# 华东理工大学 2012-2013 学年第 1 学期

# 《数理统计方法》课程考试 B 卷 2013 年 1 月

开课学院：理学院， 考试形式：闭 卷，所需时间：120 分钟

考生姓名：______ 学号：______ 任课教师：朱坤平

| 题序 | 一 | 二 | 三 | 四 | 五 | 六 | 七 | 总分 |
|---|---|---|---|---|---|---|---|---|
| 得分 | | | | | | | | |
| 评卷人 | | | | | | | | |

一. 选择题 (每小题 3 分,共 30 分)

1. 三个因子,每个因子都是 2 个水平.若考虑一级交互作用,应选取的正交表为 ( D )

（A） $L_9(3^4)$ （B）$L_{27}(3^{13})$

（C） $L_4(2^3)$ （D）$L_8(2^7)$

2. 对总体 $\xi$ 观测 4 次得到的样本观测值分别为 2, 1,1, 2, 则错误的选项是 ( B )

(A) 样本中位数 = 样本均值 （B）样本方差 = 样本极差

（C）修正样本标准差为 $\frac{\sqrt{3}}{3}$ （D）样本经验分布函数为 $F_n(x)=\begin{cases}0 & x<1\\ 0.5 & 1\le x<2\\ 1 & 2\le x\end{cases}$

3. 关于标准正态分布的分位数 $u_\alpha$ （$0<\alpha<1$, $\Phi(.)$为标准正态分布的分布函数）,错误的选项是（ A ）

（A） $u_\alpha+u_{1-\alpha}=1$ （B）$\Phi(u_\alpha)=\alpha$

（C） $\Phi(-u_\alpha)=1-\alpha$ （D）若$\alpha_1<\alpha_2$,则 $u_{\alpha_1}\le u_{\alpha_2}$

4. 设 $(X_1,X_2,\cdots,X_n)$ 是总体 $\xi$ 的样本，$\xi\sim N(\mu,\sigma_0{}^2)$，其中 $\sigma_0{}^2$ 已知，参数 $\mu$ 的置信水平为1-$\alpha$ 的置信区间 L 包含 $\mu_0$，则显著性水平 $\alpha$ 下,对原假设 $H_0:\mu=\mu_0$ 的假设检验,有（ C ）

（A）不能确定是否接受 $H_0$ （B）拒绝 $H_0$

（C）接受 $H_0$ （D）犯第二类错误的概率为 1-$\alpha$

5. 设总体 $X \sim N(0,1)$， $X_1$，$X_2,\cdots$，$X_6$ 为样本，又设

$Y=(X_1+X_3+X_5)^2+(X_2+X_4+X_6)^2$，且 $CY \sim \chi^2$ 分布，则 C 的值为（ C ）

(A) 1　　(B) $\frac{1}{2}$　　(C) $\frac{1}{3}$　　(D) $\frac{1}{6}$

6. 设( $X_1,X_2,\cdots,X_n$ )为总体 $N(\mu,\sigma^2)$ ( $\mu$ 未知)的一个样本，$\overline{X}$ 为样本均值，则总体方差 $\sigma^2$ 的无偏估计量的是( B )

（A） $\hat{\sigma}_1^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$　　（B）$\hat{\sigma}_2^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$

（C） $\hat{\sigma}_3^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2$　　（D）$\hat{\sigma}_4^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\mu)^2$

7. 在显著性水平 $\alpha$ 下的某假设检验，如果原假设 $\mathrm{H}_0$ 被拒绝，则（ A ）

（A） 在比 $\alpha$ 更大的显著性水平下检验，原假设 $\mathrm{H}_0$ 一定被拒绝

（B） 在比 $\alpha$ 更大的显著性水平下检验，原假设 $\mathrm{H}_0$ 一定被接受

（C） 在比 $\alpha$ 更大的显著性水平下检验，不能确定原假设 $\mathrm{H}_0$ 是否被拒绝

（D） 在比 $\alpha$ 更小的显著性水平下检验，原假设 $\mathrm{H}_0$ 一定被拒绝

8. 设 $(X_1,X_2,\cdots,X_m)$，$(Y_1,Y_2,\cdots,Y_n)$ 分别是取自相互独立的正态总体

$\xi \sim N(\mu_1,\sigma_1^2)$ 和 $\eta \sim N(\mu_2,\sigma_2^2)$ 的样本，则（ D ）

（A） $\overline{X}$，$\overline{Y}$，$S_x^2$，$S_y^2$ 相互独立　　（B） $\frac{\overline{X}-\mu_1}{S_x^*}\sqrt{m} \sim t$（m-1）

（C） $\frac{S_x^{*2}/\sigma_1^2}{S_y^{*2}/\sigma_2^2} \sim F(m-1,n-1)$　　（D） 以上选项全对

9. 可以克服线性回归中的复共线性问题的方法是（ A ）

（A） 减少自变元个数　　（B） 增加自变元个数

（C） 减少样本容量　　（D） 增加样本容量

10. 在多元线性回归中, 若减少了自变元的个数,则（ C ）

（A） 总离差平方和会增加　　（B） 回归平方和会增加

（C） 残差平方和会增加　　（D） 判定系数 $R^2$ 会提高

二.（12 分）设总体 $X$ 的概率分布列为：

| $X$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P$ | $(1-\lambda)^2$ | $2\lambda(1-\lambda)$ | $(1-\lambda)^2$ | $2\lambda-1$ |

其中 $\lambda$ $(0<\lambda<\frac{1}{2})$ 是未知参数. 令 $p=1-\lambda$, $(X_1, X_2, ..., X_n)$ 为总体 $X$ 的样本.

1) 若样本观测值为：(1， 3， 0， 2， 3， 3， 1， 3)

求 $p$ 的矩法估计值和 $p$ 的极大似然估计值;

2) 证明 $p$ 的矩法估计量是 $p$ 的无偏估计

解: 1） $EX=0\cdot p+2p(1-p)+2p^2+3(1-2p)=3-4p$

$\overline{X}=\frac{1}{8}\sum_{i=1}^{8}X_i=2$

令 $EX=\overline{X}$，得 $p$ 的矩法估计 $\hat{p}=\frac{1}{4}$ ---4”

$$L(p)=\prod_{i=1}^{8}P(X_i=x_i)=P(X=0)[P(X=1)]^2P(X=2)[P(X=3)]^4$$

$$=4p^6(1-p)^2(1-2p)^4$$

$$\ln(L(P))=\ln 4+6\ln p+2\ln(1-p)+4\ln(1-2p)$$

令 $[\ln(L(p))]'=\frac{6}{p}-\frac{2}{1-p}-\frac{8}{1-2p}=0$

$\Rightarrow 12p^2-14p+3=0$ 得到 $p_1=\frac{7-\sqrt{13}}{12}$， $p_1=\frac{7+\sqrt{13}}{12}$ （舍去）

所以 $p$ 的极大似然估计为 $\hat{p}=\frac{7-\sqrt{13}}{12}=0.2828$ ---4”

2) 由 1)知 $p$ 的矩法估计： $\hat{p}=\frac{3-\overline{X}}{4}$ (由 $3-4p=\overline{X}$ 得)

故 $E\hat{p}=\frac{3-E\overline{X}}{4}=\frac{3}{4}-\frac{1}{4}\mathrm{EX}=\frac{3}{4}-\frac{1}{4}(3-4p)=p$,

故 $p$ 的矩法估计是无偏的 ---4”

三.(10 分) 据某校的一份关于是否想去电影院买票看影片<泰囧>的抽样调查:

| 性别 \ 意向 | 想去 | 不想去 |
|---|---|---|
| 男 | 32 | 118 |
| 女 | 20 | 130 |

在显著性水平 $\alpha$=0.05 下,能否认为学生是否想看这个电影的意向与性别有关?

解: $H_0$: 意向与性别独立

$$\chi^2 = n(\sum_{i,j}\frac{n^2_{ij}}{n_{i.}n_{.j}}-1) = 300(\frac{32^2}{52*150}+\frac{118^2}{248*150}+\frac{20^2}{52*150}+\frac{130^2}{248*150}-1) = 3.3498$$

$$\chi^2 < \chi^2_{0.95}(1) = 3.841 .$$

故接受 $H_0$, 可以认为意向与性别无关. -----1'+6'+3'

四.(14 分)对 5 块小麦试验田的施肥量 $x$ 和亩产量 $y$ 分别进行测量,得到数据如下:

| | | | | | |
|---|---|---|---|---|---|
| 施肥量 $x_i$ | 7.7 | 8.0 | 8.4 | 8.8 | 9.6 |
| 亩产量 $y_i$ | 128 | 131 | 134 | 146 | 158 |

设有 $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ , $\varepsilon_i \sim N(0,\sigma^2)$ ($i = 1, 2, \cdots, 5$),

且 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_5$ 相互独立。

1)求 $\beta_0, \beta_1$ 的最小二乘估计 $\hat{\beta}_0, \hat{\beta}_1$ ,并写出回归方程;

2)计算残差平方和 $SS_e$ ,估计的标准差 $\hat{\sigma}$ ,样本相关系数 $r$ ;

3)检验 $H_0$: $\beta_1 = 0$ (显著水平 $\alpha = 0.05$)

解: $n=5$ , $\bar{x}=8.5$ , $L_{xx}=2.2$ , $\bar{y}=139.4$ , $L_{yy}=619.2$ , $L_{xy}=36.3$ 。

1) $\hat{\beta}_1 = \frac{L_{xy}}{L_{xx}} = \frac{36.3}{2.2} = 16.5$ , $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x} = 139.4 - 16.5\times 8.5 = -0.85$ 。

所以,回归方程为 $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x = -0.85 + 16.5x$ 。 ---4'

2) $SS_e = L_{yy} - \hat{\beta}_1 L_{xy} = 619.2 - 16.5\times 36.3 = 20.25$ ,

$$\hat{\sigma}=\sqrt{\frac{SS_e}{n-2}}=\sqrt{\frac{20.25}{5-2}}=\sqrt{6.75}=2.598$$ ，

$$r=\frac{L_{xy}}{\sqrt{L_{xx}L_{yy}}}=\frac{36.3}{\sqrt{2.2\times 619.2}}=0.98351$$ 。 ---6'

3）（方法一）用 $t$ 分布检验：

$$T=\frac{\hat{\beta}_1}{\hat{\sigma}}\sqrt{L_{xx}}=\frac{16.5}{2.598}\times\sqrt{2.2}=9.420$$ 。

对 $\alpha=0.05$，得 $t_{1-\alpha/2}(n-2)=t_{0.975}(3)=3.1824$，

因为 $|T|=|9.420|=9.420>3.1824$， 所以拒绝 $H_0$：$\beta_1=0$ ，

说明自变量 $x$ 与因变量 $y$ 之间有显著的统计线性相关关系。

（方法二）用 F 检验:

$$F=\frac{SSR/1}{SSe/(n-2)}=\frac{SST-SSe}{SSe/(n-2)}=\frac{619.2-20.25}{20.25/3}=88.7333$$

$$F>F_{1-\alpha}(1,n-2)=F_{0.95}(1,3)=10.1$$

所以拒绝 $H_0$：$\beta_1=0$ ，

说明自变量 $x$ 与因变量 $y$ 之间有显著的统计线性相关关系。 -----4'

五.(12分)设某种针织品在70°C和80°C时的强度为 $\xi\sim N(\mu_1,\sigma_1^2)$ 和 $\eta\sim N(\mu_2,\sigma_2^2)$，

在70°C时，抽取6个样品，测得样本均值 $\overline{X}=20.3$ ，**修正样本标准差** $S_x{*}=0.57$ ；

在80°C时，抽取5个样品，测得样本均值 $\overline{Y}=19.5$ ，**修正样本标准差** $S_y{*}=0.75$ 。

（1）分别求出参数 $\sigma_1^2$ 和 $\mu_2$ 置信水平为95% 的置信区间;

（2）能否认为70°C和80°C时针织品强度的方差相等（显著水平 $\alpha=0.05$）？

解 1） 参数 $\sigma_1^2$ 置信水平为95% 的置信区间为: ------3'+3'

$$\left[\frac{(n-1)S_x^{*^2}}{\chi^2_{1-\frac{\alpha}{2}}(n-1)},\quad \frac{(n-1)S_x^{*^2}}{\chi^2_{\frac{\alpha}{2}}(n-1)}\right]=[\frac{5*0.57^2}{12.833},\quad \frac{5*0.57^2}{0.831}]=[0.1266,\quad 1.9549]$$

$\mu_2$ 置信水平为95% 的置信区间为:

$$\overline{Y}\pm t_{1-\frac{\alpha}{2}}(n-1)\frac{S_y^*}{\sqrt{n}}=19.5\pm t_{0.975}(4)\frac{0.75}{\sqrt{5}}=[18.5687,\qquad 20.4313]$$

2）问题相当于要检验 $H_0$：$\sigma_1^2=\sigma_2^2$ --------6’

$$F=\frac{S_x{*}^2}{S_y{*}^2}=\frac{0.57^2}{0.75^2}=0.5776 \quad 。$$

对$\alpha=0.05$，查$F$分布表，可得

$$F_{1-\alpha/2}(m-1,n-1)=F_{0.975}(5,4)=9.36 \quad ,$$

$$F_{\alpha/2}(m-1,n-1)=\frac{1}{F_{1-\alpha/2}(n-1,m-1)}=\frac{1}{F_{0.975}(4,5)}=\frac{1}{7.39}=0.135 \quad ,$$

因为 $0.135<F=0.5776<9.36$ ，所以接受 $H_0$：$\sigma_1^2=\sigma_2^2$ 。

六.（10分）随机抽取甲、乙、丙三厂生产的某型号锂电池各三件，测量它们充满电后在手机上使用的待机时间，得到数据如下：

| 生产厂 | 锂电池的待机时间（单位：小时） |
|---|---|
| 甲厂 | 36 35 31 |
| 乙厂 | 40 38 42 |
| 丙厂 | 39 43 47 |

问这三厂生产的锂电池其待机时间是否有显著差异（$\alpha=0.05$）？

**解** 设三家厂生产的锂电池的待机时间分别为相互独立的 $\xi_i\sim N(\mu_i,\sigma^2)$，$i=1,2,3$ 。

检验三种锂电池的待机时间是否有显著差异，相当于要检验假设 $H_0$：$\mu_1=\mu_2=\mu_3$

计算结果见下表：

| 水平 | 观测值（使用寿命） | $y_i$ | $SS_i$ |
|---|---|---|---|
| $A_1$ | 36 35 31 | 34 | 14 |
| $A_2$ | 40 38 42 | 40 | 8 |
| $A_3$ | 39 43 47 | 43 | 32 |
| | | $SS_A=126$ | $SS_e=54$ |

方差分析表为：

| 来源 | 平方和 | 自由度 | 均方 | $F$ 值 | 分位数 |
|---|---|---|---|---|---|
| $A$ | $SS_A=126$ | $r-1=2$ | 63 | $F_A=7$ | $F_{0.95}(2,6)=5.14$ |
| 误差 | $SS_e=54$ | $n-r=6$ | 9 | | |

| | | | | | |
|---|---|---|---|---|---|
| 总和 | $SS_T=180$ | $n-1=8$ | | | |

因为$F_A=7>5.14=F_{0.95}(2,6)$ ，所以拒绝 $H_0$： $\mu_1=\mu_2=\mu_3$ ，

结论是：三家厂生产的电池使用寿命有显著的差异。 ------5*2’

七.（12 分）在某细沙机上进行断头率的测定，试验锭子总数 300 个，测得结果如下：

| 各锭的断头数 | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| 对应的锭数 | 120 | 90 | 60 | 30 |

问各锭子的断头数是否服从泊松分布（$\alpha$ =0.05）？

解 $H_0$：锭子的断头数 X 服从泊松分

$\hat{\lambda}=\bar{x}=1$ --------2’

作分点 [-0.5，0.5）, [0.5，1.5) , [1.5, 2.5) , [2.5, +∞）则：

$$\hat{p}_1=P\{X\in[-0.5,0.5)\}=P\{X=0\}=\frac{\hat{\lambda}^0}{0!}e^{-\hat{\lambda}}=e^{-1}$$

$$\hat{p}_2=P\{X\in[0.5,1.5)\}=P\{X=1\}=\frac{\hat{\lambda}^1}{1!}e^{-\hat{\lambda}}=e^{-1}$$

$$\hat{p}_3=P\{X\in[1.5,2.5)\}=P\{X=2\}=\frac{\hat{\lambda}^2}{2!}e^{-\hat{\lambda}}=e^{-1}/2$$

$$\hat{p}_4=1-\hat{p}_1-\hat{p}_2-\hat{p}_3=1-2.5e^{-1}$$ -------4*1’

$$\chi^2=\frac{1}{\mathrm{n}}\sum_{\mathrm{i=1}}^{\mathrm{r}}\frac{n_k^2}{\hat{p}_k}-n\sim\chi^2(r-m-1)$$

$$\chi^2=\frac{1}{300}[\frac{120^2}{\mathrm{e}^{-1}}+\frac{90^2}{\mathrm{e}^{-1}}+\frac{60^2}{0.5\mathrm{e}^{-1}}+\frac{30^2}{1-2.5\mathrm{e}^{-1}}]-300=6.48$$ ---------4’

而 $\chi^2_{1-\alpha}(r-m-1)=\chi^2_{0.95}(2)=5.991$

$\chi^2$〉$\chi^2_{0.95}(2)=5.991$，故拒绝原假设，认为锭子的断头数不服从泊松分布。

-----------2’