概率论与数理统计

Probability and

Mathematical Statistics

第四章 数字特征





- 4.1 随机变量的数学期望
- 4.2 随机变量的方差
- 4.3 随机变量的矩
- 4.4 协方差和相关系数



4.1 随机变量的数学期望(Expectation)



- 一 D.R.V.的数学期望
 - 二 C.R.V.的数学期望
- 三 R.V.函数的数学期望

四 数学期望的性质



一 D.R.V.的数学期望

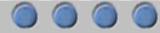


引	例
---	---

X	170	165	168
P	2/6	3/6	1/6
人数	2	3	1

平均身高=(170×2+165×3+168×1)/6

 $=170\times2/6+165\times3/6+168\times1/6$





1 定义 设D.R.V.X的分布列为 $P(X=x_i)=p_i, i=1,2...$

若 $\sum_{i} |x_{i}| p_{i} < +\infty$,称 $EX = \sum_{i} x_{i} p_{i}$ 为X的数学期望(均值).

2 常见D.R.V.的期望

例 $X\sim B(1,p)$, EX=p.

 \emptyset $X \sim B(n,p)$, EX = np.

例 $X \sim P(\lambda)$, $EX = \lambda$.

例 X~超几何分布,EX=nM/N.

例 $X\sim$ 几何分布,EX=1/p.



例 设 $X\sim B(1,p)$,求E(X)。

$$\mathbf{f} \mathbf{F} \qquad E(X) = 1 \times p + 0 \times (1 - p) = p$$

例 设 $X\sim P(\lambda)$,求E(X)。

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

二 C.R.V.的数学期望



2. 常见C.R.V.的期望

 $M \times U[a,b], E(X)=(a+b)/2.$

例 $X \sim E(\lambda)$, $E(X) = 1/\lambda$.

 $M \sim N(\mu, \sigma^2), E(X) = \mu$.

例 X服从柯西分布,密度函数为 $f(x) = \frac{1}{\pi(1+x^2)}, x \in \mathbf{R}.$ 则E(X)不存在.



4

 $m{\theta}$. 设随机变量X的概率密度为 $f(x) = \frac{1}{\pi(1+x^2)}, -\infty < x < \infty$, 证明X不存在数学期望。

:.由定义, X不存在数学期望。



例设 $X\sim E(\lambda)$,求E(X)(平均寿命)。

$$\mathbf{F} E(X) = \int_0^\infty x \lambda e^{-\lambda x} dx = -x e^{-\lambda x} \Big|_0^\infty + \int_0^{+\infty} e^{-\lambda x} dx = \frac{1}{\lambda}$$

例 设 $X \sim N(\mu, \sigma^2)$, 求E(X)。

$$\mathbf{P} E(X) = \int_{-\infty}^{+\infty} \frac{x}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \quad = \frac{1}{-\infty} \int_{-\infty}^{+\infty} \frac{\sigma u + \mu}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$$

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

三随机变量函数的期望 --定理法



设y=g(x)为一通常函数,令Y=g(X),其中X为R.V., Y也为R.V..

问题:已知X的分布,如何求Y的期望?

设z=g(x,y)为一通常函数,令Z=g(X,Y),其中(X,Y)为R.V., Z也为R.V..

问题:已知(X,Y)的分布,如何求Z的期望?



- 1.设g(x)为函数,Y=g(X)是R.V.,
- (1) 若D.R.V. X的分布列为 $P(X = x_i) = p_i, i = 1, 2, \cdots$ $\Rightarrow E[g(X)] = \sum_i g(x_i) p_i \cdot \overline{A} \sum_i |g(x_i)| p_i \psi \dot{a}.$
- (2) 若C.R.V. X的分布密度为 $f_X(x)$,

$$\Rightarrow E[g(X)] = \int_{-\infty}^{+\infty} g(x)f(x)dx$$
. 若此式绝对收敛。

- 2. 设g(x,y)为函数,Z=g(X,Y)是R.V.,
- (1) 若D.R.V. (X,Y)的分布列为 $P(X = x_i, Y = y_j) = p_{ij}$, $\Rightarrow E[g(X,Y)] = \sum_{i} \sum_{j} g(x_i, y_j) p_{ij}, \quad \text{若左式绝对收敛}.$
- (2) 若C.R.V. (X,Y)的分布密度为f(x,y),

$$\Rightarrow E[g(X,Y)] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x,y) f(x,y) dx dy.$$
 若左式绝对收敛。

$E[g(X)] = \sum_{i} g(x_i) p_i$



例 R.V.X的分布列为 $\frac{X}{P}$ 1/3 1/6 1/2 求E(X+1), E(X²).

解
$$E(X+1) = (-1+1) \times \frac{1}{3} + (0+1) \times \frac{1}{6} + (1+1) \times \frac{1}{2} = \frac{7}{6}$$
.

$$E(X^2) = (-1)^2 \times \frac{1}{3} + 0^2 \times \frac{1}{6} + 1^2 \times \frac{1}{2} = \frac{5}{6}.$$





例设某种商品每周的需求量X~U[10,30],而经销商进货数量为区间[10,30]中的某一整数,经销商每销售一单位商品可获利500元,若供大于求则削价处理,每处理1单位商品亏损100元;若供不应求,则可从外部调剂供应,此时每销售1单位商品仅获利300元,为使商店所获利润的期望值不少于9280元,试确定最少进货量。

解:设进货量为a,则利润额为

$$g(X) = \begin{cases} 500 X - 100 \cdot (a - X) = 600 X - 100 a & \mathbf{10} \le X < a \\ 500 a + 300 \cdot (X - a) = 300 X + 200 a & a < X \le \mathbf{30} \end{cases}$$

$$E[g(X)] = \int_{-\infty}^{+\infty} g(x) f(x) dx = \int_{10}^{30} g(x) \frac{1}{30 - 10} dx$$

$$= \int_{10}^{a} \frac{1}{20} (600x - 100a) dx + \int_{a}^{30} \frac{1}{20} (300x + 200a) dx$$

$$= -7.5a^{2} + 350a + 5250 \ge 9280 \Rightarrow 20\frac{2}{3} \le a \le 26 \Rightarrow a = 21 \quad 13$$



例设二维随机变量(X,Y)的联合分布律为

求随机变量 $Z = \sin \frac{\pi(X+Y)}{2}$ 的数学期望。

解:
$$E(Z) = E[\sin \frac{\pi(X+Y)}{2}]$$

解:
$$E(Z) = E[\sin \frac{\pi(X+Y)}{2}]$$

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

$$= \sin\frac{\pi(0+0)}{2} \times 0.1 + \sin\frac{\pi(0+1)}{2} \times 0.25 + \sin\frac{\pi(0+2)}{2} \times 0.15$$

$$+\sin\frac{\pi(1+0)}{2}\times0.15+\sin\frac{\pi(1+1)}{2}\times0.2+\sin\frac{\pi(1+2)}{2}\times0.15=0.25.$$



例 (X,Y)在区域A上服从均匀分布,A是由 $x + \frac{y}{2} = 1$,

x=0, y=0围成, 求E(X), E(Y), E(X+Y), E(XY).

解
$$f(x,y) = \begin{cases} 1, & 0 < x < 1, 0 < y < 2(1-x) \\ 0, & 其它 \end{cases}$$

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

$$E(Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} y f(x,y) dx dy = \int_0^2 \int_0^{1-\frac{y}{2}} y dx dy = \frac{2}{3}$$

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

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

$$E(XY) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xyf(x,y)dxdy = \int_{0}^{2} dy \int_{0}^{1-\frac{y}{2}} xydx = \int_{0}^{2} \frac{y}{2} (1-\frac{y}{2})^{2} dx = \frac{1}{6} \neq \frac{1}{3} \times \frac{2}{3}$$

四 数学期望的性质

- 1 E(c)=c, c为常数。
- 2 线性性: 対 $a,b \in \mathbb{R}$, E(aX+bY)=aE(X)+bE(Y)
 - 一般, $E(\sum_{i} k_i X_i) = \sum_{i} k_i E(X_i)$.
- 3 X与Y独立 $\Rightarrow E(XY) = E(X)E(Y)$.
- 一般, 若 X_1 , …, X_n 独立, $E(X_1 … X_n) = E(X_1) … E(X_n)$,
- 4 单调性: *X*≥0,则*EX*≥0.
- 故 若 $X_1 \ge X_2$, $E(X_1) \ge E(X_2)$; $E|X| \ge |EX|$.
- 5 Cauchy-Schwarz不等式: $E^2(XY) \leq E(X^2)E(Y^2)$.



例 $X \sim B(n, p)$, 求E(X).

解:
$$X_i = \begin{cases} 1, & \text{\hat{x}} \text{$\hat{x}$$

则 $X_i \sim (0,1)$ 分布, $E(X_i) = p$,

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





 $i = 1, 2, \dots, n$.

例 X~超几何分布,求E(X).

解:
$$X_i = \begin{cases} 1, & \text{\hat{x}} \text{$\hat{x}$$

则
$$P(X_i=1)=M/N$$
, $E(X_i)=M/N$,

$$X = \sum_{i=1}^{n} X_i \Rightarrow E(X) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} \frac{M}{N} = n \frac{M}{N}.$$





例 流水作业线上生产的每一个产品不合格的概率为p,当生产出k个不合格产品时即停工检修一次,求两次检修之间产品总数X的期望。

解:记 X_i 为第i-1个次品出现后开始,到第i个次品生产为止,生产产品的数目, $i=1,2,\cdots,k$.

则 X_i ~几何分布, $E(X_i)=1/p$,

$$X = \sum_{i=1}^{k} X_i \Rightarrow E(X) = \sum_{i=1}^{k} E(X_i) = \sum_{i=1}^{k} \frac{1}{p} = \frac{k}{p}.$$





小结

▶ 根据期望的定义,性质和定理,熟练 掌握随机变量及其函数的期望的计算

▶ 熟记常见随机变量的期望

4.2 随机变量的方差(Variance)



- 一方差的概念和计算
- 二方差的性质
- 三常见分布的方差



一方差的概念和计算



定义 若 $E(X-EX)^2$ 存在,则称其为R.V.X的方差。

记为 $DX=E(X-EX)^2$.

 \sqrt{DX} 称为X的标准差。

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

= $EX^2 - 2(EX)EX + (EX)^2$



二 方差的性质



1. 切比雪夫不等式

设R.V.X 的期望和方差均存在,则对任意 $\varepsilon>0$,有

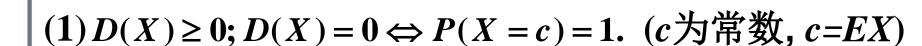
$$P(|X-EX|\geq \varepsilon) \leq \frac{DX}{\varepsilon^2}$$

$$\Leftrightarrow P(|X-EX|<\varepsilon)\geq 1-\frac{DX}{\varepsilon^2}.$$

2. 任意实数c, $E(X-c)^2 \ge E(X-EX)^2 = DX$.



3. 方差的基本性质



(2)
$$D(cX) = c^2 D(X)$$
.

(3)
$$D(X+Y) = D(X) + D(Y) + 2E[(X-EX)(Y-EY)]$$

特别,
$$D(X + Y)$$
 X与Y独立 $D(X) + D(Y)$ $D(X+c)=DX$

思考: 任意R.V.X与常数c独立。

$$(4)X,Y$$
 $\stackrel{\text{def}}{=}$ $D(XY) = D(X)D(Y) + (EY)^2D(X) + (EX)^2D(Y).$



三 常用分布的方差

(1)常见R.V.的方差

例 $X \sim B(n, p)$, D(X) = npq.

例 $X \sim P(\lambda)$, $D(X) = \lambda$.

例 X~几何分布, $D(X)=q/p^2$.

例 $X \sim U[a,b]$, $D(X) = (b-a)^2/12$.

例 $X \sim E(\lambda)$, $D(X) = 1/\lambda^2$.

例 $X \sim N(\mu, \sigma^2)$, $D(X) = \sigma^2$.



(2)常用R.V.函数的期望和方差

例 R.V.X,
$$E(X) = \mu$$
, $D(X) = \sigma^2$. X 的标准化 $Y = \frac{X - \mu}{\sigma}$ 则 $EY = 0$, $DY = 1$.

例 设
$$X_i$$
 ($i=1,2,...,n$)独立同分布($i.i.d.$),且 $EX_i = \mu$,
$$D(X_i) = \sigma^2 \cdot \quad \diamondsuit \overline{X} = \frac{1}{n} \sum_{i=1}^n X_i, 则$$

$$E\overline{X} = \underline{\mu} \quad , \quad D\overline{X} = \underline{\sigma}^2 \quad .$$

4.3 随机变量的矩

一 矩的定义 $\mathbf{R.V.} X$,若 $E|X|^k < \infty$ 。则称

 EX^k —— k阶原点矩

 $E/X/^k$ —— k阶原点绝对矩

若 $E|X-EX|^k<\infty$. 则称

 $E(X-EX)^k$ —— k阶中心矩

 $E/X-EX/^k$ —— k阶中心绝对矩

例
$$X \sim N(\mu, \sigma^2)$$
, 则 $E(X-EX)^k = \begin{cases} 0, & k$ 为奇数,
$$\sigma^k(k-1)!!, k$$
为偶数.

4.4 协方差和相关系数



$$D(X+Y)=D(X)+D(Y)+2E[(X-EX)(Y-EY)].$$

若X,Y独立,则E(X-EX)(Y-EY)=0.

问题: 若上式成立, X,Y是否独立?



一 协方差(covariance)



1.定义 R.V.(X,Y), 若E|(X-EX)(Y-EY)|存在,则称 E(X-EX)(Y-EY)

为(X,Y)的协方差,记为Cov(X,Y).

 $\stackrel{\text{?}}{=}$ 1. Cov(X,Y) = E(XY) - (EX)(EY).

2. $Cov(X,Y) = 0 \Leftrightarrow X,Y$ 不相关。

独立。文不相关

反例: $X\sim N(0,1), Y=X^2.X, Y$ 不相关,且不独立。



2. 协方差的性质 (设以下式子均存在)

- (1) Cov(X,Y)=Cov(Y,X)
- (2) Cov(X+Y,Z)=Cov(X,Z)+Cov(Y,Z)
- (3) Cov(aX,bY)=abCov(X,Y)
- (4)X,Y独立 $\Longrightarrow Cov(X,Y)=0$
- (5)D(X+Y)=DX+DY+2Cov(X,Y)

$$D(k_0 + k_1 X_1 + \dots + k_n X_n) = \sum_{i=1}^n D(k_i X_i) + \sum_{1 \le i < j \le n} 2Cov(k_i X_i, k_j X_j).$$

(6)
$$Cov(X,Y) = 0 \Leftrightarrow E(XY) = EX \cdot EY$$

 $\Leftrightarrow D(X+Y) = DX + DY$

二相关系数



1. 定义 R.V.(X,Y), 称下式 为X,Y的相关系数,

$$\rho_{XY} = \frac{Cov(X,Y)}{\sqrt{DX}\sqrt{DY}}$$
简记 ρ . 无量纲

注 $\rho=0$,称X,Y不相关。

- 2. 性质
- (1) $|\rho| \le 1$;
- (2) $|\rho|=1 \Leftrightarrow P(Y=aX+b)=1$, 其中a,b为常数。



例 设 $(X,Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$,求 ρ_{XY} .

解
$$Cov(X,Y) = \int_{v=\frac{y-\mu_{2}}{\sigma_{2}}}^{+\infty} \int_{-\infty}^{+\infty} \frac{\sigma_{1}u\sigma_{2}v}{2\pi\sqrt{1-\rho^{2}}} e^{-\frac{u^{2}-2\rho uv+v^{2}}{2(1-\rho^{2})}} du dv$$

$$= \int_{-\infty}^{+\infty} \frac{\sigma_{1}\sigma_{2}v}{\sqrt{2\pi}} e^{\frac{v^{2}}{2}} \left[\int_{-\infty}^{+\infty} \frac{u}{\sqrt{2\pi}\sqrt{1-\rho^{2}}} e^{-\frac{(u-\rho v)^{2}}{2(\sqrt{1-\rho^{2}})^{2}}} du \right] dv$$

$$= \int_{-\infty}^{+\infty} \frac{\sigma_{1}\sigma_{2}}{\sqrt{2\pi}} (\rho v) v e^{-\frac{v^{2}}{2}} dv = \sigma_{1}\sigma_{2}\rho$$

$$\therefore \rho_{XY} = \frac{\sigma_{1}\sigma_{2}\rho}{\sigma_{1}\sigma_{2}} = \rho.$$

注 对二维正态分布, X 与Y 独立 \longleftrightarrow X 与Y 不相关

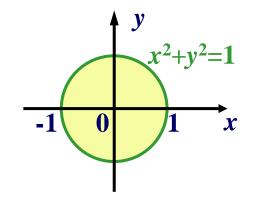


例 设(X,Y)的概率密度为
$$f(x,y) = \begin{cases} 1/\pi, & x^2 + y^2 \le 1, \\ 0, & \sharp 它, \end{cases}$$

X,Y是否相关?

$$\mathbf{F}$$
 $Cov(X,Y) = EXY - EXEY$

$$EX = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x f(x, y) dx dy$$
$$= \int_{-1}^{1} dy \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} x \cdot \frac{1}{\pi} dx = 0.$$



同理,EY=0.

$$EXY = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xyf(x,y) dxdy = \int_{-1}^{1} dy \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} xy \cdot \frac{1}{\pi} dx = 0.$$

$$\therefore Cov(X,Y)=0(\rho=0)$$
,故 X 与 Y 不相关.

■ 相关系数的性质的证明:

- 1. $|\rho_{XY}| \leq 1$
- 2. $|\rho_{XY}|=1 \Leftrightarrow$ 存在常数a,b,使P(Y=a+bX)=1

证明:考虑以X的线性函数a+bX来近似表示Y.

以均方误差
$$e(a,b) = E\{[Y - (a+bX)]^2\}$$
来衡量。

e(a,b)越小,a+bX与Y的近似程度越好。

下面来求最佳近似式: $e(a_0,b_0) = \min_{a,b} e(a,b)$

计算得:
$$e(a,b) = E(Y^2) + b^2 E(X^2) + a^2 - 2bE(XY) + 2abE(X) - 2aE(Y)$$

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

此时
$$e(a_0,b_0) = E\{[Y - (a_0 + b_0 X)]^2\}$$

= $D[Y - (a_0 + b_0 X)] + \{E[Y - (a_0 + b_0 X)]\}^2$

$$= D(Y) - \frac{[Cov(X,Y)]^2}{D(X)} = (1 - \rho_{XY}^2)D(Y)$$

$$a_0 = E(Y) - b_0 E(X)$$

$$b_0 = \frac{Cov(X, Y)}{D(X)}$$

$$\boxplus e(a_0, b_0) \ge 0 \Rightarrow 1 - \rho_{XY}^2 \ge 0 \Rightarrow |\rho_{XY}| \le 1$$

 $= D(Y - b_0 X) = D(Y) + b_0^2 D(X) - 2b_0 Cov(X, Y)$

2.
$$|\rho_{XY}| = 1 \Leftrightarrow E\{[Y - (a_0 + b_0 X)]^2\} = 0$$

$$\Leftrightarrow D[Y - (a_0 + b_0 X)] = 0 \perp E[Y - (a_0 + b_0 X)] = 0$$

$$\Leftrightarrow P\{Y - (a_0 + b_0 X) = 0\} = 1$$





小结

- 根据方差的定义,性质和定理,熟练掌握随机变量的方差的计算
- ▶ 熟记常见随机变量的方差
- 根据协方差及相关系数的定义,性质,熟练 掌握协方差(相关系数)的计算,会判断随机 变量的相关性
- → 记住随机变量矩的定义



一. 填空题

练



1. - 整数等可能地在1到10中取值,以<math>X记除得尽这一整数的正整数的个数,则EX = , DX= 。

2.设总体 $X \sim E(\lambda), X_1, X_2, \cdots, X_n$ 为总体X的样本, $\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i, 则 E(\overline{X}^2) = \underline{\hspace{1cm}}$



二. 选择题

练

1. 设 $X\sim N(3,4)$, $Y\sim E(0.2)$,则下列结论错误的是(B)

A.
$$E(X + Y) = 8$$

A.
$$E(X+Y) = 8$$
 B. $D(X+Y) = 29$

$$C.E(X^2 + Y^2) = 63$$

$$C.E(X^2 + Y^2) = 63$$
 $D.E(\frac{X}{2} + \frac{Y}{5} - \frac{5}{2}) = 0$

2. 设R.V.X与Y独立同分布, 记*U=X-Y,V=X+Y*, 则U

A.不独立

B.独立

C.相关系数为0

D.相关系数不为0



三 解答题

练

1.(10') 设随机变量(X,Y)在矩形G={(x,y): 0<x<2, 0<y<1}上服从均匀分布,记

$$U = \{ egin{aligned} 0, & \ddot{\Xi}X \leq Y, \\ 1, & \ddot{\Xi}X > Y, \end{aligned} & V = \{ egin{aligned} 0, & \ddot{\Xi}X \leq 2Y, \\ 1, & \ddot{\Xi}X > 2Y, \end{aligned} \end{cases}$$



(1)求(U,V)的联合分布; (2)求U与V的相关系数。

(提示: D.R.V.的分布的计算,相 关系数的计算)



解答题

练

2.(8')设随机变量(X,Y)的联合概率密度函数为

$$f(x,y) = \{ 1/\pi, x^2 + y^2 \le 1, \\ 0,$$
其它。

问*X*与*Y*是否相互独立,是否不相关?并说明理由。

(提示:二维C.R.V.的边缘密度的计算,独立性的证明)



解答题

练

3.(10')设 $X_1, X_2, \dots, X_n (n > 2)$ 为来自总体N(0,1)的样本, \bar{X} 为样本均值,记 $Y_i = X_i - \bar{X}, i = 1, \dots, n$. 求 DY_i 和 $Cov(Y_1, Y_n)$.

Z

$$Cov(X_i, \bar{X}) = \frac{DX_i}{n}$$
.

(提示: 方差和协方差的计算)