第 6 章　参数估计

6.1　参数的点估计

设总体 X 服从某已知分布，如 $N(\mu,\sigma^2)$、$e(\lambda)$ 等，但是其中的一个或多个参数为未知，怎样根据抽取的样本估计未知参数的值，就是参数的估计问题.

首先讨论分布参数的点估计.

设总体 X 的分布中含有未知参数 θ，从总体 X 中抽取样本 $X_1,X_2,\cdots,X_n$，构造某个统计量 $\hat{\theta}(X_1,X_2,\cdots,X_n)$ 作为参数 θ 的估计，则称 $\hat{\theta}(X_1,X_2,\cdots,X_n)$ 为参数 θ 的**点估计量**；若样本 $X_1,X_2,\cdots,X_n$ 的观测值为 $x_1,x_2,\cdots,x_n$，则称 $\hat{\theta}(x_1,x_2,\cdots,x_n)$ 为参数 θ 的**点估计值**.

例如，人的身高 $X\sim N(\mu,\sigma^2)$，一个样本为 $X_1,X_2,\cdots,X_n$，则 $\bar{X}=\frac{1}{n}(X_1+\cdots+X_n)$ 为 n 个人的平均身高，近似认为总体均值 μ 为 $\bar{X}$，即 $\hat{\mu}=\bar{X}$. 用 $\bar{X}$ 来估计 μ，这里 $\hat{\mu}$ 不是真值，而是估计值.

若总体的分布中含有 m（$m>1$）个未知参数，则需构造 m 个统计量作为相应 m 个未知参数的点估计量.

下面介绍两种常用的求未知参数点估计量的方法.

一、矩估计法

设总体 X 的分布中含有未知参数 $\theta_1,\theta_2,\cdots,\theta_m$，假定总体 X 的 1、2、…，m 阶原点矩都存在，一般来说，它们都是 $\theta_1,\theta_2,\cdots,\theta_m$ 的函数，即

$$v_k(X)=E(X^k)=v_k\ (\theta_1,\theta_2,\cdots,\theta_m)\quad (k=1,2,\cdots,m).$$

从总体 X 中抽取样本 $X_1,X_2,\cdots,X_n$，取样本 k 阶原点矩 $V_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k$ 作为总体 X 的 k 阶原点矩 $v_k(X)$ 的估计量，即

$$\hat{v}_k(X)=\frac{1}{n}\sum_{i=1}^{n}X_i^k\qquad (k=1,2,\cdots,m).$$

由此得方程组

$$\begin{cases} v_1(q_1,q_2,\mathbf{L},q_m)=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i\ , \\ v_2(q_1,q_2,\mathbf{L},q_m)=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i^{\ 2}\ , \\ \qquad\qquad \text{M} \\ v_m(q_1,q_2,\mathbf{L},q_m)=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i^{\ m}. \end{cases}$$

求这个方程组的解，得

$$\begin{cases} \hat{q}_1=\hat{q}_1(X_1,X_2,\mathbf{L},X_n), \\ \hat{q}_2=\hat{q}_2(X_1,X_2,\mathbf{L},X_n), \\ \qquad\qquad \text{M} \\ \hat{q}_m=\hat{q}_m(X_1,X_2,\mathbf{L},X_n). \end{cases}$$

它们分别是未知参数$q_1,q_2,\mathbf{L},q_m$的估计量，称为**矩估计量**．如果已知样本观测值为$x_1,x_2,\mathbf{L},x_n$，则矩估计量的观测值

$$\begin{cases} \hat{q}_1=\hat{q}_1(x_1,x_2,\mathbf{L},x_n), \\ \hat{q}_2=\hat{q}_2(x_1,x_2,\mathbf{L},x_n), \\ \qquad\qquad \text{M} \\ \hat{q}_m=\hat{q}_m(x_1,x_2,\mathbf{L},x_n). \end{cases}$$

分别是未知参数$q_1,q_2,\mathbf{L},q_m$的**矩估计值**．这种求未知参数的点估计的方法称为**矩估计法**.

例 6.1.1　设总体X在区间$[0,q]$上服从均匀分布，其中$q>0$是未知参数，如果取得样本观测值为$x_1,x_2,\mathbf{L},x_n$，求q的矩估计值.

解　因为总体X的概率密度

$$f(x;q)=\begin{cases} \dfrac{1}{q}, & (0<x<q), \\ 0, & (\text{其他}). \end{cases}$$

其中只有一个未知参数q，所以只需考虑总体X的一阶原点矩

$$v_1(X)=E(X)=\int_0^q \frac{x}{q}\mathrm{d}x=\frac{q}{2}\ ,$$

用样本一阶原点矩$V_1=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i$作为$v_1(X)$的估计量，即有

$$\frac{q}{2}=\frac{1}{n}\sum_{i=1}^{n}X_i\ ,$$

由此解得q的矩估计量是

$$\hat{q} = \frac{2}{n}\sum_{i=1}^{n} 2X_i = 2\overline{X} ,$$

而 q 的矩估计值就是

$$\hat{q} = \frac{2}{n}\sum_{i=1}^{n} 2x_i = 2\overline{x} .$$

例 6.1.2 设总体 X 服从正态分布 $N(m, s^2)$，其中 m 及 s^2 都是未知参数，如果取得样本观测值为 $x_1, x_2, \mathbf{L}, x_n$，求 m 及 s^2 的矩估计值.

解 因为总体 X 的分布中有两个未知参数，所以应该考虑一、二阶原点矩，于是有

$$v_1(x) = E(X) = m ,$$

$$v_2(x) = E(X^2) = D(X) + [E(X)]^2 = s^2 + m^2 .$$

于是，按矩估计法得方程组

$$\begin{cases} m = \dfrac{1}{n}\sum\limits_{i=1}^{n} X_i , \\ s^2 + m^2 = \dfrac{1}{n}\sum\limits_{i=1}^{n} X_i^2 . \end{cases}$$

解得 m 及 s^2 的矩估计量为

$$\begin{cases} \hat{m} = \dfrac{1}{n}\sum\limits_{i=1}^{n} X_i = \overline{X}, \\ \hat{s}^2 = \dfrac{1}{n}\sum\limits_{i=1}^{n} X_i^2 - \overline{X}^2 = \dfrac{1}{n}\sum\limits_{i=1}^{n} (X_i - \overline{X})^2 . \end{cases}$$

而 m 及 s^2 的矩估计值就是

$$\begin{cases} \hat{m} = \dfrac{1}{n}\sum\limits_{i=1}^{n} x_i = \overline{x}, \\ \hat{s}^2 = \dfrac{1}{n}\sum\limits_{i=1}^{n} (x_i - \overline{x})^2 = s\%^2 . \end{cases}$$

由此可见，总体均值 $E(X)$ 的矩估计值是样本均值 $\overline{x}$；总体方差 $D(X)$ 的矩估计值就是样本二阶中心矩 $s\%^2$. 从解题过程易知，这个结论不仅对正态总体成立，而且无论总体 X 服从什么分布，只要总体的均值与方差存在，上述结论都是成立的.

矩估计法的优点是直观、简便；特别是，正如前面已经指出的，对总体的均值与方差进行估计时，并不一定要知道总体服从什么分布，但是，矩估计法对于那些原点矩不存在的总体是不适用的.

例 6.1.3 设 $X \sim U(a,b)$，一个样本为 $X_1, X_2, \mathbf{L}, X_n$，求参数 a、b 的矩估计.

解 因为

$$\begin{cases} E(X)=\dfrac{a+b}{2} \Rightarrow b=2E(X)-a\ , \\ D(X)=\dfrac{(b-a)^2}{12} \Rightarrow \dfrac{(E(X)-a)^2}{3}=D(X). \end{cases}$$

故

$$a=E(X)-\sqrt{3D(X)},\ \ b=E(X)+\sqrt{3D(X)}\ ,$$

则

$$\hat{a}=\hat{E}(X)-\sqrt{3\hat{D}(X)}=\overline{X}-\sqrt{\frac{3}{n}\sum_{i=1}^{n}(X_i-\overline{X})}\ ,$$

$$\hat{b}=\hat{E}(X)+\sqrt{3\hat{D}(X)}=\overline{X}+\sqrt{\frac{3}{n}\sum_{i=1}^{n}(X_i-\overline{X})}\ .$$

二、最大似然估计法

（1）设总体 X 是离散型随机变量，概率函数是 $p(x;q)$，其中 q 是未知参数. 从总体 X 中抽取样本 $X_1,X_2,\mathbf{L},X_n$， 如果得到的样本观测值为 $x_1,x_2,\mathbf{L},x_n$， 则表明随机事件 $X_1=x_1,X_2=x_2,\mathbf{L},X_n=x_n$ 发生了． 因为随机变量 $X_1,X_2,\mathbf{L},X_n$ 相互独立，并且与总体 X 有相同的概率函数，所以上述 n 个相互独立的随机事件的交的概率为

$$P(\mathop{\mathbf{I}}_{i=1}^{n}(X_i=x_i))=\prod_{i=1}^{n}P(X_i=x_i)=\prod_{i=1}^{n}p(x_i;q)\ ,$$

设

$$L(q)=\prod_{i=1}^{n}p(x_i;q)\ , \tag{6.1.1}$$

函数 $L(q)$ 称为**似然函数**，对于已取定的 $x_1,x_2,\mathbf{L},x_n$， 它是未知参数 q 的函数.

最大似然估计法的直观想法就是：如果抽样的结果得到样本观测值 $x_1,x_2,\mathbf{L},x_n$， 则应当这样选取参数 q 的值，使这组样本观测值出现的可能性最大，也就是使似然函数 $L(q)$ 达到最大值，从而求得参数 q 的估计值 $\hat{q}$． 利用最大似然估计法求得的参数估计值称为**最大似然估计值**.

求未知参数 q 的最大似然估计值的问题，就是求似然函数 $L(q)$ 的最大值点的问题．当似然函数 $L(q)$ 可导时，这个问题可以通过解下面的似然方程

$$\frac{dL(q)}{dq}=0 \tag{6.1.2}$$

来解决. 因为 $\ln L(q)$ 是 $L(q)$ 的增函数，所以 $\ln L(q)$ 与 $L(q)$ 在 q 的同一值处取得最大值．因此，也可以将上面方程换成下面的方程

$$\frac{d\ln L(q)}{dq}=0\ . \tag{6.1.3}$$

解上面两个方程，就可以得到参数q 的最大似然估计值.

例 6.1.4 设总体X服从泊松分布$P(l)$，其中$l>0$ 为未知参数．如果取得样本观测值为$x_1,x_2,\mathbf{L},x_n$，求参数l 的最大似然估计值.

解 有概率函数

$$p(x;l)=\frac{l^x}{x!}\mathrm{e}^{-l}\qquad(x=0,1,2,\cdots),$$

所以，按（6.1.1），似然函数为

$$L(l)=\prod_{i=1}^{n}\left(\frac{l^{x_i}}{x_i!}\mathrm{e}^{-l}\right)=\frac{l^{\sum\limits_{i=1}^{n}x_i}}{\prod\limits_{i=1}^{n}(x_i!)}\mathrm{e}^{-nl}\ ,$$

取对数，得

$$\ln L(l)=\left(\sum_{i=1}^{n}x_i\right)\ln l-\sum_{i=1}^{n}\ln(x_i!)-nl\ ,$$

于是，按（6.1.3）得到方程

$$\frac{d\ln L(l)}{dl}=\frac{1}{l}\sum_{i=1}^{n}x_i-n=0\,,$$

由此解得l 的最大似然估计值为

$$\hat{l}=\frac{1}{n}\sum_{i=1}^{n}x_i=\overline{x}\,.$$

（2）设总体X是连续随机变量，概率密度为$f(x;q)$，其中q是未知参数．从总体X中抽取样本为$X_1,X_2,\mathbf{L},X_n$，如果得到的样本观测值为$x_1,x_2,\mathbf{L},x_n$，则随机变量X_i落在点x_i的领域（设其长度为Δx_i）内的概率近似等于$f(x_i;q)\ \Delta x_i$（$i=1,2,\mathbf{L},n$）．因为这n个随机事件相互独立，所以它们的交的概率近似等于$\prod\limits_{i=1}^{n}f(x_i;q)\Delta x_i$. 按最大似然估计法，应该选择未知参数$q$ 的值使得概率$\prod\limits_{i=1}^{n}f(x_i;q)\Delta x_i$ 达到最大值. 因为Δx_i（$i=1,2,\mathbf{L},n$）与q 无关，所以代替公式(6.1.1)，取似然函数为

$$L(q)=\prod_{i=1}^{n}f(x_i;q)\,,\qquad(6.1.4)$$

再按上述方法求参数q 的最大似然估计值.

例 6.1.5 设总体X服从指数分布$\mathrm{e}(l)$，概率密度为

$$f(x;l)=\begin{cases}l\,\mathrm{e}^{-lx}, & (x>0);\\ 0, & (x\leqslant 0).\end{cases}$$

其中$l>0$为未知参数．如果取得样本观测值为$x_1,x_2,\mathbf{L},x_n$，求参数l 的最大似然

估计值.

解 按公式（6.1.4），似然函数为

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

取对数，得

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

按（6.1.3）得方程

$$\frac{d\ln L(l)}{d\,l}=\frac{n}{l}-\sum_{i=1}^{n}x_i=0,$$

由此解得 l 的最大似然估计为

$$\hat{l}=\frac{n}{\sum_{i=1}^{n}x_i}=\frac{1}{\overline{x}}.$$

当总体 X 的分布中含有多个未知参数 $q_1,q_2,\mathbf{L},q_m$ 时，最大似然估计法也是适用的．这时，得到的似然函数 L 是这些参数的多元函数 $L(q_1,q_2,\mathbf{L},q_m)$．为了求似然函数 L 的最大值，代替方程（6.1.2）或（6.1.3），有方程组：

$$\frac{\partial L}{\partial q_j}=0\quad(j=1,2,\mathbf{L},m);\tag{6.1.5}$$

或

$$\frac{\partial\ln L}{\partial q_j}=0\quad(j=1,2,\mathbf{L},m).\tag{6.1.6}$$

解上面方程组（6.1.5）或（6.1.6），就可以得到参数 $q_1,q_2,\mathbf{L},q_m$ 的最大似然估计值.

例 6.1.6 设总体 X 服从正态分布 $N(m,s^2)$，其中 m 及 $s>0$ 都是未知参数，如果取得样本观测值为 $x_1,x_2,\mathbf{L},x_n$，求参数 m 及 s 的最大似然估计值.

解 似然函数为

$$L(m,s)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}s}\mathrm{e}^{-\frac{(x_i-m)^2}{2s^2}}$$

$$=\left(\frac{1}{\sqrt{2\pi}s}\right)^{n}\mathrm{e}^{-\frac{1}{2s^2}\sum_{i=1}^{n}(x_i-m)^2},$$

取对数，得

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

对 μ 及 σ 求偏导数，并让其等于零，得

$$\begin{cases} \dfrac{\partial \ln L}{\partial \mu} = \dfrac{1}{\sigma^2}\displaystyle\sum_{i=1}^{n}(x_i-\mu)=0; \\ \dfrac{\partial \ln L}{\partial \sigma} = \dfrac{1}{\sigma^3}\displaystyle\sum_{i=1}^{n}(x_i-\mu)^2-\dfrac{n}{\sigma}=0. \end{cases}$$

解此方程组，即得 μ 及 σ 得最大似然估计值是

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

$$\hat{\sigma} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2} = \tilde{\sigma}.$$

我们指出，最大似然估计具有下述性质：

设参数 θ 的函数 $\varphi=\varphi(\theta)$ 具有单值反函数 $\theta=\theta(\varphi)$，如果 $\hat{\theta}$ 是 θ 的最大似然估计，则 $\hat{\varphi}=\varphi(\hat{\theta})$ 是 $\varphi(\theta)$ 的最大似然估计.

因此，由例 6.1.6 可知，正态总体的方差 $\sigma^2(\sigma>0)$ 的最大似然估计值

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2 = \tilde{\sigma}^2.$$

所以，正态总体的均值及方差的最大似然估计与矩估计是相同的.

6.2 衡量点估计量好坏的标准

上一节学习了两种总体分布中求未知参数点估计的方法，它们是矩估计法和最大似然估计法．对于同一个未知参数，用不同的估计法得到的点估计量一般是不相同的，那么哪一个估计量更好呢？为此需要建立判别估计量好坏的标准，而参数 θ 的所谓“最佳估计量” $\hat{\theta}(X_1,X_2,\cdots,X_n)$ 应当是在某种意义下最接近于 θ.

最佳估计量 $\hat{\theta}(X_1,X_2,\cdots,X_n)$ 应符合下列标准：

一、无偏性

设参数 θ 的估计量 $\hat{\theta}=\hat{\theta}(X_1,X_2,\cdots,X_n)$ 的数学期望存在且等于 θ，即

$$E(\hat{\theta})=\theta,$$

则称 $\hat{\theta}$ 是参数 θ 的**无偏估计量**.

设样本观测值为 $x_1,x_2,\cdots,x_n$，则称 $\hat{\theta}(x_1,x_2,\cdots,x_n)$ 为参数 θ 的**无偏估计值**.

显然，用参数 θ 的无偏估计量 $\hat{\theta}$ 代替参数 θ 时所产生的误差的数学期望为零，即不含有系统误差.

例 6.2.1 设总体 X 的均值 $E(X)=\mu$，方差 $D(X)=\sigma^2$，证明

（1）样本均值 $\overline{X}=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i$ 是总体均值 m 的无偏估计量.

（2）样本方差 $S^2=\dfrac{1}{n-1}\sum\limits_{i=1}^{n}(X_i-\overline{X})^2$ 是总体方差 s^2 的无偏估计量.

证：（1）因为样本 $X_1,X_2,\mathbf{L},X_n$ 相互独立，且与总体 X 服从相同分布，所以有

$$E(X_i)=m,D(X_i)=s^2\quad(i=1,\ 2,\ \mathbf{L},n),$$

由于

$$E(\overline{X})=E\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n}E\left(\sum_{i=1}^{n}X_i\right)=\frac{1}{n}\sum_{i=1}^{n}E(X_i)=\frac{1}{n}\sum_{i=1}^{n}m=\frac{1}{n}\cdot nm=m\ ,$$

所以样本均值 $\overline{X}$ 是总体均值 m 的无偏估计量.

（2）因为

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

利用关于方差的定理得

$$E(X_i^2)=D(X_i)+[E(X_i)]^2=s^2+m^2\quad(i=1,2,\mathbf{L},n),$$

$$\begin{aligned}E(\overline{X}^2)&=D(\overline{X})+[E(\overline{X})]^2\\&=D\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)+m^2=\frac{1}{n^2}\sum_{i=1}^{n}D(X_i)+m^2\\&=\frac{1}{n^2}\cdot ns^2+m^2=\frac{s^2}{n}+m^2,\end{aligned}$$

于是，利用关于数学期望的定理得

$$\begin{aligned}E(S^2)&=E\left[\frac{1}{n-1}\left(\sum_{i=1}^{n}X_i^2-n\overline{X}^2\right)\right]\\&=\frac{1}{n-1}\left[\sum_{i=1}^{n}E(X_i^2)-nE(\overline{X}^2)\right]\\&=\frac{1}{n-1}\left[n(s^2+m^2)-n\left(\frac{s^2}{n}+m^2\right)\right]=s^2,\end{aligned}$$

所以，S^2 是 s^2 的无偏估计量：

$$\hat{s}^2=S^2\ .$$

应当指出，同一参数 q 的无偏估计量不是唯一的．例如，有 $E(X_i)=m$，这表明任一样本 $X_i\,(i=1,2,\ldots,n)$ 都是总体均值 m 的无偏估计量．在参数 q 的许多无偏估计量中，当然是以对 q 的平均偏差较小者为好，也就是说，较好的估计量应当有尽可能小的方差．为此，引进点估计的第二个标准.

二、有效性

设 $\hat{\theta}_1=\hat{\theta}_1(X_1,X_2,\cdots,X_n)$ 与 $\hat{\theta}_2=\hat{\theta}_2(X_1,X_2,\cdots,X_n)$ 都是参数 θ 的无偏估计量，如果

$$D(\hat{\theta}_1)<D(\hat{\theta}_2)$$

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

当样本容量 n 一定时，若 θ 的所有无偏估计量中，$\hat{\theta}$ 的方差 $D(\hat{\theta})$ 最小，则称 $\hat{\theta}$ 是参数 θ 的**有效估计量**.

例 6.2.2 证明样本均值 $\overline{X}$ 作为总体均值 μ 的估计量较个别样本 X_i $(i=1,\ 2,\ \cdots,n)$ 有效.

证：由例 6.2.1 知，$\overline{X}$ 与 X_i 都是总体均值 μ 的无偏估计量，即

$$E(\overline{X})=\mu\,,\quad E(X_i)=\mu\quad (i=1,\ 2,\ \cdots,n),$$

又

$$D(\overline{X})=D\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n^2}\sum_{i=1}^{n}D(X_i)=\frac{1}{n^2}n\sigma^2=\frac{\sigma^2}{n}\,,\quad D(X_i)=\sigma^2\quad (i=1,\ 2,\ \cdots,n).$$

所以当 $n\geqslant 2$ 时，$D(\overline{X})<D(X_i)$，故样本均值 $\overline{X}$ 作为总体均值 μ 的估计量较个别样本 X_i $(i=1,\ 2,\ \cdots,n)$ 有效.

例 6.2.3 从总体 X 中抽取样本 X_1,X_2,X_3，证明下列三个统计量

$$\hat{\theta}_1=\frac{X_1}{2}+\frac{X_2}{3}+\frac{X_3}{6}\,,\quad \hat{\theta}_2=\frac{X_1}{2}+\frac{X_2}{4}+\frac{X_3}{4}\,,\quad \hat{\theta}_3=\frac{X_1}{3}+\frac{X_2}{3}+\frac{X_3}{3}$$

都是总体均值 $E(X)=\mu$ 的无偏估计量，并确定哪个估计量更有效.

证：

$$E(\hat{\theta}_1)=E\left(\frac{X_1}{2}+\frac{X_2}{3}+\frac{X_3}{6}\right)=\frac{\mu}{2}+\frac{\mu}{3}+\frac{\mu}{6}=\mu\,,$$

$$E(\hat{\theta}_2)=E\left(\frac{X_1}{2}+\frac{X_2}{4}+\frac{X_3}{4}\right)=\frac{\mu}{2}+\frac{\mu}{4}+\frac{\mu}{4}=\mu\,,$$

$$E(\hat{\theta}_3)=E\left(\frac{X_1}{3}+\frac{X_2}{3}+\frac{X_3}{3}\right)=\frac{\mu}{3}+\frac{\mu}{3}+\frac{\mu}{3}=\mu\,.$$

所以三个统计量都是总体均值 μ 的无偏估计量.

$$D(\hat{\theta}_1)=D\left(\frac{X_1}{2}+\frac{X_2}{3}+\frac{X_3}{6}\right)=\frac{\sigma^2}{4}+\frac{\sigma^2}{9}+\frac{\sigma^2}{36}=\frac{28}{72}\sigma^2\,,$$

$$D(\hat{\theta}_2)=D\left(\frac{X_1}{2}+\frac{X_2}{4}+\frac{X_3}{4}\right)=\frac{\sigma^2}{4}+\frac{\sigma^2}{16}+\frac{\sigma^2}{16}=\frac{27}{72}\sigma^2\,,$$

$$D(\hat{\theta}_3)=D\left(\frac{X_1}{3}+\frac{X_2}{3}+\frac{X_3}{3}\right)=\frac{\sigma^2}{9}+\frac{\sigma^2}{9}+\frac{\sigma^2}{9}=\frac{24}{72}\sigma^2\,,$$

由于 $D(\hat{q}_3)=\dfrac{24}{72}s^2$ 的值最小，所以 $\hat{q}_3$ 是三个估计量中最有效估计量.

还应指出，估计量 $\hat{q}(X_1,X_2,\mathbf{L},X_n)$ 是与样本容量 n 有关的，为了明确起见，不妨记作 $\hat{q}_n$．希望当 n 越大时，对 q 的估计越精确．于是引进点估计的第三个标准.

三、一致性

如果当 $n\to\infty$ 时，$\hat{q}_n$ 按概率收敛于 q，即对于任意给定的正数 e，有

$$\lim_{n\to\infty}P(|\hat{q}_n-q|<e)=1,$$

则称 $\hat{q}_n$ 是参数 q 的**一致估计量**.

例 6.2.4 设总体 X 的均值 $E(X)=m$，方差 $D(X)=s^2$，证明样本均值 $\overline{X}=\dfrac{1}{n}\sum_{i=1}^{n}X_i$ 是总体均值 m 的一致估计量.

证：因为样本 $X_1,X_2,\mathbf{L},X_n$ 相互独立，且与总体 X 服从相同的分布，所以

$$E(X_i)=m,D(X_i)=s^2\quad(i=1,\ 2,\ \mathbf{L},n),$$

于是，由切比雪夫定理知：

$$\lim_{n\to\infty}P\left(\left|\frac{1}{n}\sum_{i=1}^{n}X_i-\frac{1}{n}\sum_{i=1}^{n}E(X_i)\right|<e\right)=\lim_{n\to\infty}P(|\overline{X}-m|<e)=1.$$

所以 $\overline{X}$ 是 m 的一致估计量.

此外，还可以证明：样本方差 S^2 是总体方差 s^2 的一致估计量.

综上所述，对于未知参数 q 的估计量，可以运用无偏性、有效性、一致性来判断其优劣，以便选择出较好的估计量.

6.3　参数的区间估计

前面两节介绍了点估计的概念．点估计只是给出了待估参数或参数函数的值是多少，但无法回答如估计误差有多大、在允许可靠范围之内最大估计误差是多少等这样的问题，这正是下面内容要回答的问题.

区间估计就是将一个未知参数或参数函数值估在一个区间范围之内，例如一个人的年龄，可以估计为 30 岁，这就是点估计；但也可估计其年龄在 29 岁到 31 岁之间，这种估计就是区间估计．从直观上说，后者给人的印象要比前者更为可信，因为后者已经把可能出现的误差考虑在内.

先看下面两个例子.

例 6.3.1 某农作物的平均亩产量 X（单位：kg）服从正态分布 $N(m,100^2)$，今随机抽取 100 亩进行试验，观察其亩产量值 $x_1,x_2,\mathbf{L},x_{100}$，基此算出 $\overline{x}$ =500kg，

因此 m 的点估计值为 500．由于抽样的随机性，m 的真值与 $\overline{x}$ 的值总有误差，希望以 95%的可靠度估计 $\overline{x}$ 与 m 的最大误差是多少？

因为，$\overline{X} \sim N\left(m, \dfrac{s^2}{n}\right)$，从而存在 $c>0$，使得

$$P(|\overline{X}-m| \leqslant c)=0.95,$$

因此，这个 c 就是可允许的最大误差.

注意：事件 $|\overline{X}-m| \leqslant c$ 等价于 $m \in [\overline{X}-c, \overline{X}+c]$，这就是说：随机区间 $[\overline{X}-c, \overline{X}+c]$ 覆盖未知参数 m 有 95%的机会，习惯上称此随机区间为 m 的区间估计.

例 6.3.2 某工厂欲对出厂的一批电子器件的平均寿命进行估计，随机地抽取 n 件产品进行试验，通过对试验的数据的加工得出该批产品是否合格的结论？并要求此结论的可信程度为 95%，应该如何来加工这些数据？

对于“可信程度”如何定义，下面再说，但从常识可以知道，通常对于电子元器件的寿命指标往往是一个范围，而不必是一个很准确的数．因此，在对这批电子元器件的平均寿命估计时，寿命的准确值并不是最重要的，重要的是所估计的寿命是否能以很高的可信程度处在合格产品的指标范围内，这里可信程度是很重要的，它涉及到使用这些电子器件的可靠性．因此，若采用点估计，不一定能达到应用的目的，这就需要引入区间估计.

区间估计粗略地说是用两个统计量 $\hat{q}_1$ 和 $\hat{q}_2$ $(\hat{q}_1 \leqslant \hat{q}_2)$ 所决定的区间 $[\hat{q}_1, \hat{q}_2]$ 作为参数 q 取值范围的估计．显然，一般地这样说是没有多大的意义的，首先，这个估计必须有一定的精度，即是说 $\hat{q}_2-\hat{q}_1$ 不能太大，太大不能说明任何问题；其次，这个估计必须有一定的可信程度，因此 $\hat{q}_2-\hat{q}_1$ 又不能太小，太小难以保证这一要求．比如从区间[1,100]去估计某人的岁数，虽然绝对可信，却不能带来任何有用的信息；反之，若用区间[30,31]去估计某人的岁数，虽然提供了关于此人年龄的信息，却很难使人相信这一结果的正确性．希望既能得到较高的精度，又能得到较高的可信程度，但在获得的信息一定（如样本容量固定）的情况下，这两者显然是不可能同时达到最理想的状态．通常是采取将可信程度固定在某一需要的水平上，求得精度尽可能高的估计区间.

下面给出区间估计的正式的定义.

设 $X_1, X_2, \mathbf{L}, X_n$ 是来自总体 $f(x,q)$ 的样本，对于参数 q，如果有两个统计量 $\hat{q}_1=\hat{q}_1(X_1, X_2, \mathbf{L}, X_n)$，$\hat{q}_2=\hat{q}_2(X_1, X_2, \mathbf{L}, X_n)$，满足对给定的 $a \in (0,1)$，有

$$P\{\hat{q}_1 \leqslant q \leqslant \hat{q}_2\}=1-a,$$

则称区间 $[\hat{q}_1, \hat{q}_2]$ 是 q 的一个**区间估计**或**置信区间**，$\hat{q}_1$，$\hat{q}_2$ 分别称为**置信下限**、**置信上限**，$1-a$ 称为**置信水平**．一旦样本有观测值 $x_1, x_2, \mathbf{L}, x_n$，则称相应的

$$[\hat{q}_1(x_1, x_2, \mathbf{L}, x_n),\ \hat{q}_2(x_1, x_2, \mathbf{L}, x_n)]$$

为**置信区间的观察值**.

这里的置信水平，就是对可信程度的度量．置信水平为$1-a$，在实际上可以这样来理解如取$1-a=95\%$，就是说若对某一参数q取 100 个容量为n的样本，用相同方法做 100 个置信区间．$[\hat{q}_1^{(k)},\hat{q}_2^{(k)}]$($k=1,2,\mathbf{L},100$)，其中有 95 个区间包含真参数$q$．因此，当实际上只做一次区间估计时，有理由认为它包含真参数．这样判断当然也可能犯错误，但犯错误的概率只有 5%.

下面来讨论一下区间估计的一般步骤.

（1）设欲估参数为q，先取q的一个点估计$\hat{q}$，它满足两点：一是它较前面提出的标准应该是一个“好的”估计量，二是它的分布形式应该已知，且只依赖未知参数q.

（2）所求的区间考虑为$\hat{q}$的一个邻域$[\hat{q}-a,\hat{q}+b]$，$a,b>0$，使得对于$0<a<1$，

$$P\{\hat{q}-a\leqslant q\leqslant\hat{q}+b\}=1-a\,,\tag{6.3.1}$$

且一般要求$a+b$尽可能小．为确定a、b，须用解不等式的方法将（6.3.1）式中的随机事件变成类似于下述等价形式：

$$\{\hat{q}-a\leqslant q\leqslant\hat{q}+b\}=\{-g(a)\leqslant T(X_1,X_2,\mathbf{L}\,X_n;q)\leqslant g(b)\}.\tag{6.3.2}$$

其中，$g(x)$为可逆的x的已知函数，$T=T(X_1,X_2,\mathbf{L},X_n;q)$的分布与$q$无关且已知，一般其分位点应有表可查，这是关键的一步．于是就可得出$g(a)$、$g(b)$为某个分位点，如$g(a)=c$、$g(b)=d$.

（3）从$g(a)$和$g(b)$的表达式中解出a、b即可.

区间估计涉及到抽样分布，对于一般分布的总体，其抽样分布的计算通常有些困难，因此，我们将主要研究正态总体参数的区间估计问题.

6.4 正态总体均值与方差的区间估计

设$X_1,X_2,\mathbf{L},X_n$为$N(m,s^2)$的样本，对给定的置信水平$1-a$（$0<a<1$），分别研究参数m与s^2的区间估计.

一、正态总体均值m的区间估计

在上述前提下，求m的置信水平为$1-a$的区间估计．考虑m的点估计为$\overline{X}$和$\overline{X}\sim N\left(m,\dfrac{s^2}{n}\right)$，确定$a>0$、$b>0$使

$$P(A)=P\{\overline{X}-a\leqslant m\leqslant\overline{X}+b\}=1-a\,,$$

且使区间长$a+b$尽可能小．下面分两种情况.

1. s^2已知

变换事件A，使A表成式（6.3.2）的形式：

$$A=\left\{\frac{-a}{s/\sqrt{n}}\leqslant\frac{m-\overline{X}}{s/\sqrt{n}}\leqslant\frac{b}{s/\sqrt{n}}\right\}$$

这里$U(m)=\dfrac{m-\overline{X}}{s/\sqrt{n}}\sim N(0,1)$，为使$P(A)=1-a$，又要尽量使$a+b$最小，亦即使$\dfrac{a+b}{s/\sqrt{n}}$最小，如图 6-1 所示，从$N(0,1)$密度函数的特点来看（对称、原点附近密度最大，往两边密度减小），只有取$\dfrac{a}{s/\sqrt{n}}=\dfrac{b}{s/\sqrt{n}}=m_{\frac{a}{2}}$，即$a=b=m_{\frac{a}{2}}s/\sqrt{n}$，从而所求的区间是

$$\left[\overline{X}-\frac{u_{\frac{a}{2}}s}{\sqrt{n}},\overline{X}+\frac{u_{\frac{a}{2}}s}{\sqrt{n}}\right].\qquad（6.4.1）$$

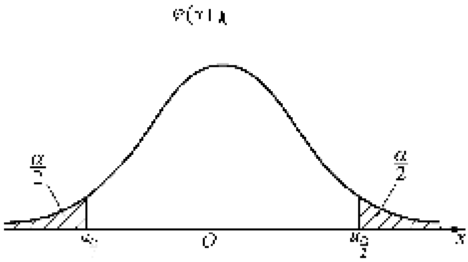

图 6-1

例 6.4.1　从一批零件中，抽取 9 个零件，测得其直径（单位：mm）为：

19.7　20.1　19.8　19.9　20.2　20.3

设零件直径服从正态分布$N(m,s^2)$，且已知s =0.21mm，求这批零件直径的均值m的置信水平为 0.95 及 0.99 的置信区间.

解　有$\overline{x}\approx 20.01$mm.

如果置信水平$1-a$ =0.95，则a =0.05，查正态分布表得$u_{\frac{a}{2}}=u_{0.025}=1.96$. 由此得

$$\frac{s_0}{\sqrt{n}}u_{\frac{a}{2}}=\frac{0.21}{\sqrt{9}}\times 1.96\approx 0.14\text{ mm},$$

所以，置信区间为

$$(19.87,20.15)\text{ mm}.$$

如果置信水平$1-a$ =0.99，则a =0.01，查附表 2 得$u_{\frac{a}{2}}=u_{0.005}=t_{0.005}(\infty)=2.58$. 由此得

$$\frac{s_0}{\sqrt{n}}u_{\frac{a}{2}}=\frac{0.21}{\sqrt{9}}\times 2.58\approx 0.18\ \text{mm},$$

所以，置信区间为

$$(19.83,20.19)\ \text{mm}.$$

从这个例子，可以看到，当样本容量 n 一定时，为了提高区间估计的可靠性，应当取较高的置信水平，但这时求出的置信区间也较大，这样就使估计的精确度降低；如果要提高估计的精确度，则应当缩小置信区间，然而对应的置信水平也将随之降低．由此可见，区间估计与置信水平有密切的关系.

2. s^2 未知

将事件 A 变换成式（6.3.2）的形式

$$A=\left\{\frac{-a}{S/\sqrt{n-1}}\leqslant\frac{m-\overline{X}}{S/\sqrt{n-1}}\leqslant\frac{b}{S/\sqrt{n-1}}\right\}.$$

已知 $T(m)=\dfrac{m-\overline{X}}{S/\sqrt{n-1}}\sim t(n-1)$，为使 $P(A)=1-a$，且区间尽量短，与 $N(0,1)$ 情形一样，只有取 $\dfrac{a}{S/\sqrt{n-1}}=\dfrac{b}{S/\sqrt{n-1}}=t_{\frac{a}{2}}(n-1)$．因此所求区间为

$$\left[\overline{X}-\frac{t_{\frac{a}{2}}(n-1)S}{\sqrt{n-1}},\ \overline{X}+\frac{t_{\frac{a}{2}}(n-1)S}{\sqrt{n-1}}\right].\qquad（6..4.2）$$

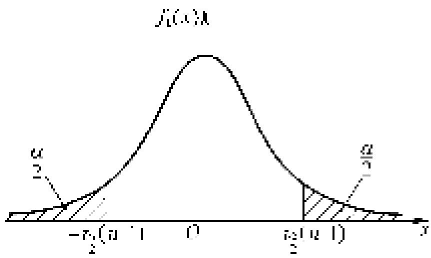

图 6-2

例 6.4.2 在上面例 6.4.1 中，设 s 未知，求零件直径的均值 m 的置信水平为 0.95 的置信区间.

解 有

$$\bar{x}\approx 20.01\ \text{mm},\quad s\approx 0.203\ \text{mm},$$

已给置信水平 $1-a$ =0.95，则 a =0.05；自由度 $k=9-1=8$；查 t 分布表得 $t_{\frac{a}{2}}(1-a)=t_{0.025}(8)=2.31$．由此得

$$\frac{s}{\sqrt{n}}t_{\frac{a}{2}}=\frac{0.203}{\sqrt{9}}\times 2.31\approx 0.16\ \text{mm},$$

所以，置信区间为

$$(19.85,20.17)\ \text{mm}.$$

例 6.4.3 一批零件尺寸服从 $N(\mu,\sigma^2)$，对 μ 进行区间估计(σ^2 未知)，要求估计精度不低于 $2d$，置信水平保持为 $1-\alpha$，问至少要抽取多少件产品作为样本？

解 显然，此处要求

$$P\{\overline{X}-d \leqslant \mu \leqslant \overline{X}+d\}=1-\alpha,$$

由式（6.4.2）知 $d=t_{1-\frac{\alpha}{2}}(n-1)S/\sqrt{n-1}$，

故

$$n=\left(\frac{t_{1-\frac{\alpha}{2}}(n-1)S}{d}\right)^2+1,$$

上式不是 n 的显式，但对于具体数值，可采取“试算法”来确定 n．一般是先对 S^2 作个大致估计（可以由以往的经验确定），然后用试算的方式确定适合上面方程的 n．例如若估计出 $S^2\approx 200$，又已知 $d=10, \alpha=0.05$，来试算 n：

取 $n=10$, $t_{0.975}(9)=2.2622$, $t_{0.975}^2(9)\times 2+1\approx 11.24$，

取 $n=11$，$t_{0.975}(10)=2.2281$, $t_{0.975}^2(10)\times 2+1\approx 10.93$，

显然，如果任正整数不可能严格满足上面方程的话，则应取使方程左边大于右边的最小的 n，因此应该取 $n=11$．

二、方差 σ^2 的区间估计

1. $\mu=\mu_0$ 已知

样本函数

$$\chi^2=\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\mu_0)^2 \sim \chi^2(n).$$

χ^2 分布的分布曲线是不对称的，对于已给的置信水平 $1-\alpha$，要想找到最短的置信区间是困难的．因此，仿照上述分布曲线为对称的情形，选取这样的的区间 $(\chi_{1-\frac{\alpha}{2}}^{\ 2}(n), \chi_{\frac{\alpha}{2}}^{\ 2}(n))$，使得

$$P\left(\chi^2 \geqslant \chi_{1-\frac{\alpha}{2}}^{\ 2}(n)\right)=1-\frac{\alpha}{2},\quad P\left(\chi^2 \geqslant \chi_{\frac{\alpha}{2}}^{\ 2}(n)\right)=\frac{\alpha}{2},$$

是合理的如图 6-3 所示．于是，有

$$P\left(\chi_{1-\frac{\alpha}{2}}^{\ 2}(n)<\chi^2<\chi_{\frac{\alpha}{2}}^{\ 2}(n)\right)=1-\alpha,$$

$$P\left(c_{1-\frac{a}{2}}{}^2(n) < \frac{1}{s^2}\sum_{i=1}^{n}(X_i - m_0)^2 < c_{\frac{a}{2}}{}^2(n) \right) = 1-a \text{ ，}$$

即

$$P\left(\frac{\sum_{i=1}^{n}(X_i - m_0)^2}{c_{\frac{a}{2}}{}^2(n)} < s^2 < \frac{\sum_{i=1}^{n}(X_i - m_0)^2}{c^2_{1-\frac{a}{2}}(n)} \right) = 1-a \text{ .}$$

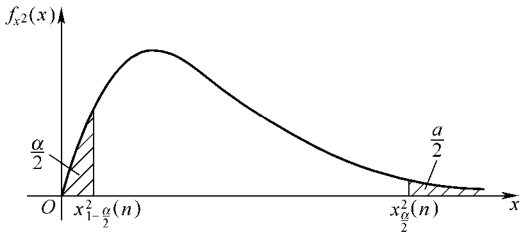

图 6-3

上式表明，总体方差s^2的置信水平为$1-a$的置信区间是

$$\left(\frac{\sum_{i=1}^{n}(X_i - m_0)^2}{c_{\frac{a}{2}}{}^2(n)}, \ \frac{\sum_{i=1}^{n}(X_i - m_0)^2}{c^2_{1-\frac{a}{2}}(n)} \right).$$

由此得到总体标准差s的置信水平为$1-a$的置信区间是

$$\left(\sqrt{\frac{\sum_{i=1}^{n}(X_i - m_0)^2}{c_{\frac{a}{2}}{}^2(n)}}, \ \sqrt{\frac{\sum_{i=1}^{n}(X_i - m_0)^2}{c^2_{1-\frac{a}{2}}(n)}} \right).$$

2. m 未知

样本函数

$$c^2 = \frac{(n-1)S^2}{s^2} \sim c^2(n-1) \text{，}$$

与上面类似，对于已给的置信水平$1-a$，有

$$P\left(c_{1-\frac{a}{2}}{}^2(n-1) < \frac{(n-1)S^2}{s^2} < c_{\frac{a}{2}}{}^2(n-1) \right) = 1-a \text{ ，}$$

即

$$P\left(\frac{(n-1)S^2}{c_{\frac{a}{2}}^{\ 2}(n-1)} < s^2 < \frac{(n-1)S^2}{c_{1-\frac{a}{2}}^{\ 2}(n-1)}\right) = 1-a\ .$$

上式表明，总体方差s^2的置信水平为$1-a$ 的置信区间是

$$\left(\frac{(n-1)S^2}{c_{\frac{a}{2}}^{\ 2}(n-1)},\quad \frac{(n-1)S^2}{c_{1-\frac{a}{2}}^{\ 2}(n-1)}\right).$$

由此得到总体标准差s 的置信水平为$1-a$ 的置信区间是

$$\left(\sqrt{\frac{(n-1)S^2}{c_{\frac{a}{2}}^{\ 2}(n-1)}}, \sqrt{\frac{(n-1)S^2}{c_{1-\frac{a}{2}}^{\ 2}(n-1)}}\right).$$

例 6.4.4 在上面例 6.4.1 中，求零件直径的方差s^2的置信水平为 0.95 的置信区间，如果：

（1）已知零件直径的均值$m = 20$ mm；

（2）m 未知.

解 （1）已知$m = 20$，不难计算

$$\sum_{i=1}^{9}(x_i - 20)^2 = 0.33\,\text{mm}^2.$$

已给置信水平$1-a$ =0.95，则a =0.05；自由度k =9；查c^2分布表，

$$c_{1-\frac{a}{2}}^{\ 2}(n) = c_{0.975}^{\ 2}(9) = 2.70,\qquad c_{\frac{a}{2}}^{\ 2}(n) = c_{0.025}^{\ 2}(9) = 19.0.$$

所以，所求置信区间为

$$\left(\frac{0.33}{19.0}, \frac{0.33}{2.70}\right),$$

即

$$(0.0174, 0.1222)\quad \text{mm}^2.$$

（2）m 未知，有

$$s^2 \approx 0.0411\,\text{mm}^2.$$

已给置信水平$1-a$ =0.95，则a =0.05；自由度k =9；查附表得

$$c_{1-\frac{a}{2}}^{\ 2}(n-1) = c_{0.975}^{\ 2}(8) = 2.18,\quad c_{\frac{a}{2}}^{\ 2}(n-1) = c_{0.025}^{\ 2}(8) = 17.5.$$

所以，所求置信区间为

$$\left(\frac{8\times 0.0411}{17.5},\frac{8\times 0.0411}{2.18}\right),$$

即

$$(0.0188,\ 0.1508)\quad \mathrm{mm}^2.$$

6.5　两个正态总体均值差及方差比的置信区间

现在讨论关于两个正态总体的均值差及方差比的区间估计问题.

设总体 X 服从正态分布 $N(m_1,s_1^2)$，从其中抽取容量为 n_1 的样本，得到样本观测值为 $x_1,x_2,\mathbf{L},x_{n_1}$；又设总体 Y 服从正态分布 $N(m_2,s_2^2)$，从其中抽取容量为 n_2 的样本，得到样本的观测值为 $y_1,y_2,\mathbf{L},y_{n_2}$；求两个总体的均值差 m_1-m_2 及方差比 $\dfrac{s_1^2}{s_2^2}$ 的置信水平为 $1-a$ 的置信区间.

一、两个正态总体均值差的区间估计

1.　s_1^2 及 s_2^2 已知

当 s_1^2 及 s_2^2 已知时，有

$$\overline{X}\sim N\left(m_1,\frac{s_1^2}{n_1}\right),\quad \overline{Y}\sim N\left(m_2,\frac{s_2^2}{n_2}\right),\quad \overline{X}-\overline{Y}\sim N\left(m_1-m_2,\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right),$$

得

$$\frac{(\overline{X}-\overline{Y})-(m_1-m_2)}{\sqrt{\dfrac{s_1^2}{n_1}+\dfrac{s_2^2}{n_2}}}\sim N(0,1),$$

对于已知置信水平 $1-a$，则有

$$P\left(\frac{|(\overline{X}-\overline{Y})-(m_1-m_2)|}{\sqrt{\dfrac{s_1^2}{n_1}+\dfrac{s_2^2}{n_2}}}<u_{\frac{a}{2}}\right)=1-a,$$

即

$$P\left(|(\overline{X}-\overline{Y})-(m_1-m_2)|<\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}\,u_{\frac{a}{2}}\right)=1-a,$$

所以两个总体均值差 m_1-m_2 的 $1-a$ 置信区间为

$$\left(\overline{X}-\overline{Y}-\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}\,u_{\frac{a}{2}},\ \overline{X}-\overline{Y}+\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}\,u_{\frac{a}{2}}\right).\qquad（6.5.1）$$

2. s_1^2 及 s_2^2 未知，假定 $s_1^2=s_2^2$

样本函数

$$T=\frac{(\overline{X}-\overline{Y})-(m_1-m_2)}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\sim t(n_1+n_2-2)\,,$$

其中

$$S_w=\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}\,,$$

对于已知的置信水平 $1-a$ ，有

$$P\left(\frac{|(\overline{X}-\overline{Y})-(m_1-m_2)|}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}<t_{\frac{a}{2}}(n_1+n_2-2)\right)=1-a\,,$$

即

$$P\left(|(\overline{X}-\overline{Y})-(m_1-m_2)|<\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\cdot S_w\cdot t_{\frac{a}{2}}(n_1+n_2-2)\right)=1-a\,,$$

故可得两个总体均值差 m_1-m_2 的置信水平为 $1-a$ 的置信区间为

$$\left(\overline{X}-\overline{Y}-\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\cdot S_w\cdot t_{\frac{a}{2}}(n_1+n_2-2),\overline{X}-\overline{Y}+\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\cdot S_w\cdot t_{\frac{a}{2}}(n_1+n_2-2)\right).\quad(6.5.2)$$

例 6.5.1 为了估计磷肥对某种农作物的增产作用，分别各选 10 块土地，分别做施肥和不施肥的试验，设施肥的亩产量 $X\sim N(m_1,s_1^2)$ ，不施肥的亩产量 $Y\sim N(m_2,s_2^2)$. 测得如下数据：$n_1=n_2=10$，$\overline{x}=600$，$\overline{y}=540$ ，$\sum_{i=1}^{10}(x_i-\overline{x})^2=6400$ ，$\sum_{i=1}^{10}(y_i-\overline{y})^2=2400$ ，取置信水平为 95%，求施肥和不施肥的平均亩产之差 m_1-m_2 的置信区间.

解 由题设知 $\overline{x}=600$，$\overline{y}=540$ ，并求得

$$(n_1-1)S_1^2=\sum_{i=1}^{n_1}(x_i-\overline{x})^2=6400\,,\quad (n_2-1)S_2^2=\sum_{i=1}^{n_2}(y_i-\overline{y})^2=2400\,,$$

$$S_w=\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}=\sqrt{\frac{6400+2400}{10+10-2}}=22.11\,,$$

由 $a=0.05$ ，查表得 $t_{\frac{a}{2}}(18)=2.10$. 故 m_1-m_2 的置信区间为

$$\left(\bar{x}-\bar{y}-\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\cdot S_w\cdot t_{\frac{a}{2}},\ \bar{x}-\bar{y}+\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\cdot S_w\cdot t_{\frac{a}{2}}\right)$$

$$=\left(600-540-\sqrt{\frac{1}{10}+\frac{1}{10}}\cdot 22.11\times 2.10, 600-540+\sqrt{\frac{1}{10}+\frac{1}{10}}\cdot 22.11\times 2.10\right)$$

$$=(60-20.7646,60+20.7646)$$

$$=(39.2354,80.7646)$$

因为置信区间的下限 39.235470，所以在置信水平为$1-a=0.95$下，认为施肥的平均产量m_1大于不施肥的平均亩产量m_2.

二、两个正态总体方差比的区间估计

当m_1和m_2都未知时，有

$$\frac{(n_1-1)S_1^2}{s_1^{\ 2}}\sim c^2(n_1-1)，\quad \frac{(n_2-1)S_2^2}{s_2^{\ 2}}\sim c^2(n_2-1)，$$

因此

$$\frac{S_1^2/s_1^{\ 2}}{S_2^2/s_2^{\ 2}}\sim F(n_1-1,n_2-1).$$

对于置信水平$1-a$，如图 6-4 所示，有

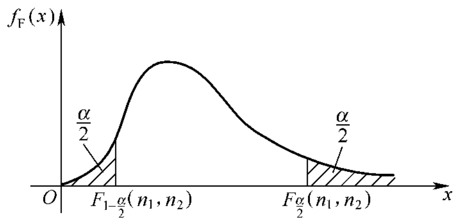

图 6-4

$$P\left(F_{1-\frac{a}{2}}(n_1-1,n_2-1)<\frac{S_1^2/S_2^2}{s_1^{\ 2}/s_2^{\ 2}}<F_{\frac{a}{2}}(n_1-1,n_2-1)\right)=1-a\ ，$$

即

$$P\left(\frac{S_1^2}{S_2^2}\cdot\frac{1}{F_{a/2}}<\frac{s_1^{\ 2}}{s_2^{\ 2}}<\frac{S_1^2}{S_2^2}\cdot\frac{1}{F_{1-a/2}}\right)=1-a\ .$$

所以，两个总体方差比$\frac{s_1^{\ 2}}{s_2^{\ 2}}$的置信水平为$1-a$的置信区间为

$$\left(\frac{S_1^2}{S_2^2}\cdot\frac{1}{F_{a/2}(n_1-1,n_2-1)},\ \frac{S_1^2}{S_2^2}\cdot\frac{1}{F_{1-a/2}(n_1-1,n_2-1)}\right). \qquad (6.5.3)$$

例 6.5.2 已知两个正态总体 $X \sim N(m_1, s_1^2)$，$Y \sim N(m_2, s_2^2)$，其中 m_1 和 m_2 未知，分别测得有关数据为 $n_1 = 4, S_1^2 = 4.8$，$n_2 = 5, S_2^2 = 4.0$．试求方差比 $\dfrac{s_1^2}{s_2^2}$ 的置信水平为 90%的置信区间.

解 由题设知

$$\frac{S_1^2}{S_2^2} = \frac{4.8}{4.0} = 1.2，\quad F_{\frac{a}{2}}(n_1 - 1, n_2 - 1) = F_{0.05}(3,4) = 6.59$$

$$F_{1-\frac{a}{2}}(n_1 - 1, n_2 - 1) = F_{0.95}(3,4) = \frac{1}{F_{0.05}(4,3)} = \frac{1}{9.12}$$

由（6.5.3）式知，所求 $\dfrac{s_1^2}{s_2^2}$ 的置信区间为

$$\left(1.2 \times \frac{1}{6.59},\ 1.2 \times 9.12\right) = (0.182, 10.944).$$

因为置信区间下限 0.185<1，而上限 10.944>1，所以在置信水平 90%下，可以认为两个总体的方差没有明显差异.

6.6 单侧置信区间

前面所讨论的置信区间都是双侧的，但在实际问题中，有时只需要讨论单侧置信上限或单侧置信下限就可以了．例如，对于家用电器的使用寿命来说，当然希望使用寿命越长越好，我们关心的是一批家用电器的平均寿命 m 的下限；又如，对于产品的次品率来说，当然希望次品率越低越好，我们关心的是一批产品的次品率的上限.

定义 6.6.1 设总体 X 的分布中含有未知参数 q，对于给定的置信水平 $1-a$，若存在统计量 $\hat{q}_{下}$，使得

$$P(q > \hat{q}_{下}) = 1-a，$$

则称 $\hat{q}_{下}$ 为 q 的置信水平为 $1-a$ 的单侧置信下限.

类似地，若存在统计量 $\hat{q}_{上}$，使得

$$P(q < \hat{q}_{上}) = 1-a，$$

则称 $\hat{q}_{上}$ 为 q 的置信水平为 $1-a$ 的单侧置信上限.

设总体 $X \sim N(m, s^2)$，其中 m、s^2 都未知，由已学过的定理知

$$\frac{\overline{X} - m}{S/\sqrt{n}} \sim t(n-1)，\quad \frac{(n-1)S^2}{s^2} \sim c^2(n-1)，$$

故有

$$P\left(\frac{\overline{X}-m}{S/\sqrt{n}}<t_a(n-1)\right)=1-a\ ,$$

即

$$P\left(m>\overline{X}-\frac{S}{\sqrt{n}}t_a(n-1)\right)=1-a\ ,$$

所以，m 的置信水平为$1-a$ 的单侧置信下限为

$$\hat{m}_{下}=\overline{X}-\frac{S}{\sqrt{n}}t_a(n-1)\ . \qquad (6.6.1)$$

同理知，有

$$P\left(\frac{(n-1)S^2}{s^2}>c^2_{1-a}(n-1)\right)=1-a\ ,$$

即

$$P\left(s^2<\frac{(n-1)S^2}{c^2_{1-a}(n-1)}\right)=1-a\ ,$$

所以得s^2 的置信水平为$1-a$ 的单侧置信上限为

$$\hat{s}^2_{上}=\left(\frac{(n-1)S^2}{c^2_{1-a}(n-1)}\right). \qquad (6.6.2)$$

例 6.6.1 已知电子元件的寿命X(小时)服从正态分布$N(m,s^2)$,其中m 和s^2 都未知，随机抽取 6 个元件测试，得有关数据$\overline{X}=4563.2,\ S^2=1024$．已给置信水平为 0.95，试分别求$m$ 的单侧置信下限和s^2 的单侧置信上限.

解 由题设知$\overline{X}=4563.2,\ S=32$．由（6.6.1）式知$m$ 的置信水平为$1-a$ 的单侧置信下限为

$$\hat{m}_{下}=\overline{X}-\frac{S}{\sqrt{n}}t_a(n-1)\ ,$$

其中$t_a(n-1)=t_{0.05}(5)=2.02$ ，代入得

$$\hat{m}_{下}=4563.2-\frac{32}{\sqrt{6}}\times 2.02=4563.2-26.39=4536.81\ ;$$

查表得

$$c^2_{1-a}(n-1)=c^2_{0.95}(5)=1.145\ ,$$

由（6.6.2）式知s^2 的置信水平为$1-a$ 的单侧置信上限为

$$\hat{s}^2_{上}=\left(\frac{(n-1)S^2}{c^2_{1-a}(n-1)}\right)=\frac{5\times 1024}{1.145}=4471.62\ .$$

概率故事

社会舆论调查

即使我下定决心，我仍充满了犹豫.

——奥斯卡 • 列文托（Oscar Levant）

过去，当权者们利用侦探系统来查明公众的观点．或许，由此所收集的信息帮助他们形成公众政策，制定和实施法律．现代的社会舆论调查的历史，是从盖洛普民意调查的第一个报告开始的．今天，社会舆论调查在报纸和其他新闻媒介中已经扮演了一个非常重要的角色．他们收集公众对各种社会、政治和经济问题上的信息，出版摘要报告．这样的舆论调查在民主政治社会中能起到积极的作用．他们可以告诉政治领导人和官僚们什么是公众的需要，什么是公众的爱好．他们也向公众报告新闻来通告公众的想法，或许可在某个重要的问题上帮助公众明确表现他们的观点.

通常以某种特定的统计形式宣布公众舆论调查的结果同时需要一定的解释．例如，播音员说:“赞成总统外交政策的人占 42%，正负误差界限为 4%.”代替给出单个数字，这里播音员给出一个区间(42-4,42+4) = (38,46)．这是如何得到的？如何解释呢？

假设所有美国成人中，实际赞成总统外交政策的比率为数值 T. 为了了解 T 的大小，必须接触每一个美国成人，得到他们对“你赞成总统的外交政策吗？”这样问题的反应．如果必须要得到一个限时的、迅速的答案，这是不可能的．最好的方法是求出一个最接近于 T 的估计值．新闻媒介对某一数量的“任意选择的个体”进行电话采访，得到他们的答案．如果接触了数量为 n 的个体，其中有 m 个人回答“赞成”，则 T 的估计值可为 $100 \times (m/n)$. 当然，这样的估计是存在一定的误差的，因为所取的仅仅是某个集合中的样本（美国成人中很小的一部分）．如果接触另外的 n 个人，可能得到不同的估计值．如何求出估计值的误差呢？基于两个统计学家内曼和阿 • 皮尔森发展起来的一个理论，可以根据算出一个数字 e，使得 T 的真实值有很高的概率，一般为 95%（或 99%），落于区间 $(100 \times (m/n)-e, 100 \times (m/n)+e)$ 之内．也就是说，这个区间不包含真实值的事件，等价于在装有 5 个（或 1 个）白球，95 个（或 99 个）黑球的口袋中随机地抽取一球，抽得白球这样一个几乎很少发生的事件.

社会舆论调查的有效性，基于所选择个体的“代表性”．十分明显的是，调查的结果是依赖于所选择个体所属的政治团体的（民主党或共和党）．即便假设所选择的个人的政治所属是没有偏差的，如果有些个体不回答问题，有些又恰恰属于某些特别的政治团体，则结果也会不同．任何调查中，都有不同程度的不回答者，

这种场合要评价误差是困难的，除非有更多的可利用的信息.

习题六

1．设 $X_1,X_2,\mathbf{L},X_n$ 是取自总体 X 的一个样本，在下列情形下，试求总体参数的矩估计与最大似然估计：

（1） $X\sim B(1,p)$，其中 p 未知， $0<p<1$；

（2） $X\sim E(l)$，其中 l 未知， $l>0$.

2．设总体 X 服从“0-1”分布：

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

如果取得样本观测值为 $x_1,x_2,\mathbf{L},x_n$（ $x_i=0$或1），求参数 p 的矩估计值与最大似然估计值.

3．设总体 X 服从几何分布：

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

如果取得样本观测值为 $x_1,x_2,\mathbf{L},x_n$，求参数 p 的矩估计值与最大似然估计值.

4．设 $X_1,X_2,\mathbf{L},X_n$ 是取自总体 X 的一个样本，其中 X 服从参数为 l 的泊松分布，其中 l 未知， $l>0$，求 l 的矩估计值与最大似然估计值．如果得到一组样本观测值如表 6-1 所示.

表 6-1

X	0	1	2	3	4
频数	17	20	10	2	1

求参数 l 的矩估计值与最大似然估计值.

5．已知某种灯泡的寿命 $X\sim N(m,s^2)$，但 m 和 s^2 未知，今随机抽取 5 只灯泡测得寿命（单位：小时）分别为

1623　1527　1287　1432　1591

求 m 和 s^2 的估计值.

6．设总体在区间 $[a,b]$ 上服从均匀分布，但 a、b 未知，现抽取样本 X_1,X_2,X_3,X_4,X_5，测得一组观测值（1,3,0,4,-2），试用矩估计法估计 a、b .

7．设总体 X 的密度函数为 $f(x,l)=\begin{cases} l\,\mathrm{e}^{-lx}, & (x>0),\\ 0, & (x\leqslant 0).\end{cases}$，其中参数 l 未知而待定，抽样得样本观测值 $x_1,x_2,\mathbf{L},x_{n_1}$，求参数 l 的极大似然估计.

8，若 $X\sim N(m,s^2)$，其中 m 和 s^2 均未知，设样本一组观测值为 $x_1,x_2,\mathbf{L},x_{n_1}$，用极大似然估计法估计 m 和 s^2 .

9．已知某种电子元件的使用寿命（从开始使用到初次失效为止）服从指数分

布 $f(x)=\lambda e^{-\lambda x}$ $(x>0,\lambda>0)$，今随机抽取 250 个元件，测得寿命数据如表 6-2 所示（单位：小时）.

表 6-2

寿命时间（小时）	元件数（个）
0～100	39
100～200	58
200～300	47
300～400	33
400～500	25
500～600	22
600～700	11
700～800	6
800～900	7
900～1000	2
合计	250

试采用极大似然估计法估计该指数分布中的参数 λ .

10．设总体 X 服从二项分布 $B(n,p)$， n 为正整数， $0<p<1$，其中 n,p 均为未知参数， $x_1,x_2,\cdots,x_m$ 是从 X 中抽取的一个样本，试分别求 n,p 的矩估计.

11．设 x_1,x_2,x_3 是正态总体 $N(\mu,\sigma^2)$ 的一个样本，其中 $\mu=0$ 并 σ^2 未知，求 x_1,x_2,x_3 的似然函数，并对 x_1,x_2,x_3 的一个样本 2.1、2.2、2.0 估计 σ^2 .

12．设 $x_1,x_2,\cdots,x_{n_1}$ 来自指数分布 $f(x,\lambda)=\begin{cases}\dfrac{1}{\theta}e^{-\frac{x}{\theta}}, & (x\geqslant 0)\\ 0, & (x<0)\end{cases}$ 的样本，试分别用矩估计法和极大似然法求 θ 的估计量.

13．证明：如果已知总体 X 的均值 μ，则总体方差的无偏估计量为

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

其中 $X_1,X_2,\cdots,X_n$ 是从总体 X 中抽取的样本.

14．设正态总体 $N(\mu,\sigma^2)$ 中 μ 未知而 σ^2 已知，又设 x_1、x_2 是来自正态总体的一个样本，问下列样本函数中哪个是统计量？在统计量中，哪些是 μ 的无偏估计？哪个是最佳无偏估计？

（1） $\mu_1=\dfrac{1}{3}x_1+\dfrac{2}{3}x_2$；　　（2） $\mu_2=\dfrac{1}{3}(x_1+\mu)$；

（3）$m_3=\frac{1}{2}x_1+\frac{1}{2}x_2$；　　　　（4）$m_4=\sum_{i=1}^{2}\frac{x_i}{s}$.

15．设 X_1,X_2,X_3 为总体 X 的样本，证明

$$\hat{m}_1=\frac{1}{6}X_1+\frac{1}{3}X_2+\frac{1}{2}X_3,$$

$$\hat{m}_2=\frac{2}{5}X_1+\frac{1}{5}X_2+\frac{2}{5}X_3,$$

都是总体均值 m 的无偏估计，并进一步判断哪一个估计更有效.

16．设总体 $X\sim N(m,0.09)$，测得一组样本的观测值为 12.6，13.4，12.8，13.2，求 m 的置信度为 0.95 的置信区间.

17．对某型号飞机的飞行速度进行了 15 次试验，测得最大飞行速度（米/秒）为

422.2	417.2	425.6	420.3	425.8
423.1	418.7	428.2	438.3	434.0
412.3	431.5	413.5	441.3	423.0

根据长期经验，最大飞行速度可以认为是服从正态分布的，试利用上述数据对最大飞行速度的期望值进行区间估计（置信度 0.95）.

18．对某种材料的强度只有下限的要求，已知该材料的强度 $X\sim N(m,s^2)$，但 m、s^2 均未知，今进行 5 次测试，得样本均值和样本均方差分别为 $\overline{X}$ ＝1160kg/cm^2，s ＝99.75kg/cm^2，现求 m 的 0.99 单侧置信区间 $(\hat{q},+\infty)$.

19．正态总体 $X\sim N(m,s^2)$ 中抽取一组样本容量为 $n=25$ 的样本，计算得样本均值 $\bar{x}=79$，样本方差为 $s^2=112.36$，试求（1）已知 $s^2=10^2$；（2）s 未知两种情况分别求总体均值 m 的置信度为 0.95 的置信区间.

20．为确定某种液体的浓度，取 4 个独立的测定值，其平均值 $\bar{x}=8.38\%$，样本标准差 $s=0.03\%$，设被测总体近似地服从正态分布 $N(m,s^2)$，求总体均值 m 的置信度为 95%的置信区间.

21．从一批钉子中随机地抽取 16 枚，测得其长度（单位：cm）为

2.14	2.10	2.13	2.15	2.13	2.12	2.13	2.10
2.15	2.12	2.14	2.10	2.13	2.11	2.14	2.11

设钉长服从正态分布 $N(m,s^2)$，试求（1）已知 s =0.1（cm）;（2）s 未知，两种情况分别求总体均值 m 的 90%的置信区间.

22．某食品加工厂有甲乙两条加工猪肉罐头的生产线．设罐头质量服从正态分布并假设甲生产线与乙生产线互不影响．从甲生产线抽取 10 只罐头测得其平均质量 $\bar{x}=501$g，已知其总体标准差 $s_1=5$g；从乙生产线抽取 20 只罐头测得其平均质量 $\bar{y}=498$g，已知其总体标准差 $s_2=4$g，求甲乙两条猪肉罐头生产线生产罐头

质量的均值差的$m_1 - m_2$的双侧 0.99 置信区间.

23．为了比较甲乙两种显像管的使用寿命X和Y，随机地抽取甲乙两种显像管各 10 只，测得数据$x_1, x_2, \mathbf{L}, x_{10}$和$y_1, y_2, \mathbf{L}, y_{10}$（单位：$10^4 h$）且由此算得

$$\overline{x} = 2.33, \quad \overline{y} = 0.75, \quad \sum_{i=1}^{10}(x_i - \overline{x})^2 = 27.5, \quad \sum_{i=1}^{10}(y_i - \overline{y})^2 = 19.2 .$$

假定两种显像管的使用寿命均服从正态分布，且由生产过程知道它们的方差相等，试求两个总体均值之差$m_1 - m_2$的双侧 0.95 置信区间.

24．从汽车轮胎厂生产的某种轮胎中抽取 10 个样品进行磨损试验，直至轮胎行驶到磨坏为止，测得它们的行驶路程（km）如下：

41 250	41 010	42 650	38 970	40 200
42 550	43 500	40 400	41 870	39 800

设汽车轮胎行驶路程服从正态分布$N(m, s^2)$，求：

（1）m 的置信水平为 0.95 的单侧置信下限；

（2）s 的置信水平为 0.95 的单侧置信上限.