

中国科学技术大学

2018—2019学年第一学期考试试卷

考试科目 概率论与数理统计(B) 得分 _____

所在系 _____ 姓名 _____ 学号 _____

考试时间: 2019 年1月9日上午8:30-10:30; 使用简单计算器

一、(30分, 每小题3分) 填空题或单选题, 答案可以直接写在试卷上.

(1) 已知 10 台洗衣机中有 7 台一等品和 3 台二等品, 现已售出一台, 在余下的 9 台中随机抽取 2 台后发现均为一等品, 则原先售出的一台为二等品的概率为_____.

(2) 设随机变量 X 的概率密度 $f(x) = Ae^{-x^2+x}$, $-\infty < x < \infty$, 则常数 $A =$ _____.

(3) 设随机变量 X 与 Y 相互独立, 且它们的取值范围分别是 $\{1, 2\}$ 和 $\{1, 2, 3\}$. 已知

$$P(Y = 1) = \frac{1}{6}, \quad P(X = 1, Y = 2) = P(Y = 1, X = 2) = \frac{1}{8},$$

则 $P(Y = 3) =$ _____.

(4) 设随机变向量 (X, Y) 的分布函数为 $\Phi(2x)\Phi(y-1)$, 其中 $\Phi(x)$ 为标准正态分布函数, 则 (X, Y) 服从二元正态分布()

(A) $N(0, 1; \frac{1}{4}, 1; 0)$ (B) $N(0, -1; \frac{1}{4}, 1; 0)$ (C) $N(0, 1; 4, 1; 0)$ (D) $N(0, -1; 4, 1; 0)$

(5) 设随机变量 X 与 Y 相互独立, 且 X 服从参数为 2 的泊松分布, Y 服从区间 $[-3, 3]$ 上的均匀分布, 则它们的乘积的方差 $\text{Var}(XY) =$ _____.

(6) 设 X_1, X_2, X_3, X_4 是来自标准正态总体的简单随机样本, $a > 0$ 为某个常数. 若已知

$$Y = a\left(\frac{1}{2}X_1^2 + \frac{1}{2}X_2^2 + \frac{1}{2}X_3^2 + \frac{1}{2}X_4^2 + X_1X_2 + X_3X_4\right)$$

服从 χ_n^2 分布, 则 $n + a =$ _____.

(7) 已知随机变量 X 服从 $F_{3,4}$ 分布. 设对给定的 $\alpha (0 < \alpha < 1)$, 实数 $F_{3,4}(\alpha)$ 满足 $P(X > F_{3,4}(\alpha)) = \alpha$. 若有 $P(X \leq x) = 1 - \alpha$, 则 x 等于()

(A) $\frac{1}{F_{4,3}(1-\alpha)}$ (B) $\frac{1}{F_{3,4}(1-\alpha)}$ (C) $F_{4,3}(\alpha)$ (D) $F_{4,3}(1-\alpha)$

(8) 设 X_1, X_2, \dots, X_n 是来自均匀总体 $U[-\theta, \theta]$ 的简单随机样本, 则参数 θ 的极大似然估计量 $\hat{\theta}$ 为()

(A) $\max_{1 \leq i \leq n} X_i$ (B) $\max_{1 \leq i \leq n} |X_i|$ (C) $-\min_{1 \leq i \leq n} X_i$ (D) $-\min_{1 \leq i \leq n} |X_i|$

(9) 设 $X_1, X_2, \dots, X_n (n > 2)$ 是来自正态总体 $N(\mu, \sigma^2)$ 的简单随机样本, 且 \bar{X} 为样本均值. 若统计量 $T = c(X_1 + X_n - 2\bar{X})^2$ 为 σ^2 的无偏估计, 则常数 $c =$ _____.

(10) 已知两个正态总体 X_1 和 X_2 分别为 $N(\mu_1, \sigma_1^2)$ 和 $N(\mu_2, \sigma_2^2)$, 为了检验总体 X_1 的均值大于 X_2 的均值, 则应作检验的假设为()

(A) $H_0: \mu_1 > \mu_2 \leftrightarrow H_1: \mu_1 \leq \mu_2$ (B) $H_0: \mu_1 \geq \mu_2 \leftrightarrow H_1: \mu_1 < \mu_2$
(C) $H_0: \mu_1 < \mu_2 \leftrightarrow H_1: \mu_1 \geq \mu_2$ (D) $H_0: \mu_1 \leq \mu_2 \leftrightarrow H_1: \mu_1 > \mu_2$

二、(24分) 设二维随机向量 (X, Y) 的联合密度函数为

$$f(x, y) = \frac{1}{5}(2x + y), \quad 0 \leq x \leq 2, 0 \leq y \leq 1.$$

- (1) 分别求 X 和 Y 的边缘密度函数 $f_X(x)$ 和 $f_Y(y)$;
- (2) 在给定 $X = 1$ 的条件下, 求 Y 在点 $y = 0.5$ 处的概率密度 $f_{Y|X}(0.5|1)$;
- (3) 求 X 和 Y 的协方差 $\text{Cov}(X, Y)$;
- (4) 求随机变量 $Z = \max\{X, Y\}$ 的密度函数 $f_Z(z)$.

三、(12分) 某药厂试制了一种新药, 声称对贫血患者的治疗有效率达到 80%. 医药监管部门随机抽取 200 个贫血患者进行此药的临床试验, 若至少有 152 人用药有效, 就批准此药的生产. 试利用中心极限定理, 求解如下问题:

- (1) 若该药的有效率确实达到 80%, 此药被批准生产的概率大约是多少?
- (2) 若监管部门的方案是 200 个人中要有 160 人用药有效才批准, 这对药厂是否公平? 需说明理由.

四、(18分) 已知总体 X 的密度函数为

$$f(x; \theta) = \begin{cases} -\theta^x \ln \theta, & x > 0; \\ 0, & x \leq 0, \end{cases}$$

其中 $0 < \theta < 1$ 为未知参数. 设 X_1, X_2, \dots, X_n 为来自该总体的一组简单随机样本.

- (1) 求 θ 的矩估计量 $\hat{\theta}$;
- (2) 求 $h(\theta) = (\ln \theta)^{-1}$ 的极大似然估计量 \hat{h}_θ ;
- (3) 试求实数 a , 使得 \hat{h}_θ 依概率收敛到 a , 即对任何 $\varepsilon > 0$, 都有 $\lim_{n \rightarrow \infty} P(|\hat{h}_\theta - a| \geq \varepsilon) = 0$.

五、(8分) 为比较A和B两种型号步枪子弹的枪口速度, 随机地抽取A型子弹 10 发, 得到枪口速度的平均值为 $\bar{x} = 500(\text{m/s})$, 样本标准差 $s_1 = 1.10(\text{m/s})$; 随机地抽取B型子弹 20 发, 得到枪口速度的平均值为 $\bar{y} = 496(\text{m/s})$, 样本标准差 $s_2 = 1.20(\text{m/s})$. 假设A和B型号子弹的枪口速度分别近似服从方差相等的正态分布 $N(\mu_1, \sigma^2)$ 和 $N(\mu_2, \sigma^2)$. 试在显著性水平 $\alpha = 0.05$ 下检验假设 $H_0: \mu_1 - \mu_2 = 5 \leftrightarrow H_1: \mu_1 - \mu_2 \neq 5$.

六、(8分) 某机构为了研究鼻咽癌是否与血型有关, 随机调查了一些患者和健康人, 得到的数据如下:

	A	B	O	AB
患者	64	86	130	20
健康人	125	138	210	26

请你根据所学的统计知识给出适当的结论(显著性水平设为 $\alpha = 0.05$).

附录: 上分位数表

$$u_{0.025} = 1.96, u_{0.05} = 1.645, \Phi(1.414) = 0.9214;$$

$$t_{28}(0.025) = 2.0484, t_{28}(0.05) = 1.7011, t_{29}(0.025) = 2.0452, t_{29}(0.05) = 1.6991;$$

$$\chi_3^2(0.95) = 0.3518, \chi_3^2(0.05) = 7.8147.$$

参考答案

一. (每小题3分)

$3/8$; $\frac{1}{\sqrt{\pi}}e^{-\frac{1}{4}}$; $1/3$; A; 18; 3; A; B; $\frac{n}{2(n-2)}$; D.

二. (1) (6分) 由边缘密度和联合密度的关系可知,

$$\begin{aligned}f_X(x) &= \int_0^1 f(x, y)dy = \frac{2}{5}x + \frac{1}{10}, \quad 0 \leq x \leq 2; \\f_Y(y) &= \int_0^2 f(x, y)dx = \frac{2}{5}y + \frac{4}{5}, \quad 0 \leq y \leq 1.\end{aligned}$$

(2) (6分) 由边缘密度、条件密度和联合密度的关系可知,

$$f_{Y|X}(0.5|1) = \frac{f(1, 0.5)}{f_X(1)} = 1.$$

(3) (6分) 由

$$E[XY] = \int_0^1 \int_0^2 xyf(x, y)dxdy = \frac{2}{3},$$

及 $EX = \frac{19}{15}$ 和 $EY = \frac{8}{15}$, 可知 $\text{Cov}(X, Y) = E[XY] - EXEY = -\frac{2}{225}$.

(4) (6分) 由

$$F_Z(z) = P(Z \leq z) = \iint_{x, y \leq z} f(x, y)dxdy$$

可知,

$$F_Z(z) = \begin{cases} 0, & z < 0; \\ \frac{3}{10}z^3, & 0 \leq z < 1; \\ \frac{1}{5}z^2 + \frac{1}{10}z, & 1 \leq z < 2; \\ 1, & z \geq 2. \end{cases}$$

从而, 所求密度函数为

$$f_Z(z) = \begin{cases} \frac{9}{10}z^2, & 0 \leq z < 1; \\ \frac{2}{5}z + \frac{1}{10}, & 1 \leq z < 2. \end{cases}$$

三. (1) (6分) 所求概率为 $\Phi(\sqrt{2}) = 0.9214$.

(2) (6分) 不公平. 对药厂而言, 在治疗有效率达到80%的情况下被批准的概率大约为 $\Phi(0) = 0.5$, 这相当于用掷硬币的方式来决定是否得到批准.

四. (1) (6分) 矩估计量 $\hat{\theta} = \exp\{-\frac{1}{\bar{X}}\}$, 其中 \bar{X} 为样本均值;

(2) (6分) 参数 θ 的极大似然估计量同样为 $\exp\{-\frac{1}{\bar{X}}\}$, 从而 $h(\theta)$ 的极大似然估计量为 $\hat{h}_\theta = -\bar{X}$;

(3) (6分) 由弱大数律可知, 所求的实数 $a = -EX = \frac{1}{\ln \theta}$.

五. (8分) 两样本 t 检验, 其检验统计量为

$$T = \frac{\bar{X} - \bar{Y} - \mu_0}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}.$$

代入数据计算可知, $s_w = 1.169, t = -2.209$. 由于 $|t| > t_{28}(0.025) = 2.0484$, 在显著性水平 $\alpha = 0.05$ 下我们应该拒绝原假设 H_0 .

六. (8分) 拟合优度联列表检验. 原假设为鼻咽癌与血型无关, 而其检验统计量为

$$\chi^2 = \sum_{i=1}^2 \sum_{j=1}^4 \frac{(nn_{ij} - n_{i \cdot} n_{\cdot j})^2}{nn_{i \cdot} n_{\cdot j}}.$$

代入数据计算可知, $\chi^2 = 1.921 < \chi_3^2(0.05) = 7.8147$. 故在显著性水平 $\alpha = 0.05$ 下我们不能拒绝原假设, 即可以认为鼻咽癌与血型无关.