厦门大学《概率统计》课程

期末试题·答案

自律督导部

考试日期：　2012.1　(A)　　信息学院自律督导部整理

1. (15 分) 一生产线生产的产品成箱包装，每箱的重量是随机的.假设每箱平均重 50 千克，方差为 25 千克，若用最大载重量为 5 吨的汽车承运，试利用中心极限定理说明每辆车最多可以装多少箱，才能保障不超载的概率大于 0.977. $(\Phi(2)=0.997)$

解： 设每箱产品的重量为 x_i，依题意 $E(x_i)=50$，$D(x_i)=25$，每辆车最多可以装 Y 箱，才能保障不超载的概率大于 0.977，由中心极限定理，则

$$P(\sum_{i=1}^{Y} X_i \le 5000) \approx \Phi(\frac{5000-50Y}{\sqrt{25Y}}) > 0.977 = \Phi(2)$$

故 $\dfrac{5000-50Y}{\sqrt{25Y}} > 2$。

解得　$Y=98$

2. (15 分) 设 $X \sim N(\mu_1, \sigma^2), Y \sim N(\mu_2, \sigma^2)$，并且相互独立，基于分别来自总体 X 和 Y 容量相应为 9 和 11 的简单随机样本，得样本均值 $\bar{X}$ 和 $\bar{Y}$，样本方差 S_X^2 和 S_Y^2。

记　$\xi=\dfrac{1}{2}(S_X^2+S_Y^2), \eta=\dfrac{1}{18}(8S_X^2+10S_Y^2)$，证明：(1) 统计量 S_X^2, S_Y^2, ξ，η 都是 σ^2 的无偏估计量；(2) η 在四个估计量 S_X^2, S_Y^2, ξ，η 中方差最小.

证：(1) 对任意总体，样本方差都是总体方差的无偏估计，所以 $ES_X^2=ES_Y^2=\sigma^2$,这是因为

$$ES^2=E[\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2]=\frac{1}{n-1}E[\sum_{i=1}^{n}X_i^2-n\bar{X}^2]=\frac{1}{n-1}[\sum_{i=1}^{n}EX_i^2-nE\bar{X}^2]$$

$$=\frac{n}{n-1}[(\sigma^2-\mu^2)-(\frac{\sigma^2}{n}+\mu^2)]=\sigma^2,$$

$$E\xi=\frac{1}{2}(ES_X^2+ES_Y^2)=\sigma^2, E\eta=\frac{1}{18}(8ES_X^2+10ES_Y^2)=\sigma^2;$$

(2) 利用 $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$ 和 $D(\chi^2(n))=2n$ 结论，得

$$DS_X^2=\frac{\sigma^4}{8^2}D(\frac{8S_X^2}{\sigma^2})=\frac{\sigma^4}{8^2}\times 2\times 8=\frac{\sigma^4}{4}$$，(其中 n=9)

$$DS_Y^2=\frac{\sigma^4}{10^2}D(\frac{10S_Y^2}{\sigma^2})=\frac{\sigma^4}{100}\times 2\times 10=\frac{\sigma^4}{5}$$，(其中 n=11)

$$D\xi=\frac{1}{4}(DS_X^2+DS_Y^2)=\frac{1}{4}(\frac{\sigma^2}{4}+\frac{\sigma^2}{5})=\frac{9\sigma^4}{80},$$

$$D\eta=\frac{1}{18^2}(\frac{64\sigma^4}{4}+\frac{100\sigma^4}{5})=\frac{\sigma^4}{9}$$

于是比较四个估计量 S_X^2,S_Y^2,ξ，η 的方差知，η 的方差最小.

3.（15 分）测量某种溶液中的水分，从它的 10 个测定值得出 $\bar{x}$=0.452(%),s=0.037(%).设测定值总体为正态，μ 为总体均值，σ 为总体标准差，试在水平 α=0.05 下检验.

（1）H_0：$\mu\geq 0.5(\%)$；H_1：$\mu<0.5(\%)$.（2）H_0：$\sigma\geq 0.04(\%)$；H_1：$\sigma<0.04(\%)$.

【解】（1）

$$\begin{aligned}
&\mu_0=0.5;n=10,\alpha=0.05,t_\alpha(n-1)=t_{0.05}(9)=1.8331,\\
&\bar{x}=0.452,s=0.037,\\
&t=\frac{\bar{x}-\mu_0}{s/\sqrt{n}}=\frac{(0.452-0.5)}{0.037}\times\sqrt{10}=-4.10241,\\
&t<-t_{0.05}(9)=-1.8331.
\end{aligned}$$

所以拒绝 H_0，接受 H_1.

(2)

$$\begin{aligned}
&\sigma_0^2=(0.04)^2,n=10,\alpha=0.05,\chi_{1-\alpha}^2=\chi_{0.95}^2(9)=3.325,\\
&\bar{x}=0.452,s=0.037,\\
&\chi^2=\frac{(n-1)s^2}{\sigma_0^2}=\frac{9\times 0.037^2}{0.04^2}=7.7006,\\
&\chi^2>\chi_{0.95}^2(9).
\end{aligned}$$

所以接受 H_0，拒绝 H_1.

4. (15 分)　两台机床加工同一种零件，分别各取 8 个零件，量其长度得 $\bar{x}=81.625$，$\bar{y}=75.875$，$S_1^2=145.60$，$S_2^2=102.13$，假定零件长度服从正态分布，

（1）检验假设 $H_0:\sigma_1^2=\sigma_2^2$？ $(\alpha=0.05)$

（2）若认为两总体方差未知但相等，试求 $\mu_1-\mu_2$ 在置信度为 0.95 下的置信区间.

解（1）问题是检验假设　$H_0:\sigma_1^2=\sigma_2^2$

选统计量 F 并计算其值

$$F=\frac{S_1^2}{S_2^2}=\frac{145.60}{102.13}=1.4256,$$

对给定的 $\alpha=0.05$ 查 F 分布表得临界值 $F_{\alpha/2}(7,7)=F_{0.025}(7,7)=4.99$，

$F_{0.975}(7,7)=\frac{1}{4.99}=0.2$.

因 $F_{0.975}(7,7)=0.2<1.4256=F<4.99=F_{0.025}(7,7)$ 故接受 H_0，即无显著差异.

（2） 此题是在 $\sigma_1^2=\sigma_2^2$ 的条件下求 $\mu_1-\mu_2$ 的置信区间.

$\mu_1-\mu_2$ 的置信区间为

$$(\bar X-\bar Y-t_{\alpha/2}(n_1+n_2-2)S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}},\ \bar X-\bar Y+t_{\alpha/2}(n_1+n_2-2)S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}})$$

其中 $\bar X=\frac{1}{8}\sum_{i=1}^{8}X_i=81.625,\ S_1^2=\frac{1}{7}(\sum_{i=1}^{8}X_i^2-8(81.625)^2)=145.60$

$$\bar Y=\frac{1}{8}\sum_{i=1}^{8}Y_i=75.875,\ S_2^2=\frac{1}{7}(\sum_{i=1}^{8}Y_i^2-8\times(75.875)^2)=102.13$$

$$S_w=\sqrt{\frac{(8-1)\times 145.60+(8-1)\times 102.13}{14}}=11.129,\ \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}=\frac{1}{2}$$

$\alpha=0.05,\quad t_{0.025}(14)=2.1448$.

所以 $\mu_1-\mu_2$ 的置信度为 0.95 下的置信区间为 $(-6.185,\ 17.685)$.

5．(15 分)设 $X_1,X_2,\cdots,X_n$ 来自几何分布

$$P(X=k)=p(1-p)^{k-1},\ k=1,2,\cdots,\quad 0<p<1,$$

试求未知参数 p 的极大似然估计.

解 $L(x_1,\cdots,x_n;p)=\prod_{i=1}^{n}p(1-p)^{x_i-1}=p^n(1-p)^{\sum_{i=1}^{n}x_i-n}$,

$$\ln L=n\ln p+(\sum_{i=1}^{n}X_i-n)\ln(1-p),$$

$$\frac{d\ln L}{dp}=\frac{n}{p}-\frac{\sum_{i=1}^{n}X_i-n}{1-p}\ 0,$$

解似然方程 $$\frac{n}{p}=\frac{-n+\sum_{i=1}^{n}X_i}{1-p},$$

得 p 的极大似然估计 $$p=\frac{1}{\bar X}.$$

6. (10 分)设总体 X 具有密度

$$f(x;\ \theta)=\begin{cases}C^{\frac{1}{\theta}}\frac{1}{\theta}x^{-(1+\frac{1}{\theta})}, & x>C,\\ 0\ , & \text{其他}.\end{cases}$$

其中参数 $0<\theta<1,\ C$ 为已知常数，且 $C>0$，从中抽得一个样本，$X_1,X_2,\cdots,X_n$，求 θ 的矩估计

解 $\mu_1=EX=\int_C^{+\infty}C^{\frac{1}{\theta}}\frac{1}{\theta}x^{-\frac{1}{\theta}}dx=C^{\frac{1}{\theta}}\frac{1}{\theta}\frac{1}{1-\frac{1}{\theta}}x^{1-\frac{1}{\theta}}\Big|_C^{+\infty}$

$$=C^{\frac{1}{\theta}}\frac{1}{\theta-1}(-C\cdot C^{-\frac{1}{\theta}})=\frac{C}{1-\theta},$$

解出θ得 $\theta=1-\frac{C}{\mu_1},$

于是θ的矩估计为 $\theta=1-\frac{C}{\overline{X}}.$

7. (15 分) 在硝酸钠($NaNO_3$)的溶解度试验中，对不同的温度$t°C$测得溶解于 100ml 水中的硝酸钠质量Y的 9 次观测数据算得

$$\sum t=234,\ \sum y=811.8,\ \sum t^2=10144,\ \sum y^2=76317.82,\ \sum ty=24646.6$$

从理论知Y与t满足线性回归模型

(1) 求Y对t的回归方程$y=a+\hat{b}t$；

(2) 检验回归方程的显著性$(\alpha=0.01)$；($F_{0.01}(1,\ 7)=12.25$)

(3) 求Y在$t=25$℃时的预测区间（置信度为 0.95）.($t_{0.025}(7)=2.3646$)

解 计算表如下

序号	t_i	y_i	t_i^2	y_i^2	$t_i\,y_i$
1	0	66.7	0	4448.89	0
2	4	71.0	16	5041.00	284
3	10	76.3	100	5821.69	763
4	15	80.6	225	6496.36	1209
5	21	85.7	441	7344.49	1799.7
6	29	92.9	841	8630.41	2694.1
7	36	99.9	1296	9980.01	3596.4
8	51	113.6	2601	12904.96	5793.6
9	68	125.1	4624	15560.01	8506.8
Σ	234	811.8	10144	76317.82	24646.6

$$\overline{t}=26,\quad \overline{y}=90.2$$

$$L_{tt}=\sum_{i=1}^{9}t_i^2-9\overline{t}^2=10144-6084=4060,$$

$$L_{ty}=\sum_{i=1}^{9}t_i\,y_i-9\overline{t}\,\overline{y}=24646.6-21106.8=3539.8,$$

$$L_{yy}=\sum_{i=1}^{9}y_i^2-9\overline{y}^2=76317.82-73224.36=3093.46$$

$$\hat{b}=\frac{L_{ty}}{L_{tt}}=0.87187,\qquad a=\overline{y}-\hat{b}\overline{t}=67.5313,$$

$$S^2=(L_{yy}-\hat{b}L_{ty})/7=1.0307,\qquad S=1.0152$$

(1) Y对t的回归方程为 $y=67.5313+0.87187t$；

(2) 方差分析表如下

方差来源	平方和	自由度	均 方	F 值

回　　归	3086.25	1	3086.25	$\frac{3086.25}{1.03}$ =2996.36
剩　　余	7.21	7	1.03	
总　　和	3093.46	8		

查 F 分布表求出临界值 $F_{0.01}(1,\ 7)=12.25$

因 $F=2996.36>>12.25=F_{0.01}(1,\ 7)$，故方程高度显著.

（3） $y_0=67.5313+0.87187\times 25=89.3281$

$$\delta(25)=t_{\alpha/2}(n-2)\times S\times\sqrt{1+\frac{1}{n}+\frac{(t_0-\overline{t})^2}{L_{tt}}}$$

$$=2.3646\times 1.0152\times 1.05=2.53$$

Y 在 $t=25$℃时的置信度为 0.95 下的预测区间为

$$(y_0-\delta(25),\ y_0+\delta(25))=(86.79,\ 91.85)).$$