第二次习题课

助教：黄晨凌, 谢心乔

2023 年 12 月 17 日

1 第 8 次作业

1.1 随机变量数字特征

- 方差和矩：方差的定义与性质；中心距与原点矩；
- 协方差和相关系数：定义与联系；独立性与相关性；
- 其他数字特征：平均绝对差；矩母函数；特征函数；

1.2 作业：Ch4 33,34,49,52,62,63,69

33. 设二维随机变量 (X, Y) 服从二维正态分布 $N\left(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho\right)$, 其中 $\mu_1=\mu_2=1, \sigma_1^2=\sigma_2^2=0.5, \rho=0.5$, 记

$$Z=|X-Y|, \quad U=\max(X, Y), \quad V=\min(X, Y).$$

(1) 求 Z 的密度函数与期望 $\mathbb{E}[Z]$; (2) 分别求数学期望 $\mathbb{E}[U]$ 和 $\mathbb{E}[V]$.

证明. (1)

法一： 密度函数变换法

$$\begin{aligned} f(x, y) &= \frac{1}{2 \pi \sigma_1 \sigma_2 \sqrt{1-\rho^2}} \exp \left\{-\frac{1}{2\left(1-\rho^2\right)}\left[\frac{\left(x-\mu_1\right)^2}{\sigma_1^2}-\frac{2 \rho\left(x-\mu_1\right)\left(y-\mu_2\right)}{\sigma_1 \sigma_2}+\frac{\left(y-\mu_2\right)^2}{\sigma_2^2}\right]\right\} \\ &= \frac{2}{\sqrt{3} \pi} \exp \left\{-\frac{4}{3}\left[(x-1)^2-(x-1)(y-1)+(y-1)^2\right]\right\}. \end{aligned}$$

$$令 \left\{\begin{array}{l} S=X-Y \\ T=Y \end{array} \Rightarrow\left\{\begin{array}{l} X=S+T \\ Y=T \end{array} \Rightarrow J=\left|\begin{array}{ll} 1 & 1 \\ 0 & 1 \end{array}\right|=1 .\right.\right.$$

$$f(s,t)=f_{X,Y}(s+t,t)\cdot|J|=\frac{2}{\sqrt{3}\pi}\exp\left\{-\frac{4}{3}\left[\left(t+\frac{s-2}{2}\right)^2+\frac{3s^2}{4}\right]\right\}$$

$$f_S(s)=\int_{-\infty}^{\infty}f(s,t)d=\frac{2}{\sqrt{3}\pi}e^{-s^2}\int_{-\infty}^{\infty}e^{-\frac{4}{3}\left(t+\frac{s-2}{2}\right)^2}dt$$

$$=\frac{2}{\sqrt{3}\pi}e^{-t^2}\frac{\sqrt{3}\cdot\sqrt{\pi}}{2}=\frac{1}{\sqrt{\pi}}e^{-t^2},t\in\mathbb{R}.$$

$S\sim N\left(0,\frac{1}{2}\right)$. 而 $Z=|S|$

$$\begin{aligned}F_z(z)=P(z\leqslant z)&=P(|S|\leqslant z)\quad(z<0\text{ 时为 }0)\\&=P(-z\leqslant S\leqslant z)\\&=P\left(-z/\sqrt{\frac{1}{2}}\leqslant S/\sqrt{\frac{1}{2}}\leqslant z/\sqrt{\frac{1}{2}}\right)\\&=\Phi(\sqrt{2}z)-\Phi(-\sqrt{2}z)=2\Phi(\sqrt{2}z)-1.\end{aligned}$$

$$f_z(z)=F_z'(z)=2\Phi'(\sqrt{2}z)\cdot\sqrt{2}=2\sqrt{2}\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac{(\sqrt{2}z)^2}{2}}=\frac{2}{\sqrt{\pi}}e^{-z^2}\cdot\mathbb{1}_{\{z\geqslant 0\}}.$$

$$EZ=\int_0^{\infty}zf_z(z)dz=\int_0^{\infty}\frac{2z}{\sqrt{\pi}}e^{-z^2}dz=\frac{1}{\sqrt{\pi}}\int_0^{\infty}e^{-z^2}dz^2$$

$$=\frac{1}{\sqrt{\pi}}-e^{-z^2}\Big|_0^{\infty}=\frac{1}{\sqrt{\pi}}.$$

法二： 正态分布性质：⋆ 正态分布的线性组合仍然为正态分布

故，计算 $E(X-Y)=0$, $Var(X-Y)=Var(X)+Var(Y)-2Cov(X,Y)=0.5$, $S=X-Y\sim N(0,0.5)$.

$$EZ=\int_{-\infty}^{\infty}|\,s\,|\,f_S(s)ds=\int_0^{\infty}\frac{2s}{\sqrt{\pi}}e^{-s^2}ds=\frac{1}{\sqrt{\pi}}.$$

(2)

$$U=\max\{X,Y\}=\frac{X+Y+|\,X-Y\,|}{2}\quad V=\min\{X,Y\}=\frac{X+Y-|X-Y|}{2}$$

$$EU=1+\frac{1}{2\sqrt{\pi}}\quad EV=1-\frac{1}{2\sqrt{\pi}}.$$

□

34. 假设有 $n(n\geqslant 3)$ 个不同的盒子与 m 个相同的小球, 每个小球独立地以概率 p_k 落入

第 k 个盒子 $(k=1,2,\ldots,n)$. 分别以 $X_1, X_2, \ldots, X_n$ 表示落入各个盒子的球数. 试求

(1) $\mathbb{E}\left[X_2 \mid X_1=k\right]$ 和 $\operatorname{Var}\left(X_2 \mid X_1=k\right)$.

(2) $\mathbb{E}\left[X_1+X_2\right]$ 和 $\operatorname{Var}\left(X_1+\cdots+X_k\right), k=1,\ldots,n$.

证明. 注：1、多项分布的边际分布为多项分布；2、多项分布的条件分布为多项分布

(1)

$$P\left(X_2=x \mid X_1=k\right)=\frac{P\left(X_2-x, X_1-k\right)}{P\left(X_1-k\right)}=\frac{(m-k)!}{x!(m-x-k)!}\left(\frac{p_2}{1-p_1}\right)^x\left(\frac{1-p_1-p_2}{1-p_1}\right)^{m-x-k}$$

$$\begin{aligned}
& X_2 \mid X_1=k \sim B\left(m-k, \frac{p_2}{1-p_1}\right) \\
& E\left[X_2 \mid X_1=k\right]=(m-k) \cdot \frac{p_2}{1-p_1} \\
& \operatorname{Var}\left(X_2 \mid X_1=k\right)=(m-k) \frac{p_2}{1-p_1}\left(1-\frac{p_2}{1-p_1}\right)
\end{aligned}$$

(2)

$$\begin{aligned}
X_i \sim B\left(m, p_i\right) & \\
E\left[X_1+X_2\right] & =E X_1+E X_2=m \cdot p_1+m \cdot p_2=m\left(p_1+p_2\right) \\
P\left(X_1+X_2=k\right) & =\sum_{x=0}^k P\left(X_1=x_1 x_2=k-x\right) \\
& =\sum_{x=0}^k \frac{m!}{x!(k-x)!(m-k)!} p_1^x p_2^{k-x}\left(1-p_1-p_2\right)^{m-k} \\
& =\frac{m!}{(m-k)!}\left(-p_1-p_2\right)^{m-k} \sum_{x=0}^k \frac{1}{x!(k-x)!} p_1^x p_2^{k-x} \\
& =\frac{m!}{(m-k)!\cdot k!}\left(1-p_1-p_2\right)^{m-k} \sum_{x=0}^k \frac{k!}{x!(k-x)!} p_1^x p_2^{k-x} \\
& =\frac{m!}{(m-k)!k!}\left(1-p_1-p_2\right)^{m-k}\left(p_1+p_2\right)^k
\end{aligned}$$

$X_1+X_2 \sim B\left(m, p_1+p_2\right)$

类似我们有 $x_1+\cdots+x_l \sim B\left(m, p_1+\cdots+p_l\right)$

$$\operatorname{Var}\left(X_1+\cdots+X_k\right)=m \cdot\left(p_1+\cdots+p_k\right)\left(1-p_1-\cdots-p_k\right).$$

□

49. 设随机变量 X 与 Y 相互独立, 且 $X \sim N(3,2), Y \sim U(1,2), Z = 2X$, 令 $W = X - Y + Z - 1$, 则 $\mathrm{Var}(W) =$

证明. 注：方差的性质

$$\begin{aligned}
&\mathrm{Var}(W) = \mathrm{Var}(X - Y + Z - 1) = \mathrm{Var}(3X - Y - 1) \\
&= \mathrm{Var}(3X) + \mathrm{Var}(-(Y+1)) + 2\,\mathrm{Cov}(3X, -(Y+1)) \\
&= 9\,\mathrm{Var}(X) + \mathrm{Var}(Y) + 0 = 9 \cdot 2 + \frac{1}{12} = \frac{217}{12}
\end{aligned}$$

□

52. 设 X_1 和 X_2 是独立的指数随机变量, 均值分别为 1 和 2 . 定义

$$Y = \min\{X_1, X_2\} \text{ 和} Z = \max\{X_1, X_2\}.$$

求 (1)$\mathbb{E}[Y]$ 和 $\mathbb{E}[Z]$; (2) $\mathrm{Var}(Y)$ 和 $\mathrm{Var}(Z)$.

证明.

$$P(Y > y) = P(\min\{X_1, X_2\} > y) = P(X_1 > y, X_2 > y) = e^{-y} \cdot e^{-\frac{y}{2}} = e^{-\frac{3}{2}y}, \quad y > 0$$

故，$Y \sim \exp(3/2)$

$$P(Z \leqslant z) = P(X_1 \leqslant z, X_2 \leqslant z) = (1 - e^{-z})(1 - e^{-\frac{3z}{2}}), \quad z > 0$$

故，$f_Z(z) = e^{-z} + 1/2e^{-z/2} - 3/2e^{-\frac{3z}{2}}$.

$$\begin{aligned}
EY &= \int_0^\infty y \cdot \frac{3}{2} e^{-\frac{3}{2}y} dy = \frac{2}{3} \int_0^\infty y e^{-y} dy \\
&= \frac{2}{3}(-1 - y)e^{-y}\Big|_0^\infty = \frac{2}{3}(0 - (-1)) = \frac{2}{3} \\
EZ &= \int_0^\infty z \cdot \left(e^{-z} + \frac{1}{2}e^{-\frac{z}{2}} - \frac{3}{2}e^{-\frac{3}{2}z}\right) dz \\
&= 1 + 2 - \frac{2}{3} = \frac{7}{3}
\end{aligned}$$

$$\begin{aligned}
\operatorname{Var} Y &= \left(\frac{2}{3}\right)^2 = \frac{4}{9} \\
EZ^2 &= \int_0^\infty z^2 \left(e^{-z} + \frac{1}{2}e^{-\frac{z}{2}} - \frac{3}{2}e^{-\frac{3}{2}z}\right) dz \\
&= \frac{82}{9} \\
\operatorname{Var} Z &= EZ^2 - (EZ)^2 = \frac{82}{9} - \left(\frac{7}{3}\right)^2 = \frac{33}{9} = \frac{11}{3}
\end{aligned}$$

□

62. 掷两颗均匀骰子, 以 X 表示第一颗骰子郑出的点数, Y 表示两颗骰子所郑出的点数中的最大值.

(1) 求 X, Y 的数学期望与方差.

(2) 求 $\operatorname{Cov}(X, Y)$.

证明. 62. (1)

$$\begin{aligned}
EX &= \sum_{i=1}^{6} i \cdot \frac{1}{6} = \frac{1}{6}(1+2+3+4+5+6) = \frac{7}{2} \\
EX^2 &= \sum_{i=1}^{6} i^2 \cdot \frac{1}{6} = \frac{1}{6}(1+4+9+16+25+36) = \frac{91}{6} \\
\operatorname{Var} X &= EX^2 - (EX)^2 = \frac{35}{12}
\end{aligned}$$

Y	1	2	3	4	5	6
P	$\frac{1}{36}$	$\frac{1}{12}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{1}{4}$	$\frac{11}{36}$

$$EY = 1 \cdot \frac{1}{36} + 2 \cdot \frac{1}{12} + 3 \cdot \frac{5}{36} + 4 \cdot \frac{7}{36} + 5 \cdot \frac{1}{4} + 6 \cdot \frac{11}{36} = \frac{161}{36}$$

$$EY^2 = 1 \cdot \frac{1}{36} + 4 \cdot \frac{1}{12} + 9 \cdot \frac{5}{36} + 16 \cdot \frac{7}{36} + 25 \cdot \frac{1}{4} + 36 \cdot \frac{11}{36} = \frac{791}{36}$$

$$\operatorname{Var} Y = EY^2 - (EY)^2 = \frac{791}{36} - \left(\frac{161}{36}\right)^2 = \frac{2555}{1296}$$

(2)

$$
\begin{aligned}
E(XY) &= \sum_{i=1}^{6}\sum_{j=1}^{6} \max\{i,j\}\cdot i\cdot\frac{1}{36} \\
&= \sum_{i=1}^{6}\left(\sum_{j=1}^{i} i^2\cdot\frac{1}{36}+\sum_{j=i+1}^{6} ij\cdot\frac{1}{36}\right) \\
&= \frac{1}{36}\sum_{i=1}^{6}\left(i^3+i\cdot\frac{(6+i+1)(6-i)}{2}\right) \\
&= \frac{1}{36}(1+20+8+36+27+45+64+44+125+30+216+0) \\
&= \frac{616}{36}=\frac{154}{9} \\
\operatorname{Cov}(X,Y) = E(XY)-EXEY &= \frac{154}{9}-\frac{7}{2}\cdot\frac{161}{36}=\frac{105}{72}=\frac{35}{24}
\end{aligned}
$$

□

63. 设随机变量 X,Y 相互独立, 具有共同分布 $N(\mu,\sigma^2)$. 设 α,β 为两个常数.

(1) 求 $\operatorname{Cov}(\alpha X+\beta Y,\alpha X-\beta Y)$.

(2) 当 α,β 取何值时, $\alpha X+\beta Y$ 与 $\alpha X-\beta Y$ 相互独立.

证明. (1) $\operatorname{Cov}(\alpha X+\beta Y,\alpha X-\beta Y)=\alpha^2\operatorname{Var}X-\beta^2\operatorname{Var}Y=(\alpha^2-\beta^2)\,\sigma^2$.

(2) $\alpha^2-\beta^2=0\Leftrightarrow(\alpha-\beta)(\alpha+\beta)=0\Leftrightarrow\alpha=\beta$ 或 $\alpha+\beta=0$ 当 $\operatorname{Cov}(\alpha X+\beta Y\ \alpha X-\beta Y)=0$ 时，两者相互独立. □

69. 投资组合是将总资本按一定比例分配于各种投资, 以分散和降低风险, 所谓风险通常以方差来度量. 现假设某两种投资的回报率 X,Y 都是随机变量, 投资的风险 (即方差) 为 $\operatorname{Var}(X)=\operatorname{Var}(Y)=\sigma^2$. 假设 $\rho_{XY}=-0.5$, 即两种投资呈负相关. 记投资组合中两种投资的比例分别为 π 和 $1-\pi$, 则投资组合的回报率为 $Z=\pi X+(1-\pi)Y$.

(1) 试证明该投资组合 Z 的风险小于将所有资本投资于其中一个的风险.

(2) 求使得投资组合风险最小的分配比例 π.

证明. (1)

$$\begin{aligned}
\operatorname{Var} Z &= \operatorname{Cov}(\pi X+(1-\pi)Y, \pi X+(1-\pi)Y) \\
&= \operatorname{Var}(\pi X)+2\operatorname{Cov}(\pi X,(1-\pi)Y)+\operatorname{Var}((1-\pi)Y) \\
&= \pi^2 \operatorname{Var} X+2\pi(1-\pi)\operatorname{Cov}(X,Y)+(1-\pi)^2 \operatorname{Var} Y \\
&= \pi^2\cdot\sigma^2+2\pi\cdot(1-\pi)(-0.5)\cdot\sigma^2+(1-\pi)^2\sigma^2 \\
&= \left(\pi^2+(1-\pi)^2-\pi(1-\pi)\right)\sigma^2=\left(3\pi^2-3\pi+1\right)\sigma^2
\end{aligned}$$

由于，$\pi\in[0,1]$. Var 在 0,1 上取最大。此时 $\operatorname{Var} Z=\sigma^2$, 故 $\operatorname{Var} Z<\sigma^2=\operatorname{Var} X=\operatorname{Var} Y\quad,\quad \pi\in(0,1)$.

(2) $\operatorname{Var} Z=(3\pi^2-3\pi+1)\,\sigma^2$. 最小值在 $\pi=\frac{1}{2}$ 处取到. 当 $\pi=\frac{1}{2}$ 时 $\operatorname{Var} Z=\frac{1}{4}\sigma^2$ □

2 第九次作业

2.1 大数定理与中心极限定理

- 大数定理
- 中心极限定理

2.2 作业：Ch4 81,83；Ch5 4,9,15,16

81 (1) 设随机变量 X 与 Y 独立, 均服从泊松分布, 参数分别为 λ 与 μ. 对任何给定的非负整数 $k\leqslant m$, 求 $\mathbb{P}(X=k\mid X+Y=m)$ 及 $\mathrm{E}(X\mid X+Y=m)$.

(2) 设随机变量 X 与 Y 独立, 均服从二项分布 $B(n,p)$, 对任何给定的非负整数 $k\leqslant m$, 求 $\mathbb{P}(X=k\mid X+Y=m)$ 及 $\mathbb{E}(X\mid X+Y=m)$.

证明. 注：1、独立 Poisson 分布和仍为 Poisson 分布；2、独立二项分布和仍然为二项分布

(1)

$X \sim P(\lambda), Y \sim P(\mu). X, Y$ 独立 $\therefore X+Y \sim P(\lambda+\mu)$

$$\begin{aligned} P(X=k \mid X+Y=m) &= \frac{P(X=k, X+Y=m)}{P(X+Y=m)} \\ &= \frac{P(X=k)P(Y=m-k)}{P(X+Y=m)} \\ &= \frac{m!}{k!(m-k)!} \cdot \left(\frac{\lambda}{\lambda+\mu}\right)^k \cdot \left(\frac{\mu}{\lambda+\mu}\right)^{m-k}, k=0,\ldots,m \end{aligned}$$

故 $X \mid X+Y=m \sim B(m, \frac{\lambda}{\lambda+\mu})$, 则 $E(X \mid X+Y=m) = \frac{m\lambda}{\lambda+\mu}$

(2)

$$X, Y \text{ 独立. } X, Y \sim B(n,p) \quad \therefore X+Y \sim B(2n,p)$$

$$\begin{aligned} P(X=k \mid X+Y=m) &= P(X=k, X+Y=m) \\ &= \frac{P(X=k) \cdot P(Y=m-k)}{P(X+Y=m)} \\ &= \frac{C_n^k \cdot C_n^{m-k}}{C_{2n}^m} \quad, k=0,\ldots,m \end{aligned}$$

易知 $P(Y=k \mid X+Y=m) = P(X=k \mid X+Y=m)$

$$\therefore E[Y \mid X+Y=m] = E[X \mid X+Y=m]$$

而 $E[X+Y \mid X+Y=m] = m = E[X \mid X+Y=m] + E[Y \mid X+Y=m]$
$\therefore E[X \mid X+Y=m] = \frac{m}{2}$ □

83 设随机变量 X 与 Y 分别是均值为 $1/\lambda$ 和 $1/\mu$ 的独立的指数随机变量.

(1) 证明在条件 $X > Y$ 下, 随机变量 $\min\{X,Y\}$ 和 $X-Y$ 是相互独立的.

(2) 证明对任意正数 $c > 0$,

$$\mathbb{E}[\min\{X,Y\} \mid X > Y+c] = \mathbb{E}[\min\{X,Y\} \mid X > Y] = \mathbb{E}[\min\{X,Y\}] = \frac{1}{\lambda+\mu}$$

证明. X, Y 独立. $\quad f(x)=\lambda e^{-\lambda x} I(x>0) \quad g(y)=\mu e^{-\mu y} I(y>0)$. (1)

令 U $=\min\{X, Y\}, \quad V=X-Y$.

$$F_U(u \mid V>0)=P(U \leqslant u \mid V>0) \quad =1-\frac{\int_u^{+\infty} \int_y^{+\infty} f(x) g(y) d x d y}{\int_0^{\infty} \int_y^{+\infty} f(x) g(y) d x d y}$$

$$=1-P(U>u \mid V>0)$$

$$=1-P(Y>u \mid V>0)=1-\frac{\frac{\mu}{\lambda+\mu} e^{-(\lambda+\mu) u}}{\frac{\mu}{\lambda+\mu}}=1-e^{-(\lambda+\mu) u} \quad(u \geqslant 0)$$

$$\begin{aligned}
& F_V(v \mid V>0)=P(V \leqslant v \mid V>0) \\
& =\frac{P(X-Y \leqslant v, X-Y>0)}{P(X-Y>0)} \\
& =\frac{\int_0^{\infty} \int_y^{y+v} f(x) g(y) d x d y}{P(v>0)}=\frac{\frac{\mu}{\lambda+\mu}\left(1-e^{-\lambda v}\right)}{\frac{\mu}{x+\mu}} \\
& \qquad =1-e^{-\lambda v} \quad(v \geqslant 0)
\end{aligned}$$

$$\begin{aligned}
& F_{U, V}(u, v \mid V>0)=P(U \leqslant u, V \leqslant v \mid V>0) \\
& =P(V \leqslant v \mid V>0)-P(U>u, V \leqslant v \mid V>0) \\
& =1-e^{-\lambda^v}-\frac{P(Y>u, V \leqslant v, V>0)}{P(V>0)} \\
& =1-e^{-\lambda v}-\frac{\int_u^{\infty} \int_y^{y+v} f(x) g(y) d x d y}{P(v>0)} \\
& =1-e^{-\lambda v}-\frac{\frac{\mu}{\lambda+\mu}\left(1-e^{-\lambda v}\right) \cdot e^{-(\lambda+\mu) u}}{\frac{\mu}{\lambda+\mu}} \\
& =1-e^{-\lambda v}-\left(1-e^{-\lambda v}\right) e^{-(\lambda+\mu) u} \\
& =\left(1-e^{-\lambda v}\right)\left(1-e^{-(\lambda+\mu) u}\right) \\
& \therefore F_{U, v}(u, v \mid v>0)=F_v(u \mid v>0) \cdot F_v(v \mid v>0)
\end{aligned}$$

$\therefore$ 在$X>Y$ 下。$\min\{X, Y\}$ 与$X-Y$ 独立。

(2). 证明:

$$\begin{aligned} F_U(u \mid V>c) &= P(U \leqslant u \mid V>c) \quad \forall c>0. \\ &= 1-P(U>u \mid V>c) \\ &= 1-P(Y>u \mid V>c) \\ &= 1-\frac{P(Y>u, V>c)}{P(V>c)} \\ &= 1-\frac{\int_u^\infty \int_{c+y}^\infty f(x)g(y)dxdy}{\int_0^\infty \int_{c+y}^\infty f(x)g(y)dxdy} \\ &= 1-\frac{\frac{e^{-\lambda+\mu}}{\lambda+\mu}\cdot e^{-(\lambda t\mu)u}}{\frac{e^{-\lambda c\mu}}{\lambda+\mu}} \\ &= 1-e^{-(\lambda+\mu)u}(u \geqslant 0) \end{aligned}$$

因此 $min\{X,Y\}$ 与 $\min\{X,Y\} \mid V>0, \min\{X,Y\} \mid V>c$ 同分布服从 $\exp(\lambda+\mu)$.

$$\therefore E[\min\{X,Y\}] = \int_0^\infty u\cdot(\lambda+\mu)e^{-(\lambda+\mu)u}du = \frac{1}{\lambda+\mu}$$

$$\therefore E[\min\{X,Y\}] = E[\min\{X,Y\} \mid V>0] = E[\min\{X,Y\} \mid V>c]$$

$$= \frac{1}{\lambda+\mu}.$$

□

4 设 $(X_1,Y_1)\ldots,(X_n,Y_n),\ldots$ 是一列独立同分布的二维随机变量, 均值均为 2 , 方差均为 2 , 设对任意的 $n\in\mathbb{N}, \operatorname{Cov}(X_n,Y_n)=1$, 记

$$Z_n = \frac{X_1Y_1+\cdots+X_nY_n}{n}, n\in\mathbb{N}$$

则 Z_n 依概率收敛于

证明.

$$EXY = \operatorname{Cov}(X,Y)+EXEY = 1+2+2=5$$

由大数定律知: $Z_n \xrightarrow{P} EXY=5.$ □

9 设某高校共有 1000 人, 选课系统中设定一门课必须有 50 人选中方可开课, 假设每人

选择某一门课程的概率是 4%, 且每个学生选择与否相互独立, 求该课程可开课的概率.

证明. X : 选课：$X=1, p=0.04$ ．不选课：$X=0, 1-p=0.96$. $X\sim B(1,p)$. $X_1,\dots,X_{1000,\dots}\sim B(1,p)$ ，独立，$Y=\sum_{i=1}^{n}X_i\sim B(n,p)$.

由 CLT 知. $\frac{Y-EY}{\sqrt{\operatorname{Var} Y}}\xrightarrow{d}N(0,1)\quad(n\to\infty)$

$$P(\text{ 该课程可开课})=P\left(Y_{1000}\geqslant 50\right)=P\left(\frac{Y_{1000}-40}{\sqrt{38.4}}\geqslant\frac{50-40}{\sqrt{38.4}}\right)$$
$$=1-\Phi(1.61)=1-0.9463=0.0537.$$

□

15 某种计算机在进行加法时, 要对每个加数进行取整. 设每次取整的误差相互独立且服从 $(-0.5,0.5)$ 上的均匀分布.

(1) 若现要进行 1500 次加法运算, 求误差总和的绝对值超过 15 的概率.

(2) 若要保证误差总和的绝对值不超过 10 的概率不小于 0.90 , 至多只能进行多少次加法运算?

证明. (1) 读差 $X_i\sim U(0.5,0.5)$ $EX_i=0, \operatorname{Var} X_i=\frac{1}{12}$. $i=1,2,\dots$ 独立 $S_n=\sum_{i=1}^{n}X_i$

$$P\left(\left|S_n\right|>15\right)=1-P\left(\left|S_n\right|\leqslant 15\right)=2-2\Phi(15/\sqrt{(1500/12)})$$
$$=0.1802.$$

$(n=3000=1500*2$ 时 $P\left(\left|S_n\right|>15\right)=0.3244)$

(2)

$$P\left(|S_n| \leqslant 0\right) \geqslant 0.90$$

$$P\left(-\frac{10}{\sqrt{\frac{n}{12}}} \leqslant \frac{S_n}{\sqrt{\frac{n}{12}}} \leqslant \frac{10}{\sqrt{\frac{n}{12}}}\right) = 2\Phi\left(\frac{20\sqrt{3}}{\sqrt{n}}\right) - 1 \geqslant 0.90.$$

$$\Phi\left(\frac{20\sqrt{3}}{\sqrt{n}}\right) \geqslant 0.95. \quad \frac{20\sqrt{3}}{\sqrt{n}} \geqslant \frac{1.65+1.64}{2.}$$

$$n \leqslant 443.45$$

运算次数 $\leqslant 443(221)$

□

16 设某生产线上组装每件产品的时间服从指数分布, 平均需要 10 分钟, 且各产品的组装时间是相互独立的.

(1) 试求组装 100 件产品需要 15 小时至 20 小时的概率.

(2) 保证有 95% 的可能性, 问 16 小时内最多可以组装多少件产品.

证明. (1) $X_i \sim \mathrm{Exp}(\lambda)$. $\quad EX_i = 10 \quad \therefore \lambda = \frac{1}{10} \quad \operatorname{Var} X_i = 100$

$$S_n = \sum_{i=1}^{n} X_i \quad P\left(15\times 60 \leqslant S_n \leqslant 20\times 60\right) = P\left(\frac{900-1000}{100} \leqslant \frac{S_n-1000}{100} \leqslant \frac{1200-1000}{100}\right)$$

$$= P\left(-1 \leqslant \frac{S_n-1000}{100} \leqslant 2\right) = \Phi(2) - \Phi(-1) = 0,8185$$

(2)

$$P\left(S_n \leqslant 16\times 60\right) \geqslant 0.95$$

$$P\left(\frac{S_n-10n}{10\sqrt{n}} \leqslant \frac{960-10n}{10\sqrt{n}}\right) = P\left(\frac{S_n-10n}{10\sqrt{n}} \leqslant \frac{96}{\sqrt{n}} - \sqrt{n}\right) \geqslant 0.95$$

$$\Phi\left(\frac{96}{\sqrt{n}} - \sqrt{n}\right) \geqslant 0.95.$$

$$\frac{96}{\sqrt{n}} - \sqrt{n} \geqslant 1.645 \quad n \leqslant 81.$$

最多组装 81 件。 □

3 第十次作业

3.1 抽样分布及点估计

- 三大抽样分布：F 分布，t 分布，χ^2 分布
- 重要的三个定理及四条推论
- 点估计：矩估计，极大似然估计

3.2 作业：Ch6 15,16,19；Ch7 2，3（3），4（2）（5），5

15. 设 X_1, X_2, X_3, X_4 是来自正态总体 $N(0, 2^2)$ 的简单随机样本, 令 $T = a(X_1 - \ 2X_2)^2 + b(3X_3 - 4X_4)^2$. 试求 a, b 使统计量 T 服从 χ^2 分布.

证明. $\frac{1}{\sqrt{20}}(X_1 - 2X_2) \sim N(0,1) \quad \frac{1}{10}(3X_3 - 4X_4) \sim N(0,1)$ 且相互独立。
$\left(\frac{1}{\sqrt{20}}(X_1 - 2X_2)\right)^2 + \left(\frac{1}{10}(3X_3 - 4X_4)\right)^2 \sim \chi^2(2). \quad a = \frac{1}{20}, \quad b = \frac{1}{100}$ □

16. 设 $X_1, X_2, \ldots, X_9$ 为独立同分布的正态随机变量, 记

$$Y_1 = \frac{1}{6}(X_1 + \ldots + X_6), \quad Y_2 = \frac{1}{3}(X_7 + X_8 + X_9), \quad S^2 = \frac{1}{2}\sum_{i=7}^{9}(X_i - Y_2)^2.$$

试求 $Z = \frac{\sqrt{2}(Y_1 - Y_2)}{S}$ 的分布.

证明. $X_1, X_2, \ldots, X_9 \overset{iid}{\sim} N(\mu, \sigma^2)$.

$$Y_1 = \frac{X_1 + \cdots + X_6}{6}, Y_2 = \frac{1}{3}(X_7 + X_8 + X_9), S^2 = \frac{1}{2}\sum_{i=1}^{9}(X_i - Y_2)^2.$$

Y_1 与 Y_2 独立, Y_1 与 S^2 独立, Y_2 与 S^2 独立。

$$Y_1 \sim N\left(\mu, \frac{\sigma^2}{6}\right), \quad Y_2 \sim N\left(\mu, \frac{\sigma^2}{3}\right). \quad \frac{2S^2}{\sigma^2} \sim X_2^2$$

$$Y_1 - Y_2 \sim N\left(0, \frac{\sigma^2}{2}\right) \text{ 且} Y_1 - Y_2, S^2 \text{ 独立。}$$

$$\frac{\sqrt{2}}{\sigma}(Y_1-Y_2)\sim N(0,1).$$

$$\frac{\frac{\sqrt{2}(Y_1-Y_2)}{\sigma}}{\sqrt{\frac{2S^2}{\sigma^2}/2}}=\frac{\sqrt{2}(Y_1-Y_2)}{S}\sim t_2$$

□

19. 设 $X_1,\ldots,X_n$ 是从两点分布 $B(1,p)$ 中抽取的简单样本, $0<p<1$, 记 $\bar{X}$ 为样本均值, 求 $S_n^2=\frac{1}{n}\sum_{i=1}^n\left(X_i-\bar{X}\right)^2$ 的期望.

证明.

$$\begin{aligned}
ES_n^2&=\frac{1}{n}E\left[\sum_{i=1}^n\left(X_i-\bar{X}\right)^2\right]=\frac{1}{n}E\left[\sum_{i=1}^n X_i^2-2\sum_{i=1}^n X_i\bar{X}+n\bar{X}^2\right]\\
&=\frac{1}{n}E\left[\sum_{i=1}^n X_i^2-n\bar{X}^2\right]=EX_1^2-E\bar{X}^2\\
&=\left(1-\frac{1}{n}\right)EX_1^2-\frac{1}{n^2}\sum_{i\neq j}EX_iEX_j=\left(1-\frac{1}{n}\right)p-\frac{n^2-n}{n^2}p^2=\left(1-\frac{1}{n}\right)\cdot p(1-p)
\end{aligned}$$

□

2. 设总体 X 的概率分布如右表, 其中 $0<p_1,p_2<1$ 为未知参数. 现从此总体中抽出一样本量为 n 的简单随机样本, 其中 1 出现了 n_1 次, 2 出现了 n_2 次, 3 出现了 n_3 次. 试求 p 的矩估计.

X	1	2	3
P	p_1	p_2	$1-p_1-p_2$

证明.

$$\left.\begin{aligned}
EX&=p_1+2p_2+3(1-p_1-p_2)=3-2p_1-p_2.\\
EX^2&=p_1+4p_2+9(1-p_1-p_2)=9-8p_1-5p_2
\end{aligned}\right\}\Rightarrow\left\{\begin{aligned}
p_1&=\tfrac{EX^2-5EX+6}{2}\\
p_2&=4EX-EX^2-3
\end{aligned}\right.$$

$$a_1=\bar{X}=\frac{n_1+2n_2+3n_3}{n}$$

$$a_2=\frac{n_1+4n_2+9n_3}{n}$$

$$\therefore\left\{\begin{aligned}
\hat{p}_1&=\tfrac{a_2-5a_1+6}{2}\\
\hat{p}_2&=4a_1-a_2-3
\end{aligned}\right.\Rightarrow\left\{\begin{aligned}
\hat{p}_1&=\tfrac{n_1}{n}\\
\hat{p}_2&=\tfrac{n_2}{n}.
\end{aligned}\right.$$

□

3(3). 设 $X_1, \ldots, X_n$ 是总体 X 的一个简单随机样本, 试求在 X 具有下列概率分布时参数 θ 的矩估计.

(3) $p(x;\theta) = (x-1)\theta^2(1-\theta)^{x-2}, \quad x = 2, 3, \ldots; 0 < \theta < 1.$

证明.

$$
\begin{aligned}
&P(x;\theta) = (x-1)\theta^2(1-\theta)^{x-2} \cdot x = 2, 3, \ldots, 0 < \theta < 1 \\
&EX = \sum_{k=2}^{\infty} k \cdot P(x=k) = \sum_{k=2}^{\infty} k \cdot (k-1)\theta^2(1-\theta)^{k-2} \\
&\quad = \theta^2 \sum_{k=2}^{\infty} k \cdot (k-1)(1-\theta)^{k-2} = \theta^2 \left(\sum_{k=2}^{\infty}(1-\theta)^k\right)'' = \theta^2 \cdot 2\theta^{-3} = \frac{2}{\theta} \\
&\hat{\theta} = \frac{2}{\bar{X}}, \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i
\end{aligned}
$$

□

4. 设 $X_1, \ldots, X_n$ 是总体 X 的一个简单随机样本, 试求在 X 具有下列概率密度时参数 θ 的矩估计.

(2) $f(x;\theta) = \begin{cases} (\theta+1)x^{\theta}, & 0 < x < 1, \theta > 0 \\ 0 & \text{其他.} \end{cases}$

(5) $p(x;\theta) = \begin{cases} \frac{6x}{\theta^3}(\theta - x), & 0 < x < \theta \\ 0 & \text{其他} \end{cases}$

证明. (2)

$$
\begin{aligned}
&EX = \int_0^1 (\theta+1)x^{\theta+1}dx = \frac{\theta+1}{\theta+2} \\
&\hat{\theta} = \frac{1}{1-\bar{X}} - 2
\end{aligned}
$$

(5).

$$
\begin{aligned}
EX &= \int_0^\theta \frac{6}{\theta^3} x^2(\theta - x)dx \\
&\overset{t=\frac{x}{\theta}}{=} 6\theta \int_0^1 t^2(1-t)dt = 6\theta\,\mathrm{Beta}(3.2) = \frac{\theta}{2}.
\end{aligned}
$$

$\hat{\theta} = 2\bar{X}.$

□

5. 总体 X 的概率密度函数为

$$
f(x) = \begin{cases} \frac{4x^2}{\theta^3\sqrt{\pi}} e^{-\frac{x^2}{\theta^2}} & x \geqslant 0 \\ 0 & \text{其他}. \end{cases}
$$

设 $X_1, X_2, \ldots, X_n$ 是取自总体 X 的简单随机样本.

(1) 求 θ 的矩估计量 $\hat{\theta}$.

(2) 求 $\hat{\theta}$ 的方差.

证明. (1)

$$
\begin{aligned}
EX &= \int_{-\infty}^{\infty} x f(x) dx = \int_0^\infty x \cdot \frac{4x^2}{\theta^3\sqrt{\pi}} e^{-\frac{x^2}{\theta^2}} dx \\
&= \int_0^\infty \frac{2x^2}{\theta^3\sqrt{\pi}} e^{-\frac{x^2}{\theta^2}} dx^2 = \int_0^\infty \frac{2t}{\theta^3\sqrt{\pi}} e^{-\frac{t}{\theta^2}} dt \\
&= \frac{2\theta}{\sqrt{\pi}}.
\end{aligned}
$$

$$
\theta = \frac{\sqrt{\pi}EX}{2}
$$

$$
\hat{\theta} = \frac{\sqrt{\pi}\bar{X}}{2}, \bar{X} = \frac{1}{n}\sum_{i=1}^n X_i
$$

(2)

$$
\begin{aligned}
\operatorname{Var}(\hat{\theta}) &= \operatorname{Var}\left(\frac{\sqrt{\pi}}{2}\cdot\frac{1}{n}\sum_{i=1}^{n}x_i\right)\\
&= \frac{\pi}{4n^2}n\operatorname{Var}X = \frac{\pi}{4n}\operatorname{Var}X\\
EX^2 &= \int_{-\infty}^{\infty}x^2 f(x)dx\\
&= \int_0^{\infty}x^2\cdot\frac{4x^2}{\theta^3\sqrt{\pi}}e^{-\frac{x^2}{\theta^2}}dx\\
&\overset{u=\frac{x^2}{\theta^2}}{=} \int_0^{\infty}\frac{4\theta t^4}{\sqrt{\pi}}e^{-t^2}\cdot\theta dt\\
&= \frac{2\theta^2}{\sqrt{x}}\int_0^{\infty}u^{\frac{3}{2}}e^{-u}du\\
&= \frac{2\theta^2}{\sqrt{\lambda}}\Gamma\left(\frac{5}{2}\right)\\
&= \frac{2\theta^2}{\sqrt{\lambda}}\cdot\frac{3}{4}\sqrt{\lambda} = \frac{3}{2}\theta^2\\
\operatorname{Var}X &= EX^2-(EX)^2 = \frac{3}{2}\theta^2-\frac{4\theta^2}{\pi} = \left(\frac{3}{2}-\frac{4}{\pi}\right)\theta^2\\
\operatorname{Var}(\hat{\theta}) &= \frac{\pi}{4n}\cdot\left(\frac{3}{2}-\frac{4}{\pi}\right)\theta^2 = \left(\frac{3\pi}{8n}-\frac{1}{n}\right)\theta^2 = \left(\frac{3\pi}{8}-1\right)\frac{\theta^2}{n}
\end{aligned}
$$

□

4 补充

1. 设随机变量 $X \sim N(0,1)$, 而对任一实数 x, 在 $X=x$ 条件下, $Y \sim N(x,1)$.

(1) 试求随机变量 Y 的密度函数 $f_Y(y)$, 并指出 Y 服从何种分布.

(2) 试求条件期望 $\mathrm{E}[XY \mid X=x]$.

(3) 试求 X 和 Y 的相关系数.

(4) 试求常数 a, 使得随机变量 $aX+Y$ 和 $aX-Y$ 相互独立.

2. 设一列随机变量 $Y_1, Y_2, \ldots, Y_n$ 满足

$$Y_i = \beta x_i + \varepsilon_i, \quad i=1,2,\ldots,n,$$

其中 $x_1, x_2, \ldots, x_n$ 为给定非负常数且不全相等, $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ 相互独立且均服从正态分布 $N(0, \sigma^2)$, 而 β 和 σ^2 为两个未知参数.

(1) 根据 $Y_1, Y_2, \ldots, Y_n$ 的分布, 请写出似然函数 $L(\beta, \sigma^2)$.

(2) 试求 β 的极大似然估计量 $\hat{\beta}$, 并证明 $\hat{\beta}$ 为 β 的一个无偏估计.

(3) 证明 $\hat{\beta}^* = \sum_{i=1}^n Y_i / \sum_{i=1}^n x_i$ 也为 β 的一个无偏估计, 并比较 $\hat{\beta}$ 和 $\hat{\beta}^*$ 哪个更有效.

注:x_i 为常数，根据正态分布线性和仍然为正态分布可以得到 Y_i 为独立不同分布的正态分布。

3. 设随机变量 X 服从参数为 λ 的指数分布, 则它的熵 $H(X) =$

4. 设二维随机向量 (X, Y) 服从平面上以 $(0,0), (0,1)$ 和 $(1,1)$ 为顶点的三角形区域上的均匀分布.

(1) 求随机向量 (X, Y) 的概率密度函数 $f(x, y)$.

(2) 设随机变量 $U, V \sim U(0,1)$ 且相互独立, 以记号 $\stackrel{d}{=}$ 表示同分布, 证明:

$$X \stackrel{d}{=} \min\{U, V\}, \quad Y \stackrel{d}{=} \max\{U, V\}.$$

(3) 对任一 $0 < x < 1$, 在条件 $X = x$ 下, Y 是否服从均匀分布? 请通过计算说明.

(4) 求随机变量 X 和 Y 的相关系数 $\rho_{X,Y}$.