

《概率论与数理统计》期末考试卷(A) 参考答案

一、填空题〔每小题 5 分，共计 25 分〕

1、设 A, B, C 是三个随机事件，则事件 $\overline{A} \cup \overline{B} \cup \overline{C}$ 表达的涵义是 三随机事件中至少有一个不发生。

2、设随机变量 X 与 Y 相互独立，且 $X \sim N(0, 1/3)$ ， $Y \sim N(0, 2/3)$ ，则

$$E[|X - Y|] = \underline{\sqrt{2/\pi}}。$$

3、已知 $X \sim N(\mu, \sigma^2)$ ， $x_1, \dots, x_n$ 来自总体 X 的样本 ($n > 1$)，则 $\sum_{i=1}^n \left(\frac{x_i - \mu}{\sigma} \right)^2$

服从的分布为 $\chi^2(n)$ 。

4、已知 $X \sim N(\mu, \sigma^2)$ ， $x_1, \dots, x_n$ 来自总体 X 的样本 ($n > 1$)，其样本均值和

样本方差分别记为 $\bar{x}$ 和 s^2 。 μ 的置信度为 $1 - \alpha$ 的右侧置信区间是

$$\left(-\infty, \bar{x} + t_{1-\alpha}(n-1) \frac{s}{\sqrt{n}} \right] ; \text{ 相应的所用枢轴量为 } t = \frac{\bar{x} - \mu}{s/\sqrt{n}} \sim t(n-1)$$

(注：指明分布)。

5、概率的定义主要有 统计学定义和公理化定义两种定义法；在本课程学习中主要采用 公理化定义。

二、〔计 10 分〕设有甲、乙、丙三人三人比赛，规定：每局两人比赛，胜者与第三人比赛，依次循环，直到有一人连胜两局为止，此人即为冠军。假设每局比赛双方取胜的概率都是 0.5，且局与局之间的胜负结果是相互独立的。现甲、乙两人率先比赛，试求各人获得冠军的概率。

解：记甲、乙、丙在第 i 局获胜的事件分别为 A_i 、 B_i 和 C_i ，有条件知：

$$P(A_i) = P(B_i) = P(C_i) = 0.5, \quad i = 1, 2, \dots$$

现甲、乙两人率先比赛，可分为甲先胜和乙先胜两种情形。若甲先胜，且甲获得冠军，则相应的事件为：

$$A_1 A_2, A_1 C_2 B_3 A_4 A_5, A_1 C_2 B_3 A_4 C_5 B_6 A_7 A_8, \dots$$

且这些事件是互不相容的，由局与局之间的胜负结果是相互独立的条件知甲获得冠军的概率为

$$\begin{aligned} & P(A_1 A_2) + P(A_1 C_2 B_3 A_4 A_5) + P(A_1 C_2 B_3 A_4 C_5 B_6 A_7 A_8) + \dots \\ &= \sum_{k=1}^{+\infty} 0.5^{3k-1} = \frac{0.25}{1-0.5^3} = \frac{2}{7}。 \end{aligned}$$

若乙先胜，且甲获得冠军，则相应的事件为：

$$B_1 C_2 A_3 A_4, B_1 C_2 A_3 B_4 C_5 A_6 A_7, B_1 C_2 A_3 B_4 C_5 A_6 B_7 C_8 A_9 A_{10}, \dots$$

类似地，可得甲获得冠军的概率为

$$P(B_1C_2A_3A_4) + P(B_1C_2A_3B_4C_5A_6A_7) + P(B_1C_2A_3B_4C_5A_6B_7C_8A_9A_{10}) + \dots$$

$$= \sum_{k=1}^{+\infty} 0.5^{3k+1} = \frac{0.5^4}{1-0.5^3} = \frac{1}{14}。$$

从而，甲获得冠军的概率为 $\frac{2}{7} + \frac{1}{14} = \frac{5}{14}。$

进一步，有对称性可知：乙获得冠军的概率也为 $\frac{5}{14}。$ 因此，丙获得冠军的概率为 $1 - \frac{5}{14} \times 2 = \frac{2}{7}。$

三、【计 14 分】设二维随机变量 (X, Y) 的联合概率密度函数是

$$p(x, y) = \begin{cases} k \cdot x \cdot (1-y), & 0 < x^2 < y < x < 1 \\ 0, & \text{其它} \end{cases}$$

其中 k 为常数，且 $k > 0$ ，试求：

- (1) 常数 k ；
- (2) 判别随机变量 X 与 Y 是否相互独立；
- (3) 设 $Z = X + Y$ ，求 Z 的概率密度函数 $p_Z(z)$ 。

解：(1) 由 $1 = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} p(x, y) dy = k \int_0^1 x dx \int_{x^2}^x (1-y) dy = \frac{k}{24}$ 得： $k = 24。$

$$(2) \quad p_X(x) = \int_{-\infty}^{+\infty} p(x, y) dy$$

$$= \begin{cases} \int_{x^2}^x 24x(1-y) dy = 12x^2(x^3 - 3x + 2), & 0 < x < 1 \\ 0, & \text{其它} \end{cases}$$

$$p_Y(y) = \int_{-\infty}^{+\infty} p(x, y) dx = \begin{cases} \int_y^{\sqrt{y}} 24x(1-y) dx = 12y(1-y)^2, & 0 < y < 1 \\ 0, & \text{其它} \end{cases}。$$

因 $p_X(x) \cdot p_Y(y) \neq p(x, y)$ ，则随机变量 X 与 Y 是不独立的。

$$(3) \quad p_Z(z) = \int_{-\infty}^{+\infty} p(x, z-x) dx$$

$$= \begin{cases} \int_{z/2}^{(\sqrt{1+4z}-1)/2} 24x(1+x-z) dx & 0 < z < 2 \\ = 2 - 6z - 15z^2 + 2z^3 + (10z-2)\sqrt{1+4z}, & \\ 0, & \text{其它} \end{cases}。$$

四、【11 分】设 $X \sim \text{Exp}(\lambda^{-1})$ ， $x_1, \dots, x_n$ 来自总体 X 的样本 ($n > 1$)，试求

参数 λ 的最大似然估计量, 并判断它是否是相合估计和无偏估计。

解: 由 $X \sim \text{Exp}(\lambda^{-1})$, 可得似然函数为

$$L(x_1, x_2, \dots, x_n; \lambda) = \prod_{i=1}^n \lambda^{-1} \exp(-x_i/\lambda) = \lambda^{-n} \exp(-\lambda^{-1} \sum_{i=1}^n x_i)$$

即 $\ln L(x_1, x_2, \dots, x_n; \lambda) = -n \ln \lambda - \lambda^{-1} \sum_{i=1}^n x_i$ 。由

$$\frac{d \ln L}{d \lambda} = -\frac{n}{\lambda} + \frac{1}{\lambda^2} \sum_{i=1}^n x_i = 0$$

解得 $\hat{\lambda} = \bar{x}$, 此即为参数 λ 的最大似然估计。由于

$$E(\hat{\lambda}) = E(\bar{x}) = \frac{1}{n} E\left(\sum_{i=1}^n x_i\right) = \frac{1}{n} \cdot n \cdot \frac{1}{\lambda^{-1}} = \lambda$$

因此, $\hat{\lambda} = \bar{x}$ 是参数 λ 的无偏估计。

又因为 $\lim_{n \rightarrow \infty} E(\hat{\lambda}) = \lambda$, 且 $\lim_{n \rightarrow \infty} \text{Var}(\hat{\lambda}) = \lim_{n \rightarrow \infty} \text{Var}(\bar{x}) = \lim_{n \rightarrow \infty} \frac{\lambda^2}{n} = 0$, 因此 $\hat{\lambda} = \bar{x}$ 也是参数 λ 的相合估计。

五、【计 10 分】已知机器 A 和 B 都生产同一型号的钢管, 记它们生产的钢管内径分别为 X 和 Y 。现从机器 A 和 B 生产的钢管中随机抽出 19 根和 13 根进行检测, 测得它们的样本方差分别为 $s_1^2 = 0.34$ 和 $s_2^2 = 0.29$ 。假设 X 和 Y 是相互独立, 且都服从正态分布。试问是否能认为两台机器生产的钢管内径的稳定程度 (即方差) 无显著差异? (取 $\alpha = 0.1$)

解: 1) 提出假设 $H_0: \sigma_1^2 = \sigma_2^2$, $H_1: \sigma_1^2 \neq \sigma_2^2$;

2) 构造检验统计量: $F = \frac{s_1^2}{s_2^2}$

3) 对于显著水平 $\alpha = 0.1$, 双侧检验的拒绝域为 $F \leq F_{\alpha/2}(m-1, n-1)$ 或

$$F \geq F_{1-\alpha/2}(m-1, n-1), \quad \text{此处 } m=19, n=13,$$

$$F_{0.95}(18, 12) = 2.57, F_{0.05}(18, 12) = 1/2.34, \text{ 由题意得拒绝域为}$$

$$W = \{F \leq 0.4274 \text{ 或 } F \geq 2.57\}$$

4) 由样本值 $s_1^2 = 0.34$ 和 $s_2^2 = 0.29$ 。算得 $F = \frac{s_1^2}{s_2^2} = \frac{0.34}{0.29} = 1.172$;

5) 未落入拒绝域, 所以接受 H_0 , 认为两台机器生产的钢管内径的稳定程度 (即方差) 无显著差异。

六、【计 15 分】某工厂在分析产量与成本关系时, 选取 10 个生产小时作样本, 测得数据如下:

$$\sum_{i=1}^{10} x_i = 777, \sum_{i=1}^{10} y_i = 1629, \sum_{i=1}^{10} x_i^2 = 70903,$$

$$\sum_{i=1}^{10} y_i^2 = 267723, \sum_{i=1}^{10} x_i y_i = 131124.$$

假定成本 y 与产量 x 间具有近似线性关系，试求：（1） y 对 x 的线性回归方程；（2）对建立的回归方程进行显著性检验（ $\alpha = 0.05$ ）；（3）当 $x = 80.5$ ，计算 y 的预测区间（ $\alpha = 0.05$ ）。

解：（1） $l_{xx} = \sum_{i=1}^{10} x_i^2 - 10\bar{x}^2 = 70903 - 10 \times \left(\frac{777}{10}\right)^2 \approx 10530.1$,

$$l_{xy} = \sum_{i=1}^{10} x_i y_i - 10\bar{x}\bar{y} = 131124 - 10 \times \left(\frac{777}{10}\right) \times \left(\frac{1629}{10}\right) \approx 4550.7,$$

$$l_{yy} = \sum_{i=1}^{10} y_i^2 - 10\bar{y}^2 = 267723 - 10 \times \left(\frac{1629}{10}\right)^2 \approx 2358.9$$

$$\hat{\beta}_1 = \frac{l_{xy}}{l_{xx}} \approx \frac{4550.7}{10530.1} \approx 0.4322,$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \approx 162.9 - 0.4322 \times 77.7 = 129.32$$

则 y 关于 x 的回归直线方程为： $\hat{y} = 129.32 + 0.4322x$ 。

（2）对回归方程显著性检验，经计算有：

$$S_T = l_{yy} = 2358.9 \quad f_T = 9$$

$$S_R = \hat{\beta}_1^2 l_{xx} = 0.4322^2 \times 10530.1 = 1966.99 \quad f_R = 1$$

$$S_e = S_T - S_R = 2358.9 - 1966.99 = 391.91 \quad f_e = 8$$

对回归方程进行显著性检验（ $\alpha = 0.05$ ）

（2）由题意得拒绝域为 $F \geq F_{1-\alpha}(1, 8)$ ，查表 $F_{0.95}(1, 8) = 5.32$ 。由样本值

$$F = \frac{S_R}{S_e/8} = \frac{1966.99}{391.91/8} = 40.152, \text{ 因为, } F \geq F_{1-\alpha}(1, 8), \text{ 落入拒绝域}$$

因此，在显著水平 $\alpha = 0.05$ 下建立的回归方程是显著的。

（3）当 $x = 80.5$ 时， $\hat{y}_0 = 164.11$ ，在 $\alpha = 0.05$ 时， $t_{0.975}(8) = 2.3060$ ，

$$\hat{\sigma} = \sqrt{\frac{391.91}{8}} = 6.99, \text{ 则}$$

$$\delta = 6.99 \times 2.3060 \times \sqrt{1 + \frac{1}{10} + \frac{(80.5 - 77.7)^2}{10530.1}} = 16.91$$

从而 y_0 的概率为 0.95 的预测区间 $164.11 \pm 16.91 = (147.2, 181.02)$ 。

七、【计 10 分】设随机变量 $X \sim Ga(\alpha, \lambda)$ ，证明：当 $\alpha \rightarrow +\infty$ 时，随机变量

$Y_\alpha = (\lambda X - \alpha)/\sqrt{\alpha}$ 依分布收敛于标准正态变量，即

$$\lim_{\alpha \rightarrow +\infty} P(Y_\alpha \leq y) = \Phi(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-t^2/2} dt$$

其中随机变量 X 的特征函数为 $\varphi_X(t) = (1 - it/\lambda)^{-\alpha}$ 。

证明：由 $\varphi_X(t) = (1 - it/\lambda)^{-\alpha}$ ，则由特征函数的性质知 $Y_\alpha = (\lambda X - \alpha)/\sqrt{\alpha}$ 的特征函数为： $\varphi_{Y_\alpha}(t) = e^{-i\sqrt{\alpha}t} (1 - it/\sqrt{\alpha})^{-\alpha}$ 。从而

$$\ln \varphi_{Y_\alpha}(t) = -i\sqrt{\alpha}t - \alpha \ln(1 - it/\sqrt{\alpha}) = -\frac{t^2}{2} + o\left(\frac{1}{\sqrt{\alpha}}\right)$$

因此， $\lim_{\alpha \rightarrow +\infty} \varphi_{Y_\alpha}(t) = e^{-\frac{1}{2}t^2}$ 。由于 $e^{-\frac{1}{2}t^2}$ 是标准正态分布的特征函数，故由特征函数的唯一性定理知：结论成立。

八、【计 5 分】东汉末年，曹操和袁绍之间爆发了著名的官渡之战。战争初期，曹操实力远弱于袁绍，节节后退。在无法继续后退的地方，即官渡，曹操据险坚守，陷入苦战，危在旦夕。但在此时，袁绍的谋士许攸深夜投奔曹操。曹操听从许攸的计谋，绕道前往乌巢，烧了袁绍的军粮。事件的结果：袁绍军心大乱，曹操以弱胜强。面对这一结果：有人说许攸的叛变是随机事件；也有人说许攸的叛变是必然事件——其必然性是由袁绍的用人特点和许攸的性格特点共同决定的。

在科学研究中，你该如何认定一件事是随机事件？

答：要紧紧扣以下两点展开说明：（1）随机事件是随机试验中样本空间的子集；（2）它要具有随机性特点，即在试验之前，其发生与否不能准确预知，但在大量重复试验中具有某种统计规律。具体回答这里略去。