河南大学 2020 2021 学年第二学期期末考试
《概率论与数理统计 A》试卷 A 卷解析

一、 选择题（每小题 **3** 分，共 **30** 分）

1. 设 $P(A)=\frac{1}{4}, P(B|A)=\frac{1}{3}, P(A|B)=\frac{1}{2}$，则 $P(A\cup B)$ 等于（ **D** ） (A) $\frac{1}{12}$ (B) $\frac{1}{6}$ (C) $\frac{1}{4}$ (D) $\frac{1}{3}$

   解析：

$$P(AB)=P(A)P(B|A)=\frac{1}{4}\times\frac{1}{3}=\frac{1}{12}$$

$$P(B)=\frac{P(AB)}{P(A|B)}=\frac{\frac{1}{12}}{\frac{1}{2}}=\frac{1}{6}$$

$$P(A\cup B)=P(A)+P(B)-P(AB)=\frac{1}{4}+\frac{1}{6}-\frac{1}{12}=\frac{4}{12}=\frac{1}{3}$$

2. 设 $X\sim N(\mu,\sigma^2)$，则 $Z=\frac{X-\mu}{\sigma}$ 服从的分布是（ **A** ） (A) $N(0,1)$ (B) $N(0,\sigma^2)$ (C) $N(\mu,1)$ (D) $N(\mu,\sigma^2)$

   解析： 根据正态分布的性质，若 $X\sim N(\mu,\sigma^2)$，则其标准化变量 $Z=\frac{X-\mu}{\sigma}$ 服从标准正态分布 $N(0,1)$。

3. 设随机变量 (X,Y) 具有概率密度 $f(x,y)=\begin{cases}2e^{\{-(2x+y)\}}x>0 \\ y>0 \\ 0 \quad 其他\end{cases}$，则 $P\{Y\le X\}=$（ **A** ） (A) $\frac{1}{3}$ (B) $\frac{1}{4}$ (C) $\frac{2}{5}$ (D) $\frac{4}{5}$

   解析：

$$\begin{aligned}P\{Y\le X\}&=\int_0^\infty\int_0^x 2e^{\{-(2x+y)\}}\mathrm{d}y\mathrm{d}x\\&=\int_0^\infty 2e^{\{-2x\}}\left[-e^{\{-y\}}\right]_0^x\mathrm{d}x\\&=\int_0^\infty 2e^{\{-2x\}}\left(1-e^{\{-x\}}\right)\mathrm{d}x\\&=\int_0^\infty\left(2e^{\{-2x\}}-2e^{\{-3x\}}\right)\mathrm{d}x\\&=\left[-e^{\{-2x\}}+\frac{2}{3}e^{\{-3x\}}\right]_0^\infty=0-\left(-1+\frac{2}{3}\right)=\frac{1}{3}\end{aligned}$$

4. 对于二维随机变量 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，则 X 与 Y 相互独立的充分必要条件是（ **B** ） (A) $\rho=-1$ (B) $\rho=0$ (C) $\rho=0.5$ (D) $\rho=1$

   解析： 对于二维正态分布，随机变量 X 与 Y 相互独立的充要条件是它们不相关，即相关系数 $\rho=0$。

5. 设随机变量 X 与 Y 独立同分布，且 X 的分布函数为 $F(x)$，则 $Z=\min\{X,Y\}$ 的分布函数为（ **C** ） (A) $F^2(z)$ (B) $2F(z)$ (C) $1-[1-F(z)]^2$ (D) $F(z)[1-F(z)]$

   解析：

$$\begin{aligned}F_{Z(z)}&=P\{\min(X,Y)\le z\}=1-P\{\min(X,Y)>z\}\\&=1-P\{X>z,Y>z\}\\&=1-P\{X>z\}P\{Y>z\}\\&=1-[1-F(z)]^2\end{aligned}$$

6. 设 $E(X)=-2,E(Y)=2,D(X)=1,D(Y)=4,\rho_{XY}=-0.5$，则由切比雪夫不等式可得 $P\{|X+Y|\geq 6\}\leq$（ **D** ） (A) $\frac{1}{2}$ (B) $\frac{1}{4}$ (C) $\frac{1}{8}$ (D) $\frac{1}{12}$

   解析： 首先计算 $X+Y$ 的期望和方差：

$$E(X+Y)=E(X)+E(Y)=-2+2=0$$

$$\begin{aligned}D(X+Y)&=D(X)+D(Y)+2\rho\sqrt{D(X)D(Y)}\\&=1+4+2(-0.5)(1\times 2)=5-2=3\end{aligned}$$

   由切比雪夫不等式 $P\{|Z-E(Z)|\geq\varepsilon\}\leq\frac{D(Z)}{\varepsilon^2}$：

$$P\{|X+Y-0|\geq 6\}\leq\frac{3}{6^2}=\frac{3}{36}=\frac{1}{12}$$

7. 设随机变量 $X\sim U(a,b)$，其中 a 已知，b 未知，X_1,X_2,X_3,X_4 是样本，则下列选项可以作为统计量的是（ **C** ） (A) $\frac{1}{b}\sum(X_i+a)$ (B) $\frac{1}{b}\sum\left(X_i-\overline{X}\right)$ (C) $\sum\left(X_i-a\right)^2$ (D) $\sum\left(X_i-b\right)^2$

   解析： 统计量是样本的函数，且不含有任何未知参数。选项 A, B, D 中均含有未知参数 b，故只有 C 是统计量。

8. 设 $\chi^2\sim\chi^2(10)$，则 $E(\chi^2)+D(\chi^2)$ 等于（ **C** ） (A) 10 (B) 20 (C) 30 (D) 40

   解析： 对于 $\chi^2(n)$ 分布，期望 $E(\chi^2)=n$，方差 $D(\chi^2)=2n$。 本题中 $n=10$，故 $E(\chi^2)+D(\chi^2)=10+20=30$。

9. 设总体 X 的均值为 μ，方差为 σ^2，$X_1,...,X_n$ 是样本，则 $D\left(\overline{X}\right)$ 等于（ **D** ） (A) μ (B) σ^2 (C) $\frac{\mu}{n}$ (D) $\frac{\sigma^2}{n}$

   解析： 样本均值的方差公式：$D\left(\overline{X}\right)=D\left(\frac{1}{n}\sum X_i\right)=\frac{1}{n^2}\sum D(X_i)=\frac{n\sigma^2}{n^2}=\frac{\sigma^2}{n}$。

10. 设 $X_1,...,X_n$ 为总体 $N(\mu,\sigma^2)$ 的样本，$\overline{X},S^2$ 分别是样本均值和样本方差，若 μ 未知，σ^2 已知，则 μ 的置信水平为 $1-\alpha$ 的置信区间为（ **B** ） (A) $\left(\overline{X}\pm\frac{\sigma}{\sqrt{n-1}}z_{\frac{\alpha}{2}}\right)$ (B) $\left(\overline{X}\pm\frac{\sigma}{\sqrt{n}}z_{\frac{\alpha}{2}}\right)$ (C) $\left(\overline{X}\pm\frac{\sigma}{\sqrt{n-1}}z_\alpha\right)$ (D) $\left(\overline{X}\pm\frac{\sigma}{\sqrt{n}}z_\alpha\right)$

    解析： 当方差 σ^2 已知时，构造统计量 $Z=\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\sim N(0,1)$，由此得置信区间。

二、 填空题（每小题 **3** 分，共 **15** 分）

1. 已知 $P\left(\overline{A}\right)=0.3, P(B)=0.4, P\left(A\overline{B}\right)=0.5$，条件概率 $P\left(B \mid A\cup\overline{B}\right)=\underline{0.25}$。

   解析：

$$P(A)=1-P\left(\overline{A}\right)=0.7$$

$$P(AB)=P(A)-P\left(A\overline{B}\right)=0.7-0.5=0.2$$

$$P\left(A\cup\overline{B}\right)=P(A)+P\left(\overline{B}\right)-P\left(A\overline{B}\right)=0.7+0.6-0.5=0.8$$

$$P\left(B \mid A\cup\overline{B}\right)=\frac{P\left(B\cap\left(A\cup\overline{B}\right)\right)}{P\left(A\cup\overline{B}\right)}=\frac{P(AB)}{0.8}=\frac{0.2}{0.8}=0.25$$

2. 设 X 与 Y 是随机变量，且 $P\{X\geq 0, Y\geq 0\}=\frac{3}{7}, P\{X\geq 0\}=P\{Y\geq 0\}=\frac{4}{7}$，则 $P\{\max(X,Y)\geq 0\}=\underline{5/7}$。

   解析： $\max(X,Y)\geq 0$ 等价于 $X\geq 0$ 或 $Y\geq 0$。

$$P\{X\geq 0\cup Y\geq 0\}=P\{X\geq 0\}+P\{Y\geq 0\}-P\{X\geq 0, Y\geq 0\}=\frac{4}{7}+\frac{4}{7}-\frac{3}{7}=\frac{5}{7}$$

3. 设平面区域 D 由曲线 $y=\frac{1}{x}$ 及直线 $y=0, x=1, x=e^2$ 所围成，二维随机变量 (X,Y) 在区域 D 上服从均匀分布，则 (X,Y) 关于 X 的边缘概率密度在 $x=2$ 处的值为 $\underline{1/4}$。

   解析： 区域面积 $S=\int_1^{\{e^2\}}\frac{1}{x}\mathrm{d}x=[\ln x]_1^{\{e^2\}}=2$。 联合密度 $f(x,y)=\frac{1}{2}$ (在 D 内)。 边缘密度 $f_{X(x)}=\int_0^{\{\frac{1}{x}\}}\frac{1}{2}\mathrm{d}y=\frac{1}{2x}, 1\leq x\leq e^2$。 当 $x=2$ 时，$f_{X(2)}=\frac{1}{2\times 2}=\frac{1}{4}$。

4. 设随机变量 X_1, X_2 的概率密度分别为 $f_1(x)=2e^{\{-2x\}}(x>0), f_2(x)=4e^{\{-4x\}}(x>0)$。若 X_1 和 X_2 相互独立，则 $E(X_1X_2)=\underline{1/8}$。

   解析： $X_1\sim\mathrm{Exp}(2)\to E(X_1)=\frac{1}{2}$；$X_2\sim\mathrm{Exp}(4)\to E(X_2)=\frac{1}{4}$。 由于独立，$E(X_1X_2)=E(X_1)E(X_2)=\frac{1}{2}\times\frac{1}{4}=\frac{1}{8}$。

5. 设样本 $X_1,...,X_6$ 来自总体 $N(0,1)$，$Y=(X_1+X_2+X_3)^2+(X_4+X_5+X_6)^2$，若 CY 服从 χ^2 分布，则常数 $C=\underline{1/3}$。

   解析： $X_1+X_2+X_3\sim N(0,3)\to\frac{X_1+X_2+X_3}{\sqrt{3}}\sim N(0,1)$。 其平方 $\frac{(X_1+X_2+X_3)^2}{3}\sim\chi^2(1)$。 同理 $\frac{(X_4+X_5+X_6)^2}{3}\sim\chi^2(1)$。 故 $\frac{Y}{3}=\frac{(X_1+X_2+X_3)^2+(X_4+X_5+X_6)^2}{3}\sim\chi^2(2)$，即 $C=\frac{1}{3}$。

三、 计算题（每小题 **11** 分，共 **44** 分）

1. 假设箱中有一个球，只知道不是白球就是红球。现在将两个白球放进箱中，然后从箱中随意取出一个球，结果是白球，求箱中原来是白球的概率？

   解： 设 H_1 为“箱中原球为白球”，H_2 为“箱中原球为红球”。 由题意知 $P(H_1) = P(H_2) = \frac{1}{2}$。 设 A 为“取出的是白球”。

   - 若 H_1 发生，箱中有 3 个白球，取出白球概率 $P(A|H_1) = \frac{3}{3} = 1$。
   - 若 H_2 发生，箱中有 2 白 1 红，取出白球概率 $P(A|H_2) = \frac{2}{3}$。

   由贝叶斯公式：

$$P(H_1|A) = \frac{P(H_1)P(A|H_1)}{P(H_1)P(A|H_1) + P(H_2)P(A|H_2)} = \frac{\frac{1}{2} \times 1}{\frac{1}{2} \times 1 + \frac{1}{2} \times \frac{2}{3}} = \frac{1}{1 + \frac{2}{3}} = \frac{3}{5}$$

2. 设随机变量 X 具有概率密度 $f_{X(x)} = \begin{cases} ke^{\{-x\}} x>0 \\ 0 \qquad x\le 0 \end{cases}$。 求：(1) 常数 k；(2) 随机变量 $Y = e^X$ 的密度函数 $f_{Y(y)}$。

   解： (1) 由密度函数的归一化性质：

$$\int_0^\infty ke^{\{-x\}} \mathrm{d}x = k\left[-e^{\{-x\}}\right]_0^\infty = k(0 - (-1)) = k = 1$$

   (2) $Y = e^X$ 是单调增函数。当 $x > 0$ 时，$y > 1$。 反函数 $x = \ln y$，导数 $\mathrm{d}\frac{x}{\mathrm{d}}y = \frac{1}{y}$。

$$f_{Y(y)} = f_{X(\ln y)} \cdot |\mathrm{d}\frac{x}{\mathrm{d}}y| = e^{\{-\ln y\}} \cdot \frac{1}{y} = \frac{1}{y^2}, \quad y > 1$$

   故 $f_{Y(y)} = \begin{cases} \frac{1}{y^2} y>1 \\ 0 \ \text{其他} \end{cases}$。

3. 已知二维随机变量 (X, Y) 在 G 上服从均匀分布，G 由直线 $x - y = 0, x + y = 2$ 与 $y = 0$ 所围成，求： (1) 边缘概率密度 $f_{X(x)}$ 和 $f_{Y(y)}$；(2) 条件概率密度 $f_{X|Y}(x|y)$。

   解： 区域 G 的顶点为 $(0,0), (2,0), (1,1)$，面积 $S = \frac{1}{2} \times 2 \times 1 = 1$。 联合密度 $f(x, y) = 1$ (在 G 内)。
   (1) 边缘密度：

$$f_{X(x)} = \begin{cases} \int_0^x 1\mathrm{d}y = x & 0 \le x \le 1 \\ \int_0^{\{2-x\}} 1\mathrm{d}y = 2 - x 1 < x \le 2 \\ 0 & \text{其他} \end{cases}$$

$$f_{Y(y)} = \int_y^{\{2-y\}} 1\mathrm{d}x = (2 - y) - y = 2 - 2y, \quad 0 \le y \le 1$$

   (2) 条件密度：

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_{Y(y)}} = \frac{1}{2 - 2y}, \quad y \le x \le 2 - y, (0 \le y < 1)$$

4. 设二维随机变量 (X, Y) 具有概率密度 $f(x, y) = \frac{1}{8}(x + y), 0 \le x \le 2, 0 \le y \le 2$。求 X 和 Y 的相关系数 ρ_{XY}。

   解： 由于对称性，$E(X) = E(Y), D(X) = D(Y)$。

$$E(X) = \int_0^2 \int_0^2 x \cdot \frac{1}{8}(x + y)\mathrm{d}x\mathrm{d}y = \frac{1}{8}\int_0^2 \left(\frac{8}{3} + 2y\right)\mathrm{d}y = \frac{1}{8}\left(\frac{16}{3} + 4\right) = \frac{7}{6}$$

$$E(X^2) = \int_0^2 \int_0^2 x^2 \cdot \frac{1}{8}(x + y)\mathrm{d}x\mathrm{d}y = \frac{1}{8}\int_0^2 \left(4 + \frac{8}{3}y\right)\mathrm{d}y = \frac{1}{8}\left(8 + \frac{16}{3}\right) = \frac{5}{3}$$

$$D(X) = \frac{5}{3} - \left(\frac{7}{6}\right)^2 = \frac{60}{36} - \frac{49}{36} = \frac{11}{36}$$

$$E(XY) = \frac{1}{8}\int_0^2 \left[\frac{1}{2}x^2y + \frac{1}{3}xy^2\right]_0^2 \mathrm{d}y = \frac{1}{8}\int_0^2 \left(2y + \frac{4}{3}y^2\right)\mathrm{d}y = \frac{1}{8}\left[y^2 + \frac{4}{9}y^3\right]_0^2 = \frac{1}{8}\left(4 + \frac{32}{9}\right) = \frac{17}{18}$$

$$\mathrm{Cov}(X,Y) = E(XY) - E(X)E(Y) = \frac{17}{18} - \left(\frac{7}{6}\right)^2 = \frac{34}{36} - \frac{49}{36} = -\frac{15}{36} = -\frac{5}{12}$$

$$\rho_{XY} = \frac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)D(Y)}} = \frac{-\frac{15}{36}}{\frac{11}{36}} = -\frac{15}{11}$$

四、 证明题（共 **11** 分）

1. 设总体 X 的概率密度为 $f(x;\theta)=\frac{1}{\theta}x^{\{\frac{1-\theta}{\theta}\}},0<x<1,\theta>0$。 (1) 求 θ 的最大似然估计量 $\hat{\theta}$；(2) 证明 $\hat{\theta}$ 是 θ 的无偏估计量。

证： (1) 似然函数：

$$L(\theta)=\prod_{i=1}^{n}f(x_i;\theta)=\theta^{\{-n\}}\prod_{i=1}^{n}x_i^{\{\frac{1-\theta}{\theta}\}}$$

对数似然函数：

$$\ln L(\theta)=-n\ln\theta+\frac{1-\theta}{\theta}\sum_{i=1}^{n}\ln x_i$$

对 θ 求导并令其为 0：

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\ln L(\theta)=-\frac{n}{\theta}-\frac{1}{\theta^2}\sum_{i=1}^{n}\ln x_i=0$$

解得最大似然估计量：

$$\hat{\theta}=-\frac{1}{n}\sum_{i=1}^{n}\ln X_i$$

(2) 计算 $E(\ln X)$：

$$\begin{aligned}E(\ln X)&=\int_0^1\ln x\cdot\frac{1}{\theta}x^{\{\frac{1-\theta}{\theta}\}}\mathrm{d}x\\&\text{令}\quad t=\ln x,x=e^t,\mathrm{d}x=e^t\mathrm{d}t\\&=\int_{\{-\infty\}}^{0}t\cdot\frac{1}{\theta}e^{\left\{\frac{t(1-\theta)}{\theta}\right\}}e^t\mathrm{d}t\\&=\frac{1}{\theta}\int_{\{-\infty\}}^{0}te^{\{\frac{t}{\theta}\}}\mathrm{d}t\\&\text{分部积分得：}=\frac{1}{\theta}\left[\theta te^{\{\frac{t}{\theta}\}}-\theta^2e^{\{\frac{t}{\theta}\}}\right]_{\{-\infty\}}^{0}=\frac{1}{\theta}(0-\theta^2)=-\theta\end{aligned}$$

因此：

$$E\left(\hat{\theta}\right)=E\left(-\frac{1}{n}\sum\ln X_i\right)=-\frac{1}{n}\sum E(\ln X_i)=-\frac{1}{n}\cdot n(-\theta)=\theta$$

故 $\hat{\theta}$ 是 θ 的无偏估计量。