习题 1 绪论

1、试举出一个有限总体的例子. 并指出其概率分布.

略

2、试举出一个无限总体的例子. 并指出其概率分布.

略

3、一个总体有N个元素，其指标分别为$a_1 > a_2 > \cdots > a_N$，指定自然数$M < N, n < N$,

并设$m = nM/N$为整数. 在$(a_1, \cdots, a_M)$中不放回地随机抽出m个，在$(a_{M+1}, \cdots, a_N)$中不放回

地随机抽出$n-m$个. 写出所得样本的分布.

解:

设所得样本为$(X_1, \cdots, X_n)$,则前m个个体$(X_1, \cdots, X_m)$在$(a_1, \cdots, a_M)$中不放回抽取

共有$A_M^m = \dfrac{M!}{m!}$种等可能结果，后$n-m$个个体$(X_{m+1}, \cdots, X_n)$在$(a_{M+1}, \cdots, a_N)$中抽取

共有$A_{N-M}^{n-m} = \dfrac{(N-M)!}{(n-m)!}$种等可能结果.故$(X_1, \cdots, X_n)$共有$A_M^m \cdot A_{N-M}^{n-m}$种等可能结果.

故$P(X_1 = x_1, \cdots, X_m = x_m, X_{m+1} = x_{m+1}, \cdots, X_n = x_n) = \dfrac{1}{A_M^m \cdot A_{N-M}^{n-m}} = \dfrac{m!\,(n-m)!}{M!\,(N-M)!}$.

4、一物体的重量a未知，有两架天平可用，其随机误差分别服从正态分布$N(0, \sigma_1^2)$和

$N(0.\sigma_2^2), \sigma_1^2$和$\sigma_2^2$都未知. 先把物件在第一架天平上称两次得$X_1, X_2$, 再在第二架天平上称

两次得$X_3 X_4$, 然后视$|X_1 - X_2| \leqslant |X_3 - X_4|$或否而在第一架天平 或第二架天平上再称

$n-4$次得$X_5, \cdots, X_n$. 写出$(X_1, \cdots, X_n)$的密度.

解:

由题意可知$X_1, X_2 \sim N(a, \sigma_1^2), f_X(x) = \dfrac{1}{\sqrt{2\pi}\,\sigma_1} e^{-\frac{(x-a)^2}{\sigma_1^2}}$,

$X_3, X_4 \sim N(a, \sigma_2^2), f_X(x) = \dfrac{1}{\sqrt{2\pi}\,\sigma_2} e^{-\frac{(x-a)^2}{\sigma_2^2}}$

当$|X_1 - X_2| \leqslant |X_3 - X_4|$时$X_i \sim N(a, \sigma_1^2), f_{X_i}(x) = \dfrac{1}{\sqrt{2\pi}\,\sigma_1} e^{-\frac{(x-a)^2}{\sigma_1^2}}$

当$|X_1 - X_2| \geqslant |X_3 - X_4|$时$X_i \sim N(a, \sigma_2^2), f_{X_i}(x) = \dfrac{1}{\sqrt{2\pi}\,\sigma_2} e^{-\frac{(x-a)^2}{\sigma_2^2}}$,

又因为$X_1, \cdots, X_n$相互独立,

故$f_{X_1, \cdots, X_n}(x_1, \cdots, x_n) =$

$$\begin{cases} \dfrac{1}{\left(\sqrt{2\pi}\right)^n \sigma_1^{n-2}\sigma_2^2}\exp\left\{-\dfrac{\sum\limits_{i=1}^{2}(x_i-a)^2+\sum\limits_{k=5}^{n}(x_k-a)^2}{\sigma_1^2}-\dfrac{\sum\limits_{j=3}^{4}(x_j-a)^2}{\sigma_2^2}\right\}, |X_1-X_2|\leqslant|X_3-X_4| \\ \dfrac{1}{\left(\sqrt{2\pi}\right)^n \sigma_1^{2}\sigma_2^{n-2}}\exp\left\{-\dfrac{\sum\limits_{i=1}^{2}(x_i-a)^2}{\sigma_1^2}-\dfrac{\sum\limits_{j=3}^{n}(x_j-a)^2}{\sigma_2^2}\right\} , |X_1-X_2|\geqslant|X_3-X_4| \end{cases}$$

5、设总体X服从两点分布$b(1,p)$ (即$P(X=1)=p,P(x=0)=1-p$),其中p是未知参数，

$\boldsymbol{X}=(X_1,\cdots,X_5)$为从此总体中抽取的简单样本.

（1）写出样本空间$\mathscr{X}$和$\boldsymbol{X}$的概率分布.

（2）指出$X_1+X_2,\min\limits_{1\leqslant i\leqslant 5}X_i,X_5+2p,X_5-E(X_1),(X_5-X_1)^2/D(X_1)$哪些是统计量，哪些不是，为什么？

（3）设样本观察值$(X_1,\cdots,X_n)$中有m个1，$n-m$个0，求此样本的经验分布函数.

解：

(1)、$\mathscr{X}=\{(X_1,\cdots,X_5)$：$X_i=0$或$1$，$i=1,2,3,4,5\},\sum\limits_{i=1}^{5}X_i\sim b(5,p)$

(2)、X_1+X_2和$\min\limits_{1\leqslant i\leqslant 5}X_i$是统计量，$X_5+2p,X_5-E(X_1),(X_5-X_1)^2/D(X_1)$不是统计量，因为它们的值还和未知参数$p$有关.

(3)、$F_{(n)}(x)=\begin{cases} 0, & x<0 \\ \dfrac{n-m}{n}, & 0\leqslant x<1 \\ 1, & 1\leqslant x \end{cases}$

6、设$a\neq 0$和b皆为常数，令$y_i=ax_i+b,i=1,2,\cdots,n$.试证明$y_1,\cdots,y_n$的样本均值$\bar{y}$,

样本方差S_y^2和$x_1,\cdots,x_n$的样本均值$\bar{x}$,样本方差S_x^2之间存在下列关系：

$$\bar{y}=a\bar{x}+b,\quad S_y^2=a^2S_x^2.$$

根据这个结果，利用适当的变换，求下列一组数据的样本均值和样本方差.

480, 550，500, 590, 510, 560, 490, 600, 580.

解：

$$\bar{y}=\frac{1}{n}\sum_{i=1}^{n}y_i=\frac{1}{n}\left(a\sum_{i=1}^{n}x_i+nb\right)=a\bar{x}+b$$

$$S_y^2=\frac{1}{n-1}\sum_{i=1}^{n}\left(y_i-\bar{y}\right)^2=\frac{1}{n-1}\sum_{i=1}^{n}\left(ax_i+b-a\bar{x}-b\right)^2=\frac{a^2}{n-1}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2=a^2S_x^2$$

则对此样本数据作下列变换$y_i=10x_i+500, i=1,2,\cdots,9$

则$x_i, i=1,2,\cdots,9$依次为$-2,5,0,9,1,6,\ -1,10,8$，则$\bar{x}=4$，$S_x^2=21$

故$\bar{y}=10\bar{x}+500=540$，$S_y^2=100S_x^2=2100$，故样本均值为540，样本方差为2100.

7、设容量为 10 的样本观察值为

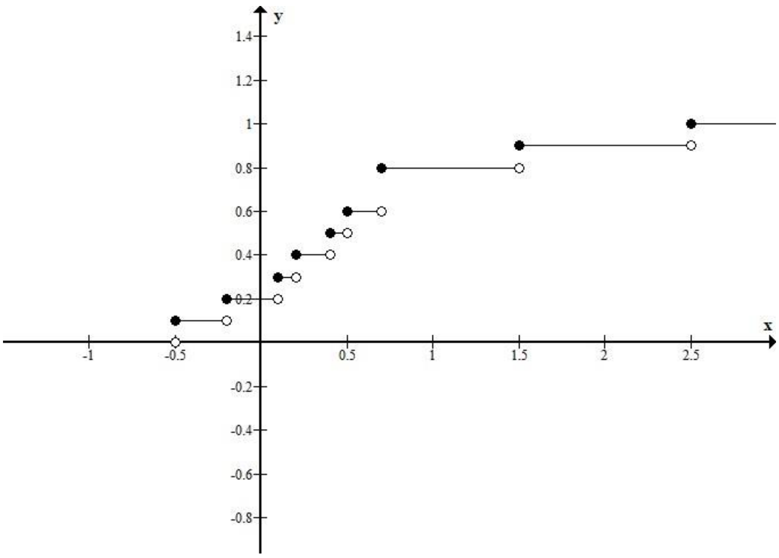
0.5, 0.7, 0.2, 0.7, 0.4, 2.5, 1.5，−0.2，−0.5, 0.1.

试绘出经验分布函数的图形.

解：

$$经验分布函数f(x)=\begin{cases}0, & x<-0.5\\ 0.1, & -0.5\leqslant x<0.2\\ 0.2, & -0.2\leqslant x<0.1\\ 0.3, & 0.1\leqslant x<0.2\\ 0.4, & 0.2\leqslant x<0.4\\ 0.5, & 0.4\leqslant x<0.5\\ 0.6, & 0.5\leqslant x<0.7\\ 0.8, & 0.7\leqslant x<1.5\\ 0.9, & 1.5\leqslant x<2.5\\ 1, & 2.5\leqslant x\end{cases}$$

图形如图所示：



8、设$X_1,\cdots,X_n$是取自总体分布函数为$F(x)$的样本，$F_n(x)$为其经验分布函数.证明对任意给定的实数x,

$$P\left(\lim_{n\to\infty}F_n(x)=F(x)\right)=1$$

$$P\left(\lim_{n\to\infty}F_n(x+0)=F(x+0)\right)=1$$

解：

令随机变量$Y_i=I_{(-\infty,\ x]}(X_i),i=1,\cdots,n$,则易见$Y_1,\cdots,Y_n\ i.i.d.\sim b(1,F(x))$,即$Y_1,\cdots,Y_n$为独立同分布的伯努利随机变量序列，$P(Y_i=0)=1-F(x),P(Y_i=1)=F(x)$,且有$F_n(x)=\frac{1}{n}\sum_{i=1}^{n}Y_i$,

由$Borel$强大数定律可知$F_n(x)\xrightarrow{a.s.}F(x)$,也即 $P\left(\lim_{n\to\infty}F_n(x)=F(x)\right)=1$

又因为分布函数右连续，故$P\left(\lim_{n\to\infty}F_n(x+0)=F(x+0)\right)=1$

9、在正态总体$N(50,6^2)$中抽取容量为36的简单样本，求样本均值$\overline{X}$落在50.6和51.8之间的概率.

解：

$\overline{X}\sim N(50,1),P\left(50.6<\overline{X}<51.8\right)=P\left(0.6<\overline{X}-50<1.8\right)=\Phi(1.8)-\Phi(0.6)=0.2384.$

10、设$X_1,\cdots,X_{100}$是取自总体$X\sim N(\mu,1)$的样本，试确定常数c,使得对任意的$\mu>0$,

$P\left(\left|\overline{X}\right|<c\right)$都不超过0.05.

解：

由题意可知$\overline{X}\sim N\left(\mu,\frac{1}{100}\right)$，则$P\left(\left|\overline{X}\right|<c\right)=P\left(\frac{-c-\mu}{1/10}<\frac{\overline{X}-\mu}{1/10}<\frac{c-\mu}{1/10}\right)$

$=\Phi(10(c-\mu))-\Phi(-10(c+\mu))\leqslant 0.05.$

又因为$\mu>0$,故$\Phi(10(c-\mu))-\Phi(-10(c+\mu))<\Phi(10c)-\Phi(-10c)$

$=2\Phi(10c)-1\leqslant 0.05$,得$c\leqslant 0.006.$

11、设$X_1,\cdots,X_n$是取自总体$X\sim N(\mu,\sigma^2)$的简单样本，假如要以99.7%的概率保证偏差$\left|\overline{X}-\mu\right|<0.1$,试问在$\sigma^2=0.5$时，样本容量$n$应取多大?

解：

$P\left(\left|\overline{X}-\mu\right|<0.1\right)=P\left(\frac{\left|\overline{X}-\mu\right|}{\sigma/\sqrt{n}}<\frac{0.1\sqrt{n}}{\sigma}\right)=2\Phi\left(\frac{0.1\sqrt{n}}{\sigma}\right)-1=0.997,$

即$\Phi\left(\frac{0.1\sqrt{n}}{\sigma}\right)=0.9985$,解得$n=441$

12、利用切比雪夫不等式求一枚均匀硬币需抛掷多少次才能使样本均值$\overline{X}$落在 0.4 和 0.6 之间的概率至少为 0.9（此处$X_i=1$表示抛掷硬币出现正面，否则$X_i=0,i=1,\cdots,n$）？若用中心极限定理计算这个问题，需抛掷的次数又是多少？

解：

切比雪夫不等式：$\forall \varepsilon > 0, P(|X - E(X)| \geqslant \varepsilon) \leqslant \dfrac{D(X)}{\varepsilon^2}$

$X_i \sim b\left(1, \dfrac{1}{2}\right)$,则$\sum\limits_{i=1}^{n} X_i \sim b\left(n, \dfrac{1}{2}\right), n\overline{X} \sim b\left(n, \dfrac{1}{2}\right), E\left(n\overline{X}\right) = 0.5n, D\left(n\overline{X}\right) = 0.25n,$

则$P\left(0.4 < \overline{X} < 0.6\right) = P\left(0.4n < n\overline{X} < 0.6n\right)$

$= P\left(\left|n\overline{X} - 0.5n\right| < 0.1n\right) > 1 - \dfrac{0.25n}{(0.1n)^2} = 1 - \dfrac{25}{n} \geqslant 0.9, n \geqslant 250.$

中心极限定理：当$n \to \infty$时，$\dfrac{n\overline{X} - 0.5n}{\sqrt{0.25n}} \xrightarrow{\mathscr{L}} N(0,1)$,即$\dfrac{\overline{X} - 0.5}{\sqrt{0.25/n}} \xrightarrow{\mathscr{L}} N(0,1).$

则$P\left(0.4 < \overline{X} < 0.6\right) = P\left(-\dfrac{0.1}{\sqrt{0.25/n}} < \dfrac{\overline{X} - 0.5}{\sqrt{0.25/n}} < \dfrac{0.1}{\sqrt{0.25/n}}\right) = 2\Phi\left(\dfrac{0.1}{\sqrt{0.25/n}}\right) - 1 \geqslant 0.9$

得$n \geqslant 68.$

习题 2 抽样分布及若干预备知识

1、设从正态总体 (20,9) 中分别抽取容量为 10 和 15 的两组独立样本，记这两组样本的样本均值和样本方差分别为 $\overline{X}$, $\overline{Y}$ 和 S_X^2, S_Y^2.

（1）求两样本均值差的绝对值大于 0.3 的概率.（2）求 $9S_X^2+14S_Y^2$ 大于 164 的概率.

解：（1）由推论 2.2.2 可得 $\overline{X}\sim N(20,0.9)$，$\overline{Y}\sim N(20,0.6)$

又因为两组样本独立，所以 $\overline{X}-\overline{Y}\sim N(0,1.5)$，故

$$
\begin{aligned}
&P\left(\left|\overline{X}-\overline{Y}\right|>0.3\right)\\
&=P\left(\left|\frac{\overline{X}-\overline{Y}}{\sqrt{1.5}}\right|>\frac{0.3}{\sqrt{1.5}}\right)\\
&=2\left(1-\Phi\left(\frac{\sqrt{6}}{10}\right)\right)=0.806
\end{aligned}
$$

即两样本均值差的绝对值大于 0.3 的概率为 0.806

（2）由定理 2.3.3 可知 $9S_X^2/9\sim\chi_9^2$，$14S_Y^2/9\sim\chi_{14}^2$，故

$9S_X^2/9+14S_Y^2/9\sim\chi_{23}^2$, 故

$$
\begin{aligned}
&P(9S_X^2+14S_Y^2>164)\\
&=P\left(\frac{9S_X^2+14S_Y^2}{9}>\frac{164}{9}\right)\\
&=P\left(\chi_{23}^2>\frac{164}{9}\right)\\
&=0.75
\end{aligned}
$$

即 $9S_X^2+14S_Y^2$ 大于 164 的概率为 0.75.

2、设 $X_1,\ldots,X_n$ 和 $Y_1,\ldots,Y_n$ 是分别从正态总体 $N(\mu,\sigma^2)$ 中抽取的两组简单样本，且二者相互独立. 令 $\overline{X}$ 和 $\overline{Y}$ 分别为这两组样本的样本均值，试确定样本大小 n 的近似值，使得 $P\left(\left|\overline{X}-\overline{Y}\right|>\sigma\right)\approx 0.01$

解：由已知得 $\overline{X}\sim N(\mu,\sigma^2/n)$，$\overline{Y}\sim N(\mu,\sigma^2/n)$

故 $\overline{X}-\overline{Y}\sim N(0,2\sigma^2/n)$，则 $P\left(\left|\overline{X}-\overline{Y}\right|>\sigma\right)=P\left(\frac{\left|\overline{X}-\overline{Y}\right|}{\sqrt{2/n}\,\sigma}>\frac{\sigma}{\sqrt{2/n}\,\sigma}\right)$

也即 $P\left(\left|\overline{X}-\overline{Y}\right|>\sigma\right)=2\left(1-\Phi\left(\sqrt{n/2}\right)\right)$

即$2\left(1-\Phi\left(\sqrt{n/2}\right)\right)=0.01$，$\Phi\left(\sqrt{n/2}\right)=0.995$，查阅表得 n=13.2098,取其近似值 13.

3、设$X_1,\ldots,X_n$为从下列总体中抽取的简单样本，利用特征函数试求样本均值$\overline{X}$的分布：

（1）正态总体$N(a,\sigma^2)$；

（2）参数为λ的 Poisson 总体$P(\lambda)$；

（3）参数为λ的指数分布.

解：（1）设$Y\sim N(0,1)$,则$\varphi_Y(t)=e^{-\frac{t^2}{2}}$

又因为$X=\sigma Y+a$，故$\varphi_X(t)=E(X)=E(e^{it\sigma Y+ita})=e^{ita-\frac{\sigma^2}{2}t^2}$.

$$\varphi_{\overline{X}}(t)=\varphi_{\sum\limits_{i=1}^{n}\frac{X_i}{n}}(t)=\prod_{i=1}^{n}\varphi_{\frac{X_i}{n}}(t)=\prod_{i=1}^{n}\varphi_{Xi}\left(\frac{t}{n}\right)=\prod_{i=1}^{n}e^{\frac{ita}{n}-\frac{\sigma^2}{2}\left(\frac{t}{n}\right)^2}=e^{ita-\frac{\sigma^2}{2n}t^2}$$

这正是正态分布$N\left(a,\dfrac{\sigma^2}{n}\right)$的特征函数，再由特征函数的唯一性定理可知

$\overline{X}\sim N\left(a,\dfrac{\sigma^2}{n}\right)$.

（2）$P(\lambda)$的特征函数$\varphi_X(t)=\sum\limits_{k=0}^{\infty}e^{ikt}\dfrac{\lambda^k}{k!}=\sum\limits_{k=0}^{\infty}\dfrac{(\lambda e^{it})^k}{k}e^{-\lambda}=e^{\lambda(e^{it}-1)}$，

$$\varphi_{\overline{X}}(t)=\prod_{i=1}^{n}\varphi_{\frac{X_i}{n}}(t)=\prod_{i=1}^{n}\varphi_{Xi}\left(\frac{t}{n}\right)=\prod_{i=1}^{n}e^{\lambda\left(e^{\frac{it}{n}}-1\right)}=e^{n\lambda\left(e^{\frac{it}{n}}-1\right)}$$

故$\varphi_{n\overline{X}}(t)=e^{n\lambda(e^{it}-1)}$，又因为$P(n\lambda)$的特征函数也为$\varphi_{n\overline{X}}(t)=e^{n\lambda(e^{it}-1)}$，由特征函数的唯一性定理可知，$n\overline{X}\sim P(n\lambda)$，

故$P\left(n\overline{X}=k\right)=\dfrac{(n\lambda)^k}{k!}e^{-n\lambda}\Longrightarrow P\left(\overline{X}=\dfrac{k}{n}\right)=\dfrac{(n\lambda)^k}{k!}e^{-n\lambda}\quad k=0,1,2\ldots$

（3）$Exp(\lambda)$的特征函数$\varphi_X(t)=\left(1-\dfrac{it}{\lambda}\right)^{-1}$，故

$\varphi_{n\overline{X}}(t)=\prod\limits_{i=1}^{n}\varphi_{X_i}(t)=\left(1-\dfrac{it}{\lambda}\right)^{-n}$，即$n\overline{X}\sim Ga(n,\lambda)$即

$$f_{n\overline{X}}(x)=\frac{\lambda^n}{\Gamma(n)}x^{n-1}e^{-\lambda x}\Longrightarrow f_{\overline{X}}(x)=\frac{(n\lambda)^n}{\Gamma(n)}x^{n-1}e^{-n\lambda x}$$

4、设$X_1,\ldots,X_n$是从两点分布$b(1,p)$中抽取的简单样本，$0<p<1$，$\overline{X}$和S_n^2为样本均值和样

本方差，求$S_n^2=\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2/n$的分布.

解：$S_n^2=\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2/n=\frac{1}{n}\left(\sum_{i=1}^{n}X_i^2-n\overline{X}^2\right)=\frac{1}{n}\sum_{i=1}^{n}X_i^2-\overline{X}^2$

$X_i=\begin{cases}1 & 第i个成立\\ 0 & 第i个不成立\end{cases}$ 故$X_i{}^2=\begin{cases}1 & 第i个成立\\ 0 & 第i个不成立\end{cases}$

故$\frac{1}{n}\sum_{i=1}^{n}X_i^2=\overline{X}$，即$S_n^2=\overline{X}-\overline{X}^2$

又因为$P\left(\overline{X}=\frac{k}{n}\right)=C_n^k p^k(1-p)^{n-k}$,

故$P\left(S_n^2=\frac{k}{n}-\frac{k^2}{n^2}\right)=P\left(\overline{X}=\frac{k}{n},\frac{n-k}{n}\right)$

$$=\begin{cases}C_n^k p^k(1-p)^{n-k}+C_n^{n-k}p^{n-k}(1-p)^k, k=0,\cdots,\left[\frac{n}{2}\right] & n为奇数\\ \begin{cases}C_n^k p^k(1-p)^{n-k}+C_n^{n-k}p^{n-k}(1-p)^k，& k=0,1\cdots,\frac{n}{2}-1\\ C_n^{n/2}p^{n/2}(1-p)^{n/2}, & k=\frac{n}{2}\end{cases} & ,n为偶数\end{cases}$$

5、设X_1,X_2为取自正态总体$X\sim N(0,\sigma^2)$的一个样本，证明统计量X_1/X_2和$\sqrt{X_1^2+X_2^2}$是相互独立的

解：

令$U=\frac{X_1}{X_2},V=\sqrt{X_1^2+X_2^2}$,即$u=\frac{x_1}{x_2},v=\sqrt{x_1^2+x_2^2}$

也即①$x_1=\frac{uv}{\sqrt{1+u^2}},x_2=\frac{v}{\sqrt{1+u^2}}\ (x_2>0)$

②$x_1=\frac{-uv}{\sqrt{1+u^2}},x_2=\frac{-v}{\sqrt{1+u^2}}\ (x_2<0)$

由此可得雅可比行列式

$$J'=\begin{vmatrix}\frac{\partial u}{\partial x_1} & \frac{\partial u}{\partial x_2}\\ \frac{\partial v}{\partial x_1} & \frac{\partial v}{\partial x_2}\end{vmatrix}=\begin{vmatrix}\frac{1}{x_2} & -\frac{x_1}{x_2^2}\\ \frac{x_1}{\sqrt{x_1^2+x_2^2}} & \frac{x_2}{\sqrt{x_1^2+x_2^2}}\end{vmatrix}=\left(\frac{1}{\sqrt{x_1^2+x_2^2}}\right)\left(1+\left(\frac{x_1}{x_2}\right)^2\right)=\frac{1+u^2}{v}$$

故$J=\begin{vmatrix}\frac{\partial x_1}{\partial u} & \frac{\partial x_1}{\partial v}\\ \frac{\partial x_2}{\partial u} & \frac{\partial x_2}{\partial v}\end{vmatrix}=\frac{1}{J'}=\frac{v}{1+u^2}$

因此在情况①下

$$f_{U,V}(u,v)=f_{X_1,X_2}\left(\frac{uv}{\sqrt{1+u^2}},\ \frac{v}{\sqrt{1+u^2}}\right)\cdot|J|=\frac{1}{2\pi\sigma^2}e^{-\frac{v^2}{2\sigma^2}}\cdot\frac{v}{1+u^2}\,(v>0)$$

同理可得情况②下

$$f_{U,V}(u,v)=f_{X_1,X_2}\left(\frac{-uv}{\sqrt{1+u^2}},\ \frac{-v}{\sqrt{1+u^2}}\right)\cdot|J|=\frac{1}{2\pi\sigma^2}e^{-\frac{v^2}{2\sigma^2}}\cdot\frac{v}{1+u^2}\,(v>0)$$

综上所述U,V的联合概率密度为上述两个概率密度之和，也即

$$f_{U,V}(u,v)=\frac{1}{\pi\sigma^2}e^{-\frac{v^2}{2\sigma^2}}\cdot\frac{v}{1+u^2}\,(u\in R,v>0)$$

$f_{U,V}(u,v)$可分离变量，故$\dfrac{X_1}{X_2}$与$\sqrt{X_1^2+X_2^2}$独立.

注：也可利用$Basu$定理，$\sqrt{X_1^2+X_2^2}$是σ^2的充分完全统计量，$\dfrac{X_1}{X_2}$的分布与σ^2无关，故$\sqrt{X_1^2+X_2^2}$和$\dfrac{X_1}{X_2}$独立.

6、设$X_1,\ldots,X_n\ i.i.d.\sim F,(X_{(1),}X_{(2)},\cdots X_{(n)})$是其次序统计量，已知$P(X_{(m)}<x)=\sum_{i=m}^{n}\binom{n}{i}(F(x))^i(1-F(x))^{n-i}$，证明下列恒等式：

$$\sum_{i=m}^{n}\binom{n}{i}(F(x))^i(1-F(x))^{n-i}=m\binom{n}{m}\int_0^{F(x)}t^{m-1}(1-t)^{n-m}dt.$$

解：

构造关于p的函数$g(p)$,

$$g(p)=\sum_{i=m}^{n}\binom{n}{i}p^i(1-p)^{n-i}-m\binom{n}{m}\int_0^{p}t^{m-1}(1-t)^{n-m}dt$$

$$g'(p)=\sum_{i=m}^{n}\binom{n}{i}i\,p^{i-1}(1-p)^{n-i}-\sum_{i=m}^{n}\binom{n}{i}(n-i)p^i(1-p)^{n-i-1}-m\binom{n}{m}p^{m-1}(1-p)^{n-m}$$

$$=\sum_{i=m}^{n}\binom{n-1}{i-1}np^{i-1}(1-p)^{n-i}-\sum_{i=m}^{n-1}\binom{n-1}{i}n\,p^i(1-p)^{n-i-1}-m\binom{n}{m}p^{m-1}(1-p)^{n-m}$$

$$=\sum_{i=m-1}^{n-1}\binom{n-1}{i}np^i(1-p)^{n-i-1}-\sum_{i=m}^{n-1}\binom{n-1}{i}n\,p^i(1-p)^{n-i-1}-m\binom{n}{m}p^{m-1}(1-p)^{n-m}$$

$$=\binom{n-1}{m-1}np^{m-1}(1-p)^{n-m}-m\binom{n}{m}p^{m-1}(1-p)^{n-m}$$

$$=\binom{n}{m}mp^{m-1}(1-p)^{n-m}-m\binom{n}{m}p^{m-1}(1-p)^{n-m}=0$$

注：$\binom{n}{i}i=\binom{n-1}{i-1}n,\binom{n}{i}(n-i)=\binom{n-1}{i}n,$

故$g(p)$为一个常数，又因为$g(0)=0$,所以$g(p)=0$,因此

$$\sum_{i=m}^{n}\binom{n}{i}(F(x))^{i}(1-F(x))^{n-i}=m\binom{n}{m}\int_{0}^{F(x)}t^{m-1}(1-t)^{n-m}dt.$$

7、设$r.v.\ X$的分布函数为$F(x)$,密度函数为$f(x)$,$X_1,\cdots X_n\ i.i.d.\sim f$,证明

$$\underset{a<x_1<\cdots<x_n<b}{\int\cdots\int} f(x_1)\cdots f(x_n)dx_1\cdots dx_n=\frac{1}{n!}[F(b)-F(a)]^n.$$

次序统计量$(X_{(1)},\cdots X_{(n)})$的联合分布由式（2.3.4）给出，$X_{(m)}$的分布可视为$(X_{(1)},\cdots X_{(n)})$

联合分布的边缘分布，利用这一事实和上述恒等式求出$X_{(m)}$的密度函数。

解：

可以注意到$\int_a^b\cdots\int_a^b f(x_1)\cdots f(x_n)dx_1\cdots dx_n=[F(b)-F(a)]^n$,

对于$1,\cdots,n$的任一排列，$i_1,\cdots i_n$,都有

$$\underset{a<x_{i_1}<\cdots x_{i_n}<b}{\int\cdots\int} f(x_1)\cdots f(x_n)dx_1\cdots dx_n=\underset{a<x_1<\cdots<x_n<b}{\int\cdots\int} f(x_1)\cdots f(x_n)dx_1\cdots dx_n$$

故$\int_a^b\cdots\int_a^b f(x_1)\cdots f(x_n)dx_1\cdots dx_n=\sum_{\text{所有}i_1,\cdots i_n\text{排列}}\underset{a<x_{i_1}<\cdots x_{i_n}<b}{\int\cdots\int} f(x_1)\cdots f(x_n)dx_1\cdots dx_n$

$$=n!\underset{a<x_1<\cdots<x_n<b}{\int\cdots\int} f(x_1)\cdots f(x_n)dx_1\cdots dx_n$$

故 $\underset{a<x_1<\cdots<x_n<b}{\int\cdots\int} f(x_1)\cdots f(x_n)dx_1\cdots dx_n=\frac{1}{n!}[F(b)-F(a)]^n$

下面求$f_m(x)$的密度函数

由式(2.3.4)可知次序统计量$(X_{(1)},\cdots X_{(n)})$的联合密度

$g(x_{(1)},\cdots x_{(n)})=n!f(x_{(1)})\cdots f(x_{(n)})$

故$f_m(x)=\underset{-\infty<x_{(1)}<\cdots<x_{(n)}<+\infty}{\int\cdots\int} n!f(x_{(1)})\cdots f(x_{(m-1)})f(x)f(x_{(m+1)})\cdots f(x_{(n)})dx_{(1)}\cdots dx_{(m-1)}dx_{(m+1)}\cdots dx_{(n)}$

$$=n!f(x)\underset{-\infty<x_{(1)}<\cdots<x_{(m-1)}<x}{\int\cdots\int} f(x_{(1)})\cdots f(x_{(m-1)})dx_{(1)}\cdots dx_{(m-1)}\underset{x<x_{(m+1)}<\cdots<x_{(n)}<+\infty}{\int\cdots\int} f(x_{(m+1)})\cdots f(x_{(n)})dx_{(m+1)}\cdots dx_{(n)}$$

$$=n!f(x)\cdot\frac{1}{(m-1)!}\cdot[F(x)-F(-\infty)]^{m-1}\cdot\frac{1}{(n-m)!}\cdot[F(+\infty)-F(x)]^{n-m}$$

$$=m\cdot\frac{n!}{m!\cdot(n-m)!}[F(x)]^{m-1}[1-F(x)]^{n-m}f(x)$$

$$=m\binom{n}{m}[F(x)]^{m-1}[1-F(x)]^{n-m}f(x)$$

8、设$X_{(1)},\cdots,X_{(n)}$为从均匀分布$U(0,1)$中抽取的次序统计量，

（1）样本容量 n 为多大时，才能使$P(X_{(n)}\geqslant 0.99)\geqslant 0.95$?

（2）求极差$R_n=X_{(n)}-X_{(1)}$的密度函数；

（3）证明统计量$Z=2n(1-R_n)$的极限分布为χ_4^2.

解：$X_{(1)},\cdots,X_{(n)}\sim U(0,1)$

$$F(x)=\begin{cases}0 & x\in(-\infty,0]\\ x & x\in(0,1)\\ 1 & x\in[1,+\infty)\end{cases}\qquad f(x)=\begin{cases}1 & x\in(0,1)\\ 0 & 其他\end{cases}$$

（1）

$F_n(x)=P(X_{(n)}\leqslant x)=x^n$

$P(X_{(n)}\geqslant 0.99)=1-P(X_{(n)}<0.99)=1-0.99^n\geqslant 0.95$

解得$n\geqslant 299$，故n至少为299时才能使$P(X_{(n)}\geqslant 0.99)\geqslant 0.95$

（2）

$$f_{1,n}(x,y)=\begin{cases}n(n-1)[F(y)-F(x)]^{n-2}f(x)f(y) & x<y\\ 0 & x>y\end{cases}$$

$$f_{1,n}(x,y)=\begin{cases}n(n-1)(y-x)^{n-2} & x<y\\ 0 & x>y\end{cases}$$

令$\begin{cases}r=y-x\\ z=x\end{cases}$ 则$\begin{cases}x=z\\ y=z+r\end{cases}$ $J=\dfrac{\partial(x,y)}{\partial(r,z)}=1$

则$f_{1,n}(r,z)=n(n-1)r^{n-2}\quad(0<r<1,0<z<1-r)$

$$f_{R_n}(r)=\begin{cases}\displaystyle\int_0^{1-r}n(n-1)r^{n-2}dz=n(n-1)r^{n-2}(1-r), & 0<r<1\\ 0, & 其他\end{cases}$$

（3）$R_n=1-\dfrac{Z_n}{2n}$

$$f_{Z_n}(z)=f_{R_n}\left(1-\frac{z}{2n}\right)\left|\frac{d\left(1-\frac{z}{2n}\right)}{dz}\right|=\frac{1}{2n}f_{R_n}\left(1-\frac{z}{2n}\right)$$

$$=\frac{1}{2n}n(n-1)\left(1-\frac{z}{2n}\right)^{n-2}\frac{z}{2n}=\frac{z}{4n}(n-1)\left(1-\frac{z}{2n}\right)^{n-2}\quad(0<z<2n)$$

$$\lim_{n\to\infty}f_{Z_n}(z)=\lim_{n\to\infty}\frac{z}{4n}(n-1)\left(1-\frac{z}{2n}\right)^{n-2}=\frac{z}{4}e^{-\frac{z}{2}}=\frac{\left(\frac{1}{2}\right)^2}{\Gamma(2)}z^{2-1}e^{-\frac{z}{2}}\sim Ga\left(2,\frac{1}{2}\right)\sim\chi_4^2$$

9、设总体$X\sim N(0,1)$，令为其分布函数.从这总体中获取一组样本观察值：$X_1=0$，

$X_2=0.2, X_3=0.25, X_4=-0.3, X_5=-0.1, X_6=2, X_7=0.15, X_8=1, X_9=-0.7$, $X_{10}=-1$.

（1）求上述样本的经验分布函数；

（2）计算$E\{F(X_{(6)})\}, D\{F(X_{(6)})\}, X_{(6)}$为容量是 10 的次序统计量.

（3）计算容量 n=10 的样本中次序统计量$X_{(6)}$的分布函数在 0.2 处的值.

解：

（1）将样本值有序排列：-1，-0.7，-0.3，-0.1, 0, 0.15, 0.2, 0.25, 1, 2

$$F_{(n)}(x)=\begin{cases}0 & x<-1\\ \frac{1}{10} & -1\leqslant x<-0.7\\ \frac{2}{10} & -0.7\leqslant x<-0.3\\ \frac{3}{10} & -0.3\leqslant x<-0.1\\ \frac{4}{10} & -0.1\leqslant x<0\\ \frac{5}{10} & 0\leqslant x<0.15\\ \frac{6}{10} & 0.15\leqslant x<0.2\\ \frac{7}{10} & 0.2\leqslant x<0.25\\ \frac{8}{10} & 0.25\leqslant x<1\\ \frac{9}{10} & 1\leqslant x<2\\ 1 & 2\leqslant x\end{cases}$$

（2）$X_{(6)}$的密度函数为$f_6(x)=6\binom{10}{6}(F(x))^{6-1}(1-F(x))^4 f(x)$

$$E\{F(X_{(6)})\}=\int_{-\infty}^{+\infty}6\binom{10}{6}(F(x))^6(1-F(x))^4 f(x)dx=\int_{-\infty}^{+\infty}6\binom{10}{6}(F(x))^6(1-F(x))^4 dF(x)$$

$$\xlongequal{F(x)=t}\int_0^1 6\binom{10}{6}t^{7-1}(1-t)^{5-1}dt=6\binom{10}{6}\frac{\Gamma(7)\Gamma(5)}{\Gamma(12)}=\frac{6}{11}.$$

$$E\{F^2(X_{(6)})\}=\int_{-\infty}^{+\infty}6\binom{10}{6}(F(x))^7(1-F(x))^4 f(x)dx=\int_{-\infty}^{+\infty}6\binom{10}{6}(F(x))^7(1-F(x))^4 dF(x)$$

$$\xlongequal{F(x)=t}\int_0^1 6\binom{10}{6}t^{8-1}(1-t)^{5-1}dt=6\binom{10}{6}\frac{\Gamma(8)\Gamma(5)}{\Gamma(13)}=\frac{7}{22}.$$

$$D\{F(X_{(6)})\}=E\{F^2(X_{(6)})\}-E^2\{F(X_{(6)})\}=\frac{5}{242}.$$

（3）$F(X_{(6)})=P(X_{(6)}<x)=\sum_{i=6}^{10}\binom{10}{i}(F(x))^{i}(1-F(x))^{10-i}$

$x=0.2$时，$F(x)=0.5793$,故$F(X_{(6)})=\sum_{i=6}^{10}\binom{10}{i}(0.5793)^{i}(0.4207)^{10-i}=0.5804$

10、设总体 X 服从威布尔分布，其分布函数为

$$F(x)=\begin{cases}1-e^{-(x/\beta)^{\alpha}}, & x\geqslant 0,\\ 0, & x<0,\end{cases}$$

其中α为形状参数，β为刻度参数.（$X_{(1)},\cdots X_{(n)}$）为从此总体中抽取的简单样本，试证

$Y=\min\{X_{(1)},\cdots,X_{(n)}\}$仍服从威布尔分布并指出$Y$的分布的形状参数和刻度参数是什么？

解：

$F_Y(y)=P(Y\leqslant y)=1-[1-F_X(y)]^{n}=1-e^{-n(x/\ \beta)^{\alpha}}$

$=1-e^{-\left(x/n^{-\frac{1}{\alpha}}\beta\right)^{\alpha}}$，$y\geqslant 0$

故形状参数为α，刻度参数为$n^{-\frac{1}{\alpha}}\beta$.

11、设$X_1,\cdots,X_n$ $i.i.d.\ \sim N(a,\sigma^2)$，且$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$,$S_n^2=\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2$,又设

$X_{n+1}\sim N(a,\sigma^2)$且与$X_1,\cdots,X_n$独立，试求统计量$\frac{X_{n+1}-\overline{X}}{S_n}\sqrt{\frac{n-1}{n+1}}$的分布.

解：

X_{n+1}与$\overline{X}$独立，故$X_{n+1}-\overline{X}\sim N\left(0,\frac{n+1}{n}\sigma^2\right)$

故$\frac{X_{n+1}-\overline{X}}{\sqrt{\frac{n+1}{n}}\sigma}\sim N(0,1)$,又因为$\frac{nS_n^2}{\sigma^2}\sim\chi_{n-1}^2$,故

$$\frac{\frac{X_{n+1}-\overline{X}}{\sqrt{\frac{n+1}{n}}\sigma}}{\sqrt{\frac{\frac{nS_n^2}{\sigma^2}}{n-1}}}=\frac{X_{n+1}-\overline{X}}{S_n}\sqrt{\frac{n-1}{n+1}}\sim t(n-1)$$

12、设$X_1,\cdots,X_n$ $i.i.d\ \sim N(\mu_1,\sigma^2),Y_1,\cdots,Y_n$ $i.i.d\ \sim N(\mu_2,\sigma^2)$，且$X_1,\cdots,X_n$和$Y_1,\cdots,Y_n$

相互独立，$\overline{X}$和$\overline{Y}$分别表示它们的样本均值，S_{1m}^2和S_{2n}^2定义类似上题中的S_n^2. α和β是两

个给定的实数，试求$T=\frac{\alpha\left(\overline{X}-\mu_1\right)+\beta\left(\overline{Y}-\mu_2\right)}{\sqrt{\frac{mS_{1m}^2+nS_{2n}^2}{n+m-2}\cdot\left(\frac{\alpha^2}{m}+\frac{\beta^2}{n}\right)}}$的分布.

解

$$\overline{X}-\mu_1\sim N\left(0,\ \frac{\sigma^2}{m}\right),\overline{Y}-\mu_2\sim N\left(0,\frac{\sigma^2}{n}\right)$$

故$\alpha\left(\overline{X}-\mu_1\right)+\beta\left(\overline{Y}-\mu_2\right)\sim N\left(0,\ \left(\frac{\alpha^2}{m}+\frac{\beta^2}{n}\right)\sigma^2\right)$

因此$\dfrac{\alpha\left(\overline{X}-\mu_1\right)+\beta\left(\overline{Y}-\mu_2\right)}{\sqrt{\frac{\alpha^2}{m}+\frac{\beta^2}{n}}\cdot\sigma}\sim N(0,1)$

又因为$\dfrac{mS_{1m}^2}{\sigma^2}\sim\chi_{m-1}^2,\dfrac{nS_{2n}^2}{\sigma^2}\sim\chi_{n-1}^2$

故$\dfrac{mS_{1m}^2+nS_{2n}^2}{\sigma^2}\sim\chi_{n+m-2}^2$

因此$T=\dfrac{\alpha\left(\overline{X}-\mu_1\right)+\beta\left(\overline{Y}-\mu_2\right)}{\sqrt{\frac{mS_{1m}^2+nS_{2n}^2}{n+m-2}\cdot\left(\frac{\alpha^2}{m}+\frac{\beta^2}{n}\right)}}\sim t(n+m-2)$

13、设$X_1,\cdots,X_n$ $i.i.d.$ $\sim N(a,\sigma^2),\overline{X}=\dfrac{1}{n}\sum_{i=1}^{n}X_i,S=\dfrac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2$,记$\xi=\left(X_1-\overline{X}\right)/S$. 试找出$\xi$与$t$分布的联系，因而定出$\xi$的密度.

解：

作正交变换$Y=AX$

$$A=\begin{pmatrix} \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \cdots & \frac{1}{\sqrt{n}} \\ \frac{n-1}{\sqrt{n(n-1)}} & -\frac{1}{\sqrt{n(n-1)}} & \cdots & -\frac{1}{\sqrt{n(n-1)}} \\ c_{31} & c_{32} & \cdots & c_{3n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix}$$

故有$Y_1=\dfrac{1}{\sqrt{n}}\sum_{i=1}^{n}X_i=\sqrt{n}\,\overline{X}$

$$Y_2=\sqrt{\frac{n}{n-1}}\left(X_1-\overline{X}\right)$$

$$\sum_{i=1}^{n}y_i^2=Y^TY=X^TA^TAX=X^TX=\sum_{i=1}^{n}X_i^2$$

$$f(x_1,\cdots,x_n)=(2\pi\sigma^2)^{-\frac{n}{2}}e^{-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}}=(2\pi\sigma^2)^{-\frac{n}{2}}e^{-\frac{\sum_{i=1}^{n}(x_i^2-2\mu x_i+\mu^2)}{2\sigma^2}}$$

$$=(2\pi\sigma^2)^{-\frac{n}{2}}e^{-\frac{\sum_{i=1}^{n}y_i^2-2\sqrt{n}y_1\mu+n\mu^2}{2\sigma^2}}=(2\pi\sigma^2)^{-\frac{n}{2}}e^{-\frac{\sum_{i=2}^{n}y_i^2+\left(y_1-\sqrt{n}\mu\right)^2}{2\sigma^2}}$$

因此$y_2,\cdots,y_n$独立同分布于$N(0,\sigma^2)$,且$(n-1)\cdot s^2=\sum\limits_{i=2}^{n}y_i^2$.

令 $t=\dfrac{y_2}{\sqrt{y_3^2+\cdots+y_n^2/(n-2)}}=\dfrac{y_2}{\sqrt{[(n-1)s^2-y_2^2]/(n-2)}}$,

$t\sim t(n-2)$,而

$$\xi=\left(x_i-\bar{x}\right)/s=\frac{n-1}{\sqrt{n}}\frac{y_2}{\sqrt{(n-1)s^2}}$$

$$=\frac{n-1}{\sqrt{n}}\frac{\dfrac{y_2}{\sqrt{[(n-1)s^2-y_2^2]/(n-2)}}}{\sqrt{\dfrac{(n-1)s^2}{[(n-1)s^2-y_2^2]/(n-2)}}}=\frac{n-1}{\sqrt{n}}\frac{t}{\sqrt{n-2+t^2}}$$

此式即为ξ与t分布的联系，因而可以据此定出ξ的密度.

14、设$X_1,\cdots,X_n$独立，$X_i\sim N(0,\sigma_i^2),i=1,\cdots,n,$定义

$$\xi=\sum_{i=1}^{n}\frac{(X_i-Z)^2}{\sigma_i^2},\qquad \text{其中}Z=\sum_{i=1}^{n}\frac{X_i}{\sigma_i^2}\Big/\sum_{i=1}^{n}\frac{1}{\sigma_i^2},$$

求ξ的分布（提示：作适当的正交变换）.

解:

$$\xi=\sum_{i=1}^{n}\frac{X_i^2-2ZX_i+Z^2}{\sigma_i^2}=\sum_{i=1}^{n}\frac{X_i^2}{\sigma_i^2}-2Z\sum_{i=1}^{n}\frac{X_i}{\sigma_i^2}+Z^2\sum_{i=1}^{n}\frac{1}{\sigma_i^2}$$

$$=\sum_{i=1}^{n}\frac{X_i^2}{\sigma_i^2}-Z\sum_{i=1}^{n}\frac{X_i}{\sigma_i^2}=\sum_{i=1}^{n}\frac{X_i^2}{\sigma_i^2}-\left(\sum_{i=1}^{n}\frac{X_i}{\sigma_i^2}\right)^2\Big/\sum_{i=1}^{n}\frac{1}{\sigma_i^2}$$

$$=\sum_{i=1}^{n}\frac{X_i^2}{\sigma_i^2}-\left(\sum_{i=1}^{n}\frac{X_i}{\sigma_i}\cdot\frac{1}{\sigma_i\sqrt{\sum\limits_{i=1}^{n}\frac{1}{\sigma_i^2}}}\right)^2.$$

作下列正交变换

$$\begin{pmatrix}Y_1\\Y_2\\\vdots\\Y_n\end{pmatrix}=\begin{pmatrix}\dfrac{1}{\sigma_1\sqrt{\sum\limits_{i=1}^{n}\frac{1}{\sigma_i^2}}} & \cdots & \dfrac{1}{\sigma_n\sqrt{\sum\limits_{i=1}^{n}\frac{1}{\sigma_i^2}}}\\ a_{21} & \cdots & a_{2n}\\ \vdots & \vdots & \vdots\\ a_{n1} & \cdots & a_{nn}\end{pmatrix}\begin{pmatrix}\dfrac{X_1}{\sigma_1}\\\dfrac{X_2}{\sigma_2}\\\vdots\\\dfrac{X_n}{\sigma_n}\end{pmatrix}$$

则有

$$Y_1^2+Y_2^2+\cdots+Y_n^2=\frac{X_1^2}{\sigma_1}+\frac{X_2^2}{\sigma_2}+\cdots+\frac{X_n^2}{\sigma_n}$$

$$Y_1=\sum_{i=1}^{n}\frac{X_i}{\sigma_i}\cdot\frac{1}{\sigma_i\sqrt{\sum_{i=1}^{n}\frac{1}{\sigma_i^2}}}$$

故

$$\xi=\sum_{i=2}^{n}Y_i^2,$$

又因为$\frac{X_i}{\sigma_i}\sim N(0,1)$,故$Y_i\sim N(0,1)$,因此$\xi\sim\chi_{n-1}^2$.

15、若X服从自由度为 n 的χ^2分布，试证

（1）X的特征函数为$\varphi(t)=(1-2it)^{-\frac{n}{2}}$;

（2）$EX=n,D(X)=2n$;

（3）如果$X_1,\cdots,X_k$相互独立且$X_i\sim\chi_{n_i}^2$,则$\sum_{i=1}^{k}X_i\sim\chi_n^2$,此处$n=\sum_{i=1}^{n}n_i$.

解：

（1）$f_X(x)=\frac{1}{2^{n/2}\Gamma(n/2)}x^{n/2-1}e^{-x/2},\ x>0$

$$\varphi(t)=E(e^{itX})=\int_0^{+\infty}e^{itx}\frac{1}{2^{n/2}\Gamma(n/2)}x^{n/2-1}e^{-x/2}dx=\frac{1}{2^{n/2}(1/2-it)^{n/2}}\int_0^{+\infty}\frac{(1/2-it)^{n/2}}{\Gamma(n/2)}x^{n/2-1}e^{-(1/2-it)x}dx$$
$$=(1-2it)^{-n/2}$$

（2）$\varphi^{(k)}(0)=i^kE(X^k)$

$\varphi'(t)=-\frac{n}{2}(1-2it)^{-\frac{n}{2}-1}(-2i)\quad \varphi'(0)=ni=iEX$

$EX=n$

$\varphi''(t)=-\frac{n}{2}\left(-\frac{n}{2}-1\right)(1-2it)^{-\frac{n}{2}-2}(-2i)^2$

$\varphi''(0)=-n(n+2)=i^2E(X^2)$　故$E(X^2)=n^2+2n$

因此$DX=E(X^2)-(EX)^2=2n$

（3）$\varphi_{X_i}(t)=(1-2it)^{-\frac{n_i}{2}}$

则$\varphi_{\sum_{i=1}^{n}X_i}(t)=\prod_{i=1}^{n}(1-2it)^{-\frac{n_i}{2}}=(1-2it)^{-\frac{\sum_{i=1}^{n}n_i}{2}}=(1-2it)^{-\frac{n}{2}}$

由特征函数的唯一性定理可知$\sum_{i=1}^{k}X_i\sim\chi_n^2$

16、设$X_1,\cdots,X_n$是从总体$X\sim\chi_m^2$中抽取的大小为 n 的简单样本，试求样本均值$\overline{X}$的概率

分布.

解：

由 15（3）可知$\sum_{i=1}^{n} X_i \sim \chi_{nm}^2$

$$故 n\overline{X} \sim g_{nm}(x) = \begin{cases} \dfrac{1}{2^{\frac{nm}{2}}\Gamma\left(\frac{nm}{2}\right)} x^{\frac{nm}{2}-1} e^{-\frac{x}{2}}, & x>0 \\ 0, & 其它 \end{cases}$$

$$令 Y = \overline{X} = \frac{n\overline{X}}{n}$$

$$g_Y(y) = g_{nm}(ny)\left|\frac{d(ny)}{dy}\right| = \begin{cases} \dfrac{n}{2^{\frac{nm}{2}}\Gamma\left(\frac{nm}{2}\right)} (ny)^{\frac{nm}{2}-1} e^{-\frac{ny}{2}}, & y>0 \\ 0, & 其它 \end{cases}$$

17、计算自由度为 n 的χ^2分布 的变异系数和峰度.

解：

χ_n^2的特征函数$\varphi(t) = (1-2it)^{-\frac{n}{2}}$

由15(2)可知$\mu_1 = E(X) = n,\ v_2 = D(X) = 2n$

故变异系数$v = \dfrac{\sqrt{v_2}}{\mu_1} = \sqrt{\dfrac{2}{n}}$

$$\varphi''(t) = -\frac{n}{2}\left(-\frac{n}{2}-1\right)(1-2it)^{-\frac{n}{2}-2}(-2i)^2$$

$\varphi''(0) = -n(n+2) = i^2E(X^2)$ 故$\mu_2 = E(X^2) = n^2+2n$

$$\varphi'''(t) = -n(n+2)\left(-\frac{n+4}{2}\right)(-2i)(1-2it)^{-\frac{n}{2}-3}$$

$\varphi'''(0) = -in(n+2)(n+4) = i^3E(X^3)$ 故$\mu_3 = E(X^3) = n(n+2)(n+4)$

同理得$\varphi^{(4)}(0) = n(n+2)(n+4)(n+6) = i^4E(X^4)$ 故$\mu_4 = E(X^4) = n(n+2)(n+4)(n+6)$

$$v_4 = \sum_{i=0}^{4}\binom{4}{i}\mu_i(-\mu_1)^{4-i} = \mu_4 - 4\mu_3\mu_1 + 6\mu_2{\mu_1}^2 - 3{\mu_1}^4 = 12n^2+48n$$

故峰度$\beta_k = \dfrac{v_4}{{v_2}^2} - 3 = \dfrac{12}{n}$

18、设$X \sim \chi_{2n}^2$，Y服从参数为λ的$Poisson$分布$P(\lambda)$,则有$P(X<2\lambda) = P(Y \geq n)$.

解：

$$P(X<2\lambda) = \frac{1}{2^n\Gamma(n)}\int_0^{2\lambda} x^{n-1}e^{-\frac{x}{2}}dx,\ \ P(Y\geq n) = 1-\sum_{i=0}^{n-1}\frac{\lambda^i}{i!}e^{-\lambda}$$

$$令 \frac{x}{2} = t, 则 P(X<2\lambda) = \frac{1}{2^n\Gamma(n)}\int_0^{\lambda} 2^n t^{n-1}e^{-t}dt = \frac{1}{(n-1)!}\int_0^{\lambda} t^{n-1}e^{-t}dt$$

$$\int_0^{\lambda} t^{n-1}e^{-t}dt = -\int_0^{\lambda} t^{n-1}de^{-t} = -t^{n-1}e^{-t}\Big|_0^{\lambda} + \int_0^{\lambda}(n-1)t^{n-2}e^{-t}dt$$

$$= -\lambda^{n-1}e^{-\lambda} - (n-1)t^{n-2}e^{-t}\Big|_0^{\lambda} + \int_0^{\lambda}(n-1)(n-2)t^{n-3}e^{-t}dt$$

$$= -\lambda^{n-1}e^{-\lambda} - (n-1)\lambda^{n-2}e^{-\lambda} + \int_0^{\lambda}(n-1)(n-2)t^{n-3}e^{-t}dt$$

$$\cdots = -\lambda^{n-1}e^{-\lambda} - (n-1)\lambda^{n-2}e^{-\lambda} - (n-1)(n-2)\lambda^{n-3}e^{-\lambda}\cdots - (n-1)!\,e^{-\lambda}$$

故$P(X<2\lambda) = \dfrac{1}{(n-1)!}\displaystyle\int_0^{\lambda} t^{n-1}e^{-t}dt = 1 - \sum_{i=0}^{n-1}\frac{\lambda^i}{i!}e^{-\lambda} = P(Y \geqslant n)$，故得证

19、设$\xi \sim \chi_n^2, r$为常数，计算$E(\xi^r)$和$D(\xi^r)$,并指出r取哪些值时这些量才存在.又问ξ的密度最大值在哪一点达到？

解：

$$E(\xi^r) = \int_0^{+\infty} \xi^r \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}\xi^{\frac{n}{2}-1}e^{-\frac{\xi}{2}}d\xi$$

$$= \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}\int_0^{+\infty}\xi^{\frac{n+2r}{2}-1}e^{-\frac{\xi}{2}}d\xi$$

$$= \frac{2^{\frac{n}{2}+r}\Gamma\left(\frac{n}{2}+r\right)}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}\int_0^{+\infty}\frac{1}{2^{\frac{n}{2}+r}\Gamma\left(\frac{n}{2}+r\right)}\xi^{\frac{n+2r}{2}-1}e^{-\frac{\xi}{2}}d\xi$$

$$= \frac{2^r\Gamma\left(\frac{n}{2}+r\right)}{\Gamma\left(\frac{n}{2}\right)}\left(\text{当且仅当}\frac{n}{2}+r>0,\text{即}r>-\frac{n}{2}\text{时存在}\right)$$

$$D(\xi^r) = E(\xi^{2r}) - E^2(\xi^r) = \frac{2^{2r}\Gamma\left(\frac{n}{2}+2r\right)}{\Gamma\left(\frac{n}{2}\right)} - \frac{2^{2r}\left[\Gamma\left(\frac{n}{2}+r\right)\right]^2}{\left[\Gamma\left(\frac{n}{2}\right)\right]^2}\left(r>-\frac{n}{4}\text{时存在}\right)$$

$$f(\xi) = \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}\xi^{\frac{n}{2}-1}e^{-\frac{\xi}{2}}$$

$$f'(\xi) = \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}\xi^{\frac{n}{2}-2}e^{-\frac{\xi}{2}}\left(\frac{n}{2}-1-\frac{\xi}{2}\right) = 0,\text{得}\xi = n-2$$

即ξ的密度最大值在$n-2$处达到.

21、若$r.v.\ X$服从自由度 n 的 t 分布，$n>1$,证明对$r<n, E(X^r)$存在且

$$E(X^r) = \begin{cases} n^{\frac{r}{2}}\cdot\dfrac{\Gamma\left(\frac{r+1}{2}\right)\Gamma\left(\frac{n-r}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{1}{2}\right)}, & r\text{为偶数},\\ 0, & r\text{为奇数}.\end{cases}$$

解：

$t_n(x)=\dfrac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{n\pi}}\left(1+\dfrac{x^2}{n}\right)^{-\frac{n+1}{2}}$,$t_n(x)$为偶函数，故$r$为奇数时，$E(X^r)=0$.

r为偶数时，$E(X^r)=E\left[(X^2)^{\frac{r}{2}}\right]$,又$t^2\sim F(1,n)$

已知若$X\sim Be(a,b),f(x)=\dfrac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^{a-1}(1-x)^{b-1}$,$0<x<1$

若$\xi\sim Be(a,b)$，则$\eta=\dfrac{b\xi}{a(1-\xi)}\sim F(2a,2b)$

故$t^2=\dfrac{\frac{n}{2}\xi}{\frac{1}{2}(1-\xi)}$，其中$\xi\sim Be\left(\dfrac{1}{2},\dfrac{n}{2}\right)$

则$E(X^r)=\displaystyle\int_0^1\left[\dfrac{\frac{n}{2}\xi}{\frac{1}{2}(1-\xi)}\right]^{\frac{r}{2}}\cdot f_\xi(\xi)d\xi$

$=n^{\frac{r}{2}}\dfrac{\Gamma\left(\frac{1+n}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{n}{2}\right)}\displaystyle\int_0^1\xi^{\frac{1+r}{2}-1}(1-\xi)^{\frac{n-r}{2}-1}d\xi$

$=n^{\frac{r}{2}}\dfrac{\Gamma\left(\frac{1+n}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{n}{2}\right)}\cdot\dfrac{\Gamma\left(\frac{1+r}{2}\right)\Gamma\left(\frac{n-r}{2}\right)}{\Gamma\left(\frac{1+n}{2}\right)}$

$=n^{\frac{r}{2}}\cdot\dfrac{\Gamma\left(\frac{r+1}{2}\right)\Gamma\left(\frac{n-r}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{1}{2}\right)}$，故得证.

22、设$r.v.\ \xi_n\sim t_n$,计算ξ_n的方差$D(\xi_n)$,并说明n取何值时，$D(\xi_n)$才存在.

解：

$D(\xi_n)=E({\xi_n}^2)-E^2(\xi_n)$

$E(\xi_n)=0$,由21题可知

$E({\xi_n}^2)=n\cdot\dfrac{\Gamma\left(\frac{2+1}{2}\right)\Gamma\left(\frac{n-2}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{1}{2}\right)}$,$n>2$时存在

即$n\geqslant3$时，$E({\xi_n}^2)=n\cdot\dfrac{\Gamma\left(\frac{3}{2}\right)\Gamma\left(\frac{n}{2}-1\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{1}{2}\right)}=\dfrac{n}{2}\cdot\dfrac{\Gamma\left(\frac{n}{2}-1\right)}{\Gamma\left(\frac{n}{2}\right)}=\dfrac{n}{n-2}$

故$n\geqslant3$时，$D(\xi_n)$存在且$D(\xi_n)=\dfrac{n}{n-2}$

23、证明当$n\to\infty$时t_n分布趋于标准正态分布.

解：

$$t_n(x)=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{n\pi}}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}},\quad x\in R$$

$$n=2k\text{时，}\ \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{n\pi}}=\frac{\Gamma\left(\frac{1}{2}+k\right)}{\Gamma(k)\sqrt{2k\pi}}=\frac{\frac{(2k-1)!!}{2^k}\sqrt{\pi}}{\sqrt{2k\pi}\,(k-1)!}=\frac{(2k-1)!!}{(2k)!!}\sqrt{\frac{k}{2}}$$

$$=\sqrt{\frac{k}{2}\cdot\frac{1}{k\pi}}=\frac{1}{\sqrt{2\pi}}\left(\text{注：}\frac{(2n-1)!!}{(2n)!!}\approx\frac{1}{\sqrt{n\pi}}\right)$$

$$n=2k+1\text{时，}\ \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{n\pi}}=\frac{\Gamma(k+1)}{\sqrt{(2k+1)\pi}\,\Gamma\left(\frac{1}{2}+k\right)}=\frac{k!}{\sqrt{(2k+1)\pi}\frac{(2k-1)!!}{2^k}\sqrt{\pi}}$$

$$=\frac{\sqrt{2k+1}}{\pi}\frac{(2k)!!}{(2k+1)!!}=\frac{\sqrt{2k+1}}{\pi}\cdot\sqrt{k\pi}=\frac{1}{\sqrt{2\pi}}$$

又因为$\lim\limits_{n\to\infty}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}=e^{-\frac{x^2}{2}}$

故$\lim\limits_{n\to\infty}t_n(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}=\varphi(x)$，故得证

24、设$r.v.\ X$服从参数为α，p的 Gamma 分布$\Gamma(p,\alpha)$，求证

（1）$\Gamma(p,\alpha)$的特征函数为$\varphi(t)=\left(\frac{\alpha}{\alpha-it}\right)^p$;

（2）$E(X)=\frac{p}{\alpha}$，$D(X)=\frac{p}{\alpha^2}$;

（3）如果$X_i\sim\Gamma(p_i,\alpha),i=1,2,\cdots,k$,且$X_1,\cdots,X_n$相互独立，记$p=\sum\limits_{i=1}^{n}p_i$,

则$\sum\limits_{i=1}^{n}X_i\sim\Gamma(p,\alpha)$;

（4）若取$\alpha=\frac{1}{2}$，$p=\frac{n}{2}$,则$\Gamma\left(\frac{n}{2},\frac{1}{2}\right)$是$\chi_n^2$分布.

解：

（1）$X\sim\Gamma(p,\alpha)$,则$f(x)=\frac{\alpha^p}{\Gamma(p)}x^{p-1}e^{-\alpha x},x>0$

$$\varphi(t)=\int_0^{+\infty}e^{itx}\frac{\alpha^p}{\Gamma(p)}x^{p-1}e^{-\alpha x}dx=\frac{\alpha^p}{(\alpha-it)^p}\int_0^{+\infty}\frac{(\alpha-it)^p}{\Gamma(p)}x^{p-1}e^{-(\alpha-it)x}dx=\left(\frac{\alpha}{\alpha-it}\right)^p$$

（2）$\varphi'(t)=(-p)\left(-\frac{i}{\alpha}\right)\left(1-\frac{it}{\alpha}\right)^{-p-1}$

$\varphi'(0)=iE(X)=\frac{p}{\alpha}i$,故$E(X)=\frac{p}{\alpha}$

$\varphi''(t)=\frac{p(1+p)}{\alpha^2}i^2\left(1-\frac{it}{\alpha}\right)^{-p-2}$

$\varphi''(0)=i^2E(X^2)=\dfrac{p(1+p)}{\alpha^2}i^2$,故$E(X^2)=\dfrac{p(1+p)}{\alpha^2}$

则$D(X)=E(X^2)-E^2(X)=\dfrac{p}{\alpha^2}$

（3）$\varphi_{\sum\limits_{i=1}^{n}X_i}(t)=\prod\limits_{i=1}^{n}\left(\dfrac{\alpha}{\alpha-it}\right)^{p_i}=\left(\dfrac{\alpha}{\alpha-it}\right)^{\sum\limits_{i=1}^{n}p_i}$

即$\sum\limits_{i=1}^{n}X_i\sim\Gamma(p,\alpha),p=\sum\limits_{i=1}^{n}p_i$

（4）$\Gamma\left(\dfrac{n}{2},\dfrac{1}{2}\right)$的特征函数$\varphi(t)=(1-2it)^{-\frac{n}{2}}$

由特征函数的唯一性定理知$\Gamma\left(\dfrac{n}{2},\dfrac{1}{2}\right)$是$\chi_n^2$分布.

26,、某电子元件寿命服从指数分布，其密度函数为

$$f(x)=\begin{cases}\dfrac{1}{\lambda}e^{-x/\lambda}, & x>0\\ 0, & x\leqslant 0\end{cases}$$

从这批产品中抽取 n 个作寿命试验，规定到第r个$(0<r\leqslant n)$电子元件失效时就停止试验.

这样获得前r个次序统计量$X_{(1)}\leqslant X_{(2)}\leqslant\cdots\leqslant X_{(r)}$和$n$个电子元件总试验时间

$T=\sum\limits_{i=1}^{r}X_{(i)}+(n-r)X_{(r)}$,证明$2T/\lambda$服从自由度为$2r$的$\chi^2$分布，即$2T/\lambda\sim\chi_{2r}^2$.

解：

$$f(x_{(1)},x_{(2)},\cdots,x_{(n)})=\begin{cases}n!\lambda^{-n}e^{-\frac{1}{\lambda}\sum\limits_{i=1}^{n}x_{(i)}}, & x_{(1)}<x_{(2)}<\cdots<x_{(n)}\\ 0, & \text{其它}\end{cases}$$

$T=\sum\limits_{i=1}^{r}X_{(i)}+(n-r)X_{(r)}=x_{(1)}+x_{(2)}+\cdots+x_{(r-1)}+(n-r+1)x_{(r)}$

令$Z_i=X_{(i)}-X_{(i-1)}$，$i=1,2,\cdots,n$　设$X_{(0)}=0$

则$X_{(i)}=\sum\limits_{i=1}^{n}Z_i$

则$J=\begin{vmatrix}1&0&0&\cdots&0\\1&1&0&\cdots&0\\1&1&1&\cdots&0\\\vdots&\vdots&\vdots&&\vdots\\1&1&1&\cdots&1\end{vmatrix}=1$

$g(z_1,z_2,\ldots,z_n)=f\left(z_1,z_1+z_2,\ldots,\sum\limits_{k=1}^{i}z_k,\ldots,\sum\limits_{k=1}^{n}z_k\right)|J|$

$$=\begin{cases} n!\lambda^{-n}e^{-\frac{1}{\lambda}\sum\limits_{i=1}^{n}\sum\limits_{k=1}^{i}z_k}, & z_1,z_2,\dots,z_n>0 \\ 0, & \text{其它} \end{cases}$$

$$=\begin{cases} n!\lambda^{-n}e^{-\frac{1}{\lambda}\sum\limits_{k=1}^{n}(n-k+1)z_k}, & z_1,z_2,\dots,z_n>0 \\ 0, & \text{其它} \end{cases}$$

$$=\begin{cases} \prod\limits_{i=1}^{n}\dfrac{n-i+1}{\lambda}e^{-\frac{1}{\lambda}(n-i+1)z_i}, & z_1,z_2,\dots,z_n>0 \\ 0, & \text{其它} \end{cases}$$

故$Z_i\sim Exp\left(\dfrac{n-i+1}{\lambda}\right)$,则由推论2.4.5可知

$$2T/\lambda=\frac{2\sum\limits_{i=1}^{r}(n-i+1)z_i}{\lambda}\sim\chi^2_{2r}.$$

27、设总体X服从双指数分布，其分布函数为

$$F(x)=\begin{cases} 1-\exp\left\{-\dfrac{x-\mu}{\sigma}\right\}, & x>\mu, \\ 0, & x\leqslant\mu. \end{cases}$$

其中$-\infty<\mu<+\infty, 0<\sigma<+\infty, X_{(1)}\leqslant\cdots\leqslant X_{(n)}$是样本$X_1,\cdots,X_n$的次序统计量.

试证明$\dfrac{2(n-i+1)}{\sigma}(X_{(i)}-X_{(i-1)})$服从自由度为2的$\chi^2$分布$(i=2,\cdots,n)$.

解：

$$f(x_{(1)},x_{(2)},\cdots,x_{(n)})=\begin{cases} n!\sigma^{-n}e^{-\frac{1}{\sigma}\left(\sum\limits_{i=1}^{n}x_{(i)}-n\mu\right)}, & x_{(1)}<x_{(2)}<\cdots<x_{(n)} \\ 0, & \text{其它} \end{cases}$$

令$Z_i=X_{(i)}-X_{(i-1)}$，则$X_{(i)}=\sum\limits_{i=1}^{n}Z_i$

$$\text{则}J=\begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 1 & 1 & \cdots & 1 \end{vmatrix}=1$$

$$g(z_1,z_2,\dots,z_n)=f\left(z_1,z_1+z_2,\dots,\sum_{k=1}^{i}z_k,\dots,\sum_{k=1}^{n}z_k\right)|J|$$

$$=\begin{cases} n!\sigma^{-n}e^{-\frac{1}{\sigma}\left(\sum\limits_{k=1}^{n}(n-k+1)z_k-n\mu\right)}, & z_1,z_2,\dots,z_n>0 \\ 0, & \text{其它} \end{cases}$$

$$=\begin{cases}\prod_{k=1}^{n}\dfrac{n-k+1}{\sigma}e^{-\frac{1}{\sigma}((n-i+1)z_k-\mu)}, & z_1,z_2,\dots,z_n>0\\ 0, & \text{其它}\end{cases}$$

可知$Z_i\sim Exp\left(\dfrac{n-i+1}{\sigma},\mu\right)$,则$2\dfrac{n-i+1}{\sigma}Z_i\sim\chi_2^2$

28、设$X_1\sim\Gamma(\alpha_1,\lambda),X_2\sim\Gamma(\alpha_2,\lambda)$,且$X_1$与$X_2$独立，则$Y_1=X_1+X_2$,与$Y_2=X_1/(X_1+X_2)$亦独立，且$Y_1\sim\Gamma(\alpha_1+\alpha_2,\lambda)$，$Y_2=Be(\alpha_1,\alpha_2)$.

解：

$$\varphi_{X_1}(t)=\left(1-\frac{it}{\lambda}\right)^{-\alpha_1},\varphi_{X_2}(t)=\left(1-\frac{it}{\lambda}\right)^{-\alpha_2}$$

则$\varphi_{X_1+X_2}(t)=\left(1-\dfrac{it}{\lambda}\right)^{-(\alpha_1+\alpha_2)}$

由特征函数的唯一性定理可知$Y_1=X_1+X_2\sim\Gamma(\alpha_1+\alpha_2,\lambda)$

即$f_{Y_1}(y_1)=\dfrac{\lambda^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1+\alpha_2)}y_1^{\alpha_1+\alpha_2-1}e^{-\lambda y_1},y_1>0$

$$f(x_1,x_2)=\frac{\lambda^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)}x_1^{\alpha_1-1}x_2^{\alpha_2-1}e^{-\lambda(x_1+x_2)},x_1,x_2>0$$

令$\begin{cases}u=\dfrac{x_1}{x_1+x_2}\\ v=x_1\end{cases}$,解得$\begin{cases}x_1=v\\ x_2=\dfrac{v}{u}-v\end{cases}$,则$J=\dfrac{\partial(x_1,x_2)}{\partial(u,v)}=\dfrac{v}{u^2}$

$$f(u,v)=\frac{\lambda^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)}v^{\alpha_1+\alpha_2-2}\left(\frac{1}{u}-1\right)^{\alpha_2-1}e^{-\lambda\frac{v}{u}}\frac{v}{u^2}$$

则$f_U(u)=\displaystyle\int_{-\infty}^{+\infty}f(u,v)dv=\frac{\lambda^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\int_{-\infty}^{+\infty}v^{\alpha_1+\alpha_2-1}e^{-\lambda\frac{v}{u}}dv(1-u)^{\alpha_2-1}u^{-\alpha_2-1}$

$$\xrightarrow{\lambda\frac{v}{u}=t}\frac{\lambda^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\int_{-\infty}^{+\infty}t^{\alpha_1+\alpha_2-1}e^{-t}dt\frac{u^{\alpha_1+\alpha_2}}{\lambda^{\alpha_1+\alpha_2}}(1-u)^{\alpha_2-1}u^{-\alpha_2-1}$$

$$=\frac{\Gamma(\alpha_1+\alpha_2)\lambda^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\int_{-\infty}^{+\infty}\frac{1}{\Gamma(\alpha_1+\alpha_2)}t^{\alpha_1+\alpha_2-1}e^{-t}dt\frac{u^{\alpha_1+\alpha_2}}{\lambda^{\alpha_1+\alpha_2}}(1-u)^{\alpha_2-1}u^{-\alpha_2-1}$$

$$=\frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)}u^{\alpha_1-1}(1-u)^{\alpha_2-1}$$

故$Y_2\sim Be(\alpha_1,\alpha_2)$

由$f(y_1,y_2)=f(x_1,x_2)\left|\dfrac{\partial(x_1,x_2)}{\partial(y_1,y_2)}\right|=\dfrac{y_1\lambda^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)}(y_1y_2)^{\alpha_1-1}(y_1-y_1y_2)^{\alpha_2-1}e^{-\lambda y_1}$

$$=\frac{\lambda^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1+\alpha_2)}y_1^{\alpha_1+\alpha_2-1}e^{-\lambda y_1}\cdot\frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)}y_2^{\alpha_1-1}(1-y_2)^{\alpha_2-1}=f_{Y_1}(y_1)f_{Y_2}(y_2)$$

即Y_1与Y_2独立

29、证明定理 2.4.2.

解：

$$X\sim N(0,1),f_X(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}},Y\sim\chi_n^2,f_Y(y)=\frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}y^{\frac{n}{2}-1}e^{-\frac{y}{2}}$$

$$令Z=\sqrt{Y/n},则f_Z(z)=f_Y(nz^2)\left|\frac{d(nz^2)}{dz}\right|=\frac{2nz}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}(nz^2)^{\frac{n}{2}-1}e^{-\frac{nz^2}{2}}$$

$$令\begin{cases}u=\dfrac{x}{z}\\v=x\end{cases},则\begin{cases}x=v\\z=\dfrac{v}{u}\end{cases},J=\frac{\partial(x,z)}{\partial(u,v)}=\begin{vmatrix}0&1\\-\dfrac{v}{u^2}&\dfrac{1}{u}\end{vmatrix}=\frac{v}{u^2}$$

$$f(u,v)=\frac{1}{\sqrt{2\pi}}e^{-\frac{v^2}{2}}\frac{2nv}{u2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}\left(n\frac{v^2}{u^2}\right)^{\frac{n}{2}-1}e^{-\frac{nv^2}{2u^2}-\frac{v^2}{2}}\frac{v}{u^2}$$

$$=\frac{n^{\frac{n}{2}}}{2^{\frac{n}{2}-1}\sqrt{2\pi}\,\Gamma\left(\frac{n}{2}\right)}\frac{v^n}{u^{n+1}}e^{-\frac{nv^2}{2u^2}-\frac{v^2}{2}}$$

$$故f_U(u)=\frac{n^{\frac{n}{2}}}{2^{\frac{n}{2}-1}\sqrt{2\pi}\,\Gamma\left(\frac{n}{2}\right)}\int_{-\infty}^{+\infty}\frac{v^n}{u^{n+1}}e^{-\frac{nv^2}{2u^2}-\frac{v^2}{2}}dv$$

$$\xrightarrow{\frac{nv^2}{2u^2}+\frac{v^2}{2}=t}\frac{n^{\frac{n}{2}}}{2^{\frac{n}{2}-1}\sqrt{2\pi}\,\Gamma\left(\frac{n}{2}\right)}\int_0^{+\infty}\frac{2^{\frac{n-1}{2}}}{(n+u^2)^{\frac{n+1}{2}}}t^{\frac{n-1}{2}}e^{-t}dt=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{n\pi}}\left(1+\frac{u^2}{n}\right)^{-\frac{n+1}{2}}$$

得证.

30、证明定理 2.4.3.

解：

$$X\sim\chi_m^2,f_X(x)=\frac{1}{2^{\frac{m}{2}}\Gamma\left(\frac{m}{2}\right)}x^{\frac{m}{2}-1}e^{-\frac{x}{2}},Y\sim\chi_n^2,f_Y(y)=\frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}y^{\frac{n}{2}-1}e^{-\frac{y}{2}}$$

$$U=X/m,V=Y/n$$

$$f_U(u)=f_X(mu)m=\frac{m}{2^{\frac{m}{2}}\Gamma\left(\frac{m}{2}\right)}(mu)^{\frac{m}{2}-1}e^{-\frac{mu}{2}}=\frac{m^{\frac{m}{2}}}{2^{\frac{m}{2}}\Gamma\left(\frac{m}{2}\right)}u^{\frac{m}{2}-1}e^{-\frac{mu}{2}}$$

$$f_V(v)=f_Y(nv)n=\frac{n}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}(nv)^{\frac{n}{2}-1}e^{-\frac{nv}{2}}=\frac{n^{\frac{n}{2}}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}v^{\frac{n}{2}-1}e^{-\frac{nv}{2}}$$

$$f(u,v)=\frac{m^{\frac{m}{2}}}{2^{\frac{m}{2}}\Gamma\left(\frac{m}{2}\right)}\frac{n^{\frac{n}{2}}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}u^{\frac{m}{2}-1}v^{\frac{n}{2}-1}e^{-\frac{mu}{2}}e^{-\frac{nv}{2}}$$

$$F=\frac{U}{V}$$

$$f_F(f)=\int_0^{+\infty}f(vf,v)vdv=\int_0^{+\infty}\frac{m^{\frac{m}{2}}}{2^{\frac{m}{2}}\Gamma\left(\frac{m}{2}\right)}\frac{n^{\frac{n}{2}}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}(vf)^{\frac{m}{2}-1}v^{\frac{n}{2}-1}e^{-\frac{mvf}{2}}e^{-\frac{nv}{2}}vdv$$

$$=\frac{m^{\frac{m}{2}}}{2^{\frac{m}{2}}\Gamma\left(\frac{m}{2}\right)}\frac{n^{\frac{n}{2}}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}f^{\frac{m}{2}-1}\int_0^{+\infty}v^{\frac{m+n}{2}-1}e^{-\frac{mf+n}{2}v}dv$$

$$\xrightarrow{\frac{mf+n}{2}v=t}\frac{m^{\frac{m}{2}}}{2^{\frac{m}{2}}\Gamma\left(\frac{m}{2}\right)}\frac{n^{\frac{n}{2}}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}f^{\frac{m}{2}-1}\int_0^{+\infty}\left(\frac{2}{mf+n}\right)^{\frac{m+n}{2}}t^{\frac{m+n}{2}-1}e^{-t}dt$$

$$=\frac{m^{\frac{m}{2}}}{2^{\frac{m}{2}}\Gamma\left(\frac{m}{2}\right)}\frac{n^{\frac{n}{2}}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}f^{\frac{m}{2}-1}\left(\frac{2}{mf+n}\right)^{\frac{m+n}{2}}\Gamma\left(\frac{m+n}{2}\right)$$

$$=\frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}m^{\frac{m}{2}}n^{\frac{n}{2}}f^{\frac{m}{2}-1}(mf+n)^{-\frac{m+n}{2}},f>0$$

得证.

31、记$r.v.\xi\sim F_{m,n}$,，求$E(\xi)$和$D(\xi)$，并指出 m, n 取什么值时这些量才存在？

解

$$\xi=\frac{\frac{U}{m}}{\frac{W}{n}},\text{其中}U\sim\chi_m^2,W\sim\chi_n^2$$

$$E(\xi^r)=\frac{n^r}{m^r}E(v^r)E(w^{-r}),\text{由19题可知}$$

$$E(v^r)=\frac{2^r\Gamma\left(\frac{m}{2}+r\right)}{\Gamma\left(\frac{m}{2}\right)}\left(r>-\frac{m}{2}\right),\quad E(w^{-r})=\frac{\Gamma\left(\frac{n}{2}-r\right)}{2^r\Gamma\left(\frac{n}{2}\right)}\left(r<\frac{n}{2}\right)$$

$$E(\xi^r)=\frac{n^r}{m^r}\frac{\Gamma\left(\frac{m}{2}+r\right)\Gamma\left(\frac{n}{2}-r\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}\left(-\frac{m}{2}<r<\frac{n}{2}\right)$$

$$E(\xi)=\frac{n}{m}\frac{\Gamma\left(\frac{m}{2}+1\right)\Gamma\left(\frac{n}{2}-1\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}=\frac{n}{m}\frac{\frac{m}{2}}{\frac{n}{2}-1}=\frac{n}{n-2},\quad(m>-2,n>2)$$

$$E(\xi^2)=\frac{n^2}{m^2}\frac{\Gamma\left(\frac{m}{2}+2\right)\Gamma\left(\frac{n}{2}-2\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}=\frac{n^2}{m^2}\frac{\left(\frac{m}{2}+1\right)\frac{m}{2}}{\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2\right)}=\frac{n^2(m+2)}{m(n-2)(n-4)},(m>-2,n>4)$$

$$D(\xi)=E(\xi^2)-E^2(\xi)=\frac{n^2(m+2)}{m(n-2)(n-4)}-\frac{n^2}{(n-2)^2},(m>-2,n>4)$$

32、若$X\sim F_{m,n}$,记$A=F_{m,n}(1-\alpha)\left(\text{即}\int_0^A F_{m,n}(x)dx=\alpha\right)$,证明

$$F_{m,n}(1-\alpha)=\frac{1}{F_{n,m}(\alpha)}.$$

解:

$P(x \leqslant A)=\alpha, A=F_{m,n}(1-\alpha)$

$\Longrightarrow P\left(\frac{1}{x} \geqslant \frac{1}{A}\right)=\alpha \Longrightarrow P\left(\frac{1}{x}<\frac{1}{A}\right)=1-\alpha,$

又因为$X \sim F_{m,n}$,所以$\frac{1}{X} \sim F_{n,m}$,即$\frac{1}{A}=F_{n,m}(\alpha)$，故$F_{m,n}(1-\alpha)=\frac{1}{F_{n,m}(\alpha)}$.

33、设总体X的$2k$原点矩$\alpha_{2k}=E(X^{2k})<\infty$，说明样本的$k$阶原点矩$A_k=\frac{1}{n}\sum_{i=1}^{n} X_i^{\,k}$

的渐进分布为$N(\alpha_k,(\alpha_{2k}-\alpha_k^{\,2})/n)$.

解：

$E(A_k)=\alpha_k, D(A_k)=\frac{\alpha_{2k}-\alpha_k^{\,2}}{n}$,由中心极限定理可知当$n \to \infty$时，

$\sqrt{n}\frac{A_k-\alpha_k}{\sqrt{\alpha_{2k}-\alpha_k^{\,2}}} \xrightarrow{\mathscr{L}} N(0,1)$

34、设$\xi_n \sim \chi_n^{\,2}$,证明当$n \to \infty$时$(\xi-n)/\sqrt{2n} \xrightarrow{\mathscr{L}} N(0,1)$.利用这一事实给出$\chi_n^{\,2}$的$p$

分位点与$N(0,1)$的p分位点之间的一个近似关系.

解：

$E(\xi_n)=n, D(\xi_n)=2n$,由中心极限定理可知当$n \to \infty$时

$(\xi-n)/\sqrt{2n} \xrightarrow{\mathscr{L}} N(0,1)$

$P(\xi_n>\chi_n^{\,2}(p))=p \Longleftrightarrow P(\xi_n<\chi_n^{\,2}(p))=1-p=\Phi\left(\frac{\chi_n^{\,2}(p)-n}{\sqrt{2n}}\right).$

35、设$X_1,\cdots,X_n$为来自均匀分布总体$U(0,\theta)$的简单样本，根据中心极限定理，求样本

均值$\overline{X}$的渐近分布.

解：

$E\overline{X}=\frac{\theta}{2}, D\overline{X}=\frac{\theta^2}{12n}$,由中心极限定理可知当$n \to \infty$时

$\frac{\overline{X}-\frac{\theta}{2}}{\sqrt{\frac{\theta^2}{12n}}} \xrightarrow{\mathscr{L}} N(0,1)$，即$\overline{X}$的渐进分布为$N\left(\frac{\theta}{2},\frac{\theta^2}{12n}\right)$.

36、设$X_1,\cdots,X_n$为取自$Poission$总体$P(\lambda)$的样本.试证$\left(\overline{X}-\lambda\right)\Big/\sqrt{\overline{X}/n}$的极限分布

为$N(0,1)$.

解：

$E\overline{X}=\lambda,D\overline{X}=\dfrac{\lambda}{n}$,由中心极限定理可知当$n\to\infty$时

$\dfrac{\overline{X}-\lambda}{\sqrt{\lambda/n}}\xrightarrow{\mathscr{L}}N(0,1)$,由大数定律可知$\sqrt{\dfrac{\lambda}{\overline{X}}}\xrightarrow{p}1$.

再由$Slutsky$引理知$\dfrac{\overline{X}-\lambda}{\sqrt{\lambda/n}}\cdot\sqrt{\dfrac{\lambda}{\overline{X}}}=\dfrac{\overline{X}-\lambda}{\sqrt{\overline{X}/n}}\xrightarrow{\mathscr{L}}N(0,1)$.

37、设$X_1,\cdots,X_m$和$Y_1,\cdots,Y_n$分别取自正态总体$N(\mu,\sigma_1^2)$和$N(\mu,\sigma_2^2)$的样本，试证明当m,n都趋于无穷时，统计量$\left(\overline{X}-\overline{Y}\right)\Big/\sqrt{\frac{S_X^2}{m}+\frac{S_Y^2}{n}}\xrightarrow{\mathscr{L}}N(0,1)$.其中$S_X^2$和$S_Y^2$分别为两组样本的样本方差.

解

$\overline{X}$与$\overline{Y}$独立，$E\left(\overline{X}-\overline{Y}\right)=0,D\left(\overline{X}-\overline{Y}\right)=\dfrac{\sigma_1^2}{m}+\dfrac{\sigma_2^2}{n}$

由中心极限定理可知当$n\to\infty$时,$\left(\overline{X}-\overline{Y}-0\right)\Big/\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}\xrightarrow{\mathscr{L}}N(0,1)$

由辛钦大数定律可知$S_X^2\xrightarrow{P}\sigma_1^2,S_Y^2\xrightarrow{P}\sigma_2^2$,故$\sqrt{\dfrac{S_X^2}{m}+\dfrac{S_Y^2}{n}}\xrightarrow{P}\sqrt{\dfrac{\sigma_1^2}{m}+\dfrac{\sigma_2^2}{n}}$,

由$Slutsky$ 引理可知$\left(\overline{X}-\overline{Y}\right)\Big/\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}\cdot\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}\Big/\sqrt{\frac{S_X^2}{m}+\frac{S_Y^2}{n}}$

$=\left(\overline{X}-\overline{Y}\right)\Big/\sqrt{\frac{S_X^2}{m}+\frac{S_Y^2}{n}}\xrightarrow{\mathscr{L}}N(0,1)$.

38、分别列出一个单参数指数族和多参数指数族的例子

单参数指数族：$P(\lambda),B(n,p)$,多参数指数族$N(\mu,\sigma^2)$.

39、将负二项分布和指数分布写成指数族的标准形式，并求出其自然参数空间.

解：

负二项分布：$P(x;\theta)=\dbinom{x-1}{r-1}\theta^r(1-\theta)^{x-r}$, $x=r,r+1,\cdots$

$P(x;\theta)=\left(\dfrac{\theta}{1-\theta}\right)^r(1-\theta)^x\dbinom{x-1}{r-1}=\left(\dfrac{\theta}{1-\theta}\right)^re^{\ln(1-\theta)x}\dbinom{x-1}{r-1}$

自然形式：$\ln(1-\theta)=\varphi$，$\theta=1-e^{\varphi}$

$P(x;\theta)=\left(\dfrac{1-e^{\varphi}}{e^{\varphi}}\right)^re^{\varphi x}\dbinom{x-1}{r-1}=C^*(\varphi)\exp\{\varphi T(x)\}h(x)$,

$\Theta^*=(\varphi\colon-\infty<\varphi<0),h(x)=I_{(x>0)}$

指数分布：$f(x;\theta)=\theta e^{-\theta x}=C^*(\varphi)\exp\{\varphi T(x)\}h(x),\Theta^*=\{\theta\colon\theta>0\},\varphi=\theta,h(x)=I_{(x>0)}$

40、设指数族的自然形式为$f(x,\theta)=C(\theta)\exp\left\{\sum\limits_{j=1}^k\theta_jT_j(x)\right\}h(x)$,证明

$E_\theta(T_j(x))=-\dfrac{\partial\log C(\theta)}{\partial\theta_j}=-\dfrac{1}{C(\theta)}\dfrac{\partial C(\theta)}{\partial\theta_j},Cov(T_j(x),T_s(x))=-\dfrac{\partial^2\log C(\theta)}{\partial\theta_j\partial\theta_s}$.

解：

$\int_0^{+\infty} f(x,\theta)dx=1$,故$C(\theta)=\dfrac{1}{\int_0^{+\infty}\exp\left\{\sum_{j=1}^{k}\theta_j T_j(x)\right\}h(x)dx}=\dfrac{1}{p(\theta)}$

$$\frac{\partial \log C(\theta)}{\partial\theta_j}=\frac{1}{C(\theta)}\frac{\partial C(\theta)}{\partial\theta_j}=-p(\theta)\frac{\int_0^{+\infty}T_j(x)\exp\left\{\sum_{j=1}^{k}\theta_j T_j(x)\right\}h(x)dx}{(p(\theta))^2}$$

$$=-\int_0^{+\infty}T_j(x)C(\theta)\exp\left\{\sum_{j=1}^{k}\theta_j T_j(x)\right\}h(x)dx=-E_\theta(T_j(x)),$$

即$E_\theta(T_j(x))=-\dfrac{\partial\log C(\theta)}{\partial\theta_j}=-\dfrac{1}{C(\theta)}\dfrac{\partial C(\theta)}{\partial\theta_j}$

$$\frac{\partial^2\log C(\theta)}{\partial\theta_j\partial\theta_s}=-\frac{1}{C^2(\theta)}\frac{\partial C(\theta)}{\partial\theta_s}\frac{\partial C(\theta)}{\partial\theta_j}+\frac{1}{C(\theta)}\frac{\partial^2 C(\theta)}{\partial\theta_j\partial\theta_s}$$

$$=-(p(\theta))^2\left[-\frac{\int_0^{+\infty}T_s(x)\exp\left\{\sum_{j=1}^{k}\theta_j T_j(x)\right\}h(x)dx}{(p(\theta))^2}\right]\cdot$$

$$\left[-\frac{\int_0^{+\infty}T_j(x)\exp\left\{\sum_{j=1}^{k}\theta_j T_j(x)\right\}h(x)dx}{(p(\theta))^2}\right]+p(\theta)\cdot$$

$$\left[\begin{array}{c}-\dfrac{(p(\theta))^2\int_0^{+\infty}T_s(x)T_j(x)\exp\left\{\sum_{j=1}^{k}\theta_j T_j(x)\right\}h(x)dx}{(p(\theta))^4}+\\ \dfrac{2p(\theta)\int_0^{+\infty}T_s(x)\exp\left\{\sum_{j=1}^{k}\theta_j T_j(x)\right\}h(x)dx\int_0^{+\infty}T_j(x)\exp\left\{\sum_{j=1}^{k}\theta_j T_j(x)\right\}h(x)dx}{(p(\theta))^4}\end{array}\right]$$

$$=-E_\theta(T_j(x))E_\theta(T_s(x))-E_\theta(T_j(x)T_s(x))+2E_\theta(T_j(x))E_\theta(T_s(x))$$

$$=E_\theta(T_j(x))E_\theta(T_s(x))-E_\theta(T_j(x)T_s(x))=-Cov(T_j(x),T_s(x))$$

即$Cov(T_j(x),T_s(x))=-\dfrac{\partial^2\log C(\theta)}{\partial\theta_j\partial\theta_s}$.

41、证明指数族的自然参数空间为凸集.

解：

自然参数空间：$\Theta=\left\{\boldsymbol{\theta}=(\theta_1,\cdots,\theta_k):\int_{\mathscr{X}}\exp\left\{\sum_{i=1}^{k}\theta_i T_i(x)\right\}d\mu(x)<\infty\right\},d\mu(x)=h(x)dx$

设$\boldsymbol{\theta}'=(\theta_1',\cdots,\theta_k')$，$\boldsymbol{\theta}''=(\theta_1'',\cdots,\theta_k'')\in\Theta$,又设$0<\alpha<1$,即证$\boldsymbol{\theta}=\alpha\boldsymbol{\theta}'+(1-\alpha)\boldsymbol{\theta}''=(\theta_1,\cdots,\theta_k)\in\Theta$

$$\int_{\mathscr{X}}\exp\left\{\sum_{i=1}^{k}\theta_i T_i(x)\right\}d\mu(x)$$

$$=\int_{\mathscr{X}}\left[\exp\left\{\sum_{i=1}^{k}\theta_i' T_i(x)\right\}\right]^{\alpha}\left[\exp\left\{\sum_{i=1}^{k}\theta_i'' T_i(x)\right\}\right]^{1-\alpha}d\mu(x)$$

$$\leqslant\left[\int_{\mathscr{X}}\exp\left\{\sum_{i=1}^{k}\theta_i{}'T_i(x)\right\}d\mu(x)\right]^{\alpha}\left[\int_{\mathscr{X}}\exp\left\{\sum_{i=1}^{k}\theta_i{}''T_i(x)\right\}d\mu(x)\right]^{1-\alpha}$$,[注：赫尔德不等式]

由于$\boldsymbol{\theta}'$，$\boldsymbol{\theta}''\in\Theta$，所以不等式右边两个积分存在，故$\boldsymbol{\theta}\in\Theta$，即指数族的自然参数空间为凸集.

42、设$\boldsymbol{X}=(X_1,\cdots,X_n)$是从$Poisson$分布$P(\lambda)$中抽取的简单样本，从定义出发证明

$T(\boldsymbol{X})=\sum\limits_{i=1}^{n}X_i$为充分统计量，再用因子分解定理证明之.

解：

定义证明：$P(X_1=x_1,\cdots,X_n=x_n|T=t)$

$$=\frac{P(X_1=x_1,\cdots,X_n=x_n,T=t)}{P(T=t)}$$

$$=\frac{\prod\limits_{i=1}^{n-1}\frac{\lambda^{x_i}}{x_i!}\cdot\frac{\lambda^{t-\sum\limits_{i=1}^{n-1}x_i}}{\left(t-\sum\limits_{i=1}^{n-1}x_i\right)!}e^{-n\lambda}}{\sum\limits_{k_1+\cdots+k_n=t}P(x_1=k_1,\cdots,x_n=k_n)}=\frac{\prod\limits_{i=1}^{n-1}\frac{1}{x_i!}\cdot\frac{\lambda^{t}}{\left(t-\sum\limits_{i=1}^{n-1}x_i\right)!}e^{-n\lambda}}{\sum\limits_{k_1+\cdots+k_n=t}\frac{\lambda^{k_1}}{k_1!}\cdots\frac{\lambda^{k_{n-1}}}{k_{n-1}!}\frac{\lambda^{t-\sum\limits_{i=1}^{n-1}k_i}}{\left(t-\sum\limits_{i=1}^{n-1}k_i\right)!}e^{-n\lambda}}$$

$$=\frac{\prod\limits_{i=1}^{n-1}\frac{1}{x_i!}\cdot\frac{1}{\left(t-\sum\limits_{i=1}^{n-1}x_i\right)!}}{\sum\limits_{k_1+\cdots+k_n=t}\frac{1}{k_1!}\cdots\frac{1}{k_{n-1}!}\frac{1}{\left(t-\sum\limits_{i=1}^{n-1}k_i\right)!}}$$,与λ无关，故$T(\boldsymbol{X})=\sum\limits_{i=1}^{n}X_i$为充分统计量

因子分解法：

$$f(\boldsymbol{x},\lambda)=\prod_{i=1}^{n}\frac{\lambda^{x_i}}{x_i!}e^{-\lambda}=\prod_{i=1}^{n}\frac{1}{x_i!}\lambda^{\sum\limits_{i=1}^{n}x_i}e^{-n\lambda}=g(t(\boldsymbol{x}),\lambda)h(x)$$

$g(t(\boldsymbol{x}),\lambda)=\lambda^{\sum\limits_{i=1}^{n}x_i}e^{-n\lambda},h(x)=\prod\limits_{i=1}^{n}\frac{1}{x_i!}$,故$T(\boldsymbol{X})=\sum\limits_{i=1}^{n}X_i$为充分统计量

43、设$X_1,\cdots,X_n$ $i.i.d.$ $\sim$几何分布，试用两种方法证明$T(\boldsymbol{X})=\sum\limits_{i=1}^{n}X_i$为充分统计量:

（1）从定义出发，（2）用因子分解定理.

解：

(1)$P(X_1=x_1,\cdots,X_n=x_n|T=t)=\dfrac{P(X_1=x_1,\cdots,X_n=x_n,T=t)}{P(T=t)}$

$$=\frac{q^{x_1}p\cdots q^{x_{n-1}}pq^{t-\sum\limits_{i=1}^{n-1}x_i-1}p}{\sum\limits_{k_1+\cdots+k_n=t}\prod\limits_{i=1}^{n}\left(q^{k_i-1}p\right)}=\frac{q^{t-n}p^{n}}{p^{n}\sum\limits_{k_1+\cdots+k_n=t}q^{t-n}}=\frac{1}{\sum\limits_{k_1+\cdots+k_n=t}1}=\frac{1}{\binom{t-1}{n-1}}$$

$$P(X_1=x_1,\cdots,X_n=x_n|T=t)=\begin{cases}\dfrac{1}{\binom{t-1}{n-1}}, & \sum\limits_{i=1}^{n}X_i=t \\ 0, & \sum\limits_{i=1}^{n}X_i\neq t\end{cases},\text{与}p\text{无关},$$

故$T(\boldsymbol{X})=\sum\limits_{i=1}^{n}X_i$为充分统计量

(2) $f(\boldsymbol{x},p)=\prod\limits_{i=1}^{n}(q^{x_i-1}q)=q^{t-n}p^n=g(t,p)h(x)$

其中$g(t,p)=q^{t-n}p^n$，$h(x)\equiv 1$

44、设$T=T(X)$是充分统计量，又$S(X)=G(T(X))$而函数$S=G(T(X))$是一对一的

即$T_1\neq T_2\Rightarrow G(T_1)\neq G(T_2)$,则$S$也是充分统计量.

解：

$T(X)$是充分统计量，故$f(\boldsymbol{x},\theta)=g(t(\boldsymbol{x}),\theta)h(\boldsymbol{x})$，而函数$S=G(T(X))$是一对一的

故存在$T(X)=M(S(X))$,故$f(\boldsymbol{x},\theta)=g(m(s(\boldsymbol{x})),\theta)h(\boldsymbol{x})=g'(s(\boldsymbol{x}),\theta)h(\boldsymbol{x})$

即S也是充分统计量.

45、设$X_1,\cdots,X_n$ $i.i.d.$ $\sim$指数分布$f(x;\lambda)=\lambda e^{-\lambda x}I_{(x>0)}$，用因子分解定理证明$\overline{X}$

是充分统计量.

解：

$$f(\boldsymbol{x},\lambda)=\lambda^n e^{-\lambda\sum\limits_{i=1}^{n}X_i}I_{[x_1>0,\cdots,x_n>0]}=g(t(\boldsymbol{x}),\lambda)h(\boldsymbol{x})$$

$g(t(\boldsymbol{x}),\lambda)=\lambda^n e^{-\lambda\sum\limits_{i=1}^{n}X_i}$,$h(\boldsymbol{x})=I_{[x_1>0,\cdots,x_n>0]}$,故$T=\sum\limits_{i=1}^{n}X_i$是充分统计量.

又因为$\overline{X}$与$\sum\limits_{i=1}^{n}X_i$一一对应，故$\overline{X}$也是充分统计量.

46、设$X_1,\cdots,X_n$ $i.i.d.$ $\sim N(\theta,\theta^2),\theta>0$,问$\overline{X}$是否为充分统计量?

解：

$$f(x_i,\theta)=\frac{1}{\sqrt{2\pi}\,\theta}e^{-\frac{(x_i-\theta)^2}{2\theta^2}}$$

$$f(\boldsymbol{x},\theta)=\left(\sqrt{2\pi}\,\theta\right)^{-n}e^{-\frac{\sum\limits_{i=1}^{n}(x_i-\theta)^2}{2\theta^2}}=\left(\sqrt{2\pi}\,\theta\right)^{-n}e^{-\frac{\sum\limits_{i=1}^{n}x_i^2-2\theta\sum\limits_{i=1}^{n}x_i+n\theta^2}{2\theta^2}}$$

$$=\left(\sqrt{2\pi}\,\theta\right)^{-n}e^{-\frac{n}{2}}e^{-\frac{\sum\limits_{i=1}^{n}x_i^2-2\theta\sum\limits_{i=1}^{n}x_i}{2\theta^2}},\quad T(\boldsymbol{X})=\left(\sum_{i=1}^{n}X_i,\sum_{i=1}^{n}x_i^2\right)$$

由因子分解定理可知$T(\boldsymbol{X})$是充分统计量,$T(\boldsymbol{X})$与$\overline{X}$不一一对应，故$\overline{X}$不是充分统计量.

47、设$X_1,\cdots,X_m\ i.i.d.\ \sim N(a,\sigma^2),Y_1,\cdots,Y_n\ i.i.d.\ \sim N(b,\sigma^2)$且两组样本独立.记

$\overline{X}=\sum_{i=1}^{m}X_i\Big/m,\overline{Y}=\sum_{i=1}^{n}Y_i\Big/n$,而

$$S^2=\frac{1}{n+m-2}\left[\sum_{i=1}^{m}\left(X_i-\overline{X}\right)^2+\sum_{j=1}^{n}\left(Y_j-\overline{Y}\right)^2\right]$$

证明$\left(\overline{X},\overline{Y},S^2\right)$为充分完全统计量.

解：

因为两样本相互独立，所以

$$f(x_1,\cdots,x_m,y_1,\cdots,y_n)=\prod_{i=1}^{m}(2\pi\sigma^2)^{-\frac{1}{2}}\exp\left\{-\frac{(x_i-a)^2}{2\sigma^2}\right\}\prod_{k=1}^{n}(2\pi\sigma^2)^{-\frac{1}{2}}\exp\left\{-\frac{(y_k-b)^2}{2\sigma^2}\right\}$$

$$=(2\pi\sigma^2)^{-\frac{m+n}{2}}\exp\left\{-\frac{\sum_{i=1}^{m}(x_i-a)^2+\sum_{k=1}^{n}(y_k-b)^2}{2\sigma^2}\right\}$$

$$=(2\pi\sigma^2)^{-\frac{m+n}{2}}\exp\left\{-\frac{\sum_{i=1}^{m}(x_i^2-2ax_i+a^2)+\sum_{k=1}^{n}(y_k^2-2by_k+b^2)}{2\sigma^2}\right\}$$

$$=(2\pi\sigma^2)^{-\frac{m+n}{2}}\exp\left\{-\frac{\sum_{i=1}^{m}x_i^2+\sum_{k=1}^{n}y_k^2-2am\bar{x}-2bn\bar{y}+ma^2+nb^2}{2\sigma^2}\right\}$$

$=g(t_1(\boldsymbol{x}),t_2(\boldsymbol{x}),t_3(\boldsymbol{x}),a,b,\sigma^2)h(x),t_1(\boldsymbol{x})=\sum_{i=1}^{m}x_i^2+\sum_{k=1}^{n}y_k^2,t_2(\boldsymbol{x})=\bar{x},t_3(\boldsymbol{x})=\bar{y},h(x)\equiv 1,$

故$\left(\sum_{i=1}^{m}X_i^2+\sum_{k=1}^{n}Y_k^2,\overline{X},\overline{Y}\right)$是充分统计量

$f(x_1,\cdots,x_m,y_1,\cdots,y_n)=C^*(\varphi)\exp\{\varphi_1T_1(x)+\varphi_2T_2(x)+\varphi_3T_3(x)\}h(x)$

其中$\varphi_1=-\frac{1}{2\sigma^2}$，$\varphi_2=\frac{am}{\sigma^2},\varphi_3=\frac{bn}{\sigma^2}$.自然参数空间为

$\Theta^*=\{(\varphi_1,\varphi_2,\varphi_3),-\infty<\varphi_1<0,-\infty<\varphi_2<+\infty,-\infty<\varphi_3<+\infty\}$

Θ^*作为R_3的子集显然有内点，故$\left(\sum_{i=1}^{m}X_i^2+\sum_{k=1}^{n}Y_k^2,\overline{X},\overline{Y}\right)$为完全统计量

又因为$T^*(\boldsymbol{X},\boldsymbol{Y})=\left(S^2,\overline{X},\overline{Y}\right)$为$\left(\sum_{i=1}^{m}X_i^2+\sum_{k=1}^{n}Y_k^2,\overline{X},\overline{Y}\right)$的函数,

故$\left(S^2,\overline{X},\overline{Y}\right)$为充分完全统计量.

48、设$X_1,\cdots,X_n$是从下列总体中抽取的样本，其密度函数为

$f(x,\theta)=\dfrac{1}{2\theta}\exp\left\{-\dfrac{|x|}{\theta}\right\},-\infty<x<+\infty,\theta>0,$

证明$T=\sum\limits_{i=1}^{n}|x_i|$是$\theta$的充分完全统计量.

解：

$$f(x_1,\cdots,x_n,\theta)=\prod_{i=1}^{n}f(x_i,\theta)=(2\theta)^{-n}\exp\left\{-\frac{\sum\limits_{i=1}^{n}|x_i|}{\theta}\right\}=g(t(\mathbf{x}),\theta)h(x)$$

$$g(t(\boldsymbol{x}),\theta)=(2\theta)^{-n}\exp\left\{-\frac{\sum\limits_{i=1}^{n}|x_i|}{\theta}\right\},h(x)\equiv1,$$

故由因子分解定理知$T=\sum\limits_{i=1}^{n}|x_i|$是$\theta$的充分统计量.

$$f(x_1,\cdots,x_n,\theta)=(2\theta)^{-n}\exp\left\{-\frac{\sum\limits_{i=1}^{n}|x_i|}{\theta}\right\}=C(\theta)\exp\{\varphi T(\boldsymbol{x})\}h(x)$$

$\varphi=-\dfrac{1}{\theta}$,则$C(\theta)=\left(\dfrac{-2}{\theta}\right)^{-n}$,$h(x)\equiv1$,$\Theta^*=\{\varphi:-\infty<\varphi<0\}$

显然Θ^*作为R_1的子集有内点，故$T=\sum\limits_{i=1}^{n}|x_i|$是$\theta$的完全统计量.

综上所述，$T=\sum\limits_{i=1}^{n}|x_i|$是$\theta$的充分完全统计量.

49、设$X_1,\cdots,X_n$为取自下列指数分布总体的简单样本，

$$f(x,\theta)=\exp\{-(x-\theta)\},\quad x>\theta,$$

其中$-\infty<\theta<+\infty$为位置参数. 证明$T(\boldsymbol{X})=X_{(1)}$为充分完全统计量.

解：

$$f(x_1,\cdots,x_n,\theta)=\exp\left\{-\sum_{i=1}^{n}x_i+n\theta\right\}I_{(x_1,\cdots,x_n>\theta)}=g(t(x_{(1)}),\theta)h(x)$$

其中$g(t(x_{(1)}),\theta)=e^{n\theta}I_{(x_{(1)}>\theta)}$,$h(x)=\exp\left\{-\sum\limits_{i=1}^{n}x_i\right\}$,

故$T(\boldsymbol{X})=X_{(1)}$为充分统计量.

$T(\boldsymbol{X})=X_{(1)}$的密度函数为

$$g(t)=\begin{cases} ne^{-n(t-\theta)}, t>\theta \\ 0, \quad\quad 其它 \end{cases}$$

设$\varphi(t)$为t的任一实函数，满足$E_\theta\varphi(T)=0$,此即

$$\int_\theta^{+\infty}\varphi(t)e^{-n(t-\theta)}dt=0,一切-\infty<\theta<+\infty$$

$$\int_\theta^{+\infty}\varphi(t)e^{-nt}dt=0,一切-\infty<\theta<+\infty$$

对上式两边关于θ求导得，$\varphi(\theta)e^{-n\theta}=0\Leftrightarrow\varphi(\theta)=0$,一切$-\infty<\theta<+\infty$

改θ为t，即$\varphi(t)=0$故$T(\boldsymbol{X})=X_{(1)}$为完全统计量.

综上所述，故$T(\boldsymbol{X})=X_{(1)}$为充分完全统计量.

50、设$X=(X_1,\cdots,X_n)$为从均匀分布$U(-\theta/2,\theta/2),0<\theta<+\infty$中抽取的简单样本，

证明$(X_{(1)},X_{(n)})$为充分统计量，但不是完全的.

解：

$$f(x,\theta)=\frac{1}{\theta^n}I_{\left(-\frac{\theta}{2}<x_{(1)}<\cdots<x_{(n)}<\frac{\theta}{2}\right)}=g(t(x),\theta)h(x)$$

$$g(t(x),\theta)=\frac{1}{\theta^n}I_{\left(-\frac{\theta}{2}<x_{(1)}<\cdots<x_{(n)}<\frac{\theta}{2}\right)},h(x)\equiv1,$$

由因子分解定理知$(X_{(1)},X_{(n)})$为充分统计量.

令$Y_i=\dfrac{X_i}{\theta},i=1,2,\cdots,n$,则$Y_1,Y_2,\cdots,Y_n\ i.i.d.\ U\left(-\dfrac{1}{2},\dfrac{1}{2}\right)$

则$\dfrac{X_{(n)}}{X_{(1)}}=\dfrac{Y_{(n)}}{Y_{(1)}}$的分布与$\theta$无关，容易找的常数$1<a<b<+\infty$

$P(x<a)=P(x>b)>0$

$$构造\varphi(t_1,t_2)=\begin{cases} 1,\ t_2/t_1<a \\ -1,\ t_2/t_1>b=I_{(t_2/t_1<a)}-I_{(t_2/t_1>b)} \\ 0,\ 其它 \end{cases}$$

显然此时$E_\theta\varphi(t_1,t_2)=0$,又$\varphi(t_1,t_2)$，故$(X_{(1)},X_{(n)})$不是完全统计量.

51、设$X_1,\cdots,X_n\ i.i.d.\sim$均匀分布$U(\theta,2\theta),\theta>0$,证明$(X_{(1)},X_{(n)})$是充分统计量，但不是

完全统计量.

解：

与50题同理。

$f(x,\theta)=\frac{1}{\theta^n}I_{(\theta<x_{(1)}<\cdots<x_{(n)}<2\theta)}=g(t(x),\theta)h(x)$

$g(t(x),\theta)=\frac{1}{\theta^n}I_{(\theta<x_{(1)}<\cdots<x_{(n)}<2\theta)},h(x)\equiv 1,$

由因子分解定理知$(X_{(1)},X_{(n)})$为充分统计量.

令$Y_i=\frac{X_i}{\theta},i=1,2,\cdots,n$,则$Y_1,Y_2,\cdots,Y_n\ i.i.d.\ U(1,2)$

则$\frac{X_{(n)}}{X_{(1)}}=\frac{Y_{(n)}}{Y_{(1)}}$的分布与$\theta$无关，容易找的常数$1<a<b<+\infty$

$P(x<a)=P(x>b)>0$

构造$\varphi(t_1,t_2)=\begin{cases}1,\ t_2/t_1<a\\ -1,\ t_2/t_1>b\\ 0,\ 其它\end{cases}=I_{(t_2/t_1<a)}-I_{(t_2/t_1>b)}$

显然此时$E_\theta\varphi(t_1,t_2)=0$,又$\varphi(t_1,t_2)$，故$(X_{(1)},X_{(n)})$不是完全统计量.

52、设$X_1,\cdots,X_n\ i.i.d.\sim$两参数的指数分布，其密度函数为

$$f(x;\lambda;\mu)=\lambda^{-1}\exp\left\{-\frac{x-\mu}{\lambda}\right\}I_{[x>\mu]},$$

其中$-\infty<\mu<+\infty,-\infty<\lambda<+\infty$.设$X_{(1)},\cdots,X_{(n)}$是样本的次序统计量，证明：

（1）$\left(X_{(1)},\sum_{i=1}^{n}X_{(i)}\right)$是$(\lambda,\mu)$的充分统计量.*（2）$X_{(1)}$与$\sum_{i=1}^{n}(X_{(i)}-X_{(1)})$独立.

解：

$$(1)f(\mathbf{x};\lambda;\mu)=\lambda^{-n}\exp\left\{-\frac{\sum_{i=1}^{n}X_i-n\mu}{\lambda}\right\}I_{[X_{(1)}>\mu]}=g\left(t\left(X_{(1)},\sum_{i=1}^{n}X_i\right),\lambda,\mu\right)h(\mathbf{x})$$

$$g\left(t\left(X_{(1)},\sum_{i=1}^{n}X_i\right),\lambda,\mu\right)=\lambda^{-n}\exp\left\{-\frac{\sum_{i=1}^{n}X_i-n\mu}{\lambda}\right\}I_{[X_{(1)}>\mu]},h(x)\equiv 1$$

故$\left(X_{(1)},\sum_{i=1}^{n}X_{(i)}\right)$是$(\lambda,\mu)$的充分统计量.

（2）先将λ固定于λ_0，则可证$X_{(1)}$为μ的充分完全统计量，以下为证明过程

$$f(\mathbf{x};\mu)=\lambda_0{}^{-n}\exp\left\{-\frac{\sum_{i=1}^{n}X_i-n\mu}{\lambda_0}\right\}I_{[X_{(1)}>\mu]}=g\left(t(X_{(1)}),\mu\right)h(\mathbf{x})$$

其中$g\left(t(X_{(1)}),\mu\right)=\lambda_0{}^{-n}\exp\left\{\frac{n\mu}{\lambda_0}\right\}I_{[X_{(1)}>\mu]},h(\mathbf{x})=\exp\left\{-\frac{\sum_{i=1}^{n}X_i}{\lambda_0}\right\}$

故$X_{(1)}$为μ的充分统计量.

又$F(X_{(1)}<t)=1-[1-F(t)]^n$,故$f_{X_{(1)}}(t)=n[1-F(t)]^{n-1}f(t)=\frac{n}{\lambda_0}\exp\left\{-\frac{(n+1)(t-\mu)}{\lambda_0}\right\},(t>\mu)$

设$\varphi(t)$为t的任一实函数，满足$E_\mu(\varphi(T))=0$,即

$$\int_\mu^{+\infty}\varphi(t)\frac{n}{\lambda_0}\exp\left\{-\frac{(n+1)(t-\mu)}{\lambda_0}\right\}dt=0,\text{一切}\mu\in R,$$

$$\int_\mu^{+\infty}\varphi(t)\exp\left\{-\frac{(n+1)t}{\lambda_0}\right\}dt=0,\text{一切}\mu\in R,$$

则两边对μ求导得$\varphi(\mu)\exp\left\{-\frac{(n+1)}{\lambda_0}\mu\right\}=0\Leftrightarrow\varphi(\mu)=0$,一切$\mu\in R$，

即$\varphi(t)=0$，故$X_{(1)}$为μ的完全统计量.

综上所述，$X_{(1)}$为μ的充分完全统计量.

构造$Y=X-\mu$，则$f(y)=\lambda_0{}^{-1}\exp\left\{-\frac{y}{\lambda_0}\right\},y>0,Y_{(i)}-Y_{(1)}=(X_{(i)}-\mu)-(X_{(1)}-\mu)=X_{(i)}-X_{(1)}$,

即$\sum_{i=1}^{n}(X_{(i)}-X_{(1)})$的分布与$\mu$无关，故由$Basu$定理可知，$X_{(1)}$与$\sum_{i=1}^{n}(X_{(i)}-X_{(1)})$独立.

53、设$X_1,\cdots,X_n$为来自$N(a,\sigma^2)$中抽取的$i.i.d.$样本，证明$\overline{X}$与$X_{(n)}-X_{(1)}$独立.

解：

先将σ固定于σ_0，则由例 2.8.3 可知$\overline{X}$为a的完全统计量，由例 2.7.2 可知$\overline{X}$为a的充分统计量. 综上，$\overline{X}$为a的充分完全统计量. 设$Y=X-a$,则$Y\sim N(0,\sigma_0{}^2)$

$X_{(n)}-X_{(1)}=(X_{(n)}-a)-(X_{(1)}-a)=Y_{(n)}-Y_{(1)}$, 即$X_{(n)}-X_{(1)}$的分布与$a$无关，故由

$Basu$定理可知，$\overline{X}$与$X_{(n)}-X_{(1)}$独立.

54、设$\boldsymbol{X}=(X_1,\cdots,X_n)$是从$N(a,\sigma_1{}^2)$中抽取的简单样本，$\boldsymbol{Y}=(Y_1,\cdots,Y_n)$是从$N(b,\sigma_2{}^2)$中抽取的简单样本，且合样本$X_1,\cdots,X_n$；$Y_1,\cdots,Y_n$相互独立，证明

$T(\boldsymbol{X},\boldsymbol{Y})=\left(\overline{X},\overline{Y},Q_1{}^2,Q_2{}^2\right)$与$r(\boldsymbol{X},\boldsymbol{Y})=\sum_{i=1}^{n}\left(X_i-\overline{X}\right)\left(Y_i-\overline{Y}\right)\Big/\sqrt{Q_1{}^2\cdot Q_2{}^2}$独立，此处

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i\ ,\overline{Y}=\frac{1}{n}\sum_{i=1}^{n}Y_i\ ,Q_1{}^2=\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2,Q_2{}^2=\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2.$$

解：

易证明$T(\boldsymbol{X},\boldsymbol{Y})=\left(\overline{X},\overline{Y},Q_1{}^2,Q_2{}^2\right)$为$a,b,\sigma_1{}^2,\sigma_2{}^2$的充分完全统计量，

$\left(X_i-\overline{X}\right)/\sigma_1\sqrt{(n-1)/n}\sim N(0,1),\left(Y_i-\overline{Y}\right)/\sigma_2\sqrt{(n-1)/n}\sim N(0,1).$

$Q_1^2/\sigma_1^2\sim\chi_{n-1}^2,Q_2^2/\sigma_2^2\sim\chi_{n-1}^2.$

$r(\boldsymbol{X},\boldsymbol{Y})=\sum_{i=1}^{n}\left(X_i-\overline{X}\right)\left(Y_i-\overline{Y}\right)/\sqrt{Q_1{}^2\cdot Q_2{}^2}$

$=\sum_{i=1}^{n}\left(\left(X_i-\overline{X}\right)/\sigma_1\sqrt{(n-1)/n}\right)\left(\left(Y_i-\overline{Y}\right)/\sigma_2\sqrt{(n-1)/n}\right)/\sqrt{Q_1^2/(n-1)\sigma_1^2\cdot Q_1^2/(n-1)\sigma_2^2}\ \cdot\frac{1}{n}$

令$T'=\left(\left(X_i-\overline{X}\right)/\sigma_1\sqrt{(n-)1/n}\right)/\sqrt{Q_1^2/(n-1)\sigma_1^2},T''=\left(\left(Y_i-\overline{Y}\right)/\sigma_2\sqrt{(n-1)/n}\right)/\sqrt{Q_1^2/(n-1)\sigma_2^2},$

则$T'\sim t(n-1),T''\sim t(n-1)$,则$r(\boldsymbol{X},\boldsymbol{Y})=\frac{1}{n}\cdot\sum_{i=1}^{n}T_i'\cdot T_i'',$

即$r(\boldsymbol{X},\boldsymbol{Y})$与$a,b,\sigma_1{}^2,\sigma_2{}^2$无关，故由$Basu$定理可知$T(\boldsymbol{X},\boldsymbol{Y})$与$r(\boldsymbol{X},\boldsymbol{Y})$独立.

55、设$X_1,\cdots,X_n\ i.i.d.\sim N(a,\sigma^2),\overline{X}$为样本均值，$\xi=f(x_1,\cdots,x_n)$满足条件

$f(x_1+c,\ \cdots,x_n+c)=f(x_1,\cdots,x_n)$,对任何常数$c$,证明$\xi$与$\overline{X}$独立.

解：

先将σ^2固定于σ_0^2，则$t(x)$为族$\{N(a,\sigma_0^2),-\infty<a<+\infty\}$之下的充分完全统计量.

$f(X_1,\cdots,X_n)=f(X_1-a,\ \cdots,X_n-a)$,而$X_i-a\sim N(0,\sigma_0^2)$与$a$无关，故$f(X_1,\cdots,X_n)$的分布与$a$无关，有$Basu$定理可知，$\overline{X}$与$\xi$独立.

习题 3 点估计

1、设$X_1,\cdots,X_n\ i.i.d.\sim N(a,\sigma^2)$,样本方差$S^2=\sum\limits_{i=1}^{n}\left(X_i-\overline{X}\right)^2/(n-1)$是$\sigma^2$的无偏估计.

证明S^2是σ^2的强相合估计和均方相合估计.

解：

作正交变换$\boldsymbol{Y}=\boldsymbol{A}\boldsymbol{X}$,其中$\boldsymbol{Y}=(Y_1,\cdots,Y_n)^T$,$\boldsymbol{X}=(X_1,\cdots,X_n)^T$,$\boldsymbol{A}=\begin{pmatrix}\frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \cdots & \frac{1}{\sqrt{n}}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{pmatrix}$,

则$Y_1=\frac{1}{\sqrt{n}}\sum\limits_{i=1}^{n}X_i=\sqrt{n}\,\overline{X}$，$Y_1^2+\cdots+Y_n^2=X_1^2+\cdots+X_n^2$，

则$S^2=\dfrac{\sum\limits_{i=1}^{n}\left(X_i-\overline{X}\right)^2}{n-1}=\dfrac{\sum\limits_{i=1}^{n}X_i^2-n\overline{X}^2}{n-1}=\dfrac{\sum\limits_{i=2}^{n}Y_i^2}{n-1}$

又$\because Y_1,\cdots,Y_n$相互独立，$Y_i\sim N(\mu_i,\sigma^2),i=2,\cdots,n$，且$\boldsymbol{A}$的行向量正交，

$\therefore \mu_i=a\sum\limits_{k=1}^{n}a_{ik}=a\sqrt{n}\sum\limits_{k=1}^{n}\frac{1}{\sqrt{n}}\cdot a_{ik}=0$,即$Y_2,\cdots,Y_n i.i.d.\sim N(0,\sigma^2)$,

$S^2=\dfrac{\sum\limits_{i=2}^{n}Y_i^2}{n-1}$，其中$Y_i^2$独立同分布,$E(Y_i^2)=\sigma^2$，则由柯尔莫哥洛夫强大数定律可知

$P\left(\lim\limits_{n\to\infty}S^2=\sigma^2\right)=1$,即$S^2$为$\sigma^2$的强相合估计.

$E_\sigma(S^2-\sigma^2)=0,E_\sigma[(S^2-\sigma^2)^2]=D(S^2-\sigma^2)=DS^2$,

又$\because \dfrac{S^2(n-1)}{\sigma^2}\sim\chi^2(n-1),\therefore \dfrac{(n-1)^2}{\sigma^4}DS^2=2(n-1),\therefore DS^2=\dfrac{2\sigma^4}{n-1}$

故$\lim\limits_{n\to\infty}E_\sigma|S^2-\sigma^2|^2=\lim\limits_{n\to\infty}\dfrac{2\sigma^4}{n-1}=0$,即$S^2$为$\sigma^2$的均方相合估计.

2、证明$\hat\theta_n$是θ的均方相合估计的充要条件为：$\hat\theta_n$是渐进无偏的且$\lim\limits_{n\to\infty}D_\theta\left(\hat\theta_n\right)=0$，对任何

$\theta\in\Theta$.

解：

充分性：$\lim\limits_{n\to\infty}E_\theta\left|\hat\theta_n-\theta\right|^2=\lim\limits_{n\to\infty}\left[E_\theta^2\left(\hat\theta_n-\theta\right)+D_\theta\left(\hat\theta_n-\theta\right)\right]$

$\because\hat\theta_n$是渐进无偏的，$\therefore\lim\limits_{n\to\infty}E\left(\hat\theta_n\right)=\theta,\lim\limits_{n\to\infty}\left[E_\theta^2\left(\hat\theta_n-\theta\right)\right]=0$

又$\because\lim\limits_{n\to\infty}D_\theta\left(\hat\theta_n\right)=0$，$\therefore\lim\limits_{n\to\infty}D_\theta\left(\hat\theta_n-\theta\right)=0$

$\therefore\lim\limits_{n\to\infty}E_\theta\left|\hat\theta_n-\theta\right|^2=0$，即$\hat\theta_n$是$\theta$的均方相合估计，充分性得证.

必要性：$\lim\limits_{n\to\infty} E_\theta\left|\hat{\theta}_n-\theta\right|^2=\lim\limits_{n\to\infty}\left[E_\theta^2\left(\hat{\theta}_n-\theta\right)+D_\theta\left(\hat{\theta}_n-\theta\right)\right]=0$

且$E_\theta^2\left(\hat{\theta}_n-\theta\right)\geqslant 0$，$D_\theta\left(\hat{\theta}_n-\theta\right)\geqslant 0$，要使$\lim\limits_{n\to\infty} E_\theta\left|\hat{\theta}_n-\theta\right|^2=0$，

需$\lim\limits_{n\to\infty}\left[E_\theta^2\left(\hat{\theta}_n-\theta\right)\right]=0$，$\lim\limits_{n\to\infty} D_\theta\left(\hat{\theta}_n-\theta\right)=0$，

也即$\hat{\theta}_n$是渐进无偏的，$\lim\limits_{n\to\infty} D_\theta\left(\hat{\theta}_n\right)=0$，故必要性得证.

3、设$\boldsymbol{X}=(X_1,\cdots,X_n)$是 从某总体中抽取的简单样本，满足$E(X_1)=\mu<\infty, E(X_1^2)<\infty$，

证明$T(\boldsymbol{X})=2\sum\limits_{i=1}^{n} iX_i\Big/[n(n+1)]$是$\mu$的弱相合估计.

解：

$$E(T(\boldsymbol{X}))=\frac{2}{n(n+1)}\cdot E\left(\sum_{i=1}^{n} iX_i\right)=\frac{2E(X_i)}{n(n+1)}\cdot\sum_{i=1}^{n} i=\frac{2\mu}{n(n+1)}\cdot\frac{n(n+1)}{2}=\mu$$

$$D(T(\boldsymbol{X}))=D\left(\frac{2}{n(n+1)}\cdot\sum_{i=1}^{n} iX_i\right)=\frac{4DX_i}{n^2(n+1)^2}\cdot\sum_{i=1}^{n} i^2=\frac{4DX_i}{n^2(n+1)^2}\cdot\frac{n(n+1)(2n+1)}{6}$$

又因为$E(X_1)=\mu<\infty, E(X_1^2)<\infty$，故可设$DX_i=\sigma^2<\infty$，则$D(T(\boldsymbol{X}))=\dfrac{2(2n+1)\sigma^2}{3n(n+1)}$

则$\lim\limits_{n\to\infty} P(|T(\boldsymbol{X})-\mu|\geqslant\varepsilon)\leqslant\lim\limits_{n\to\infty}\dfrac{D(T(\boldsymbol{X}))}{\varepsilon^2}=\lim\limits_{n\to\infty}\dfrac{2(2n+1)\sigma^2}{3n(n+1)\varepsilon^2}=0$,

故$\lim\limits_{n\to\infty} P(|T(\boldsymbol{X})-\mu|\geqslant\varepsilon)=0$，$T(\boldsymbol{X})=2\sum\limits_{i=1}^{n} iX_i\Big/[n(n+1)]$是$\mu$的弱相合估计.

4、设$X_n\xrightarrow{a.s.}X, g(x)$为$x$的连续函数，则$g(X_n)\xrightarrow{a.s.}g(X)$.

解:

在$g(x)$的定义域上任取一点x,由$g(x)$在x处连续知，对于$\forall\varepsilon>0, \exists\delta>0$使得

$\{|g(y)-g(x)|<\varepsilon\}\supset\{|y-x|<\delta\}$

故$P(|g(X_n)-g(X)|<\varepsilon)\geqslant P\{|X_n-X|<\delta\}$

又由X_n是X的强相合估计知，存在正整数N,当$n>N$时，

$P\{|X_n-X|<\delta\}=1$，因此

$1\geqslant P(|g(X_n)-g(X)|<\varepsilon)\geqslant P\{|X_n-X|<\delta\}=1$，

$P(|g(X_n)-g(X)|<\varepsilon)=1$，

即$P\left(\lim\limits_{n\to\infty} g(X_n)=g(X)\right)=1$,证毕.

5、设总体X服从均匀分布$U(0,2\theta)$,令$X_1,\cdots,X_n$为从总体X中抽取的简单样本，

（1）证明$\hat{\theta}_1=\overline{X}$和$\hat{\theta}_2=(n+1)X_{(n)}/2n$为$\theta$的无偏估计.

（2）证明$\hat{\theta}_1$为θ的强相合估计，$\theta_2^*=X_{(n)}/2$为θ的弱相合估计.

（3）求$\hat{\theta}_1$和$\hat{\theta}_2$的方差，问哪一个更有效?

解：

$$(1)EX=\theta,DX=\frac{\theta^2}{3},故E\hat{\theta}_1=E\overline{X}=\theta$$

$$F(t)=P(X_{(n)}\leqslant t)=\prod_{i=1}^{n}P(X_i\leqslant t)=\left(\frac{t}{2\theta}\right)^n$$

$$f(t)=\frac{nt^{n-1}}{(2\theta)^n},故EX_{(n)}=\int_0^{2\theta}\frac{nt^n}{(2\theta)^n}dt=\frac{2\theta n}{n+1}$$

故$E\hat{\theta}_2=E[(n+1)X_{(n)}/2n]=\theta$,即$\hat{\theta}_1$和$\hat{\theta}_2$都是$\theta$的无偏估计.

(2)$\hat{\theta}_1=\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$,又因为$EX=\theta$,故$kolmogorov$由强大数定律可知

$P_\theta\left(\lim_{n\to\infty}\hat{\theta}_1=\theta\right)=1$,即$\hat{\theta}_1$是$\theta$的强相合估计.

$$P\left(\left|\hat{\theta}_2^*-\theta\right|\geqslant\varepsilon\right)=P(|X_{(n)}/2-\theta|\geqslant\varepsilon)=P(X_{(n)}\geqslant 2\varepsilon+2\theta\cup X_{(n)}\leqslant 2\theta-2\varepsilon)$$

$$=P(0<X_{(n)}\leqslant 2\theta-2\varepsilon)=\int_0^{2\theta-2\varepsilon}\frac{nt^{n-1}}{(2\theta)^n}dt=\left(1-\frac{\varepsilon}{\theta}\right)^n,$$

则$\lim_{n\to\infty}P\left(\left|\hat{\theta}_2^*-\theta\right|\geqslant\varepsilon\right)=\lim_{n\to\infty}\left(1-\frac{\varepsilon}{\theta}\right)^n=0$，即$\hat{\theta}_2^*$是$\theta$的弱相合估计.

$$(3)D\left(\hat{\theta}_1\right)=D\left(\overline{X}\right)=\frac{\theta^2}{3n},D\left(\hat{\theta}_2\right)=\frac{(n+1)^2}{4n^2}D(X_{(n)})$$

$$E(X_{(n)}^2)=\int_0^{2\theta}\frac{nt^{n+1}}{(2\theta)^n}=\frac{4n}{n+2}\theta^2,D\left(\hat{\theta}_2\right)=E\left(\hat{\theta}_2^2\right)-E^2\left(\hat{\theta}_2\right)=\frac{\theta^2}{n(n+2)}<D\left(\hat{\theta}_1\right)$$

故$\hat{\theta}_2$更有效.

6、设总体X服从二项分布$b(k,p)$,k是正整数，$0<p<1$,两者都是未知参数，$X_1,\cdots,X_n$

为从中抽取的简单样本，求k和p的矩估计.

解：

$$EX=kp,DX=kp(1-p),\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i,S_n^2=\frac{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}{n}$$

则$\hat{p}=1-\frac{S_n^2}{\overline{X}},\hat{k}=\left[\frac{\overline{X}^2}{\overline{X}-S_n^2}\right]$,“[ ]”为取整函数.

7、设$X_1,\cdots,X_n$为取自对数级数分布

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

的随机样本，求参数p的矩估计量.

解：

$$EX=\sum_{k=1}^{\infty}-\frac{p^k}{\ln(1-p)}=-\frac{1}{\ln(1-p)}\cdot\frac{p}{1-p}$$

$$E(X^2)=\sum_{k=1}^{\infty}-\frac{kp^k}{\ln(1-p)}=-\frac{1}{\ln(1-p)}\cdot\frac{p}{(1-p)^2}$$

则p的矩估计$\hat{p}=1-\dfrac{\sum_{i=1}^{n}X_i}{\sum_{i=1}^{n}X_i^2}$.

8、设样本$X_1,\cdots,X_n$抽自正态分布$N(0,\sigma^2)$,求σ的矩估计量:

（1）利用$E|X_1|=\sqrt{\dfrac{2}{\pi}}\sigma$;

（2）利用$\sigma=\sqrt{D(X_1)}$.

解：

(1)$E|X_1|=2\displaystyle\int_0^{+\infty}\frac{x}{\sqrt{2\pi}\sigma}e^{-\frac{x^2}{2\sigma^2}}dx=\sqrt{\frac{2}{\pi}}\sigma.E|\overline{X}|=E|X_1|=\sqrt{\frac{2}{\pi}}\sigma.$

故$\hat{\sigma}=\sqrt{\dfrac{\pi}{2}}\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}|X_i|$

(2)$\sigma=\sqrt{D(X_1)}$，故$\hat{\sigma}=\sqrt{\displaystyle\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2\Big/n}$

9、设$X_1,\cdots,X_n$是来自正态总体$N(a,\sigma^2)$,求$P(X>1)$的矩估计量.

解：

$$P(X>1)=P\left(\frac{X-a}{\sigma}>\frac{1-a}{\sigma}\right)=1-\Phi\left(\frac{1-a}{\sigma}\right)=1-\Phi\left(\frac{1-EX}{\sqrt{DX}}\right)$$

则$\hat{p}=1-\Phi\left(\dfrac{1-\overline{X}}{S_n^2}\right)$,其中$\overline{X}=\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}X_i$，$S_n^2=\displaystyle\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2\Big/n.$

10、设$X_1,\cdots,X_n$ $i.i.d.\sim$伽马分布$\Gamma(r,\lambda)$,其中r已知，求λ矩估计量，并讨论它的无偏性.

解：

$$EX=\int_0^{+\infty}x\cdot\frac{\lambda^r}{\Gamma(r)}x^{r-1}e^{-\lambda x}dx=\frac{\Gamma(r+1)}{\lambda\Gamma(r)}\int_0^{+\infty}\frac{\lambda^{r+1}}{\Gamma(r+1)}x^{r+1-1}e^{-\lambda x}dx=\frac{r}{\lambda}.$$

则λ的矩估计量$\hat{\lambda}=\dfrac{r}{\overline{X}}$,其中$\overline{X}=\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}X_i$,

$$t=\sum_{i=1}^{n}X_i\sim\Gamma(nr,\lambda),\text{则}E\hat{\lambda}=E\left(\frac{nr}{t}\right)=nr\int_0^{+\infty}\frac{1}{t}\cdot\frac{\lambda^{nr}}{\Gamma(nr)}t^{nr-1}e^{-\lambda t}dt=$$

$$\frac{\lambda nr\Gamma(nr-1)}{\Gamma(nr)}\int_0^{+\infty}\frac{\lambda^{nr-1}}{\Gamma(nr-1)}t^{(nr-1)-1}e^{-\lambda t}dt=\frac{\lambda nr}{nr-1}\neq\lambda,$$

故$\hat{\lambda}$不是λ的无偏估计，$\lim\limits_{n\to\infty} E\hat{\lambda}=\lambda$,故$\hat{\lambda}$是$\lambda$的渐进无偏估计.

11、若$X=e^{\xi}$,而$\xi\sim N(a,\sigma^2)$,则$r.v.\ X$的分布成为对数正态分布.求出X的密度函数.

设$X_1,\cdots,X_n$是从总体X抽取的简单随机样本，求a和σ^2的矩估计和极大似然估计.

解：

$$f_{\xi}(\xi)=\frac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{(\xi-a)^2}{2\sigma^2}},\xi=\ln x\text{故}f_X(x)=f_{\xi}(\ln x)\left|\frac{d\ln x}{dx}\right|=\frac{1}{\sqrt{2\pi}\,\sigma x}e^{-\frac{(\ln x-a)^2}{2\sigma^2}}$$

$$\text{故}EX=\int_0^{+\infty}\frac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{(\ln x-a)^2}{2\sigma^2}}dx\xlongequal{\ln x=t}\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}\,\sigma}e^{t}e^{-\frac{(t-a)^2}{2\sigma^2}}dt$$

$$=e^{2a+\sigma^2/2}\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{(t-a-\sigma^2)^2}{2\sigma^2}}dt=e^{a+\frac{\sigma^2}{2}}$$

同理可得$EX^2=e^{2\sigma^2+2a}$,因此$DX=EX^2-(EX)^2=e^{2a+\sigma^2}(e^{\sigma^2}-1)$

$$\text{因此}e^{\sigma^2}-1=\frac{DX}{(EX)^2}\Rightarrow\sigma^2=\ln\left[1+\frac{DX}{(EX)^2}\right],a+\frac{\sigma^2}{2}=\ln(EX)$$

$$\text{因此}a=\ln(EX)-\frac{1}{2}\ln\left[1+\frac{DX}{(EX)^2}\right]$$

$$\text{故矩估计}\hat{a}=\ln\overline{X}-\frac{1}{2}\ln\left[1+\frac{S_n^2}{\overline{X}^2}\right],\hat{\sigma}^2=\ln\left[1+\frac{S_n^2}{\overline{X}^2}\right],\text{其中}\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i,S_n^2=\frac{\sum\limits_{i=1}^{n}\left(X_i-\overline{X}\right)^2}{n}$$

$$L(\boldsymbol{x};a,\sigma^2)=(2\pi\sigma^2)^{-\frac{n}{2}}\prod_{i=1}^{n}\frac{1}{x_i}\exp\left\{-\frac{\sum\limits_{i=1}^{n}(\ln x_i-a)^2}{2\sigma^2}\right\}$$

$$\ln L(\boldsymbol{x};a,\sigma^2)=-\frac{n}{2}\ln(2\pi\sigma^2)-\sum_{i=1}^{n}\ln x_i--\frac{\sum\limits_{i=1}^{n}(\ln x_i-a)^2}{2\sigma^2}$$

$$\frac{\partial\ln L}{\partial a}=\frac{1}{\sigma^2}\sum_{i=1}^{n}(\ln x_i-a)=0,\hat{a}^*=\frac{1}{n}\sum_{i=1}^{n}\ln X_i$$

$$\frac{\partial\ln L}{\partial\sigma^2}=-\frac{n}{2\sigma^2}+\frac{1}{2\sigma^4}\sum_{i=1}^{n}(\ln x_i-a)^2,\hat{\sigma}^{*2}=\frac{1}{n}\sum_{i=1}^{n}(\ln X_i-\hat{a}^*)^2$$

12、设$X_1,\cdots,X_n$是从总体X中抽取的简单样本，X的分布为下列之一，试分别用矩估计法和极大似然法求θ的估计量：

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

$$(2)f(x,\theta)=(\theta+1)x^{\theta},\quad 0<x<1,\theta>-1;$$

$$(3)f(x,\theta)=\begin{cases}\frac{\theta^{-r}}{\Gamma(r)}x^{r-1}\exp\left\{-\frac{x}{\theta}\right\}, & x>0,\\ 0, & \text{其他，}\end{cases}\quad r>0\text{已知.}$$

解：

$$(1)EX=\int_0^1\theta(\theta+1)x^\theta(1-x)dx=\frac{\theta}{\theta+2},\text{矩估计量}\hat{\theta}=\frac{2\overline{X}}{1-\overline{X}}$$

$$L(\theta)=\prod_{i=1}^n f(x_i,\theta)=\theta^n(\theta+1)^n\left(\prod_{i=1}^n x_i\right)^{\theta-1}\prod_{i=1}^n(1-x_i)$$

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

$$\frac{\partial\ln L(\theta)}{\partial\theta}=\frac{n}{\theta}+\frac{n}{\theta+1}+\sum_{i=1}^n\ln x_i=0,$$

$$\hat{\theta}^*=\frac{-\left(2+\frac{1}{n}\sum_{i=1}^n\ln X_i\right)+\sqrt{4+\left(\frac{1}{n}\sum_{i=1}^n\ln X_i\right)^2}}{\frac{2}{n}\sum_{i=1}^n\ln X_i}\text{(另一个解不符合}\theta>0\text{,舍去)}$$

$$(2)EX=\int_0^1(\theta+1)X^{\theta+1}dx=\frac{\theta+1}{\theta+2},\hat{\theta}=\frac{2\overline{X}-1}{1-\overline{X}}$$

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

$$\frac{\partial\ln L(\boldsymbol{x},\theta)}{\partial\theta}=\frac{n}{\theta+1}+\sum_{i=1}^n\ln x_i=0,\hat{\theta}^*=-\frac{n}{\sum_{i=1}^n\ln X_i}-1$$

$$(3)EX=\int_0^{+\infty}\frac{\theta^{-r}}{\Gamma(r)}x^r\exp\left\{-\frac{x}{\theta}\right\}dx=r\theta\int_0^{+\infty}\frac{\theta^{-(r+1)}}{\Gamma(r+1)}x^{r+1-1}\exp\left\{-\frac{x}{\theta}\right\}dx=r\theta$$

$$\text{矩估计量}\hat{\theta}=\frac{\overline{X}}{r}$$

$$L(\boldsymbol{x},\theta)=\left(\frac{1}{\Gamma(r)}\right)^n\theta^{-nr}\prod_{i=1}^n x_i^{r-1}\exp\left\{-\frac{\sum_{i=1}^n x_i}{\theta}\right\}$$

$$\ln L(\boldsymbol{x},\theta)=-n\ln[\Gamma(r)]-nr\ln\theta-(r-1)\sum_{i=1}^n\ln x_i--\frac{\sum_{i=1}^n x_i}{\theta}$$

$$\frac{\partial\ln L(\boldsymbol{x},\theta)}{\partial\theta}=-\frac{nr}{\theta}+\frac{1}{\theta^2}\sum_{i=1}^n x_i=0,\hat{\theta}^*=\frac{1}{nr}\sum_{i=1}^n X_i=\frac{\overline{X}}{r}.$$

13、设总体X的密度为$\frac{1}{2\sigma}\exp\{-|x-a|/\sigma\}$,其中$\sigma>0$和$-\infty<a<+\infty$为未知参数.设$X_1,\cdots,X_n$为抽自此总体的简单随机样本，求$a$和$\sigma$的矩估计和极大似然估计.

解：

$$EX=\int_0^{+\infty}\frac{x}{2\sigma}\exp\{-|x-a|/\sigma\}dx \xlongequal{t=x-a} \int_{-\infty}^{+\infty}\frac{t+a}{2\sigma}\exp\{-|t|/\sigma\}dt$$

$$=\int_{-\infty}^{+\infty}\frac{a}{2\sigma}\exp\{-|t|/\sigma\}dt=a,\text{故}\hat{a}=\overline{X}$$

$$EX^2=\int_0^{+\infty}\frac{x^2}{2\sigma}\exp\{-|x-a|/\sigma\}dx \xlongequal{t=x-a} \int_{-\infty}^{+\infty}\frac{(t+a)^2}{2\sigma}\exp\{-|t|/\sigma\}dt$$

$$=\int_{-\infty}^{+\infty}\frac{t^2+a^2}{2\sigma}\exp\{-|t|/\sigma\}dt=2\sigma^2+a^2$$

$$\text{则}\sigma=\sqrt{\frac{EX^2-(EX)^2}{2}},\text{则}\hat{\sigma}^2=\sqrt{\frac{\frac{1}{n}\sum_{i=1}^{n}X_i-\overline{X}^2}{2}}$$

$$L(a,\sigma)=\prod_{i=1}^{n}\frac{1}{2\sigma}\exp\{-|x_i-a|/\sigma\}=\left(\frac{1}{2\sigma}\right)^n\exp\left\{-\frac{\sum_{i=1}^{n}|x_i-a|}{\sigma}\right\}$$

$$\ln L(a,\sigma)=-n\ln(2\sigma)-\frac{\sum_{i=1}^{n}|x_i-a|}{\sigma}$$

为使$\ln L(a,\sigma)$尽可能大，应使$\sum_{i=1}^{n}|x_i-a|$尽可能小，则$\hat{a}^*=m_{0.5}(X_1,\cdots,X_n$中位数)

$$\frac{\partial\ln L(a,\sigma)}{\partial\sigma}=-\frac{n}{\sigma}+\frac{1}{\sigma^2}\sum_{i=1}^{n}|x_i-a|=0,\hat{\sigma}^*=\sum_{i=1}^{n}|x_i-\hat{a}^*|\Big/n$$

14、设随机变量X的密度函数为

$$f(x,\theta)=\begin{cases}\frac{x}{\theta^2}\exp\left\{-\frac{x^2}{2\theta^2}\right\}, & x>0,\\ 0, & x\leqslant 0,\end{cases}\quad\theta>0.$$

设$X_1,\cdots,X_n$为从总体X中抽取的随机样本，求θ的MLE.

解：

$$L(\boldsymbol{x},\theta)=\frac{\prod_{i=1}^{n}x_i}{\theta^{2n}}e^{-\frac{\sum_{i=1}^{n}x_i^2}{2\theta^2}},\ln L(\boldsymbol{x},\theta)=\sum_{i=1}^{n}\ln x_i-2n\ln\theta-\frac{1}{2\theta^2}\sum_{i=1}^{n}x_i^2$$

$$\frac{\partial\ln L(\boldsymbol{x},\theta)}{\partial\theta}=-\frac{2n}{\theta}+\frac{1}{\theta^3}\sum_{i=1}^{n}x_i^2=0,\hat{\theta}=\sqrt{\frac{\sum_{i=1}^{n}X_i^2}{2n}}.$$

15、设$X_1,\cdots,X_n$是来自双参数指数分布

$$f(x,\mu,\sigma)=\frac{1}{\sigma}\exp\left\{-\frac{x-\mu}{\sigma}\right\},x\geqslant\mu$$

的简单随机样本，其中$-\infty<\mu<+\infty,\sigma>0$.记$\theta=(\mu,\sigma)$,求$\mu$，$\sigma$和$P_\theta(X_1\geqslant t)\ (t>\mu)$的矩估计和极大似然估计.

解：

$$EX=\int_{\mu}^{+\infty}\frac{x}{\sigma}\exp\left\{-\frac{x-\mu}{\sigma}\right\}dx\xlongequal{\frac{x-\mu}{\sigma}=t}\int_0^{+\infty}(\mu+\sigma t)e^{-t}dt=\mu+\sigma$$

$$EX^2=\int_{\mu}^{+\infty}\frac{x^2}{\sigma}\exp\left\{-\frac{x-\mu}{\sigma}\right\}dx\xlongequal{\frac{x-\mu}{\sigma}=t}\int_0^{+\infty}(\mu^2+2\mu\sigma t+\sigma^2t^2)e^{-t}dt=\mu^2+2\sigma\mu+2\sigma^2$$

$$\hat{\sigma}=\sqrt{EX^2-(EX)^2}=\sqrt{S_n^2},\hat{\mu}=\overline{X}-\sqrt{S_n^2}$$

$$P(X_1>t)=\int_t^{+\infty}\frac{1}{\sigma}\exp\left\{-\frac{x-\mu}{\sigma}\right\}dt=e^{\frac{\mu-t}{\sigma}},P(X_1>t)\text{的矩估计量为}e^{\frac{\hat{\mu}-t}{\hat{\sigma}}}$$

$$L(\mu,\sigma)=\frac{1}{\sigma^n}\exp\left\{-\frac{\sum_{i=1}^n x_i-n\mu}{\sigma}\right\},\ln L(\mu,\sigma)=-n\ln\sigma--\frac{\sum_{i=1}^n x_i-n\mu}{\sigma}(\mu\leqslant x)$$

$$\frac{\partial\ln L(\mu,\sigma)}{\partial\mu}=\frac{n}{\sigma}>0,\frac{\partial\ln L(\mu,\sigma)}{\partial\sigma}=-\frac{n}{\sigma}+\frac{1}{\sigma^2}\sum_{i=1}^n x_i-\frac{n}{\sigma^2}\mu=0$$

$$\hat{\mu}^*=X_{(1)},\hat{\sigma}^*=\frac{1}{n}\sum_{i=1}^n X_i-X_{(1)},P(X_1>t)\text{的极大似然估计量为}e^{\frac{\hat{\mu}^*-t}{\hat{\sigma}^*}}.$$

16、设总体X服从$Weibull$分布，密度函数为

$$f(x,\lambda)=\begin{cases}\lambda\alpha\cdot x^{\alpha-1}e^{-\lambda x^\alpha}, & x>0\\ 0, & \text{其他，}\end{cases}\quad\lambda>0,\alpha>0.$$

设$X_1,\cdots,X_n$为此总体中抽取的简单样本.若α已知，求λ的MLE.

解：

$$L(\boldsymbol{x},\lambda)=\lambda^n\alpha^n\prod_{i=1}^n x_i^{\alpha-1}e^{-\lambda\sum_{i=1}^n x_i^\alpha}$$

$$\ln L(\boldsymbol{x},\lambda)=n\ln\lambda+n\ln\alpha+(\alpha-1)\sum_{i=1}^n\ln x_i-\lambda\sum_{i=1}^n x_i^\alpha$$

$$\frac{\partial\ln L(\boldsymbol{x},\lambda)}{\partial\lambda}=\frac{n}{\lambda}-\sum_{i=1}^n x_i^\alpha=0,\text{得}\hat{\lambda}=n\Big/\sum_{i=1}^n x_i^\alpha$$

17、设$X_1,\cdots,X_n$为从总体X中抽取的随机样本，X服从均匀分布$U(\theta-1/2,\theta+1/2)$,

求θ的MLE.

解：

$f(x,\theta)=1,\theta-1/2<x<\theta+1/2,$

$L(x,\theta)=I_{[\theta-1/2<x_{(1)}<X_{(n)}<\theta+1/2]}=I_{[x_{(n)}-1/2<\theta<x_{(1)}+1/2]}$

则$[x_{(n)}-1/2<\theta<x_{(1)}+1/2]$中任意值都是$\theta$的$MLE$.

18、设$X_1,\cdots,X_n$是来自均匀分布$U(0,2\theta)$的简单随机样本，其中$0<\theta<+\infty$,求θ的MLE,它是θ的无偏估计吗？如果不是，试对它略作修改，得到θ的一个无偏估计.

解：

$f(x,\theta)=\dfrac{1}{\theta},L(x,\theta)=\dfrac{1}{\theta^n}I_{[\theta<x_{(1)}<x_{(n)}<2\theta]}$,则$\hat{\theta}=\dfrac{X_{(n)}}{2}$

记$X_{(n)}=t$,则$f(t)=\dfrac{n(t-\theta)^{n-1}}{\theta^n}$,$E(X_{(n)})=\displaystyle\int_\theta^{2\theta}\frac{nt(t-\theta)^{n-1}}{\theta^n}dt=\frac{2n+1}{n+1}\theta$

则$E\hat{\theta}=\dfrac{2n+1}{2n+2}\theta\neq\theta$,故$MLE$不是无偏估计，经过修正的$\hat{\theta}^*=\dfrac{2n+2}{2n+1}\hat{\theta}$是无偏估计.

19、设$X_1,\cdots,X_m$和$Y_1,\cdots,Y_n$分别来自总体$N(\mu_1,\sigma^2)$和$N(\mu_2,\sigma^2)$的两组独立样本，

求μ_1，μ_2和σ^2的MLE.

解：

$$L(\mu_1,\mu_2,\sigma^2)=(2\pi)^{-\frac{m+n}{2}}\sigma^{-(m+n)}\exp\left\{-\frac{\sum_{i=1}^m(x_i-\mu_1)^2}{2\sigma^2}-\frac{\sum_{j=1}^n(y_j-\mu_2)^2}{2\sigma^2}\right\}$$

$$\ln L(\mu_1,\mu_2,\sigma^2)=-\frac{m+n}{2}\ln(2\pi)-\frac{m+n}{2}\ln\sigma^2-\frac{\sum_{i=1}^m(x_i-\mu_1)^2}{2\sigma^2}-\frac{\sum_{j=1}^n(y_j-\mu_2)^2}{2\sigma^2}$$

$$\frac{\partial\ln L(\mu_1,\mu_2,\sigma^2)}{\partial\mu_1}=\frac{\sum_{i=1}^m(x_i-\mu_1)}{\sigma^2}=0,\hat{\mu}_1=\frac{1}{m}\sum_{i=1}^m X_i$$

$$\frac{\partial\ln L(\mu_1,\mu_2,\sigma^2)}{\partial\mu_2}=\frac{\sum_{j=1}^n(y_j-\mu_1)}{\sigma^2}=0,\hat{\mu}_2=\frac{1}{n}\sum_{j=1}^n Y_j$$

$$\frac{\partial\ln L(\mu_1,\mu_2,\sigma^2)}{\partial\sigma^2}=-\frac{m+n}{2\sigma^2}+\frac{\sum_{i=1}^m(x_i-\mu_1)^2+\sum_{j=1}^n(y_j-\mu_2)^2}{2\sigma^4}=0,$$

$$\hat{\sigma}^2=\frac{1}{m+n}\left[\sum_{i=1}^m\left(X_i-\overline{X}\right)^2+\sum_{j=1}^n\left(Y_j-\overline{Y}\right)^2\right]$$

20、为了估计湖中有多少条鱼，从中捞出1000条，标上记号后放回湖中，然后再捞出150条鱼，发现其中有10条鱼有记号.问湖中有多少条鱼，才能使150条鱼中出现10条带标记号的鱼的概率最大？

解：

设湖中有M条鱼，则每条鱼被标记的概率为$\dfrac{1000}{M}$，设$p=\dfrac{1000}{M}$.

设$X=\begin{cases}1,\text{带标记}\\0,\text{不带标记}\end{cases}$,则$X\sim b(1,p)$,

则题意转化为从$X\sim b(1,p)$分布中取容量150的样本，$X_1,\cdots,X_{150}.\sum\limits_{i=1}^{150}x_i=10$.

$$L(\boldsymbol{x},p)=\prod_{i=1}^{n}P(X_i=x_i)=p^{\sum\limits_{i=1}^{n}x_i}(1-p)^{n-\sum\limits_{i=1}^{n}x_i}$$

$$\ln L(\boldsymbol{x},p)=\sum_{i=1}^{n}x_i\ln p+\left(n-\sum_{i=1}^{n}x_i\right)\ln(1-p)$$

$$\frac{\partial\ln L(\boldsymbol{x},p)}{\partial p}=\frac{\sum\limits_{i=1}^{n}x_i}{p}-\frac{n-\sum\limits_{i=1}^{n}x_i}{1-p}=0,\text{解得}\hat{p}=\overline{X},\text{故}\hat{M}=\frac{1000}{\overline{X}}=\frac{1000}{\frac{10}{150}}=15000.$$

21、一个罐子中装有黑白两种球，今有放回地抽取一个大小为n的样本，其中有k个白球，求罐中黑白球之比的极大似然估计.

解：

设罐中白球所占的比例为p,

则设$X=\begin{cases}1,\text{白球}\\0,\text{黑球}\end{cases}$,则$X\sim b(1,p)$

$$L(\boldsymbol{x},p)=\prod_{i=1}^{n}P(X_i=x_i)=p^{\sum\limits_{i=1}^{n}x_i}(1-p)^{n-\sum\limits_{i=1}^{n}x_i}$$

$$\ln L(\boldsymbol{x},p)=\sum_{i=1}^{n}x_i\ln p+\left(n-\sum_{i=1}^{n}x_i\right)\ln(1-p)$$

$$\frac{\partial\ln L(\boldsymbol{x},p)}{\partial p}=\frac{\sum\limits_{i=1}^{n}x_i}{p}-\frac{n-\sum\limits_{i=1}^{n}x_i}{1-p}=0,\text{解得}\hat{p}=\overline{X}=\frac{k}{n}$$

而黑白球之比极大似然估计$\hat{p}'=\dfrac{1-\hat{p}}{\hat{p}}=\dfrac{n-k}{k}$

22、设$X_1,\cdots,X_n$为抽自下列指数分布的简单样本；

$$f(x,\theta)=\begin{cases}e^{-(x-\mu)}, & x>\mu\\0, & \text{其他,}\end{cases}$$

其中$-\infty<\mu<+\infty$.

（1）试求μ的极大似然估计$\hat{\mu}^*$，$\hat{\mu}^*$是μ的无偏估计吗？如果不是，使对它作修改，以得到μ的无偏估计$\hat{\mu}^{**}$；

（2）试求μ的矩估计$\hat{\mu}$，并证明它是μ的无偏估计.

（3）试问$\hat{\mu}^{**}$和$\hat{\mu}$哪一个有效？

解：

$$(1)L(\boldsymbol{x},\mu)=e^{-\sum_{i=1}^{n}x_i+n\mu},\mu<x_{(1)}<x_{(n)},$$

$$\ln L(\boldsymbol{x},\mu)=-\sum_{i=1}^{n}x_i+n\mu,\frac{\partial\ln L(\boldsymbol{x},\mu)}{\partial\mu}=n>0,\text{故}\hat{\mu}^*=X_{(1)}$$

令$X_{(1)}=t$,易得$f(t)=ne^{-n(t-\mu)}$

$$E(X_{(1)})=\int_{\mu}^{+\infty}nte^{-n(t-\mu)}dt=\mu+\frac{1}{n}\neq\mu,\text{故}X_{(1)}\text{不是}\mu\text{的无偏估计.}$$

修正后的$\hat{\mu}^{**}=X_{(1)}-\dfrac{1}{n}$是无偏估计.

$$(2)EX=\int_{\mu}^{+\infty}xe^{-(x-\mu)}dx=\mu+1,\text{则}\hat{\mu}=\overline{X}-1,$$

$E\hat{\mu}=E\left(\overline{X}-1\right)=\mu$，故$\hat{\mu}$是无偏估计.

$$(3)D\left(\hat{\mu}^{**}\right)=D\left(X_{(1)}-\frac{1}{n}\right)=D(X_{(1)})$$

$$E(X_{(1)}^2)=\int_{\mu}^{+\infty}nt^2e^{-n(t-\mu)}dt\xlongequal{n(\mu-t)=u}\int_0^{+\infty}\left(\mu^2+\frac{u^2}{n^2}+2\frac{\mu u}{n}\right)e^{-u}du=\mu^2+\frac{2}{n^2}+\frac{2\mu}{n}$$

$$D(X_{(1)})=E(X_{(1)}^2)-E^2(X_{(1)})=\mu^2+\frac{2}{n^2}+\frac{2\mu}{n}-\left(\mu^2+\frac{1}{n^2}+\frac{2\mu}{n}\right)=\frac{1}{n^2}$$

$$EX^2=\int_{\mu}^{+\infty}x^2e^{-(x-\mu)}dx\xlongequal{x-\mu=w}\int_0^{+\infty}(w^2+2\mu w+\mu^2)e^{-w}dw=2+2\mu+\mu^2$$

$$D\left(\hat{\mu}\right)=D\left(\overline{X}\right)=\frac{DX}{n}=\frac{2+2\mu+\mu^2-(1+2\mu+\mu^2)}{n}=\frac{1}{n},$$

故$D\left(\hat{\mu}^{**}\right)<D\left(\hat{\mu}\right)$,即$\hat{\mu}^{**}$更有效.

23、设$X_1,\cdots,X_n\ i.i.d.\sim$均匀分布$U(0,\theta)$,求$\theta$的$MLE\ \hat{\theta}_n$,证明$\hat{\theta}_n$是$\theta$的强相合估计和$r(r>0)$阶矩相合估计.提示：证强相合估计要用到$Borel-Contelli$引理：证明对任给$\varepsilon>0$,有$\sum_{i=1}^{\infty}P\left(\left|\hat{\theta}_n-\theta\right|\geq\varepsilon\right)<\infty$.

解：

$$f(x,\theta)=\begin{cases}\frac{1}{\theta},\ x\in(0,\theta)\\0,\ \text{其他}\end{cases},\quad L(\theta)=\frac{1}{\theta^n}I_{[0<x_{(n)}<\theta]},\text{则可知}\hat{\theta}_n=X_{(n)}$$

令$T=X_{(n)}$,则$f(t)=\dfrac{nt^{n-1}}{\theta^n},0<t<\theta$,则$P\left(\left|\hat{\theta}_n-\theta\right|\geqslant\varepsilon\right)=P\left(|X_{(n)}-\theta|\geqslant\varepsilon\right)$

$=P(0<X_{(n)}<\theta-\varepsilon)=\displaystyle\int_0^{\theta-\varepsilon}f(t)dt=\left(1-\frac{\varepsilon}{\theta}\right)^n$

$\forall\varepsilon>0,\displaystyle\sum_{n=1}^{\infty}P\left(\left|\hat{\theta}_n-\theta\right|\geqslant\varepsilon\right)=\sum_{n=1}^{\infty}\left(1-\frac{\varepsilon}{\theta}\right)^n=\frac{\theta-\varepsilon}{\varepsilon}<\infty,$

故由$Borel-Contelli$引理知：$P\left(\displaystyle\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty}\left|\hat{\theta}_n-\theta\right|\geqslant\varepsilon\right)=0.$

则$\hat{\theta}_n\xrightarrow{a.s.}\theta$,即$\hat{\theta}_n$是$\theta$的强相合估计.

$E\left|\hat{\theta}_n-\theta\right|^r=\displaystyle\int_0^{\theta}(\theta-t)^r\frac{nt^{n-1}}{\theta^n}dt\xlongequal{t=\theta u}n\theta^r\int_0^1(1-u)^r u^{n-1}du$

$=\dfrac{n\theta^r\Gamma(r+1)\Gamma(n)}{\Gamma(r+1+n)}\displaystyle\int_0^1\frac{\Gamma(r+1+n)}{\Gamma(r+1)\Gamma(n)}(1-u)^{r+1-1}u^{n-1}du=\frac{\theta^r\Gamma(r+1)\Gamma(n+1)}{\Gamma(r+1+n)}$

故$\displaystyle\lim_{n\to\infty}E\left|\hat{\theta}_n-\theta\right|^r=\lim_{n\to\infty}\frac{\theta^r\Gamma(r+1)\Gamma(n+1)}{\Gamma(r+1+n)}=\lim_{n\to\infty}\frac{\theta^r r!n!}{(r+n)!}=0$

故$\hat{\theta}_n$是θ的r阶相合估计.

24、设$X_1,\cdots,X_n\ i.i.d.\sim N(\sigma,\sigma^2),\sigma>0$,试求$\sigma^2$的$MLE\ \hat{\sigma}^2$，并证明$\hat{\sigma}$为$\sigma$的弱相合估计.

解：

$$L(\boldsymbol{x},\sigma)=(2\pi)^{-\frac{n}{2}}\sigma^{-n}\exp\left\{-\frac{\sum_{i=1}^{n}(x_i-\sigma)^2}{2\sigma^2}\right\},$$

$$\ln L(\boldsymbol{x},\sigma)=-\frac{n}{2}\ln(2\pi)-n\ln\sigma-\frac{\sum_{i=1}^{n}(x_i-\sigma)^2}{2\sigma^2},$$

$$\frac{\partial\ln L(\boldsymbol{x},\sigma)}{\partial\sigma}=-\frac{n}{\sigma}+\frac{\sum_{i=1}^{n}(x_i-\sigma)^2}{\sigma^3}+\frac{\sum_{i=1}^{n}(x_i-\sigma)}{\sigma^2}=0,$$

$n\sigma^2+\sigma\displaystyle\sum_{i=1}^{n}x_i-\sum_{i=1}^{n}x_i^2=0$,解得$\hat{\sigma}^2=\left(-\dfrac{\overline{X}}{2}+\sqrt{\dfrac{\overline{X}^2}{4}+\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}x_i^2}\right)^2,$

$\hat{\sigma}=-\dfrac{\overline{X}}{2}+\sqrt{\dfrac{\overline{X}^2}{4}+\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}x_i^2}$,又因为$\overline{X}\xrightarrow{P}\sigma,\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}x_i^2\xrightarrow{P}2\sigma^2,$

$\hat{\sigma}$为$\overline{X}$和$\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}x_i^2$的连续函数，故$\hat{\sigma}\xrightarrow{P}-\dfrac{\sigma}{2}+\sqrt{\dfrac{\sigma^2}{4}+2\sigma^2}=\sigma,$

即$\hat{\sigma}$是σ的弱相合估计.

25、设总体X为下列$0-1$分布族

$$P_\theta(X=1)=1-P_\theta(X=0)=\begin{cases}\theta, & \theta\text{为有理数，}\\ 1-\theta, & \theta\text{为无理数.}\end{cases}$$

设$X_1,\cdots,X_n$为从总体X中抽取的简单样本，证明θ的MLE不是相合估计.

解：

$L(\boldsymbol{x},\theta)=\prod_{i=1}^{n}P_\theta(X_i=x_i)=[A(\theta)]^{\sum_{i=1}^{n}x_i}[1-A(\theta)]^{n-\sum_{i=1}^{n}x_i}$,

其中$A(\theta)=\theta I_{[\theta\in Q]}+(1-\theta)I_{[\theta\in R-Q]}$,$Q$表示有理数，$R-Q$表示无理数.

$\ln L(\boldsymbol{x},\theta)=\sum_{i=1}^{n}x_i\ln A(\theta)+\left(n-\sum_{i=1}^{n}x_i\right)\ln[1-A(\theta)]$

$\dfrac{\partial\ln L(\boldsymbol{x},\theta)}{\partial\theta}=\dfrac{\sum_{i=1}^{n}x_i}{A(\theta)}-\dfrac{n-\sum_{i=1}^{n}x_i}{1-A(\theta)}=0$,解得$A(\theta)=\dfrac{1}{n}\sum_{i=1}^{n}x_i$,

其中$\dfrac{1}{n}\sum_{i=1}^{n}x_i$为有理数，故$\hat{\theta}^*=\overline{X}$，

由辛钦大数定律知$\overline{X}\xrightarrow{P}\begin{cases}\theta, & \theta\text{为有理数，}\\ 1-\theta, & \theta\text{为无理数.}\end{cases}$,收敛不唯一，故$\hat{\theta}^*$不是相合估计.

26、设$X_1,\cdots,X_n$是来自伽马分布$\Gamma(\alpha,\lambda)$的简单样本，试证明

（1）当$\lambda=1$时，记α的MLE为$\hat{\alpha}$，则

$$\sqrt{n}(\hat{\alpha}-\alpha)\xrightarrow{\mathscr{L}}N\left(0,\left[E\left(\frac{\partial\log g(x,\alpha)}{\partial\alpha}\right)^2\right]^{-1}\right),$$

此处$g(x,\alpha)$是参数分别为α，$\lambda=1$的伽马分布的密度函数；

（2）当α已知时，λ的MLE $\hat{\lambda}$近似服从$N(\lambda,\lambda^2/(n\alpha))$,即$\sqrt{n}\left(\hat{\lambda}-\lambda\right)\xrightarrow{\mathscr{L}}N(0,\lambda^2/\alpha)$.

解：

(1)$L(\boldsymbol{x},\alpha)=\prod_{i=1}^{n}f(x_i,\alpha)$,$\ln L(\boldsymbol{x},\alpha)=\sum_{i=1}^{n}\ln f(x_i,\alpha)$

$\dfrac{\partial\ln L(\boldsymbol{x},\alpha)}{\partial\alpha}=\sum_{i=1}^{n}\dfrac{\partial\ln f(x_i,\alpha)}{\partial\alpha}=0$,伽马分布显然满足定理3.3.2的条件(1)$-$(3)

因此$\dfrac{\partial\ln L(\boldsymbol{x},\alpha)}{\partial\alpha}=\sum_{i=1}^{n}\dfrac{\partial\ln f(x_i,\alpha)}{\partial\alpha}=0$的解$\hat{\alpha}$既是$\alpha$的$MLE$估计，也是$CAN$估计.

则由定理3.3.2可知$\sqrt{n}(\hat{\alpha}-\alpha)\xrightarrow{\mathscr{L}}N\left(0,\left[E\left(\dfrac{\partial\log g(x,\alpha)}{\partial\alpha}\right)^2\right]^{-1}\right)$.

(2)$g(x,\lambda)=\dfrac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x}$,$L(\boldsymbol{x},\lambda)=\dfrac{\lambda^{n\alpha}}{\Gamma^n(\alpha)}\prod_{i=1}^{n}x_i^{\alpha-1}e^{-\lambda\sum_{i=1}^{n}x_i}$,

$$\ln L(\boldsymbol{x},\lambda)=n\alpha\ln\lambda-n\ln[\Gamma(\alpha)]+(\alpha-1)\ln x_i-\lambda\sum_{i=1}^{n}x_i,$$

$$\frac{\partial\ln L(\boldsymbol{x},\lambda)}{\partial\lambda}=\frac{n\alpha}{\lambda}-\sum_{i=1}^{n}x_i=0,\text{得}\hat{\lambda}=n\alpha\Big/\sum_{i=1}^{n}x_i,\hat{\lambda}\text{也是}\lambda\text{的}CAN\text{估计}.$$

$$I(\theta)=E\left(\frac{\partial\log g(x,\lambda)}{\partial\lambda}\right)^2=E\left(\frac{\alpha}{\lambda}-X\right)^2=EX^2-\frac{2\alpha}{\lambda}EX+\frac{\alpha^2}{\lambda^2}$$

$$=\frac{\alpha}{\lambda^2}+\frac{\alpha^2}{\lambda^2}-\frac{2\alpha^2}{\lambda^2}+\frac{\alpha^2}{\lambda^2}=\frac{\alpha}{\lambda^2}\text{，故}\sqrt{n}\left(\hat{\lambda}-\lambda\right)\xrightarrow{\mathscr{L}}N(0,1/I(\theta))=N(0,\lambda^2/\alpha)$$

27、考察均匀分布族$\{U(0,\theta):\theta>0\}$,则不管样本容量n有多大，$g(\theta)=1/\theta$不是可估.

参数.试以$n=1$为例，证明这个结论.

解：

证明：当$n=1$时，设样本为X,假设$g(\theta)=1/\theta$可估，

则$\exists\, g(x)$,使$E_\theta[g(x)]=1/\theta$,对$\forall\theta>0$成立

即$\displaystyle\int_0^\theta g(x)\frac{1}{\theta}dx=\frac{1}{\theta}\Longrightarrow\int_0^\theta g(x)dx=1\xRightarrow{\forall\theta>0}P_\theta(g(x)=0)=1$,

与假设$E_\theta[g(x)]=1/\theta$矛盾,故$g(\theta)=1/\theta$不可估.

28、设$X_1,\cdots,X_n\ i.i.d.\sim N(\theta,1)$,证明$g(\theta)=|\theta|$没有无偏估计，即$g(\theta)$不是可估参数.

(提示：利用$g(\theta)$在$\theta=0$不可导)

解：

设$\hat{g}(x)$为无偏估计，则对$-\infty<\theta<+\infty$有

$$E_\theta(\hat{g}(x))=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\hat{g}(x)\exp\left\{-\frac{1}{2}(x-\theta)^2\right\}dx=|\theta|.$$

但由指数族的性质，知上式中间一项在$-\infty<\theta<+\infty$对θ有各级连续导数，

而$|\theta|$在$\theta=0$处不可导，因此是不可能的，$g(\theta)$为不可估参数.

29、试证明定理3.4.1 逆成立：若$\hat{g}$为$g(\theta)$的$UMVUE$,$D_\theta(\hat{g})<\infty$ 对任何$\theta\in\Theta$，

则当$E_\theta(\hat{l})=0$且$D_\theta(\hat{l})<\infty$对一切$\theta\in\Theta$成立时，必有$Cov_\theta(\hat{g},\hat{l})=0$对一切$\theta\in\Theta$.

解：

设存在$\theta_0\in\Theta$,有$Cov_{\theta_0}(\hat{g},\hat{l})\neq0$

则$D_{\theta_0}(\hat{g}+\delta\hat{l})=D_{\theta_0}(\hat{g})+\delta^2D_{\theta_0}(\hat{l})+2\delta Cov_{\theta_0}(\hat{g},\hat{l})$

$=D_{\theta_0}(\hat{g})+D_{\theta_0}(\hat{l})\left[\delta\left(\delta+2Cov_{\theta_0}(\hat{g},\hat{l})\right)\right],\exists\,\delta_0$,使$\delta_0\left(\delta_0+2Cov_{\theta_0}(\hat{g},\hat{l})\right)<0$

此时$D_{\theta_0}(\hat{g}+\delta\hat{l})<D_{\theta_0}(\hat{g})$,与$\hat{g}$为$g(\theta)$的$UMVUE$矛盾，故$Cov_\theta(\hat{g},\hat{l})=0$，对一切$\theta\in\Theta$

30、设$X_1,\cdots,X_n\ i.i.d.\sim N(a,\sigma^2)$,用零无偏估计法证明$S^2=\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2/n-1$为

σ^2的$UMVUE.$
解：

由例2.7.5可知$T=(T_1,T_2)$为$\theta=(a,\sigma^2)$的充分统计量，其中$T_1=\overline{X},T_2=\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2$
又T_1与T_2独立，且$T_1\sim N(a,\sigma^2/n),T_2/\sigma^2\sim\chi^2_{n-1}$,因此$T_1,T_2$的联合密度为

$$f_\theta(t_1,t_2)=\begin{cases}\dfrac{\sqrt{n}}{\sqrt{2\pi}\sigma}e^{-\frac{n(t_1-a)^2}{2\sigma^2}}\cdot\dfrac{1}{2^{\frac{n-1}{2}}\Gamma((n-1)/2)\sigma^{n-1}}t_2^{\frac{n-1}{2}-1}e^{-\frac{t_2}{2\sigma^2}}, & -\infty<t_1<+\infty,t_2>0\\ 0, & \text{其他}\end{cases}$$

令$h_2(T)=T_2/(n-1)$,故$h_2(T)$是$g_2(\theta)=\sigma^2$的无偏估计,且$D_\theta(h_2(T))=\dfrac{2\sigma^4}{n-1}<\infty$

设$\delta(T)=\delta(T_1,T_2)$的任一零无偏估计，则有

$$E_\theta(\delta(T))=\int_0^\infty\int_{-\infty}^{+\infty}\delta(t_1,t_2)f_\theta(t_1,t_2)dt_1dt_2=0,\text{此处}-\infty<a<+\infty,\sigma>0,\tag{1}$$

将上式两边对a求导，得

$$\int_0^\infty\int_{-\infty}^{+\infty}\delta(t_1,t_2)(t_1-a)\cdot t_2^{\frac{n-3}{2}}\exp\left\{-\frac{1}{2\sigma^2}[n(t_1-a)^2+t_2]\right\}dt_1dt_2=0\tag{2}$$

再对a求导得

$$\int_0^\infty\int_{-\infty}^{+\infty}\delta(t_1,t_2)\left[\frac{n(t_1-a)^2}{\sigma^2}-1\right]\cdot t_2^{\frac{n-3}{2}}\exp\left\{-\frac{1}{2\sigma^2}[n(t_1-a)^2+t_2]\right\}dt_1dt_2=0\tag{3}$$

将(1)式对σ求导得

$$\int_0^\infty\int_{-\infty}^{+\infty}\delta(t_1,t_2)\frac{n(t_1-a)^2+t_2-n\sigma^2}{\sigma^{n+3}}\cdot t_2^{\frac{n-3}{2}}\exp\left\{-\frac{1}{2\sigma^2}[n(t_1-a)^2+t_2]\right\}dt_1dt_2=0\tag{4}$$

由(1) (3)式得$\int_0^\infty\int_{-\infty}^{+\infty}\delta(t_1,t_2)n(t_1-a)^2\cdot t_2^{\frac{n-3}{2}}\exp\left\{-\frac{1}{2\sigma^2}[n(t_1-a)^2+t_2]\right\}dt_1dt_2=0$

故$\int_0^\infty\int_{-\infty}^{+\infty}\delta(t_1,t_2)(t_2-n\sigma^2)\cdot t_2^{\frac{n-3}{2}}\exp\left\{-\frac{1}{2\sigma^2}[n(t_1-a)^2+t_2]\right\}dt_1dt_2=0\Rightarrow$

$$\int_0^\infty\int_{-\infty}^{+\infty}\delta(t_1,t_2)\frac{t_2}{n-1}\cdot t_2^{\frac{n-3}{2}}\exp\left\{-\frac{1}{2\sigma^2}[n(t_1-a)^2+t_2]\right\}dt_1dt_2=0,$$

因此$E_\theta(h_2(T)\cdot\delta(T_1,T_2))=0,-\infty<a<+\infty,\sigma>0$
故推论3.4.1的条件满足.所以$h_2(T)=T_2/(n-1)$为$g_2(\theta)=\sigma^2$的$UMVUE.$

31、设$X_1,\cdots,X_n\ i.i.d.\sim N(0,\sigma^2),\sigma^2>0$为未知参数，试求
（1）σ^2的充分完全统计量;
（2）σ和$3\sigma^4$的$UMVUE.$
解：

$(1)f(x,\sigma^2)=\frac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{x^2}{2\sigma^2}},f(\boldsymbol{x},\sigma^2)=(2\pi)^{-\frac{n}{2}}(\sigma^2)^{-\frac{n}{2}}e^{-\frac{\sum_{i=1}^{n}x_i^2}{2\sigma^2}}=g(t(\boldsymbol{x}),\theta)h(x),$

其中$g(t(\boldsymbol{x}),\theta)=(2\pi)^{-\frac{n}{2}}(\sigma^2)^{-\frac{n}{2}}e^{-\frac{\sum_{i=1}^{n}x_i^2}{2\sigma^2}}$，$t(\boldsymbol{x})=\sum_{i=1}^{n}x_i^2,h(x)\equiv 1$,

故$T(X)=\sum_{i=1}^{n}X_i^2$是充分统计量，令$\varphi=-\frac{1}{2\sigma^2}$，$f(\boldsymbol{x},\varphi)=\left(\frac{-\varphi}{\pi}\right)^{\frac{n}{2}}e^{\varphi t(\boldsymbol{x})}h(x)$

自然参数空间$\Theta^*=\{\varphi：-\infty<\varphi<0\}$,作为$R_1$的子集有内点.

故$T(X)=\sum_{i=1}^{n}X_i^2$是σ^2的充分完全统计量.

$(2)Y=T(X)/\sigma^2\sim\chi_n^2$，$Y$的密度函数为

$$g(y)=\begin{cases}\frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}y^{\frac{n}{2}-1}e^{-\frac{1}{2}y},y>0\\ 0, \quad\quad\quad\quad 其他,\end{cases}$$

$$E\left(\frac{T}{\sigma^2}\right)^{\frac{r}{2}}=\int_0^{+\infty}\frac{y^{\frac{r}{2}}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}y^{\frac{n}{2}-1}e^{-\frac{1}{2}y}dy=\frac{2^{\frac{r}{2}}\Gamma\left(\frac{n+r}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\int_0^{+\infty}\frac{1}{2^{\frac{n+r}{2}}\Gamma\left(\frac{n+r}{2}\right)}y^{\frac{n+r}{2}-1}e^{-\frac{1}{2}y}dy$$

$$=\frac{2^{\frac{r}{2}}\Gamma\left(\frac{n+r}{2}\right)}{\Gamma\left(\frac{n}{2}\right)},故E\left(\frac{\Gamma\left(\frac{n}{2}\right)}{2^{\frac{r}{2}}\Gamma\left(\frac{n+r}{2}\right)}T^{\frac{r}{2}}\right)=\sigma^r$$

令$h_1(T)=\frac{\Gamma\left(\frac{n}{2}\right)}{2^{\frac{1}{2}}\Gamma\left(\frac{n+1}{2}\right)}T^{\frac{1}{2}}$，则$h_1(T)$是$\sigma$的无偏估计，且$T(X)$是充分完全统计量.

由$L-S$定理可知，$h_1(T)$是σ的$UMVUE$.

同理$h_2(T)=\frac{3\Gamma\left(\frac{n}{2}\right)}{2^{\frac{4}{2}}\Gamma\left(\frac{n+4}{2}\right)}T^{\frac{4}{2}}=\frac{3T^2}{n(n+2)}$是$3\sigma^4$的$UMVUE$.

32、设$X_1,\cdots,X_n\ i.i.d.\sim N(\mu,\sigma^2),-\infty<\mu<+\infty,\sigma^2>0$都是未知参数，试求参数函数

（1）$3\mu+4\sigma^2$；

（2）$\mu^2/4\sigma^2$的$UMVUE$.

解：

(1)由样本联合密度函数可知$\left(\overline{X},S^2\right)$是$\theta=(\mu,\sigma^2)$的充分完全统计量.

其中$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i,S^2=\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2,E\overline{X}=\mu,ES^2=\sigma^2,$

则由$L-S$定理可知，$3\overline{X}+4S^2$是$3\mu+4\sigma^2$的$UMVUE$.

(2)令$Y=\dfrac{n-1}{\sigma^2}S^2$,则$Y\sim\chi^2_{n-1}$

$$\mathrm{E}\left(\frac{1}{\mathrm{Y}}\right)=\int_0^{+\infty}\frac{1}{\mathrm{y}}\frac{1}{2^{\frac{\mathrm{n}-1}{2}}\Gamma\left(\frac{\mathrm{n}-1}{2}\right)}\mathrm{y}^{\frac{\mathrm{n}-1}{2}-1}\mathrm{e}^{-\frac{1}{2}\mathrm{y}}\mathrm{dy}=\frac{\Gamma\left(\frac{\mathrm{n}-3}{2}\right)}{2\Gamma\left(\frac{\mathrm{n}-1}{2}\right)}\int_0^{+\infty}\frac{1}{2^{\frac{\mathrm{n}-3}{2}}\Gamma\left(\frac{\mathrm{n}-3}{2}\right)}\mathrm{y}^{\frac{\mathrm{n}-3}{2}-1}\mathrm{e}^{-\frac{1}{2}\mathrm{y}}\mathrm{dy}$$

$$=\frac{\Gamma\left(\frac{\mathrm{n}-1}{2}-1\right)}{2\Gamma\left(\frac{\mathrm{n}-1}{2}\right)}=\frac{1}{\mathrm{n}-3},\text{则}\mathrm{E}\left(\frac{1}{\mathrm{S}^2}\right)=\frac{\mathrm{n}-1}{(\mathrm{n}-3)\sigma^2},$$

又因为$\overline{X}$与S^2独立,故$\mathrm{E}\left(\dfrac{\overline{X}^2}{S^2}\right)=\mathrm{E}\left(\overline{X}^2\right)\mathrm{E}\left(\dfrac{1}{\mathrm{S}^2}\right)=\left(\mu^2+\dfrac{\sigma^2}{\mathrm{n}}\right)\cdot\dfrac{\mathrm{n}-1}{(\mathrm{n}-3)\sigma^2}$

$$=\frac{(n-1)\mu^2}{(n-3)\sigma^2}+\frac{\mathrm{n}-1}{\mathrm{n}(\mathrm{n}-3)}.$$

因此$\mathrm{E}\left(\dfrac{\overline{X}^2}{S^2}-\dfrac{\mathrm{n}-1}{\mathrm{n}(\mathrm{n}-3)}\right)=\dfrac{(n-1)\mu^2}{(n-3)\sigma^2}$，由L－S定理可知

$\dfrac{(n-3)\overline{X}^2}{4(n-1)S^2}-\dfrac{1}{4\mathrm{n}}$是$\dfrac{\mu^2}{4\sigma^2}$的UMVUE.

33、 设$X_1,\cdots,X_n$ $i.i.d.$服从两点分布$b(1,p)$, $0<p<1$是未知参数，试求

（1）p^s的$UMVUE$.

（2）$p^s+(1-p)^{n-s}$的$UMVUE$($0<s<n$为整数).

解：

(1)由例2.7.6和例2.8.1可知$T(\boldsymbol{X})=\sum\limits_{i=1}^{n}X_i$为充分完全统计量.

设$\delta(t)$为$g(p)=p^s$的无偏估计，易知$T(X)\sim b(n,p)$,则有

$\sum\limits_{t=0}^{n}\binom{n}{t}p^t(1-p)^{n-t}\delta(t)=p^s$,令$\rho=\dfrac{p}{1-p}$,则$p=\dfrac{\rho}{1+\rho}$, $1-p=\dfrac{1}{1+\rho}$

$$\sum_{t=0}^{n}\binom{n}{t}\left(\frac{\rho}{1+\rho}\right)^t\left(\frac{1}{1+\rho}\right)^{n-t}\delta(t)=\left(\frac{\rho}{1+\rho}\right)^s\rightarrow\sum_{t=0}^{n}\binom{n}{t}\rho^t\delta(t)=\rho^s(1+\rho)^{n-s}$$

$$(1+\rho)^{n-s}=\sum_{t=0}^{n-s}\binom{n-s}{t}\rho^t,\rho^s(1+\rho)^{n-s}=\sum_{t=0}^{n-s}\binom{n-s}{t}\rho^{t+s}=\sum_{t=s}^{n}\binom{n-s}{t-s}\rho^t$$

故$\sum\limits_{t=0}^{n}\binom{n}{t}\rho^t\delta(t)=\sum\limits_{t=s}^{n}\binom{n-s}{t-s}\rho^t$，因此

$\delta(t)=0$,当$t=0,1,\cdots,s-1$. $\delta(t)=\binom{n-s}{t-s}\Big/\binom{n}{t}$, $t=s,\cdots,n$.

综上所述，$\delta(T)=\begin{cases}0, & T=0,1,\cdots,s-1\\ \binom{n-s}{t-s}\Big/\binom{n}{t}, & T=s,\cdots,n\end{cases}$.

$\delta(T)$为$g(p)=p^s$的无偏估计,且是充分完全统计量$T(\boldsymbol{X})=\sum_{i=1}^{n}X_i$的函数,

由$L-S$定理可知$\delta(T)$为p^s的$UMVUE$.

(2) $\sum_{t=0}^{n}\binom{n}{t}p^t(1-p)^{n-t}\delta(t)=p^s+(1-p)^{n-s}$,令$\rho=\dfrac{p}{1-p}$,则$p=\dfrac{\rho}{1+\rho}$,$1-p=\dfrac{1}{1+\rho}$

则$\sum_{t=0}^{n}\binom{n}{t}\left(\dfrac{\rho}{1+\rho}\right)^t\left(\dfrac{1}{1+\rho}\right)^{n-t}\delta(t)=\left(\dfrac{\rho}{1+\rho}\right)^s+\left(\dfrac{1}{1+\rho}\right)^{n-s}$

$\to\sum_{t=0}^{n}\binom{n}{t}\rho^t\delta(t)=\rho^s(1+\rho)^{n-s}+(1+\rho)^s$

$\rho^s(1+\rho)^{n-s}+(1+\rho)^s=\rho^s\sum_{t=0}^{n-s}\binom{n-s}{t}\rho^t+\sum_{t=0}^{s}\binom{s}{t}\rho^t=\sum_{t=s}^{n}\binom{n-s}{t-s}\rho^t+\sum_{t=0}^{s}\binom{s}{t}\rho^t$

即$\sum_{t=0}^{n}\binom{n}{t}\rho^t\delta(t)=\sum_{t=s}^{n}\binom{n-s}{t-s}\rho^t+\sum_{t=0}^{s}\binom{s}{t}\rho^t$，因此

$$\delta(T)=\begin{cases}\binom{s}{t}\Big/\binom{n}{t}, & T=0,1,\cdots,s-1\\ 2\Big/\binom{n}{s}, & T=s\\ \binom{n-s}{t-s}\Big/\binom{n}{t}, & T=s+1,\cdots,n\end{cases},\delta(T)\text{为}p^s+(1-p)^{n-s}\text{的}UMVUE.$$

34、设$X_1,\cdots,X_m\ i.i.d.\sim N(a,\sigma^2)$,$Y_1,\cdots,Y_n\ i.i.d.\sim N(a,2\sigma^2)$,且两组样本独立，求$a$和$\sigma^2$的$UMVUE$.

解：

$$f(\boldsymbol{x},\boldsymbol{y},\theta)=2^{-\frac{n}{2}}(2\pi)^{-\frac{m+n}{2}}\sigma^{-(m+n)}e^{-\frac{2\sum_{i=1}^{m}(x_i-a)^2+\sum_{j=1}^{n}(y_j-a)^2}{4\sigma^2}}$$

$$=2^{-\frac{n}{2}}(2\pi)^{-\frac{m+n}{2}}\sigma^{-(m+n)}e^{-\frac{2\sum_{i=1}^{m}x_i^2+\sum_{j=1}^{n}y_j^2-4a\sum_{i=1}^{m}x_i-2a\sum_{j=1}^{n}y_j-(2m+n)a^2}{4\sigma^2}}$$

$=g((t_1(\boldsymbol{x},\boldsymbol{y}),t_2(\boldsymbol{x},\boldsymbol{y})),\theta)h(\boldsymbol{x},\boldsymbol{y})$,

其中$t_1(\boldsymbol{x},\boldsymbol{y})=2\sum_{i=1}^{m}x_i+\sum_{j=1}^{n}y_j$，$t_2(\boldsymbol{x},\boldsymbol{y})=2\sum_{i=1}^{m}x_i^2+\sum_{j=1}^{n}y_j^2$,

$g((t_1(\boldsymbol{x},\boldsymbol{y}),t_2(\boldsymbol{x},\boldsymbol{y})),\theta)=2^{-\frac{n}{2}}(2\pi)^{-\frac{m+n}{2}}\sigma^{-(m+n)}e^{-\frac{2\sum_{i=1}^{m}x_i^2+\sum_{j=1}^{n}y_j^2-4a\sum_{i=1}^{m}x_i-2a\sum_{j=1}^{n}y_j-(2m+n)a^2}{4\sigma^2}}$，$h(\boldsymbol{x})\equiv1$

且令$\varphi_1=-\dfrac{a}{2\sigma^2}$,$\varphi_2=\dfrac{1}{4\sigma^2}$，自然参数空间$\Theta^*=\{(\varphi_1,\varphi_2):-\infty<\varphi_1<+\infty,0<\varphi_2<+\infty\}$

作为R_2的子集有内点，因此$(t_1(\boldsymbol{x},\boldsymbol{y}),t_2(\boldsymbol{x},\boldsymbol{y}))$是$(a,\sigma^2)$的充分完全统计量.

$E(T_1(\boldsymbol{X},\boldsymbol{Y}))=(2m+n)a$,$E(T_2(\boldsymbol{X},\boldsymbol{Y}))=(2m+n)a^2+(2m+2n)\sigma^2$,

$E(T_1^2(\boldsymbol{X},\boldsymbol{Y}))=(4m+2n)\sigma^2+(2m+n)^2a^2$,

因此$E\left(\frac{T_1(\boldsymbol{X},\boldsymbol{Y})}{2m+n}\right)=a,E\left(\frac{T_2(\boldsymbol{X},\boldsymbol{Y})-\frac{T_1{}^2(\boldsymbol{X},\boldsymbol{Y})}{2m+n}}{2m+2n-2}\right)=\sigma^2$,由$L-S$定理可知,

$\frac{T_1(\boldsymbol{X},\boldsymbol{Y})}{2m+n}$和$\frac{T_2(\boldsymbol{X},\boldsymbol{Y})-\frac{T_1{}^2(\boldsymbol{X},\boldsymbol{Y})}{2m+n}}{2m+2n-2}$分别为$a$和$\sigma^2$的$UMVUE$.

35、设$X_1,\cdots,X_n$ $i.i.d.\sim$几何分布族

$$P(X_1=i)=\theta(1-\theta)^{i-1},i=1,2,\cdots,0<\theta<1$$

（1）试证明$T=\sum_{i=1}^{n}X_i$是θ的充分完全统计量，且服从帕斯卡(负二项)分布

$$P_\theta(T=t)=\binom{t-1}{n-1}\theta^n(1-\theta)^{t-n},t=n,n+1,n+2,\cdots.$$

（2）计算$E_\theta T$,并由此求θ^{-1}的$UMVUE$.

（3）试证明

$$\psi(X_1)=\begin{cases}1,X_1=1,\\0,X_1=2,3,\cdots\end{cases}$$

是θ的无偏估计，计算$E_\theta(\psi(X_1)|T=t)$,并由此求得θ的$UMVUE$.

解：

(1)$f(\boldsymbol{x},\theta)=\theta^n(1-\theta)^{\sum_{i=1}^{n}x_i-n}=g(t(\boldsymbol{x}),\theta)h(x)$,

其中$t(\boldsymbol{x})=\sum_{i=1}^{n}x_i,g(t(\boldsymbol{x}),\theta)=\theta^n(1-\theta)^{\sum_{i=1}^{n}x_i-n},h(x)\equiv1$,故$T(\boldsymbol{X})=\sum_{i=1}^{n}X_i$是充分统计量.

$\varphi=\ln(1-\theta)$,则自然参数空间$\Theta^*=\{\varphi:-\infty<\varphi<0\}$作为$R_1$的子集有内点.

故$T(\boldsymbol{X})=\sum_{i=1}^{n}X_i$是充分完全统计量.

$P_\theta(T=t)$可以理解为有现有t个“○”将其分为n堆,

即在t个“○”的$t-1$个间隙中选取$n-1$个空隙填入“|”，○|○⋯○○|○○

即$P_\theta(T=t)=\binom{t-1}{n-1}\theta^n(1-\theta)^{t-n},t=n,n+1,n+2,\cdots.$

(2)$E_\theta T=\sum_{k=n}^{\infty}kP(T=k)=\sum_{k=n}^{\infty}k\binom{k-1}{n-1}\theta^n(1-\theta)^{k-n}\xlongequal{k\binom{k-1}{n-1}=n\binom{k}{n}}\sum_{k=n}^{\infty}n\binom{k}{n}\theta^n(1-\theta)^{k-n}$

$=\frac{1}{\theta}\sum_{k=n}^{\infty}n\binom{k+1-1}{n+1-1}\theta^{n+1}(1-\theta)^{k+1-(n+1)}=\frac{n}{\theta}$,则$\frac{T}{n}$是$\theta^{-1}$的无偏估计,

由$L-S$定理可知$\frac{T}{n}$是θ^{-1}的$UMVUE$.

(3) $E\psi(X_1)=\psi(1)P(X_1=1)=\theta$,因此$\psi(X_1)$是$\theta$的无偏估计

$$E_\theta(\psi(X_1)\,|T=t)=E_\theta\left(I_{[X_1=1]}|T=t\right)=1\cdot P(X_1=1\,|T=t)$$

$$=\frac{P\left(X_1=1,\sum_{i=2}^{n}X_i=t-1\right)}{P(T=t)}=\frac{\theta\cdot\binom{t-2}{n-2}\theta^{n-1}(1-\theta)^{t-n}}{\binom{t-1}{n-1}\theta^n(1-\theta)^{t-n}}=\frac{n-1}{t-1},$$

则$\dfrac{n-1}{T-1}$为θ的$UMVUE$.

36、设有分布族$\{\theta e^{-\theta x}I_{(0,\infty)}(x),\theta>0\}$,$X_1,\cdots,X_n$为从中抽取的简单样本，求$e^{-\theta\tau}$的$UMVUE$(注意：$e^{-\theta\tau}=P_\theta(X_1>\tau)$),此处$\tau>0$为给定的数.

解：

$f(\boldsymbol{x},\theta)=\theta^n e^{-\theta\sum_{i=1}^{n}x_i}I_{[0<x_{(1)}<x_{(n)}<\infty]}=g(t(\boldsymbol{x}),\theta)h(\boldsymbol{x})$,其中$t(\boldsymbol{x})=\sum_{i=1}^{n}x_i$

$g(t(\boldsymbol{x}),\theta)=\theta^n e^{-\theta\sum_{i=1}^{n}x_i}$,$h(\boldsymbol{x})=I_{[0<x_{(1)}<x_{(n)}<\infty]}$,因此$T(\boldsymbol{X})=\sum_{i=1}^{n}x_i$是充分统计量.

$f(\boldsymbol{x},\theta)=(-\varphi)^n\exp\{\varphi T(\boldsymbol{X})\}h(\boldsymbol{x})$,自然参数空间$\Theta^*=\{\varphi：\ -\infty<\varphi<0\}$,

作为R_1的子集有内点.故$T(\boldsymbol{X})=\sum_{i=1}^{n}x_i$是充分完全统计量.

$X\sim\Gamma(1,\theta)$,$T(\boldsymbol{X})=\sum_{i=1}^{n}x_i\sim\Gamma(n,\theta)$现设$h(\boldsymbol{T})$为$e^{-\theta\tau}$的无偏估计，则有

$$\int_0^\infty h(t)\frac{\theta^n}{\Gamma(n)}t^{n-1}e^{-\theta t}dt=e^{-\theta\tau}\int_0^\infty h(t)\frac{\theta^n}{\Gamma(n)}t^{n-1}e^{-\theta(t-\tau)}dt,$$

易见要使$E[h(\boldsymbol{T})]=e^{-\theta\tau}$,必须使

$h(t)=\dfrac{(t-\tau)^{n-1}}{t^{n-1}}I_{(\tau,\infty)}(t)$,则$\dfrac{(t-\tau)^{n-1}}{t^{n-1}}I_{(\tau,\infty)}(t)$为$e^{-\theta\tau}$的无偏估计.

由$L-S$定理可知$\dfrac{(t-\tau)^{n-1}}{t^{n-1}}I_{(\tau,\infty)}(t)$为$e^{-\theta\tau}$的$UMVUE$.

37、设$X_1,\cdots,X_n$是来自均匀分布总体$U(\theta_1,\theta_2)$的简单随机样本.

（1）试求θ_1和θ_2的MLE $\hat{\theta}_1$和$\hat{\theta}_2$；

（2）$\hat{\theta}_1$和$\hat{\theta}_2$分别是θ_1和θ_2的无偏估计吗？如果不是，请加以修正，以获得θ_1和θ_2的无偏估计$\Big($提示：确定a和b,使得$a\hat{\theta}_1+b\hat{\theta}_2$是$\theta_1$(或$\theta_2$)的无偏估计$\Big)$；

（3）若$\hat{\theta}_1$和$\hat{\theta}_2$为充分完全统计量，由(2)得到的无偏估计是否分别为$\hat{\theta}_1$和$\hat{\theta}_2$的$UMVUE$？

（4）试分别求出中点$(\theta_1+\theta_2)/2$和变程$\theta_2-\theta_1$的$UMVUE$.

解：

(1) $f(\boldsymbol{x},\theta_1,\theta_2)=\dfrac{1}{(\theta_2-\theta_1)^n}I_{[\theta_1<x_{(1)}<x_{(n)}<\theta_2]}$，则$\hat{\theta}_1=X_{(1)},\hat{\theta}_2=X_{(n)}$

故θ_1和θ_2的MLE $\hat{\theta}_1$和$\hat{\theta}_2$分别为$X_{(1)}$和$X_{(n)}$

(2)令$Y=\dfrac{X-\theta_1}{\theta_2-\theta_1}$，则$Y\sim U(0,1)$

$f_Y(y)=1,y\in(0,1)$，则$f_{Y_{(1)}}(y)=n(1-y)^{n-1}=\dfrac{\Gamma(n+1)}{\Gamma(1)\Gamma(n)}y^{1-1}(1-y)^{n-1}$

即$Y_{(1)}=\dfrac{X_{(1)}-\theta_1}{\theta_2-\theta_1}\sim Be(1,n)$，则$EY_{(1)}=\dfrac{1}{1+n}$，即$EX_{(1)}=\dfrac{n}{1+n}\theta_1+\dfrac{1}{1+n}\theta_2$

$f_{Y_{(n)}}(y)=ny^{n-1}=\dfrac{\Gamma(n+1)}{\Gamma(1)\Gamma(n)}y^{n-1}(1-y)^{1-1}$

即$Y_{(n)}=\dfrac{X_{(n)}-\theta_1}{\theta_2-\theta_1}\sim Be(n,1)$，则$EY_{(1)}=\dfrac{n}{1+n}$，即$EX_{(n)}=\dfrac{1}{1+n}\theta_1+\dfrac{n}{1+n}\theta_2$

故$\hat{\theta}_1$和$\hat{\theta}_2$不是θ_1和θ_2的无偏估计，修正得$\hat{\theta}_1{*}=\dfrac{nX_{(1)}-X_{(n)}}{n-1}$是$\theta_1$的无偏估计.

$\hat{\theta}_2{*}=\dfrac{nX_{(n)}-X_{(1)}}{n-1}$是$\theta_2$的无偏估计.

(3)若$\hat{\theta}_1$和$\hat{\theta}_2$为充分完全统计量，由(2)得到的无偏估计由$L-S$定理可知
是分别为$\hat{\theta}_1$和$\hat{\theta}_2$的$UMVUE$.

(4) $E\dfrac{X_{(1)}+X_{(n)}}{2}=\dfrac{\theta_1+\theta_2}{2}$，$E\left(\dfrac{n+1}{n-1}(X_{(n)}-X_{(1)})\right)=\theta_2-\theta_1$，则由$L-S$定理可知

$\dfrac{X_{(1)}+X_{(n)}}{2}$和$\dfrac{n+1}{n-1}(X_{(n)}-X_{(1)})$分别为中点$(\theta_1+\theta_2)/2$和变程$\theta_2-\theta_1$的$UMVUE$.

38、设$X_1,\cdots,X_n$ $i.i.d.\sim N(\theta,1)$，求θ^2的$UMVUE$.证明此$UMVUE$的方差达不到

$C-R$不等式的下界，即它不是有效估计.

解：

由例2.7.2和例2.8.3可知$T(\boldsymbol{X})=\overline{X}=\sum_{i=1}^{n}X_i\Big/n$是充分完全统计量.

易知$E\overline{X}^2=\theta^2+\dfrac{1}{n}$，故$\overline{X}^2-\dfrac{1}{n}$是$\theta^2$的无偏估计量，又是$\overline{X}$的函数

因此根据$L-S$定理可知$\overline{X}^2-\dfrac{1}{n}$是$\theta^2$的$UMVUE$.

由于正态分布是指数族，$C-R$正则条件皆成立.密度函数为

$$f(x)=\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{(x-\theta)^2}{2}\right\}$$

$Fisher$信息函数为

$$I(\theta)=E_\theta\left[\frac{\partial \log f(X,\theta)}{\partial\theta}\right]^2=E_\theta[(X-\theta)^2]=DX=1$$

故$C-R$下界为$4\theta^2/n$,而$D\left(\overline{X}^2-\frac{1}{n}\right)=D\overline{X}^2=E\overline{X}^4-\left(E\overline{X}^2\right)^2$

$E\overline{X}^2=D\overline{X}+\left(E\overline{X}\right)^2=\theta^2+\frac{1}{n},\varphi_{\overline{X}}(t)=e^{i\theta t-\frac{1}{2n}t^2}$,

$\varphi_{\overline{X}}(t)$在$t=0$处的泰勒展开为

$$\varphi_{\overline{X}}(t)=\left[1+i\theta t+\frac{(i\theta t)^2}{2!}+\frac{(i\theta t)^3}{3!}+\frac{(i\theta t)^4}{4!}+o(t^4)\right]\cdot\left[1-\frac{1}{2n}t^2+\frac{\left(-\frac{1}{2n}t^2\right)^2}{2!}+o(t^4)\right]$$

$=\cdots+\frac{1}{8n^2}t^4+\frac{\theta^2}{4n}t^4+\frac{\theta^4}{4!}t^4+o(t^4)$,则$\varphi_{\overline{X}}^{(4)}(0)=\frac{3}{n^2}+\frac{6\theta^2}{n}+\theta^4$,

又因为$\varphi_{\overline{X}}^{(4)}(0)=i^4E\overline{X}^4$,因此$E\overline{X}^4=\frac{3}{n^2}+\frac{6\theta^2}{n}+\theta^4$，则

$D\overline{X}^2=\frac{3}{n^2}+\frac{6\theta^2}{n}+\theta^4-\left(\theta^2+\frac{1}{n}\right)^2=\frac{4\theta^2}{n}+\frac{2}{n^2}$,即$D\left(\overline{X}^2-\frac{1}{n}\right)>\frac{4\theta^2}{n}$,

即θ^2的$UMVUE$ $\overline{X}^2-\frac{1}{n}$达不到$C-R$不等式的下界，不是有效估计.

39、证明均匀分布族$\mathscr{F}=\{U(0,\theta):0<\theta<\infty\}$不是$C-R$正则分布族.

解：

对于$C-R$正则条件(4)

$\frac{\partial}{\partial\theta}\int f(x,\theta)dx=\frac{\partial 1}{\partial\theta}=0,\int\frac{\partial}{\partial\theta}f(x,\theta)dx=\int_0^\theta -\frac{1}{\theta^2}dx=-\frac{1}{\theta}$,

故$\frac{\partial}{\partial\theta}\int f(x,\theta)dx\neq\int\frac{\partial}{\partial\theta}f(x,\theta)dx$，即此分布族不满足正则条件,

因此均匀分布族$\mathscr{F}=\{U(0,\theta):0<\theta<\infty\}$不是$C-R$正则分布族.

40、设$X_1,\cdots,X_n$ 为来自下列总体中抽取的简单样本,

$$f(x,\theta)=\begin{cases}\theta^{-1}e^{-x/\theta}, & x>0\\ 0, & 其他\end{cases}$$

其中$\theta>0$为未知参数，求θ的$UMVUE$，并比较此$UMVUE$的方差与θ的无偏估计

方差的$C-R$下界.

解：

易证$T(\boldsymbol{X})=\sum_{i=1}^n X_i$为充分完全统计量.

$E\frac{T}{n}=E\overline{X}=EX=\int_0^{+\infty}\frac{x}{\theta}e^{-\frac{x}{\theta}}dx=\theta$,则$h(T)=\frac{T}{n}$是$\theta$的无偏估计.

根据$L-S$定理可知，$h(T)$是θ的$UMVUE$.

指数分布属于指数族，故$C-R$正则条件都满足，

$Fisher$信息函数$I(\theta)=E\left[\dfrac{\partial \log f(X,\theta)}{\partial\theta}\right]^2=E\left(-\dfrac{1}{\theta}+\dfrac{X}{\theta^2}\right)^2=\dfrac{1}{\theta^2}$

$C-R$下界$=\dfrac{1}{nI(\theta)}=\dfrac{\theta^2}{n}$,有$D(h(T))=\dfrac{\theta^2}{n}$,

因此θ的$UMVUE$的方差与$C-R$不等式的下界相等，故是有效估计.

41、设$X_1,\cdots,X_n$ 是来自伽马分布族$\{\Gamma(\alpha,\lambda):\alpha$已知，$\lambda>0\}$的简单随机样本.

试证$\overline{X}/\alpha$是$g(\lambda)=1/\lambda$的有效估计.

解：

伽马分布属于指数族，故满足$C-R$正则条件

$I(\lambda)=E\left[\dfrac{\partial \log f(X,\lambda)}{\partial\lambda}\right]^2=E\left(\dfrac{\alpha}{\lambda}-X\right)^2=\dfrac{\alpha^2}{\lambda^2}-\dfrac{2\alpha}{\lambda}EX+EX^2=\dfrac{\alpha}{\lambda^2}$

故$C-R$下界为$\dfrac{[g'(\theta)]^2}{nI(\theta)}=\dfrac{\left(-\frac{1}{\lambda^2}\right)^2}{n\cdot\frac{\alpha}{\lambda^2}}=\dfrac{1}{n\alpha\lambda^2}$,

而$D\left(\dfrac{\overline{X}}{\alpha}\right)=\dfrac{DX}{n\alpha^2}=\dfrac{1}{n\alpha\lambda^2}$,即$\overline{X}/\alpha$的方差达到了$C-R$下界，是有效估计.

42、设$X_1,\cdots,X_n$ $i.i.d.\sim N(0,\sigma^2),\sigma^2>0$为未知参数；

（1）求σ^2的无偏估计方差的$C-R$下界；

（2）求σ^2的一致最小方差无偏估计及它的效率.

解：

(1)$f(x,\sigma^2)=\dfrac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{x^2}{2\sigma^2}},\log f(x,\sigma^2)=-\ln\sqrt{2\pi}-\dfrac{1}{2}\ln\sigma^2-\dfrac{x^2}{2\sigma^2}$

$I(\sigma^2)=E\left[\dfrac{\partial \log f(X,\sigma^2)}{\partial\sigma^2}\right]^2=E\left(-\dfrac{1}{2\sigma^2}+\dfrac{X^2}{2\sigma^4}\right)^2=\dfrac{1}{4\sigma^8}E(X^2-\sigma^2)^2$

$=\dfrac{1}{4\sigma^8}D(X^2)=\dfrac{1}{2\sigma^4}\left(\dfrac{X^2}{\sigma^2}\sim\chi_1^2,D\left(\dfrac{X^2}{\sigma^2}\right)=2,D(X^2)=2\sigma^4\right)$

故$C-R$下界为$\dfrac{[g'(\theta)]^2}{nI(\theta)}=\dfrac{2\sigma^4}{n}$.

(2)易知$T(\boldsymbol{X})=\sum\limits_{i=1}^{n}X_i$是充分完全统计量.，令$h(T)=\dfrac{T}{n},E(h(T))=\sigma^2$,

则由$L-S$定理可知$h(T)$是σ^2的$UMVUE$.

$D(h(T))=\dfrac{n\cdot 2\sigma^4}{n^2}=\dfrac{2\sigma^4}{n},e_{\hat{g}}(\theta)=\dfrac{[g'(\theta)]^2/nI(\theta)}{D(h(T))}=1$,即效率为1.

43、设$X_1,\cdots,X_n$ $i.i.d.\sim N(a,\sigma^2),a$已知，证明$\dfrac{1}{n}\sqrt{\dfrac{\pi}{2}}\sum\limits_{i=1}^{n}|X_i-a|$为$\sigma$的无偏估计量，

且有效率为$1/(\pi-2)$.

解：

令$Y=X-a$,则$Y\sim N(0,\sigma^2)$,$E\left(\frac{1}{n}\sqrt{\frac{\pi}{2}}\sum_{i=1}^{n}|X_i-a|\right)=\sqrt{\frac{\pi}{2}}E|Y|$

$$E|Y|=\int_{-\infty}^{+\infty}\frac{|y|}{\sqrt{2\pi}\,\sigma}e^{-\frac{y^2}{2\sigma^2}}dy=2\int_0^{+\infty}\frac{y}{\sqrt{2\pi}\,\sigma}e^{-\frac{y^2}{2\sigma^2}}dy=2\int_0^{+\infty}\frac{\sigma}{\sqrt{2\pi}}de^{-\frac{y^2}{2\sigma^2}}=\sqrt{\frac{2}{\pi}}\sigma$$

则$E\left(\frac{1}{n}\sqrt{\frac{\pi}{2}}\sum_{i=1}^{n}|X_i-a|\right)=\sqrt{\frac{\pi}{2}}E|Y|=\sigma$，即$\frac{1}{n}\sqrt{\frac{\pi}{2}}\sum_{i=1}^{n}|X_i-a|$是$\sigma$的无偏估计.

$$I(\sigma)=E\left[\frac{\partial\log f(X,\sigma)}{\partial\sigma}\right]^2=E\left[\frac{\partial\left(-\ln\sqrt{2\pi}-\ln\sigma-\frac{(X-a)^2}{2\sigma^2}\right)}{\partial\sigma}\right]^2$$

$$=E\left(\frac{(X-a)^2-\sigma^2}{\sigma^3}\right)^2=\frac{1}{\sigma^6}D[(X-a)^2]=\frac{2}{\sigma^2},$$

故$C-R$下界为$\frac{[g'(\theta)]^2}{nI(\theta)}=\frac{\sigma^2}{2n}$,

$$D\left(\frac{1}{n}\sqrt{\frac{\pi}{2}}\sum_{i=1}^{n}|X_i-a|\right)=\frac{\pi}{2n}D(|Y|)=\frac{\pi}{2\pi}[EY^2-E^2|Y|]=\frac{\pi-2}{2n}\sigma^2,$$

$e_{\hat{g}}(\theta)=\frac{[g'(\theta)]^2/nI(\theta)}{D(\hat{g}(\theta))}=\frac{\sigma^2}{2n}\Big/\frac{\pi-2}{2n}\sigma^2=\frac{1}{\pi-2}$,即效率为$\frac{1}{\pi-2}$.

44、设$X_1,\cdots,X_n\ i.i.d.\sim N(0,\sigma^2)$,证明

$$\hat{\sigma}=\frac{\Gamma(n/2)}{\sqrt{2}\,\Gamma((n+1)/2)}\left(\sum_{i=1}^{n}X_i^2\right)^{1/2}$$

是σ的$UMVUE$,试求其效率.

解：

易知$T(\boldsymbol{X})=\sum_{i=1}^{n}X_i^2$为充分完全统计量.

$\frac{T}{\sigma^2}\sim\chi^2_{(n)}$,则$\varphi(t)=\frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}\left(\frac{t}{\sigma^2}\right)^{\frac{n}{2}-1}e^{-\frac{t}{2\sigma^2}}$,设$\delta(t)$为$\sigma$的无偏估计，则有

$$\int_0^{\infty}\frac{\delta(t)}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}\left(\frac{t}{\sigma^2}\right)^{\frac{n}{2}-1}e^{-\frac{t}{2\sigma^2}}dt=\sigma\Rightarrow\int_0^{\infty}\frac{\delta(t)}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)\sqrt{t}}\left(\frac{t}{\sigma^2}\right)^{\frac{n+1}{2}-1}e^{-\frac{t}{2\sigma^2}}dt=1$$

$$\Rightarrow\int_0^{\infty}\frac{2^{\frac{1}{2}}\Gamma\left(\frac{n+1}{2}\right)\delta(t)}{\Gamma\left(\frac{n}{2}\right)\sqrt{t}}\cdot\frac{1}{2^{\frac{n+1}{2}}\Gamma\left(\frac{n+1}{2}\right)}\left(\frac{t}{\sigma^2}\right)^{\frac{n+1}{2}-1}e^{-\frac{t}{2\sigma^2}}dt=1$$

即$\delta(t)=\frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{2}\,\Gamma\left(\frac{n+1}{2}\right)}\sqrt{t}=\frac{\Gamma(n/2)}{\sqrt{2}\,\Gamma((n+1)/2)}\left(\sum_{i=1}^{n}X_i^2\right)^{1/2}$

故$\delta(t)$是σ的无偏估计，由$L-S$定理可知，$\delta(t)$是σ的$UMVUE$.

由42题可知$C-R$下界为$\dfrac{[g'(\theta)]^2}{nI(\theta)}=\dfrac{2\sigma^4}{n}$,

$D(\delta(t))=E[\delta(t)]^2-\sigma^2$

$$E[\delta(t)]^2=\int_0^{\infty}\frac{\delta^2(t)}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}\left(\frac{t}{\sigma^2}\right)^{\frac{n}{2}-1}e^{-\frac{t}{2\sigma^2}}dt=\int_0^{\infty}\frac{t\Gamma(n/2)}{2^{\frac{n}{2}+1}\Gamma^2\left(\frac{n+1}{2}\right)}\left(\frac{t}{\sigma^2}\right)^{\frac{n}{2}-1}e^{-\frac{t}{2\sigma^2}}dt$$

$$=\frac{\sigma^2\Gamma(n/2)\Gamma\left(\frac{n}{2}+1\right)}{\Gamma^2\left(\frac{n+1}{2}\right)}\int_0^{\infty}\frac{1}{2^{\frac{n}{2}+1}\Gamma\left(\frac{n}{2}+1\right)}\left(\frac{t}{\sigma^2}\right)^{\frac{n+2}{2}-1}e^{-\frac{t}{2\sigma^2}}dt=\frac{\sigma^2\Gamma(n/2)\Gamma\left(\frac{n}{2}+1\right)}{\Gamma^2\left(\frac{n+1}{2}\right)},$$

$$D(\delta(t))=\left(\frac{\Gamma(n/2)\Gamma\left(\frac{n}{2}+1\right)}{\Gamma^2\left(\frac{n+1}{2}\right)}-1\right)\sigma^2=\alpha\sigma^2,$$

则效率$e_{\hat{g}}(\theta)=\dfrac{[g'(\theta)]^2/nI(\theta)}{D(h(T))}=\dfrac{2}{n\alpha}\sigma^2$

45、设$X_1,\cdots,X_n$ 自下列总体中抽取的简单样本，

$$f(x,\theta)=\begin{cases}1,\ \theta-\dfrac{1}{2}\leqslant x\leqslant\theta+\dfrac{1}{2},\\ 0,\ \text{其他},\end{cases}\quad -\infty<\theta<+\infty.$$

证明样本均值$\overline{X}$及$\dfrac{1}{2}\left(\max\limits_{1\leqslant i\leqslant n}X_i+\min\limits_{1\leqslant i\leqslant n}X_i\right)$都是$\theta$的无偏估计，问何者更有效?

解：

$EX=\theta$,则$E\overline{X}=\theta$,即$\overline{X}$是θ的无偏估计，

令$Y=X+\dfrac{1}{2}-\theta$,则$Y\sim U(0,1),Y_{(1)}\sim Be(1,n),Y_{(n)}\sim Be(n,1)$,

故$EY_{(1)}=\dfrac{1}{1+n},EY_{(n)}=\dfrac{n}{1+n},$,则$EY_{(1)}+EY_{(n)}=1$,

即$EX_{(1)}+EX_{(n)}+1-2\theta=1$,即$\dfrac{1}{2}(EX_{(1)}+EX_{(n)})=\theta$

即$\overline{X}$及$\dfrac{1}{2}\left(\max\limits_{1\leqslant i\leqslant n}X_i+\min\limits_{1\leqslant i\leqslant n}X_i\right)$都是$\theta$的无偏估计.

$$D\left(\overline{X}\right)=\frac{1}{n}D(X)=\frac{1}{12n},D\left(\frac{1}{2}\left(\max_{1\leqslant i\leqslant n}X_i+\min_{1\leqslant i\leqslant n}X_i\right)\right)=\frac{1}{4}D(Y_{(1)}+Y_{(n)}-1+2\theta)$$

$$=\frac{1}{4}D(Y_{(1)})+\frac{1}{4}D(Y_{(n)})+\frac{1}{4}Cov(Y_{(1)},Y_{(n)})=\frac{n}{2(1+n)^2(n+2)}+\frac{1}{2}Cov(Y_{(1)},Y_{(n)})$$

$$R_n=Y_{(n)}-Y_{(1)}\sim Be(n-1,2),\text{则}DR_n=\frac{2(n-1)}{(n+1)^2(n+2)}=DY_{(n)}+DY_{(1)}-2Cov(Y_{(1)},Y_{(n)})$$

则$Cov(Y_{(1)},Y_{(n)})=\dfrac{1}{(n+1)^2(n+2)}$

则$D\left(\frac{1}{2}\left(\max\limits_{1\leqslant i\leqslant n} X_i+\min\limits_{1\leqslant i\leqslant n} X_i\right)\right)=\frac{1}{2(n+1)(n+2)}<\frac{1}{12n}\ (n>1$时恒成立)

则$\frac{1}{2}\left(\max\limits_{1\leqslant i\leqslant n} X_i+\min\limits_{1\leqslant i\leqslant n} X_i\right)$更有效.

46、设$X_1,X_2,X_3\ i.i.d.$ 服从均匀分布$U(0,\theta)$,试证$\frac{4}{3}\max\limits_{1\leqslant i\leqslant 3} X_i$及$4\min\limits_{1\leqslant i\leqslant 3} X_i$都是$\theta$的无偏估计量，哪个更有效?

解：

令$Y=\frac{X}{\theta}$,则$Y_{(1)}\sim Be(1,3),Y_{(n)}\sim Be(3,1)$,因此$EY_{(1)}=\frac{1}{4},EY_{(n)}=\frac{3}{4}$

因此$EX_{(1)}=\frac{1}{4}\theta,EX_{(n)}=\frac{3}{4}\theta,E\left(4\min\limits_{1\leqslant i\leqslant 3} X_i\right)=\theta,E\left(\frac{4}{3}\max\limits_{1\leqslant i\leqslant 3} X_i\right)=\theta,$

故$\frac{4}{3}\max\limits_{1\leqslant i\leqslant 3} X_i$和$4\min\limits_{1\leqslant i\leqslant 3} X_i$都是$\theta$的无偏估计.

$D(4X_{(1)})=16\theta^2 DY_{(1)}=\frac{3}{5}\theta^2,D\left(\frac{4}{3}X_{(3)}\right)=\frac{16}{9}\theta^2 DY_{(n)}=\frac{1}{15}\theta^2,$

则$D(4X_{(1)})>D\left(\frac{4}{3}X_{(3)}\right)$,因此$\frac{4}{3}\max\limits_{1\leqslant i\leqslant 3} X_i$更有效.

47、设$\hat{\theta}_1$及$\hat{\theta}_2$是参数θ的两个独立的无偏估计量，且$\hat{\theta}_1$的方差是$\hat{\theta}_2$的方差的2倍，试确定常数c_1及c_2使得$c_1\hat{\theta}_1+c_2\hat{\theta}_2$为参数$\theta$的无偏估计量，并且在所有这样的线性估计中方差最小.

解：

$E\hat{\theta}_1=E\hat{\theta}_2=\theta,D\hat{\theta}_1=2D\hat{\theta}_2=2\sigma^2,$

$E\left(c_1\hat{\theta}_1+c_2\hat{\theta}_2\right)=(c_1+c_2)\theta=\theta,$故$c_1+c_2=1$，

$D\left(c_1\hat{\theta}_1+c_2\hat{\theta}_2\right)=2c_1^2\sigma^2+c_2^2\sigma^2=(3c_1^2-2c_1+1)\sigma^2\geqslant\frac{2}{3}$

当$c_1=\frac{1}{3},c_2=\frac{2}{3}$时成立，即$\frac{1}{3}\hat{\theta}_1+\frac{2}{3}\hat{\theta}_2$在所有线性估计中方差最小.

48、设总体X的数学期望为a，$\hat{a}_1$及$\hat{a}_2$分别为a的两个无偏估计量，它们的方差分别为σ_1^2和σ_2^2，相关系数为ρ，试确定常数$c_1>0,c_2>0,c_1+c_2=1$,使得$c_1\hat{a}_1+c_2\hat{a}_2$有最小方差.

解：

$D(c_1\hat{a}_1+c_2\hat{a}_2)=c_1^2D\hat{a}_1+c_1^2D\hat{a}_2+2c_1c_2\rho\sqrt{D\hat{a}_1D\hat{a}_2}=c_1^2\sigma_1^2+c_2^2\sigma_2^2+2c_1c_2\rho\sigma_1\sigma_2$

$L(\lambda)=c_1^2\sigma_1^2+c_2^2\sigma_2^2+2c_1c_2\rho\sigma_1\sigma_2+\lambda(c_1+c_2-1)$

$\frac{\partial L(\lambda)}{\partial c_1}=2c_1\sigma_1^2+2c_2\rho\sigma_1\sigma_2+\lambda=0,$

$\frac{\partial L(\lambda)}{\partial c_2}=2c_2\sigma_2^2+2c_1\rho\sigma_1\sigma_2+\lambda=0$

$\frac{\partial L(\lambda)}{\partial \lambda}=c_1+c_2-1=0$

解得$c_1=\dfrac{\sigma_2^2-\rho\sigma_1\sigma_2}{\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2}$，$c_2=\dfrac{\sigma_1^2-\rho\sigma_1\sigma_2}{\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2}$

49、设$T_1,\cdots,T_m$为$g(\theta)$的m个独立线性无偏估计，$D(T_i)=\sigma^2<\infty,i=1,2,\cdots,m.$ 则$\overline{T}=\sum\limits_{i=1}^{m}T_i/m$为$T_1,\cdots,T_m$的线性组合类中$g(\theta)$的最小方差无偏估计（所谓线性无偏估计量，即$T_i$为$X_1,\cdots,X_n$的线性函数且$E(T_i)=g(\theta)$）.

解：

设$T_1,\cdots,T_m$无偏线性组合为$T=c_1T_1+c_2T_2+\cdots+c_mT_m$,

且$ET=(c_1+c_2+\cdots+c_m)g(\theta),(c_1+c_2+\cdots+c_m=1)$

$DT=(c_1^2+c_2^2+\cdots+c_m^2)\sigma^2,$

$L(\lambda)=c_1^2+c_2^2+\cdots+c_m^2+\lambda(c_1+c_2+\cdots+c_m-1)$

$\Rightarrow c_1=c_2=\cdots=c_m=\dfrac{1}{m},$

因此$\overline{T}=\sum\limits_{i=1}^{m}T_i/m$为$T_1,\cdots,T_m$的线性组合类中$g(\theta)$的最小方差无偏估计

50、设概率密度函数$f(x)$的核估计$f_n(x)$的定义由式(3.6.2)给出，在定理3.6.2的条件下证明$f_n(x)$的均方相合性，即证明$\lim\limits_{n\to\infty}E|f_n(x)-f(x)|^2=0.$

解：

$K(\cdot)$满足条件$(a)K(\cdot)$有界；$(b)\displaystyle\int_{-\infty}^{+\infty}|K(u)|du<\infty,$

$(c)\,|u|\,|K(u)|\to0,$当$|u|\to\infty;,f(\cdot)$在x处连续，且$\lim\limits_{n\to\infty}h_n=0,\lim\limits_{n\to\infty}nh_n=\infty.$

先证$\lim\limits_{n\to\infty}E|f_n(x)|=f(x)$:

$$Ef_n(x)=\frac{1}{nh_n}E\sum_{i=1}^{n}K\left(\frac{x-X_i}{h_n}\right)=\frac{1}{h_n}EK\left(\frac{x-X_1}{h_n}\right)$$

$$=\frac{1}{h_n}\int_{-\infty}^{+\infty}K\left(\frac{x-y}{h_n}\right)f(y)dy\overset{u=\frac{x-y}{h_n}}{=\!=\!=}\int_{-\infty}^{+\infty}K(u)f(x-h_ny)du,\text{由}K(\cdot)\text{有界可知}$$

$$|K(u)f(x-h_ny)|\leqslant c|K(u)|,\int_{-\infty}^{+\infty}|K(u)|du<\infty,$$

$$\lim_{n\to\infty}(E|f_n(x)|-f(x))=\lim_{n\to\infty}\int_{-\infty}^{+\infty}K(u)[f(x-h_ny)-f(x)]du$$

$$=\int_{-\infty}^{+\infty}K(u)\lim_{n\to\infty}[f(x-h_ny)-f(x)]du=0,\text{因此}\lim_{n\to\infty}E|f_n(x)|=f(x)$$

下证$\lim\limits_{n\to\infty}E|f_n(x)-f(x)|^2=0$

$$E|f_n(x)-f(x)|^2=E|f_n(x)-Ef_n(x)+Ef_n(x)-f(x)|^2$$

$$=Df_n(x)+E(Ef_n(x)-f(x))^2$$

$$Df_n(x)=D\left(\frac{1}{nh_n}\sum_{i=1}^{n}K\left(\frac{x-X_i}{h_n}\right)\right)=\frac{1}{nh_n^2}D\left(K\left(\frac{x-X_1}{h_n}\right)\right)$$

$$=\frac{1}{nh_n^2}\left(EK^2\left(\frac{x-X_1}{h_n}\right)-\left(EK\left(\frac{x-X_1}{h_n}\right)\right)^2\right)$$

$$=\frac{1}{nh_n^2}\left(\int_{-\infty}^{+\infty}K^2\left(\frac{x-y}{h_n}\right)f(y)dy-\left(\int_{-\infty}^{+\infty}K\left(\frac{x-y}{h_n}\right)f(y)dy\right)^2\right)$$

$$\leqslant\frac{1}{nh_n^2}\int_{-\infty}^{+\infty}K^2\left(\frac{x-y}{h_n}\right)f(y)dy\xlongequal{u=\frac{x-y}{h_n}}\frac{1}{nh_n}\int_{-\infty}^{+\infty}K^2(u)f(x-h_nu)du$$

由$\lim\limits_{|u|\to\infty}|u|\,|K(u)|\to 0$，$K(u)\sim\dfrac{1}{u^{1+\varepsilon}},\varepsilon>0$,则可知$\displaystyle\int_{-\infty}^{+\infty}K^2(u)f(x-h_nu)du$有界

又$\lim\limits_{n\to\infty}nh_n=\infty$则$\lim\limits_{n\to\infty}\dfrac{1}{nh_n}\displaystyle\int_{-\infty}^{+\infty}K^2(u)f(x-h_nu)du=0$,

故$\lim\limits_{n\to\infty}E|f_n(x)-f(x)|^2=0$得证.

51、设概率密度函数$f(x)$的核估计$f_n(x)$的定义由式(3.6.2)给出，在定理3.6.2的条件下证明$f_n(x)$的弱相合性，即证明对$\forall\varepsilon>0$有$\lim\limits_{n\to\infty}P(|f_n(x)-f(x)|\geqslant\varepsilon)=0$.

解：

由$Morkov$不等式，$\forall\varepsilon>0,P(|f_n(x)-f(x)|\geqslant\varepsilon)\leqslant\dfrac{E|f_n(x)-f(x)|^2}{\varepsilon^2}$

由50题可知$\lim\limits_{n\to\infty}E|f_n(x)-f(x)|^2=0$，

故$\lim\limits_{n\to\infty}P(|f_n(x)-f(x)|\geqslant\varepsilon)=0$，故$f_n(x)$的弱相合性成立.

习题 4 区间估计

1、对物体某指标进行5次测量，得其数据为$4.781, 4.795, 4.769, 4.792, 4.779$.设指标值服从$N(\mu,\sigma^2)$,其中$\sigma=0.01$，试求物体该指标平均值的置信系数为$0.95$的置信区间.

解：

由样本观察值得$\overline{X}=4.7832, n=5, \sigma=0.01$,查表得$u_{0.025}=1.96, \sigma u_{\alpha/2}/\sqrt{n}=0.00877$,则该指标平均值的置信系数为$0.95$的置信区间为

$$\left[\overline{X}-\frac{\sigma}{\sqrt{n}}u_{\alpha/2}, \overline{X}+\frac{\sigma}{\sqrt{n}}u_{\alpha/2}\right]=[4.7744, 4.7920]$$

2、设在上题中σ未知，作置信系数分别为0.95和0.99的置信区间.

解：

由样本观察值得$S=0.0105$,查表得$t_4(0.025)=2.7764, t_4(0.005)=4.6041$,
故$St_4(0.025)/\sqrt{n}=0.0130, St_4(0.005)/\sqrt{n}=0.0216$,
因此该指标平均值的置信系数为0.95的置信区间为

$$\left[\overline{X}-\frac{S}{\sqrt{n}}t_{n-1}(\alpha/2), \overline{X}+\frac{S}{\sqrt{n}}t_{n-1}(\alpha/2)\right]=[4.7702, 4.7962]$$

该指标平均值的置信系数为0.99的置信区间为

$$\left[\overline{X}-\frac{S}{\sqrt{n}}t_{n-1}(\alpha/2), \overline{X}+\frac{S}{\sqrt{n}}t_{n-1}(\alpha/2)\right]=[4.7616, 4.8048]$$

3、设某种电子管的使用寿命服从正态分布，从中随机抽取15个进行试验.得样本均值为1950小时，样本标准差300小时，以95%的可靠度求整批电子管平均使用寿命的置信区间.

解：

$1-\alpha=0.95$，$n=15, \overline{X}=1950, S=300$,查表得$t_{14}(0.025)=2.1448$,
故$St_{14}(0.025)/\sqrt{n}=166.1355$,故整批电子管平均使用寿命的置信系数$0.95$的置信区间为

$$\left[\overline{X}-\frac{S}{\sqrt{n}}t_{n-1}(\alpha/2), \overline{X}+\frac{S}{\sqrt{n}}t_{n-1}(\alpha/2)\right]=[1783.8645, 2116.1355]$$

4、设$X_1, \cdots, X_n$取自正态总体$N(\mu, 16)$,为使$\left[\overline{X}-1, \overline{X}+1\right]$是$\mu$的置信系数为$0.90$的置信区间，样本容量$n$至少为多少？

解：

由题意可知μ的置信区间应为

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

则$\dfrac{\sigma}{\sqrt{n}}u_{\alpha/2}=1$，其中$\sigma=4$，$\alpha=0.1, u_{\alpha/2}=1.645$,解得$n=43.2964$,

则n至少为44.

5、设样本$X_1,\cdots,X_n$取自正态总体$N(\mu,16)$，为使得μ的置信系数为$1-\alpha$的置信区间的长度不大于给定的L,样本容量n至少应为多少？

解：

μ的置信水平为$1-\alpha$的置信区间为$\left[\overline{X}-\dfrac{4}{\sqrt{n}}u_{\alpha/2},\overline{X}+\dfrac{4}{\sqrt{n}}u_{\alpha/2}\right]$，对应的区间长度

为$\dfrac{8}{\sqrt{n}}u_{\alpha/2}$，令$\dfrac{8}{\sqrt{n}}u_{\alpha/2}\leqslant L$，得$n\geqslant u_{\alpha/2}^2\dfrac{64}{L^2}$,因此样本容量$n$至少为$u_{\alpha/2}^2\dfrac{64}{L^2}$.

6、设$X_1,\cdots,X_m\ i.i.d.\sim N(a,\sigma_1^2),Y_1,\cdots,Y_n\ i.i.d.\ \sim N(ca,\sigma_2^2),c,\sigma_1^2,\sigma_2^2$已知，$a$未知，$c\neq 0$.又设样本$X_1,\cdots,X_m$，$Y_1,\cdots,Y_n$独立.

(1)找a的$UMVUE$;

(2)基于此$UMVUE$构造a的一个置信系数为$1-\alpha$的置信区间.

解：

$$(1)\,f(\boldsymbol{x},\boldsymbol{y},a)=(2\pi)^{-\frac{m+n}{2}}\sigma_1^m\sigma_2^n\exp\left\{-\frac{\sum_{i=1}^m(x_i-a)^2}{2\sigma_1^2}-\frac{\sum_{j=1}^n(y_j-ca)^2}{2\sigma_2^2}\right\}$$

$$=(2\pi)^{-\frac{m+n}{2}}\sigma_1^m\sigma_2^n\exp\left\{-\frac{\sum_{i=1}^m x_i^2-2a\sum_{i=1}^m x_i+ma^2}{2\sigma_1^2}-\frac{\sum_{j=1}^n y_j^2-2ca\sum_{j=1}^n y_j+nc^2a^2}{2\sigma_2^2}\right\}$$

$=g(T(\boldsymbol{X},\boldsymbol{Y}),a)h(\boldsymbol{X},\boldsymbol{Y})$,其中$T(\boldsymbol{X},\boldsymbol{Y})=\dfrac{1}{\sigma_1^2}\sum\limits_{i=1}^m x_i+\dfrac{c}{\sigma_2^2}\sum\limits_{j=1}^n y_j$,

$g(T(\boldsymbol{X},\boldsymbol{Y}),a)=(2\pi)^{-\frac{m+n}{2}}\sigma_1^m\sigma_2^n\exp\left\{aT(\boldsymbol{X},\boldsymbol{Y})-\dfrac{m}{2\sigma_1^2}a^2-\dfrac{nc^2}{2\sigma_2^2}a^2\right\}$,

$h(\boldsymbol{X},\boldsymbol{Y})=\exp\left\{-\dfrac{1}{2\sigma_1^2}\sum\limits_{i=1}^m x_i^2-\dfrac{1}{2\sigma_2^2}\sum\limits_{j=1}^n y_j\right\}$,故$T(\boldsymbol{X},\boldsymbol{Y})$是充分统计量，

自然参数空间$\Theta=\{a:-\infty<a<+\infty\}$作为$R_1$的子集有内点，即$T(\boldsymbol{X},\boldsymbol{Y})$是充分完全统计量

又$ET(\boldsymbol{X},\boldsymbol{Y})=\left(\dfrac{m}{\sigma_1^2}+\dfrac{nc^2}{\sigma_2^2}\right)a$,故$E\left(\dfrac{T(\boldsymbol{X},\boldsymbol{Y})}{\frac{m}{\sigma_1^2}+\frac{nc^2}{\sigma_2^2}}\right)=a$,

故$H=\dfrac{\frac{1}{\sigma_1^2}\sum\limits_{i=1}^m X_i+\frac{c}{\sigma_2^2}\sum\limits_{j=1}^n Y_j}{\frac{m}{\sigma_1^2}+\frac{nc^2}{\sigma_2^2}}$是$a$的$UMVUE$

(2)$H\sim N\left(a,\dfrac{1}{\frac{m}{\sigma_1^2}+\frac{nc^2}{\sigma_2^2}}\right)$,即$(H-a)\sqrt{\dfrac{m}{\sigma_1^2}+\dfrac{nc^2}{\sigma_2^2}}\sim N(0,1)$

$P\left(\left|(H-a)\sqrt{\dfrac{m}{\sigma_1^2}+\dfrac{nc^2}{\sigma_2^2}}\right|\leqslant u_{\alpha/2}\right)=1-\alpha$,变形得$a$的置信区间为

$$\left[H-u_{\alpha/2}\Big/\sqrt{\frac{m}{\sigma_1^2}+\frac{nc}{\sigma_2^2}},H+u_{\alpha/2}\Big/\sqrt{\frac{m}{\sigma_1^2}+\frac{nc}{\sigma_2^2}}\right]$$

7、调查10个企业的研究经费，得到的情况如下：

企业	1	2	3	4	5	6	7	8	9	10
研究经费所占比例	2	0	6	3	4	2	6	8	7	4

假设研究经费所占比例服从正态分布，求方差的区间估计，置信系数为0.95.

解：

$n-1=9,\alpha/2=0.025,1-\alpha/2=0.975$.查表得$\chi_9^2(0.025)=19.023,\chi_9^2(0.975)=2.700$

$S^2=6.4$，则方差的置信系数为0.95的置信区间为

$$\left[\frac{(n-1)S^2}{\chi_{n-1}^2(\alpha/2)},\frac{(n-1)S^2}{\chi_{n-1}^2(1-\alpha/2)}\right]=[3.0279,21.3333]$$

8、某电子产品的某一参数服从正态分布，从某天生产的产品中抽取15只，测得该参数为

$$3.0,2.7,2.9,2.8,3.1,2.6,2.5,2.8,2.4,2.9,2.7,2.6,3.2,3.0,2.8$$

试分别对该参数的均值和方差作置信系数为95%的置信区间.

解：

$n-1=14,\alpha/2=0.025,1-\alpha/2=0.975,\overline{X}=2.8,S^2=0.05$,查表得$t_{14}(0.025)=2.1448$

$\chi_{14}^2(0.025)=26.119,\chi_{14}^2(0.975)=5.629$,则均值的置信系数95%的置信区间为

$$\left[\overline{X}-\frac{S}{\sqrt{n}}t_{n-1}(\alpha/2),\overline{X}+\frac{S}{\sqrt{n}}t_{n-1}(\alpha/2)\right]=[2.6761,2.9238]$$

方差的置信系数95%的置信区间为

$$\left[\frac{(n-1)S^2}{\chi_{n-1}^2(\alpha/2)},\frac{(n-1)S^2}{\chi_{n-1}^2(1-\alpha/2)}\right]=[0.0268,0.1244]$$

9、设有3个同方差的正态总体，现从中各取一个样本，样本容量n和$Q^2=\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2$的值分别为

n	6	3	8
Q^2	40	20	50

试求共同方差σ^2的置信系数为0.95的单侧置信上限.

解：

$X_1\sim N(\mu_1,\sigma^2),X_2\sim N(\mu_2,\sigma^2),X_3\sim N(\mu_3,\sigma^2)$

$P\left(\frac{(n-1)S^2}{\sigma^2}\geqslant\chi_{n-1}^2(1-\alpha)\right)=1-\alpha$，即$P\left(\sigma^2\leqslant\frac{(n-1)S^2}{\chi_{n-1}^2(1-\alpha)}\right)=1-\alpha$

即σ^2的置信区间为单侧置信上限为$\frac{(n-1)S^2}{\chi_{n-1}^2(1-\alpha)}$，

对于$X_1,n-1=5,S^2=8,\chi_5^2(0.95)=1.145,\dfrac{(n-1)S^2}{\chi_{n-1}^2(1-\alpha)}=34.934$

对于$X_2,n-1=2,S^2=10,\chi_2^2(0.95)=0.103,\dfrac{(n-1)S^2}{\chi_{n-1}^2(1-\alpha)}=194.17$

对于$X_3,n-1=7,S^2=7.143,\chi_7^2(0.95)=2.167,\dfrac{(n-1)S^2}{\chi_{n-1}^2(1-\alpha)}=23.073$

故共同方差σ^2的单侧置信上限为194.17.

10、设$X_1,\cdots,X_n$为取自正态总体$N(\mu,\sigma^2)$的样本.为使$\left[\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2\right]^{1/2}\Big/4$成为$\sigma$的置信系数为0.95的(单侧)置信下限，样本容量$n$至少应取多少?

解:

$$P\left(\frac{(n-1)S^2}{\sigma^2}\leqslant\chi_{n-1}^2(\alpha)\right)=1-\alpha，即P\left(\sigma^2\geqslant\frac{(n-1)S^2}{\chi_{n-1}^2(\alpha)}\right)=1-\alpha$$

即σ的置信下限为$\sqrt{\dfrac{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}{\chi_{n-1}^2(\alpha)}}=\left[\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2\right]^{1/2}\Big/4$，即$\sqrt{\chi_{n-1}^2(0.05)}=4$

得n至少为10

11、随机地从A批导线中抽取4根，从B批导线中抽取5根，测量电阻（单位：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

设测试数据分别服从$N(a_1,\sigma^2)$和$N(a_2,\sigma^2)$,并且它们相互独立，a_1，a_2，σ^2均未知，求a_1-a_2的95%的置信区间.

解:

设$X\sim N(a_1,\sigma^2),Y\sim N(a_2,\sigma^2)$,由观察值得$\overline{X}=0.14125$，$\overline{Y}=0.1392$，

$S_1^2=0.00000825$，$S_2^2=0.00000520$，$\overline{X}-\overline{Y}=0.00205$，

$$S_\omega^2=\frac{(m-1)S_1^2+(n-1)S_2^2}{n+m-2}=0.00000651,S_\omega=0.00255,$$

查表得$t_{m+n-2}(\alpha/2)=t_7(0.025)=2.3646$,算得$S_\omega t_{m+n-2}(\alpha/2)\sqrt{\dfrac{1}{n}+\dfrac{1}{m}}=0.00404$

则a_1-a_2的95%的置信区间为

$$\left[\overline{X}-\overline{Y}-S_\omega t_{m+n-2}(\alpha/2)\sqrt{\frac{1}{n}+\frac{1}{m}},\overline{X}-\overline{Y}+S_\omega t_{m+n-2}(\alpha/2)\sqrt{\frac{1}{n}+\frac{1}{m}}\right]$$

$=[-0.00199,0.00609]$

12、枪弹的速度（单位：m/s）服从正态分布，为了比较两种枪弹的速度，在相同的条件下进行速度测定.算得数据如下:

$$\text{枪弹甲：} m=110, \overline{X}=2805, S_1=120.41$$
$$\text{枪弹乙：} n=100, \overline{Y}=2680, S_2=105.00$$

试求这两种枪弹的平均速度之差的置信水平近似为95%的置信区间.

解：

由于$\sigma_1^2 \neq \sigma_2^2$，故这是$Behrens-Fisher$问题，用大样本的近似方法求置信区间.

由数据算得$\overline{X}-\overline{Y}=125, \sqrt{S_1^2/m+S_2^2/n}=15.5581$,查表得$u_{0.025}=1.96$,算得

$u_{\alpha/2}\sqrt{S_1^2/m+S_2^2/n}=30.4939$,故平均速度之差的置信水平近似为95%的置信区间为

$\left[\overline{X}-\overline{Y}-u_{\alpha/2}\sqrt{S_1^2/m+S_2^2/n}, \overline{X}-\overline{Y}+u_{\alpha/2}\sqrt{S_1^2/m+S_2^2/n}\right]=[94.5061, 155.4939]$

13、某电子产品有两种型号.为了比较它们的某项参数值，分别从这两种型号电子产品中随机抽取若干个，测量该项参数值，得如下数据：

型号甲：$10.1, 10.3, 10.4, 9.7, 9.8$；

型号乙：$12.5, 12.2, 12.1, 12.0, 11.9, 11.8, 12.8.$

假设这两种型号的电子产品的该项参数值皆服从正态分布，而且它们的方差相等.试求它们的平均参数之差的置信系数为95%的置信区间.

解：

$m=5, n=7, \overline{X}=10.06, \overline{Y}=12.1857, S_1^2=0.093, S_2^2=0.1248, \overline{X}-\overline{Y}=-2.1257$

$S_\omega^2=\dfrac{(m-1)S_1^2+(n-1)S_2^2}{n+m-2}=0.1121, S_\omega=0.3348$,查表得$t_{m+n-2}(\alpha/2)=2.2281$，

算得$S_\omega t_{m+n-2}(\alpha/2)\sqrt{\dfrac{1}{n}+\dfrac{1}{m}}=0.4368$，则置信区间为

$$\left[\overline{X}-\overline{Y}-S_\omega t_{m+n-2}(\alpha/2)\sqrt{\frac{1}{n}+\frac{1}{m}}, \overline{X}-\overline{Y}+S_\omega t_{m+n-2}(\alpha/2)\sqrt{\frac{1}{n}+\frac{1}{m}}\right]$$

$=[-2.5625, -1.6889]$

14、设$X_1,\cdots,X_m$ 是来自正态总体$N(a,\sigma_1^2)$抽取的简单随机样本，$Y_1,\cdots,Y_n$是来自正态总体$N(b,\sigma_2^2)$抽取的简单随机样本，且两组样本独立.当$\sigma_2^2/\sigma_1^2=\lambda$且$\lambda$已知时，求$b-a$的置信系数为$1-\alpha$的置信区间.

解：

因为$\overline{X}\sim N\left(a,\dfrac{\sigma_1^2}{m}\right), \overline{Y}\sim N\left(b,\dfrac{\sigma_2^2}{n}\right)$,且相互独立，故$\overline{Y}-\overline{X}\sim N\left(b-a,\dfrac{\sigma_1^2}{m}+\dfrac{\sigma_2^2}{n}\right)\sim$

$N\left(b-a,\dfrac{1}{m}\sigma_1^2+\dfrac{\lambda}{n}\sigma_2^2\right)$.令

$$\xi=\frac{\left(\frac{1}{m}+\frac{\lambda}{n}\right)^{-1/2}\sigma_1^{-1}\left(\overline{Y}-\overline{X}-b+a\right)}{\sqrt{\left[\sum_{i=1}^{m}\left(X_i-\overline{X}\right)^2+\frac{1}{\lambda}\sum_{j=1}^{n}\left(Y_j-\overline{Y}\right)^2\right]\Big/[(m+n-2)\sigma_1^2]}}$$

$$=\frac{\left(\frac{1}{m}+\frac{\lambda}{n}\right)^{-1/2}\left(\overline{Y}-\overline{X}-b+a\right)}{\sqrt{\left[\sum_{i=1}^{m}\left(X_i-\overline{X}\right)^2+\frac{1}{\lambda}\sum_{j=1}^{n}\left(Y_j-\overline{Y}\right)^2\right]\Big/(m+n-2)}}$$

故$\xi\sim t(m+n-2)$,令$S(X,Y)=\sqrt{\left[\sum_{i=1}^{m}\left(X_i-\overline{X}\right)^2+\frac{1}{\lambda}\sum_{j=1}^{n}\left(Y_j-\overline{Y}\right)^2\right]\Big/(m+n-2)}$

$M=\left(\frac{1}{m}+\frac{\lambda}{n}\right)^{-1/2}$,故$b-a$水平为$1-\alpha$的置信区间为

$$\left[\overline{Y}-\overline{X}-\frac{t_{\alpha/2}}{M}S(X,Y),\overline{Y}-\overline{X}+\frac{t_{\alpha/2}}{M}S(X,Y)\right].$$

15、有甲、乙两台机床加工同样产品，分别从它们加工的产品中抽取若干个产品，分别测得产品直径后，算得如下数据：

机床甲：$m=8,\bar{x}=30.97,S_1=36.7$

机床乙：$n=7,\bar{y}=21.99,S_2=8.1$

假设产品直径服从正态分布.试求

(1) 这两台机床加工产品的平均直径之差的置信水平近似为95%的置信区间;

(2)这两台机床加工精度（即方差）之比的置信水平为90%的置信区间.

解：

(1)$S_*^2=\frac{S_1^2}{m}+\frac{S_2^2}{n}=177.734,r=S_*^4\Big/\left[\frac{S_1^4}{m^2(m-1)}+\frac{S_2^4}{n^2(n-1)}\right]=7.77\approx 8$

枢轴变量为$T_*=\frac{\overline{X}-\overline{Y}-(b-a)}{S_*}$

平均直径之差的置信水平近似为95%的置信区间为

$\left[\bar{x}-\bar{y}-S_*t_r(\alpha/2),\bar{x}-\bar{y}+S_*t_r(\alpha/2)\right]=[-21.76,39.72]$

(2) $S_1^2/S_2^2=20.529$,查表得$F_{m-1,n-1}(\alpha/2)=F_{7,6}(0.05)=4.21$,

$F_{m-1,n-1}(1-\alpha/2)=1/F_{n-1,m-1}(\alpha/2)=1/F_{6,7}(0.05)=1/3.87$,

则σ_1^2/σ_2^2的置信系数为90%的置信区间为

$\left[\frac{S_1^2}{S_2^2}\cdot\frac{1}{F_{m-1,n-1}(\alpha/2)},\frac{S_1^2}{S_\omega}\cdot F_{n-1,m-1}(\alpha/2)\right]=[4.876,79.447]$

16、有两个化验员A,B,他们独立地对某种聚合物的含氯量用相同的方法作了10次测量，其测定值的方差S^2分别为0.5419和0.6065，设σ_A^2和σ_B^2分别为A,B所测数据总体（设为正态总体）的方差，求σ_A^2/σ_B^2的95%的置信区间.

解：

$S_1^2/S_2^2=0.893$,查表得$F_{m-1,n-1}(\alpha/2)=F_{9,9}(0.025)=4.03$,
$F_{m-1,n-1}(1-\alpha/2)=1/F_{n-1,m-1}(\alpha/2)=1/F_{9,9}(0.025)=1/4.03$,
则σ_A^2/σ_B^2的置信系数为95%的置信区间为
$$\left[\frac{S_1^2}{S_2^2}\cdot\frac{1}{F_{m-1,n-1}(\alpha/2)},\frac{S_1^2}{S_2^2}\cdot F_{n-1,m-1}(\alpha/2)\right]=[0.222,3.599]$$

17、设从两台机器所生产的滚珠轴承中分别抽取容量为10个样本，算得它们直径的标准差分别为$S_1=0.042cm$和$S_2=0.035cm$，假定轴承直径服从正态分布，试求两个总体方差比的置信系数为0.99的置信区间.

解：
$S_1^2/S_2^2=1.44$,查表得$F_{m-1,n-1}(\alpha/2)=F_{9,9}(0.005)=6.54$,
$F_{m-1,n-1}(1-\alpha/2)=1/F_{n-1,m-1}(\alpha/2)=1/F_{9,9}(0.005)=1/6.54$,
则σ_1^2/σ_2^2的置信系数为99%的置信区间为
$$\left[\frac{S_1^2}{S_2^2}\cdot\frac{1}{F_{m-1,n-1}(\alpha/2)},\frac{S_1^2}{S_\omega}\cdot F_{n-1,m-1}(\alpha/2)\right]=[0.220,9.418]$$

18、设$X_1,\cdots,X_n$取自正态总体$N(\mu,\sigma^2)$的随机样本.求$\boldsymbol{\theta}=(\mu,\sigma^2)$的置信水平为$1-\alpha$的置信域（即参数$\mu$和$\sigma^2$同时的置信集）.
解：
联合置信域即寻找区域G,使得$P\{(\mu,\sigma^2)\in G\}\geqslant 1-\alpha$
枢轴变量$\sqrt{n}\dfrac{\overline{X}-\mu}{\sigma}\sim N(0,1),\dfrac{(n-1)S^2}{\sigma^2}\sim\chi_{n-1}^2$,
即$P\left\{\left|\sqrt{n}\dfrac{\overline{X}-\mu}{\sigma}\right|\leqslant a,c_1\leqslant\dfrac{(n-1)S^2}{\sigma^2}\leqslant c_2\right\}=1-\alpha$,
其中$(2\Phi(a)-1)\left(\chi_{n-1}^2(c_2)-\chi_{n-1}^2(c_1)\right)=1-\alpha$
故联合置信域为
$$\left\{(\mu,\sigma^2):\overline{X}-\frac{a\sigma}{\sqrt{n}}\leqslant\mu\leqslant\overline{X}+\frac{a\sigma}{\sqrt{n}};\frac{(n-1)S^2}{c_2}\leqslant\sigma^2\leqslant\frac{(n-1)S^2}{c_2}\right\}$$

19、设$X_1,\cdots,X_n$是来自均匀分布$U(\theta_1,\theta_2)$的简单随机样本，此处$-\infty<\theta_1<\theta_2<\infty$,
记$X_{(1)}\leqslant X_{(2)}\leqslant\cdots\leqslant X_{(n)}$为其次序统计量.求
(1)$\theta_2-\theta_1$的置信系数为$1-\alpha$的置信区间（提示：证明$[(\theta_2-\theta_1)-(X_{(n)}-X_{(1)})]/(\theta_2-\theta_1)$分布与$\theta_1$，$\theta_2$无关）；
(2)求$[X_{(1)}+X_{(n)}-(\theta_1+\theta_2)]/(X_{(n)}-X_{(1)})$的分布；
(3)求$(\theta_1+\theta_2)/2$的置信水平为$1-\alpha$的置信区间.

解：

(1)令$Y=\dfrac{X-\theta_1}{\theta_2-\theta_1}$,则$Y\sim U(0,1)$,则$Z,W=(Y_{(1)},Y_{(n)})$的联合概率密度为

$f(z,v)=n(n-1)(w-z)^{n-2}$,可推出极差$R=Y_{(n)}-Y_{(1)}$的概率密度为

$f(r)=n(n-1)r^{n-2}(1-r)$，即$Y_{(n)}-Y_{(1)}\sim Be(n-1,2)$,

$Y_{(n)}-Y_{(1)}=\dfrac{X_{(n)}-X_{(1)}}{\theta_2-\theta_1}$，所以$P\left(Be_{1-\frac{\alpha}{2}}(n-1,2)\leqslant\dfrac{X_{(n)}-X_{(1)}}{\theta_2-\theta_1}\leqslant Be_{\frac{\alpha}{2}}(n-1,2)\right)$

即$\theta_2-\theta_1$的置信系数为$1-\alpha$的置信区间为$\left[\dfrac{X_{(n)}-X_{(1)}}{Be_{\frac{\alpha}{2}}(n-1,2)},\dfrac{X_{(n)}-X_{(1)}}{Be_{1-\frac{\alpha}{2}}(n-1,2)}\right]$

(2)令$u=[x_{(1)}+x_{(n)}-(\theta_1+\theta_2)]/(x_{(n)}-x_{(1)})$，$v=x_{(n)}-x_{(1)}$

则$x_{(1)}=\dfrac{uv-v+\theta_1+\theta_2}{2},x_{(n)}=\dfrac{uv+v+\theta_1+\theta_2}{2},J=\left\|\dfrac{\partial(x_{(1)},x_{(n)})}{\partial(u,v)}\right\|=\dfrac{v}{2}$

又$f(x_{(1)},x_{(n)})=\dfrac{n(n-1)}{(\theta_2-\theta_1)^n}(x_{(n)}-x_{(1)})^{n-2}$故$g(u,v)=\dfrac{n(n-1)}{2(\theta_2-\theta_1)^n}v^{n-1}$

$\theta_1<\dfrac{uv-v+\theta_1+\theta_2}{2}<\dfrac{uv+v+\theta_1+\theta_2}{2}<\theta_2\Rightarrow\theta_1-\theta_2<uv-v<uv+v<\theta_2-\theta_1$

显然$v>0$,故有$-1<u<1$,得出$v<\dfrac{\theta_2-\theta_1}{1+u}(u>0)$或$v<\dfrac{\theta_2-\theta_1}{1-u}(u<0)$

故$h(u)=I_{\{1>u\geqslant 0\}}\displaystyle\int_0^{\frac{\theta_2-\theta_1}{1+u}}g(u,u)dv+I_{\{-1<u<0\}}\int_0^{\frac{\theta_2-\theta_1}{1-u}}g(u,v)dv=\dfrac{n-1}{2(1+|u|)^n}$

(3)对于$c>0$,有$P(|u|\leqslant c)=\displaystyle\int_0^c\dfrac{n-1}{(1+u)^n}du=1-(1+c)^{-(n-1)}$.

取$c_0=\alpha^{-\frac{1}{1-n}}-1$,则有$P(-c_0\leqslant u\leqslant c_0)=1-\alpha$,即

$$P(-c_0\leqslant[X_{(1)}+X_{(n)}-(\theta_1+\theta_2)]/(X_{(n)}-X_{(1)})\leqslant c_0)=1-\alpha,$$

故$\theta_2-\theta_1$的置信水平为$1-\alpha$的置信区间为

$$\left[\frac{X_{(n)}+X_{(1)}}{2}-\frac{c_0(X_{(n)}-X_{(1)})}{2},\frac{X_{(n)}+X_{(1)}}{2}+\frac{c_0(X_{(n)}-X_{(1)})}{2}\right]$$

20、设$X_1,\cdots,X_m\ i.i.d.\sim U(0,\theta_1),Y_1,\cdots,Y_n\ i.i.d.\sim U(0,\theta_2)$，$\theta_1>0,\theta_2>0$皆未知,且合样本独立. 求$\theta_1/\theta_2$的一个置信系数为$1-\alpha$的置信区间（提示：令$T_1=X_{(m)}$，$T_2=Y_{(n)}$,证明$(T_2\theta_1)/(T_1\theta_2)$的分布与$\theta_1$，$\theta_2$无关）.

解：

令$T_1=X_{(m)},T_2=Y_{(n)}$,则$X=\dfrac{T_1}{\theta_1}\sim Be(m,1),Y=\dfrac{T_2}{\theta_2}\sim Be(n,1)$,即

故$T=(T_2\theta_1)/(T_1\theta_2)$的分布与$\theta_1$，$\theta_2$无关.则$0<x<1,0<y<1,y=tx,0<tx<1$,

故有$t>1$时，$0<x<\dfrac{1}{t},0<t<1$时，$0<x<1$.

则$t\geqslant1$时$g(t)=\displaystyle\int_0^{\frac{1}{t}}mx^{m-1}n(tx)^{n-1}xdx=\dfrac{mn}{m+n}t^{-(m+1)}$

$0<t<1$时，$g(t)=\int_0^1 mx^{m-1}n(tx)^{n-1}xdx=\frac{mn}{m+n}t^{n-1}$,故可得其分布函数

$$F(t)=\begin{cases}\frac{mt^n}{m+n},t\in(0,1)\\1-\frac{n}{m+n}t^{-m},t\in[1,+\infty)\end{cases}.$$

$P(t_1\leqslant T\leqslant t_2)=1-\alpha$,令$F(t_1)=\frac{\alpha}{2}$,$F(t_2)=1-\frac{\alpha}{2}$,

解得$t_1=\left(\frac{m+n}{m}\frac{\alpha}{2}\right)^{1/n}$,$t_2=\left(\frac{m+n}{n}\right)^{-1/m}$，故$\theta_1/\theta_2$的一个置信系数为$1-\alpha$的置信区间为

$$\left[\frac{X_{(m)}}{Y_{(n)}}\left(\frac{m+n}{m}\frac{\alpha}{2}\right)^{1/n},\frac{X_{(m)}}{Y_{(n)}}\left(\frac{m+n}{n}\right)^{-1/m}\right].$$

21、设$X_1,\cdots,X_n$是来自均匀分布$U(\theta-1/2,\theta+1/2)$的简单随机样本，求θ的置信系数$1-\alpha$的置信区间

解：

由习题3第45题可知$\frac{X_{(1)}+X_{(n)}}{2}$为$\theta$的无偏估计，令$Z=\frac{X_{(1)}+X_{(n)}}{2}-\theta$,

下证Z为枢轴变量，$X_{(1)}$与$X_{(n)}$不独立，联合分布密度函数为

$$P(x,y)=n(n-1)(y-x)^{n-2}\quad\left(-\frac{1}{2}+\theta\leqslant x\leqslant y\leqslant\frac{1}{2}+\theta\right)$$

令$X=(X_{(1)}-\theta)/2$,$Y=(X_{(n)}-\theta)/2$,则其联合密度函数为

$P(x,y)=2^n n(n-1)(y-x)^{n-2}\left(-\frac{1}{4}\leqslant x\leqslant y\leqslant\frac{1}{4}\right)$,与$\theta$无关,

故$Z=X+Y=\frac{X_{(1)}+X_{(n)}}{2}-\theta$也与$\theta$无关，故$Z$为枢轴量

下求Z的密度函数$f(z)=\int_{-\infty}^{+\infty}P(x,z-x)dx\quad\left(-\frac{1}{4}\leqslant x\leqslant z-x\leqslant\frac{1}{4}\right)$

（即$-\frac{1}{4}\leqslant x\leqslant\frac{1}{4}$且$z-\frac{1}{4}\leqslant x\leqslant z+\frac{1}{4}$且$x\leqslant\frac{z}{2}$）

故当$-\frac{1}{2}\leqslant z\leqslant 0$,$x\in\left[-\frac{1}{4},\frac{z}{2}\right]$时$f(z)=\int_{-\frac{1}{4}}^{\frac{z}{2}}2^n n(n-1)(z-2x)^{n-2}dx=n(1+2z)^{n-1}$

当$0\leqslant z\leqslant\frac{1}{2}$,$x\in\left[z-\frac{1}{4},\frac{z}{2}\right]$时，$f(z)=\int_{z-\frac{1}{4}}^{\frac{z}{2}}2^n n(n-1)(z-2x)^{n-2}dx=n(1-2z)^{n-1}$

综上所述$f(z)=\begin{cases}n(1-|2z|)^{n-1},-\frac{1}{2}\leqslant z\leqslant\frac{1}{2}\\0,\quad\text{其他}\end{cases}$

故设常数$c\in\left[-\frac{1}{2},0\right]d\in\left[0,\frac{1}{2}\right]$,使得

$$P\left(c\leqslant\frac{X_{(1)}+X_{(n)}}{2}-\theta\leqslant d\right)=\int_c^0 n(1+2z)^{n-1}dz+\int_0^d n(1-2z)^{n-1}dz$$

$=1-\dfrac{(1+2c)^n+(1-2d)^n}{2}=1-\alpha$,故$(1+2c)^n+(1-2d)^n=2\alpha$ (*)

则置信区间为$\left[\dfrac{X_{(1)}+X_{(n)}}{2}-d,\dfrac{X_{(1)}+X_{(n)}}{2}-c\right]$,区间长度为$d-c$

$L(c,d)=d-c+\lambda((1+2c)^n+(1-2d)^n-2\alpha)$,

$\dfrac{\partial L}{\partial c}=-1+2n\lambda(1+2c)^{n-1}=0,\ \dfrac{\partial L}{\partial d}=1-2n\lambda(1-2d)^{n-1}=0,\ \dfrac{\partial L}{\partial \lambda}=(1+2c)^n+(1-2d)^n=2\alpha$

得$c=\dfrac{\alpha^{1/n}-1}{2},d=\dfrac{1-\alpha^{1/n}}{2}$时区间长度有最小值，故$\theta$的$1-\alpha$置信系数的置信区间为

$\left[\dfrac{X_{(1)}+X_{(n)}}{2}-\dfrac{1-\alpha^{1/n}}{2},\dfrac{X_{(1)}+X_{(n)}}{2}+\dfrac{1-\alpha^{1/n}}{2}\right]$

23、设$X_1,\cdots,X_n$为具有密度$f(x,\theta)=e^{-(x-\theta)}I_{[x>\theta]}$的总体中抽取的简单随机样本，此处$-\infty<\theta<+\infty,\theta$未知.

(1)证明$X_{(1)}-\theta$的分布与θ无关，并求出此分布.此处$X_{(1)}=\min\{X_{(1)},\cdots,X_{(n)}\}$（提示：记$X_i'=X_{(i)}-\theta,i=1,2,\cdots,n$.证明$X_1',\cdots,X_n'$的分布皆与$\theta$无关，又注意到$X_{(1)}-\theta=X_{(1)}'$）.

(2)求θ的置信系数为$1-\alpha$的置信区间.

解：

(1)令$Y=X-\theta$,则$g(y)=e^{-y}I_{[y>0]}$即Y的分布与θ无关，故$Y_{(1)}=X_{(1)}-\theta$的分布也和θ无关

可得$Z=Y_{(1)}$的概率函数为$g(z)=ne^{-nz}I_{[z>0]}$.

(2)设有常数$c,d>0$,使得$P(c\leqslant X_{(1)}-\theta\leqslant d)=1-\alpha$,

$P(c\leqslant X_{(1)}-\theta\leqslant d)=\int_c^d ne^{-nz}dz=e^{-nc}-e^{-nd}$,因为$g(z)$是个单减函数，故为使区间长度最短，应使$c=0$,解得$d=\dfrac{-\ln\alpha}{n}$,故$\theta$的置信水平为$1-\alpha$的置信区间为

$\left[X_{(1)}+\dfrac{\ln\alpha}{n},X_{(1)}\right]$.

24、设$X_1,\cdots,X_n$ 是来自具有密度函数

$$f(x,\theta)=\frac{\theta}{x^2}I_{[0<\theta<x<\infty]}$$

的总体的简单随机样本，求θ的置信系数为$1-\alpha$的置信区间.

解：

$L(\boldsymbol{x},\theta)=\dfrac{\theta^n}{\prod\limits_{i=1}^n x_i^2},\ln L(\boldsymbol{x},\theta)=n\ln\theta-2\sum\limits_{i=1}^n\ln x_i,\dfrac{\partial\ln L(x,\theta)}{\partial\theta}=\dfrac{n}{\theta}>0(0<\theta<x_{(1)}<x_{(n)}<\infty)$,

故θ的MLE为$X_{(1)}$,设$T=X_{(1)}$, 则$G(t)=1-\dfrac{\theta^n}{t^n},g(t)=\dfrac{n\theta^n}{t^{n+1}},(0<\theta<t<\infty)$

令$Z=\dfrac{\theta^n}{T^n}(0<z<1)$,则$P(Z\leqslant z)=P\left(\dfrac{\theta^n}{T^n}\leqslant z\right)=P\left(\dfrac{\theta}{z^{\frac{1}{n}}}\leqslant T\right)=1-(1-z)=z$

故$Z\sim U(0,1)$,设有常数$c,d\in[0,1]$，有$P\left(c\leqslant\frac{\theta^n}{X_{(1)}^n}\leqslant d\right)=P\left(c^{\frac{1}{n}}X_{(1)}\leqslant\theta\leqslant d^{\frac{1}{n}}X_{(1)}\right)$

$=d-c=1-\alpha$,为使置信区间长度最短，应使$d^{\frac{1}{n}}-c^{\frac{1}{n}}$在$d-c=1-\alpha$的条件下最小

$f(d)=d^{\frac{1}{n}}-(d+\alpha-1)^{\frac{1}{n}},f'(d)=\frac{1}{n}\left[d^{\frac{1}{n}-1}-(d+\alpha-1)^{\frac{1}{n}-1}\right]<0,$

故d取1时$d^{\frac{1}{n}}-c^{\frac{1}{n}}$有最小值，此时$c=\alpha,\theta$的置信水平为$1-\alpha$的置信区间为

$\left[\sqrt[n]{\alpha}X_{(1)},X_{(1)}\right]$.

25、某学校计划组织一次大的交友活动，为此要了解学生对该活动的支持程度，随机地对100名学生进行了解，其中有22名支持者.若记该校学生中支持这项活动的人数比例为p，并设该校学生数足够多.

(1)求p的置信系数为0.99的置信区间;

(2)若只关心p的下限，求出置信系数为0.95的置信下限.

解：

(1)设$X_i=\begin{cases}1，第i个人支持\\0，第i个人反对\end{cases}$，则$X_i\sim b(1,p),\sum\limits_{i=1}^{n}X_i\sim b(n,p)$

根据式(4.3.12)可知$\frac{\sqrt{n}(\hat{p}-p)}{\sqrt{\hat{p}(1-\hat{p})}}\xrightarrow{\mathscr{L}}N(0,1)$,其中$\hat{p}=\frac{1}{n}\sum\limits_{i=1}^{n}X_i$,

故$P\left(\left|\frac{\sqrt{n}(\hat{p}-p)}{\sqrt{\hat{p}(1-\hat{p})}}\right|\leqslant u_{\alpha/2}\right)=1-\alpha,u_{\alpha/2}=u_{0.005}=2.575$，$\hat{p}=\frac{22}{100}=0.22$

$u_{\alpha/2}\cdot\sqrt{\hat{p}(1-\hat{p})/n}=0.1067$则$p$的置信水平为0.99的置信区间为

$\left[\hat{p}-u_{\alpha/2}\cdot\sqrt{\hat{p}(1-\hat{p})/n},\hat{p}+u_{\alpha/2}\cdot\sqrt{\hat{p}(1-\hat{p})/n}\right]=[0.1133,0.3267]$.

(2)$P\left(\frac{\sqrt{n}(\hat{p}-p)}{\sqrt{\hat{p}(1-\hat{p})}}\leqslant u_{0.05}\right)=0.95\Rightarrow P\left(p\geqslant\hat{p}-u_{0.05}\cdot\sqrt{\hat{p}(1-\hat{p})/n}\right)=0.95$

又$u_{0.05}=1.645$,故p的置信水平为0.95的置信下限为

$\hat{p}-u_{0.05}\cdot\sqrt{\hat{p}(1-\hat{p})/n}=0.152$

26、设$X\sim$二项分布$b(m,p_1),Y\sim b(n,p_2),m,n$很大，$p_1,p_2$是未知参数，且$X,Y$独立，做出基于$X$，$Y$的$p_2-p_1$的置信系数近似为$1-\alpha$的置信区间.

解：

$\sum_{i=1}^{m} X_i \sim b(m,p_1), \sum_{i=1}^{n} Y_i \sim b(n,p_2)$,令$\hat{p}_1=\frac{1}{m}\sum_{i=1}^{m} X_i, \hat{p}_2=\frac{1}{n}\sum_{i=1}^{n} Y_i$

$\frac{\hat{p}_1-p_1}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{m}}} \xrightarrow{\mathscr{L}} N(0,1), \frac{\hat{p}_2-p_2}{\sqrt{\frac{\hat{p}_2(1-\hat{p}_2)}{n}}} \xrightarrow{\mathscr{L}} N(0,1), X$，$Y$独立，故

$\frac{\hat{p}_2-\hat{p}_1-(p_2-p_1)}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{m}+\frac{\hat{p}_2(1-\hat{p}_2)}{n}}} \xrightarrow{\mathscr{L}} N(0,1)$,因此$P\left(\left|\frac{\hat{p}_2-\hat{p}_1-(p_2-p_1)}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{m}+\frac{\hat{p}_2(1-\hat{p}_2)}{n}}}\right| \leqslant u_{\alpha/2}\right)=1-\alpha$,

即p_2-p_1的置信水平为$1-\alpha$的置信区间为

$$\left[\hat{p}_2-\hat{p}_1-u_{\alpha/2}\cdot\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{m}+\frac{\hat{p}_2(1-\hat{p}_2)}{n}}, \hat{p}_2-\hat{p}_1+u_{\alpha/2}\cdot\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{m}+\frac{\hat{p}_2(1-\hat{p}_2)}{n}}\right].$$

27、设$X_1,\cdots,X_m$和$Y_1,\cdots,Y_n$分别为自具有位置参数θ_1和θ_2的$Cauchy$分布总体中抽取的简单随机样本，且样本$X_1,\cdots,X_m$，$Y_1,\cdots,Y_n$相互独立，m,n 很大，求$\theta_2-\theta_1$的置信系数为$1-\alpha$的置信区间（$Cauchy$分布密度$f(x)=1/\{\pi[1+(x-\theta)^2]\}$，$-\infty<x<+\infty$. 提示：利用样本中位数及其极限定理.）.

解：

设m_x,m_y分别为$X_1,\cdots,X_m$和 $Y_1,\cdots,Y_n$的样本中位数，m,n足够大，由式(4.3.7)可知

$\frac{m_x-\theta_1}{\sqrt{\frac{\pi^2}{4m}}} \xrightarrow{\mathscr{L}} N(0,1), \frac{m_y-\theta_2}{\sqrt{\frac{\pi^2}{4n}}} \xrightarrow{\mathscr{L}} N(0,1)$，$X$，$Y$独立，故有

$\frac{m_y-m_x-(\theta_2-\theta_2)}{\sqrt{\frac{\pi^2}{4m}+\frac{\pi^2}{4n}}} \xrightarrow{\mathscr{L}} N(0,1)$，故$P\left(\left|\frac{m_y-m_x-(\theta_2-\theta_2)}{\sqrt{\frac{\pi^2}{4m}+\frac{\pi^2}{4n}}}\right|<u_{\alpha/2}\right)=1-\alpha$

故$\theta_2-\theta_1$的置信系数$1-\alpha$的置信区间为

$$\left[m_y-m_x-u_{\alpha/2}\cdot\sqrt{\frac{\pi^2}{4m}+\frac{\pi^2}{4n}}, m_y-m_x+u_{\alpha/2}\cdot\sqrt{\frac{\pi^2}{4m}+\frac{\pi^2}{4n}}\right]$$

28、电话总机在单位时间内接到呼唤次数服从$Poisson$ 分布$P(\lambda)$.观察单位时间内呼唤次数40次，获得如下数据：

接到呼唤次数	0	1	2	3	4	5	6	7
观察次数	5	10	12	8	3	2	0	0

试求λ的置信水平近似为0.95的置信区间.

解：

设$\hat{\lambda}=\frac{1}{n}\sum_{i=1}^{n} X_i=\frac{80}{40}=22$,由式(4.3.16)可知$\frac{\sqrt{n}\left(\hat{\lambda}-\lambda\right)}{\sqrt{\hat{\lambda}}} \xrightarrow{\mathscr{L}} N(0,1)$,

$u_{\alpha/2}=u_{0.025}=1.96 . n=40, \lambda$的置信水平近似为0.95的置信区间为

$\left[\hat{\lambda}-u_{\alpha/2}\sqrt{\hat{\lambda}/n}, \hat{\lambda}+u_{\alpha/2}\sqrt{\hat{\lambda}/n}\right]=[1.56,2.44]$.

29、设$X_1,\cdots,X_m$和$Y_1,\cdots,Y_n$分别是从具有参数λ_1和λ_2的指数分布（密度函数为$f(x,\lambda_i)=\lambda_i e^{-\lambda_i x}I_{[x>0]},i=1,2$）中抽取的 $i.i.d.$ 样本，且合样本独立.试确定$\lambda_2-\lambda_1$的一个信仰系数为$1-\alpha$的信仰区.（提示：利用$2\lambda_1 m\overline{X}\sim\chi^2_{2m},2\lambda_2 n\overline{Y}\sim\chi^2_{2n}$,不必算到底，指出方法即可）

解：

设$\overline{X}$,$\overline{Y}$为样本均值.$\overline{X}$是λ_1的充分统计量，$\overline{Y}$是λ_2的充分统计量，所以基于$\overline{X}$，$\overline{Y}$考虑问题，则已知$X\sim Exp(\lambda_1)$,故$X\sim Ga(1,\lambda_1)$,又由于$Gamma$分布的可加性，故$m\overline{X}\sim Ga(m,\lambda_1)$,即$m\overline{X}$的概率密度为$p(x,\lambda_1)=\dfrac{\lambda_1^m}{\Gamma(m)}x^{m-1}e^{-\lambda_1 m},0<x$,令$e_1=2\lambda_1 m\overline{X}$,则有$e_1\sim\chi^2_{2m}$,$e_1$的分布与$\lambda_1$无关,从而有$m\overline{X}=\dfrac{e_1}{2\lambda_1}$,整理得$\lambda_1=\dfrac{e_1}{2m\overline{X}}$,故有样本观察值$X$后，可得$\overline{X}$，即可由此得出$\lambda_1$的信仰分布，同理可设$e_2\sim\chi^2_{2n}$,可得$\lambda_2=\dfrac{e_2}{2n\overline{Y}}$,即得$\lambda_2$的信仰分布，又$X,Y$相互独立，故$\lambda_2-\lambda_1=\dfrac{e_2}{2n\overline{Y}}-\dfrac{e_1}{2m\overline{X}}$，设$\dfrac{e_2}{2n\overline{Y}}-\dfrac{e_1}{2m\overline{X}}$为$F_{n,m,\chi}$分布，对于$\lambda_2-\lambda_1$有

$$P\left(F_{n,m,\chi}\left(1-\frac{\alpha}{2}\right)\leqslant\lambda_2-\lambda_1\leqslant F_{n,m,\chi}\left(\frac{\alpha}{2}\right)\right)=1-\alpha,$$

即可得$\lambda_2-\lambda_1$的一个信仰系数为$1-\alpha$的信仰区间为

$$\left[F_{n,m,\chi}\left(1-\frac{\alpha}{2}\right),F_{n,m,\chi}\left(\frac{\alpha}{2}\right)\right].$$

30、设$X_1,\cdots,X_m$ $i.i.d.\sim F_1,Y_1,\cdots,Y_n$ $i.i.d.\sim F_2$,且合样本独立，仿上题的方法求

(1)若F_1,F_2分别为$N(a,\sigma_1^2)$和$N(b,\sigma_2^2),a,b,\sigma_1^2,\sigma_2^2$皆未知，求$\sigma_2^2-\sigma_1^2$的信仰系数为$1-\alpha$的信仰区间;

(2)若F_1,F_2分别是表示均匀分布$U(0,\theta_1)$和$U(0,\theta_2)$，$\theta_1>0,\theta_2>0$ 皆未知，求$\theta_2-\theta_1$的信仰系数为$1-\alpha$的信仰区间.

解：

(1)设$Q_1^2=\sum\limits_{i=1}^{m}\left(X_i-\overline{X}\right)^2,Q_2^2=\sum\limits_{j=1}^{n}\left(Y_j-\overline{Y}\right)^2$,则有$Q_1^2/\sigma_1^2\sim\chi^2_{m-1}$，$Q_2^2/\sigma_2^2\sim\chi^2_{2n}$，

设$\xi_1\sim\chi^2_{m-1},\xi_2\sim\chi^2_{n-1}$,等价地有$Q_1^2=\sigma_1^2\xi_1$，$Q_2^2=\sigma_2^2\xi_2$,移项整理得

$\sigma_1^2=Q_1^2/\xi_1$，$\sigma_2^2=Q_2^2/\xi_2$，又X,Y相互独立故$\sigma_2^2-\sigma_1^2=Q_2^2/\xi_2-Q_1^2/\xi_1$,

设$Q_2^2/\xi_2-Q_1^2/\xi_1$服从$F_{n,m,\chi}$分布，故$\sigma_2^2-\sigma_1^2$的信仰系数为$1-\alpha$的信仰区间为

$$\left[F_{n,m,\chi}\left(1-\frac{\alpha}{2}\right),F_{n,m,\chi}\left(\frac{\alpha}{2}\right)\right].$$

(2)易知$X_{(m)}$是θ_1的充分统计量，$Y_{(n)}$是θ_2的充分统计量.

$X_{(m)}$的密度函数为$p(x,\theta_1)=\dfrac{mx^{m-1}}{\theta_1^m},0\leqslant x\leqslant\theta_1$,

令$e_1=\dfrac{X_{(m)}}{\theta_1}$,则$e_1$的密度函数为$p(t)=mt^{m-1},0\leqslant t\leqslant 1$,故$e_1$分布与$\theta_1$无关,

故可得$\theta_1=\dfrac{X_{(m)}}{e_1}$,同理设$e_2=\dfrac{Y_{(n)}}{\theta_2}$,则$e_2$的分布也与$\theta_2$无关，$\theta_2=\dfrac{Y_{(n)}}{e_2}$,则

$\theta_2-\theta_1=\dfrac{Y_{(n)}}{e_2}-\dfrac{X_{(m)}}{e_1}$设$\theta_2-\theta_1$服从$F_{n,m,e}$分布，

则$\theta_2-\theta_1$的信仰系数为$1-\alpha$的信仰区间为$\left[F_{n,m,e}\left(1-\dfrac{\alpha}{2}\right),F_{n,m,e}\left(\dfrac{\alpha}{2}\right)\right]$

31、设$X_1,\cdots,X_n$为来自指数分布（密度函数为$f(x,\lambda)=\lambda e^{-\lambda x}I_{(0,\infty)}(x)$）中抽取的简单随机样本，定出$\lambda$的信仰分布，由之求出信仰系数为$1-\alpha$的最短信仰区间（提示:利用充分统计量$\overline{X}$）.

解：

由于$\overline{X}$是λ的充分统计量，故基于$\overline{X}$考虑问题，因为$2\lambda n\overline{X}\sim\chi_{2n}^2$,设$e\sim\chi_{2n}^2$,则有

$\lambda=\dfrac{e}{2n\overline{X}}$,其中$e$的概率密度为$p(t)=\dfrac{1}{2^n\Gamma(n)}t^{n-1}e^{-\frac{t}{2}},t>0$,设有常数$c,d>0$,有

$P(c\leqslant\lambda\leqslant d)=P\left(c\leqslant\dfrac{e}{2n\overline{X}}\leqslant d\right)=P\left(2cn\overline{X}\leqslant e\leqslant 2dn\overline{X}\right)=1-\alpha,Gamma$分布当$\alpha$足够大

时，越近似于正态密度函数，$\chi_{2n}^2\sim Ga\left(n,\dfrac{1}{2}\right)$,故越靠近中心$2(n-1)$,区间长度$d-c$越小，

故$P\left(2cn\overline{X}\leqslant e\leqslant 2dn\overline{X}\right)=P\left(\chi_{2n}^2\left(1-\dfrac{\alpha}{2}\right)\leqslant e\leqslant\chi_{2n}^2\left(\dfrac{\alpha}{2}\right)\right)=1-\alpha,$

故λ的信仰系数为$1-\alpha$的最短置信区间为

$$\left[\frac{\chi_{2n}^2\left(1-\frac{\alpha}{2}\right)}{2n\overline{X}},\frac{\chi_{2n}^2\left(\frac{\alpha}{2}\right)}{2n\overline{X}}\right].$$

32、设样本与形体同上题，试求出形如$c\overline{X}$的水平（$1-\beta$，$1-\gamma$）的容忍上、下限，并利用引理4.5.1求出水平（$1-\beta$，$1-\gamma$）的容忍区间.

解：

容忍上限：$P\left(P\left(X\leqslant c\overline{X}\right)\leqslant 1-\beta\right)=P\left(1-e^{-\lambda c\overline{X}}\leqslant 1-\beta\right)=P\left(2\lambda n\overline{X}\leqslant-\dfrac{2n\ln\beta}{c}\right)\geqslant 1-\gamma$

由$2\lambda n\overline{X}\sim\chi_{2n}^2$,因此$-\dfrac{2n\ln\beta}{c}=\chi_{2n}^2(\gamma)$,即$c=-\dfrac{2n\ln\beta}{\chi_{2n}^2(\gamma)}$,容忍上限为$-\dfrac{2n\overline{X}\ln\beta}{\chi_{2n}^2(\gamma)}$

容忍下限：$P\left(P\left(X\geqslant d\overline{X}\right)\leqslant 1-\beta\right)=P\left(e^{-\lambda d\overline{X}}\leqslant 1-\beta\right)=P\left(2\lambda n\overline{X}\geqslant-\dfrac{2n\ln(1-\beta)}{d}\right)\geqslant 1-\gamma$

由$2\lambda n\overline{X}\sim\chi_{2n}^2$,因此$-\dfrac{2n\ln(1-\beta)}{d}=\chi_{2n}^2(1-\gamma)$,即$d=-\dfrac{2n\ln(1-\beta)}{\chi_{2n}^2(1-\gamma)}$,

容忍下限为$-\dfrac{2n\overline{X}\ln(1-\beta)}{\chi_{2n}^2(\gamma)}$.

由引理4.5.1可知容忍区间为：$\left[-\dfrac{2n\ln(1-\beta)}{\chi^2_{2n}\left(1-\frac{\gamma}{2}\right)},-\dfrac{2n\overline{X}\ln\beta}{\chi^2_{2n}\left(\frac{\gamma}{2}\right)}\right]$.

33、经验表明，某种型号玻璃纸的横向延伸率（单位：%）服从正态分布，现从一批玻璃纸中随机抽取8个样品，测得其横向延伸率为

$$35.5,39.5,43.5,41.5,47.5,37.5,49.5,45.5.$$

试求玻璃纸的横向延伸率的水平为（0.95,0.95）的容忍区间.

解：

$n=8,\beta=0.05,\gamma=0.05,\overline{X}=42.5,S^2=24,S=4.9$,查表得$\lambda(8,0.95,0.95)=3.73$,

$T_1=\overline{X}-\lambda S=42.5-3.73\times4.9=24.223$,

$T_2=\overline{X}+\lambda S=42.5+3.73\times4.9=60.777$.

故玻璃纸的横向延伸率的水平为（0.95,0.95）的容忍区间为[24.223,60.777].

34、某种型号的圆钢的硬度（单位：kg/mm^2）服从正态分布.现从一批圆钢中随机抽取12个样品，测量其硬度，由此算得$\overline{X}=207$，$S=65$.试求圆钢的水平为（0.90,0.95）的容忍下限.

解：

查表得$\lambda=2.21$，则容忍下限为$\overline{X}-\lambda S=207-2.21\times65=63.35$.

35、(1)为使$X_{(1)}$成为$F(x)$的水平（0.95,0.99）容忍下限，样本容量n至少应为多少？

(2)为使$X_{(n)}$成为$F(x)$的水平（0.90,0.95）容忍上限，样本容量n至少应为多少？

解：

(1)可知$F(X_{(1)})=U_{(1)}$密度函数为

$$f(y)=n(1-y)^{n-1}I_{(0,1)}(y)$$

故要使

$$P\{F(X_{(1)})\leqslant\beta\}=P(U_{(1)}\leqslant\beta)=\int_0^\beta n(1-y)^{n-1}dy\geqslant1-\gamma$$

解得$n\geqslant\dfrac{\ln\gamma}{\ln(1-\beta)}=89.78\approx90$

(2)可知$F(X_{(n)})=U_{(n)}$密度函数为

$$f(z)=nz^{n-1}I_{(0,1)}(z)$$

故要使

$$P\{F(X_{(n)})\geqslant1-\beta\}=P(U_{(n)}\geqslant1-\beta)=\int_{1-\beta}^1 nz^{n-1}dz\geqslant1-\gamma$$

解得$n\geqslant\dfrac{\ln\gamma}{\ln(1-\beta)}=89.78\approx90$

36、设$X_1,\cdots,X_n$ $i.i.d.\sim$均匀分布$U(0,\theta),\theta>0$未知，试利用统计量$\max\limits_{1\leqslant i\leqslant n}X_i$ 定出水平为（$1-\beta$，$1-\gamma$）的容忍上下限.

解：

$X_{(n)}$的密度函数为$f(y)=\dfrac{ny^{n-1}}{\theta^n},0<y<\theta$

设$aX_{(n)}$为容忍上限则有

$$P\left(P\left(X\leqslant aX_{(n)}\right)\geqslant 1-\beta\right)=P\left(aX_{(n)}\geqslant 1-\beta\right)=P\left(X_{(n)}\geqslant\frac{1-\beta}{a}\right)$$

$$=\int_{\frac{1-\beta}{a}}^{\theta}\frac{ny^{n-1}}{\theta^n}dy=1-\left(\frac{1-\beta}{a\theta}\right)^n\geqslant 1-\gamma,\text{则得}a=\frac{1-\beta}{\theta\gamma^{\frac{1}{n}}},$$

故水平为（$1-\beta$，$1-\gamma$）的容忍上限为$\dfrac{1-\beta}{\theta\gamma^{\frac{1}{n}}}X_{(n)}$.

设$bX_{(n)}$为容忍下限则有

$$P\left(P\left(X\geqslant bX_{(n)}\right)\geqslant 1-\beta\right)=P\left(1-bX_{(n)}\geqslant 1-\beta\right)=P\left(X_{(n)}\leqslant\frac{\beta}{b}\right)$$

$$=\int_0^{\frac{\beta}{b}}\frac{ny^{n-1}}{\theta^n}dy=\left(\frac{\beta}{b\theta}\right)^n\geqslant 1-\gamma,\text{解得}b=\frac{\beta}{\theta(1-\gamma)^{\frac{1}{n}}},$$

故水平为（$1-\beta$，$1-\gamma$）的容忍下限为$\dfrac{\beta}{\theta(1-\gamma)^{\frac{1}{n}}}X_{(n)}$

习题 5 参数假设检验

1、设$\boldsymbol{X}=(X_1,\cdots,X_{20})$为从两点分布总体$b(1,p)$中抽取的简单样本，对未知参数$p$的检验问题为$H_0:p=0.2\leftrightarrow H_1:p\neq 0.2$,取检验函数为

$$\phi(\boldsymbol{x})=\begin{cases}1, & \sum_{i=1}^{20}x_i\geqslant 7\text{或}\sum_{i=1}^{20}x_i\leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

(1)求此检验函数在$p=0,0.1,0.2,\cdots,0.9,1$时的功效函数并作图;

(2)求检验的水平α和$p=0.10$时犯第二类错误的概率.

解:

(1)检验函数$\beta_\varphi(\theta)=E_\theta[\varphi(\boldsymbol{X})]$,

$\because X_i\sim b(1,p),\therefore \sum_{i=1}^{20}x_i\sim b(20,p)$,

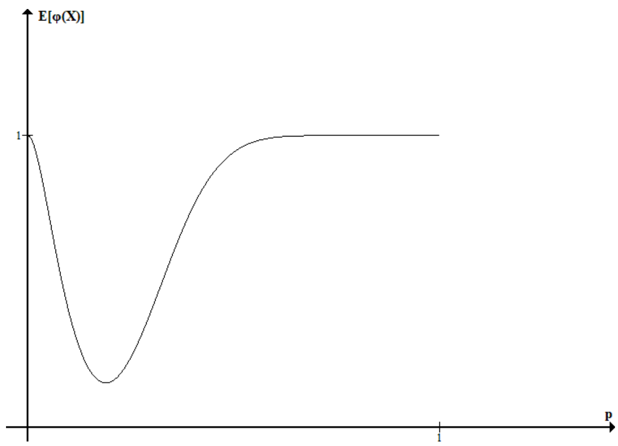
$\sum_{i=1}^{20}x_i\geqslant 7=\sum_{k=1}^{20}\binom{20}{k}p^k(1-p)^{20-k},\sum_{i=1}^{20}x_i\leqslant 1=\binom{20}{1}p(1-p)^{19}+(1-p)^{20}$,

故$E_\theta[\varphi(\boldsymbol{X})]=\sum_{k=7}^{20}\binom{20}{k}p^k(1-p)^{20-k}+\binom{20}{1}p(1-p)^{19}+(1-p)^{20}$

$p=0$时，$E_\theta[\varphi(\boldsymbol{X})]=1$,

$p=0.1$时，$E_\theta[\varphi(\boldsymbol{X})]=\sum_{k=7}^{20}\binom{20}{k}0.1^k0.9^{20-k}+\binom{20}{1}0.1(1-0.1)^{19}+(1-0.1)^{20}$

$\cdots p=1$时，$E_\theta[\varphi(\boldsymbol{X})]=1$,图像如图所示



(2)$\alpha=sup\{\beta_\varphi(p),p\in\Theta_0\}=\beta_\varphi(0.2)=E_{p=0.2}[\varphi(\boldsymbol{X})]=0.1559$,

$p=0.10$时犯第二类错误的概率为$1-\beta_\varphi(0.1)=0.6059$.

2、设$X_1,\cdots,X_n$ $i.i.d.\sim N(\mu,9)$,其中μ为未知参数，$\overline{X}$为样本均值. 设检验问题H_0:$\mu=\mu_0\leftrightarrow H_1:\mu\neq\mu_0$ 的否定域为

$$\left\{(X_1,\cdots,X_n):\left|\overline{X}-\mu_0\right|\geqslant c\right\}.$$

(1)定出常数c，使检验水平为$\alpha=0.05$;

(2)求此检验的功效函数$\beta(\mu)$;

(3)固定样本容量$n=25$，分析犯两种错误概率α和β之间的关系.

解：

(1)$\alpha=P\{$用检验φ否定了$H_0\}=E_{\mu_0}[\varphi(\boldsymbol{X})]=1\cdot P\left(\left|\overline{X}-\mu_0\right|\geqslant c\right)$

$=P\left(\frac{\left|\overline{X}-\mu_0\right|}{3/\sqrt{n}}\geqslant\frac{c\sqrt{n}}{3}\right)=0.05$,故$\frac{c\sqrt{n}}{3}=u_{0.025}$,即$c=\frac{3u_{0.025}}{\sqrt{n}}$.

(2)$\beta(\mu)=E_\mu[\varphi(\boldsymbol{X})]=P\left(\left|\overline{X}-\mu_0\right|\geqslant c\right)=$

$$P\left(\frac{\overline{X}-\mu_0-(\mu-\mu_0)}{3/\sqrt{n}}\geqslant\frac{c-(\mu-\mu_0)}{3/\sqrt{n}}\cup\frac{\overline{X}-\mu_0-(\mu-\mu_0)}{3/\sqrt{n}}\leqslant\frac{-c-(\mu-\mu_0)}{3/\sqrt{n}}\right)$$

$$=1-\Phi\left(\frac{c-(\mu-\mu_0)}{3/\sqrt{n}}\right)+\Phi\left(\frac{-c-(\mu-\mu_0)}{3/\sqrt{n}}\right)$$

(2)$n=25$时，$\alpha=\begin{cases}\beta(\mu_0),\mu=\mu_0\\0, & \mu\neq\mu_0\end{cases}$，$\beta=\begin{cases}0, & \mu=\mu_0\\1-\beta(\mu), & \mu\neq\mu_0\end{cases}$,

其中$\beta(\mu_0)=P\left(\frac{5\left|\overline{X}-\mu_0\right|}{3}\geqslant\frac{5c}{3}\right)=2-2\Phi\left(\frac{5c}{3}\right)$

$1-\beta(\mu)=\Phi\left(\frac{c-(\mu-\mu_0)}{3/5}\right)-\Phi\left(\frac{-c-(\mu-\mu_0)}{3/5}\right)$,

α越大，即c越小，则$\Phi\left(\frac{c-(\mu-\mu_0)}{3/5}\right)-\Phi\left(\frac{-c-(\mu-\mu_0)}{3/5}\right)$越小，故$\beta$越小.

3、设$X_1,\cdots,X_n$ $i.i.d.$ 服从均匀分布$U(0,\theta)$，θ为未知参数.$X_{(n)}=\max\{X_1,\cdots,X_n\}$对检验问题$H_0:\theta\leqslant\theta_0\leftrightarrow H_1:\theta>\theta_0$取检验函数为

$$\varphi(\boldsymbol{x})=\begin{cases}1, & X_{(n)}\geqslant c\\0, & \text{其他.}\end{cases}$$

(1)求检验函数$\varphi(\boldsymbol{x})$的功效函数，并证明它是θ的单调递增函数.
(2)在检验问题$H_0:\theta\leqslant 1/2\leftrightarrow H_1:\theta>1/2$中，选择什么样的$c$使检验水平恰好为$0.05$？
(3)画出$n=20$时(2)中制定的$\varphi(\boldsymbol{x})$的功效函数的粗略图形;
(4)n该多大，能使(2)中指定的检验函数$\varphi(\boldsymbol{x})$,当$\theta=3/4$功效（即功效函数在Θ_1中某一点的值）为0.98？

解：

(1)易知$X_{(n)}$的概率密度为$f(t)=\frac{nt^{n-1}}{\theta^n},0\leqslant t\leqslant\theta$,

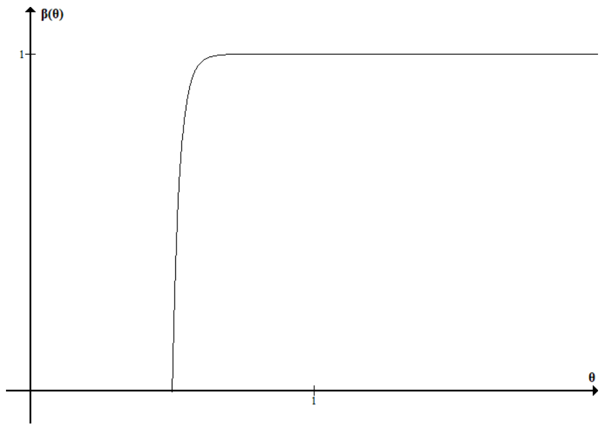
功效函数$\beta_\varphi(\theta)=E_\theta[\varphi(\boldsymbol{X})]=1\cdot P(X_{(n)}\geqslant c)=\begin{cases}1, & c<0\\\int_c^\theta\frac{nt^{n-1}}{\theta^n}dt=1-\left(\frac{c}{\theta}\right)^n, & 0\leqslant c\leqslant\theta\\0, & \theta<c\end{cases}$

$\beta_\varphi{}'(\theta)=\frac{nc^n}{\theta^{n+1}}>0$,故功效函数是$\theta$的单调递增函数.

(2)检验水平$\alpha = sup\{\beta_\varphi(\theta), \theta \leqslant 1/2\}$,又由(1)可知功效函数是$\theta$的单调递增函数

故$\alpha = \beta_\varphi(1/2) = 1-(2c)^n = 0.05$,故$c = \dfrac{\sqrt[n]{0.95}}{2}$.

(3)$n=20$时，$c = 0.4987, \beta_\varphi(\theta) = \begin{cases} 1-\dfrac{0.95}{(2\theta)^{20}}, & 0.4987 \leqslant \theta \\ 0, & \theta < 0.4987 \end{cases}$,函数图像如图所示:



(4)$\beta_\varphi(\theta) = \begin{cases} 1, & c<0 \\ 1-\dfrac{0.95}{(2\theta)^n}, & 0 \leqslant c \leqslant \theta \\ 0, & \theta < c \end{cases}$，其中$c = \dfrac{\sqrt[n]{0.95}}{2}$,$\theta = 3/4$时，

令$1-\dfrac{0.95}{(3/2)^n} = 0.98$，解得$n \approx 10$

4、设$X_1, \cdots, X_n$ 取自均匀分布$U(0,\theta)$的随机样本，其中$\theta>0$为未知参数.记$X_{(n)} = \max\{X_1, \cdots, X_n\}$.如对检验问题

$$H_0:\ \theta \geqslant 2 \leftrightarrow H_1:\ \theta < 2$$

取检验的否定域为

$$W = \{(X_1, \cdots, X_n):\ X_{(n)} \leqslant 1.5\},$$

试求此检验犯第一类错误的概率的最大值.

解:

易知$X_{(n)}$的概率密度为$f(t) = \dfrac{nt^{n-1}}{\theta^n}, 0 \leqslant t \leqslant \theta,$

则功效函数$\beta_\varphi(\theta) = E_\theta[\varphi(\boldsymbol{X})] = \begin{cases} \displaystyle\int_0^{1.5} \frac{nt^{n-1}}{\theta^n} dt = \left(\frac{1.5}{\theta}\right)^n, & 1.5 \leqslant \theta \\ 1, & 1.5 > \theta \end{cases}$

$\Theta_0: \theta \geqslant 2$，$\Theta_1: \theta < 2$，

故犯第一类错误概率$\alpha_\varphi{}^*(\theta) = \begin{cases} \beta_\varphi(\theta), & \theta \in \Theta_0 \\ 0, & \theta \in \Theta_1 \end{cases}$,

$\max \alpha_\varphi{}^*(\theta) = \max\limits_{\theta \geqslant 2} \left(\dfrac{1.5}{\theta}\right)^n = \left(\dfrac{3}{4}\right)^n$.

5、设$X_1,\cdots,X_{10}$ 为取自$Poisson\ P(\lambda)$分布的随机样本.

(1)试用直观方法求单边假设检验问题$H_0:\lambda\leqslant 0.1\leftrightarrow H_1:\lambda>0.1$ 的水平$\alpha=0.05$的检验.

(2)求此检验的功效函数$\beta(\lambda)$在$\lambda=0.05,0.2,0.3,\cdots,0.9$时的值，并画出$\beta(\lambda)$的图像.

解：易知$\overline{X}$是λ的一个优良估计，

故设$\left\{X_1,\cdots,X_n:\overline{X}-0.1>c\right\}$为水平$\alpha=0.05$检验否定域.则设常数$c>0$有

$$P\left(\overline{X}-0.1>c\right)=P\left(\sum_{i=1}^{10}X_i>10c+1\right),$$

其中$\sum_{i=1}^{10}X_i\sim P(10\lambda)$,则有功效函数

$$\beta(\lambda)=1\cdot P\left(\sum_{i=1}^{10}X_i>10c+1\right)=\sum_{k=10c+2}^{\infty}\frac{(10\lambda)^k}{k!}e^{-10\lambda}$$

$$\beta'(\lambda)=\sum_{k=10c+2}^{\infty}\frac{10(10\lambda)^{k-1}}{k!}e^{-10c}[k-10\lambda],$$

易知H_0成立时，$k>1,\lambda\leqslant 0.1$,故$\beta'(\lambda)>0$

故检验水平$\alpha=sup\{\beta(\lambda),\lambda\leqslant 0.1\}=\beta(0.1)$.

$\beta(0.1)=\sum_{k=10c+2}^{\infty}\frac{1}{k!}e^{-1}=1-\sum_{k=0}^{10c}\frac{1}{k!}e^{-1}\leqslant 0.05$($P172$、定义5.1.3)

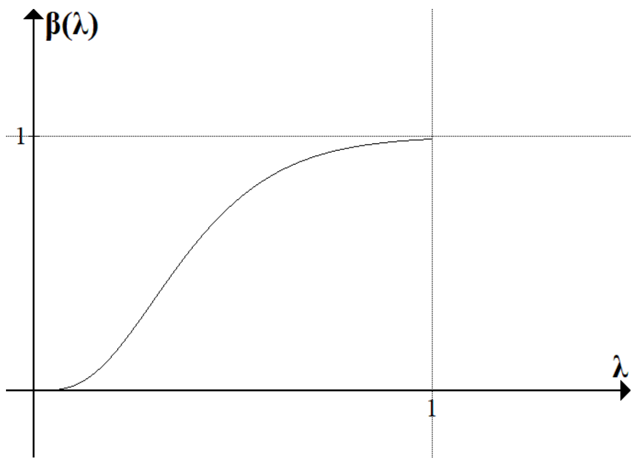
即$\sum_{k=0}^{10c+1}\frac{1}{k!}\geqslant 0.95e=2.58$,解得$c$至少为0.2，故

$D=\left\{(X_1,\cdots,X_n):\overline{X}>0.3\right\}$为水平$\alpha=0.05$的否定域确定的检验满足题意.

(2)$\beta(\lambda)=1-\sum_{k=0}^{3}\frac{(10\lambda)^k}{k!}e^{-10\lambda}$

$\beta(0.05)=1.75\times 10^{-3},\beta(0.2)=0.14,\beta(0.3)=0.35,\beta(0.4)=0.57,\beta(0.5)=0.73$

$\beta(0.6)=0.85,\beta(0.7)=0.92,\beta(0.8)=0.96,\beta(0.9)=0.98,\beta(\lambda)$如图所示：



6、设$X_1,\cdots,X_n$ 为取自两点分布$b(1,p)$的随机样本.

(1)试用直观方法求单边假设检验问题

$$H_0:p\leqslant 0.01\leftrightarrow H_1:p>0.01$$

的水平$\alpha=0.05$的检验.

(2)若上述检验的否定域为$\left\{(X_1,\cdots,X_n):\sum_{i=1}^{n}X_i\geqslant 1\right\}$,要求这个检验在$p=0.08$时犯第二类错误的概率不超过$0.10$，样本容量$n$应为多大?

解：易知$\overline{X}$是p的一个优良估计

故设$\left\{X_1,\cdots,X_n:\overline{X}>c/n\right\}$为水平$\alpha=0.05$检验否定域.则设常数$c>0$有

$$P\left(n\overline{X}>c\right)=P\left(\sum_{i=1}^{n}X_i>c\right)$$

其中$\sum_{i=1}^{n}X_i\sim b(n,p)$,则有功效函数

$$\beta(p)=1\cdot P\left(n\sum_{i=1}^{n}X_i>c\right)=\sum_{k=c+1}^{\infty}\binom{n}{k}p^k(1-p)^{n-k}=1-\sum_{k=0}^{c}\binom{n}{k}p^k(1-p)^{n-k}$$

$$\beta'(p)=-\sum_{k=0}^{c}\binom{n}{k}kp^{k-1}(1-p)^{n-k}+\sum_{k=0}^{c}\binom{n}{k}(n-k)p^k(1-p)^{n-k-1}$$

$$=-\sum_{k=1}^{c}\binom{n-1}{k-1}np^{k-1}(1-p)^{n-k}+\sum_{k=0}^{c}\binom{n-1}{k}np^k(1-p)^{n-k-1}$$

$$=-\sum_{k=0}^{c-1}\binom{n-1}{k}np^k(1-p)^{n-k-1}+\sum_{k=0}^{c}\binom{n-1}{k}np^k(1-p)^{n-k-1}$$

$=\binom{n-1}{c}np^c(1-p)^{n-c-1}>0$,即$\beta(p)$是$p$的单调增函数.

故检验水平$\alpha=sup\{\beta(p),p\leqslant 0.01\}=\beta(0.01)=\sum_{k=c+1}^{\infty}\binom{n}{k}0.01^k0.99^{n-k}\leqslant 0.05$

设有最小解c_0,则$D=\left\{(X_1,\cdots,X_n):\overline{X}>c_0/n\right\}$为水平$\alpha=0.05$的否定域的检验符合题意.

(2)犯第二类错误的概率$\beta_{\varphi}{}^*(p)=1-\beta(0.08)=\sum_{k=0}^{1}\binom{n}{k}0.08^k0.92^{n-k}\leqslant 0.1$

即$0.92^n+0.08n0.92^{n-1}\leqslant 0.1$,

$n=47$时，$0.92^n+0.08n0.92^{n-1}=0.101$

$n=48$时，$0.92^n+0.08n0.92^{n-1}=0.0945$

故n至少为48.

7、假定某切割机在正常工作时，切割每段金属棒的长度（单位：cm）服从正态分布$N(10.5,0.15^2)$.今从一批产品中随机抽取15段进行测量，其结果如下

$$10.4,10.6,10.1,10.4,10.5,10.3,10.3,10.2,$$
$$10.9,10.6,10.8,10.5,10.7,10.2,10.7.$$

试问该机工作是否正常（$\alpha=0.05$）?

解:

检验问题为

$$H_0:\mu=10.5\leftrightarrow H_1:\mu\neq 10.5.$$

故否定域为

$$D=\left\{(X_1,\cdots,X_n):\left|\frac{\sqrt{n}\left(\overline{X}-\mu_0\right)}{\sigma}\right|>u_{\alpha/2}\right\}$$

其中$n=15,\overline{X}=10.48,\mu_0=10.5,\sigma=0.15,u_{0.025}=1.96$

$$|U|=\left|\frac{\sqrt{15}\times(10.48-10.5)}{0.15}\right|=0.516<1.96$$

所以在显著性水平$\alpha=0.05$下认为没有足够理由说明该机工作不正常.

8、设某产品的指标服从正态分布，已知它的标准差$\sigma=150$，今抽了一个样本容量为26的样本，计算得样本均值为1637，问在5%的显著水平下，能否认为这批产品指标的期望值μ不大于1600.

解：

检验问题为：

$$H_0:\mu\leqslant1600\leftrightarrow H_1:\mu>1600$$

否定域为：

$$D=\left\{(X_1,\cdots,X_n):U=\frac{\sqrt{n}\left(\overline{X}-\mu_0\right)}{\sigma}>u_\alpha\right\}$$

其中$n=26,\overline{X}=1637,\mu_0=1600,\sigma=150,u_{0.05}=1.645$

$$U=\frac{\sqrt{n}\left(\overline{X}-\mu_0\right)}{\sigma}=1.258<1.645,$$

所以在5%的显著水平下有足够理由认为这批产品指标的期望值μ不大于1600.

9、某纺织厂在正常运行条件下，平局每台布机每小时经纱端头为0.973根，各台布机断头数的方差为0.162根.该厂进行工艺改革，减少经纱上浆率，在200台布机上进行试验，结果平均每台每小时经纱断头为0.994根.假定每台布机断头数服从正态分布，问新工艺上浆率能否推广（$\alpha=0.05$）？

解：

检验问题为

$$H_0:\mu\leqslant0.973\leftrightarrow H_1:\mu>0.973$$

否定域为

$$D=\left\{(X_1,\cdots,X_n)：U=\frac{\sqrt{n}\left(\overline{X}-\mu_0\right)}{\sigma}>u_\alpha\right\}$$

$n=200,\overline{X}=0.994,\mu_0=0.973,\sigma=\sqrt{0.162},u_\alpha=u_{0.05}=1.645$

$$U=\frac{\sqrt{200}\,(0.994-0.973)}{\sqrt{0.162}}=0.737865<1.645$$

故在5%的显著性水平下认为新工艺上浆率能推广.

10、根据长期经验和资料分析，某砖瓦厂生产的砖的抗断强度X服从方差为σ^2的正态分布，今从该厂所生产的一批砖中随机抽取6块，测得抗断强度（单位：kg/cm^2）如下：

$$32.56,29.66,31.64,30.00,31.87,31.03$$

问这一批砖的平均抗断强度可否认为是32.50 kg/cm^2？（$\alpha=0.05$）

解：

检验问题为

$$H_0:\mu=32.50 \leftrightarrow H_1:\mu\neq 32.50$$

否定域为

$$D=\left\{(X_1,\cdots,X_n):T=\left|\frac{\sqrt{n}\left(\overline{X}-\mu_0\right)}{S}\right|>t_{\alpha/2}(n-1)\right\}$$

$n=6,\overline{X}=31.13,S^2=1.26,\mu_0=32.50,t_{0.025}(5)=2.5706$

$$T=\left|\frac{\sqrt{6}\times(31.13-32.50)}{\sqrt{1.26}}\right|=2.98959>2.5706$$

故拒绝H_0,没有充分理由认为平均抗断强度为32.50 kg/cm^2.

11、测定某种溶液中的水分，它的10个测定值给出$\overline{X}=0.452\%$，$S=0.037\%$,设测定值总体为正态分布$N(a,\sigma^2)$.试在水平$\alpha=0.05$下检验假设

(1) H_0: $a\leqslant 0.5\% \leftrightarrow H_1$: $a>0.5\%$;

(2) H_0: $\sigma\geqslant 0.04\% \leftrightarrow H_1$: $\sigma<0.04\%$.

解：

(1)否定域为

$$D=\left\{(X_1,\cdots,X_n):T=\frac{\sqrt{n}\left(\overline{X}-\mu_0\right)}{S}>t_\alpha(n-1)\right\}$$

$n=10,\overline{X}=0.452,\mu_0=0.5,S=0.037,t_{0.05}(9)=1.8331$

$$T=\frac{\sqrt{10}(0.452-0.5)}{0.037}=-4.10241<1.8331$$

故在显著水平$\alpha=0.05$下认为没有足够拒绝H_0.

(2)否定域为

$$D=\left\{(X_1,\cdots,X_n):\frac{(n-1)S^2}{\sigma_0^2}<\chi^2_{n-1}(1-\alpha)\right\}$$

$n-1=9,S^2=0.037^2,\sigma_0^2=0.04^2,\chi_9^2(0.95)=3.325$

$$\frac{(n-1)S^2}{\sigma_0^2}=\frac{9\times 0.037^2}{0.04^2}=7.70062>3.325$$

故在显著水平$\alpha=0.05$下没有充分理由拒绝H_0.

12、设有甲乙两个实验员，对同样的试样进行分析，各人实验分析结果如下（分析结果服从正态分布）：

实验员	1	2	3	4	5	6	7	8

甲	4.3	3.2	3.8	3.5	3.5	4.8	3.3	3.9
乙	3.7	4.1	3.8	3.8	4.6	3.9	2.8	4.4

试问甲乙两实验员实验分析结果之间有无显著差异（$\alpha=0.05$）？

解：

此题是成对比较问题，设甲实验员的分析结果为$X_1,\cdots,X_8$，乙实验员的分析结果为$Y_1,\cdots,Y_8$，设$Z_i=Y_i-X_i,i=1,\cdots,8$,假定$Z_1,\cdots Z_8$为自$N(\mu,\sigma^2)$中抽取的随机样本，则检验问题为

$$H_0:\ \mu=0\leftrightarrow H_1:\ \mu\neq 0$$

否定域为

$$D=\left\{(Z_1,\cdots,Z_n):T=\left|\frac{\sqrt{n}\left(\bar{Z}-\mu_0\right)}{S}\right|>t_{n-1}(\alpha/2)\right\}$$

$n=8,\bar{Z}=-0.1,\mu_0=0,S=0.73,t_7(0.025)=2.3646$

$$T=\left|\frac{\sqrt{8}(-0.1-0)}{0.73}\right|=0.387456<2.3646$$

故在显著水平$\alpha=0.05$下没有充分理由拒绝H_0,即甲乙两实验员试验分析结果之间无显著差异.

13、有一种新安眠药，据说在一定剂量下，能比某旧安眠药平均增加睡眠时间3小时.根据以往资料，用旧安眠药时平均睡眠时间为20.8小时.为了检验新安眠药是否达到疗效，收集到一组用新安眠药的睡眠时间，分别为

$$26.7,22.0,24.1,21.0,27.2,25.0,23.4$$

假定睡眠时间服从正态分布，试在水平$\alpha=0.05$下，检验假设：新安眠药已达到疗效.

解：

设上述数据为从$X\sim N(\mu,\sigma^2)$中抽取的随机样本，则检验问题为

$$H_0:\ \mu=23.8\leftrightarrow H_1:\mu\neq 23.8$$

否定域为

$$D=\left\{(X_1,\cdots,X_n):T=\left|\frac{\sqrt{n}\left(\bar{X}-\mu_0\right)}{S}\right|>t_{n-1}(\alpha/2)\right\}$$

$n=7,\bar{X}=24.2,\mu_0=23.8,S=2.244,t_6(0.025)=2.4469$

$$T=\left|\frac{\sqrt{7}(24.2-23.8)}{2.244}\right|=0.471613<2.4469$$

故在显著水平$\alpha=0.05$下没有充分理由拒绝H_0,即新安眠药已达到疗效.

14、某工厂生产的某种细沙支数服从正态分布，其总体标准差为1.2，现从某日生产的一批产品中随机抽取 16 缕进行支数测量，求得样本标准差为 1.2，问纱的均匀度是否变劣（$\alpha=0.05$）？

解:

设上述数据为从$X\sim N(\mu,\sigma^2)$中抽取的随机样本，则检验问题为

$$H_0:\ \sigma^2\geqslant 1.2^2 \leftrightarrow H_1:\sigma^2<1.2^2$$

则否定域为

$$D=\left\{(X_1,\cdots,X_n):\frac{(n-1)S^2}{\sigma_0^2}<\chi^2_{n-1}(1-\alpha)\right\}$$

$n-1=15,S=2.1,\sigma_0^2=1.2^2,\chi^2_{15}(0.95)=7.261$

$$\frac{(n-1)S^2}{\sigma_0^2}=\frac{15\times 2.1^2}{1.2^2}=45.9375>7.261$$

故在显著水平$\alpha=0.05$下没有充分理由拒绝H_0,即纱的均匀度没有变劣.

15、某厂的一批电子产品，其寿命T服从指数分布，其密度函数为

$$f(t,\theta)=\theta^{-1}\exp\{-t/\theta\}I_{(0,\infty)}(t)$$

从以往生产情况知平均寿命$\theta=2000$小时.为检验当日生产是否稳定，任取10个产品进行寿命试验，到全部失效时试验停止.试验得失效寿命数据之和为30200.试在显著水平$\alpha=0.05$下检验假设

$$H_0:\ \theta=2000\ \leftrightarrow H_1:\theta\neq 2000.$$

解:

易知$\dfrac{2n\overline{X}}{\theta}\sim\chi^2_{2n}$,故否定域为

$$D=\left\{(X_1,\cdots,X_n):\frac{2n\overline{X}}{\theta_0}<\chi^2_{2n}(1-\alpha/2)\text{或}\frac{2n\overline{X}}{\theta_0}>\chi^2_{2n}(\alpha/2)\right\}$$

$n\overline{X}=30200,\theta_0=2000,\chi^2_{20}(0.975)=9.591,\chi^2_{20}(0.025)=34.170$

$$9.591<\frac{2n\overline{X}}{\theta_0}=\frac{2\times 30200}{2000}=30.2<34.170$$

故接受原假设.

16、设$X_1,\cdots,X_9\ i.i.d.\sim N(a,2.5^2);Y_1,\cdots,Y_{16}\ i.i.d.\sim N(b,3.4^2)$,算得$\bar{x}=49,\bar{y}=44$,在水平$\alpha=0.05$下检验

$$H_0:\ a\geqslant b\ \leftrightarrow H_1:a<b$$

解:

检验问题转化为

$$H_0:\ a-b\geqslant 0 \leftrightarrow H_1: a-b<0$$

否定域为

$$D=\left\{(X_1,\cdots,X_n): \frac{\overline{X}-\overline{Y}-\mu_0}{\sqrt{\sigma_1^2/m+\sigma_2^2/n}}<-u_\alpha\right\}$$

$\overline{X}=49$，$\overline{Y}=44, \mu_0=0, \sigma_1=2.5, \sigma_2=3.4, m=9, n=16$， $u_{0.05}=1.645$

$$\frac{49-44-0}{\sqrt{2.5^2/9+3.4^2/16}}=4.20043>-1.645$$

故接受H_0.

17、用两种方法研究冰溶解时的潜热，对冷却到$-0.72^\circ C$的冰溶解为$0^\circ C$的水的热量改变数据（单位：cal/g）：

方法1：$79.98, 80.04, 80.02, 80.04, 80.03, 80.03, 80.04$
$79.97, 80.05, 80.03, 80.02, 80.00, 80.02.$

方法2：$80.02, 79.94, 79.98, 79.97, 80.03, 79.95, 79.97, 79.97$

试问对水平$\alpha=0.05$这两种处理方法的平均性能有否显著差异？假设这两种方法测定值都是服从正态分布且方差相同.

解：

设方法1的样本为$X_1,\cdots,X_{13}$，服从正态分布$N(\mu_1,\sigma^2)$,方法2的样本为$Y_1,\cdots,Y_8$,
服从正态分布$N(\mu_2,\sigma^2)$.则检验问题为

$$H_0:\ \mu_2-\mu_1=0 \leftrightarrow H_1: \mu_2-\mu_1\neq 0$$

则否定域为

$$D=\left\{(X_1,\cdots,X_m;Y_1,\cdots,Y_n): T=\left|\frac{\overline{Y}-\overline{X}}{S_\omega}\sqrt{\frac{mn}{m+n}}\right|>t_{m+n-2}\left(\frac{\alpha}{2}\right)\right\}$$

$$m=13, n=8, \overline{X}=80.02, \overline{Y}=79.98, S_\omega^2=\frac{(m-1)S_1^2+(n-1)S_2^2}{m+n-2}=0.0012,$$

$$T=\left|\frac{79.98-80.02}{\sqrt{0.0012}}\times\sqrt{\frac{13\times 8}{13+8}}\right|=2.56966>t_{19}(0.025)=2.093$$

故拒绝原假设，在$\alpha=0.05$的水平下两种处理方法的平均性能有显著差异.

18、有甲乙两台机床加工同样产品，从这两台机床加工的产品中随机抽取若干产品，测得产品直径（单位：mm）为

机床甲：$20.5, 29.8, 19.7, 20.4, 20.1, 20.0, 19.0, 19.9$

机床乙：$19.7, 20.8, 20.5, 19.8, 19.4, 20.6, 19.2.$

假定产品直径服从正态分布，且二者方差相同，试比较甲乙两台机床加工的质量有无显著差异（$\alpha=0.05$）？

解：

设机床甲的样本为$X_1,\cdots,X_8$，服从正态分布$N(\mu_1,\sigma^2)$,机床乙的样本为$Y_1,\cdots,Y_7$,服从正态分布$N(\mu_2,\sigma^2)$.则检验问题为

$$H_0:\ \mu_2-\mu_1=0 \leftrightarrow H_1:\mu_2-\mu_1\neq 0$$

则否定域为

$$D=\left\{(X_1,\cdots,X_m;Y_1,\cdots,Y_n):T=\left|\frac{\overline{Y}-\overline{X}}{S_\omega}\sqrt{\frac{mn}{m+n}}\right|>t_{m+n-2}\left(\frac{\alpha}{2}\right)\right\}$$

$$m=8,n=7,\overline{X}=21.175,\overline{Y}=20,S_\omega^2=\frac{(m-1)S_1^2+(n-1)S_2^2}{m+n-2}=6.84,t_{13}(0.025)=2.16$$

$$T=\left|\frac{21.175-20}{\sqrt{6.84}}\times\sqrt{\frac{8\times 7}{8+7}}\right|=0.868077<2.16$$

故接受H_0,甲乙两台机床加工的质量无显著差异.

19、假设甲乙两煤矿的含灰率分别服从正态分布$N(a,\sigma^2)$和$N(b,\sigma^2)$.为检验这两个煤矿的煤含灰率有无显著差异，从两矿中各取若干份，分析结果（单位：%）为

$$\begin{aligned}&\text{甲矿：}24.3,18.8,22.7,19.3,20.4\\&\text{乙矿：}25.2,28.9,24.2,26.7,22.3,20.4\end{aligned}$$

试在水平$\alpha=0.05$之下，检验“含灰量无差异”这个假设.

解：

设甲矿的样本为$X_1,\cdots,X_5$，乙矿的样本为$Y_1,\cdots,Y_6$,则检验问题为

$$H_0:\ b-a=0 \leftrightarrow H_1:b-a\neq 0$$

则否定域为

$$D=\left\{(X_1,\cdots,X_m;Y_1,\cdots,Y_n):T=\left|\frac{\overline{Y}-\overline{X}}{S_\omega}\sqrt{\frac{mn}{m+n}}\right|>t_{m+n-2}\left(\frac{\alpha}{2}\right)\right\}$$

$$m=5,n=6,\overline{X}=21.1,\overline{Y}=24.62,S_\omega^2=7.57,t_9(0.025)=2.2622$$

$$T=\left|\frac{24.62-21.1}{\sqrt{7.57}}\times\sqrt{\frac{5\times 6}{5+6}}\right|=2.1128<2.2622$$

故接受H_0,认为含灰量无差异.

20、两台机床加工同一零件，分别取6个和9个零件，量其长度$X_i(i=1,\cdots,6)$和Y_j $(j=1,\cdots,9)$后算得$Q_1^2=\sum_{i=1}^{6}\left(X_i-\overline{X}\right)^2=1.725$，$Q_2^2=\sum_{j=1}^{9}\left(Y_j-\overline{Y}\right)^2=2.856$.假定零件长度服从正态分布，问是否可认为这两台机床加工的零件的方差无显著差异

（$\alpha=0.02$）？

解：

设$X\sim N(\mu_1,\sigma_1^2),Y\sim N(\mu_2,\sigma_2^2)$,则检验问题为

$$H_0:\ \sigma_1^2=\sigma_2^2 \leftrightarrow H_1:\sigma_1^2=\sigma_2^2$$

则拒绝域为

$$D=\left\{(X_1,\cdots,X_m;Y_1,\cdots,Y_n):F=\frac{S_2^2}{S_1^2}>F_{n-1,m-1}(\alpha/2)\text{或}F<F_{n-1,m-1}(1-\alpha/2)\right\}$$

$$m=6,n=9,S_1^2=1.725/5,S_2^2=2.856/8,F_{8,5}(0.01)=10.29,$$

$$F_{8,5}(0.99)=\frac{1}{F_{5,8}(0.01)}=0.613$$

$$F=\frac{2.856/8}{1.725/5}=1.03478, 0.613<F<10.29$$

故接受原假设，方差无显著差异.

21、两位化验员A,B对一种矿砂的含铁量各自独立地用同一种方法作了5次分析，得到样本方差分别0.4322和0.5006，若假定A,B测定值的总体都是正态分布，其方差分别为σ_A^2,σ_B^2,试在水平$\alpha=0.10$下检验$H_0:\sigma_A^2=\sigma_B^2$.

解：

$\alpha=0.10$

设$X\sim N(\mu_A,\sigma_A^2),Y\sim N(\mu_B,\sigma_B^2)$,则检验问题为

$$H_0:\ \sigma_A^2=\sigma_B^2 \leftrightarrow H_1:\sigma_A^2=\sigma_B^2$$

则拒绝域为

$$D=\left\{(X_1,\cdots,X_m;Y_1,\cdots,Y_n):F=\frac{S_B^2}{S_A^2}>F_{n-1,m-1}(\alpha/2)\text{或}F<F_{n-1,m-1}(1-\alpha/2)\right\}$$

$m=5,n=5,S_A^2=0.4322,S_B^2=0.5006,F_{4,4}(0.05)=6.39,F_{4,4}(0.95)=0.156$

$$F=\frac{S_B^2}{S_A^2}=\frac{0.5006}{0.4322}=1.15826, 0.156<F<6.39$$

故接受H_0.

22、为研究矽肺患者功能的变化情况，某医院对Ⅰ，Ⅱ期矽肺患者各33名测定其肺活量，得到Ⅰ期患者的平均数为2710ml,标准差为147ml，Ⅱ期患者的平均数为2830ml，标准差为118ml，对水平$\alpha=0.10$，试问第Ⅰ，Ⅱ期患者的肺活量有无显著差异（假定肺活

量服从正态分布）？

解：

设Ⅰ期样本为$X_1,\cdots,X_m$，$X\sim N(\mu_1,\sigma_1^2)$，Ⅱ期样本为$Y_1,\cdots,Y_n$，$Y\sim N(\mu_2,\sigma_2^2)$

首先考虑下列检验问题.

(1)H_0: $\sigma_1^2=\sigma_2^2\leftrightarrow H_1$:$\sigma_1^2\neq\sigma_2^2,\alpha=0.10$.

此检验的接受域为

$$D=\left\{(\boldsymbol{X},\boldsymbol{Y}):F_{m-1,n-1}(1-\alpha/2)<\frac{S_1^2}{S_2^2}<F_{m-1,n-1}(\alpha/2)\right\},$$

此处$m=33,n=33,\dfrac{S_1^2}{S_2^2}=\dfrac{147^2}{118^2}=1.55192,F_{32,32}(0.05)=1.81,F_{32,32}(0.95)=1/1.81$

$$\frac{1}{1.84}<\frac{S_1^2}{S_2^2}=1.55<1.84$$

故没有足够的证据否定H_0,因此接受H_0.

在接受上述检验后，可以假定$\sigma_1^2=\sigma_2^2$，进一步考虑下列检验问题.

(2)H_0':$\mu_1=\mu_2\leftrightarrow H_1'$:$\mu_1\neq\mu_2,\alpha=0.10$

此检验的否定域为

$$\{(X,Y):|T_\omega|>t_{n+m-2}(\alpha/2)\},$$

$m=33,n=33,\overline{X}=2710,\overline{Y}=2830,S_1=147,S_2=118$,因此有

$$S_\omega^2=\frac{1}{64}(32\times147^2+32\times118^2)=17766.5,S_\omega=\sqrt{17766.5}=133.291$$

由$\alpha=0.10$，查表得$t_{64}(0.05)=1.671$

$$|T_\omega|=\sqrt{\frac{mn}{m+n}}\left|\frac{\overline{Y}-\overline{X}}{S_\omega}\right|=\sqrt{\frac{33^2}{33\times2}}\left|\frac{2830-2710}{133.291}\right|=3.65698>1.671$$

故拒绝H_0,认为两期患者的肺活量有显著差异.

23、为了比较两种枪弹的速度（单位：米/秒），在相同条件下进行速度测定.算得样本均值和样本标准差如下：

$$\text{枪弹甲：}\ n_1=120,\bar{x}=2805,s_1=120.41;$$
$$\text{枪弹乙：}\ n_2=60,\bar{y}=2680,s_2=90.00.$$

问在水平$\alpha=0.05$下，两种枪弹的速度和均匀性方面是否有显著差异？

解；

用$Behrens-Fisher$大样本方法解决下列检验问题 H_0: $\mu_2-\mu_1=0\leftrightarrow H_1$:$\mu_2-\mu_1\neq0$

故否定域为

$$\{(X,Y):|U^*|>u_{\alpha/2}\},$$

其中$U^*=\overline{Y}-\overline{X}/\sqrt{S_1^2/n_1+S_2^2/n_2}=2680-2805/\sqrt{120.41^2/120+90^2/60}=-7.81523$,
$u_{0.025}=$查表得1.96,故$|U^*|>1.96$

因此否定H_0,认为两种枪弹的速度有显著差异.
再设下列检验问题：$H_0':\sigma_1^2=\sigma_2^2 \leftrightarrow H_1':\sigma_1^2\neq\sigma_2^2$
则接受域为

$$D=\left\{(\boldsymbol{X},\boldsymbol{Y}):F_{n_1-1,n_2-1}(1-\alpha/2)<\frac{S_1^2}{S_2^2}<F_{n_1-1,n_2-1}(\alpha/2)\right\},$$

其中$\dfrac{S_1^2}{S_2^2}=\dfrac{120.41^2}{90^2}=1.78995,F_{119,59}(0.025)=1.58,F_{59,119}(0.975)=\dfrac{1}{1.58}=0.632911$

$$\frac{S_1^2}{S_2^2}=1.79>1.58$$

故就拒绝原假设，两种枪弹的均匀性有显著差异.

24、在10块土地上同时试种甲、乙两个品种的农作物，假定每种作物的产量服从正态分布，并算得样本均值和样本标准差如下$\overline{X}=30.97$，$\overline{Y}=21.79,S_x=26.7,S_y=10.1$，问这两种作物的产量有无显著差异？（$\alpha=0.02$，提示：用$Behrens-Fisher$小样本方法的近似解）

解：
检验问题为

$$H_0:\ \mu_1=\mu_2 \leftrightarrow H_1:\mu_1\neq\mu_2$$

否定域为

$$D=\left\{(X,Y):|T|_*=\left|\frac{\overline{X}-\overline{Y}}{\sqrt{S_1^2/m+S_2^2/n}}\right|>t_r(\alpha/2)\right\}$$

$$r=\frac{(S_1^2/m+S_2^2/n)^2}{S_1^4/m^2(m-1)+S_2^4/n^2(n-1)}=\frac{(26.7^2/10+10.1^2/10)^2}{26.7^4/10^2\times9+10.1^4/10^2\times9}=11.524\approx12$$

查表得$t_{12}(0.01)=2.681$,由数据得

$$T_*=\left|\frac{30.97-21.79}{\sqrt{26.7^2/10+10.1^2/10}}\right|=1.01693<2.681$$

故接受H_0,即两种作物的产量无显著差异.

25、为确定肥料的效果，取1000株植物做试验，在没有施肥的100株植物中有53株长势良好，在已施肥的900株中，则有783株长势良好，问施肥的效果是否显著？($\alpha=0.01$)

解：

此题为二项分布的大样本检验，设未施肥的样本为$X_1,\cdots,X_m$，$X\sim b(1,p_1)$,
施肥的样本为$Y_1,\cdots,Y_n$，$Y\sim b(1,p_2)$,检验问题为

$$H_0:\ p_2>p_1 \leftrightarrow H_1:p_2\leqslant p_1$$

则由中心极限定理可知

$U=\dfrac{\overline{Y}-\overline{X}-(p_2-p_1)}{\sqrt{\frac{p_1(1-p_1)}{m}+\frac{p_2(1-p_2)}{n}}}\xrightarrow{\mathscr{L}}N(0,1)$，当$H_0$成立时$p_1=p_2=\hat{p}$,用合样本估计$\hat{p}$,

$$\tilde{p}=\frac{1}{m+n}\left(\sum_{i=1}^{m}X_i+\sum_{j=1}^{n}Y_j\right)$$

则有$\widetilde{U}=\dfrac{\overline{Y}-\overline{X}}{\sqrt{\hat{p}(1-\hat{p})}}\sqrt{\dfrac{mn}{m+n}}\xrightarrow{\mathscr{L}}N(0,1)$, 故否定域为

$$D=\left\{(\boldsymbol{X},\boldsymbol{Y}):\widetilde{U}<-u_\alpha\right\}$$

$m=100,n=900,\overline{X}=0.53,\overline{Y}=0.87,\hat{p}=0.836$

$$\widetilde{U}=\frac{0.87-0.83}{\sqrt{0.836\times(1-0.836)}}\sqrt{\frac{100\times900}{100+900}}=1.02484>-u_{0.01}=-2.325$$

故接受H_0,即认为施肥的效果显著.

26、设有两工厂生产的同一种产品，要检验假设H:它们的废品率p_1,p_2相同，在第一、二工厂的产品中各自独立抽取$n_1=1500$个及$n_2=1800$个，分别有废品300个及320个，问在5%水平上应接受还是拒绝H.

解：

此题为二项分布的大样本检验，设工厂一的样本为$X_1,\cdots,X_{n_1}$， $X\sim b(1,p_1)$,
工厂二的样本为$Y_1,\cdots,Y_{n_2}$， $Y\sim b(1,p_2)$,检验问题为

$$H_0:\ p_2=p_1\leftrightarrow H_1:p_2\neq p_1$$

则由中心极限定理可知

$U=\dfrac{\overline{Y}-\overline{X}-(p_2-p_1)}{\sqrt{\frac{p_1(1-p_1)}{n_1}+\frac{p_2(1-p_2)}{n_2}}}\xrightarrow{\mathscr{L}}N(0,1)$，当$H_0$成立时$p_1=p_2=\hat{p}$,用合样本估计$\hat{p}$,

$$\tilde{p}=\frac{1}{n_1+n_2}\left(\sum_{i=1}^{n_1}X_i+\sum_{j=1}^{n_2}Y_j\right)$$

则有$\widetilde{U}=\dfrac{\overline{Y}-\overline{X}}{\sqrt{\hat{p}(1-\hat{p})}}\sqrt{\dfrac{mn}{m+n}}\xrightarrow{\mathscr{L}}N(0,1)$, 故否定域为

$$D=\left\{(\boldsymbol{X},\boldsymbol{Y}):\left|\widetilde{U}\right|>u_{\alpha/2}\right\}$$

$n_1=1500,n_2=1800,\overline{X}=0.2,\overline{Y}=0.18,\hat{p}=0.19$

$$\widetilde{U}=\left|\frac{0.18-0.2}{\sqrt{0.19\times(1-0.19)}}\times\sqrt{\frac{1500\times1800}{1500+1800}}\right|=1.45826<u_{0.025}=1.96$$

故接受H_0,废品率相同.

27、设$X_1,\cdots,X_m\, i.i.d. \sim$ $Poisson$分布$P(\lambda_1)$, $Y_1,\cdots,Y_n\, i.i.d. \sim Poisson$分布$P(\lambda_2)$, 且样本$X_1,\cdots,X_m$和$Y_1,\cdots,Y_n$独立，用大样本方法求下列检验问题$H_0$: $\lambda_2-\lambda_1=0 \leftrightarrow H_1$:$\lambda_2-\lambda_1\neq 0$，检验水平$\alpha$给定.

解：

$$T_1=\sum_{i=1}^{m}X_i\sim P(m\lambda_1),T_2=\sum_{j=1}^{n}Y_j\sim P(n\lambda_2)$$

由中心极限定理可知

$$T=\frac{T_2-T_1-(\lambda_2-\lambda_1)}{\sqrt{m\lambda_1+n\lambda_2}}\xrightarrow{\mathscr{L}}N(0,1)$$

当H_0成立时，$\lambda_2=\lambda_1=\lambda_0,\lambda_0$用合样本估计，$\lambda_0=\dfrac{1}{m+n}\left(\sum\limits_{i=1}^{m}X_i+\sum\limits_{j=1}^{n}Y_j\right)$，则

$$T'=\frac{T_2-T_1-(\lambda_2-\lambda_1)}{\sqrt{\lambda_0(m+n)}}\xrightarrow{\mathscr{L}}N(0,1)$$

故否定域为

$$D=\{(\boldsymbol{X},\boldsymbol{Y}):|T'|>u_{\alpha/2}\}$$

28、$X_1,\cdots,X_m\, i.i.d.\sim b(1,p_1),Y_1,\cdots,Y_n\, i.i.d.\sim b(1,p_2)$,两组样本独立，用大样本方法求检验

$$H_0:\ p_2-p_1=p_0\leftrightarrow H_1:p_2-p_1\neq p_0$$

检验水平α和p_0给定.

解：

由中心极限定理可知

$U=\dfrac{\overline{Y}-\overline{X}-(p_2-p_1)}{\sqrt{\frac{p_1(1-p_1)}{m}+\frac{p_2(1-p_2)}{n}}}\xrightarrow{\mathscr{L}}N(0,1)$，当$H_0$成立时$p_2-p_1=p_0$,用合样本估计$\hat{p}_1,\hat{p}_2$,

$$\tilde{p}_1=\frac{1}{m+n}\left(\sum_{i=1}^{m}X_i+\sum_{j=1}^{n}Y_j-np_0\right),\hat{p}_2=\frac{1}{m+n}\left(\sum_{i=1}^{m}X_i+\sum_{j=1}^{n}Y_j+np_0\right)$$

则有$\widetilde{U}=\dfrac{\overline{Y}-\overline{X}-p_0}{\sqrt{\hat{p}_1(1-\hat{p}_1)/m+\hat{p}_2(1-\hat{p}_2)/n}}\xrightarrow{\mathscr{L}}N(0,1)$, 故否定域为

$$D=\left\{(\boldsymbol{X},\boldsymbol{Y}):\left|\widetilde{U}\right|>u_{\alpha/2}\right\}$$

29、设$X_1,\cdots,X_n\ i.i.d.\sim N(\mu,\sigma^2)$,其中$\sigma^2$已知，求检验问题：$H_0:\mu=0\leftrightarrow H_1:\mu\neq 0$水平为$\alpha$的似然比检验..

解：

似然函数为

$$f(\boldsymbol{x},\mu)=(2\pi\sigma^2)^{-\frac{n}{2}}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2\right\}$$

参数空间为

$$\Theta=\{\mu:-\infty<\mu<+\infty\}$$

零假设H_0对应的Θ的子集为

$$\Theta_0;\ \{\mu:\ \mu=0\}$$

在Θ上，μ的极大似然估计为

$$\hat{\mu}=\overline{X};$$

在Θ_0上 $\hat{\mu}=0$

故有

$$\sup_{\mu\in\Theta}f(\boldsymbol{x},\mu)=(2\pi\sigma^2)^{-\frac{n}{2}}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2\right\}$$

$$\sup_{\mu\in\Theta_0}f(\boldsymbol{x},\mu)=(2\pi\sigma^2)^{-\frac{n}{2}}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}X_i^2\right\}$$

从而有

$$\begin{aligned}\lambda(\boldsymbol{x})&=\exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2-\sum_{i=1}^{n}X_i^2\right]\right\}\\&=\exp\left\{\frac{n\overline{X}^2}{2\sigma^2}\right\}=\exp\left\{\frac{1}{2}|T(\boldsymbol{X})|^2\right\}\end{aligned}$$

其中$T(\boldsymbol{X})=\dfrac{\sqrt{n}\,\overline{X}}{\sigma}\sim N(0,1)$，$\lambda(\boldsymbol{x})$是$|T(\boldsymbol{X})|$的严格增函数，故检验的否定域为

$D=\{(X_1,\cdots,X_n):\lambda(\boldsymbol{x})>c\}=\{\boldsymbol{X}:|T|>c'\}$,令$P(|T|>c'|H_0)=\alpha$,因此

$$\varphi(\boldsymbol{x})=\begin{cases}1,\left|\sqrt{n}\,\overline{X}/\sigma\right|>u_{\alpha/2}\\0,\left|\sqrt{n}\,\overline{X}/\sigma\right|\leqslant u_{\alpha/2}\end{cases}$$

是检验问题的一个水平为α的似然比检验.

30、设$X_1,\cdots,X_n$ $i.i.d.\sim N(\mu,\sigma^2),-\infty<\mu<+\infty,\sigma^2>0$都是未知参数，试求检验问题$H_0:\sigma^2\leqslant\sigma_0^2\leftrightarrow H_1:\sigma^2>\sigma_0^2$水平为$\alpha$的似然比检验.

解：

令$\boldsymbol{\theta}=(\mu,\sigma^2)$,则$\boldsymbol{\theta}$的似然函数为

$$f(\boldsymbol{x},\boldsymbol{\theta})=(2\pi\sigma^2)^{-\frac{n}{2}}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2\right\}$$

参数空间为

$$\Theta=\{\boldsymbol{\theta}=(\mu,\sigma^2):-\infty<\mu<+\infty,\sigma^2>0\}$$

零假设H_0对应的Θ的子集为

$$\Theta_0=\{\theta=(\mu,\sigma^2):-\infty<\mu<+\infty,0<\sigma^2\leqslant\sigma_0^2\}$$

故有

$$L_{\Theta}(\boldsymbol{x})=\sup_{\boldsymbol{\theta}\in\Theta}f(\boldsymbol{x},\boldsymbol{\theta})=\left(\frac{2\pi e}{n}\right)^{-n/2}\left(\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2\right)^{-n/2},$$

而

$$L_{\Theta_0}(\boldsymbol{x})=\sup_{\boldsymbol{\theta}\in\Theta_0}(2\pi\sigma^2)^{-\frac{n}{2}}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2\right\}$$

记$g(\sigma^2)=(\sigma^2)^{-\frac{n}{2}}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2\right\}$,

则$g'(\sigma^2)=\left(\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2-n\sigma^2\right)(\sigma^2)^{-\frac{n}{2}-2}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2\right\}$,

记$\hat{\sigma}^2=\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2\Big/n$，当$\sigma^2\leqslant\hat{\sigma}^2$时，$g'(\sigma^2)>0,g(\sigma^2)$关于$\sigma^2$单调增，当$\sigma^2>\hat{\sigma}^2$时，$g'(\sigma^2)>0$,故$g(\sigma^2)$关于$\sigma^2$单调减.因此

（1）当$\sigma_0^2<\hat{\sigma}^2$时，若$H_0$成立，$g(\sigma^2)$在$\sigma^2=\sigma_0^2$处取得最大，故有

$$L_{\Theta_0}(\boldsymbol{x})=(2\pi\sigma_0^2)^{-\frac{n}{2}}\exp\left\{-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2\right\}$$

（2）当$\sigma_0^2\geqslant\hat{\sigma}^2$时，若$H_0$成立，$g(\sigma^2)$在$\sigma^2=\hat{\sigma}^2$处取得最大，故有

$$L_{\Theta_0}(\boldsymbol{x})=\left(\frac{2\pi e}{n}\right)^{-\frac{n}{2}}\left(\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2\right)^{-\frac{n}{2}}$$

因此有

$$L_{\Theta_0}(\boldsymbol{x})=\begin{cases}(2\pi\sigma_0^2)^{-\frac{n}{2}}\exp\left\{-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2\right\}, & \sigma_0^2<\hat{\sigma}^2\\ \left(\frac{2\pi e}{n}\right)^{-\frac{n}{2}}\left(\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2\right)^{-\frac{n}{2}}, & \sigma_0^2\geqslant\hat{\sigma}^2\end{cases}$$

故

$$\lambda(\boldsymbol{x})=\begin{cases}\left(\frac{n}{e}\right)^{-\frac{n}{2}}\left[\frac{1}{\sigma_0^2}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2\right]^{-n/2}\exp\left\{\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2\right\}, & \sigma_0^2<\hat{\sigma}^2\\ 1, & \sigma_0^2\geqslant\hat{\sigma}^2\end{cases}$$

令$\xi=T[\boldsymbol{X}]=\frac{1}{\sigma_0^2}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2,g(\xi)=\xi^{-\frac{n}{2}}e^{\xi/2},g'(\xi)=(\xi-n)\frac{\xi^{-\frac{n}{2}-1}}{2}e^{\xi/2}$,又$\sigma_0^2<\hat{\sigma}^2$时,
$\xi=T[\boldsymbol{X}]>n$,故$g'(\xi)>0,g(\xi)$关于ξ递增，故拒绝域为
$\bar{D}=\{\boldsymbol{X}:\lambda(\boldsymbol{x})>c\}=\{\boldsymbol{X}:T(\boldsymbol{X})>c'\}=\left\{\boldsymbol{X}:\frac{1}{\sigma_0^2}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2>c'\right\}$,又因为$H_0$
成立时$\frac{1}{\sigma_0^2}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2\sim\chi_{n-1}^2$,令$P(T(\boldsymbol{X})>c'|H_0)=\alpha$,故$c'=\chi_{n-1}^2(\alpha)$.
因此检验问题的水平为α的似然比检验为

$$\varphi(\boldsymbol{x})=\begin{cases}0, & \frac{1}{\sigma_0^2}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2\leqslant\chi_{n-1}^2(\alpha)\\ 1, & \frac{1}{\sigma_0^2}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2>\chi_{n-1}^2(\alpha)\end{cases}$$

31、设$X_1,\cdots,X_m$ $i.i.d.\sim N(\mu_1,\sigma^2),Y_1,\cdots,Y_n$ $i.i.d.\sim N(\mu_2,\sigma^2)$,且样本$X_1,\cdots,X_m$，和$Y_1,\cdots,Y_n$独立，求下列检验问题水平为$\alpha$的似然比检验：

$$H_0:\mu_2-\mu_1=0\leftrightarrow H_1:\mu_2-\mu_1\neq 0.$$

解：
参数空间

$$\Theta=\{(\mu_1,\mu_2,\sigma^2):-\infty<\mu_1,\mu_2<+\infty,\sigma^2>0\},$$

零假设H_0对应的Θ的子集为

$$\Theta_0=\{(\mu_1,\mu_2,\sigma^2):-\infty<\mu_2=\mu_1<+\infty,\sigma^2>0\}$$

似然函数为

$$L(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\theta})=f(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\theta})=(2\pi\sigma^2)^{-\frac{m+n}{2}}\exp\left\{-\frac{\sum_{i=1}^{m}(x_i-\mu_1)^2+\sum_{j=1}^{n}(y_j-\mu_2)^2}{2\sigma^2}\right\}$$

对数似然函数为

$$\ln L(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\theta})=-\frac{m+n}{2}[\ln(2\pi)+\ln\sigma^2]-\frac{\sum_{i=1}^{m}(x_i-\mu_1)^2+\sum_{j=1}^{n}(y_j-\mu_2)^2}{2\sigma^2}$$

先求$L_\Theta(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\theta})$

$$\frac{\partial \ln L(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\theta})}{\partial \sigma^2}=-\frac{m+n}{2\sigma^2}+\frac{\sum_{i=1}^{m}(x_i-\mu_1)^2+\sum_{j=1}^{n}(y_j-\mu_2)^2}{2\sigma^4}=0$$

$$\frac{\partial \ln L(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\theta})}{\partial \mu_1}=\left(\sum_{i=1}^{m}x_i-m\mu_1\right)\Big/\sigma^2=0$$

$$\frac{\partial \ln L(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\theta})}{\partial \mu_2}=\left(\sum_{j=1}^{n}y_j-n\mu_2\right)\Big/\sigma^2=0$$

解得$\hat{\sigma}^2=\left(\sum_{i=1}^{m}(x_i-\mu_1)^2+\sum_{j=1}^{n}(y_j-\mu_2)^2\right)\Big/m+n$,

$\mu_1=\sum_{i=1}^{m}x_i\Big/m=\overline{X},\mu_2=\sum_{j=1}^{n}y_j\Big/n=\overline{Y}.$

故$L_{\Theta}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\theta})=\left(\frac{2\pi e}{m+n}\right)^{-\frac{m+n}{2}}\left(\sum_{i=1}^{m}\left(x_i-\bar{x}\right)^2+\sum_{j=1}^{n}\left(y_j-\bar{y}\right)^2\right)^{-\frac{m+n}{2}}$

H_0成立时，设$\mu_1=\mu_2=\mu$对数似然函数为

$$\ln L(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\theta})=-\frac{m+n}{2}[\ln(2\pi)+\ln\sigma^2]-\frac{\sum_{i=1}^{m}(x_i-\mu)^2+\sum_{j=1}^{n}(y_j-\mu)^2}{2\sigma^2}$$

$$\frac{\partial \ln L(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\theta})}{\partial \sigma^2}=-\frac{m+n}{2\sigma^2}+\frac{\sum_{i=1}^{m}(x_i-\mu)^2+\sum_{j=1}^{n}(y_j-\mu)^2}{2\sigma^4}=0$$

$$\frac{\partial \ln L(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\theta})}{\partial \mu}=\frac{1}{2\sigma^2}\left[\sum_{i=1}^{m}x_i+\sum_{j=1}^{n}y_j-(m+n)\mu\right]=0$$

解得$\hat{\sigma}^2=\left(\sum_{i=1}^{m}(x_i-\mu)^2+\sum_{j=1}^{n}(y_j-\mu)^2\right)\Big/m+n,\hat{\mu}=\frac{1}{m+n}\left(\sum_{i=1}^{m}x_i+\sum_{j=1}^{n}y_j\right)$

故$L_{\Theta_0}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\theta})=\left(\frac{2\pi e}{m+n}\right)^{-\frac{m+n}{2}}\left(\sum_{i=1}^{m}\left(x_i-\hat{\mu}\right)^2+\sum_{j=1}^{n}\left(y_j-\hat{\mu}\right)^2\right)^{-\frac{m+n}{2}}$

则$\lambda(\boldsymbol{x},\boldsymbol{y})=\left(\frac{\sum_{i=1}^{m}(x_i-\hat{\mu})^2+\sum_{j=1}^{n}(y_j-\hat{\mu})^2}{\sum_{i=1}^{m}\left(x_i-\bar{x}\right)^2+\sum_{j=1}^{n}\left(y_j-\bar{y}\right)^2}\right)^{\frac{m+n}{2}}$

$$=\left(\frac{\sum_{i=1}^{m}\left(x_i-\bar{x}\right)^2+m\left(\bar{x}-\hat{\mu}^2\right)^2+\sum_{j=1}^{n}\left(y_j-\bar{y}\right)^2+n\left(\bar{y}-\hat{\mu}\right)^2}{\sum_{i=1}^{m}\left(x_i-\bar{x}\right)^2+\sum_{j=1}^{n}\left(y_j-\bar{y}\right)^2}\right)^{\frac{m+n}{2}}$$

$$=\left(1+\frac{m\left(\bar{x}-\frac{m\bar{x}+n\bar{y}}{m+n}\right)^{2}+n\left(\bar{y}-\frac{m\bar{x}+n\bar{y}}{m+n}\right)^{2}}{(m+n-2)S_{\omega}^{2}}\right)^{\frac{m+n}{2}},\left(S_{\omega}^{2}=\frac{\sum_{i=1}^{m}\left(x_{i}-\bar{x}\right)^{2}+\sum_{j=1}^{n}\left(y_{j}-\bar{y}\right)^{2}}{m+n-2}\right)$$

$$=\left(1+\frac{mn}{m+n}\cdot\frac{\left(\bar{x}-\bar{y}\right)^{2}}{(m+n-2)S_{\omega}^{2}}\right)^{\frac{m+n}{2}}=\left(1+\frac{T^{2}}{m+n-2}\right)^{\frac{m+n}{2}},T=\sqrt{\frac{mn}{m+n}}\frac{\overline{Y}-\overline{X}}{S_{\omega}^{2}}$$

易知$T\sim t_{m+n-2}$,故$P(|T|>c|H_0)=\alpha$,故可得检验问题的水平为α的似然比检验

$$\lambda(\boldsymbol{x})=\begin{cases}1,|T|>t_{m+n-2}(\alpha/2)\\0,|T|\leqslant t_{m+n-2}(\alpha/2)\end{cases}$$

32、设$X_1,\cdots,X_m\ i.i.d.\sim N(\mu_1,\sigma_1^2),Y_1,\cdots,Y_n i.i.d.\ \sim N(\mu_2,\sigma_2^2)$,且$X_1,\cdots,X_m$与$Y_1,\cdots,Y_n$独立，试求$H{:}\sigma_1^2=\sigma_2^2\leftrightarrow K{:}\sigma_1^2\neq\sigma_2^2$的似然比检验.（取检验水平为$\alpha$）

解:

似然比函数为

$$L(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\theta})=(2\pi)^{-\frac{m+n}{2}}(\sigma_1^2)^{-\frac{m}{2}}(\sigma_2^2)^{-\frac{n}{2}}\exp\left\{-\frac{\sum_{i=1}^{m}(x_i-\mu_1)^2}{2\sigma_1^2}-\frac{\sum_{j=1}^{n}(y_j-\mu_2)^2}{2\sigma_2^2}\right\}$$

$$\ln L(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\theta})=-\frac{m+n}{2}\ln(2\pi)-\frac{m}{2}\ln\sigma_1^2-\frac{n}{2}\ln\sigma_2^2-\frac{\sum_{i=1}^{m}(x_i-\mu_1)^2}{2\sigma_1^2}-\frac{\sum_{j=1}^{n}(y_j-\mu_2)^2}{2\sigma_2^2}$$

参数空间为

$$\Theta=\{(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2){:}-\infty<\mu_1,\mu_2<+\infty,\sigma_1^2>0,\sigma_2^2>0\}$$

零假设H对应的Θ的子集为

$$\Theta_0=\{(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2){:}-\infty<\mu_1,\mu_2<+\infty,\sigma_1^2=\sigma_2^2=\sigma^2>0\}$$

计算可得

$$L_{\Theta}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\theta})=(2\pi e)^{-\frac{m+n}{2}}\left(\frac{1}{m}\sum_{i=1}^{m}\left(x_i-\bar{x}\right)^2\right)^{-\frac{m}{2}}\left(\frac{1}{n}\sum_{j=1}^{n}\left(y_j-\bar{y}\right)^2\right)^{-\frac{n}{2}}$$

$$L_{\Theta_0}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\theta})=\left(\frac{2\pi e}{m+n}\right)^{-\frac{m+n}{2}}\left(\sum_{i=1}^{m}\left(x_i-\bar{x}\right)^2+\sum_{j=1}^{n}\left(y_j-\bar{y}\right)^2\right)^{-\frac{m+n}{2}}$$

从而有

$$\lambda(\boldsymbol{x},\boldsymbol{y})=\frac{m^{\frac{m}{2}}n^{\frac{n}{2}}}{(m+n)^{\frac{m+n}{2}}}\left(1+\frac{\sum_{j=1}^{n}\left(y_j-\bar{y}\right)^2}{\sum_{i=1}^{m}\left(x_i-\bar{x}\right)^2}\right)^{\frac{m}{2}}\left(1+\frac{\sum_{i=1}^{m}\left(x_i-\bar{x}\right)^2}{\sum_{j=1}^{n}\left(y_j-\bar{y}\right)^2}\right)^{\frac{n}{2}}$$

令$T=\dfrac{\sum_{j=1}^{n}\left(y_j-\bar{y}\right)^2}{\sum_{i=1}^{m}\left(x_i-\bar{x}\right)^2}$,$g(t)=(1+t)^{\frac{m}{2}}\left(1+\dfrac{1}{t}\right)^{\frac{n}{2}}=(1+t)^{\frac{m+n}{2}}t^{-\frac{n}{2}}$

则$g'(t)=(1+t)^{\frac{m+n}{2}-1}t^{-\frac{n}{2}-1}\left(\dfrac{m}{2}t-\dfrac{n}{2}\right)$,当$0<t<\dfrac{n}{m}$时，$g(t)$是关于$t$的单减函数，

当$\dfrac{n}{m}\leqslant t<\infty$时，$g(t)$是关于$t$的单减函数.故接受域为$\overline{D}=\{(\boldsymbol{X},\boldsymbol{Y}):\lambda(\boldsymbol{x},\boldsymbol{y})<c\}$

$\{(\boldsymbol{X},\boldsymbol{Y}):\lambda(\boldsymbol{x},\boldsymbol{y})<c'\}=\{(\boldsymbol{X},\boldsymbol{Y}):k_1<T<k_2\}$,其中

$$\begin{cases} g(k_1)=g(k_2)=c' \\ P(k_1<T<k_2|H_0)=1-\alpha \end{cases}$$

又H_0成立时，$\dfrac{m-1}{n-1}T=\dfrac{m-1}{n-1}\dfrac{\sum_{j=1}^{n}\left(Y_j-\overline{Y}\right)^2}{\sum_{i=1}^{m}\left(X_i-\overline{X}\right)^2}\sim F_{n-1,m-1}$

故检验问题的水平为α的似然比检验为

$$\varphi(\boldsymbol{x},\boldsymbol{y})=\begin{cases} 1,\dfrac{m-1}{n-1}T<F_{n-1,m-1}(1-\alpha/2)\text{或}\dfrac{m-1}{n-1}T>F_{n-1,m-1}(\alpha/2) \\ 0,F_{n-1,m-1}(1-\alpha/2)<\dfrac{m-1}{n-1}T<F_{n-1,m-1}(\alpha/2) \end{cases}$$

33、从k个具有同方差的正态总体中各取一容量为n的样本，导出检验假设“所有平均数都为0”水平为α的似然比检验，证明此检验的似然比是F变量的严格单调增函数.

解：

设$X_i\sim N(\mu_i,\sigma^2),i=1,\cdots,k$.则样本为$X_{i_j},i=1,\cdots,k,j=1,\cdots,n$.

$$f(\boldsymbol{x},\boldsymbol{\theta})=(2\pi\sigma^2)^{-\frac{kn}{2}}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{k}\sum_{j=1}^{n}(x_{i_j}-\mu_i)^2\right\}$$

参数空间为

$$\Theta=\{(\mu_1,\cdots,\mu_k,\sigma^2):-\infty<\mu_i<+\infty,i=1,\cdots,k,\sigma^2>0\}$$

零假设H_0对应的Θ的子集为

$$\Theta_0=\{(\mu_1,\cdots,\mu_k,\sigma^2):\mu_1=\cdots=\mu_k=0,\sigma^2>0\}$$

计算可得

$$L_{\Theta}(\boldsymbol{x},\boldsymbol{\theta})=\left(\frac{2\pi e}{kn}\right)^{-\frac{kn}{2}}\left(\sum_{i=1}^{k}\sum_{j=1}^{n}\left(x_{i_j}-\bar{x}_i\right)^2\right)^{-\frac{kn}{2}}$$

$$L_{\Theta_0}(\boldsymbol{x},\boldsymbol{\theta})=\left(\frac{2\pi e}{kn}\right)^{-\frac{kn}{2}}\left(\sum_{i=1}^{k}\sum_{j=1}^{n}x_{i_j}^2\right)^{-\frac{kn}{2}}$$

从而有

$$\lambda(\boldsymbol{x})=\left[\left(\sum_{i=1}^{k}\sum_{j=1}^{n}x_{i_j}^2\right)\Big/\sum_{i=1}^{k}\sum_{j=1}^{n}\left(x_{i_j}-\bar{x}\right)^2\right]^{\frac{kn}{2}}$$

令$T[\boldsymbol{X}]=\left(\sum_{i=1}^{k}\sum_{j=1}^{n}x_{i_j}^2\right)\Big/\sum_{i=1}^{k}\sum_{j=1}^{n}\left(x_{i_j}-\bar{x}\right)^2$,则$\lambda(\boldsymbol{x})$是关于$T$的增函数，则接受域为

$\overline{D}=\{\boldsymbol{X}:\lambda(\boldsymbol{x})<c\}=\{\boldsymbol{X}:T(\boldsymbol{X})<c'\}$,而$H_0$成立时，$\dfrac{kn-k}{kn}T\sim F_{kn,kn-k}$,故此检验的似然比是$F$变量的严格单调增函数.

34、在题32的正态两样本情况下，给定$c>0$,找出$H_0:\sigma_1^2=c\sigma_2^2\leftrightarrow H_0:\sigma_1^2\neq c\sigma_2^2$的水平为$\alpha$的似然比检验.

解：

似然比函数为

$$L(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\theta})=(2\pi)^{-\frac{m+n}{2}}(\sigma_1^2)^{-\frac{m}{2}}(\sigma_2^2)^{-\frac{n}{2}}\exp\left\{-\frac{\sum_{i=1}^{m}(x_i-\mu_1)^2}{2\sigma_1^2}-\frac{\sum_{j=1}^{n}(y_j-\mu_2)^2}{2\sigma_2^2}\right\}$$

$$\ln L(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\theta})=-\frac{m+n}{2}\ln(2\pi)-\frac{m}{2}\ln\sigma_1^2-\frac{n}{2}\ln\sigma_2^2-\frac{\sum_{i=1}^{m}(x_i-\mu_1)^2}{2\sigma_1^2}-\frac{\sum_{j=1}^{n}(y_j-\mu_2)^2}{2\sigma_2^2}$$

参数空间为

$$\Theta=\{(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2):-\infty<\mu_1,\mu_2<+\infty,\sigma_1^2>0,\sigma_2^2>0\}$$

零假设H_0对应的Θ的子集为

$$\Theta_0=\{(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2):-\infty<\mu_1,\mu_2<+\infty,\sigma_1^2=c\sigma_2^2=\sigma^2>0\}$$

计算可得

$$L_{\Theta}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\theta})=(2\pi e)^{-\frac{m+n}{2}}\left(\frac{1}{m}\sum_{i=1}^{m}\left(x_i-\bar{x}\right)^2\right)^{-\frac{m}{2}}\left(\frac{1}{n}\sum_{j=1}^{n}\left(y_j-\bar{y}\right)^2\right)^{-\frac{n}{2}}$$

$$L_{\Theta_0}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\theta})=\left(\frac{2\pi e}{m+n}\right)^{-\frac{m+n}{2}}c^{\frac{n}{2}}\left(\sum_{i=1}^{m}\left(x_i-\bar{x}\right)^2+c\sum_{j=1}^{n}\left(y_j-\bar{y}\right)^2\right)^{-\frac{m+n}{2}}$$

从而有

$$\lambda(\boldsymbol{x})=\frac{m^{\frac{m}{2}}n^{\frac{n}{2}}}{(m+n)^{\frac{m+n}{2}}}(1+T)^{\frac{m}{2}}\left(1+\frac{1}{T}\right)^{\frac{n}{2}},\text{其中}T=c\frac{\sum_{j=1}^{n}\left(y_j-\bar{y}\right)^2}{\sum_{i=1}^{m}\left(x_i-\bar{x}\right)^2}$$

则$g'(t)=(1+t)^{\frac{m+n}{2}-1}t^{-\frac{n}{2}-1}\left(\frac{m}{2}t-\frac{n}{2}\right)$,当$0<t<\frac{n}{m}$时，$g(t)$是关于$t$的单减函数，

当$\frac{n}{m}\leqslant t<\infty$时，$g(t)$是关于$t$的单减函数.故接受域为$\overline{D}=\{(\boldsymbol{X},\boldsymbol{Y}):\lambda(\boldsymbol{x},\boldsymbol{y})<c\}$

$\{(\boldsymbol{X},\boldsymbol{Y}):\lambda(\boldsymbol{x},\boldsymbol{y})<c'\}=\{(\boldsymbol{X},\boldsymbol{Y}):k_1<T<k_2\}$,其中

$$\begin{cases}g(k_1)=g(k_2)=c'\\P(k_1<T<k_2|H_0)=1-\alpha\end{cases}$$

又H_0成立时，$\frac{m-1}{n-1}T=\frac{m-1}{n-1}\frac{\sum_{j=1}^{n}\left(Y_j-\overline{Y}\right)^2}{\sum_{i=1}^{m}\left(X_i-\overline{X}\right)^2}\sim F_{n-1,m-1}$

故检验问题的水平为α的似然比检验为

$$\varphi(\boldsymbol{x},\boldsymbol{y})=\begin{cases}1,\frac{m-1}{n-1}T<F_{n-1,m-1}(1-\alpha/2)\text{或}\frac{m-1}{n-1}T>F_{n-1,m-1}(\alpha/2)\\0,F_{n-1,m-1}(1-\alpha/2)<\frac{m-1}{n-1}T<F_{n-1,m-1}(\alpha/2)\end{cases}$$

35、设$X_1,\cdots,X_n\ i.i.d.\sim$指数分布$Exp(\lambda)$,求(1)$H_0:\lambda=\lambda_0\leftrightarrow H_1:\lambda\neq\lambda_0$和(2)
$H_0:\lambda\leqslant\lambda_0\leftrightarrow H_1:\lambda>\lambda_0$的水平为$\alpha$的似然比检验.

解：

（1）参数λ的似然函数为

$$L(\boldsymbol{x},\lambda)=\begin{cases}\lambda^n\exp\left\{-\lambda n\bar{x}\right\},\text{当}x_1,\cdots,x>0,\\0,\qquad\text{其他}\end{cases}$$

参数空间和H_0对应的参数空间的子集分别为$\Theta=(0,\infty)$和$\Theta_0=\{\lambda:\lambda=\lambda_0\}$.
由于λ的极大似然估计为$\hat{\lambda}=\overline{X}^{-1}$，故有

$$L_{\Theta}(\boldsymbol{x},\lambda)=\overline{X}^{-n}e^{-n},$$
$$L_{\Theta_0}(\boldsymbol{x},\lambda)=\lambda_0^n\exp\left\{-\lambda_0 n\bar{x}\right\}$$

记$t=T(\boldsymbol{x})=\sum_{i=1}^{n}x_i=n\bar{x}$,则似然比为

$$\lambda(\boldsymbol{x})=\frac{n^n}{\lambda_0^n e^n t^n}e^{\lambda_0 t}$$

令$g(t)=t^{-n}e^{\lambda_0 t}$,易见当$t\to 0$和$t\to\infty$时$g(t)\to\infty$,$g(t)$的形状与图5.3.1类似.因此

$$\lambda(\boldsymbol{x})=\begin{cases}1,当\lambda(\boldsymbol{x})>c\\0,当\lambda(\boldsymbol{x})\leqslant c\end{cases}=\begin{cases}1,当t<k_1或t>k_2,\\0,其他\end{cases}$$

又当H_0成立时$2\lambda_0 T(\boldsymbol{X})\sim\chi^2_{2n}$,故可得检验水平为$\alpha$的似然比检验为

$$\varphi(\boldsymbol{x})=\begin{cases}1,当t<\dfrac{1}{2\lambda_0}\chi^2_{2n}\left(1-\dfrac{\alpha}{2}\right)或t>\dfrac{1}{2\lambda_0}\chi^2_{2n}\left(\dfrac{\alpha}{2}\right)\\0,其他\end{cases}$$

（2）

$L_\Theta(\boldsymbol{x},\lambda)$与上一问等同，对于$L_{\Theta_0}(\boldsymbol{x},\lambda)$,令$g(t)=t^n\exp\left\{-tn\bar{x}\right\}$,$g'(t)=\left(1-\bar{x}t\right)nt^{n-1}e^{-tn\bar{x}}$,

故当$\lambda_0<\bar{x}^{-1}$时，$L_{\Theta_0}(\boldsymbol{x},\lambda)=\lambda_0^n\exp\left\{-\lambda_0 n\bar{x}\right\}$,$\lambda_0\geqslant\bar{x}^{-1}$时，$L_{\Theta_0}(\boldsymbol{x},\lambda)=L_\Theta(\boldsymbol{x},\lambda)$，故

$$若记t=n\bar{x},\lambda(\boldsymbol{x})=\begin{cases}\dfrac{n^n}{\lambda_0^n e^n t^n}e^{\lambda_0 t},\lambda_0<\bar{x}^{-1}\\1,\qquad\qquad\lambda_0\geqslant\bar{x}^{-1}\end{cases}$$

令$g(t)=t^{-n}e^{\lambda_0 t}$,则$g'(t)=(\lambda_0 t-n)t^{-n-1}e^{\lambda_0 t}$,当$\lambda_0<\bar{x}^{-1}$,即$g'(t)<0$,故检验的否定域为$D=(\boldsymbol{X}:\lambda(\boldsymbol{x})>c)=(\boldsymbol{X}:T<c')$,即$P(T<c'|H_0)=\alpha$,又$2\lambda_0 n\overline{X}\sim\chi^2_{2n}$,故

$$\varphi(\boldsymbol{x})=\begin{cases}1,n\overline{X}<\dfrac{1}{2\lambda_0}\chi^2_{2n}(1-\alpha)\\0,n\overline{X}\geqslant\dfrac{1}{2\lambda_0}\chi^2_{2n}(1-\alpha)\end{cases}$$

为此检验问题检验水平为α的似然比检验.

36、设$X_1,\cdots,X_n$为取自下列指数分布总体的样本：

$$f(x,\mu)=\exp\{-(x-\mu)\},x\geqslant\mu,-\infty<\mu<+\infty.$$

求检验问题$H_0:\mu=\mu_0\leftrightarrow H_1:\mu\neq\mu_0$的水平为$\alpha$的似然比检验.

解：

似然函数为

$$L(\boldsymbol{x},\mu)=\exp\left\{-\sum_{i=1}^n x_i+n\mu\right\}$$

参数空间为$\Theta=\{\mu:-\infty<\mu<+\infty\}$，$H_0$对应的$\Theta$的子集为$\Theta_0=\{\mu:\mu=\mu_0\}$

易知$X_{(1)}$为μ的MLE. 故

$$L_{\Theta}(\boldsymbol{x},\lambda)=\exp\left\{-\sum_{i=1}^{n}x_i+nx_{(1)}\right\}$$
$$L_{\Theta_0}(\boldsymbol{x},\lambda)=\exp\left\{-\sum_{i=1}^{n}x_i+n\mu_0\right\}$$

由此可得

$$\lambda(\boldsymbol{x})=\exp\{n(x_{(1)}-\mu_0)\}$$

$\lambda(\boldsymbol{x})$是关于$X_{(1)}$的增函数， 故否定域为$D=\{\boldsymbol{X}:\lambda(\boldsymbol{x})>c\}=\{\boldsymbol{X}:X_{(1)}>c'\}$

即$P(X_{(1)}>c'|H_0)=\alpha$，设$T=X_{(1)}$,则密度函数$g(t)=n\exp\{-n(t-\mu_0)\}$,故
$\int_{c'}^{+\infty}n\exp\{-n(t-\mu_0)\}dt=\alpha$,则$e^{-n(c'-\mu_0)}=\alpha$,解得$c'=\mu_0-\dfrac{\ln\alpha}{n}$.

故检验问题的检验水平为α的似然比检验为

$$\varphi(\boldsymbol{x})=\begin{cases}1, X_{(1)}>\mu_0-\dfrac{\ln\alpha}{n}\\ 0, X_{(1)}\leqslant\mu_0-\dfrac{\ln\alpha}{n}\end{cases}$$

37、设$X_1,\cdots,X_m$为从指数分布$f(x,\theta_1)=\theta_1^{-1}\exp\{-x/\theta_1\}$抽取的简单样本，$Y_1,\cdots,Y_n$从指数分布$f(x,\theta_2)=\theta_2^{-1}\exp\{-x/\theta_2\}$抽取的简单样本,且两组样本独立，$\theta_1$和$\theta_2$为未知参数.求检验问题$H_0:\theta_1=\theta_2\leftrightarrow H_1:\theta_1\neq\theta_2$的似然比检验.

解:
联合似然函数为

$$L(\boldsymbol{x},\boldsymbol{y},\theta_1,\theta_2)=\theta_1^{-m}\theta_2^{-n}\exp\left\{-\sum_{i=1}^{m}x_i\Big/\theta_1-\sum_{j=1}^{n}y_j\Big/\theta_2\right\}$$

参数空间为

$$\Theta=\{(\theta_1,\theta_2):-\infty<\theta_1,\theta_2<+\infty\}$$

零假设H_0对应的参数空间Θ_0的子集为

$$\Theta_0=\{(\theta_1,\theta_2):-\infty<\theta_1=\theta_2=\theta<+\infty\}$$

计算可得

$$L_{\Theta}(\boldsymbol{x},\boldsymbol{y},\theta_1,\theta_2)=\left(\bar{x}\right)^{-m}\left(\bar{y}\right)^{-n}e^{-m-n}$$

$$L_{\Theta_0}(\boldsymbol{x},\boldsymbol{y},\theta_1,\theta_2)=\left(\frac{\sum\limits_{i=1}^{m}x_i+\sum\limits_{j=1}^{n}y_j}{m+n}\right)^{-m-n}\mathrm{e}^{-m-n}$$

故$\lambda(\boldsymbol{x})=\dfrac{1}{(m+n)^{m+n}}(1+T)^m\left(1+\dfrac{1}{T}\right)^n,T=\sum\limits_{j=1}^{n}Y_j\Big/\sum\limits_{i=1}^{m}X_i$

令$g(t)=(1+t)^m\left(1+\dfrac{1}{t}\right)^n,g'(t)=(mt-n)(1+t)^{m+n-1}t^{-n-1}$,即$g(t)$先减后增,

故$D=\{\boldsymbol{X}:\lambda(\boldsymbol{x})>c\}=\{\boldsymbol{X}:T<k_1\cup T>k_2\}$,故$P(T<k_1\cup T>k_2|H_0)=\alpha$,

又因为H_0成立时,$T\dfrac{m}{n}=\dfrac{2\theta\sum\limits_{j=1}^{n}Y_j}{2\theta\sum\limits_{i=1}^{m}X_i}\cdot\dfrac{2m}{2n}\sim F_{2n,2m}$,故

$$k_1=\frac{n}{m}F_{2n,2m}\left(1-\frac{\alpha}{2}\right),k_2=\frac{n}{m}F_{2n,2m}\left(\frac{\alpha}{2}\right)$$

故检验问题的检验水平为α的似然比检验为

$$\varphi(\boldsymbol{x})=\begin{cases}1,T<\dfrac{n}{m}F_{2n,2m}\left(1-\dfrac{\alpha}{2}\right)\text{或}T>\dfrac{n}{m}F_{2n,2m}\left(\dfrac{\alpha}{2}\right)\\0,\text{其他}\end{cases}$$

38、设$X_1,\cdots,X_n\ i.i.d.\sim B(1,p)$,用大样本方法求检验问题$H:p=p_0\leftrightarrow K:p\neq p_0$的似然比检验(取水平为$\alpha$).

解:

似然函数为$L(\boldsymbol{x},p)=p^{\sum\limits_{i=1}^{n}x_i}(1-p)^{n-\sum\limits_{i=1}^{n}x_i}$

参数空间为$\Theta=\{p:p>0\}$,其维数为1,零假设对应的Θ的子空间为$\Theta_0=\{p:p=p_0\}$,其维数为0,易知

$$L_{\Theta}(\boldsymbol{x},p)=\overline{X}^{\sum\limits_{i=1}^{n}x_i}\left(1-\overline{X}\right)^{n-\sum\limits_{i=1}^{n}x_i}$$

$$L_{\Theta_0}(\boldsymbol{x},p)=p_0^{\sum\limits_{i=1}^{n}x_i}(1-p_0)^{n-\sum\limits_{i=1}^{n}x_i}$$

因此可得

$$\lambda(\boldsymbol{x})=\left(\frac{\overline{X}}{p_0}\right)^{\sum\limits_{i=1}^{n}x_i}\left(\frac{1-\overline{X}}{1-p_0}\right)^{n-\sum\limits_{i=1}^{n}x_i}$$

令$Y(\boldsymbol{x})=2\ln\lambda(\boldsymbol{x})=\sum\limits_{i=1}^{n}x_i\left(\ln\overline{X}-\ln p_0\right)+\left(n-\sum\limits_{i=1}^{n}x_i\right)\left[\ln\left(1-\overline{X}\right)-\ln(1-p_0)\right]$

故当零假设成立且$n\to\infty$时

$$Y(\boldsymbol{x})\xrightarrow{\mathscr{L}}\chi_1^2$$

故得到大样本检验有水平近似为α的否定域

$$D=\{\boldsymbol{X}:Y(\boldsymbol{x})>\chi_1^2(\alpha)\}$$

39、设$X_1,\cdots,X_n$ $i.i.d.$服从成功概率为p的几何分布，用大样本方法检验问题H_0: $p=p_0\leftrightarrow H_1:p\neq p_0$的检验水平为$\alpha$的似然比检验.

解：

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

参数空间为$\Theta=\{p:p>0\}$,其维数为1，零假设对应的Θ的子空间为$\Theta_0=\{p:p=p_0\}$,其维数为0,易知

$$L_\Theta(\boldsymbol{x},p)=\overline{X}^n\left(1-\overline{X}\right)^{\sum_{i=1}^n x_i-n}$$

$$L_{\Theta_0}(\boldsymbol{x},p)=p_0^n(1-p_0)^{\sum_{i=1}^n x_i-n}$$

因此可得

$$\lambda(\boldsymbol{x})=\left(\frac{\overline{X}}{p_0}\right)^n\left(\frac{1-\overline{X}}{1-p_0}\right)^{\sum_{i=1}^n x_i-n}$$

令$Y(\boldsymbol{x})=2\ln\lambda(\boldsymbol{x})=n\left(\ln\overline{X}-\ln p_0\right)+\left(\sum_{i=1}^n x_i-n\right)\left[\ln\left(1-\overline{X}\right)-\ln(1-p_0)\right]$

故当零假设成立且$n\to\infty$时

$$Y(\boldsymbol{x})\xrightarrow{\mathscr{L}}\chi_1^2$$

故得到大样本检验有水平近似为α的否定域

$$D=\{\boldsymbol{X}:Y(\boldsymbol{x})>\chi_1^2(\alpha)\}$$

40、设$0<\alpha<1$,φ是检验问题$H_0:\theta=\theta_0\leftrightarrow K:\theta=\theta_1$的水平为$\alpha$的$UMPT$,且假定$\beta=E_K[\varphi(X)]<1$,则$1-\varphi$是检验问题$K$对$H$的水平为$1-\beta$的$UMPT$.

解：

φ是检验问题$H_0:\theta=\theta_0\leftrightarrow K:\theta=\theta_1$的水平为$\alpha$的$UMPT$，则$\exists c\geqslant 0, 0\leqslant r\leqslant 1$,使

(i)

$$\varphi(\boldsymbol{x})=\begin{cases}1\,,f(\boldsymbol{x},\theta_1)/f(\boldsymbol{x},\theta_0)>c,\\ r,f(\boldsymbol{x},\theta_1)/f(\boldsymbol{x},\theta_0)=c,\\ 0\,,f(\boldsymbol{x},\theta_1)/f(\boldsymbol{x},\theta_0)<c.\end{cases}$$

(ii)

$$E_{\theta_0}[\varphi(X)]=\alpha$$

下证明$c\neq0$,利用反证法

设$R^+=\{x{:}\,f(\boldsymbol{x},\theta_1)/f(\boldsymbol{x},\theta_0)>c\}=\{x{:}f(\boldsymbol{x},\theta_1)>c\}$，因此

$$\begin{aligned}E_{\theta_1}[\phi(X)]&=\int_{\chi}\phi(x)f(x,\theta_1)dx=\int_{R^+}\phi(x)f(x,\theta_1)dx\\&=\int_{R^+}f(x,\theta_1)dx=\int_{\chi}f(x,\theta_1)dx=1.\end{aligned}$$

与$E_{\theta_1}[\phi(X)]<1$相矛盾，故$c\neq0\,,c>0$,故$1/c\in(0,+\infty)$.

则对于检验问题K对H,有$c'=1/c,r'=1-r$

$$\varphi'(\boldsymbol{x})=1-\varphi(\boldsymbol{x})=\begin{cases}1\,,f(\boldsymbol{x},\theta_0)/f(\boldsymbol{x},\theta_1)>c'\\ r',f(\boldsymbol{x},\theta_0)/f(\boldsymbol{x},\theta_1)=c'\\ 0\,,f(\boldsymbol{x},\theta_0)/f(\boldsymbol{x},\theta_1)<c'\end{cases}$$

且

$$E_{\theta_1}[1-\varphi(X)]=1-\beta$$

故$1-\varphi$是检验问题K对H的水平为$1-\beta$的$UMPT$

41、设$\boldsymbol{X}=(X_1,\cdots,X_n)$为自总体分布族$\{f(x,\theta),\theta\in\Theta\}$中抽取的随机样本，且$T=T(X_1,\cdots,X_n)$为$\theta$的充分统计量，则由$NP$引理确定的检验函数$\varphi(\boldsymbol{x})$为充分统计量$T$的函数.

解：

利用因子分解定理，$T(\boldsymbol{X})$为充分统计量故

$$f(\boldsymbol{x},\theta)=g(t(\boldsymbol{x}),\theta)h(\boldsymbol{x})$$

对于检验问题$H_0{:}\theta=\theta_0\leftrightarrow H_1{:}\theta=\theta_1$,有

$$f(\boldsymbol{x},\theta_0)=g(t(\boldsymbol{x}),\theta_0)h(\boldsymbol{x})$$
$$f(\boldsymbol{x},\theta_1)=g(t(\boldsymbol{x}),\theta_1)h(\boldsymbol{x})$$

故有

$$f(\boldsymbol{x},\theta_1)/f(\boldsymbol{x},\theta_0)=g(t(\boldsymbol{x}),\theta_1)/g(t(\boldsymbol{x}),\theta_0)$$

故检验函数

$$\varphi(\boldsymbol{x})=\begin{cases}1, g(t(\boldsymbol{x}),\theta_1)/g(t(\boldsymbol{x}),\theta_0)>c,\\ r, g(t(\boldsymbol{x}),\theta_1)/g(t(\boldsymbol{x}),\theta_0)=c,\\ 0, g(t(\boldsymbol{x}),\theta_1)/g(t(\boldsymbol{x}),\theta_0)<c.\end{cases}$$

检验函数$\varphi(\boldsymbol{x})$为充分统计量T的函数.

42、设$X_1,\cdots,X_n$ $i.i.d.$服从均匀分布$U(0,\theta)$，试求

(1)检验问题$H_0:\theta=1\leftrightarrow H_1:\theta=2$检验水平为$\alpha$的$UMPT$;

(2)检验问题$H_0':\theta=1\leftrightarrow H_1':\theta=1/2$检验水平为$\alpha$的$UMPT$.

解：

(1)服从均匀分布的样本$\boldsymbol{X}$的密度函数和似然比分别为

$$f(\boldsymbol{x},\theta)=\frac{1}{\theta^n}I_{(0,\theta)}(x_{(n)}),$$

$$\lambda(\boldsymbol{x})=\frac{f(\boldsymbol{x},\theta_1)}{f(\boldsymbol{x},\theta_0)}=\begin{cases}\dfrac{1}{2^n}, 0<x_{(n)}<1,\\ \infty, x_{(n)}\geqslant 1.\end{cases}$$

$\lambda(\boldsymbol{x})$关于$T(\boldsymbol{x})=x_{(n)}$非降，故由$NP$引理，可知水平为$\alpha$的$UMPT$函数有形式

$$\varphi(\boldsymbol{x})=\begin{cases}1, x_{(n)}>c,\\ 0, x_{(n)}\leqslant c.\end{cases}$$

$T=X_{(n)}$的密度函数为

$$g_\theta(t)=\frac{nt^{n-1}}{\theta^n}I_{(0,\theta)}(t).$$

故当H_0成立时，$g(t)=nt^{n-1}I_{(0,1)}(t)$,故有

$$E_{\theta_0}[\varphi(\boldsymbol{X})]=\int_0^\infty\varphi(t)g(t)dt=\int_c^1 nt^{n-1}dt=1-c^n=\alpha$$

故得$c=\sqrt[n]{1-\alpha}$,因此

$$\varphi(\boldsymbol{x})=\begin{cases}1, x_{(n)}>\sqrt[n]{1-\alpha},\\ 0, x_{(n)}\leqslant\sqrt[n]{1-\alpha}.\end{cases}$$

为一个水平为α的$UMPT$.

(2)与(1)类似求出似然比为

$$\lambda(\boldsymbol{x})=\frac{f(\boldsymbol{x},\theta_1)}{f(\boldsymbol{x},\theta_0)}=\begin{cases}2^n,0<x_{(n)}<1/2,\\0,x_{(n)}\geqslant 1/2.\end{cases}$$

$\lambda(\boldsymbol{x})$关于$T(\boldsymbol{x})=x_{(n)}$非增，故由$NP$引理，可知水平为$\alpha$的$UMPT$函数有形式

$$\varphi(\boldsymbol{x})=\begin{cases}1,x_{(n)}<c,\\0,x_{(n)}\geqslant c.\end{cases}$$

$T=X_{(n)}$的密度函数为

$$g_\theta(t)=\frac{nt^{n-1}}{\theta^n}I_{(0,\theta)}(t).$$

故当H_0'成立时，$g(t)=nt^{n-1}I_{(0,1)}(t)$,故有

$$E_{\theta_0}[\varphi(\boldsymbol{X})]=\int_0^\infty\varphi(t)g(t)dt=\int_0^c nt^{n-1}dt=c^n=\alpha$$

故得$c=\sqrt[n]{\alpha}$,因此

$$\varphi(\boldsymbol{x})=\begin{cases}1,x_{(n)}<\sqrt[n]{\alpha},\\0,x_{(n)}\geqslant\sqrt[n]{\alpha}.\end{cases}$$

为一个水平为α的$UMPT$.

43、设$X_1,\cdots,X_n\ i.i.d.\sim N(\mu,1)$,对水平$\alpha$，试求检验问题$H_0:\mu\geqslant\mu_0\leftrightarrow H_1:\mu<\mu_0$的$UMPT$,其中$\mu_0$和$\alpha$给定.

解：

正态分布为指数族分布，样本密度为

$$\begin{aligned}f(\boldsymbol{x},\mu)&=(2\pi)^{-\frac{n}{2}}\exp\{-n\mu^2/2\}\exp\left\{n\mu\bar{x}\right\}\exp\left\{-\sum_{i=1}^n x_i^2\Big/2\right\}\\&=c(\mu)\exp\{Q(\mu)T(\boldsymbol{x})\}h(\boldsymbol{x}),\end{aligned}$$

其中$c(\mu)=(2\pi)^{-n/2}\exp\{-n\theta^2/2\},h(\boldsymbol{x})=\exp\left\{-\sum_{i=1}^n x_i^2/2\right\},T(\boldsymbol{x})=\bar{x},Q(\mu)=n\mu$为$\mu$的严格增函数，由定理5.4.3可知水平为$\alpha$的$UMPT$由下式给出：

$$\varphi(\boldsymbol{x})=\begin{cases}1,&T(\boldsymbol{x})<c,\\0,&T(\boldsymbol{x})\geqslant c.\end{cases}$$

由于$T(\boldsymbol{X})=\overline{X}\sim N(\mu,1/n)$,故$\sqrt{n}\left(\overline{X}-\mu\right)\sim N(0,1)$,令

$$\alpha=E_{\mu_0}[\varphi(\boldsymbol{X})]=P_{\mu_0}(T(\boldsymbol{X})<c)=P_{\mu_0}\left(\sqrt{n}\left(\overline{X}-\mu_0\right)<\sqrt{n}(c-\mu_0)\right).$$

可知$\sqrt{n}(c-\mu_0)=u_{1-\alpha}$,即$c=\mu_0+u_{1-\alpha}/\sqrt{n}$.故水平为$\alpha$的$UMPT$为

$$\varphi(\boldsymbol{x})=\begin{cases}1, & \bar{x}<\mu_0+u_{1-\alpha}/\sqrt{n}\\ 0, & \bar{x}\geqslant\mu_0+u_{1-\alpha}/\sqrt{n}\end{cases}$$

44、设$X_1,\cdots,X_n$ $i.i.d.\sim Poisson$分布$P(\lambda)$，对水平α，试求检验问题$H_0:\lambda\geqslant\lambda_0\leftrightarrow H_1:\lambda<\lambda_0$的$UMPT$,其中$\lambda_0$和$\alpha$给定.

解：

泊松分布为指数族.样本$\boldsymbol{X}$的密度函数为

$$f(\boldsymbol{x},\lambda)=\frac{\lambda^{T(\boldsymbol{x})}e^{-n\lambda}}{\prod_{i=1}^{n}x_i!}=c(\lambda)\exp\{Q(\lambda)T(\boldsymbol{x})\}h(\boldsymbol{x}),$$

其中$c(\lambda)=e^{-n\lambda}$,$T(\boldsymbol{x})=\sum_{i=1}^{n}x_i$,$h(\boldsymbol{x})=1\Big/\prod_{i=1}^{n}x_i!$，$Q(\lambda)=\log\lambda$为$\lambda$的严格增函数，故

$$\varphi(\boldsymbol{x})=\begin{cases}1, & T(\boldsymbol{x})<c,\\ r, & T(\boldsymbol{x})=c,\\ 0, & T(\boldsymbol{x})>c.\end{cases}$$

因为$T(\boldsymbol{X})=\sum_{i=1}^{n}x_i\sim P(n\lambda)$,$c$由下列不等式确定

$$\alpha_1=\sum_{k=0}^{c-1}\frac{(n\lambda_0)^k e^{-n\lambda_0}}{k!}\leqslant\alpha\leqslant\sum_{k=0}^{c}\frac{(n\lambda_0)^k e^{-n\lambda_0}}{k!}$$

取r为

$$r=\frac{(\alpha-\alpha_1)c!}{(n\lambda_0)^c e^{-n\lambda_0}}$$

则必有

$$E_{\lambda_0}[\varphi(\boldsymbol{X})]=P_{\lambda_0}(T(\boldsymbol{X})<c)+r\cdot P_{\lambda_0}(T(\boldsymbol{X})=c)=\alpha$$

故上述检验$\varphi(\boldsymbol{x})$为水平为α的$UMPT$.

45、设$X_1,\cdots,X_n$ $i.i.d.\sim N(0,\sigma^2)$,对水平α，试求检验问题$H_0:\sigma^2\leqslant\sigma_0^2\leftrightarrow H_1:\sigma^2>\sigma_0^2$的$UMPT$,其中$\sigma_0^2$和$\alpha$给定.

解：

正态分布为指数族分布，样本密度为

$$f(\boldsymbol{x},\sigma^2)=(2\pi\sigma^2)^{-\frac{n}{2}}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}x_i^2\right\}$$
$$=c(\sigma^2)\exp\{Q(\sigma^2)T(\boldsymbol{x})\}h(\boldsymbol{x})$$

其中$c(\sigma^2)=(2\pi\sigma^2)^{-n/2}$,，$T(\boldsymbol{x})=\sum_{i=1}^{n}x_i^2,h(\boldsymbol{x})\equiv 1,Q(\sigma^2)=-1/2\sigma^2$为$\sigma^2$的严格增函数，由定理5.4.2可知水平为$\alpha$的$UMPT$为

$$\varphi(\boldsymbol{x})=\begin{cases}1, & T(\boldsymbol{x})<c,\\ 0, & T(\boldsymbol{x})\geqslant c.\end{cases}$$

由于$T(\boldsymbol{X})/\sigma^2=\sum_{i=1}^{n}X_i^2\Big/\sigma^2\sim\chi_n^2$,令

$$\alpha=E_{\sigma_0^2}[\varphi(\boldsymbol{X})]=P_{\sigma_0^2}[T(\boldsymbol{X})>c]=P_{\sigma_0^2}\left[\sum_{i=1}^{n}x_i^2\Big/\sigma_0^2>c/\sigma_0^2\right]$$

故可知$c/\sigma_0^2=\chi_n^2(\alpha)$,故$c=\chi_n^2(\alpha)\sigma_0^2$,故水平为$\alpha$的$UMPT$为

$$\varphi(\boldsymbol{x})=\begin{cases}1, & \sum_{i=1}^{n}x_i^2>\chi_n^2(\alpha)\sigma_0^2,\\ 0, & \sum_{i=1}^{n}x_i^2>\chi_n^2(\alpha)\sigma_0^2.\end{cases}$$

46、设$X_1,\cdots,X_n\ i.i.d.\sim b(1,p)$,对水平$\alpha$，试求检验问题$H_0:p\geqslant 1/2\leftrightarrow H_1:p<1/2$的$UMPT$,其中$\alpha$给定.

解：

二项分布族为指数族，故概率分布为

$$f(\boldsymbol{x},p)=p^{\sum_{i=1}^{n}x_i}(1-p)^{n-\sum_{i=1}^{n}x_i}=c(p)\exp\{Q(p)T(\boldsymbol{x})\}h(\boldsymbol{x})$$

其中$c(p)=(1-p)^n,T(\boldsymbol{x})=\sum_{i=1}^{n}x_i,h(\boldsymbol{x})\equiv 1,Q(p)=\ln(p/1-p)$为$p$的严格单增函数，故由定理5.4.2可知

$$\varphi(\boldsymbol{x})=\begin{cases}1, & T(\boldsymbol{x})<c,\\ r, & T(\boldsymbol{x})=c,\\ 0, & T(\boldsymbol{x})>c.\end{cases}$$

已知$T(\boldsymbol{X})\sim b(n,p)$，故$c$由下列不等式决定

$$\alpha_1=\sum_{k=c+1}^{n}\binom{n}{k}p^k(1-p)^{n-k}\leqslant\alpha\leqslant\sum_{k=c}^{n}\binom{n}{k}p^k(1-p)^{n-k}$$

取r为

$$r=\frac{\alpha-\alpha_1}{\binom{n}{c}p^c(1-p)^{n-c}}.$$

则必有

$$E_{p_0}[\varphi(\boldsymbol{X})]=P_{p_0}(T(\boldsymbol{X})<c)+r\cdot P_{p_0}(T(\boldsymbol{X})=c)=\alpha$$

由此得到上述检验$\varphi(\boldsymbol{x})$为水平为α的$UMPT$.

47、设k为已知的自然数.为检验一事件的概率$p\leqslant p_0$是否成立，将试验独立地重复下去，直到该事件发生k次为止.以X记到那时为止的试验次数.

(1)证明X的分布为

$$P(X=x)=\binom{x-1}{k-1}p^k(1-p)^{x-k},\quad x=k,k+1,\cdots;$$

(2)求$H_0:p\leqslant p_0\leftrightarrow H_1:p>p_0$水平为$\alpha$的$UMPT$,其中$p_0$和$\alpha$给定.

解：

(1)可以理解为前$x-1$次试验服从二项分布发生了$k-1$次，第x次发生则

$$P(X=x)=\binom{x-1}{k-1}p^k(1-p)^{x-k},\quad x=k,k+1,\cdots;$$

(2)负二项分布为指数族分布，故概率密度为

$$f(\boldsymbol{x},p)=\prod_{i=1}^{n}\binom{x_i-1}{k-1}p^k(1-p)^{x_i-k}=c(p)\exp\{Q(p)T(\boldsymbol{x})\}h(\boldsymbol{x})$$

其中$c(p)=p^{nk}(1-p)^{-nk},T(\boldsymbol{x})=\sum\limits_{i=1}^{n}x_i,h(\boldsymbol{x})=\prod\limits_{i=1}^{n}\binom{x_i}{k},Q(p)=\ln(1-p)$是$p$的严格减函数，故

$$\varphi(\boldsymbol{x})=\begin{cases}1, & T(\boldsymbol{x})<c,\\ r, & T(\boldsymbol{x})=c,\\ 0, & T(\boldsymbol{x})>c.\end{cases}$$

记(1)中分布为$Pascal(k)$,则$\sum\limits_{i=1}^{n}x_i\sim Pascak(nk)$,故$c$可由下式给定，

$$\alpha_1=\sum_{i=nk}^{c-1}\binom{i-1}{nk-1}p^{nk}(1-p)^{i-nk}\leqslant\alpha\leqslant\sum_{i=nk}^{c}\binom{i-1}{nk-1}p^{nk}(1-p)^{i-nk}$$

取r为

$$r=\frac{\alpha-\alpha_1}{\binom{c-1}{nk-1}p^{nk}(1-p)^{c-nk}}.$$

则必有

$$E_{p_0}[\varphi(\boldsymbol{X})]=P_{p_0}(T(\boldsymbol{X})<c)+r\cdot P_{p_0}(T(\boldsymbol{X})=c)=\alpha$$

由此得到上述检验$\varphi(\boldsymbol{x})$为水平为α的$UMPT$.

48、设$X_1,\cdots,X_n$为自指数分布总体$Exp(\lambda)$中抽取的随机样本，$\lambda>0$为未知参数.

求$H_0:\lambda\geqslant\lambda_0\leftrightarrow H_1:\lambda<\lambda_0$的水平为$\alpha$的$UMPT$,其中$\lambda_0$和$\alpha$给定.

解：

指数分布为指数族分布，故联合密度函数为

$$f(\boldsymbol{x},\lambda)=\lambda^n\exp\left\{-\lambda\sum_{i=1}^{n}x_i\right\}I_{(0,\infty)}(x_{(1)})=c(\lambda)\exp\{Q(\lambda)T(\boldsymbol{x})\}h(\boldsymbol{x})$$

其中$c(\lambda)=\lambda^n,T(\boldsymbol{x})=\sum\limits_{i=1}^{n}x_i,h(\boldsymbol{x})\equiv I_{(0,\infty)}(x_{(1)}),Q(\lambda)=-\lambda$是$\lambda$的严格单调减函数，故

$$\varphi(\boldsymbol{x})=\begin{cases}1, & T(\boldsymbol{x})>c,\\ 0, & T(\boldsymbol{x})\leqslant c.\end{cases}$$

其中$2\lambda\sum\limits_{i=1}^{n}x_i\sim\chi^2_{2n}$,故

$$\alpha=E_{\lambda_0}[\varphi(\boldsymbol{X})]=P_{\lambda_0}(T(\boldsymbol{X})>c)=P_{\lambda_0}(2\lambda_0T(\boldsymbol{X})>2\lambda_0c)$$

故$2\lambda_0c=\chi^2_{2n}(\alpha),c=\chi^2_{2n}(\alpha)/2\lambda_0$,故

$$\varphi(\boldsymbol{x})=\begin{cases}1, & T(\boldsymbol{x})>\chi^2_{2n}(\alpha)/2\lambda_0,\\ 0, & T(\boldsymbol{x})\leqslant\chi^2_{2n}(\alpha)/2\lambda_0.\end{cases}$$

故上述检验$\varphi(\boldsymbol{x})$为水平为α的$UMPT$.

49、设$X_1,\cdots,X_n$表示受试的n个元件寿命，他们皆服从指数分布$Exp(\lambda)$.而实际上只试验到有r个失效时即停止.这r个寿命从小到大排列为$X_{(1)}\leqslant\cdots\leqslant X_{(r)}$.试求$H_0:\lambda\leqslant\lambda_0\leftrightarrow H_1:\lambda>\lambda_0$的基于$X_{(1)}\cdots X_{(r)}$水平为$\alpha$的$UMPT$（提示：由例3.3.5可证$T=\sum\limits_{i=1}^{r}x_{(r)}+(n-r)X_{(r)}$为充分统计量，且$2\lambda T\sim\chi^2_{2r}$）.

解：

由例3.3.5可知，令$t_i=X_{(i)}$，$X_{(1)}\cdots X_{(r)}$的联合密度函数为

$$f(\boldsymbol{t},\lambda)=\frac{n!}{(n-r)!}\lambda^r\exp\left\{-\lambda\left(\sum_{i=1}^r t_i+(n-r)t_r\right)\right\}I_{(0,\infty)}(t_1)$$
$$=c(\lambda)\exp\{Q(\lambda)T(\boldsymbol{t})\}h(\boldsymbol{t})$$

其中$c(\lambda)=\dfrac{n!}{(n-r)!}\lambda^r, T(\boldsymbol{t})=\sum\limits_{i=1}^r t_i+(n-r)t_r, h(\boldsymbol{t})\equiv I_{(0,\infty)}(t_1), Q(\lambda)=-\lambda$是$\lambda$的严格单调减函数，故

$$\varphi(\boldsymbol{t})=\begin{cases}1, & T(\boldsymbol{t})<c,\\ 0, & T(\boldsymbol{t})\geqslant c.\end{cases}$$

由习题2.26可知$2\lambda T\sim\chi^2_{2r}$,故

$$\alpha=E_{\lambda_0}[\varphi(\boldsymbol{T})]=P_{\lambda_0}(T(\boldsymbol{T})<c)=P_{\lambda_0}(2\lambda_0T(\boldsymbol{T})<2\lambda_0c)$$

解得$2\lambda_0c=\chi^2_{2r}(1-\alpha)$， $c=\chi^2_{2r}(1-\alpha)/2\lambda_0$

$$\varphi(\boldsymbol{t})=\begin{cases}1, & T(\boldsymbol{t})<\chi^2_{2r}(1-\alpha)/2\lambda_0,\\ 0, & T(\boldsymbol{t})\geqslant\chi^2_{2r}(1-\alpha)/2\lambda_0.\end{cases}$$

故上述检验$\varphi(\boldsymbol{t})$为水平为α的$UMPT$.

50、设$\varphi(\boldsymbol{x})$为一检验函数，而$T=T(\boldsymbol{x})$为充分统计量.证明$E_\theta[\varphi(\boldsymbol{X})|T]$也是检验函数，且其功效函数与$\varphi$同.这个结果说明了什么问题？（提示：由充分性知$E_\theta[\varphi(\boldsymbol{X})|T]$与$\theta$无关，再由$\varphi$是检验函数，证明$E_\theta[\varphi(\boldsymbol{X})|T]$也是检验函数.）

解：

记$h(T)=E_\theta[\varphi(\boldsymbol{X})|T]$,由于$T$是充分统计量，故给定$T$时，$X$的条件分布与$\theta$无关，所以$h(T)$与$\theta$无关，由于$\varphi$是检验函数，故$\varphi(\boldsymbol{x})$取值于$[0,1]$，故$h(T)$也取值于$[0,1]$，故$E_\theta[\varphi(\boldsymbol{X})|T]$也是检验函数，又功效函数

$\beta_\theta(h(T))=E_\theta[h(T)]=E_\theta\{E_\theta[\varphi(\boldsymbol{X})|T]\}=E_\theta[\varphi(\boldsymbol{X})]=\beta_\theta[\varphi(\boldsymbol{X})]$,故$h(T)$的功效函数与$\varphi$相同，这个结果说明欲找$UMPT$，只需在关于$T$的函数中寻找.

51、设$X_1,\cdots,X_m\ i.i.d.\sim N(a,1), Y_1,\cdots,Y_n\ i.i.d.\ \sim N(b,1)$，且样本$X_1,\cdots,X_m$和$Y_1,\cdots,Y_n$独立，证明$H_0{:}b\leqslant a\leftrightarrow H_1{:}b>a$水平为$\alpha$的$UMPT$存在，否定域为$\overline{Y}-\overline{X}>c$（提示：利用上题结果）

解：

先考虑以下假设$H_0''{:}a=b=\mu=\dfrac{ma_1+nb_1}{m+n}\leftrightarrow H_1''{:}a=a_1,b=b_1$（$b_1>a_1$）

$$f_0(\boldsymbol{x},\boldsymbol{y})=(2\pi)^{-\frac{m+n}{2}}\exp\left\{-\frac{\sum_{i=1}^{m}(x_i-\mu)^2+\sum_{j=1}^{n}(y_j-\mu)^2}{2}\right\}$$

$$=(2\pi)^{-\frac{m+n}{2}}\exp\left\{-\frac{\sum_{i=1}^{m}x_i^2-2\mu\sum_{i=1}^{m}x_i+m\mu^2+\sum_{j=1}^{n}y_j^2-2\mu\sum_{j=1}^{n}y_j+n\mu^2}{2}\right\}$$

$$f_1(\boldsymbol{x},\boldsymbol{y})=(2\pi)^{-\frac{m+n}{2}}\exp\left\{-\frac{\sum_{i=1}^{m}(x_i-a_1)^2+\sum_{j=1}^{n}(y_j-b_1)^2}{2}\right\}$$

$$=(2\pi)^{-\frac{m+n}{2}}\exp\left\{-\frac{\sum_{i=1}^{m}x_i^2-2a_1\sum_{i=1}^{m}x_i+ma_1^2+\sum_{j=1}^{n}y_j^2-2b_1\sum_{j=1}^{n}y_j+nb_1^2}{2}\right\}$$

故似然比可表示为

$$\lambda(\boldsymbol{x},\boldsymbol{y})=\exp\left\{\frac{2(a_1-\mu)\sum_{i=1}^{m}x_i+m(\mu^2-a_1^2)+2(b_1-\mu)\sum_{j=1}^{n}y_j+n(\mu^2-b_1^2)}{2}\right\}$$

$$=\exp\left\{\frac{mn(b_1-a_1)}{m+n}\left(\bar{y}-\bar{x}\right)-\frac{mn}{m+n}(b_1-a_1)^2\right\}$$

故检验函数为

$$\varphi(\boldsymbol{x},\boldsymbol{y})=\begin{cases}1,\lambda(\boldsymbol{x},\boldsymbol{y})>c\\0,\lambda(\boldsymbol{x},\boldsymbol{y})\leqslant c\end{cases}=\begin{cases}1,\overline{Y}-\overline{X}>c'\\0,\overline{Y}-\overline{X}\leqslant c'\end{cases}$$

为上述检验问题的检验函数，由于$\varphi(\boldsymbol{x},\boldsymbol{y})$与$a_1,b_1$无关，故上述检验函数也是

$H_0':a=b\leftrightarrow H_1':b>a$的水平为$\alpha$的$UMPT$,由于$\overline{X}$和$\overline{Y}$是充分统计量，

利用上题的结论欲求$H_0:b\leqslant a\leftrightarrow H_1:b>a$的检验函数，只$H_0$条件下$\beta_\varphi(a,b)$关于$b-a$

递增，$\beta_\varphi(a,b)=P\left(\overline{Y}-\overline{X}>c'\right)=P\left(\frac{\overline{Y}-\overline{X}-(b-a)}{\sqrt{\frac{1}{n}+\frac{1}{m}}}>\frac{c'-(b-a)}{\sqrt{\frac{1}{n}+\frac{1}{m}}}\right)$

$=1-\Phi\left(\frac{c'-(b-a)}{\sqrt{\frac{1}{n}+\frac{1}{m}}}\right)$,故$\beta_\varphi(a,b)$关于$b-a$递增.$b-a=0$时，$\sup\limits_{b\leqslant a}\beta_\varphi(a,b)=\alpha$，

解得$c=u_\alpha\sqrt{\frac{1}{n}+\frac{1}{m}}$,故题中检验问题的检验水平为$\alpha$的检验函数为

$$\varphi(\boldsymbol{x},\boldsymbol{y})=\begin{cases}1, & \overline{Y}-\overline{X}>u_{\alpha}\sqrt{\dfrac{1}{n}+\dfrac{1}{m}}\\ 0, & \overline{Y}-\overline{X}\leqslant u_{\alpha}\sqrt{\dfrac{1}{n}+\dfrac{1}{m}}\end{cases}$$

52、设$X_1,\cdots,X_n$ $i.i.d.\sim$均匀分布$U(0,\theta)$,对水平α，试求检验问题$H_0{:}\theta\leqslant\theta_0\leftrightarrow H_1{:}\theta>\theta_0$的$UMPT$,且求出该$UMPT$的功效函数.

解：

由例5.4.3可知$H_0{:}\theta=\theta_0\leftrightarrow H_1{:}\theta>\theta_0$的水平为$\alpha$的$UMPT$为

$$\varphi(\boldsymbol{x})=\begin{cases}1, x_{(n)}>\theta_0\sqrt[n]{1-\alpha}\\ 0, x_{(n)}\leqslant\theta_0\sqrt[n]{1-\alpha}\end{cases}$$

则功效函数$\beta_{\theta}[\varphi(\boldsymbol{x})]=P(X_{(n)}>c)=\int_c^{\theta}\dfrac{nt^{n-1}}{\theta^n}dt=1-\dfrac{c^n}{\theta^n}$是$\theta$的单调增函数，

故$\sup\limits_{\theta\leqslant\theta_0}\beta_{\theta}[\varphi(\boldsymbol{x})]=\beta_{\theta_0}[\varphi(\boldsymbol{x})]$，故上述检验函数也是题中所求的检验函数.

53、设$X_1,\cdots,X_n$ $i.i.d.\sim$均匀分布$U(\theta,2\theta),\theta>0$,指定$\theta_0>0$,问$H_0{:}\theta\leqslant\theta_0\leftrightarrow H_1{:}\theta>\theta_0$水平为$\alpha$的$UMPT$是否存在?此处$\theta_0$和$\alpha$给定.

解：

先考虑以下假设

$$H_0'{:}\theta=\theta_0\leftrightarrow H_1'{:}\theta=\theta_1,\theta_1>\theta_0.$$

似然函数分别为

$$f(\boldsymbol{x},\theta_0)=\frac{1}{\theta_0^n}I_{(\theta_0<x_{(1)}<x_{(n)}<2\theta_0)}$$
$$f(\boldsymbol{x},\theta_1)=\frac{1}{\theta_1^n}I_{(\theta_1<x_{(1)}<x_{(n)}<2\theta_1)}$$

似然比为

$$\lambda(\boldsymbol{x})=\begin{cases}\dfrac{\theta_0^n}{\theta_1^n},\theta_1<x_{(1)}<x_{(n)}<2\theta_0,\\ 0,\ \theta_0<x_{(1)}<x_{(n)}<\theta_1\text{或}\theta_0<x_{(1)}<\theta_1<x_{(n)}<2\theta_0\\ \infty,\text{其他}.\end{cases}$$

是随机化检验，故$UMPT$有以下形式

$$\varphi(\boldsymbol{x})=\begin{cases}1,\lambda(\boldsymbol{x})>\theta_0^n/\theta_1^n,\\ r,\lambda(\boldsymbol{x})=\theta_0^n/\theta_1^n.\\ 0,\lambda(\boldsymbol{x})<\theta_0^n/\theta_1^n\end{cases}$$

$(T,Y)=(X_{(1)},X_{(n)})$的密度函数为

$$g_\theta(t,y)=n(n-1)\frac{(y-t)^{n-2}}{\theta^n}I_{(\theta<t<y<2\theta)}$$

当H_0成立时，有

$$E_{\theta_0}[\varphi(\boldsymbol{X})]=r\int_{\theta_1}^{2\theta_0}\int_{\theta_1}^{y}n(n-1)\frac{(y-t)^{n-2}}{\theta_0^n}dtdy=r\left(2-\frac{\theta_1}{\theta_0}\right)^n$$

又因为$\beta(\theta)$与θ_1无关，故它不是

$$H_0{:}\theta=\theta_0\leftrightarrow H_1{:}\theta>\theta_0$$

的水平为α的$UMPT$.故题中所求的$UMPT$不存在.

54、设$X_1,\cdots,X_n\ i.i.d.$服从下列分布:$f(x)=e^{-(x-\theta)}I_{(\theta,\infty)}(x)$,试求检验问题$H_0{:}\theta=\theta_0\leftrightarrow H_1{:}\theta\neq\theta_0$水平为$\alpha$的$UMPT$.

解：

先考虑以下假设

$$H_0'{:}\theta=\theta_0\leftrightarrow H_1'{:}\theta=\theta_1,\theta_1>\theta_0.$$

$\boldsymbol{X}$的密度函数和似然比为

$$f(\boldsymbol{x},\theta_1)=e^{-\sum\limits_{i=1}^{n}x_i+n\theta_1}I_{(\theta_1,\infty)}(x_{(1)}),f(\boldsymbol{x},\theta_0)=e^{-\sum\limits_{i=1}^{n}x_i+n\theta_0}I_{(\theta_0,\infty)}(x_{(1)})$$

$$\lambda(\boldsymbol{x})=\begin{cases}e^{n(\theta_1-\theta_2)}，&\infty>x_{(1)}>\theta_1，\\0，&\theta_1\geqslant x_{(1)}>\theta_0.\\\infty,&\theta_0\geqslant x_{(1)}.\end{cases}$$

则$UMPT$有如下形式

对于假设$H_0{:}\theta=\theta_0\leftrightarrow H_1{:}\theta=\theta_1,\theta_1<\theta_0$,有

$$\lambda(\boldsymbol{x})=\begin{cases}e^{n(\theta_1-\theta_0)},x_{(1)}>\theta_0\\\infty,\qquad x_{(1)}\leqslant\theta_0\end{cases}$$

则$UMPT$有如下形式

$$\varphi(\boldsymbol{x})=\begin{cases}1,\lambda(\boldsymbol{x})>c,\\0,\lambda(\boldsymbol{x})\leqslant c.\end{cases}$$

$T=X_{(1)}$的密度函数为$g(t)=ne^{-n(t-\theta)}I_{(\theta,\infty)}(t)$，

当$c\geqslant e^{n(\theta_1-\theta_0)}$时，$E_{\theta_0}[\varphi(\boldsymbol{X})]=\int_{-\infty}^{\theta_0}\varphi(x)g_{\theta_0}(t)dt=0$,不符合题意，故$c<e^{n(\theta_1-\theta_0)}$.

故当H_0成立时有

$$\beta(\theta)=E_{\theta_0}[\varphi(\boldsymbol{X})]=\int_{d}^{+\infty}\varphi(x)g_{\theta_0}(t)dt+\int_{-\infty}^{\theta_0}\varphi(x)g_{\theta_0}(t)dt=\int_{d}^{+\infty}ne^{-n(t-\theta_0)}dt=\alpha.$$

$d=\theta_0-\ln\alpha/n.$

则$UMPT$有如下形式

$$\varphi_2(\boldsymbol{x})=\begin{cases}1, x_{(1)}<\theta_0\text{或}x_{(1)}>d,\\ 0,\text{其他}.\end{cases}$$

55、设$X_1,\cdots,X_n\ i.i.d.\sim$均匀分布$U(0,\theta)$,指定$\theta_0$.证明$H_0:\theta=\theta_0\leftrightarrow H_1:\theta\neq\theta_0$水平为$\alpha$的$UMPT$存在（提示：取$\theta_1>\theta_0$和$\theta_1<\theta_0$分别考虑检验问题$H_0:\theta=\theta_0\leftrightarrow H_1:\theta\neq\theta_0$）.

解：

由例5.4.3可知，检验问题

$$H_0:\theta=\theta_0\leftrightarrow H_1:\theta=\theta_1,\quad \theta_1>\theta_0.$$

似然比为

$$\lambda(\boldsymbol{x})=\begin{cases}\left(\dfrac{\theta_0}{\theta_1}\right)^n, 0<x_{(n)}<\theta_0,\\ \infty,\qquad \theta_0<x_{(n)}<\infty.\end{cases}$$

对于检验问题

$$H_0:\theta=\theta_0\leftrightarrow H_1:\theta=\theta_1,\quad 0<\theta_1<\theta_0$$

似然比为

$$\lambda(\boldsymbol{x})=\begin{cases}\left(\dfrac{\theta_0}{\theta_1}\right)^n, 0<x_{(n)}<\theta_1,\\ 0,\qquad \theta_1<x_{(n)}<\theta_0,\\ \infty,\qquad \theta_0<x_{(n)}<\infty\end{cases}$$

则$UMPT$有如下形式

$$\varphi(\boldsymbol{x})=\begin{cases}1,\lambda(\boldsymbol{x})>c,\\ 0,\lambda(\boldsymbol{x})\leqslant c.\end{cases}$$

$T=X_{(n)}$的密度函数为

$$g(t)=\frac{nt^{n-1}}{\theta^n}I_{(0,\theta)}(t)$$

当$c\geqslant\dfrac{\theta_0^n}{\theta_1^n}$时，$E_{\theta_0}[\varphi(\boldsymbol{X})]=\int_{\theta_0}^{\infty}\varphi(x)g_{\theta_0}(t)dt=0$,不符合题意，故$c<\dfrac{\theta_0^n}{\theta_1^n}$.

则

$$E_{\theta_0}[\varphi(\boldsymbol{X})]=\int_0^c\varphi(x)g_{\theta_0}(t)dt+\int_{\theta_0}^{\infty}\varphi(x)g_{\theta_0}(t)dt=\int_0^c\frac{nt^{n-1}}{\theta_0^n}dt=\frac{c^n}{\theta_0^n}=\alpha,$$

故$c=\theta_0\alpha^{\frac{1}{n}}$,与$\theta_1$无关，故检验问题$H_0:\theta=\theta_0\leftrightarrow H_1:\theta\neq\theta_0$水平为$\alpha$的$UMPT$为

$$\varphi(\boldsymbol{x})=\begin{cases}1, x_{(n)}<\theta_0\alpha^{\frac{1}{n}}\text{或}x_{(1)}>\theta_0,\\0,\text{其他}\end{cases}$$

56、设$X_1,\cdots,X_n\ i.i.d.\sim N(\mu,\sigma^2)$,$\mu$和$\sigma^2$皆未知，利用假设检验方法导出$\sigma^2$的置信系数为$1-\alpha$的置信区间和置信上、下限.

解：对于$H_0{:}\sigma_1^2=\sigma_2^2\leftrightarrow H_1{:}\sigma_1^2\neq\sigma_2^2$,令$F=\dfrac{S_1^2}{S_2^2}\cdot\dfrac{\sigma_2^2}{\sigma_1^2}$,则有接受域

检验问题$H_0{:}\sigma^2=\sigma_0^2\leftrightarrow H_1{:}\sigma^2\neq\sigma_0^2$的水平为$\alpha$的接受域为

$$\overline{D}=\left\{\boldsymbol{X}{:}\chi_{n-1}^2\left(1-\frac{\alpha}{2}\right)\leqslant\frac{(n-1)S^2}{\sigma_0^2}\leqslant\chi_{n-1}^2\left(\frac{\alpha}{2}\right)\right\}$$

H_0成立时$\sigma^2=\sigma_0^2$,故可解得置信区间为

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

同理可解得置信上、下限为$\dfrac{(n-1)S^2}{\chi_{n-1}^2(1-\alpha)}$，$\dfrac{(n-1)S^2}{\chi_{n-1}^2(\alpha)}$

57、设$X_1,\cdots,X_m\ i.i.d.\sim N(\mu_1,\sigma_1^2)$,$Y_1,\cdots,Y_n\ i.i.d.\sim N(\mu_2,\sigma_2^2)$,且样本$X_1,\cdots,X_m$和$Y_1,\cdots,Y_n$独立.利用假设检验方法求出$\sigma_1^2/\sigma_2^2$的置信系数为$1-\alpha$的置信区间和置信上、下限.

解：

对于检验问题$H_0{:}\sigma_1^2/\sigma_2^2=c\leftrightarrow H_1{:}\sigma_1^2/\sigma_2^2\neq c$，$\dfrac{S_2^2}{S_1^2}\cdot c\sim F_{n-1,m-1}$,接受域为

$$\overline{D}=\left\{(\boldsymbol{X},\boldsymbol{Y}){:}F_{n-1,m-1}\left(1-\frac{\alpha}{2}\right)\leqslant\frac{S_2^2}{S_1^2}\cdot c\leqslant F_{n-1,m-1}\left(\frac{\alpha}{2}\right)\right\}$$

故置信区间为

$$\left[\frac{S_1^2}{S_2^2}\cdot F_{n-1,m-1}(1-\alpha/2),\frac{S_1^2}{S_2^2}\cdot F_{n-1,m-1}(\alpha/2)\right].$$

对于检验问题$H_0{:}\sigma_1^2/\sigma_2^2\leqslant c\leftrightarrow H_1{:}\sigma_1^2/\sigma_2^2>c$，接受域为

$$\overline{D}=\left\{(\boldsymbol{X},\boldsymbol{Y}){:}\frac{S_2^2}{S_1^2}\leqslant\frac{F_{n-1,m-1}(\alpha)}{c}\right\}\Rightarrow 1-\alpha=P\left(c\leqslant\frac{S_1^2}{S_2^2}\cdot F_{n-1,m-1}(\alpha)\right)$$

故置信上限为$\dfrac{S_1^2}{S_2^2}\cdot F_{n-1,m-1}(\alpha)$

同理可得置信下限为$\dfrac{S_1^2}{S_2^2}\cdot F_{n-1,m-1}(1-\alpha)$.

习题 6 非参数假设检验

1、在8块土地上同时试种甲、乙两种品种作物，每块分成大小、形状一样的两小块，随机地将其中一小块派给甲，另一小块给乙.它们的产量列于下表：

编号	1	2	3	4	5	6	7	8
品种甲	209	200	177	169	159	187	169	198
品种乙	151	168	14	164	166	176	169	188

（1）假定甲、乙两种作物产量的差服从正态分布，试在$\alpha=0.05$下判定，甲品种是否是对乙品种的改良（取取$\alpha=0.05$）.

（2）去掉（1）中正态假定，分别用符号检验法、符号秩和检验法判别甲品种是否是对乙品种的改良（取$\alpha=0.05$）

解：

（1）设品种甲样本为$X_1,\cdots,X_n$,品种乙样本为$Y_1,\cdots,Y_n$，设$Z_i=X_i-Y_i,i=1,\cdots,8$,$Z_1,\cdots,Z_n\ i.i.d.\sim N(\mu,\sigma^2)$,则相当于对下列检验问题假设检验

$$H_0:\mu>0\leftrightarrow H_1:\mu\leqslant 0$$

易知其检验统计量为$T=\sqrt{n}\,\overline{Z}/S,T\sim t_{n-1}$,则否定域为

$$\overline{D}=\{\boldsymbol{Z}:T<-t_{n-1}(\alpha)\}$$

计算得$n=8,\overline{Z}=17.375,S=21.27,t_7(0.05)=1.8946$,故有

$$T=\sqrt{n}\,\overline{Z}/S=2.31>-1.8946$$

故接受H_0,即甲品种是对乙品种的改良.

（2）①符号检验法

设检验问题为

$$H_0:\theta>\frac{1}{2}\leftrightarrow H_1:\theta\leqslant\frac{1}{2}$$

故检验的水平为α的否定域为

$$\{n_+\leqslant c\}$$

$Z_1,\cdots,Z_8$中，$n_+=6,n_-=1$,故$n=7$,

$$p=\sum_{i=0}^{6}\binom{7}{i}\left(\frac{1}{2}\right)^7=0.992>0.05$$

故接受H_0,即甲品种是对乙品种的改良.

②符号秩和检验法

编号	1	2	3	4	5	6	7	8
甲	209	200	177	169	159	187	169	198
乙	151	168	147	164	166	176	169	188
符号	+	+	+	+	−	+	0	+
$\lvert Z_i\rvert$	58	32	30	5	7	11	0	10
秩	[7]	[6]	[5]	[1]	2	[4]		[3]

检验问题为

$$H_0\text{:甲不是乙的改良}\leftrightarrow H_1\text{:甲是乙的改良}$$

H_0成立时，W^+应该很大，故检验的否定域为

$$\{W^+>c\}$$

本题中，$n=7,\alpha=0.05$，查表得$c=25$,而题中$W^+=26$，故拒绝H_0，即甲品种是对乙品种的改良.

2、从某铜矿东西两段各抽取容量为10的样本，随机配成10对，数据如下表：

对子	1	2	3	4	5	6	7	8	9	10
东段含铜量	28	20	4	32	8	12	16	48	8	20
西段含铜量	20	11	13	10	45	15	11	13	25	8

用符号检验法，在水平$\alpha=0.10$下，检验“东西段含铜量无显著差异”之假设.

解：
设东段样本为$X_1,\cdots,X_n$,西段样本为$Y_1,\cdots,Y_n$，设$Z_i=X_i-Y_i,i=1,\cdots,10$， 则检验问题为

$$H_0\text{:东西段含铜量无显著差异}\leftrightarrow K\text{:东西段含铜量有显著差异}$$

则否定域的形式为

$$D=\{n_+\geqslant c\text{或}n_+\leqslant d\}$$

$n=10,\alpha=0.10$,查表得$c=9,d=n-c=1$,而$n_+=6$在接受域，故接受H_0

3、比较新旧两种饲料对猪的催肥效果（以饲养一段时间后猪的增重表示）.由于猪的体质，特别猪的初试重量对猪的催肥有相当重要的影响，所以取体质尽可能相同的两头猪做对比试验.下表是9个对比试验的记录（增重x_i和y_i以斤为单位）：

试验号	1	2	3	4	5	6	7	8	9
x_i(新饲料)	35	24	24	33	30	24	29	37	21

y_i(旧饲料)	30	26	24	29	28	25	26	30	27

问新饲料能否推广（用符号检验和符号秩和检验，$\alpha=0.05$）？

解：

①符号检验法

设检验问题为

$$H_0\text{:新饲料能推广} \leftrightarrow K\text{:新饲料不能推广}$$

则检验问题的否定域有如下形式

$$D=\{n_+\leqslant d\}$$

由题中可得$n_+=5, n_-=3$,故$n=n_++n_-=8$,查表得$\alpha=0.05$时，$c=7$时，有$P(n_+\geqslant c)=\alpha$,故$d=n-c=1$，而$n_+=5$在接受域内，故接受H_0.

②符号秩和检验

设检验问题为

$$H_0\text{:新饲料能推广} \leftrightarrow K\text{:新饲料不能推广}$$

成立时W^+应当很大，故检验的否定域为

$$\{W^+\leqslant d\}$$

试验号	1	2	3	4	5	6	7	8	9
x_i(新饲料)	35	24	24	33	30	24	29	37	21
y_i(旧饲料)	30	26	24	29	28	25	26	30	27
符号	+	−	0	+	+	−	+	+	−
$\|Z_i\|$	5	2	0	4	2	1	3	7	6
秩	[6]	3		[5]	[2]	1	[4]	[8]	7

由题可得$n=8, \alpha=0.05$,查表得$c=31$时，$P(W^+\geqslant c)=\alpha$，故$d=n(n+1)/2-c=5$,

本题$W^+=\dfrac{2+3}{2}+4+5+6+8=25.5>5$,在接受域内，故新饲料能推广.

4、甲、乙两厂都生产电视机显像管.从甲厂抽取8只，从乙厂抽取10只，测出其寿命（单位：月）如下：

甲	32	25	40	31	35	29	37	39		
乙	41	39	36	47	45	34	48	44	30	33

试在水平$\alpha=0.10$下用$Wilcoxon$秩和检验法，检验“两厂显像管平均寿命相同”之假设.

解：

设X和Y分别表示甲乙两厂显像管寿命，分布函数分别为$F(x),F(x-\theta)$,则检验问题为

$$H_0:\theta=0\leftrightarrow H_1:\theta\neq 0$$

故否定域的形式为

$$\{W\leqslant d\text{或}W\geqslant c\}$$

$n=8,m=10,\alpha=0.10$,查表得$c=96,d=n(m+n-1)-c=56$,即否定域为

$$D=\{(X,Y):W\leqslant 56\text{或}W\geqslant 96\}$$

25	29	30	31	32	33	34	35	36	37	39	39	40	41	44	45	47	48
$\underline{1}$	$\underline{2}$	3	$\underline{4}$	$\underline{5}$	6	7	$\underline{8}$	9	$\underline{10}$	$\underline{11}$	12	$\underline{13}$	14	15	16	17	18

下划线表示甲寿命的秩，则秩和为

$$W=1+2+4+5+8+10+11.5+13=54.5<56$$

故拒绝H_0,认为两厂显像管平均寿命不相同.

5、用两种材料的灯丝制造灯泡，设甲材料灯泡的寿命分布为$F(x)$,乙材料灯泡寿命为$F(x-\theta)$.今分别就两种材料灯泡随机抽取若干个进行试验，得到寿命（单位：h）数据如下表：

甲材料	1610	1650	1680	1710	1750	1720	1800
乙材料	1580	1600	1640	1630	1700		

试用$Wilcoxon$秩和检验法检验两种材料对灯泡寿命的影响有无显著差异（$\alpha=0.05$）？

解：

检验问题为

$$H_0:\theta=0\leftrightarrow H_1:\theta\neq 0$$

故否定域的形式为

$$\{W\leqslant d\text{或}W\geqslant c\}$$

$n=7,m=5,\alpha=0.05$,查表得$c=45,d=n(m+n-1)-c=20$,即否定域为

$$D=\{(X,Y):W\leqslant 20\text{或}W\geqslant 45\}$$

1580	1600	1610	1630	1640	1650	1680	1700	1710	1720	1750	1800
$\underline{1}$	$\underline{2}$	3	$\underline{4}$	$\underline{5}$	6	7	$\underline{8}$	9	10	11	12

下划线表示乙材料的秩，则秩和为

$$W=1+2+4+5+8=20\leqslant 20$$

故拒绝H_0,认为两种材料对灯泡寿命的影响有显著差异.

6、分别在甲、乙两厂生产的电子元器件中抽取150个和130个进行寿命试验.算出乙工厂样本秩和$W=20251$.试在水平$\alpha=0.05$下检验假设

$$H_0\text{:这两厂所生产的电子元器件的寿命无显著差异.}$$

解：

$m=150, n=130, \alpha=0.05, W=20251$,由于$m,n$较大，故采用大样本方法.查表得否定域为

$$D=\{(X,Y):|W^*|\geqslant u_{0.025}=1.96\}$$

其中

$$|W^*|=\left|\frac{W-n(n+m+1)/2}{\sqrt{mn(m+n+1)/12}}\right|=\left|\frac{20251-130\times(280+1)/2}{\sqrt{130\times150\times281/12}}\right|=2.939\geqslant1.96$$

故拒绝H_0,认为这两厂所生产的电子元器件的寿命无显著差异.

7、掷骰子300次，结果如下表：

点数	1	2	3	4	5	6
频数	43	49	56	45	66	41

试在水平$\alpha=0.05$下检验假设：此骰子是均匀的.

解：

骰子掷到每一面的概率为1/6，若记$X=i$表示骰子掷到第i面，$i=1,\cdots,6$.检验问题为

$$H_0:P(X=1)=\frac{1}{6},P(X=2)=\frac{1}{6},P(X=3)=\frac{1}{6},$$
$$P(X=4)=\frac{1}{6},P(X=5)=\frac{1}{6},P(X=6)=\frac{1}{6}$$

令$v_1=43, v_2=49, v_3=56, v_4=45, v_5=66, v_6=41, n=300$,则

$$k_0=\frac{(43-300\times1/6)^2}{300\times1/6}+\frac{(49-300\times1/6)^2}{300\times1/6}+\frac{(56-300\times1/6)^2}{300\times1/6}$$
$$+\frac{(45-300\times1/6)^2}{300\times1/6}+\frac{(66-300\times1/6)^2}{300\times1/6}+\frac{(41-300\times1/6)^2}{300\times1/6}=8.96$$

又因为

$$\chi_5^2(0.05)=11.071,\text{而}k_0<11.071$$

故否定H_0,认为骰子不是均匀的.

8、64只某种杂交的几内亚猪后代，34只红的，10只黑的，20只白的.根据遗传学的模型，它们之间的比例应为$9:3:4$，问以上数据是否在$\alpha=0.05$水平下与模型相符.

解：

后代为红的、黑的、白的概率分别为9/16，3/16，4/16，若记$X=i$分别表示后代为红、黑、白，$i=1,2,3$.则检验问题为

$$H_0:P(X=1)=\frac{9}{16},P(X=2)=\frac{3}{16},P(X=3)=\frac{4}{16}.$$

令$v_1=34,v_2=10,v_3=20$,则

$$k_0=\frac{(34-64\times 9/16)^2}{64\times 9/16}+\frac{(10-64\times 3/16)^2}{64\times 3/16}+\frac{(20-64\times 4/16)^2}{64\times 4/16}=1.44$$

又因为

$$k_0=1.44<\chi_2^2(0.05)=5.991$$

故接受H_0, 以上数据在$\alpha=0.05$水平下与模型相符.

9、在生产同一产品的甲、乙两厂中各抽出n个产品组成n对，让n个人每人评判一对.各人评判是“甲好”“乙好”“甲乙一样”这三种之一.试根据评判结果，用χ^2检验法检验“甲、乙两厂产品质量无差别“这一假设.

解：

设$X=1$表示“甲好”，$X=2$表示“乙好”，则检验问题为

$$H_0:P(X=1)=\frac{1}{2},P(X=2)=\frac{1}{2}$$

v_1表示$X=1$事件出现的频数，v_2表示$X=2$事件出现的频数，$n'=v_1+v_2$，则

$$k_0=\frac{(v_1-n'\times 1/2)^2}{n'\times 1/2}+\frac{(v_2-n'\times 1/2)^2}{n'\times 1/2}$$

则若

$$k_0>\chi_1^2(\alpha)\text{则否定}H_0,\text{否则接受}H_0.$$

10、某种配偶的后代按体格的属性分为三类，各类的个数为$10,53,46$.根据某种遗传学模型，其相对频率的比应为$p^2:2p(1-p):(1-p)^2$.问以上数据是否在$\alpha=0.05$水平下与模型相符.

解：

设$X=i$分别表示后代的三类属性，$i=1,2,3$.则检验问题为

$$H_0:P(X=1)=p^2,P(X=2)=2p(1-p),P(X=3)=(1-p)^2.$$

$v_1=10,v_2=53,v_3=46$,则可求出p的MLE为

$$\hat{p}=\frac{2v_1+v_2}{2v_1+2v_2+2v_3}=\frac{20+53}{2\times 109}=\frac{73}{218}$$

故检验问题转化为

$$H_0:P(X=1)=\hat{p}^2,P(X=2)=2\hat{p}(1-\hat{p}),P(X=3)=(1-\hat{p})^2.$$

则

$$\widetilde{K}_n^*=\frac{(10-109\times\hat{p}^2)^2}{109\times\hat{p}^2}+\frac{(53-109\times 2\hat{p}(1-\hat{p}))^2}{109\times 2\hat{p}(1-\hat{p})}+\frac{(46-109\times(1-\hat{p})^2)^2}{109\times(1-\hat{p})^2}=0.9134$$

拟合优度为

$$p\left(\tilde{k}_0^*\right)=P\left(\widetilde{K}_n^*>\tilde{k}_0^*|H_0\right)\approx P(\chi_1^2>0.9134)<0.25<\alpha$$

故否定H_0,以上数据在$\alpha=0.05$水平下与模型不相符.

11、用运动手枪对100个靶子各射击10发子弹，只记录命中与否，射击结果的记录如下表：

命中数	0	1	2	3	4	5	6	7	8	9	10
靶数	0	2	4	10	22	26	18	12	4	2	0

用χ^2检验法检验射击结果是否服从二项分布（取$\alpha=0.05$）.

解：

先求p的MLE,

$$\hat{p}=\overline{X}=\frac{2\times1+4\times2+10\times3+22\times4+26\times5+18\times6+12\times7+4\times8+2\times9}{10\times100}=0.5$$

则检验问题为

$$H_0:\text{射击结果服从二项分布}\,b(10,0.5)$$

对数据重新分组

0-2	3	4	5	6	7	8-10
6	10	22	26	18	12	6

故有

$$\begin{aligned}n\hat{p}_0^*&=100\times0.5^{10}\times(C_{10}^0+C_{10}^1+C_{10}^2)=5.47,\\ n\hat{p}_i^*&=100\times0.5^{10}\times C_{10}^i,i=3,\cdots,7,\\ n\hat{p}_8^*&=100\times0.5^{10}\times(C_{10}^8+C_{10}^9+C_{10}^{10})=5.47.\end{aligned}$$

从而有

$$\widetilde{K}_n^*=\frac{(6-5.47)^2}{5.47}+\frac{(10-n\hat{p}_3^*)^2}{n\hat{p}_3^*}+\cdots+\frac{(12-n\hat{p}_7^*)^2}{n\hat{p}_7^*}+\frac{(6-5.47)^2}{5.47}=0.856$$

拟合优度为

$$p\left(\tilde{k}_0^*\right)=P\left(\widetilde{K}_n^*>\tilde{k}_0^*|H_0\right)\approx P(\chi_5^2>0.856)>0.975>\alpha$$

故接受H_0,射击结果服从二项分布.

12、观察某城市每日交通情况，210天的记录如下表：

事故数	0	1	2	3	4	≥5
天数	109	65	22	3	4	7

初步推测每日发生的事故数服从$Poisson$分布，试用χ^2试验法试验之（$\alpha=0.05$）.

解：

先求λ的MLE,

$\hat{\lambda}=\overline{X}=\frac{65\times1+22\times2+3\times3+4\times4+5\times7}{210}=0.805$

则检验问题为

$$H_0\text{:每日发生的交通事故服从}P\left(\hat{\lambda}\right)$$

对数据重新分组

事故数	0	1	2	⩾3
天数	109	65	22	14

$$\widetilde{K}_n^*=\frac{(109-210\times e^{-0.805})^2}{210\times e^{-0.805}}+\cdots+\frac{\left(14-210\times\left(e^{0.805}-1-\frac{0.805}{1!}-\frac{0.805^2}{2!}\right)e^{-0.805}\right)^2}{210\times\left(e^{0.805}-1-\frac{0.805}{1!}-\frac{0.805^2}{2!}\right)e^{-0.805}}=7.74$$

拟合优度为

$$p\left(\tilde{k}_0^*\right)=P\left(\widetilde{K}_n^*>\tilde{k}_0^*|H_0\right)\approx P(\chi_2^2>10.42)\approx0.025<\alpha$$

故否定H_0,每日发生的事故数不服从$Poisson$分布.

13、卢瑟福观察了每1/8分钟内一放射性物质放射的粒子数.他观察了2612次，结果如下表：

粒子数	0	1	2	3	4	5	6	7	8	9	10	11
频数	57	203	383	525	532	408	273	139	49	27	10	6

问以上数据是否在$\alpha=0.10$水平下与$Poisson$分布相符.

解：

解：

对数据重新分组

粒子数	0	1	2	3	4	5	6	7	8-11
频数	57	203	383	525	532	408	273	139	92

先求λ的MLE,

$\hat{\lambda}=\overline{X}=\frac{1\times203+2\times383+\cdots+10\times10+11\times6}{2612}=3.88$

则检验问题为

$$H_0\text{:每日发生的交通事故服从}P\left(\hat{\lambda}\right)$$

$$\widetilde{K}_n^*=\sum_{i=0}^{8}\frac{(v_i-np_i^*)^2}{np_i^*}=18.34$$

拟合优度为

$$p\left(\tilde{k}_0^*\right)=P\left(\widetilde{K}_n^*>\tilde{k}_0^*|H_0\right)\approx P(\chi_7^2>18.34)\approx 0.01<\alpha$$

故否定H_0,粒子数不服从$Poisson$分布.

14、下表列出了某型号标准容量为100μ的电介容器100只的实际电容量的数据，问该电介电容器的电容量的分布是否为正态分布（$\alpha=0.05$）？

电容量	101	102	103	104	105	106	107
电容个数	1	2	3	3	7	16	13
电容量	108	109	110	111	102	103	104
电容个数	17	11	9	10	3	4	1

解：

设$r.v.X$表示电容器的电容量，则检验问题为

$$H_0\text{:存在}\mu_0\text{，}\sigma_0^2\text{使得}X\text{服从正态分布}N(\mu_0,\sigma_0^2).$$

μ和σ^2的MLE为

$$\hat{\mu}^*=\overline{X}=\frac{101+102\times 2+103\times 3+\cdots+113\times 4+114}{100}=107.85$$

$$\hat{\sigma}_*^2=\frac{1}{100}\sum_{i=1}^{100}\left(x_i-\bar{x}\right)^2=\frac{1}{100}[1\times(101-107.85)^2+\cdots+1\times(114-107.85)^2]=7.0725$$

将数据重新分组

$(-\infty,104.5)$	$[104.5,105.5)$	$[105.5,106.5)$	$[106.5,107.5)$	$[107.5,108.5)$
9	7	16	13	17
$[108.5,109.5)$	$[109.5,110.5)$	$[110.5,111.5)$	$[111.5,+\infty)$	
11	9	10	8	

列表为

i	区间	v_i	$\hat{p}_i^*$	$n\hat{p}_i^*$	$\left(v_i-n\hat{p}_i^*\right)^2$	$\frac{\left(v_i-n\hat{p}_i^*\right)^2}{n\hat{p}_i^*}$
1	$(-\infty,104.5)$	9	0.104	10.4	1.96	0.188
2	$[104.5,105.5)$	7	0.086	8.6	2.56	0.298
3	$[105.5,106.5)$	16	0.117	11.7	18.49	1.580
4	$[106.5,107.5)$	13	0.142	14.2	1.44	0.101
5	$[107.5,108.5)$	17	0.148	14.8	4.84	0.327
6	$[108.5,109.5)$	11	0.136	13.6	6.76	0.497
7	$[109.5,110.5)$	9	0.108	10.8	3.24	0.3
8	$[110.5,111.5)$	10	0.074	7.4	6.76	0.508
9	$[111.5,+\infty)$	8	0.085	8.5	0.25	0.029

$\sum$	——	100	1	100	——	3.828

故$\widetilde{K}_n^* = 3.828$，故拟合优度为

$$p\left(\tilde{k}_0^*\right) = P\left(\widetilde{K}_n^* > \tilde{k}_0^* | H_0\right) \approx P(\chi_6^2 > 3.828) \approx 0.75 > \alpha$$

因此接受H_0.

15、试证明2×2列联表的检验统计量为

$$K^* = \frac{n(n_{11}n_{22} - n_{12}n_{21})^2}{n_{1\cdot}n_{2\cdot}n_{\cdot 1}n_{\cdot 2}}.$$

解：

$$\begin{aligned} K^* &= n\left(\sum_{i=1}^{2}\sum_{j=1}^{2}\frac{n_{ij}^2}{n_{i\cdot}n_{\cdot j}} - 1\right) = n\left(\frac{n_{11}^2}{n_{1\cdot}n_{\cdot 1}} + \frac{n_{21}^2}{n_{2\cdot}n_{1\cdot}} + \frac{n_{12}^2}{n_{1\cdot}n_{\cdot 2}} + \frac{n_{22}^2}{n_{2\cdot}n_{\cdot 2}} - 1\right) \\ &= n\left(\frac{n_{11}^2 n_{2\cdot}n_{\cdot 2} + n_{21}^2 n_{1\cdot}n_{\cdot 2} + n_{12}^2 n_{2\cdot}n_{\cdot 1} + n_{22}^2 n_{1\cdot}n_{\cdot 1} - n_{1\cdot}n_{2\cdot}n_{\cdot 1}n_{\cdot 2}}{n_{1\cdot}n_{2\cdot}n_{\cdot 1}n_{\cdot 2}}\right) \end{aligned}$$

要证

$$K^* = \frac{n(n_{11}n_{22} - n_{12}n_{21})^2}{n_{1\cdot}n_{2\cdot}n_{\cdot 1}n_{\cdot 2}}$$

需证

$$n_{11}^2 n_{2\cdot}n_{\cdot 2} + n_{21}^2 n_{1\cdot}n_{\cdot 2} + n_{12}^2 n_{2\cdot}n_{\cdot 1} + n_{22}^2 n_{1\cdot}n_{\cdot 1} - n_{1\cdot}n_{2\cdot}n_{\cdot 1}n_{\cdot 2} = (n_{11}n_{22} - n_{12}n_{21})^2$$

$$\text{等式右边} = n_{11}^2 n_{22}^2 + n_{12}^2 n_{21}^2 - 2n_{11}n_{12}n_{21}n_{22}$$

$$\begin{aligned} &\text{等式左边} = n_{11}^2(n_{21}+n_{22})(n_{12}+n_{22}) + n_{21}^2(n_{11}+n_{12})(n_{12}+n_{22}) + n_{12}^2(n_{21}+n_{22})(n_{11}+n_{21}) \\ &+ n_{22}^2(n_{11}+n_{12})(n_{11}+n_{21}) - (n_{11}+n_{12})(n_{21}+n_{22})(n_{11}+n_{21})(n_{12}+n_{22}) \\ &= \cancel{n_{11}^2 n_{21} n_{12}} + \cancel{n_{11}^2 n_{21} n_{22}} + \cancel{n_{11}^2 n_{22} n_{12}} + \cancel{n_{11}^2 n_{22}^2} + \cancel{n_{21}^2 n_{11} n_{12}} + \cancel{n_{21}^2 n_{11} n_{22}} + \cancel{n_{21}^2 n_{12}^2} + \cancel{n_{21}^2 n_{12} n_{22}} \\ &+ \cancel{n_{12}^2 n_{21} n_{11}} + n_{12}^2 n_{21}^2 + \cancel{n_{12}^2 n_{22} n_{11}} + \cancel{n_{12}^2 n_{22} n_{21}} + n_{22}^2 n_{11}^2 + \cancel{n_{22}^2 n_{11} n_{21}} + \cancel{n_{22}^2 n_{12} n_{11}} + \cancel{n_{22}^2 n_{12} n_{21}} \\ &- \cancel{n_{11}^2 n_{21} n_{12}} - \cancel{n_{11}^2 n_{21} n_{22}} - \cancel{n_{21}^2 n_{11} n_{12}} - \cancel{n_{21}^2 n_{11} n_{22}} - \cancel{n_{11}^2 n_{22} n_{12}} - \cancel{n_{11}^2 n_{22}^2} - n_{11}n_{12}n_{21}n_{22} - \cancel{n_{22}^2 n_{11} n_{21}} \\ &- \cancel{n_{12}^2 n_{11} n_{21}} - n_{11}n_{12}n_{21}n_{22} - \cancel{n_{12}^2 n_{21}^2} - \cancel{n_{21}^2 n_{12} n_{22}} - \cancel{n_{12}^2 n_{11} n_{21}} - \cancel{n_{22}^2 n_{11} n_{12}} - \cancel{n_{12}^2 n_{21} n_{22}} - \cancel{n_{22}^2 n_{12} n_{21}} \\ &= n_{22}^2 n_{11}^2 + n_{12}^2 n_{21}^2 - 2n_{11}n_{12}n_{21}n_{22} = \text{等式右边} \end{aligned}$$

故得证.

16、为研究慢性气管炎与吸烟量的关系，调查272人结果如下：

吸烟量/（支/日）	0-9	10-19	20 以上
患者人数	22	98	25
健康者人数	22	89	16

试问慢性气管炎与吸烟量是否有关（$\alpha = 0.05$）.

解：

检验问题为

$$H_0\text{:气管炎与吸烟量无关}$$

检验统计量为

$$K^* = n\left(\sum_{i=1}^{2}\sum_{j=1}^{3}\frac{n_{ij}^2}{n_{i\cdot}n_{\cdot j}} - 1\right)$$

其中$n=272, n_{1\cdot}=145, n_{2\cdot}=127, n_{\cdot 1}=44, n_{\cdot 2}=187, n_{\cdot 3}=41$,

故$K^* = 272\times\left(\frac{22^2}{145\times 44}+\frac{98^2}{145\times 187}+\frac{25^2}{145\times 41}+\frac{22^2}{127\times 44}+\frac{89^2}{127\times 187}+\frac{16^2}{127\times 41}-1\right)=1.22294$

而$\chi_2^2(0.05)=5.991$,即$K^*<\chi_2^2(0.05)$,故接受H_0.

17、在500人身上试验某种血清预防感冒的作用.把他们在一年中的记录与另外500名未用血清处理的人作比较，结果如下表：

	未感冒	感冒一次	感冒两次以上
处理	252	145	103
未处理	224	136	140

未感冒、感冒一次，感冒两次以上的人所占的比例，是否在与未用血清处理的人中，具有这些情况的人所占比例相等（$\alpha=0.05$）？

解：

检验问题为

$$H_0\text{:未用血清和用血清的人感冒情况比例相等}$$

检验统计量为

$$K^* = n\left(\sum_{i=1}^{2}\sum_{j=1}^{3}\frac{n_{ij}^2}{n_{i\cdot}n_{\cdot j}} - 1\right)$$

其中$n=1000, n_{1\cdot}=500, n_{2\cdot}=500, n_{\cdot 1}=476, n_{\cdot 2}=281, n_{\cdot 3}=243$,

故$K^*=1000\times$（$\frac{252^2}{500\times 476}+\frac{145^2}{500\times 281}+\frac{103^2}{500\times 243}+\frac{224^2}{500\times 476}+$

$\frac{136^2}{500\times 281}+\frac{140^2}{500\times 243}-1$）$=7.56906$

$$\text{而}\chi_2^2(0.05)=5.991,\text{即}K^*>\chi_2^2(0.05)$$

故拒绝H_0.未用血清和用血清的人感冒情况比例不等

18、使用柯尔莫哥洛夫检验法检验下述25个数据（已按大小排列）：

-2.46， -2.11， -1.23， -0.99， -0.42， -0.39， -0.21， -0.15， -0.10，
-0.07， -0.02， 0.27， 0.40， 0.42， 0.44， ,0.70， 0.81， 0.88，
1.07， 1.39， 1.40， 1.47， 1.62， 1.64， 1.76.

试问上述25个数据是否与$N(0,1)$相符（$\alpha=0.05$）？

解：

检验问题为

$$H_0\text{:样本来自的总体分布为}F(x)=F_0(x)$$

其中$F_0(x)$为正态分布$N(0,1)$.

柯尔莫哥洛夫检验

i	$x_{(i)}$	$F_0(x_{(i)})$	$(i-1)/n$	i/n	δ_i
1	-2.46	0.0069	0	0.04	0.0331
2	-2.11	0.0174	0.04	0.08	0.0626
3	-1.23	0.1093	0.08	0.12	0.0293
4	-0.99	0.1611	0.12	0.16	0.0411
5	-0.42	0.3372	0.16	0.20	0.1772
6	-0.39	0.3483	0.20	0.24	0.1483
7	-0.21	0.4168	0.24	0.28	0.1768
8	-0.15	0.4404	0.28	0.32	0.1604
9	-0.10	0.4602	0.32	0.36	0.1402
10	-0.07	0.4721	0.36	0.40	0.1121
11	-0.02	0.4920	0.40	0.44	0.0920
12	0.27	0.6064	0.44	0.48	0.1664
13	0.40	0.6554	0.48	0.52	0.1754
14	0.42	0.6628	0.52	0.56	0.1428
15	0.44	0.6700	0.56	0.60	0.1100
16	0.70	0.7580	0.60	0.64	0.1580
17	0.81	0.7910	0.64	0.68	0.1510
18	0.88	0.8106	0.68	0.72	0.1306
19	1.07	0.8577	0.72	0.76	0.1377
20	1.39	0.9177	0.76	0.80	0.1577
21	1.40	0.9192	0.80	0.84	0.1192
22	1.47	0.9292	0.84	0.88	0.0892
23	1.62	0.9474	0.88	0.92	0.0674
24	1.64	0.9495	0.92	0.96	0.0295
25	1.76	0.9608	0.96	1	0.0392

由表可得$D_n=\max\{\delta_i,i=1,\cdots,25\}=0.1772,n=25,\alpha=0.05$,查表得$D_{25,0.05}=0.264$
$D_n<D_{25,0.05}$,不足以否定H_0,故接受H_0,认为以上数据与正态分布$N(0,1)$相符.

19、在π的前800位小数的数字中，$0,1,2$，$\cdots$，9相应出现了$74,92,83,79,80,73,77,75,76,91$次.试用柯氏检验法检验这些数据与$[0,10]$上的均匀分布相适合的假设（$\alpha=0.05$）.

解：

检验问题为

$$H_0\text{:样本来自的总体分布为}F(x)=F_0(x)$$

其中$F_0(x)$为均匀分布$U(0,10)$.

柯尔莫哥洛夫检验

i	$x_{(i)}$	$F_0(x_{(i)})$	$F_n(x_{(i-1)})$	$F_n(x_{(i)})$	δ_i
1	0	0	0	0.0925	0.0925
2	1	0.1	0.0925	0.2075	0.1075
3	2	0.2	0.2075	0.31125	0.11125
4	3	0.3	0.31125	0.41	0.11
5	4	0.4	0.41	0.51	0.11
6	5	0.5	0.51	0.60125	0.10125
7	6	0.6	0.60125	0.6975	0.0975
8	7	0.7	0.6975	0.79125	0.09125
9	8	0.8	0.79125	0.88625	0.08625
10	9	0.9	0.88625	1	0.1

由表可得$D_n=\max\{\delta_i,i=1,\cdots,10\}=0.11125,n=10,\alpha=0.05$,查表得$D_{25,0.05}=0.409$
$D_n<D_{25,0.05}$,不足以否定H_0,故接受H_0,认为以上数据与均匀分布$U(0,10)$相符.

20、抽查用克矽平治疗的矽肺患者10名，得到他们治疗前后血红蛋白的差（单位：g%）如下：

$$2.7,-1.2,-1.0,0,0.07,2.0,3.7,-0.6,0.8,-0.3$$

用W检验法检验“治疗前后血红蛋白差服从正态分布”这一假设（取$\alpha=0.05$）.

解：

将数据填入表中

i	$x_{(i)}$	$x_{(11-i)}$	$x_{(11-i)}-x_{(i)}$	a_i
1	-1.2	3.7	4.9	0.5739
2	-1.0	2.7	3.7	0.3291
3	-0.6	2.0	2.6	0.2141
4	-0.3	0.8	1.1	0.1224
5	0	0.07	0.07	0.0399

经计算得

$$\sum_{i=1}^{10} x_{(i)}=6.17, \bar{x}=\frac{1}{10} \sum_{i=1}^{10} x_{(i)}=0.617$$

$$\sum_{i=1}^{10}\left(x_{(i)}-\bar{x}\right)^{2}=24.70, \sum_{i=1}^{5} a_{i}\left(x_{(11-i)}-x_{(i)}\right)=4.15$$

于是有

$$W=\frac{\left[\sum_{i=1}^{5} a_{i}\left(x_{(11-i)}-x_{(i)}\right)\right]^{2}}{\sum_{i=1}^{10}\left(x_{(i)}-\bar{x}\right)^{2}}=\frac{4.15^{2}}{24.70}=0.697$$

查表得$W_{0.05}=0.842>0.697$，故拒绝H_0.

21、为检验一批地砖的抗压强度是否服从正态分布，从中抽取20块，得抗压强度（由小到大排列）为

$$57,62,66,67,74,76,77,80,81,86$$
$$87,89,91,94,95,96,97,103,109,112.$$

试用W检验法检验“这些数据是否与正态分布相符”（取$\alpha=0.05$）.

解：

将数据填入表中

i	$x_{(i)}$	$x_{(21-i)}$	$x_{(21-i)}-x_{(i)}$	a_i
1	57	112	55	0. 4734
2	62	109	47	0. 3211
3	66	103	37	0. 2565
4	67	97	30	0. 2085
5	74	96	22	0. 1686
6	76	95	19	0. 1334
7	77	94	17	0. 1013
8	80	91	11	0. 0711
9	81	89	8	0. 0422
10	86	87	1	0. 0140

经计算得

$$\sum_{i=1}^{20} x_{(i)}=1699, \bar{x}=\frac{1}{10} \sum_{i=1}^{10} x_{(i)}=84.95$$

$$\sum_{i=1}^{20}\left(x_{(i)}-\bar{x}\right)^{2}=4436.95, \sum_{i=1}^{10} a_{i}\left(x_{(11-i)}-x_{(i)}\right)=60.3443$$

于是有

$$W=\frac{\left[\sum_{i=1}^{10} a_i\left(x_{(11-i)}-x_{(i)}\right)\right]^2}{\sum_{i=1}^{20}\left(x_{(i)}-\bar{x}\right)^2}=\frac{60.3443^2}{4436.95}=0.82$$

查表得$W_{0.05}=0.905>0.82$，故拒绝H_0.

22、下表数据是200个零件的直径（单位：cm）：

直径	2.25	2.35	2.45	2.55	2.65	2.75	2.85	2.95
频数	3	4	5	11	12	17	19	26
直径	3.05	3.15	3.25	3.35	3.45	3.55	3.65	3.75
频数	24	22	19	13	13	7	3	2

试问这批数据是否与正态分布相符（用D检验，取$\alpha=0.05$）.

解：

检验问题为

$$H_0:\text{这批数据与正态分布相符}$$

$$D=\sum_{i=1}^{n}\left(i-\frac{n+1}{2}\right)X_{(i)}\Bigg/\left[n^3\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2\right]^{1/2}=0.0074$$

令

$$Y=\frac{\sqrt{n}\,(D-0.28209479)}{0.02998598}.$$

否定域为$Y\leqslant Y_{\alpha/2}$或$Y\geqslant Y_{1-\alpha/2}$,$Y_{0.025}=-2.39$,$Y_{0.9725}=1.5$,

$$Y=\frac{0.0074-0.2821}{0.0021}=-130.81\leqslant-2.39$$

故拒绝H_0,这批数据与正态分布不相符.

习题 7 Bayes 方法和统计决策理论

1、设参数θ的先验分布为贝塔（$Bata$）分布$Be(\alpha,\beta)$,若从先验信息中获得其均值和方差分别为1/3与1/45，试确定该先验分布.

解：

$$均值=\frac{\alpha}{\alpha+\beta}=\frac{1}{3}，方差=\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}=\frac{1}{45}$$

解得$\alpha=3,\beta=6$,故θ的先验分布为$Be(3,6)$.

2、设θ的先验分布是伽马分布，其均值为10，方差为5，试确定θ的先验分布.

解：

$$均值=\frac{\alpha}{\lambda}=10，方差=\frac{\alpha}{\lambda^2}=5$$

解得$\alpha=20$，$\lambda=2$，故先验分布为$Ga(20,2)$.

3、设θ是一批产品的不合格率，已知它不是0.1就是0.2，且其先验分布为$\pi(0.1)=0.7$, $\pi(0.2)=0.3$.假如从这批产品中随机抽取8个进行检查，发现有2个不合格，求θ的后验分布.

解：

设事件A为随机抽取8个进行检查，有2个不合格

$$令\theta_1=0.1，\theta_2=0.2，则\pi(\theta_1)=0.7,\pi(\theta_2)=0.3$$

$$P(A|\theta_1)=28(0.1)^2(0.9)^6=0.148803$$
$$P(A|\theta_2)=28(0.2)^2(0.8)^6=0.293601$$

故根据全概率公式有

$$P(A)=P(A|\theta_1)\pi(\theta_1)+P(A|\theta_2)\pi(\theta_2)=0.19224$$

故后验分布为

$$\pi(\theta_1|A)=\frac{P(A|\theta_1)}{P(A)}=0.774,$$

$$\pi(\theta_1|A)=\frac{P(A|\theta_2)}{P(A)}=0.226.$$

4、设一卷磁带上的缺陷数服从$poisson$分布$P(\lambda)$,其中λ可取1.0和1.5中的一个，又设λ的先验分布为$\pi(1.0)=0.4,\pi(1.5)=0.6$.假如检查一卷磁带发现3个缺陷，求$\lambda$的后验分布.

解：

设事件A为检查一卷磁带发现3个缺陷.

$$令\theta_1=1.0，\theta_2=1.5，则\pi(\theta_1)=0.4,\pi(\theta_2)=0.6$$

$$P(A|\theta_1)=\left(1+1+\frac{1}{2!}+\frac{1}{3!}\right)e^{-1}=0.981012$$

$$P(A|\theta_2)=\left(1+\frac{1.5}{1!}+\frac{1.5^2}{2!}+\frac{1.5^3}{31}\right)e^{-1.5}=0.833139$$

故根据全概率公式有

$$P(A)=P(A|\theta_1)\pi(\theta_1)+P(A|\theta_2)\pi(\theta_2)=0.89229$$

故后验分布为

$$P(\theta_1|A)=\frac{P(A|\theta_1)\pi(\theta_1)}{P(A)}=0.43977$$

$$P(\theta_2|A)=\frac{P(A|\theta_2)\pi(\theta_2)}{P(A)}=0.56023$$

5、考虑一个试验，对给定的θ，试验结果X有如下密度函数：

$$p(x|\theta)=2x/\theta^2,\ \ 0<x<\theta<1.$$

（1）假如θ的先验分布是（$0,1$）上的均匀分布，试求θ的后验分布；

（2）假如θ的先验密度是$\pi(\theta)=3\theta^2, 0<\theta<1$,试求$\theta$的后验分布.

解：

（1）先验分布为

$$\pi(\theta)=1, 0<\theta<1.$$

则后验密度为

$$\pi(\theta|x)=\frac{p(x|\theta)\pi(\theta)}{\int_0^1 p(x|\theta)\pi(\theta)d\theta}=\frac{1/\theta^2}{\int_{x_{(n)}}^1 1/\theta^2 d\theta}=\frac{x_{(n)}}{\theta^2(1-x_{(n)})}.$$

（2）后验密度为

$$\pi(\theta|x)=\frac{p(x|\theta)\pi(\theta)}{\int_0^1 p(x|\theta)\pi(\theta)d\theta}=\frac{2x/\theta^2\cdot 3\theta^2}{\int_0^1 2x/\theta^2\cdot 3\theta^2 d\theta}=1$$

故后验分布为$U(0,1)$.

6、对下列每个分布中的未知参数使用$Fisher$信息量决定$Jeffreys$无信息先验：

（1）$Poisson$分布$P(\lambda)$;

（2）二项分布$b(n,p)$,其中n已知;

（3）$Gamma$分布$\Gamma(\alpha,\lambda)$，其中α已知.

解：

（1）

$$P(x=k)=\frac{\lambda^k}{k!}e^{-\lambda}, P(\boldsymbol{x}|\lambda)=\frac{\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!}e^{-\lambda n}$$

对数似然函数为

$$l(\lambda|\boldsymbol{x}) = \sum_{i=1}^{n} x_i \ln\lambda - \sum_{i=1}^{n} \ln x_i! - \lambda n.$$

故有

$$I(\lambda) = E_{X|\lambda}\left\{-\frac{\partial^2 l}{\partial\lambda^2}\right\} = E\left\{\frac{\sum_{i=1}^{n} x_i}{\lambda^2}\right\} = \frac{n}{\lambda}$$

故$\pi(\lambda) = [I(\lambda)]^{1/2} = \sqrt{\frac{n}{\lambda}}$.

（2）

$$P(\boldsymbol{x}|\theta) = \prod_{i=1}^{n} \binom{n}{x_i} \theta^{x_i} (1-\theta)^{n-x_i}$$

对数似然函数为

$$l(\theta|\boldsymbol{x}) = \sum_{i=1}^{n} \ln\binom{n}{x_i} + \sum_{i=1}^{n} x_i \ln\theta + \left(n^2 - \sum_{i=1}^{n} x_i\right)\ln(1-\theta)$$

故有

$$I(\theta) = E\left\{-\frac{\partial^2 l}{\partial\theta^2}\right\} = E\left\{\frac{\sum_{i=1}^{n} x_i}{\theta^2} + \frac{n^2 - \sum_{i=1}^{n} x_i}{(1-\theta)^2}\right\} = \frac{n^2}{\theta} + \frac{n^2}{1-\theta}$$

故$\pi(\theta) = [I(\theta)]^{1/2} = n\sqrt{\frac{1}{\theta} + \frac{1}{1-\theta}}$

（3）

$$f(\boldsymbol{x}|\lambda) = \prod_{i=1}^{n} \frac{\lambda^\alpha}{\Gamma(\alpha)} x_i^{\alpha-1} e^{-\lambda x_i} = \left(\frac{\lambda^\alpha}{\Gamma(\alpha)}\right)^n \prod_{i=1}^{n} x_i^{\alpha-1} e^{-\lambda\sum_{i=1}^{n} x_i}$$

对数似然函数为

$$l(\lambda|\boldsymbol{x}) = n\alpha\ln\lambda - n\ln\Gamma(\alpha) + (\alpha-1)\sum_{i=1}^{n} \ln x_i - \lambda\sum_{i=1}^{n} x_i$$

故有

$$I(\theta) = E\left\{-\frac{\partial^2 l}{\partial\theta^2}\right\} = E\left\{\frac{n\alpha}{\lambda^2}\right\} = \frac{n\alpha}{\lambda^2}$$

故$\pi(\theta) = [I(\theta)]^{1/2} = \frac{\sqrt{n\alpha}}{\lambda}$.

7、设随机变量X服从负二项分布，其概率分布为

$$f(x|p)=\binom{x-1}{k-1}p^{k}(1-p)^{x-k},\ x=k,k+1,\cdots.$$

证明其成功概率p的共轭先验分布族为$Beta$分布族.

解：

取先验分布为$Beta$分布，即

$$\pi(p)=\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}p^{a-1}(1-p)^{b-1}.$$

故有

$$\begin{aligned}\pi(p|x)&=\frac{\binom{x-1}{k-1}p^{k}(1-p)^{x-k}\cdot\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}p^{a-1}(1-p)^{b-1}}{\int_0^1\binom{x-1}{k-1}p^{k}(1-p)^{x-k}\cdot\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}p^{a-1}(1-p)^{b-1}dp}\\&=\frac{p^{k+a-1}(1-p)^{x-k+b-1}}{\int_0^1 p^{k+a-1}(1-p)^{x-k+b-1}dp}\\&=\frac{\Gamma(k+a+x-k+b)}{\Gamma(k+a)\Gamma(x-k+b)}p^{k+a-1}(1-p)^{x-k+b-1}\end{aligned}$$

即后验分布是$Beta$分布$Be(k+a,x-k+b)$.先验分布和后验分布都是$Beta$分布，故成功概率p的共轭先验分布族为$Beta$分布族.

8、设随机变量X服从参数为λ的指数分布$Exp(\lambda)$,证明指数分布中参数λ的共轭先验分布族为$Gamma$分布族.

解：

取先验分布为$Gamma(\alpha,\beta)$分布，即

$$\pi(\lambda)=\frac{\beta^{\alpha}}{\Gamma(\alpha)}\lambda^{\alpha-1}e^{-\beta\lambda}$$

故有

$$\begin{aligned}\pi(\lambda|x)&=\frac{\lambda e^{-\lambda x}\frac{\beta^{\alpha}}{\Gamma(\alpha)}\lambda^{\alpha-1}e^{-\beta\lambda}}{\int_0^1\lambda e^{-\lambda x}\frac{\beta^{\alpha}}{\Gamma(\alpha)}\lambda^{\alpha-1}e^{-\beta\lambda}d\lambda}=\frac{\lambda^{\alpha}e^{-(x+\beta)\lambda}}{\int_0^1\lambda^{\alpha}e^{-(x+\beta)\lambda}}\\&=\frac{(x+\beta)^{\alpha+1}}{\Gamma(\alpha+1)}\lambda^{\alpha}e^{-(x+\beta)\lambda}\end{aligned}$$

即后验分布是$Gamma$分布$Ga(\alpha+1,x+\beta)$.先验分布和后验分布都是$Gamma$分布，故参数λ的共轭先验分布族为$Gamma$分布族.

9、设随机变量X服从指数型分布，其密度函数为$f(x|\theta)=\exp\{a(\theta)b(x)+c(\theta)+d(x)\}$,其中$a(\theta),c(\theta)$是$\theta$的函数，$b(x),d(x)$是$x$的函数.证明先验分布$h(x)=Ae^{k_1a(\theta)+k_2c(\theta)}$是参数$\theta$的共轭先验分布，其中$A$为常数，$k_1,k_2$是与$\theta$无关的常数.

解：

取先验分布$h(\theta)=Ae^{k_1a(\theta)+k_2c(\theta)}$，则后验分布

$$h(\theta|x)\propto f(x|\theta)h(\theta)\propto\exp\{a(\theta)b(x)+c(\theta)\}\cdot\exp\{k_1a(\theta)+k_2c(\theta)\}\propto e^{[k_1+b(x)]a(\theta)+[k_2+1]c(\theta)}$$

易知上式右边是“$Ae^{[k_1+b(x)]a(\theta)+[k_2+1]c(\theta)}$的核”，故得证.

10、设随机变量X的密度函数为

$$f(x|\lambda)=\begin{cases}\lambda^{-1}e^{-x/\lambda}, & 0<x<\infty.\\ 0, & x\leqslant 0.\end{cases}$$

证明参数λ的共轭先验分布族为逆$Gamma$分布族.

解：

设λ服从逆$Gamma$分布$\Gamma^{-1}(\alpha,\beta)$，故

$$\pi(\lambda)=\frac{\beta^{\alpha}}{\Gamma(\alpha)}\lambda^{-(\alpha+1)}e^{-\frac{\beta}{\lambda}}.$$

则后验分布

$$\pi(\lambda|x)\propto f(x|\lambda)\pi(\lambda)\propto\lambda^{-1}e^{-x/\lambda}\lambda^{-(\alpha+1)}e^{-\beta/\lambda}=\lambda^{-(\alpha+2)}e^{-(x+\beta)/\lambda}$$

易知上式右边是$\Gamma^{-1}(\alpha+1,\beta+x)$的核，故得证.

11、设$\boldsymbol{X}=(X_1,\cdots,X_n)$是从均匀分布$U(0,\theta)$中抽取的简单样本，又假设$\theta$的先验分布是$Pareto$分布，其密度函数为

$$\pi(\theta)=\begin{cases}\dfrac{\alpha\theta_0^{\alpha}}{\theta^{\alpha+1}}, & \theta>\theta_0,\\ 0, & \theta\leqslant\theta_0.\end{cases}$$

其中$\theta_0>0,\alpha>0$.证明$Pareto$分布是$U(0,\theta)$中端点θ的共轭先验分布.

解：

$$f(\boldsymbol{x}|\theta)=\frac{1}{\theta^n}\propto\frac{1}{\theta^n}$$

$$\pi(\theta|\boldsymbol{x})\propto f(\boldsymbol{x}|\theta)\pi(\theta)\propto\frac{1}{\theta^n}\cdot\frac{1}{\theta^{\alpha+1}}\propto\frac{1}{\theta^{n+\alpha+1}}$$

故有

$$\pi(\theta|\boldsymbol{x})=\begin{cases}\dfrac{(n+\alpha)\theta_0^{n+\alpha}}{\theta^{n+\alpha+1}},\theta>\theta_0,\\0, \qquad\qquad \theta\geqslant\theta_0.\end{cases}$$

故$Pareto$分布是$U(0,\theta)$中端点θ的共轭先验分布.

12、从一批产品中抽检100个，发现3个不合格品，假定产品不合格率θ的先验分布为$Beta$分布$Be(2,200)$,求θ的后验分布.

解：

设事件A为抽检100个，发现3个不合格品，则

$$p(x|\theta)=\binom{100}{3}\theta^3(1-\theta)^{97}$$

$$\pi(\theta|x)=\frac{p(x|\theta)\pi(\theta)}{\int_0^1 p(x|\theta)\pi(\theta)d\theta}=\frac{\theta^3(1-\theta)^{97}\cdot\theta(1-\theta)^{199}}{\int_0^1\theta^4(1-\theta)^{296}d\theta}=\frac{\Gamma(302)}{\Gamma(5)\Gamma(297)}\theta^4(1-\theta)^{296}$$

故θ的后验分布为$Beta(5,297)$.

13、设$X_1,\cdots,X_n$为从正态总体$N(\theta,2^2)$中抽取随机样本，又设θ的先验分布为正态分布.

（1）若样本容量$n=100$,证明不管先验标准差为多少，后验标准差一定小于1/5.

（2）若θ的先验分布的标准差为1，要使后验方差不超过0.1，最少要抽取样本容量多大的样本?

解：

（1）

设先验分布为$N(\mu,\tau^2)$,则

$$\pi(\theta|\bar{x})\sim N(\mu_n(\bar{x}),\eta_n^2).$$

其中

$$\eta_n^2=\frac{\sigma^2\tau^2}{n\tau^2+\sigma^2}$$

故$n=100$时，

$$\eta_n^2=\frac{4\tau^2}{4+100\tau^2}<\frac{4\tau^2}{100\tau^2}=\frac{1}{25},\text{故}\eta_n<\frac{1}{5}.$$

（2）

$$\eta_n^2=\frac{\sigma^2\tau^2}{n\tau^2+\sigma^2}=\frac{4}{n+4}<0.1$$

解得

$$n>36.$$

14、设随机变量X服从几何分布$p(x|\theta)=P(X=x|\theta)=\theta(1-\theta)^{x-1},x=1,2,\cdots$,其中参

数θ的先验分布为均匀分布$U(0,1)$.

（1）若只对X作一次观察，观察值为3，求θ的后验期望估计；
（2）若对X作三次观察，观察值为$3,2,5$，求θ的后验期望估计.

解：;
（1）

$$p(x|\theta)=\theta(1-\theta)^2,\pi(\theta)=1$$

$$\pi(\theta|x)\propto p(x|\theta)\pi(\theta)\propto\theta(1-\theta)^2\sim Be(2,3)$$

故后验期望估计为

$$E(\theta|x)=\hat{\theta}_E=\frac{2}{2+3}=0.4$$

（2）

$$p(x|\theta)=\theta^3(1-\theta)^7,\pi(\theta)=1,$$

$$\pi(\theta|x)\propto p(x|\theta)\pi(\theta)\propto\theta^3(1-\theta)^7\sim Be(4,8)$$

故后验期望估计为

$$E(\theta|x)=\hat{\theta}_E=\frac{4}{4+8}=\frac{1}{3}$$

15、对正态总体$N(\theta,1)$作三次观察，获得样本的具体观察值为$2,4,3$.若θ的先验分布为正态分布$N(3,1)$,求θ的可信系数为0.95的可信区间.

解：

设先验分布为$N(\mu,\tau^2)$,则

$$\pi(\theta|\bar{x})\sim N(\mu_n(\bar{x}),\eta_n^2).$$

其中

$$\mu_n(\bar{x})=\frac{\sigma^2}{\sigma^2+n\tau^2}\mu+\frac{n\tau^2}{\sigma^2+n\tau^2}\bar{x},\eta_n^2=\frac{\sigma^2\tau^2}{n\tau^2+\sigma^2}$$

解得

$$\mu_n(\bar{x})=\frac{1}{1+3}\cdot3+\frac{3}{1+3}\cdot3=3,\eta_n^2=\frac{1\times1}{3+1}=0.25$$

故θ的可信系数为0.95的可信区间为

$$[\mu_n(\bar{x})-\eta_n u_{\alpha/2},\mu_n(\bar{x})+\eta_n u_{\alpha/2}]=[2.02,3.98]$$

16、接到船运来的一大批零件，从中抽检5件，看有多少次品，假设零件中的次品数X服从二项分布$b(5,\theta)$.从以往各批的情况中已知θ的先验分布为贝塔分布$Be(1,9)$,若观测值$X=0$,求θ的可信系数为95%且长度最短的可信区间.

解：由例7.2.9可知θ的后验分布为$Be(x+a,n-x+b)$,即$Be(1,14)$.

可信区间满足以下条件

$$P(c\leqslant\theta\leqslant d)=1-\alpha$$

故有

$$\int_c^d 14(1-\theta)^{13}d\theta=(1-c)^{14}-(1-d)^{14}=1-\alpha$$

$g(\theta)=14(1-\theta)^{13}$是关于$\theta$的单调减函数.故$c=0$,时$d-c$最小，即$d=1-\sqrt[14]{\alpha}$时，可信区间长度最短.

17、设X为从总体中抽取的样本，总体密度$f(x|\theta)$和θ的先验密度$\pi(\theta)$分别为

$$f(x|\theta)=\begin{cases}e^{-(x-\theta)}, & x>\theta,\\ 0, & x\geqslant\theta.\end{cases}\quad \pi(\theta)=\begin{cases}e^{-\theta}, & \theta>0,\\ 0, & \theta\leqslant 0.\end{cases}$$

求检验问题$H_0:\theta\leqslant 1\leftrightarrow H_1:\theta>1$.

解：

$$\pi(\theta|x)=\frac{f(x|\theta)\pi(\theta)}{\int_0^x f(x|\theta)\pi(\theta)d\theta}=\frac{e^{-x}}{\int_0^x e^{-x}d\theta}=x$$

$$\alpha_0=P(\theta\in\Theta_0)=\int_0^1 e^{-x}d\theta=e^{-x}$$

$$\alpha_1=P(\theta\in\Theta_1)=\int_1^x e^{-x}d\theta=(x-1)e^{-x}$$

$\dfrac{\alpha_0}{\alpha_1}=\dfrac{1}{x-1}$,故当$x<2$时，肯定$H_0$,$x\geqslant 2$时，拒绝$H_0$.

18、设$X_1,\cdots,X_m\ i.i.d.\sim N(a,1)$,$Y_1,\cdots,Y_n\ i.i.d.\sim N(b,1)$,其中$a,b$为参数，又样本$X_1,\cdots,X_m$，$Y_1,\cdots,Y_n$独立.令$\theta=(a,b)$的先验分布为：$a,b$独立，$a\sim N(\mu_1,\tau_1^2)$,$b\sim N(\mu_2,\tau_2^2)$.求检验问题$H_0:a-b\leqslant 0\leftrightarrow H_1:a-b>0$.

解：

a,b的后验分布分别为

$$\pi(a|\boldsymbol{x})\sim N(\mu_m(\bar{x}),\eta_m^2),\pi(b|\boldsymbol{y})\sim N\left(\mu_n(\bar{y}),\eta_n^2\right)$$

其中

$$\mu_m(\bar{x})=\frac{1}{1+m\tau_1^2}\mu_1+\frac{m\tau_1^2}{1+m\tau_1^2}\bar{x},\eta_n^2=\frac{\tau_1^2}{m\tau_1^2+1}$$

$$\mu_n(\bar{y})=\frac{1}{1+n\tau_2^2}\mu_2+\frac{n\tau_2^2}{1+n\tau_2^2}\bar{y},\eta_n^2=\frac{\tau_2^2}{n\tau_2^2+1}$$

因为a,b独立，故

$$\pi(a,b|\boldsymbol{x},\boldsymbol{y})=\frac{1}{2\pi\eta_m\eta_n}\exp\left\{-\frac{(a-\mu_m(\bar{x}))^2}{2\eta_m^2}-\frac{\left(b-\mu_n(\bar{y})\right)^2}{2\eta_n^2}\right\}$$

因此

$$\alpha_0=P(\theta\in\Theta_0|\boldsymbol{x},\boldsymbol{y})=\int_{-\infty}^{+\infty}\int_{-\infty}^{b}\frac{1}{2\pi\eta_m\eta_n}\exp\left\{-\frac{(a-\mu_m(\bar{x}))^2}{2\eta_m^2}-\frac{\left(b-\mu_n(\bar{y})\right)^2}{2\eta_n^2}\right\}dadb$$

$$\alpha_1=P(\theta\in\Theta_1|\boldsymbol{x},\boldsymbol{y})=\int_{-\infty}^{+\infty}\int_{b}^{+\infty}\frac{1}{2\pi\eta_m\eta_n}\exp\left\{-\frac{(a-\mu_m(\bar{x}))^2}{2\eta_m^2}-\frac{\left(b-\mu_n(\bar{y})\right)^2}{2\eta_n^2}\right\}dadb$$

若$\dfrac{\alpha_0}{\alpha_1}>1$，则接受$H_0$,否则拒绝$H_0$.

19、设$X_1,\cdots,X_n$是来自正态总体$N(\mu,1)$的一个样本，其中未知参数μ的先验分布是$N(0,1)$,在平方损失函数即$L\left(\mu,\hat{\mu}\right)=\left(\mu-\hat{\mu}\right)^2$下求$\mu$的$Bayes$估计.

解：

已知若先验分布为$N(\mu,\tau^2)$,则

$$\pi(\mu|\bar{x})\sim N(\mu_n(\bar{x}),\eta_n^2).$$

其中

$$\mu_n(\bar{x})=\frac{\sigma^2}{\sigma^2+n\tau^2}\mu+\frac{n\tau^2}{\sigma^2+n\tau^2}\bar{x},\eta_n^2=\frac{\sigma^2\tau^2}{n\tau^2+\sigma^2}$$

故平方损失下μ的$Bayes$估计是

$$\hat{\mu}_B(x)=E(\mu|x)=\mu_n(\bar{x})=0+\frac{1}{1/n+1}\bar{x}=\frac{n}{n+1}\overline{X}.$$

20、设$X_1,\cdots,X_n$是来自参数为θ的几何分布（参见题14）的一个样本，假如未知参数θ的先验分布为$Beta$分布$Be(\alpha,\beta)$.在平方损失函数下求θ的$Bayes$估计.

解：

$$p(\boldsymbol{x}|\theta)=\theta^n(1-\theta)^{\sum_{i=1}^n x_i-n},\pi(\theta)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}$$

$$\pi(\theta|\boldsymbol{x})=\frac{p(\boldsymbol{x}|\theta)\pi(\theta)}{\int_0^1 p(\boldsymbol{x}|\theta)\pi(\theta)d\theta}=\frac{\theta^{n+\alpha-1}(1-\theta)^{\sum_{i=1}^n x_i-n+\beta-1}}{\int_0^1\theta^{n+\alpha-1}(1-\theta)^{\sum_{i=1}^n x_i-n+\beta-1}}$$

$$=\frac{\Gamma\left(\alpha+\beta+\sum_{i=1}^n x_i\right)}{\Gamma(n+\alpha)\Gamma\left(\sum_{i=1}^n x_i-n+\beta\right)}\theta^{n+\alpha-1}(1-\theta)^{\sum_{i=1}^n x_i-n+\beta-1}$$

故平方损失函数下θ的$Bayes$估计为

$$\hat{\theta}_B(x) = E(\theta|x) = \frac{n+\alpha}{\alpha+\beta+\sum_{i=1}^{n} x_i}.$$

21、设$X_1,\cdots,X_n$为抽自参数为θ的$Poisson$分布的一个样本，假如未知参数θ的先验分布是伽马分布，其密度为

$$h(\theta) = \begin{cases} \theta e^{-\theta}, \theta > 0, \\ 0, \quad \theta \leqslant 0. \end{cases}$$

在平方损失函数下，求θ的$Bayes$估计.

解：

$$p(\boldsymbol{x}|\theta) = \frac{\theta^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!} e^{-n\theta}$$

$$h(\theta|\boldsymbol{x}) = \frac{p(\boldsymbol{x}|\theta)h(\theta)}{\int_0^{+\infty} p(\boldsymbol{x}|\theta)h(\theta)d\theta} = \frac{\theta^{1+\sum_{i=1}^{n} x_i} e^{-(n+1)\theta}}{\int_0^{+\infty} \theta^{1+\sum_{i=1}^{n} x_i} e^{-(n+1)\theta} d\theta} = \frac{(n+1)^{\sum_{i=1}^{n} x_i + 2}}{\Gamma\left(\sum_{i=1}^{n} x_i + 2\right)} \theta^{1+\sum_{i=1}^{n} x_i} e^{-(n+1)\theta}$$

故θ的后验分布为$Ga\left(\sum_{i=1}^{n} x_i + 2, n+1\right)$，故$\theta$的$Bayes$估计为

$$\hat{\theta}_B(x) = E(\theta|x) = \frac{\sum_{i=1}^{n} x_i + 2}{n+1}.$$

22、某产品的寿命服从指数分布$Exp(\theta)$.对n个这种产品进行寿命试验，获得了一个样本$X_1,\cdots,X_n$，假如未知参数θ的先验分布为$Gamma$分布$\Gamma(\beta,\alpha)$.在平方损失下，试求θ和$1/\theta$的$Bayes$估计.

解：

$$f(\boldsymbol{x}|\theta) = \theta^n e^{-\theta\sum_{i=1}^{n} x_i}, \pi(\theta) = \frac{\alpha^\beta}{\Gamma(\beta)} \theta^{\beta-1} e^{-\alpha\theta}$$

$$\pi(\theta|\boldsymbol{x}) = \frac{f(\boldsymbol{x}|\theta)\pi(\theta)}{\int_0^{+\infty} f(\boldsymbol{x}|\theta)\pi(\theta)d\theta} = \frac{\theta^{n+\beta-1} e^{-\left(\alpha+\sum_{i=1}^{n} x_i\right)\theta}}{\int_0^{+\infty} \theta^{n+\beta-1} e^{-\left(\alpha+\sum_{i=1}^{n} x_i\right)\theta} d\theta}$$

$$= \frac{\left(\alpha+\sum_{i=1}^{n} x_i\right)^{n+\beta}}{\Gamma(n+\beta)} \theta^{n+\beta-1} e^{-\left(\alpha+\sum_{i=1}^{n} x_i\right)\theta}$$

即θ的后验分布为$Ga\left(n+\beta, \alpha+\sum_{i=1}^{n} x_i\right)$,则$\theta$的$Bayes$估计为

$$\hat{\theta}_B(x)=E(\theta|x)=\frac{n+\beta}{\alpha+\sum_{i=1}^{n}x_i}$$

$1/\theta$的$Bayes$估计为

$$\widehat{(1/\theta)}_B=E(1/\theta|x)=\frac{n+\beta}{\alpha+\sum_{i=1}^{n}x_i-1}$$

23、设$X_1,\cdots,X_n$是从均匀分布$\{U(0,\theta),\theta>0\}$中抽取的简单样本.设θ的先验分布为$U(0,a),a>0$已知.求平方损失下θ的$Bayes$估计.

解：

$f(\boldsymbol{x}|\theta)=\dfrac{1}{\theta^n}I_{(0,\theta)}(x_{(n)}),\pi(\theta)=\dfrac{1}{a}I_{(0,a)}(\theta)$

$$\pi(\theta|\boldsymbol{x})=\frac{f(\boldsymbol{x}|\theta)\pi(\theta)}{\int_0^{+\infty}f(\boldsymbol{x}|\theta)\pi(\theta)d\theta}=\frac{\frac{1}{\theta^n}}{\int_{x_{(n)}}^{a}\frac{1}{\theta^n}d\theta}=\frac{n-1}{[x_{(n)}^{1-n}-a^{1-n}]\theta^n}$$

故θ的$Bayes$估计为

$$\hat{\theta}_B(x)=E(\theta|x)=\int_{x_{(n)}}^{a}\frac{n-1}{[x_{(n)}^{1-n}-a^{1-n}]\theta^{n-1}}d\theta=\frac{n-1}{n}\cdot\frac{x_{(n)}^{-n}-a^{-n}}{x_{(n)}^{1-n}-a^{1-n}}.$$

24、设$X_1,\cdots,X_n$是来自正态总体$N(0,\tau)$的一个样本，其中方差τ的先验分布为逆$Gamma$分布，即

$$h(\tau)=\begin{cases}\dfrac{\alpha^\beta}{\Gamma(\beta)}\tau^{-(\beta+1)}e^{-\frac{\alpha}{\tau}},\tau>0,\\ 0, \qquad\qquad \tau\leqslant 0.\end{cases}$$

在加权平方损失函数$L(\tau,\hat{\tau})=(\tau-\hat{\tau})^2/\tau^2$下，求$\tau$的$Bayes$估计.

解：

$$f(\boldsymbol{x}|\tau)=\left(\frac{1}{\sqrt{2\pi\tau}}\right)^n e^{-\frac{\sum_{i=1}^{n}x_i^2}{2\tau}}$$

$$h(\tau|\boldsymbol{x})=\frac{f(\boldsymbol{x}|\tau)h(\tau)}{\int_0^{+\infty}f(\boldsymbol{x}|\tau)h(\tau)d\tau}=\frac{\tau^{-\left(\frac{n}{2}+\beta+1\right)}e^{-\frac{\alpha+\frac{1}{2}\sum_{i=1}^{n}x_i^2}{\tau}}}{\int_0^{+\infty}\tau^{-\left(\frac{n}{2}+\beta+1\right)}e^{-\frac{\alpha+\frac{1}{2}\sum_{i=1}^{n}x_i^2}{\tau}}}$$

$$=\frac{\left(\alpha+\frac{1}{2}\sum_{i=1}^{n}x_i^2\right)^{\frac{n}{2}+\beta}}{\Gamma\left(\frac{n}{2}+\beta\right)}\tau^{-\left(\frac{n}{2}+\beta+1\right)}e^{-\frac{\alpha+\frac{1}{2}\sum_{i=1}^{n}x_i^2}{\tau}}$$

故τ的后验分布为$\Gamma^{-1}\left(\frac{n}{2}+\beta,\alpha+\frac{1}{2}\sum_{i=1}^{n}x_i^2\right)$，设$k=\frac{n}{2}+\beta,\lambda=\alpha+\frac{1}{2}\sum_{i=1}^{n}x_i^2,$.

$$\begin{aligned}E(\tau w(\tau)|\boldsymbol{x})=E(\tau^{-1}|\boldsymbol{x})&=\int_0^{+\infty}\frac{\lambda^k}{\Gamma(k)}\tau^{-(k+2)}e^{-\frac{\lambda}{\tau}}d\tau\\&=\frac{\Gamma(k+1)}{\lambda\Gamma(k)}\int_0^{+\infty}\frac{\lambda^{k+1}}{\Gamma(k+1)}\tau^{-(k+2)}e^{-\frac{\lambda}{\tau}}d\tau\\&=\frac{k}{\lambda}\end{aligned}$$

$$\begin{aligned}E(w(\tau)|\boldsymbol{x})=E(\tau^{-2}|\boldsymbol{x})&=\int_0^{+\infty}\frac{\lambda^k}{\Gamma(k)}\tau^{-(k+3)}e^{-\frac{\lambda}{\tau}}d\tau\\&=\frac{\Gamma(k+2)}{\lambda^2\Gamma(k)}\int_0^{+\infty}\frac{\lambda^{k+2}}{\Gamma(k+2)}\tau^{-(k+3)}e^{-\frac{\lambda}{\tau}}d\tau\\&=\frac{k(k+1)}{\lambda^2}\end{aligned}$$

因此得到

$$\hat{\tau}_B=\frac{E(\tau w(\tau)|\boldsymbol{x})}{E(w(\tau)|\boldsymbol{x})}=\frac{\lambda}{k+1}=\frac{\alpha+\frac{1}{2}\sum_{i=1}^{n}x_i^2}{\frac{n}{2}+\beta+1}.$$

25、设$X\sim b(n,\theta),\theta\sim Be(\alpha,\beta)$,在损失函数

$$L(a,\theta)=\frac{(\theta-a)^2}{\theta(1-\theta)}$$

下，求θ的$Bayes$估计.

解：

$$\begin{aligned}\pi(\theta|x)&=\frac{f(x|\theta)\pi(\theta)}{\int_0^1 f(x|\theta)\pi(\theta)d\theta}=\frac{\theta^{x+\alpha-1}(1-\theta)^{n-x+\beta-1}}{\int_0^1\theta^{x+\alpha-1}(1-\theta)^{n-x+\beta-1}d\theta}\\&=\frac{\Gamma(n+\alpha+\beta)}{\Gamma(x+\alpha)\Gamma(n-x+\beta)}\theta^{x+\alpha-1}(1-\theta)^{n-x+\beta-1}\end{aligned}$$

即θ的后验分布为$Be(x+\alpha,n-x+\beta)$,即$k=x+\alpha,\lambda=n-x+\beta$,则

$$\begin{aligned}E(\theta w(\theta)|x)&=\int_0^1\frac{\Gamma(k+\lambda)}{\Gamma(k)\Gamma(\lambda)}\theta^{k-1}(1-\theta)^{\lambda-2}d\theta\\&=\frac{\Gamma(\lambda-1)\Gamma(k+\lambda)}{\Gamma(\lambda)\Gamma(k+\lambda-1)}\int_0^1\frac{\Gamma(k+\lambda-1)}{\Gamma(k)\Gamma(\lambda-1)}\theta^{k-1}(1-\theta)^{\lambda-2}d\theta\\&=\frac{k+\lambda-1}{\lambda-1}\end{aligned}$$

$$E(w(\theta)|x)=\int_0^1\frac{\Gamma(k+\lambda)}{\Gamma(k)\Gamma(\lambda)}\theta^{k-2}(1-\theta)^{\lambda-2}d\theta$$
$$=\frac{\Gamma(k+\lambda)\Gamma(k-1)\Gamma(\lambda-1)}{\Gamma(k)\Gamma(\lambda)\Gamma(k+\lambda-2)}\int_0^1\frac{\Gamma(k+\lambda-2)}{\Gamma(k-1)\Gamma(\lambda-1)}\theta^{k-2}(1-\theta)^{\lambda-2}d\theta$$
$$=\frac{(k+\lambda-1)(k+\lambda-2)}{(k-1)(\lambda-1)}$$

因此得到

$$\hat{\theta}_B=\frac{E(\theta w(\theta)|x)}{E(w(\theta)|x)}=\frac{k-1}{k+\lambda-2}=\frac{x+\alpha-1}{n+\alpha+\beta-2}.$$

26、接到船运来的一大批零件，从中抽检5件.假设其中不合格品数$X\sim b(5,\theta)$,又从以往各批的先验信息中确定θ的先验分布为$Be(1,9)$.若观察值为$x=0$,在下列损失函数下求检验问题：

$$H_0:0\leqslant\theta\leqslant0.15\leftrightarrow H_1:\theta>0.15.$$

设行动a_i表示接受$H_i,i=0,1$.

(1) $L(a_0,\theta)=\begin{cases}0,\theta\leqslant0.15,\\1,\theta>0.15,\end{cases}\qquad L(a_1,\theta)=\begin{cases}2,\theta\leqslant0.15,\\0,\theta>0.15.\end{cases}$

(2) $L(a_0,\theta)=\begin{cases}0,\theta\leqslant0.15,\\1,\theta>0.15,\end{cases}\qquad L(a_1,\theta)=\begin{cases}0.15-\theta,\theta\leqslant0.15,\\0,\qquad\theta>0.15.\end{cases}$

解：$p(x|\theta)=\binom{5}{x}\theta^x(1-\theta)^{5-x},\pi(\theta)=9(1-\theta)^8$

$$\pi(\theta|x)=\frac{p(x|\theta)\pi(\theta)}{\int_0^1p(x|\theta)\pi(\theta)d\theta}$$
$$=\frac{\theta^x(1-\theta)^{13-x}}{\int_0^1\theta^x(1-\theta)^{13-x}d\theta}$$
$$=\frac{\Gamma(15)}{\Gamma(x+1)\Gamma(14-x)}\theta^x(1-\theta)^{13-x}$$

$x=0$,故$\theta|x\sim Be(1,14)$

（1）$R(a_0|x)=\int\limits_{\Theta_1}\pi(\theta|x)d\theta=\int_{0.15}^1 14(1-\theta)^{13}d\theta=0.103$

$R(a_1|x)=\int_0^{0.15}2\times14(1-\theta)^{13}d\theta=1.794,$

$R(a_0|x)<R(a_1|x)$,故接受H_0,

（2）$R(a_0|x)=\int_{0.15}^1 14(1-\theta)^{13}d\theta=0.103$

$R(a_1|x)=\int_0^{0.15}14(0.15-\theta)(1-\theta)^{13}d\theta=0.011$

$R(a_0|x)>R(a_1|x)$,故否定H_0,接受H_1.

27、在上题中若取损失函数为$0-1$损失，其他不变，考虑检验问题

$$H_0:0\leqslant\theta\leqslant0.15\leftrightarrow H_1:\theta>0.15.$$

解：

$$P(\Theta_0|x)=P(0\leqslant\theta\leqslant0.15\,|x)=\int_0^{0.15}14(1-\theta)^{13}d\theta=0.897$$
$$P(\Theta_1|x)=P(\theta>0.15\,|x)=\int_{0.15}^{1}14(1-\theta)^{13}d\theta=0.103$$

由于$P(\Theta_0|x)>P(\Theta_1|x)$,故接$H_0$.

28、设随机变量X服从二项分布$b(n,\theta),0<\theta<1$,证明$d(x)=x/n$在损失函数$L(d,\theta)=(d-\theta)^2/[\theta(1-\theta)]$下是$\theta$的极小极大估计.（提示：先验分布取均匀分布）

解：

取先验分布为$U(0,1)$,则

$$p(x|\theta)=C_n^x\theta^x(1-\theta)^{n-x},\pi(\theta)=1,$$

则θ的后验分布为

$$\pi(\theta|x)=\frac{p(x|\theta)\pi(\theta)}{\int_0^1p(x|\theta)\pi(\theta)d\theta}=\frac{\theta^x(1-\theta)^{n-x}}{\int_0^1\theta^x(1-\theta)^{n-x}d\theta}=\frac{\Gamma(n+2)}{\Gamma(x+1)\Gamma(n-x+1)}\theta^x(1-\theta)^{n-x}$$

$$E(\theta w(\theta)|x)=\int_0^1\frac{\Gamma(n+2)}{\Gamma(x+1)\Gamma(n-x+1)}\theta^x(1-\theta)^{n-x-1}d\theta=\frac{n+1}{n-x}$$
$$E(w(\theta)|x)=\int_0^1\frac{\Gamma(n+2)}{\Gamma(x+1)\Gamma(n-x+1)}\theta^{x-1}(1-\theta)^{n-x-1}d\theta=\frac{n(n+1)}{x(n-x)}$$

故θ的$Bayes$估计为

$$\hat{\theta}_B=\frac{E(\theta w(\theta)|x)}{E(w(\theta)|x)}=\frac{x}{n}$$

风险函数为

$$R(d,\theta)=E\left[\frac{\left(\frac{X}{n}-\theta\right)^2}{\theta(1-\theta)}\right]=\frac{1}{\theta(1-\theta)}E\left(\frac{X^2}{n^2}-2\frac{\theta X}{n}+\theta^2\right)$$

其中$EX=n\theta,EX^2=n\theta(1-\theta)+n^2\theta^2$

故

$$R(d,\theta)=\frac{1}{\theta(1-\theta)}E\left(\frac{X^2}{n^2}-2\frac{\theta X}{n}+\theta^2\right)=\frac{1}{n}$$

$R(d,\theta)$为常数，由定理7.5.1可知$\frac{x}{n}$是$Minimax$估计.

29、设$X_1,\cdots,X_n$为自正态总体$N(0,\sigma^2)$中抽取的$i.i.d.$样本，令$\tau=\sigma^2$的先验分布为无信息先验，即$\pi(\tau)=(1/\tau)I_{(0,\infty)}(\tau)$,取损失函数为$L(d,\tau)=(d-\tau)^2/\tau^2$,求$\tau$的贝叶斯估计$\hat{\tau}_B$,并证明$\hat{\tau}_B$是$\tau$的$Minimax$估计.

解：

$$\pi(\tau|\boldsymbol{x})=\frac{\left(\frac{1}{\sqrt{2\pi\tau}}\right)^n e^{-\frac{\sum_{i=1}^n x_i^2}{2\tau}}\frac{1}{\tau}}{\int_0^{+\infty}\left(\frac{1}{\sqrt{2\pi\tau}}\right)^n e^{-\frac{\sum_{i=1}^n x_i^2}{2\tau}}\frac{1}{\tau}d\tau}=\frac{\left(\sum_{i=1}^n x_i^2\Big/2\right)^{\frac{n}{2}}}{\Gamma(n/2)}\tau^{-(n/2+1)}e^{-\frac{\sum_{i=1}^n x_i^2/2}{\tau}}$$

即后验分布为逆$Gamma$分布，

$$E(\tau w(\tau)|\boldsymbol{x})=\int_0^{+\infty}\frac{1}{\tau}\pi(\tau|\boldsymbol{x})d\tau=\frac{\Gamma(n/2+1)}{\sum_{i=1}^n x_i^2\Big/2\cdot\Gamma(n/2)}$$

$$E(w(\tau)|\boldsymbol{x})=\int_0^{+\infty}\frac{1}{\tau^2}\pi(\tau|\boldsymbol{x})d\tau=\frac{\Gamma(n/2+2)}{\left(\sum_{i=1}^n x_i^2\Big/2\right)^2\cdot\Gamma(n/2)}$$

故

$$\hat{\tau}_B=\frac{E(\tau w(\tau)|\boldsymbol{x})}{E(w(\tau)|\boldsymbol{x})}=\frac{\sum_{i=1}^n x_i^2\Big/2\cdot\Gamma(n/2+1)}{\Gamma(n/2+2)}=\frac{1}{n+2}\sum_{i=1}^n x_i^2$$

风险函数为

$$R(d,\tau)=E\left[\left(\frac{1}{n+2}\sum_{i=1}^n x_i^2-\tau\right)^2\Big/\tau^2\right]=E\left(\frac{1}{n+2}\sum_{i=1}^n x_i^2\Big/\tau-1\right)^2$$

因为

$$\sum_{i=1}^n x_i^2\Big/\tau\sim\chi_n^2$$

故

$$\begin{aligned}R(d,\tau)&=E\left[\frac{1}{(n+2)^2}\left(\sum_{i=1}^n x_i^2\Big/\tau\right)^2-\frac{2}{n+2}\sum_{i=1}^n x_i^2\Big/\tau+1\right]\\&=\frac{n^2+2n}{(n+2)^2}-\frac{2n}{n+2}+1=2\end{aligned}$$

$R(d,\tau)$为常数，由定理7.5.1可知$\hat{\tau}_B$是$Minimax$估计.

30、设$X_1,\cdots,X_n$ $i.i.d.\sim N(\theta,1)$,要估计θ.令损失函数为平方损失，取估计量为

$\hat{\theta}_n = c_1 X_1 + \cdots + c_n X_n$.证明若$c_1 + \cdots + c_n = 1$,则除非$c_1 = \cdots = c_n = 1/n$,$\hat{\theta}_n$是不可容许的.

解：

$c_1 X_1 + \cdots + c_n X_n \sim N\left(\theta, \sum_{i=1}^{n} c_i^2\right)$

$R(\delta,\theta) = E(c_1 X_1 + \cdots + c_n X_n - \theta)^2 = E(c_1 X_1 + \cdots + c_n X_n)^2 + \theta^2 - 2\theta E(c_1 X_1 + \cdots + c_n X_n)$

$= \sum_{i=1}^{n} c_i^2 + \theta^2 + \theta^2 - 2\theta^2 = \sum_{i=1}^{n} c_i^2 \geqslant \frac{\left(\sum_{i=1}^{n} c_i\right)^2}{n} = \frac{1}{n}$

当且仅当$c_1 = \cdots = c_n = 1/n$时成立，且$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} x_i \in \Theta$，故

若$c_1 + \cdots + c_n = 1$,当$c_1, \cdots, c_n$不为$c_1 = \cdots = c_n = 1/n$情况时，存在δ_1一致优于δ，即δ不可容许.