厦门大学《概率统计 I》期末试卷

_________学院_________系_________年级_________专业

主考教师：　　　　试卷类型：（B 卷）　2018.1.9

$\Phi(1.64)=0.95, \chi^2_{0.025}(24)=39.364, \chi^2_{0.05}(3)=7.81, \chi^2_{0.975}(24)=12.401, t_{0.025}(3)=3.1824,$

$t_{0.01}(18)=2.55, t_{0.05}(4)=2.132, t_{0.025}(3)=3.1824, F_{0.005}(9,9)=6.54, F_{0.05}(2,37)=3.23$

一、（10 分）假设一条生产线生产的产品合格率是 0.8.要使一批产品的合格率达到在 76%与 84%之间的概率不小于 90%，问这批产品至少要生产多少件？$\Phi(1.64)=0.95,$

【解】 令 $X_i \begin{cases} 1, \text{若第}i\text{个产品是合格品}, \\ 0, \text{其他情形}. \end{cases}$

而至少要生产 n 件，则 $i=1,2,\ldots,n$,且

X_1，X_2，…，X_n 独立同分布，$p=P\{X_i=1\}=0.8$.

现要求 n,使得

$$P\{0.76 \le \frac{\sum_{i=1}^{n} X_i}{n} \le 0.84\} \ge 0.9.$$

即

$$P\{\frac{0.76n-0.8n}{\sqrt{n\times 0.8\times 0.2}} \le \frac{\sum_{i=1}^{n} X_i - 0.8n}{\sqrt{n\times 0.8\times 0.2}} \le \frac{0.84n-0.8n}{\sqrt{n\times 0.8\times 0.2}}\} \ge 0.9$$

由中心极限定理得

$$\Phi\left(\frac{0.84n-0.8n}{\sqrt{0.16n}}\right)-\Phi\left(\frac{0.76n-0.8n}{\sqrt{0.16n}}\right) \ge 0.9,$$

整理得 $\Phi\left(\frac{\sqrt{n}}{10}\right) \ge 0.95$, 查表 $\frac{\sqrt{n}}{10} \ge 1.64$,

$n \ge 268.96$,　　故取 $n=269$.

二、（12 分）（1）在某学校中，随机抽取 25 名同学测量身高数据，假设所测身高近似服从正态分布，算得平均身高为 170 cm，标准差为 12cm，求该校学生身高标准差σ的 95%的置信区间。

（2）制造某种产品的单件平均工时服从正态分布，现从中抽取 5 件，记录它们的制造工时（小时）如下：6.3,6.6,6.9,7.1,6.2，给定置信水平为 0.95，求其单件平均工时的单侧置信上限。

【解】

$$(1)\left(\sqrt{\frac{(n-1)s^2}{\chi^2_{0.025}(24)}},\sqrt{\frac{(n-1)s^2}{\chi^2_{0.975}(24)}}\right)=(9.34,16.69)。$$

$$(2)\ \overline{\mathrm{x}}=6.62,\ \mathrm{s}^2=0.147, t_{0.05}(4)=2.132, \overline{\mathrm{x}}+t_{0.05}(4)\frac{s}{\sqrt{5}}=6.99.$$

三、（20 分）设总体X的分布密度为$f(x)=\begin{cases}\dfrac{1}{\theta}e^{-\frac{x-\mu}{\theta}}, & x\ge\mu \\ 0, & 其它\end{cases}$$(\theta,\mu均未知)$.

$X_1,X_2,\cdots,X_{\mathrm{n}}$是取自总体$X$的简单随机样本，求

（1）求$\hat{\theta}_{ME}$和$\hat{\mu}_{\mathrm{ME}}$的矩估计量；

（2）求$\hat{\theta}_{MLE}$和$\hat{\mu}_{\mathrm{MLE}}$的最大似然估计量。

解：(1) $EX=\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x=\int_{\mu}^{+\infty}x\frac{1}{\theta}e^{-\frac{x-\mu}{\theta}}\mathrm{d}x=\mu+\theta,$

$$EX^2=\int_{-\infty}^{+\infty}x^2f(x)\mathrm{d}x=\int_{\mu}^{+\infty}x^2\frac{1}{\theta}e^{-\frac{x-\mu}{\theta}}\mathrm{d}x=\mu^2+2\theta(\mu+\theta).$$

$$\hat{\theta}_{ME}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2}, \hat{\mu}_{ME}=\bar{X}-\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2}。$$

$$(2) L(\mu,\theta)=\theta^{-n}e^{-\sum_{i=1}^{n}\frac{(x_i-\mu)}{\theta}}=\theta^{-n}e^{-\frac{1}{\theta}\sum_{i=1}^{n}x_i+\frac{n\mu}{\theta}}, x_{(1)}\ge\mu. L(\mu,\theta)关于\mu单调递增，\hat{\mu}_{\mathrm{MLE}}=\mathrm{x}_{(1)}$$

$$InL=-nIn\theta-\frac{1}{\theta}\sum_{i=1}^{n}(x_i-x_{(1)}), \hat{\theta}_{\mathrm{MLE}}=\bar{x}-x_{(1)}。$$

四、（14 分）某灯泡厂在采用一项新工艺的前后,分别抽取 10 个灯泡进行寿命试验.计算得到：采用新工艺前灯泡寿命的样本均值为 2460 小时，样本标准差为 56 小时；采用新工艺后灯泡寿命的样本均值为 2550 小时，样本标准差为 48 小时。设灯泡的寿命服从正态分布，是否可以认为采用新工艺后灯泡的平均寿命有显著提高（$\alpha=0.01$）？

解：(1)$H_{01}:\sigma_1^2=\sigma_2^2$，$H_{11}:\sigma_1^2\neq\sigma_2^2, f_0=\dfrac{s_1^2}{s_2^2}\approx 1.36<\mathrm{F}_{0.005}(9,9)=6.54$。

接受原假设，认为两总体的方差没有显著差异。

(2)$H_{02}:\mu_1\geq\mu_2$，$H_{12}:\mu_1<\mu_2$

取检验统计量$T=\dfrac{\bar{X}-\bar{Y}}{S_\omega\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$，其中$S_\omega=\sqrt{\dfrac{(n_1-1)\mathrm{s}_1^2+(n_2-1)\mathrm{s}_2^2}{n_1+n_2-2}}$

当原假设H_0为真时，$\mathrm{T}\sim \mathrm{t}(n_1+n_2-2)=\mathrm{t}(18)$ ，其拒绝域为$\mathrm{W}=\{\mathrm{t}_0<\mathrm{t}_{0.01}(18)\}$。依题意有,

$$\mathrm{s}_\omega=\sqrt{\frac{9\times(56^2+48^2)}{18}}=52.15.$$

$t=\dfrac{\bar{X}-\bar{Y}}{S_\omega\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}=\dfrac{2460-2550}{52.15\sqrt{\dfrac{1}{10}+\dfrac{1}{10}}}\approx -3.86<-t_{0.01}(18)=-2.55$,所以拒绝$H_{02}$，即认为采用新工艺后灯泡寿命明显提高。

五、（12 分）植物学家 G.J.Mendel 做豌豆试验时考虑豌豆的颜色和形状，一共有四种组合：（黄，圆），（黄，皱），（绿，圆），（绿，皱）。按 Mendel 理论，这四类应有 9:3:3:1 的比例，在一次具体观察中，发现这 4 类的观察数分别为 315,101,108 和 32. 在显著性水平为 0.05 下检验比例 9:3:3:1 的正确性。

【解】零假设为：

$$H_0: p_1=\frac{9}{16}, p_2=\frac{3}{16}, p_3=\frac{3}{16}, p_4=\frac{1}{16}$$

	p_i	n_i	np_i	n_i-np_i	$\frac{(n_i-np_i)^2}{np_i}$
1	9/16	315	312.75	2.25	0.0162
2	3/16	101	104.25	-3.25	0.1013
3	3/16	108	104.25	3.75	0.1349
4	1/16	32	34.75	2.75	0.2176
合计	1	556	556	0	χ_0^2=0.47

$\chi^2_{0.05}(3)=7.81>\chi_0^2$=0.47, 故接受原假设，即由观察数据可认为在显著性水平为 0.05 下 9:3:3:1 的比例是正确的。

六、（12 分）某年级有三个班，他们进行了一次概率统计的考试，现从三个班中各随机地抽取了一些学生，其成绩记录如下：

班级	成绩
I	73 89 82 43 80 73 66 60 45 93 36 77
II	88 78 48 91 51 85 74 56 77 31 78 62 76 96 80
III	68 79 56 91 71 71 87 41 59 68 53 79 15

设各个总体服从正态分布，且方差相等，试在α=0.05 下检验各班级的平均分数有无显著差异。

解：设 I，II，III 班的平均分数为μ_1,μ_2,μ_3,

$$H_0:\mu_1=\mu_2=\mu_3$$

$$S_T=13685.1, S_E=13349.75, S_A=335.35$$

$$F=\frac{37\times 335.35}{2\times 13349.75}=0.4647<3.23=F_{0.05}(2,37),$$

接受H_0，认为各班平均分数无显著差异。

七（12 分）某职工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表:

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

假设 y 与 x 之间存在近似的线性关系。

(1) 求经验回归方程 $\hat{y}=\widehat{a}+\widehat{b}x$;

(2) 检验线性关系的显著性($\alpha=0.05$).

【解】（1） $\overline{x}=30,\ S_{xx}=40,\ \overline{y}=1052,\ S_{yy}=54649.2,$

$$S_{xy}=\sum_{i=1}^{n}x_i y_i-n\overline{x}\,\overline{y}=1478,$$

故 $\widehat{b}=S_{xy}/S_{xx}=36.95,\quad \widehat{a}=\overline{y}-\widehat{b}\overline{x}=-11.3,$

经验回归方程 $\hat{y}=-11.3+36.95x.$

(2)法（一）

$$\widehat{\sigma}=\sqrt{\frac{s_{yy}-\widehat{b}s_{xy}}{n-2}}=\sqrt{\frac{54649.2-36.95\times1478}{3}}=3.52$$

$|t|=\dfrac{|\widehat{b}|}{\widehat{\sigma}}\sqrt{s_{xx}}=\dfrac{|36.95|}{3.52}\sqrt{40}=66.39>t_{0.025}(3)=3.1824$.故回归效果是显著的。

法（二）

$$S_{回}=\hat{\beta}_1L_{xy}=(-0.826)\times(-5.93)=4.898,\quad S_{剩}=L_{yy}-\hat{\beta}_1L_{xy}=1.682,$$

$$F_0=\frac{S_{回}}{S_{剩}/(n-2)}=8\times\frac{4.898}{1.682}=23.297,$$

$\alpha=0.05,\ F_{0.05}(1,8)=5.32.$ 因 $F_0>F_{0.05}(1,8)$, 故回归效果是显著的。

八、（8 分）（1）设$(X_1,X_2,\cdots,X_n)$为总体 X 的一个简单随机样本，$\overline{X}$是样本均值。

$E(X)=\mu, D(X)=\sigma^2$,求$\rho_{X_i-\overline{X},X_j-\overline{X}}$（$i\neq j$）。

（2）总体 X 服从$[0,\theta]$上的均匀分布，$\theta>0$为未知参数，$(X_1,X_2,\cdots,X_n)$为总体 X 的一个简单随机样本。试比较参数θ的两个无偏估计量$\hat{\theta}_1=2\bar{X}$与$\hat{\theta}_2=\frac{n+1}{n}\max\{X_1,X_2,\cdots,X_n\}$的有效性。

解：

（1）$\operatorname{cov}(X_i-\overline{X},X_j-\overline{X})=\operatorname{cov}(X_i,X_j)-\operatorname{cov}(X_i,\overline{X})-\operatorname{cov}(X_j,\overline{X})+D(\overline{X})$

$$=\operatorname{cov}(X_i,X_j)-2\operatorname{cov}(X_i,\overline{X})+\frac{\sigma^2}{n}=\operatorname{cov}(X_i,X_j)-\frac{\sigma^2}{n}=\begin{cases}\frac{n-1}{n}\sigma^2, & i=j\\ -\frac{\sigma^2}{n} & i\neq j\end{cases}$$

$$\rho_{X_i-\overline{X},X_j-\overline{X}}=\frac{-\frac{\sigma^2}{n}}{\frac{n-1}{n}\sigma^2}=\frac{-1}{n-1}。$$

（2） $D(X)=\frac{\theta^2}{12}, D(\hat{\theta}_1)=4D(\bar{X})=\frac{\theta^2}{3n}$

$$D(\hat{\theta}_2)=(\frac{n+1}{n})^2\frac{n\theta^2}{(n+1)^2(n+2)}=\frac{\theta^2}{n(n+2)}$$

$n\geq 2, n(n+2)>3n, D(\hat{\theta}_2)<D(\hat{\theta}_1), \hat{\theta}_2$比 $\hat{\theta}_1$有 效 。