

《概率论与数理统计》期末考试卷(B) 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 A, B, C 三个事件中恰好有两个发生的事件可

表示为 $\overline{A}BC \cup A\overline{B}C \cup AB\overline{C}$ 。

2、设随机变量 X 与 Y 相互独立, 且 $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$, 则

$Z = aX + bY \sim$ $N(a\mu_1 + b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2)$ (其中 a 和 b 为常数, 且 $ab \neq 0$)。

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 。 σ^2 的置信度为 $1-\alpha$ 的双侧置信区间是

$$\left[\frac{(n-1)s^2}{\chi_{1-\alpha/2}^2(n-1)}, \frac{(n-1)s^2}{\chi_{\alpha/2}^2(n-1)} \right]; \text{ 相应的所用枢轴量为 } \underline{\frac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1)}$$

(注: 指明分布)。

5、本教程中参数估计方法主要有 点估计和区间估计。

二、〔计 12 分〕有两箱零件: 第一箱装 50 件, 其中 20 件为一等品; 第二箱装 30 件, 其中 18 件为一等品。现从两箱中随意挑出一箱, 然后从该箱中先后任取两个零件 (注: 不放回抽样)。试求:

(1) 第一次取出的零件是一等品的概率;

(2) 已知第一次取出的是一等品的条件下, 第二次取出的零件仍是一等品的概率。

解: 记取出第 i 箱的事件为 A_i , 第 j 次取出的是一等品的事件分别为 B_j ,

$j=1, 2$ 。有条件知:

$$P(A_1) = P(A_2) = 0.5$$

$$(1) P(B_1) = P(A_1)P(B_1|A_1) + P(A_2)P(B_1|A_2)$$

$$= 0.5 \times \frac{C_{20}^1}{C_{50}^1} + 0.5 \times \frac{C_{18}^1}{C_{30}^1} = 0.5;$$

(2) 由

$$\begin{aligned}
 P(B_1 B_2) &= P(A_1)P(B_1 | A_1)P(B_2 | A_1 B_1) + P(A_2)P(B_1 | A_2)P(B_2 | A_2 B_1) \\
 &= 0.5 \times \frac{C_{20}^1}{C_{50}^1} \times \frac{C_{19}^1}{C_{49}^1} + 0.5 \times \frac{C_{18}^1}{C_{30}^1} \times \frac{C_{17}^1}{C_{29}^1} = \frac{3601}{14210}。
 \end{aligned}$$

因此, $P(B_2 | B_1) = \frac{P(B_1 B_2)}{P(B_1)} = \frac{3601}{7105}$ (或 0.5068)。

三、【计 10 分】设随机变量 $X \sim N(0,1)$, 试求 $Y = X^2$ 的概率密度函数 $P_Y(y)$ 。

解: 由 $X \sim N(0,1)$ 知: $Y = X^2$ 的取值范围为 $[0, +\infty)$ 。

$\forall y \in R$, 有: 当 $y \leq 0$ 时, $F_Y(y) = 0$;

当 $y > 0$ 时, $F_Y(y) = P(Y = X^2 \leq y) = P(-\sqrt{y} \leq X \leq \sqrt{y})$

$$= \Phi(\sqrt{y}) - \Phi(-\sqrt{y}) = 2\Phi(\sqrt{y}) - 1$$

因此, $F_Y(y) = \begin{cases} 2\Phi(\sqrt{y}) - 1 & y > 0 \\ 0 & y \leq 0 \end{cases}$ 。于是,

$$p_Y(y) = F'_Y(y) = \begin{cases} \frac{e^{-\frac{y}{2}}}{\sqrt{2\pi}}, & y > 0 \\ 0, & y \leq 0 \end{cases}$$

四、【12 分】设随机变量 X 的期望和方差都存在, 分别记为 μ 和 σ^2 。

$x_1, \dots, x_n$ 是来自总体 X 的样本 ($n > 1$), 试求参数 μ 和 σ^2 的矩估计量, 并判断它是否是无偏估计。

解: 由题给条件, 可令:
$$\begin{cases} \mu = E(X) = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \\ \mu^2 + \sigma^2 = E(X^2) = \overline{x^2} = \frac{1}{n} \sum_{i=1}^n x_i^2 \end{cases}$$

解得:
$$\begin{cases} \hat{\mu} = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \\ \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2 \end{cases}$$

由于

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

$$E(\hat{\sigma}^2) = E\left(\frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2\right) = \frac{n-1}{n} \cdot \sigma^2 \neq \sigma^2$$

因此, $\hat{\mu} = \bar{x}$ 是参数 μ 的无偏估计, $\hat{\sigma}^2$ 是参数 σ^2 的有偏估计。

五、【计 10 分】为比较两个小麦品种的产量, 选择 18 块条件相似的试验田, 采用相同的耕作方法做试验。现测得播种甲品种 X 的 8 块试验田的产量和播种乙品种 Y 的 10 块试验田的产量 (单位: kg/亩), 计算得它们的样本均值和方差分别是 $\bar{x} = 569.38$, $S_x^2 = 2140.55$; $\bar{y} = 487.00$,

$S_y^2 = 3256.22$ 。假设 X 和 Y 是相互独立, 且都服从正态分布。试检验: 在

显著性水平 $\alpha = 0.05$ 下, 两个小麦品种的产量的均值是否相等?

解: 不妨设 $\sigma_1^2 = \sigma_2^2 = \sigma^2$

1) 提出假设 $H_0: \mu_1 = \mu_2$, $H_1: \mu_1 \neq \mu_2$;

2) 构造检验统计量: $t = \frac{\bar{x} - \bar{y}}{S_w \sqrt{\frac{1}{m} + \frac{1}{n}}} \sim t(m+n-2)$, 其中

$$S_w = \sqrt{\frac{(m-1)S_x^2 + (n-1)S_y^2}{m+n-2}}$$

3) 对于显著水平 $\alpha = 0.05$, 双侧检验的拒绝域为 $F \leq F_{\alpha/2}(m-1, n-1)$ 或 $|t| \geq t_{1-\alpha/2}(m+n-2)$, 此处 $m=8, n=10$, $t_{0.975}(16) = 2.1199$, 由题意得拒绝域为 $W = \{t \mid |t| > 2.1199\}$ 。

4) 由样本值 $\bar{x} = 569.38$, $\bar{y} = 487.00$, $S_x^2 = 2140.55$ 和 $S_y^2 = 3256.22$, 算得 $|t| = \frac{569.38 - 487}{S_w \sqrt{\frac{1}{8} + \frac{1}{10}}} = 3.3009$;

5) 因落入拒绝域, 所以拒绝 H_0 , 接受 H_1 , 即认为两个小麦品种的产量的均值是不相等的。

六、【计 14 分】现收集了 16 组金钢中的碳含量 x 与强度 y 的数据, 整理数据如下:

$$\bar{x} = 0.125, \bar{y} = 45.789, l_{xx} = 0.3024, l_{xy} = 25.522, l_{yy} = 2432.457$$

假定 Y 对 X 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 $x = 0.20$, 计算 y 的预测区间 ($\alpha = 0.05$)。

解: (1) 由 $\hat{\beta}_1 = \frac{l_{xy}}{l_{xx}} = \frac{25.522}{0.3024} \approx 84.3981$,

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 45.789 - 84.3981 \times 0.125 = 35.2392$$

则 y 关于 x 的回归直线方程为: $\hat{y} = 35.2392 + 84.3981x$ 。

(2) 对回归方程显著性检验, 经计算有:

$$S_T = l_{yy} = 2432.457, \quad f_T = 15$$

$$S_R = \hat{\beta}_1^2 l_{xx} = 84.3981^2 \times 0.3024 = 2154.0081, \quad f_R = 1$$

$$S_e = S_T - S_R = 2432.457 - 2154.0081 = 278.4489, \quad f_e = 14$$

由题意得拒绝域为 $F \geq F_{1-\alpha}(1, 14)$, 查表 $F_{0.95}(1, 14) = 4.60$ 。由计算得

$$\hat{F} = \frac{S_R}{S_e/14} = \frac{2154.0081}{278.4489/14} = 108.3004$$

由于 $\hat{F} \geq F_{1-\alpha}(1, 8)$, 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 $x = 0.2$ 时, $\hat{y} = 52.1188$ 。在 $\alpha = 0.05$ 时, $t_{0.975}(14) = 2.1448$,

$$\hat{\sigma} = \sqrt{\frac{278.4489}{14}} = 4.4597, \quad \text{则}$$

$$\delta = 4.4597 \times 2.1448 \times \sqrt{1 + \frac{1}{16} + \frac{(0.2 - 0.125)^2}{0.3024}} = 9.9455$$

从而 $\hat{y}$ 的置信度为 0.95 的预测区间

$$52.1188 \pm 9.9455 = (42.1733, 62.0643)。$$

七、【计 12 分】设随机变量 $X \sim Ga(\alpha, \lambda)$, 其特征函数为

$$\varphi_X(t) = (1 - it/\lambda)^{-\alpha}$$

试求: (1) 利用特征函数计算 $E(X)$ 和 $Var(X)$;

(2) 利用特征函数证明: Γ -分布对参数 α 具有可加性。

证明: (1) 由 $\varphi_X(t) = (1 - it/\lambda)^{-\alpha}$, 则

$$\varphi'_X(t) = \frac{\alpha i}{\lambda} (1 - it/\lambda)^{-\alpha-1}, \quad \varphi''_X(t) = -\frac{\alpha(\alpha+1)}{\lambda^2} (1 - it/\lambda)^{-\alpha-2}$$

于是, 有

$$E(X) = \frac{\varphi'_X(0)}{i} = \frac{\alpha}{\lambda}, \quad D(X) = -\varphi''_X(0) + [\varphi'_X(0)]^2 = \frac{\alpha}{\lambda^2}。$$

(2) 假设 $X_1 \sim Ga(\alpha_1, \lambda)$, $X_2 \sim Ga(\alpha_2, \lambda)$, 且 X_1 与 X_2 独立。由特征函数的性质知:

$$\varphi_{X_1+X_2}(t) = \varphi_{X_1}(t)\varphi_{X_2}(t) = (1 - it/\lambda)^{-\alpha_1} \times (1 - it/\lambda)^{-\alpha_2} = (1 - it/\lambda)^{-(\alpha_1+\alpha_2)}$$

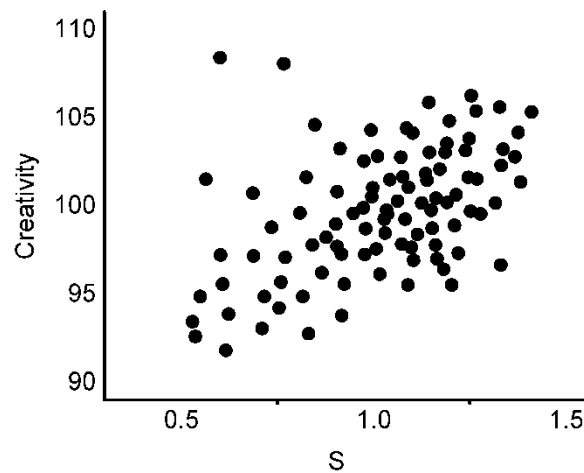
由随机变量的特征函数唯一性知: $X_1 + X_2 \sim Ga(\alpha_1 + \alpha_2, \lambda)$, 即结论成立。

八、【计 5 分】人的大脑, 表面是一层灰色的褶皱状“皮质”, 里面充满了“白

质”。不同的皮质位置 i ，执行不同的功能。白质的功能是传递信息，使得皮质的任意两个位置有可能相互交换信息。如今，运用核磁共振技术，可以实时检测大脑皮质的活性变化。有人测量了大脑皮层活跃程度的熵 S （其定义为 $S = -\sum_i p_i \ln p_i$ ，其中 p_i 表示大脑皮层的某位置 i 处于活跃状态的概率。

值 S 越大，则表明大脑运转的随机性越强）与创造性思维能力 **Creativity** 的关系，结果如下图所示。

试问： S 与 **Creativity** 的关系是正相关还是负相关？请说明理由或列出研究计划。



答：正相关。具体回答这里略去。