

《概率论与数理统计》模拟试卷 3 参考答案

一、单选题.

1. 设 A, B 为任意两事件, 且 $A \subset B$, $P(B) > 0$, 则下列成立的是(A).

- A. $P(A) < P(A|B)$ B. $P(A) \leq P(A|B)$
C. $P(A) > P(A|B)$ D. $P(A) \geq P(A|B)$

2. 设 $\Phi(x)$ 为标准正态分布函数, $X_i = \begin{cases} 1, & A \text{ 发生;} \\ 0, & A \text{ 不发生;} \end{cases} \quad i = 1, 2, \dots, 100$, 且

$P(A) = 0.8$, $X_1, X_2, \dots, X_{100}$ 相互独立. 令 $Y = \sum_{i=1}^{100} X_i$, 则由中心极限定理知 Y 的分布

函数 $F(y)$ 近似于 (B).

- A. $\Phi(y)$ B. $\Phi(\frac{y-80}{4})$ C. $\Phi(16y+80)$ D. $\Phi(4y+80)$

3. (X, Y) 是二维随机向量, 与 $\text{Cov}(X, Y) = 0$ 不等价的是(D).

- A. $E(XY) = E(X)E(Y)$ B. $D(X+Y) = D(X) + D(Y)$
C. $D(X-Y) = D(X) + D(Y)$ D. X 和 Y 相互独立

4. 设随机变量 X 的密度函数为 $f(x)$, 且 $f(x) = f(-x)$, $F(x)$ 是 X 的分布函数, 则对于任意实数 a , 有 (A).

- A. $F(-a) = 1 - \int_0^a f(x)dx$ B. $F(-a) = \frac{1}{2} - \int_0^a f(x)dx$
C. $F(-a) = F(a)$; D. $F(-a) = 2F(a) - 1$

5. 设总体 $X \sim N(\mu, 1)$, (X_1, X_2) 是取自于总体 X 的一个简单随机样本, 则下列 μ 的无偏估计量中哪一个最有效 (即方差最小) (C).

- A. $\hat{\mu}_1 = \frac{2}{3}X_1 + \frac{1}{3}X_2$ B. $\hat{\mu}_2 = \frac{1}{4}X_1 + \frac{3}{4}X_2$
C. $\hat{\mu}_3 = \frac{1}{2}X_1 + \frac{1}{2}X_2$ D. $\hat{\mu}_4 = \frac{2}{5}X_1 + \frac{3}{5}X_2$

二、填空题.

1. 某人连续向一目标射击, 每次命中目标的概率为 $3/4$, 他连续射击直到命中为止, 则

射击次数为 3 的概率是 $\frac{3}{64}$.

2. 设随机变量 $X \sim N(110, 144)$, 求 $P\{100 < X < 120\} = \underline{0.5934}$ ($\Phi(\frac{5}{6}) = 0.7967$)

3. 设随机变量 X 的分布函数为: $F(x) = \begin{cases} 0 & x \leq 1 \\ a \ln x & 1 < x \leq e \\ 1 & x > e \end{cases}$. 则常数 $a = \underline{1}$.

4. 设随机变量的概率分布为 $P(X = k) = \frac{k+1}{10}$, $k = 0, 1, 2, 3$, 则 $E(X) = \underline{2}$.

5. $X_1, X_2, \dots, X_n$ 是取自总体 $N(\mu, \sigma^2)$ 的简单样本, 则 $\frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \bar{X})^2 = \underline{\chi^2(n-1)}$.

三、证明题.

1. 设 $X \sim \pi(\lambda_1)$, $Y \sim \pi(\lambda_2)$, 且 X, Y 相互独立, 求证 $X+Y \sim \pi(\lambda_1 + \lambda_2)$.

证: $\because X \sim \pi(\lambda_1), Y \sim \pi(\lambda_2)$,

$$\therefore P(X = k_1) = \frac{\lambda_1^{k_1}}{k_1!} e^{-\lambda_1}, P(Y = k_2) = \frac{\lambda_2^{k_2}}{k_2!} e^{-\lambda_2}, k_1, k_2 \in Z^+.$$

又 X, Y 相互独立, 则对 $\forall n \in Z^+$ 有

$$\begin{aligned} P(X+Y=n) &= \sum_{i=1}^n P(X=i, Y=n-i) = \sum_{i=1}^n P(X=i)P(Y=n-i) \\ &= \sum_{i=1}^n \frac{\lambda_1^i}{i!} e^{-\lambda_1} \frac{\lambda_2^{n-i}}{(n-i)!} e^{-\lambda_2} = e^{-\lambda_1-\lambda_2} \sum_{i=1}^n \frac{n!}{i!(n-i)!} \frac{1}{n!} \lambda_1^i \lambda_2^{n-i} \\ &= e^{-\lambda_1-\lambda_2} \frac{1}{n!} \sum_{i=1}^n \frac{n!}{i!(n-i)!} \lambda_1^i \lambda_2^{n-i} = \frac{(\lambda_1 + \lambda_2)^n}{n!} e^{-\lambda_1-\lambda_2}. \end{aligned}$$

所以 $X+Y \sim \pi(\lambda_1 + \lambda_2)$.

2. 设 (X, Y) 有联合分布密度函数 $f(x, y) = \begin{cases} 12e^{-(4x+3y)}, & x > 0, y > 0, \\ 0 & \text{其他.} \end{cases}$,

证明 X, Y 相互独立.

证: 因为

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy = \begin{cases} \int_0^{+\infty} 12e^{-(4x+3y)} dy = 4e^{-4x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$
$$f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx = \begin{cases} \int_0^{+\infty} 12e^{-(4x+3y)} dx = 3e^{-3y}, & y > 0 \\ 0, & y \leq 0 \end{cases}$$

即对 $\forall x, y \in R$ 都有 $f(x, y) = f_X(x)f_Y(y)$, 所以 X, Y 相互独立.

四、计算题.

1. 某工厂由甲、乙、丙三个车间生产同一种产品, 每个车间的产量分别占全厂的 25%, 35%, 40%, 各车间产品的次品率分别为 5%, 4%, 2%. 求: (1) 全厂产品的次品率 (2) 若任取一件产品发现是次品, 此次品是甲车间生产的概率是多少?

解: 设事件 A_1, A_2, A_3 分别表示取到的产品由甲、乙、丙三个车间生产, 事件 B 表示取到的是次品, 则根据题意有

$$P(A_1) = 0.25, P(A_2) = 0.35, P(A_3) = 0.4,$$

$$P(B|A_1) = 0.05, P(B|A_2) = 0.04, P(B|A_3) = 0.02.$$

由全概率公式得

$$P(B) = P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + P(A_3)P(B|A_3)$$

$$= 0.25 \times 0.05 + 0.35 \times 0.04 + 0.4 \times 0.02 = 0.0345.$$

由贝叶斯公式得

$$P(A_1|B) = \frac{P(A_1B)}{P(B)} = \frac{P(A_1)P(B|A_1)}{P(B)}$$

$$= \frac{0.25 \times 0.05}{0.0345} = \frac{25}{69} \approx 0.362.$$

2. 设总体 X 的概率分布为 $P\{X=x\} = p^x(1-p)^{1-x}, x=0,1$, 设 $X_1, X_2, X_3, \dots, X_n$ 为总体 X 的一组简单随机样本, 求参数 p 的矩估计和最大似然估计.

解: 似然函数为

$$L(p) = \prod_{i=1}^n P(X_i = x_i, p) = \prod_{i=1}^n p^{x_i} (1-p)^{1-x_i} = p^{\sum_{i=1}^n x_i} (1-p)^{n-\sum_{i=1}^n x_i} \quad (x_i = 0, 1)$$

对数似然函数为

$$\ln L(p) = \left(\sum_{i=1}^n x_i\right) \ln p + \left(n - \sum_{i=1}^n x_i\right) \ln(1-p),$$

$$\text{令 } \frac{d \ln L(p)}{dp} = \frac{1}{p} \sum_{i=1}^n x_i - \frac{1}{1-p} \left(n - \sum_{i=1}^n x_i\right) = 0, \text{ 得 } \hat{p}_{MLE} = \frac{1}{n} \sum_{i=1}^n x_i = \bar{X}.$$

$$\text{又 } E(X) = \sum_{x_i=0,1} x_i p^{x_i} (1-p)^{1-x_i} = p,$$

$$\text{令 } \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i = E(X),$$

$$\text{得 } \hat{\theta}_{\text{矩}} = \bar{X}.$$

3. 正常人的脉搏平均为 72 次/分，今对某种疾病患者 9 人，测得其脉搏为（次/分）：
68 65 77 70 64 69 72 62 71. 设患者的脉搏次数 X 服从正态分布，经计算得其标准差为 4.583. 试在显著水平 $\alpha=0.1$ 下，检测患者的脉搏与正常人的脉搏有无显著差异？（已知： $t_{0.05}(9)=1.833$, $t_{0.05}(8)=1.860$ $z_{0.05}=1.645$, $z_{0.025}=1.960$ ）

解：设 $X \sim N(\mu, \sigma^2)$ ，其中 μ, σ^2 未知，提出假设

$$H_0: \mu = \mu_0 = 72, H_1: \mu \neq \mu_0 = 72.$$

取检验统计量为

$$t = \frac{\bar{X} - \mu_0}{S / \sqrt{n}},$$

拒接域为

$$|t| \geq t_{0.05}(8) = 1.860,$$

代入数据计算得

$$t = \left| \frac{68.67 - 72}{4.583 / \sqrt{9}} \right| \approx 2.182 > 1.86,$$

所以在该显著水平下拒接原假设，即认为检测患者的脉搏与正常人的脉搏有显著差异.

4. 随机抽取某种炮弹 9 发做实验，测得炮口速度的样本标准 $S = 3(m/s)$ ，设炮口速度服从正态分布，求这种炮弹的炮口速度的方差 σ^2 的置信度为 0.95 的置信区间.

(已知： $\chi^2_{0.025}(8)=17.535$, $\chi^2_{0.975}(8)=2.18$; $\chi^2_{0.025}(9)=19.02$, $\chi^2_{0.975}(9)=2.7$)

解：设表示炮口速度 $X \sim N(\mu, \sigma^2)$ 且 μ, σ^2 未知， $X_i (i=1, 2, \dots, n)$ 表示 X 的样本，

$\bar{X} = \sum_{i=1}^n X_i$ 和 $S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$ 分别为样本均值和方差.

取枢轴量为 $\chi^2 = \frac{(n-1)S^2}{\sigma^2}$, 则 $\chi^2 \sim \chi^2(n-1)$.

又由

$$P\{\chi_{0.025}^2(n-1) < \frac{(n-1)S^2}{\sigma^2} < \chi_{0.975}^2(n-1)\} = 0.95$$

得方差 σ^2 的置信度为 0.95 的置信区间为

$$\left(\frac{(n-1)S^2}{\chi_{0.975}^2(n-1)}, \frac{(n-1)S^2}{\chi_{0.025}^2(n-1)} \right),$$

将数据 $n=9, S=3$ 代入得所求置信区间为 (4.106, 33.028).

5. 为研究某一化学反应过程中, 温度 $x(^{\circ}\text{C})$ 对产品得率 $Y(\%)$ 的影响, 测得数据如下:

温度 $x(^{\circ}\text{C})$	100	110	120	130	140
得率 $Y(\%)$	45	51	54	61	66

求 Y 关于 x 的线性回归方程.

解: 设产品得率 y 关于温度 x 的直线回归方程为 $y = a + bx$,

记

$$S_{xx} = \sum_{i=1}^5 (x_i - \bar{x})^2 = \sum_{i=1}^5 x_i^2 - \frac{1}{5} \left(\sum_{i=1}^5 x_i \right)^2,$$

$$S_{xy} = \sum_{i=1}^5 (x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^5 x_i y_i - \frac{1}{5} \left(\sum_{i=1}^5 x_i \right) \left(\sum_{i=1}^5 y_i \right),$$

$$S_{yy} = \sum_{i=1}^5 (y_i - \bar{y})^2 = \sum_{i=1}^5 y_i^2 - \frac{1}{5} \left(\sum_{i=1}^5 y_i \right)^2,$$

$$\text{其中, } \bar{x} = \frac{1}{5} \sum_{i=1}^5 x_i, \quad \bar{y} = \frac{1}{5} \sum_{i=1}^5 y_i$$

代入数据计算得

$$\bar{x} = 120, \bar{y} = 55.4, \quad \hat{b} = \frac{S_{xy}}{S_{xx}} = \frac{520}{1000} = 0.52, \hat{a} = \bar{y} - \hat{b}\bar{x} = -7$$

所以产品得率 y 关于温度 x 的直线回归方程为 $y = -7 + 0.52x$.