中 国 科 学 技 术 大 学

2023—2024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: 2024 年 6 月 21 日下午 14:30–16:30; 可使用简单计算器

装 ………… 订 ………… 线

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) 设事件 A 和 B 相互独立且 $\mathrm{P}(A)=\mathrm{P}(B)$, 若 $\mathrm{P}(A\cup B)=0.64$, 则 $\mathrm{P}(A)=$______.

(2) 已知现在的电子邮件里有一半是垃圾邮件, 而某软件能以 99% 的概率正确识别垃圾邮件, 但也会以 5% 的概率误将常规邮件识别为垃圾邮件. 若一封电子邮件被该软件识别为垃圾邮件, 则它其实为常规邮件的概率是______.

(3) 设 x, y 为任意给定的实数, 对独立的随机变量 X, Y, 下列中不一定正确的是( )

(A) 若 $\mathrm{P}(Y=y)>0$, 则 $\mathrm{P}(X\le x|Y=y)=\mathrm{P}(X\le x)$ (B) X^2 与 Y^2 相互独立

(C) 随机事件 $\{X=x\}$ 与 $\{Y=y\}$ 相互独立 (D) X 与 Y 不相关

(4) 将 n 根短绳的 $2n$ 个端头随机两两连接, 以 a_n 表示最终所连成圈的个数的期望(注意圈也可由多根绳子构成), 则 $\lim_{n\to\infty}(a_n/\ln n)=$ ( )

(A) 0 (B) $\frac{1}{2}$ (C) 1 (D) 2

(5) 设 X, Y 为相互独立且服从均匀分布 $U(0,1)$ 的随机变量. 若分别记随机变量 $Z=\max\{X,Y\}, W=\min\{X,Y\}$, 则协方差 $\mathrm{Cov}(Z,W)=$______.

(6) 设随机变量 X_n 服从自由度为 n 的 χ^2 分布, 则当 $n\to\infty$ 时, X_n/n 依概率收敛到 ______.

(7) 设 $X_1, X_2, \cdots, X_n$ 为来自正态总体 $N(0,\sigma^2)$ 的简单随机样本, 且以 $\overline{X}$ 和 S^2 分别表示样本均值和样本方差, 则下列统计量中服从 F 分布的是( )

(A) $(n-1)\overline{X}^2/S^2$ (B) $n\overline{X}^2/S^2$ (C) $(n+1)\overline{X}^2/S^2$ (D) $\overline{X}^2/S^2$

(8) 对一给定总体的某个参数的无偏估计, 下述表述一定正确的是( )

(A) 无偏估计优于有偏估计 (B) 无偏估计的个数至多有限个

(C) 无偏估计一定存在 (D) 若总体期望存在但未知, 则样本均值为无偏估计

(9) 已知一设备生产出的零件的长度服从正态分布 $N(\mu,\sigma^2)$, 其中 μ, σ^2 均未知. 现从中随机抽取 16 个产品, 测得样本均值和样本标准差分别为 20 厘米和 0.1 厘米, 则方差 σ^2 的置信系数为 0.9 的置信区间为__________(精确到小数点后三位).

(10) 设 Z_n 为服从 χ_n^2 分布的随机变量, 若在 3×5 列联表独立性检验中, 由样本求得的检验统计量的值为 z_0, 则该假设检验的 p 值大约为( ).

(A) $\mathrm{P}(Z_8\ge z_0)$ (B) $\mathrm{P}(Z_8\le z_0)$ (C) $\mathrm{P}(Z_{14}\ge z_0)$ $\mathrm{P}(Z_{14}\le z_0)$

二、 (18分) 设随机向量 (X,Y) 的联合概率密度函数为 $f(x,y)=xe^{-x(y+1)}, x,y>0.$

(1) 分别求 X 和 Y 的边缘密度函数 $f_1(x)$ 和 $f_2(y)$.

(2) 通过计算说明在给定 $X=x\,(x>0)$ 的条件下 Y 服从指数分布, 并指出参数.

(3) 计算期望 $\mathrm{E}Y, \mathrm{E}[Y|X]$ 和 $\mathrm{E}[XY]$.

三、 (18分) 设随机变量 X 和 Y 相互独立, 且分别服从参数为 λ 和 μ 的指数分布. 记随机变量 $Z=\min\{X,Y\}$.

(1) 试求概率 $\mathrm{P}(Z=X)$.

(2) 试求随机变量 $W=\max\{X-Y,0\}$ 的分布函数 $F_W(w)$.

(3) 问 Z 是否与示性变量 $I_{\{X<Y\}}$ (亦即与随机事件 $\{X<Y\}$) 独立? 需写出证明过程.

四、 (20分) 设 $X_1,X_2,\cdots,X_n$ 为抽自总体 X 的一组简单随机样本, 且 X 的概率密度函数为 $f(x;\theta)=-\theta^x\ln\theta\cdot I_{(0,\infty)}(x)$, 其中 $0<\theta<1$ 为一未知参数.

(1) 试求 θ 的矩估计 $\hat{\theta}$.

(2) 试求 $h(\theta)=(\ln\theta)^{-1}$ 的最大似然估计 $\hat{h}$.

(3) 上一小题中, 估计量 $\hat{h}$ 是否为无偏估计? 证明你的结论.

(4) 试求常数 $b>0$, 使得对任一实数 x, 均有

$$\lim_{n\to\infty}\mathrm{P}\left(\frac{\sqrt{n}\big(\hat{h}-(\ln\theta)^{-1}\big)}{b}\le x\right)=\Phi(x)$$

成立, 其中 $\Phi(\cdot)$ 为标准正态分布函数.

五、 (14分) 从甲乙两班中随机抽取部分同学, 获得他们的《概率论与数理统计》期末考试成绩如下:

甲	85	72	94	69	88	87	64	74	63	86	99	91	82
乙	93	96	70	82	71	99	84	95	78	68	77	88	

假设甲班和乙班的学生成绩独立地分别服从 $N(\mu_1,195)$ 和 $N(\mu_2,120)$分布, 其中 μ_1,μ_2 均未知.

(1) 在显著性水平 $\alpha=0.05$ 下, 对如下检验问题做出决策:

$$H_0:\mu_1=\mu_2 \longleftrightarrow H_1:\mu_1-\mu_2=-1.8.$$

(2) 在 H_1 成立条件下计算上述检验的功效函数值.

附录 上分位数: $u_{0.025}=1.96$, $u_{0.05}=1.645$, $u_{0.1}=1.285$, $u_{0.2}=0.845$, $u_{0.3}=0.525$, $t_{23}(0.05)=1.7139$, $t_{23}(0.025)=2.0687$, $t_{24}(0.05)=1.7109$, $t_{24}(0.025)=2.0639$, $\chi^2_{15}(0.05)=24.996$, $\chi^2_{15}(0.95)=7.261$, $\chi^2_{15}(0.10)=22.307$, $\chi^2_{15}(0.90)=8.547$

(完)

参考答案

一、 每小题 3 分.

$\frac{2}{5}$　　$\frac{5}{104}$　　D　　B　　$\frac{1}{36}$　　1　　B　　D　　$[0.006, 0.021]$　　A

二、 每小题 6 分.

(1) 由联合密度和边缘密度的关系可知(注: 变量取值范围缺失或出错，一处扣 1 分),

$$f_1(x) = \int_0^\infty xe^{-x(y+1)}\mathrm{d}y = e^{-x}, \quad x > 0,$$
$$f_2(y) = \int_0^\infty xe^{-x(y+1)}\mathrm{d}x = \frac{1}{(y+1)^2}, \quad y > 0.$$

(2) 由条件密度函数公式和 (1), 可知在 $X = x > 0$ 的条件下 Y 的条件密度函数为

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_1(x)} = \frac{xe^{-x(y+1)}}{e^{-x}} = xe^{-xy}, \quad y > 0,$$

即此时 Y 的条件分布是参数为 x 的指数分布.

(3) 由 (1) 可知, $\mathrm{E}Y = \int_0^\infty \frac{y}{(y+1)^2}\mathrm{d}y = \infty$, 即期望 $\mathrm{E}Y$ 不存在. 而由 (2) 及指数分布的性质可知, 对任一 $x > 0$, 有 $\mathrm{E}[Y|X = x] = \frac{1}{x}$, 故 $\mathrm{E}[Y|X] = \frac{1}{X}$. 再由全期望公式,

$$\mathrm{E}[XY] = \mathrm{E}[\mathrm{E}(XY|X)] = \mathrm{E}[X\mathrm{E}(Y|X)] = \mathrm{E}\Big[X \cdot \frac{1}{X}\Big] = 1.$$

上述结果也可由 $\mathrm{E}[XY] = \int_0^\infty\int_0^\infty xyf(x,y)\mathrm{d}x\mathrm{d}y$ 直接计算得出.

三、 每小题 6 分.

(1) 易知 $\mathrm{P}(Z = X) = \mathrm{P}(X < Y)$, 故由连续型全概率公式可知

$$\mathrm{P}(Z = X) = \int_0^\infty \mathrm{P}(Y > x|X = x)f_X(x)\mathrm{d}x = \int_0^\infty \lambda e^{-(\lambda+\mu)x}\mathrm{d}x = \frac{\lambda}{\lambda+\mu}.$$

(2) 首先由 (1) 可知, $\mathrm{P}(W = 0) = \mathrm{P}(X < Y) = \frac{\lambda}{\lambda+\mu}$. 而对任一 $w > 0$,

$$\mathrm{P}(W > w) = \mathrm{P}(X - Y > w) = \int_0^\infty \mathrm{P}(X > y + w)\mu e^{-\mu y}\mathrm{d}y$$
$$= \int_0^\infty \mu e^{-\lambda(y+w)-\mu y}\mathrm{d}y = \frac{\mu}{\lambda+\mu}e^{-\lambda w}.$$

从而, 分布函数 $F_W(w)$ 为 (注意下式的完整性, 不完整至少扣 1 分)

$$F_W(w) = \begin{cases} 0, & w < 0; \\ \frac{\lambda}{\lambda+\mu}, & w = 0; \\ 1 - \frac{\mu}{\lambda+\mu}e^{-\lambda w}, & w > 0. \end{cases}$$

(3) 独立. 我们只需证明对任一 $z > 0$, 事件 $\{Z > z\}$ 与 $\{X < Y\}$ 相互独立即可. 由

$$\mathrm{P}(X < Y, Z > z) = \mathrm{P}(z < X < Y) = \frac{\lambda}{\lambda+\mu}e^{-(\lambda+\mu)z} = \mathrm{P}(X < Y) \cdot \mathrm{P}(Z > z),$$

即知结论成立. (结论 2 分, 证明过程 4 分.)

四、 每小题 5 分.

(1) 由于总体期望

$$\mathrm{E}X = \int_0^\infty -x\theta^x \ln\theta \mathrm{d}x = -\frac{1}{\ln\theta},$$

可知 θ 的矩估计 $\hat{\theta} = \exp\{-1/\overline{X}\}$, 其中 $\overline{X}$ 为样本均值.

(2) 先求 θ 的最大似然估计. 由似然函数 $L(\theta) = -\theta^{\sum_{i=1}^n X_i}(\ln\theta)^n$ 知对数似然函数为

$$l(\theta) = -\ln\theta\sum_{i=1}^n X_i - n\ln\ln\theta.$$

令

$$\frac{\partial l}{\partial\theta} = -\frac{1}{\theta}\sum_{i=1}^n X_i - \frac{n}{\theta\ln\theta} = 0,$$

再化简可得 $\overline{X} + \frac{1}{\ln\theta} = 0$. 由此可知, θ 的最大似然估计也为 $\exp\{-1/\overline{X}\}$. 从而, $h(\theta) = (\ln\theta)^{-1}$ 的最大似然估计 $\hat{h} = -\overline{X}$.

(3) 为无偏估计. 这是因为由 (1) 和 (2) 可知,

$$\mathrm{E}[\hat{h}] = -\mathrm{E}[\overline{X}] = -\mathrm{E}X = (\ln\theta)^{-1}.$$

(4) 由独立同分布场合下的中心极限定理(或 Lindeberg-Lévy CLT)可知 $b = \sqrt{\mathrm{Var}(X)}$. 而由

$$\mathrm{E}[X^2] = \int_0^\infty -x^2\theta^x \ln\theta \mathrm{d}x = \frac{2}{\ln^2\theta}$$

及 (1) 可知, $\mathrm{Var}(X) = 1/(\ln^2\theta)$. 故 $b = -(\ln\theta)^{-1}$.

五、 每小题 7 分.

(1) 首先通过计算可以得到两班成绩的样本均值分别为 $\bar{x} = 81.0769$, $\bar{y} = 83.4167$. 在原假设 H_0 成立条件下, 检验统计量

$$T = \frac{\bar{x}-\bar{y}}{\sqrt{\frac{195}{13}+\frac{120}{12}}} = \frac{\bar{x}-\bar{y}}{5} \sim N(0,1).$$

故在显著性水平 $\alpha = 0.05$ 下, 拒绝域为 $T < -u_{0.05} = -1.645$. 代入数值后计算可得 $T = -0.4679 > -1.645$, 从而我们应接受 H_0.

(2) 在备择假设 H_1 成立下, 检验统计量 $T \sim N(-1.8, 1)$, 故功效函数值为

$$\begin{aligned}
\mathrm{P}_{H_1}(T < -u_{0.05}) &= \mathrm{P}_{H_1}\Big(\frac{\bar{x}-\bar{y}}{5} < -u_{0.05}\Big) \\
&= \mathrm{P}_{H_1}\Big(\frac{\bar{x}-\bar{y}+1.8}{5} < \frac{1.8}{5} - 1.645\Big) \\
&= \Phi(-1.285) \\
&= 1 - \Phi(1.285) \\
&= 0.1.
\end{aligned}$$