第五章 数理统计初步

1.(Ch5-4) 设 X_1, X_2, X_3, X_4 是来自正态总体 $N(0,3^2)$ 的随机样本，$X=a(X_1-2X_2)^2+b(2X_3-3X_4)^2$。试确定 a、b 使统计量 X 服从 χ^2 分布，并指出其自由度。

分析：依题意，要使统计量 X 服从 χ^2 分布，则必需使 $a^{1/2}(X_1-2X_2)$ 及 $b^{1/2}(2X_3-3X_4)$ 服从标准正态分布。由相互独立的正态随机变量的性质知 $a^{1/2}(X_1-2X_2)\sim N(0,(9a+36a))$，从而解得 $a=1/45$。同理得 $b=1/117$。

解：X_1, X_2, X_3, X_4 是来自正态总体 $N(0,3^2)$ 的随机样本，所以，X_1, X_2, X_3, X_4 相互独立，由相互独立的正态随机变量的性质知

$a^{1/2}(X_1-2X_2)\sim N(0,(9a+36a))$，即 $a^{1/2}(X_1-2X_2)\sim N(0,\ 45a)$，

当 $a=1/45$ 时，$a^{1/2}(X_1-2X_2)\sim N(0,\ 1)$，

同理，当 $a=1/117$ 时，$b^{1/2}(2X_3-3X_4)\sim N(0,\ 1)$，

即，当 $a=1/45$，$a=1/117$ 时，$X\sim\chi^2(2)$，自由度为 2。

2.(Ch5-5) 设 X 和 Y 独立同分布 $N(0,\ 3^2)$，$X_1, X_2,\cdots, X_9$ 和 $Y_1, Y_2,\cdots, Y_9$ 分别是来自 X 和 Y 的简单抽样，试确定统计量 $U=\dfrac{X_1+\cdots+X_9}{\sqrt{Y_1^2+\cdots+Y_9^2}}$ 所服从的分布。

解：$U=\dfrac{X_1+\cdots+X_9}{\sqrt{Y_1^2+\cdots+Y_9^2}}=\dfrac{(X_1/3+\cdots+X_9/3)\ /3}{\sqrt{(Y_1^2/9+\cdots+Y_9^2/9)\ /9}}\sim t(9)$。

3.(Ch5-6) 设随机变量 $X\sim t(n)\ (n>1)$，试确定统计量 $Y=\dfrac{1}{X^2}$ 所服从的分布。

分析：先由 t 分布的定义知 $X=\dfrac{U}{\sqrt{V/n}}$，其中 $U\sim N(0,1), V\sim\chi^2(n)$，再将其代入 $Y=\dfrac{1}{X^2}$，然后利用 F 分布的定义即可。

解：由题设知，$X=\dfrac{U}{\sqrt{V/n}}$，其中 $U\sim N(0,1), V\sim\chi^2(n)$，于是

$Y=\dfrac{1}{X^2}=\dfrac{V/n}{U^2}=\dfrac{V/n}{U^2/1}$，这里 $U^2\sim\chi^2(1)$，根据 F 分布的定义知 $Y=\dfrac{1}{X^2}\sim F(n,1)$。

4.(Ch5-7) 设总体 X 服从正态分布 $N(0,\ 2^2)$，而 $X_1, X_2,\cdots, X_{15}$ 是来自总体

X的简单随机样本，试确定随机变量$Y=\frac{X_1^2+\cdots+X_{10}^2}{2(X_{11}^2+\cdots+X_{15}^2)}$所服从的分布。

解：由于$(X_1^2/4+\cdots+X_{10}^2/4)/10\sim\chi^2(10)$，

$(X_{11}^2/4+\cdots+X_{15}^2/4)/5\sim\chi^2(5)$

$$Y=\frac{X_1^2+\cdots+X_{10}^2}{2(X_{11}^2+\cdots+X_{15}^2)}=\frac{(X_1^2/4+\cdots+X_{10}^2/4)/10}{(X_{11}^2/4+\cdots+X_{15}^2/4)/5}$$

故$Y=\frac{X_1^2+\cdots+X_{10}^2}{2(X_{11}^2+\cdots+X_{15}^2)}\sim F(10,\quad 5)$。

5．设总体$X\sim N(\mu,\quad\sigma^2)$，$X_1,X_2,\cdots,X_n$是来自的样本，记$\overline{X}$为样本均值，$S^2$为样本方差，$X_{n+1}$是对$X$的又一独立观测值，试证明统计量$T=\frac{X_{n+1}-\overline{X}}{S}\sqrt{\frac{n}{n+1}}$服从$t$分布，自由度为$n-1$。

解：由于$\overline{X}\sim N(\mu,\quad\sigma^2/n)$，$X_{n+1}\sim N(\mu,\sigma^2)$，从而

$X_{n+1}-\overline{X}\sim N(0,(n+1)\sigma^2/n)$，因此有$U=\frac{X_{n+1}-\overline{X}}{\sigma}\sqrt{\frac{n}{n+1}}\sim N(0,1)$，

$V=\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$，由于$\overline{X}$与$S^2$相互独立，$X_1,X_2,\cdots,X_n$与 X_{n+1}相互独立，所以$\overline{X}$，X_{n+1}，S^2相互独立，从而$X_{n+1}-\overline{X}$与S^2相互独立，故

$$T=\frac{X_{n+1}-\overline{X}}{S}\sqrt{\frac{n}{n+1}}=\frac{U}{\sqrt{V/(n-1)}}\sim t(n-1)$$

6．设$X_1,X_2,\cdots,X_{n+m}$（$n>m$）独立同分布，且有有限方差，试求$Y=\sum_{i=1}^{n}X_i$与$Z=\sum_{k=1}^{n}X_{m+k}$的相关系数。

解：设$EX_k=a$，$DX_k=\sigma^2$，则

$$Cov(Y,Z)=E(Y-EY)(Z-EZ)$$
$$=E[\sum_{i=1}^{n}(X_i-a)][\sum_{k=1}^{n}(X_{m+k}-a)$$
$$=E[\sum_{k=1}^{n-m}(X_{m+k}-a)^2]=(n-m)\sigma^2$$

又$DY=DZ=n\sigma^2$，所以

$$\rho_{YZ}=\frac{Cov(Y,Z)}{\sqrt{DY}\sqrt{DZ}}=\frac{n-m}{n}。$$

7．设$X_1,X_2,\cdots X_8$是来自正态总体$N(\theta,\theta^2)$的样本，记$\overline{X}_1=\frac{1}{4}\sum_{i=1}^{4}x_i$，

$\overline{X}_2=\frac{1}{4}\sum_{i=5}^{8}x_i$ ，$S_1^2=\frac{1}{4}\sum_{i=1}^{4}(X_i-\overline{X}_1)^2$ ，$S_2^2=\frac{1}{4}\sum_{i=5}^{8}(X_i-\overline{X}_2)^2$ ，已知参数$\theta>0$，求 $\frac{\overline{X}_1+\overline{X}_2-2\theta}{\sqrt{S_1^2+S_2^2}}\sqrt{3}$ 的分布。

解：由条件知$X_i\sim N(\theta,\theta^2)$，从而$\overline{X}_i\sim N(\theta,\ \frac{\theta^2}{4})$

$$\overline{X}_1+\overline{X}_2\sim N(2\theta,\ \frac{\theta^2}{2})$$

故
$$\frac{\overline{X}_1+\overline{X}_2-2\theta}{\sqrt{\theta^2/2}}\sim N(0,\ 1)$$

又
$$\frac{4S_1^2}{\theta^2}\sim\chi^2(3)\qquad\qquad \frac{4S_2^2}{\theta^2}\sim\chi^2(3)$$

由独立性
$$\frac{4S_1^2+4S_2^2}{\theta^2}\sim\chi^2(6)$$

故
$$\frac{\dfrac{\overline{X}_1+\overline{X}_2-2\theta}{\sqrt{\theta^2/2}}}{\sqrt{\dfrac{4S_1^2+4S_2^2}{\theta^2}/6}}=\frac{\overline{X}_1+\overline{X}_2-2\theta}{\sqrt{S_1^2+S_2^2}}\sqrt{3}\sim t(6)$$

8．设$X_1,X_2,\cdots,X_n$为来自正态总体的一个样本，求下述各总体的密度函数中的未知参数的矩估计及极大似然估计。

（1） $f(x,\theta)=\begin{cases}(\theta+1)x^{\theta},0<x<1,\\ 0,\qquad\qquad \text{其它},\end{cases}$

其中$\theta>-1$为未知参数。

（2） $f(x)=\begin{cases}\dfrac{x}{\theta^2}e^{-x^2/2\theta^2}, & x>0,\\ \quad 0, & \text{其它},\end{cases}$

其中，$\theta>0,\theta$为未知参数。

分析：矩估计法和极大似然估计法是点估计的两种常用方法，所谓矩估计法就是用样本的某种矩作为总体的相应矩的估计，因此需要首先计算（或已知）总体的某（几）种矩，由于本例只涉及一个未知参数，故只要知道总体的某一种矩即可。极大似然估计可依据内容提要中的四个步骤来完成，其关键是正确构造似然函数。

解：(1) 矩估计：由于

$$EX=\int_{-\infty}^{+\infty}xf(x)dx=\int_0^1(\theta+1)x^{\theta+1}dx=\frac{\theta+1}{\theta+2}$$

设 $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ 为样本均值，令 $\frac{\theta+1}{\theta+2}=\overline{X}$，解得未知参数 θ 的矩估计量为 $\hat{\theta}=\frac{2\overline{X}-1}{1-\overline{X}}$。

极大似然估计：设 x_1，x_2，$\cdots$，x_n 为 X_1，X_2，$\cdots$，X_n 观测值，构造似然函数

$$L(\theta)=\prod_{i=1}^{n}(\theta+1)x_i^{\theta}=(\theta+1)^n(\prod_{i=1}^{n}x_i)^{\theta}$$

$$\ln L=n\ln(\theta+1)+\theta\sum_{i=1}^{n}\ln x_i$$

令

$$\frac{d\ln L}{d\theta}=\frac{n}{\theta+1}+\sum_{i=1}^{n}\ln x_i=0$$

解得 θ 的极大似然估计量为 $\hat{\theta}=-1-\frac{n}{\sum_{i=1}^{n}\ln X_i}$。

(2) 矩估计

$$EX=\int_0^{+\infty}x\frac{x}{\theta^2}e^{-x^2/2\theta^2}dx=-\int_0^{+\infty}xde^{-x^2/2\theta^2}=\theta\sqrt{\frac{\pi}{2}}$$

令 $EX=\theta\sqrt{\frac{\pi}{2}}=\overline{X}$，解得 $\hat{\theta}=\overline{X}\sqrt{\frac{2}{\pi}}$。

极大似然估计：构造似然函数

$$L(\theta)=\prod_{i=1}^{n}\frac{x_i}{\theta^2}e^{-x_i^2/2\theta^2}=\frac{1}{\theta^{2n}}\prod_{i=1}^{n}x_i e^{-x_i^2/2\theta^2}$$

$$\ln L=-\ln\theta^{2n}+\sum_{i=1}^{n}\ln x_i-\sum_{i=1}^{n}\frac{x_i^2}{2\theta^2}$$

令 $$\frac{d\ln L}{d\theta}=-2n/\theta+\frac{1}{\theta^3}\sum_{i=1}^{n}x_i^2=0$$

解得 $$\hat{\theta}=\sqrt{\frac{1}{2n}\sum_{i=1}^{n}x_i^2}$$。

9.(Ch5-12)设 X_1，X_2，$\cdots$，X_n 为总体 X 的一个样本，且 X 服从参数为 m, p 的二项分布，求 p 的极大似然估计量。

解：设 x_1，x_2，$\cdots$，x_n 为 X_1，X_2，$\cdots$，X_n 观测值，则构造似然函数

$$L\ (p)=\prod_{i=1}^{n}p(x_i,p)=\prod_{i=1}^{n}C_m^{x_i}p^{x_i}(1-p)^{m-x_i}$$

$$=\prod_{i=1}^{n}C_m^{x_i}\times p^{\sum_{i=1}^{n}x_i}\times(1-p)^{nm-\sum_{i=1}^{n}x_i}$$

$$\ln L=\sum\ln C_m^{x_i}+\ln p\sum_{i=1}^{n}x_i+(nm-\sum_{i=1}^{n}x_i)\ln(1-p)$$

$$\frac{d\ln L}{dp}=\frac{\sum_{i=1}^{n}x_i}{p}-\frac{nm-\sum_{i=1}^{n}x_i}{1-p}$$

令 $\frac{d\ln L}{dp}=0$，解得 $p=\frac{1}{n}\sum_{i=1}^{n}x_i/m$，

因此 p 的极大似然估计量为

$$\hat{p}=\overline{X}/m。$$

10.(Ch5-15) 设某种元件的使用寿命 X 的概率密度为

$$f(x;\theta)=\begin{cases}2e^{-2(x-\theta)}, & x>\theta\\ 0, & x\le\theta\end{cases},$$

其中 $\theta>0$ 为未知参数。又设 x_1，x_2,$\cdots$，x_n 是 X 的一组样本观察值，求 θ 的极大似然估计值。

解：构造似然函数 $L(\theta)=\prod_{i=1}^{n}2e^{-2(x_i-\theta)}=2ne^{-2\sum_{i=1}^{n}(x_i-\theta)}$

$$\ln L=\ln 2n-2\sum_{i=1}^{n}(x_i-\theta)\ ,$$

$$\frac{d\ln L}{d\theta}=2n\ （与参数\theta无关）,$$

由条件，当 $x>\theta$ 时，$f(x)=2e^{-2(x-\theta)}$ （$\theta>0$），所以当 $\theta=\min(x_1,x_2,\cdots,x_n)$ 时，似然函数 L 取得最大值，从而知 $\hat{\theta}=\min(x_1,x_2,\cdots,x_n)$。

11．设总体 X 的概率密度为

$$f(x)=\begin{cases}2e^{-2(x-\theta)}, x>\theta,\\ 0, \quad x\le\theta,\end{cases}$$

其中$\theta>0$是未知参数．从总体 X 中抽取简单随机样本$X_1,X_2,\cdots,X_n$，记$\hat{\theta}=\min(X_1,X_2,\cdots,X_n)$

(1)求总体 X的分布函数 $F(x)$;。

(2)求统计量$\hat{\theta}$的分布函数$F_{\hat{\theta}}(x)$；

(3)如果用$\hat{\theta}$作为θ的估计量，讨论它是否具有无偏性．

分析：求分布函数 $F(x)$是基本题型；求统计量$\hat{\theta}$的分布函数$F_{\hat{\theta}}(x)$，可作为多维相互独立且同分布的随机变量函数求分布函数，直接用定义即可；是否具有无偏性，只需检验$E\hat{\theta}=\theta$是否成立．

解：(1)

$$F(x)=\int_{-\infty}^{x}f(t)dt=\begin{cases}1-e^{-2(x-\theta)}, x>\theta,\\ \quad 0, \qquad x\le\theta.\end{cases}$$

(2) $F_{\hat{\theta}}(x)=P\{\hat{\theta}\le x\}=P\{\min(X_1,X_2,\cdots,X_n)\le x\}$

$$=1-P\{\min(X_1,X_2,\cdots,X_n)>x\}$$

$$=1-P\{X_1>x,X_2>x,\cdots,X_n>x\}$$

$$=1-[1-F(x)]^n$$

$$=\begin{cases}1-e^{-2n(x-\theta)}, x>\theta,\\ \quad 0, \qquad x\le\theta.\end{cases}$$

(3) $\hat{\theta}$概率密度为

$$f_{\hat{\theta}}(x)=\frac{dF_{\hat{\theta}}(x)}{dx}=\begin{cases}2ne^{-2n(x-\theta)}, x>\theta,\\ \quad 0, \qquad x\le\theta.\end{cases}$$

因为 $E\hat{\theta}=\int_{-\infty}^{+\infty}xf_{\hat{\theta}}(x)dx=\int_{\theta}^{+\infty}2nxe^{-2n(x-\theta)}dx$

$$=\theta+\frac{1}{2n}\ne\theta\text{。}$$

所以$\hat{\theta}$作为θ的估计量不具有无偏性。

评注:本题表面上是一数理统计问题，实际上考查了求分布函数、随机变量的函数求分布和概率密度以及数学期望的计算等多个知识点．将数理统计的概念与随机变量求分布与数字特征结合起来是一种典型的命题形式．

12.(Ch3-16)设总体X的概率分布为

X	0	1	2	3
P	θ^2	$2\theta(1-\theta)$	θ^2	$1-2\theta$

其中θ $(0<\theta<\frac{1}{2})$是未知参数，利用总体X的如下样本值 3，1，3，0，3，1，2，3，求θ的矩估计值和极大似然估计值。

解：$EX = 0\times\theta^2 + 1\times 2\theta(1-\theta) + 2\times\theta^2 + 3\times(1-2\theta) = 3-4\theta$

$$\bar{x} = \frac{1}{8}(3+1+3+0+3+1+2+3) = 2$$

令 $EX = \bar{x}$，即 $3-4\theta = 2$，解得 θ 的矩估计值为 $\hat{\theta} = 1/4$。

对于给定的样本值，似然函数为

$$L(\theta) = 4\theta^6(1-\theta)^2(1-2\theta)^4$$

$$\ln L(\theta) = \ln 4 + 6\ln\theta + 2\ln(1-\theta) + 4\ln(1-2\theta)$$

$$\frac{d\ln L(\theta)}{d\theta} = \frac{6}{\theta} - \frac{2}{1-\theta} - \frac{8}{1-2\theta} = \frac{6-28\theta+24\theta^2}{\theta(1-\theta)(1-2\theta)}$$

令 $\frac{d\ln L(\theta)}{d\theta} = 0$，解得

$$\theta_{1,2} = \frac{7\pm\sqrt{13}}{12}$$

因 $\frac{7+\sqrt{13}}{12} > \frac{1}{2}$ 不合题意，所以 θ 的极大似然估计值为

$$\hat{\theta} = \frac{7-\sqrt{13}}{12}。$$

13.(Ch5-20) 随机的从 A 批导线中抽取 4 根，又从 B 批导线中抽取 5 根，测得电阻（Ω）为

A 批导线：0.143，0.142，0.143，0.137，

B 批导线：0.140，0.142，0.136，0.138，0.140，

设测定数据分别来自正态总体 $N(\mu_1,\ \sigma^2)$、$N(\mu_2,\ \sigma^2)$，且两样本相互独立。又 μ_1、μ_2、σ^2 均为未知，试求 $\mu_1-\mu_2$ 的置信度为 95%的置信区间。

解:取样本函数

$$T = \frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{S_w\sqrt{\frac{1}{m}+\frac{1}{n}}} \sim t(m+n-2)$$

又知 $1-\alpha = 0.95$，得 $\alpha = 0.05$，$m+n-2 = 7$,查表得 $\lambda = 2.3646$，经计算 $\bar{X}_A = 0.14125$，$S_A^2 = 8.25\times10^{-6}$，$\bar{X}_A = 0.1392$，$S_A^2 = 5.29\times10^{-6}$

由 $S_W^2 = \frac{(m-1)S_X^2+(n-1)S_Y^2}{m+n-2}$，得 $S_W \approx 0.00255$,从而得 $\mu_A-\mu_B$ 的 0.95 的置信区间为（-0.002，0.006）。

14.(Ch5-21) 设两位化验员 A、B 独立地对某种聚合物含氯量用同样的方法

各作 10 次测定，其测定的样本方差依次为$S_A^2=0.5419$，$S_B^2=0.6065$，设σ_A^2、σ_B^2分别为A、B所测定的测定值总体的方差，两总体均服从正态分布。试求方差比σ_A^2/σ_B^2的置信度为 95%的置信区间。

解：取样本函数

$$F=\frac{S_1^2 / S_2^2}{\sigma_1^2 / \sigma_2^2}\sim F(9,9)$$

查表λ_2=4.03，$\lambda_1=1/4.03$，得方差比σ_1^2 / σ_2^2的置信度为 0.90 的置信区间是（0.222,3.601）。

15.(Ch5-23) 原铸造成品率的平均值为 83.8%，今换用便宜的原料，成品率抽样数据（%）如下：83.9，84.6，82.4，84.1，84.9，82.9，85.2，83.3，82.0，83.5，问原料代用后，成品率是否发生了变化？（α=0.05）

分析：依题意，可认为成品率这样的计量值数据服从正态分布，因此该问题即为方差未知的情况下，检验成品率的平均值是否仍为 83.8%。

解：依题意需要检验假设$H_0:\mu=83.8$，故选取统计量：

$$T=\frac{\overline{X}-83.8}{S/\sqrt{10}}\sim t\ (10-1)\quad (H_0:\mu=83.8\text{成立时})$$

对于给定的显著性水平α=0.05，由$P\{T>\lambda\}=\alpha/2=0.025$，查$t$ (9) 分布表得临界值λ=2.262。经计算得$\bar{x}=83.68$，$S^2=\dfrac{10.116}{9}$，

$$t=\frac{83.68-83.8}{\sqrt{10.116/(9\times 10)}}\approx -0.36\text{。}$$

由于$|t|=0.36<2.262$，从而应接受假设H_0。即原料代用后，成品率无显著变化。

16.(Ch5-25) 机床厂某日从两台机器生产的同一零件中，分别抽取若干个样品测量的长度如下

第一台机器：6.2，5.7，6.5，6.0，6.3，5.8，5.7，6.0，6.0，5.8，6.0；

第一台机器：5.6，5.9，5.6，5.7，6.0，5.8，5.7，5.5，5.5。

问这两台机器的加工精度有无显著差异（α=0.05）？

分析：依题意，可认为样品测量这样的计量值数据服从正态分布，因此比较两台机器的加工精度有无显著差异，即为检验假设$H_0:\sigma_1^2=\sigma_2^2$成立与否。

解：依题意，需检验假设 $H_0:\sigma_1^2=\sigma_2^2$，

故取统计量 $F=\dfrac{S_1^2}{S_2^2}$。

则当 H_0 成立时， $F\sim F(10,8)$。

对于给定的显著性水平 $\alpha=0.05$，由

$$P\{F>\lambda_2\}=\frac{\alpha}{2}=0.025$$

查 $F(10,8)$ 分布表，得临界值 $\lambda_2=4.30$。

由 $$P\{\frac{1}{F}>\lambda_0=\frac{1}{\lambda_1}\}=\frac{\alpha}{2}=0.025$$

查 $F(8,10)$ 分布表，得 $\lambda_0=3.85$，从而确定临界值 $\lambda_1=0.26$。

而 $$\overline{x}=6.0,\quad S_1^2=0.064,\quad \overline{y}=5.7,\quad S_2^2=0.03,$$

$$F=\frac{S_1^2}{S_2^2}=\frac{0.064}{0.03}\approx 2.133,$$

由于 $\lambda_1=0.26<F=2.133<\lambda_2=4.30$,故应接受假设 H_0,即认为两台机器的加工精度无显著差异。

17.(Ch5-27) 假设 0.50，1.25，0.80，2.00 是来自总体 X 的简单随机样本值。已知 $Y=\ln x$ 服从正态分布 $N(\mu,1)$。(1) 求 X 的数学期望 EX （记 EX 为 b）；(2) 求 μ 的置信度为 0.95 的置信区间；(3) 利用上述结果求 b 的置信度为 0.95 的置信区间。

解：(1) $$EX=Ee^Y=\int_{-\infty}^{+\infty}e^y\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(y-\mu)^2}dy=e^{\mu}\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(y-\mu)^2}de^{y-\mu}$$

$$\overset{y-\mu=t}{=}e^{\mu}\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t^2}e^t dt=e^{\mu}\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(t-1)^2}e^{\frac{1}{2}}dt$$

$$\overset{z=t-1}{=}e^{\mu+\frac{1}{2}}\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}dz=e^{\mu+\frac{1}{2}}$$

(2) 由 X： 0.50， 1.25， 0.80， 2.00 得

$Y=\ln x$： -0.693，0.223，-0.223，0.693。取样本函数

$$U=\frac{\overline{Y}-\mu}{\sqrt{1/4}}\sim N(0,1)$$

对给定$\alpha=1-0.95=0.05$查标准正态分布表得临界值$\lambda=1.96$，又$\overline{Y}=0$，从而得μ的置信度为 0.95 的置信区间为

$$(\overline{Y}-\lambda\sqrt{\frac{1}{4}},\overline{Y}+\lambda\sqrt{\frac{1}{4}})=(-0.98,0.98)$$

（3）由上述结果知

$$P\{-0.98<\mu<0.98\}=0.95$$

即 $$P\{e^{-0.98+0.5}<e^{\mu+0.5}<e^{0.98+0.5}\}=0.95$$

亦即 $$P\{e^{-0.98+0.5}<b<e^{0.98+0.5}\}=P\{0.619<b<4.39\}=0.95$$

故b的置信度为 0.95 的置信区间为$(0.619,\quad 4.39)$。

18.(Ch5-28) 设总体X服从正态分布$N(\mu,\quad\sigma^2)$ $(\sigma>0)$，从总体中抽取简单随机样本$X_1,X_2,\cdots,X_{2n}$ $(n>2)$，其样本均值为$\overline{X}=\frac{1}{2n}\sum_{i=1}^{2n}X_i$，求统计量$Y=\sum_{i=1}^{n}(X_i+X_{n+i}-2\overline{X})^2$的数学期望$EY$。

解法 1：考虑$(X_1+X_{n+1}),(X_2+X_{n+2}),\cdots,(X_n+X_{2n})$，将其视为取自总体$N(2\mu,\quad 2\sigma^2)$的简单随机样本，则其样本均值、样本方差分别为

$$\frac{1}{n}\sum_{i=1}^{n}(X_i+X_{n+i})=\frac{1}{n}\sum_{i=1}^{2n}X_i=2\overline{X}\quad,\qquad\qquad\frac{1}{n-1}Y$$

由于$E(\frac{1}{n-1}Y)=2\sigma^2$，所以$EY=(n-1)(2\sigma^2)=2(n-1)\sigma^2$

解法 2：记$\overline{X}'=\frac{1}{n}\sum_{i=1}^{n}X_i$，$\overline{X}''=\frac{1}{n}\sum_{i=1}^{n}X_{n+i}$，显然有$2\overline{X}=\overline{X}'+\overline{X}''$。因此

$$EY=E[\sum_{i=1}^{n}(X_i+X_{n+i}-2\overline{X})^2]=E\{\sum_{i=1}^{n}[(X_i-\overline{X}')+(X_{n+i}-\overline{X}'')]^2\}$$

$$=E\{\sum_{i=1}^{n}[(X_i-\overline{X}')^2+2(X_i-\overline{X}')(X_{n+i}-\overline{X}'')+(X_{n+i}-\overline{X}'')^2]\}$$

$$=E[\sum_{i=1}^{n}[(X_i-\overline{X}')^2]+0\quad E[\sum_{i=1}^{n}[(X_{n+i}-\overline{X}'')^2]$$

$$=(n-1)\sigma^2+(n-1)\sigma^2=2(n-1)\sigma^2。$$

19.(Ch5-29)设随机变量 X 的分布函数为 $F(x,\alpha,\beta)=\begin{cases}1-(\alpha/x)^{\beta}, & x>\alpha,\\ 0, & x\le\alpha,\end{cases}$

其中参数 $\alpha>0,\beta>1$。设 $X_1,X_2,\cdots,X_n$ 为来自总体 X 的简单随机样本，求:

(1) 当 $\alpha=1$ 时，求未知参数 β 的矩估计量；

(2) 当 $\alpha=1$ 时,求未知参数 β 的最大似然估计量；

(3) 当 $\beta=2$ 时，求未知参数 α 的最大似然估计量。

分析：本题是一个常规题型，只要注意求连续型总体未知参数的矩估计和最大似然估计都须已知密度函数，从而先由分布函数求导得密度函数。

解： 当 $\alpha=1$ 时，X 的概率密度为

$$f(x,\beta)=\begin{cases}\dfrac{\beta}{x^{\beta+1}}, & x>1,\\ 0, & x\le 1,\end{cases}$$

(1) 由于

$$EX=\int_{-\infty}^{+\infty}xf(x;\beta)dx=\int_{1}^{+\infty}x\cdot\frac{\beta}{x^{\beta+1}}dx=\frac{\beta}{\beta-1},$$

令 $\dfrac{\beta}{\beta-1}=\overline{X}$， 解得 $\beta=\dfrac{\overline{X}}{\overline{X}-1}$，

所以，参数 β 的矩估计量为 $\beta=\dfrac{\overline{X}}{\overline{X}-1}$。

(2) 对于总体 X 的样本值 $x_1,x_2,\cdots,x_n$，似然函数为

$$L(\beta)=\prod_{i=1}^{n}f(x_i;\alpha)=\begin{cases}\dfrac{\beta^n}{(x_1x_2\cdots x_n)^{\beta+1}}, & x_i>1(i=1,2,\cdots,n),\\ 0, & \text{其他}.\end{cases}$$

当 $x_i>1(i=1,2,\cdots,n)$ 时，$L(\beta)>0$，取对数得

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

对 β 求导数，得

$$\frac{d[\ln L(\beta)]}{d\beta}=\frac{n}{\beta}-\sum_{i=1}^{n}\ln x_i\ ,$$

令 $\dfrac{d[\ln L(\beta)]}{d\beta}=\dfrac{n}{\beta}-\sum\limits_{i=1}^{n}\ln x_i=0$， 解得 $\beta=\dfrac{n}{\sum\limits_{i=1}^{n}\ln x_i}$，

于是 β 的最大似然估计量为

$$\hat{\beta}=\frac{n}{\sum\limits_{i=1}^{n}\ln x_i}\text{。}$$

（3）当 $\beta=2$ 时，X 的概率密度为

$$f(x,\beta)=\begin{cases}\dfrac{2\alpha^2}{x^3}, & x>\alpha, \\ 0, & x\le\alpha,\end{cases}$$

对于总体 X 的样本值 $x_1,x_2,\cdots,x_n$，似然函数为

$$L(\alpha)=\prod_{i=1}^{n}f(x_i;\alpha)=\begin{cases}\dfrac{2^n\alpha^{2n}}{(x_1x_2\cdots x_n)^3}, & x_i>\alpha(i=1,2,\cdots,n), \\ 0, & \text{其他}.\end{cases}$$

当 $x_i>\alpha(i=1,2,\cdots,n)$ 时，α 越大，$L(\alpha)$ 越大，即 α 的最大似然估计值为

$$\hat{\alpha}=\min\{x_1,x_2,\cdots,x_n\},$$

于是 α 的最大似然估计量为

$$\hat{\alpha}=\min\{X_1,X_2,\cdots,X_n\}。$$