考试时间2019年1月22日(9:30-11:30)

改完卷已放假,查卷时间只能在下学期开学初

---试卷答题请步骤完整(要写上公式)---

1)考试可以使用计算器

2)可带半张A4自己整理的公式纸

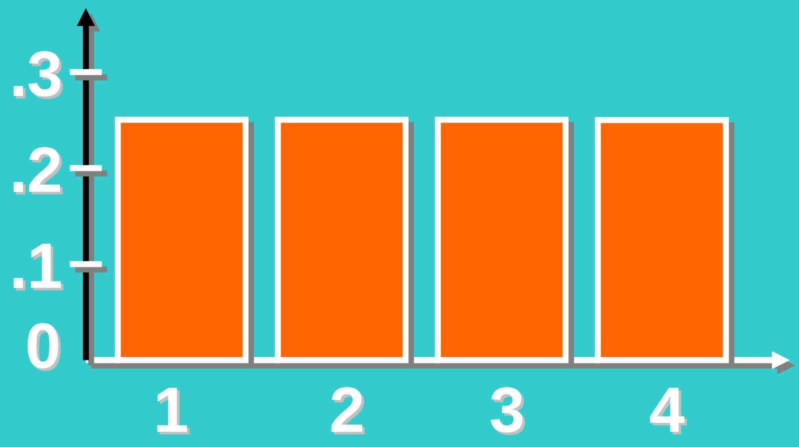
3)考前的21日晚18:00-20:00在A教二楼有答疑

4)样卷见附件,以下为作业及答疑常见问题,与考试无关

总体分布,样本的经验分布, 联合分布,抽样分布

例:设一个总体，含有4个元素(个体)，即总体单位数N=4。4 个个体分别为x_1=1、x_2=2、x_3=3 、x_4=4 。总体X的分布及期望和方差如下

X	1	2	3	4
P	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$



$$E(X)=\mu=\sum_{i=1}^{N}x_i p_i=\frac{1}{4}(1+2+3+4)=2.5$$

$$D(X)=\sigma^2=\sum_{i=1}^{N}(x_i-\mu)^2 p_i=1.25$$

现从总体中抽取$n=2$的简单随机样本，在重复抽样条件下，共有$4^2=16$个样本。样本(X1,X2)所有的结果为:

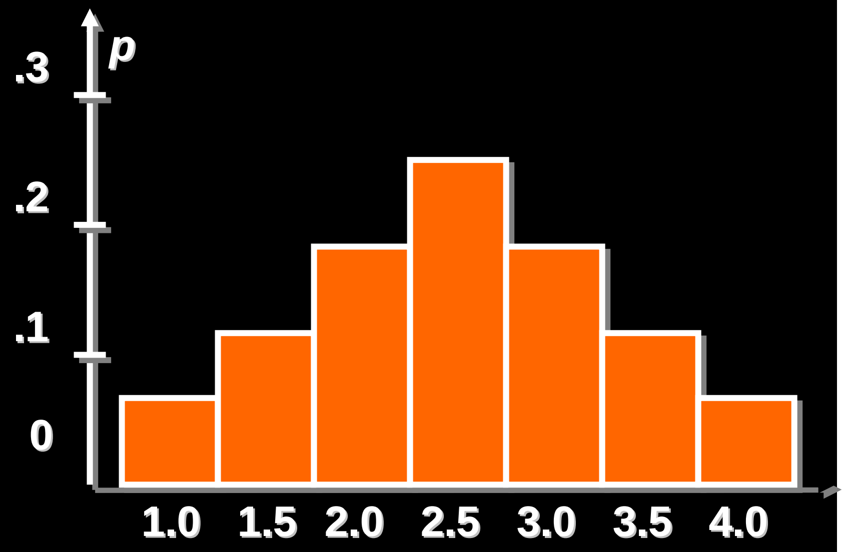
所有可能的n = 2 的样本(共16个)				
第一个观察值	第二个观察值			
	1	2	3	4
1	(1,1)	(1,2)	(1,3)	(1,4)
2	(2,1)	(2,2)	(2,3)	(2,4)
3	(3,1)	(3,2)	(3,3)	(3,4)
4	(4,1)	(4,2)	(4,3)	(4,4)

计算出各样本的均值，如下表。并给出样本均值的抽样分布:

$\bar{X}$	1	1.5	2	2.5	3	3.5	4
P	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{3}{16}$	$\frac{4}{16}$	$\frac{3}{16}$	$\frac{2}{16}$	$\frac{1}{16}$

16个样本的均值

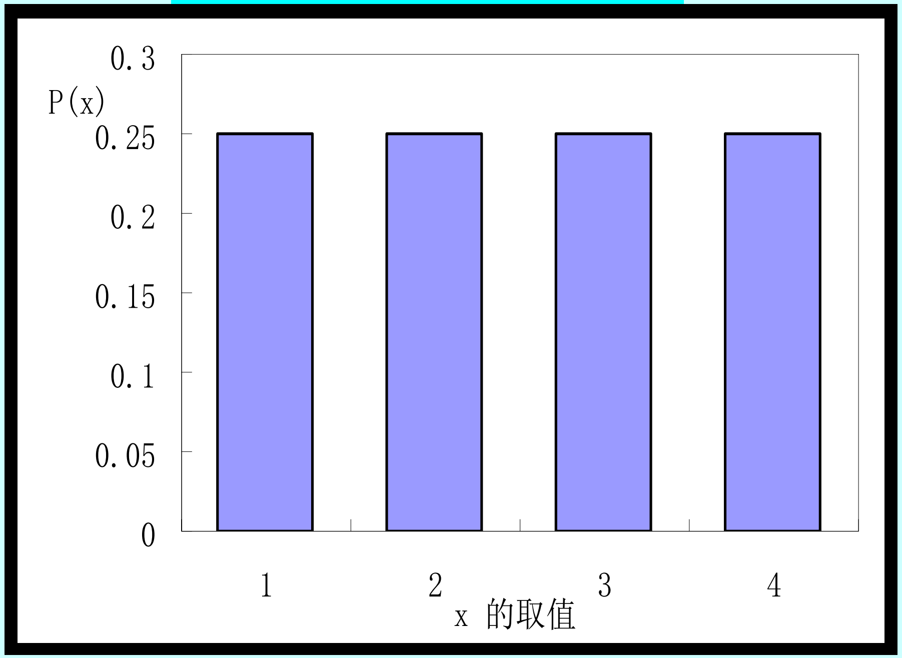
第一个观察值	第二个观察值			
	1	2	3	4
1	1.0	1.5	2.0	2.5
2	1.5	2.0	2.5	3.0
3	2.0	2.5	3.0	3.5
4	2.5	3.0	3.5	4.0



样本均值的抽样分布

总体的期望和方差　与　样本均值的期望和方差　的比较

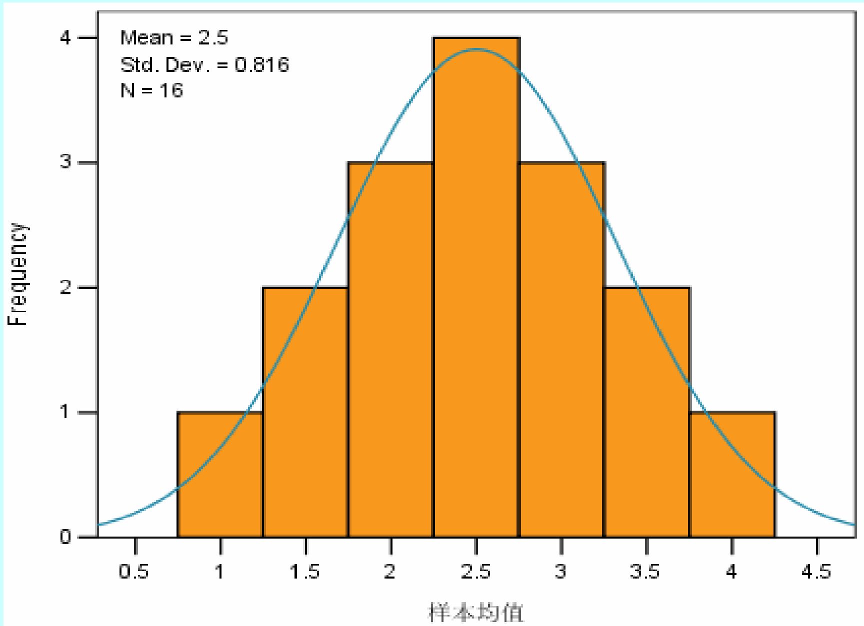
总体分布



$$\mu = 2.5$$

$$\sigma^2 = 1.25$$

样本均值分布



$$\mu_{\bar{x}} = 2.5$$

$$\sigma_{\bar{x}}^2 = 0.625 = 1.25 / 2 = \sigma^2 / n$$

2.8 设 $(X_1,X_2,\cdots,X_m)$ 是 $\xi\sim N(0,\sigma^2)$ 的样本，$(Y_1,Y_2,\cdots,Y_n)$ 是 $\eta\sim N(0,\sigma^2)$ 的样本，两个样本相互独立，证明：（1）$\dfrac{\sum_{i=1}^{m}X_i^2+\sum_{j=1}^{n}Y_j^2}{\sigma^2}\sim\chi^2(m+n)$；（2）$\dfrac{\sum_{i=1}^{m}X_i}{\sqrt{\sum_{j=1}^{n}Y_j^2}}\sqrt{\dfrac{n}{m}}\sim t(n)$。

证 （1）因为 $(X_1,X_2,\cdots,X_m)$ 是 $\xi\sim N(0,\sigma^2)$ 的样本，所以 $X_i\sim N(0,\sigma^2)$，$i=1,2,\cdots,m$，$X_1,X_2,\cdots,X_m$ 相互独立。即有 $\dfrac{X_i}{\sigma}\sim N(0,1)$，$i=1,2,\cdots,m$，$\dfrac{X_1}{\sigma},\dfrac{X_2}{\sigma},\cdots,\dfrac{X_m}{\sigma}$ 相互独立。

由 χ^2 分布定义可知 $\dfrac{\sum_{i=1}^{m}X_i^2}{\sigma^2}=\sum_{i=1}^{m}\left(\dfrac{X_i}{\sigma}\right)^2\sim\chi^2(m)$。同理可证 $\dfrac{\sum_{j=1}^{n}Y_j^2}{\sigma^2}\sim\chi^2(n)$。

而且由于两个样本相互独立，所以 $\dfrac{\sum_{i=1}^{m}X_i^2}{\sigma^2}$ 与 $\dfrac{\sum_{j=1}^{n}Y_j^2}{\sigma^2}$ 相互独立。

因此，由 χ^2 分布的可加性（定理 2.3）可知 $\dfrac{\sum_{i=1}^{m}X_i^2+\sum_{j=1}^{n}Y_j^2}{\sigma^2}\sim\chi^2(m+n)$

（2） 因为 $(X_1,X_2,\cdots,X_m)$ 是 $\xi\sim N(0,\sigma^2)$ 的样本，所以 $X_i\sim N(0,\sigma^2)$，$i=1,2,\cdots,m$ ，而且相互独立。因此有 $\sum_{i=1}^{m}X_i\sim N(0,m\sigma^2)$，所以 $\dfrac{\sum_{i=1}^{m}X_i-0}{\sqrt{m\sigma^2}}=\dfrac{1}{\sigma\sqrt{m}}\sum_{i=1}^{m}X_i\sim N(0,1)$ 。

同时，在上面（1）中已经证得 $\dfrac{\sum_{j=1}^{n}Y_j^{\,2}}{\sigma^2}\sim\chi^2(n)$，而且由于两个样本相互独立，所以 $\dfrac{1}{\sigma\sqrt{m}}\sum_{i=1}^{m}X_i$ 与 $\dfrac{\sum_{j=1}^{n}Y_j^{\,2}}{\sigma^2}$ 相互独立。

因此，由 t 分布的定义可知 $\dfrac{\sum_{i=1}^{m}X_i}{\sqrt{\sum_{j=1}^{n}Y_j^2}}\sqrt{\dfrac{n}{m}}=\dfrac{\dfrac{1}{\sigma\sqrt{m}}\sum_{i=1}^{m}X_i}{\sqrt{\dfrac{\sum_{j=1}^{n}Y_j^{\,2}}{\sigma^2}\Big/n}}\sim t(n)$ 。

分析 根据 t 分布的定义，$\dfrac{\boxed{N(0,1)}}{\sqrt{\boxed{\chi^2(n)}\big/n}}\sim t(n)$ 。。

3.5 设总体 ξ 服从 $[a,b]$ 上的均匀分布，概率密度为 $\varphi(x)=\begin{cases}1/(b-a) & a\le x\le b\\ 0 & \text{其他}\end{cases}$，其中，$a<b$ 是未知参数，$(X_1,X_2,\cdots,X_n)$ 是 ξ 的样本，求：（1） a,b 的矩法估计；（2） a,b 的极大似然估计。

解 （1） $E\xi=\int_{-\infty}^{+\infty}x\varphi(x)\mathrm{d}x=\int_a^b x\dfrac{1}{b-a}\mathrm{d}x=\dfrac{a+b}{2}$，$E(\xi^2)=\int_{-\infty}^{+\infty}x^2\varphi(x)\mathrm{d}x=\int_a^b x^2\dfrac{1}{b-a}\mathrm{d}x=\dfrac{b^2+ab+a^2}{3}$。

解方程 $\begin{cases}\dfrac{\hat a+\hat b}{2}=\hat{E}\xi=\overline{X} & (1)\\[2mm] \dfrac{\hat b^2+\hat a\hat b+\hat a^2}{3}=\hat{E}(\xi^2)=\overline{X^2} & (2)\end{cases}$。

$(2)-(1)^2$：$\dfrac{(\hat b-\hat a)^2}{12}=\dfrac{\hat b^2+\hat a\hat b+\hat a^2}{3}-(\dfrac{\hat a+\hat b}{2})^2=\overline{X^2}-\overline{X}^2=\dfrac{1}{n}\sum_{i=1}^n X_i^2-\overline{X}^2=S^2$，

两边开方：$\dfrac{\hat b-\hat a}{2\sqrt3}=\pm\sqrt{S^2}=\pm S$，即 $\dfrac{\hat b-\hat a}{2}=\pm\sqrt3\,S\quad(3)$，

$(1)-(3)$得：$\hat a=\dfrac{\hat a+\hat b}{2}-\dfrac{\hat b-\hat a}{2}=\overline{X}\mp\sqrt3\,S$，$(1)+(3)$得：$\hat b=\dfrac{\hat a+\hat b}{2}+\dfrac{\hat b-\hat a}{2}=\overline{X}\pm\sqrt3\,S$，

即有 $\begin{cases}\hat a=\overline{X}-\sqrt3\,S\\ \hat b=\overline{X}+\sqrt3\,S\end{cases}$ 或 $\begin{cases}\hat a=\overline{X}+\sqrt3\,S\\ \hat b=\overline{X}-\sqrt3\,S\end{cases}$。因为$a<b$，第二组解应该舍去，所以矩法估计为 $\begin{cases}\hat a=\overline{X}-\sqrt3\,S\\ \hat b=\overline{X}+\sqrt3\,S\end{cases}$。

(2) 似然函数

$$L=\prod_{i=1}^{n}\varphi(x_i)=\begin{cases}\prod_{i=1}^{n}\dfrac{1}{b-a} & a\le x_i\le b\ ,i=1,2,\cdots,n\\ 0 & \text{其他}\end{cases}=\begin{cases}\dfrac{1}{(b-a)^n} & a\le\min_i x_i\le\max_i x_i\le b\\ 0 & \text{其他}\end{cases}。$$

当 $L\neq 0$ 时，对 L 取对数，得到 $\ln L=-n\ln(b-a)$ 。

求导，列方程 $\begin{cases}\dfrac{\partial\ln L}{\partial a}=\dfrac{n}{b-a}=0\\ \dfrac{\partial\ln L}{\partial b}=-\dfrac{n}{b-a}=0\end{cases}$ ，显然无解。

但从 $L=\dfrac{1}{(b-a)^n}$ 可看出，它要达到最大，a 要尽可能大，但 a 不能大于 $\min_i x_i$，b 要尽可能小，但 b 不能小于 $\max_i x_i$，可见，当且仅当 $a=\min_i x_i$，$b=\max_i x_i$ 时 $L=\dfrac{1}{(b-a)^n}$ 取到最大值。

所以极大似然估计为 $\begin{cases}\hat{a}_L=\min_i X_i\\ \hat{b}_L=\max_i X_i\end{cases}$ 。

3.10　设总体 ξ 的概率分布为

ξ	0	1	2
$P\{\xi=k\}$	$1-3\theta$	θ	2θ

其中，θ（$0<\theta<\frac{1}{3}$） 是未知参数，利用总体 ξ 的如下样本观测值

1，0，1，2，1，

求 θ 的矩法估计值和极大似然估计值。

解　（1） $E\xi=\sum_{k=0}^{2}kP\{\xi=k\}=0\times(1-3\theta)+1\times\theta+2\times2\theta=5\theta$ ，$\bar{x}=\frac{1+0+1+2+1}{5}=\frac{5}{5}=1$ 。

解方程 $5\hat{\theta}=\hat{E}\xi=\bar{x}=1$ ，得到 θ 的矩法估计值 $\hat{\theta}=\frac{1}{5}$ 。

（2） 似然函数 $L=\prod_{i=1}^{n}P\{\xi=x_i\}=(1-3\theta)^1\times\theta^3\times(2\theta)^1=2(1-3\theta)\theta^4$ 。$\ln L=\ln 2+\ln(1-3\theta)+4\ln\theta$ ，

解方程 $\frac{\mathrm{d}\ln L}{\mathrm{d}\theta}=-\frac{3}{1-3\theta}+\frac{4}{\theta}=\frac{-3\theta+4-12\theta}{(1-3\theta)\theta}=\frac{4-15\theta}{(1-3\theta)\theta}=0$ 得到 θ 的极大似然估计值 $\hat{\theta}_L=\frac{4}{15}$ 。

4.1 设总体 $\xi \sim N(\mu,\sigma^2)$，已知其中 $\sigma=\sigma_0$，$(X_1,X_2,\cdots,X_n)$ 是 ξ 的样本，μ 的置信水平为 $1-\alpha$ 的置信区间为 $[\underline{\theta},\bar{\theta}]$，其中 $\underline{\theta},\bar{\theta}=\bar{X}\mp u_{1-\alpha/2}\dfrac{\sigma_0}{\sqrt{n}}$；检验 H_0：$\mu=\mu_0$ 的统计量 $U=\dfrac{\bar{X}-\mu_0}{\sigma_0}\sqrt{n}$。

证明：在显著水平 α 下，拒绝假设 H_0：$\mu=\mu_0$ 的充分必要条件是 $\mu_0\notin[\underline{\theta},\bar{\theta}]$。

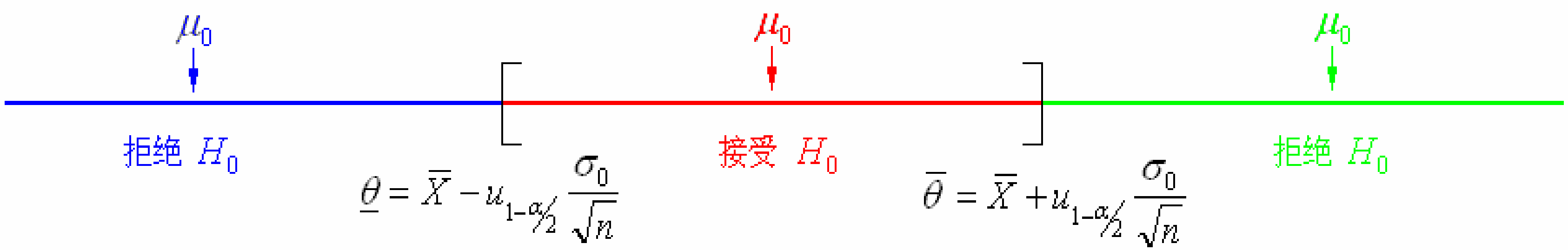
这说明参数的区间估计与参数的假设检验有着一一对应的关系，可以用参数的区间估计代替参数的假设检验。

证 在显著水平 α 下，拒绝假设 H_0：$\mu=\mu_0$ 的充分必要条件是 $|U|>u_{1-\alpha/2}$，其中 $U=\dfrac{\bar{X}-\mu_0}{\sigma_0}\sqrt{n}$，

$$\text{即}\ \left|\frac{\bar{X}-\mu_0}{\sigma_0}\sqrt{n}\right|>u_{1-\alpha/2}\text{，也就是}\ |\bar{X}-\mu_0|>u_{1-\alpha/2}\frac{\sigma_0}{\sqrt{n}}\text{，它等价于}$$

$$\mu_0<\bar{X}-u_{1-\alpha/2}\frac{\sigma_0}{\sqrt{n}}\ \text{或}\ \mu_0>\bar{X}+u_{1-\alpha/2}\frac{\sigma_0}{\sqrt{n}}\text{，}$$

$$\text{即}\ \mu_0\notin[\underline{\theta},\bar{\theta}]\text{，其中}\ \underline{\theta}=\bar{X}-u_{1-\alpha/2}\frac{\sigma_0}{\sqrt{n}}\text{，}\bar{\theta}=\bar{X}+u_{1-\alpha/2}\frac{\sigma_0}{\sqrt{n}}\text{。}$$



4.2 设$(X_1, X_2, ..., X_{16})$是取自正态总体$N(\mu, 1)$的样本，在显著性水平$\alpha = 0.05$下检验$H_0: \mu = 0$, $H_1: \mu \neq 0$.若正态总体的期望的真值为$\mu = 0.08$，求该检验犯第二类错误的概率.

解 犯第二类错误的概率为

$$\beta = P\{接受H_0 \mid H_0为假\} = P\{统计量观测值落入接受域 \mid \mu = 0.08\}$$

$$= P\left\{\frac{\bar{x} - \mu_0}{\sigma_0}\sqrt{n} \in \omega_0 \middle| \mu = 0.08\right\} = P\left\{\left|\frac{\bar{x} - 0}{1}\sqrt{16}\right| \le \mu_{1-\frac{\alpha}{2}} \middle| \mu = 0.08\right\}$$

$$= P\{-1.96 \le 4\bar{x} \le 1.96 \mid \mu = 0.08\} = P\left\{\frac{-1.96}{4} \le \bar{x} \le \frac{1.96}{4} \middle| \mu = 0.08\right\}$$

$$= P\left\{\frac{\frac{-1.96}{4} - 0.08}{\sqrt{\frac{1^2}{16}}} \le \frac{\bar{x} - 0.08}{\sqrt{\frac{1^2}{16}}} \le \frac{\frac{1.96}{4} - 0.08}{\sqrt{\frac{1^2}{16}}} \middle| \mu = 0.08\right\}$$

$$= P\left\{-1.96 - 0.32 \le \frac{\bar{x} - 0.08}{\sqrt{\frac{1^2}{16}}} \le 1.96 - 0.32 \middle| \mu = 0.08\right\}$$

$$= \Phi(1.64) - \Phi(-2.28) = \Phi(1.64) - [1 - \Phi(2.28)] = \Phi(1.64) + \Phi(2.28) - 1$$

$$= 0.9495 + 0.9887 - 1 = 0.9382$$

4.15 某厂从用新、旧工艺生产的灯泡中各取n只测试其寿命，设新、旧工艺生产的灯泡寿命分别为$\xi \sim N\left(\mu_1, \sigma_1^2\right)$和$\eta \sim N\left(\mu_2, \sigma_2^2\right)$，其样本分别记为$\left(X_1, X_2, \ldots, X_n\right)$和$\left(Y_1, Y_2, \ldots, Y_n\right)$.令$Z_i = X_i - Y_i \left(i = 1, 2, \ldots, n\right)$; Z_i的均值为$\bar{Z}$;样本修正方差记为S_z^{*2}.(1)证明$\dfrac{\bar{Z} - \left(\mu_1 - \mu_2\right)}{S_z^*}\sqrt{n} \sim t\left(n-1\right)$;(2)求显著性水平$\alpha$下，$H_0: \mu_1 = \mu_2$的拒绝域.

4.21 从某厂生产的布匹中抽查 50 匹，查得布匹上的疵点数如下：

疵点数	0	1	2	3	≥4
频数	20	16	8	6	0

问:是否可以认为每匹布上的疵点数 ξ 服从 Poisson(普阿松)分布?(显著水平 $\alpha=0.05$)

$$\hat{p}_4=\hat{P}\{2.5<\xi<+\infty\}=1-\hat{P}\{\xi\le 2.5\}=1-\hat{p}_1-\hat{p}_2-\hat{p}_3$$

$$=1-0.36788-0.36788-0.18394=0.08030\text{ 。}$$

$$\chi^2=\frac{1}{n}\sum_{k=1}^{4}\frac{n_k^2}{\hat{p}_k}-n=\frac{1}{50}\times\left(\frac{20^2}{0.36788}+\frac{16^2}{0.36788}+\frac{8^2}{0.18394}+\frac{6^2}{0.08030}\right)-50=1.589\text{ 。}$$

对 $\alpha=0.05$，$1-\alpha=0.95$，自由度 $r-m-1=4-1-1=2$ ，查 χ^2 分布表，可得

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

4.22 从某车床生产的滚珠中，抽取 50 颗，测得它们的直径（单位：mm）落在各区间中的频数为：

区间	(14.14,14.51]	(14.51,14.88]	(14.88,15.25]	(15.25,15.62]	(15.62,15.99]
频数	6	8	20	11	5

已知滚珠直径的样本均值为 $\overline{X}=15.078$ ，样本标准差为 $S=0.428154$ 。

问：滚珠的直径 ξ 是否服从正态分布？（显著水平 $\alpha=0.05$ ）

$$\hat{\mu}=\overline{X}=15.078\ ,\ \hat{\sigma}=S=0.428154\ 。$$

作分点 $-\infty<14.51<14.88<15.25<15.62<+\infty$，把 ξ 的取值范围分成 5 个区间.

总体 ξ 落在各个区间 $(a_{k-1},a_k]$ 中的概率的估计值 $\hat{p}_k$ 可由下式求出：

$$\hat{p}_k=\hat{P}\{a_{k-1}<\xi\le a_k\}=\Phi\left(\frac{a_k-\hat{\mu}}{\hat{\sigma}}\right)-\Phi\left(\frac{a_{k-1}-\hat{\mu}}{\hat{\sigma}}\right)\ 。$$

用本题的数据代入，得计算结果如下：

$(a_{k-1},a_k]$	$(-\infty$, 14.51]	(14.51, 14.88]	(14.88, 15.25]	(15.25, 15.62]	(15.62, $+\infty$)
n_k	6	8	20	11	5
$\hat{p}_k$	0.11682	0.18778	0.26234	0.23495	0.19811

$$\chi^2=\frac{1}{n}\sum_{k=1}^{r}\frac{n_k^2}{\hat{p}_k}-n$$

$$=\frac{1}{50}\times\left(\frac{6^2}{0.09232}+\frac{8^2}{0.22956}+\frac{20^2}{0.33418}+\frac{11^2}{0.24117}+\frac{5^2}{0.10277}\right)-50=2.214\ 。$$

对 $\alpha=0.05$，$r-m-1=2$，查 χ^2 分布表，可得 $\chi^2_{1-\alpha}(r-m-1)=\chi^2_{0.95}(2)=5.991$.

4.26 为研究儿童智力发展与营养的关系，抽查了 950 名小学生，得到统计数据如下：

	智商			
	<80	80~89	90~99	≥100
营养良好	245	228	177	219
营养不良	31	27	13	10

问：儿童的智力发展是否与营养状况有关？（显著水平 $\alpha=0.05$）

$$\chi^2=n\left(\sum_{i=1}^{r}\sum_{j=1}^{s}\frac{n_{ij}^{\ 2}}{n_{i\cdot}\,n_{\cdot j}}-1\right)$$

$$=950\times\left(\frac{245^2}{869\times276}+\frac{228^2}{869\times255}+\frac{177^2}{869\times190}+\frac{219^2}{869\times229}\right.$$

$$\left.+\frac{31^2}{81\times276}+\frac{27^2}{81\times255}+\frac{13^2}{81\times190}+\frac{10^2}{81\times229}-1\right)=9.751$$ 。

对显著水平 $\alpha=0.05$，自由度 $(r-1)(s-1)=(2-1)\times(4-1)=3$，查 χ^2 分布表，可得分位数 $\chi^2_{1-\alpha}((r-1)(s-1))=\chi^2_{0.95}(3)=7.815$ 。

5.7 设 $y_i=\beta_0+\beta_1 x_i+\varepsilon_i$，$\varepsilon_i \sim N(0,\sigma^2)$，$i=1,2,\cdots,n$， $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 相互独立，$\hat{\beta}_0,\hat{\beta}_1$ 是 β_0,β_1 的最小二乘估计。证明：$\mathrm{Cov}(\hat{\beta}_0,\hat{\beta}_1)=0$ 的充分必要条件是 $\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i=0$。

证 因为

$$\mathrm{Cov}(\hat{\beta}_0,\hat{\beta}_1)=\mathrm{Cov}(\bar{y}-\hat{\beta}_1\bar{x},\hat{\beta}_1)=\mathrm{Cov}(\bar{y},\hat{\beta}_1)-\mathrm{Cov}(\hat{\beta}_1,\hat{\beta}_1)\bar{x}$$

$$=0-D(\hat{\beta}_1)\bar{x}=-\frac{\sigma^2}{L_{xx}}\bar{x}$$ 。

5.11 多元线性回归模型中，若先根据变量y和x_i $(i = 1, 2, ..., m)$的观测值

$(y_1, y_2, ..., y_n)^T$和$(x_{1i}, x_{2i}, ..., x_{ni})^T$对变量“标准化”，即令

$$x_i^* = \frac{x_i - \bar{x}_i}{\sqrt{L_{ii}}} \ (i = 1, 2, ..., m); \ y^* = \frac{y - \bar{y}}{\sqrt{L_{yy}}}; \ \hat{y}^* = \frac{\hat{y} - \bar{y}}{\sqrt{L_{yy}}}, \text{其中}$$

$$L_{ii} = \sum_{k=1}^{n} \left(x_{ki} - \bar{x}_i\right)^2; \ \bar{x}_i = \frac{1}{n}\sum_{k=1}^{n} x_{ki}; \ L_{yy} = \sum_{k=1}^{n} \left(y_i - \bar{y}\right).$$

此时再求y^*关于$x_i^*(i = 1, 2, \cdots, m)$的回归称为标准回归.

（1）证明标准回归方程的常数项为零，即$\hat{y}^* = \sum_{i=1}^{m} d_i x_i^*$

（2）证明标准回归的总离差平方和$\tilde{SS}_T = \sum_{i=1}^{n} \left(y_i^* - \bar{y}^*\right)^2 = 1.$

证明（1）设y关于$x_1, x_2, ..., x_m$的线性回归方程为

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + ... + \hat{\beta}_m x_m$$

于是有

$$\frac{\hat{y} - \bar{y}}{\sqrt{L_{yy}}} = \frac{\hat{\beta}_0 + \hat{\beta}_1 x_1 + ... + \hat{\beta}_m x_m - \bar{y}}{\sqrt{L_{yy}}}$$

即

$$\hat{y}^* = \frac{\hat{\beta}_0 + \hat{\beta}_1 \left(x_1^* \sqrt{L_{11}} + \bar{x}_1\right) + ... + \hat{\beta}_m \left(x_m^* \sqrt{L_{mm}} + \bar{x}_m\right) - \bar{y}}{\sqrt{L_{yy}}}$$

$$= \frac{\hat{\beta}_0 + \hat{\beta}_1 \bar{x}_1 + ... + \hat{\beta}_m \bar{x}_m - \bar{y}}{\sqrt{L_{yy}}} + \hat{\beta}_1 \sqrt{\frac{L_{11}}{L_{yy}}} x_1^* + ... + \hat{\beta}_m \sqrt{\frac{L_{mm}}{L_{yy}}} x_m^*$$

$$= 0 + d_1 x_1^* + ... + d_m x_m^*$$

注：上式用到$\bar{y} = \hat{\beta}_0 + \hat{\beta}_1 \bar{x}_1 + ... + \hat{\beta}_m \bar{x}_m$

此外，从上式还可得$d_i = \hat{\beta}_i \sqrt{\frac{L_{ii}}{L_{yy}}}, i = 1, 2, ..., m$

（2）$\tilde{SS}_T = \sum_{i=1}^{n}\left(y_i^* - \bar{y}^*\right)^2 = \sum_{i=1}^{n}\left(y_i^* - 0\right)^2 = \sum_{i=1}^{n}\left(\frac{y_i - \bar{y}}{\sqrt{L_{yy}}}\right)^2$

$$= \frac{1}{L_{yy}} \sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2 = \frac{L_{yy}}{L_{yy}} = 1。$$

6.4 对某厂早，中，晚三班的产量统计如下：

班次	产量
早班	279　334　303　338　198
中班	229　274　310
晚班	210　285　117

问在显著性水平 $\alpha = 0.05$ 下能否认为不同班次的产量无显著性差异？

解 方差分析的前提为早、中、晚班的产量均服从正态分布，相互独立且方差相等，

$\xi_i \sim N\left(\mu_i, \delta^2\right), i = 1, 2, 3.$

$$H_0: \mu_1 = \mu_2 = \mu_3 \qquad F_A = \frac{SS_A/(r-1)}{SS_e/(n-r)} = \frac{14365.53/2}{30453.2/8} = 1.8869$$

查表 $F_{1-\alpha}(r-1, n-r) = F_{0.95}(2, 8) = 4.4589$; $F_A < F_{1-\alpha}(r-1, n-r)$

故接受 H_0，即在显著性水平 $\alpha = 0.05$ 下认为不同班次产量无显著性差异。

方差分析						
差异源	SS	df	MS	F	P-value	F crit
组间	14365.53	2	7182.764	1.886899	0.213154	4.45897
组内	30453.2	8	3806.65			
总计	44818.73	10				

课件及习题全解见课程主页:

http://59.78.108.56/msta/sltj/

建模用到相关数据分析的SPSS操作演示见

http://59.78.108.56/msta/MYCOURSES/SJFX.HTM

考前课程复习的在线答疑(密码你懂的):

"华理-数理统计学习群"

认真复习

祝你考出好成绩！