浙江大学 2024 – 2025 学年 春夏 学期

《概率论与数理统计》期末考试试卷解答

By Texas

一、填空题

1.

答案：$\frac{5}{12}$

这题比较特殊，画图分析过后可以简化计算过程，得到如下

$$P(A\overline{B}\,\overline{C}) = P(A) - P(AC) = \frac{1}{4} - \frac{1}{12} = \frac{1}{6}$$
$$P(\overline{A}\,B\,\overline{C}) = P(B) - P(BC) = \frac{1}{4} - \frac{1}{12} = \frac{1}{6}$$
$$P(\overline{A}\,\overline{B}\,C) = P(C) - P(AC) - P(BC) = \frac{1}{4} - \frac{1}{12} - \frac{1}{12} = \frac{1}{12}$$

设题目所求事件为 Q，即“恰有一个事件发生”，则

$$P(Q) = P(A\overline{B}\,\overline{C}) + P(\overline{A}\,B\,\overline{C}) + P(\overline{A}\,\overline{B}\,C) = \frac{1}{6} + \frac{1}{6} + \frac{1}{12} = \frac{5}{12}$$

2.

答案：$\frac{25}{64}$；$\frac{1}{64}$

第二代为 0 个时需要对第一代分三类进行讨论，运用全概率公式即可；而第二代有 4 个，则第一代必有 2 个，故无需进行讨论

$$\begin{aligned} P(X_2=0) =& P(X_1=0)P(X_2=0|X_1=0) + P(X_1=1)P(X_2=0|X_1=1) \\ &+ P(X_1=2)P(X_2=0|X_1=2) \\ =& \frac{1}{4}\cdot 1 + \frac{1}{2}\cdot\frac{1}{4} + \frac{1}{4}\cdot\left(\frac{1}{4}\right)^2 \\ =& \frac{25}{64}\text{；} \\ P(X_2=4) =& P(X_1=2)P(X_2=2|X_1=2) = \frac{1}{4}\cdot\left(\frac{1}{4}\right)^2 = \frac{1}{64} \end{aligned}$$

3.

答案：$F(1,1)$；5

$$E(2X-3Y-10) = 2E(X) - 3E(Y) - 10 = 2\cdot 5 - 3\cdot 0 - 10 = 10 - 10 = 0$$
$$Var(2X-3Y-10) = 4Var(X) + 9Var(Y) + 2Cov(2X, -3Y) = 4\cdot 9 + 9\cdot 4 = 72$$

$$\therefore 2X - 3Y - 10 \sim N(0, 72)$$

$$E(2X+3Y-10) = 2E(X) + 3E(Y) - 10 = 2\cdot 5 + 3\cdot 0 - 10 = 10 - 10 = 0$$
$$Var(2X+3Y-10) = 4Var(X) + 9Var(Y) + 2Cov(2X,3Y) = 4\cdot 9 + 9\cdot 4 = 72$$

$$\therefore 2X + 3Y - 10 \sim N(0, 72)$$
$$\therefore \left(\frac{2X-3Y-10}{2X+3Y-10}\right)^2 = \left(\frac{\frac{2X-3Y-10}{\sqrt{72}}}{\frac{2X+3Y-10}{\sqrt{72}}}\right)^2 \sim F(1,1)$$

因为 n 足够大，所以 $1\sim n$ 与 $n+1\sim 2n$ 其实并没有区别，再由大数定律得

$$\frac{1}{2n}\sum_{i=n+1}^{2n} 2X_i = \frac{1}{2n}\sum_{i=1}^{n} 2X_i = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} E(X) = 5$$

4.

答案：0.2

已知 $t(n) = \dfrac{Z}{\sqrt{\chi^2/n}}$, 所以 $t(n)^2 = \dfrac{Z^2}{\chi^2/n} \sim F(1,n)$

$$\therefore P(Y > c^2) = P(X^2 > c^2) = P(X < -c \text{ or } X > c) = 2P(0 < X < c) = 2\times 0.1 = 0.2$$

5.

答案：是 ; $\dfrac{1}{10}$

由题意我们可以先计算出

$$E(\bar{X}^2) = Var(\bar{X}) + E(\bar{X})^2 = \frac{\sigma^2}{n} + \mu^2$$

又因为 S^2 是 σ^2 的无偏估计，故

$$E(T) = E\left(\bar{X}^2 - \frac{S^2}{n}\right) = E(\bar{X}^2) - \frac{1}{n}E(S^2) = \frac{\sigma^2}{n} + \mu^2 - \frac{1}{n}\cdot\sigma^2 = \mu^2$$

所以 T 是 μ^2 的无偏估计

$$\because \bar{X} \sim N(0, \frac{1}{5})$$
$$\therefore 5\bar{X}^2 \sim \chi^2(1)$$
$$\therefore Var(\bar{X}^2) = Var\left(\frac{1}{5}\chi^2(1)\right) = \frac{1}{5^2}Var\left(\chi^2(1)\right) = \frac{2}{25}$$
$$\because \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$
$$\therefore Var(S^2) = Var\left(\frac{\sigma^2}{n-1}\chi^2\right) = \left(\frac{\sigma^2}{n-1}\right)^2 Var\left(\chi^2(n-1)\right) = \frac{2\sigma^4}{n-1} = \frac{1}{2}$$
$$\therefore Var(T) = Var\left(\bar{X}^2 - \frac{S^2}{n}\right) = Var(\bar{X}^2) + Var\left(\frac{S^2}{n}\right) = \frac{2}{25} + \frac{1}{25}\cdot\frac{1}{2} = \frac{1}{10}$$

6.

答案：3.635 ; 1.5

构造枢轴量

$$\begin{aligned}\frac{\bar{X}-\mu}{S/\sqrt{n}} &= \frac{\bar{X}-\mu}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2}/\sqrt{n}} \\ &= \frac{\bar{X}-\mu}{\sqrt{\frac{1}{n-1}\left(\sum_{i=1}^{n}X_i^2-n\bar{X}^2\right)}/\sqrt{n}} \\ &= \frac{5-\mu}{\sqrt{\frac{1}{9}\times(300-10\times 25)}/\sqrt{10}} \\ &= \frac{5-\mu}{\sqrt{\frac{1}{9}\times 50}/\sqrt{10}} \\ &= \frac{5-\mu}{\sqrt{5/9}}\end{aligned}$$

由题意

$$P\left\{\frac{5-\mu}{\sqrt{5/9}} \le t(9)_{0.05}\right\} = 1-5\%$$

解得

$$\mu \ge 5 - t(9)_{0.05}\cdot\sqrt{\frac{5}{9}} = 5-1.833\cdot\sqrt{\frac{5}{9}} \approx 3.635$$

所以 μ 的置信下限 $\hat{\mu}_L = 3.635$

由枢轴量的构造可以知道

$$\begin{aligned}t(9)_{\alpha/2} &= \frac{\bar{X}-\hat{\mu}_L}{S/\sqrt{n}} \\ &= \frac{5-\left(5-\frac{\sqrt{5}}{2}\right)}{5/9} \\ &= \frac{3}{2}\end{aligned}$$

所以 $t_{\alpha/2}(9) = 1.5$

第二空本来问的好像是具体的 α 是多少，但显然我们就算翻书查表也找不到结果，所以就把题目改到只要算出分位数对应的值就行了

7.

答案： $\frac{1}{6}$ ； $\frac{n-1}{(n+1)^2}$

因为 X_i 彼此之间是独立的，所以对于 X_i 来说，只有 $X_{i-1}X_i, X_iX_{i+1}$ 与 X_i 之间不是独立的

且

$$\begin{aligned}E(X_i) &= 0\cdot P(X_i=0)+1\cdot P(X_i=1)\\&= 0\cdot\frac{1}{i}+1\cdot 1-\frac{1}{i}=1-\frac{1}{i}\end{aligned}$$

$$\begin{aligned}Var(X_i) &= E(X_i^2)-[E(X_i)]^2\\&= 0^2\cdot P(X_i=0)+1^2\cdot P(X_i=1)-\left(1-\frac{1}{i}\right)^2\\&= 0^2\cdot\frac{1}{i}+1^2\cdot\left(1-\frac{1}{i}\right)-\left(1-\frac{1}{i}\right)^2\\&= 1-\frac{1}{i}-\left(1-\frac{1}{i}\right)^2\\&= \frac{1}{i}-\frac{1}{i^2}\end{aligned}$$

$$\begin{aligned}Cov(X_i,X_{i\pm1}X_i) &= E(X_{i\pm1}X_i^2)-E(X_{i\pm1})\,[E(X_i)]^2\\&= E(X_{\pm i})\left(E(X_i^2)-[E(X_i)]^2\right)\\&= E(X_{\pm i})Var(X_i)\end{aligned}$$

所以由协方差的线性性质可以得知

$$\begin{aligned}Cov(X_2,S) &= Cov(X_2,X_1X_2)+Cov(X_2,X_2X_3)\\&= E(X_1)Var(X_2)+E(X_3)Var(X_2)\\&= \left(1-\frac{1}{1}\right)\left(\frac{1}{2}-\frac{1}{4}\right)+\left(1-\frac{1}{3}\right)\left(\frac{1}{2}-\frac{1}{4}\right)\\&= 0\cdot\frac{1}{4}+\frac{2}{3}\cdot\frac{1}{4}\\&= \frac{1}{6}\end{aligned}$$

$$\begin{aligned}Cov(X_{n+1},S) &= Cov(X_{n+1},X_nX_{n+1})\\&= E(X_n)Var(X_{n+1})\\&= \left(1-\frac{1}{n}\right)\left(\frac{1}{n+1}-\frac{1}{(n+1)^2}\right)\\&= \left(\frac{n-1}{n}\right)\cdot\left(\frac{n}{(n+1)^2}\right)\\&= \frac{n-1}{(n+1)^2}\end{aligned}$$

二、计算题

1.

(1)

解：X 的分布律为

$$P(X=k)=\frac{1}{3}\left(\frac{2}{3}\right)^{k-1},k=1,2,3,\dots$$

(2)

解：

$$\begin{aligned}P(Y=1)&=\frac{1}{3}\\P(Y=2)&=\frac{2}{3}\cdot\frac{1}{2}=\frac{1}{3}\\P(Y=3)&=\frac{2}{3}\cdot\frac{1}{2}\cdot 1=\frac{1}{3}\end{aligned}$$

Y 的分布律为

Y	1	2	3
P	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

2.

(1)

解：

$$\begin{aligned}P(X<2Y)&=\iint\limits_{x<2y}f(x,y)\,dx\,dy\\&=\int_0^1\int_{x/2}^{x}8xy\,dy\,dx\\&=\int_0^1 8x\left(\frac{x^2}{2}-\frac{x^2}{8}\right)dx\\&=\int_0^1 3x^3\,dx\\&=\frac{3}{4}\end{aligned}$$

(2)

解：

$$
\begin{aligned}
f_X(x) &= \int_{-\infty}^{\infty} f(x,y)\,dy \\
&= \int_0^1 8xy\,dy \\
&= 8x\left[\frac{y^2}{2}\right]_0^x \\
&= 8x\cdot\frac{x^2}{2} \\
&= 4x^3
\end{aligned}
$$

$$
\begin{aligned}
f_Y(y) &= \int_{-\infty}^{\infty} f(x,y)\,dx \\
&= \int_y^1 8xy\,dx \\
&= 8y\left[\frac{x^2}{2}\right]_y^1 \\
&= 8y\cdot\left(\frac{1}{2}-\frac{y^2}{2}\right) \\
&= 4y(1-y^2)
\end{aligned}
$$

所以

$$
f_X(x) = \begin{cases} 4x^3\,, & 0 < x < 1, \\ 0\,, & \text{else.} \end{cases}
$$

$$
f_Y(y) = \begin{cases} 4y(1-y^2)\,, & 0 < y < 1, \\ 0\,, & \text{else.} \end{cases}
$$

$$
\because f(x,y) \neq f_X(x) f_Y(y)
$$

所以 X 和 Y 不独立

(3)

解：

$$\begin{aligned}
E(X) &= \int_0^1 x f_X(x)\,dx \\
&= \int_0^1 x \cdot 4x^3\,dx \\
&= 4\int_0^1 x^4\,dx \\
&= \frac{4}{5}
\end{aligned}$$

$$\begin{aligned}
E(Y) &= \int_0^1 y f_Y(y)\,dy \\
&= \int_0^1 y \cdot 4y(1-y^2)\,dy \\
&= 4\left(\int_0^1 y^2\,dy - \int_0^1 y^4\,dy\right) \\
&= 4\left(\frac{1}{3}-\frac{1}{5}\right) \\
&= \frac{8}{15}
\end{aligned}$$

$$\begin{aligned}
E(XY) &= \int_0^1\int_0^1 xyf(x,y)\,dx\,dy \\
&= \int_0^1 8y^2\int_0^y x^2\,dx\,dy \\
&= \frac{8}{3}\int_0^1 y^5\,dy \\
&= \frac{4}{9}
\end{aligned}$$

$$\begin{aligned}
Cov(X,Y) &= E(XY) - E(X)E(Y) \\
&= \frac{4}{9} - \frac{4}{5}\cdot\frac{8}{15} \\
&= \frac{4}{225}
\end{aligned}$$

$$\because Cov(X,Y) \neq 0$$

所以 X 和 Y 相关

3.

解：

设实际乘机人数为 X，售卖票数为 n，则 $X \sim B(n,p)$，其中 $p = 90\%$

因为 n 较大，所以可以用正态分布 $N(np, np(1-p))$ 近似 X 的分布

先对 X 进行标准化处理

$$\frac{X-np}{\sqrt{np(1-p)}}$$

由题意得：$\alpha=0.05$, 也就是要 X 超过 200 的概率小于等于 0.05 , 所以当 $X=200$ 时 ,

$$\Phi\left(\frac{200-np}{\sqrt{np(1-p)}}\right)\geq 1-\alpha$$

将 $z_\alpha=z_{0.05}=1.645$ 代入，得到

$$\frac{200-np}{\sqrt{np(1-p)}}\geq 1.645$$

解得

$$n\leq 214.2$$

因为 n 必须是整数，所以 n 的最大值为 214, 即最多可以售卖 214 张票

4.

(1)

解：

$$L(\theta)=\prod_{i=1}^{n}f(x_i;\theta)=\prod_{i=1}^{n}\frac{1}{\theta}=\frac{1}{\theta^n}$$

$\because$ 要使 $L(\theta)$ 尽量大，且 $L(\theta)$ 在 $\theta>0$ 上单调递减

$\therefore\theta$ 的取值要尽量小

$\because\theta\geq x_{(n)}$

$\therefore\hat{\theta}=x_{(n)}$ 是 θ 的最大似然估计

由最大似然估计的不变性

$$\widehat{Var(X)}=\frac{\hat{\theta}^2}{12}=\frac{x_{(n)}^2}{12}$$

(2)

解：

已知 $F_X(t)=P(X\leq t)=\frac{t}{\theta}$，所以 $X_{(n)}$ 的累积分布函数

$$\begin{aligned}F_{X_{(n)}}(t)&=P(X_{(n)}\leq t)\\&=P(\max\{X_1,X_2,\ldots,X_n\}\leq t)\\&=P(X_1\leq t,X_2\leq t,\ldots,X_n\leq t)\\&=[P(X\leq t)]^n\\&=\left(\frac{t}{\theta}\right)^n,\quad \text{其中}0\leq t\leq\theta\end{aligned}$$

所以 $X_{(n)}$ 的概率密度函数

$$\begin{aligned}f_{X_{(n)}}(t)&=\frac{d}{dt}F_{X_{(n)}}(t)\\&=\frac{nt^{n-1}}{\theta^n}\qquad\text{其中}0\leq t\leq\theta\end{aligned}$$

$X_{(n)}$ 的数学期望：

$$\begin{aligned}
E\left(X_{(n)}\right) &= \int_{-\infty}^{\infty} t f_{X_{(n)}}(t)\,dt \\
&= \int_0^{\theta} t\left(\frac{nt^{n-1}}{\theta^n}\right) dt \\
&= \frac{n}{\theta^n}\int_0^{\theta} t^n\,dt \\
&= \frac{n}{\theta^n}\cdot\frac{\theta^{n+1}}{n+1} \\
&= \frac{n\theta}{n+1}
\end{aligned}$$

又因为

$$E\left(T_c\right) = E\left(cX_{(n)}\right) = cE\left(X_{(n)}\right) = c\cdot\frac{n\theta}{n+1} = \theta$$

解得

$$c = \frac{n+1}{n}$$

(3)

解：

我们先求 $X_{(n)}^2$ 的均值

$$\begin{aligned}
E(X_{(n)}^2) &= \int_{-\infty}^{\infty} t^2 f_{X_{(n)}}(t)\,dt \\
&= \int_0^{\theta} t^2\left(\frac{nt^{n-1}}{\theta^n}\right) dt \\
&= \frac{n}{\theta^n}\int_0^{\theta} t^{n+1}\,dt \\
&= \frac{n}{\theta^n}\cdot\frac{\theta^{n+2}}{n+2} \\
&= \frac{n\theta^2}{n+2}
\end{aligned}$$

由均方误差的定义得

$$\begin{aligned}
Mse\left(T_c\right) &= E\left[\left(T_c-\theta\right)^2\right] \\
&= E\left[\left(cX_{(n)}-\theta\right)^2\right] \\
&= E\left[c^2X_{(n)}^2 - 2c\theta X_{(n)} + \theta^2\right] \\
&= c^2E(X_{(n)}^2) - 2c\theta E(X_{(n)}) + \theta^2
\end{aligned}$$

可以看出，$Mse\left(T_c\right)$ 是一个关于 c 的二次函数。所以当 $Mse\left(T_c\right)$ 取最小值时，有

$$\begin{aligned}
c &= -\frac{b}{2a} \\
&= -\frac{-2\theta E\left(X_{(n)}\right)}{2E\left(X_{(n)}^2\right)} \\
&= \frac{\theta E\left(X_{(n)}\right)}{E\left(X_{(n)}^2\right)} \\
&= \frac{n+2}{n+1}
\end{aligned}$$

5.

(1)

解：
确定统计量

$$F = \frac{S_1^2/\sigma^2}{S_2^2/\sigma^2} = \frac{S_1^2}{S_2^2}$$

由题意可知为双侧检验，所以拒绝域

$$F > F_{\alpha/2}(7,9) \text{ 或} F < F_{1-\alpha/2}(7,9)$$

查表可得

$$F_{0.025}(7,9) = 4.20$$

又因为

$$F_{0.975}(7,9) = \frac{1}{F_{0.025}(9,7)} \approx \frac{1}{4.20} \approx 0.238$$

所以拒绝域为

$$F > 4.20 \text{ 或} F < 0.238$$

计算检验统计量的值

$$\begin{aligned} F = \frac{s_1^2}{s_2^2} &= \frac{0.000212}{0.000093} \\ &= \frac{212}{93} \\ &\approx 2.2796 \end{aligned}$$

未落入拒绝域内，所以可以接受原假设 H_0，即认为 $\sigma_1^2 = \sigma_2^2$

(2)

解：
由题目的已知条件，我们可以这样构造枢轴量

$$\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{S_w\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2)$$

其中

$$\begin{aligned} S_w^2 &= \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \\ &= \frac{(8-1)(0.000212) + (10-1)(0.000093)}{8 + 10 - 2} \\ &= 0.0001450625 \end{aligned}$$

由置信水平为 95%，则 $\alpha = 0.05$ 。由于双侧区间，所以需要 $\alpha/2 = 0.025$
因为自由度 $n_1 + n_2 - 2 = 8 + 10 - 2 = 16$。查找 t- 分布表可得

$$t_{0.025}(16) = 2.120$$

置信区间的公式为

$$(\bar{x}-\bar{y}) \pm t_{\alpha/2}(16) \cdot S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

代入题目数据得置信区间为 $(0.010, 0.034)$

由于 $\mu_1-\mu_2=0$ 未落在置信区间内，我们可以认为 $\mu_1-\mu_2>0$，即 $\mu_1>\mu_2$。因此，在 95% 的置信水平下，两位作家使用中短单词的平均比例存在显著差异，作家甲的平均比例显著高于作家乙。