

西安交通大学考试题

成绩

课程 概率论与数理统计 (A 卷)

学 院 _____ 考试日期 2023 年 6 月 24 日

专业班号 _____

姓 名 _____ 学 号 _____ 期中 ☐ 期末 ☒

注: $\Phi(x)$ 是标准正态分布的分布函数。

一、单项选择题 (每小题 3 分, 共 15 分)

1. 已知 $P(A)=0.4$, $P(\bar{A} \cup B)=0.7$ 则 $P(\bar{B}|A) = (\quad)$ 。
A. 0.25 B. 0.5 C. 0.75 D. $\frac{4}{7}$
2. 设 X, Y 独立同分布, $P\{X=-1\}=P\{X=1\}=0.5$, 下列各式成立的是 ($\quad$)。
A. $P\{X=Y\}=1$ B. $P\{X=Y\}=0.5$
C. $P\{XY=1\}=0.25$ D. $P\{X+Y=0\}=0.25$
3. 设 $X \sim U(0,2)$, $Y \sim P(3)$, 且 X, Y 相互独立, 则 $D(3X-Y-2) = (\quad)$ 。
A. 0 B. 2 C. 4 D. 6
4. 设随机变量 X 与 Y 的相关系数为 1, $Z = X+Y$, X 与 Z 的相关系数为 ($\quad$)。
A. 1 B. 0 C. $\frac{1}{2}$ D. -1
5. 设总体 $X \sim N(12,4)$, $(X_1, X_2, \dots, X_n)$ 为来自 X 的样本, $X_{(n)}$ 是样本的极大值, 则 $P\{X_{(n)} > 16\} = (\quad)$ 。
A. $[\Phi(1)]^n$ B. $1-[\Phi(1)]^n$ C. $[\Phi(2)]^n$ D. $1-[\Phi(2)]^n$

二、填空题 (每小题 3 分, 共 15 分)

1. 袋中有 10 个红球和 5 个白球, 每次从袋中任取一球, 无放回地取两次, 则取出的两球是同色的概率为_____。
2. 设 X 的概率密度为 $f(x) = \begin{cases} e^{-x}, & x > 0 \\ 0, & \text{其他} \end{cases}$, 则 $Y = e^X$ 的概率密度为_____。
3. 设随机变量 $X \sim U(1,4)$, 现对 X 进行独立重复观测三次, 则至少有两次观测值小于 3 的概率为_____。
4. 设随机变量 X 服从参数为 1 的泊松分布, 则 $E(|X-1|) = \underline{\hspace{2cm}}$ 。

5. 设 $(X_1, X_2, \dots, X_{10})$ 是来自总体 $N(0, 4)$ 的样本, 则统计量 $\frac{4(X_1^2 + X_2^2)}{\sum_{i=3}^{10} X_i^2}$ 服从_____分布。

(请写出分布类型及自由度)

三、(10 分) 某人要购买一部手机, 现等可能地从甲、乙两种型号中选择购买。已知甲型号手机出厂是合格品的概率为 0.95, 乙型号手机出厂是合格品的概率为 0.99。

(1) 购买的手机是合格品的概率是多少?

(2) 若购买的手机不是合格品, 则该手机是甲型号的概率是多少?

四、(14 分) 设随机向量 (X, Y) 服从区域 $D = \{(x, y) | 0 < x < y, 0 < y < 1\}$ 上的均匀分布。

求: (1) X 与 Y 是否相互独立? 为什么? (2) $f_{X|Y}(x|y)$; (3) $P\left\{X > \frac{1}{4} \middle| Y = \frac{1}{2}\right\}$ 。

五、(10 分) 设随机变量 X 与 Y 相互独立, 且 $X \sim U(0, 2)$, $Y \sim \text{Exp}(1)$ 。求:

(1) $Z = \frac{X}{Y}$ 的概率密度; (2) $P\{X \geq Y\}$ 。

六、(8 分) 对一物体的长度 (单位: cm), 进行重复测量, 用这些测量值的平均值作为长度真值的估计值。假设各测量值是独立同分布的随机变量, 其期望是长度真值 d , 方差是 4。试用中心极限定理估计, 至少需要重复测量多少次才能使得平均值与真值的差的绝对值小于 0.5 的概率不小于 0.95? ($\Phi(1.65) = 0.95$, $\Phi(1.96) = 0.975$)

七、(10 分) 设 X 的概率密度为 $f(x; \theta) = \begin{cases} \sqrt{\theta} x^{\sqrt{\theta}-1}, & 0 < x < 1 \\ 0, & \text{其他} \end{cases}$, 其中 $\theta > 0$ 未知,

$(X_1, X_2, \dots, X_n)$ 是来自 X 的样本, 求 (1) θ 的矩估计量; (2) θ 的最大似然估计量。

八、(10 分) 某公司生产一种电动工具, 公司在新型号的产品中使用了一种新材料, 为了考察新型号产品的使用寿命 (单位: h) 是否比老型号有提高, 从老型号产品中随机抽取 8 台, 测得平均使用寿命 $\bar{x} = 1010$, 样本标准差 $s_1 = 60$; 从新型号产品中抽取 10 台, 测得平均使用寿命 $\bar{y} = 1090$, 样本标准差 $s_2 = 54$ 。假定新、老产品的使用寿命均服从正态分布, 且方差相等。

(1) 新型号产品的使用寿命是否比老型号有显著提高 ($\alpha = 0.05$)?

(2) 求新型号产品的平均使用寿命的置信度为 0.95 的双侧置信区间。

九、(8 分) 设 $(X_1, X_2, \dots, X_n)$ 是来自总体 $N(\mu, \sigma^2)$ 的样本。(1) 若 $T = C \sum_{i=1}^{n-1} (X_{i+1} - X_i)^2$ 是 σ^2

的无偏估计量, 求常数 C 。(2) $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ 是 σ^2 的相合估计量吗? 请说明理由。

参考数据: $u_{0.05} = 1.65$, $u_{0.025} = 1.96$, $t_{0.05}(16) = 1.7459$, $t_{0.025}(16) = 2.1199$,
 $t_{0.05}(9) = 1.8331$, $t_{0.025}(9) = 2.2622$, $t_{0.05}(7) = 1.8946$, $t_{0.025}(7) = 2.3646$

西安交通大学本科生课程考试试题标准答案与评分标准

课程名称：概率论与数理统计(A卷) 课时：48 考试时间：2023 年 6 月 24 日

一、1、C 2、B 3、D 4、A 5、D

二、1、 $\frac{11}{21}$ 2、 $f(y) = \begin{cases} \frac{1}{y^2}, & y > 1 \\ 0, & \text{其他} \end{cases}$ 3、 $\frac{20}{27}$ 4、 $2e^{-1}$ 5、 $F(2,8)$

三、(10分) $A = \{\text{购买甲型号手机}\}$, $B = \{\text{购买乙型号手机}\}$, $C = \{\text{购买的是合格品}\}$, 则
 $P(A) = P(B) = 0.5$, $P(C|A) = 0.95$, $P(C|B) = 0.99$.

(1) $P(C) = P(A)P(C|A) + P(B)P(C|B) = 0.5 \cdot 0.95 + 0.5 \cdot 0.99 = 0.97$. (5分)

(2) $P(\bar{C}) = 1 - P(C) = 0.03$, $P(\bar{C}|A) = 0.05$, $P(A|\bar{C}) = \frac{P(A)P(\bar{C}|A)}{P(\bar{C})} = \frac{0.5 \cdot 0.05}{0.03} = \frac{5}{6}$ (10分)

四、(14分) (1) $f(x, y) = \begin{cases} 2, & 0 < x < y < 1 \\ 0, & \text{其他} \end{cases}$, (2分)

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy = \begin{cases} \int_x^1 2 dy, & 0 < x < 1 \\ 0, & \text{其他} \end{cases} = \begin{cases} 2(1-x), & 0 < x < 1 \\ 0, & \text{其他} \end{cases} \quad (4 \text{分})$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx = \begin{cases} \int_0^y 2 dx, & 0 < y < 1 \\ 0, & \text{其他} \end{cases} = \begin{cases} 2y, & 0 < y < 1 \\ 0, & \text{其他} \end{cases} \quad (6 \text{分})$$

在 $0 < x < y < 1$ 上, $f(x, y) \neq f_X(x)f_Y(y)$, 所以 X, Y 不独立. (8分)

(2) 当 $0 < y < 1$ 时, $f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)} = \begin{cases} \frac{1}{y}, & 0 < x < y \\ 0, & \text{其他} \end{cases}$, (11分)

(3) $f_{X|Y}(x|\frac{1}{2}) = \begin{cases} 2, & 0 < x < \frac{1}{2} \\ 0, & \text{其他} \end{cases}$, $P\left\{X > \frac{1}{4} \middle| Y = \frac{1}{2}\right\} = \int_{\frac{1}{4}}^{\frac{1}{2}} f_{X|Y}(x|\frac{1}{2}) dx = \int_{\frac{1}{4}}^{\frac{1}{2}} 2 dx = \frac{1}{2}$. (14分)

五、(10分) (1) $f_X(x) = \begin{cases} \frac{1}{2} & 0 \leq x \leq 2 \\ 0 & \text{其他} \end{cases}$, $f_Y(y) = \begin{cases} e^{-y} & y > 0 \\ 0 & \text{其他} \end{cases}$ (2分)

$Z = \frac{X}{Y}$ 的概率密度 $f(z) = \int_{-\infty}^{+\infty} f_X(yz)f_Y(y)|y|dy$. (4分)

当 $z \leq 0$ 时, $f(z) = 0$; 当 $z > 0$ 时, $f(z) = \int_0^{\frac{2}{z}} \frac{1}{2} ye^{-y} dy = \frac{1}{2} - \left(\frac{1}{z} + \frac{1}{2}\right)e^{-\frac{2}{z}}$. (7分)

或令 $Z = \frac{X}{Y}$, $W = Y$, 则 (Z, W) 的概率密度为

$$f_{(Z,W)}(z,w) = f_X(zw)f_Y(w) \left| \frac{\partial(x,y)}{\partial(z,w)} \right| = \frac{1}{2} e^{-w} w, \quad w > 0, 0 \leq zw \leq 2 \quad (4 \text{ 分})$$

$Z = \frac{X}{Y}$ 的边缘概率密度为

$$f_Z(z) = \int_{-\infty}^{+\infty} f_{(Z,W)}(z,w) dw = \int_0^{\frac{2}{z}} \frac{1}{2} e^{-w} w dw = \frac{1}{2} - \left(\frac{1}{z} + \frac{1}{2} \right) e^{-\frac{2}{z}}, \quad z > 0 \quad (7 \text{ 分})$$

$$(2) \quad P\{X \geq Y\} = \iint_{x \geq y} f_X(x)f_Y(y) dx dy = \int_0^2 dx \int_0^{\frac{x}{2}} \frac{1}{2} e^{-y} dy = \frac{1}{2}(1 + e^{-2}). \quad (10 \text{ 分})$$

六、(8 分) 设需测量 n 次, 记 X_i 为第 i 次测量值, 则 $E(X_i) = d$, $D(X_i) = 4$. (2 分)

$$E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = d, D\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{4}{n}. \quad (3 \text{ 分})$$

$$P\left\{-0.5 \leq \frac{1}{n} \sum_{i=1}^n X_i - d \leq 0.5\right\} = P\left\{-\frac{0.5n}{2\sqrt{n}} \leq \frac{\sum_{i=1}^n X_i - nd}{2\sqrt{n}} \leq \frac{0.5n}{2\sqrt{n}}\right\} \approx 2\Phi\left(\frac{\sqrt{n}}{4}\right) - 1 \geq 0.95 \quad (7 \text{ 分})$$

$$\frac{\sqrt{n}}{4} \geq 1.96, \text{ 解得 } n \geq 61.47, \text{ 故应测量 } 62 \text{ 次}. \quad (8 \text{ 分})$$

$$\text{七、(10 分) (1) } E(X) = \int_0^1 x \sqrt{\theta} x^{\sqrt{\theta}-1} dx = \frac{\sqrt{\theta}}{\sqrt{\theta}+1}, \quad \frac{\sqrt{\hat{\theta}}}{\sqrt{\hat{\theta}}+1} = \bar{X}, \text{ 解得 } \hat{\theta} = \left(\frac{\bar{X}}{1-\bar{X}}\right)^2. \quad (5 \text{ 分})$$

$$(2) \text{ 似然函数 } L(\theta) = \prod_{i=1}^n f(x_i) = \begin{cases} \theta^{\frac{n}{2}} (x_1 x_2 \cdots x_n)^{\sqrt{\theta}-1}, & 0 < x_1, x_2, \cdots, x_n < 1 \\ 0 & \text{其它} \end{cases}, \quad (7 \text{ 分})$$

$$\text{在 } 0 < x_1, x_2, \cdots, x_n < 1 \text{ 时, } \ln L(\theta) = \frac{n}{2} \ln \theta + (\sqrt{\theta} - 1) \sum_{i=1}^n \ln x_i \quad (8 \text{ 分})$$

$$\frac{d}{d\theta} \ln L(\theta) = \frac{n}{2} \frac{1}{\theta} + \frac{1}{2\sqrt{\theta}} \sum_{i=1}^n \ln x_i = 0, \quad (9 \text{ 分})$$

$$\text{得 } \theta = \left(\frac{1}{n} \sum_{i=1}^n \ln x_i\right)^{-2}, \quad \theta = \left(\frac{1}{n} \sum_{i=1}^n \ln X_i\right)^{-2}. \quad (10 \text{ 分})$$

八、(10分) (1) 设老产品服从 $N(\mu_1, \sigma^2)$, 新产品服从 $N(\mu_2, \sigma^2)$.

检验假设 $H_0: \mu_1 = \mu_2, H_1: \mu_1 < \mu_2$ 或 $H_0: \mu_1 \geq \mu_2, H_1: \mu_1 < \mu_2$, (1分)

$$\text{检验统计量 } T = \frac{\bar{X} - \bar{Y}}{S_w \sqrt{1/n_1 + 1/n_2}} \stackrel{\mu_1 = \mu_2}{\sim} t(n_1 + n_2 - 2), \quad S_w = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}, \quad (2 \text{ 分})$$

拒绝域为 $t \leq -t_\alpha(n_1 + n_2 - 2)$, (3分)

$$n_1 = 8, n_2 = 10, s_1 = 60, s_2 = 54, \alpha = 0.05, \quad t_\alpha(n_1 + n_2 - 2) = t_\alpha(16) = 1.7459, \quad t = -2.9743, \quad (4 \text{ 分})$$

落入拒绝域中, 拒绝 H_0 , 即认为新型号产品的使用寿命比老型号有显著提高。(5分)

若 $H_0: \mu_1 \leq \mu_2, H_1: \mu_1 > \mu_2$, 拒绝域为 $t \geq t_\alpha(n_1 + n_2 - 2)$, 应接受 H_0 。

$$(2) \text{ 枢轴量 } T = \frac{\bar{Y} - \mu_2}{S_2 / \sqrt{n}} \sim t(n_2 - 1), \quad (7 \text{ 分}) \quad \mu_2 \text{ 的置信度为 } 0.95 \text{ 的双侧置信区间为}$$

$$\left(\bar{Y} - \frac{S_2}{\sqrt{n_2}} t_{\alpha/2}(n_2 - 1), \bar{Y} + \frac{S_2}{\sqrt{n_2}} t_{\alpha/2}(n_2 - 1) \right), \quad (9 \text{ 分}) \quad \text{得置信区间 } (1051.37, 1128.63). \quad (10 \text{ 分})$$

$$\text{九、(8分) (1) } E(T) = C \cdot \sum_{i=1}^{n-1} E(X_{i+1} - X_i)^2 = (n-1)C \cdot E(X_{i+1} - X_i)^2 \quad (2 \text{ 分})$$

$$\text{而 } E(X_{i+1} - X_i)^2 = D(X_{i+1} - X_i) + [E(X_{i+1} - X_i)]^2 = 2\sigma^2 \quad (3 \text{ 分})$$

$$\text{由 } (n-1)C \cdot 2\sigma^2 = \sigma^2, \text{ 解得 } C = \frac{1}{2(n-1)}. \quad (4 \text{ 分})$$

$$(2) \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1), \text{ 所以 } E(S^2) = \sigma^2, \quad D(S^2) = \frac{2}{n-1} \sigma^4, \text{ 由切比雪夫大不等式,}$$

$$0 \leq P\{|S^2 - \sigma^2| \geq \varepsilon^2\} \leq \frac{D(S^2)}{\varepsilon^2} = \frac{2\sigma^4}{(n-1)\varepsilon^2}, \text{ 故 } S^2 \text{ 是 } \sigma^2 \text{ 的相合估计量。} \quad (8 \text{ 分})$$

或, $\lim_{n \rightarrow \infty} E(S^2 - \sigma^2)^2 = \lim_{n \rightarrow \infty} D S^2 = 0$, 所以 S^2 是 σ^2 的均方相合估计量, 也是相合估计量。

$$\text{或, } S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n-1} \left(\sum_{i=1}^n X_i^2 - n\bar{X}^2 \right) = \frac{n}{n-1} \left(\frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}^2 \right)$$

$$\xrightarrow{P} 1 \cdot (\mu^2 + \sigma^2 - \mu^2) = \sigma^2.$$

期末考试模拟题（九）2023.2

一、单项选择题（每小题 3 分，共 30 分）

1、某人每次射击命中目标的概率为 p ，他独立地进行了四次射击，至少命中一次的概率是 $\frac{80}{81}$ ，则概率 p 为（ ）。

- A. $\frac{8}{81}$ B. $\frac{1}{2}$ C. $\frac{1}{3}$ D. $\frac{2}{3}$

2、在某港口，甲船每天在 8 点到 9 点间的任意时刻到达，而乙船在 7:30 到 8:30 间的任意时刻到达，则事件“两船到达时间相差不超过 20 分钟”的概率为（ ）。

- A. $\frac{2}{3}$ B. $\frac{1}{3}$ C. $\frac{4}{9}$ D. $\frac{5}{9}$

3、某人决定将一笔钱投资于房产、股票和期货之一，他选择这三种投资渠道的概率依次为 0.5, 0.25 和 0.25。据有关信息显示，现阶段投资这三种渠道获利的概率分别为 0.875, 0.75 和 0.875，则他投资获利的概率为（ ）。

- A. $\frac{27}{32}$ B. $\frac{5}{32}$ C. $\frac{5}{16}$ D. $\frac{11}{16}$

4、设随机变量 $X \sim N(2, \sigma^2)$ 且 $P(2 < X < 4) = 0.3$ ，则概率 $P(X \leq 0)$ 为（ ）。

- A. 0.3 B. 0.2 C. 0.7 D. 0.5

5、设随机变量 X 服从期望为 2 的指数分布，则 $P\{X > 3 | X > 1\} =$ （ ）。

- A. e^{-1} B. e^{-2} C. $e^{-\frac{3}{2}}$ D. e^{-4}

6、设 (X, Y) 的概率密度为 $f(x, y) = \begin{cases} 2e^{-(2x+y)}, & x > 0, y > 0, \\ 0, & \text{其它.} \end{cases}$ ，则 $P\{Y \leq X\} =$ （ ）。

- A. $\frac{3}{5}$ B. $\frac{1}{5}$ C. $\frac{1}{3}$ D. $\frac{5}{3}$

7、随机变量 X 和 Y 相互独立， X 服从正态分布 $N(0, 1)$ ， Y 服从正态分布 $N(1, 2)$ ，则随机变量 $X - 2Y$ 服从的分布是（ ）。

- A. $N(-2, -3)$ B. $N(-2, 5)$ C. $N(-2, 8)$ D. $N(-2, 9)$

8、设随机变量 X 与 Y 相互独立， X 服从参数为 2 的指数分布， Y 服从区间 $(1, 3)$ 内的均匀分布，则方差 $D(2X - Y)$ 为（ ）。

- A. 2 B. $\frac{7}{6}$ C. $\frac{3}{4}$ D. $\frac{4}{3}$

9、若随机变量 X 服从参数为 2 的泊松分布， $X_1, \dots, X_n$ 是来自总体 X 的样本，

则当 $n \rightarrow \infty$ 时, $Y_n = \frac{1}{n} \sum_{i=1}^n X_i^2$ 依概率收敛于 ()。

- A. 2 B. 4 C. 6 D. 8

10、设总体 $X \sim N(2, 3^2)$, $X_1, X_2, \dots, X_n$ ($n > 1$) 为来自 X 的样本, $\bar{X}$ 和 S^2 分别表示样本均值和样本方差, 则下面结果正确的是 ()。

- A. $\frac{\bar{X}-2}{3} \sim N(0,1)$ B. $\frac{(n-1)S^2}{9} \sim \chi^2(n-1)$
C. $\frac{nS^2}{9} \sim \chi^2(n)$ D. $\frac{\bar{X}-2}{S/\sqrt{n}} \sim t(n)$

二、填空题 (每小题 3 分, 共 15 分)

1、已知 $P(\bar{A})=0.3$, $P(B)=0.4$, $P(\bar{A}B)=0.5$, 则 $P(B-(A \cup \bar{B}))=$ _____。

2、设 X 在 $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ 内服从均匀分布, 则 $Y = \tan X$ 的概率密度 $f(y)=$ _____。

3、若随机变量 X_1, X_2 相互独立且分别服从参数为 λ_1, λ_2 的泊松分布, 若

$P\{X_1 + X_2 > 0\} = 1 - e^{-1}$, 则 $X_1 + X_2$ 的分布律为_____。

4、设两个随机变量 X, Y , $E(X) = -2$, $E(Y) = 4$, $D(X) = 4$, $D(Y) = 9$, $\rho(X, Y) = -0.5$
则 $E(3X^2 - 2XY + Y^2 - 3) =$ _____。

5、设总体 $X \sim N(\mu, 3^2)$, $X_1, X_2, \dots, X_n$ 为取自总体的一个样本, $\bar{X}$ 为样本均值, 要使 $E(\bar{X} - \mu)^2 \leq 0.1$ 成立, 则样本容量 n 至少应取_____。

三、(12 分) 设随机变量 X 与 Y 相互独立, 且都服从分布 $B(1, p)$, $0 < p < 1$ 。令

$$Z = \begin{cases} 1, & X+Y \text{ 为偶数} \\ 0, & X+Y \text{ 为奇数} \end{cases}$$

(1) 求 Z 的分布律;

(2) 求 (X, Z) 的分布律, 并判断 p 取何值时 X 与 Z 独立?

四、(12 分) 某生产线一次加工产品的合格率为 0.5, 不合格品立即再加工, 再加工的合格率仍为 0.5, 剩下的为废品。已知: 合格品每件可获利 80 元, 再加工费用为 20 元, 而废品每件总亏损 20 元。

(1) 为保证每天总利润的期望不低于 5 万元, 问他们至少要加工多少件产品?

(2) 如果想每天利润多于 5 万的概率大于 0.977, 利用中心极限定理估计至少要加工多少件? ($\Phi(2.0) = 0.977$, $\Phi(x)$ 表示标准正态分布的分布函数)

五、(12 分) 若总体 X 的概率密度为 $f(x) = \begin{cases} \frac{1}{\lambda} e^{-\frac{1}{\lambda}(x-\theta)}, & x \geq \theta \\ 0, & x < \theta \end{cases}$, 其中 θ, λ 是参数,

$X_1, X_2, \dots, X_n$ 是来自 X 的样本。(1) 若 θ 已知, 求参数 λ 的矩估计量;

(2) 若 λ 已知, 求参数 θ 的最大似然估计量。

六、(9分) 设 $X_1, X_2, \dots, X_6$ 是来自正态总体 $N(0, 4)$ 的样本。令

$$Y = aX_1^2 + b(2X_2 - X_3)^2 + c(3X_4 - 2X_5 - X_6)^2,$$

(1) 若 $Y \sim \chi^2(3)$, 求常数 a, b, c ;

(2) 证明随机变量 $Z = \frac{\sqrt{2}}{\sqrt{3}} \frac{X_1 + X_2 + X_3}{|X_5 - X_6|}$ 服从自由度为1的 t 分布。

七、(10 分) 某作物的产量近似服从正态分布 $N(\mu, \sigma^2)$, 现发现新的种子, 取 25 块样板田做实验, 发现平均亩产为 1864 公斤, 样本标准差为 50 公斤。

(1) 求方差 σ^2 的置信度为 0.95 的双侧置信区间。

(2) 此新种子的平均亩产量是不是以显著性水平 $\alpha = 0.05$ 显著大于 1800 公斤?

参 考 数 据 : $u_{0.05} = 1.65, u_{0.025} = 1.96$, $t_{0.05}(25) = 1.708, t_{0.05}(24) = 1.711$,

$t_{0.025}(25) = 2.060, t_{0.025}(24) = 2.064$, $\chi_{0.025}^2(25) = 40.646, \chi_{0.975}^2(25) = 13.120$,

$\chi_{0.05}^2(25) = 37.652, \chi_{0.95}^2(25) = 14.611$, $\chi_{0.025}^2(24) = 39.364, \chi_{0.975}^2(24) = 12.401$,

$\chi_{0.05}^2(24) = 36.415, \chi_{0.95}^2(24) = 13.848$.

期末考试模拟题（九）答案 2023.2

一、选择题（每小题 3 分，共 30 分）

1、D 2、B 3、A 4、B 5、A 6、C 7、D 8、D 9、C 10、B

二、填空题（每小题 3 分，共 15 分）

1、0.2 2、 $f(y) = \frac{1}{\pi(1+y^2)}, -\infty < y < +\infty$ 3、 $\frac{1}{k!}e^{-1}, k=0,1,2,\dots$

4、68 5、90

三、（12分）（1）由于 X 与 Y 相互独立，故 Z 的分布律：

$$P(Z=0) = P(X=0, Y=1) + P(X=1, Y=0) = 2p(1-p)$$

$$P(Z=1) = P(X=0, Y=0) + P(X=1, Y=1) = p^2 + (1-p)^2$$

（2） (X, Z) 的分布律

$Z \backslash X$	0	1	
0	$p(1-p)$	$(1-p)^2$	$1-p$
1	$p(1-p)$	p^2	p
	$2p(1-p)$	$p^2 + (1-p)^2$	

当 $p=0.5$ 时， X 与 Z 独立。

四、（12分）（1）设至少要加工 n 件产品，设 X_i 为加工第 i 件产品的利润，则 X_i 的分布为

X_i	80	60	-20
P	0.5	0.25	0.25

$E(X_i) = 50$ ，每天总利润的期望 $E(\sum_{i=1}^n X_i) = 50n \geq 50000$ ， $n \geq 1000$ ， n 至少为 1000。

（2）仍设至少要加工 n 件产品， $D(X_i) = 1700$ ， $D(\sum_{i=1}^n X_i) = 1700n$ 。由中心极限定理，

得 $\frac{\sum_{i=1}^n X_i - 50n}{\sqrt{1700n}} \sim N(0, 1)$ （近似）。

$$P(\sum_{i=1}^n X_i \geq 50000) = P(\frac{\sum_{i=1}^n X_i - 50n}{\sqrt{1700n}} \geq \frac{50000 - 50n}{\sqrt{1700n}}) > 0.977,$$

$$\Phi\left(\frac{50n-50000}{\sqrt{1700n}}\right) > 0.977. \text{ 查表, 即 } \frac{50n-50000}{\sqrt{1700n}} \geq 2, \text{ 得 } n=1054.$$

五、(12 分) (1) $EX = \int_{\theta}^{+\infty} x \frac{1}{\lambda} e^{-\frac{1}{\lambda}(x-\theta)} dx = \lambda + \theta, \quad \hat{\lambda} = \bar{X} - \theta.$

(2) 似然函数
$$L(\lambda, \theta) = \begin{cases} \frac{1}{\lambda^n} e^{-\frac{1}{\lambda} \sum_{i=1}^n (x_i - \theta)} & x_1, x_2, \dots, x_n > \theta \\ 0 & \text{其它} \end{cases},$$

由 $x_1, x_2, \dots, x_n > \theta$, 得 θ 的最大似然估计为 $\theta = \min\{x_1, x_2, \dots, x_n\} = x_{(1)}, \quad \theta = X_{(1)}.$

六、(9 分) (1) $Y \sim \chi^2(3)$, 则 $\sqrt{a}X_1, \sqrt{b}(2X_2 - X_3), \sqrt{c}(3X_4 - 2X_5 - X_6) \sim N(0, 1)$,
故 $D(\sqrt{a}X_1) = 4a = 1, D(\sqrt{b}(2X_2 - X_3)) = 20b = 1, D(\sqrt{c}(3X_4 - 2X_5 - X_6)) = 56c = 1,$

由此得 $a = \frac{1}{4}, b = \frac{1}{20}, c = \frac{1}{56}.$

(2) $\frac{1}{\sqrt{12}}(X_1 + X_2 + X_3) \sim N(0, 1), \frac{1}{8}(X_5 - X_6)^2 \sim \chi^2(1)$, 且他们相互独立, 故

$$\sqrt{\frac{2}{3}} \frac{X_1 + X_2 + X_3}{|X_5 - X_6|} \sim t(1).$$

七、(10 分) (1) σ^2 的置信度为 $1-\alpha$ 的置信区间为 $\left(\frac{(n-1)S^2}{\chi_{\frac{\alpha}{2}}^2(n-1)}, \frac{(n-1)S^2}{\chi_{1-\frac{\alpha}{2}}^2(n-1)} \right).$

$n=25, s=50, \alpha=0.05, \chi_{0.025}^2(24)=39.364, \chi_{0.975}^2(24)=12.401,$

方差 σ^2 的置信度为 0.95 的置信区间为 (1524.235, 4838.32).

(2) 设 $H_0: \mu=1800, H_1: \mu>1800$

取检验统计量 $T = \frac{\bar{X}-1800}{S} \sqrt{n}$, 在 H_0 为真时 $T \sim t(n-1).$

$n=25, s=50, \alpha=0.05$, 及 $t_{\alpha}(24)=1.711$, 给出拒绝域: $t \geq 1.711. \bar{x}=1864,$
 $t=6.4 \geq 1.711.$ 所以拒绝 H_0 , 认为此新种子的亩产量显著大于 1800 公斤.

期末考试模拟题（八）2022.6

一、单项选择题（本大题共 5 小题，每小题 3 分，共 15 分）

1. 设 $X \sim N(-1, \sigma^2)$ 且 $P(-3 < X < -1) = 0.4$, 则 $P(X \geq 1) =$ ().
(A) 0.1 (B) 0.2 (C) 0.3 (D) 0.4
2. 设随机变量 X 服从参数为 2 的指数分布, 即 $X \sim \text{Exp}(2)$, 以 Y 表示对 X 的三次独立重复观察中随机事件 $\{X > 1/2\}$ 出现的次数, 则 $P(Y = 2) =$ ().
(A) $e^{-2}(1 - e^{-1})$ (B) $3e^{-1}(1 - e^{-2})$
(C) $3e^{-2}(1 - e^{-1})$ (D) $e^{-1}(1 - e^{-2})$
3. 设随机变量 Z 的分布函数为 $F(z) = \begin{cases} 1 - e^{-3z} - e^{-4z} + e^{-7z}, & z \geq 0, \\ 0, & \text{其他.} \end{cases}$, 则 $E(Z) =$ ().
(A) $\frac{7}{12}$ (B) $\frac{37}{84}$ (C) $-\frac{7}{12}$ (D) $-\frac{37}{84}$
4. 设随机变量 X 的期望、方差都存在, 并且 $E(X) = 7$, $D(X) = 4$. 随机变量 X_1, X_2, X_3 相互独立、都与 X 有相同的分布, 下面说法正确的是 ().
(A) $D(X_1 + X_2) = D(2X) = 4D(X) = 16$ (B) $D(X_1 + X_2) = 2D(X) = 8$
(C) $P(X_1 = X_2 = X_3) = 0$ (D) $X_1 = X_2 = X_3$
5. 在假设检验中, 不拒绝原假设意味着 ().
(A) 原假设肯定是正确的 (B) 原假设肯定是错误的
(C) 没有证据证明原假设是正确的 (D) 没有充分证据证明原假设是错误的

二、填空题（本大题共 5 小题，每小题 3 分，共 15 分）

6. 设事件 A, B, C 满足 $B \subset A$, $B \subset C$, $P(A) = 0.8$, $P(AC) = 0.6$, $P(A - B) = 0.5$, 则 $P(\overline{A}BC) =$ _____.
7. 设连续型随机变量 X 的分布函数为 $F_X(x)$, $Y = 2X^2 + 1$, 则 Y 的分布函数 $F_Y(y) =$ _____.
8. 设 1 小时内进入某图书馆的读者人数服从 Poisson 分布, 已知 1 小时内无人进入该图书馆的概率为 0.01, 则 1 小时内至少有 2 个读者进入该图书馆的概率为 _____.
9. 设 $(X_1, X_2, \dots, X_n, \dots)$ 相互独立且同分布的随机变量序列, 其概率密度为 $f(x) = \begin{cases} 5e^{-5x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$. $n \rightarrow +\infty$ 时, $\frac{1}{n} \sum_{i=1}^n X_i^2$ 以概率收敛于 _____.
10. 设两个总体 X 与 Y 相互独立且具有相同的分布 $N(1, \sigma^2)$, $(X_1, X_2, \dots, X_5)$ 和 $(Y_1, Y_2, \dots, Y_5)$ 分别为来自这两个总体的简单样本, $\bar{X}$ 、 $\bar{Y}$ 分别为两个样本均值,

$S_1^2 = \frac{1}{4} \sum_{k=1}^5 (X_k - \bar{X})^2$, $S_2^2 = \frac{1}{4} \sum_{k=1}^5 (Y_k - \bar{Y})^2$, 则统计量 $\frac{5(\bar{X}-\bar{Y})^2}{S_1^2+S_2^2}$ 服从的分布为 _____.

三、计算与证明（本大题共 8 小题，共 70 分）

11. (10 分) 某企业流水线生产的产品按 100 件装箱，该企业出厂的检验标准是从每箱产品中抽取 10 件进行检验，若没发现不合格产品就通过检验，否则开箱逐个检验. 据统计每箱产品中的次品数不超过 4 件，每箱产品中有 i 件次品的概率如下表所示

i	0	1	2	3	4
P	0.1	0.2	0.4	0.2	0.1

(1) 请帮检验部门估计每箱产品的通过率;

(2) 假设按照这个检验方法某箱产品通过了检验，该箱产品中依然有 2 件次品的概率有多大?

12. (10 分) 设二维随机变量 (X, Y) 的概率密度为

$$f(x, y) = \begin{cases} 1, & 0 < x < 1, 0 < y < 2x \\ 0, & x \leq 0 \end{cases},$$

令 $Z = 2X - Y$, (1) 求 Z 的分布函数, (2) 求 $P(Y < 1/2 | X < 1/2)$.

13. (8 分) 设随机变量 X 的概率密度为 $f_X(x) = \begin{cases} 2e^{-2x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$

而且当随机变量 $X = x (x > 0)$ 条件下, Y 的条件概率密度为

$$f_{Y|X}(y|x) = \begin{cases} e^{-(y-x)}, & y > x \\ 0, & \text{其他} \end{cases}$$

(1) 求条件概率 $P(Y > 2 | X = 1)$; (2) 求 $P(X + Y < 2)$.

14. (6 分) 设总体的二阶矩存在, $X_1, X_2, \dots, X_n$ 是来自该总体的简单随机样本, $\bar{X}$ 为样本均值, 令 $Y = X_i - \bar{X}$, $Z = X_j - \bar{X}$, 求 Y 与 Z 的相关系数.

15. (10 分) 据调查某社区 400 个家庭中, 每个家庭购买车辆数为 0, 1, 2 的概率如下表

0	1	2
0.05	0.8	0.15

请利用中心极限定理近似计算:

(1) 假设各个家庭购买的车辆数相互独立, 求购买的车辆数超过 450 部的概率,

(2) 求购买 1 部车的家庭数不多于 340 的概率.

($\Phi(2.5) = 0.9938$, $\Phi(1.15) = 0.8749$)

16. (10 分) 设总体 X 的概率密度 $f(x; \theta) = \begin{cases} \frac{x}{\theta^2} e^{-\frac{x}{\theta}}, & x > 0 \\ 0, & x \leq 0 \end{cases}$, 其中 $\theta > 0$ 为未知参数, 从总体中抽取简单样本 $X_1, X_2, \dots, X_n$.

(1) 求 θ 的矩估计量 $\hat{\theta}_1$, 并判断是否为无偏估计量;

(2) 求 θ 的最大似然估计量 $\hat{\theta}_2$.

17. (10 分) 假设某型号彩色显像管的寿命服从正态分布, 并且已知标准差 $\sigma = 40$ 小时。从这批显像管中随机抽取 100 只, 算得其平均寿命为 10000 小时, 试求

(1) 显像管平均寿命 μ 的置信度为 0.99 的置信区间;

(2) 若显像管的平均寿命不小于 10100 小时被认为合格, 试在显著性水平

$\alpha = 0.005$ 下检验这批显像管是否合格? ($\Phi(2.576) = 0.995$, $\Phi(2.33) = 0.99$)

18. (6 分) 设总体 X 的分布函数为 $F(x)$, $(X_1, X_2, \dots, X_n)$ 为来自总体 X 的简单随机样本, $X_{(1)} = \min(X_1, X_2, \dots, X_n)$, $X_{(n)} = \max(X_1, X_2, \dots, X_n)$. 证明:

(1) 当 $x \leq y$ 时, $P\{X_{(1)} > x, X_{(n)} \leq y\} = (F(y) - F(x))^n$.

(2) 随机变量 $(X_{(1)}, X_{(n)})$ 的分布函数为

$$G(x, y) = \begin{cases} F^n(y) - (F(y) - F(x))^n, & x \leq y \\ F^n(y), & x > y \end{cases}.$$

期末考试模拟题（八）答案 2022.6

一、 A C B B D

二、 1. 0.3 2. $F_X\left(\sqrt{\frac{y-1}{2}}\right) - F_X\left(-\sqrt{\frac{y-1}{2}}\right)$ 3. $1 - 0.01(1 + 2\ln 10)$ 或 0.9439

4. $\frac{2}{25}$ 5. $F(1, 8)$

三、 11. 设A表示“按照检验标准任取一箱产品能够通过检验”， B_i 表示“箱中有*i*件次品”，

$i = 0, 1, 2, 3, 4$. 则 $P(A|B_i) = \frac{C_{100-i}^{10}}{C_{100}^{10}}, i = 0, 1, 2, 3, 4$.

(1) 由全概率公式, $P(A) = \sum_{i=0}^4 P(B_i)P(A|B_i) \approx 0.814$;

(2) 由贝叶斯公式, $P(B_2|A) = \frac{P(B_2)P(A|B_2)}{P(A)} \approx 0.397$.

12. (1) $F_Z(z) = P(Z \leq z) = P(2X - Y \leq z) = \iint_{2x-y \leq z} f(x, y) dx dy$

$$= \begin{cases} 0, & z \leq 0; \\ z - \frac{z^2}{4}, & 0 < z < 2; \\ 1, & z \geq 2. \end{cases}$$

$$(2) P(Y \leq \frac{1}{2} | X \leq \frac{1}{2}) = \frac{P(X \leq \frac{1}{2}, Y \leq \frac{1}{2})}{P(X \leq \frac{1}{2})} = \frac{\iint_{x \leq \frac{1}{2}, y \leq \frac{1}{2}} f(x, y) dx dy}{\int_0^{\frac{1}{2}} 2x dx} = \frac{3}{4}$$

(注: 此小题比较简单, 因为是均匀分布, 可不积分直接算出答案)

13. (1) $f_{Y|X}(y|x=1) = \begin{cases} e^{-(y-1)}, & y > 1; \\ 0, & \text{其他.} \end{cases}$

$$P(Y > 2 | X = 1) = \int_2^{+\infty} e^{-(y-1)} dy = e^{-1}$$

(2) $f(x, y) = f_X(x)f_{Y|X}(y|x) = \begin{cases} 2e^{-(x+y)}, & x > 0, y > x; \\ 0, & x \leq 0. \end{cases}$

$$P(X + Y < 2) = \int_0^1 dx \int_x^{2-x} 2e^{-(x+y)} dy = 1 - 3e^{-2}$$

14. $X_1, X_2, \dots, X_n$ 相互独立, 假设总体的方差为 $D(X)$, 则 $D(X_i) = D(X), D(\bar{X}) = \frac{D(X)}{n}$,

$$\text{cov}(Y, Z) = \text{cov}(X_i - \bar{X}, X_j - \bar{X}) = \begin{cases} \frac{n-1}{n} D(X), & i = j \\ -\frac{1}{n} D(X), & i \neq j \end{cases}$$

$$D(Y) = D(Z) = \frac{n-1}{n} D(X), \quad \rho = \frac{\text{cov}(Y, Z)}{\sqrt{D(Y)}\sqrt{D(Z)}} = \begin{cases} 1, & i = j \\ -\frac{1}{n-1}, & i \neq j \end{cases} \dots$$

15. (1) 设 X_k 为第 k 个家庭购买的车辆数, $k = 1, 2, \dots, 400$, 则 $E(X_k) = 1.1, D(X_k) = 0.19$.

设 X 为 400 个家庭购买的车辆数, 则 $X = \sum_{k=1}^{400} X_k \sim N(400 \times 1.1, 400 \times 0.19)$ (近似),

$$P(X \geq 450) \approx 1 - \Phi\left(\frac{450-400 \times 1.1}{\sqrt{400 \times 0.19}}\right) = 1 - \Phi(1.1471) = 0.1251.$$

(2) 设Y为400个家庭购买一部车的家庭数, 则 $Y \sim B(400, 0.8)$.

$$Y \sim N(400 \times 0.8, 400 \times 0.8 \times 0.2) \text{ (近似)}$$

$$P(X \leq 340) \approx \Phi\left(\frac{340-400 \times 0.8}{\sqrt{400 \times 0.8 \times 0.2}}\right) = \Phi(2.5) = 0.9938.$$

$$16. (1) E(X) = \int_{-\infty}^{+\infty} xf(x)dx = \int_0^{+\infty} x \cdot \frac{x}{\theta^2} e^{-\frac{x}{\theta}} dx = \theta \int_0^{+\infty} t^2 e^{-t} dt = 2\theta$$

令 $\bar{X} = E(X)$, 所以矩估计量 $\hat{\theta}_1 = \frac{\bar{X}}{2}$,

由于 $E(\hat{\theta}_1) = \frac{1}{2}E(\bar{X}) = \frac{1}{2}E(X) = \theta$, 所以无偏...

$$(2) L(\theta) = \prod_{i=1}^n \left(\frac{x_i}{\theta^2} e^{-\frac{x_i}{\theta}}\right) = \left(\prod_{i=1}^n x_i\right) \theta^{-2n} e^{-\frac{1}{\theta} \sum_{i=1}^n x_i}, \ln L(\theta) = \ln\left(\prod_{i=1}^n x_i\right) - 2n \ln \theta - \frac{1}{\theta} \sum_{i=1}^n x_i,$$

$$\text{令 } \frac{d}{d\theta} \ln L(\theta) = 0, \text{ 得 } -\frac{2n}{\theta} + \frac{1}{\theta^2} \sum_{i=1}^n x_i = 0, \text{ 解得极大似然估计量为 } \hat{\theta}_2 = \frac{\bar{X}}{2}.$$

17. (1) 设X表示彩色显像管的寿命, 由题设 $X \sim N(\mu, 40^2)$, 方差已知 $\sigma^2 = 40^2$, 对均值 μ 的估计为 $\left(\bar{x} - \frac{\sigma}{\sqrt{n}} u_{\frac{\alpha}{2}}, \bar{x} + \frac{\sigma}{\sqrt{n}} u_{\frac{\alpha}{2}}\right)$

$$= \left(10000 - 2.576 \times \frac{40}{\sqrt{100}}, 10000 + 2.576 \times \frac{40}{\sqrt{100}}\right) = (9989.70, 10010.3)$$

(2) $H_0: \mu \geq 10100, H_1: \mu < 10100$, (6分) 检验统计量 $U = \frac{\bar{x} - 10100}{\sigma/\sqrt{n}} \sim N(0, 1)$,

$$\text{拒绝域为 } \{U \leq -u_{\alpha}\} = \{U \leq -u_{0.005}\} = \{U \leq -2.576\}$$

$$\text{检验统计量的样本值 } u = \frac{\bar{x} - 10100}{\sigma/\sqrt{n}} = \frac{10000 - 10100}{40/\sqrt{100}} = -25 < -2.576$$

拒绝原假设, 认为生产的彩色显像管不合格.

$$18. (1) P\{X_{(1)} > x, X_{(n)} \leq y\} = P(x < X_1 \leq y, x < X_2 \leq y, \dots, x < X_n \leq y)$$

$$= \prod_{i=1}^n P(x < X_i \leq y) = (F(y) - F(x))^n$$

(2) 当 $x \leq y$ 时, $(X_{(1)}, X_{(n)})$ 的分布函数

$$\begin{aligned} G(x, y) &= P(X_{(1)} \leq x, X_{(n)} \leq y) = P(X_{(n)} \leq y) - P(\{X_{(1)} > x\} \cap \{X_{(n)} \leq y\}) \\ &= P(\max(X_1, X_2, \dots, X_n) \leq y) - P(\{X_{(1)} > x\} \cap \{X_{(n)} \leq y\}) \\ &= P(X_1 \leq y, X_2 \leq y, \dots, X_n \leq y) - P(\{X_{(1)} > x\} \cap \{X_{(n)} \leq y\}) \\ &= F^n(y) - (F(y) - F(x))^n \end{aligned}$$

当 $x > y$ 时, $P(\{X_{(1)} > x\} \cap \{X_{(n)} \leq y\}) = 0$

$$\begin{aligned} G(x, y) &= P(X_{(1)} \leq x, X_{(n)} \leq y) = P(X_{(n)} \leq y) - P(\{X_{(1)} > x\} \cap \{X_{(n)} \leq y\}) \\ &= P(X_{(n)} \leq y) = F^n(y) \end{aligned}$$

$$\text{所以 } G(x, y) = \begin{cases} F^n(y) - (F(y) - F(x))^n, & x \leq y; \\ F^n(y), & x > y. \end{cases}$$

期末考试模拟题（七）2022.2

备用数据: $\Phi(\cdot)$: 标准正态分布函数, $t_{\alpha}(n)$ 和 $\chi_{\alpha}^2(n)$: 相应分布的上侧 α 分位数.

$$\Phi(1)=0.841, \Phi(1.645)=0.95, \Phi(1.96)=0.975, \Phi(2)=0.977;$$

$$t_{0.025}(8)=2.306, t_{0.05}(8)=1.860; \chi_{0.05}^2(8)=15.507, \chi_{0.95}^2(8)=2.733.$$

一、完成下列各题（每小题 6 分，共 30 分）

1. 设 A, B, C 为三个随机事件, 已知 A 发生时 B 必然发生, B 与 C 互不相容, 且 $P(A)=0.2, P(B)=0.4, P(C)=0.3$, 求 $P(A \cup B \cup C)$; 判断 A 与 C 是否独立? 说明理由.

2. 设 X 的概率密度函数 $f(x) = \begin{cases} 0.4, & -1 < x < 0, \\ cx, & 0 < x < 1, \\ 0, & \text{其他.} \end{cases}$ 求 (1) c 的值; (2) X 的分布函数.

3. 设随机变量 X 服从均值为 2 的指数分布, 求 (1) X 的方差; (2) $P(X < 3 | X > 1)$.

4. 设随机变量 X 与 Y 相互独立, 同服从均值为 2 的泊松分布, 求:

$$(1) P\{X + Y = 2\};$$

$$(2) P\{\min(X, Y) = 1\}.$$

5. 设总体 $X \sim N(\theta, \theta^2)$, $X_1, \dots, X_n$ 是来自 X 的简单随机样本, $\bar{X}, S^2$ 分别是其样本均值和样本方差, a, b 是实数, 设 $T(a, b) = a\bar{X}^2 + bS^2$, 求: (1) $T(a, b)$ 是 θ^2 的无偏估计量的充要条件; (2) $T(0, 1)$ 的方差.

二. (10 分) 设随机变量 X 与 Y 相互独立, X 服从区间 $(0, 2)$ 上的均匀分布, $Y \sim$

$B(2, 0.5)$, 令 $Z = XY$, 求: (1) Z 的数学期望与方差; (2) Z 的分布函数.

三. (8 分) 设随机变量 X 的概率密度为 $f(x) = \begin{cases} \frac{x}{4}, & 1 < x < 3, \\ 0, & \text{其他.} \end{cases}$, 对 X 独立重复观察

396 次, 观察值为 $X_1, \dots, X_{396}$, 记 $Y = X_1 + \dots + X_{396}$. 试由中心极限定理, 求

$P\{Y < 880\}$ 的近似值.

四. (12 分) 设 (X, Y) 的概率密度为 $f(x, y) = \begin{cases} \frac{1}{8}(2 - xy), & 0 < x < 2, -1 < y < 1, \\ 0, & \text{其他.} \end{cases}$

(1) 求 $P(X + Y > 1)$; (2) 求条件概率密度 $f_{X|Y}(x|y)$;

(3) 判断 X 与 $|Y|$ 是否独立? 说明理由.

五.(10 分) 设 (X,Y) 服从二维正态分布, $X \sim N(2, 4)$, $Y \sim N(1, 9)$, X 与 Y 的相关系数为 ρ . (1) 若 $\rho=0$, 求 $P(2X > Y-2)$, 问 a 取何值时, $aX+Y$ 与 $X-2Y$ 独立? (2) 若 $\rho=0.5$, 求 $2X-Y$ 的概率密度.

六.(10 分) 设总体 X 的分布律为 $P(X=0)=(1-\theta)^2$, $P(X=1)=2\theta(1-\theta)$, $P(X=2)=\frac{2}{3}\theta^2$, $P(X=3)=\frac{1}{3}\theta^2$, 未知参数 $0<\theta<1$, 从总体中抽取容量为25的简单随机样本, 观测到“0”, “1”, “2”, “3”出现的次数分别为3, 13, 7, 2, 分别求 θ 的矩估计值和最大似然估计值.

七.(10 分) 设总体 X 的概率密度 $f(x;\theta)=\begin{cases} \frac{2x}{\theta^2}, & 0 < x \leq \theta, \\ 0, & \text{其他.} \end{cases}$, 未知参数 $\theta > 0$.

$X_1, \dots, X_n$ 为来自 X 的样本, $\bar{X}$ 是样本均值. (1) 判断 $\frac{9}{4}\bar{X}^2$ 是否为 θ^2 的无偏估计量, 说明理由; (2) 判断 $\frac{9}{4}\bar{X}^2$ 是否为 θ^2 的相合(一致)估计量, 说明理由.

八.(10 分) 设一工厂生产的某种型号无缝钢管的内径 $X \sim N(\mu, \sigma^2)$, 从某日生产的钢管中随机抽出9根, 测得内径(单位: cm)的样本均值为53.84, 样本标准差为1.32. (1) 在显著水平0.05下检验 $H_0: \sigma^2=1, H_1: \sigma^2>1$. (2) 求 μ 的置信水平为0.95的置信区间.(保留两位小数)

期末考试模拟题（七）答案 2022.2

备用数据: $\Phi(\cdot)$: 标准正态分布函数, $t_{\alpha}(n)$ 和 $\chi_{\alpha}^2(n)$: 相应分布的上侧 α 分位数.

$$\Phi(1)=0.841, \Phi(1.645)=0.95, \Phi(1.96)=0.975, \Phi(2)=0.977;$$

$$t_{0.025}(8)=2.306, t_{0.05}(8)=1.860; \chi_{0.05}^2(8)=15.507, \chi_{0.95}^2(8)=2.733.$$

一、完成下列各题（每小题 6 分，共 30 分）

1. 设 A, B, C 为三个随机事件, 已知 A 发生时 B 必然发生, B 与 C 互不相容, 且 $P(A)=0.2, P(B)=0.4, P(C)=0.3$, 求 $P(A \cup B \cup C)$; 判断 A 与 C 是否独立? 说明理由.

解: $P(A \cup B \cup C) = P(B \cup C) = P(B) + P(C) = 0.7$, $P(AC) = 0 \neq P(A)P(C)$, 所以 A 与 C 不独立.

2. 设 X 的概率密度函数 $f(x) = \begin{cases} 0.4, & -1 < x < 0, \\ cx, & 0 < x < 1, \\ 0, & \text{其他.} \end{cases}$ 求(1) c 的值; (2) X 的分布函数.

$$\text{解: (1) } 1 = \int_{-\infty}^{+\infty} f(x)dx = 0.4 + \frac{1}{2}c, \quad c = 1.2.$$

$$(2) F(x) = \int_{-\infty}^x f(t)dt = \begin{cases} 0, & x < -1, \\ 0.4(x+1), & -1 \leq x < 0, \\ 0.4 + 0.6x^2, & 0 \leq x < 1, \\ 1, & x \geq 1. \end{cases}$$

3. 设随机变量 X 服从均值为 2 的指数分布, 求(1) X 的方差; (2) $P(X < 3 | X > 1)$.

$$\text{解: (1) } X \sim \text{Exp}(\lambda), \lambda = 0.5, \text{Var}(X) = 1/\lambda^2 = 4$$

$$(2) P(X < 3 | X > 1) = P(X < 2) = 1 - e^{-1}.$$

4. 设随机变量 X 与 Y 相互独立, 同服从均值为 2 的泊松分布, 求:

$$(1) P\{X + Y = 2\}; \quad (2) P\{\min(X, Y) = 1\}.$$

$$\text{解: (1) } X + Y \sim P(4), P(X + Y = 2) = 8e^{-4};$$

$$(2) P(\min(X, Y) = 1) = P(X = 1, Y \geq 1) + P(X > 1, Y = 1) = 4e^{-2}(1 - 2e^{-2})$$

5. 设总体 $X \sim N(\theta, \theta^2)$, $X_1, \dots, X_n$ 是来自 X 的简单随机样本, $\bar{X}, S^2$ 分别是其样本均值和样本方差, a, b 是实数, 设 $T(a, b) = a\bar{X}^2 + bS^2$, 求: (1) $T(a, b)$ 是 θ^2 的无偏估计量的充要条件; (2) $T(0, 1)$ 的方差.

$$\text{解: (1) } E(T(a, b)) = E(a\bar{X}^2 + bS^2) = a(\theta^2 + \frac{1}{n}\theta^2) + b\theta^2 = \theta^2, \Leftrightarrow \frac{n+1}{n}a + b = 1,$$

$$(2) \frac{(n-1)S^2}{\theta^2} \sim \chi^2(n-1), \text{Var}\left[\frac{(n-1)S^2}{\theta^2}\right] = 2(n-1), \Rightarrow \text{Var}(T(0,1)) = \text{Var}(S^2) = \frac{2}{n-1} \theta^4$$

二. (10 分) 设随机变量 X 与 Y 相互独立, X 服从区间 $(0, 2)$ 上的均匀分布, $Y \sim B(2, 0.5)$, 令 $Z = XY$, 求: (1) Z 的数学期望与方差; (2) Z 的分布函数.

解: (1) $E(X)=1, \text{Var}(X)=1/3, E(X^2)=4/3; E(Y)=1, \text{Var}(Y)=1/2, E(Y^2)=3/2$,
所以 $E(Z) = E(X)E(Y)=1, \text{Var}(Z) = E(X^2)E(Y^2) - [E(X)E(Y)]^2 = 1$;

$$(2) X \text{ 的分布函数 } F(x) = \begin{cases} 0, & x < 0, \\ \frac{x}{2}, & 0 \leq x < 2, \\ 1, & x \geq 2. \end{cases}$$

Y 的分布律 $P(Y=0)=1/4, P(Y=1)=1/2, P(Y=2)=1/4$.

$$F_Z(z) = P(XY \leq z) = P(Y=0)P(0 \leq z) + P(Y=1)P(X \leq z) + P(Y=2)P(2X \leq z)$$

$$= \begin{cases} 0, & z < 0, \\ \frac{4+5z}{16}, & 0 \leq z < 2, \\ \frac{12+z}{16}, & 2 \leq z < 4, \\ 1, & z \geq 4. \end{cases}$$

三. (8 分) 设随机变量 X 的概率密度为 $f(x) = \begin{cases} \frac{x}{4}, & 1 < x < 3, \\ 0, & \text{其他.} \end{cases}$, 对 X 独立重复观察

396 次, 观察值为 $X_1, \dots, X_{396}$, 记 $Y = X_1 + \dots + X_{396}$. 试由中心极限定理, 求

$P\{Y < 880\}$ 的近似值.

$$\text{解: } E(X) = \int_1^3 \frac{x^2}{4} dx = \frac{13}{6}, E(X^2) = \int_1^3 \frac{x^3}{4} dx = 5, \text{Var}(X) = \frac{11}{36}.$$

根据中心极限定理, $Y = X_1 + \dots + X_{396} \stackrel{\text{近似}}{\sim} N(858, 121)$.

$$P(Y \leq 880) \approx \Phi\left(\frac{880-858}{11}\right) = \Phi(2) = 0.977.$$

四. (12 分) 设 (X, Y) 的概率密度为 $f(x, y) = \begin{cases} \frac{1}{8}(2-xy), & 0 < x < 2, -1 < y < 1, \\ 0, & \text{其他.} \end{cases}$

(1) 求 $P(X+Y > 1)$; (2) 求条件概率密度 $f_{X|Y}(x|y)$;

(3) 判断 X 与 Y 是否独立? 说明理由.

解: (1) $P(X+Y>1) = \iint_{x+y>1} f(x,y)dxdy = \int_0^2 dx \int_{1-x}^1 \frac{2-xy}{8} dy = \frac{5}{12},$

(2) Y 的边际密度函数 $f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)dx = \begin{cases} \int_0^2 \frac{2-xy}{8} dx = \frac{2-y}{4}, & -1 < y < 1, \\ 0, & \text{其他.} \end{cases}$

当 $-1 < y < 1$ 时, $f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} = \begin{cases} \frac{2-xy}{2(2-y)}, & 0 < x < 2, \\ 0, & \text{其他.} \end{cases}$

(3) 计算得, 对一切的 x, y , $P(X \leq x, |Y| \leq y) = P(X \leq x)P(|Y| \leq y)$ 成立, 所以 X 与 $|Y|$ 相互独立.

其中当 $0 < x < 2, 0 < y < 1$ 时, $P(X \leq x, |Y| \leq y) = \int_0^x du \int_{-y}^y \frac{2-uv}{8} dv = \frac{xy}{2},$

$P(X \leq x) = \int_0^x du \int_{-1}^1 \frac{2-uv}{8} dv = \frac{x}{2}, P(|Y| \leq y) = \int_0^2 du \int_{-y}^y \frac{2-uv}{8} dv = y.$

五. (10 分) 设 (X, Y) 服从二维正态分布, $X \sim N(2, 4), Y \sim N(1, 9), X$ 与 Y 的相关系数为 ρ . (1) 若 $\rho = 0$, 求 $P(2X > Y - 2)$, 问 a 取何值时, $aX + Y$ 与 $X - 2Y$ 独立? (2) 若 $\rho = 0.5$, 求 $2X - Y$ 的概率密度.

解: (X, Y) 服从正态分布, 且 $X \sim N(2, 4), Y \sim N(1, 9),$

(1) $\rho = 0$, X 与 Y 相互独立, 所以 $2X - Y + 2 \sim N(5, 25),$

$P(2X > Y - 2) = P(2X - Y + 2 > 0) = \Phi(1) = 0.841,$

$aX + Y$ 与 $X - 2Y$ 独立的充要条件是 $aX + Y$ 与 $X - 2Y$ 不相关, 即 $\text{Cov}(aX + Y, X - 2Y) = 4a - 18 = 0$, 得 $a = 9/2;$

(2) $\rho = 0.5, E(X) = 3, \text{Var}(2X - Y) = 4 \times 4 + 9 - 4 \times 3 = 13, 2X - Y \sim N(3, 13),$

$$f(x) = \frac{1}{\sqrt{26\pi}} e^{-\frac{(x-3)^2}{26}}, -\infty < x < \infty$$

六. (10 分) 设总体 X 的分布律为 $P(X = 0) = (1 - \theta)^2, P(X = 1) = 2\theta(1 - \theta),$

$P(X = 2) = \frac{2}{3}\theta^2, P(X = 3) = \frac{1}{3}\theta^2,$ 未知参数 $0 < \theta < 1$, 从总体中抽取容量为 25 的简单随机样本, 观测到“0”, “1”, “2”, “3”出现的次数分别为 3, 13, 7, 2, 分别求 θ 的矩估计值和最大似然估计值.

解: 矩估计: $\mu_1 = E(X) = \frac{\theta^2}{3} + 2\theta, \frac{1}{3}\hat{\theta}^2 + 2\hat{\theta} = \bar{X}, \bar{x} = \frac{33}{25},$ 解得矩估计值为 $\hat{\theta} = 0.6.$

最大似然估计: 似然函数 $L(\theta) = [(1 - \theta)^2]^3 [2\theta(1 - \theta)]^{13} [\frac{2}{3}\theta^2]^7 [\frac{1}{3}\theta^2]^2 = c\theta^{31}(1 - \theta)^{19}$

$$\ln L(\theta) = \ln c + 31 \cdot \ln \theta + 19 \ln(1 - \theta), \quad \frac{d}{d\theta} \ln L(\theta) = \frac{31}{\theta} - \frac{19}{1 - \theta} = 0 \Rightarrow \hat{\theta} = 0.62$$

七.(10 分) 设总体 X 的概率密度 $f(x; \theta) = \begin{cases} \frac{2x}{\theta^2}, & 0 < x \leq \theta, \\ 0, & \text{其他.} \end{cases}$, 未知参数 $\theta > 0$.

$X_1, \dots, X_n$ 为来自 X 的样本, $\bar{X}$ 是样本均值. (1) 判断 $\frac{9}{4} \bar{X}^2$ 是否为 θ^2 的无偏估计量, 说明理由; (2) 判断 $\frac{9}{4} \bar{X}^2$ 是否为 θ^2 的相合(一致)估计量, 说明理由.

$$\text{解: (1) } E(\bar{X}) = E(X) = \int_0^\theta \frac{2x^2}{\theta^2} dx = \frac{2}{3} \theta, \quad \text{Var}(\bar{X}) > 0,$$

$$E\left(\frac{9}{4} \bar{X}^2\right) = \frac{9}{4} \{ \text{Var}(\bar{X}) + [E(\bar{X})]^2 \} = \frac{9}{4} \text{Var}(\bar{X}) + \theta^2 > \theta^2,$$

即 $\frac{9}{4} \bar{X}^2$ 不是 θ^2 的无偏估计量.

(2) 根据辛钦大数定律, 当 $n \rightarrow +\infty$ 时, $\bar{X} \xrightarrow{P} E(X) = \frac{2}{3} \theta$, 由依概率收敛的性质,

当 $n \rightarrow +\infty$ 时, $\frac{9}{4} \bar{X}^2 \xrightarrow{P} \frac{9}{4} \left(\frac{2}{3} \theta\right)^2 = \theta^2$, 所以 $\frac{9}{4} \bar{X}^2$ 是 θ^2 的相合(一致)估计量.

八.(10 分) 设一工厂生产的某种型号无缝钢管的内径 $X \sim N(\mu, \sigma^2)$, 从某日生产的钢管中随机抽出 9 根, 测得内径(单位: cm)的样本均值为 53.84, 样本标准差为 1.32. (1) 在显著水平 0.05 下检验 $H_0: \sigma^2 = 1, H_1: \sigma^2 > 1$. (2) 求 μ 的置信水平为 0.95 的置信区间.(保留两位小数)

解: (1) 检验 $H_0: \sigma^2 = 1, H_1: \sigma^2 > 1$, 检验统计量为 $\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$, 拒绝域为

$$\chi^2 \geq \chi_{0.05}^2(8),$$

计算得 $\chi^2 = 13.9392$, 而 $\chi_{0.05}^2(8) = 15.507$, 未落在拒绝域内, 所以不拒绝原假设.

(2) μ 的置信水平为 0.95 的置信区间为 $(\bar{X} \pm \frac{S}{\sqrt{9}} t_{0.025}(8))$, 计算得 (52.83, 54.85)。

概率统计期末考试模拟题（六） 2021.6

(标准正态分布函数 $\Phi(x)$: $\Phi(1.50) = 0.9332$, $\Phi(2.5) = 0.9938$,

$\chi_{0.975}^2(14) = 5.629$, $\chi_{0.025}^2(14) = 26.119$, $\chi_{0.975}^2(15) = 6.262$, $\chi_{0.025}^2(15) = 27.488$

$t_{0.025}(14) = 2.1448$, $t_{0.025}(15) = 2.1314$, $t_{0.05}(14) = 1.7613$, $t_{0.05}(15) = 1.7531$)

一. 填空题 (每小题 3 分, 共 30 分)

1. 一射手对同一目标独立地进行四次射击, 若至少命中一次的概率为 $\frac{80}{81}$, 则该

射手进行一次射击的命中率为_____.

2. 随机变量 X 和 Y 独立同分布于参数为 λ ($\lambda > 0$) 的泊松分布, 令

$U = 2X + Y$, $V = 2X - Y$, 则 U 和 V 的相关系数为_____.

3. 随机变量 $X \sim N(\mu, 4^2)$, $Y \sim N(\mu, 5^2)$, 设 $p_1 = P(X \leq \mu - 4)$, $p_2 = P(Y \geq \mu + 5)$,

则比较 p_1 _____ p_2 (填 $\leq$, $\geq$ 或 $=$).

4. 设有事件 A, B , $P(\bar{A}) = 0.3$, $P(B) = 0.4$, $P(A\bar{B}) = 0.5$, 则 $P(B | A \cup \bar{B}) =$ _____.

5. 随机变量 $X \sim N(\mu, \sigma^2)$, 则 $E[e^X] =$ _____.

6. 设 $(X_1, X_2, \dots, X_n)$ 是来自正态总体 $N(\mu, \sigma^2)$ 的样本, 其中 μ, σ^2 是未知参数, 样本均值记为 $\bar{X}$, 样本方差为 S^2 , 则参数 μ 的置信水平为 $1 - \alpha$ 的置信区间长度为_____.

7. 设总体 $X \sim N(0, 1)$, $(X_1, X_2, \dots, X_6)$ 是来自 X 的样本, 设

$Y = (X_1 + X_2 + X_3)^2 + (X_4 + X_5 + X_6)^2$, 若 $CY \sim \chi^2(2)$, 则 $C =$ _____.

8. 设随机变量 X 服从参数为 9 的泊松分布, $Y \sim N(9, 4)$, 且 X, Y 相互独立, 则

根据切比雪夫不等式有: $P(|X - Y| \geq 4) \leq$ _____.

9. 设随机变量 X 的概率密度为 $f_X(x) = \begin{cases} 1, & 0 < x < 1, \\ 0, & \text{其它} \end{cases}$, 则 $Y = -2 \ln X$ 的概率密度

$f_Y(y) =$ _____.

10. 设 X 与 Y 相互独立且均服从参数为 1 的指数分布, 则

$P\{1 < \min(X, Y) \leq 2\} =$ _____.

二. (8 分) 已知男人中有 5% 是色盲患者, 女人中有 0.25% 是色盲患者, 今从男、女人人数相等的人群中随机挑选一人, 发现恰好是色盲患者, 问此人是男性的概率是多大?

三. (8 分) 某箱装 100 件产品, 其中一、二、三等品分别为 80, 10 和 10 件. 现从中随机取一件, 定义三个随机变量 X_1, X_2, X_3 如下:

$$X_i = \begin{cases} 1, & \text{抽到 } i \text{ 等品} \\ 0, & \text{其他} \end{cases}, i = 1, 2, 3. \text{ 求 } X_1 \text{ 和 } X_2 \text{ 的相关系数.}$$

四. (14 分) 二维随机变量 (X, Y) 的概率密度为

$$f(x, y) = \begin{cases} x, & 0 \leq x \leq 2, \max\{0, x-1\} \leq y \leq \min\{1, x\} \\ 0, & \text{其他} \end{cases}.$$

求 (1) 边缘概率密度 $f_X(x), f_Y(y)$; (2) 协方差 $Cov(X, Y)$;

(3) 对 $y \in (0, 1)$, 求条件概率密度 $f_{X|Y}(x|y)$.

五. (10 分) 设 X 和 Y 是两个相互独立的随机变量, X 服从区间 $(0, 4)$ 上的均匀分布, Y 服从参数为 1 的指数分布, 求 $Z = X + 2Y$ 的分布函数.

六. (10 分) 从一台车床加工的一批轴料中取 15 件测量其椭圆度, 计算得椭圆度的样本标准差 $s = 0.025$, 问该批轴料椭圆度的方差与规定的 $\sigma^2 = 0.0004$ 有

无显著差别 ($\alpha = 0.05$, 椭圆度服从正态分布)?

七. (10 分) 有一批建筑房屋用的木柱, 其中 80% 的长度不小于 3m, 现从这批木柱中随机地取出 100 根, 问其中至少有 30 根短于 3m 的概率是多少?

八. (10 分) 设总体 X 的概率密度为

$$f(x) = \begin{cases} \theta, & 0 \leq x < 1; \\ 1 - \theta, & 1 \leq x \leq 2; \\ 0, & \text{其他} \end{cases}$$

其中 $\theta > 0$ 为未知参数. $(X_1, X_2, \dots, X_n)$ 是来自总体 X 的样本, 记 N 是样本观测值中小于 1 的个数, 求 θ 的矩估计量 $\hat{\theta}_M$ 和极大似然估计量 $\hat{\theta}_L$, 并试计算极大似然估计量 $\hat{\theta}_L$ 的无偏性.

概率统计期末考试模拟题（六）答案

一. (每题 3 分)

$$1. \frac{2}{3} \quad 2. \frac{3}{5} \quad 3. = \quad 4. 0.25 \quad 5. e^{\mu + \frac{\sigma^2}{2}} \quad 6. \frac{2S}{\sqrt{n}} t_{\frac{\alpha}{2}}(n-1)$$

$$7. \frac{1}{3} \quad 8. \frac{13}{16} \quad 9. f_Y(y) = \begin{cases} \frac{1}{2} e^{-\frac{1}{2}y}, & y > 0 \\ 0, & y \leq 0 \end{cases} \quad 10. e^{-2} - e^{-4}$$

二. (8 分) 解: 设事件 $M = \{\text{男人}\}$, $S = \{\text{色盲患者}\}$, 题意有 $P(M) = P(\bar{M}) = 0.5$ -

$$P(M|S) = \frac{P(S|M)P(M)}{P(S|M)P(M) + P(S|\bar{M})P(\bar{M})} = \frac{0.05 \times 0.5}{0.05 \times 0.5 + 0.0025 \times 0.5} = \frac{20}{21} \quad (= 0.9524)$$

三. (8 分) 解: $X_1 \sim B(1, 0.8)$, $EX_1 = 0.8$, $D(X_1) = 0.16$

$X_2 \sim B(1, 0.1)$, $EX_2 = 0.1$, $D(X_2) = 0.09$ -- 2 分 $X_1 X_2 \sim B(1, 0)$, $E[X_1 X_2] = 0$

$$\text{Cov}(X_1, X_2) = E[X_1 X_2] - E[X_1]E[X_2] = -0.08$$

$$\rho = \frac{\text{Cov}(X_1, X_2)}{\sqrt{\text{Var}(X_1)\text{Var}(X_2)}} = -\frac{0.08}{\sqrt{0.16 \times 0.09}} = -\frac{2}{3}$$

四. (14 分) 解: (1) $f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy$, $0 < x \leq 1$ 时, $f_X(x) = \int_0^x x dy = x^2$,

$1 < x \leq 2$ 时, $f_X(x) = \int_{x-1}^1 x dy = 2x - x^2$, x 属于其他情况时, $f_X(x) = 0$.

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx, \quad 0 < y \leq 1 \text{ 时}, f_Y(y) = \int_y^{y+1} x dx = \frac{2y+1}{2},$$

y 属于其他情况时, $f_Y(y) = 0$.

$$(2) \quad EX = \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^1 x^3 dx + \int_1^2 x(2x - x^2) dx = \frac{7}{6},$$

$$EY = \int_{-\infty}^{\infty} y f_Y(y) dy = \int_0^1 y(y + \frac{1}{2}) dy = \frac{7}{12},$$

$$E[XY] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f(x, y) dx dy = \int_0^1 dy \int_y^{y+1} x^2 y dy = \frac{3}{4},$$

$$\text{Cov}(X, Y) = E[XY] - E[X]E[Y] = \frac{5}{72}.$$

$$(3) \quad y \in (0, 1) \text{ 时}, f_{X|Y}(x|y) = \begin{cases} \frac{f(x, y)}{f_Y(y)} = \frac{2x}{2y+1}, & x \in (y, y+1) \\ 0, & \text{其他} \end{cases}$$

五. (10 分) 解: $f(x, y) = f_X(x)f_Y(y) = \begin{cases} \frac{1}{4}e^{-y}; & 0 < x < 4, y > 0 \\ 0; & \text{其他} \end{cases}$

$$F_Z(z) = P(X + 2Y \leq z) = \int \int_{\{x+2y \leq z\}} f(x, y) dx dy,$$

- $z \leq 0$ 时, $F_Z(z) = 0$;

$$0 < z \leq 4 \text{ 时, } F_Z(z) = \int_0^z dx \int_0^{\frac{z-x}{2}} \frac{1}{4} e^{-y} dy = \frac{1}{4}(z - 2 + 2e^{-\frac{z}{2}}),$$

$$z > 4 \text{ 时, } F_Z(z) = \int_0^4 dx \int_0^{\frac{z-x}{2}} \frac{1}{4} e^{-y} dy = 1 - \frac{1}{2}[e^2 - 1]e^{-\frac{z}{2}},$$

六. (10 分) 解 由题意, 建立假设 $H_0: \sigma^2 = 0.0004 \leftrightarrow H_1: \sigma^2 \neq 0.0004$.

$$\text{当 } H_0 \text{ 成立时, 统计量 } \chi^2 = \frac{(n-1)S^2}{\sigma^2} = \frac{(n-1)S^2}{0.0004} \sim \chi^2(n-1),$$

$$\text{拒绝域为 } W = \{\chi^2 > \chi_{\frac{\alpha}{2}}^2(n-1) \text{ 或 } \chi^2 < \chi_{1-\frac{\alpha}{2}}^2(n-1)\}$$

$$n=15, \chi_{0.975}^2(14) = 5.629, \chi_{0.025}^2(14) = 26.119,$$

$$\text{计算 } \chi^2 = \frac{14 \times 0.025^2}{0.0004} = 21.875 \in (5.629, 26.119), \text{ 不在拒绝域中, 因此接受原假设, 即认为}$$

为该批轴料的椭圆度的方差和规定的无明显差别。

七. (10 分) 解: X_i 表示木柱短于 3 米, 则 $X_1, X_2, \dots, X_{100}$ 独立同分布且 $X_i \sim B(1, 0.2)$,

$$E[X_i] = 0.2, \text{Var}(X_i) = 0.16, \text{---} 2 \text{ 分, } 100 \text{ 根中短于 3 米的根数为 } \sum_{i=1}^{100} X_i, \text{ 因此}$$

$$\begin{aligned} P\left\{\sum_{i=1}^{100} X_i \geq 30\right\} &= P\left\{\frac{\sum_{i=1}^{100} X_i - E\left(\sum_{i=1}^{100} X_i\right)}{\sqrt{\text{Var}\left(\sum_{i=1}^{100} X_i\right)}} \geq \frac{30 - E\left(\sum_{i=1}^{100} X_i\right)}{\sqrt{\text{Var}\left(\sum_{i=1}^{100} X_i\right)}}\right\} = P\left(\frac{\sum_{i=1}^{100} X_i - 20}{4} \geq \frac{30 - 20}{4}\right) \\ &\approx 1 - \Phi(2.5) = 1 - 0.9938 = 0.0062 \end{aligned}$$

$$\text{八. (10 分) 解: } EX = \theta \int_0^1 x dx + (1 - \theta) \int_1^2 x dx = \frac{3-2\theta}{2},$$

$$\text{令 } \frac{3-2\hat{\theta}}{2} = \bar{X}, \text{ 得到 } \theta \text{ 的矩估计为 } \hat{\theta}_M = \frac{3-2\bar{X}}{2}$$

似然函数 $L(\theta) = \prod_{i=1}^n f(X_i) = \theta^N (1 - \theta)^{n-N}$ ---1 分,

$$\ln L(\theta) = N \ln(\theta) + (n - N) \ln(1 - \theta), \quad \frac{\partial \ln L}{\partial \theta}(\theta) = \frac{N}{\theta} - \frac{n - N}{1 - \theta},$$

$$\text{令 } \frac{\partial \ln L}{\partial \theta}(\theta) = 0, \text{ 得到 } \hat{\theta}_L = \frac{N}{n}. \text{---} 3 \text{ 分}$$

由于 $P(X \leq 1) = \theta$, 随机变量 N 服从二项分布, $N \sim B(n, \theta)$.

因此 $EN = \theta n$,

推出 $E\hat{\theta}_L = \theta$. 最大似然估计是 θ 的无偏估计。

概率统计期末考试模拟题（五） 2021.1

一、单项选择题（每小题 3 分，共 15 分）

1. 下列说法错误的是（ ）。
- (A) A, B 互为对立事件, 则 A, B 一定互不相容
(B) A, B 互不相容, 且 $P(A), P(B)$ 均不为 0, 则 A, B 一定不是相互独立的
(C) 同一个试验的基本事件（即由单个样本点构成的随机事件）一定互不相容
(D) 同一个试验的基本事件一定相互独立
2. 连续型随机变量 X 的密度函数 $f(x)$ 必满足条件（ ）。
- (A) $0 \leq f(x) \leq 1$ (B) $\int_{-\infty}^{+\infty} f(x)dx = 1$
(C) 是单调不减函数 (D) 是连续函数
3. 下列说法错误的是（ ）。
- (A) 若 (X, Y) 服从二维正态分布, 则 X 和 Y 都服从一维正态分布;
(B) 若 (X, Y) 服从二维正态分布, 则 $X+Y$ 服从一维正态分布;
(C) 若 X 和 Y 都服从一维正态分布, 且 X, Y 相互独立, 则 $X+Y$ 服从一维正态分布;
(D) 若 X 和 Y 都服从一维正态分布, 且 $\rho_{XY} = 0$, 则 X, Y 相互独立。
4. 设总体 X 的均值 μ 、方差 σ^2 均未知, $(X_1, X_2, \dots, X_n)$ 为来自该总体的样本, $\bar{X}$ 为样本均值, 则 σ^2 的矩估计量是（ ）。

- (A) $\bar{X}$ (B) $\frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2$
(C) $\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$ (D) $\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$

5. 设 (X_1, X_2) 是来自总体 X 的样本, 则总体均值 μ 最有效的无偏估计是（ ）。

- (A) $\hat{\mu}_1 = \frac{1}{2}X_1 + \frac{1}{2}X_2$ (B) $\hat{\mu}_2 = \frac{1}{3}X_1 + \frac{2}{3}X_2$
(C) $\hat{\mu}_3 = \frac{1}{4}X_1 + \frac{3}{4}X_2$ (D) $\hat{\mu}_4 = \frac{2}{5}X_1 + \frac{3}{5}X_2$

二、填空题（每小题 3 分，共 15 分）

6. 若 $P(A \cup B) = \frac{3}{5}, P(\bar{A} \cup B) = \frac{9}{10}$, 则 $P(B) = \underline{\hspace{2cm}}$ 。

7. 随机变量 X 的分布函数为 $F(x) = \begin{cases} 0, & x < 1 \\ 0.4, & 1 \leq x < 2 \\ 0.5, & 2 \leq x < 3 \\ 1, & x \geq 3 \end{cases}$, 则

$P\{X=1\} = \underline{\hspace{2cm}}$, $P\{X=2\} = \underline{\hspace{2cm}}$,

$P\{X=3\} = \underline{\hspace{2cm}}$ 。

8. 设随机变量 $X \sim N(\frac{1}{2}, 2)$, 以 Y 表示对 X 的三次独立重复观察中 $\{X \leq \frac{1}{2}\}$ 出现的次数, 则 $P\{Y=2\} = \underline{\hspace{2cm}}$ 。(结果用分数表示)

9. 设某种病菌在人群中的带菌率为 10%, 检测时带菌者呈阳性、阴性反应的概率分别为 0.95 和 0.05, 不带菌者呈阳性、阴性反应的概率分别为 0.01 和 0.99。今某人检测一次呈阳性, 则他是带菌者的概率为 $\underline{\hspace{2cm}}$ 。(结果用分数表示)

10. 设两总体 $X \sim N(0, 1), Y \sim N(0, 4)$ 相互独立, X_1, X_2, X_3, X_4 和 $Y_1, Y_2, Y_3, \dots, Y_9$ 分别为来自总体 X 和 Y 的样本, 则统计量 $Z = \frac{3(X_1 + X_2 + X_3 + X_4)}{\sqrt{\sum_{i=1}^9 Y_i^2}}$ 服从分布 $\underline{\hspace{2cm}}$ 。(给出分布类型和参数)

三、判断题: 判断以下说法正确与否, 正确的打 $\checkmark$, 错误的打 $\times$ (每小题 2 分, 共 10 分)

11. 若 $P(A \cup B) = P(A) + P(B)$, 则 $AB = \emptyset$ 。()

12. $P(A|B)$ 的含义是事件 A, B 都发生的概率。()

13. 取值范围为连续区间的随机变量就是连续型随机变量。()

14. 伯努利大数定律为实际中用频率近似代替概率提供了理论依据。()

15. 假设检验可能犯弃真错误和纳伪错误, 在样本容量一定时, 不能同时减小犯这两类错误的概率。()

四、解答题 (每小题 10 分, 共 60 分)

16. 加法器在做加法运算时, 根据四舍五入原则先对每个数取整后再运算, 这样产生的误差服从区间 $[-0.5, 0.5]$ 上的均匀分布。问: 要使误差总和的绝对值不超过 10 的概率大于 0.95, 最多能有多少个数相加? ($\Phi(0.95) = 0.829$,

$\Phi(0.05) = 0.52$, $\Phi(1.645) = 0.95$, $\Phi(1.96) = 0.975$)

17. 设随机变量 X, Y 的概率密度分别为

$$f_X(x) = \begin{cases} \lambda_1 e^{-\lambda_1 x}, & x > 0 \\ 0, & x \leq 0 \end{cases}, f_Y(y) = \begin{cases} \lambda_2 e^{-\lambda_2 y}, & y > 0 \\ 0, & y \leq 0 \end{cases}, (\lambda_1 > 0, \lambda_2 > 0, \lambda_1 \neq \lambda_2),$$

且 X, Y 相互独立。

(1) 求 $Z = X + Y$ 的概率密度;

(2) 根据(1)中结论回答指数分布是否具有可加性, 并列举两种具有可加性的分布类型。

18. 设随机变量 (X, Y) 的概率密度为

$$f(x, y) = \begin{cases} c, & |y| < x, 0 < x < 2 \\ 0, & \text{其他} \end{cases},$$

(1) 求常数 c ; (2) 求边缘概率密度 $f_X(x), f_Y(y)$;

(3) 问 X 与 Y 是否相互独立? 是否不相关? 请说明理由。

19. 设某校女生的身高服从正态分布 $N(\mu, \sigma^2)$, 今从该校随机抽取 9 名女生, 计算身高样本均值和样本标准差为: $\bar{x} = 162.67\text{cm}$, $s = 4.20\text{cm}$, 求身高方差 σ^2 的置信度为 0.95 的置信区间。

$$(\chi_{0.025}^2(8) = 17.535, \chi_{0.975}^2(8) = 2.18; \chi_{0.025}^2(9) = 19.02, \chi_{0.975}^2(9) = 2.7)$$

$$20. \text{ 设总体 } X \text{ 的密度函数为 } f(x) = \begin{cases} \frac{1}{\sigma} x e^{-\frac{x^2}{2\sigma}} & x \geq 0 \\ 0 & x < 0 \end{cases}, \text{ 其中 } \sigma > 0 \text{ 未知,}$$

$(X_1, X_2, \dots, X_n)$ 为来自 X 的样本, 求 σ 的极大似然估计量。

21. 已知人每分钟的脉搏次数服从正态分布, 正常人的平均脉搏为 72 次/每分钟, 现测得 9 例砒剂中毒患者的脉搏, 算得平均次数为 66.4 次, 样本方差为 5.92^2 。试问: 中毒患者与正常人的平均脉搏有无显著差异? ($\alpha = 0.05$)

$$(\Phi(1.96) = 0.975, \Phi(1.65) = 0.95, t_{0.025}(8) = 2.306, t_{0.05}(8) = 1.8595, t_{0.025}(9) = 2.2622, t_{0.05}(9) = 1.8331)$$

概率统计期末考试模拟题（五）答案

一、单项选择题（每小题 3 分，共 15 分） D B D C A

二、填空题 6. 0.5 7. 0.4; 0.1; 0.5 8. 3/8

9. 95/104 10. $t(9)$

三、判断题 X X X $\checkmark$ $\checkmark$

四、16. 设 X_i 为第 i 个数四舍五入产生的误差，则 $X_i \sim U(-0.5, 0.5)$ ， $i=1, 2, \dots, n$ 相互独立。

$$\text{误差总和为 } \sum_{i=1}^n X_i, \quad E\left(\sum_{i=1}^n X_i\right) = 0, \quad D\left(\sum_{i=1}^n X_i\right) = \frac{n}{12},$$

$$\begin{aligned} P\{|\sum_{i=1}^n X_i| \leq 10\} &= P\left\{-\frac{10}{\sqrt{n/12}} \leq \frac{\sum_{i=1}^n X_i}{\sqrt{n/12}} \leq \frac{10}{\sqrt{n/12}}\right\} \\ &\approx 2\Phi\left(\frac{10}{\sqrt{n/12}}\right) - 1 \geq 0.95 \Rightarrow \frac{10}{\sqrt{n/12}} \geq 1.96 \end{aligned}$$

解得 $n \leq 312$ 故最多能有 312 个这样的数相加。

$$17. \quad f_X(x)f_Y(z-x) = \begin{cases} \lambda_1 \lambda_2 e^{(\lambda_2 - \lambda_1)x - \lambda_2 z}, & x > 0, z > x \\ 0, & \text{其他} \end{cases}$$

$$\begin{aligned} f_Z(z) &= f_{X+Y}(z) = \int_{-\infty}^{+\infty} f_X(x)f_Y(z-x)dx \\ &= \begin{cases} 0, & z \leq 0 \\ \int_0^z \lambda_1 \lambda_2 e^{(\lambda_2 - \lambda_1)x - \lambda_2 z} dx = \frac{\lambda_1 \lambda_2}{\lambda_2 - \lambda_1} (e^{-\lambda_1 z} - e^{-\lambda_2 z}), & z > 0 \end{cases} \end{aligned}$$

(2) 指数分布不具有可加性，
泊松分布、正态分布具有可加性。

$$18. \quad (1) \quad \iint_D f(x, y) dx dy = 1, \quad \text{解得 } c = \frac{1}{4}$$

$$(2) \quad f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy = \begin{cases} \int_{-x}^x \frac{1}{4} dy, & 0 < x < 2 \\ 0, & \text{其它} \end{cases} = \begin{cases} \frac{x}{2}, & 0 < x < 2 \\ 0, & \text{其它} \end{cases}$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx = \begin{cases} \int_{|y|}^2 \frac{1}{4} dx, & |y| < 2 \\ 0, & \text{其它} \end{cases} = \begin{cases} \frac{1}{4}(2 - |y|), & |y| < 2 \\ 0, & \text{其它} \end{cases}$$

(3) 显然，在面积非 0 的区域 $D = \{(x, y) \mid |y| < x, 0 < x < 2\}$ 上，有

$$f(x, y) \neq f_X(x)f_Y(y), \quad \text{所以 } X \text{ 与 } Y \text{ 不相互独立。}$$

$$\text{又因为 } E(Y) = \int_{-\infty}^{+\infty} y f_Y(y) dy = \int_{-2}^2 y \cdot \frac{1}{4}(2 - |y|) dy = 0$$

$$E(XY) = \iint_{R^2} xyf(x, y) dx dy = \int_0^2 dx \int_{-x}^x \frac{1}{4} xy dy = 0$$

故 $\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = 0$ ，因此 X 与 Y 不相关。

19. 因女生身高 $X \sim N(\mu, \sigma^2)$ ，故 $\chi^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$

$$P\{\chi_{0.975}^2(n-1) \leq \frac{(n-1)S^2}{\sigma^2} \leq \chi_{0.025}^2(n-1)\} = 0.95$$

解得 σ^2 的置信区间为: $\left(\frac{(n-1)S^2}{\chi_{0.025}^2(n-1)}, \frac{(n-1)S^2}{\chi_{0.975}^2(n-1)} \right)$

代入数据得置信区间为 $\left(\frac{8 \times 4.2^2}{17.535}, \frac{8 \times 4.2^2}{2.180} \right)$ 即 $(8.048, 64.734)$.

20. 设样本观测值为 $(x_1, x_2, \dots, x_n)$

$$\text{似然函数 } L(x_i, \sigma) = \left(\frac{1}{\sigma} \right)^n e^{-\sum_{i=1}^n \frac{x_i^2}{2\sigma^2}} x_1 x_2 \cdots x_n,$$

$$\ln L(x_i, \sigma) = -n \ln \sigma + \sum_{i=1}^n \ln x_i - \sum_{i=1}^n \frac{x_i^2}{2\sigma^2}$$

$$\frac{\partial \ln L(x_i, \sigma)}{\partial \sigma} = \frac{-n}{\sigma} + \sum_{i=1}^n \frac{x_i^2}{2\sigma^2} = 0,$$

$$\sigma = \frac{1}{2n} \sum_{i=1}^n x_i^2, \text{ 所求极大似然估计量为 } \hat{\sigma} = \frac{1}{2n} \sum_{i=1}^n X_i^2$$

21. 设中毒患者的脉搏次数 $X \sim N(\mu, \sigma^2)$ ， $H_0: \mu = \mu_0 = 72$ ， $H_1: \mu \neq \mu_0 = 72$

$$\text{当 } H_0 \text{ 成立时, } T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t(n-1)$$

$$H_0 \text{ 的拒绝域为 } W = \{ |t| > t_{\alpha/2}(n-1) \}$$

其中 $\alpha = 0.05, n = 9, \bar{x} = 66.4, s = 5.92$ 代入

$$|t| = \left| \frac{66.4 - 72}{5.92/\sqrt{9}} \right| \approx 2.838 > t_{0.025}(8) = 2.306$$

故拒绝 H_0 ，认为中毒患者与正常人的平均脉搏有显著差异。

概率统计期末模拟题（四）2020.6

一、选择题（每小题 3 分，共 15 分）

1. 设 A, B 是两个随机事件, 且 $0 < P(A) < 1, P(B) > 0, P(B|A) = P(B|\bar{A})$, 则必有 ().
 (A) $P(A|B) = P(\bar{A}|B)$ (B) $P(A|B) \neq P(\bar{A}|B)$
 (C) $P(AB) = P(A)P(B)$ (D) $P(AB) \neq P(A)P(B)$
2. 设随机变量 $X \sim N(\mu, 2^2), Y \sim N(\mu, 4^2)$, 记 $p_1 = P\{X \leq \mu - 2\}, p_2 = P\{Y \geq \mu + 4\}$, 则 ().
 (A) 对任意实数 μ , 有 $p_1 = p_2$ (B) 对任意实数 μ , 有 $p_1 < p_2$
 (C) 对任意实数 μ , 有 $p_1 > p_2$ (D) 对个别 μ 的取值, 有 $p_1 = p_2$

3. 设随机变量 X 的分布函数为 $F(x) = \begin{cases} 0, & x < 0 \\ 1/2, & 0 \leq x < 1 \\ 1 - e^{-x}, & x \geq 1 \end{cases}$, 则 $P\{X=1\} = ()$.
 (A) 0 (B) $\frac{1}{2}$ (C) $\frac{1}{2} - e^{-1}$ (D) $1 - e^{-1}$

4. 设随机变量 X 的取值范围是 $(-1, 1)$, 则以下可作为 X 概率密度的是 ().

(A) $f(x) = \begin{cases} x, & 0 \leq x < 1 \\ 0, & \text{其他} \end{cases}$ (B) $f(x) = \begin{cases} 2, & -1 \leq x < 1 \\ 0, & \text{其他} \end{cases}$
 (C) $f(x) = \begin{cases} 2x, & -1 \leq x < 1 \\ 0, & \text{其他} \end{cases}$ (D) $f(x) = \begin{cases} \frac{3}{2}x^2, & -1 \leq x < 1 \\ 0, & \text{其他} \end{cases}$

5. 下列说法不正确的是 ().

- (A) 二维正态分布的边缘分布是一维正态分布
 (B) 二维均匀分布的边缘分布是一维均匀分布
 (C) 二维正态分布的条件分布是一维正态分布
 (D) 二维均匀分布的条件分布是一维均匀分布

二、填空题（每小题 3 分，共 15 分）

1. 某出版社调查读者阅读习惯, 设 A 表示喜欢艺术类, B 表示喜欢文学类, C 表示喜欢科技类, 比例如下表。

事件	A	B	C	AB	AC	BC	ABC
比例	0.14	0.23	0.37	0.08	0.09	0.13	0.05

现随机抽取一位读者, 发现该读者喜欢科技类, 则他不喜欢艺术类和文学类的概率是_____。

2. 请列举三种具有可加性（再生性）的分布：_____、_____、_____。

3. 关于随机变量 $X_1, X_2, \dots, X_n$ 的相互独立、两两独立、两两不相关, 一般情况下, 三者的关系是：_____ $\Rightarrow$ _____ $\Rightarrow$ _____。

4. 设随机变量 $X \sim U(0,1)$, 则由切比雪夫不等式可得 $P\{|X - \frac{1}{2}| \geq \frac{1}{\sqrt{3}}\} \leq$ _____。

5. 设总体 $X \sim N(\mu, 4)$, 则样本容量至少为_____时, 才能以 95% 的把握保证 μ 的置信区间长度不大于 4? ($u_{0.025} = 1.96, u_{0.05} = 1.65$)

三、计算题（每题 10 分，共 70 分）

1. 设 X 和 Y 的分布律如下，且 $P\{XY=0\}=1$,

X	-1	0	1
p	1/4	1/2	1/4

Y	0	1
p	1/2	1/2

(1) 求 (X,Y) 的联合分布律；(2) 判断 X,Y 的独立性。

2. 随机变量 (X, Y) 的联合概率密度为 $f(x, y) = \begin{cases} e^{-y}, & y > x, x > 0; \\ 0, & \text{其它} \end{cases}$, 求

(1) $f_{X|Y}(x|y)$; (2) $E(XY)$ (化为二次 (累次) 积分即可, 不必计算结果)。

3. 设 $P\{X=0\}=P\{X=1\}=\frac{1}{2}$, $Y \sim U(0,1)$, 且 X, Y 相互独立, 求 $X+Y$ 的概率密度函数。

4. 独立重复地抛掷一枚均匀硬币 10000 次, 每次观察出现正面还是反面, 用中心极限定理估算: 在 99% 的把握之下, 能够保证正面出现的频率与概率的误差控制在什么范围之内?

($u_{0.01}=2.33, u_{0.005}=2.58$) (10 分)

5. 设 X_1, X_2 是来自总体 $X \sim N(0, \sigma^2)$ 的样本, 试讨论:

(1) $X_1 + X_2$ 与 $X_1 - X_2$ 是否相互独立? (2) $Y = \frac{(X_1 + X_2)^2}{(X_1 - X_2)^2}$ 服从什么分布?

6. 设测量误差服从零均值的正态分布, 进行 4 次独立测量, 各次误差为 0.1, -0.06, 0.04, -0.12。求测量误差方差的矩估计值和极大似然估计值。

7. 设某班车在两点间固定路线上的运行时间 $X \sim N(\mu, \sigma^2)$ (单位: 分钟)。现测得该班车 16 次运行时间, 计算得 $\bar{x} = 41.5, s^2 = 7.8$ 。在 0.1 的显著性水平下, 能否认为班车的平均运行时间为 40 分钟? ($u_{0.05}=1.645, u_{0.1}=1.28, t_{0.05}(15)=1.7531, t_{0.05}(16)=1.7459$)

概率统计期末模拟题（四）答案 2020.6

一、选择题（每小题 3 分，共 15 分）

1. 设 A, B 是两个随机事件, 且 $0 < P(A) < 1, P(B) > 0, P(B|A) = P(B|\bar{A})$, 则必有 (C)
(A) $P(A|B) = P(\bar{A}|B)$ (B) $P(A|B) \neq P(\bar{A}|B)$
(C) $P(AB) = P(A)P(B)$ (D) $P(AB) \neq P(A)P(B)$
2. 设随机变量 $X \sim N(\mu, 2^2), Y \sim N(\mu, 4^2)$, 记 $p_1 = P\{X \leq \mu - 2\}, p_2 = P\{Y \geq \mu + 4\}$, 则 (A)
(A) 对任意实数 μ , 有 $p_1 = p_2$ (B) 对任意实数 μ , 有 $p_1 < p_2$
(C) 对任意实数 μ , 有 $p_1 > p_2$ (D) 对个别 μ 的取值, 有 $p_1 = p_2$

3. 设随机变量 X 的分布函数为 $F(x) = \begin{cases} 0, & x < 0 \\ 1/2, & 0 \leq x < 1 \\ 1 - e^{-x}, & x \geq 1 \end{cases}$, 则 $P\{X=1\} =$ (C)

(A) 0 (B) $\frac{1}{2}$ (C) $\frac{1}{2} - e^{-1}$ (D) $1 - e^{-1}$

4. 设随机变量 X 的取值范围是 $(-1, 1)$, 则以下可作为 X 概率密度的是 (D)

(A) $f(x) = \begin{cases} 1, & 0 \leq x < 1 \\ 0, & \text{其他} \end{cases}$ (B) $f(x) = \begin{cases} 2, & -1 \leq x < 1 \\ 0, & \text{其他} \end{cases}$
(C) $f(x) = \begin{cases} 2x, & -1 \leq x < 1 \\ 0, & \text{其他} \end{cases}$ (D) $f(x) = \begin{cases} \frac{3}{2}x^2, & -1 \leq x < 1 \\ 0, & \text{其他} \end{cases}$

5. 下列说法不正确的是 (B)

(A) 二维正态分布的边缘分布是一维正态分布
(B) 二维均匀分布的边缘分布是一维均匀分布
(C) 二维正态分布的条件分布是一维正态分布
(D) 二维均匀分布的条件分布是一维均匀分布

二、填空题（每小题 3 分，共 15 分）

1. 某出版社调查读者阅读习惯, 设 A 表示喜欢艺术类, B 表示喜欢文学类, C 表示喜欢科技类, 比例如下表。

事件	A	B	C	AB	AC	BC	ABC
比例	0.14	0.23	0.37	0.08	0.09	0.13	0.05

现随机抽取一位读者, 发现该读者喜欢科技类, 则他不喜欢艺术类和文学类的概率是 $\frac{20}{37}$ 。

2. 请列举三种具有可加性（再生性）的分布：二项分布、泊松分布、正态分布。

3. 关于随机变量 $X_1, X_2, \dots, X_n$ 的相互独立、两两独立、两两不相关, 一般情况下, 三者的关系是：相互独立 $\Rightarrow$ 两两独立 $\Rightarrow$ 两两不相关。

4. 设随机变量 $X \sim U(0, 1)$, 则由切比雪夫不等式可得 $P\{|X - \frac{1}{2}| \geq \frac{1}{\sqrt{3}}\} \leq \frac{1}{4}$ 。

5. 设总体 $X \sim N(\mu, 4)$, 则样本容量至少为 4 时, 才能以 95% 的把握保证 μ 的置信区间长度不大于 4? ($u_{0.025} = 1.96, u_{0.05} = 1.65$)

三、计算题（每题 10 分，共 70 分）

1. 设 X 和 Y 的分布律如下，且 $P\{XY=0\}=1$,

X	-1	0	1
p	1/4	1/2	1/4

Y	0	1
p	1/2	1/2

(1) 求 (X,Y) 的联合分布律；(2) 判断 X,Y 的独立性。

解：(1)

$$\begin{aligned} 1 &= P\{XY=0\} = P\{X=0 \cup Y=0\} \\ &= P\{X=0\} + P\{Y=0\} - P\{X=0, Y=0\} \\ &= \frac{1}{2} + \frac{1}{2} - P\{X=0, Y=0\} \end{aligned}$$

故 $P\{X=0, Y=0\}=0$ 。

又由 $P\{XY \neq 0\}=0$ 知 $P\{X=-1, Y=1\}=P\{X=1, Y=1\}=0$ ，由此可得

$$P\{X=-1, Y=0\} = P\{X=-1\} - P\{X=-1, Y=1\} = \frac{1}{4},$$

$$P\{X=0, Y=1\} = P\{X=0\} - P\{X=0, Y=0\} = \frac{1}{2},$$

$$P\{X=1, Y=0\} = P\{X=1\} - P\{X=1, Y=1\} = \frac{1}{4},$$

$$(2) \text{易知 } 0 = P\{X=0, Y=0\} \neq P\{X=0\}P\{Y=0\} = \frac{1}{4}$$

故 X,Y 不相互独立。

2. 随机变量 (X, Y) 的联合概率密度为 $f(x, y) = \begin{cases} e^{-y}, & y > x, x > 0; \\ 0, & \text{其它} \end{cases}$ ，求

(1) $f_{X|Y}(x|y)$ ；(2) $E(XY)$ (化为二次（累次）积分即可，不必计算结果)。

$$\text{解：(1) } f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx = \begin{cases} 0, & y \leq 0 \\ \int_0^y e^{-y} dx = ye^{-y}, & y > 0 \end{cases}$$

$$\text{当 } y > 0 \text{ 时, } f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)} = \begin{cases} \frac{1}{y}, & 0 < x < y \\ 0, & \text{其它} \end{cases}, \text{ 其余时候条件密度不存在。}$$

$$(2) E(XY) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xyf(x, y) dx dy = \int_0^{+\infty} \left[\int_0^y xye^{-y} dx \right] dy = \int_0^{+\infty} \left[\int_x^{+\infty} xye^{-y} dy \right] dx$$

3. 设 $P\{X=0\}=P\{X=1\}=\frac{1}{2}$ ， $Y \sim U(0,1)$ ，且 X, Y 相互独立，求 $X+Y$ 的概率密度函数。

解：令 $Z=X+Y$ ，则

$$\begin{aligned} F_Z(z) &= P\{Z \leq z\} \\ &= P\{X+Y \leq z\} \\ &= P\{X+Y \leq z, X=0\} + P\{X+Y \leq z, X=1\} \\ &= P\{X+Y \leq z | X=0\}P\{X=0\} + P\{X+Y \leq z | X=1\}P\{X=1\} \\ &= \frac{P\{Y \leq z\}}{2} + \frac{P\{Y \leq z-1\}}{2} \end{aligned}$$

$$= \begin{cases} 0+0=0, & z < 0 \\ \frac{z}{2}+0=\frac{z}{2}, & 0 \leq z < 1 \\ \frac{1}{2}+\frac{z-1}{2}=\frac{z}{2}, & 1 \leq z < 2 \\ \frac{1}{2}+\frac{1}{2}=1, & z \geq 2 \end{cases} = \begin{cases} 0, & z < 0 \\ \frac{z}{2}, & 0 \leq z < 2 \\ 1, & z \geq 2 \end{cases}$$

$$f_z(z) = F'_z(z) = \begin{cases} \frac{1}{2}, & 0 < z < 2 \\ 0, & \text{其他} \end{cases}, \text{故 } Z=X+Y \sim U(0,2).$$

4. 独立重复地抛掷一枚均匀硬币 10000 次，每次观察出现正面还是反面，用中心极限定理估算：在 99% 的把握之下，能够保证正面出现的频率与概率的误差控制在什么范围之内？

($u_{0.01} = 2.33, u_{0.005} = 2.58$) (10 分)

解：设正面出现的次数为 X ，则 $X \sim B(10000, \frac{1}{2})$ ，正面出现的频率为 $\frac{X}{10000}$ ，

要求 δ 使 $P\{|\frac{X}{10000} - \frac{1}{2}| < \delta\} = 0.99$

$$E(X) = 5000, D(X) = 2500,$$

$$P\{5000 - 10000\delta < X < 5000 + 10000\delta\} = P\{\frac{-10000\delta}{\sqrt{2500}} < \frac{X - 5000}{\sqrt{2500}} < \frac{10000\delta}{\sqrt{2500}}\}$$

由棣莫佛拉普拉斯中心极限定理， $\frac{X - 5000}{\sqrt{2500}}$ 近似服从标准正态分布，故

$$\approx \Phi(200\delta) - \Phi(-200\delta) = 2\Phi(200\delta) - 1 = 0.99$$

$$\Rightarrow \Phi(200\delta) = 0.995 \Rightarrow 200\delta = 2.58 \Rightarrow \delta = 0.0129$$

可以 99% 的把握保证正面出现的频率和概率的误差控制在 0.0129 的范围内。

5. 设 X_1, X_2 是来自总体 $X \sim N(0, \sigma^2)$ 的样本，试讨论：

(1) $X_1 + X_2$ 与 $X_1 - X_2$ 是否相互独立？(2) $Y = \frac{(X_1 + X_2)^2}{(X_1 - X_2)^2}$ 服从什么分布？

解样本 X_1, X_2 相互独立且同服从 $N(0, \sigma^2)$ 分布，故 $(X_1, X_2)^T$ 服从二维正态分布 $\mathbf{X} \sim N(\mathbf{0}, \sigma^2; \mathbf{0}, \sigma^2, \mathbf{0})$

(1) 因 $\begin{pmatrix} X_1 + X_2 \\ X_1 - X_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ ， $(X_1 + X_2, X_1 - X_2)^T$ 是二维正态随机变量

$(X_1, X_2)^T$ 的线性变换，故仍服从二维正态分布。又

$$\text{Cov}(X_1 + X_2, X_1 - X_2) = \text{Cov}(X_1, X_1) - \text{Cov}(X_1, X_2) + \text{Cov}(X_1, X_2) - \text{Cov}(X_2, X_2)$$

$$= D(X_1) - \text{Cov}(X_1, X_2) + \text{Cov}(X_1, X_2) - D(X_2) = \sigma^2 - \sigma^2 = 0$$

故 $X_1 + X_2$ 与 $X_1 - X_2$ 不相关，由 $(X_1 + X_2, X_1 - X_2)^T$ 服从二维正态知 $X_1 + X_2$ 与 $X_1 - X_2$ 相互独立。

(2)

$$E(X_1 + X_2) = 0, E(X_1 - X_2) = 0,$$

$$D(X_1 + X_2) = 2\sigma^2, D(X_1 - X_2) = 2\sigma^2$$

根据 χ^2 分布结构定理

$$\chi_1^2 = \frac{(X_1 + X_2)^2}{2\sigma^2} \sim \chi^2(1), \quad \chi_2^2 = \frac{(X_1 - X_2)^2}{2\sigma^2} \sim \chi^2(1)$$

由第一问 $X_1 + X_2$ 与 $X_1 - X_2$ 的独立性知 χ_1^2 与 χ_2^2 也相互独立, 再由 F 分布结构定理知

$$Y = \frac{(X_1 + X_2)^2}{(X_1 - X_2)^2} = \frac{(X_1 + X_2)^2 / 2\sigma^2}{(X_1 - X_2)^2 / 2\sigma^2} \sim F(1, 1)$$

6. 设测量误差服从零均值的正态分布, 进行 4 次独立测量, 各次误差为 0.1, -0.06, 0.04, -0.12。求测量误差方差的矩估计值和极大似然估计值。

解: 设测量误差 $X \sim N(0, \sigma^2)$,

(1) 矩估计: 总体方差 σ^2 的矩估计量为样本二阶中心矩 $\frac{1}{4} \sum_{i=1}^4 (X_i - \bar{X})^2$ (答案不唯一),

代入样本观测值 0.1, -0.06, 0.04, -0.12 得矩估计值 $\hat{\sigma}^2 = 0.0073$.

(2) 极大似然估计: x_1, x_2, x_3, x_4 为样本观测值。

似然函数为

$$L(\sigma) = \prod_{i=1}^4 \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x_i^2}{2\sigma^2}} = \frac{1}{4\pi^2\sigma^4} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^4 x_i^2} = \frac{1}{4\pi^2\sigma^4} e^{-\frac{0.0148}{\sigma^2}}$$

令 $\frac{d \ln L}{d\sigma} = 0$, 得极大似然估计值 $\hat{\sigma}^2 = 0.0074$ 。

7. 设某班车在两点间固定路线上的运行时间 $X \sim N(\mu, \sigma^2)$ (单位: 分钟)。现测得该班车 16 次运行时间, 计算得 $\bar{x} = 41.5, s^2 = 7.8$ 。在 0.1 的显著性水平下, 能否认为班车的平均运行时间为 40 分钟? ($u_{0.05} = 1.645, u_{0.1} = 1.28, t_{0.05}(15) = 1.7531, t_{0.05}(16) = 1.7459$)

解: 检验 $H_0: \mu = 40; H_1: \mu \neq 40$ 。

$$\text{当 } H_0 \text{ 成立时, 统计量 } T = \frac{\bar{X} - 40}{S/4} \sim t(15)$$

由条件得: $\bar{x} = 41.5, s^2 = 7.8$,

$$|t| = \left| \frac{\bar{x} - 40}{s/4} \right| \approx 2.15 > t_{0.05}(15) = 1.7531$$

在拒绝域内, 故拒绝 H_0 , 不可认为该班车的平均运行时间为 40 分钟。

概率统计期末模拟题（三） 2020.1

一、填空题（每空 3 分，共 18 分）

1. 设 A, B 互不相容, $P(A) = 0.3, P(B) = 0.5$, 则 $P(\bar{B}|\bar{A}) =$ _____.
2. 设 X 服从二项分布 $B(2, p)$, Y 的分布律为 $P\{Y = k\} = (1-p)^{k-1}p, k=1, 2, \dots$,
若已知 $P(X \geq 1) = \frac{5}{9}$, 则 $D(Y) =$ _____.
3. 设随机变量 (X, Y) 服从区域 $G = \{(x, y) | x^2 + y^2 \leq 1\}$ 上的均匀分布, 则
 $P\left(\min(X, Y) \geq -\frac{1}{2}, \max(X, Y) \leq \frac{1}{2}\right) =$ _____.
4. 若 X_1, X_2, X_3 独立同分布, $P(X_n = 1) = \frac{1}{4}, P(X_n = -1) = \frac{3}{4}, Y_n = X_n X_{n+1}, n = 1, 2,$
则 $P(Y_2 = 1 | Y_1 = 1) =$ _____.
5. 若有一个容量为 11 的样本, 其样本均值为 2, 样本二阶原点矩为 10, 则样本
方差为_____.
6. 设 $(X_1, X_2, \dots, X_5)$ 是取自总体 $N(0, 1)$ 的样本, 统计量 $T = \frac{X_1 + X_2}{\sqrt{C(X_3 + X_4)^2 + X_5^2}}$, 则当
 $C =$ _____时, T 服从 $t(2)$ 分布.

二、单项选择题（每题 3 分，共 15 分）

1. 设随机变量 X 的分布函数 $F(x) = 0.4\Phi(x) + 0.6\Phi(x-1)$, $\Phi(x)$ 为标准正态分布函数, 则 $E(X) =$ ().
(A) 0 (B) 0.4 (C) 0.5 (D) 0.6
2. 将长为 1 米的木棍随机截成两段, 则此两段长度的相关系数等于 ().
(A) -1 (B) 0 (C) 0.5 (D) 1
3. 设随机变量 $X_i \sim \begin{pmatrix} -1 & 0 & 1 \\ 1/4 & 1/2 & 1/4 \end{pmatrix}, i = 1, 2$, 且满足 $P(X_1 + X_2 = 0) = 1$, 则
 $P(X_1 = X_2) =$ ().
(A) 0 (B) $\frac{1}{4}$ (C) $\frac{1}{2}$ (D) 1
4. 设 $\{X_n\}_{n=1}^{+\infty}$ 独立同分布于 Poisson 分布 $P(\lambda)$, $S_n = \sum_{i=1}^n X_i$, 记 $\Phi(x)$ 为标准正态分布函数, 则有 ().
(A) $\lim_{n \rightarrow \infty} P\left(\frac{S_n - n\lambda}{n\lambda} \leq x\right) = \Phi(x)$ (B) $\lim_{n \rightarrow \infty} P\left(\frac{S_n - n\lambda}{\sqrt{n\lambda}} \leq x\right) = \Phi(x)$
(C) $\lim_{n \rightarrow \infty} P\left(\frac{S_n - n\lambda}{\sqrt{n\lambda}} \leq x\right) = \Phi(x)$ (D) $\lim_{n \rightarrow \infty} P\left(\frac{S_n - \lambda}{\sqrt{n\lambda}} \leq x\right) = \Phi(x)$

5 下面说法中正确的是 ().

- ① 样本均值是总体期望的无偏估计
- ② 样本 k 阶原点矩是总体 k 阶原点矩的无偏估计
- ③ 样本方差是总体方差的无偏估计
- ④ 样本 k 阶中心矩是总体 k 阶中心矩的无偏估计

(A) ①②③ (B) ①③④ (C) ①②④ (D) ①②③④

三、(10 分) 设某地区移动、电信、联通的用户比例为 4:3:2, 一份对运营商的抽样调查数据表明: 移动、电信、联通的好评率分别为 80%、60%、70%. 现从这些数据资料中任取一位用户的评价,

- (1) 求该评价为好评的概率;
- (2) 若该评价是好评, 求该用户是电信用户的概率.

四、(10 分) 设 $F(x) = \begin{cases} 0, & x < -1, \\ a + b \arcsin x, & -1 \leq x < 1, \\ 1, & x \geq 1, \end{cases}$

- (1) a, b 满足什么条件时, $F(x)$ 为某随机变量的分布函数?
- (2) 若 $F(x)$ 为连续型随机变量的分布函数, 求 a, b .

五、(10 分) 设 X 服从参数为 λ 的指数分布, $Y = [X] + 1$, $[x]$ 为不超过 x 的最大整数. (1) 求 Y 的分布; (2) 求在已知 $Y = 2$ 的条件下, X 的条件概率密度.

六、(12 分) 设随机变量 ξ, η 独立同分布, $\xi \sim U\left(-\frac{\pi}{3}, \frac{\pi}{3}\right)$.

- (1) 求 $\xi + \eta$ 的概率密度;
- (2) 若 $X = \xi \cos \eta, Y = \xi \sin \eta$, 问 X, Y 是否不相关?

七、(10 分) 设总体 X 的密度

$$f(x) = \begin{cases} a|x - \mu|, & \mu - \theta \leq x \leq \mu + \theta, \\ 0, & \text{其它,} \end{cases}, (X_1, X_2, \dots, X_n) \text{ 为取自总体 } X \text{ 的样本.}$$

- (1) 求 a ;
- (2) 若 μ 已知, 求 θ 的矩估计量.

八、(15 分) 设 $(X_1, X_2, \dots, X_n)$ 是取自总体 $X \sim N(\mu, \sigma^2)$ 的样本.

- (1) 若 μ 已知, σ^2 未知, 求 σ^2 的极大似然估计量;
- (2) 若 μ 已知, 由 (1) 构造 σ 的置信水平为 95% 的双侧置信区间;
- (3) 若已知 $\sigma = 2$, 考虑如下的假设检验问题,

$$H_0: \mu = 2 \quad \text{vs} \quad H_1: \mu = 3$$

检验由拒绝域 $W = \{\bar{x} \geq 2.8\}$ 确定, 当 $n = 36$ 时, 求检验犯两类错误的概率.

(用 $\Phi(\cdot)$ 表示, 不用计算)

概率统计期末模拟题（三）答案 2020.1

一、填空题（每空 3 分，共 18 分） 1. $\frac{2}{7}$ 2. $\frac{6}{9}$ 3. $\frac{1}{\pi}$ 4. $\frac{0.7}{9}$ 5. $\frac{6.6}{9}$ 6. $\frac{0.5}{9}$

二、单项选择题（每题 3 分，共 15 分） D A C C A

三、（10 分）记 $A_1 = \{\text{移动用户}\}, A_2 = \{\text{电信用户}\}, A_3 = \{\text{联通用户}\}, B = \{\text{好评}\},$

（1）由全概率公式， $P(B) = \sum_{i=1}^3 P(A_i)P(B|A_i) = \frac{4}{9} \times 0.8 + \frac{3}{9} \times 0.6 + \frac{2}{9} \times 0.7 = \frac{32}{45}.$

（2）由 Bayes 公式， $P(A_2|B) = \frac{P(A_2)P(B|A_2)}{P(B)} = \frac{\frac{3}{9} \times 0.6}{\frac{32}{45}} = \frac{9}{32}.$

四、（10 分）（1）由定义， $F(x)$ 已满足右连续性和 $F(-\infty) = 0, F(+\infty) = 1$ ，还需要满足单调不减性，所以 $b \geq 0, a - \frac{\pi}{2}b \geq 0, a + \frac{\pi}{2}b \leq 1.$

（2）若 $F(x)$ 为连续型随机变量的分布函数，则 $a - \frac{\pi}{2}b = 0, a + \frac{\pi}{2}b = 1, a = \frac{1}{2}, b = \frac{1}{\pi}.$

五、（10 分）（1） Y 的取值为自然数，对 $k = 1, 2, \dots,$

$$P(Y = k) = P(k-1 \leq X < k) = \int_{k-1}^k \lambda e^{-\lambda x} dx = (e^{-\lambda})^{k-1} (1 - e^{-\lambda}),$$

所以， $Y \sim G(1 - e^{-\lambda}).$

（2）在已知 $Y = 2$ 的条件下， X 取值于区间 $[1, 2)$ ，对 $x \in [1, 2)$,

$$P(X \leq x | Y = 2) = \frac{P(1 \leq X \leq x)}{P(Y=2)} = \frac{e^{-\lambda} - e^{-\lambda x}}{e^{-\lambda} - e^{-2\lambda}},$$

$$\therefore f(x|Y=2) = \begin{cases} \frac{\lambda e^{-\lambda x}}{e^{-\lambda} - e^{-2\lambda}}, & x \in [1, 2), \\ 0, & \text{其它.} \end{cases}$$

六、（12 分）（1）由 $f_{\xi+\eta}(z) = \int_{-\infty}^{+\infty} f_{\xi}(x)f_{\eta}(z-x)dx,$

要使 $f_{\xi}(x)f_{\eta}(z-x) \neq 0$ ，必须 $-\frac{\pi}{3} < x < \frac{\pi}{3}, -\frac{\pi}{3} < z-x < \frac{\pi}{3}$ ，所以

对于 $z < -\frac{2\pi}{3}$ 或 $z > \frac{2\pi}{3}, f_{\xi+\eta}(z) = 0,$

对于 $-\frac{2\pi}{3} \leq z \leq 0, f_{\xi+\eta}(z) = \int_{-\frac{\pi}{3}}^{z+\frac{\pi}{3}} \left(\frac{3}{2\pi}\right)^2 dx = \left(\frac{3}{2\pi}\right)^2 \left(z + \frac{2\pi}{3}\right),$

对于 $0 < z \leq \frac{2\pi}{3}, f_{\xi+\eta}(z) = \int_{z-\frac{\pi}{3}}^{\frac{\pi}{3}} \left(\frac{3}{2\pi}\right)^2 dx = \left(\frac{3}{2\pi}\right)^2 \left(\frac{2\pi}{3} - z\right).$

（2）因为 $E(X) = E(\xi)E(\cos \eta) = 0$ ，（9 分）所以

$$\text{cov}(X, Y) = E(XY) = E(\xi^2)E\left(\frac{1}{2}\sin 2\eta\right) = \frac{\left(\frac{2\pi}{3}\right)^2}{12} \int_{-\frac{\pi}{3}}^{\frac{\pi}{3}} \frac{1}{2}\sin 2x \frac{3}{2\pi} dx = 0,$$

所以 X, Y 不相关。（12 分）

七、（10 分）（1）由 $\int_{\mu-\theta}^{\mu+\theta} a|x-\mu|dx = a \int_0^{\theta} 2xdx = a\theta^2 = 1$ ，有 $a = \frac{1}{\theta^2}.$

$$(2) \mu \text{ 已知, } E(X) = \mu, D(X) = \int_{\mu-\theta}^{\mu+\theta} (x-\mu)^2 \frac{1}{\theta^2} |x-\mu| dx = \frac{\theta^2}{2},$$

$$\text{令 } \tilde{S}^2 = \frac{\theta^2}{2}, \text{ 得 } \theta \text{ 的矩估计 } \hat{\theta}_M = \sqrt{2\tilde{S}} = \sqrt{\frac{2}{n} \sum_{i=1}^n (X_i - \bar{X})^2}. \quad (\text{注: 估计不唯一})$$

八、(15 分) (1) μ 已知, 极大似然函数

$$L(\sigma^2) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x_i-\mu)^2}{2\sigma^2}\right\} = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2\right\},$$

$$\text{对数极大似然函数 } \ln L(\sigma^2) = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2,$$

$$\text{求导 } \frac{d(\ln L(\sigma^2))}{d(\sigma^2)} = -\frac{n}{2} \frac{1}{\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^n (x_i - \mu)^2,$$

$$\text{令 } \frac{d(\ln L(\sigma^2))}{d(\sigma^2)} = 0, \text{ 得 } \sigma^2 \text{ 的极大似然估计 } \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2,$$

$$\text{所以, } \sigma^2 \text{ 的极大似然估计量 } \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2.$$

$$(2) \text{ 因为 } \sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma}\right)^2 \sim \chi^2(n),$$

$$\text{所以, } P\left(\chi_{1-\frac{\alpha}{2}}^2(n) \leq \frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma^2} \leq \chi_{\frac{\alpha}{2}}^2(n)\right) = 1 - \alpha, \text{ 有}$$

$$P\left(\sqrt{\frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi_{\frac{\alpha}{2}}^2(n)}} \leq \sigma \leq \sqrt{\frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi_{1-\frac{\alpha}{2}}^2(n)}}\right) = 1 - \alpha,$$

$$\sigma \text{ 的置信水平为 95\% 的双侧置信区间 } \left[\sqrt{\frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi_{0.025}^2(n)}}, \sqrt{\frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi_{0.975}^2(n)}} \right].$$

$$(3) \text{ 在 } H_0 \text{ 成立的条件下, } \bar{X} \sim N(2, \frac{4}{36}), \text{ 检验犯第一类错误的概率}$$

$$\alpha = P(\bar{X} \geq 2.8 | H_0) = P\left(\frac{\bar{X} - 2}{\sqrt{\frac{4}{36}}} \geq \frac{2.8 - 2}{\sqrt{\frac{4}{36}}}\right) = 1 - \Phi(2.4),$$

$$\text{在 } H_1 \text{ 成立的条件下, } \bar{X} \sim N(3, \frac{4}{36}), \text{ 检验犯第二类错误的概率}$$

$$\beta = P(\bar{X} < 2.8 | H_1) = P\left(\frac{\bar{X} - 3}{\sqrt{\frac{4}{36}}} < \frac{2.8 - 3}{\sqrt{\frac{4}{36}}}\right) = \Phi(-0.6) = 1 - \Phi(0.6).$$

概率统计期末模拟题（二）2019.6

(标准正态分布函数 $\Phi(x)$: $\Phi(1.56) = 0.941$, $\Phi(1.645) = 0.95$, $\Phi(1.96) = 0.975$;

$$\chi_{0.975}^2(9) = 2.70, \chi_{0.025}^2(9) = 19.02, \chi_{0.975}^2(10) = 3.25, \chi_{0.025}^2(10) = 20.48,$$

$$t_{0.025}(9) = 2.26, t_{0.025}(10) = 2.23, t_{0.05}(9) = 1.83, t_{0.05}(10) = 1.81)$$

一、完成下列各题（共 36 分）

1. (5 分) 某大厦有四部电梯, 已知某时刻 T , 每部电梯正在运行的概率为 0.7, 求 T 时刻至少两部电梯运行的概率.

2. (5 分) 随机变量 X 的分布列为

X	0	$\pi/2$	π
P	0.2	0.5	0.3

求 $D(\sin X)$.

3. (5 分) 已知 $P(A) = 0.7$, $P(B) = 0.4$, $P(\overline{AB}) = 0.8$, 求 $P(A|A \cup \overline{B})$.

4. (6 分) 成箱出售的玻璃杯, 每箱 20 只. 设每箱含 0, 1, 2 只残次品的概率分别为 0.8, 0.1, 0.1. 顾客购买时, 售货员随意取一箱, 而顾客随意取四只检查, 若无残次品则买下, 否则退回. 现售货员随意取一箱玻璃杯, 求顾客买下的概率 (保留小数点后三位小数).

5. (7 分) 随机变量 X, Y 满足 $EX = 1$, $EY = 2$, $D(X) = 1$, $D(Y) = 4$, $\rho_{XY} = 0.6$,

设 $Z = (2X - Y + 1)^2$, 求 $E(Z)$.

6. (8 分) 设总体 $X \sim N(0, \sigma^2)$, 且 $x_1, \dots, x_{10}$ 为样本观测值, 样本方差 $s^2 = 2$.

分别求 σ^2 和 $\text{Var}(\frac{X^2}{\sigma^3})$ 的置信水平为 0.95 的置信区间.

二、(10 分) 二维随机变量 (X, Y) 的联合密度函数为

$$f(x, y) = \begin{cases} \frac{1}{8}(x+y), & 0 \leq x \leq 2, 0 \leq y \leq 2; \\ 0, & \text{其他} \end{cases}$$

求 $Z = X + Y$ 的概率密度.

三、(10 分) 将 2 个球随机地放入 3 个盒子, 设 X 为第一个盒子内放入的球数, Y 为有球的盒子的个数, 求 (1) (X, Y) 的联合分布列; (2) $E(X)$, $D(Y)$.

四、(10 分) 二维随机变量 (X, Y) 的联合密度函数为

$$f(x, y) = \begin{cases} e^{-y}, & 0 \leq x \leq y \\ 0, & \text{其他} \end{cases}$$

(1) 求 $P(X+Y \leq 1)$; (2) X, Y 是否独立?

五、(10 分) 某种零件的尺寸标准为 $\sigma=5.2$ ，对一批这类零件检查 9 件得到平均尺寸数据 (毫米)： $\bar{x}=26.56$ 。设零件尺寸服从正态分布，问这批零件的平均尺寸能否认为是 26 毫米？ ($\alpha=0.05$)

六、(10 分) 一条生产线生产的产品合格率为 0.8，要使一批产品的合格率在 76% 与 84% 之间的概率不小于 90%，问至少要生产多少件产品？

七、(14 分) 设总体 X 的概率密度函数为 $f(x)=\begin{cases} \frac{1}{k\theta}, & \theta < x < (k+1)\theta; \\ 0, & \text{其他} \end{cases}$ 其中 k 为已知正数， $\theta > 0$ 为一个待估参数。 $X_1, \dots, X_n$ 为来自总体的样本。

(1) 求 θ 的最大似然估计量和矩估计量；(2) θ 的最大似然估计是否无偏？

概率统计期末模拟题（二）答案 2019.6

一、（5 分）1.某大厦有四部电梯，已知某时刻 T，每部电梯正在运行的概率为 0.7,求 T 时刻至少两部电梯运行的概率。

解： 设 X 为正在运行的电梯的个数，则 $X \sim b(4, 0.7)$ 。

$$P(X \geq 2) = 1 - P(X = 0) - P(X = 1) = 1 - 0.3^4 - 4 \times 0.3^3 \times 0.7 = 0.9163$$

2.（5 分）随机变量 X 的分布列为

X	0	$\pi/2$	π
P	0.2	0.5	0.3

求 $\text{Var}(\sin X)$.

解： $Y = \sin X$, $P(Y=0)=0.5$, $P(Y=1)=0.5$, $E[Y]=0.5$;

$$E[Y^2]=0.5, \quad \text{Var}(Y)=0.25.$$

3.（5 分）已知 $P(A)=0.7$, $P(B)=0.4$, $P(\overline{AB})=0.8$, 求 $P(A|A \cup \overline{B})$.

解： $P(A \cap \overline{B}) = P(A) - P(A \cap B) = P(A) - 1 + P(\overline{AB}) = 0.5$,

$$P(A|A \cup \overline{B}) = \frac{P(A \cap (A \cup \overline{B}))}{P(A \cup \overline{B})} = \frac{P(A)}{P(A) + P(\overline{B}) - P(A \cap \overline{B})} = \frac{0.7}{0.7 + 0.6 - 0.5} = \frac{7}{8}.$$

4.（6 分）成箱出售的玻璃杯，每箱 20 只。设每箱含 0, 1, 2 只残次品的概率分别为 0.8, 0.1, 0.1。顾客购买时，售货员随意取一箱，而顾客随意取四只检查，若无残次品则买下，否则退回。现售货员随意取一箱玻璃杯，求顾客买下的概率（保留小数点后三位小数）。

解：设事件 A 为售货员随意取一箱玻璃杯，顾客买下；

事件 C_i 为箱中有 i 个残次品， $i=0, 1, 2$ ，由全概率公式，

$$P(A) = \sum_{i=0}^2 P(A|C_i)P(C_i) = 1 \times 0.8 + \frac{C_{19}^4}{C_{20}^4} \times 0.1 + \frac{C_{18}^4}{C_{20}^4} \times 0.1 \approx 0.943$$

5.（7 分）随机变量 X, Y, 满足 $E[X]=1$, $E[Y]=2$, $\text{Var}[X]=1$, $\text{Var}[Y]=4$, $\rho_{XY} = 0.6$,

设 $Z = (2X - Y + 1)^2$, 求 $E[Z]$.

解： $E[X^2] = \text{Var}(X) + (E[X])^2 = 2$ （1 分）， $E[Y^2] = \text{Var}(Y) + (E[Y])^2 = 8$,

$$E[XY] = \text{Cov}(X, Y) + E[X]E[Y] = \rho_{XY} \sqrt{\text{Var}(X)} \sqrt{\text{Var}(Y)} + 2 = 3.2,$$

$$\begin{aligned} EZ &= E(2X - Y + 1)^2 = 4E[X^2] + E[Y^2] + 1 + 4E[X] - 4E[XY] - 2E[Y] \\ &= 8 + 8 + 1 + 4 - 4 \times 3.2 - 4 = 4.2 \end{aligned}$$

6. (8分) 设总体 $X \sim N(0, \sigma^2)$, 且 $x_1, \dots, x_{10}$ 为样本观测值, 样本方差 $s^2 = 2$ 。

分别求 σ^2 和 $\text{Var}(\frac{X^2}{\sigma^3})$ 的置信水平为 0.95 的置信区间。

解: σ^2 的置信水平为 0.95 的置信区间为

$$(\frac{(n-1)s^2}{\chi_{0.025}^2(9)}, \frac{(n-1)s^2}{\chi_{0.975}^2(9)}) = (\frac{18}{\chi_{0.025}^2(9)}, \frac{18}{\chi_{0.975}^2(9)}) = (0.9462, 6.6667),$$

由于 $\frac{X^2}{\sigma^2} \sim \chi^2(1)$, $\text{Var}(\frac{X^2}{\sigma^3}) = \frac{1}{\sigma^2} \text{Var}(\frac{X^2}{\sigma^2}) = \frac{2}{\sigma^2}$ 是 σ^2 的减函数, 因此 $\text{Var}(\frac{X^2}{\sigma^3})$

的置信水平为 0.95 的置信区间为 $(\frac{2}{6.6667}, \frac{2}{0.9462}) = (0.3000, 2.1137)$ 。

二、(10分) 二维随机变量 (X, Y) 的联合密度函数为

$$f(x, y) = \begin{cases} \frac{1}{8}(x+y), & 0 \leq x \leq 2, 0 \leq y \leq 2; \\ 0, & \text{其他} \end{cases}$$

求 $Z = X + Y$ 的概率密度。

解: $f_Z(z) = \int_{-\infty}^{+\infty} f(x, z-x) dx,$

当 $z < 0$ 或 $z > 4$, $f_Z(z) = 0$;

当 $0 \leq z < 2$, $f_Z(z) = \int_0^z \frac{1}{8} z dx = \frac{1}{8} z^2$;

当 $2 \leq z < 4$, $f_Z(z) = \int_{z-2}^2 \frac{1}{8} z dx = \frac{1}{8} z(4-z)$ 。

三、(10分) 将 2 个球随机地放入 3 个盒子, 设 X 为第一个盒子内放入的球数, Y 为有球的盒子的个数, (1) 求 (X, Y) 的联合分布列; (2) 求 $E[X]$, $\text{Var}(Y)$ 。

解: (1) 随机变量 X 可能的取值为 0, 1, 2; 随机变量 Y 可能的取值为 1, 2。

$$\begin{aligned} P(X=0, Y=1) &= \frac{2}{9}, P(X=0, Y=2) = \frac{2}{9}, P(X=1, Y=1) = 0, \\ P(X=1, Y=2) &= \frac{4}{9}, P(X=2, Y=1) = \frac{1}{9}, P(X=2, Y=2) = 0 \end{aligned}$$

(2) X 的分布列 $P(X=0) = \frac{4}{9}, P(X=1) = \frac{4}{9}, P(X=2) = \frac{1}{9}, E[X] = 2/3,$

Y 的分布列为 $P(Y=1) = \frac{1}{3}, P(Y=2) = \frac{2}{3},$

$$\text{Var}(Y) = E[Y^2] - (E[Y])^2 = \frac{1}{3} + \frac{8}{3} - \left(\frac{1}{3} + \frac{4}{3}\right)^2 = \frac{2}{9}.$$

四、(10 分) 二维随机变量 (X, Y) 的联合密度函数为 $f(x, y) = \begin{cases} e^{-y}, 0 \leq x \leq y; \\ 0, \text{其他} \end{cases}$.

(1) 求 $P(X + Y \leq 1)$; (2) X, Y 是否独立?

解: (1)

$$P(X + Y \leq 1) = \int_0^{\frac{1}{2}} dx \int_x^{1-x} e^{-y} dy \quad (2\text{分}) = \int_0^{\frac{1}{2}} (e^{-x} - e^{x-1}) dx = (1 - e^{-\frac{1}{2}})^2$$

$$(2) \quad f_X(x) = \begin{cases} \int_x^{\infty} e^{-y} dy = e^{-x}, x > 0; \\ 0, \text{其他} \end{cases}, \quad f_Y(y) = \begin{cases} \int_0^y e^{-y} dx = ye^{-y}, y > 0; \\ 0, \text{其他} \end{cases}$$

$f_X(x) f_Y(y) \neq f(x, y)$, 因此 X, Y 不独立。

五、(10 分) 某种零件的尺寸标准为 $\sigma = 5.2$, 对一批这类零件检查 9 件得到平均尺寸数据 (毫米): $\bar{x} = 26.56$. 设零件尺寸服从正态分布, 问这批零件的平均尺寸能否认为是 26 毫米? ($\alpha = 0.05$)

解: 原假设 $H_0: \mu_0 = 26$, H_0 成立时, 样本均值 $\bar{X} \sim N(26, \frac{5.2^2}{9})$,

$$H_0 \text{ 的拒绝域 } W = \{ |\frac{\bar{x} - 26}{5.2} \times 3| > u_{\frac{\alpha}{2}} \},$$

已知 $u_{0.025} = 1.96$, 而 $|\frac{0.56}{5.2} \times 3| < 1.96$, 因此接受原假设。

六、(10 分) 一条生产线生产的产品合格率为 0.8, 要使一批产品的合格率在 76% 与 84% 之间的概率不小于 90%, 问至少要生产多少件产品?

解: 设随机变量 $X_i = \begin{cases} 1, \text{第 } i \text{ 个产品合格;} \\ 0, \text{其他} \end{cases}$ 则 X_i 服从伯努利分布 $b(1, 0.8)$.

$E[X_i] = 0.8, \text{Var}(X_i) = 0.16$ ---- 2 分。求 n 使得 $P(0.76 \leq \frac{\sum_{i=1}^n X_i}{n} \leq 0.84) \geq 0.9$

而由中心极限定理, 样本为独立同分布随机变量序列,

$$P(0.76 \leq \frac{\sum_{i=1}^n X_i}{n} \leq 0.84) = P\left(\frac{0.76n - E[\sum_{i=1}^n X_i]}{\sqrt{\text{Var}(\sum_{i=1}^n X_i)}} \leq \frac{\sum_{i=1}^n X_i - E[\sum_{i=1}^n X_i]}{\sqrt{\text{Var}(\sum_{i=1}^n X_i)}} \leq \frac{0.84n - E[\sum_{i=1}^n X_i]}{\sqrt{\text{Var}(\sum_{i=1}^n X_i)}}\right)$$

$$= P\left(\frac{0.76n-0.8n}{0.4\sqrt{n}} \leq \frac{\sum_{i=1}^n X_i - 0.8n}{0.4\sqrt{n}} \leq \frac{0.84n-0.8n}{0.4\sqrt{n}}\right) \\ \approx \Phi(0.1\sqrt{n}) - \Phi(-0.1\sqrt{n}) = 2\Phi(0.1\sqrt{n}) - 1$$

$\Phi(0.1\sqrt{n}) \geq 0.95$, 则 $0.1\sqrt{n} \geq 1.645$, $n \geq 270.6$, 故取 $n=271$.

七、(14 分) 设总体 X 的概率密度函数为 $f(x) = \begin{cases} \frac{1}{k\theta}, & \theta < x < (k+1)\theta; \\ 0, & \text{其他} \end{cases}$ 其中

k 为已知正数, $\theta > 0$ 为一个待估参数。 $X_1, \dots, X_n$ 为来自总体的样本。

(1) 求 θ 的最大似然估计和矩估计; (2) θ 的最大似然估计是否无偏?

解: (1) 极大似然函数 $L(\theta) = \left(\frac{1}{k\theta}\right)^n$, 取对数 $\ln L(\theta) = -n(\ln k + \ln \theta)$,

$\frac{d \ln L(\theta)}{d\theta} = -\frac{n}{\theta} < 0$, $L(\theta)$ 为单调减函数, 由于次序统计量

$\theta \leq X_{(1)} \leq \dots \leq X_{(n)} \leq (k+1)\theta$, $\frac{X_{(n)}}{k+1} \leq \theta \leq X_{(1)}$, 因此 θ 的极大似然估计为

$$\hat{\theta} = \frac{X_{(n)}}{k+1}.$$

$E[X] = \int_{\theta}^{(k+1)\theta} \frac{x}{k\theta} dx = \frac{\theta(k+2)}{2}$, 因此 θ 的矩估计为 $\hat{\theta}_1 = \frac{2\bar{X}}{k+2}$.

(2) 最大次序统计量的分布函数 $F_n(x) = P(X_{(n)} \leq x) = F(x)^n$,

最大次序统计量的密度函数

$$f_n(x) = nF(x)^{n-1} f(x) = \begin{cases} n\left(\frac{x-\theta}{k\theta}\right)^{n-1} \frac{1}{k\theta} = n \frac{(x-\theta)^{n-1}}{(k\theta)^n}, & \theta \leq x \leq (k+1)\theta; \\ 0, & \text{其他} \end{cases}$$

$$E[\hat{\theta}] = \int_{\theta}^{(k+1)\theta} \frac{nx(x-\theta)^{n-1}}{(k+1)(k\theta)^n} dx = \frac{n}{(k+1)(k\theta)^n} \int_{\theta}^{(k+1)\theta} x(x-\theta)^{n-1} dx = \theta - \frac{k\theta}{(k+1)(n+1)}$$

$E[\hat{\theta}] \neq \theta$, 故不是 θ 的无偏估计。

概率统计期末模拟题（一）2019.1

备用数据： $\Phi(\bullet)$:标准正态分布函数， $t_{\alpha}(n)$:自由度为 n 的 t 分布的上 α 分位数.

$\Phi(1) = 0.841$, $\Phi(1.645) = 0.95$, $\Phi(1.96) = 0.975$, $\Phi(2) = 0.977$;

$t_{0.025}(8) = 2.306$, $t_{0.025}(16) = 2.120$, $t_{0.05}(8) = 1.860$, $t_{0.05}(16) = 1.746$.

一、完成下列各题（每小题 6 分，共 30 分）

1. 某学院一、二、三、四年级各有 100 名学生，其中男生分别有 80、70、85、75 名，从该学院随机选一名学生， A 表示选到的是一年级学生， B 表示选到的是男生，则 $\overline{A \cup B}$ 表示什么？其概率是多少？

2. 某小区有 $n(n \geq 4)$ 个业主申请停车位，实际停车位有 $m(m+4 < n)$ 个，管理员采用抽签方法确定停车位的使用权，求：(1) 前 4 个排队者中至少有 1 个人抽中的概率；(2) 第 4 个排队者抽中的概率.

3. 某热线电话在 t 小时内接到的呼叫次数 X_t 服从参数为 $t/2$ 的泊松分布，设在两个不重叠的时间段接到的呼叫次数相互独立.

(1) 求 8 点到 12 点至少接到 1 次呼叫的概率；

(2) 若已知 8 点到 12 点至少接到 1 次呼叫，求在 8 点到 10 点没有接到呼叫的概率.

4. 二维随机变量 (X, Y) 在以 $(1, 0)$, $(1, 1)$, $(0, 1)$ 为顶点的三角形区域服从均匀分布. 求：(1) $P(X+Y < 1.5)$ 的值；(2) (X, Y) 的分布函数值 $F(1.5, 0.5)$.

5. 设总体 $X \sim N(0, \sigma^2)$, $X_1, \dots, X_4$ 是 X 的简单随机样本， $\bar{X}$ 是样本均值.

(1) 求 $P(|\bar{X}| > \sigma)$ ；(2) 若 $\frac{cX_1}{\sqrt{X_2^2 + X_3^2 + X_4^2}} \sim t(3)$ ，求 c 的值.

二. (10 分) 设随机变量 X 与 Y 相互独立， X 服从指数分布， $E(X) = 2$ ， $Y \sim B(1, 0.4)$,

令 $U = \frac{1}{2}e^{-\frac{X}{2}}$ ， $V = X + Y$ ，分别求 U, V 的分布函数.

三. (12 分) 设随机变量 X 的概率密度函数 $f(x) = \begin{cases} 6x(1-x), & 0 < x < 1, \\ 0, & \text{其他.} \end{cases}$ 对 X 独立

重复观察 100 次，观察值为 $X_1, \dots, X_{100}$ ，记 Y 为 $\{X_i < 0.5\} (i=1, \dots, 100)$ 出现的次数，

$Z = X_1 + \dots + X_{100}$. 求：(1) $\sum_{i=1}^{60} X_i$ 与 $\sum_{i=41}^{100} X_i$ 的相关系数；(2) $P(Y > 45)$ 的近似值；

(3) Z 的近似分布(要求写出参数).

四. (12 分) 设 (X, Y) 的联合密度函数为 $f(x, y) = \begin{cases} \frac{3y}{2}, & |x| < y < 1, \\ 0, & \text{其他.} \end{cases}$

- (1) 求 $P(X+Y>1)$ 的值; (2) 求条件密度函数 $f_{x|y}(x|y)$ 和 $P(X>0.2|Y=0.5)$ 的值;
(3) 判断 X 与 Y 是否相关? 说明理由.

五. (10 分) 超市中有两种商品, 它们每月销售量之间存在相关性, 设商品甲月销售件数(单位: 千件) $X \sim N(5.25, 0.64)$, 商品乙月销售件数(单位: 千件) $Y \sim N(2.53, 0.25)$, 且 (X, Y) 服从正态分布, X 与 Y 的相关系数 $\rho = 0$. 求: (1) 一个月中这两种商品销售总量超过 7.5 千件的概率; (2) 一个月中甲销售量超过乙销售量 2 倍的概率.(注: 结果用 $\Phi(\cdot)$ 表示)

六. (16 分) 设总体 X 的概率密度函数 $f(x; \alpha, \beta) = \begin{cases} \alpha x^{\alpha-1} / \beta^\alpha, & 0 < x < \beta, \\ 0, & \text{其他.} \end{cases}$ 未知参数 $\alpha > 1, \beta > 0$, $X_1, \dots, X_n$ 为 X 的简单随机样本. (1) 若 $\alpha = 3$, 求 β 的矩估计量 $\hat{\beta}$, 并判断 $\hat{\beta}$ 是否为 β 的无偏估计量, 说明理由; (2) 若 $\beta = 3$, 求 α 的最大似然估计量 $\hat{\alpha}$, 并判断 $\hat{\alpha}$ 是否为 α 的相合(一致)估计量, 说明理由.

七. (10 分) 某培训班声称经过培训, 学员掷实心球的成绩能够提高 1 米以上. 为检验他们的说法是否符合实际, 随机选出 9 人, 记录他们培训前和培训后的成绩如下 (单位: 米):

培训前成绩 x_i	9.63 7.61 6.28 8.32 5.40 5.82 6.89 8.17 5.78	$\bar{x} = 7.10$	$s_x^2 = 2.0352$
培训后成绩 y_i	10.24 8.75 7.05 8.91 6.51 6.58 8.02 9.08 6.32	$\bar{y} = 7.94$	$s_y^2 = 1.938$
差值 $z_i = y_i - x_i$	0.61 1.14 0.77 0.59 1.11 0.76 1.13 0.91 0.54	$\bar{z} = 0.84$	$s_z^2 = 0.058575$

设差值数据来自正态分布, 在显著水平 0.05 下, 检验 H_0 : 平均成绩提高 1 米及以上, H_1 : 平均成绩提高不到 1 米.

概率统计期末模拟题（一）2019.1 答案

一. 1. 解: $\overline{A \cup B}$ 表示选到二、三、四年级的女生, 其概率为 $7/40$.

2. 解: (1) $1 - C_{n-m}^4 / C_n^4$, (2) m/n .

3. 解: (1) $X_4 \sim P(2), P(X_4 \geq 1) = 1 - P(X_4 = 0) = 1 - e^{-2}$,

$$(2) \frac{P(X_2 = 0)P(Y_2 \geq 1)}{P(X_4 \geq 1)} = \frac{1}{1+e}.$$

4. 解: $P(X+Y < 1.5) = 3/4$, $F(1.5, 0.5) = 1/4$.

5. 解: (1) $\bar{X} \sim N(0, \sigma^2/4), P(|\bar{X}| > \sigma) = 2(1 - \Phi(2)) = 0.046$,

(2) $\frac{X_1}{\sigma} \sim N(0, 1), \frac{X_2^2 + X_3^2 + X_4^2}{\sigma^2} \sim \chi^2(3)$, 且相互独立,

$$\frac{\sqrt{3}X_1}{\sqrt{X_2^2 + X_3^2 + X_4^2}} \sim t(3), \Rightarrow c = \sqrt{3}.$$

二. 解: X 的分布函数 $F(x) = \begin{cases} 1 - e^{-\frac{x}{2}}, & x > 0, \\ 0, & x \leq 0. \end{cases}$ Y 的分布律

$$P(Y=0)=0.6, P(Y=1)=0.4.$$

$$F_U(u) = P(U \leq u) = P\left(\frac{1}{2}e^{-\frac{X}{2}} \leq u\right), \text{ 当 } u < 0 \text{ 时, } F_U(u) = 0, \text{ 当 } u \geq 1/2 \text{ 时, } F_U(u) = 1,$$

$$\text{当 } 0 \leq u < 1/2 \text{ 时, } F_U(u) = P\left(\frac{1}{2}e^{-\frac{X}{2}} \leq u\right) = P(X \geq -2\ln(2u)) = 2u. \text{ 即}$$

$$F_U(u) = \begin{cases} 0, & u < 0, \\ 2u, & 0 \leq u < \frac{1}{2}, \\ 1, & u \geq \frac{1}{2}. \end{cases}$$

根据全概率公式,

$$F_V(v) = P(X+Y \leq v) = P(Y=0)P(X \leq v) + P(Y=1)P(X \leq v-1)$$

$$= \begin{cases} 0, & v < 0, \\ 0.6(1 - e^{-\frac{v}{2}}), & 0 \leq v < 1, \\ 1 - 0.6e^{-\frac{v}{2}} - 0.4e^{-\frac{v-1}{2}}, & v \geq 1. \end{cases}$$

三. 解: 记 $\text{Var}(X) = \sigma^2$,

$$(1) \text{Cov}(\sum_{i=1}^{60} X_i, \sum_{i=41}^{100} X_i) = 20\sigma^2, \text{Var}(\sum_{i=1}^{60} X_i) = \text{Var}(\sum_{i=41}^{100} X_i) = 60\sigma^2,$$

所以, $\sum_{i=1}^{60} X_i$ 与 $\sum_{i=41}^{100} X_i$ 的相关系数为 $\frac{\text{Cov}(\sum_{i=1}^{60} X_i, \sum_{i=41}^{100} X_i)}{\sqrt{\text{Var}(\sum_{i=1}^{60} X_i) \text{Var}(\sum_{i=41}^{100} X_i)}} = \frac{1}{3};$

$$(2) P(X < 0.5) = \int_0^{0.5} 6x(1-x)dx = 0.5, Y \sim B(100, 0.5), \text{ 根据中心极限定理,}$$

$$Y \overset{\text{近似}}{\sim} N(50, 25), P(Y > 45) \approx 1 - \Phi(-1) = \Phi(1) = 0.841;$$

$$(3) E(X) = \int_0^1 6x^2(1-x)dx = 0.5, E(X^2) = \int_0^1 6x^3(1-x)dx = 0.3, \text{Var}(X) = 0.05$$

$$\text{根据中心极限定理 } Z = X_1 + \dots + X_{100} \overset{\text{近似}}{\sim} N(50, 5).$$

$$\text{四. 解: (1) } P(X+Y > 1) = \int_{0.5}^1 dy \int_{1-y}^y \frac{3y}{2} dx = \frac{5}{16},$$

$$(2) Y \text{ 的边际密度函数 } f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx = \begin{cases} \int_{-y}^y \frac{3y}{2} dx = 3y^2, & 0 < y < 1, \\ 0, & \text{其他.} \end{cases}$$

$$\text{当 } 0 < y < 1 \text{ 时, } f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)} = \begin{cases} \frac{1}{2y}, & -y < x < y, \\ 0, & \text{其他.} \end{cases} \text{ 为均匀分布密度函数,}$$

$$\text{所以 } P(X > 0.2 | Y = 0.5) = 0.3;$$

$$(3) \text{ 计算得, } E(X)=0, E(XY)=0, \text{ 所以 } \text{Cov}(X, Y)=0, \text{ 因此 } X \text{ 与 } Y \text{ 不相关.}$$

$$\text{五. 解: } X \sim N(5.25, 0.64), Y \sim N(2.53, 0.25), \rho = 0.$$

$$(1) X+Y \sim N(7.78, 0.89), P(X+Y > 7.5) = \Phi(0.2968);$$

$$(2) X-2Y \sim N(0.19, 1.64), P(X > 2Y) = P(X-2Y > 0) = \Phi(0.1484).$$

六. 解: (1) $f(x; 3, \beta) = \begin{cases} 3x^2/\beta^3, & 0 < x < \beta, \\ 0, & \text{其他.} \end{cases}$

$$\mu_1 = E(X) = \int_0^\beta x \cdot 3x^2/\beta^3 dx = \frac{3\beta}{4}, \quad \hat{\mu}_1 = \bar{X},$$

所以, β 的矩估计量 $\hat{\beta} = \frac{4\bar{X}}{3}$; $E(\hat{\beta}) = \frac{4E(\bar{X})}{3} = \beta$, $\hat{\beta}$ 是 β 的无偏估计量;

(2) $f(x; \alpha, 3) = \begin{cases} \alpha x^{\alpha-1}/3^\alpha, & 0 < x < 3, \\ 0, & \text{其他.} \end{cases}$

似然函数 $L(\alpha) = \prod_{i=1}^n f(x_i; \alpha, 3) = \frac{\alpha^n (x_1 \dots x_n)^{\alpha-1}}{3^{n\alpha}}$

对数似然函数 $l(\alpha) = n \ln \alpha + (\alpha - 1)(\ln x_1 + \dots + \ln x_n) - n\alpha \ln 3$

$$\frac{d}{d\alpha} l(\alpha) = \frac{n}{\alpha} + \ln x_1 + \dots + \ln x_n - n \ln 3 = 0$$

所以, α 的最大似然估计量

$$\hat{\alpha} = \frac{-n}{\ln X_1 + \dots + \ln X_n - n \ln 3} = \frac{-1}{\frac{1}{n}(\ln X_1 + \dots + \ln X_n) - \ln 3};$$

根据辛钦大数定律, 当 $n \rightarrow +\infty$ 时, $\frac{1}{n}(\ln X_1 + \dots + \ln X_n) \xrightarrow{P} E(\ln X)$,

$$E(\ln X) = \int_0^3 \ln x \cdot \alpha x^{\alpha-1}/3^\alpha dx \stackrel{t=\ln x}{=} \int_{-\infty}^{\ln 3} \alpha t e^{\alpha t}/3^\alpha dt = \ln 3 - \frac{1}{\alpha},$$

所以 $\hat{\alpha} = \frac{-1}{\frac{1}{n}(\ln X_1 + \dots + \ln X_n) - \ln 3} \xrightarrow{P} \alpha$, $\hat{\alpha}$ 是 α 的相合(一致)估计量.

七. 解: 根据题意, 差值数据来自正态总体 $Z \sim N(\mu, \sigma^2)$,

要检验 $H_0: \mu \geq 1, H_1: \mu < 1$,

该检验的拒绝域为 $t = \frac{\bar{Z} - 1}{S_z/\sqrt{9}} \leq -t_{0.05}(8)$,

计算得 $t = -1.98 < -t_{0.05}(8) = -1.86$,

落在拒绝域内, 所以拒绝原假设, 即认为培训后平均成绩提高不到 1 米.