院
班 级
学 号
姓 名

密 封 线

东北大学考试试卷

2007 — 2008 学年第 1 学期

课程名称： 应用数理统计

一	二	三	四	五	六	七	总分

说明：

1. 样本均值与样本方差分别定义为

$$\overline{X}=\frac{1}{n}\sum_{k=1}^{n}X_k,\ S^2=\frac{1}{n-1}\sum_{k=1}^{n}(X_k-\overline{X})^2$$

2. 分位点统一为上分位点，即随机变量 X 的上 α $(0<\alpha<1)$ 分位点 Q_α 是指：

$$P(X>Q_\alpha)=\alpha \qquad \overline{X}_1\sim N(\mu,\frac{\sigma^2}{n})$$

一.（10 分）设 $\overline{X}_1$ 与 $\overline{X}_2$ 是从同一正态总体 $N(\mu,\sigma^2)$ 独立抽取的容量相同的两个样本均值。试确定其中一个样本的样本容量 n 的取值范围，使得 $P(|\overline{X}_1-\overline{X}_2|>\sigma)\le 0.01$。

$\overline{X}_1-\overline{X}_2\sim N(0,\frac{2\sigma^2}{n})$

总体 $\sim N(\mu,\sigma^2)$，$\overline{X}_2\sim N(\mu,\sigma^2)$；$\overline{X}_1\sim N(\mu,\frac{\sigma^2}{n})$

$\overline{X}_2\sim N(\mu,\frac{\sigma^2}{n})$

$\overline{X}_1-\overline{X}_2\sim N(0,\frac{2\sigma^2}{n})$

$\dfrac{\overline{X}_1-\overline{X}_2}{\sqrt{\frac{2\sigma^2}{n}}}\sim N(0,1)$

$P(|\overline{X}_1-\overline{X}_2|>\sigma)$

$=P\left(\dfrac{|\overline{X}_1-\overline{X}_2|}{\sqrt{\frac{2\sigma^2}{n}}}>\sqrt{\frac{n}{2}}\right)\le 0.01 \Rightarrow P(|U|<\sqrt{\frac{n}{2}})>0.99$

$P\left(\left|\dfrac{\overline{X}_1-\overline{X}_2}{\sqrt{\frac{2\sigma^2}{n}}}\right|>\dfrac{\sigma}{\sqrt{\frac{2\sigma^2}{n}}}=\sqrt{\frac{n}{2}}\right)\le 0.01$ $\sqrt{\frac{n}{2}}\ge 2.58$

$P(|U|<\sqrt{\frac{n}{2}})=2\Phi(\sqrt{\frac{n}{2}})-1\ge 0.99$ $\Phi(\sqrt{\frac{n}{2}})\ge 0.995$ $n\ge 14$

二.（15 分）设总体 $X\sim U(0,\theta)$，其中 $0<\theta<\infty$ 是未知参数，从总体 X 中抽取简单随机样本 $X_1,\ldots,X_5$，设 $X_{(1)}=\min\{X_1,\ldots,X_5\}$，$X_{(5)}=\max\{X_1,\ldots,X_5\}$。

(1) 由 θ 的极大似然估计 $\boxed{X_{(5)}}$ 构造 θ 的一个无偏估计 $\hat\theta_1$；

(2) 由 $\boxed{X_{(1)}}$ 构造 θ 的另外一个无偏估计 $\hat\theta_2$；

(3) 问 θ 的两个无偏估计 $\hat\theta_1$ 和 $\hat\theta_2$ 哪个更有效？

$E(X_{(5)}^2)=\int_0^\theta x^2\cdot 5(1-\frac{x}{\theta})^4\frac{1}{\theta}dx$

$=\frac{5}{\theta^5}\int_0^\theta x^6 dx=\frac{5}{7}\theta$

(1)(2) $X\sim U(0,\theta)$，$f(x)=\begin{cases}\frac{1}{\theta} & x\in(0,\theta)\\ 0 & 其它\end{cases}$

$X_{(1)}=1-[1-F(x)]^5$

$X_{(5)}=F^5(x)$

$F(x)=\begin{cases}0 & x<0\\ \frac{x}{\theta} & 0<x<\theta\\ 1 & x>\theta\end{cases}$

$\begin{cases}f_1(x)=X_{(1)}'=5[1-F(x)]^4 f(x)\\ f_5(x)=X_{(5)}'=5F^4(x)f(x)\end{cases}$

$E(X_{(5)})=\int_{-\infty}^{+\infty}x\cdot 5F^4(x)f(x)dx=\int_0^\theta x\cdot 5(\frac{x}{\theta})^4\frac{1}{\theta}dx=\frac{5}{\theta^5}\int_0^\theta x^5$

$E(\frac{6}{5}X_{(5)})=\theta$. $E(X_{(1)})=\int_0^\theta x\cdot 5(1-\frac{x}{\theta})^4\frac{1}{\theta}dx$ 换元 $1-\frac{x}{\theta}=t$

$\hat\theta_1=\frac{6}{5}X_{(5)}$ $=\int_0^1 5(1-t)t^4\theta dt=\frac{1}{6}\theta$. $E(6X_{(1)})=\theta$

(3) $E(\hat\theta_1)=\theta$；$D(\hat\theta_1)=D(\frac{6}{5}X_{(5)})=\frac{36}{25}D(X_{(5)})=\frac{36}{25}[E(X_{(5)}^2)-E^2(X_{(5)})]$ $\hat\theta_2=$

$=\frac{36}{25}[\frac{5}{7}\theta^2-(\frac{5}{6}\theta)^2]=\frac{1}{35}\theta^2$

$E(\hat\theta_2)=\theta$；$D(\hat\theta_2)=D(6X_{(1)})=36D(X_{(1)})=36[E(X_{(1)}^2)-E^2(X_{(1)})]$

学院	
班级	
学号	
姓名	

密 封 线

答案

三（15 分）设 $X_1,\dots,X_n$ 为来自总体 $X\sim N(\mu_1,\sigma_1^2)$ 的样本，$Y_1,\dots,Y_m$ 为来自总体 $Y\sim N(\mu_2,\sigma_2^2)$ 的样本，且两个样本独立。请分别给出在下面两种条件：

(1) σ_1^2 与 σ_2^2 已知时；

(2) σ_1^2 与 σ_2^2 未知，但是 $\sigma_1^2=\sigma_2^2$ 时；$\mu_1-3\mu_2$ 的置信度为 0.95 的双侧置信区间。

1) $\bar{X}\sim N(\mu_1,\frac{\sigma_1^2}{n})$，$\bar{Y}\sim N(\mu_2,\frac{\sigma_2^2}{m})$，$3\bar{Y}\sim N(3\mu_2,\frac{9\sigma_2^2}{n})$

$\bar{X}-3\bar{Y}\sim N(\mu_1-3\mu_2,\frac{\sigma_1^2}{n}+\frac{9\sigma_2^2}{m})$

$\frac{(\bar{X}-3\bar{Y})-(\mu_1-3\mu_2)}{\sqrt{\frac{\sigma_1^2}{n}+\frac{9\sigma_2^2}{m}}}\sim N(0,1)$　$(\bar{X}-3\bar{Y}\pm u_{1-\frac{0.05}{2}}\sqrt{\frac{\sigma_1^2}{n}+\frac{9\sigma_2^2}{m}})$

2) $\frac{\bar{X}-3\bar{Y}-(\mu_1-3\mu_2)}{\sqrt{\frac{1}{n}+\frac{9}{m}}\,\sigma}\sim N(0,1)$

$\frac{(n-1)S_1^2+(m-1)S_2^2}{\sigma^2}\sim\chi^2(m+n-2)$

$\frac{\bar{X}-3\bar{Y}-(\mu_1-3\mu_2)}{\sqrt{\frac{1}{n}+\frac{9}{m}}\,\sigma}\Big/\sqrt{\frac{(n-1)S_1^2+(m-1)S_2^2}{\sigma^2(m+n-2)}}\sim t(m+n-2)$

$\bar{X}-3\bar{Y}\pm t_{1-\frac{0.05}{2}}\sqrt{\frac{(n-1)S_1^2+(m-1)S_2^2}{m+n-2}}\sqrt{\frac{1}{n}+\frac{9}{m}}$

$t_{1-\frac{0.05}{2}}$

四.（15 分）设总体 $X\sim N(\mu_1,\sigma^2)$，总体 $Y\sim N(\mu_2,\sigma^2)$，从总体 X 抽取样本 $X_1,\dots,X_n$，从总体 Y 抽取样本 $Y_1,\dots,Y_m$，且两个样本独立。考虑如下假设检验问题：

$$H_0:c\mu_1+d\mu_2=\delta\leftrightarrow H_1:c\mu_1+d\mu_2\neq\delta,$$

其中 $c\neq0,d\neq0,\delta$ 都是已知常数，求在显著性水平为 α 时此检验的拒绝域。

$\bar{X}\sim N(\mu_1,\frac{\sigma^2}{n})$，$\bar{Y}\sim N(\mu_2,\frac{\sigma^2}{m})$；$c\bar{X}\sim N(c\mu_1,\frac{c^2\sigma^2}{n})$，$d\bar{Y}\sim N(d\mu_2, \dots$

$c\bar{X}+d\bar{Y}\sim N[(c\mu_1+d\mu_2),\frac{c^2\sigma^2}{n}+\frac{d^2\sigma^2}{m}]$

构造统计量

$\frac{c\bar{X}+d\bar{Y}-(c\mu_1+d\mu_2)}{\sqrt{\frac{c^2\sigma^2}{n}+\frac{d^2\sigma^2}{m}}}\sim N(0,1)$ ①

$\frac{(n-1)S_1^2+(m-1)S_2^2}{\sigma^2}\sim\chi^2(m+n-2)$ ②

①/② $\sim t(\ )$

$T=\frac{c\bar{X}+d\bar{Y}-(c\mu_1+d\mu_2)}{\sqrt{\frac{c^2\sigma^2}{n}+\frac{d^2\sigma^2}{m}}}\Big/\sqrt{\frac{(n-1)S_1^2+(m-1)S_2^2}{\sigma^2(m+n-2)}}\sim t(m+n-2)$

$c\mu_1+d\mu_2=\delta$

拒绝域 $\chi_0=\{|T|>c\}=\left\{\left|\frac{(c\bar{X}+d\bar{Y}-\delta)\sqrt{m+n-2}}{\sqrt{(\frac{c^2}{n}+\frac{d^2}{m})[(n-1)S_1^2+(m-1)S_2^2]}}\right|>t_{1-\frac{\alpha}{2}}(m+n-2)\right\}$

学 院
班 级
学 号
姓 名

密 封 线

五.（15 分）掷一颗骰子 60 次，结果如下

点数	1	2	3	4	5	6
出现次数	7	9	11	8	12	13

试在显著性水平为 0.05 下检验这颗骰子是否均匀。检验是否均匀分布

解：$m=6,\ n=60,\ p_i=\frac{1}{6}$

$\chi_0=\left\{\sum_{i=1}^{6}\frac{(\nu_i-np_i)^2}{np_i}>\chi^2_{1-\alpha}(6-1)\right\}$

$\sum_{i=1}^{6}\frac{(\nu_i-10)^2}{10}=\frac{1}{10}[(-3)^2+(9-10)^2+(11-10)^2+(8-10)^2+(12-10)^2+(13-10)^2]$

$=2.8$

$\chi^2_{1-\alpha}(6-1)=\chi^2_{1-0.05}(5)=\chi^2_{0.95}(5)=11.07$

$2.8<11.07$

∴ 拒绝域不成立，所以接受是均匀的。

$\sum(x_i-\bar{x})^2=$

$n_iS_i^2=n_i\cdot\frac{1}{n_i}\sum(x_i-\bar{x})^2=$ 偏差平方和

六.（15 分）为了了解生产某种电子设备的公司在过去三年中的科研经费投入（分为低、中、高三档）对当年生产能力提高的影响，调查了共计 27 家生产该设备的公司，并对当年生产能力较之三年前的提高量做了评估，得到数据如下表所示：

经费投入	生产能力提高量											
低	7.6	8.2	6.8	5.8	6.9	6.6	6.3	6.7	7.0			
中	6.7	8.1	9.4	8.6	7.8	7.7	8.9	7.9	8.3	8.7	7.1	8.4
高	8.5	9.7	10.1	7.8	9.6	9.5						

假设当科研经费投入水平不同时，生产能力提高量均服从正态分布，且方差相等。则由上面的样本数据判断科研经费的高低是否对生产能力提高量有影响，取显著性水平 $\alpha=0.05$。(提示：低档的偏差平方和是 5.3，中档的偏差平方和是 6.3，高档的偏差平方和是 3.7，总平方和是 35.5)

解：用方差分析　$H_0:\ a_1=a_2=a_3$

$\chi_0=\left\{\frac{(n-r)Q_A}{(r-1)Q_e}>F_{1-\alpha}(r-1,n-r)\right\}$ （$r=3$ 组，$n=27$ 家）$\alpha=0.05$

$\frac{(n-r)Q_A}{(r-1)Q_e}=\frac{24Q_A}{2Q_e}=12\frac{Q_A}{Q_e}=\frac{12\times20.2}{15.3}\approx3.4$ 拒绝

$9\times5.3+6.3\times12+5.7\times6$

$Q_e=\sum_{i=1}^{3}n_iS_i^2=5.3+6.3+3.7=15.3$　$Q=35.5$

$Q_A=Q-Q_e=35.5-15.3=20.2$

$F_{0.95}(2,24)=3.4$

∴ $a_1=a_2=a_3$ 相等不成立，所以科研经费的高低对生产能力提高有影响。

院		班级		学号		姓名

密　封　线

七.(15分)已知随机变量 y 关于 x 具有线性关系 $y=\beta_0+\beta_1 x+\varepsilon, \varepsilon\sim N(0,\sigma^2)$，现有5组相互独立的观测数据：

x	2	2	4	5	7	4
y	2	3	7	8	10	6

求

(1) β_0、β_1、σ^2 的最小二乘估计；

(2) 检验 $H_0:\beta_1=0 \leftrightarrow H_1:\beta_1\neq 0$；取显著性水平 $\alpha=0.05$；

(3) 若接受(2)中的原假设，请说明 y 和 x 之间有什么样的结论；若拒绝(2)中的原假设，请给出当 $x_0=5$ 时，y_0 的置信度为95%的双侧置信区间。

解 1) $X=\begin{pmatrix}1&2\\1&2\\1&4\\1&5\\1&7\end{pmatrix}$ $Y=\begin{pmatrix}2\\3\\7\\8\\10\end{pmatrix}$. $X'X\begin{pmatrix}\hat\beta_0\\\hat\beta_1\end{pmatrix}=X'Y$

$L_{xy}=\sum_{t=1}^{5}(x_t-4)(y_t-6)$. $\bar{x}=4$ $\bar{y}=6$

$=(2-4)(2-6)+(2-4)(3-6)+(4-4)(7-6)+(5-4)(8-6)+(7-4)(10-6)$

$=28$

$L_{xx}=\sum_{t=1}^{5}(x_t-\bar{x})^2=18$

$L_{yy}=\sum_{t=1}^{5}(y_t-\bar{y})^2=46$

$\hat\beta=\begin{pmatrix}\bar{y}-\bar{x}L_{xy}/L_{xx}\\L_{xy}/L_{xx}\end{pmatrix}=\begin{pmatrix}6-\frac{4\times 28}{18}\\\frac{28}{18}\end{pmatrix}=\begin{pmatrix}-\frac{2}{9}\\\frac{14}{9}\end{pmatrix}=\begin{pmatrix}\hat\beta_0\\\hat\beta_1\end{pmatrix}$

$\hat\sigma_e^2=\frac{Q_e}{n-k-1}=\frac{L_{yy}-\hat\beta_1^2L_{xx}}{5-1-1}=\frac{22}{27}$

由 $\bar{x}=4$
$\bar{y}=6$
$L_{xy}=28$
$L_{xx}=18$
$L_{yy}=46$
~~$\hat\beta_0=\bar{y}-\hat\beta_1$~~
$\hat\beta_1=\frac{28}{18}=\frac{14}{9}$
$\hat\beta_0=6-\frac{14}{9}\times 4=-\frac{2}{9}$

4

2) $\left|\frac{\hat\beta_1}{\sqrt{Q_e/(n-k-1)}}\right|$ ... $t_{1-\alpha/2}(n-k-1)$

$U_R=\hat\beta_1^2 L_{xx}$ 左边 $=13.4$ 右边 $=10.01$ (拒绝)

$\beta_1\neq 0$.

3) $\hat{y}_1=\bar{y}+\hat\beta_1(x-\bar{x})-\hat\sigma_e t_{1-\alpha/2}(n-2)\sqrt{1+\frac{1}{n}+\frac{(x-\bar{x})^2}{L_{xx}}}$

$\hat{y}_2=\bar{y}+\hat\beta_1(x-\bar{x})+\hat\sigma_e t_{1-\alpha/2}(n-2)\sqrt{1+\frac{1}{n}+\frac{(x-\bar{x})^2}{L_{xx}}}$