第四章 随机变量的数字特征

修贤超

https://xianchaoxiu.github.io

目录

数学期望

- 定义: 离散型随机变量 X 分布律为 $P\{X = x_k\} = p_k$, $k = 1, 2, \cdots$, 若级数
$$\sum_{k=1}^{\infty} x_k p_k$$
绝对收敛, 则随机变量 X 的数学期望为
$$E(X) = \sum_{k=1}^{\infty} x_k p_k$$

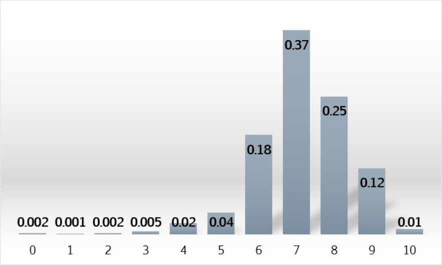
- 定义: 连续型随机变量 X 概率密度为 $f(x)$, 若积分
$$\int_{-\infty}^{\infty} x f(x) dx$$
绝对收敛, 则随机变量 X 的数学期望为
$$E(X) = \int_{-\infty}^{\infty} x f(x) dx$$

- 某医院当新生儿诞生时，医生依据婴儿身体状况进行评分，分数为随机变量 X，按以往资料其分布律如表所示，试求 X 的数学期望 $E(X)$

X	0	1	2	3	4	5	6	7	8	9	10
p_k	0.002	0.001	0.002	0.005	0.02	0.04	0.18	0.37	0.25	0.12	0.01

解答

$$
\begin{aligned}
E(X) &= 0\times 0.002 + 1\times 0.001 + 2\times 0.002 \\
&\quad + 3\times 0.005 + 4\times 0.0 + 5\times 0.04 \\
&\quad + 6\times 0.18 + 7\times 0.37 + 8\times 0.25 \\
&\quad + 9\times 0.12 + 10\times 0.01 \\
&= 7.15
\end{aligned}
$$



例 3

- 某车站$8:00$-$9:00$、$9:00$-$10:00$ 各有一辆车到站，但到站时间随机，两者到站时间相互独立，如下表所示．一旅客 $8:20$ 到站，求其候车时间的数学期望

到站时间	$8:10$	$8:30$	$8:50$	$9:10$	$9:30$	$9:50$
概率	$1/6$	$3/6$	$2/6$	$1/6$	$3/6$	$2/6$

解答 设旅客候车时间为 X，分布律为

X	10	30	50	70	90
p_k	$3/6$	$2/6$	$1/6\times1/6$	$1/6\times3/6$	$1/6\times2/6$

$$\begin{aligned}E(X)&=10\times3/6+30\times2/6+50\times1/6\times1/6\\&+70\times1/6\times3/6+90\times1/6\times2/6\\&=27.22\end{aligned}$$

例 4

- 某商店销售某家电, 先使用后付款, 按使用寿命 X 收费. 若 $X1$, 一台付款 1500 元. 若 $1 < X \le 2$, 一台付款 2000 元. 若 $2 < X \le 3$, 一台付款 2500 元. 若 $X > 3$, 一台付款 3000 元. 设寿命 X 服从指数分布

$$f(x) = \begin{cases} \frac{1}{10} e^{-x/10}, & x > 0 \\ 0, & x \le 0 \end{cases}$$

  求每台家电收费 Y 的数学期望

  解答 先求寿命 X 落在各个时间区间的概率, 即

$$P\{X \le 1\} = \int_0^1 \frac{1}{10} e^{-x/10} dx = 1 - e^{-0.1} = 0.0952$$

$$P\{1 < X \le 2\} = \int_1^2 \frac{1}{10} e^{-x/10} dx = e^{-0.1} - e^{-0.2} = 0.0861$$

例 4

$$P\{2 < X \leq 3\} = \int_2^3 \frac{1}{10} e^{-x/10} dx = e^{-0.2} - e^{-0.3} = 0.0779$$

$$P\{X > 3\} = \int_3^\infty \frac{1}{10} e^{-x/10} dx = e^{-0.3} = 0.7408$$

每台家电收费 Y 的分布律为

Y	1500	2000	2500	3000
p_k	0.0952	0.0861	0.0779	0.7408

于是数学期望为

$$E(Y) = \sum_{k=1}^{4} y_k p_k = 2732.15$$

例 5

- 在群体中普查某种疾病，可用两种方法抽验 N 人血液: (1) 分别抽验，共需验 N 次; (2) k 人一组混合验血，若成阳性，再分别抽验，一组人共需验 $k+1$ 次. 个人呈阳性概率为 p，且相互独立. 采用第二种方案，求合适的 k 值

解答 由于阳性概率 p，阴性概率为 $q=1-p$. 设以 k 人一组混合验血，阴性概率 q^k，阳性概率 $1-q^k$. 设每人化验次数为随机变量 X，其分布律为

X	$1/k$	$(k+1)/k$
p_k	q^k	$1q^k$

X 的数学期望为

$$E(X)=\frac{1}{k}q^k+\frac{k+1}{k}(1-q^k)=1-q^k+\frac{1}{k}$$

于是 N 个人平均需化验次数为 $N(1-q^k+\frac{1}{k})$

例 6

- 设随机变量 $X \sim \pi(\lambda)$, 求 $E(X)$

  解答　X 的分布律为

$$P(X=k)=\frac{\lambda^k e^{-\lambda}}{k!},\ k=0,1,\cdots,\ \lambda>0$$

  X 的数学期望为

$$\begin{aligned}E(X)&=\sum_{k=0}^{\infty}k\frac{\lambda^k e^{-\lambda}}{k!}=\lambda e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^{k-1}}{(k-1)!}\\&=\lambda e^{-\lambda}e^{\lambda}=0.0952\end{aligned}$$

- 泊松分布 $X \sim \pi(\lambda)$ 的数学期望 $E(X)=\lambda$

例 7

- 设随机变量 $X \sim U(a,b)$, 求 $E(X)$

  解答 X 的概率密度为

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

  X 的数学期望为

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

- 均匀分布 $X \sim U(a,b)$ 的数学期望位于区间的中点

课堂练习 1

- 设随机变量 X 服从指数分布, 求 $E(X)$

课堂练习 1

- **设随机变量 X 服从指数分布, 求 $E(X)$**

  解答 X 的概率密度为

$$f(x) = \begin{cases} \frac{1}{\theta} e^{\frac{-x}{\theta}}, & x > 0 \\ 0, & \text{其他} \end{cases}, \quad \theta > 0$$

  X 的数学期望为

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} x f(x) dx = \int_0^{\infty} \frac{x}{\theta} e^{\frac{-x}{\theta}} dx = \int_0^{\infty} (-x) e^{\frac{-x}{\theta}} d\left(\frac{-x}{\theta}\right) \\ &= \int_0^{\infty} (-x) d\left(\frac{-x}{\theta}\right) = -x e^{\frac{-x}{\theta}} |_0^{\infty} + \int_0^{\infty} e^{\frac{-x}{\theta}} dx \\ &= -\theta e^{\frac{-x}{\theta}} |_0^{\infty} = \theta \end{aligned}$$

- **指数分布的数学期望 $E(X) = \theta$**

课堂练习 2

- 设随机变量 X 服从正态分布 $X \sim N(\mu, \sigma^2)$, 求 $E(X)$

课堂练习 2

- 设随机变量 X 服从正态分布 $X \sim N(\mu, \sigma^2)$, 求 $E(X)$

  解答 X 的概率密度为

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

  X 的数学期望为

$$\begin{aligned}
E(X) &= \int_{-\infty}^{\infty} xf(x)dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} d(x-\mu) \\
&= \int_{-\infty}^{\infty} \frac{y+\mu}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2\sigma^2}} dy = \int_{-\infty}^{\infty} \frac{y}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2\sigma^2}} dy + \int_{-\infty}^{\infty} \frac{\mu}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2\sigma^2}} dy \\
&= \frac{-\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2\sigma^2}} d\left(\frac{y^2}{2\sigma^2}\right) + \mu \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2\sigma^2}} dy = 0 + \mu = \mu
\end{aligned}$$

- 正太分布 $X \sim N(\mu, \sigma^2)$ 的数学期望 $E(X) = \mu$

数学期望

- 定理: 设 Y 是随机变量 X 的函数, $Y = g(X)$ 且 g 是连续函数
  - 如果 X 是离散型随机变量, 分布律为 $P\{X = x_k\} = p_k,\ k = 1, 2, \cdots$, 若级数 $\sum_{k=1}^{\infty} g(x_k)p_k$ 绝对收敛, 则有

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

  - 如果 X 是连续型随机变量, 概率密度为 $f(x)$, 若积分 $\int_{-\infty}^{\infty} g(x)f(x)dx$ 绝对收敛, 则有

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

- 当我们求 $E(Y)$ 时, 不必算出 Y 的分布律或概率密度函数

例 8

- 设风速 V 在 $(0, a)$ 上服从均匀分布, 即具有概率密度

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

由设飞机机翼受到正压力 W 是 V 的函数: $W = kV^2$, 求 W 的数学期望

解答　根据上述定理有

$$\begin{aligned} E(W) &= \int_{-\infty}^{\infty} kv^2 f(v) dv \\ &= \int_0^a kv^2 \frac{1}{a} dv = \frac{1}{3} ka^2 \end{aligned}$$

数学期望

- 定理：设 Z 是随机变量 X, Y 的函数，$Z = g(X, Y)$ 且 g 是连续函数
  - 如果 X, Y 是离散型随机变量，分布律为 $P\{X = x_i, Y = y_j\} = p_{ij}$，$i, j = 1, 2, \cdots$，则有

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

  - 如果 X, Y 是连续型随机变量，概率密度为 $f(x, y)$，则有

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

- 右边项应满足绝对收敛

例 9

- 设随机变量 (X, Y) 具有概率密度

$$f(x,y)=\begin{cases}\frac{3}{2x^3y^2}, & \frac{1}{x}<y<x,\ \ x>1\\ 0, & \text{其他}\end{cases}$$

由数学期望 $E(Y), E(\frac{1}{XY})$

解答 根据上述定理有

$$E(Y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}yf(x,y)dxdy=\int_{1}^{\infty}\int_{1/x}^{x}\frac{3y}{2x^3y^2}dydx=3\int_{1}^{\infty}\frac{ln(x)}{x^3}dx=\frac{3}{4}$$

$$E\left(\frac{1}{XY}\right)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{xy}f(x,y)dxdy=\int_{1}^{\infty}\int_{1/x}^{x}\frac{3y}{2x^4y^3}dydx$$

$$=\frac{3}{4}\int_{1}^{\infty}\left(\frac{1}{x^2}-\frac{1}{x^6}\right)dx=\frac{3}{5}$$

例 10

- 某公司规划新产品, 已知售出一件获利 m 元, 积压一件损失 n 元, 预测销售量 Y 服从指数分布, 其概率密度为

$$f_Y(y) = \begin{cases} \frac{1}{\theta} e^{-y/\theta}, & y > 0 \\ 0, & \text{其他} \end{cases}$$

若要获得利润数学期望最大, 应生产多少件产品

解答 设生产 x 件, 获利为 Q, 则

$$Q = Q(m) = \begin{cases} mY - n(x - Y), & Y < x \\ mx, & Y \geq x \end{cases}$$

于是 Q 的数学期望为

$$\begin{aligned} E(Q) &= \int_0^\infty Q f_Y(y) dy \\ &= \int_0^x [my - n(x-y)] \frac{1}{\theta} e^{-y/\theta} dy + \int_x^\infty mx \frac{1}{\theta} e^{-y/\theta} dy \\ &= (m+n)\theta - (m+n)\theta e^{-x/\theta} - nx \end{aligned}$$

求极值

$$\frac{d}{dx} E(Q) = (m+n)\theta - (m+n)\theta e^{-x/\theta} - nx = 0 \quad \Rightarrow \quad x = -\theta ln \left(\frac{n}{m+n} \right)$$

例如, $\theta = 10000$, $m = 500$, $n = 2000$ 得到

$$x = -10000 ln \left(\frac{2000}{500 + 2000} \right) = 2231.4$$

性质

- 设 C 为常数, 则 $E(C) = C$
- 设 X 为随机变量, C 为常数, 则有 $E(CX) = CE(X)$
- 设 X, Y 为随机变量, 则有 $E(X+Y) = E(X) + E(Y)\}$

  证明

$$\begin{aligned}
E(X+Y) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f(x,y)dxdy \\
&= \int_{-\infty}^{\infty} x\int_{-\infty}^{\infty} f(x,y)dydx + \int_{-\infty}^{\infty} y\int_{-\infty}^{\infty} f(x,y)dxdy \\
&= \int_{-\infty}^{\infty} xf_X(x)dx + \int_{-\infty}^{\infty} yf_Y(y)dy \\
&= E(X) + E(Y)
\end{aligned}$$

性质

- 设 X, Y 为相互独立的随机变量, 则有 $E(XY) = E(X)E(Y)$

  证明

$$\begin{aligned}
E(XY) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)dxdy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_X(x)f_Y(y)dxdy \\
&= \int_{-\infty}^{\infty} xf_X(x)\int_{-\infty}^{\infty} yf_Y(y)dydx = \int_{-\infty}^{\infty} xf_X(x)E(Y)dx \\
&= E(Y)\int_{-\infty}^{\infty} xf_X(x)dx \\
&= E(X)E(Y)
\end{aligned}$$

例 12

- 机场大巴载 20 人至 10 个车站陆续下车，到站无人下车则不停车，以 X 表示停车次数，求 $E(X)$

  解答 引入随机变量

$$X_i = \begin{cases} 0, & \text{在第 } i \text{ 站无人下车} \\ 1, & \text{在第 } i \text{ 站有人下车} \end{cases}, \quad i = 1, 2, \cdots, 10$$

  易知 $X = X_1 + X_2 + \cdots + X_{10}$. 根据题意，每个人不在第 i 站下车的概率是 $\frac{9}{10}$，所有人都不在第 i 站下车的概率是 $\left(\frac{9}{10}\right)^{20}$. 也就是

$$P\{X_i = 0\} = \left(\frac{9}{10}\right)^{20}, \quad P\{X_i = 1\} = 1 - \left(\frac{9}{10}\right)^{20}, \; i = 1, 2, \cdots, 10$$

因此

$$E(X_i) = 0 \times \left(\frac{9}{10}\right)^{20} + 1 \times (1 - \left(\frac{9}{10}\right)^{20}) = 1 - \left(\frac{9}{10}\right)^{20},\ i = 1, 2, \cdots, 10$$

进而

$$\begin{aligned} E(X) &= E(X_1 + X_2 + \cdots + X_{10}) \\ &= E(X_1) + E(X_2) + \cdots + E(X_{10}) \\ &= 10\left[1 - \left(\frac{9}{10}\right)^{20}\right] = 8.784 \end{aligned}$$

例 13

- **设一电流中电流 $I(A)$ 和电阻 $R(\Omega)$ 是两个相互独立的随机变量，其概率密度分别为**

$$g(i)=\begin{cases}2i, & 0\le i\le 1\\ 0, & \text{其他}\end{cases},\quad h(r)=\begin{cases}\frac{r^2}{9}, & 0\le r\le 3\\ 0, & \text{其他}\end{cases}$$

试求电压 $U=IR$ 的均值

解答　**易知**

$$\begin{aligned}E(U)&=E(IR)=E(I)E(R)\\&=\left[\int_{-\infty}^{\infty} ig(i)di\right]\left[\int_{-\infty}^{\infty} rh(r)dr\right]\\&=\left(\int_0^1 2i^2di\right)\left(\int_0^3 \frac{r^3}{9}dr\right)=\frac{3}{2}(V)\end{aligned}$$

目录

方差

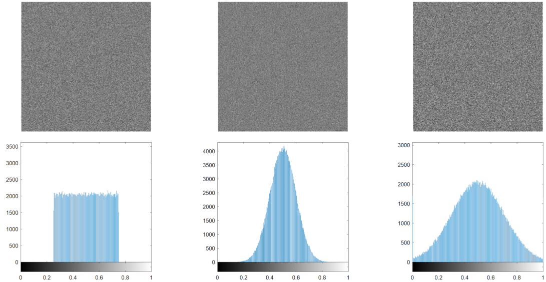
- **定义**: 设 X 是随机变量, 则 X 的**方差**为

$$D(X) = \mathrm{Var}(X) = E\{[X - E(X)]^2\}$$

标准差为

$$\sigma(X) = \sqrt{D(X)}$$

- 方差表示随机变量取值相对于均值的分散程度



方差

- 离散型随机变量 X 分布律为 $P\{X = x_k\} = p_k$, 则 X 的方差为

$$D(X) = \sum_{k=1}^{\infty} [x_k - E(X)]^2 p_k$$

- 连续型随机变量 X 概率密度为 $f(x)$, 则 X 的方差为

$$D(X) = \int_{-\infty}^{\infty} [x_k - E(X)]^2 f(x) dx$$

- 随机变量 X 的方差为

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

例 1

- 设随机变量 X 具有数学期望 $E(X)=\mu$, 方差 $D(X)=\sigma^2\neq 0$, 记

$$X^*=\frac{X-\mu}{\sigma}$$

则

$$E(X^*)=E(\frac{X-\mu}{\sigma})=\frac{1}{\sigma}[E(X)-\mu]=0$$

$$D(X^*)=E[(X^*)^2]-[E(X^*)]^2=E\left[\left(\frac{X-\mu}{\sigma}\right)^2\right]$$

$$=\frac{1}{\sigma^2}E[(X-\mu)^2]=\frac{1}{\sigma^2}\sigma^2=1$$

称 X^* 为 X 的标准化变量

例 2

- 设随机变量 X 具有 $(0-1)$ 分布, 其分布律为

$$P\{X=1\}=p,\quad P\{X=0\}=1-p$$

求 $D(X)$

解答

$$E(X)=0\times(1-p)+1\times p=p$$
$$E(X^2)=0^2\times(1-p)+1^2\times p=p$$

于是

$$\begin{aligned}D(X)&=E(X^2)-[E(X)]^2\\&=p-p^2=p(1-p)\end{aligned}$$

- $(0-1)$ 分布的数学期望为 p, 方差为 $p(1-p)$

- 设随机变量 $X \sim \pi(\lambda)$, 求 $D(X)$

  解答 随机变量 X 的分布律为

$$P\{X=k\}=\frac{\lambda^k e^{-\lambda}}{k!},\ k=0,1,\cdots,\ \lambda>0$$

  已知 $E(X)=\lambda$, 而

$$\begin{aligned} E(X^2) &= E[X(X-1)+X] = E[X(X-1)]+E(X) \\ &= \sum_{k=0}^{\infty} k(k-1)\frac{\lambda^k e^{-\lambda}}{k!}+\lambda = \lambda^2 e^{-\lambda}\sum_{k=2}^{\infty}\frac{\lambda^{k-2}}{(k-2)!}+\lambda \\ &= \lambda^2 e^{-\lambda}e^{\lambda}+\lambda = \lambda^2+\lambda \end{aligned}$$

  于是 $D(X)=E(X^2)-[E(X)]^2=\lambda$

- 泊松分布 $X \sim \pi(\lambda)$ 的数学期望和方差都为 λ

例 4

- 设随机变量 $X \sim U(a,b)$, 求 $D(X)$

  解答　随机变量 X 的概率密度为

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

  已知 $E(X)=\frac{a+b}{2}$, 则方差为

$$\begin{aligned} D(X) &= E(X^2)-[E(X)]^2 \\ &= \int_a^b x^2\frac{1}{b-a}dx-\left(\frac{a+b}{2}\right)^2=\frac{(b-a)^2}{12} \end{aligned}$$

- 均匀分布 $X \sim U(a,b)$ 的数学期望为 $\frac{a+b}{2}$, 方差为 $\frac{(b-a)^2}{12}$

例 5

- 设随机变量 X 服从指数分布, 其概率密度为

$$f(x) = \begin{cases} \frac{1}{\theta} e^{-x/\theta}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

其中 $\theta > 0$, 求 $E(X)$, $D(X)$

解答 随机变量 X 的期望和方差分别为

$$E(X) = \int_{-\infty}^{\infty} x f(x) = \int_{0}^{\infty} x \frac{1}{\theta} e^{-x/\theta} dx = \theta$$

$$\begin{aligned} D(X) &= E(X^2) - [E(X)]^2 \\ &= \int_{-\infty}^{\infty} x^2 f(x) dx - \theta^2 = \int_{0}^{\infty} \frac{x^2}{\theta} e^{-x/\theta} dx - \theta^2 = 2\theta^2 - \theta^2 = \theta^2 \end{aligned}$$

- 指数分布的数学期望为 θ, 方差为 θ^2

性质

- 设 C 为常数, 则

$$D(C) = 0$$

- 设 X 为随机变量, C 为常数, 则有

$$D(CX) = C^2 D(X), \quad D(X + C) = D(X)$$

- 设 X, Y 为随机变量, 则有

$$D(X + Y) = D(X) + D(Y) + 2E\{[X - E(X)][Y - E(Y)]\}$$

- 设 X, Y 为相互独立的随机变量, 则有

$$D(X + Y) = D(X) + D(Y)$$

- 设 $D(X) = 0$ 的充分必要条件是 X 以概率 1 取常数 $E(X)$, 即

$$P\{X = E(X)\} = 1$$

例 6

- 设随机变量 $X \sim b(n,p)$, 求 $E(X)$, $D(X)$

  解答 随机变量 X 是 n 重伯努利试验中事件 A 发生的次数, 引入随机变量

$$X_k = \begin{cases} 1, & \text{第 } k \text{ 次试验 } A \text{ 发生} \\ 0, & \text{第 } k \text{ 次试验 } A \text{ 不发生} \end{cases}$$

  已知 $E(X_k) = p, D(X_k) = p(1-p)$, 则

$$E(X) = E\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} E(X_k) = np$$

$$D(X) = D\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} D(X_k) = np(1-p)$$

- 伯努利分布 $X \sim b(n,p)$ 的数学期望为 np, 方差为 $np(1-p)$

例 7

- 设随机变量 $X \sim N(\mu, \sigma^2)$, 求 $E(X)$, $D(X)$

解答 $f(x)$ 的概率密度为

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

于是

$$E(X) = 0$$

$$D(X) = E\{[X - E(X)]^2\} = E\{[X - \mu]^2\} = \int_{-\infty}^{\infty} (x-\mu)^2 f(x) dx$$

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

- 正态分布 $X \sim N(\mu, \sigma^2)$ 的数学期望为 μ, 方差为 σ^2

例 8

- 设活塞直径为随机变量 $X \sim N(22.40, 0.03^2)$, 气缸直径为随机变量 $Y \sim N(22.50, 0.04^2)$, 两者相互独立, 任取一只活塞和一只气缸, 求活塞能装入气缸的概率

解答 由于 $X - Y$ 是随机变量, 有

$$X - Y \sim N(-0.10, 0.05^2)$$

于是

$$\begin{aligned} P(X < Y) &= P\{X - Y < 0\} \\ &= P\left\{\frac{(X - Y) - (-0.1)}{0.05} < \frac{0 - (-0.1)}{0.05}\right\} \\ &= \Phi\left(\frac{0.10}{0.05}\right) = \Phi(2) = 0.9972 \end{aligned}$$

切比雪夫不等式

- **定理**: 设随机变量 X 具有数学期望 $E(X)=\mu$, 方差 $D(X)=\sigma^2$, 则对于任意正数 ε, 有下列不等式成立

$$P\{|X-\mu|\geq\varepsilon\}\leq\frac{\sigma^2}{\varepsilon^2}$$

证明

$$\begin{aligned}P\{|X-\mu|\geq\varepsilon\}&=\int_{|X-\mu|\geq\varepsilon}f(x)dx\leq\int_{|X-\mu|\geq\varepsilon}\frac{|X-\mu|^2}{\varepsilon^2}f(x)dx\\&=\frac{1}{\varepsilon^2}\int_{-\infty}^{\infty}|X-\mu|^2f(x)dx=\frac{\sigma^2}{\varepsilon^2}\end{aligned}$$

- 切比雪夫不等式也可写作 $P\{|X-\mu|<\varepsilon\}\geq1-\frac{\sigma^2}{\varepsilon^2}$
- 切比雪夫不等式给出了在随机变量分布未知, 只知道其数学期望与方差情况下, 概率 $P\{|X-\mu|\geq\varepsilon\}$ 估计值的界限

目录

协方差及相关系数

- 定义: 设 X_1 和 X_2 是随机变量, 则其协方差为

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

相关系数为

$$\rho = \frac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$$

Plant Ecology **171:** 209–220, 2004.

Pines and oaks in the restoration of Mediterranean landscapes of Spain: New perspectives for an old practice – a review

Juli G. Pausas[1,2], Carme Bladé[1], Alejandro Valdecantos[1], José P. Seva[1], David Fuentes[1], José A. Alloza[1], Alberto Vilagrosa[1], Susana Bautista[2], Jordi Cortina[2] & Ramon Vallejo[1,*]
[1]*CEAM Centro de Estudios Ambientales del Mediterráneo, C. Charles Darwin 14, Parco Tecnològico, 46980 Paterna, València, SPAIN;* [2] *Departament d'Ecologia, Universitat d'Alacant, Alacant, SPAIN; *Author for correspondence (e-mail ramonv@ceam.es)*

Key words: Fire, *Pinus halepensis*, Plantation, *Quercus ilex*, Reforestation, Seedling survival

Abstract

Pines have been extensively used for land restoration in the Mediterranean basin and in other parts of the world, since the late 19^{th} century. The theoretical basis supporting pine utilisation was its stress-tolerant and pioneer features, and their attributed role of facilitating the development of late-successional hardwoods in the long-term. In the present work, the use of pines and hardwoods in forest restoration is discussed in the frame of the current disturbance regime and social demands for Mediterranean forests. Large pine plantations have recently disappeared because of their sensitivity to fire (e.g., *Pinus nigra*) or because of the short fire-intervals (e.g., *Pinus halepensis*). Combined pine and oak plantations are proposed for degraded land restoration on the basis of the complementary

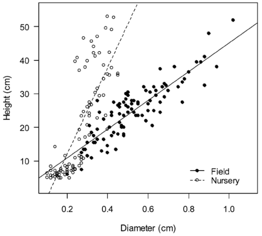


Figure 5. Height - diameter relationships for 1-year-old *Pinus halepensis* seedlings growing the nursery (open circles, dashed line) just before outplanting and for the same seedlings 18 months after planting in the field (i.e. seedlings of 2.5 years old) (black

协方差

- $\mathrm{Cov}(X,Y)=D(X)$
- $\mathrm{Cov}(X,Y)=\mathrm{Cov}(Y,X)$
- $\mathrm{Cov}(X_1+X_2,Y)=\mathrm{Cov}(X_1,Y)+\mathrm{Cov}(X_2,Y)$
- $\mathrm{Cov}(aX,bY)=ab\mathrm{Cov}(X,Y)$
- $\mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y)$

  证明

$$
\begin{aligned}
&\mathrm{Cov}(X,Y)=E\{[X-E(X)][Y-E(Y)]\}\\
&=E\{XY-YE(X)-XE(Y)+E(X)E(Y)\}\\
&=E(XY)-E(Y)E(X)-E(X)E(Y)+E(X)E(Y)\\
&=E(XY)-E(X)E(Y)
\end{aligned}
$$

协方差

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

  证明

$$
\begin{aligned}
&D(X,Y) = E[(X+Y)^2] - [E(X+Y)]^2\} \\
&= E[X^2 + 2XY + Y^2] - [E(X) + E(Y)] \\
&= E(X^2) + 2E(XY) + E(Y^2) - [E(X)]^2 - 2E(X)E(Y) - [E(Y)]^2 \\
&= \{E(X^2) - [E(X)]^2\} + \{E(Y^2) - [E(Y)]^2\} - 2\{E(XY) - E(X)E(Y)\} \\
&= D(X) + D(Y) + 2\mathrm{Cov}(X,Y)
\end{aligned}
$$

相关系数

- 用直线 $Y = a + bX$ 拟合数据, 残差的均方值为

$$e = E\{[Y - (a + bX)]^2\}$$

参数 a 和 b 使残差取得最小值

$$\begin{cases} \frac{\partial e}{\partial a} = 2a + 2bE(X) - 2E(Y) = 0 \\ \frac{\partial e}{\partial b} = 2bE(X^2) - 2E(XY) + 2aE(X) = 0 \end{cases}$$

求得

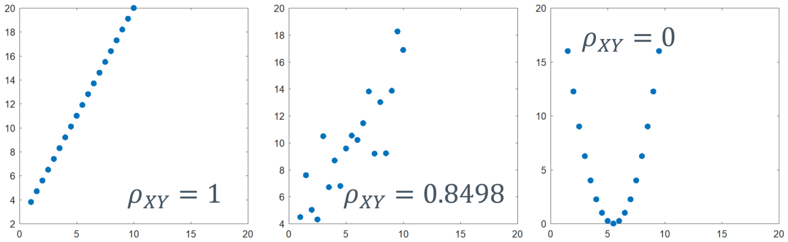
$$\begin{cases} a = E(Y) - E(X)\frac{\mathrm{Cov}(X,Y)}{D(X)} \\ b = \frac{\mathrm{Cov}(X,Y)}{D(X)} \end{cases}$$

于是得到

$$\min_{a,b} \; e = (1 - \rho_{XY}^2)D(Y)$$

相关系数

- **定理**: 设 X 和 Y 是随机变量, 则
  - $|\rho_{XY}| \leq 1$
  - $|\rho_{XY}| = 1$ 的充分必要条件是存在 a 和 b, 使 $P\{Y = a + bX\} = 1$



- ρ_{XY} 是 X 和 Y 线性相关程度的度量, $|\rho_{XY}| = 1$ 表示 X 和 Y 存在线性关系, $\rho_{XY} = 0$ 表示 X 和 Y 不线性相关

例 1

- 设 (X,Y) 的分布律如表所示, 求 X 和 Y 的关系

	$X=-2$	$X=-1$	$X=1$	$X=2$	$P\{Y=j\}$
$Y=1$	0	1/4	1/4	0	1/2
$Y=4$	1/4	0	0	1/4	1/2
$P\{X=i\}$	1/4	1/4	1/4	1/4	1

解答 易知 $P\{X=-2,Y=1\}=0$, $P\{X=-2\}P\{Y=1\}=1/8$, 则

$$P\{X=-2,Y=1\}\neq P\{X=-2\}P\{Y=1\} \quad\Rightarrow\quad \text{不相互独立}$$

因 $\mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y)=0$, 则

$$\rho_{XY}=0 \quad\Rightarrow\quad \text{不线性相关}$$

例 2

- 设二维随机变量 (X, Y) 服从正态分布, 概率密度为

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}exp\left\{\frac{-1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\}$$

求 X 和 Y 的相关系数

解答　已知 (X, Y) 的边缘密度为

$$f_X(x)=\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}},\quad f_Y(y)=\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}$$

求得 $\mathrm{Cov}(X,Y)=\rho\sigma_1\sigma_2$ 且 $\rho_{XY}=\rho$

目录

矩

- **定义**: 设 (X,Y) 是二维随机变量
  - X 的 k 阶**原点矩** (k 阶矩)

$$E(X^k),\ \ k=1,2,\cdots$$

  - X 的 k 阶**中心矩**

$$E\{[X-E(X)]^k\},\ \ k=2,3,\cdots$$

  - X 和 Y 的 $k+l$ 阶**混合矩**

$$E(X^kY^l),\ \ k,l=1,2,\cdots$$

  - X 和 Y 的 $k+l$ 阶**混合中心矩**

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

协方差矩阵

- **定义**: 二维随机变量 (X_1, X_2) 的协方差矩阵

$$\begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$$

其中

- $c_{11} = E\{[X_1 - E(X_1)]^2\}$
- $c_{12} = E\{[X_1 - E(X_1)]\}E\{[X_2 - E(X_2)]\}$
- $c_{21} = E\{[X_2 - E(X_2)]\}E\{[X_1 - E(X_1)]\}$
- $c_{22} = E\{[X_2 - E(X_2)]^2\}$

协方差矩阵

- **定义**: n 维随机变量 $(X_1, X_2, \cdots, X_n)$ 的协方差矩阵

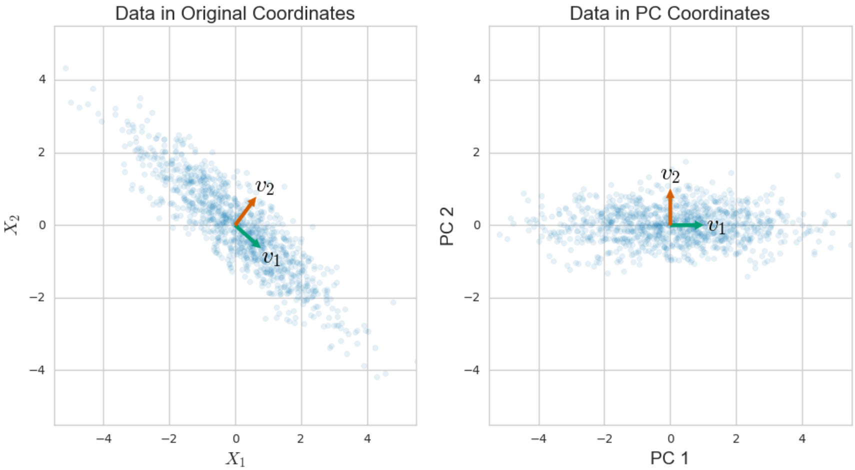
$$\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix}$$

其中

$$\begin{aligned} c_{ij} &= \mathrm{Cov}(X_i, X_j) \\ &= E\{[X_i - E(X_i)]\}E\{[X_j - E(X_j)]\}, \ i, j = 1, 2, \cdots, n \end{aligned}$$

主成分分析

- 协方差矩阵 $\Rightarrow$ 特征值分解



n 维正态分布

- **定义**: n 维正态分布随机变量 $(X_1, X_2, \cdots, X_n)$ 的概率密度

$$f(x_1, x_2, \cdots, x_n) = \frac{1}{(2\pi)^{n/2}[det(C)]^{1/2}} exp\left\{-\frac{1}{2}(X-\mu)^\top C^{-1}(X-\mu)\right\}$$

其中

- $X = (x_1, x_2, \cdots, x_n)^\top$
- $\mu = [E(X_1), E(X_2), \cdots, E(X_n)]^\top$
- C 是 $(X_1, X_2, \cdots, X_n)$ 的协方差矩阵

性质

- n 维正态分布随机变量 $(X_1, X_2, \cdots, X_n)$ 的每个分量为 X_i 为正态随机变量; $X_1, X_2, \cdots, X_n$ 为相互独立的正态随机变量, 则 $(X_1, X_2, \cdots, X_n)$ 为 n 维正态随机变量
- n 维随机变量 $(X_1, X_2, \cdots, X_n)$ 服从 n 维正态分布的充分必要条件是 $X_1, X_2, \cdots, X_n$ 的任意线性组合服从一维正态分布
- 若 $(X_1, X_2, \cdots, X_n)$ 服从 n 维正态分布, $Y_1, Y_2, \cdots, Y_k$ 是 X_j 的线性函数, 则 $(Y_1, Y_2, \cdots, Y_k)$ 也服从多维正态分布
- 设 $(X_1, X_2, \cdots, X_n)$ 服从 n 维正态分布, 则 $X_1, X_2, \cdots, X_n$ 相互独立等价于 $X_1, X_2, \cdots, X_n$ 两两不相关

小结

- **考试内容**
  - 随机变量的数学期望、方差、标准差及其性质、随机变量函数的数学期望、矩、协方差、相关系数及其性质
- **考试要求**
  - 理解随机变量数字特征（数学期望、方差、标准差、矩、协方差、相关系数）的概念，会运用数字特征的基本性质，并掌握常用分布的数字特征
  - 会求随机变量函数的数学期望

Q&A

Thank you!

感谢您的聆听和反馈