



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：(B 卷)

解题过程中可能用到以下数据：

$\Phi(0.6) = 0.7257$, $\Phi(1.65) = 0.9500$, $\Phi(1.96) = 0.9750$, $\Phi(2) = 0.977$, $\Phi(3) = 0.9987$
 $\chi^2_{0.05}(5) = 11.070$, $\chi^2_{0.05}(6) = 12.592$, $\chi^2_{0.025}(24) = 39.364$, $\chi^2_{0.025}(25) = 40.646$, $\chi^2_{0.05}(24) = 36.415$,
 $\chi^2_{0.05}(25) = 37.652$, $\chi^2_{0.975}(24) = 12.401$, $\chi^2_{0.975}(25) = 13.120$, $\chi^2_{0.95}(24) = 13.848$, $\chi^2_{0.95}(25) = 14.611$,
 $t_{0.025}(5) = 2.571$, $t_{0.05}(5) = 2.015$, $t_{0.025}(6) = 2.4469$, $t_{0.05}(6) = 1.9432$, $t_{0.025}(7) = 2.3646$,
 $t_{0.05}(7) = 1.8946$, $t_{0.025}(18) = 2.1009$, $t_{0.025}(19) = 2.0930$, $t_{0.025}(20) = 2.0860$, $t_{0.025}(24) = 2.0639$,
 $t_{0.05}(24) = 1.7109$, $t_{0.025}(25) = 2.0595$, $t_{0.05}(25) = 1.7081$, $F_{0.05}(3,6) = 4.76$, $F_{0.05}(4,6) = 4.53$,
 $F_{0.05}(8,7) = 3.73$, $F_{0.025}(8,7) = 4.9$, $F_{0.05}(9,8) = 3.39$, $F_{0.025}(9,8) = 4.36$

分数	阅卷人

1、(8分) 一生产线生产的产品成箱包装，每箱的重量是随机的，假设每箱平均重 为 50 kg，标准差为 5 kg，若用最大载重量为 5 吨的汽车承运，试利用中心极限定理说明每辆车最多可以装多少箱，才能保证不超载的概率大于0.977.

解 记 X_i 为第 i 箱的重量, 由题意知 X_1, X_2, \dots , 独立同分布, $EX_i = 50$, $DX_i = 5$.

设装载 k 箱符合要求, 则

由中心极限定理

$$\frac{\sum_{i=1}^k X_i - 50k}{5\sqrt{k}} \sim N(0,1)$$

$$P\left(\sum_{i=1}^k X_i \leq 5000\right) = P\left(\frac{\sum_{i=1}^k X_i - 50k}{5\sqrt{k}} \leq \frac{5000 - 50k}{5\sqrt{k}}\right) \geq 0.977.$$

则 $\frac{5000 - 50k}{5\sqrt{k}} \geq 2$. 解得: $k \leq 98.0199$.

即最多可装 98 箱.

分数	阅卷人

2、(8分) 设总体 $X \sim N(20, 3/2)$ ，从 X 中分别抽取容量为 10, 15 的两个相互独立的样本，求两样本均值之差的绝对值大于 0.3 的概率.

解： 设 \bar{X} , \bar{Y} 分别为两样本均值. 由抽样定理可知

$$\bar{X} \sim N(20, 3/20), \quad \bar{Y} \sim N(20, 3/30),$$

且相互独立, 于是

$$\bar{X} - \bar{Y} \sim N(0, 3/20 + 3/30) = N(0, 1/4)$$

于是所求概率为

$$P(|\bar{X} - \bar{Y}| > 0.3) = P\left(\left|\frac{\bar{X} - \bar{Y}}{1/2}\right| > 0.6\right) = 2(1 - \Phi(0.6)) = 2(1 - 0.7257) = 0.5486.$$

分数	阅卷人

3、(8分) 设总体 X 的概率密度为

$$f(x) = \begin{cases} \theta x^{-(\theta+1)}, & x > 1 \\ 0, & \text{其它} \end{cases}$$

其中 $\theta > 0$ 。设 X_1, X_2, \dots, X_n 为来自总体 X 的简单随机样本。求参数 θ 的最大似然估计。

解：设 x_1, x_2, \dots, x_n 为样本观测值，似然函数为

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

对数似然函数为

$$\ln(L(\theta)) = n \ln(\theta) - (\theta + 1) \sum_{i=1}^n \ln(x_i)$$

令 $(\ln(L(\theta)))' = 0$ 得

$$\frac{n}{\theta} - \sum_{i=1}^n \ln(x_i) = 0$$

解得 $\hat{\theta} = \frac{n}{\sum_{i=1}^n \ln(x_i)}$

故 θ 的最大似然估计量为

$$\hat{\theta} = \frac{n}{\sum_{i=1}^n \ln(X_i)}$$

分数	阅卷人

4、(12分) 在某大学中, 随机抽取 25 名男同学测量身高数据, 算得平均高为 170 cm, 标准差为 12 cm, 试分别求该大学全体男同学平均身高 μ 和身高标准差 σ 的 0.95 置信区间(假设所测身高服从正态分布)。

解 由题设知身高 $X \sim N(\mu, \sigma^2)$, $n = 25$, $\bar{x} = 170$, $S = 12$, $\alpha = 0.05$

(1) 先求 μ 的置信区间(σ^2 未知). 取统计量 $T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t(n-1)$, 可得 μ 的置信度为 $1 - \alpha$ 的置信区间为

$$\left[\bar{x} - \frac{S}{\sqrt{n}} t_{\alpha/2}(n-1), \bar{x} + \frac{S}{\sqrt{n}} t_{\alpha/2}(n-1) \right]$$

查表得 $t_{\alpha/2}(24) = t_{0.025}(24) = 2.0639$. 将数据代入得置信区间为 (165.05, 174.95)。

(2) σ^2 的置信区间(μ 未知). 取统计量

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

由 $P(\chi^2 > \chi^2_{\alpha/2}(n-1)) = P(\chi^2 < \chi^2_{1-\alpha/2}(n-1)) = \alpha/2$, 查 χ^2 分布表, 所以 $\chi^2_{1-\alpha/2}(n-1) = \chi^2_{0.975}(24) = 12.401$, 和 $\chi^2_{\alpha/2}(n-1) = \chi^2_{0.025}(24) = 39.364$, 得参数 σ^2 的置信度为 $1 - \alpha = 0.95$ 的置信区间为

$$\left[\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}, \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)} \right] \quad (87.80, 278.69)$$

的 0.95 的置信区间为 $(\sqrt{87.80}, \sqrt{278.69}) = (9.34, 16.69)$

分数	阅卷人

5、(10 分) 已知某类材料的强度 X 服从正态分布 $N(\mu, \sigma^2)$ 且 $EX = 52$. 今改变配方, 利用新配方生产材料, 从新生产的材料中抽取 7 根, 测得其强度为(单位: M Pa) 52.45, 48.51, 56.02, 51.53, 49.02, 53.38, 54.04, 问用新配方生产的材料强度的均值是否有显著提高? ($\alpha = 0.05$)

解 本题是单侧检验问题, 欲检验的假设为 $H_0: \mu = \mu_0 = 52$, $H_1: \mu > 52$, 由于 σ^2 未知选取统计量 $T = \frac{\bar{X} - \mu_0}{S} \sqrt{n}$. 由所给数

据求得

$$\bar{x} = \frac{1}{7} \sum_{i=1}^7 x_i = 52.14, \quad s^2 = \frac{1}{6} \sum_{i=1}^7 (x_i - \bar{x})^2 = (2.7)^2$$

$$t = \frac{\bar{x} - 52}{2.7} \sqrt{7} = \frac{52.14 - 52}{2.7} \sqrt{7} = 0.14$$

由 $\alpha = 0.05$, 自由度 $n - 1 = 6$, 查 t 分布表得 $t_{0.05}(6) = 1.9432$, 因 $t = 0.14 < t_{0.05}(6) = 1.9432$, 所以接受 H_0 , 即可认为新配方生产的材料强度没有显著提高.

分数	阅卷人

6、（8分）设某零件厂生产的零件的直径服从正态分布。为了减小方差，提高生产精度，工厂试用了新工艺。现分别抽测了采用新、旧工艺生产的零件的直径（单位：mm），结果如下：

旧工艺的样品数量为9件，样本方差为0.1950（mm²）；

新工艺的样品数量为8件，样本方差为0.0486（mm²）。

试问抽样结果是否有充分理由支持工厂采用新工艺（ $\alpha = 0.05$ ）？

解 设旧、新工艺下零件的直径分别服从 $N(\mu_1, \sigma_1^2), N(\mu_2, \sigma_2^2)$ 。

建立假设

$$H_0: \sigma_1^2 \leq \sigma_2^2 \quad H_1: \sigma_1^2 > \sigma_2^2$$

这里 $n_1 = 9, n_2 = 8$. 给定 $\alpha = 0.05$, 查附表 $F_{0.05}(8, 7) = 3.73$, 利用样本计算

$$F = \frac{s_1^2}{s_2^2} = 4.01.$$

由于 $F = 4.01 > F_{0.05}(8, 7) = 3.73$, 故拒绝 H_0 , 即认为抽样结果可以支持工厂采用新工艺.

分数	阅卷人

7、(10分) 考虑抛骰子试验, 若独立重复进行60次, 点数 1, 2, 3, 4, 5, 6 出现的次数如下: 5, 5, 7, 13, 15, 15. 使用 χ^2 拟合检验法检验该骰子是否为均匀的? (显著性水平 $\alpha=0.05$), 即检验假设

$$H_0: p_i = 1/6, \quad i=1,2,\dots,6, \quad \text{其中 } p_i \text{ 为点数 } i \text{ 出现的概率.}$$

解: 在 H_0 成立下,

$$\text{检验统计量 } \chi^2 = \left(\sum_{i=1}^6 \frac{n_i^2}{n p_i} \right) - n \sim \chi^2(5)$$

$$\text{拒绝域为 } \chi^2 \geq \chi_{0.05}^2(5) = 11.07$$

$$\begin{aligned} \chi^2 \text{ 的观察值为 } \chi^2 &= 2.5 + 2.5 + 4.9 + 16.9 + 22.5 + 22.5 - 60 \\ &= 11.8 \geq 11.07 \end{aligned}$$

因此拒绝 H_0

分数	阅卷人

8、(12 分) 某食品公司对一种食品设计了四种新包装.为了考察哪种包装最受顾客欢迎,选了 10 个地段繁华程度相似且规模相近的商店做试验,其中两种包装各指定了两个商店销售,另两个包装各指定三个商店销售.在试验期内各店货架摆放的位置和空间都相同,营业员的促销方法也基本相同,经过一段时间,记录其销售量数据.问:

四种包装是否存在显著差异? ($\alpha = 0.05$)

四种包装的销售量

A ₁	12	18	
A ₂	14	12	13
A ₃	19	17	21
A ₄	24	30	

解: 由上表求得各水平下数据和与总和, 列于下表中:

四种包装销售量的计算过程

水平	n_i	T_i	T_i^2/n_i	$\sum_{j=1}^{n_j} x_{ij}^2$
A1	2	30	450	468
A2	3	39	507	509
A3	3	57	1083	1091
A4	2	54	2458	1476
和	$n=10$	$T=180$	$\sum_{i=1}^4 \frac{T_i^2}{n_i} = 3498$	$\sum_{j=1}^4 \sum_{i=1}^{n_j} x_{ij}^2 = 3544$

求得各类偏差平方和为:

$$S_T = 3544 - 3240 = 304, \quad S_A = 3498 - 3240 = 258$$

$$S_E = S_T - S_A = 304 - 258 = 46$$

列出方差分析表:

四种包装销售量的方差分析表

方差来源	平方和	自由度	均方和	F 比
效应	258	3	86	11.22
误差	46	6	7.67	
总和	304	9		

在 $\alpha = 0.05$ 时, 查得心 $F_{0.05}(3,6)=4.76$, $F=11.22>4.76$, 故样本落入拒绝域, 即认为四种包装的销售量有显著差异, 这说明不同包装受欢迎的程度不同

分数	阅卷人

9、(12分) 在某国, 某地区被称呼为“霾都”. 假设得到最近一次雾霾爆发期间该地区周边的20个城市的数据: 这些城市与该地区的距离(单位为百公里)为 x , 污染指数为 y (取值为0到20之间, 值越大, 污染越严重). 计算得到如下结果:

$$\sum_{i=1}^{20} x_i = 200, \quad \sum_{i=1}^{20} x_i^2 = 2500, \quad \sum_{i=1}^{20} y_i = 150, \quad \sum_{i=1}^{20} y_i^2 = 2775, \quad \sum_{i=1}^{20} x_i y_i = 1000$$

(1) 试建立污染指数 y 对距离 x 的线性回归方程 $\hat{y} = \hat{a} + \hat{b}x$.

(2) 检验回归效果是否显著(显著性水平 $\alpha = 0.05$)? (若显著, 则可以支持该地区被称呼为“霾都”的说法.)

$$\text{解: (i)} \quad S_{xx} = \sum_{i=1}^{20} x_i^2 - \frac{1}{20} \left(\sum_{i=1}^{20} x_i \right)^2 = 500$$

$$S_{yy} = \sum_{i=1}^{20} y_i^2 - \frac{1}{20} \left(\sum_{i=1}^{20} y_i \right)^2 = 1650$$

$$S_{xy} = \sum_{i=1}^{20} x_i y_i - \frac{1}{20} \sum_{i=1}^{20} x_i \sum_{i=1}^{20} y_i = -500$$

$$\Rightarrow \hat{b} = \frac{S_{xy}}{S_{xx}} = -1$$

$$\hat{a} = \frac{1}{20} \sum_{i=1}^{20} y_i - \left(\frac{1}{20} \sum_{i=1}^{20} x_i \right) \hat{b} = 17.5$$

$$\text{回归方程为 } \hat{y} = 17.5 - x$$

$$(ii) \quad \hat{\sigma}^2 = \frac{1}{n-2} (S_{yy} - \hat{b} S_{xy}) = 63.89$$

$$\text{检验统计量 } t = \frac{\hat{b}}{\hat{\sigma}} \sqrt{S_{xx}}, \text{ 拒绝域 } |t| > t_{0.025}(18)$$

$$\text{计算得 } |t| \approx 2.8$$

$$\text{临界值 } t_{0.025}(18) = 2.1009$$

因此拒绝 H_0 . 也即回归效果显著,

支持该地区为“霾都”的说法.

分数	阅卷人

10、(12分) 已知两个总体 X, Y 均服从指数分布, 参数分别为 θ_1, θ_2 .

现分别得到两个总体的两个独立样本 X_1, X_2, \dots, X_{n_1} 和 Y_1, Y_2, \dots, Y_{n_2} .

记其样本均值分别为 \bar{X} 和 \bar{Y} . 引入枢轴量 $W = \frac{\bar{X}/\theta_1}{\bar{Y}/\theta_2}$.

(1) 试求 W 的分布(已知: $2X_i/\theta_1 \sim \chi^2(2)$, $2Y_i/\theta_2 \sim \chi^2(2)$);

(2) 从上述枢轴量的角度构造参数 θ_1/θ_2 的 $1-\alpha$ 水平的单侧置信上限;

(3) 从上述枢轴量的角度构造如下假设检验问题的拒绝域(显著性水平为 α):

$$H_0: \theta_1 = \theta_2; H_1: \theta_1 \neq \theta_2.$$

解: (i) $\sum_{i=1}^{n_1} 2X_i/\theta_1 \sim \chi^2(2n_1), \sum_{i=1}^{n_2} 2Y_i/\theta_2 \sim \chi^2(2n_2)$

$$\text{枢轴量 } W = \frac{\bar{X}/\theta_1}{\bar{Y}/\theta_2} = \frac{(\sum_{i=1}^{n_1} 2X_i/\theta_1)/(2n_1)}{(\sum_{i=1}^{n_2} 2Y_i/\theta_2)/(2n_2)} \sim F(2n_1, 2n_2)$$

$$\begin{aligned} \text{(ii)} \quad 1-\alpha &= P\left\{ \frac{\bar{X}/\bar{Y}}{\theta_1/\theta_2} > F_{1-\alpha}(2n_1, 2n_2) \right\} \\ &= P\left\{ \theta_1/\theta_2 < \frac{\bar{X}/\bar{Y}}{F_{1-\alpha}(2n_1, 2n_2)} \right\} \end{aligned}$$

$$\text{因此置信上限为 } \frac{\bar{X}/\bar{Y}}{F_{1-\alpha}(2n_1, 2n_2)}$$

$$\text{(iii)} \quad \text{检验统计量 } W_0 = \frac{\bar{X}}{\bar{Y}}$$

$$\text{拒绝域为 } \bar{X}/\bar{Y} \geq F_{\frac{\alpha}{2}}(2n_1, 2n_2)$$

$$\text{或 } \bar{X}/\bar{Y} \leq F_{1-\frac{\alpha}{2}}(2n_1, 2n_2)$$