

清华大学试题专用纸

考试课程：概率论与数理统计 考试时间：2021 年 6 月 16 日 8: 00—10: 00

姓名_____学号 **20** 班级_____.

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. 18 世纪时，提出了一个关于投针的著名概率问题的学者是_____。

布丰，蒲丰，布冯，Buffon 等均可

2. 利用快速排序（quicksort）算法对序列“3, 2, 1”进行排序，则平均比较次数为_____。

解：以 3, 1 为基准平均比较次数为 3 次，以 2 为基准比较次数为 2，所以平均比较次数为 **8/3**。

3. 设随机变量 $X \sim Ge\left(\frac{1}{3}\right)$ ，则 $E\left(\left(\frac{4}{3}\right)^X\right) = \underline{\hspace{2cm}}$ 。

$$\begin{aligned}\text{解：} E\left(\left(\frac{4}{3}\right)^X\right) &= \sum_{k=1}^{\infty} \left(\frac{4}{3}\right)^k p(1-p)^{k-1} = \sum_{k=1}^{\infty} \left(\frac{4}{3}\right)^k \frac{1}{3} \left(\frac{2}{3}\right)^{k-1} \\ &= \frac{4}{3} \cdot \frac{1}{3} \sum_{k=1}^{\infty} \left(\frac{4}{3} \cdot \frac{2}{3}\right)^{k-1} = \frac{4}{3} \cdot \frac{1}{3} \cdot \frac{9}{9} \cdot \sum_{k=1}^{\infty} \frac{1}{9} \cdot \left(1 - \frac{1}{9}\right)^{k-1} = \frac{4}{3} \cdot \frac{1}{3} \cdot \frac{9}{9} \cdot \frac{1}{1 - \frac{1}{9}} = \frac{4}{3} \cdot \frac{1}{3} \cdot \frac{9}{8} = \frac{4}{3} \cdot \frac{3}{8} = \frac{1}{2}\end{aligned}$$

4. 设随机变量 $X \sim Exp(1)$ ， $Y \sim b\left(2, \frac{1}{3}\right)$ ， X, Y 相互独立，则 $P(XY \leq 2\ln 2) = \underline{\hspace{2cm}}$ 。

解：利用全概率公式

$$\begin{aligned}P(XY \leq 2\ln 2) &= P(Y=0)P(XY \leq 2\ln 2|Y=0) + P(Y=1)P(XY \leq 2\ln 2|Y=1) \\ &\quad + P(Y=2)P(XY \leq 2\ln 2|Y=2) \\ &= P(Y=0) + P(Y=1)P(X \leq 2\ln 2) + P(Y=2)P(X \leq \ln 2) \\ &= \left(1 - \frac{1}{3}\right)^2 + 2 \times \frac{1}{3} \times \left(1 - \frac{1}{3}\right) \times (1 - e^{-2\ln 2}) + \left(\frac{1}{3}\right)^2 \times (1 - e^{-\ln 2}) = \frac{5}{6}\end{aligned}$$

5. 若 Y_1 服从参数为 $(8, 1/2)$ 的负二项分布， Y_2 服从参数为 $(6, 1/2)$ 的负二项分布，且 Y_1 与 Y_2 的相关系数为 $-1/\sqrt{3}$ ，

则 $Var(Y_1) = \underline{\hspace{2cm}}$ ， $Var(Y_1 + Y_2) = \underline{\hspace{2cm}}$ 。

解： $Var(Y_1) = Var(X_1 + \dots + X_8) = 8Var(X) = \mathbf{16}$ ， $Var(Y_2) = 6Var(X) = 12$

$$\text{Var}(Y_1 + Y_2) = \text{Var}(Y_1) + \text{Var}(Y_2) + 2 \cdot \rho \cdot \sqrt{\text{Var}(Y_1) \cdot \text{Var}(Y_2)} = 16 + 12 - 2 \cdot 8 = 12$$

6. 从1,2两个数中等可能地随机取1个数记为 X ，再从 $1, \dots, X$ 中等可能地随机取1个数记为 Y 。则条件期望 $E(X|Y)$ 为_____，条件方差 $\text{Var}(Y|X)$ 为_____。

解： X, Y 的联合分布列为

$$P(X=1, Y=1) = \frac{1}{2}, P(X=1, Y=2) = 0, P(X=2, Y=1) = P(X=2, Y=2) = \frac{1}{4}$$

$$E(X|Y=1) = \frac{4}{3} \left(1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} \right) = \frac{4}{3}, E(X|Y=2) = \frac{4}{1} \left(1 \cdot 0 + 2 \cdot \frac{1}{4} \right) = 2,$$

$$E(X|Y) \sim \begin{pmatrix} 4/3 & 2 \\ 3/4 & 1/4 \end{pmatrix} \text{ 或 } \frac{2}{3}(Y+1),$$

$$X=1 \text{ 时, } Y \text{ 为常数 } 1, \text{ 所以 } \text{Var}(Y|X=1) = 0; Y|X=2 \sim \begin{pmatrix} 1 & 2 \\ 1/2 & 1/2 \end{pmatrix}, \text{Var}(Y|X=2) = \frac{5}{2} - \left(\frac{3}{2} \right)^2 = \frac{1}{4}$$

$$\text{Var}(Y|X) \sim \begin{pmatrix} 0 & 1/4 \\ 1/2 & 1/2 \end{pmatrix} \text{ 或 } \frac{1}{4}(X-1)$$

7. 设总体 $X \sim N(\mu, 0.5^2)$ ， X_1, X_2, \dots, X_n 为简单随机样本，要使 μ 的置信系数96%的双侧置信区间长度不超过0.4，则样本容量 n 至少要达到_____。

$$\text{解: } X \sim N(\mu, 0.5^2) \Rightarrow \bar{X} \sim N\left(\mu, \frac{0.5^2}{n}\right) \Rightarrow \frac{\bar{X} - \mu}{0.5/\sqrt{n}} \sim N(0, 1).$$

$$P\left\{\left|\frac{\bar{X} - \mu}{0.5/\sqrt{n}}\right| \leq u_{0.98}\right\} = 0.96 \Rightarrow P\left(\bar{X} - \frac{u_{0.98}}{2\sqrt{n}} \leq \mu \leq \bar{X} + \frac{u_{0.98}}{2\sqrt{n}}\right)$$

$$\text{置信区间长度为 } 2 \cdot \frac{u_{0.98}}{2\sqrt{n}} = \frac{2}{\sqrt{n}}, \text{ 令 } \frac{2}{\sqrt{n}} \leq 0.4, \text{ 得到 } n \geq 25$$

8. 对参数 p 做假设检验， $H_0: p \in \left(0, \frac{1}{2}\right], H_1: p \in \left(\frac{1}{2}, 1\right]$ ，检验统计量 $X \sim b(4, p)$ ，要实现真实水平为0.05的检验，拒绝域应为_____。

$$\text{解: 拒绝域的形式为 } \{X \geq c\}, \text{ 当 } p = \frac{1}{2} \text{ 时, } P(X=4) = \frac{1}{16} > 0.05,$$

$$\alpha(p) = E_p(\Phi(X)) \leq r \cdot P\left(X=4 \mid p = \frac{1}{2}\right) = r \cdot \frac{1}{16} = 0.05 \Rightarrow r = 0.8,$$

当 $X=4$ 时，以概率0.8拒绝原假设。

二. (10 分) 两台车床加工同样的零件, 第一台出现不合格品的概率是 0.1, 第二台出现不合格品的概率是 0.16, 加工出来的零件放在一起, 并且已知第一台车床加工的零件数是第二台加工的零件数的 2 倍。

(1) 求任取一个零件是合格品的概率; (2) 如果取出的零件是合格品, 求它是由第二台车床加工的概率。

解: 设事件 A 为第一台机床加工, 事件 B 为产品合格

$$(1) \text{ 由全概率公式 } P(B) = P(A)P(B|A) + P(\bar{A})P(B|\bar{A}) = \frac{2}{3} \times 0.90 + \frac{1}{3} \times (1 - 0.16) = 0.88$$

$$(2) \text{ 由贝叶斯公式 } P(\bar{A}|B) = \frac{P(B|\bar{A})P(\bar{A})}{P(B)} = \frac{\frac{1}{3} \times 0.84}{0.88} = \frac{28}{88} = \frac{7}{22}.$$

三. (12 分) 已知随机变量 $X \sim U(0, 2)$,

(1) 计算 $Var(X)$; (2) 计算 $E(\min\{X^2, 1\})$; (3) 计算 $Var(\min\{X^2, 1\})$

$$\text{解: } (1) E(X) = 1, \quad E(X^2) = \int_{-\infty}^{+\infty} x^2 p(x) dx = \int_0^2 x^2 \frac{1}{2} dx = \frac{4}{3}$$

$$Var(X) = E(X^2) - E(X)^2 = \frac{1}{3};$$

$$(2) E(\min\{X^2, 1\}) = \int_{-\infty}^{+\infty} \min\{x^2, 1\} \cdot p(x) dx = \int_0^2 \min\{x^2, 1\} \cdot \frac{1}{2} dx$$

$$= \int_0^1 x^2 \cdot \frac{1}{2} dx + \int_1^2 1 \cdot \frac{1}{2} dx = \frac{1}{6} + \frac{1}{2} = \frac{2}{3}$$

$$(3) E\left(\left(\min\{X^2, 1\}\right)^2\right) = \int_0^2 \left(\min\{x^2, 1\}\right)^2 \cdot \frac{1}{2} dx$$

$$= \int_0^1 x^4 \cdot \frac{1}{2} dx + \int_1^2 1 \cdot \frac{1}{2} dx = \frac{1}{10} + \frac{1}{2} = \frac{3}{5}$$

$$Var(\min\{X^2, 1\}) = E\left(\left(\min\{X^2, 1\}\right)^2\right) - E(\min\{X^2, 1\})^2 = \frac{3}{5} - \left(\frac{2}{3}\right)^2 = \frac{7}{45}.$$

四. (12 分) 设 X 、 Y 为相互独立的随机变量, 均服从 $Exp(1)$, $Z = 2X + Y$, $W = \frac{Y}{2X + Y}$,

(1) 求联合密度 $p_{Z,W}(z, w)$; (2) 求边缘密度 $p_W(w)$; (3) 求 X 、 Z 的协方差。

$$\text{解: } (1) p_{X,Y}(x, y) = e^{-x-y} I_{x>0, y>0}$$

$$\begin{cases} Z = 2X + Y \\ W = \frac{Y}{2X + Y} \end{cases} \Rightarrow \begin{cases} x = \frac{z - zw}{2}, \\ y = zw \end{cases}, \quad J = \frac{\partial(x, y)}{\partial(z, w)} = \begin{vmatrix} \frac{1-w}{2} & -z \\ w & z \end{vmatrix} = \frac{z}{2}$$

$$p_{Z,W}(z, w) = p_{X,Y}\left(\frac{z-zw}{2}, zw\right) \cdot |J| = e^{-\frac{(z-zw)^2}{2} - zw} \cdot \frac{z}{2} = \frac{1}{2} e^{-\frac{z-zw}{2}} \cdot z \cdot I_{z>0, 0<w<1}$$

$$(2) \quad p_W(w) = \int_0^{+\infty} p_{Z,W}(z, w) dz = \int_0^{+\infty} \frac{1}{2} e^{-\frac{z-zw}{2}} \cdot z \cdot dz$$

$$= \int_0^{+\infty} \frac{1}{2} e^{-\frac{(1+w)}{2}z} \cdot z \cdot dz = \frac{1}{2} \cdot \frac{2}{1+w} \int_0^{+\infty} z \cdot \frac{1+w}{2} \cdot e^{-\frac{(1+w)}{2}z} dz$$

$$= \frac{1}{2} \cdot \frac{2}{1+w} \cdot \frac{2}{1+w} = \frac{2}{(1+w)^2} I_{0<w<1}$$

$$(3) \quad \text{Cov}(X, Z) = \text{Cov}(X, 2X + Y) = 2\text{Cov}(X, X) + \text{Cov}(X, Y) = 2\text{Var}(X) = 2$$

五. (12分) 已知 X_1, X_2, \dots, X_n 为来自于总体 X 的简单随机样本, 总体 X 的概率分布函数为

$$F(x) = \begin{cases} e^{-\frac{x^2}{\theta}}, & x \leq 0 \\ 1, & x > 0 \end{cases}, \quad \text{其中 } \theta > 0 \text{ 为未知参数,}$$

(1) 求参数 θ 的矩估计量;

(2) 求参数 θ 的最大似然估计量.

$$\text{解: (1) 密度函数 } p(x) = F'(x) = \begin{cases} -\frac{2}{\theta} x e^{-\frac{x^2}{\theta}}, & x < 0 \\ 0, & \text{其他} \end{cases}$$

$$E(X) = \int_{-\infty}^{+\infty} x \cdot p(x) dx = \int_{-\infty}^0 x \cdot \left(-\frac{2}{\theta} x e^{-\frac{x^2}{\theta}} \right) dx = -\int_{-\infty}^0 \frac{2}{\theta} x^2 e^{-\frac{x^2}{\theta}} dx = -\int_{-\infty}^{+\infty} \frac{1}{\theta} x^2 e^{-\frac{x^2}{\theta}} dx$$

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

$$\text{令 } \bar{X} = E(X) = -\frac{1}{2} \sqrt{\pi\theta}, \quad \text{得到参数 } \theta \text{ 的矩估计量为 } \bar{\theta} = \frac{4\bar{X}^2}{\pi}.$$

$$(2) \text{ 似然函数 } L(x_1, x_2, \dots, x_n; \theta) = \prod_{k=1}^n p(x_k) = \left(\frac{2}{\theta}\right)^n \cdot e^{\frac{-1}{\theta} \sum_{k=1}^n x_k^2} \cdot \prod_{k=1}^n (-x_k)$$

$$\begin{aligned} \text{对数似然函数 } \ln L(x_1, x_2, \dots, x_n; \theta) &= n \ln\left(\frac{2}{\theta}\right) - \frac{1}{\theta} \sum_{k=1}^n x_k^2 + \sum_{k=1}^n \ln(-x_k) \\ &= n \ln 2 - n \ln \theta - \frac{1}{\theta} \sum_{k=1}^n x_k^2 + \sum_{k=1}^n \ln(-x_k) \end{aligned}$$

$$\text{关于参数 } \theta \text{ 求导, } \frac{d \ln}{d \theta} = -\frac{n}{\theta} + \frac{1}{\theta^2} \sum_{k=1}^n x_k^2,$$

$$\text{令 } \frac{d \ln L}{d \theta} = 0, \text{ 得 } \theta = \frac{1}{n} \sum_{k=1}^n x_k^2, \text{ 导数在该点左侧大于 } 0, \text{ 右侧小于 } 0,$$

所以似然函数在该点处取得最大值.

$$\text{参数 } \theta \text{ 的最大似然估计量 } \hat{\theta} = \frac{1}{n} \sum_{k=1}^n X_k^2.$$

六. (8 分) X_1, X_2, \dots, X_n 是来自总体 $X \sim N(\mu_1, \sigma^2)$ 的样本, Y_1, Y_2, \dots, Y_m 是来自总体 $Y \sim N(\mu_2, c \cdot \sigma^2)$ 的样本, σ^2 未知. $n=6, m=6, c=0.5, \bar{x}=52, \bar{y}=40, s_X^2=96, s_Y^2=52$. 求参数 $\mu_1 - \mu_2$ 的 90% 置信水平的双侧置信区间.

解: 根据正态总体的性质有

$$\frac{(n-1)s_X^2}{\sigma^2} \sim \chi^2(n-1), \quad \frac{(m-1)s_Y^2}{c \cdot \sigma^2} \sim \chi^2(m-1),$$

$$\frac{(n-1)s_X^2}{\sigma^2} + \frac{(m-1)s_Y^2}{c \cdot \sigma^2} \sim \chi^2(n+m-2),$$

构造 t 分布枢轴量

$$\frac{\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sigma \sqrt{\frac{1}{n} + \frac{c}{m}}}}{\frac{1}{\sigma} \sqrt{\left((n-1)s_X^2 + \frac{(m-1)s_Y^2}{c} \right) / (n+m-2)}} = \frac{\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{1}{n} + \frac{c}{m}}}}{\sqrt{\left((n-1)s_X^2 + \frac{(m-1)s_Y^2}{c} \right) / (n+m-2)}} \sim t(n+m-2)$$

$$\text{这里 } n=6, m=6, c=0.5, \sqrt{\frac{1}{n} + \frac{c}{m}} = 0.5$$

$$s_X^2 = 96, s_Y^2 = 52, \sqrt{\left((n-1)s_X^2 + \frac{(m-1)s_Y^2}{c} \right) / (n+m-2)} = 10$$

$$t \text{ 分布枢轴量为 } \frac{\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{0.5}}{10} = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{5} \sim t(10)$$

$$P\left(\left|\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{5}\right| \leq t_{0.95}(10)\right) = 0.9$$

参数 $\mu_1 - \mu_2$ 的 90% 单侧置信区间为 $(\bar{X} - \bar{Y} - 5 \cdot t_{0.95}(10), \bar{X} - \bar{Y} + 5 \cdot t_{0.95}(10))$, 即 $(3, 21)$ 。

7. (16 分) 设 X_1, \dots, X_{50} 是来自总体 $N(\mu, 2)$ 的样本。对期望进行假设检验, $H_0: \mu \leq 0, H_1: \mu > 0$, 若取拒绝域为

$$\{(x_1, \dots, x_{50}) : \bar{x} > 0.33\},$$

(1) 求此检验犯第一类错误的概率的最大值;

(2) 若 $\mu = -0.27$ 时, 此检验犯错误时是第几类错误, 并求犯此类错误的概率;

(3) 当 $\mu = 0.68$ 时, 此检验犯错误时是第几类错误, 并求犯此类错误的概率;

(4) $\bar{x} = 0.2$ 的 p 值。

解: (1) 当 $\mu = 0$ 时, 第一类错误达到最大, 此时第一类错误为

$$P(\bar{X} > 0.33 | \mu = 0) = P\left(\frac{\bar{X}}{\sqrt{\frac{2}{50}}} > 1.65 | \mu = 0\right) = 1 - \Phi(1.65) = 0.05,$$

(2) 当 $\mu = -0.27$ 时, 犯错误是第一类错误(拒真), 此时第一类错误为

$$P(\bar{X} > 0.33 | \mu = -0.27) = P\left(\frac{\bar{X} - (-0.27)}{\sqrt{\frac{2}{50}}} > \frac{0.33 - (-0.27)}{\sqrt{\frac{2}{50}}} | \mu = 0\right) = 1 - \Phi(3) = 0.001$$

(3) 当 $\mu = 0.68$ 时, 犯错误是第二类错误(受伪), 此时第二类错误为

$$P(\bar{X} \leq 0.33 | \mu = 0.68) = P_{\mu=0.68}\left(\frac{\bar{X} - 0.68}{\sqrt{\frac{2}{50}}} \leq \frac{0.33 - 0.68}{0.2}\right) = \Phi(-1.75) = 1 - \Phi(1.75) = 0.04$$

$$(4) p\text{值} = P(\bar{X} > 0.2 | \mu = 0) = P_{\mu=0}\left(\frac{\bar{X}}{\sqrt{\frac{2}{50}}} > \frac{0.2}{0.2}\right) = 1 - \Phi(1) = 0.16$$

备注 1. 本考卷的样本均为简单随机样本, 样本均值 $\bar{X} = \frac{X_1 + \cdots + X_n}{n}$, 样本方差为 $S^2 = \frac{1}{n-1} \sum_{k=1}^n (X_k - \bar{X})^2$

备注 2. 正态总体的样本均值和样本方差相互独立, 且 $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$, 其中 n 为样本容量

备注 3. 指数分布随机变量 $X \sim \text{Exp}(\lambda)$, $F_X(x) = (1 - e^{-\lambda x}) I_{x>0}$, $E(X) = 1/\lambda$, $\text{Var}(X) = 1/\lambda^2$

备注 4. 已知参数为 p 的几何分布随机变量的方差为 $\frac{1-p}{p^2}$ 。

备注 5. 分布函数和分位数 (题目解答要严格按照下面给出数值进行计算)

$\Phi(1.28) = 0.9$, $\Phi(1.44) = 0.925$, $\Phi(1.65) = 0.95$, $\Phi(1.96) = 0.975$,

$\Phi(1) = 0.84$, $\Phi(1.25) = 0.89$, $\Phi(1.5) = 0.93$, $\Phi(1.75) = 0.96$, $\Phi(2) = 0.98$, $\Phi(3) = 0.999$

$P(t(9) > 1.38) = 0.1$, $P(t(10) > 1.37) = 0.1$, $P(t(11) > 1.36) = 0.1$, $P(t(12) > 1.35) = 0.1$

$P(t(9) > 1.83) = 0.05$, $P(t(10) > 1.80) = 0.05$, $P(t(11) > 1.79) = 0.05$, $P(t(12) > 1.78) = 0.05$