Теория вероятностей и математическая статистика

«Многомерные распределения вероятностей»

Тюрнева Т.Г.,

доцент ИМЭИ ИГУ

Определения

- Упорядоченный набор (X_1, X_2, …X_n) случайных величин X_i ($i=\overline{1,n}$) ,заданных на одном и том же ПЭС, называется ***n – мерной случайной величиной* или *системой n случайных величин.***
- Одномерные случайные величины $\mathbf{X_1, X_2, \ldots X_n}$ называются
- ***компонентами* или *составляющими***

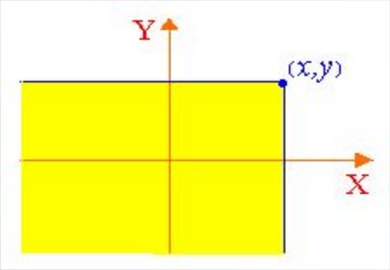
  n – мерной случайной величины (X_1, X_2 , …X_n) .
- Упорядоченная пара (X,Y) двух случайных величин X и Y называется ***двумерной случайной*** величиной или ***системой двух одномерных*** случайных величин X и Y.

Общий план исследования двумерного распределения вероятностей

1. Составить закон распределения вероятностей (X,Y).
2. Найти законы распределения и числовые характеристики случайных величин X и Y.
3. Установить зависимы или независимы с.в. X и Y.
4. Составить ковариационную и корреляционную матрицы.
5. Описать регрессии величины X на Y и величины Y на X.
6. Построить наилучшие в среднем квадратическом оценки величины X по Y и величины Y по X.
7. Проверить формулу полного математического ожидания.

Функция распределения. Свойства функции распределения

Функцией распределения F(*x*, *y*) *двумерной случайной величины* (X, Y) *называется вероятность совместного выполнения двух неравенств:* $X<x,\ Y<y$: $F(x, y) = P(X<x, Y<y)$.



1. Функция распределения – величина неотрицательная, не превышающая единицы:

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

2. Функция распределения есть неубывающая функция по каждому аргументу.

3. Если хотя бы один из аргументов стремится к – ∞ *, то функция распределения стремится к нулю:*

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

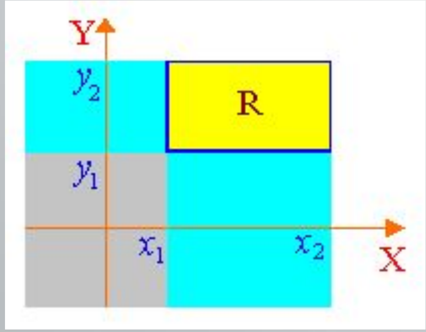
Функция распределения. Свойства функции распределения

4. Если оба аргумента стремятся к + ∞, то функция распределения стремится к единице: $F(+\infty,+\infty)=1$

5. Если один из аргументов обращается в +∞, то функция распределения F*(x,y) становится функцией распределения, соответствующей другому аргументу:* $F(x,+\infty)=F_1(x)=P(X<x)$ $F(+\infty,y)=F_2(y)=P(Y<y)$

6. Вероятность попадания случайной величины {X, Y} в прямоугольник со сторонами R=($x_1 < X < x_2$, $y_1 < Y < y_2$):

$$P(\{X,Y\}\in R)=F(x_2,y_2)-F(x_2,y_1)-F(x_1,y_2)+F(x_1,y_1)$$



- *Двумерная случайная величина* (X,Y) *называется* ***дискретной***, *если каждая из случайных величин* X *и* Y *является дискретной.*

- Если дискретная величина X может принимать только значения $x_1, \ldots, x_n$, а случайная величина Y – значения $y_1, \ldots, y_m$, то двумерный случайный вектор (X,Y) может принимать только пары значений (x_i, y_j).

- Закон распределения системы двух ДСВ описывается ***матрицей распределения***, т.е. прямоугольной таблицей, в которой записаны все вероятности p_{ij} того, что двумерная ДСВ примет значение (x_i, y_j).

$$\sum \sum p_{ij} = 1.$$

- Зная матрицу распределения системы двух ДСВ легко найти законы распределения отдельных случайных величин, входящих в систему:

$$p_i = \mathrm{P}(X = x_i) = \sum_{j=1}^{m} p_{ij}, \quad q_j = \mathrm{P}(Y = y_j) = \sum_{i=1}^{n} p_{ij}$$

- *Двумерная случайная величина* {X,Y} *называется* ***непрерывной***, *если каждая из случайных величин* X *и* Y *является непрерывной.* Система двух НСВ обычно описывается ***плотностью распределения***:

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

Свойства плотности распределения f(x,y): $f(x,y) \geq 0,$

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

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

$$F_1(x) = F(x, +\infty) = \int_{-\infty}^{x} dx \int_{-\infty}^{+\infty} f(x, y)\,dy$$

$$F_2(y) = F(+\infty, y) = \int_{-\infty}^{y} dy \int_{-\infty}^{+\infty} f(x, y)\,dx$$

$$f_1(x) = \int_{-\infty}^{+\infty} f(x, y)\,dy, \qquad f_2(y) = \int_{-\infty}^{+\infty} f(x, y)\,dx$$

Дискретные двумерные распределения вероятностей

Задача

Дважды бросается игральная кость.

Случайные величины:

X – число появлений шестерки,

Y – число появлений нечетной цифры.

1. Составить закон распределения вероятностей (X,Y):
   таблица распределения;
   функция распределения.

2. Найти законы распределения и числовые характеристики случайных величин X и Y.

y_j \ x_i	**0**	**1**	**2**	**P(Y= y_j)**
0				
1			0	
2		0	0	
P(X= x_i)				**1**

Установить зависимы или независимы с.в. X и Y

- События A и B называются независимыми, если

$$P(AB) = P(A)\cdot P(B).$$

- Случайные величины X и Y называются **независимыми**, если закон распределения каждой из них не зависит от того, какое значение приняла другая.

- 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

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

Установить зависимы или независимы с.в. X и Y

Пусть X и Y

дискретные случайные величины

- Случайные величины X и Y независимы тогда и только тогда, когда для любых значений x_i и y_j выполнено

$$P(X = x_i, Y = y_j) = P(X = x_i) \cdot P(Y = y_j).$$

Пусть X и Y

непрерывные случайные величины

- Случайные величины X и Y независимы тогда и только тогда, когда

$$f_{XY}(x,y) = f_X(x) \cdot f_Y(y).$$

Ковариация и коэффициент корреляции

- Ковариацией (смешанный второй центральный момент, корреляционный момент) случайных величин X и Y называют число
- $\mathrm{cov}(X,Y) = M((X-MX)\cdot(Y-MY))$
- $\mathrm{cov}(X,Y) = MXY-MX\cdot MY.$

$$MXY=MX\cdot MY+\mathrm{cov}(X,Y)$$

$$D(X+Y) = DX + DY + 2\mathrm{cov}(X,Y)$$

Свойства ковариации

1. Ковариация не меняется при перестановке случайных величин:

$\operatorname{cov}(X,Y) = \operatorname{cov}(Y,X)$.

2. Если C =const, то $\operatorname{cov}(X,C) = 0$.

3. Если случайные величины X и Y независимы, то

$\operatorname{cov}(X,Y) = 0$.

- 4. $\operatorname{cov}(X,X) = DX, \quad \operatorname{cov}(Y,Y) = DY$.

5. Ковариация линейна по каждому из своих аргументов:

$$\operatorname{cov}(C_1X_1 + C_2X_2, Y) = C_1\operatorname{cov}(X_1,Y) + C_2\operatorname{cov}(X_2,Y),$$

где C_1, C_2 =const.

6. $|\operatorname{cov}(X,Y)| \leq \sqrt{DX \cdot DY}$

Коэффициент корреляции

- Коэффициентом корреляции величин X и Y называют отношение ковариации к произведению средних квадратических отклонений этих величин

$$\rho_{XY} = \frac{cov(X,Y)}{\sigma_X \sigma_Y}$$

- Коэффициент корреляции—безразмерная величина, причем $|\rho_{XY}| \leq 1$.
- Коэффициент корреляции служит для оценки тесноты **линейной связи**
- между X и Y:

чем ближе абсолютная величина коэффициента корреляции к единице, тем связь сильнее;

чем ближе абсолютная величина коэффициента корреляции к нулю, тем связь слабее.

- Коэффициент корреляции равен 1 тогда и только тогда, когда случайные величины линейно связаны.

Коэффициент корреляции

- Если коэффициент корреляции равен нулю, то величины называют **некоррелированными**.
- Из независимости двух величин следует их некоррелированность, но из некоррелированности еще нельзя сделать вывод о независимости этих величин.
- Для некоторых распределений понятия независимости и некоррелированности являются эквивалентными.
- В частности, если случайные величины X и Y имеют **нормальное распределение** и $\rho_{XY} = 0$, то они независимы.

Ковариационная и корреляционная матрицы

- $\begin{pmatrix} DX & cov(X,Y) \\ cov(X,Y) & DY \end{pmatrix}$
-
- $\begin{pmatrix} 1 & \rho_{XY} \\ \rho_{XY} & 1 \end{pmatrix}$

$$\rho_{XY} = \frac{cov(X,Y)}{\sigma_X \sigma_Y}$$

Регрессии величины X на Y и величины Y на X

- Условное математическое ожидание случайной величины Y при условии, что X приняла одно из своих возможных значений;
- Функция регрессии величины Y на X;
- Условное математическое ожидание случайной величины Y при условии X - случайная величина;
- Наилучшая в среднем квадратическом оценка величины Y по величине X;
- Формула полного математического ожидания.

Условное распределение

- Условная вероятность: P(A/B) = $\frac{P(AB)}{P(A)}$.
- Условное распределение: с.в. Y

$$P(Y = y_j/X = x_i) = \frac{P(Y=y_j, X=x_i)}{P(X=x_i)}.$$

-
- Условное математическое ожидание случайной величины Y при условии, что X приняла одно из своих возможных значений x_i, называется действительное число $\widetilde{x}_i = M(Y/X = x_i) = \sum y_j\, P(Y = y_j/X = x_i)$,

$i = 0,1,2.$

Функция регрессии

- Условным математическим ожиданием случайной величины Y при условии X называется **случайная величина** $\tilde{X}$, обозначаемая также M(Y/X), возможные значения которой $\tilde{x}_i = M(Y/X = x_i)$, а соответствующие вероятности равны $P(\tilde{X} = \tilde{x}_i) = P(X = x_i), i = 0,1,2.$
- *Функция* $M(Y/X = x) = \psi(x)$, *заданная на множестве значений случайной величины X называется регрессией величины Y на X.*

!!! Характеризует изменение среднего значения с.в. Y при изменении значений с.в. X.

Наилучшая в среднем квадратическом оценка величины Y по величине X

Корреляционная зависимость

Линейная

нелинейная

Y

$$g(X) = aX + b$$

$g(X)$ – среднеквадратическая регрессия Y на X

$$g(X) = m_y + \rho\frac{\sigma_y}{\sigma_x}(X\text{-}m_x)$$

Формула полного математического ожидания

$$\tilde{x}_i = M(Y/X = x_i)$$

$$P(\tilde{X} = \tilde{x}_i) = P(X = x_i),\ i = 0,1,2$$

M(M(Y/X) = M(Y)

случайная величина

Непрерывные двумерные распределения вероятностей

Задача

- Двумерная случайная величина (X,Y) задана плотностью
- совместного распределения на D
- $f(x, y) = \begin{cases} Cxy, & \text{при } (x, y) \in D \\ 0, & \text{при } (x, y) \notin D \end{cases}$ $D\colon\ y < -x, y < 2, x < 0$

Двумерное нормальное распределение вероятностей

$$\phi(x,y) = \frac{1}{2\pi\,\sigma_x\,\sigma_y\sqrt{1-r^2}}\,e^{-\frac{1}{2(1-r^2)}\left[\frac{(x-m_x)^2}{\sigma_x^2}-\frac{2r(x-m_x)(y-m_y)}{\sigma_x\,\sigma_y}+\frac{(y-m_y)^2}{\sigma_y^2}\right]}$$

r - коэффициент корреляции случайных величин X и Y

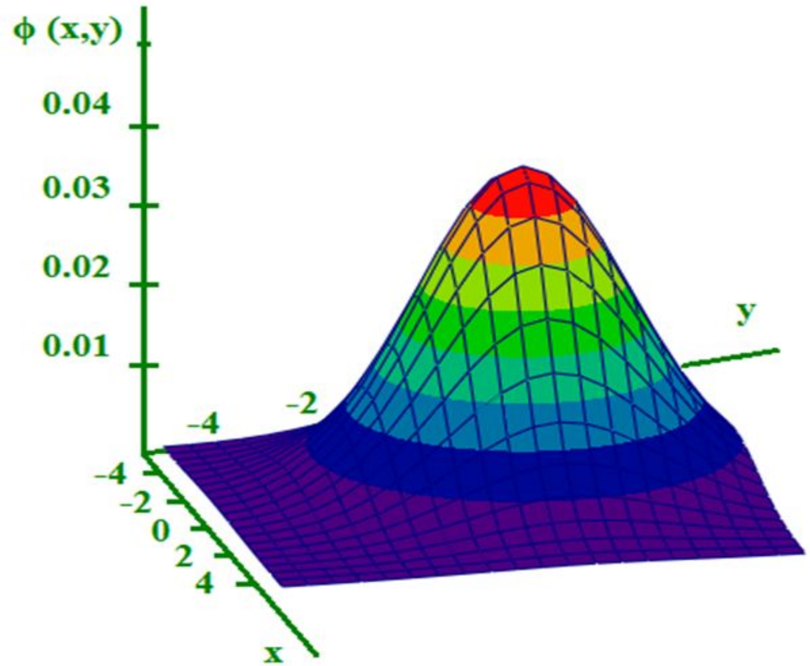
σx - среднее квадратическое отклонение случайной величины X

σy - среднее квадратическое отклонение случайной величины Y

mx- математическое ожидание случайной величины X

my - математическое ожидание случайной величины Y

$$\phi(x,y) = \frac{1}{2\pi\,\sigma_x\,\sigma_y}\, e^{-\frac{(x - m_x)^2}{2\sigma_x^2} - \frac{(y - m_y)^2}{2\sigma_y^2}}$$



$r = 0$

$\sigma x = 2$

$\sigma y = 2$

$mx = -1$

$my = 1$

Домашняя контрольная работа

<u>Тема</u>: ***Многомерные случайные величины.***

1. Монету бросают три раза, отмечая результат каждого бросания знаком плюс или знаком минус в зависимости от того, что выпало – герб или решка соответственно. Пусть X – число выпавших гербов, а Y – число перемен знака в образовавшейся последовательности плюсов и минусов.
    1. Записать з.р. случайного вектора (X, Y).
    2. Записать з.р. X и з.р. Y. Найти их числовые характеристики.
    3. Определить зависимые или независимые величины X и Y.
    4. Составить ковариационную и корреляционную матрицы случайного вектора (X, Y).
    5. Найти условные распределения и записать функции регрессии.

2. Для случайного вектора (X, Y) с ковариационной матрицей

$$\begin{pmatrix} 2 & -3 \\ -3 & 6 \end{pmatrix}$$

   Вычислите:
   D (X+Y); D (X-Y); D (2X+3Y -2).

3. Заданы следующие характеристики двумерного НОРМАЛЬНОГО вектора: MX=-2, MY=3 и ковариационная матрица

$$\begin{pmatrix} 16 & 12 \\ 12 & 25 \end{pmatrix}$$

   Записать выражение для плотности распределения вероятностей f(x,y).
   Написать уравнение регрессии Y на X.