

《概率论与数理统计》模拟试卷 2 参考答案

一、单选题.

1. 设 A, B 为两个随机事件, 且 $A \subset B$, 则下列等式不一定正确的是 (B).
A. $P(B-A) = P(B) - P(A)$ B. $P(AB) = P(A)P(B)$
C. $P(B\bar{A}) = P(B) - P(AB)$ D. $P(A \cup B) = P(B)$
2. 已知随机变量 X 的概率密度为 $f_X(x)$, 令 $Y = -2X + 3$, 则 Y 的概率密度 $f_Y(y)$ 为 (B).
A. $-\frac{1}{2}f_X(-\frac{y-3}{2})$ B. $\frac{1}{2}f_X(-\frac{y-3}{2})$ C. $-\frac{1}{2}f_X(-\frac{y+3}{2})$ D. $\frac{1}{2}f_X(-\frac{y+3}{2})$
3. 若随机变量 X 与 Y 相互独立, 则 (C).
A. $D(X-3Y) = DX - 9DY$ B. $D(XY) = DX DY$
C. $E(XY) = EX \cdot EY$ D. $P\{Y = aX + b\} = 1$
4. 设总体 X 的数学期望 $E(X) = \mu$, 方差 $D(X) = \sigma^2$, X_1, X_2, X_3, X_4 是来自总体 X 的简单随机样本, 则下列 μ 的估计量中最有效的是 (D).
A. $\frac{1}{6}X_1 + \frac{1}{6}X_2 + \frac{1}{3}X_3 + \frac{1}{3}X_4$ B. $\frac{1}{3}X_1 + \frac{1}{3}X_2 + \frac{1}{3}X_3$
C. $\frac{3}{5}X_1 + \frac{4}{5}X_2 - \frac{1}{5}X_3 - \frac{1}{5}X_4$ D. $\frac{1}{4}X_1 + \frac{1}{4}X_2 + \frac{1}{4}X_3 + \frac{1}{4}X_4$
5. 在假设检验中, H_0 为原假设, H_1 为备择假设, 下列说法正确的是 (D).
A. 拒绝 H_0 则认为 H_0 是错的
B. H_1 不真时接受 H_1 称为犯第二类错误
C. 设 $P\{\text{拒绝}H_0 | H_0\text{真}\} = \alpha$, $P\{\text{接受}H_0 | H_0\text{不真}\} = \beta$, 则 $\alpha + \beta = 1$
D. α 、 β 的意义同 C, 当样本容量一定时, α 变大时则 β 变小

二、填空题.

1. 设一次试验中事件 A 发生的概率为 p , 现重复进行 n 次独立试验, 则事件 A 至多发生一次的概率为 $(1-p)^n + np(1-p)^{n-1} = (1-p)^{n-1}(1-p+np)$.
2. 设随机变量 $X \sim N(2, 9)$, 且 $P\{X \geq 2\} = P\{X \leq c+2\}$, 则 $c = \underline{0}$.

3. 若随机变量 $X \sim b(400, 0.8)$, 根据中心极限定理, $P\{X \leq 320\} \approx \frac{1}{2}$.

4. 设随机变量 $X \sim U[0, 2]$, $Y = 2X + 1$, 则 $EY = 3$.

5. 设总体 $X \sim N(\mu, \sigma^2)$, $X_1, X_2, \dots, X_n$ 为总体 X 的样本, 设 $\bar{X}, S^2$ 分别为样本均值和样本方差, z_α 为标准正态分布的上 α 分位点, 则当 σ^2 已知时, 总体均值 μ 的置信水平为 $1-\alpha$ 的置信区间为 $\bar{X} \pm \frac{\sigma}{\sqrt{n}} z_{\alpha/2}$.

三、证明题.

1. 设总体 $X \sim N(\mu, \sigma^2)$, $X_1, X_2, \dots, X_n$ 为总体 X 的样本, 设 $\bar{X}, S^2$ 分别为样本均值和样本方差, 证明 $\bar{X}, S^2$ 分别为 μ 与 σ^2 的无偏估计.

$$\text{证明: } E(\bar{X}) = \frac{1}{n} \sum_{i=1}^n EX_i = \frac{1}{n} n\mu = \mu \quad D(\bar{X}) = \frac{1}{n^2} \sum_{i=1}^n DX_i = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}$$

$$\begin{aligned} E(S^2) &= E \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n-1} E(\sum_{i=1}^n X_i^2 - n\bar{X}^2) \\ &= \frac{1}{n-1} [\sum_{i=1}^n (DX_i + (EX_i)^2) - n(D\bar{X} + E(\bar{X})^2)] \\ &= \frac{1}{n-1} (\sum_{i=1}^n (\sigma^2 + \mu^2) - n(\frac{\sigma^2}{n} + \mu^2)) \\ &= \frac{1}{n-1} (n\sigma^2 + n\mu^2 - \sigma^2 - n\mu^2) = \sigma^2 \end{aligned}$$

2. 设随机变量 (X, Y) 具有分布律 $P\{X = x, Y = y\} = p^2(1-p)^{x+y-2}$. 其中 $0 < p < 1$, x, y 均为正整数. 证明 X, Y 相互独立.

证明: 对任意正整数 x, y 有

$$\begin{aligned} P\{X = x\} &= \sum_{y=1}^{+\infty} p^2(1-p)^{x+y-2} = p^2(1-p)^{x-2} \sum_{y=1}^{+\infty} (1-p)^y \\ &= p^2(1-p)^{x-2} \frac{1-p}{1-(1-p)} = p(1-p)^{x-1}. \end{aligned}$$

$$\text{同理可得 } P\{Y = y\} = \sum_{x=1}^{+\infty} p^2(1-p)^{x+y-2} = p(1-p)^{y-1}.$$

即 $P\{X = x, Y = y\} = P\{X = x\}P\{Y = y\}$, 所以 X, Y 独立.

四、计算题.

1. 设考生的报名表来自三个地区, 分别有 10 份, 15 份, 25 份, 其中女生的分别为 3 份, 7 份, 5 份. 随机不放回地从一地区中先后任取两份报名表, 求: (1) 后取的那份报名表是女生的概率 p ; (2) 已知后取到的报名表是男生的, 而先取的那份报名表是女生的概率 q .

解: A_i 表示第 i 次取到女生的报名表 ($i=1, 2$), 则 $A_1, \bar{A}_1$ 为样本空间的一个划分, 由全概率公式得

$$\begin{aligned} p = P(A_2) &= P(A_1)P(A_2 | A_1) + P(\bar{A}_1)P(A_2 | \bar{A}_1) \\ &= \frac{15}{50} \times \frac{14}{49} + \frac{35}{50} \times \frac{15}{49} = 0.3. \end{aligned}$$

又由贝叶斯公式得

$$q = P(A_1 | \bar{A}_2) = \frac{P(A_1 \bar{A}_2)}{P(\bar{A}_2)} = \frac{P(A_1)P(\bar{A}_2 | A_1)}{1 - P(A_2)} = \frac{\frac{15}{50} \times \frac{35}{49}}{1 - \frac{3}{10}} = \frac{15}{49} \approx 0.306.$$

2. 设总体 X 的概率密度为

$$f(x) = \begin{cases} \sqrt{\theta} x^{\sqrt{\theta}-1}, & 0 < x < 1, \\ 0, & \text{其他.} \end{cases}$$

其中参数 $\theta > 0$ 未知, $X_1, X_2, \dots, X_n$ 为来自总体的一个样本, 求参数 θ 的矩估计量和最大似然估计量.

解: 似然函数为

$$L(\theta) = \prod_{i=1}^n f(x_i, \theta) = \prod_{i=1}^n \theta^{\frac{1}{2}} x_i^{\sqrt{\theta}-1} = \theta^{\frac{n}{2}} (\prod_{i=1}^n x_i)^{\sqrt{\theta}-1} (0 < x_i < 1),$$

对数似然函数为

$$\ln L(\theta) = \frac{n}{2} \ln \theta + (\sqrt{\theta} - 1) \ln \prod_{i=1}^n x_i,$$

$$\text{令 } \frac{d \ln L(\theta)}{d\theta} = \frac{n}{2\theta} + \frac{\ln \prod_{i=1}^n x_i}{2\sqrt{\theta}} = 0, \text{ 得 } \hat{\theta}_{MLE} = \frac{\theta^2}{(\ln \prod_{i=1}^n x_i)^2} = \frac{\theta^2}{(\sum_{i=1}^n \ln x_i)^2}.$$

$$\text{又 } E(X) = \int_0^1 x \sqrt{\theta} x^{\sqrt{\theta}-1} dx = \int_0^1 \sqrt{\theta} x^{\sqrt{\theta}} dx = \frac{\sqrt{\theta}}{\sqrt{\theta}+1} x^{\sqrt{\theta}+1} \Big|_0^1 = \frac{\sqrt{\theta}}{\sqrt{\theta}+1},$$

$$\text{令 } \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i = E(X), \text{ 得 } \hat{\theta}_{矩} = \left(\frac{\bar{X}}{1-\bar{X}} \right)^2.$$

3. 按规定, 某灯泡厂生产的灯泡平均寿命不得小于 1120 小时, 现从一批新生产的灯泡中抽取 9 个样本, 测得其平均寿命为 1070 小时, 样本标准差 $S=109$ 小时, 假设灯泡寿命服从正态分布. 问在 $\alpha=0.05$ 显著性水平下, 检测该批灯泡是否符合要求.

(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 表示灯泡的寿命, 则 $X \sim N(\mu, \sigma^2)$, 其中 μ, σ^2 未知, 提出假设

$$H_0: \mu \geq \mu_0 = 1120, H_1: \mu < \mu_0 = 1120.$$

$$\text{取检验统计量为: } t = \frac{\bar{X} - \mu_0}{S / \sqrt{n}},$$

$$\text{拒接域为: } t < -t_{0.05}(8) = -1.860,$$

$$\text{代入数据计算得: } t = \frac{1070 - 1120}{109 / \sqrt{9}} \approx -1.38 > -t_{0.05}(8),$$

所以在该显著水平下不能拒接原假设, 即认为该批灯泡符合要求.

4. 设 (X, Y) 有联合分布密度函数

$$f(x, y) = \begin{cases} Ae^{-(4x+3y)}, & x > 0, y > 0, \\ 0 & \text{其他.} \end{cases}$$

求 (1) 常数 A ; (2) 判断 X 与 Y 是否相互独立, 为什么?

(3) 求概率 $P\{0 < X < 1, 0 < Y < 1\}$.

解: (1) 由 $\int_0^{+\infty} \int_0^{+\infty} Ae^{-(4x+3y)} dx dy = 1$ 得 $A=12$.

$$(2) f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy = \begin{cases} \int_0^{+\infty} 12e^{-(4x+3y)} dy = 4e^{-4x}, & x > 0, \\ 0, & x \leq 0 \end{cases},$$

类似地有

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

所以对 $\forall x, y$ 有 $f(x, y) = f_X(x)f_Y(y)$ ，即 X, Y 相互独立.

$$(3) \quad P(0 < X < 1, 0 < Y < 1) = \int_0^1 e^{-4x} dx \int_0^1 3e^{-3y} dy = (1 - e^{-4})(1 - e^{-3}).$$

5. 某房地产投资公司出售五个楼盘面积与售价资料如下：

楼盘面积 x （百平方米）	9	15	10	11	10
售价 y （千元）	36	80	44	55	35

试建立售价 y 对楼盘面积 x 的直线回归方程？

解：设售价 y 对楼盘面积 x 的直线回归方程为

$$y = a + bx,$$

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

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

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

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

代入数据计算得

$$\bar{x} = 11, \bar{y} = 50, \quad \hat{b} = \frac{S_{xy}}{S_{xx}} = \frac{169}{22} \approx 7.68, \hat{a} = \bar{y} - \hat{b}\bar{x} \approx -34.5$$

所以售价 y 对楼盘面积 x 的直线回归方程为

$$y = -34.5 + 7.68x.$$