Министерство образования и науки Российской Федерации
Ивановский государственный химико-технологический университет

Системы случайных величин

Методические указания для самостоятельной работы студентов

Составитель В.А. Бобкова

Иваново 2010

Составитель В.А. Бобкова

УДК 519.2

Системы случайных величин: метод. указания для самостоятельн. работы ст-тов / Сост. В. А. Бобкова; Иван. гос. хим.-технол. ун-т. – Иваново, 2010-28 с.

Методические указания посвящены одному из разделов курса «Теория вероятностей и математическая статистика», а именно: системам случайных величин. Дано понятие системы случайных величин, описаны способы задания систем дискретных и непрерывных случайных величин. Рассмотрены понятия зависимости и независимости случайных величин, условные законы распределения, числовые характеристики зависимости. Отдельно рассмотрены системы нормально распределенных случайных величин. Приведены графические иллюстрации и примеры решения задач.

Методические указания предназначены для самостоятельной работы студентов всех направлений подготовки.

Библиогр.: 4 назв.

Рецензент доктор технических наук, профессор А. Н. Лабутин
(Ивановский государственный химико-технологический университет)

1. Основные сведения о системах случайных величин и о способах их задания

1.1. Понятие о системе случайных величин

В практических применениях теории вероятностей приходится иметь дело с задачами, в которых результат опыта описывается не одной случайной величиной, а двумя или более случайными величинами, образующими систему или случайный вектор.

Например, успеваемость наудачу взятого студента характеризуется несколькими оценками, полученными им в ходе экзаменационной сессии; на урожайность данной сельскохозяйственной культуры влияют погодные условия, применяемые удобрения, характер почвы, качество посевного материала и так далее.

Случайным вектором (n-мерной случайной величиной, системой n случайных величин) называют упорядоченный набор из n случайных величин (X_1, X_2, … , X_n).

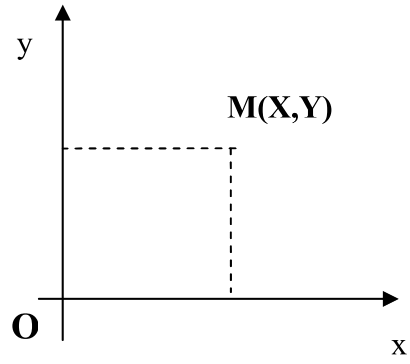
Одномерные случайные величины $\boldsymbol{X}_1$, $\boldsymbol{X}_2$, … , $\boldsymbol{X}_n$ называются ***компонентами*** или ***составляющими*** n-мерной случайной величины ($\boldsymbol{X}_1$, $\boldsymbol{X}_2$, … , $\boldsymbol{X}_n$). Их можно рассматривать как координаты случайной точки или случайного вектора $\overline{X} = (X_1, X_2, ..., X_n)$ в пространстве n измерений.

Системы случайных величин могут быть ***дискретными, непрерывными*** и ***смешанными*** в зависимости от типа случайных величин, образующих систему. В первом случае компоненты этих систем являются дискретными случайными величинами, во втором – непрерывными случайными величинами, в третьем – случайными величинами разных типов.

Рассмотрим сначала наиболее простой случай – систему, состоящую из двух случайных величин (двумерную случайную величину).

Пример: станок-автомат штампует стальные плитки. Контролируемыми размерами являются длина X и ширина Y. Имеем двумерную случайную величину (X,Y).

Геометрически двумерную случайную величину (X,Y) можно истолковать либо как случайную точку M(X,Y) на плоскости (то есть как точку со случайными координатами), либо как случайный вектор **OM** (рис. 1 и рис. 2).



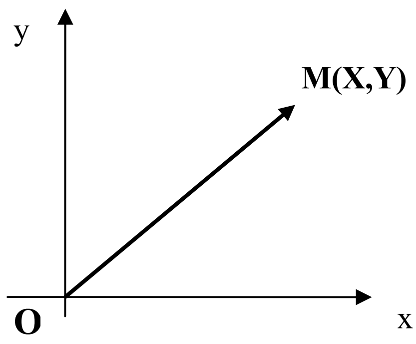


Рис.1. Рис. 2

1.2.Функция распределения вероятностей двумерной случайной величины и её свойства

Универсальной формой задания двумерной случайной величины является функция распределения (или «интегральная функция»).

Функцией распределения двумерной случайной величины (X,Y) называют вероятность совместного выполнения двух неравенств {X<x} и {Y<y}:

$$F(x, y) = p\{X < x; Y < y\}. \qquad (1)$$

Геометрическая интерпретация: если двумерную случайную величину (X,Y) рассматривать как случайную точку в прямоугольной декартовой системе координат, то функция распределения F(x,y) есть вероятность попадания случайной точки (X,Y) в бесконечный квадрант с вершиной в точке (x,y), лежащий левее и ниже её (рис.3):

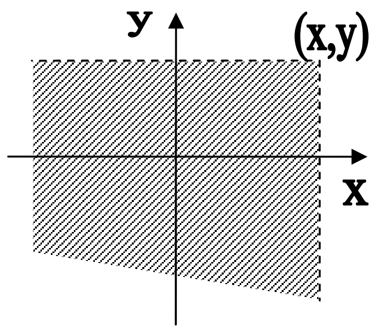


Рис. 3. Геометрическая интерпретация функции распределения двумерной случайной величины

В аналогичной интерпретации функция распределения первой компоненты X случайного вектора – обозначим её $F_1(x)$ – представляет собой вероятность попадания случайной точки в полуплоскость, ограниченную справа абсциссой x (рис. 4); функция распределения величины Y – $F_2(y)$ – вероятность попадания в полуплоскость, ограниченную сверху ординатой y (рис. 5):

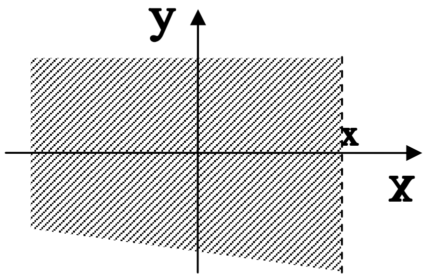


Рис. 4. Геометрическая интерпретация функции распределения первой компоненты $F_1(x)$ двумерной случайной величины (X,Y).

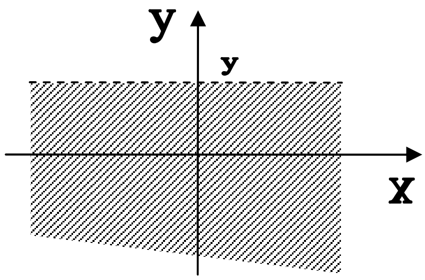


Рис. 5. Геометрическая интерпретация функции распределения второй компоненты $F_2(y)$ двумерной случайной величины (X,Y).

Пример 1. Найти вероятность того, что в результате испытания составляющая X двумерной случайной величины (X,Y) примет значение X<2, при этом составляющая Y примет значение Y<3, если известна функция распределения системы

$$F(x,y)=(\frac{1}{\pi}arctg\frac{x}{2}+\frac{1}{2})\cdot(\frac{1}{\pi}arctg\frac{y}{3}+\frac{1}{2})$$

Решение: $F(x,y)=P\{X<x;Y<y\}$

Тогда

$$p\{X<2;Y<3\}=F(2,3)=(\frac{1}{2}arctg\frac{2}{2}+\frac{1}{2})\cdot(\frac{1}{\pi}arctg\frac{3}{3}+\frac{1}{2})=$$

$$=(\frac{1}{\pi}\cdot\frac{\pi}{4}+\frac{1}{2})\cdot(\frac{1}{\pi}\cdot\frac{\pi}{4}+\frac{1}{2})=\frac{3}{4}\cdot\frac{3}{4}=\frac{9}{16}=0{,}5625$$

Ответ: 0,5625.

Свойства функции распределения двумерной случайной величины

1) $$0\le F(x,y)\le 1, \qquad (2)$$

так как это вероятность.

2) F(x,y) есть неубывающая функция своих аргументов, то есть

$$F(x_2,y)\ge F(x_1,y) \text{ при } x_2>x_1, \qquad (3)$$

$$F(x,y_2)\ge F(x,y_1) \text{ при } y_2>y_1. \qquad (4)$$

Доказательство. При увеличении какого-либо из аргументов (x,y) заштрихованная на рис.1 область увеличивается, значит, вероятность попадания в неё случайноё точки (X,Y) не может уменьшаться.

3) Если хотя бы один из аргументов обращается в -∞, то функция распределения F(x,y) равна нулю:

$$F(x,-\infty)=F(-\infty,y)=F(-\infty,-\infty)=0. \qquad (5)$$

Доказательство. События $\{X<-\infty\}$, $\{Y<-\infty\}$ и их произведение невозможны, следовательно, вероятности этих событий равны нулю.

4) Если оба аргумента равны +∞, то функция распределения F(x,y) равна единице:

$$F(+\infty,+\infty)=1. \qquad (6)$$

Доказательство. Событие $\{X<+\infty\}\cdot\{Y<+\infty\}$ достоверно, следовательно, его вероятность равна единице.

5) При одном из аргументов, равном +∞, функция распределения двумерного вектора превращается в функцию распределения компоненты, соответствующей другому аргументу:

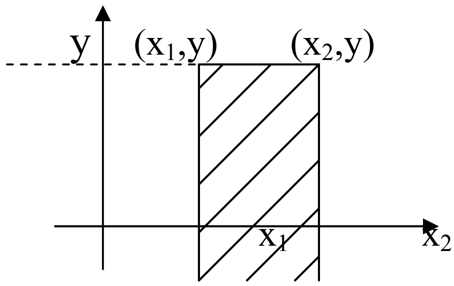
$$F(x,+\infty) = F_1(x); \quad F(+\infty, y) = F_2(y). \qquad (7)$$

Доказательство: Так как событие $Y<+\infty$ достоверно, то $F(x, +\infty)$ определяет вероятность события $X<x$, то есть представляет собой функцию распределения составляющей X.

Итак, зная совместное распределение двух случайных величин, можно найти одномерные распределения каждой из этих случайных величин, однако обратное, в общем случае, неверно.

Вероятность попадания случайной точки в полуполосу

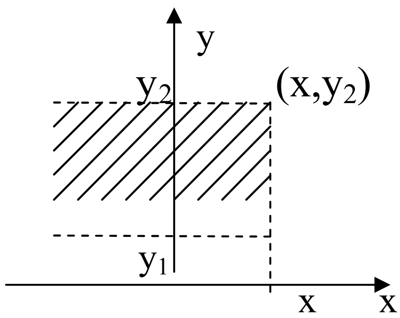
Используя функцию распределения системы случайных величин X и Y, найдем вероятность того, что в результате испытания случайная точка попадет в полуполосу $x_1<X<x_2$ и $Y<y$



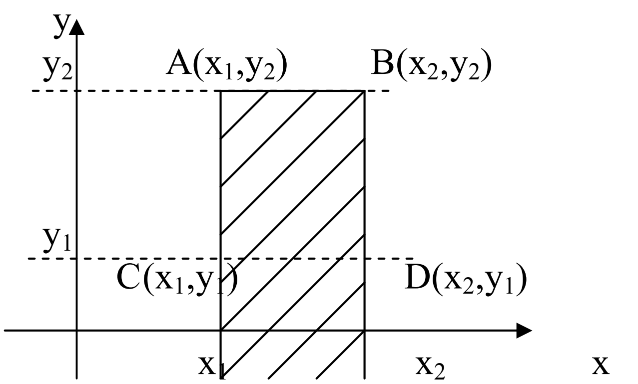
$p\{x_1<X<x_2, Y<y\} =$
$= p(x<x_2, Y<y) - p\{X<x_1, Y<y\}=$
$= F(x_2,y) - F(x_1,y)$

x

Аналогично,
$p\{X<x, y_1<Y<y_2\} = F(x,y_2) - F(x,y_1)$.
Таким образом, вероятность попадания случайной точки в полуполосу равна приращению функции
(x,y_1)
распределения по одному из аргументов.



Вероятность попадания случайной точки в прямоугольник



Рассмотрим прямоугольник со сторонами, параллельными координатным осям. Пусть уравнения сторон: $X=x_1$; $X=x_2$; $Y=y_1$; $Y=y_1$.
Тогда $p\{x_1<X<x_2, y_1<Y<y_2\} = p\{x_1<X<x_2, Y< y_2\} - p\{x_1<X<x_2, Y< y_1\} =$
$= [F(x_2,y_2)- F(x_1,y_2)]- [F(x_2,y_1)- F(x_1,y_1)]$.

Итак,

$$p\{x_1 < X < x_2, y_1 < Y < y_2\} = [F(x_2, y_2) - F(x_1, y_2)] - [F(x_2, y_1) - F(x_1, y_1)] \qquad (8)$$

Пример 2. Найти вероятность попадания случайной точки (X,Y) в прямоугольник, ограниченный прямыми $x=\pi/6$; $x=\pi/2$; $y=\pi/4$; $y=\pi/3$, если $F(x,y)=\sin(x)\sin(y)$ $(0\le x<\pi/2;\ 0\le y<\pi/2)$.

Решение:

$$P(\frac{\pi}{6} < x < \frac{\pi}{2}, \frac{\pi}{4} < y < \frac{\pi}{3}) = [F(\frac{\pi}{2}, \frac{\pi}{3}) - F(\frac{\pi}{6}, \frac{\pi}{3})] - [F(\frac{\pi}{2}, \frac{\pi}{4}) - F(\frac{\pi}{6}, \frac{\pi}{4})] =$$

$$= [\sin\frac{\pi}{2}\sin\frac{\pi}{3} - \sin\frac{\pi}{6}\sin\frac{\pi}{3}] - [\sin\frac{\pi}{2}\sin\frac{\pi}{4} - \sin\frac{\pi}{6}\sin\frac{\pi}{4}] =$$

$$= \left(\frac{\sqrt{3}}{2} - \frac{\sqrt{3}}{4}\right) - \left(\frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{4}\right) = \frac{(\sqrt{3} - \sqrt{2})}{4}.$$

Ответ: $\frac{(\sqrt{3} - \sqrt{2})}{4}$.

1.3. Закон распределения вероятностей дискретной двумерной случайной величины

Законом распределения вероятностей дискретной двумерной случайной величины называется совокупность всех возможных значений этой случайной величины, то есть пар чисел (x_i, y_i), и их вероятностей $p(x_i, y_i)$:

$$p_{ij} = p(x_i, y_i) = p\{X = x_i; Y = y_i\},\ i = \overline{1,n},\ j = \overline{1,m}$$

Здесь n, m – число возможных значений случайных величин X и Y (n и m могут быть конечными или бесконечными).

Обычно закон распределения задают в виде таблицы. Пусть $x_1, x_2, \dots, x_n$ – все возможные значения случайной величины X, $y_1, y_2, \dots, y_m$ – все возможные значения случайной величины Y. Тогда закон распределения может быть представлен в виде следующей таблицы:

Таблица 1

y / x	y_1	y_2	…	y_j	…	y_m
x_1	$p(x_1,y_1)$	$p(x_1,y_2)$	…	$p(x_1,y_j)$	…	$P(x_1,y_m)$
x_2	$p(x_2,y_1)$	$p(x_2,y_2)$	…	…	…	…
…	…	…	…	…	…	…
x_i	$p(x_i,y_1)$	…	…	$p(x_i,y_j)$	…	…
…	…	…	…	…	…	…
x_n	$p(x_n,y_1)$	…	…	…	…	$p(x_n,y_m)$

В клетке на пересечении строки x_i и столбца y_j указана вероятность $p(x_i,y_j)$ того, что двумерная случайная величина (X,Y) примет значение (x_i,y_j).

Так как события $\{X = x_i; Y = y_j\}$, $i = \overline{1,n}$, $j = \overline{1,m}$ образуют полную группу событий, то сумма вероятностей во всех клетках равна единице:

$$\sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij} = 1 \qquad (9)$$

Зная закон распределения двумерной случайной величины, можно найти законы распределения каждой из составляющих. Действительно, так как, например, события $\{X=x_1,Y=y_1\}$,$\{X=x_1,Y=y_2\}$, … , $\{X=x_1,Y=y_m\}$ несовместны, то вероятность $p(x_1)$ того, что одномерная случайная величина X примет значение x_1, по теореме сложения вероятностей несовместных событий равна

$$p_1 = p\{X = x_1\} = p(x_1, y_1) + p(x_1, y_2) + \ldots + p(x_1, y_m). \qquad (10)$$

В общем случае имеем:

$$p_i = p\{X = x_i\} = \sum_{j=1}^{m} p(x_i, y_j) = \sum_{j=1}^{m} p_{ij}\ . \qquad (11)$$

Аналогично можно записать:

$$p_j = p\{Y = y_j\} = \sum_{i=1}^{n} p_{ij}\ . \qquad (12)$$

Пример 3. Качество продукции характеризуется двумя случайными величинами: X и Y. Закон распределения случайного вектора (X,Y) представлен в таблице:

Таблица 2

x_i \ y_j	0	0,1	0,2	0,3	p_i
5	0,2	0,1	0,05	0,05	0,4
6	0	0,15	0,15	0,1	0,4
7	0	0	0,1	0,1	0,2
p_j	0,2	0,25	0,3	0,25	$\sum_i \sum_j p_{ij} = 1$

Найдём закон распределения координат X и Y случайного вектора. Вероятность события $\{X = x_i\} = p_i$ есть сумма вероятностей, находящихся в i-ой строке. Вероятности p_i находятся в последнем столбце таблицы.

Ряд распределения случайной величины X имеет вид:

Таблица 3

x_i	5	6	7

p_i	0,4	0,4	0,2

Ряд распределения Y находим, вычисляя суммы элементов столбцов таблицы 2. Эти вероятности p_j находятся в последней строке таблицы 2.

Ряд распределения случайной величины Y имеет вид:

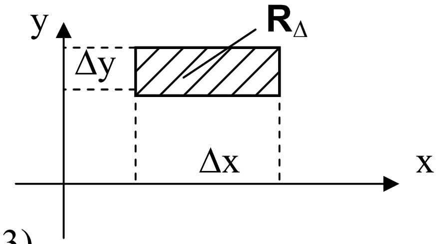
Таблица 4

y_i	0	0,1	0,2	0,3
p_j	0,2	0,25	0,3	0,25

1.4. Плотность распределения вероятностей непрерывной двумерной случайной величины и её свойства

Распределение многомерных непрерывных случайных величин обычно характеризуют плотностью распределения.

Плотностью распределения двумерной непрерывной случайной величины называют предел отношения вероятности попадания случайной величины в малый прямоугольник к площади этого прямоугольника, когда оба его размера стремятся к нулю:



$$f(x,y) = \lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \frac{p\{(x,y) \in R_\Delta\}}{\Delta x \cdot \Delta y} = \frac{\partial^2 F(x,y)}{\partial x \partial y}$$

(13)

Плотность распределения двумерной непрерывной случайной величины вычисляется как вторая смешанная частная производная от функции распределения.

Геометрически плотность распределения вероятностей $f(x,y)$ системы двух случайных величин (X,Y) представляет собой некоторую поверхность, называемую ***поверхностью распределения*** (рис. 6):

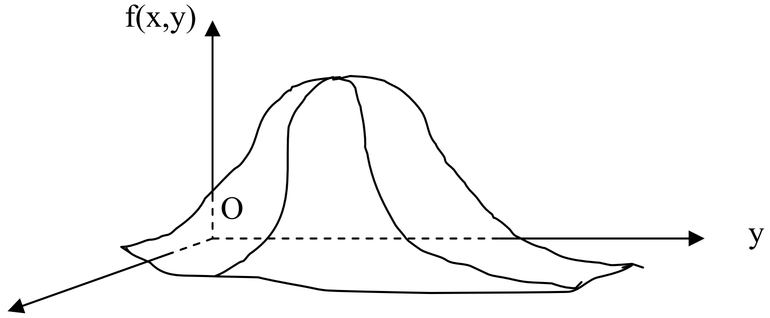


Рис. 6. Поверхность распределения системы двух случайных величин

Пример 4. Найдите плотность распределения вероятностей системы двух случайных величин (X,Y) по заданной функции распределения

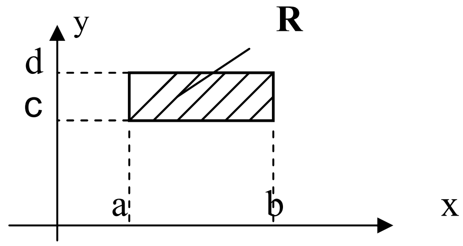
$$F(x,y)=\left(1-e^{-9x}\right)\left(1-e^{-3y}\right).$$

Решение: $f(x,y)=\dfrac{\partial^2 F(x,y)}{\partial x\partial y}=\dfrac{\partial}{\partial y}\left(\dfrac{\partial}{\partial x}\left(\left(1-e^{-9x}\right)\left(1-e^{-3y}\right)\right)\right)=$

$$=\frac{\partial}{\partial y}\left(\left(1-e^{-3y}\right)\cdot\frac{\partial}{\partial x}\left(1-e^{-9x}\right)\right)==\frac{\partial}{\partial y}\left(9e^{-9x}\cdot\left(1-e^{-3y}\right)\right)=9e^{-9x}\cdot\frac{\partial}{\partial y}\left(1-e^{-3y}\right)=9e^{-9x}\cdot 3e^{-3y}=27e^{-9x}e^{-3y}.$$

Ответ: $f(x,y)=27e^{-9x}e^{-3y}$.

Используя плотность распределения, можно записать формулу для вычисления вероятности попадания двумерной случайной величины (X,Y) в прямоугольник **R**, ограниченный прямыми $x=a,\ x=b,\ y=c,\ y=d$:



$$p\{(X,Y)\in R\}=\int_c^d\int_a^b f(x,y)dxdy$$

(14)

Свойства плотности распределения двумерной случайной величины

1) Плотность распределения двумерного случайного вектора есть функция ***неотрицательная:***

$$f(x,y)\ge 0.$$

(15)

Действительно, по определению плотность распределения есть предел отношения двух неотрицательных величин: вероятности попадания в прямоугольник и площади прямоугольника – и, следовательно, отрицательной быть не может.

2) ***Вероятность попадания случайной точки (X,Y) в область D*** равна двойному интегралу от плотности по области D:

$$p\{(X,Y)\in D\}=\iint\limits_{D} f(x,y)dxdy \qquad (16)$$

Действительно, разбив область D на прямоугольники и применив к каждому из них равенство (13), получаем, по теореме сложения вероятностей, при стремлении к нулю площадей прямоугольников (то есть при $dx\to 0$ и $dy\to 0$), формулу (15). Геометрически эта вероятность изображается объёмом цилиндрического тела, ограниченного сверху поверхностью распределения $f(x,y)$ и опирающегося на область D.

3) ***Функция распределения*** двумерной случайной величины выражается через плотность распределения следующим образом:

$$F(x,y)=\int\limits_{-\infty}^{x}\int\limits_{-\infty}^{y} f(x,y)dxdy \qquad (17)$$

Эта формула следует из (13), так как F(x,y) есть вероятность попадания в прямоугольник, ограниченный абсциссами $-\infty, x$ и ординатами $-\infty, y$.

4) ***Условие нормировки***: двойной несобственный интеграл в бесконечных пределах от плотности вероятности двумерной случайной величины равен единице:

$$\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} f(x,y)dxdy=1. \qquad (18)$$

Действительно, этот интеграл есть вероятность попадания во всю плоскость xOy, то есть вероятность достоверного события. Геометрически свойство 4 означает, что ***объём тела, ограниченного поверхностью распределения и плоскостью xOy, равен единице.***

5) ***Плотности распределения компонент*** двумерного случайного вектора могут быть получены по формулам:

$$f_1(x)=\int\limits_{-\infty}^{\infty} f(x,y)dy, \qquad (19)$$

$$f_2(y)=\int\limits_{-\infty}^{\infty} f(x,y)dx. \qquad (20)$$

Доказательство. Принимая во внимание, что

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y} f(x,y)dxdy \quad \text{и} \quad F_1(x)=F(x,\infty),$$

найдем

$$F_1(x)=\int_{-\infty}^{x}\int_{-\infty}^{\infty} f(x,y)dxdy\,.$$

Продифференцировав обе части этого равенства по x, получим формулу (19):

$$f_1(x)=\frac{dF_1}{dx}=\int_{-\infty}^{\infty} f(x,y)dy\,.$$

Аналогично выводится формула (20).

Пример 5. Двумерная случайная величина (X,Y) задана плотностью распределения вероятностей $f(x,y)=\dfrac{A}{(1+x^2)(1+y^2)}$.

Найти: 1) A; 2) $p\{0<x<1,\ 0<y<2\}$; 3) $f_1(x)$ и $f_2(y)$.

Решение: 1) Постоянную A найдём, используя условие нормировки:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{A}{(1+x^2)(1+y^2)}dxdy=1, \qquad A\int_{-\infty}^{\infty}\frac{dx}{1+x^2}\int_{-\infty}^{\infty}\frac{dy}{1+y^2}=1\ ,$$

$$A\cdot arctgx\Big|_{-\infty}^{\infty}\cdot arctgy\Big|_{-\infty}^{\infty}=1\ , \qquad A\cdot\pi^2=1\ .$$

Следовательно, $A=\dfrac{1}{\pi^2}$.

2) Используя формулу (14), находим:

$$p\{0<X<1\ ,0<Y<2\}=\int_0^2\int_0^1 f(x,y)dxdy=\int_0^2\int_0^1\frac{1}{\pi^2(1+x^2)(1+y^2)}dxdy=\frac{1}{\pi^2}\int_0^2\left(\int_0^1\frac{dx}{1+x^2}\right)\frac{dy}{1+y^2}=$$

$$=\frac{1}{\pi^2}\int_0^2(arctg1-arctg0)\frac{dy}{1+y^2}=\frac{1}{\pi^2}\cdot\frac{\pi}{4}\int_0^2\frac{dy}{1+y^2}=\frac{1}{4\pi}\cdot(arctg2-arctg0)=\frac{1}{4\pi}\cdot 1{,}107149=0{,}088104.$$

Можно было сначала по формуле (17) найти функцию распределения

$$F(x,y)=\frac{1}{\pi^2}\int_{-\infty}^{y}\left(\int_{-\infty}^{x}\frac{dx}{1+x^2}\right)\frac{dy}{1+y^2}=\frac{1}{\pi^2}\int_{-\infty}^{y}\left(arctgx\Big|_{-\infty}^{x}\right)\frac{dy}{1+y^2}=\frac{1}{\pi^2}\left(arctgx+\frac{\pi}{2}\right)\cdot\int_{-\infty}^{y}\frac{dy}{1+y^2}=$$

$$= \frac{1}{\pi^2}\left(arctgx + \frac{\pi}{2}\right) \cdot arctgy\Big|_{-\infty}^{y} = \frac{1}{\pi^2}\left(arctgx + \frac{\pi}{2}\right)\left(arctgy + \frac{\pi}{2}\right)$$

и затем воспользоваться формулой (8):

$$p\{0 < X < 1\ ,0 < Y < 2\} = [F(1;2) - F(0;2)] - [F(1;0) - F(0;0)] = (0{,}607224 - 0{,}399190) -$$

$$-(0.350059 - 0.230130) = 0{,}088104.$$

3) $$f_1(x) = \int_{-\infty}^{\infty} \frac{1}{\pi^2} \cdot \frac{dy}{(1+x^2)(1+y^2)} = \frac{1}{\pi^2(1+x^2)} \int_{-\infty}^{\infty} \frac{dy}{1+y^2} = \frac{1}{\pi^2(1+x^2)} \cdot arctgy\Big|_{-\infty}^{\infty} =$$

$$= \frac{1}{\pi^2(1+x^2)} \cdot \left(\frac{\pi}{2} + \frac{\pi}{2}\right) = \frac{1}{\pi^2(1+x^2)} \cdot \pi = \frac{1}{\pi(1+x^2)}\ ,$$

$$f_2(y) = \int_{-\infty}^{\infty} \frac{1}{\pi^2} \cdot \frac{dx}{(1+x^2)(1+y^2)} = \ \dots\ = \frac{1}{\pi(1+y^2)}\ .$$

Ответ:

1) $A = \frac{1}{\pi^2}$; 2) $p\{0 < X < 1\ ,0 < Y < 2\} = 0{,}088104$; 3) $f_1(x) = \frac{1}{\pi(1+x^2)}$; $f_2(y) = \frac{1}{\pi(1+y^2)}$.

1.5. Система n случайных величин

Перейдём теперь к системе n случайных величин или к n-мерной случайной величине.

Функция распределения n-мерной случайной величины - это вероятность совместного выполнения неравенств вида: $\{X_i<x_i\}$:

$$F(x_1,\ x_2,\ \dots\ ,\ x_n) = p\{\ X_1<x_1;\ X_2<x_2;\ \dots\ ;\ X_n<x_n\}. \qquad (21)$$

Свойства функции распределения аналогичны двумерному случаю, в частности:

$$F(x_1,\ x_2,\ \dots\ ,\ x_n) \in [0;1];\ F(\infty,\ \infty,\ \dots\ ,\ \infty) = 1;\ F_i(x_i) = F(\infty,\ \dots\ ,\ x_i,\ \dots\ ,\ \infty). \qquad (22)$$

Закон распределения дискретной n-мерной случайной величины - это совокупность всех возможных значений, которые может принимать n-мерная случайная величина, и их вероятностей:

$$p_{i_1 i_2 \dots i_n} = p\{X_1 = x_{i_1}; X_2 = x_{i_2};\ \dots\ ; X_n = x_{i_n}\}\ . \qquad (23)$$

Для непрерывной n-мерной случайной величины вводят понятие плотности распределения.

Плотность распределения n-мерной случайной величины $(X_1, X_2,\ \dots\ , X_n)$ – это n-ная смешанная частная производная функции распределения $F(x_1, x_2,\ \dots\ , x_n)$, взятая один раз по каждому аргументу:

$$f(x_1, x_2, \ldots, x_n) = \frac{\partial^n F(x_1, x_2, \ldots, x_n)}{\partial x_1 \partial x_2 \ldots \partial x_n}. \qquad (24)$$

При этом: $f(x_1, x_2, \ldots, x_n) \geq 0$,

(25)

$$\underbrace{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty}}_{n} f(x_1, x_2, \ldots, x_n) dx_1 dx_2 \ldots dx_n = 1,$$

(26)

$$p\{(X_1, X_2, \ldots, X_n) \in D\} = \int\int\ldots\int_{(D)} f(x_1, x_2, \ldots, x_n) dx_1 dx_2 \ldots dx_n . \qquad (27)$$

Функция распределения выражается через плотность ***n***-кратным интегралом

$$F(x_1, x_2, \ldots, x_n) = \underbrace{\int_{-\infty}^{x_1}\int_{-\infty}^{x_2}\ldots\int_{-\infty}^{x_n}}_{n} f(x_1, x_2, \ldots, x_n) dx_1 dx_2 \ldots dx_n. \qquad (28)$$

2. Зависимость и независимость случайных величин

Ранее было показано, как, зная закон распределения вероятностей системы двух случайных величин (двумерного вектора), определить законы распределения отдельных случайных величин, входящих в систему (то есть компонент вектора).

Естественно, возникает вопрос: нельзя ли по законам распределения вероятностей отдельных величин, входящих в систему, найти закон распределения системы? Оказывается, ***в общем случае этого сделать нельзя***. Это можно сделать только в том случае, когда случайные величины X и Y независимы.

a. Независимые случайные величины

Случайные величины X и Y называются ***независимыми***, если события {X<x} и {Y<y} независимы при любых x и y. В противном случае величины называются ***зависимыми***.

Теорема. Для того, чтобы случайные величины X и Y были независимы, необходимо и достаточно, чтобы функция распределения системы (X,Y) была равна произведению функций распределения составляющих:

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

(29)

Доказательство

а) Необходимость. Пусть X и Y независимы. Тогда события A={X<x} и B={Y<y} независимы, следовательно, вероятность совмещения этих событий равна произведению их вероятностей:

$$\underbrace{p\{X<x,Y<y\}}_{F(x,y)}=p(A\cdot B)=p(A)\cdot p(B)=\underbrace{p\{X<x\}}_{F_1(x)}\cdot\underbrace{p\{Y<y\}}_{F_2(y)},$$

то есть $$F(x,y)=F_1(x)\cdot F_2(y).$$

б) Достаточность. Пусть $F(x,y)=F_1(x)\cdot F_2(y)$. Отсюда

$$p\{X<x,Y<y\}=p\{X<x\}\cdot p\{Y<y\},$$

то есть вероятность совмещения событий {X<x} и {Y<y} равна произведению вероятностей этих событий. Следовательно, случайные величины X и Y независимы.

Следствие 1. Необходимым и достаточным условием независимости двух непрерывных случайных величин, образующих систему (X,Y), является равенство

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

(30)

Следствие 2. Необходимым и достаточным условием независимости двух дискретных случайных величин, образующих систему (X,Y), является равенство

$$p\{X=x_i,Y=y_j\}=p\{X=x_i\}\cdot p\{Y=y_j\} \qquad (31)$$

для любых $i=1,\ 2,\ \ldots\ ,\ n,\quad j=1,\ 2,\ \ldots\ ,\ m$

Пример 6. Плотность распределения системы непрерывных случайных величин (X,Y) имеет вид:

$$f(x,y)=\frac{1}{\pi^2(x^2+y^2+x^2y^2+1)}.$$

Определить, зависимы или независимы случайные величины X и Y.
Решение: Разлагая знаменатель на множители, получаем:

$$f(x,y)=\frac{1}{\pi(1+x^2)}\cdot\frac{1}{\pi(1+y^2)}.$$

Из того, что функция f(x,y) распалась на произведение двух функций, из которых одна зависит только от x, а другая только от y, заключаем, что величины X и Y независимы.

Пример 7. Определить, зависимы или независимы дискретные случайные величины X и Y, если закон распределения случайного вектора (X,Y) представлен в таблице 2 (см. пример 3 на стр.8).
Решение: Из таблицы 2 имеем $p\{X=5,Y=0\}=0{,}2$; из таблицы 3 имеем $p\{X=5\}=0{,}4$; из таблицы 4 имеем: $p\{Y=0\}=0{,}2$.

Поскольку $p\{X=5,Y=0\}=0{,}2\neq p\{X=5\}\cdot p\{Y=0\}=0{,}2\cdot 0{,}4=0{,}08$, можно сделать вывод, что величины X и Y зависимы.

Понятие независимости можно обобщить на случай ***n*** величин. Случайные величины $X_1, X_2, \ldots, X_n$ называются ***независимыми в совокупности (или взаимно независимыми)***, если события $\{X_1<x_1\}, \{X_2<x_2\}, \ldots, \{X_n<x_n\}$ независимы в совокупности (взаимно независимы) при любых $(x_1, x_2, \ldots, x_n)$.

Необходимым и достаточным условием взаимной независимости n случайных величин является равенство:

$$F_{X_1,X_2,\ldots,X_n}(x_1,\ x_2,\ \ldots,\ x_n) = F_{X_1}(x_1)\cdot F_{X_2}(x_2)\cdot\ldots\cdot F_{X_n}(x_n)\ ; \qquad (32)$$

для ***n*** непрерывных случайных величин

$$f_{X_1,X_2,\ldots,X_n}(x_1,\ x_2,\ \ldots,\ x_n) = f_{X_1}(x_1)\cdot f_{X_2}(x_2)\cdot\ldots\cdot f_{X_n}(x_n)\ . \qquad (33)$$

2.2. Условные законы распределения

Обратимся теперь к зависимым величинам. Вероятностная зависимость между случайными величинами часто встречается на практике. Если случайные величины X и Y находятся в вероятностной зависимости, это не означает, что с изменением величины X величина Y изменяется вполне определенным образом; это лишь означает, что с изменением величины X величина Y имеет тенденцию также изменяться (например, возрастать или убывать с ростом X). Эта тенденция соблюдается лишь в общих чертах, и в каком-то отдельном случае от неё возможны отступления. Примеры случайных величин, находящихся в вероятностной зависимости: рост и возраст ребенка; затраты и прибыль при производстве определенной продукции; затраты на рекламу и объем продаваемой продукции.

Для того, чтобы полностью описать систему, недостаточно знать распределение каждой из составляющих; нужно ещё знать зависимость между величинами, входящими в систему. Эта зависимость характеризуется с помощью ***условных законов распределения.***

Условным законом распределения одной из случайных величин, входящих в систему (X,Y), называется её закон распределения, найденный при условии, что другая случайная величина приняла определённое значение (или попала в какой-то интервал).

Пусть (X,Y) – дискретная двумерная случайная величина и

$$p_{ij} = p\{X = x_i,\ Y = y_j\}\ , i = 1,\ 2,\ \ldots,\ n,\quad j = 1,\ 2,\ \ldots,\ m.$$

В соответствии с определением условных вероятностей событий*), условная вероятность того, что случайная величина X примет значение x_i при условии $Y = y_j$, определяется равенством

$$p(x_i|y_j) = p\{X = x_i|Y = y_j\} = \frac{p\{X = x_i, Y = y_j\}}{p\{Y = y_j\}}, i = 1,\ 2,\ \ldots,\ n,\quad j = 1,\ 2,\ \ldots,\ m \qquad (34)$$

Совокупность вероятностей (34), то есть $p(x_1|y_j)$, $p(x_2|y_j)$, ... , $p(x_n|y_j)$, представляет собой условный закон распределения случайной величины X при условии $Y = y_j$. Сумма условных вероятностей $\sum_{i=1}^{n} p(x_i|y_j) = 1$.

Аналогично определяются условная вероятность и условный закон распределения случайной величины Y при условии $X = x_i$:

$$p(y_j|x_i) = p\{Y = y_j | X = x_i\} = \frac{p\{X = x_i, Y = y_j\}}{p\{X = x_i\}}, i = 1,\ 2,\ ...\ ,\ n, \quad j = 1,\ 2,\ ...\ ,\ m.$$

(35)

Пример 8. Пусть закон распределения двумерного случайного вектора (X,Y) задан таблицей 2 (стр. 8). Найти условный закон распределения случайной величины X при Y =0,1.

Решение. С учетом формулы (34) имеем:

$$p\{X = 5|Y = 0{,}1\} = \frac{p\{X = 5, Y = 0{,}1\}}{p\{Y = 0{,}1\}} = \frac{0{,}1}{0{,}25} = 0{,}4;$$

(значение $p\{Y = 0{,}1\} = 0{,}25$ взято из безусловного закона распределения случайной величины Y, приведенного в таблице 4 на стр. 9).

---

*) Пусть A и B – случайные события. Тогда вероятность их совместного появления равна $p(A \cdot B) = p(A) \cdot p(B|A) = p(B) \cdot p(A|B)$, где $p(B|A)$ - условная вероятность события B при условии, что событие A произошло; $p(A|B)$ - условная вероятность события A при условии, что событие B произошло. Тогда $p(B|A) = \frac{p(A \cdot B)}{p(A)}$, $p(A|B) = \frac{p(A \cdot B)}{p(B)}$.

$$p\{X = 6|Y = 0{,}1\} = \frac{p\{X = 6, Y = 0{,}1\}}{p\{Y = 0{,}1\}} = \frac{0{,}15}{0{,}25} = 0{,}6;$$

$$p\{X = 7|Y = 0{,}1\} = \frac{p\{X = 7, Y = 0{,}1\}}{p\{Y = 0{,}1\}} = \frac{0}{0{,}25} = 0.$$

Таким образом, условный закон распределения случайной величины X при Y =0,1 таков:

Таблица 5

X	5	6	7

$p_{Y=0,1}$	0,4	0,6	0

Сравнивая найденный условный закон распределения случайной величины X с безусловным законом её распределения (таблица 3 на стр. 8), видим, что они различны. Следовательно, случайные величины X и Y находятся в вероятностной зависимости.

Пусть теперь (X,Y) – непрерывная двумерная случайная величина с плотностью $f(x,y)$; $f_1(x)$ и $f_2(y)$- плотности распределения соответственно случайной величины X и случайной величины Y.

Условной плотностью распределения составляющей X при условии Y=y называют отношение плотности совместного распределения к плотности распределения составляющей Y:

$$f_X(x\,|\,y)=\frac{f(x,y)}{f_2(y)}=\frac{f(x,y)}{\int\limits_{-\infty}^{\infty}f(x,y)dx}\ ,\ f_2(y)\neq 0. \tag{36}$$

Аналогично определяется условная плотность распределения составляющей Y при условии X=x:

$$f_Y(y\,|\,x)=\frac{f(x,y)}{f_1(x)}=\frac{f(x,y)}{\int\limits_{-\infty}^{\infty}f(x,y)dy}\ ,\ f_1(x)\neq 0. \tag{37}$$

Из (36) и (37) получим:

$$f(x,y)=f_2(y)\cdot f_X(x\,|\,y)=f_1(x)\cdot f_Y(y\,|\,x). \tag{38}$$

Таким образом, плотность распределения системы двух непрерывных случайных величин равна произведению плотности одной составляющей на условную плотность другой составляющей.

Как и любая плотность распределения, условные плотности обладают следующими свойствами:

$$f_X(x\,|\,y)\geq 0;\quad f_Y(y\,|\,x)\geq 0;\quad \int\limits_{-\infty}^{\infty}f_X(x\,|\,y)dx=1;\quad \int\limits_{-\infty}^{\infty}f_Y(y\,|\,x)dy=1. \tag{39}$$

Пример 9. Непрерывный вектор (X,Y) равномерно распределен в круге с радиусом 1, то есть

$$f(x,y)=\begin{cases}\dfrac{1}{\pi}, если\ \ x^2+y^2\leq 1,\\ 0, если\ \ x^2+y^2>1.\end{cases}$$

Найти условные плотности распределения компонент этого вектора.

Решение

Условную плотность составляющей X при $|x| \le \sqrt{1-y^2}$ найдём по формуле (36):

$$f_2(y) = \int_{-\infty}^{\infty} f(x,y)dx = \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} \frac{1}{\pi}dx = \frac{1}{\pi}\cdot \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} dx = \frac{1}{\pi}\cdot x\Bigg|_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} = \frac{2}{\pi}\sqrt{1-y^2}\ ;$$

$$f_X(x\,|\,y) = \frac{f(x,y)}{f_2(y)} = \frac{\frac{1}{\pi}}{\frac{2}{\pi}\sqrt{1-y^2}} = \frac{2}{\pi}\sqrt{1-y^2}\ .$$

Так как $f(x,y)=0$ при $x^2+y^2>1$, то $f_X(x\,|\,y)=0$ при $|x|>\sqrt{1-y^2}$.

Аналогично находим:

$$f_1(x) = \frac{2}{\pi}\sqrt{1-x^2}\ ; \qquad f_Y(y\,|\,x) = \frac{f(x,y)}{f_1(x)} = \frac{2}{\pi}\sqrt{1-x^2} \quad \text{при} \quad |y| \le \sqrt{1-x^2}$$

Итак, искомые условные плотности распределения составляющих системы (X,Y) имеют вид:

$$f_X(x\,|\,y) = \begin{cases} \dfrac{1}{2\sqrt{1-y^2}},\ \text{если } x^2+y^2 \le 1 \\ 0,\ \text{если } x^2+y^2>1 \end{cases}; \qquad f_Y(y\,|\,x) = \begin{cases} \dfrac{1}{2\sqrt{1-x^2}},\ \text{если } x^2+y^2 \le 1 \\ 0,\ \text{если } x^2+y^2>1 \end{cases}.$$

Для независимых случайных величин условная плотность распределения совпадает с безусловной плотностью распределения. Действительно,

$$f_X(x\,|\,y) = \frac{f(x,y)}{f_2(y)} = \frac{f_1(x)\cdot f_2(y)}{f_2(y)} = f_1(x).$$

(40)

Аналогично $$f_Y(y\,|\,x) = f_2(y).$$

(41)

Степень зависимости между случайными величинами обычно оценивают с помощью числовых характеристик зависимости.

2.3. Числовые характеристики зависимости (ковариация, корреляция)

Основными числовыми характеристиками случайного вектора являются моменты.

Моментом порядка k,s случайного вектора (X,Y) называют математическое ожидание произведения X^k на Y^s:

$$\mu_{ks} = M(X^k \cdot Y^s) = \begin{cases} \sum\limits_i \sum\limits_j x_i^k y_j^s p_{ij}, \ \textit{если } (X,Y) \textit{ дискретен} \\ \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} x^k y^s f(x,y)dxdy, \ \textit{если } (X,Y) \textit{ непрерывен} \end{cases} \quad (42)$$

Центральным моментом порядка k,s случайного вектора (X,Y) называют математическое ожидание произведения k-ой и s-ой степени центрированных величин:

$$M(X^k \cdot Y^s) = \begin{cases} \sum\limits_i \sum\limits_j (x_i - m_x)(y_j - m_y)p_{ij}, \ \textit{если } (X,Y) \textit{ дискретен} \\ \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} (x-m_x)^k (y-m_y)^s f(x,y)dxdy, \textit{если } (X,Y) \textit{ непрерывен} \end{cases} \quad (43)$$

Здесь $m_x = M(X); m_y = M(Y)$.

Соответственно суммарному порядку k+s моменты классифицируются на первые, вторые и так далее.

На практике обычно применяют моменты первого и второго порядков. Первые моменты представляют собой математические ожидания величин X и Y:

$$M(X^1Y^0) = M(X) = m_X; \qquad M(X^0Y^1) = M(Y) = m_Y. \quad (44)$$

Совокупность математических ожиданий представляет собой характеристику положения случайного вектора (X,Y). Геометрически это координаты некоторой «средней» точки, вокруг которой происходит рассеивание (X,Y).

Вторые несмешанные центральные моменты случайного вектора – это дисперсии величин X и Y:

$$M[(X-m_x)^2 \cdot (Y-m_y)^0] = D(X); \qquad M[(X-m_x)^0 \cdot (Y-m_y)^2] = D(Y). \quad (45)$$

Они характеризуют рассеивание (разброс) случайной точки (X,Y) вокруг центра рассеивания (m_X, m_Y).

Особую роль для характеристики случайного вектора играет второй смешанный центральный момент, называемый ***ковариацией***.

Ковариация случайных величин X и Y – это математическое ожидание произведения центрированных величин

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

$$= \begin{cases} \sum\limits_i \sum\limits_j (x_i - m_X)\cdot(y_j - m_Y)\cdot p_{ij}, \text{ если величины X и Y дискретны,} \\ \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} (x-m_X)\cdot(y-m_Y)\cdot f(x,y)dxdy, \text{ если величины X и Y непрерывны} \end{cases} \quad (46)$$

Ковариация характеризует зависимость случайных величин X и Y.

Теорема. Если случайные величины X и Y независимы, то

$$\mathrm{cov}(X,Y)=0. \qquad (47)$$

Доказательство следует из свойств математического ожидания: если случайные величины X и Y независимы, то их отклонения тоже независимы; математическое ожидание произведения независимых величин равно произведению их математических ожиданий, следовательно,

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

так как $M(X-m_X)=0,\ M(Y-m_Y)=0$.

Таким образом, если $\mathrm{cov}(X,Y)\neq 0$, то случайные величины X и Y зависимы. В этом случае ($\mathrm{cov}(X,Y)\neq 0$) их называют ***коррелированными***. Однако из того, что $\mathrm{cov}(X,Y)=0$, не следует независимость X и Y. В этом случае ($\mathrm{cov}(X,Y)=0$) случайные величины называют ***некоррелированными***. Из независимости вытекает некоррелированность; обратное, вообще говоря, неверно. Другими словами, некоррелированность является необходимым, но не достаточным условием независимости случайных величин X и Y.

В качестве примера рассмотрим случайный вектор (X,Y), равномерно распределенный в круге с радиусом 1 и с центром в начале координат и имеющий следующую плотность распределения вероятностей:

$$f(x,y)=\begin{cases}\dfrac{1}{\pi}, при \quad x^2+y^2\le 1\\ 0, при\, |x|>1\end{cases}$$

Убедимся, что случайные величины X и Y зависимы:

$$f_1(x)=\int_{-\infty}^{\infty} f(x,y)dy=\begin{cases}\dfrac{2}{\pi}\sqrt{1-x^2}, при \quad |x|\le 1;\\ 0, при \quad |x|>1;\end{cases}$$

$$f_2(y)=\int_{-\infty}^{\infty} f(x,y)dx=\begin{cases}\dfrac{2}{\pi}\sqrt{1-y^2}, при \quad |y|\le 1;\\ 0, при \quad |y|>1;\end{cases}$$

Так как $f(x,y)\neq f_1(x)\cdot f_2(y)$, то случайные величины X и Y зависимы.

Вычислим ковариацию этих величин. Сначала найдем математические ожидания: $m_x=\int_{-\infty}^{\infty} xf_1(x)dx=\int_{-1}^{1} x\cdot\dfrac{2}{\pi}\sqrt{1-x^2}dx=0,$ так как подынтегральная функция нечетная, а отрезок интегрирования симметричен относительно начала координат. Аналогично получаем, что $m_Y=0$. Тогда ковариация

$$\mathrm{cov}(X,Y)=M\left[(X-m_x)\cdot(Y-m_y)\right]=M(X\cdot Y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\cdot f(x,y)dx=$$

$$=\frac{1}{\pi}\int_{-1}^{1} ydy\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} xdx=\frac{1}{\pi}\int_{-1}^{1} ydy\cdot\left.\frac{x^2}{2}\right|_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}}=\frac{1}{\pi}\int_{-1}^{1} ydy\left(\frac{1-y^2}{2}-\frac{1-y^2}{2}\right)=$$

$$=\frac{1}{\pi}\int_{-1}^{1}\frac{1-y^2-1+y^2}{2}\cdot ydy=0.$$

Таким образом, случайные величины X и Y некоррелированы, но они не являются независимыми.

Ковариация характеризует не только степень зависимости случайных величин, но и их рассеивание вокруг точки (m_X, m_Y). Так, если величина X очень мало отклоняется от своего математического ожидания, то ковариация будет мала, несмотря на наличие зависимости между X и Y. В качестве числовой характеристики зависимости (а не рассеивания) случайных величин X и Y используют безразмерную характеристику – коэффициент корреляции.

Коэффициентом корреляции случайных величин X и Y называют их нормированную ковариацию:

$$r(X,Y)=\frac{\operatorname{cov}(X,Y)}{\sigma(X)\cdot\sigma(Y)}=\frac{\operatorname{cov}(X,Y)}{\sqrt{D(X)}\cdot\sqrt{D(Y)}}. \tag{48}$$

Теорема. Если случайные величины X и Y связаны линейной зависимостью Y=aX+b, то коэффициент корреляции:

$$r(X,Y)=\begin{cases}+1, & \text{если} \quad a>0,\\ -1, & \text{если} \quad a<0.\end{cases}$$

(49)

При доказательстве этой теоремы используются свойства математического ожидания и дисперсии случайных величин:

$$m_y=M(aX+b)=M(aX)+M(b)=aM(X)+b=am_x+b;$$

$$\operatorname{cov}(X,Y)=M[(X-m_x)\cdot(Y-m_y)]=M\left[(X-m_X)\cdot\left((aX+b)-(am_X+b)\right)\right]=$$

$$=M[(X-m_x)\cdot(aX+b-am_x-b)]=M\left[(X-m_X)\cdot(aX-am_X)\right]=$$

$$=M\left[(X-m_X)\cdot a(X-m_X)\right]=aM[(X-m_x)^2]=aD(X);$$

$$D(Y)=D(aX+b)=D(aX)+D(b)=a^2D(X)+0=a^2D(X);$$

$$r(X,Y)\overset{def}{=}\frac{\operatorname{cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}=\frac{aD(X)}{\sqrt{D(X)}\sqrt{a^2D(X)}}=\frac{aD(X)}{\sqrt{a^2}\sqrt{D^2(X)}}=$$

$$\underset{D(X)>0}{\underset{\uparrow}{=}}\frac{aD(X)}{|a|\cdot D(X)}=\frac{a}{|a|}=\begin{cases}+1, & \text{если} \quad a>0;\\ -1, & \text{если} \quad a<0.\end{cases}$$

Можно показать, что обратное утверждение также верно.

Возникает вопрос: в каких пределах находится значение коэффициента корреляции? Ответ на него даёт следующее свойство коэффициента корреляции: ***величина коэффициента корреляции заключена в пределах [-1,1]***.

Если коэффициент корреляции r(X,Y)>0, то говорят о ***положительной корреляции*** случайных величин X и Y; если же r(X,Y)<0, то об их ***отрицательной корреляции.***

Положительная корреляция означает, что при возрастании одной из случайных величин другая имеет тенденцию в среднем возрастать. Например, вес и рост человека связаны положительной корреляцией.

Отрицательная корреляция означает, что при возрастании одной из случайных величин другая имеет тенденцию в среднем убывать. Например, время, потраченное на регулировку прибора, и количество неисправностей, обнаруженных при работе прибора, связаны отрицательной корреляцией.

В качестве характеристики зависимости системы n случайных величин (n-мерного случайного вектора) используют ковариационную матрицу.

Ковариационная матрица случайного вектора (X_1, X_2, … , X_n) – это матрица, состоящая из элементов

$$k_{ij} = \mathrm{cov}(X_i, X_j) = M[(X_i - m_{X_i})(X_j - m_{X_j})],\ i, j = \overline{1,n}. \qquad (50)$$

Очевидно, что $k_{ij} = k_{ji}$, то есть ковариационная матрица симметрична.

По главной диагонали ковариационной матрицы стоят дисперсии случайных величин X_1, X_2, …X_n. При этом, если случайные величины X_1, X_2, …X_n некоррелированы, то ковариационная матрица имеет диагональный вид:

$$\begin{pmatrix} D(X_1) & 0 & \dots & 0 \\ 0 & D(X_2) & \dots & 0 \\ \dots & \dots & \dots & 0 \\ 0 & 0 & \dots & D(X_n) \end{pmatrix}$$

(51)

3. Нормальное распределение системы случайных величин

3.1. Двумерное нормальное распределение

В теории вероятностей и её приложениях большую роль играет двумерное нормальное распределение. Плотность двумерной нормальной случайной величины (X,Y) имеет вид

$$f(x,y)=\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-r^2}}\exp\left\{-\frac{1}{2(1-r^2)}\left[\frac{(x-m_X)^2}{\sigma^2{}_X}-\frac{2r(x-m_X)(y-m_Y)}{\sigma_X\cdot\sigma_Y}+\frac{(y-m_Y)^2}{\sigma^2{}_Y}\right]\right\} \qquad (52)$$

Здесь m_X, m_Y - математические ожидания величин X и Y; σ_X, σ_Y - средние квадратичные отклонения величин X и Y; r – коэффициент корреляции величин X и Y.

Предположим, что случайные величины X и Y не коррелированы, то есть r=0. Тогда имеем:

$$f(x,y)=\frac{1}{2\pi\sigma_X\sigma_Y}\exp\left\{-\frac{(x-m_X)^2}{2\sigma^2{}_X}-\frac{(y-m_Y)^2}{2\sigma^2{}_Y}\right\}=$$

$$=\frac{1}{\sigma_X\sqrt{2\pi}}\exp\left\{-\frac{(x-m_X)^2}{2\sigma^2{}_X}\right\}\cdot\frac{1}{\sigma_Y\sqrt{2\pi}}\exp\left\{-\frac{(y-m_Y)^2}{2\sigma^2{}_Y}\right\}=f_1(x)\cdot f_2(y)$$

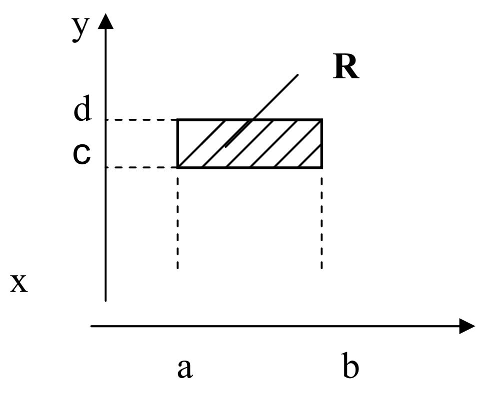
(53)

Получили, что плотность распределения системы двух случайных величин (X,Y) равна произведению плотностей распределения компонент X и Y, а это значит, что X и Y – независимые случайные величины.

Таким образом, доказана следующая **теорема**: ***из некоррелированности нормально распределенных случайных величин следует их независимость***. Поскольку из независимости любых случайных величин следует их некоррелированность, то можно сделать вывод, что термины «некоррелированные» и «независимые» величины для случая нормального распределения эквивалентны.

Приведём формулы для вероятности попадания нормально распределённой двумерной случайной величины в различные области на плоскости.

Пусть случайный вектор (X,Y), компоненты которого независимы, распределён по нормальному закону (53). Тогда вероятность попадания случайной точки (X,Y) в прямоугольник **R,** стороны которого параллельны координатным осям, равна



$$p\{(X,Y)\in R\}=\iint_R f(x,y)dxdy=\int_a^b\int_c^d f(x,y)dydx=$$

$$=\int_a^b\frac{1}{\sigma_X\sqrt{2\pi}}\exp\left\{-\frac{(x-m_X)^2}{2\sigma_X^2}\right\}dx\cdot$$

(54)

$$\cdot\int_c^d\frac{1}{\sigma_Y\sqrt{2\pi}}\exp\left\{-\frac{(y-m_Y)^2}{2\sigma_Y^2}\right\}dy=$$

$$=\left[\Phi\left(\frac{b-m_X}{\sigma_X}\right)-\Phi\left(\frac{a-m_X}{\sigma_X}\right)\right]\cdot\left[\Phi\left(\frac{d-m_Y}{\sigma_Y}\right)-\Phi\left(\frac{c-m_Y}{\sigma_Y}\right)\right],$$

где $\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_0^x \exp\left(-\frac{z^2}{2}\right)dz$ - функция Лапласа. Эта функция табулирована.

Пусть плотность распределения нормального закона системы случайных величин (X,Y) задана в виде (52). Ясно, что данная плотность сохраняет постоянное значение на эллипсах:

$$\frac{(x-m_X)^2}{\sigma^2{}_X}-\frac{2r(x-m_X)(y-m_Y)}{\sigma_X\cdot\sigma_Y}+\frac{(y-m_Y)^2}{\sigma^2{}_Y}=C^2\ , \qquad (55)$$

где C – постоянная; на этом основании такие эллипсы носят название **эллипсов равных вероятностей**. Можно показать, что вероятность попадания точки (X,Y) внутрь эллипса равной вероятности равна

$$p(C)=1-\exp\left[-\frac{C^2}{2(1-r^2)}\right]$$

(56)

Пример 10. Случайные величины X и Y независимы и нормально распределены с $m_X=m_Y=0; D(X)=D(Y)=1$. Найти вероятность того, что случайная точка (X,Y) попадет в кольцо $k=\{(x,y):4\le x^2+y^2\le 9\}$.

Решение: Так как случайные величины X и Y независимы, то они не коррелированы и, следовательно, r = 0. Подставляя $m_x=m_y=0, \sigma_x=\sigma_y=1, r=0$ в (С), получаем

$$x^2+y^2=C^2,$$

то есть эллипс равной вероятности выродился в круг равной вероятности. Тогда

$$P\{(X,Y)\in k\}=P(3)-P(2)=\left(1-\exp\left(-\frac{9}{2}\right)\right)-\left(1-\exp\left(-\frac{4}{2}\right)\right)=e^{-2}-e^{-4,5}\approx 0.1242.$$

Ответ: 0,1242.

3.2. Общий случай n-мерного нормального распределения

Плотность нормального распределения системы ***n*** случайных величин имеет вид:

$$f(x_1,x_2,x_3,\ \dots\ ,x_n)=\frac{\sqrt{|c|}}{(2\pi)^{n/2}}\exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}c_{ij}(x_i-m_{X_i})(x_j-m_{X_j})\right\}, \qquad (57)$$

где $|C|$ - определитель матрицы C - обратной к ковариационной матрице; m_{X_i} - математическое ожидание случайной величины X_i - i-той компоненты ***n***-мерного нормального случайного вектора.

Из общего выражения вытекают все формы нормального закона для любого числа измерений и для любых видов зависимости между случайными величинами. В частности, при ***n*** = 2 ковариационная матрица имеет вид:

$$k = \begin{Vmatrix} \sigma_X^2 & \sigma_X \sigma_Y r \\ \sigma_X \sigma_Y r & \sigma_Y^2 \end{Vmatrix},$$

(58)

её определитель $|k| = \sigma_X^2 \cdot \sigma_Y^2 \cdot (1-r^2)$; матрица С, обратная к ковариационной матрице, имеет вид

$$C = \begin{Vmatrix} \dfrac{1}{\sigma_X^2 \cdot (1-r^2)} & -\dfrac{r}{\sigma_X \cdot \sigma_Y \cdot (1-r^2)} \\ -\dfrac{r}{\sigma_X \cdot \sigma_Y \cdot (1-r^2)} & \dfrac{1}{\sigma_Y^2 \cdot (1-r^2)} \end{Vmatrix}.$$

(59)

Подставляя $|C|$ и элементы матрицы С в общую формулу (57), получаем формулу для нормального распределения на плоскости (52).

Если случайные величины $X_1, X_2, \ldots, X_n$ независимы, то плотность распределения системы $(X_1,\ X_2, \ldots,\ X_n)$ равна

$$f(x_1, x_2, \ \ldots\ ,\ x_n) = \frac{1}{(2\pi)^{n/2} \sigma_1 \sigma_2 \ldots \sigma_n} \exp\left\{ -\frac{1}{2} \sum_{i=1}^{n} \frac{\left(x_i - m_{x_i}\right)^2}{\sigma_i^2} \right\}. \qquad (60)$$

При n = 2 эта формула принимает вид (53).

3.2. Функции от нормально распределенных случайных величин. Распределения хи-квадрат, Стьюдента, Фишера- Снедекора

Рассмотрим общий случай: линейную функцию от нормально распределенных аргументов. Пусть дан n-мерный нормально распределенный случайный вектор $(X_1,\ X_2, \ \ldots\ ,\ X_n)$, случайная величина Y представляет собой линейную функцию от этих величин:

$$Y = \sum_{i=1}^{n} a_i X_i + b.$$

(61)

Можно показать, что случайная величина Y также распределена нормально с параметрами

$$m_Y = \sum_{i=1}^{n} a_i m_i + b; \qquad (62)$$

$$\sigma_Y^2 = \sum_{i=1}^{n} a_i^2 \sigma_i^2 + 2\sum_{i<j} a_i \cdot a_j \cdot r_{ij} \cdot \sigma_i \cdot \sigma_j, \qquad (63)$$

где m_i – математическое ожидание случайной величины X_i, σ_i^2 - дисперсия случайной величины X_i, r_{ij} - коэффициент корреляции между X_i и X_j.

Пример 11. Записать плотность распределения случайной величины $Y = X_1 + 2X_2 + 3$, если случайные величины X_1 и X_2 имеют нормальное распределение с параметрами $m_1 = 0$, $m_2 = 2$, $\sigma_1 = \sigma_2 = 1$, их коэффициент корреляции $r_{12} = 1$.

Решение. По условию задачи имеем: n=2; $a_1 = 1$; $a_2 = 2$; $b = 3$. Используя формулу (62), получаем: $m_Y = a_1 m_1 + a_2 m_2 + b = 1\cdot 0 + 2\cdot 2 + 3 = 7$. Используя формулу (63), получаем: $\sigma_Y^2 = a_1^2 \cdot \sigma_1^2 + a_2^2 \cdot \sigma_2^2 + 2a_1 a_2 r_{12}\sigma_1\sigma_2 = 1^2 \cdot 1^2 + 2^2 \cdot 1^2 + 2\cdot 1\cdot 2\cdot 1\cdot 1\cdot 1 = 9$.

Тогда искомая функция распределения случайной величины Y имеет вид:

$$f(Y) = \frac{1}{\sigma_Y\sqrt{2\pi}} e^{-\frac{(y-m_Y)^2}{2\sigma_Y^2}} = \frac{1}{3\sqrt{2\pi}} e^{-\frac{(y-7)^2}{18}}.$$

Пусть $(X_1,\ X_2, \ldots ,\ X_n)$- независимые случайные величины, подчиняющиеся нормальному распределению с нулевым математическим ожиданием и единичной дисперсией, то есть стандартному нормальному распределению. Распределение случайной величины, являющейся суммой квадратов этих величин

$$\chi_n^2 = \sum_{i=1}^{n} X_i^2 .$$

(64)

называется “***распределением ХИ - квадрат с n степенями свободы***”.

Плотность распределения ХИ – квадрат с n=2 степенями свободы равна

$$f(x) = \begin{cases} \frac{1}{2} e^{-\frac{x}{2}},\ \textit{при}\ x > 0; \\ 0,\ \textit{при}\ x \le 0. \end{cases}$$

(65)

Плотность ХИ – квадрат распределения с n степенями свободы имеет вид:

$$f(x)=\begin{cases}\dfrac{x^{\frac{n}{2}-1}\cdot e^{-\frac{x}{2}}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}\ ,\ при\ x>0;\\ 0\ ,\ при\ x\le 0,\end{cases}$$

(66)

где $\Gamma(\lambda)=\int_0^{\infty}t^{\lambda-1}e^{-t}dt$ - гамма-функция Эйлера. С возрастанием числа степеней свободы распределение χ^2 приближается к нормальному закону распределения (при ***n***>30 распределение χ^2 практически не отличается от нормального). Математическое ожидание χ^2 - распределения с n степенями свободы равно ***n***, а дисперсия равна 2***n***.

Распределение Стьюдента с n степенями свободы St(n) определяется как распределение случайной величины

$$T_n=\frac{Z}{\sqrt{\frac{1}{n}\chi_n^2}},$$

(67)

где Z – стандартная нормальная величина, независимая от χ_n^2 распределения.

Плотность распределения Стьюдента с n степенями свободы имеет вид:

$$f(x)=\frac{1}{\sqrt{n\pi}}\cdot\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\cdot\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}},$$

(68)

Математическое ожидание при $n\ge 2$равно 0, дисперсия при $n>2$ равна $\frac{n}{n-2}$. При $n\to\infty$ распределение Стьюдента приближается к нормальному (уже при ***n***>30 почти совпадает с нормальным распределением).

Распределением Фишера-Снедекора (или F-распределением) с n_1 и n_2 степенями свободы называется распределение случайной величины

$$F(n_1, n_2) = \frac{\chi^2_{n_1}}{n_1} : \frac{\chi^2_{n_2}}{n_2},$$

(69)

где $\chi^2_{n_1}$ и $\chi^2_{n_2}$ - случайные величины, имеющие χ^2 - распределение с n_1 и n_2 степенями свободы, соответственно.

Плотность распределения Фишера-Снедекора при $x > 0$ имеет вид:

$$f(x) = \frac{\Gamma\left(\frac{n_1 + n_2}{2}\right) \cdot x^{\frac{n_1}{2} - 1}}{\Gamma\left(\frac{n_1}{2}\right) \cdot \Gamma\left(\frac{n_2}{2}\right) \cdot (1 + x)^{\frac{n_1 + n_2}{2}}}.$$

(70)

Математическое ожидание при $n > 2$ равно $\frac{n}{n-2}$, дисперсия при $n > 4$ равна $\frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$. При $n \to \infty$ F - распределение стремится к нормальному закону.

Распределения ХИ - квадрат, Стьюдента, Фишера - Снедекора используются в математической статистике.

Список литературы

1. Гмурман В.Е. / Теория вероятностей и математическая статистика. – М.: «Высшая школа”, 1977.
2. Гмурман В.Е. / Руководство к решению задач по теории вероятностей и математической статистике. – М.: ”Высшая школа”, 1975.
3. Вентцель Е.С. / Теория вероятностей. – М.: Гос.изд-во физ.-мат. лит-ры, 1962.
4. Письменный Д.Т. Конспект лекций по теории вероятностей и математической статистике. – М.: Айрис-пресс, 2004.

Содержание

Составитель Бобкова Вера Александровна

Системы случайных величин

Методические указания для самостоятельной работы студентов

Редактор Г.В.Куликова

Подписано в печать 02.03.2010. Формат 60x84 $1/16$. Бумага писчая. Усл.печ.л.1,63.

Уч.-изд.л.1,81. Тираж 50 экз.

ГОУ ВПО Ивановский государственный химико-технологический университет

Отпечатано на полиграфическом оборудовании кафедры экономики и финансов ГОУ ВПО «ИГХТУ»

153000, г.Иваново, пр. Ф.Энгельса, 7