厦门大学《概率统计I》课程试卷

________学院______系 ______年级______专业

主考教师：______ 试卷类型：（A卷）

有可能使用的分位数：$t_{0.05}(8) = 1.860$, $t_{0.025}(8) = 2.306$, $t_{0.025}(9) = 2.262$, $t_{0.025}(13) = 2.1604$, $t_{0.025}(25) = 2.060$, $t_{0.025}(26) = 2.056$, $F_{0.05}(2,13) = 3.81$, $\chi^2_{0.05}(10) = 18.307$.

1.

分数	阅卷人

(10分) 假设$X_1, X_2, \cdots, X_n$是两两不相关的随机变量，有相同的数学期望μ和方差σ^2，

(i) $S_n = X_1 + X_2 + \cdots + X_n$，计算$S_n/n$的数学期望和方差；

(ii) $T_n = X_1 - X_2 + \cdots + (-1)^{n-1}X_n$，计算$T_n$的数学期望和方差；

(iii) 计算S_n与T_n的协方差。

(i)

$$ES_n = \sum_{i=1}^n EX_i = n\mu, \qquad E\left(\frac{S_n}{n}\right) = \mu, \tag{2分}$$

$$DS_n = \sum_{i=1}^n DX_i = n\sigma^2, \qquad D\left(\frac{S_n}{n}\right) = \frac{\sigma^2}{n}. \tag{2分}$$

(ii)

$$ET_n = \sum_{i=1}^n (-1)^{n-1} EX_i = \sum_{i=1}^n (-1)^{n-1}\mu = \begin{cases} 0, & \text{n为偶数} \\ \mu, & \text{n为奇数} \end{cases} \tag{2分}$$

$$DT_n = \sum_{i=1}^n DX_i = n\sigma^2. \tag{2分}$$

(iii)

$$\begin{aligned} Cov(S_n, T_n) &= \sum_{i=1}^n Cov(S_n, (-1)^{i-1}X_i) \\ &= \sum_{i=1}^n (-1)^{i-1} Cov(S_n, X_i) \\ &= \sum_{i=1}^n (-1)^{i-1}\sigma^2 \\ &= \begin{cases} 0, & \text{n为偶数} \\ \sigma^2, & \text{n为奇数} \end{cases} \end{aligned}$$

(2分)

2.

分数	阅卷人

(8分) 一位职工每天乘坐公交车上班，如果每天用于等车的时间服从均值为5 min的指数分布，估算他在324个工作日中用于上班的等车时间之和大于24h的概率。（结果用正态分布的分布函数表示）

假设每天的等车时间为X_i，根据题意，$X_i \sim Exp(5)$，$i = 1, 2, \cdots, n$, $n = 324$。
所以，

$$EX_i = 5, \quad DX_i = 25. \tag{2分}$$

根据中心极限定理，

$$\frac{1}{n}\sum_{i=1}^{n} X_i \sim N(5, 25/n), \quad n = 324. \tag{2分}$$

因此，

$$\begin{aligned} P\left\{\sum_{i=1}^{n} X_i \geq 24*60\right\} &= P\left\{\frac{\frac{1}{324}\sum_{i=1}^{n} X_i - 5}{5/\sqrt{n}} \geq \frac{\frac{1}{324}*24*60-5}{5/\sqrt{324}}\right\} \\ &\approx 1-\Phi\left(\frac{24*60-5*324}{5\sqrt{324}}\right) \\ &= 1-\Phi(-2) = \Phi(2). \end{aligned}$$

(4分)

3.

分数	阅卷人

(10分) 假设随机变量X服从几何分布，

$$P\{X=k\}=(1-p)^{k-1}p, \qquad k=1,2,\cdots$$

$X_1,X_2,\cdots,X_n$是来自总体X的一组样本。计算p的矩估计和极大似然估计。

(i) 矩估计：

$$EX=\sum_{k=1}^{\infty}k(1-p)^{k-1}p \qquad (2分)$$

$$=-\left(\sum_{k=1}^{\infty}(1-p)^{k}\right)'p$$

$$=-\left(\frac{1}{1-(1-p)}\right)'p=\frac{1}{p}, \qquad (1分)$$

因此，

$$\hat{p}=\bar{X}^{-1}. \qquad (2分)$$

(ii) 最大似然估计：

$$L(p)=\prod_{i=1}^{n}P\{X=x_i\}=\prod_{i=1}^{n}(1-p)^{X_i-1}p=p^n(1-p)^{\sum_{i=1}^{n}X_i-n}, \qquad (2分)$$

$$\ln L(p)=n\ln p+\left(\sum_{i=1}^{n}X_i-n\right)\ln(1-p),$$

求导，

$$\frac{d\ln L(p)}{dp}=\frac{n}{p}+\frac{n-\sum_{i=1}^{n}X_i}{1-p}=0, \qquad (1分)$$

因此，

$$\hat{p}=\bar{X}^{-1}. \qquad (2分)$$

4.

分数	阅卷人

(15分) 抽样调查了5mm玻璃样本量为$n=9$的样本，得到数据（单位：mm）：

4.8　4.1　4.4　4.4　4.2　4.5　4.1　4.9　4.2

在显著性水平0.05之下，

(i) 能否认为5mm玻璃厚度μ达到标准？

(ii) 能够认为$\mu \geq 4.8mm$？

(iii) 在置信水平0.95下，计算玻璃平均厚度的单侧置信上、下限。

首先计算样本均值和样本方差：$\bar{X}=4.4, \qquad S^2=0.085.$

(i) 考虑假设检验问题：$H_0:\mu=5mm, \qquad H_1:\mu\neq 5mm.$ 检验统计量

$$T=\frac{\bar{X}-5}{S/\sqrt{9}}=\frac{4.4-5}{\sqrt{0.085}/3}=-6.1739. \tag{2分}$$

拒绝域

$$\{|T|\geq t_{0.025}(8)\}. \tag{2分}$$

由于$t_{0.025}(8)=2.306$，拒绝原假设，认为玻璃厚度不达标。

(1分)

(ii) 考虑假设检验问题：$H_0:\mu\geq 4.8mm, \qquad H_1:\mu<4.8mm.$ 检验统计量

$$T=\frac{\bar{X}-4.8}{S/\sqrt{9}}=\frac{4.4-4.8}{\sqrt{0.085}/3}=-4.1160. \tag{2分}$$

拒绝域

$$\{T\leq -t_{0.05}(8)\} \tag{2分}$$

由于$t_{0.05}(8)=1.860$，拒绝原假设，认为玻璃厚度$\mu<4.8mm$。

(1分)

(iii) 枢轴量为

$$\frac{\bar{X}-\mu}{S/\sqrt{n}}, \tag{1分}$$

因此，单侧置信上限为

$$\bar{X}+t_{0.05}(8)\frac{S}{\sqrt{9}}=4.5808, \tag{2分}$$

单侧置信下限为

$$\bar{X}-t_{0.05}(8)\frac{S}{\sqrt{9}}=4.2192. \tag{2分}$$

5.

分数	阅卷人

(12分) 某高校为分析不同专业的人才培养情况，对毕业生进行问卷调查，调查的满意度是他们对自己毕业后的工作、收入等指标的平均，最高为6分，最低为1分。以下是调查结果：

专业	满意度					
A	4.5	4.2	4.6	4.1	4.1	4.3
B	4.5	3.9	4.1	4.7	3.8	
C	4.4	4.3	5.2	4.9	5.2	

假设不同专业学生打分服从正态分布，并且其波动性不存在显著差异，在显著性水平0.05之下，

(i) 能否认为A、B、C专业满意度存在显著差异？

(ii) 能否认为A、B专业满意度存在显著差异？

记A、B、C三个专业学生评分的样本均值为$\bar{X}_1, \bar{X}_2, \bar{X}_3$，样本方差为$S_1^2, S_2^2, S_3^2$，所有样本的平均值为$\bar{X}$，可以计算出

$$\bar{X}_1 = 4.3, \quad \bar{X}_2 = 4.2, \quad \bar{X}_3 = 4.8, \quad \bar{X} = 4.425,$$

$$S_1^2 = 0.22/5, \quad S_2^2 = 0.6/4, \quad S_3^2 = 0.74/4.$$

(i)考虑假设检验问题：$H_0: \mu_A = \mu_B = \mu_C, \quad H_1: \mu_A, \mu_B, \mu_C$不全相等，

(2分)

根据平方和分解式，$S_T = S_A + S_E$,

$$S_E = \sum_{j=1}^{3} \sum_{i=1}^{n_j} (X_{ij} - \bar{X}_j)^2 = 5 * S_1^2 + 4 * S_2^2 + 4 * S_3^2 = 1.56,$$

$$S_A = \sum_{j=1}^{3} \sum_{i=1}^{n_j} (\bar{X}_j - \bar{X})^2 = 6 * 0.125^2 + 5 * 0.225^2 + 5 * 0.375^2 = 1.05,$$

检验统计量

$$F = \frac{S_A/2}{S_E/13} = 4.375. \qquad (3分)$$

拒绝域$\{F \geq F_{0.05}(2, 13)\}$。由于$F_{0.05}(2, 13) = 3.81$，因此，拒绝原假设，认为A、B、C专业满意度存在显著差异。

(1分)

(ii) 考虑假设检验问题：$H_0: \mu_A - \mu_B = 0, \quad H_1: \mu_A - \mu_B \neq 0$，

(2分)

解法一：

$$S_w = \sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}} = 0.3018,$$

检验统计量

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - 0}{\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} S_w} = 0.5472, \qquad (3分)$$

拒绝域$\{|T| \geq t_{0.025}(9)\}$。由于$t_{0.025}(9) = 2.262$，因此，不能拒绝原假设。

(1分)

解法二：

$$\hat{\sigma} = \sqrt{\bar{S}_E} = \sqrt{\frac{S_E}{16-3}} = 0.3464,$$

检验统计量

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - 0}{\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \cancel{\bar{S}_E}\hat{\sigma}} = 0.4767, \qquad (3分)$$

拒绝域$\{|T| \geq t_{0.025}(13)\}$。由于$t_{0.025}(13) = 2.1604$，因此，不能拒绝原假设。

(1分)

6.

分数	阅卷人

(10分) 掷一枚骰子120次，在显著性水平α下给出判断骰子是否均匀的规则。

假设$X_1, X_2, \cdots, X_n$为120次试验的结果，考虑假设检验问题

$$H_0: P\{X=k\}=\frac{1}{6}, k=1,2,\cdots,6. \qquad H_1: \exists k, P\{X=k\}\neq\frac{1}{6} \tag{2分}$$

记

$$f_k=\sum_{i=1}^{120} I_{\{X_i=k\}},\ k=1,2,\cdots,6.$$

检验统计量

$$\chi^2=\sum_{i=1}^{6}\frac{(f_k-np_k)^2}{np_k}=\sum_{i=1}^{6}\frac{(f_k-20)^2}{20}=\sum_{k=1}^{6}\frac{f_k^2}{20}-120. \tag{4分}$$

拒绝域

$$\{\chi^2\geq\chi^2_\alpha(5)\}=\left\{\sum_{i=1}^{6}(f_k-20)^2\geq 20*\chi^2_\alpha(5)\right\}. \tag{2分}$$

给定显著性水平α，统计各个点数出现的频率值f_k，判断是否落在拒绝域。
如果落入拒绝域，则认为骰子不是均匀的；如果未落入拒绝域，则认为骰子是均匀的。

(2分)

7.

分数	阅卷人

(10分)

以x和Y分别表示人的身高和臂长，测量了27名男生的身高x_i和臂长Y_i。经过计算，

$$\bar{x}=174.7037,\quad \bar{Y}=172.4815,\quad S_{xx}=1165.6424,\quad S_{xY}=1083.8516,\quad S_{YY}=1774.7338.$$

(i) 计算Y关于x的回归方程$\hat{y}=\hat{a}+\hat{b}x$

(ii) b的置信水平为0.95的置信区间。

(i)

$$\hat{b}=\frac{S_{xY}}{S_{xx}}=0.9298, \qquad (2分)$$

$$\hat{a}=\bar{Y}-\hat{b}\bar{x}=10.0420. \qquad (2分)$$

(ii)由于

$$Q_e=S_{YY}-\hat{b}^2S_{xx},\qquad \hat{\sigma^2}=\frac{Q_e}{n-2}=30.6787, \qquad (2分)$$

$$\hat{b}\sim N\left(b,\frac{\sigma^2}{S_{xx}}\right),\qquad \frac{Q_e}{\sigma^2}\sim\chi^2(27-2)$$

枢轴量

$$\frac{\hat{b}-b}{\hat{\sigma}/\sqrt{S_{xx}}} \qquad (2分)$$

$t_{0.025}(25)=2.060$，因此b的置信区间为

$$\left(\hat{b}\pm t_{0.025}(25)\frac{\hat{\sigma}}{\sqrt{S_{xx}}}\right)=(0.5957,1.2639) \qquad (2分)$$

8.

分数	阅卷人

(10分) 甲有8万元可以投资两个项目。项目A需要投资至少5万，成功概率为0.8，失败概率为0.2，成功后收回本金并获利50%，失败将损失2万元。项目B需要投资至少6万，成功概率为0.6，失败概率为0.4，成功后收回本金并获利70%，失败将损失3万元。假设甲总是将手中的资金全部用于投资，且只能对各项目投资一次。

(i) 先投资A项目，然后再投资B项目，求平均收益；

(ii) 先投资B项目，然后再投资A项目，求平均收益；

(iii) 应该选择哪种决策。

用A、B分别表示投资项目A、B成功，用X表示投资的收益。
如果8万元投资给A项目，投资收益分布为：$P\{X=4\}=0.8,\quad P\{X=-2\}=0.2.$
如果8万元投资给B项目，投资收益分布为：$P\{X=5.6\}=0.6,\quad P\{X=-3\}=0.4.$

(i) 先投资A项目，然后再投资B项目，总收益分布为

$$\begin{aligned}
P(AB) &= P\{X=12*0.7+4\}=P\{X=12.4\}=0.48,\\
P(A\bar{B}) &= P\{X=-3+4\}=P\{X=1\}=0.32,\\
P(\bar{A}B) &= P\{X=6*0.7-2\}=P\{X=2.2\}=0.12,\\
P(\bar{A}\bar{B}) &= P\{X=-2-3\}=P\{X=-5\}=0.08,
\end{aligned}$$

所以，这种方案的期望收益是$EX=12.4*0.48+1*0.32+2.2*0.12-5*0.08=6.136.$

(4分)

(ii) 先投资B项目，然后再投资A项目，总收益分布为

$$\begin{aligned}
P(BA) &= P\{X=13.6*0.5+5.6\}=P\{X=12.4\}=0.48,\\
P(\bar{B}A) &= P\{X=5*0.5-3\}=P\{X=-0.5\}=0.32,\\
P(B\bar{A}) &= P\{X=-2+5.6\}=P\{X=3.6\}=0.12,\\
P(\bar{B}\bar{A}) &= P\{X=-2-3\}=P\{X=-5\}=0.08,
\end{aligned}$$

所以，这种方案的期望收益是$EX=12.4*0.48-0.5*0.32+3.6*0.12-5*0.08=5.824.$

(4分)

(iii) 应该选择第一种方案。

(2分)

9.

分数	阅卷人

(15分) 假设随机变量X的密度函数为

$$f(x) = \frac{1}{2\sigma} exp\left\{-\frac{|x-\mu|}{\sigma}\right\},$$

其中$\sigma > 0$。$X_1, X_2, \cdots, X_n$是来自总体X的一组样本。

(i) 计算μ和σ 的矩估计；

(ii) 证明σ 的矩估计是一个相合估计。

(i) 计算前两阶矩，

$$EX = \int x f(x)\, dx = \int \frac{x}{2\sigma} \exp\left\{-\frac{|x-\mu|}{\sigma}\right\} dx \qquad (2分)$$

$$\begin{aligned} &= \frac{1}{2\sigma}\int t \exp\left\{-\frac{|t|}{\sigma}\right\} dt + \frac{\mu}{2\sigma}\int \exp\left\{-\frac{|t|}{\sigma}\right\} dt \\ &= \frac{\mu}{\sigma}\int_0^\infty e^{-t/\sigma} dt = \mu, \end{aligned}$$

(2分)

$$EX^2 = \int x^2 f(x)\, dx = \int \frac{x^2}{2\sigma} \exp\left\{-\frac{|x-\mu|}{\sigma}\right\} dx \qquad (2分)$$

$$\begin{aligned} &= \frac{1}{\sigma}\int_0^\infty t^2 e^{-t/\sigma} dt + \frac{\mu^2}{\sigma}\int_0^\infty e^{-t/\sigma} dt \\ &= 2\sigma^2 + \mu^2. \end{aligned}$$

(2分)

因此，

$$\hat{\mu} = \bar{X}, \qquad (2分)$$

$$\hat{\sigma} = \sqrt{\frac{1}{2}\left(\frac{1}{n}\sum_{i=1}^n X_i^2 - \bar{X}^2\right)} = \sqrt{\frac{1}{2n}\sum_{i=1}^n (X_i - \bar{X})^2}. \qquad (2分)$$

(ii) 由于$EX_1^2 = 2\sigma^2 + \mu^2, EX_1 = \mu$，根据大数定理，

$$\frac{1}{n}\sum_{i=1}^n X_i^2 \longrightarrow EX_1^2 = 2\sigma^2 + \mu^2, \text{ in pr.} \qquad (1分)$$

$$\bar{X} = \sum_{i=1}^{n} X_i \longrightarrow EX = \mu, \ \text{in pr.} \qquad (1分)$$

所以，

$$\hat{\sigma} = \sqrt{\frac{1}{2}\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2\right)} \longrightarrow \sqrt{\frac{1}{2}(2\sigma^2 + \mu^2 - \mu^2)} = \sigma, \ \text{in pr.} \qquad (1分)$$