西安交通大学考试题数理统计 2000 年

一、填空

1. 设 $X_1, X_2, \cdots X_{10}$ 是来自正态总体 $N(0,4)$ 的样本，

$$c \cdot \frac{\sum_{i=1}^{5} X_i}{\sqrt{\sum_{j=1}^{5}(X_j - \frac{1}{5}\sum_{i=1}^{5} X_i)^2 + \sum_{i=6}^{10} X_i^2}} \sim t(m)$$

则 $c = 70$ ，$m = 49$

2. 用 $\Phi(x)$ 表示标准正态分布 $N(0,1)$ 的分布函数，则 $\Phi(-x)$ 与 $\Phi(x)$ 的关系为 $\Phi(-x) = 1 - \Phi(x)$

3. 已知 $T \sim t(n)$，则 $T^2 \sim F(1,n)$

4. 设总体 X 的概率密度为 $f(x;\theta)$，则参数 θ 估计的费歇（Fisher）信息量 $I(\theta) = -E\left[\frac{\partial^2 \ln f(x;\theta)}{\partial \theta^2}\right]$

二、选择题（填 A,B,C,D，有几个正确填几个，若都不正确，则填 E）

1. 设 $(X_1, X_2, \cdots X_{20})$ 是来自正态总体 $N(\mu, \sigma^2)$ 的样本，统计量 $S = \sum_{i=1}^{10}(2X_{2i} - X_{2i-1})^2$，则 （C）

A. $\frac{1}{5\sigma^2} S \sim \chi^2(9)$　　B. $\frac{1}{3\sigma^2} S \sim \chi^2(9)$

C. $\frac{1}{5\sigma^2} S \sim \chi^2(10)$　　D. $\frac{1}{3\sigma^2} S \sim \chi^2(10)$

2. 独立地分别从两总体 X 和 Y 中抽得大小各为 m 和 n 的样本，其样本均值分别为 $\bar{X}$ 和 $\bar{Y}$，则 $D(\bar{X} - \bar{Y}) =$ （D）

A. $\frac{D(X)}{m^2} - \frac{D(Y)}{n^2}$　　B. $\frac{D(X)}{m^2} + \frac{D(Y)}{n^2}$

C. $\frac{D(X)}{m} - \frac{D(Y)}{n}$　　D. $\frac{D(X)}{m} + \frac{D(Y)}{n}$

3. 设 $X_1, X_2, \cdots X_n$ 是来自正态总体 $N(\mu, \sigma^2)$ 的样本，其中 μ 已知，而 σ^2 未

知，则下列是统计量的是 （AD）

A. $\overline{X}+X_6$　　B. $\frac{1}{\sigma^2}\sum_{i=1}^{n}X_i^2$

C. $\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2$　　D. $(X_2-\mu)^2+2X_8+\mu$

4. 设 $F\sim F(6,8)$，则 （AD）

A. $P\{F<F_{1-\alpha}(6,8)\}=\alpha$　　B. $P\{|F|>\frac{1}{F_\alpha(8,6)}\}=\alpha$

C. $P\{|F|>F_{1-\frac{\alpha}{2}}(6,8)\}=\alpha$　　D. $P\{F<\frac{1}{F_\alpha(8,6)}\}=\alpha$

三、为比较 A、B 两型灯泡的寿命，随机抽取 A 型灯泡 5 只，测得平均寿命 $\overline{x}$=1000（小时），标准差 s_A=28（小时）；随机抽取 B 型等泡 7 只，测得平均寿命 $\overline{y}$=980（小时），标准差 s_B=32（小时），设总体都是正态的，试在显著性水平之下 $(\alpha=0.05)$ 检验两总体寿命分布是否相同。

解：

依题意，需要检验假设为

$$H_{01}:\sigma_x=\sigma_y \leftrightarrow H_{02}:\sigma_x\neq\sigma_y$$
$$H_{02}:\mu_x=\mu_y \leftrightarrow H_{02}:\mu_x\neq\mu_y$$

i）对于假设 H_{01}，由于 μ_x,μ_y 未知，则可选取

$$F=\frac{S_x^{*2}}{S_y^{*2}}$$

作为检验统计量，当 H_{01} 成立时，

$$F=\frac{S_x^{*2}}{S_y^{*2}}\sim F(n_x-1,n_y-1)$$

且当 H_{01} 成立时，F 过分偏大将不利于 H_{01} 的成立，则其拒绝域的形式为

$$W=\{F\le k_1\cup F\ge k_2\}$$

对于给定的水平 $\alpha=0.05$，于是 $k_1=F_{0.975}(n_x-1,n_y-1),k_2=F_{0.025}(n_x-1,n_y-1)$

对于本题，

$$n_x = 5, S_x^{*2} = n_x \cdot s_A{}^2 / (n_x - 1) = 980, n_y = 7, S_y^{*2} = n_y \cdot s_B{}^2 / (n_y - 1) = 7 \times 32^2 / 6$$

则

$$F = \frac{S_x^{*2}}{S_y^{*2}} = \frac{980 \times 6}{7 \times 32^2} = 0.8203$$

查表得，$F_{0.025}(4,6) = 6.23, F_{0.975}(4,6) = 1/F_{0.025}(6,4) = 1/9.2 = 0.0892$。由于 $0.0892 < F < 6.23$，则接受假设，即认为 $\sigma_x = \sigma_y$。

ii）对于假设 H_{02}，由于 $\sigma_x = \sigma_y$，且未知，则可选取

$$T = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{1}{n_x} + \frac{1}{n_y}} \cdot \sqrt{\frac{(n_x - 1)S_x^{*2} + (n_y - 1)S_y^{*2}}{n_x + n_y - 2}}}$$

作为检验统计量，当 H_{02} 成立时，

$$T = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{1}{n_x} + \frac{1}{n_y}} \cdot \sqrt{\frac{(n_x - 1)S_x^{*2} + (n_y - 1)S_y^{*2}}{n_x + n_y - 2}}} \sim t(n_x + n_y - 2)$$

且当 H_{01} 成立时，T 过分偏大将不利于 H_{02} 的成立，则其拒绝域的形式为

$$W = \{|T| \geq k\}$$

对于给定的水平 $\alpha = 0.05$，于是 $k = t_{0.025}(n_x - 1, n_y - 1)$

对于本题，

$$n_x = 5, S_x^{*2} = n_x \cdot s_A{}^2 / (n_x - 1) = 980, n_y = 7, S_y^{*2} = n_y \cdot s_B{}^2 / (n_y - 1) = 7 \times 32^2 / 6$$

则 $T = 1.0258$，查表得，$T_{0.025}(10) = 2.2281$。由于 $1.0258 < T$，则拒绝假设，即认为 $\mu_x \neq \mu_y$。

综上所述，两总体寿命分布不相同。

四、想要考察特定一群人的收入与其花在书籍报纸上的支出有无关系，把收入分成高、中、低三档，书报上的支出分别为多、少两档。设随机抽查了 201 名，结果如下表：

收入 / 书报上支出	低	中	高
少	63	37	60
多	16	17	8

试在水平$\alpha = 0.05$之下检验收入与书报上支出有无关联。

解：

依题意，有两个指标，记X为书报上的支出情况，记Y为收入情况，且记A_1, A_2分别为书报上的支出多、少两档；B_1, B_2, B_3分别为收入高、中、低三档。则需要检验的假设为$H_0: p_{ij} = p_{i\cdot} \cdot p_{\cdot j}, i = 1,2; j = 1,2,3 \leftrightarrow H_1$:至少存在某组$(i, j)$，使得$p_{ij} \neq p_{i\cdot} \cdot p_{\cdot j}$。其中$p_{ij} = P\{X = A_i, Y = B_j\}, p_{i\cdot} = P\{X = A_i\}, p_{\cdot j} = P\{Y = B_j\}$。

于是，

$$k = \frac{\sum_{j=1}^{3}\sum_{i=1}^{2}(n_{ij} - \frac{n_{i\cdot}n_{\cdot j}}{n})^2}{\frac{n_{i\cdot}n_{\cdot j}}{n}} = 6333.7$$

查表得，$\chi^2_{0.05}(2) = 5.991$。由于$k > 5.991$，则拒绝假设，即认为收入与书报上支出有关联。

五、在硝酸钠（$NaNO_3$）的溶解度试验中，测得在不同温度$x\ (^\circ C)$下，溶解于9份水中的硝酸钠份数y的数据如下表：

x：	0	4	10	15	21	29	36	51	68
y：	66.7	71.0	76.3	80.6	85.7	92.9	99.4	113.6	125.1

假设y与x之间有线性关系，在正态假定下，求y在$x_0 = 25$的置信度为95%的预测区间。

解：

依题意，可计算得

$$\hat{b} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{n}(x_i - \overline{x})^2} = 0.8706$$

$$\hat{a} = \overline{y} - \hat{b}\overline{x} = 67.5078$$

则y与x之间有线性关系为$y = 0.8706x + 67.5078$。

于是，当$x_0 = 25$时，$y_0 = 89.2728$。

y置信度为95%的预测区间为

$$(y_0 \mp t_{0.025}(n-2)\cdot\sigma^{*2} \Big/ \sqrt{1+1/n+(x_0-\bar{x})\Big/\sum_{i=1}^{n}(x_i-\bar{x})})$$

则 y 在 $x_0=25$ 的置信度为 95%的预测区间为

六、设自一大批产品中随机抽出 200 个产品，发现其中 120 个是一等品，求这大批产品的一等品率的 95%置信区间。

解：

$$U=\frac{\sqrt{n}(\bar{X}-p)}{\sqrt{p(1-p)}}\sim N(0,1)$$

七、今有两台测量合金材料中某种金属含量的光谱仪，为鉴定他们的测量准确性有无显著差异，对 9 件含该金属分别为 $x_1,x_2,\cdots,x_9$（不等）的合金材料进行测量，第一台测量结果服从正态分布 $N(x_t+\delta_1,\sigma_1^2)$， $t=1,2,\cdots,9$；第二台测量结果服从 $N(x_t+\delta_2,\sigma_2^2)$， $t=1,2,\cdots,9$。测得的 9 对观测值如下：

第一台：	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
第二台：	0.10	0.21	0.52	0.32	0.78	0.59	0.68	0.77	0.89

问能否认为第一台的测量值比第二台显著偏大（$\alpha=0.05$）？

八、设 $X_1,X_2,\cdots X_n$ 是来自总体 X 的样本（$n>2$），总体的概率密度为

$$f(x;\lambda,a)=\begin{cases}\lambda e^{-\lambda(x-a)},当x\geq a\\ 0\quad ,当x<a\end{cases}\quad（参数\ \lambda>0）$$

1. 设 $\lambda=\frac{1}{\sqrt{2}}$，试求参数 a 的最大似然估计。
2. 设 $a=0$，试求参数 λ 的矩估计。
3. 设 $a=0$，试推导 $2n\lambda\bar{X}$ 服从的分布，其中 $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ 。
4. 设 $a=0$，试计算 $E(\hat{\lambda})$，并求 k，使 $\hat{\lambda}^*=k\hat{\lambda}$ 为 λ 的无偏估计。

解：

1. 先得到似然函数

$$L(a)=\prod_{i=1}^{n}\lambda e^{-\lambda(x_i-a)}=\lambda^n e^{-\lambda\sum_{i=1}^{n}(x_i-a)}=\lambda^n e^{-\lambda\sum_{i=1}^{n}x_i+n\lambda a} \qquad x_i\ge a$$

由于

2.