P170 习题 7

1. 从某地区 11 岁的男小学生中随机抽取 9 人，测得其身高和体重值格式为（身高，体重），数据如下

（160,43），（157,40），（153,42），（158,49），（157,45）

（154,42），（154,41），（163,46），（154,45）

用数字特征法分别对身高 X 和体重 Y 的均值和方差进行估计.

解答：身高的均值和方差的估计为

$$\mu_1=\overline{x}=\frac{1}{9}\sum_{i=1}^{9}x_i=\frac{1}{9}(160+157+153+158+157+154+154+163+154)=156.7$$

$$\sigma_1^2=s_1^2=\frac{1}{8}\sum_{i=1}^{9}(x_i-\overline{x})^2=11$$

体重的均值和方差的估计为

$$\mu_2=\overline{y}=\frac{1}{9}\sum_{i=1}^{9}y_i=43.7$$

$$\sigma_2^2=s_2^2=\frac{1}{8}\sum_{i=1}^{9}(y_i-\overline{y})^2=8$$

注：计算时可以将数据输入 Excel，然后通过命令“=average(??:??)” “=var(??:??)”获得样本均值和样本方差。（其中(??:??)表示输入数据所在的单元格范围）

2. 若总体 X 服从参数为λ的指数分布，$X_1,X_2,\cdots,X_n$是该总体的一个样本，求未知参数λ的一个矩估计量

解答 1：总体的一阶原点矩$E(X)=\dfrac{1}{\lambda}$

样本的一阶原点矩$\overline{X}=\dfrac{1}{n}\sum_{i=1}^{n}X_i$

由矩估计法得$\dfrac{1}{\lambda}=\overline{X}$，因此$\lambda=\dfrac{1}{\overline{X}}$

解答 2：总体的二阶中心矩$D(X)=\dfrac{1}{\lambda^2}$

样本的二阶中心矩 $S_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$

由矩估计法得 $\frac{1}{(\lambda)^2} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$ ，因此 $\lambda = \sqrt{\frac{n}{\sum_{i=1}^{n}(X_i - \overline{X})^2}}$

3. 若总体 X 服从几何分布

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

$X_1, X_2, \cdots, X_n$ 是总体 X 的一个容量为 n 的样本，求未知参数 p 的矩估计和最大似然估计量

解答：（矩估计）

总体的一阶原点矩 $E(X) = \sum_{k=1}^{\infty} k(1-p)^{k-1} p = \frac{1}{p}$

样本的一阶原点矩 $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

由矩估计法得 $\frac{1}{p} = \overline{X}$ ，因此 $p = \frac{1}{\overline{X}}$ 50 分

（最大似然估计）

似然函数为 $L(p) = \prod_{i=1}^{n}(1-p)^{x_i - 1} p = p^n (1-p)^{\sum_{i=1}^{n} x_i - n}$

对数似然函数为 $\ln L(p) = n \ln p + \left(\sum_{i=1}^{n} x_i - n\right)\ln(1-p)$

两边对参数 p 求导得 $\frac{d \ln L(p)}{dp} = \frac{n}{p} + \left(\sum_{i=1}^{n} x_i - n\right)\left(-\frac{1}{1-p}\right)$

解 $\frac{d \ln L(p)}{dp} = \frac{n}{p} + \left(\sum_{i=1}^{n} x_i - n\right)\left(-\frac{1}{1-p}\right) = 0$ 得 $p = \frac{n}{\sum_{i=1}^{n} x_i} = \frac{1}{\overline{x}}$

因此 p 的最大似然估计量为 $p = \frac{1}{\overline{X}}$ 50 分

4. 若总体 X 服从二项分布 $B(N,p)$，其中 N 已知，$X_1,X_2,\cdots,X_n$ 是总体 X 的一个容量为 n 的样本，求未知参数 p 的矩估计和最大似然估计量

解答：（矩估计）

总体的一阶原点矩 $E(X)=Np$

样本的一阶原点矩 $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$

由矩估计法得 $Np=\overline{X}$，因此 $p=\frac{\overline{X}}{N}$

（最大似然估计）

似然函数为 $L(p)=\prod_{i=1}^{n}C_N^{x_i}p^{x_i}(1-p)^{N-x_i}=\left(\prod_{i=1}^{n}C_N^{x_i}\right)p^{\sum_{i=1}^{n}x_i}(1-p)^{Nn-\sum_{i=1}^{n}x_i}$

对数似然函数为 $\ln L(p)=\ln\left(\prod_{i=1}^{n}C_N^{x_i}\right)+\left(\sum_{i=1}^{n}x_i\right)\ln p+\left(Nn-\sum_{i=1}^{n}x_i\right)\ln(1-p)$

两边对参数 p 求导得 $\frac{d\ln L(p)}{dp}=\frac{\sum_{i=1}^{n}x_i}{p}+\left(Nn-\sum_{i=1}^{n}x_i\right)\left(-\frac{1}{1-p}\right)$

解 $\frac{d\ln L(p)}{dp}=\frac{\sum_{i=1}^{n}x_i}{p}+\left(Nn-\sum_{i=1}^{n}x_i\right)\left(-\frac{1}{1-p}\right)=0$ 得 $p=\frac{\sum_{i=1}^{n}x_i}{Nn}=\frac{\bar{x}}{N}$

因此 p 的最大似然估计量为 $p=\frac{\overline{X}}{N}$

5. 设 $X_1,X_2,\cdots,X_n$ 是从总体 X 抽得的一个简单随机样本，总体 X 的概率密度函数为

$$p(x,\theta)=\begin{cases}\theta x^{\theta-1} & 0<x<1\\ 0 & else\end{cases}$$

其中 $\theta>0$ 是未知参数，试用矩估计法和最大似然法估计参数 θ

解答：（矩估计）

总体的一阶原点矩 $E(X)=\int_{-\infty}^{+\infty}xp(x,\theta)dx=\int_0^1 x\theta x^{\theta-1}dx=\frac{\theta}{\theta+1}$

样本的一阶原点矩 $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$

由矩估计法得 $\frac{\theta}{\theta+1}=\overline{X}$ ，因此 $\theta=\frac{\overline{X}}{1-\overline{X}}$

（最大似然估计）

似然函数为 $L(\theta)=\prod_{i=1}^{n}\theta x_i^{\theta-1}=\theta^n\left(\prod_{i=1}^{n}x_i\right)^{\theta-1}$

对数似然函数为 $\ln L(\theta)=n\ln\theta+(\theta-1)\ln\left(\prod_{i=1}^{n}x_i\right)$

两边对参数 θ 求导得 $\frac{d\ln L(\theta)}{d\theta}=\frac{n}{\theta}+\ln\left(\prod_{i=1}^{n}x_i\right)$

解 $\frac{d\ln L(\theta)}{d\theta}=\frac{n}{\theta}+\ln\left(\prod_{i=1}^{n}x_i\right)=0$ 得 $\theta=\frac{-n}{\ln\left(\prod_{i=1}^{n}x_i\right)}=\frac{-n}{\sum_{i=1}^{n}\ln x_i}$

因此 θ 的最大似然估计量为 $\theta=\frac{-n}{\sum_{i=1}^{n}\ln X_i}$

6. 设 $X_1,X_2,\cdots,X_n$ 是从总体 X 抽得的一个简单随机样本，总体 X 的概率密度函数为

$$p(x,\theta)=\begin{cases}\frac{1}{\theta}e^{-\frac{x}{\theta}} & x>0,\theta>0\\ 0 & else\end{cases}$$

试用最大似然法估计参数 θ

解答：似然函数为 $L(\theta)=\prod_{i=1}^{n}\frac{1}{\theta}e^{-\frac{x_i}{\theta}}=\frac{1}{\theta^n}e^{-\frac{\sum_{i=1}^{n}x_i}{\theta}}$ 30 分

对数似然函数为 $\ln L(\theta)=-n\ln\theta-\frac{\sum_{i=1}^{n}x_i}{\theta}$

两边对参数 θ 求导得 $\frac{d\ln L(\theta)}{d\theta}=-\frac{n}{\theta}+\frac{\sum_{i=1}^{n}x_i}{\theta^2}$ 30 分

解 $\frac{d\ln L(\theta)}{d\theta}=-\frac{n}{\theta}+\frac{\sum_{i=1}^{n}x_i}{\theta^2}=0$ 得 $\theta=\frac{\sum_{i=1}^{n}x_i}{n}=\overline{x}$

因此 θ 的最大似然估计量为 $\theta=\overline{X}$ 40 分

7. 设总体 X 服从正态分布 $N(\mu,1)$ ，X_1,X_2,X_3 为其一个样本，说明下列三个统计量中哪些是无偏估计量，并说明无偏估计中哪个方差最小

（1）$\mu_1=\frac{1}{5}X_1+\frac{3}{10}X_2+\frac{1}{2}X_3$

（2）$\mu_2=\frac{1}{3}X_1+\frac{1}{4}X_2+\frac{5}{12}X_3$

（3）$\mu_3=\frac{1}{3}X_1+\frac{1}{5}X_2+\frac{1}{12}X_3$

解答：由于 $E(\mu_1)=E(\frac{1}{5}X_1+\frac{3}{10}X_2+\frac{1}{2}X_3)=\frac{\mu}{5}+\frac{3}{10}\mu+\frac{\mu}{2}=\mu$

$E(\mu_2)=E(\frac{1}{3}X_1+\frac{1}{4}X_2+\frac{5}{12}X_3)=\frac{\mu}{3}+\frac{\mu}{4}+\frac{5}{12}\mu=\mu$

$E(\mu_3)=E(\frac{1}{3}X_1+\frac{1}{5}X_2+\frac{1}{12}X_3)=\frac{\mu}{3}+\frac{\mu}{5}+\frac{\mu}{12}=\frac{37}{60}\mu$

因此 μ_1,μ_2 是 μ 的无偏估计

$D(\mu_1)=D(\frac{1}{5}X_1+\frac{3}{10}X_2+\frac{1}{2}X_3)=\frac{1}{25}+\frac{9}{100}+\frac{1}{4}=\frac{19}{50}$

$D(\mu_2)=D(\frac{1}{3}X_1+\frac{1}{4}X_2+\frac{5}{12}X_3)=\frac{1}{9}+\frac{1}{16}+\frac{25}{144}=\frac{25}{72}$

由于 $\frac{19}{50}>\frac{25}{72}$ ，即 $D(\mu_1)>D(\mu_2)$ ，因此 μ_2 比 μ_1 有效

8. 设某种水稻的亩产量服从正态分布 $N(\mu,\sigma^2)$ ，随机抽取 9 亩试验田，测得亩产量（单位：千克）如下

510,485,505,505,490,495,520,515,490

试求均值 μ 置信度为 99%的置信区间

解答：计算得$\bar{x}=501.67$，$s^2=150$

方差σ^2未知，均值μ的置信度为$1-\alpha$的置信区间为

$$\left[\bar{X}-t_{\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}},\bar{X}+t_{\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}}\right]$$

查表得$t_{\frac{0.01}{2}}(8)=3.355$，代入$n=9$，$\bar{x}=501.67$，$s^2=150$得$\mu$的置信度为 99%的置信区间为$[488.0,515.4]$

9. 设某种电子管的使用寿命服从正态分布，从中随机抽取 16 个进行检验，得平均使用寿命为 1950 小时，标准差$s=300$小时，试分别求

（1）整批电子管平均使用寿命的置信度为 95%的置信区间；

（2）使用寿命的标准差的置信度为 95%的置信区间

解答：由题目条件，$\bar{x}=1950$，$s=300$

（1）方差σ^2未知，均值μ的置信度为$1-\alpha$的置信区间为

$$\left[\bar{X}-t_{\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}},\bar{X}+t_{\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}}\right]$$

查表得$t_{\frac{0.05}{2}}(15)=2.131$，代入$n=16$，$\bar{x}=1950$，$s=300$得整批电子管平均使用寿命的置信度为 95%的置信区间为$[1790.175,2109.825]$50 分

说明：教材后面答案是按$t_{\frac{0.05}{2}}(15)=2.1315$计算出来的，所查的表精度不同而已。

（2）均值μ未知，方差σ^2的置信度为$1-\alpha$的置信区间为

$$\left[\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)},\frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}\right]$$

查表得$\chi^2_{\frac{0.05}{2}}(15)=27.488,\chi^2_{1-\frac{0.05}{2}}(15)=6.262$，代入$n=16$，$s=300$得整批电子管的使用寿命的方差的置信度为 95%的置信区间为$[49112.34,215586.07]$

从而整批电子管的使用寿命的标准差的置信度为 95%的置信区间为$[221.61,464.31]$50 分

10. 设总体 $X \sim N(\mu,100)$，若置信度为 95%时，μ 的置信区间长度为 5，则样本容量 n 至少为多少？若置信度为 99%，样本容量 n 至少为多少？

解答：总体方差 $\sigma^2=100$ 已知时，均值的置信度为 $1-\alpha$ 的置信区间为

$$\left[\overline{X}-u_{\frac{\alpha}{2}}\frac{10}{\sqrt{n}},\overline{X}+u_{\frac{\alpha}{2}}\frac{10}{\sqrt{n}}\right]$$

置信区间长度为 $2\cdot u_{\frac{\alpha}{2}}\frac{10}{\sqrt{n}}=u_{\frac{\alpha}{2}}\frac{20}{\sqrt{n}}$

（1）当 $\alpha=0.05$ 时，查表得 $u_{\frac{0.05}{2}}=1.96$

由 $u_{\frac{\alpha}{2}}\frac{20}{\sqrt{n}}\le 5$ 解得 $n\ge 61.4656$，因此样本容量至少为 62

（2）当 $\alpha=0.01$ 时，查表得 $u_{\frac{0.05}{2}}=2.57$

由 $u_{\frac{\alpha}{2}}\frac{20}{\sqrt{n}}\le 5$ 解得 $n\ge 105.6784$，因此样本容量至少为 106

11. 假设某地旅游者的消费额 X 服从正态分布 $X \sim N(\mu,\sigma^2)$，且标准差 $\sigma=500$ 元，μ 未知，今要对该地旅游者的平均消费额 μ 加以估计，为了能以 95%的置信度来相信这种估计绝对误差小于 50 元，问至少要调查多少名游客？

解答：总体标准差 $\sigma=500$ 已知时，均值的置信度为 $1-\alpha$ 的置信区间为

$$\left[\overline{X}-u_{\frac{\alpha}{2}}\frac{500}{\sqrt{n}},\overline{X}+u_{\frac{\alpha}{2}}\frac{500}{\sqrt{n}}\right]$$ 50 分

这种估计的绝对误差不超过 $u_{\frac{\alpha}{2}}\frac{500}{\sqrt{n}}$

当 $\alpha=0.05$ 时，查表得 $u_{\frac{0.05}{2}}=1.96$

由 $u_{\frac{\alpha}{2}}\frac{500}{\sqrt{n}}<50$ 解得 $n>384.16$，因此至少要调查 385 名游客 50 分

12. 已知某种果树产量服从正态分布 $N(218,\sigma^2)$ ，随机抽取 6 棵计算其产量为（单位：公斤）

221,191,202,205,256,236

试以 95%的置信度估计果树常量的方差

解答：总体均值 $\mu=218$ 已知时，方差的置信度为 $1-\alpha$ 的置信区间为

$$\left[\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\frac{\alpha}{2}}(n)},\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\frac{\alpha}{2}}(n)}\right]$$

由题意， $n=6,\alpha=0.05$ ，查表得 $\chi^2_{\frac{0.05}{2}}(6)=14.449,\chi^2_{1-\frac{0.05}{2}}(6)=1.237$

计算得 $\sum_{i=1}^{6}(x_i-218)^2=2931$ ，因此果树常量方差的置信度为 95%的置信区间为 $[202.85,2369.44]$

13. 已知某种木材横纹抗压力的实验值服从正态分布，对 10 个试件做横纹抗压力试验得到数据如下（单位：kg/cm2）

482,493,457,510,446,435,418,394,496,480

试以 95%的置信度对该木材横纹抗压力的方差进行区间估计

解答：均值 μ 未知，方差 σ^2 的置信度为 $1-\alpha$ 的置信区间为

$$\left[\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)},\frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}\right]$$

查表得 $\chi^2_{\frac{0.05}{2}}(9)=19.023,\chi^2_{1-\frac{0.05}{2}}(9)=2.700$ ，计算得 $s^2=1411.878$

因此该种木材横纹抗压力的方差的置信度为 95%的置信区间为 $[667.98,4706.26]$

说明：教材上的答案是按照样本数据为“482,493,457,471,510,446,435,418,394,496”计算出来的，在这组数据下 $s^2=1382$

14.（本题为两个正态总体的估计问题，不在学习范围内）

15.（本题为两个正态总体的估计问题，不在学习范围内）

16.（本题为两个正态总体的估计问题，不在学习范围内）

17.（本题为非正态总体的估计问题，不在学习范围内）

18.（本题为非正态总体的估计问题，不在学习范围内）