§1 点估计

设总体X的分布函数$F(x;\theta)$的形式为已知，θ是待估参数。$X_1\cdots X_n$是X的一个样本，$x_1\cdots x_n$是相应的样本值。

点估计问题：

构造一个适当的统计量$\theta(X_1,\cdots,X_n)$，用它的观察值$\hat{\theta}(x_1,\cdots,x_n)$来估计未知参数$\theta$。

我们称$\theta(X_1,\cdots,X_n)$为θ的<u>估计量</u>；称$\hat{\theta}(x_1,\cdots,x_n)$为$\theta$<u>估计值</u>。

1. 矩估计法

设X为连续型随机变量，其概率密度为$f(x;\theta_1,\cdots,\theta_k)$，

X为离散型随机变量，其分布列为$P\{X=x\}=P(x;\theta_1,\cdots,\theta_k)$，

其中$\theta_1,\cdots,\theta_k$是待估参数,，$X_1,\cdots,X_n$为来自X的样本。

设 $EX^l=\mu_l, l=1,2,\cdots,k$.存在。

则 $A_l=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i^l$

令 $A_l=\mu_l,\quad l=1,\cdots,k$

这里是包含k个未知参数$\theta_1,\cdots,\ \theta_k$的联立方程组，

从中解出方程组的解$\hat{\theta}_1,\cdots,\ \hat{\theta}_k$。

用$\hat{\theta}_1,\cdots,\ \hat{\theta}_k$分别作为$\theta_1,\cdots,\ \theta_k$的估计量，这种求估计量的方法称为矩估计法。

这种估计量称为矩估计量；矩估计量的观察值称为矩估计值。

例 1 设某炸药厂一天中发生着火现象的次数X服从参数为λ的泊松分布，λ未知，有以下样本值；试估计参数λ（用矩法）。

着火的次数k	0	1	2	3	4	5	6	
发生k次着火天数n_k	75	90	54	22	6	2	1	$\sum = 250$

解： $\mu_1 = EX = \lambda \quad A_1 = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$

令 $\overline{X} = \lambda,$

则 $\hat{\lambda} = \overline{x} = \frac{1}{250}(0\times 75 + 1\times 90 + \cdots + 6\times 1) = 1.22$

返回主目录

所以 $\overline{X}=\lambda$, 估计值$\hat{\lambda}=1.22$。

例2. 设总体$X\sim U[a,b]$, a,b未知；$X_1,\cdots,X_n$是一个样本；

求：a,b的矩估计量。

解： $\mu_1=EX=\dfrac{a+b}{2}$,

$$\mu_2=EX^2=DX+(EX)^2=\frac{(b-a)^2}{12}+\frac{(a+b)^2}{4}$$

令 $$\frac{a+b}{2}=A_1=\frac{1}{n}\sum_{i=1}^{n}X_i$$

$$\frac{(b-a)^2}{12}+\frac{(a+b)^2}{4}=A_2=\frac{1}{n}\sum_{i=1}^{n}X_i^2$$

即 $a+b=2A_1,\quad b-a=\sqrt{12(A_2-A_1^2)}$

解得：$\hat{a}=A_1-\sqrt{3(A_2-A_1^2)}=\bar{X}-\sqrt{\dfrac{3}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2}$

$$\hat{b}=A_1+\sqrt{3(A_2-A_1^2)}=\bar{X}+\sqrt{\dfrac{3}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2}$$

例3. 设总体X的均值μ，方差σ都存在，且$\sigma^2>0$,

但μ，σ^2未知，又设$X_1,\cdots,X_n$是一个样本；

求：μ,σ^2的矩估计量。

解：$\mu_1=EX=\mu,$

$$\mu_2=EX^2=DX+(EX)^2=\sigma^2+\mu^2$$

令 $\mu_1=A_1,\quad \mu_2=A_2,$

即 $\mu=A_1,\quad \sigma^2+\mu^2=A_2,$

所以 $\hat{\mu}=A_1=\overline{X},$

$$\hat{\sigma}^2=A_2-A_1^2=\frac{1}{n}\sum_{i=1}^{n}X_i^2-\overline{X}^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$$

返回主目录

特别，若 $X \sim N(\mu,\sigma^2)$，μ,σ^2未知；

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

2. 极大似然估计法

(1).若总体X属离散型，其分布律$P\{X=x\}=p(x;\theta),\theta\in\Theta$的形式为已知，$\theta$为待估参数，$\Theta$是$\theta$可能取值的范围。

设$X_1,\cdots,X_n$是来自X的样本；则$X_1,\cdots,X_n$的联合分布律

$$\prod_{i=1}^{n}p(x_i;\theta)$$

又设$x_1,\cdots,x_n$是$X_1,\cdots,X_n$的一个样本值；

易知样本$X_1,\cdots,X_n$取$x_1,\cdots,x_n$的概率，亦即

事件$\{X_1=x_1,\cdots,X_n=x_n\}$发生的概率为：

返回主目录

$$L(\theta)=L(x_1,\cdots,x_n;\theta)=\prod_{i=1}^{n}p(x_i;\theta),\ \theta\in\Theta. \quad (1.1)$$

它是θ的函数。$L(\theta)$称为样本的<u>似然函数</u>。

由极大似然估计法：固定$x_1,\cdots,x_n$；挑选使概率$L(x_1,\cdots,x_n;\theta)$达到最大的参数$\hat{\theta}$，作为θ的估计值，即取$\hat{\theta}$使得：

$$L(x_1,\cdots,x_n;\hat{\theta})=\max_{\theta\in\Theta}L(x_1,\cdots,x_n;\theta) \quad (1.2)$$

$\hat{\theta}$与$x_1,\cdots,x_n$有关，记为$\hat{\theta}(x_1,\cdots,x_n)$；

称其为参数θ的<u>极大似然估计值</u>。

$\hat{\theta}(X_1,\cdots,X_n)$称为参数$\theta$的<u>极大似然估计量</u>。

(2).若总体X属连续型，其概率密度$f(x;\theta),\theta\in\Theta$的形式已知，$\theta$为待估参数;

则$X_1,\cdots,X_n$的联合密度：

$$\prod_{i=1}^{n} f(x_i;\theta)$$

设$x_1,\cdots,x_n$是相应$X_1,\cdots,X_n$的一个样本值，则随机点$(X_1,\cdots,X_n)$落在$(x_1,\cdots,x_n)$的邻域（边长分别为$dx_1,\cdots,dx_n$的n维立方体）内的概率近似为：

$$\prod_{i=1}^{n} f(x_i;\theta)dx_i \qquad (1.3)$$

我们取θ的估计值$\hat{\theta}$，使概率(1.3)取到最大值。

但$\prod_i dx_i$不随θ而变，故只需考虑：

$$L(\theta)=L(x_1,\cdots,x_n;\theta)=\prod_{i=1}^{n} f(x_i;\theta), \qquad (1.4)$$

的最大值，这里$L(\theta)$称为样本的似然函数。

若 $$L(x_1,\cdots,x_n;\hat{\theta})=\max_{\theta\in\Theta} L(x_1,\cdots,x_n;\theta)$$

则称$\hat{\theta}(x_1,\cdots,x_n)$为$\theta$的极大似然估计值。

称$\hat{\theta}(X_1,\cdots,X_n)$为$\theta$的极大似然估计量。

一般，$p(x;\theta), f(x;\theta)$关于θ可微，故θ可由下式求得：

$$\frac{dL(\theta)}{d\theta}=0.$$

又因$L(\theta)$与$\ln L(\theta)$在同一θ处取到极值，因此θ的极大似然估计θ也可从下述方程解得：

$$\frac{d}{d\theta}\ln L(\theta)=0. \tag{1.5}$$

若母体的分布中包含多个参数，

即可令 $\dfrac{\partial L}{\partial \theta_i}=0, i=1,\cdots,k.$或$\dfrac{\partial \ln L}{\partial \theta_i}=0, i=1,\cdots,k.$

解k个方程组求得$\theta_1,\cdots,\theta_k$的极大似然估计值。

例4. 设$X \sim B(1,p)$; $X_1,\cdots,X_n$是来自X的一个样本，试求参数p的极大似然估计量。

解：设$x_1,\cdots,x_n$是一个样本值。X的分布律为：

$$P\{X=x\}=p^x(1-p)^{1-x}, \quad x=0,1;$$

故似然函数为

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

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

令 $$\frac{d}{dp}\ln L(p)=\frac{\sum_{i=1}^{n}x_i}{p}-\frac{n-\sum_{i=1}^{n}x_i}{1-p}=0.$$

返回主目录

解得p的极大似然估计值

$$\hat{p} = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$$

p的极大似然估计量为

$$\hat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}$$

--------它与矩估计量是相同的。

例5. 设$X \sim N(\mu,\sigma^2)$; μ,σ^2为未知参数，$x_1,\cdots,x_n$

是来自X的一个样本值，

求：μ,σ^2的极大似然估计量。

解：X的概率密度为：

$$f(x;\mu,\sigma^2)=\frac{1}{\sqrt{2\pi}\sigma}\exp\{-\frac{1}{2\sigma^2}(x-\mu)^2\}$$

似然函数为：

$$L(\mu,\sigma^2)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}\sigma}\exp\{-\frac{1}{2\sigma^2}(x_i-\mu)^2\}$$

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

返回主目录

令 $\begin{cases} \dfrac{\partial \ln L}{\partial \mu} = 0 \\ \dfrac{\partial \ln L}{\partial \sigma^2} = 0 \end{cases}$ 即: $\begin{cases} \dfrac{1}{\sigma^2}[\sum\limits_{i=1}^{n} x_i - n\mu] = 0 \\ -\dfrac{n}{2\sigma^2} + \dfrac{1}{(2\sigma^2)^2}\sum\limits_{i=1}^{n}(x_i - \mu)^2 = 0 \end{cases}$

解得: $\hat{\mu} = \dfrac{1}{n}\sum\limits_{i=1}^{n} x_i = \bar{x}$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

例6. 设$X \sim U[a,b]$;a,b未知，$x_1,\cdots,x_n$是一个样本值，求：a,b的极大似然估计量。

解：设 $x_{(1)} = \min(x_1,\cdots,x_n), x_{(n)} = \max(x_1,\cdots,x_n)$,

X的概率密度为:

$$f(x;a,b) = \begin{cases} \dfrac{1}{b-a}, a \le x \le b; \\ 0 \qquad ,\text{其它} \end{cases}$$

因为$a \le x_1,\cdots,x_n \le b$,等价于$a \le x_{(1)}, x_{(n)} \le b$,

$$L(a,b) = \begin{cases} \dfrac{1}{(b-a)^n}, a \le x_{(1)}, b \ge x_{(n)}; \\ 0 \qquad ,\text{其它} \end{cases}$$

返回主目录

对于满足$a \le x_{(1)}, b \ge x_{(n)}$的任意$a,b$有

$$L(a,b) = \frac{1}{(b-a)^n} \le \frac{1}{(x_{(n)} - x_{(1)})^n}$$

即：$L(a,b)$在$a = x_{(1)}, b = x_{(n)}$时，取最大值$(x_{(n)} - x_{(1)})^n$

故a,b的极大似然估计值为：

$$\hat{a} = x_{(1)} = \min x_i, \quad \hat{b} = x_{(n)} = \max x_i,$$

故a,b的极大似然估计量为：

$$\hat{a} = \min X_i, \quad \hat{b} = \max X_i,$$

性质：设θ的函数$u=u(\theta),\theta\in\Theta$具有单值反函数，

$\hat{\theta}$是θ的极大似然估计；

则$\hat{u}=u(\hat{\theta})$是$u(\theta)$的极大似然估计。

例：$\sigma^2=\dfrac{1}{n}\sum\limits_{i=1}^{n}(X_i-\overline{X})^2$是$\sigma$的极大似然估计

$u=u(\sigma^2)=\sqrt{\sigma^2}$有单值反函数$\sigma^2=u^2,(u\geq 0)$

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

是σ的极大似然估计

§2 估计量的标准

1. 无偏性：若$\hat{\theta}=\hat{\theta}(X_1,\cdots,X_n)$的数学期望存在，

且$E\hat{\theta}=\theta$.

则称$\hat{\theta}$是θ的无偏估计量。

2. 有效性：若$\hat{\theta}_1=\hat{\theta}_1(X_1,\cdots,X_n)$，$\hat{\theta}_2=\hat{\theta}_2(X_1,\cdots,X_n)$

都是θ的无偏估计量；若$\mathrm{D}(\hat{\theta}_1)<\mathrm{D}(\hat{\theta}_2)$.

则称$\hat{\theta}_1$较$\hat{\theta}_2$有效。

3. 一致性：若$\hat{\theta}=\hat{\theta}_1(X_1,\cdots,X_n)$为参数$\theta$的估计量

若对于任意$\theta\in\Theta$，当$n\longrightarrow\infty$时$\hat{\theta}\xrightarrow{p}\theta$.

则称$\hat{\theta}$是θ的一致估计。

返回主目录

§3 区间估计

区间估计要求根据样本给出未知参数的一个范围，并保证真参数以指定的较大概率属于这个范围。

1. 置信区间与置信度

定义：设总体X含一待估参数θ；对于样本$x_1,\cdots,x_n$，找出统计量$\theta_i=\theta_i(x_1,\cdots,x_n)(i=1,2),\theta_1<\theta_2$，使得：

$$P\{\theta_1\le\theta\le\theta_2\}=1-\alpha,\quad(0<\alpha<1)$$

称区间$[\theta_1，\theta_2]$为θ的<u>置信区间</u>，$1-\alpha$为该区间的<u>置信度</u>。

区间$[\theta_1，\theta_2]$是一个随机区间；$1-\alpha$给出该区间含真值θ的可靠程度。α表示该区间不包含真值 θ的可能性。

返回主目录

例如：若$\alpha=5\%$，即置信度为$1-\alpha=95\%$.

这时重复抽样100次，则在得到的100个区间中包含θ真值的有95个左右，不包含θ真值的有5个左右。

通常，采用95%的置信度，有时也取99%或90%

2. 均值的区间估计

设$x_1,\cdots,x_n$为总体$X\sim N(\mu,\sigma^2)$的一个样本。

在置信度$1-\alpha$下，来确定μ的置信区间$[\theta_1, \theta_2]$。

(1). 已知方差，估计均值

设已知方差$\sigma^2=\sigma_0^2$，且知道$\overline{x}=\dfrac{1}{n}\sum\limits_{i=1}^{n}x_i$是$\mu$的一个点估计，又知道$u=\dfrac{\overline{x}-\mu}{\sigma_0/\sqrt{n}}\sim N(0,1)$。

返回主目录

对于给定的置信度$1-\alpha$，查正态分布表，找出临界值λ_1，λ_2，使得：

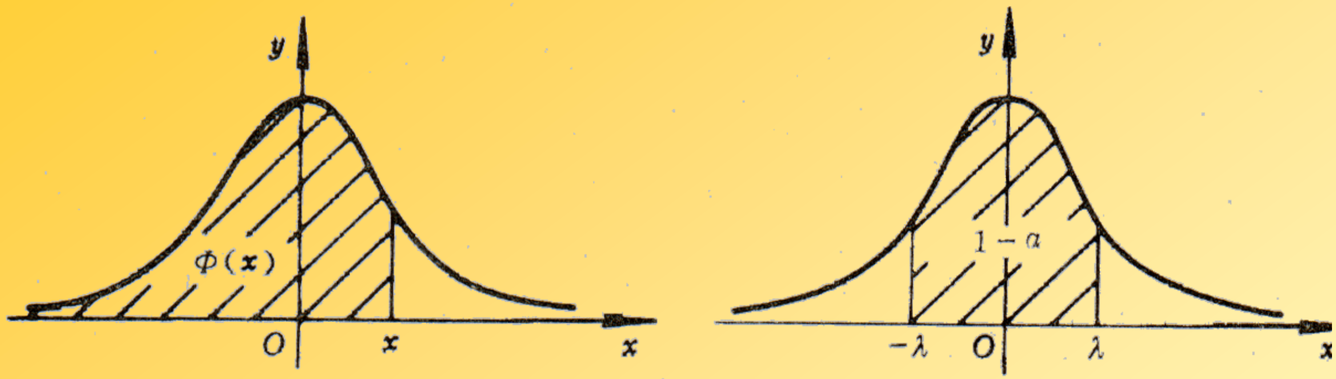
$$P\{\lambda_1 \le u \le \lambda_2\} = 1-\alpha.$$

由此可找出无穷多组λ_1，λ_2；通常我们取对称区间$[-\lambda,\lambda]$,使：

$$\mathrm{P}\{|\mathrm{u}| \le \lambda\} = 1\text{-}\alpha$$

即：$$\mathrm{P}\{-\lambda \le \frac{\overline{\mathrm{X}}-\mu}{\sigma_0/\sqrt{n}} \le \lambda\} = 1\text{-}\alpha$$

由正态分布表的构造，由$P\{|t|\le\lambda\}=1-\alpha$，可知：



查正态分布表$\Phi(\lambda)=1-\alpha/2$,找出λ，得：

$$-\lambda\le\frac{(\overline{x}-\mu)\sqrt{n}}{\sigma_0}\le\lambda$$

推得，随机区间：

$$[\overline{x}-\lambda\frac{\sigma_0}{\sqrt{n}},\overline{x}+\lambda\frac{\sigma_0}{\sqrt{n}}]$$

它以$1-\alpha$的概率包含μ。

返回主目录

例6. 已知幼儿身高服从正态分布，现从5~6岁的幼儿中随机地抽查了9人，其高度分别为：

115,120,131,115,109,115,115,105,110cm;

假设标准差$\sigma_0 = 7$，置信度为95%;

试求总体均值μ的置信区间。

解：已知$\sigma_0 = 7, n = 9, \alpha = 0.05.$ 由样本值算得：

$$\bar{x} = \frac{1}{9}(115 + 120 + \cdots + 110) = 115.$$

查正态分布表得临界值$\lambda = 1.96$，由此得置信区间：

$$\left[115 - 1.96 \times 7/\sqrt{9}\,,\, 115 + 1.96 \times 7/\sqrt{9}\right]$$

$$= [110.43\,,\, 119.57]$$

(2). 未知方差，估计均值

由于未知方差σ^2，这时可用样本方差：

$$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

而选取样本函数： $t=\dfrac{\bar{x}-\mu}{S/\sqrt{n}}$

则随机变量t服从n-1个自由度的t分布。

对于给定的$1-\alpha$，查t分布表，得临界值λ_1与λ_2，使得：

$$P\{\lambda_1\le t\le\lambda_2\}=1-\alpha,$$

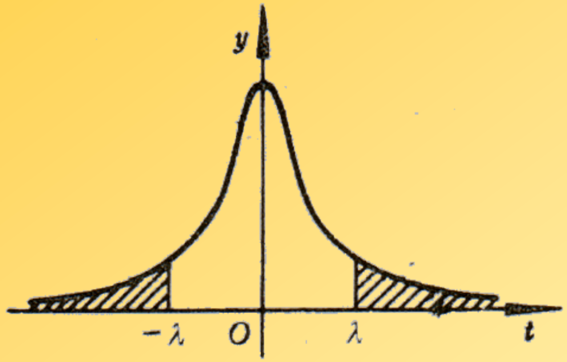
我们仍然取成对称区间$[-\lambda,\lambda]$，使得：

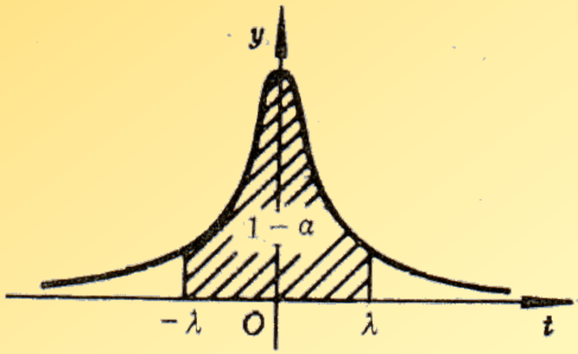
$$P\{|t|\le\lambda\}=1-\alpha,$$

即 $P\{-\lambda \le \dfrac{\overline{x}-\mu}{S/\sqrt{n}} \le \lambda\} = 1-\alpha,$

由t分布表的构造，比较$P\{|t|> \lambda\} = \alpha$与

$P\{|t| \le \lambda\} = 1-\alpha$，可知：





查t分布表$t(n-1,\alpha/2)$，找出λ.

其中，n是样本容量，n-1是表中自由度；由此得：

$$-\lambda \le \frac{\overline{x}-\mu}{S/\sqrt{n}} \le \lambda$$

推得，随机区间：

$$[\bar{x}-\lambda\frac{S}{\sqrt{n}},\bar{x}+\lambda\frac{S}{\sqrt{n}}]$$

它以$1-\alpha$的概率包含μ。

例7. 用仪器测量温度，重复测量7次，测得温度分别为：115,120,131,115,109,115,115；设温度

$X \sim N(\mu,\sigma^2)$。

在置信度为95%时，试求温度的真值所在范围。

解：设μ是温度的真值，X是测量值。

已知$n=7,\alpha=0.05$. 由样本值算得：

$$\bar{x}=112.8,\quad S^2=1.29.$$

查$t(6,0.025)$得临界值$\lambda = 2.447$。由此得置信区间：

$$\left[112.8 - 2.447\sqrt{\frac{1.29}{7}},\quad 112.8 + 2.447\sqrt{\frac{1.29}{7}}\right]$$

$$= [111.75,\quad 113.85]$$

3. 方差的区间估计

设$x_1,\cdots,x_n$为总体$X \sim N(\mu,\sigma^2)$的一个样本。

我们知道$S^2 = \dfrac{1}{n-1}\sum\limits_{i=1}^{n}(x_i - \bar{x})^2$是$\sigma^2$的一个点估计

并且知道样本函数：$\varpi = \dfrac{(n-1)S^2}{\sigma^2}$服从$n-1$个

自由度的χ^2分布。

返回主目录

对于给定的$1-\alpha$，查χ^2分布表，得临界值λ_1与λ_2，使得：

$$P\{\lambda_1 \le \varpi \le \lambda_2\} = 1-\alpha,$$

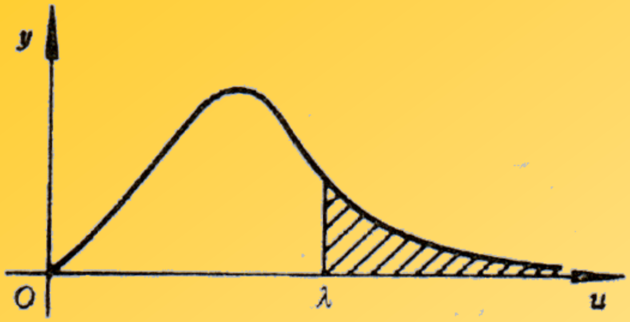
由于χ^2分布无对称性，我们采用使概率对称的区间：

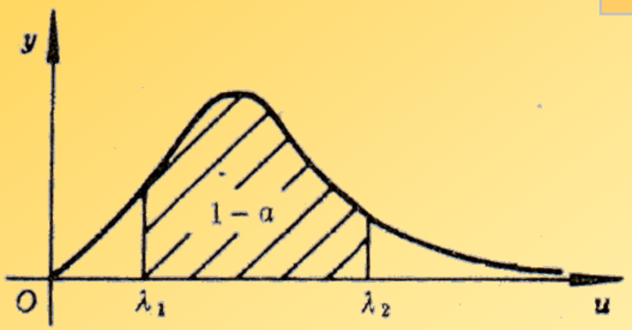
$$P\{\varpi < \lambda_1\} = P\{\varpi > \lambda_2\} = \alpha/2,$$

即 $$P\{\lambda_1 \le \frac{(n-1)S^2}{\sigma^2} \le \lambda_2\} = 1-\alpha,$$

由χ^2分布表的构造，比较$P\{\chi^2 > \lambda\} = \alpha$与

$P\{\lambda_1 \le \varpi \le \lambda_2\} = 1-\alpha$，可知：





查$\chi^2(n-1,\alpha/2)$分布表,找出λ_2,.而查$\chi^2(n-1,1-\alpha/2)$分布表,找出λ_1。

其中，n是样本容量，n-1是表中自由度；由此得：

$$\lambda_1 \le \frac{(n-1)S^2}{\sigma^2} \le \lambda_2$$

推得：$\dfrac{(n-1)S^2}{\lambda_2} \le \sigma^2 \le \dfrac{(n-1)S^2}{\lambda_1}$

这就是说，随机区间：

$$\left[\frac{(n-1)S^2}{\lambda_2},\ \frac{(n-1)S^2}{\lambda_1}\right]$$

以$1-\alpha$的概率包含σ^2，而随机区间

$$\left[\sqrt{\frac{n-1}{\lambda_2}}S,\ \sqrt{\frac{n-1}{\lambda_1}}S\right]$$

以$1-\alpha$的概率包含σ.

例8. 设某机床加工的零件长度 $X \sim N(\mu,\sigma^2)$,
今抽查16个零件，测得长度（单位：mm）如下：

12.15, 12.12, 12.01, 12.08,
12.09, 12.16, 12.03, 12.01,
12.06, 12.13, 12.07, 12.11,
12.08, 12.01, 12.03, 12.06,

在置信度为95%时，试求总体方差 σ^2 的置信区间。

解：已知 $n = 16, \alpha = 0.05$. 由样本值算得：

$$S^2 = 0.00244.$$

查 $\chi^2(15,0.975)$ 得 $\lambda_1 = 6.26$; 查 $\chi^2(15,0.025)$ 得 $\lambda_2 = 27.5$.

由此得置信区间：

$$\left[\frac{15\times 0.00244}{27.5},\quad \frac{15\times 0.00244}{6.26}\right] = [0.0013, 0.0058]$$

返回主目录

1 给出了点估计的概念，要掌握矩估计法、极大似然估计法。

2 了解估计量的评选标准（无偏性、有效性、一致性）。

作业: P_{173} $1, 2, 3, 6, 7.$