中 国 科 学 技 术 大 学

2022—2023学年第二学期考试试卷

考试科目 概率论与数理统计 得分 ____________

所在院系 ________ 姓名 ____________ 学号 ____________

考试时间: 2023 年 6 月 27 日下午 14:30–16:30; 可使用简单计算器

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) 设 A 和 B 是一随机试验的两个互斥结果, 且 $\mathrm{P}(A)=p, \mathrm{P}(B)=q, 0<p+q\leq 1$. 现独立重复这一试验时, 则结果 A 比 B 先出现的概率为______.

(2) 设随机变量 X 的概率密度函数 $f(x)$ 满足 $f(2+x)=f(2-x)$ 及 $\int_0^2 f(x)\mathrm{d}x=0.3$, 则 $\mathrm{P}(X\leq 0)=$ ( )

(A) 0.1 (B) 0.2 (C) 0.3 (D) 0.4

(3) 设 X 和 Y 分别为离散型和连续型随机变量且相互独立. 则 $X+Y$ 的类型为 ( ).

(A) 连续型 (B) 离散型 (C) 非离散型也非连续型 (D) 以上皆有可能

(4) 掷一枚均匀的骰子, 欲使 1 至 6 点均出现至少一次, 则平均需要掷______次.

(5) 在单位圆盘上独立地随机取两点 A 和 B, 则以 A 为圆心, 线段 AB 为半径的圆完全落在该单位圆盘内的概率是( )

(A) $\frac{1}{2}$ (B) $\frac{1}{3}$ (C) $\frac{1}{4}$ (D) $\frac{1}{6}$

(6) 设 $X_1, X_2, \cdots, X_n$ 是来自指数总体 $\mathrm{Exp}(2)$ 的一组简单随机样本, 则当 $n\to\infty$ 时, 统计量 $Y_n=\sum_{i=1}^n X_i^2/n$ 依概率收敛到 ( )

(A) $\frac{1}{2}$ (B) $\frac{1}{4}$ (C) 2 (D) 4

(7) 设 X_1, X_2, X_3, X_4, X_5 是来自正态总体 $N(0,\sigma^2)$ 的一组简单随机样本, 若统计量 $T=a(X_1+X_2)/\sqrt{X_3^2+X_4^2+X_5^2}$ 服从 t 分布, 则常数 $a=$ ( )

(A) $\frac{1}{2}$ (B) $\frac{\sqrt{2}}{2}$ (C) $\frac{\sqrt{6}}{2}$ (D) $\frac{\sqrt{6}}{6}$

(8) 已知一批零件的长度 X (单位: cm)服从正态分布, 从中随机抽取 25 个零件, 得到样本的平均长度为 36 cm, 样本方差为 4 cm^2, 则总体均值 $\mathrm{E}X$ 的置信水平为 0.95 的置信区间为______________ (精确到小数点后三位).

(9) 从电信公司用户每月话费账单中随机抽取 25 张, 算得平均消费为 32.8 元, 标准差为 20.8 元. 设用户每月话费服从均值为 μ 的正态分布, 对假设检验 $H_0: \mu\leq 30 \leftrightarrow H_1: \mu>30$, 则在显著性水平 $\alpha=0.5$ 下我们____拒绝原假设 (填“可以”或“不能”).

(10) 设一总体分布函数为 $F(x;\theta)$, 其中 θ 为一未知参数, 而 θ_0 为一已知常数. 若一假设检验 $H_0: \theta=\theta_0 \leftrightarrow H_1: \theta\neq\theta_0$ 在显著性水平 α 下的拒绝域为 $|\overline{X}-\theta_0|>1$, 这里 $\overline{X}$ 为样本均值, 则 θ 的置信系数为 $1-\alpha$ 的置信区间为_________.

装 订 线

二、 (20分) 设 $X_1, X_2, \cdots, X_n$ 为一列独立的随机变量, 且同分布于参数为 λ 的指数分布 $\mathrm{Exp}(\lambda)$. 记 $Y = \min(X_1, X_2, \cdots, X_n)$.

(1) 求 Y 的分布.

(2) 求概率 $\mathrm{P}(X_1 = Y)$.

(3) 求随机向量 (X_1, Y) 的联合分布函数 $F(x, y)$.

(4) 对任一实数 $y > 0$, 求条件期望 $\mathrm{E}[X_1 | Y = y]$.

三、 (14分) 设随机变量 X, Y, Z 为相互独立的标准正态随机变量, 且记

$$U = \frac{X - Y}{\sqrt{2}}, \quad V = \frac{X + Y - 2Z}{\sqrt{6}}, \quad W = \frac{X + Y + Z}{\sqrt{3}}.$$

(1) 试求随机向量 (U, V, W) 的联合概率密度函数 $f(u, v, w)$.

(2) 随机变量 U, V, W 是否相互独立? 需说明理由.

四、 (21分) 设某种电子器件的寿命 (单位: 小时) T 服从双参数指数分布, 即其概率密度函数为

$$f(t) = \frac{1}{\theta} \exp\left\{ -\frac{t - c}{\theta} \right\} I_{(c,\infty)}(t),$$

其中参数 $\theta, c > 0$. 现自一批这种器件中随机地取 n 件进行寿命试验, 记录它们的失效时间依次为 $t_1 < t_2 < \cdots < t_n$.

(1) 求 θ 和 c 的矩估计值.

(2) 求 θ 和 c 的极大似然估计值.

(3) 若已知参数 $\theta = \frac{1}{2}$, 则 c 的极大似然估计是否无偏? 若是, 证明之; 若否, 修正之.

五、 (15分) 某篮球爱好者每天进行 5 组投篮练习, 每组练习均包含 4 次投篮, 完成后他记录下每组投中的次数. 经过 80 天后, 得到如下数据:

投中数	0	1	2	3	4
频数	25	118	139	93	25

结合你所学的统计知识, 我们能否认为该篮球爱好者每组投中的次数服从二项分布 (显著性水平 $\alpha = 0.05$)?

附录 上分位数: $t_{24}(0.025) = 2.064$, $t_{24}(0.05) = 1.711$

$\chi_3^2(0.05) = 7.815$, $\chi_3^2(0.95) = 0.352$, $\chi_4^2(0.05) = 9.488$, $\chi_4^2(0.95) = 0.711$

(完)

参考答案

一、 每小题 3 分.

$\frac{p}{p+q}$; B; A; 14.7; D; A; C; $(35.174, 36.826)$; 不能; $[\overline{X}-1, \overline{X}+1]$.

二、 每小题 5 分.

(1) 对任一 $y>0$, 由 $\mathrm{P}(Y>y)=[\mathrm{P}(X_1>y)]^n=\mathrm{e}^{-n\lambda y}$, 可知 $Y\sim \mathrm{Exp}(n\lambda)$, 即 Y 服从参数为 $n\lambda$ 的指数分布.

(2) 由对称性可知 $\mathrm{P}(X_1=Y)=\cdots=\mathrm{P}(X_n=Y)$, 且各项之和为 1. 从而所求概率为

$$\mathrm{P}(X_1=Y)=\frac{1}{n}.$$

(3) 当 $0<y<x$ 时, 由独立性可知

$$\begin{aligned}\mathrm{P}(X_1\le x, Y>y)&=\mathrm{P}(y<X_1\le x, X_2>y,\cdots,X_n>y)\\&=(\mathrm{e}^{-\lambda y}-\mathrm{e}^{-\lambda x})\mathrm{e}^{-(n-1)\lambda y}\\&=\mathrm{e}^{-n\lambda y}-\mathrm{e}^{-\lambda[x+(n-1)\lambda y]}.\end{aligned}$$

从而,

$$\begin{aligned}\mathrm{P}(X_1\le x, Y\le y)&=\mathrm{P}(X_1\le x)-\mathrm{P}(X_1\le x, Y>y)\\&=1-\mathrm{e}^{-\lambda x}-\mathrm{e}^{-n\lambda y}+\mathrm{e}^{-\lambda[x+(n-1)\lambda y]}.\end{aligned}$$

而当 $0<x\le y$ 时, $\mathrm{P}(X_1\le x, Y\le y)=\mathrm{P}(X_1\le x)=1-\mathrm{e}^{-\lambda x}$. 综上可知, 随机向量 (X_1,Y) 的联合分布函数为

$$F(x,y)=\begin{cases}1-\mathrm{e}^{-\lambda x}-\mathrm{e}^{-n\lambda y}+\mathrm{e}^{-\lambda[x+(n-1)\lambda y]}, & 0<y<x;\\ 1-\mathrm{e}^{-\lambda x}, & 0<x\le y;\\ 0, & \text{其它}.\end{cases}$$

(4) 由 (2) 及指数分布的无记忆性可知

$$\mathrm{E}[X_1|Y=y]=y\cdot\frac{1}{n}+\Big(y+\frac{1}{\lambda}\Big)\frac{n-1}{n}=y+\frac{n-1}{n\lambda}.$$

三、 每小题 7 分.

(1) 由计算可知 Jacob 行列式绝对值为 1, 且 $u^2+v^2+w^2=x^2+y^2+z^2$, 故由密度变换公式可知随机向量 (U,V,W) 的概率密度函数为

$$f(u,v,w)=\frac{1}{(2\pi)^{3/2}}\exp\Big\{-\frac{u^2+v^2+w^2}{2}\Big\}.$$

(2) 由上式中的概率密度函数 $f(u,v,w)$ 可分离变量即知随机变量 U,V,W 相互独立.

四、 每小题 7 分.

(1) 首先, 可计算得

$$\mathrm{E}T = \theta + c, \quad \mathrm{Var}(T) = \theta^2.$$

然后通过联立方程

$$\mathrm{E}T = \bar{t} = \frac{1}{n}\sum_{i=1}^{n} t_i, \quad \mathrm{Var}(T) = s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(t_i - \bar{t})^2,$$

可得 α 和 c 的矩估计值分别为$\hat{\theta}_1 = s, \quad \hat{c}_1 = \bar{t} - s.$

注: 在本小题中, 可将样本方差 s^2 替换成样本中心二阶矩 $m_2 = \frac{1}{n}\sum_{i=1}^{n}(t_i - \bar{t})^2$.

(2) 可得似然函数为

$$L(\theta, c) = \frac{1}{\theta^n}\exp\Big\{ -\frac{1}{\theta}\sum_{i=1}^{n}(t_i - c)\Big\}, \quad c \le t_1 < t_2 < \cdots < t_n.$$

由于该函数关于 c 单调递增, 可知参数 c 的极大似然估计值为 $\hat{c}_2 = t_1$. 然后, 可求得对数似然函数为

$$l(\alpha, c) = -n\ln\theta - \frac{1}{\theta}\sum_{i=1}^{n}(t_i - c).$$

故通过令 $\frac{\partial l(\theta,c)}{\partial\theta} = 0$, 可得 θ 的极大似然估计值为 $\hat{\theta}_2 = \bar{t} - t_1$.

(3) 否. 若 $\theta = 1/2$, 则可求得估计量 $\hat{c}_2 = T_1$ 的概率密度函数为

$$g(t) = 2n\mathrm{e}^{-2n(t-c)} I_{(c,\infty)}(t).$$

由此可求得 $\mathrm{E}[\hat{c}_2] = c + \frac{1}{2n}$, 从而 $\hat{c}_2$ 不是 c 的无偏估计, 但可修正为 $\hat{c}_2^* = t_1 - \frac{1}{2n}$.

五、 拟合优度检验. 原假设为每组投中的次数服从二项分布 $B(4, p)$, 其中 $0 < p < 1$ 为某个常数. 我们先估计参数 p, 利用极大似然估计或者矩估计, 均有

$$\hat{p} = \frac{1}{80 \times 5 \times 4}(25 + 118 \times 1 + 139 \times 2 + 93 \times 3 + 25 \times 4) = \frac{775}{1600} = 0.4844.$$

由此可计算得出投中次数0, 1, 2, 3, 4的理论频数分别为 $28.3, 106.2, 149.7, 93.8, 22.0$. 从而, 由检验统计量

$$\chi^2 = \frac{(25 - 28.3)^2}{28.3} + \cdots + \frac{(25 - 22.0)^2}{22.0} = 2.877 < \chi_3^2(0.05) = 7.815,$$

故我们应接受每组投中次数服从二项分布的说法.

步骤分: 估计 p, 5 分; 计算出检验统计量的值, 5 分; 临界值, 3分; 统计决策, 2 分.