第 3 章　参数估计

参数估计(Parameter Estimation)是大致判断出总体分布的类型后,根据从总体中抽取的样本参数(均值或方差)推断总体分布中包含的未知相应参数的方法。它是统计推断的一种基本形式,是数理统计学的一个重要分支,分为点估计和区间估计两部分。

3.1　点　估　计

点估计依据样本来估计总体分布中所含的未知参数或未知参数的函数。通常它们是总体的某个特征值,如数学期望、方差和相关系数等。点估计问题就是要构造一个只依赖于样本的量,作为未知参数或未知参数的函数的估计值。例如,设一批产品的废品率为 θ。为了估计 θ,从这批产品中随机抽出 n 个做检查,以 m 标记其中的废品个数,用$\frac{m}{n}$估计 θ,就是一个点估计。构造点估计常用的方法有以下几种。

① 矩估计法。用样本矩估计总体矩,如用样本均值估计总体均值。

② 最大似然估计法。1912 年由英国统计学家 R. A. 费希尔提出,利用样本分布密度构造似然函数来求出参数的最大似然估计。

③ 最小二乘法。主要用于线性统计模型中的参数估计问题。

④ 贝叶斯估计法。基于贝叶斯学派的观点而提出的估计法。

由于可以用来估计未知参数的估计量很多,于是产生了怎样选择一个优良估计量的问题。首先必须对优良性定出准则,这种准则不是唯一的,可以根据实际问题和理论研究的方便进行选择。优良性准则有两大类:一类是小样本准则,即在样本大小固定时的优良性准则;另一类是大样本准则,即在样本大小趋于无穷时的优良性准则。最重要的小样本优良性准则是无偏性及与此相关的一致最小方差无偏估计,其次有容许性准则、最小化最大准则、最优同变准则等。大样本优良性准则有相合性、最优渐近正态估计和渐近有效估计等。

进行点估计时,由于不同样本算得的 μ 的估计值是不同的。因此,还希望根据所给的样本确定一个随机区间,使其包含参数真值的概率达到指定的要求,这就需要做区间估计。

3.2　区 间 估 计

区间估计是依据抽取的样本,根据一定的正确度与精确度的要求,构造出适当的区间,作为总体分布的未知参数或参数的函数的真值所在范围的估计。例如人们常说的有百分之

多少的把握保证某值在某个范围内，即是区间估计的最简单的应用。1934 年统计学家 J. 奈曼创立了一种严格的区间估计理论。求置信区间常用的有利用已知的抽样分布求置信区间、利用区间估计与假设检验的联系求置信区间、利用大样本理论求置信区间三种方法。

利用已知的抽样分布求置信区间可以在重复抽样（Sampling with Replication）条件下进行，也可以在不重复抽样（Sampling without Repeating）条件下进行。重复抽样又称放回式抽样，即每次从总体中抽取的样本单位，经检验之后又重新放回总体，参加下次抽样，这种抽样的特点是总体中每个样本单位被抽中的概率是相等的。不重复抽样亦称不放回式抽样，即每次从总体中抽取的样本单位，经检验之后不再放回总体，在下次抽样时不会再次抽到前面已抽中过的样本单位，总体每经一次抽样，其样本单位数就减少一个，因此每个样本单位在各次抽样中被抽中的概率是不同的。该种抽样过程为：从总体 N 个单位中要抽取一个容量为 n 的样本，每次从总体中抽取一个单位，连续进行 n 次抽选，构成一个样本，但每次抽选一个单位就不再放回。不重置抽样的样本是由 n 次连续抽选的结果组成，实质上等于一次同时从总体中抽 n 个单位组成一个样本。连续 n 次抽选的结果不是相互独立的，第一次抽选的结果影响下一次抽样，每抽一次，总体的单位数就少一个。因此，每个单位的选中或下次选中的机会是不同的。由于不重复抽样，每个样本单位最多只有一次被抽中的机会，而且随着抽中单位的不断增多，剩下的单位被抽中的机会不断增大，所以不重复抽样的误差小于重复抽样的误差。

区间估计的主要内容如下所示。

- 区间估计
  - 一个总体
    - 均值的区间估计
      - 重复抽样条件下
        - 方差已知时
        - 方差未知时
      - 不重复抽样条件下
        - 方差已知时
        - 方差未知时
    - 方差的区间估计
    - 比率的区间估计
  - 两个总体
    - 均值差的区间估计
      - 方差已知时
      - 方差未知且相等时
      - 方差未知且任意时
    - 方差比的区间估计

3.2.1 一个总体均值的区间估计

1. 方差已知时的区间估计

1）理论基础

（1）求置信区间时需要的定义。

① 标准正态分布单侧上 α 分位点

若随机变量 X 服从标准正态分布，则满足式子 $P(X>z_\alpha)=\alpha$ 或 $P(X<-z_\alpha)=\alpha$ 的 z_α 或 $-z_\alpha$ 叫标准正态分布的单侧上 α 分位点，如图 3-1 所示。

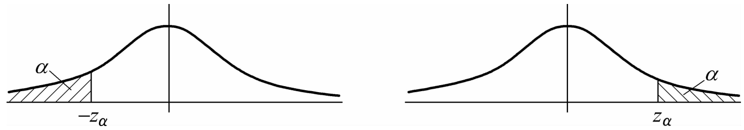

图 3-1 标准正态分布单侧上 α 分位点图解

② 标准正态分布双侧上 α 分位点

若随机变量 X 服从标准正态分布，则满足式子 $P(|X|>z_{\alpha/2})=\alpha$ 的 $z_{\alpha/2}$ 和 $-z_{\alpha/2}$ 叫标准正态分布的双侧上 α 分位点，如图 3-2 所示。

(2) 求置信区间时需要的定理。

若随机变量 $X\sim N(\mu,\sigma^2)$，则有如下定理成立。

① 重复抽样条件下

$$Z=\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$$

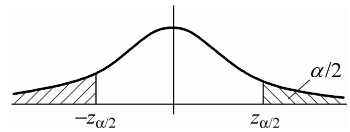

图 3-2 标准正态分布双侧上 α 分位点图解

② 不重复抽样条件下

$$Z=\frac{\bar{x}-\mu}{\sigma/\sqrt{n}\sqrt{\dfrac{N-n}{N-1}}}\sim N(0,1)$$

2) 置信区间的求解

(1) 重复抽样条件下均值的$(1-\alpha)\%$的置信区间。

① 单侧置信区间的求解

因为$\dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}}$服从标准正态分布，所以 $P\left(\dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}}>Z_\alpha\right)=\alpha$ 成立或 $P\left(\dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}}<-Z_\alpha\right)=\alpha$ 成立，即等价于 $P\left(\dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}}\leqslant Z_\alpha\right)=1-\alpha$ 成立或 $P\left(\dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}}\geqslant -Z_\alpha\right)=1-\alpha$ 成立，所以令$\dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}}\leqslant Z_\alpha$或$\dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}}\geqslant -Z_\alpha$ 成立，解这两个不等式，就可以得到重复抽样条件下均值的两个$(1-\alpha)\%$的单侧置信区间。

单侧右置信区间：$\mu\geqslant\bar{x}-z_\alpha\dfrac{\sigma}{\sqrt{n}}$

单侧左置信区间：$\mu\leqslant\bar{x}+z_\alpha\dfrac{\sigma}{\sqrt{n}}$

② 双侧置信区间的求解

因为 $\dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}}$ 服从标准正态分布，所以 $P\left(\left|\dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}}\right|>Z_{\alpha/2}\right)=\alpha$ 成立，即等价于 $P\left(\left|\dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}}\right|\leqslant Z_{\alpha/2}\right)=1-\alpha$ 成立，所以令$\left|\dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}}\right|\leqslant Z_{\alpha/2}$成立，解这个绝对值不等式，就可以得到重复抽样条件下均值的$(1-\alpha)\%$双侧置信区间。

$$\bar{x}-z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\leqslant\mu\leqslant\bar{x}+z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

单双侧置信区间中的$\frac{\sigma}{\sqrt{n}}=\sigma_{\bar{X}}$称为均值的标准误差，又称为抽样平均误差，表示任何一个分布函数的标准差，是原来分布函数标准差的$\sqrt{n}$分之一，或者说$\bar{X}$分布的方差，就是X分布方差的n分之一。

(2) 不重复抽样条件下均值的$(1-\alpha)\%$的置信区间

① 单侧置信区间的求解

因为$\dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}\sqrt{\frac{N-n}{N-1}}}$服从标准正态分布，所以$P\left(\dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}\sqrt{\frac{N-n}{N-1}}}>Z_\alpha\right)=\alpha$成立或$P\left(\dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}\sqrt{\frac{N-n}{N-1}}}<-Z_\alpha\right)=\alpha$成立，即等价于$P\left(\dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}\sqrt{\frac{N-n}{N-1}}}\leqslant Z_\alpha\right)=1-\alpha$成立或$P\left(\dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}\sqrt{\frac{N-n}{N-1}}}\geqslant -Z_\alpha\right)=1-\alpha$成立，所以令$\dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}\sqrt{\frac{N-n}{N-1}}}\leqslant Z_\alpha$或$\dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}\sqrt{\frac{N-n}{N-1}}}\geqslant -Z_\alpha$成立，解这两个不等式，就可以得到重复抽样条件下均值的两个$(1-\alpha)\%$的单侧置信区间。

单侧右区间：$\mu\geqslant\bar{x}-z_\alpha\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$

单侧左区间：$\mu\leqslant\bar{x}+z_\alpha\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$

② 双侧置信区间的求解

因为$\dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}\sqrt{\frac{N-n}{N-1}}}$服从标准正态分布，所以$P\left(\left|\dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}\sqrt{\frac{N-n}{N-1}}}\right|>Z_{\alpha/2}\right)=\alpha$成立，即等价于$P\left(\left|\dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}\sqrt{\frac{N-n}{N-1}}}\right|\leqslant Z_{\alpha/2}\right)=1-\alpha$成立，所以令$\left|\dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}\sqrt{\frac{N-n}{N-1}}}\right|\leqslant Z_{\alpha/2}$成立，解这个绝对值不等式，就可以得到重复抽样条件下均值的$(1-\alpha)\%$双侧置信区间。

$$\bar{x}-z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}\leqslant\mu\leqslant\bar{x}+z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

例 3.1 设某厂生产的灯泡寿命$X\sim N(\mu,100^2)$，现随机抽取5只，测量其寿命如下：1455、1502、1370、1610、1430，试以95%的概率保证程度推断该厂灯泡的平均使用寿命为多少？

解：由题可知，这是一个总体方差已知时均值的双侧区间估计。因为N未知，所以采用重复抽样条件下的置信区间求解。因为重复抽样条件下均值的$(1-\alpha)\%$的双侧置信区间为：

$$\bar{x}-z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\leqslant\mu\leqslant\bar{x}+z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

因为资料未分组，$\sigma=100$，$n=5$

所以$\bar{x}=\dfrac{\sum x}{n}=\dfrac{1}{5}(1455+1502+1370+1610+1430)=1473.4$

又因为 $\alpha=0.05$，　所以查表可知 $z_{\alpha/2}=1.96$

所以$\bar{x}-z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}=1473.4-1.96\times\dfrac{100}{\sqrt{5}}=1385.746$(小时)

$\bar{x}+z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}=1473.4+1.96\times\dfrac{100}{\sqrt{5}}=1561.054$(小时)

所以有 95% 的把握认为该厂灯泡平均使用寿命为(1385.746,1561.054)。

例 3.2　设某厂生产的灯泡寿命 $X\sim N(\mu,100^2)$，现随机抽取 5 只，测量其寿命如下：1455、1502、1370、1610、1430，试以 95%的概率保证程度推断该厂灯泡的平均使用寿命不低于多少？

解：由题可知，这是一个总体方差已知时均值的单侧右区间估计。因为 N 未知，所以采用重复抽样条件下的置信区间求解。因为重复抽样条件下均值的$(1-\alpha)$%的单侧右置信区间为：

$$\mu\geqslant\bar{x}-z_{\alpha}\frac{\sigma}{\sqrt{n}}$$

因为资料未分组，$\sigma=100$，$n=5$

所以$\bar{x}=\dfrac{\sum x}{n}=\dfrac{1}{5}(1455+1502+1370+1610+1430)=1473.4$

又因为 $\alpha=0.05$，　所以查表可知 $z_{\alpha}=1.645$

所以$\bar{x}-z_{\alpha}\dfrac{\sigma}{\sqrt{n}}=1473.4-1.645\times\dfrac{100}{\sqrt{5}}=1399.833$(小时)

所以有 95% 的把握认为该厂灯泡平均使用寿命不低于 1399.833 小时。

2. 方差未知时的区间估计

1) 理论基础

(1) 求置信区间时需要的定义。

① t 分布单侧上 α 分位点

若随机变量 X 服从 t 分布，当自由度小于 45 时，满足式子 $P(X>t_{\alpha}(n))=\alpha$ 或 $P(X<-t_{\alpha}(n))=\alpha$ 的 $t_{\alpha}(n)$或$-t_{\alpha}(n)$叫 t 分布的单侧上 α 分位点；当自由度大于 45 时，t 分布单侧上 α 分位点转化为标准正态分布单侧上 α 分位点，如图 3-3 所示。

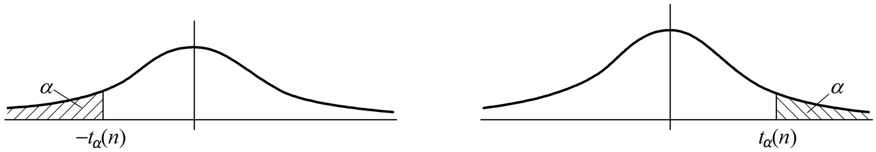

图 3-3　t 分布单侧上 α 分位点图解

② t 分布双侧上 α 分位点

若随机变量 X 服从 t 分布，当自由度小于 45 时，满足式子 $P(|X|>t_{\alpha/2}(n))=\alpha$ 的 $t_{\alpha/2}(n)$和$-t_{\alpha/2}(n)$叫 t 分布的双侧上 α 分位点。当自由度大于 45 时，t 分布双侧上 α 分位点

转化为标准正态分布双侧上 α 分位点，如图 3-4 所示。

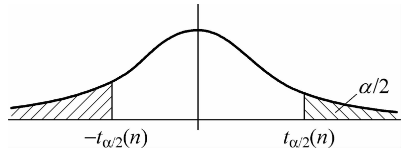

图 3-4　t 分布双侧上 α 分位点图解

（2）求置信区间时需要的定理。

若随机变量 $X \sim N(\mu, \sigma^2)$ 则有如下定理成立。

① 重复抽样条件下

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} \sim t(n-1)$$

② 不重复抽样条件下

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}\sqrt{\frac{N-n}{N-1}}} \sim t(n-1)$$

2）方差未知时置信区间的求解

（1）重复抽样条件下均值的(1−α)%的置信区间。

① 单侧置信区间的求解

因为$\frac{\bar{x}-\mu}{s/\sqrt{n}}$服从 t 分布，所以 $P\left(\frac{\bar{x}-\mu}{s/\sqrt{n}} > t_\alpha(n-1)\right) = \alpha$ 成立或 $P\left(\frac{\bar{x}-\mu}{s/\sqrt{n}} < -t_\alpha(n-1)\right) = \alpha$ 成立，即等价于 $P\left(\frac{\bar{x}-\mu}{s/\sqrt{n}} \leqslant t_\alpha(n-1)\right) = 1-\alpha$ 成立或 $P\left(\frac{\bar{x}-\mu}{s/\sqrt{n}} \geqslant -t_\alpha(n-1)\right) = 1-\alpha$ 成立，所以令$\frac{\bar{x}-\mu}{s/\sqrt{n}} \leqslant t_\alpha(n-1)$成立或$\frac{\bar{x}-\mu}{s/\sqrt{n}} \geqslant -t_\alpha(n-1)$成立，解这两个不等式就可以得到结果。

自由度小于 45 时重复抽样条件下均值的两个(1−α)%的单侧置信区间。

单侧右置信区间：$\mu \geqslant \bar{x} - t_\alpha(n-1)\frac{s}{\sqrt{n}}$

单侧左置信区间：$\mu \leqslant \bar{x} + t_\alpha(n-1)\frac{s}{\sqrt{n}}$

自由度大于 45 时重复抽样条件下均值的两个(1−α)%的单侧置信区间。

单侧右置信区间：$\mu \geqslant \bar{x} - z_\alpha \frac{s}{\sqrt{n}}$

单侧左置信区间：$\mu \leqslant \bar{x} + z_\alpha \frac{s}{\sqrt{n}}$

② 双侧置信区间的求解

因为 $\frac{\bar{x}-\mu}{s/\sqrt{n}}$ 服从 t 分布，所以 $P\left(\left|\frac{\bar{x}-\mu}{s/\sqrt{n}}\right| > t_{\alpha/2}(n-1)\right) = \alpha$ 成立，即等价于 $P\left(\left|\frac{\bar{x}-\mu}{s/\sqrt{n}}\right| \leqslant t_{\alpha/2}(n-1)\right) = 1-\alpha$ 成立，所以令 $\left|\frac{\bar{x}-\mu}{s/\sqrt{n}}\right| \leqslant t_{\alpha/2}(n-1)$成立，解这个绝对值不等式就可以得到结果。

自由度小于 45 时重复抽样条件下均值的(1−α)%双侧置信区间为：

$$\bar{x} - t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}} \leqslant \mu \leqslant \bar{x} + t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}}$$

自由度大于 45 时重复抽样条件下均值的(1−α)%双侧置信区间为：

$$\bar{x} - z_{\alpha/2}\frac{s}{\sqrt{n}} \leqslant \mu \leqslant \bar{x} + z_{\alpha/2}\frac{s}{\sqrt{n}}$$

单双侧置信区间中的$\frac{s}{\sqrt{n}}$称为均值的标准误差，又称为抽样平均误差或均值标准误、标准误，表示任何一个分布函数的标准差，是原来分布函数标准差的$\sqrt{n}$分之一，或者说$\overline{X}$分布的方差，就是X分布方差的n分之一。

(2) 不重复抽样条件下均值的$(1-\alpha)\%$的置信区间。

① 单侧置信区间的求解

因为$\dfrac{\bar{x}-\mu}{s/\sqrt{n}\sqrt{\frac{N-n}{N-1}}}$服从$t$分布，所以$P\left(\dfrac{\bar{x}-\mu}{s/\sqrt{n}\sqrt{\frac{N-n}{N-1}}}>t_{\alpha}(n-1)\right)=\alpha$成立或$P\left(\dfrac{\bar{x}-\mu}{s/\sqrt{n}\sqrt{\frac{N-n}{N-1}}}<-t_{\alpha}(n-1)\right)=\alpha$成立，即等价于$P\left(\dfrac{\bar{x}-\mu}{s/\sqrt{n}\sqrt{\frac{N-n}{N-1}}}\leqslant t_{\alpha}(n-1)\right)=1-\alpha$或$P\left(\dfrac{\bar{x}-\mu}{s/\sqrt{n}\sqrt{\frac{N-n}{N-1}}}\geqslant -t_{\alpha}(n-1)\right)=1-\alpha$成立，所以令$\dfrac{\bar{x}-\mu}{s/\sqrt{n}\sqrt{\frac{N-n}{N-1}}}\leqslant t_{\alpha}(n-1)$或$\dfrac{\bar{x}-\mu}{s/\sqrt{n}\sqrt{\frac{N-n}{N-1}}}\geqslant -t_{\alpha}(n-1)$成立，解这两个不等式就可以得到结果。

自由度小于45时重复抽样条件下均值的两个$(1-\alpha)\%$的单侧置信区间。

单侧右置信区间：$\mu\geqslant\bar{x}-t_{\alpha}(n-1)\frac{s}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$

单侧左置信区间：$\mu\leqslant\bar{x}+t_{\alpha}(n-1)\frac{s}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$

自由度大于45时重复抽样条件下均值的两个$(1-\alpha)\%$的单侧置信区间。

单侧右置信区间：$\mu\geqslant\bar{x}-z_{\alpha}\frac{s}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$

单侧左置信区间：$\mu\leqslant\bar{x}+z_{\alpha}\frac{s}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$

② 双侧置信区间的求解

因为$\dfrac{\bar{x}-\mu}{s/\sqrt{n}\sqrt{\frac{N-n}{N-1}}}$服从$t$分布，所以$P\left(\left|\dfrac{\bar{x}-\mu}{s/\sqrt{n}\sqrt{\frac{N-n}{N-1}}}\right|>t_{\alpha/2}(n-1)\right)=\alpha$成立，即等价于$P\left(\left|\dfrac{\bar{x}-\mu}{s/\sqrt{n}\sqrt{\frac{N-n}{N-1}}}\right|\leqslant t_{\alpha/2}(n-1)\right)=1-\alpha$成立，所以令$\left|\dfrac{\bar{x}-\mu}{s/\sqrt{n}\sqrt{\frac{N-n}{N-1}}}\right|\leqslant t_{\alpha/2}(n-1)$成立，解这个绝对值不等式就可以得到结果。

自由度小于45时重复抽样条件下均值的$(1-\alpha)\%$双侧置信区间为：

$$\bar{x}-t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}\leqslant\mu\leqslant\bar{x}+t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

自由度大于45时重复抽样条件下均值的$(1-\alpha)\%$双侧置信区间为：

$$\bar{x}-z_{\alpha/2}\frac{s}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}\leqslant\mu\leqslant\bar{x}+z_{\alpha/2}\frac{s}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

例 3.3 设某厂生产的灯泡寿命 $X\sim N(\mu,\sigma^2)$，现随机抽取 5 只，测量其寿命如下：1455、1502、1370、1610、1430，试以 95%的概率保证程度推断该厂灯泡的平均使用寿命为多少？

解：由题可知，这是一个总体方差已知时均值的双侧区间估计。因为 N 未知，所以采用重复抽样条件下的置信区间求解。因为自由度小于 45，所以重复抽样条件下均值的$(1-\alpha)$%的双侧置信区间为：

$$\bar{x}-t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}}\leqslant\mu\leqslant\bar{x}+t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}}$$

因为资料未分组，$\sigma=100, n=5$

所以 $\bar{x}=\dfrac{\sum x}{n}=\dfrac{1}{5}(1455+1502+1370+1610+1430)=1473.4$

$$S=\sqrt{\frac{\sum(x-\bar{x})^2}{n-1}}=\sqrt{\frac{\sum(x-1473.4)^2}{5-1}}=89.988$$

又因为 $\alpha=0.05$，所以查表可知 $t_{\alpha/2}(5-1)=t_{0.025}(4)=2.7764$

所以 $\bar{x}-t_{\alpha/2}(n-1)\dfrac{s}{\sqrt{n}}=1361.67$(小时)

$\bar{x}+t_{\alpha/2}(n-1)\dfrac{s}{\sqrt{n}}=1585.13$(小时)

所以有 95% 的把握认为该厂灯泡平均使用寿命为(1361.67，1585.13)。

例 3.4 设某厂生产的灯泡寿命 $X\sim N(\mu,\sigma^2)$，现随机抽取 5 只，测量其寿命如下：1455、1502、1370、1610、1430，试以 95%的概率保证程度推断该厂灯泡的平均使用寿命不低于多少？

解：由题可知，这是一个总体方差已知时均值的单侧右区间估计。因为 N 未知，所以采用重复抽样条件下的置信区间求解。因为重复抽样条件下均值的$(1-\alpha)$%的单侧右置信区间为：

$$\mu\geqslant\bar{x}-t_{\alpha}(n-1)\frac{s}{\sqrt{n}}$$

因为资料未分组，$\sigma=100, n=5$

所以 $\bar{x}=\dfrac{\sum x}{n}=\dfrac{1}{5}(1455+1502+1370+1610+1430)=1473.4$

又因为 $\alpha=0.05$，所以查表可知 $t_{\alpha}(5-1)=t_{0.05}(4)=2.1318$

因为 $s=\sqrt{\dfrac{\sum(x-\bar{x})^2}{n-1}}=\sqrt{\dfrac{\sum(x-1473.4)^2}{5-1}}=89.988$

所以 $\bar{x}-t_{0.05}(4)\dfrac{s}{\sqrt{n}}=1473.4-2.1318\times\dfrac{89.988}{\sqrt{5}}=1387.61$

所以有 95% 的把握认为该厂灯泡平均使用寿命不低于 1387.61 小时。

该题的 SPSS 操作过程如下。

(1) 在定义变量标签窗口定义“灯泡使用寿命”变量。

(2) 在录入数据窗口依次录入 1455、1502、1370、1610、1430 这 5 个数据，如图 3-5 所示。

(3) 选择 Analyze 下拉菜单中的 Descriptive Statistics 子菜单，选择 Explore 模块，如图 3-6 所示。

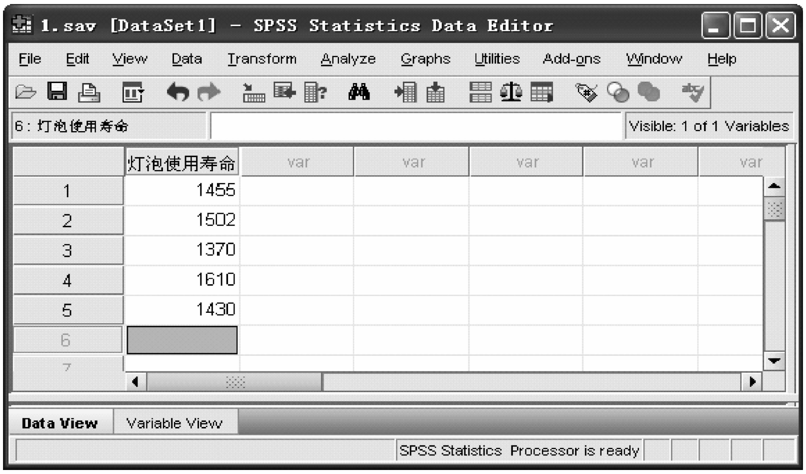

图 3-5　SPSS 定义变量窗口

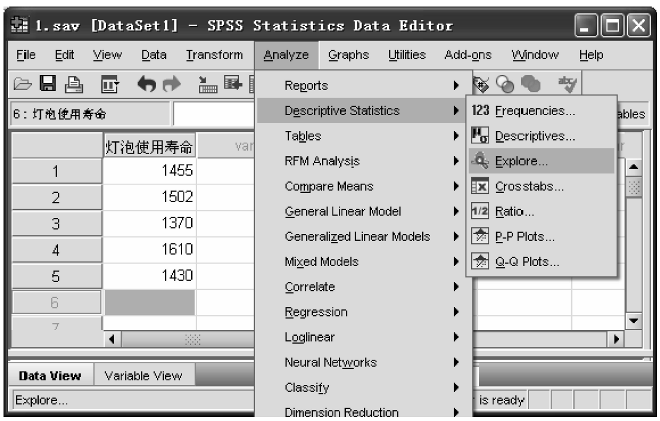

图 3-6　SPSS 的 Analyze 界面

(4) 选择 Explore 模块后，会出现探索分析对话框，然后将“灯泡使用寿命”从左边框选中移入右边框，在 Display 栏选择 Statistics，如图 3-7 所示。

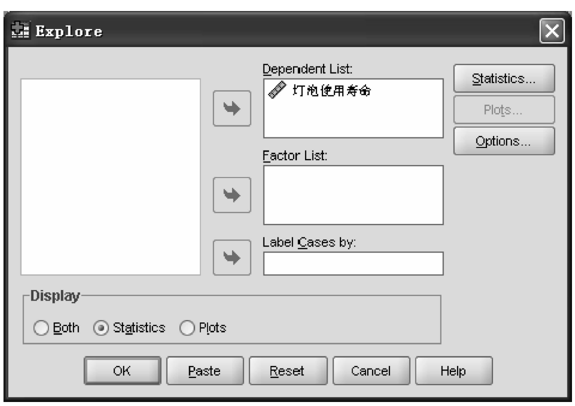

图 3-7　Explore 对话框

然后单击 OK,SPSS 就会将区间估计的结果输出,如表 3-1 和表 3-2 所示。

表 3-1 Case Processing Summary

	Cases					
	Valid		Missing		Total	
灯泡使用寿命	N	Percent	N	Percent	N	Percent
	5	100.0%	0	0.0%	5	100.0%

表 3-2 Descriptives

			Statistic	Std. Error
灯泡使用寿命	Mean		1473.40	40.244
	95% Confidence Interval for Mean	Lower Bound	1361.67	
		Upper Bound	1585.13	
	5% Trimmed Mean		1471.56	
	Median		1455.00	
	Variance		8097.800	
	Std. Deviation		89.988	
	Minimum		1370	
	Maximum		1610	
	Range		240	
	Interquartile Range		156	
	Skewness		0.788	0.913
	Kurtosis		0.899	2.000

由表 3-1 可知,灯泡使用寿命样本数据有 5 个,缺失值为 0,全部为有效数据。

由表 3-2 可知:① 因为 $n=5>45$,所以所有灯泡使用寿命 95% 的置信区间为$\left(\bar{x}-\frac{s}{\sqrt{n}}z_{\alpha/2},\bar{x}+\frac{s}{\sqrt{n}}z_{\alpha/2}\right)$,

又因为由表可知$\bar{x}=\frac{\sum x}{n}=\frac{1}{5}(1455+1502+1370+1610+1430)=1473.4$,

所以所有学生考试成绩 95%的置信区间为(1361.67,1585.13)。

② 因为$\bar{x}=1473.4$,

$M_e=L+\frac{\frac{\sum f}{2}-s}{f}\times d=1455$,不大致相同,又因为峰度为 0.899,不接近于 0;偏度为 0.788 不接近于 0,所以该 5 个灯泡使用寿命数据不服从正态分布。

③ 偏度为 0.788>0,表示正偏差数值较大,说明数据位于均值右边的比位于左边的少,直观表现为右边的尾部相对于比左边的尾部要长,因为有少数变量值很大,使曲线右侧尾部拖得很长;峰度为 0.899>0,表示比标准正态分布高峰更加陡峭,为尖顶峰分布。

④ 100%的数据落入范围(1370,1610),它们之间的距离为 240;50%的数据落入范围(Q_1,Q_3),它们之间的距离为 156。