数理统计复习题解答

一、填空

1、设 $X_1,\cdots,X_n$ 是来自总体 $X\sim U(-1,1)$ 的简单随机样本，则 $E\overline{X}=$____________，

$D\overline{X}=$____________。

解：$E\overline{X}=EX=0$，$D\overline{X}=\dfrac{DX}{n}=\dfrac{1}{3n}$

2、设 $X_1,\cdots,X_n$ 是来自总体 $X\sim N(\mu,1)$ 的简单随机样本，则样本均值 $\overline{X}\sim$______，

$\sum\limits_{i=1}^{n}(X_i-\overline{X})^2\sim$__________，而 $\sum\limits_{i=1}^{n}(X_i-\mu)^2\sim$_________，且其期望为______，方差为________。

解：$\overline{X}\sim N(\mu,\dfrac{1}{n})$，$\sum\limits_{i=1}^{n}(X_i-\overline{X})^2\sim\chi^2(n-1)$，$\sum\limits_{i=1}^{n}(X_i-\mu)^2\sim\chi^2(n)$，

$E\sum\limits_{i=1}^{n}(X_i-\mu)^2=E\chi^2(n)=n$，$D\sum\limits_{i=1}^{n}(X_i-\mu)^2=D\chi^2(n)=2n$

3、设 $T\sim t(n)$，若 $P(|T|>\lambda)=\alpha$，则 $P(T\le\lambda)=$____________。

解：利用 t 分布的对称性得：$P(T\le-\lambda)=P(T\ge\lambda)$

由 $P(|T|>\lambda)=\alpha$ 得 $P(T\ge\lambda)=\dfrac{\alpha}{2}$

所以 $P(T\le\lambda)=1-P(T>\lambda)=1-\dfrac{\alpha}{2}$

4、设 $X_1,\cdots,X_n$ 是来自总体 $X\sim b(1,p)$ 的简单随机样本，则参数 p 的无偏估计为________，实际频数 $n\hat{p}\sim$___________。

解：参数 p 的无偏估计为样本均值 $\overline{X}$，$n\hat{p}=X_1+\cdots+X_n\sim b(n,p)$

5、设 $X_1,\cdots,X_n$ 是来自总体 $X\sim N(\mu,\sigma^2)$ 的简单随机样本，其中 μ 未知，则检验假设 $H_0:\sigma=\sigma_0$ 的检验统计量为________________，在 H_0 下服从________。

解：检验统计量为 $\dfrac{(n-1)S^2}{\sigma_0^2}$，在 H_0 下服从 $\chi^2(n-1)$。

6、在假设检验中显著性水平α是用来控制犯第一类错误的概率，第一类错误是指____。

解：第一类错误是指原假设为真，但是作出的判断是拒绝原假设。

7、设$X_1,\cdots,X_n$是来自总体$X\sim P(\lambda)$的简单随机样本，则样本$X_1,\cdots,X_n$的联合分布律为____________________________。

解：$P(X_1=x_1,\cdots,X_n=x_n)=\dfrac{\lambda^{x_1+\cdots+x_n}}{x_1!\cdots x_n!}e^{-n\lambda}$

8、从某一总体中抽取一样本得观察值为：15，25，30，40，50，则样本均值的观察值为___________，样本方差的观察值为____________。

解：$\overline{X}=\dfrac{1}{5}\left(15+25+30+40+50\right)=32$，$S^2=\dfrac{1}{4}\left(15^2+25^2+30^2+40^2+50^2-5\times\overline{X}^2\right)=182.5$

二、设某大公司的员工在上、下班上所化时间$X\sim N\left(\mu,25\right)$，该公司为了解员工在上、下班上所化平均时间，现抽取了一个简单随机样本$X_1,\cdots,X_n$，问应抽取多大样本容量n才能以 99%概率保证样本均值与真正均值μ的差的绝对值不超过 1 分钟。

解：因为总体$X\sim N\left(\mu,25\right)$，所以$\overline{X}\sim N\left(\mu,\dfrac{25}{n}\right)$。

又 $P\left(|\overline{X}-\mu|<1\right)=P\left(\mu-1<\overline{X}<\mu+1\right)=2\Phi\left(\dfrac{\sqrt{n}}{5}\right)-1>0.99$

由此得：$\dfrac{\sqrt{n}}{5}>2.575$，即$n>165.77$，所以至少抽 166 个。

三、设总体$X\sim N\left(\mu,25\right)$，从这总体中抽取二个独立样本，样本容量分别为 20 和 12，$\overline{X}_1,\overline{X}_2$分别是这二个样本的样本均值，求$P\left(\left|\overline{X}_1-\overline{X}_2\right|\geq 1\right)$。

解：因为总体$X\sim N\left(\mu,25\right)$，所以$\overline{X}_1\sim N\left(\mu,\dfrac{25}{20}\right)$，$\overline{X}_2\sim N\left(\mu,\dfrac{25}{12}\right)$。

因此$\overline{X}_1-\overline{X}_2\sim N(0,\dfrac{10}{3})$。

$P\left(\left|\overline{X}_1-\overline{X}_2\right|\geq 1\right)=1-P\left(\left|\overline{X}_1-\overline{X}_2\right|<1\right)=2-2\Phi\left(\dfrac{\sqrt{30}}{10}\right)$

四、设总体 X 的密度函数为 $f(x)=\begin{cases}\theta\left(\theta+1\right)x^{\theta-1}(1-x) & 0<x<1\\ 0 & other\end{cases}$, $X_1,\cdots,X_n$ 是来自其的简单随机样本，求参数 θ 的矩估计。

解：因为 $EX=\int_{-\infty}^{+\infty}xf(x)dx=\int_0^1\theta\left(\theta+1\right)x^{\theta}(1-x)dx=\dfrac{\theta}{\theta+2}$，所以 $\theta=\dfrac{2EX}{1-EX}$.

从而参数 θ 的矩估计为 $\hat{\theta}=\dfrac{2\overline{X}}{1-\overline{X}}$。

五、设总体 X 的密度函数为 $f(x)=\begin{cases}2\theta xe^{-\theta x^2} & x>0\\ 0 & x\le 0\end{cases}$，$X_1,\cdots,X_n$ 是来自其的简单随机样本，求参数 θ 的最大似然估计。

解：似然函数为 $L(\theta)=f(x_1)\cdots f(x_n)=2\theta x_1e^{-\theta x_1^2}\cdots 2\theta x_ne^{-\theta x_n^2}=2^n\theta^n x_1\cdots x_ne^{-\theta(x_1^2+\cdots+x_n^2)}$

对数似然函数为

$\ln L(\theta)=n\ln 2+n\ln\theta+\ln x_1+\cdots+\ln x_n-\theta(x_1^2+\cdots+x_n^2)$

所以 $\dfrac{\partial\ln L(\theta)}{\partial\theta}=\dfrac{n}{\theta}-(x_1^2+\cdots+x_n^2)=0$

从而参数 θ 的最大似然估计为 $\hat{\theta}=\dfrac{n}{x_1^2+\cdots+x_n^2}$

六、设 $X_{i1},\cdots,X_{in_i}$ 是来自总体 $X_i\sim N(\mu_i,\sigma_i^2)$ 的简单随机样本($i=1,2$)，$\overline{X}_i,S_i^2$ ($i=1,2$)是这二个样本的样本均值和样本方差，$S_w^2=\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$。如果 $\sigma_1=\sigma_2=\sigma$，证明（1）S_1^2,S_2^2,S_w^2 都是参数 σ^2 的无偏估计；（2）S_w^2 比 S_1^2,S_2^2 都有效。

证明：（1）因为 $ES_1^2=DX=\sigma^2$，$ES_2^2=DX=\sigma^2$，

$$ES_w^2=\frac{(n_1-1)ES_1^2+(n_2-1)ES_2^2}{n_1+n_2-2}=\frac{(n_1-1)\sigma^2+(n_2-1)\sigma^2}{n_1+n_2-2}=\sigma^2$$

所以 S_1^2,S_2^2,S_w^2 都是参数 σ^2 的无偏估计。

（2）注意到 $\dfrac{(n_1-1)S_1^2}{\sigma^2}\sim\chi^2(n_1-1),\dfrac{(n_2-1)S_2^2}{\sigma^2}\sim\chi^2(n_2-1)$ 及 S_1^2,S_2^2 独立，则

$$\frac{(n_1+n_2-2)S_w^2}{\sigma^2}=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{\sigma^2}\sim\chi^2(n_1+n_2-2)$$

从而 $D\left[\frac{(n_1-1)S_1^2}{\sigma^2}\right]=2(n_1-1), D\left[\frac{(n_2-1)S_2^2}{\sigma^2}\right]=2(n_2-1)$

$$D\left[\frac{(n_1+n_2-2)S_w^2}{\sigma^2}\right]=2(n_1+n_2-2)$$

由此得：$D\left(S_1^2\right)=\frac{2}{(n_1-1)}\sigma^4, D\left(S_2^2\right)=\frac{2}{(n_2-1)}\sigma^4$，$D\left(S_w^2\right)=\frac{2}{(n_1+n_2-2)}\sigma^4$

因此 $D\left(S_w^2\right)<D\left(S_1^2\right), D\left(S_2^2\right)$，即 S_w^2 比 S_1^2, S_2^2 都有效。

七、设钢轴的直径 $X\sim N\left(\mu, 0.012^2\right)$，根据设计要求钢轴的直径应该是 0.15，现抽检了 75 件，测得均值为 0.154，问 $\alpha=0.1$ 下这批产品是否合格？

解：已知 $n=75, \bar{X}=0.154$

需要检验的假设是：$H_0:\mu=0.15,\quad v.s.\quad H_1:\mu\neq 0.15$

由于方差已知，所以检验统计量为 $U=\frac{\bar{X}-0.15}{0.012}\sqrt{75}\overset{H_0}{\sim}N(0,1)$

检验的拒绝域是 $W=\{|U|>u_{0.05}=1.645\}$

检验统计量的值为 $U=\frac{0.154-0.15}{0.012}\sqrt{75}=\frac{5\sqrt{5}}{3}\in W$，

所以拒绝原假设，即认为这批产品不合格。

八、某种片剂药物中成份 A 的含量规定为 10%，现抽检 5 个片剂，测得成份 A 的含量，经计算得：$\bar{X}=0.1005, S=0.0059$，设成份 A 的含量 $X\sim N\left(\mu,\sigma^2\right)$，问在 $\alpha=0.05$ 下成份 A 的含量是否符合标准？

解：$H_0:\mu=0.1,\quad v.s.\quad H_1:\mu\neq 0.1$

由于方差未知，所以检验统计量为 $T=\frac{\bar{X}-0.1}{S}\sqrt{5}\overset{H_0}{\sim}t(4)$

检验的拒绝域是 $W=\{|T|>t_{0.025}(4)=2.7764\}$

检验统计量的值为$T=\dfrac{0.1005-0.1}{0.0059}\sqrt{5}=1.895\notin W$，

所以接受原假设，即认为这批产品符合标准。

九、某自动机床加工同类型的轴承，现从两个不同班次的产品中各抽检了 5 根轴承，并测定它们的直径，经计算得：

A 班：$X\sim N(\mu_1,\sigma_1^2)$， $\bar{X}=2.0648, S_1^2=1.07\times10^{-5}$

B 班：$Y\sim N(\mu_2,\sigma_2^2)$， $\bar{Y}=2.0594, S_2^2=5.3\times10^{-6}$

试根据抽样结果说明两个班次生产的产品直径有无显著差异？

解：（1）先检验方差是否相等

$H_0:\sigma_1^2=\sigma_2^2 \quad v.s. \quad H_1:\sigma_1^2\neq\sigma_2^2$

检验统计量为$F=\dfrac{S_1^2}{S_2^2}\overset{H_0}{\sim}F(4,4)$

拒绝域为$W=\{F>F_{0.025}(4,4)=9.6045\}\cup\{F<F_{0.975}(4,4)=1/F_{0.025}(4,4)\}$

检验统计是的值为$F=2.0189\notin W$，即认为方差相等。

（2）在方差相等基础上检验均值是否相等

$H_0:\mu_1=\mu_2 \quad v.s. \quad H_1:\mu_1\neq\mu_2$

检验统计量为$T=\dfrac{\bar{X}-\bar{Y}}{S_w}\sqrt{\dfrac{n_1n_2}{n_1+n_2}}\overset{H_0}{\sim}t(8)$

拒绝域为$W=\{|T|>t_{0.025}(8)=2.3060\}$

检验统计是的值为$T=3.8184\in W$，

即认为两个班次生产的产品直径有显著差异。