2003 届硕士研究生 应用数理统计试题（签） 20040106

解答一律写在答题本上，试题（签）必须和答题本一起交回！

一、填空（记5分×10=50分）

1. 设总体 $X\sim N(\mu,\sigma^2)$，来自 X 的一组样本记为 $X_1,X_2,\cdots,X_n$，记

$\overline{X}_n=\frac{1}{n}\sum_{i=1}^{n}X_i$，$S_n^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X}_n)^2$，又从 X 中抽取另一个样本 X_{n+1}，则统计量

$X_{n+1}-\overline{X}_n\sim(0,\frac{n+1}{n}\sigma^2)$

$\sqrt{\frac{n}{n+1}}\frac{X_{n+1}-\overline{X}_n}{S_n}$ 服从的抽样分布是 ____$t(n-1)$____。

2. 设总体 $X\sim N(\mu,\sigma^2)$，来自 X 的一组样本记为 $X_1,X_2,\cdots,X_n$，其中 σ^2 已知。

$\overline{X}_n=\frac{1}{n}\sum_{i=1}^{n}X_i$，$S_n^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X}_n)^2$，则 $\lim\limits_{n\to\infty}P(\frac{(n-1)(S_n^2-\sigma^2)}{\sigma^2\sqrt{2n-2}}\le x)=$ ________

3. 若母体 X 在区间 $[\alpha,\alpha+\theta]$ 上均匀分布，来自 X 的一组样本为 $\xi_i,(i=1,2,\Lambda,n)$，统计量 $\alpha=\underset{1\le i\le n}{Min}\,\xi_i$，$\beta=\underset{1\le i\le n}{Max}\,\xi_i$，那么参数 α，θ 的极大似然估计分别为：

$\hat{\alpha}=$ ____α____，$\hat{\theta}=$ ____$\beta-\alpha$____

$D(m\hat\theta_1+n\hat\theta_2)=$

$2m^2+(1-m)$

$E(m\hat\theta_1)+E(n\hat\theta_2)=m+n=1$

$m+n=1$

4. 已知参数 θ 的两个独立的无偏估计统计量为 $\hat{\theta}_1,\hat{\theta}_2$，而且 $D\hat{\theta}_1=2D\hat{\theta}_2$，由 $\hat{\theta}_1,\hat{\theta}_2$ 的线性组合构成的无偏估计统计量为 $m\hat{\theta}_1+n\hat{\theta}_2$（$m,n$ 为常数），要使所有这样的线性估计量中的方差最小，该统计量应该为 ____$\frac{1}{3}\hat\theta_1+\frac{2}{3}\hat\theta_2$____

2.5

5. 设总体 $X\sim N(\mu,10^2)$，若使总体均值 μ 对于置信概率 95% 的置信区间的长度不大于 5，问至少要取 ____62____ 个样本。[$\Phi(1.96)=0.975$，$\Phi(1.645)=0.95$]

$(\frac{19.6}{2.5})$

6. 已知统计量 $X\sim t(n)$，那么统计量 $\frac{1}{X^2}\sim$ ____$F(n,1)$____

7. 设总体 $X\sim N(\mu,\sigma^2)$，来自 X 的一组简单随机样本记为 $X_1,X_2,\cdots,X_n$，其中 μ,σ^2 均已知，

则 $E[\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2]^2=$ ____σ^4____

$\frac{n+2}{n}\sigma^4$ $E\left(\frac{n S^2}{\sigma^2}\right)^2=n^2-1$

$X_i-\mu\sim N(0,\sigma^2)$

$D(\sum X_i-\mu)^2$ $\frac{n\sigma^2}{n}$ $\frac{\sigma^4}{n^2}$ $D(3)+E^2(3)$

$\frac{X_i-\mu}{\sigma}\sim N(0,1)$ χ^2 $E(t^2)=[D(t)+E^2(t)]$ $2(n-1)+(n-1)^2$

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

$\frac{\sum(X_i-\mu)^2}{\sigma^2}\sim\chi^2(n)$ $E(t^2)=\frac{2n+n^2}{n^2}\sigma^4$ n^2-1

$E\left[\frac{1}{n}\sum(X_i-\mu)^2\right]=\sigma^2$

从一批某种型号的电子管中抽取容量为 10 的子样，计算得到修正标准差 $s^*=45$（小时）。假设这批电子管的寿命服从正态分布，那么这批电子管寿命标准差 σ 对于置信概率为 0.95 的单侧置信上限为________。

9. 假定每次试验时，出现事件 A 的概率 p 相同但未知，若在 100 次独立试验中，事件 A 出现了 10 次，那么概率 p 对置信概率为 95% 的置信区间是________。 $0.1\pm\frac{\sqrt{0.1\times0.9}}{10}\times1.96$

10. 在假设检验理论中常犯的两类错误：第一类（弃真）错误，概率为 α；第二类（纳伪）错误，概率为 β。根据下表中的事实与决策，在表中填写上适当的概率：（共四项）

决策 \ 事实	H_0 为真	H_0 为假
接受 H_0		β
拒绝 H_0	α	

二、解答下列各题（应该有必要的解题步骤，共 50 分）。

11 某医院用光电比色计检验尿汞时，得尿汞含量（mg / L）与消光系数读数的结果如下：

尿汞含量 x_i	2	4	6	8	10
消光系数 y_i	64	138	205	285	360

128　552　1230　2280　3600

P223, 1.

已知它们之间服从线性模型：$E(Y_i)=\alpha+\beta x_i$，试求 α 和 β 的最小二乘估计，并写出其经验回归直线方程。（本题 7 分）

12. 若 $\hat{y}=\hat{\alpha}+\hat{\beta}x$ 是 n 组数据 (x_i,y_i)（$i=1,2,\cdots,n$）的经验回归直线方程，记 $\hat{y}_i=\hat{\alpha}+\hat{\beta}x_i$（$i=1,2,\cdots,n$），求证：$\sum_{i=1}^{n}(y_i-\bar{y})^2=\sum_{i=1}^{n}(y_i-\hat{y}_i)^2+\sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2$。（本题 6 分）

P220

13 某灯泡厂在采用一项新工艺前后，分别取 10 个灯泡进行寿命试验，得到数据如下：新工艺前：$\bar{x}_1=2460$（小时），$s_1=56$（小时）；新工艺后：$\bar{x}_2=2550$（小时），$s_2=48$（小时）。若设灯泡的寿命服从正态分布，是否可以认为采用新工艺后使得灯泡的寿命有显著的提高。（取 $\alpha=0.01$）（本题 9 分）

$\mu_1<\mu_2$　σ^2 未知

解：假设检验：$H_0:\mu_1-\mu_2\le0$　$H_1:\mu_1-\mu_2>0$

由题意：$\bar{X}_1-\bar{X}_2\sim N(0,(\frac{1}{n_1}+\frac{1}{n_2})\sigma^2)$

$\frac{\bar{X}_1-\bar{X}_2}{\sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\sim N(0,1)$

由题意：$\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{\sigma^2}\sim\chi^2(n_1+n_2-2)$

$\sim t(n_1+n_2-2)$

14. 检查产品质量的时候，在生产过程中每次抽取 10 个产品来检查，抽查 100 次，得到每 10 个产品中的次品数的分布如下：

$H_0: X\sim B(n,p)=B(n,0.1)$;　P168

每 10 个产品的次品数 K	0	1	2	3	4	5	≥5	总计
频数 m_i	32	45	17	4	1	1	0	100

利用分布律的假设检验理论来检验：生产过程中出现的次品的概率 p 是否可以认为是不变的。

$E(\xi)=np=1=\bar{x}$

$\frac{0\times32+1\times45+\cdots+5\times1+6\times0}{100}=1$

$p=\frac{1}{10}=0.1$

$P_i=C_n^i\,0.1^i\,0.9^{n-i}$

$P_0=C_{10}^0\cdot0.1^0\cdot0.9^{10}$　$P_1=C_{10}^1\cdot0.1\cdot0.9^9$

$P_2=C_{10}^2\cdot0.1^2\cdot0.9^8$　$P_3=C_{10}^3\cdot0.1^3\cdot0.9^7$

$P_4=C_{10}^4\cdot0.1^4\cdot0.9^6$　$P_5=C_{10}^5\cdot0.1^5\cdot0.9^5$

试是说每次抽查的10个产品中的次品数K的概率分布为

$P(\xi=K)=C_{10}^{K}p^{K}(1-p)^{10-K}, K=0,1,2,\cdots$ （取$\alpha=0.05$）（[illegible]）

15 为了考察收缩率与总拉伸倍数以及它们之间的交互作用对合成纤维的弹性的影响，各取四个水平，并且在每个水平的配合下重复试验两次，其数据如下：

总拉伸倍数 / 收缩率	B_1	B_2	B_3	B_4	行平均
A_1	71, 73 (72)	72, 73 (72.5)	73, 75 (74)	75, 77 (76)	$\bar{x}_{1.}=73.625$
A_2	73, 75 (74)	74, 76 (75)	77, 78 (77.5)	74, 74 (74)	$\bar{x}_{2.}=75.125$
A_3	73, 76 (74.5)	77, 79 (78)	74, 75 (74.5)	74, 73 (73.5)	[illegible]
A_4	73, 75 (74)	72, 73 (72.5)	71, 70 (70.5)	69, 69 (69)	[illegible]
列平均	$\bar{x}_{.1}=73.625$	$\bar{x}_{.2}=74.5$	$\bar{x}_{.3}=74.125$	$\bar{x}_{.4}=73.125$	[illegible]

括号里的数是两次试验数据的平均。假设合成纤维的弹性服从正态分[illegible]倍数以及它们的交互作用对合成纤维的弹性的影响是否显著？并指出[illegible]

表格中的32个数据的修正标准差$S^*=2.4111$，取$\alpha=0.01$.（本题10分）

16 乙酰苯磺化试验。试验目的：提高收得率。因素与水平为：

A：反应温度（度）；$A_1=50$，$A_2=70$.

B：反应时间（小时），$B_1=1$，$B_2=2$.

C：硫酸浓度（%），$C_1=17$，$C_2=27$.

D：操作条件，D_1=搅拌，D_2=不搅拌.

方差分析

根据以往的经验已经知道：只有因素 A, B 之间才可能存在交互作用。因此采用如下的正交表 $L_8(2^7)$ 的表头设计，用方差分析法找出对收得率有显著影响的因素。（取显著性水平$\alpha=10\%$）

注：$F_{0.1}(1,2)=8.53$，$F_{0.1}(2,1)=49.5$，$F_{0.1}(1,1)=39.86$.（本题9分）.