

第四章 随机变量的数字特征

- 数学期望
- 方差
- 协方差和相关系数
- 矩、协方差矩阵

4.1数学期望

一.数学期望的定义

数学期望——描述随机变量取值的平均特征

一般意义下变量 x 平均数的计算:

x	x_1	x_2	\dots	x_k	\dots
n_k	n_1	n_2	\dots	n_k	\dots

如何计算变量 x 的平均数?

数学期望定义

解： 变量 x 的平均数为其取值总数/总次数：

$$\frac{x_1 \times n_1 + x_2 \times n_2 + \cdots + x_k \times n_k + \cdots}{n_1 + n_2 + \cdots + n_k + \cdots}$$

即：

$$x_1 \times f_1 + x_2 \times f_2 + \cdots + x_k \times f_k + \cdots$$

于是，离散随机变量的平均数类似的定义：

$$x_1 \times p_1 + x_2 \times p_2 + \cdots + x_k \times p_k + \cdots$$


离散型随机变量的期望

定义 1. 若 $X \sim P\{X=x_k\}=p_k, k=1,2,\dots,n, (\dots)$

且
$$\sum_k |x_k p_k| < \infty$$

则称
$$E(X) = \sum_k x_k p_k$$

为r.v.X的数学期望，简称期望或均值。



例1 掷一颗均匀的骰子，以 X 表示掷得的点数，求 X 的数学期望。

$$E(X) = \sum_{k=1}^6 k \frac{1}{6} = \frac{7}{2}$$

连续型随机变量的期望

定义 2 若 $X \sim f(x)$, $-\infty < X < \infty$, $\int_{-\infty}^{\infty} |xf(x)| dx < \infty$

则称 $E(X) = \int_{-\infty}^{\infty} xf(x) dx$

为 X 的数学期望。

例2. 若随机变量X服从拉普拉斯分布，其密度函数为

$$f(x) = \frac{1}{2\lambda} \exp\left\{-\frac{|x-\mu|}{\lambda}\right\} \quad \text{试求} E(X).$$

解：

$$E(X) = \int_{-\infty}^{\infty} \frac{x}{2\lambda} \exp\left\{-\frac{|x-\mu|}{\lambda}\right\} dx$$

$$\stackrel{\text{令 } t = \frac{x-\mu}{\lambda}}{=} \int_{-\infty}^{\infty} \frac{\lambda t + \mu}{2\lambda} \exp\{-|t|\} \lambda dt$$

$$= \int_{-\infty}^{\infty} 0.5\lambda t \exp\{-|t|\} dt + \int_{-\infty}^{\infty} 0.5\mu \exp\{-|t|\} dt$$

$$= \int_0^{\infty} \mu \exp\{-t\} dt = \mu$$

二.几个重要r.v.的期望

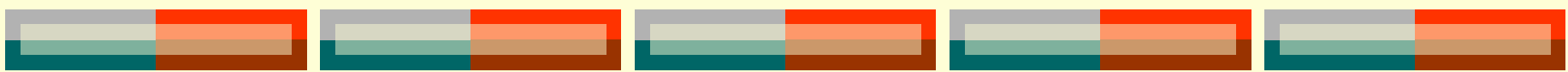
1. 0-1分布的数学期望

$$\begin{array}{ccc} X & 1 & 0 \\ p_k & p & 1-p \end{array} \Rightarrow E(X) = p$$

2. 二项分布 $B(n, p)$

$$P\{X = k\} = C_n^k p^k (1-p)^{n-k} \quad k = 0, 1, \dots, n$$

$$E(X) = np$$



$$E(X) = \sum_{k=0}^n k \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$

泊松分布

$$= \sum_{k=1}^n \frac{n!}{(k-1)!(n-k)!} p^k (1-p)^{n-k}$$

$$(a+b)^n = \sum_{k=0}^n C_n^k a^k b^{n-k}$$

$$= np \sum_{k=1}^n \frac{(n-1)!}{(k-1)!(n-k)!} p^{k-1} (1-p)^{n-1-(k-1)}$$

$$\underline{\underline{\text{令 } l = k - 1}} \quad np \sum_{l=0}^{n-1} C_{n-1}^l p^l (1-p)^{n-1-l} = np$$



3. 泊松分布

$$X \sim P\{X = k\} = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \dots$$

$$e^{\lambda} = \sum_{l=0}^{\infty} \frac{\lambda^l}{l!}$$

$$E(X) = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda$$

4. 均匀分布U(a, b)

$$X \sim f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{其他} \end{cases}$$

$$E(X) = \int_a^b \frac{x}{b-a} dx = \frac{a+b}{2}$$

5. 指数分布

分部积分法

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & x \leq 0 \end{cases} \quad \int_a^b u \, dv = uv \Big|_a^b - \int_a^b v \, du$$

$$E(X) = \int_0^{\infty} x \lambda e^{-\lambda x} dx = - \int_0^{\infty} x de^{-\lambda x}$$

$$= -xe^{-\lambda x} \Big|_0^{\infty} + \int_0^{\infty} e^{-\lambda x} dx = \frac{1}{\lambda}$$

6. 正态分布 $N(\mu, \sigma^2)$

$$X \sim f(x) = \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty$$

$$E(X) = \int_{-\infty}^{\infty} \frac{x}{\sqrt{2\pi} \sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

$$\underline{\underline{\text{令 } t = \frac{x - \mu}{\sigma} \int_{-\infty}^{\infty} \frac{\sigma t + \mu}{\sqrt{2\pi} \sigma} e^{-\frac{t^2}{2}} \sigma dt}}$$

$$= \int_{-\infty}^{\infty} \frac{\sigma t}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt + \mu \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \mu$$

三. 随机变量函数的期望

定理1 若 $X \sim P\{X=x_k\}=p_k, k=1,2,\dots$, 则 $g(X)$ 的期望 $E(g(X))$ 为

$$E[g(X)] = \sum_{k=1}^{\infty} g(x_k) p_k$$

推论: 若 $(X, Y) \sim P\{X=x_i, Y=y_j\}=p_{ij}$,

$i, j=1, 2, \dots$, 则 $g(X, Y)$ 的期望

$$E[g(X, Y)] = \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} g(x_i, y_j) p_{ij}$$

例1 设随机变量 (X, Y) 的分布律如下，求 $E(XY)$

$\begin{array}{c} Y \\ \diagdown \\ X \end{array}$	1	2
0	0.15	0.15
1	0.45	0.25

$$\begin{aligned} \text{解: } E(XY) &= 0 \times 1 \times 0.15 + 0 \times 2 \times 0.15 \\ &\quad + 1 \times 1 \times 0.45 + 1 \times 2 \times 0.25 \\ &= 0.95 \end{aligned}$$

连续型随机变量函数的期望

定理2 若 $X \sim f(x)$, $-\infty < X < \infty$, 则 $g(X)$ 的期望

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x) dx$$

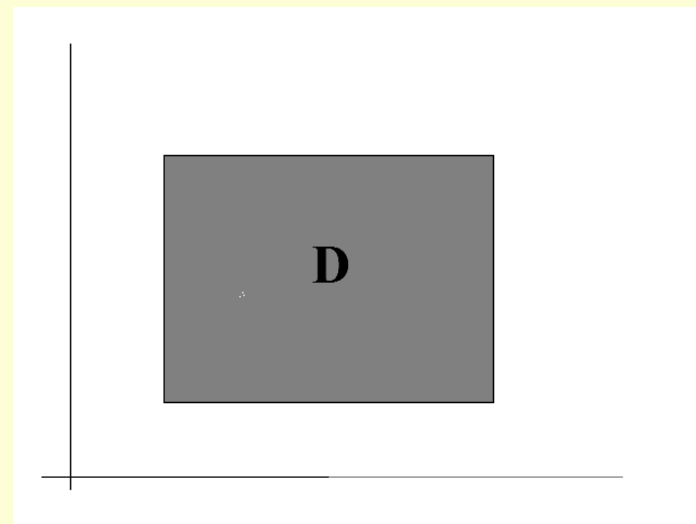
推论 若 $(X, Y) \sim f(x, y)$, $-\infty < X < \infty$, $-\infty < Y < \infty$, 则 $g(X, Y)$ 的期望

$$\begin{aligned} E[g(X, Y)] \\ = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y) dx dy \end{aligned}$$

二维连续型随机变量函数的期望

- 若积分区域为

$$(1) \mathbf{D} = \{(\mathbf{x}, \mathbf{y}) \mid \mathbf{a} \leq \mathbf{x} \leq \mathbf{b}, \mathbf{c} \leq \mathbf{y} \leq \mathbf{d}\}$$



则

$$\iint_{\mathbf{D}} \mathbf{g}(\mathbf{x}, \mathbf{y}) \mathbf{f}(\mathbf{x}, \mathbf{y}) \mathrm{d}\mathbf{x} \mathrm{d}\mathbf{y}$$

$$= \int_{\mathbf{a}}^{\mathbf{b}} \left[\int_{\mathbf{c}}^{\mathbf{d}} \mathbf{g}\mathbf{f} \, \mathrm{d}\mathbf{y} \right] \mathrm{d}\mathbf{x}$$

$$= \int_{\mathbf{c}}^{\mathbf{d}} \left[\int_{\mathbf{a}}^{\mathbf{b}} \mathbf{g}\mathbf{f} \, \mathrm{d}\mathbf{x} \right] \mathrm{d}\mathbf{y}$$

- 若积分区域为

(2) $\mathbf{D} =$

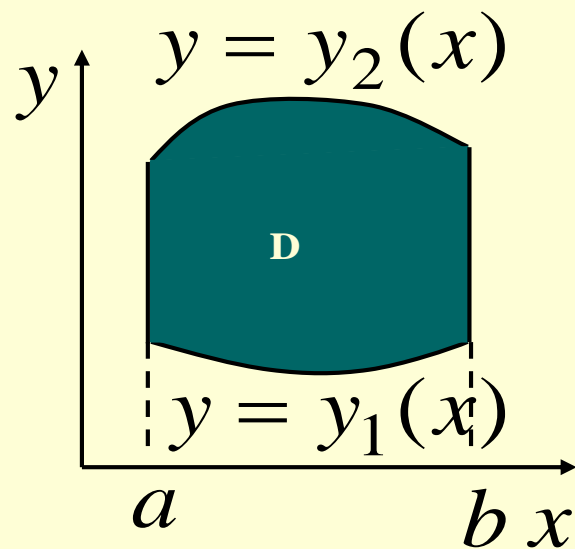
$$\{(\mathbf{x}, \mathbf{y}) \mid \mathbf{a} \leq \mathbf{x} \leq \mathbf{b}, \mathbf{y}_1(\mathbf{x}) \leq \mathbf{y} \leq \mathbf{y}_2(\mathbf{x})\}$$

$$\iint_{\mathbf{D}} \mathbf{g}(\mathbf{x}, \mathbf{y}) \mathbf{f}(\mathbf{x}, \mathbf{y}) \mathrm{d}\mathbf{x} \mathrm{d}\mathbf{y}$$

则

$$= \int_a^b \mathrm{d}\mathbf{x} \int_{y_1(\mathbf{x})}^{y_2(\mathbf{x})} \mathbf{g}\mathbf{f} \mathrm{d}\mathbf{y}$$

$$= \int_a^b \left[\int_{y_1(\mathbf{x})}^{y_2(\mathbf{x})} \mathbf{g}\mathbf{f} \mathrm{d}\mathbf{y} \right] \mathrm{d}\mathbf{x}$$



- 若积分区域为

(3) **D**

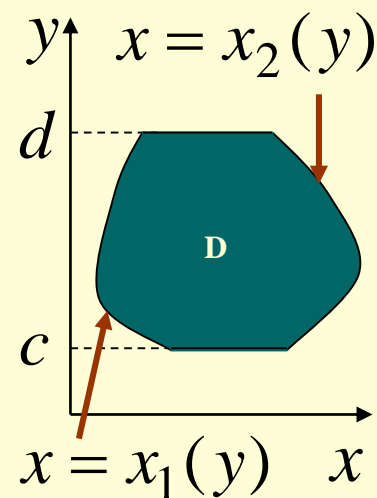
$$= \{(\mathbf{x}, \mathbf{y}) \mid \mathbf{c} \leq \mathbf{y} \leq \mathbf{d}, \mathbf{x}_1(\mathbf{y}) \leq \mathbf{x} \leq \mathbf{x}_2(\mathbf{y})\}$$

$$\iint_{\mathbf{D}} \mathbf{g}(\mathbf{x}, \mathbf{y}) \mathbf{f}(\mathbf{x}, \mathbf{y}) \mathrm{d}\mathbf{x} \mathrm{d}\mathbf{y}$$

则

$$= \int_{\mathbf{c}}^{\mathbf{d}} \mathrm{d}\mathbf{y} \int_{\mathbf{x}_1(\mathbf{y})}^{\mathbf{x}_2(\mathbf{y})} \mathbf{g}\mathbf{f} \mathrm{d}\mathbf{x}$$

$$= \int_{\mathbf{c}}^{\mathbf{d}} \left[\int_{\mathbf{x}_1(\mathbf{y})}^{\mathbf{x}_2(\mathbf{y})} \mathbf{g}\mathbf{f} \mathrm{d}\mathbf{x} \right] \mathrm{d}\mathbf{y}$$




例1 设 X 服从 $N(0, 1)$ 分布, 求 $E(X^2), E(X^3), E(X^4)$

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

$$E(X^2) = \int_{-\infty}^{\infty} \frac{x^2}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = - \int_{-\infty}^{\infty} \frac{x}{\sqrt{2\pi}} de^{-\frac{x^2}{2}}$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 1$$


$$E(X^3) = \int_{-\infty}^{\infty} \frac{x^3}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 0$$

$$E(X^4) = \int_{-\infty}^{\infty} \frac{x^4}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

$$= - \int_{-\infty}^{\infty} \frac{x^3}{\sqrt{2\pi}} de^{-\frac{x^2}{2}}$$

$$= 3 \int_{-\infty}^{\infty} \frac{x^2}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 3$$

例2 长途汽车起点站于每时的10分、30分、55分发车，设乘客不知发车时间，于每小时的任意时刻随机地到达车站，求乘客的平均候车时间



解:设乘客于某时 X 分到达车站,候车时间为 Y ,则

$$Y = g(X) = \begin{cases} 10 - X & 0 \leq X < 10 \\ 30 - X & 10 \leq X < 30 \\ 55 - X & 30 \leq X < 55 \\ 70 - X & 55 \leq X < 60 \end{cases} \quad f_X(x) = \begin{cases} \frac{1}{60} & 0 < x < 60 \\ 0 & \text{others} \end{cases}$$

$$\therefore E(Y) = \frac{1}{60} \int_0^{60} g(x) dx$$

$$= 10 \text{分} 25 \text{秒}$$

四.数学期望的性质

1. $E(c)=c, c$ 为常数;

2. $E(cX)=cE(X)$, c 为常数;

3. $E(X+Y)=E(X)+E(Y)$;

$$E\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n E(X_i)$$

$$E\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i E(X_i)$$

4. 若 X 与 Y 独立, 则 $E(XY)=E(X)E(Y)$.

若 X_1, \dots, X_n 独立, 则 $E\left(\prod_{i=1}^n X_i\right) = \prod_{i=1}^n E(X_i)$

例1 若 $X \sim B(n, p)$, 求 $E(X)$

解: 设

$$X_i = \begin{cases} 1 & \text{第 } i \text{ 次试验事件 } A \text{ 发生} \\ 0 & \text{第 } i \text{ 次试验事件 } A \text{ 不发生} \end{cases}$$

则 $E(X_i) = p$ $X = \sum_{i=1}^n X_i$


$$E(X) = \sum_{i=1}^n E(X_i) = np$$

例2 设某种疾病的发病率为1%，在1000个人中普查这种疾病，为此要化验每个人的血。方法是，每100个人一组，把从100个人抽来的血混在一起化验，如果混合血样呈阴性，则通过，如果混合血样呈阳性，则再分别化验该组每个人的血样。求平均化验次数。（此病不传染）

解：设 X_j 为第 j 组的化验次数， $j = 1, \dots, 10$

X 为1000人的化验次数，则

X_j	1	101
p_j	$(99\%)^{100}$	$1 - (99\%)^{100}$


$$EX_j = 0.99^{100} + (101)(1 - 0.99^{100})$$

$$E(X) = E\left(\sum_{j=1}^{10} X_j\right) = \sum_{j=1}^{10} E(X_j)$$

$$= 10[0.99^{100} + (101)(1 - 0.99^{100})]$$

$$= 1000 \cdot \left[1 + \frac{1}{100} - 0.99^{100}\right]$$

$$\approx 644$$

第一节 期望小结

知识点

- 1、一维随机变量的期望；
- 2、一维随机变量函数的期望；
- 3、二维随机变量函数的期望；
- 4、期望的性质；

考点

- 1、会求一维随机变量的期望；
- 2、会求一维随机变量函数的期望；
- 3、会求二维随机变量函数的期望；
- 4、会利用期望的性质求随机变量函数的期望；

4.2 方差

一. 定义与性质

方差描述了随机变量的取值与其均值的偏离程度。

1. **定义** 若 $E(X)$, $E(X^2)$ 存在, 则称

$$E[X-E(X)]^2$$

为r. v. X 的**方差**, 记为 $D(X)$, 或 $\text{Var}(X)$.

称 $\sigma(X) = \sqrt{D(X)}$ 为r. v. X 的**标准差或均方差**

2. **推论** $D(X) = E(X^2) - [E(X)]^2$

例1： 设随机变量X的概率密度为


$$f(x) = \begin{cases} 1+x & -1 < x < 0 \\ 1-x & 0 \leq x < 1 \\ 0 & \text{其它} \end{cases}$$

1) 求 $D(X)$, 2) 求 $D(X^2)$

$$\text{解: (1) } E(X) = \int_{-1}^0 x(1+x)dx + \int_0^1 x(1-x)dx = 0$$

$$E(X^2) = \int_{-1}^0 x^2(1+x)dx + \int_0^1 x^2(1-x)dx = \frac{1}{6}$$

$$\therefore D(X) = E(X^2) - [E(X)]^2 = \frac{1}{6}$$


$$(2) \quad D(X^2) = E(X^4) - [E(X^2)]^2$$

$$E(X^4) = \int_{-1}^0 x^4(1+x)dx + \int_0^1 x^4(1-x)dx = \frac{1}{15}$$

$$D(X^2) = \frac{1}{15} - \left(\frac{1}{6}\right)^2 = \frac{7}{180}$$

$$f(x) = \begin{cases} 1+x & -1 < x < 0 \\ 1-x & 0 \leq x < 1 \\ 0 & \text{其它} \end{cases}$$

3. 方差的性质

(1) $D(c)=0$. 反之, 若 $D(X)=0$, 则存在常数 C ,
使 $P\{X=C\}=1$, 且 $C=E(X)$;

(2) $D(aX)=a^2D(X)$, a 为常数; $D(-X)=D(X)$

(3) 若 X, Y 独立, 则 $D(X+Y)=D(X)+D(Y)$;

若 X_1, \dots, X_n 独立, 则 $D\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n D(X_i)$

二.几个重要r.v.的方差


1. 二项分布 $B(n, p)$:

解： 设 $X_i = \begin{cases} 1 & \text{第}i\text{次试验事件}A\text{发生} \\ 0 & \text{第}i\text{次试验事件}A\text{不发生} \end{cases}$

则

$$D(X_i) = E(X_i^2) - [E(X_i)]^2 = p - p^2 = p(1-p)$$

$$D(X) = \sum_{i=1}^n D(X_i) = np(1-p)$$



2. 泊松分布 $p(\lambda)$: $E(X) = \lambda$ $\left(\sum_{n=1}^{\infty} f_n(x) \right)' = \sum_{n=1}^{\infty} f_n'(x)$


$$X \sim P\{X = k\} = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \dots$$

$$E(X^2) = \sum_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=1}^{\infty} \frac{k \lambda^k e^{-\lambda}}{(k-1)!}$$

$$\text{由于 } E(X) = \sum_{k=1}^{\infty} \frac{\lambda^k e^{-\lambda}}{(k-1)!} = \lambda \quad \text{即} \quad \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} = \lambda e^{\lambda}$$

两边对 λ 求导,得

$$\sum_{k=1}^{\infty} \frac{k \lambda^{k-1}}{(k-1)!} = (1 + \lambda) e^{\lambda} \quad \text{即} \quad \sum_{k=1}^{\infty} \frac{k \lambda^k e^{-\lambda}}{(k-1)!} = \lambda(1 + \lambda)$$
$$\therefore D(X) = \lambda$$



1. 二项分布 $B(n, p)$: $D(X) = np(1-p)$

2. 泊松分布 $p(\lambda)$: $D(X) = \lambda$

3. 均匀分布 $U(a, b)$: $D(X) = \frac{(b-a)^2}{12}$.

4. 指数分布: $D(X) = \frac{1}{\lambda^2}$

5. 正态分布 $N(\mu, \sigma^2)$: $D(X) = \sigma^2$.

思考:

1. 请给出一个离散型随机变量 X 和一个连续型随机变量 Y , 使它们的期望都是 2, 方差都是 1。

三.切比雪夫不等式

若**r.v.X**的期望和方差存在，则对任意 $\varepsilon > 0$ ，
有

$$P\{|X - E(X)| \geq \varepsilon\} \leq \frac{D(X)}{\varepsilon^2}$$

这就是著名的切比雪夫(Chebyshev)不等式。

它有以下等价的形式：

$$P\{|X - E(X)| < \varepsilon\} \geq 1 - \frac{D(X)}{\varepsilon^2}$$



已知某种股票每股价格 X 的平均值为1元，标准差为0.1元，求 a ，使股价大于等于 $1+a$ 元或小于等于 $1-a$ 元的概率小于10%。

解：由切比雪夫不等式

$$P\{|X - 1| \geq a\} \leq \frac{0.01}{a^2};$$

令 $\frac{0.01}{a^2} < 0.1$

$$P\{|X - E(X)| \geq \varepsilon\} \leq \frac{D(X)}{\varepsilon^2}$$

$$\Rightarrow a^2 > 0.1 \quad \Rightarrow a > 0.32$$

4.3 协方差，相关系数

一.协方差定义与性质

1.协方差定义 若r.v. X 的期望 $E(X)$ 和 Y 的期望 $E(Y)$ 存在, 则称

$$\text{Cov}(X, Y) = E\{[X - E(X)][Y - E(Y)]\}$$

为 X 与 Y 的**协方差**,


易见 $\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$

当 $\text{Cov}(X, Y) = 0$ 时, 称 X 与 Y 不相关。



“ X 与 Y 独立”和“ X 与 Y 不相关”有何关系？

X 与 Y 独立 \rightarrow X 与 Y 不相关 X 与 Y 相关 \rightarrow X 与 Y 不独立



2.协方差性质

(1) $\text{Cov}(X, Y) = \text{Cov}(Y, X);$


(2) $\text{Cov}(X, X) = D(X); \text{Cov}(X, c) = 0$

(3) $\text{Cov}(aX, bY) = ab\text{Cov}(X, Y)$, 其中a, b为 常数

(4) $\text{Cov}(X+Y, Z) = \text{Cov}(X, Z) + \text{Cov}(Y, Z);$

(5) $\text{Cov}(a_1X + b_1Y + c_1, a_2X + b_2Y + c_2)$

$$= a_1a_2D(X) + b_1b_2D(Y) + (a_1b_2 + a_2b_1)\text{Cov}(X, Y)$$


$$(5) \text{Cov}(a_1X + b_1Y + c_1, a_2X + b_2Y + c_2)$$


$$=a_1a_2D(X) + b_1b_2D(Y) + (a_1b_2 + a_2b_1)COV(X, Y)$$

$$D(X+C)=D(X)$$

$$(6) D(X+Y)=D(X)+D(Y)+2\text{Cov}(X, Y).$$

$$D(X-Y)=D(X)+D(Y)-2\text{Cov}(X, Y).$$

$$(7) D(aX + bY + C) = a^2D(X) + b^2D(Y) + 2abCOV(X, Y)$$



EX: 设随机变量 $X \sim B(12, 0.5)$, $Y \sim N(0, 1)$,

$\text{Cov}(X, Y) = -1$, 求 $V = 4X + 3Y + 1$ 与 $W = -2X + 4Y$

的方差与协方差

答: $D(X) = 3$, $D(Y) = 1$

$$D(V) = 16D(X) + 9D(Y) + 24\text{Cov}(X, Y) = 33$$

$$D(W) = 4D(X) + 16D(Y) - 16\text{Cov}(X, Y) = 44$$

$$\text{Cov}(V, W) = \text{Cov}(4X + 3Y, -2X + 4Y)$$

$$= -8D(X) + 16\text{Cov}(X, Y) - 6\text{Cov}(Y, X) + 12D(Y)$$

$$= -22$$

二.相关系数

1. 定义: 若r.v. X, Y 的方差和协方差均存在, 且 $DX>0, DY>0$, 则

$$\rho_{XY} = \frac{\text{cov}(X, Y)}{\sqrt{DX} \sqrt{DY}}$$

称为 X 与 Y 的**相关系数**.

注: 若记 $X^* = \frac{X - E(X)}{\sqrt{DX}}$

称为 X 的标准化, 易知 $EX^*=0$, $DX^*=1$.且

$$\rho_{XY} = \rho_{X^*Y^*}$$

2.相关系数的性质

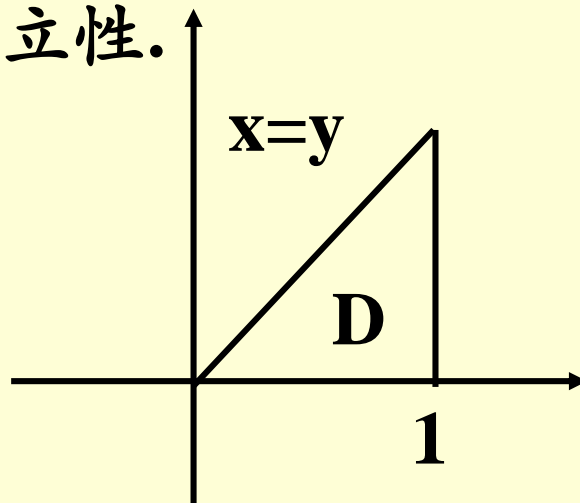
- (1) $|\rho_{XY}| \leq 1$;
- (2) $|\rho_{XY}| = 1 \Leftrightarrow$ 存在常数 a, b 使 $P\{Y = aX + b\} = 1$;
- (3) X 与 Y 不相关 $\Leftrightarrow \rho_{XY} = 0$.


EX

1. 设 (X, Y) 服从区域 $D: 0 < x < 1, 0 < y < x$ 上的均匀分布, 求 X 与 Y 的相关系数与独立性.

解

$$f(x, y) = \begin{cases} 2 & (x, y) \in D \\ 0 & \text{others} \end{cases}$$




$$E(X) = \int_0^1 2x dx \int_0^x dy = \frac{2}{3}$$

$$D(X) = \int_0^1 2x^2 dx \int_0^x dy - \frac{4}{9} = \frac{1}{18}$$

$$E(Y) = \int_0^1 2dx \int_0^x y dy = \frac{1}{3}$$

$$D(Y) = \int_0^1 2dx \int_0^x y^2 dy - \frac{1}{9} = \frac{1}{18}$$

$$E(XY) = \int_0^1 2x dx \int_0^x y dy = \frac{1}{4}$$

$$COV(X, Y) = E(XY) - E(X)E(Y) = \frac{1}{36}$$

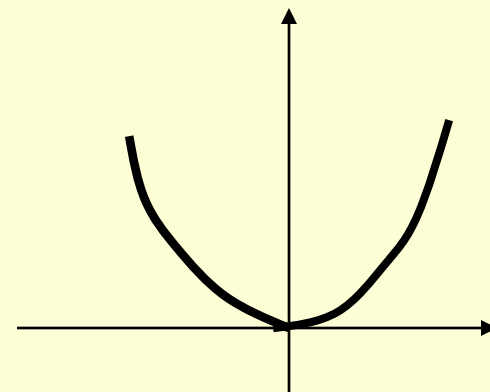
$$\rho_{XY} = \frac{COV(X, Y)}{\sqrt{D(X)D(Y)}} = \frac{1}{2}$$

X与Y相关且不独立。

EX2

1) $X \sim U(0,1), Y = X^2$, 求 ρ_{XY}

2) $X \sim U(-1,1), Y = X^2$, 求 ρ_{XY}



解1)


$$E(X) = \frac{1}{2}, E(Y) = \frac{1}{3}, E(XY) = \frac{1}{4}, D(X) = \frac{1}{12}, D(Y) = \frac{4}{45}$$

$$\rho_{XY} = \frac{\frac{1}{12}}{\sqrt{\frac{1}{12} \times \frac{4}{45}}} \approx 0.968$$

$$2) E(X) = 0, E(XY) = 0$$

$$\rho_{XY} = 0$$

以上EX的结果说明了什么？



例3 设 $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, 则 $\rho_{XY} = \rho$.

可见, 若 (X, Y) 服从二维正态分布, 则 X 与 Y 独立的充分必要条件是 X 与 Y 不相关。

4.4 矩、协方差矩阵

1. K阶原点矩 $E(X^k)$, $k=1, 2, \dots$

2. K阶中心矩 $E[X-E(X)]^k$, $k=2, 3, \dots$

3. K+I阶混合原点矩

$$E(X^k Y^l), k, l=1, 2, \dots;$$

4. K+I阶混合中心矩

$$E\{[X-E(X)]^k[Y-E(Y)]^l\}, k, l=1, 2, \dots;$$

4.5 协方差矩阵

1.定义 设 X_1, \dots, X_n 为 n 个r.v., 记 $c_{ij} = \text{Cov}(X_i, X_j)$, $i, j=1, 2, \dots, n$. 则称由 c_{ij} 组成的矩阵为随机变量 X_1, \dots, X_n 的协方差矩阵 \mathbf{C} 。即

$$\mathbf{C} = (c_{ij})_{n \times n} = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \dots & \dots & \dots & \dots \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{bmatrix}$$



本节要求

知识点

- 1、随机变量的方差；
- 2、二维随机变量协方差与相关系数；
- 3、方差与协方差的性质；
- 4、矩与协方差矩阵；

考点

- 1、会求随机变量的方差；
- 2、会求二维随机变量协方差与相关系数；
- 3、会利用方差与协方差的性质求随机变量的方差与协方差；
- 4、会求矩与协方差矩阵

切比雪夫不等式

小结

