# 四川大学期末考试试题 (闭卷)

**(2021 — 2022 学年第 1 学期)　A 卷**

| | | | | |
|---|---|---|---|---|
| 课程号: 201162040 | 课序号: 01,02 | 课程名称: 概率论 | 任课教师: 常寅山, 彭雪 | 成绩: |
| 适用专业年级: 2020 级等 | 学生人数: | 印题份数: 300 | 学号: | 姓名: |

**考 生 承 诺**

我已认真阅读并知晓《四川大学考场规则》和《四川大学本科学生考试违纪作弊处分规定（修订）》，郑重承诺：

1. 已按要求将考试禁止携带的文具用品或与考试有关的物品放置在指定地点；
2. 不带手机进入考场；
3. 考试期间遵守以上两项规定，若有违规行为，同意按照有关条款接受处理。

考生签名:

**注意: 所有答案一律写在答题纸上, 写清题目编号. 每一张答题纸上写明姓名和学号.**

1. (10 分) 证明 $P(\bigcup_{n=1}^{+\infty} A_n) \le \sum_{n=1}^{+\infty} P(A_n)$.

   证明. 令 $B_n = \cup_{m=1}^{n} A_m$, 则 $B_n$ 是单调不减的集合列. 令 $C_n = B_n \setminus B_{n-1}$, 其中, $B_0$ 被定义为空集. (4 分) 那么, $\cup_{n\ge1} A_n = \cup_{n\ge1} C_n$. (2 分) 而由于 $\{C_n\}_{n\ge1}$ 两两互斥, 所以

   $$P(\cup_{n\ge1} C_n) = \sum_{n\ge1} P(C_n).(2 \text{ 分})$$

   注意到 $C_n \subset A_n$, 所以, $P(C_n) \le P(A_n)$. 综上可得结论. (2 分)

2. (10 分) 假设 $0 < P(A) < 1$, $0 < P(B) < 1$ 并且 $P(A|B) = 1$. 证明 $P(B^c|A^c) = 1$.

   证明. 由于 $P(A|B) = 1$, 所以, $P(AB) = P(B)$. (2 分) 所以, $P(A^cB) = P(B) - P(AB) = 0$. (3 分) 所以, $P(A^c) = P(A^cB^c) + P(A^cB) = P(A^cB^c)$. (3 分) 最后, 根据条件概率的定义, $P(B^c|A^c) = P(A^cB^c)/P(A^c) = 1$. (2 分)

3. (20 分) 设随机变量 $X$ 与 $Y$ 独立同分布, 都服从标准正态分布 $N(0,1)$. 考虑 $X = R\cos\Theta$, $Y = R\sin\Theta$, 其中, $R \ge 0$, $\Theta \in [0, 2\pi)$.

   证明: $R$ 和 $\Theta$ 相互独立.

证明. 考虑变换 $x = r\cos\theta$, $y = r\sin\theta$. 其中, $r \geq 0$, $v \in [0, 2\pi)$. (2 分) 变换的雅可比为 $J(r,\theta) = \frac{\partial(x,y)}{\partial(r,\theta)} = r$. (2 分) 根据独立性和题意, $X$ 和 $Y$ 的联合密度函数为

$$f_{X,Y}(x,y) = f_X(x) f_Y(y) = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}}. \text{ (4 分)}$$

所以, $R$ 和 $\Theta$ 的联合密度函数为

$$\begin{aligned} f_{R,\Theta}(r,\theta) &= f_{X,Y}(r\cos\theta, r\sin\theta) J(r,\theta) \\ &= \frac{r}{2\pi} e^{-\frac{r^2}{2}}, r \geq 0, \theta \in [0, 2\pi). \text{ (4 分)} \end{aligned}$$

进一步, 我们可以求出 $R$ 的边缘密度函数:

$$f_R(r) = \int_{-\infty}^{+\infty} f_{R,\Theta}(r,\theta)\, \mathrm{d}\theta = r e^{-r^2/2}, r \geq 0. \text{ (2 分)}$$

类似的, 可以求出 $\Theta$ 的边缘密度函数

$$f_\Theta(\theta) = \int_{-\infty}^{+\infty} f_{R,\Theta}(r,\theta)\, \mathrm{d}r = \frac{1}{2\pi}, \theta \in [0, 2\pi). \text{ (2 分)}$$

最后, 由于

$$f_{R,\Theta}(r,\theta) = f_R(r) f_\Theta(\theta),$$

所以, $R$ 和 $\Theta$ 相互独立. (4 分)

4. (15 分) 证明下列结论:

(a) 假设 $\mathrm{Var}(X) = 0$, 证明 $P(X = E(X)) = 1$.

(b) 考虑三维随机向量 $(X_1, X_2, X_3)$ 及其协方差矩阵 $\Sigma = (\Sigma_{i,j})_{i,j=1,2,3}$, 其中, $\Sigma_{i,j} = \mathrm{Cov}(X_i, X_j)$. 假设 $\det(\Sigma) = 0$. 证明存在实数 $a_1, a_2, a_3, b$ 使得 $P(a_1X_1 + a_2X_2 + a_3X_3 = b) = 1$.

(c) 证明满足 (b) 问中条件的 $(X_1, X_2, X_3)$ 不可能是三维连续型随机向量.

证明,

(a) 根据 Chebyshev 不等式, $\forall k \geq 1$,

$$P(|X - EX| > 1/k) \leq k^2 \mathrm{Var}(X) = 0. \text{ (2 分)}$$

利用概率的连续性, 我们有

$$P(|X - EX| > 0) = \lim_{k\to\infty} P(|X - EX| > 1/k) = 0. \text{ (2 分)}$$

最后, $P(X = EX) = 1 - P(|X - EX| > 0) = 1 - 0 = 1$. (1 分)

(b) 由于 $\det\Sigma = 0$, 所以, 存在 $a_1, a_2, a_3$ 使得

$$\sum_{i,j=1,2,3} \Sigma_{i,j} a_i a_j = 0. \text{ (2 分)}$$

但是,

$$\begin{aligned}\sum_{i,j=1,2,3} \Sigma_{i,j} a_i a_j &= \sum_{i,j=1,2,3} a_i a_j \operatorname{Cov}(X_i, X_j) \\ &= \operatorname{Cov}(\sum_{i=1}^{3} a_i X_i, \sum_{j=1}^{3} a_j X_j) \\ &= \operatorname{Var}(a_1 X_1 + a_2 X_2 + a_3 X_3). \text{ (2 分)}\end{aligned}$$

根据 (a) 问的结果,

$$P(a_1X_1 + a_2X_2 + a_3X_3 = a_1EX_1 + a_2EX_2 + a_3EX_3) = 1. \text{ (1 分)}$$

(c) 反证法. 假设 $(X_1, X_2, X_3)$ 是连续型随机向量, 则存在密度函数 $f(x_1, x_2, x_3)$. (2 分) 并且,

$$\begin{aligned}&P(a_1X_1 + a_2X_2 + a_3X_3 - b = 0) \\ &\quad = \int_{a_1x_1+a_2x_2+a_3x_3-b=0} f(x_1, x_2, x_3)\, \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3 = 0. \text{ (2 分)}\end{aligned}$$

这和 (b) 问矛盾, 所以, (b) 问中的 $(X_1, X_2, X_3)$ 不可能是连续型随机向量. (1 分)

5. (10 分) 假设有栋楼有 $n$ 层. 现在有 $N$ 个人在 1 层等电梯, 其中, $N$ 服从参数为 $\lambda$ 的泊松分布. 每个等电梯的人等可能地在第 2 层至第 $n$ 层下电梯. 并且假设这些人彼此独立. 假设电梯最后停在第 $X$ 层. (如果 $N = 0$, $X = 1$; 如果 $N \geq 1$, $X$ 是最后下电梯的那个人下电梯的楼层.)

(a) 对于 $k = 1, 2, \ldots, n$, 计算 $P(X \leq k)$.

(b) 计算 $P(X = k)$.

解.

(a) 利用全概率公式, 对于 $k \geq 1$,

$$P(X \leq k) = \sum_{m=0}^{+\infty} P(X \leq k | N = m) P(N = m). \text{ (2 分)}$$

注意到 $P(N = m) = \frac{\lambda^m}{m!} e^{-\lambda}$. (2 分) 另外, 注意到

$$P(X \leq k | N = m) = P(m\text{个人都不去}k\text{层以上}) = \left(\frac{k-1}{n-1}\right)^m. \text{ (2 分)}$$

综上,

$$P(X \le k) = \sum_{m=0}^{+\infty} \left(\frac{k-1}{n-1}\right)^m \frac{\lambda^m}{m!} e^{-\lambda} = e^{-\lambda \frac{n-k}{n-1}}.(2\text{ 分})$$

(b) 对于 $k \ge 2$, $P(X = k) = P(X \le k) - P(X \le k-1) = e^{-\lambda \frac{n-k}{n-1}} - e^{-\lambda \frac{n+1-k}{n-1}}$. 而 $P(X=1) = P(X \le 1) = e^{-\lambda}$. (2 分)

6. (18 分) 设一个种群有 $n$ 个个体. 每一个个体的寿命独立同分布于随机变量 $Z$, 并且 $Z$ 服从参数为 $\mu$ 的指数分布. 对任意一个个体而言, 当寿命 $Z = t$ 时, 该个体生育的后代数目服从参数为 $\lambda t$ 的泊松分布. 设各个体之间的寿命, 生育能力是相互独立的. 设 $X_i$ 表示该群体中第 $i$ 个个体生育的后代数, 令 $S_n = X_1 + X_2 + \cdots + X_n$, 即 $S_n$ 表示该群体的所有个体生育的后代总数.

请证明: 当 $n \to +\infty$ 时, $\frac{S_n - n\frac{\lambda}{\mu}}{\sqrt{n(\frac{\lambda^2}{\mu^2} + \frac{\lambda}{\mu})}}$ 依分布收敛于标准正态分布.

证明. 我们先计算 $X_1$ 的期望和方差. 利用重期望公式, 我们有

$$EX_1 = E(E(X_1|Z)).\ (2\text{ 分})$$

根据题设, $E(X_1|Z=t) = \lambda t$. 故, $E(X_1|Z) = \lambda Z$. (2 分) 而 $E(Z) = 1/\mu$. 故, $EX_1 = E(E(X_1|Z)) = E(\lambda Z) = \lambda/\mu$. (2 分) 接下来, 注意到

$$\mathrm{Var}(E(X_1|Z)) = \mathrm{Var}(\lambda Z) = \lambda^2 \mathrm{Var}(Z).\ (2\text{ 分})$$

而 $Z$ 服从参数是 $\mu$ 的指数分布, 所以, $\mathrm{Var}(Z) = 1/\mu^2$. 故, $\mathrm{Var}(E(X_1|Z)) = \lambda^2/\mu^2$. (2 分) 然后, 给定 $Z = t$, $X_1$ 的条件分布是参数为 $\lambda t$ 的泊松分布, 所以, 我们有 $\mathrm{Var}(X_1|Z=t) = \lambda t$. 故, $\mathrm{Var}(X_1|Z) = \lambda Z$. (2 分) 所以, $E(\mathrm{Var}(X_1|Z)) = E(\lambda Z) = \lambda/\mu$. (2 分) 接着, 根据 $\mathrm{Var}(X_1) = \mathrm{Var}(E(X_1|Z)) + E(\mathrm{Var}(X_1|Z))$, 我们有

$$\mathrm{Var}(X_1) = \lambda^2/\mu^2 + \lambda/\mu.\ (2\text{ 分})$$

最后, 根据 Lindeberg-Lévy 中心极限定理即可得证. (2 分)

7. (10 分) 设随机变量 $X$ 服从参数为 $\alpha, \lambda$ 的 $\Gamma$ 分布 (伽马分布). 请用特征函数法或者矩母函数法证明: 当 $\alpha \to +\infty$ 时, 随机变量 $\frac{\lambda X - \alpha}{\sqrt{\alpha}}$ 依分布收敛于标准正态分布. (提示: 取对数, 利用泰勒展开.)

证明. 我们计算 $Y = \frac{\lambda X - \alpha}{\sqrt{\alpha}}$ 的矩母函数. 注意到

$$\begin{aligned} M_Y(t) = E(e^{tY}) &= e^{-t\sqrt{\alpha}} E(e^{\frac{t\lambda}{\sqrt{\alpha}} X}) \\ &= e^{-t\sqrt{\alpha}} \left(\frac{\lambda}{\lambda - \frac{\lambda t}{\sqrt{\alpha}}}\right)^{\alpha} = e^{-\alpha \ln(1 - t/\sqrt{\alpha}) - t\sqrt{\alpha}}.\ (4\text{ 分}) \end{aligned}$$

利用 Taylor 展开, 我们有

$$M_Y(t) = e^{\alpha(t/\sqrt{\alpha}+t^2/(2\alpha)+o(1/\alpha))-t\sqrt{\alpha}} = e^{t^2/2+o(1)} \stackrel{\alpha\to+\infty}{\to} e^{t^2/2}. \text{ (4 分)}$$

而 $e^{t^2/2}$ 正好是标准正态分布的矩母函数. 利用矩母函数的连续性定理, 得证. (2 分)

8. (7 分) 把 $(0,1)$ 区间划为为 $n$ 个彼此不相交的长度为 $p_1, p_2, \ldots, p_n$ 的区间. 此划分的熵被定义为

$$h = -\sum_{i=1}^{n} p_i \ln p_i.$$

假设 $X_1, X_2, \ldots$ 是一列独立同分布的随机变量, 共有分布为 $(0,1)$ 上的均匀分布. 对任意正整数 $m$, 用 $Z_m(i)$ 表示随机变量 $X_1, X_2, \ldots, X_m$ 中取值落在第 $i$ 个区间的个数. 令

$$R_m = \prod_{i=1}^{n} p_i^{Z_m(i)}.$$

证明：当 $m \to +\infty$ 时, $m^{-1} \ln R_m$ 几乎处处 (即" 几乎必然") 收敛到 $-h$.

证明. 令 $Y_j = \ln p_{X_j} = \sum_{i=1}^n \ln p_i 1_{X_j=i}$. (2 分) 则, $\ln R_m = \sum_{j=1}^m Y_j$. (1 分) 而, $E(Y_j) = \sum_{i=1}^n p_i \ln p_i = -h$. (2 分) 最后, 对独立同分布随机变量序列 $\{Y_j\}_{j\geq 1}$ 应用强大数定律即可. (2 分)