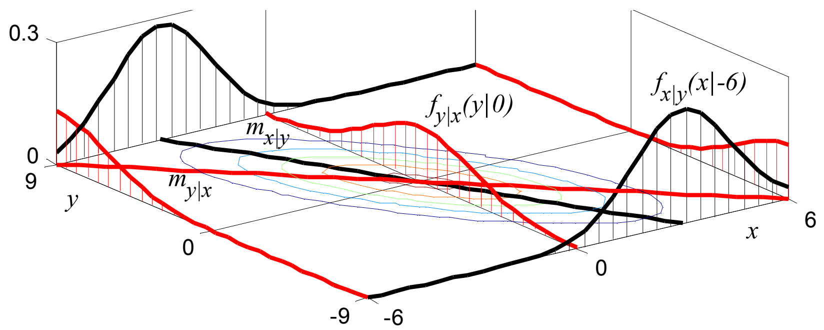



Рис. 6.6. Условные распределения с нормальной регрессией

## *Статистическое моделирование системы двух СВ*

**Статистическое оценивание параметров системы**

Чтобы разыграть реализации двух зависимых СВ, согласно правилу умножения плотностей нужно получить случайную реализацию *rand* одной из них (допустим $X$) по ее закону распределения $f_X(x)$ , а затем – другой, но уже по условному распределению $f_{y|x}(y|rand)$.

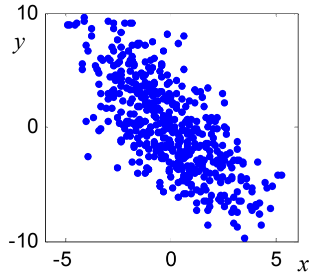



Рис. 6.7. Разброс точек с отрицательной корреляцией

В качестве примера разыграем реализации системы ($X$, $Y$) из предыдущего примера, воспользовавшись объектами X и Yx. Сравним вычисленные характеристики $r$, $\sigma_y$ и $m_y$ с их статистическими оценками, полученными с помощью функций CorrelCoef (Листинг 4.3), std (стандартное отклонение) и mean (среднее арифметическое):

```
>> X=Norm_1(0,2); Yx=Norm_1(0,3);
>> N=10000;xs=Gen(X,N);ys= Gen(X,N);
>> for i=1:N ys(i)=Gen(setval(Yx,-1.5*xs(i),[]));end
>> r=CorrelCoef([xs, ys]),sY = std(ys),mY = mean(ys)
r =   -0.7017  sY =    4.2383   mY =    0.0086
```

При большом числе испытаний $N$=10000 результат оказался очень близок к точным значениям ЧХ r=-0.7071, sY = 4.2426 и $m_y$ = 0. Выведем первые 500 точек (рис. 6.7):

```
>> plot(xs(1:500),ys(1:500),'.')
```

Расположение точек характерно для сильной отрицательной корреляции. Посмотрим, насколько увеличится погрешность оценок по меньшему объему статистики:

```
>> r=CorrelCoef(xs(1:500),ys(1:500)),sY = std(ys(1:500)),mY = mean(ys(1:500))
r =   -0.7363  sY =    4.4120  mY =    0.0971
>> r=CorrelCoef(xs(1:50),ys(1:50)),sY = std(ys(1:50)),mY = mean(ys(1:50))
r =   -0.7671  sY =    4.4228  mY =    0.2457
```

Статистические оценки достаточно близки к теоретическим параметрам распределений, чтобы ими можно оперировать как точными при большом числе испытаний порядка десятков тысяч.

**Программа верификации кода MATLAB**

```
clear all
X=Norm_1(0,2);Y=Norm_1(0,3);x=Net(X,30,4,1);y=Net(Y,30,4,1);
[xx,yy]=meshgrid(x,y); Z=f(X,xx).*f(Y,yy); surf(xx,yy,Z), Trap2(Z,x,y)
x1=find(xx==-2);  f1=Trap(Z(x1),y), hold on,plot3(xx(x1),y,Z(x1),'r')
f1=f(X,-2)
plot3(xx(x1),y,Z(x1)/f1*0.2,'k')
y1=find(yy==-4);  plot3(x,ones(size(x))*(-4),Z(y1)/f(Y,-4)*0.1,'k')

clear all
n=9;p=[0.3 0.4]; [I J]=meshgrid([0:n],[0:n]); P=p_Polynom(p,n,[I(:) J(:)]);
P= reshape(P,size(I)); surf(I,J,P)
sum(P),Ber(p(1),n)
sum(P,2)'-Ber(p(2),n)

clear all
X=Norm_1(0,2); Yx=Norm_1(0,3); x=Net(X,30,3);y=Net(Yx,30,3);
[xx,yy]=meshgrid(x,y); Z=f(X,xx).* f(Yx,yy+1.5*xx); surf(xx,yy,Z), hold on
for i=[5,10,25] x1=find(xx==x(i));f_1=Z(x1)/f(X,x(i))*0.2; plot3(xx(x1), yy,
f_1,'r'),end

clear all
X=Norm_1(0,2); Yx=Norm_1(0,3); x=Net(X,30,3);y=Net(Yx,30,3);
[xx,yy]=meshgrid(x,y); Z=f(X,xx).* f(Yx,yy+1.5*xx);
r=-1/sqrt(2), sX=2;sY=3/sqrt(1-r^2), sXy=sqrt(2), Xy=Norm_1(0, sXy);
y_1=-9;f_1= f(Xy,x-y_1*r*sX/sY);plot3(x,y_1*ones(size(x)),f_1, 'k'),hold on
for i=1:length(x) plot3([x(i) x(i)],[y_1 y_1],[0 f_1(i)],'k'), end
x_1=-6;f_1= f(Yx,y-x_1*r*sY/sX);plot3(x_1*ones(size(y)),y,f_1, 'r')
for i=1:length(y) plot3([x_1 x_1],[y(i) y(i)],[0 f_1(i)],'r'), end
plot3(x,x*r*sY/sX,zeros(size(x)),'r', y*r*sX/sY,y,zeros(size(y)),'k'),
contour(xx,yy,Z), grid

clear all
X=Norm_1(0,2); Yx=Norm_1(0,3);
N=10000;xs=Gen(X,N);ys= Gen(X,N);
for i=1:N ys(i)=Gen(setval(Yx,-1.5*xs(i),[]));end
r=CorrelCoef([xs, ys]),sY = std(ys),mY = mean(ys)
plot(xs(1:500),ys(1:500),'.')
r=CorrelCoef(xs(1:500),ys(1:500)),sY = std(ys(1:500)),mY = mean(ys(1:500))
r=CorrelCoef(xs(1:50),ys(1:50)),sY = std(ys(1:50)),mY = mean(ys(1:50))
```

## Контрольные вопросы

1. Объясните вероятностный смысл функции распределения системы двух СВ.
2. Плотность распределения системы двух непрерывных СВ.
3. Каков смысл элемента вероятностей непрерывной системы двух СВ?
4. Как связаны функция распределения и плотность распределения системы двух непрерывных СВ?
5. Какова связь между совместным и частными законами распределения независимых и зависимых СВ?
6. Правило умножения плотностей независимых и зависимых СВ. Каков смысл условного закона распределения? Как получить условный закон распределения, если известен совместный закон?
7. Основное свойство матрицы распределения. как получить ряд распределения одной из СВ по матрице распределения системы.
8. Какому закону подчиняется сумма строк и сумма столбцов матрицы распределения полиномиального закона?
9. Как вычислить условное МО одной из СВ системы по совместной плотности распределения?
10. Логическая связь между фактами зависимости (независимости) и коррелированности (некоррелированности) двух СВ? В каких случаях зависимые СВ могут быть некоррелированными?
11. Свойства линейной и нормальной регрессии.