

2015-2016(2)信计专业

《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 A, B, C 中至少有一个不发生的事件可表示为 $\overline{A \cup B \cup C}$ 或 $\overline{ABC}$ 。
- 2、已知随机事件 A, B 满足： $P(A) = 1/2$ ， $P(B) = 1/3$ ， $P(\overline{A} | B) = 1/6$ ，则 $P(A | A \cup B) = \underline{0.9}$ 。
- 3、设 $X \sim \chi^2(m)$ ， $Y \sim \chi^2(n)$ ， X 与 Y 相互独立，则 $Z = \frac{X/m}{Y/n}$ 服从的分布为 $\underline{F(m, n)}$ 。
- 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 。在 μ 未知的条件下， σ^2 的置信度为 $1 - \alpha$ 的双侧置信区间是 $\left[\frac{(n-1)s^2}{\chi_{1-\alpha/2}^2(n-1)}, \frac{(n-1)s^2}{\chi_{\alpha/2}^2(n-1)} \right]$ 。
- 5、假设每个人生日的月份是随机的。现某班级有 30 位同学，则至少有两位同学在同一月份过生日的概率是 $\underline{1}$ 。

二、【计 10 分】有 m ($m > 1$) 个人相互传球，球开始时由甲传出。假设每次传球者等可能的把球传给其余的 $m-1$ 个人。求第 n 次传球时仍有甲传出的概率。

解：记第 n 次由甲传球的事件为 A_n ，相应的概率为 p_n ($n = 1, 2, 3, \dots$)。显然，当 $n = 1$ ，有 $p_1 = p(A_1) = 1$ 。一般地，当 $n > 1$ 时，有

$$\begin{aligned} p_n &= p(A_n) = p(A_{n-1}) \cdot p(A_n | A_{n-1}) + p(\overline{A_{n-1}}) \cdot p(A_n | \overline{A_{n-1}}) \\ &= \frac{1}{m-1} [1 - p_{n-1}] \end{aligned}$$

因此可得 p_n 的递推公式如下：

$$\begin{cases} p_n = \frac{1}{m-1} [1 - p_{n-1}] \\ p_1 = 1, n = 2, 3, \dots \end{cases}$$

容易求得： $p_n = \frac{1}{m} \left[1 - \left(-\frac{1}{m-1} \right)^{n-1} \right]$ ($n = 2, 3, \dots$)。

三、【计 12 分】设随机变量 X 和 Y 相互独立，且都服从参数为 λ 的指数分布。令

$$\begin{cases} U = X + Y \\ V = X / Y \end{cases}$$

试求：(1) (U, V) 的联合概率密度函数 $p_{U,V}(u, v)$ ；(2) $U = X + Y$ 的概率密

度函数 $p_U(u)$; (3) 判断随机变量 U 与 V 是否独立。

解: (1) 由 $\begin{cases} u = x + y \\ v = x / y \end{cases}$, 可得 $\begin{cases} x = uv / (v + 1) \\ y = u / (v + 1) \end{cases}$ 。于是

$$J = \frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} v / (v + 1) & u / (v + 1)^2 \\ 1 / (v + 1) & -u / (v + 1)^2 \end{vmatrix} = -\frac{u}{(v + 1)^2}$$

因此, (U, V) 的联合概率密度函数为

$$p_{U,V}(u, v) = p_{X,Y}(x(u, v), y(u, v)) \cdot |J| = \begin{cases} \frac{\lambda^2 u e^{-\lambda u}}{(v + 1)^2}, & u \geq 0, v > 0 \\ 0, & \text{其它} \end{cases}$$

$$(2) \quad p_U(u) = \int_{-\infty}^{+\infty} p_{U,V}(u, v) dv = \begin{cases} \lambda^2 u e^{-\lambda u}, & u \geq 0 \\ 0, & u < 0 \end{cases}$$

(3) 类似 (2) 可求: 由公式可得:

$$p_V(v) = \int_{-\infty}^{+\infty} p_{U,V}(u, v) du = \begin{cases} \frac{1}{(v + 1)^2}, & v > 0 \\ 0, & v \leq 0 \end{cases}$$

因为 $p_{U,V}(u, v) = p_U(u)p_V(v)$, 故随机变量 U 与 V 独立。

四、【计 12 分】设 $x_1, \dots, x_n$ 是来自总体 $X \sim U(0, \theta)$ 的一个样本。试求参数 θ 的矩估计和最大似然估计量, 并讨论它们是否为无偏估计?

解: 由于 $X \sim U(0, \theta)$, 因此 $E(X) = \theta / 2$, 令

$$\theta / 2 = E(X) = \bar{x}$$

可得参数 θ 的矩估计为: $\hat{\theta}_1 = 2\bar{x}$ 。因 $E(\hat{\theta}_1) = E(2\bar{x}) = 2 \cdot \theta / 2 = \theta$, 因此, $\hat{\theta}_1 = 2\bar{x}$ 是参数 θ 的无偏估计。

由 $X \sim U(0, \theta)$, 即 $p(x) = \begin{cases} 1/\theta, & 0 < x < \theta \\ 0, & \text{其他} \end{cases}$, 则似然函数为

$$L(x_1, x_2, \dots, x_n; \lambda) = \prod_{i=1}^n p(x_i, \lambda) = \frac{1}{\theta^n}$$

由于 $L(x_1, x_2, \dots, x_n; \lambda) = \frac{1}{\theta^n} \leq \frac{1}{[\max\{x_1, \dots, x_n\}]^n}$, 因此 θ 的最大似然估计

量为 $\hat{\theta}_2 = \max\{x_1, \dots, x_n\}$ 。

由于 $X \sim U(0, \theta)$, 其分布函数为

$$F(x) = \begin{cases} 0, & x \leq 0 \\ x/\theta, & 0 < x < \theta, \\ 1, & x \geq \theta \end{cases}$$

因此 $Y = \max\{x_1, \dots, x_n\}$ 的概率密度函数为

$$P_Y(y) = nF^{n-1}(y) \cdot P(y) = \begin{cases} \frac{n \cdot y^{n-1}}{\theta^n}, & 0 < x < \theta \\ 0, & \text{其它} \end{cases}$$

因 为 $E(\hat{\theta}_2) = E(Y) = \int_0^\theta y \cdot \frac{n \cdot y^{n-1}}{\theta^n} dy = \frac{n}{n+1} \theta \neq \theta$, 所 以

$\hat{\theta}_2 = \max\{x_1, \dots, x_n\}$ 不是参数 θ 的无偏估计。

五、【计 14 分】已知某维尼纶纤维的耐水性能与其“缩醇化度”密切相关，且它的甲醇浓度 x 与“缩醇化度” y 间具有近似的线性关系。现测得 7 组数据，整理如下： $\bar{x} = 24.00$, $\bar{y} = 28.99$, $l_{xx} = 112.0000$, $l_{xy} = 29.6000$, $l_{yy} = 9.0729$ 。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 $x = 25.4$ ，计算 y 的预测值。

解：(1) 由 $\hat{\beta}_1 = l_{xy}/l_{xx} = 0.2643$, $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 22.6468$ ，因此， y 对 x 的线性回归方程为： $\hat{y} = 22.6468 + 0.2643x$ ；

(2) 由 $S_R = \hat{\beta}_1^2 l_{xx} = 7.8237$, $S_e = S_T - S_R = l_{yy} - S_R = 1.2492$ ，可得下列方差分析表

来源	平方和	自由度	均方	F 比
回 归	$S_R = 7.8237$	$f_R = 1$	$MS_R = 7.8237$	31.3148
残 差	$S_e = 1.2492$	$f_e = 5$	$MS_e = 0.2498$	
总 计	$S_T = 9.0729$	$f_T = 6$		

由于 $\hat{F} = 31.3148 > F_{0.95}(1, 5) = 6.61$ ，因此 (1) 中回归方程是显著的；

(3) 当 $x = 25.4$ ， y 的预测值为 $\hat{y} = 29.3600$ 。

六、【计 11 分】已知正常成年男 (X) 和女 (Y) 所含红血球 (单位: $10^4/\text{mm}^3$) 都服从正态分布，且 $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$ 。为研究某地正常成年男和女所含红血球的差异，今随机抽取该地成年男性 15 人和女性 7 人进行检测，测得样本均值和方差如下： $\bar{x} = 465.13$, $\bar{y} = 422.16$, $S_x^2 = 54.80^2$, $S_y^2 = 49.20^2$ 。假定 $\sigma_1^2 = \sigma_2^2 = \sigma^2$ ，试检验：该地区正常成年男女所含红血球的平均值是否有差异? (取 $\alpha = 0.05$)

解：检验假设： $H_0: \mu_1 = \mu_2$ vs $H_1: \mu_1 \neq \mu_2$

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

其中 $m = 15$, $n = 7$, $S_w^2 = \frac{(m-1)S_x^2 + (n-1)S_y^2}{m + n - 2}$ 。

当 $\alpha = 0.05$ 时, 拒绝域为

$$W = \{t \mid |t| > t_{1-\alpha/2}(m+n-2)\} = \{t \mid |t| > 2.0860\}$$

由 $\bar{x} = 465.13$, $\bar{y} = 422.16$, $S_x^2 = 54.80^2$, $S_y^2 = 49.20^2$, 可得:

$$S_w = \sqrt{\frac{14 \times 54.80^2 + 6 \times 49.20^2}{20}} = 53.182$$

$$|\hat{t}| = \left| \frac{465.13 - 422.16}{S_w \sqrt{1/15 + 1/7}} \right| = 1.7652 \notin W, \text{ 因此, 可以认为该地区正常成年男}$$

女所含红血球的平均值没有差异。

七、【计 10 分】 设随机变量 $X \sim Ga(\alpha, \lambda)$, 其特征函数为

$$\varphi_X(t) = \left(1 - \frac{it}{\lambda}\right)^{-\alpha}$$

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。

解: (1) 由 $\varphi'_X(t) = \frac{\alpha i}{\lambda} \left(1 - \frac{it}{\lambda}\right)^{-\alpha-1}$, $\varphi''(t) = -\frac{\alpha(\alpha+1)}{\lambda^2} \left(1 - \frac{it}{\lambda}\right)^{-\alpha-2}$, 可

得: $E(X) = \frac{\varphi'_X(0)}{i} = \frac{\alpha}{\lambda}$, $E(X^2) = -\varphi''_X(0) = \frac{\alpha(\alpha+1)}{\lambda^2}$ 。因此, 有:

$$E(X) = \frac{\alpha}{\lambda}, \quad Var(X) = E(X^2) - E^2(X) = \frac{\alpha}{\lambda^2}。$$

证明: (2) 不妨设 $X \sim Ga(\alpha, \lambda)$, $Y \sim Ga(\beta, \lambda)$, 且 X 与 Y 独立。由特征函数的性质可得 $X+Y$ 的特征函数为

$$\varphi_{X+Y}(t) = \varphi_X(t) \cdot \varphi_Y(t) = \left(1 - \frac{it}{\lambda}\right)^{-\alpha} \cdot \left(1 - \frac{it}{\lambda}\right)^{-\beta} = \left(1 - \frac{it}{\lambda}\right)^{-(\alpha+\beta)}$$

由随机分布特征函数的唯一性知: $\left(1 - \frac{it}{\lambda}\right)^{-(\alpha+\beta)}$ 所对应的分布正是参数为

$\alpha + \beta$ 和 λ 的 Γ -分布, 即 $X + y \sim Ga(\alpha + \beta, \lambda)$, 故结论成立。

八、【计 6 分】 试简述本课程 CH4 “大数定律和中心极限定理”的主要内容, 并说明在本课程中的作用。

解: CH4 “大数定律和中心极限定理”的主要内容包括三个部分: (1) 随机变量的特征函数及其性质; (2) 大数定律; (3) 中心极限定理。在本课程中起到“承上启下”的作用, 即大数定律是对前三章概率论部分的深化和总结, 而中心极限定理是对把握后四章, 即数理统计部分内容, 至关重要。