样题二解答

一 填空题（每空 3 分，共 30 分；答案均写在试卷上，注意标清题号）

1. 事件 A, B 相互独立, 且 $P\left(\bar{A}\bar{B}\right)=P\left(AB\right)=0.24$, $P\left(A\right)\le P\left(B\right)$, 则 $P(A)=$ ____________________。

$$P\left(\bar{A}\bar{B}\right)=P\left(AB\right)\Rightarrow\left(1-P\left(A\right)\right)\left(1-P\left(B\right)\right)=P\left(A\right)P\left(B\right)\Rightarrow P\left(A\right)+P\left(B\right)=1$$

$$P\left(A\right),P\left(B\right)=0.4,0.6\Rightarrow P\left(A\right)=0.4$$

2. 随机变量 $X\sim Ge\left(\frac{1}{3}\right)$, Y 服从参数为 $\left(8,\frac{1}{3}\right)$ 的负二项分布, 已知 $E\left(X^2\right)=15$, 则 $Var\left(X\right)=$ ____________,

$E\left(X^2\middle|X>2\right)=$ ____________, $E\left(X^2\middle|2<X<5\right)=$ ____________, $Var\left(Y\right)=$ ____________。

$$Var\left(X\right)=E\left(X^2\right)-E\left(X\right)^2=15-9=6$$

$$E\left(X^2\middle|X>2\right)=E\left(\left(X+2\right)^2\right)=E\left(X^2\right)+4E\left(X\right)+4=15+12+4=31$$

$$E\left(X^2\middle|2<X<5\right)=\frac{9P\left(X=3\right)+16P\left(X=4\right)}{P\left(X=3\right)+P\left(X=4\right)}=\frac{9p\left(1-p\right)^2+16p\left(1-p\right)^3}{p\left(1-p\right)^2+p\left(1-p\right)^3}=\frac{9+16\left(1-p\right)}{1+1-p}=\frac{9+\frac{32}{3}}{5/3}=\frac{59}{5}$$

$$Var\left(Y\right)=Var\left(Y\right)=Var\left(X_1+\cdots+X_8\right)=8Var\left(X\right)=48$$

3. 将一枚均匀的硬币独立地抛掷 100 次，记正面次数为 X，利用中心极限定理估计 $P\left(40\le X\le 60\right)\approx$ ____________，利用切比雪夫不等式得到关于 $P\left(40\le X\le 60\right)$ 的估计为____________________。

$$X\sim b\left(100,0.5\right),\quad X\overset{\bullet}{\sim}N\left(50,5^2\right)$$

$$P\left(40\le X\le 60\right)=P\left(\frac{40-50}{5}\le\frac{X-50}{5}\le\frac{60-50}{5}\right)\approx\Phi\left(2\right)-\Phi\left(-2\right)=2\Phi\left(2\right)-1$$

$$P\left(40\le X\le 60\right)=P\left(39.5\le X\le 60.5\right)=P\left(\frac{39.5-50}{5}\le\frac{X-50}{5}\le\frac{60.5-50}{5}\right)\approx\Phi\left(2.1\right)-\Phi\left(-2.1\right)=2\Phi\left(2.1\right)-1$$

$$P\left(40\le X\le 60\right)=P\left(\left|X-50\right|\le 10\right)=1-P\left(\left|X-50\right|>10\right)\ge 1-\frac{25}{10^2}=\frac{3}{4},\quad P\left(40\le X\le 60\right)\ge\frac{3}{4}$$

4. 从正态总体 $N\left(100,8^2\right)$ 中抽取样本容量为 16 的样本，样本均值为 $\overline{X}$，若 $P\left(\left|\overline{X}-100\right|<k\right)=0.90$，$k=$ ________。

$\bar{X} \sim N\left(100, 2^2\right), \quad P\left(\left|\bar{X}-100\right|<k\right)=P\left(\left|\frac{\bar{X}-100}{2}\right|<\frac{k}{2}\right)=0.90 \Rightarrow \frac{k}{2}=u_{0.95} \Rightarrow k=2u_{0.95}$

5. X_1, X_2 为来自正态总体 $N\left(0, \sigma^2\right)$ 的样本，则 $P\left(X_1+X_2 \leq \left|X_1-X_2\right|\right)=$ ____________________________。

$$X_1+X_2 \sim N\left(0, 2\sigma^2\right), X_1-X_2 \sim N\left(0, 2\sigma^2\right)$$

$$Cov\left(X_1+X_2, X_1-X_2\right)=0,$$

$$P\left(X_1+X_2 \leq \left|X_1-X_2\right|\right)=P\left(\frac{X_1+X_2}{\sqrt{2}\sigma} \leq \left|\frac{X_1-X_2}{\sqrt{2}\sigma}\right|\right)$$

$$Y=\frac{\frac{X_1+X_2}{\sqrt{2}\sigma}}{\left|\frac{X_1-X_2}{\sqrt{2}\sigma}\right|} \sim t(1), \quad P\left(X_1+X_2 \leq \left|X_1-X_2\right|\right)=P\left(Y \leq 1\right)=F_{t(1)}(1)=0.75$$

$$Z=\frac{\left(\frac{X_1+X_2}{\sqrt{2}\sigma}\right)^2}{\left(\frac{X_1-X_2}{\sqrt{2}\sigma}\right)^2} \sim F(1,1), \quad P\left(\left|X_1+X_2\right| \leq \left|X_1-X_2\right|\right)=P\left(Z \leq 1\right)$$

$$\begin{aligned} P\left(\left|X_1+X_2\right| \leq \left|X_1-X_2\right|\right) &= P\left(-\left|X_1-X_2\right| \leq X_1+X_2 \leq \left|X_1-X_2\right|\right) \\ &= P\left(X_1+X_2 \leq \left|X_1-X_2\right|\right)-P\left(X_1+X_2<-\left|X_1-X_2\right|\right) \\ &= 2P\left(X_1+X_2 \leq \left|X_1-X_2\right|\right)-1=P\left(Z \leq 1\right) \end{aligned}$$

$$\Rightarrow P\left(X_1+X_2 \leq \left|X_1-X_2\right|\right)=\frac{P\left(Z \leq 1\right)+1}{2}=\frac{F_Z(1)+1}{2}=0.75$$

6. $X_1, \cdots, X_n$ 为来自正态总体 $N\left(0, \sigma^2\right)$ 的样本，样本均值 $\bar{X}=\frac{X_1+\cdots+X_n}{n}$，样本方差为 $S^2=\frac{1}{n-1}\sum_{k=1}^{n}\left(X_k-\bar{X}\right)^2$，已知 $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$，且 $E\left(X_1^{\,4}\right)=3\sigma^4$，则 $E\left(\left(S^2-\sigma^2\right)^2\right)=$ ________________________。

$$E\left(\left(S^2-\sigma^2\right)^2\right)=Var\left(S^2\right), \quad Var\left(\frac{(n-1)S^2}{\sigma^2}\right)=Var\left(Y_1^{\,2}+Y_{n-1}^{\,2}\right)=(n-1)Var\left(Y_1^{\,2}\right)=2(n-1),$$

$$Var\left(\frac{(n-1)S^2}{\sigma^2}\right)=2(n-1)=\frac{(n-1)^2}{\sigma^4}Var\left(S^2\right)=2(n-1) \Rightarrow Var\left(S^2\right)=\frac{2\sigma^4}{n-1}$$

二. (10 分) 钥匙掉了，掉在宿舍里、掉在教室里、掉在路上的概率分别是 50%, 30%和 20%，而掉在上述三处地方被找到的概率分别为 0.8, 0.3 和 0.1。

（1）试求找到钥匙的概率；

（2）若钥匙已找到，求最初是掉在教室里的概率。

解：设掉在宿舍里、掉在教室里、掉在路上分别为事件 A_1, A_2, A_3，找到为事件 B

（1）$P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + P(B|A_3)P(A_3) = 0.8\times0.5 + 0.3\times0.3 + 0.1\times0.2 = 0.51$

（2）$P(A_2|B) = \dfrac{P(B|A_2)P(A_2)}{P(B)} = \dfrac{0.3\times0.3}{0.51} = \dfrac{9}{51} = \dfrac{3}{17}$

三. (8 分)随机变量 X 的密度函数为 $p(x) = \frac{1}{2}e^{-|x|}$，$x\in R$，求随机变量 $Y = |X|$ 的分布函数、密度函数和期望。

当 $y<0$ 时，$F_Y(y) = 0$,

当 $y\geq0$ 时，$F_Y(y) = P(Y\leq y) = P(|X|\leq y) = P(-y\leq X\leq y) = \int_{-y}^{y}\frac{1}{2}e^{-|x|}dx$

$$= \int_0^y e^{-x}dx = 1 - e^{-y}$$

当 $y<0$ 时，$p_Y(y) = 0$,　　　当 $y\geq0$ 时，$p_Y(y) = \dfrac{dF_Y(y)}{dy} = e^{-y}$

$E(Y) = \int_{-\infty}^{+\infty} p_Y(y)dy = \int_0^{+\infty} e^{-y}dy = 1$。

四. (10 分)随机变量 X_1 以等可能取值为 0 和 1，X_2 以等可能取值为 $0,1,2$，X_1 和 X_2 相互独立

（1）求 $Y_1 = X_1 - 2X_2$，$Y_2 = X_1 + 2X_2$ 的联合分布；

（2）计算相关系数 $\rho(Y_1, Y_2)$。

（1）$(X_1 = 0, X_2 = 0) \to (Y_1 = 0, Y_2 = 0), (X_1 = 0, X_2 = 1) \to (Y_1 = -2, Y_2 = 2)$

$(X_1 = 0, X_2 = 2) \to (Y_1 = -4, Y_2 = 4), (X_1 = 1, X_2 = 0) \to (Y_1 = 1, Y_2 = 1)$

$(X_1 = 1, X_2 = 1) \to (Y_1 = -1, Y_2 = 3), (X_1 = 1, X_2 = 2) \to (Y_1 = -3, Y_2 = 5)$

$P(Y_1 = 0, Y_2 = 0) = \dfrac{1}{6}$，$P(Y_1 = -2, Y_2 = 2) = \dfrac{1}{6}$

$P(Y_1=-4,Y_2=4)=\frac{1}{6}$， $P(Y_1=1,Y_2=1)=\frac{1}{6}$

$P(Y_1=-1,Y_2=3)=\frac{1}{6}$， $P(Y_1=-3,Y_2=5)=\frac{1}{6}$

（2） $E(X_1)=\frac{1}{2},E(X_1{}^2)=\frac{1}{2}$， $E(X_2)=1,E(X_2{}^2)=\frac{5}{3}$

$$E(X_1X_2)=\frac{1}{6}(1+2)=\frac{1}{2}$$

$E(Y_1)=E(X_1-2X_2)=-\frac{3}{2}$， $E(Y_2)=E(X_1+2X_2)=\frac{5}{2}$

$Cov(Y_1,Y_2)=E\left((X_1-2X_2)(X_1+2X_2)\right)-E(Y_1)E(Y_2)=E(X_1{}^2)-4E(X_2{}^2)+\frac{15}{4}=-\frac{29}{12}$

$Var(Y_1)=E\left((X_1-2X_2)^2\right)-\frac{9}{4}=E(X_1{}^2)+4E(X_2{}^2)-4E(X_1,X_2)=\frac{43}{6}-2-\frac{9}{4}=\frac{35}{12}$

$Var(Y_2)=E\left((X_1+2X_2)^2\right)-\frac{25}{4}=\frac{43}{6}+2-\frac{25}{4}=\frac{35}{12}$

$$\rho(Y_1,Y_2)=\frac{Cov(Y_1,Y_2)}{\sqrt{Var(Y_1)Var(Y_2)}}=\frac{-\frac{29}{12}}{\sqrt{\frac{35}{12}\frac{35}{12}}}=-\frac{29}{35}。$$

五.（10分）已知$(X,Y)\sim N(0,0,1,1,0)$，求(1) $E(X|X+Y)$；（2） $E(X^2|X+Y=1)$。

$Cov(X+Y,X-Y)=0$ ，所以X,Y相互独立，

（1） $E(X|X+Y)=E\left(\frac{X+Y}{2}+\frac{X-Y}{2}\middle|X+Y\right)=E\left(\frac{X+Y}{2}\middle|X+Y\right)+E\left(\frac{X-Y}{2}\middle|X+Y\right)=\frac{X+Y}{2}+0=\frac{X+Y}{2}$

或根据对称性$E(X|X+Y)=E(Y|X+Y)$， $E(X|X+Y)+E(Y|X+Y)=E(X+Y|X+Y)=X+Y$

所以$E(X|X+Y)=\frac{X+Y}{2}$

（2） $E(X^2|X+Y=1)=E\left(\left(\frac{X+Y}{2}+\frac{X-Y}{2}\right)^2\middle|X+Y=1\right)$

$$=E\left(\left(\frac{X+Y}{2}\right)^2+2\left(\frac{X+Y}{2}\right)\left(\frac{X-Y}{2}\right)+\left(\frac{X-Y}{2}\right)^2\middle|X+Y=1\right)$$

$$=\frac{1}{2}+2\frac{1}{2}E\left(\frac{X-Y}{2}\right)+E\left(\left(\frac{X-Y}{2}\right)^2\right)$$

$$=\frac{1}{2}+\frac{1}{4}\left(E\left(X^2\right)-2E\left(XY\right)+E\left(Y^2\right)\right)=\frac{1}{2}+\frac{1}{2}=1$$

六.（8 分）有一枚不均匀的硬币，设掷出正面的概率为 p，掷出反面的概率为 $1-p$。另有一枚公平的刻有 1-6 点的六面的骰子。首先抛掷硬币，直到首次出现正面停止，记抛掷次数为 N；然后将骰子抛掷 N 次，记 N 次抛掷的总点数为 S。求 $E(S)$ 和 $Var(S)$。

$$N \sim Ge(p),\quad S=X_1+X_2+\cdots+X_N,$$

对任意正整数 k，$X_k \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix}$，$E\left(X_k\right)=\frac{7}{2}, E\left(X_k{}^2\right)=\frac{91}{6}$

$$E(S)=E\left(\sum_{k=1}^{N}X_k\right)=E\left(E\left(\sum_{k=1}^{N}X_k \middle| N\right)\right)=\sum_{n=1}^{+\infty}E\left(\sum_{k=1}^{N}X_k \middle| N=n\right)P(N=n)$$

$$=\sum_{n=1}^{+\infty}E\left(\sum_{k=1}^{n}X_k\right)P(N=n)=\sum_{n=1}^{+\infty}\frac{7}{2}n\cdot P(N=n)=\frac{7}{2}E(N)=\frac{7}{2p}\text{。}$$

$$E\left(S^2\right)=E\left(E\left(S^2 \middle| N\right)\right)=\sum_{n=1}^{+\infty}E\left(\left(\sum_{k=1}^{N}X_k\right)^2 \middle| N=n\right)P(N=n)=\sum_{n=1}^{+\infty}E\left(\left(\sum_{k=1}^{n}X_k\right)^2\right)P(N=n)$$

$$E\left(\left(\sum_{k=1}^{n}X_k\right)^2\right)=E\left(\sum_{k=1}^{n}X_k{}^2\right)+E\left(\sum_{i\neq j,1\le i,j\le n}X_iX_j\right)=nE\left(X_1{}^2\right)+n(n-1)E\left(X_1X_2\right)$$

$$=n\cdot\frac{91}{6}+n(n-1)E\left(X_1\right)E\left(X_2\right)=\frac{91n}{6}+\frac{49n(n-1)}{4}$$

$$E\left(S^2\right)=\sum_{n=1}^{+\infty}\left(\frac{91n}{6}+\frac{49n(n-1)}{4}\right)P(N=n)=\left(\frac{91}{6}-\frac{49}{4}\right)E(N)+\frac{49}{4}E\left(N^2\right)$$

$$Var(S)=E\left(S^2\right)-E(S)^2=\left(\frac{91}{6}-\frac{49}{4}\right)E(N)+\frac{49}{4}E\left(N^2\right)-\frac{49}{4}E(N)^2$$

$$=\left(\frac{91}{6}-\frac{49}{4}\right)E(N)+\frac{49}{4}Var(N)=\left(\frac{91}{6}-\frac{49}{4}\right)\frac{1}{p}+\frac{49}{4}\left(\frac{1}{p^2}-\frac{1}{p}\right)=\frac{49}{4p^2}-\frac{28}{3p}$$

七．（12 分）设总体 $X \sim U(-\theta,\theta)$，参数 $\theta > 0$ 未知，$X_1, X_2, \cdots, X_n$ 是来自 X 的简单随机样本，

（1）设 $Y = \max(|X_1|,|X_2|,\cdots,|X_n|)$，证明 Y 的密度函数为 $p_Y(y) = \begin{cases} \dfrac{n}{\theta^n} y^{n-1}, & y \in [0,\theta] \\ 0, & 其他 \end{cases}$；

（2）求参数 θ 的矩估计量 $\hat{\theta}_1$ 和极大似然估计量 $\hat{\theta}_2$；

（3）判断估计量 $\hat{\theta}_1$ 和 $\hat{\theta}_2$ 是否无偏，如果不是无偏，是否可以做无偏校正。

（1）当 $y < 0$ 时，$F_Y(y) = 0$，$p_Y(y) = 0$；当 $y > \theta$ 时，$F_Y(y) = 1$；当 $0 \le y \le \theta$ 时，

$$F_Y(y) = P(T \le y) = P\left(\max(|X_1|,|X_2|,\cdots,|X_n|) \le y\right) = P\left(|X_1| \le y, |X_2| \le y, \cdots, |X_n| \le y\right)$$

$$= P(|X_1| \le y)^n = P(-y \le X_1 \le y)^n = \left(\frac{y}{\theta}\right)^n$$

$$p_Y(y) = \frac{dF_Y(y)}{dy} = \frac{n}{\theta^n} y^{n-1}$$

$$p_Y(y) = \begin{cases} \dfrac{n}{\theta^n} y^{n-1}, & y \in [0,\theta] \\ 0, & 其他 \end{cases}$$

（2）$Var(X) = \dfrac{\theta^2}{3}$，令 $s^2 = Var(X) = \dfrac{\theta^2}{3}$，解得参数 θ 的矩估计量 $\hat{\theta}_1 = \sqrt{3}s$

似然函数 $L(x_1,\cdots,x_n;\theta) = \prod_{k=1}^{n} p(x_k;\theta) = \left(\dfrac{1}{2\theta}\right)^n$，$-\theta \le x_1,\cdots,x_n \le \theta$

当 $\theta = \max(|x_1|,|x_2|,\cdots,|x_n|)$ 时，似然函数 $L(x_1,\cdots,x_n;\theta)$ 只达到最大

所以参数 θ 的极大似然估计量 $\hat{\theta}_2 = \max(|X_1|,|X_2|,\cdots,|X_n|)$。

（3）$E(\hat{\theta}_1)^2 = E(\sqrt{3}s)^2 = E(3s^2) - Var(\sqrt{3}s) < 3E(s^2) = \theta^2$

$\hat{\theta}_1$ 不是参数 θ 的无偏估计，且不易进行无偏校正。

$$E(\hat{\theta}_2) = E(Y) = \int_0^\theta y \frac{n}{\theta^n} y^{n-1} dy = \frac{n}{n+1}\theta$$

$\hat{\theta}_2$ 不是参数 θ 的无偏估计，可做无偏校正，$\dfrac{n+1}{n}\hat{\theta}_2$ 是参数 θ 的无偏估计。

八. (12 分) 设某工厂生产一种产品，它的一个指标参数服从正态分布 $N\left(\mu,3^2\right)$，μ 只能取整数，$\mu=10$ 为优级。利用样本 $X_1,\cdots,X_n$ 对参数 μ 做如下假设检验，$H_0:\mu=10$ VS $H_1:\mu\neq 10$，$\bar{x}$ 为检验统计量，显著性水平 $\alpha=0.1$。

（1）写出 $n=36$ 时，拒绝域的范围；（2）证明当样本容量无限增大时，第二类错误趋向于 0；

（3）计算 $n=36$ 条件下，$\bar{x}=11$ 的 p 值。

（1）$\mu=10$ 时，检验统计量 $\bar{x}\sim N\left(10,\left(\frac{1}{2}\right)^2\right)$

$$P\left(\left|\frac{\bar{x}-10}{\frac{1}{2}}\right|>u_{0.95}\right)=0.1$$，所以拒绝域为 $\left\{\bar{x}:\left|\bar{x}-10\right|>\frac{1}{2}u_{0.95}\right\}$

即 $\left\{\bar{x}:\bar{x}>10.82\text{或}\bar{x}<10.82\right\}$　或　$\left\{\bar{x}:\bar{x}>10+\frac{u_{0.95}}{2}\text{或}\bar{x}<10-\frac{u_{0.95}}{2}\right\}$

（2）样本容量为 n 时，拒绝域为 $\left\{\bar{x}:\left|\bar{x}-10\right|>\frac{3}{\sqrt{n}}u_{0.95}\right\}$，$\left\{\bar{x}:\bar{x}>10+\frac{3u_{0.95}}{\sqrt{n}}\text{或}\bar{x}<10-\frac{3u_{0.95}}{\sqrt{n}}\right\}$

当 $\mu=11$ 时，第二类错误 $\beta=P\left(10-\frac{3}{\sqrt{n}}u_{0.95}\leq\bar{x}\leq 10+\frac{3}{\sqrt{n}}u_{0.95}\,\middle|\,\mu=11\right)$

$$\beta=P\left(\frac{-1-\frac{3}{\sqrt{n}}u_{0.95}}{3/\sqrt{n}}\leq\frac{\bar{x}-11}{3/\sqrt{n}}\leq\frac{-1+\frac{3}{\sqrt{n}}u_{0.95}}{3/\sqrt{n}}\right)$$

$$=P\left(\frac{-\sqrt{n}}{3}-u_{0.95}\leq\frac{\bar{x}-11}{3/\sqrt{n}}\leq\frac{-\sqrt{n}}{3}+u_{0.95}\right)\to 0$$

（3）$p=P\left(\left|\bar{x}-10\right|\geq 1\right)=P\left(\left|\frac{\bar{x}-10}{\frac{1}{2}}\right|\geq 2\right)=2\left(1-\Phi\left(2\right)\right)\approx 0.05$

备注 1. 解答中标准正态随机变量的分布函数和密度函数分别可用$\Phi(x)$和$\varphi(x)$表示

备注 2. $\Phi(1.28)=0.9$, $\Phi(1.44)=0.925$, $\Phi(1.65)=0.95$, $\Phi(1.96)=0.975$, $\Phi(2.33)=0.99$

备注 3. 正态、χ^2、t等分布所需取值，均用（下侧）分位数表示，例如$X\sim t(n)$，则$P\left(X<t_{\alpha}(n)\right)=\alpha$

备注 4. $t_{0.75}(1)=1, t_{0.75}(2)=0.79, t_{0.8}(1)=1.38, t_{0.8}(2)=1.06, F_{0.5}(1,1)=1, F_{0.5}(1,2)=0.67, F_{0.75}(1,1)=5.83$