Lec7: 区间估计(一): 置信区间

张伟平

2011 年 3 月 28 日

1 区间估计的基本概念

一、参数的区间估计问题

使用点估计$\hat{g}(X)$估计$g(\theta)$的缺点是: 单从所给出的估计值上, 无法看出它的精度有多大. 当然你可以定义某种指标, 如估计的均方误差之类去刻画它的精度, 但也还是间接的. 更直接的方法是指出了一个误差限$d(\mathbf{X})$, 而把估计写成$\hat{g}(\mathbf{X}) \pm d(\mathbf{X})$的形式. 这实际上就是一种区间估计, 即估计$g(\theta)$ 的取值在$[\hat{g}(\mathbf{X}) - d(\mathbf{X}),\ \hat{g}(\mathbf{X}) + d(\mathbf{X})]$ 之内. 将其一般化, 给出区间估计的下列定义.

定义1 设有一个参数分布族$\mathscr{F} = \{f(x,\theta),\ \theta \in \Theta\}$, $g(\theta)$ 是定义在参数空间Θ 的一个已知函数, $\mathbf{X} = (X_1, \cdots, X_n)$ 是从分布族中某总体$f(x,\theta)$ 中抽取的样本, 令$\hat{g}_1(\mathbf{X})$ 和$\hat{g}_2(\mathbf{X})$ 为定义在样本空间$\mathscr{X}$ 上, 取值在Θ 上的两个统计量, 且$\hat{g}_1(\mathbf{X}) \le \hat{g}_2(\mathbf{X})$, 则称随机区间$[\hat{g}_1(\mathbf{X}),\ \hat{g}_2(\mathbf{X})]$ 为$g(\theta)$ 的一个*区间估计*(Interval estimation).

根据这个定义, 从形式上看, 任何一个满足条件$\hat{g}_1 \le \hat{g}_2$ 的统计量$\hat{g}_1$, $\hat{g}_2$ 都可构成$g(\theta)$ 的一个区间估计$[\hat{g}_1, \hat{g}_2]$. 既然一个未知参数的区间估计有很多种, 如何从中挑选一个好的区间估计呢? 这就涉及到评价一个区间估计优劣的标准问题. 评价一个区间估计优劣的标准有两个要素: 可靠性与精确度(也称精度). *可靠性*是指待估参数$g(\theta)$ 被包含在$[\hat{g}_1, \hat{g}_2]$ 内的可能性有多大. 可能性越大, 可靠性越高. *精确度*可由随机区间的平均长度来度量. 长度越短, 精确度越高.

不言而喻, 我们希望所作的区间估计既有高的可靠性, 又有高的精确度. 但这二者往往是彼此矛盾的, 不可能同时都很高. 当样本大小固定时, 若精确度提高了, 可靠性就降低了; 反之, 若可靠度提高了, 则精确度就降低了.

如何构造尽可能高的可靠性和高精确度的区间估计呢? 通常采用的方法是在保证一定可靠度的前提下选择精确度尽可能高的区间估计. 这就是著名统计学家Neyman 提出的一种妥协方案.

当然, 如果在应用中人们要求可靠性和精度都很高, 则必须加大样本容量, 也就是说要多做一些试验, 才可能实现.

二、置信区间

为书写简单计, 本节以下假定被估计的$g(\theta)$ 就是θ 自身, 这与一般情况没有原则区别.

1. 置信度

设$\mathbf{X}$ 为样本, $[\hat{\theta}_1(\mathbf{X}),\hat{\theta}_2(\mathbf{X})]$ 是θ 的一个区间估计. 由于θ 是未知的, 且样本是随机的, 我们不能保证在任何情况下(即对任何具体的样本值), 区间$[\hat{\theta}_1,\hat{\theta}_2]$ 必定包含θ , 而只能以一定的概率保证它. 希望随机区间$[\hat{\theta}_1,\hat{\theta}_2]$ 包含θ 的概率$P_\theta(\hat{\theta}_1 \le \theta \le \hat{\theta}_2)$ 越大越好. 这个概率就是我们前面所说的可靠性, 数理统计学上称这个概率为置信度. 一般说来, 这个概率与θ 有关, 假如一个区间估计对某个$\theta_1 \in \Theta$ 其置信度大, 而对另一个$\theta_2 \in \Theta$ 其置信度小, 那么这种区间估计的适应性要差一些, 不能认为是一个好的区间估计. 若对参数空间Θ 中的任一θ, 其置信度都很大, 则此种区间估计就是一种好的区间估计. 因此有如下定义.

定义2　设随机区间$[\hat{\theta}_1,\hat{\theta}_2]$ 为参数θ 的一个区间估计, 则称置信度在参数空间Θ 上的下确界

$$\inf_{\theta\in\Theta} P_\theta\left(\hat{\theta}_1 \le \theta \le \hat{\theta}_2\right)$$

为该区间估计的置信系数(Confidence coefficient)

显然, 一个区间估计的置信度越大越好. 为了计算置信度和置信系数, 需要利用统计量的精确分布或渐近分布. 可见抽样分布在评价和构造区间估计中发挥重要作用.

2. 精确度

精确度的概念我们在前面已说过. 精确度的标准不止一个. 这里介绍其中最常见的一个标准, 即随机区间$[\hat{\theta}_1,\hat{\theta}_2]$ 的平均长度$E_\theta(\hat{\theta}_2-\hat{\theta}_1)$. 平均长度越短, 精确度越高, 这也是符合实际的一项要求. 为说明精确度和置信度及其关系, 请看下例.

例1　设样本$\mathbf{X}=(X_1,\cdots,X_n)$ 来自正态总体$N(\mu,\sigma^2)$, 其中$-\infty<\mu<\infty$, $\sigma^2>0$. μ 和σ^2 的估计量分别是样本均值$\bar{X}$ 和样本方差$S^2=\frac{1}{n-1}\sum_{i=1}^n(X_i-\bar{X})^2$, 我们用$[\bar{X}-kS/\sqrt{n},\ \bar{X}+kS/\sqrt{n}]$ 作为总体均值μ 的区间估计. 考虑其置信度和精确度.

解　上述区间估计的置信度为

$$\begin{aligned}&P_\mu\left(\bar{X}-kS/\sqrt{n}\le\mu\le\bar{X}+kS/\sqrt{n}\right)=P_\mu\left(\left|\sqrt{n}(\bar{X}-\mu)/S\right|\le k\right)\\&=P(|T|\le k),\end{aligned}$$

其中$T=\sqrt{n}(\bar{X}-\mu)/S\sim t_{n-1}$, 其分布与$\mu$ 无关, 因而区间估计的置信系数为$P(|T|\le k)$. 显然k 越大, 区间的置信系数越大, 区间就越可靠.

由于$(n-1)S^2/\sigma^2\sim\chi^2_{n-1}$, 所以区间的平均长度为

$$l_k=2kE(s)/\sqrt{n}=\frac{2\sqrt{2}k\sigma\,\Gamma(n/2)}{\sqrt{n(n-1)}\,\Gamma((n-1)/2)}.$$

显然,k 越大,区间也越长,也就越不精确.

由此例可以看到, 在样本容量n 给定后, 为了提高置信度, 需要增加k 值, 从而放大了区间, 降低了精确度. 反过来, 为了提高精确度, 需要减小k 值, 从而缩短了区间, 降低了置信度. 置信度与精确度互相制约着. 如前所述, 面对这一矛盾, 著名统计学家Neyman 建议采取如下方案: 在保证置信系数达到指定要求的前提下, 尽可能提高精确度. 这一建议导致引入如下置信区间的概念,由于是Neyman 建议的, 通常也称置信区间为Neyman 置信区间.

定义3 设$[\hat{\theta}_1(\mathbf{X}), \hat{\theta}_2(\mathbf{X})]$ 是参数θ 的一个区间估计, 若对给定的$0 < \alpha < 1$, 有

$$P_\theta\left(\hat{\theta}_1(\mathbf{X}) \le \theta \le \hat{\theta}_2(\mathbf{X})\right) \ge 1 - \alpha, \qquad \text{对一切 } \theta \in \Theta,$$

则称$[\hat{\theta}_1(\mathbf{X}), \hat{\theta}_2(\mathbf{X})]$ 是θ 的置信水平 (Confidence level)为$1-\alpha$ 的置信区间 (Confidence interval).

置信区间有如下频率解释: 设$\alpha = 0.05$, 则$1 - \alpha = 0.95$, 若把置信区间 $[\hat{\theta}_1, \hat{\theta}_2]$ 反复使用多次, 如使用100 次, 平均大约有95 次随机区间$[\hat{\theta}_1, \hat{\theta}_2]$ 包含真参数θ , 大约平均有5 次随机区间$[\hat{\theta}_1, \hat{\theta}_2]$ 不包含θ . 当使用次数充分大时, 频率接近于置信系数.

三、置信限

在一些实际问题中, 人们感兴趣的有时仅仅是未知参数的置信上限或置信下限. 例如一种新材料的强度, 我们关心它最低不少于多少; 一个工厂的废品率, 我们关心它最高不超过多少等等. 这也是一种区间估计, 称之为置信下限或置信上限, 定义如下:

定义4 设$\hat{\theta}_L(\mathbf{X})$ 和$\hat{\theta}_U(\mathbf{X})$ 是定义在样本空间$\mathscr{X}$ 上, 在参数空间Θ 上取值的两个统计量, 若对给定的$0 < \alpha < 1$, 有

$$P_\theta\left(\hat{\theta}_L(\mathbf{X}) \le \theta\right) \ge 1 - \alpha, \quad \text{对一切 } \theta \in \Theta,$$
$$P_\theta\left(\theta \le \hat{\theta}_U(\mathbf{X})\right) \ge 1 - \alpha, \quad \text{对一切 } \theta \in \Theta,$$

则分别称$\hat{\theta}_L(\mathbf{X})$ 和$\hat{\theta}_U(\mathbf{X})$ 是θ 的置信水平为$1-\alpha$ 的(单侧) 置信下限 (Lower confidence limit)和置信上限 (Upper confidence limit). 上式左端概率在参数空间Θ 上的下确界分别称为置信下、上限的置信系数.

显然, 对置信下限$\hat{\theta}_L$ 而言, 若$E(\hat{\theta}_L)$ 越大, 则置信下限的精确度越高; 对置信上限$\hat{\theta}_U$ 而言, 若$E(\hat{\theta}_U)$ 越小, 其精确度越高.

容易看出, 单侧置信上、下限都是置信区间的特例. 因此寻求置信区间的方法可以毫不困难地用来求单侧置信上、下限. 单侧置信限与双侧置信限之间存在着一个简单的联系, 下述的引理告诉我们, 在有了单侧置信上、下限后, 也不难求得置信区间.

引理1 设$\hat{\theta}_L(\mathbf{X})$ 和$\hat{\theta}_U(\mathbf{X})$ 分别是参数θ 的置信水平为$1 - \alpha_1$ 和$1 - \alpha_2$ 的单侧置信下、上限, 且对任何样本$\mathbf{X}$, 都有$\hat{\theta}_L(\mathbf{X}) \le \hat{\theta}_U(\mathbf{X})$, 则$[\hat{\theta}_L(\mathbf{X}), \hat{\theta}_U(\mathbf{X})]$ 是θ 的置信水平为$1 - (\alpha_1 + \alpha_2)$ 的双侧置信区间.

证 在引理的假设下, 下列三个事件

$$\left\{\hat{\theta}_L(\mathbf{X}) \le \theta \le \hat{\theta}_U(\mathbf{X})\right\}, \quad \left\{\theta < \hat{\theta}_L(\mathbf{X})\right\}, \quad \left\{\theta > \hat{\theta}_U(\mathbf{X})\right\}$$

是互不相容的, “三个事件之并”为“必然事件”. 再考虑到

$$P_\theta\left(\theta < \hat{\theta}_L(\mathbf{X})\right) = 1 - P_\theta\left(\hat{\theta}_L(\mathbf{X}) \le \theta\right) \le \alpha_1,$$
$$P_\theta\left(\theta > \hat{\theta}_U(\mathbf{X})\right) = 1 - P_\theta\left(\theta \le \hat{\theta}_U(\mathbf{X})\right) \le \alpha_2,$$

因此有

$$P_\theta\left(\hat{\theta}_L(\mathbf{X}) \le \theta \le \hat{\theta}_U(\mathbf{X})\right) = 1 - P_\theta\left(\theta < \hat{\theta}_L(\mathbf{X})\right) - P_\theta\left(\theta > \hat{\theta}_U(\mathbf{X})\right)$$

$$\geq 1-(\alpha_1+\alpha_2)\,.$$

引理得证.

四、置信域

以上讨论的置信区间和置信上、下限都是假定参数θ 是一维的, 可以将其推广到参数θ 是k维($k\geq 2$) 的情形, 就得如下定义的置信域.

定义5 设有一个参数分布族$\mathscr{F}=\{f(x,\theta),\,\theta\in\Theta\}$, Θ是参数空间. 其中$\theta=(\theta_1,\cdots,\theta_k)\in\Theta\subset R_k$, $k\geq 2$. $\mathbf{X}=(X_1,\cdots,X_n)$ 是来自分布族中某总体$f(x,\theta)$ 的样本. 若$S(\mathbf{X})$ 满足

(i) 对任一样本$\mathbf{X}$, $S(\mathbf{X})$ 是Θ 的一个子集;

(ii)对给定的$0<\alpha<1$, $P_\theta\big(\theta\in S(\mathbf{X})\big)\geq 1-\alpha$, 一切$\theta\in\Theta$;

则称$S(\mathbf{X})$ 是θ 的置信水平为$1-\alpha$ 的置信域(Confidence region) 或置信集, 而$\inf\limits_{\theta\in\Theta}P_\theta\big(\theta\in S(\mathbf{X})\big)$称为置信系数.

在多维场合, 置信域$S(\mathbf{X})$ 的形状可以是各种各样的, 但实用上只限于一些规则的几何图形, 如其各面与坐标平面平行的长方体、球、椭球等. 特别当置信集是长方体(其面与坐标平面平行) , 则称其为联合置信区间.

五、构造区间估计的方法

目前应用最广泛的区间估计的形式是Neyman 的置信区间. 本章第二节和第三节将介绍这一方法, 这一方法的关键是基于点估计去构造枢轴变量, 因此也称为枢轴变量法. 另外一种构造区间估计的重要方法是利用假设检验构造置信区间, 它与枢轴变量法同属于一个理论体系, 即Neyman 的关于置信区间和假设检验的理论. 利用假设检验构造置信区间的方法将在下一章有专门的一节介绍.

本章的最后两节将介绍区间估计的其它两种方法, 即Fisher 的信仰推断方法和容忍区间和容忍限.

用Bayes 方法求区间估计的内容将放在本书的最后一章介绍.

2 枢轴变量法—正态总体参数的置信区间

一、引言

这个方法的基本要点, 就是在参数的点估计基础上, 去找它的置信区间. 由于点估计是由样本决定的, 是最有可能接近真参数θ之值. 因此, 围绕点估计值的区间, 包含真参数值的可能性也就要大一些. 请看下面的例子, 是如何构造置信区间的.

例1 设$\mathbf{X}=(X_1,\cdots,X_n)$ 是从总体$N(\mu,\sigma^2)$ 中抽取的简单随机样本, 此处σ^2 已知, 求μ的置信系数为$1-\alpha$ 的置信区间和置信上、下限.

解 显然, μ 的一个良好的点估计是$\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$, 其分布为$\bar{X}\sim N(\mu,\sigma^2/n)$, 将其标准化得

$$U=\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma}\sim N(0,1),$$

其分布与μ 无关. 由于正态分布的对称性, 可得

$$P_\mu\left(\left|\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma}\right|\le u_{\alpha/2}\right)=1-\alpha,$$

此处$u_{\alpha/2}$ 为标准正态分布的上侧$\alpha/2$ 分位数. 经不等式等价变形, 可知

$$P_\mu\left(\bar{X}-\frac{\sigma}{\sqrt{n}}u_{\alpha/2}<\mu<\bar{X}+\frac{\sigma}{\sqrt{n}}u_{\alpha/2}\right)=1-\alpha.$$

因此$\left[\bar{X}-\frac{\sigma}{\sqrt{n}}u_{\alpha/2},\bar{X}+\frac{\sigma}{\sqrt{n}}u_{\alpha/2}\right]$ 为μ 的置信系数$1-\alpha$ 的置信区间.

由本例可知构造置信区间的步骤如下:

1. 找待估参数μ 的一个良好点估计. 此例中这个点估计是$T(\mathbf{X})=\bar{X}$.

2. 构造一个$T(\mathbf{X})$ 和μ 的函数$\varphi(T,\mu)$, 使其满足:

   (i) 其表达式与待估参数μ 有关;

   (ii) 其分布与待估参数μ 无关.

则称随机变量$\varphi(T,\mu)$ 为枢轴变量. 本例中这一变量即为$U=\sqrt{n}(\bar{X}-\mu)/\sigma$, 它的表达式与$\mu$ 有关, 但其分布$N(0,1)$ 与μ 无关. 因此U 为枢轴变量.

3. 对给定的$0<\alpha<1$, 决定两个常数a 和b, 使得

$$P_\mu\left(a\le\varphi(T,\mu)\le b\right)=1-\alpha.$$

解括号中的不等式得到$\hat{\mu}_L(\mathbf{X})\le\mu\le\hat{\mu}_U(\mathbf{X})$, 则有

$$P_\mu\left(\hat{\mu}_L(\mathbf{X})\le\mu\le\hat{\mu}_U(\mathbf{X})\right)=1-\alpha.$$

这表明$\left[\hat{\mu}_L(\mathbf{X}),\hat{\mu}_U(\mathbf{X})\right]$ 是μ 的置信水平为$1-\alpha$ 的置信区间.

例4.2.1中的μ 的置信区间$\left[\bar{X}-\sigma u_{\alpha/2}/\sqrt{n},\bar{X}+\sigma u_{\alpha/2}/\sqrt{n}\right]$ 就是通过上述三个步骤获得的. 其中最关键的步骤是第二步, 即构造枢轴变量$\varphi(T,\mu)$, 这个变量一定和μ 的一个良好的点估计有关.这种构造置信区间的方法称为*枢轴变量法*.

二、单个正态总体参数的置信区间

正态分布$N(\mu,\sigma^2)$ 是常用的分布. 寻求它的两个参数μ 和σ^2 的置信区间是实际中常遇到的的问题, 下面将分几种情况分别加以讨论. 这里总假设$\mathbf{X}=(X_1,\cdots,X_n)$ 是从正态总体$N(\mu,\sigma^2)$ 抽取的简单随机样本. 记

$$\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i,\quad S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2,$$

即$\bar{X}$ 和S^2 分别为样本均值和样本方差.

1. σ^2 已知, 求μ 的置信区间

这就是例4.2.1讨论过的问题. μ 的置信系数为$1-\alpha$的置信区间为

$$\Big[\bar{X}-\frac{\sigma}{\sqrt{n}}u_{\alpha/2}, \bar{X}+\frac{\sigma}{\sqrt{n}}u_{\alpha/2}\Big]. \tag{2.1}$$

区间的长度为$l_n=2\sigma u_{\alpha/2}/\sqrt{n}$. 由此可以看出

(1)样本容量n 越大,该区间越短,精确度就越高.

(2) σ 越大, 则l_n 越大, 精确度越低. 这是因为方差越大, 随机影响也就越大, 精确度就会低下来.

(3)置信系数$1-\alpha$ 越大, 则α 越小,从而$u_{\alpha/2}$ 就越大,l_n 越长,精确度就越低.

由此可见在σ 和α 固定的情形下, 要提高精确度, 只有增加样本容量. 例如, 置信系数$1-\alpha$固定, 要使上述置信区间的长度$l_n\le l_0$, l_0 为给定的常数, 则$n\ge\big[\big(2\sigma u_{\alpha/2}/l_0\big)^2\big]$, 其中$[x]$ 表示实数x 的整数部分.

例2 设某车间生产零件的长度$X\sim N(\mu,0.09)$, 若得到一组样本观察值为

$$12.6, \quad 13.4, \quad 12.8, \quad 13.2$$

求零件平均长度μ 的95%的置信区间.

解 由样本观察值算得$\bar{X}=13$, $n=4$, $\sigma=0.3$, 查表求得$u_{0.025}=1.96$, $\sigma u_{\alpha/2}/\sqrt{n}=0.3\times1.96/2=0.294$, 由公式(2.1)可知$\mu$ 的95%的置信区间

$$\Big[\bar{x}-\frac{\sigma}{\sqrt{n}}u_{\alpha/2}, \bar{x}+\frac{\sigma}{\sqrt{n}}u_{\alpha/2}\Big]=[12.71, 13.29].$$

2. σ^2 未知, 求μ 的置信区间

在这种情况下, μ 的良好的点估计仍为$\bar{X}$, 基于$\bar{X}$ 构造枢轴变量

$$T=\frac{\sqrt{n}(\bar{X}-\mu)}{S}.$$

由推论2.4.2可知$T\sim t_{n-1}$. 可见T 的表达式与μ 有关, 而其分布与μ 无关. 故T 为枢轴变量. 由于t 分布关于原点对称, 令

$$P_\mu(|T|\le c)=P\bigg(-c\le\frac{\sqrt{n}(\bar{X}-\mu)}{S}\le c\bigg)=1-\alpha,$$

则$c=t_{n-1}(\alpha/2)$. 将括号中的不等式经过等价变形得μ 的置信系数为$1-\alpha$ 的置信区间为

$$\Big[\bar{X}-\frac{S}{\sqrt{n}}t_{n-1}(\alpha/2),\ \bar{X}+\frac{S}{\sqrt{n}}t_{n-1}(\alpha/2)\Big], \tag{2.2}$$

此处$t_{n-1}(\alpha/2)$ 是自由度$n-1$ 的t 分布的上侧$\alpha/2$ 分位数.

例3 为测得某种溶液中的甲醛浓度, 取样得4个独立测定值的平均值$\bar{X}=8.34\%$, 样本标准差$S=0.03\%$, 并设被测总体近似服从正态分布, 求总体均值μ 的95%的置信区间.

解　因为$1-\alpha=0.95$, $n=4$,查表得$t_{n-1}(\alpha/2)=t_3(0.025)=3.182$, 故有$St_{n-1}(\alpha/2)/\sqrt{n}=0.03\times 3.182/2=0.0477$, $\bar{X}=8.34$, 由公式(2.2)可知μ 得置信系数为95%的置信区间为

$$\left[\bar{X}-\frac{S}{\sqrt{n}}t_{n-1}(\alpha/2),\,\bar{X}+\frac{S}{\sqrt{n}}t_{n-1}(\alpha/2)\right]=[8.292\%,\,8.388\%].$$

3. μ **已知, 求**σ^2 **的置信区间**

当μ 已知时, σ^2 的一个良好的无偏估计为$S_n^2=\frac{1}{n}\sum_{i=1}^n(X_i-\mu)^2$, 且$nS^2/\sigma^2\sim\chi_n^2$. 则取$T=nS_n^2/\sigma^2$ 为枢轴变量, 其表达式与σ^2 有关, 但其分布与σ^2 无关, 找c_1 和c_2 使得

$$P_{\sigma^2}\Big(c_1\le nS_n^2/\sigma^2\le c_2\Big)=1-\alpha.$$

满足上式要求的c_1 和c_2 有无穷多对, 其中有一对c_1和c_2 , 使区间的长度最短, 但这样一对c_1 和c_2 不易求得, 且表达式复杂, 应用不方便. 一般令c_1 和c_2 满足下列要求

$$P_{\sigma^2}\Big(nS_n^2/\sigma^2<c_1\Big)=\alpha/2,\quad P_{\sigma^2}\Big(nS_n^2/\sigma^2>c_2\Big)=\alpha/2.$$

由χ^2分布的上侧分位数表可知$c_1=\chi_n^2(1-\alpha/2)$, $c_2=\chi_n^2(\alpha/2)$, 即有

$$P_{\sigma^2}\Big(\chi_n^2(1-\alpha/2)\le nS_n^2/\sigma^2\le\chi_n^2(\alpha/2)\Big)=1-\alpha.$$

最后利用不等式的等价变形, 得到σ^2 的置信系数为$1-\alpha$ 的置信区间为

$$\left[\frac{nS_n^2}{\chi_n^2(\alpha/2)},\,\frac{nS_n^2}{\chi_n^2(1-\alpha/2)}\right]=\left[\frac{\sum_{i=1}^n(X_i-\mu)^2}{\chi_n^2(\alpha/2)},\,\frac{\sum_{i=1}^n(X_i-\mu)^2}{\chi_n^2(1-\alpha/2)}\right],\tag{2.3}$$

此处$\chi_n^2(\alpha/2)$ 和$\chi_n^2(1-\alpha/2)$ 分别是自由度为n 的χ^2分布的上侧$\alpha/2$ 和$1-\alpha/2$ 分位数.

例4　为了解一台测量长度的仪器的精度, 对一根长30 mm 的标准金属棒进行了6次测量, 结果(单位:mm)是

$$30.1,\ 29.9,\ 29.8,\ 30.3,\ 30.2,\ 29.6$$

假如测量值服从正态分布$N(30,\sigma^2)$, 要求σ^2 的置信水平为0.95的置信区间.

解　此处$n=6$, $\mu=30$,易得出$\sum_{i=1}^6(X_i-\mu)^2=0.35$, $\alpha=0.05$, 查表得$\chi_6^2(0.025)=14.4494$, $\chi_6^2(0.975)=1.2375$, 由公式(2.3)可算得

$$\widehat{\sigma}_L^2=\sum_{i=1}^n(X_i-\mu)^2\Big/\chi_n^2(\alpha/2)=0.35/14.4494=0.0242,$$
$$\widehat{\sigma}_U^2=\sum_{i=1}^n(X_i-\mu)^2\Big/\chi_n^2(1-\alpha/2)=0.35/1.2375=0.2828.$$

因此σ^2 的置信水平为95%的置信区间为$\left[\widehat{\sigma}_L^2,\widehat{\sigma}_U^2\right]=[0.0242,\,0.2828]$.

4. μ **未知, 求**σ^2 **得置信区间**

记$\theta=(\mu,\sigma^2)$. 此时$S^2=\frac{1}{n-1}\sum_{i=1}^n(X_i-\bar{X})^2$ 是σ^2 的良好估计, 它是无偏的. 且由定理2.2.3可知$(n-1)S^2/\sigma^2\sim\chi_{n-1}^2$. 取$T=(n-1)S^2/\sigma^2$ 为枢轴变量, 其表达式与σ^2 有关, 而其分布与σ^2 无关. 找d_1 和d_2, 使得

$$P_\theta\Big(d_1\le(n-1)S^2/\sigma^2\le d_2\Big)=1-\alpha.$$

类似于3中确定c_1 和c_2 的理由和方法, 取$d_1=\chi^2_{n-1}(1-\alpha/2)$, $d_2=\chi^2_{n-1}(\alpha/2)$, 故有

$$P_\theta\Big(\chi^2_{n-1}(1-\alpha/2)\le(n-1)S^2/\sigma^2\le\chi^2_{n-1}(\alpha/2)\Big)=1-\alpha.$$

最后再利用不等式的等价变形, 得出σ^2 的置信系数为$1-\alpha$ 的置信区间为

$$\left[\frac{(n-1)S^2}{\chi^2_n(\alpha/2)},\frac{(n-1)S^2}{\chi^2_n(1-\alpha/2)}\right]=\left[\frac{\sum_{i=1}^n(X_i-\bar X)^2}{\chi^2_{n-1}(\alpha/2)},\ \frac{\sum_{i=1}^n(X_i-\bar X)^2}{\chi^2_{n-1}(1-\alpha/2)}\right].\tag{2.4}$$

若把这个随机区间的两个端点开平方，得到σ的置信系数为$1-\alpha$ 的置信区间如下：

$$\left[\Big(\sum_{i=1}^n(X_i-\bar X)^2\Big/\chi^2_{n-1}(\alpha/2)\Big)^{1/2},\ \Big(\sum_{i=1}^n(X_i-\bar X)^2\Big/\chi^2_{n-1}(1-\alpha/2)\Big)^{1/2}\right].$$

例5　求例3中总体方差σ^2及σ的置信系数为95%的置信区间.

解　如同例3, $n-1=3$, $\alpha/2=0.025$, $1-\alpha=0.975$.查表求得$\chi^2_3(0.025)=9.348$, $\chi^2_3(0.975)=0.216$, $S^2=0.0009$,由公式(2.4)可知σ^2的置信系数为95%的置信区间为

$$\big[(n-1)S^2/\chi^2_n(\alpha/2),\ (n-1)S^2/\chi^2_n(1-\alpha/2)\big]=[0.00029,\ 0.0125],$$

σ 的置信系数为95% 的置信区间为

$$\begin{aligned}&\left[\Big((n-1)S^2/\chi^2_{n-1}(\alpha/2)\Big)^{1/2},\ \Big((n-1)S^2/\chi^2_{n-1}(1-\alpha/2)\Big)^{1/2}\right]\\&=[0.017,\ 0.112].\end{aligned}$$

5. 二维参数$\theta=(\mu,\sigma^2)$ 的置信水平为$1-\alpha$的置信域

在正态分布情况下, $\bar X$ 和S^2 分别为μ 和σ^2 的无偏估计, 且为充分统计量, 它们之间还相互独立. 取枢轴变量为

$$\begin{aligned}&\sqrt{n}(\bar X-\mu)/\sigma\sim N(0,1),\\&(n-1)S^2/\sigma^2\sim\chi^2_{n-1}.\end{aligned}$$

对给定的置信水平$1-\alpha$, 可以通过标准正态分布和自由度为$n-1$ 的χ^2 分布找出三个数: c, d_1 和d_2, 使得

$$\begin{aligned}&P_\theta\left(|\sqrt{n}(\bar X-\mu)|/\sigma\le c\right)=\sqrt{1-\alpha},\\&P_\theta\left(d_1\le(n-1)S^2/\sigma^2\le d_2\right)=\sqrt{1-\alpha}.\end{aligned}$$

令$\sqrt{1-\alpha}=1-\gamma$, 则取$c=u_{\gamma/2}$, $d_1=\chi^2_{n-1}(1-\gamma/2)$, $d_2=\chi^2_{n-1}(\gamma/2)$. 由于$\bar X$ 和S^2 独立, 有

$$P\left((\bar X-\mu)^2\le\sigma^2u^2_{\gamma/2}/n,\ \ \frac{(n-1)S^2}{\chi^2_{n-1}(\gamma/2)}\le\sigma^2\le\frac{(n-1)S^2}{\chi^2_{n-1}(1-\gamma/2)}\right)=1-\alpha.$$

所以$\theta=(\mu,\sigma^2)$ 的置信水平为$1-\alpha$ 的置信域为

$$\left\{(\mu,\sigma^2):(\bar X-\mu)^2\le\sigma^2u^2_{\gamma/2}/n,\ \ \frac{(n-1)S^2}{\chi^2_{n-1}(\gamma/2)}\le\sigma^2\le\frac{(n-1)S^2}{\chi^2_{n-1}(1-\gamma/2)}\right\}.$$

三、两个正态总体参数的置信区间

设$X_1, \cdots, X_m$ 是自正态总体$N(a, \sigma_1^2)$ 抽取的简单随机样本, $Y_1, \cdots, Y_n$ 是自正态总体$N(b, \sigma_2^2)$ 抽取的简单随机样本, 且合样本独立. 设$\bar{X}, \bar{Y}$ 和S_X^2, S_Y^2 分别为这两组样本的样本均值和样本方差, 其中$S_1^2 = \frac{1}{m-1}\sum_{i=1}^m (X_i - \bar{X})^2$, $S_2^2 = \frac{1}{n-1}\sum_{i=1}^n (Y_i - \bar{Y})^2$. 下面分两种情况讨论两个正态总体均值差和方差比的置信区间问题.

1. 均值差$b-a$ 的置信区间

分下列几种情况:

(1) 当$m = n$ 时, 令$Z_i = Y_i - X_i$, $i = 1, 2, \cdots, n$, 且记$\tilde{\mu} = b - a$, $\tilde{\sigma}^2 = \sigma_1^2 + \sigma_2^2$, 则有

$$Z_i \sim N(\tilde{\mu}, \tilde{\sigma}^2), \quad i = 1, 2, \cdots, n.$$

这就转化为单个正态总体当$\tilde{\sigma}^2$ 未知, 求其均值$\tilde{\mu}$ 的置信区间问题. 显见$\bar{Z} = \bar{Y} - \bar{X}$ 是$\tilde{\mu}$ 的一个良好的无偏估计, 枢轴变量

$$T_z = \sqrt{n}(\bar{Z} - \tilde{\mu})/S_z \sim t_{n-1},$$

此处$S^2{}_z = \frac{1}{n-1}\sum_{i=1}^n (Z_i - \bar{Z})^2$, T_z 的表达式与$\tilde{\mu} = b - a$ 有关, 但其分布与$\tilde{\mu}$ 无关, 因此取T_z为枢轴变量. 由前面已讨论过的情形的结果, 可知$\tilde{\mu} = b - a$ 的置信系数为$1 - \alpha$ 的置信区间为

$$\left[\bar{Z} - \frac{S_z}{\sqrt{n}}\, t_{n-1}(\alpha/2), \bar{Z} + \frac{S_z}{\sqrt{n}}\, t_{n-1}(\alpha/2)\right]. \tag{2.5}$$

(2) 当σ_1^2 和σ_2^2 已知时, 我们知道$\bar{Y} - \bar{X}$ 为$b - a$ 的一个良好的无偏估计, 枢轴变量

$$T_U = \frac{\bar{Y} - \bar{X} - (b - a)}{\sqrt{\sigma_1^2/m + \sigma_2^2/n}} \sim N(0, 1),$$

故有

$$P_{a,b}\left(\left|\frac{\bar{Y} - \bar{X} - (b - a)}{\sqrt{\sigma_1^2/m + \sigma_2^2/n}}\right| \le u_{\alpha/2}\right) = 1 - \alpha.$$

再用不等式的等价变形得到$b - a$ 的置信系数为$1 - \alpha$ 的置信区间为

$$\left[\bar{Y} - \bar{X} - u_{\alpha/2}\sqrt{\sigma_1^2/m + \sigma_2^2/n},\ \bar{Y} - \bar{X} + u_{\alpha/2}\sqrt{\sigma_1^2/m + \sigma_2^2/n}\,\right]. \tag{2.6}$$

(3) 当$\sigma_1^2 = \sigma_2^2 = \sigma^2$ 未知时, 令

$$\begin{aligned} S_\omega^2 &= \frac{1}{m+n-2}\Big[(m-1)S_1^2 + (n-1)S_2^2\Big] \\ &= \frac{1}{m+n-2}\left[\sum_{i=1}^m (X_i - \bar{X})^2 + \sum_{i=1}^n (Y_i - \bar{Y})^2\right], \end{aligned}$$

显然$\bar{Y} - \bar{X}$ 是$b - a$ 的无偏估计, 由推论2.4.3可知枢轴变量

$$T = \frac{\bar{Y} - \bar{X} - (b - a)}{S_\omega}\sqrt{\frac{mn}{m+n}}\ \sim\ t_{n+m-2},$$

它的表达式与$b-a$ 有关, 但其分布与$b-a$ 无关, 取T 为枢轴变量. 故有

$$P\left(\left|\frac{\bar{Y}-\bar{X}-(b-a)}{S_\omega}\right|\sqrt{\frac{mn}{m+n}} \le t_{m+n-2}(\alpha/2)\right)=1-\alpha,$$

由不等式的等价变形, 可得$b-a$ 的置信系数为$1-\alpha$ 的置信区间为

$$\left[\bar{Y}-\bar{X}-S_\omega t_{m+n-2}\left(\frac{\alpha}{2}\right)\sqrt{\frac{1}{m}+\frac{1}{n}}, \bar{Y}-\bar{X}+S_\omega t_{m+n-2}\left(\frac{\alpha}{2}\right)\sqrt{\frac{1}{m}+\frac{1}{n}}\right], \tag{2.7}$$

其中$t_{m+n-2}(\alpha/2)$ 为自由度$m+n-2$ 的t分布的上侧$\alpha/2$ 分位数.

(4) 当$\sigma_2^2/\sigma_1^2=\lambda$, λ 已知时, 基于$b-a$ 的无偏估计$\bar{Y}-\bar{X}$ 的枢轴变量及其分布为

$$T=\sqrt{\frac{mn(m+n-2)}{m\lambda+n}}\cdot\frac{\bar{Y}-\bar{X}-(b-a)}{\sqrt{Q_1^2+Q_2^2/\lambda}} \ \sim\ t_{n+m-2},$$

此处$Q_1^2=(m-1)S_1^2=\sum_{i=1}^m(X_i-\bar{X})^2$, $Q_2^2=(n-1)S_2^2=\sum_{i=1}^n(Y_i-\bar{Y})^2$. 故有

$$P\left(\sqrt{\frac{mn(m+n-2)}{m\lambda+n}}\cdot\left|\frac{\bar{Y}-\bar{X}-(b-a)}{\sqrt{Q_1^2+Q_2^2/\lambda}}\right| \le t_{m+n-2}(\alpha/2)\right)=1-\alpha.$$

由不等式的等价变形, 得到$b-a$ 的置信系数$1-\alpha$ 的置信区间为

$$\left[\bar{Y}-\bar{X}-t_{m+n-2}(\alpha/2)\sqrt{\frac{m\lambda+n}{mn(m+n-2)}}\cdot\sqrt{Q_1^2+Q_2^2/\lambda},\right.$$
$$\left.\bar{Y}-\bar{X}+t_{m+n-2}(\alpha/2)\sqrt{\frac{m\lambda+n}{mn(m+n-2)}}\cdot\sqrt{Q_1^2+Q_2^2/\lambda}\right].$$

(5) 当$\sigma_1^2\ne\sigma_2^2$ 且皆未知时, 要求$b-a$ 的置信区间问题. 这是著名的*Behrens-Fisher* 问题. 它是Behrens 在1929年从实际应用提出的问题, 它的几种特殊情况如上所述, 已获圆满解决, 但一般情况至今还有文献在讨论. Fisher首先研究了这个问题, 并对一般情况给出近似解法. 随后许多著名统计学家, 如Scheffe 和Welch 等也研究过这个问题. 至今还得不出简单、精确的解法. 只提出一些近似的解法. 下面给出两种近似结果.

(i) 当m 与n 都充分大时可用大样本方法, 由于

$$\frac{\bar{Y}-\bar{X}-(b-a)}{\sqrt{\sigma_1^2/m+\sigma_2^2/n}}\sim N(0,1), \tag{2.8}$$

且$S_1^2\xrightarrow{P}\sigma_1^2$, $S_2^2\xrightarrow{P}\sigma_2^2$, 将(2.8)中的$\sigma_1^2$ 和σ_2^2 分别用S_1^2 和S_2^2 代入, 得

$$\frac{\bar{Y}-\bar{X}-(b-a)}{\sqrt{S_1^2/m+S_2^2/n}}\xrightarrow{\mathscr{L}}N(0,1).$$

因此, 当m, n 充分大时$b-a$ 的置信系数近似为$1-\alpha$ 的置信区间是

$$\left[\bar{Y}-\bar{X}-u_{\alpha/2}\sqrt{S_1^2/m+S_2^2/n},\ \bar{Y}-\bar{X}+u_{\alpha/2}\sqrt{S_1^2/m+S_2^2/n}\right]. \tag{2.9}$$

(ii) 一般情形, 即m 和n 都不是充分大的情形. 令

$$S_*^2=S_1^2/m+S_2^2/n.$$

取枢轴变量为

$$T = \frac{\bar{Y} - \bar{X} - (b - a)}{S_*}.$$

在一般情况下, T 已不服从t 分布. 但与具有适当自由度r 的t 分布很接近. 其中r 由下列公式确定

$$r = S_*^4 \Bigg/ \left[\frac{S_1^4}{m^2(m-1)} + \frac{S_2^4}{n^2(n-1)} \right].$$

r 一般不为整数, 可取与之最近的整数代替. 于是,我们近似地有

$$T = \frac{\bar{Y} - \bar{X} - (b - a)}{S_*} \ \sim \ t_r.$$

运用与前面类似步骤可得$b - a$ 的置信系数近似为$1 - \alpha$ 的置信区间为

$$\left[\bar{Y} - \bar{X} - S_* \cdot t_r(\alpha/2),\ \bar{Y} - \bar{X} + S_* \cdot t_r(\alpha/2) \right]. \tag{2.10}$$

这是由Welch 在1938年给出的Behrens-Fisher 问题的一个近似解法.

例6 某公司利用两条自动化流水线灌装矿泉水. 现从生产线上抽取样本$X_1, \cdots, X_{12}$和$Y_1, \cdots, Y_{17}$, 它们是每瓶矿泉水的体积(毫升). 算得样本均值$\bar{X} = 501.1$ 和$\bar{Y} = 499.7$; 样本方差$S_1^2 = 2.4$, $S_2^2 = 4.7$. 假设这两条流水线所装的矿泉水的体积分别服从正态分布$N(a, \sigma^2)$和$N(b, \sigma^2)$, 给定置信系数0.95 , 试求$b - a$ 的置信区间.

解 $\bar{Y} - \bar{X} = -1.4$, $S_\omega^2 = \frac{(m-1)S_1^2 + (n-1)S_2^2}{n+m-2} = \frac{11\times 2.4 + 16 \times 4.7}{12+17-2} = 3.763$, 于是$S_\omega = 1.94$, 查表求得$t_{m+n-2}(\alpha/2) = t_{27}(0.025) = 2.05$, 算得

$$S_\omega t_{m+n-2}(\alpha/2)\sqrt{\frac{1}{n} + \frac{1}{m}} = 1.94 \times 2.05 \sqrt{\frac{1}{12} + \frac{1}{17}} = 1.50.$$

因此$b - a$ 的95%的置信区间按公式(2.7)算得[-2.9, 0.1].

例7 欲比较甲、乙两种棉花品种的优劣. 现假设用它们纺出的棉纱强度服从$N(a, \sigma_1^2)$和$N(b, \sigma_2^2)$, $\sigma_1^2 \neq \sigma_2^2$. 试验者从这两批棉纱中分别抽取样本$X_1, \cdots, X_{120}$ 和$Y_1, \cdots, Y_{60}$, 其均值分别为$\bar{X} = 3.32$, $\bar{Y} = 3.76$, $S_1^2 = 2.18$, $S_2^2 = 5.76$. 试给出$a - b$ 的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$\bar{X} - \bar{Y} = -0.44$, $\sqrt{S_1^2/m + S_2^2/n} = \sqrt{2.18/120 + 5.76/60} = 0.338$, 查表求得$u_{0.025} = 1.96$, 算得

$$u_{\alpha/2}\sqrt{S_1^2/m + S_2^2/n} = 1.96 \times 0.338 = 0.662,$$

按公式(2.9)可得$a - b$ 的置信系数近似为0.05的置信区间为$[-1.102, 0.222]$.

2. 方差比σ_1^2/σ_2^2 的置信区间

(i)若a 和b 已知, 记$S_1^2 = \frac{1}{m}\sum_{i=1}^m (X_i - a)^2$, $S_2^2 = \frac{1}{n}\sum_{i=1}^n (Y_i - a)^2$, 显见$mS_1^2/\sigma_1^2 \sim \chi_m^2$, $nS_2^2/\sigma_2^2 \sim \chi_n^2$, 且$S_1^2$ 为σ_1^2 的无偏估计, S_2^2 为σ_2^2 的无偏估计, 故枢轴变量可取为

$$F = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} = \frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \ \sim \ F_{m,n}.$$

F 的表达式与σ_1^2/σ_2^2 无关.找c_1 和c_2, 使得

$$P\left(c_1 \le \frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \le c_2\right) = 1 - \alpha.$$

满足上述要求的c_1 和c_2 有无穷多对, 其中存在一对使区间长度最短, 但这样一对c_1 和c_2 不但不易求得, 而且表达式复杂, 应用起来不方便. 下列方法确定的c_1 和c_2 虽然不能使置信区间的精度最高, 但c_1 和c_2 容易求得, 且表达式简单, 使用方便. 即令

$$P\left(\frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} < c_1\right) = \alpha/2, \quad P\left(\frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} > c_2\right) = \alpha/2,$$

查自由度为m, n 的F 分布上侧分位数表易得$c_2 = F_{m,n}(\alpha/2)$, $c_1 = F_{m,n}(1-\alpha/2)$, 因此有

$$P\left(F_{m,n}(1-\alpha/2) \le \frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \le F_{m,n}(\alpha/2)\right) = 1 - \alpha.$$

再利用不等式的等价变形, 得到σ_1^2/σ_2^2 的置信系数为$1-\alpha$ 的置信区间为

$$\left[\frac{S_1^2}{S_2^2} \cdot \frac{1}{F_{m,n}(\alpha/2)}, \frac{S_1^2}{S_2^2} \cdot \frac{1}{F_{m,n}(1-\alpha/2)}\right]. \tag{2.11}$$

注意到由于α 较小, $F_{m,n}(1-\alpha/2)$ 在F 分布表上查不到, 利用F 分布的如下性质

$$F_{m,n}(1-\alpha/2) = 1/F_{n,m}(\alpha/2),$$

可以将其通过查$F_{n,m}(\alpha/2)$ 算得.

(ii)若a 和b 未知, 显然$(m-1)S_1^2/\sigma_1^2 \sim \chi^2_{m-1}$, $(n-1)S_2^2/\sigma_2^2 \sim \chi^2_{n-1}$, 且$S_X^2$ 和S_2^2 分别为σ_1^2 和σ_2^2 的无偏估计, 故由推论2.4.4可知枢轴变量及其分布为

$$F = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} = \frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \ \sim \ F_{m-1,n-1}.$$

F 的表达式与σ_1^2/σ_2^2 无关. 找d_1 和d_2 , 使得

$$P\left(d_1 \le \frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \le d_2\right) = 1 - \alpha.$$

类似于(i)中确定c_1 和c_2 的理由和方法, 取$d_1 = F_{m-1,n-1}(1-\alpha/2)$, $d_2 = F_{m-1,n-1}(\alpha/2)$, 故有

$$P\left(F_{m-1,n-1}(1-\alpha/2) \le \frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \le F_{m-1,n-1}(\alpha/2)\right) = 1 - \alpha.$$

最后利用不等式的等价变形, 得到σ_1^2/σ_2^2 的置信系数为$1-\alpha$ 的置信区间为

$$\left[\frac{S_1^2}{S_2^2} \cdot \frac{1}{F_{m-1,n-1}(\alpha/2)}, \frac{S_1^2}{S_2^2} \cdot \frac{1}{F_{m-1,n-1}(1-\alpha/2)}\right]. \tag{2.12}$$

由于当α较小时(如$\alpha = 0.01$, $\alpha = 0.05$), $F_{m-1,n-1}(1-\alpha/2)$ 在F 分布表中无法查到, 利用F 分布的如下性质(第二章中的一个习题)

$$F_{m-1,n-1}(1-\alpha/2) = 1/F_{n-1,m-1}(\alpha/2),$$

可通过查$F_{n-1,m-1}(\alpha/2)$将其算得.

例8 求例7中方差比σ_1^2/σ_2^2 的置信系数为90%的置信区间.

解 $S_1^2/S_2^2 = 2.18/5.76 = 0.378$, 查表求得$F_{m-1,n-1}(\alpha/2) = F_{119,59}(0.05) = 1.47$, $F_{m-1,n-1}(1-\alpha/2) = 1/F_{n-1,m-1}(\alpha/2) = 1/F_{59,119}(0.05) = 1/1.43$, 由公式(2.12)得$\sigma_1^2/\sigma_2^2$ 的置信系数为90%的置信区间为

$$\left[\frac{S_1^2}{S_2^2}\cdot\frac{1}{F_{m-1,n-1}(\alpha/2)},\ \frac{S_1^2}{S_2^2}\cdot F_{n-1,m-1}(\alpha/2)\right] = [0.257,\ 0.541].$$

3 枢轴变量法—非正态总体参数的置信区间

利用枢轴变量法构造置信区间的方法, 在上一节已介绍, 本节将用这一方法讨论几个非正态总体参数的置信区间的问题. 若枢轴变量的精确分布容易求得, 可用小样本方法获得精确的置信区间; 若枢轴变量的精确分布不易求得, 或若其精确分布虽可以求得,但表达式很复杂, 使用不方便, 则可用枢轴变量的极限分布, 来构造有关参数的近似的置信区间.

一、小样本方法

1. 指数分布参数的置信区间

设$X_1,\cdots,X_n$ 为从指数分布$EP(\lambda)$中抽取的简单随机样本, 其密度函数为$f(x,\lambda) = \lambda e^{-\lambda x}I_{[x>0]}, \lambda>0$, 要求$\lambda$ 的置信系数为$1-\alpha$ 的置信区间.

因为$\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$ 是$1/\lambda$ 的一个无偏估计(且是UMVUE), 设想λ 的置信区间可通过$\bar{X}$表示. 枢轴变量可取位$2\lambda n\bar{X}$, 由推论2.4.5可知

$$2\lambda n\bar{X} = 2\lambda\sum_{i=1}^n X_i \sim \chi^2_{2n}. \tag{3.1}$$

确定a, b 使得

$$P(a \le 2\lambda n\bar{X} \le b) = 1-\alpha.$$

使上式成立的a, b 有很多对, 其中存在一对, 使置信区间的长度最短, 但那对a, b 表达式复杂, 不易求得, 应用上也不方便. 通常采用下列方法, 令

$$P\big(2\lambda n\bar{X} < a\big) = \alpha/2, \quad P\big(2\lambda n\bar{X} > b\big) = \alpha/2.$$

由(3.1)可知,$a = \chi^2_{2n}(1-\alpha/2)$, $b = \chi^2_{2n}(\alpha/2)$, 这样找到的$a$, b 虽不能使置信区间的精确度最高, 但表达式简单, 可通过查χ^2 分布的上侧α 分位数表求得, 应用上很方便. 因此有

$$P\left(\chi^2_{2n}(1-\alpha/2) \le 2\lambda n\bar{X} \le \chi^2_{2n}(\alpha/2)\right) = 1-\alpha.$$

利用不等式的等价变形, 可得λ 的置信系数为$1-\alpha$ 的置信区间为

$$\left[\chi^2_{2n}(1-\alpha/2)/2n\bar{X},\ \chi^2_{2n}(\alpha/2)/2n\bar{X}\right]. \tag{3.2}$$

同理可求得λ 的置信系数为$1-\alpha$ 的置信下、上限分别为:

$$\chi^2_{2n}(1-\alpha)/2n\bar{X},\ \ 和\ \ \chi^2_{2n}(\alpha)/2n\bar{X}. \tag{3.3}$$

例9 设某电子产品的寿命服从指数分布$EP(\lambda)$, 现从此分布的一批产品中抽取容量为9的样本, 测得寿命为(单位: 千小时)

$$15,\ 45,\ 50,\ 53,\ 60,\ 65,\ 70,\ 83,\ 90$$

求平均寿命$1/\lambda$ 的置信系数为90%的置信区间和置信上、下限.

解 $n=9$,由样本算得：$\bar{X}=59,\ 2n\bar{X}=1062$. 查表求得:

$$\chi^2_{18}(0.05)=28.869,\quad \chi^2_{18}(0.95)=9.390,$$

$$\chi^2_{18}(0.10)=25.989,\quad \chi^2_{18}(0.90)=10.865,$$

则$g(\lambda)=1/\lambda$ 的置信系数为90%的置信区间由下式确定

$$P_\lambda\left(\chi^2_{18}(0.95)\le 2n\lambda\bar{X}\le \chi^2_{18}(0.05)\right)=0.90\,.$$

解括号内不等式得$g(\lambda)=1/\lambda$ 的90%的置信区间为

$$\left[2n\bar{X}/\chi^2_{18}(0.05),\ 2n\bar{X}/\chi^2_{18}(0.95)\right]=[36.787,\ 113.099].$$

类似方法求得$g(\lambda)=1/\lambda$ 的置信系数为90%的置信上、下限$\hat{g}_U$ 和$\hat{g}_L$ 分别为

$$\hat{g}_U=2n\bar{X}/\chi^2_{18}(0.90)=97.745(\text{千小时}),$$

$$\hat{g}_L=2n\bar{X}/\chi^2_{18}(0.10)=40.863(\text{千小时}).$$

2. 均匀分布参数的置信区间

设$\mathbf{X}=(X_1,\cdots,X_n)$ 为自均匀分布总体$U(0,\theta)$ 中抽取的简单随机样本, 求θ 的置信系数为$1-\alpha$ 的置信区间.

设$T(\mathbf{X})=X_{(n)}=\max\{X_1,\cdots,X_n\}$, $T(\mathbf{X})$ 为充分统计量,可知$\frac{n+1}{n}T$ 是θ 的无偏估计(也是UMVUE的). 设想枢轴变量一定与T 有关. 由于$X_i/\theta\sim U(0,1), i=1,2,\cdots,n$, 取$T/\theta$为枢轴变量，其表达式与$\theta$有关,但其分布与$\theta$ 无关, 其密度函数为

$$T/\theta\sim f(t)=nt^{n-1}I_{[0<t<1]}. \tag{3.4}$$

确定$c_1,\ c_2,\ 0<c_1<c_2\le 1$ 使得

$$\begin{aligned}P_\theta\ (c_1<T/\theta<c_2)&=P\left(T/c_2<\theta<T/c_1\right)\\&=\int_{c_1}^{c_2}nt^{n-1}dt=c_2^n-c_1^n=1-\alpha,\end{aligned} \tag{3.5}$$

于是$\left[T/c_2,\ T/c_1\right]$为$\theta$ 的置信系数为$1-\alpha$的置信区间.要在$0<c_1<c_2\le 1$的范围内取$c_1,\ c_2$,使(3.5)成立，并使$\frac{1}{c_1}-\frac{1}{c_2}$尽可能小(以使置信区间最短). 不难证明:这要求$c_2=1,\ c_1=\sqrt[n]{\alpha}$. 因此$\theta$ 的置信系数为$1-\alpha$ 的置信区间为

$$\left[T/c_2,\ T/c_1\right]=\left[T,\ T/\sqrt[n]{\alpha}\right].$$

二、大样本方法

1. Cauchy分布位置参数的置信区间

设$X_1, \cdots, X_n$ 为从Cauchy 分布总体中抽取的简单随机样本, Cauchy 分布有密度函数

$$f(x,\theta) = \frac{1}{\pi[1+(x-\theta)^2]}, -\infty < x < +\infty, -\infty < \theta < +\infty,$$

要求位置参数θ 的置信区间.

以m_n 记$X_1, \cdots, X_n$ 的样本中位数, θ 是总体的中位数, 因此m_n 作为θ 的点估计是合适的. 由于Cauchy 分布关于θ 对称, 故$m_n - \theta$ 的分布与从$\theta = 0$ 的Cauchy 分布中抽取的大小为n 的样本中位数的分布相同. 因此$m_n - \theta$ 的分布与θ 无关, 可作为枢轴变量, 其分布密度记为$f_n(x)$. 当n 为奇数时$f_n(x)$ 的表达式利用公式(**??**)可直接写出来. 当n 为偶数时要复杂些, 但原则上求其表达式并无困难. 找到$f_n(x)$ 后, 要确定c, 使得

$$P_\theta(|m_n - \theta| \le c) = \int_{-c}^{c} f_n(x)dx = 1-\alpha, \tag{3.6}$$

由此得出θ 的置信系数为$1-\alpha$ 的置信区间为$[m_n - c, m_n + c]$.

但是, 因为$f_n(x)$ 的表达式很复杂, 要由(3.6)决定c 不容易. 但根据(**??**)式, 并注意到$f(\xi_{1/2}) = 1/\pi$, 可知当$n \to \infty$ 时有

$$\sqrt{n}(m_n - \theta) \xrightarrow{\mathscr{L}} N(0, \pi^2/4), \tag{3.7}$$

亦即$2\sqrt{n}(m_n - \theta)/\pi \xrightarrow{\mathscr{L}} N(0,1)$, 因此有

$$P\left(2\sqrt{n}|m_n - \theta|/\pi \le u_{\alpha/2}\right) \approx 1-\alpha,$$

此处$u_{\alpha/2}$ 是标准正态分布的上侧$\alpha/2$ 分位数. 利用不等式的等价变形可知, 当n 充分大时θ 的置信系数近似为$1-\alpha$ 的置信区间为

$$\left[m_n - \frac{\pi}{2\sqrt{n}}u_{\alpha/2}, m_n + \frac{\pi}{2\sqrt{n}}u_{\alpha/2}\right].$$

2. 二项分布总体参数的置信区间

设$\mathbf{X} = (X_1, \cdots, X_n)$ 为自两点分布$b(1,p)$ 中抽取的简单随机样本, 要求p 的置信区间.

令$S_n = \sum_{i=1}^n X_i$, 可知$S_n \sim b(n,p)$, 即二项分布. 当n 充分大时, 利用中心极限定理, 有

$$\frac{S_n - np}{\sqrt{np(1-p)}} \xrightarrow{\mathscr{L}} N(0,1), \quad n \to \infty\text{时}. \tag{3.8}$$

这表明当n 充分大时, 随机变量$T = (S_n - np)\big/\sqrt{np(1-p)}$ 的渐近分布是$N(0,1)$, 与未知参数p 无关. 我们取这样的T 为枢轴变量. 因此n 充分大时有

$$P(|T| \le u_{\alpha/2}) = P\left(\left|\frac{S_n - np}{\sqrt{np(1-p)}}\right| \le u_{\alpha/2}\right) \approx 1-\alpha. \tag{3.9}$$

现在要把不等式$|T| \le u_{\alpha/2}$ 改写成$c_1 \le p \le c_2$ 的形状, $c_1 = c_1(\mathbf{x})$, $c_2 = c_2(\mathbf{x})$ 与p 无关. 若能得到这样的c_1, c_2, 则由(3.9) 有

$$p(c_1(\mathbf{x}) \le p \le c_2(\mathbf{x})) = P(|T| \le u_{\alpha/2}) \approx 1-\alpha. \tag{3.10}$$

故$[c_1(\mathbf{x}), c_2(\mathbf{x})]$ 可作为p 的置信系数近似为$1-\alpha$ 的置信区间. 记$\hat{p} = S_n/n$, $\lambda = u_{\alpha/2}$ (3.9)式括号中不等式等价于

$$(\hat{p}-p)^2 \le \lambda^2 p(1-p)/n,$$

将不等式两边展开, 整理合并同类项, 得到p 的二次三项式

$$p^2(n+\lambda) - p(2n\hat{p}+\lambda^2) + n\hat{p}^2 \le 0.$$

上式左端p 的二次三项式辨别式大于0, 且平方项系数为正, 故满足上述不等式的p 介于其两个不相等的正根之间, 记这两个根为c_1, c_2, $c_1 < c_2$, 它们是

$$c_1,\ c_2 = \frac{n}{n+\lambda^2}\left[\hat{p} + \frac{\lambda^2}{2n} \pm \lambda\sqrt{\frac{\hat{p}(1-\hat{p})}{n} + \frac{\lambda^4}{4n^2}}\ \right],\ \ c_1\text{相应于“}-\text{”号}. \tag{3.11}$$

因此$[c_1(\mathbf{x}), c_2(\mathbf{x})]$ 就是p 的置信系数近似为$1-\alpha$ 的置信区间.

在实用中, 可采用下列更简单的方法: 由$\hat{p} = S_n/n \xrightarrow{P} p$ (即依概率收敛到p)及(3.8), 因此有

$$\sqrt{\frac{p(1-p)}{\hat{p}(1-\hat{p})}} \xrightarrow{P} 1, \qquad \frac{\hat{p}-p}{\sqrt{p(1-p)/n}} \xrightarrow{\mathscr{L}} N(0,1).$$

将上述两式左变相乘, 按照依分布收敛的性质, 有

$$\frac{\hat{p}-p}{\sqrt{\hat{p}(1-\hat{p})/n}} = \frac{\hat{p}-p}{\sqrt{p(1-p)/n}} \cdot \frac{\sqrt{p(1-p)}}{\sqrt{\hat{p}(1-\hat{p})}} \xrightarrow{\mathscr{L}} N(0,1),$$

即

$$\frac{\hat{p}-p}{\sqrt{\hat{p}(1-\hat{p})/n}} \xrightarrow{\mathscr{L}} N(0,1). \tag{3.12}$$

故我们可取$T = \hat{p}-p/\sqrt{\hat{p}(1-\hat{p})/n}$ 作为枢轴变量, 其极限分布与p 无关. 令

$$P\left(\left|\frac{\hat{p}-p}{\sqrt{\hat{p}(1-\hat{p})/n}}\right| \le u_{\alpha/2}\right) \approx 1-\alpha.$$

利用不等式等价变形, 得到p 的置信系数近似为$1-\alpha$ 的置信区间为

$$\left[\hat{p} - u_{\alpha/2}\cdot\sqrt{\hat{p}(1-\hat{p})/n},\ \hat{p} + u_{\alpha/2}\cdot\sqrt{\hat{p}(1-\hat{p})/n}\right]. \tag{3.13}$$

3. Poisson 分布参数的置信区间

设$\mathbf{X} = (X_1, \cdots, X_n)$ 为抽自Poisson 总体$P(\lambda)$ 的简单随机样本, 要求λ 的置信区间.

记$S_n = \sum_{i=1}^n X_i$, 则S_n 为服从参数为$n\lambda$ 的Poisson 分布, 即

$$P(S_n = k) = \frac{e^{-n\lambda}(n\lambda)^k}{k!},\ \ k = 0, 1, 2, \cdots.$$

当n 充分大时, 由中心极限定理可知

$$\frac{S_n - n\lambda}{\sqrt{n\lambda}} \xrightarrow{\mathscr{L}} N(0,1), \quad n \to \infty\text{时}, \tag{3.14}$$

将随机变量$T = (S_n - n\lambda)\big/\sqrt{n\lambda}$ 作为枢轴变量, 其极限分布与未知参数λ 无关. 令

$$P(|T| \leq u_{\alpha/2}) = P\left(\left|\frac{S_n - n\lambda}{\sqrt{n\lambda}}\right| \leq u_{\alpha/2}\right) \approx 1 - \alpha.$$

仿照前一段的方法得到λ 的置信区间$[d_1, d_2]$,其中$d_1 = d_1(\mathbf{x})$, $d_2 = d_2(\mathbf{x})$ 为二次方程

$$(\hat{\lambda} - \lambda)^2 = \lambda u_{\alpha/2}/n$$

的两根, 此处$\hat{\lambda} = S_n/n$. 即有

$$d_1,\ d_2 = \hat{\lambda} + u_{\alpha/2}/(2n) \pm u_{\alpha/2}\sqrt{\frac{u_{\alpha/2}^2}{4n^2} + \frac{\hat{\lambda}}{n}},\ \ d_1\text{相应于“}-\text{”号}. \tag{3.15}$$

因此λ 的置信系数近似为$1-\alpha$ 的置信区间为$[d_1(\mathbf{x}), d_2(\mathbf{x})]$.

实用上, 可采用下列更简单的方法: 类似于上一段获得极限分布(3.12), 的原理, 可知

$$\left(\hat{\lambda} - \lambda\right)\Big/\sqrt{\hat{\lambda}/n} \xrightarrow{\mathscr{L}} N(0,1). \tag{3.16}$$

令$T = \left(\hat{\lambda} - \lambda\right)\Big/\sqrt{\hat{\lambda}/n}$ 为枢轴变量, 其极限分布与未知参数λ 无关. 给定置信系数$1-\alpha$, 则有

$$P\left(\left|\hat{\lambda} - \lambda\right|\Big/\sqrt{\hat{\lambda}/n} \leq u_{\alpha/2}\right) \approx 1 - \alpha.$$

由不等式的等价变形, 得到λ 的置信系数近似为$1-\alpha$ 的置信区间为

$$\left[\hat{\lambda} - u_{\alpha/2}\sqrt{\hat{\lambda}/n},\ \hat{\lambda} + u_{\alpha/2}\sqrt{\hat{\lambda}/n}\,\right].$$

4. 一般情形

用渐近分布寻求未知参数的近似置信区间在很多场合是切实可行的. 例如, 在§3.3 中, 我们曾在一般条件下证明了θ 的极大似然估计$\hat{\theta}^* = \hat{\theta}^*(X_1, \cdots, X_n)$ 有渐近正态分布: $\sqrt{n}(\hat{\theta}^* - \theta) \xrightarrow{\mathscr{L}} N(0, 1/I(\theta))$, 即$\hat{\theta}^*$渐近服从$N(\theta, \sigma^2(\theta))$. 此处$\sigma^2(\theta) = 1\big/[nI(\theta)]$, 其中

$$I(\theta) = E_\theta\left[\left(\frac{\partial \log f(x,\theta)}{\partial\theta}\right)^2\right]$$

是Fisher 信息量, $f(x,\theta)$是总体的密度函数. 当n很大时, 如果$\sigma^2(\hat{\theta}^*)$ 依概率收敛到$\sigma^2(\theta)$, 我们常用$\hat{\theta}^*$去代替$\sigma^2(\theta)$ 中的θ, 由此可得θ的置信系数近似为$1-\alpha$的置信区间

$$\left[\hat{\theta}^* - u_{\alpha/2}\cdot\sigma(\hat{\theta}^*),\ \hat{\theta}^* + u_{\alpha/2}\cdot\sigma(\hat{\theta}^*)\right].$$

此处$u_{\alpha/2}$是标准正态分布的上侧$\alpha/2$分位数.

5. 非参数情形

设某总体有均值θ, 方差σ^2, θ 和σ^2 皆未知. 从这一总体中抽取简单随机样本$X_1, \cdots, X_n$, 要求θ 的置信区间.

因为对总体分布没有作任何假定, 这是非参数型的分布族. 要找出适合小样本情形的枢轴变量是不可能的. 但是, 若n 充分大, 则由中心极限定理可知$\sqrt{n}(\bar{X}-\theta)/\sigma \xrightarrow{\mathscr{L}} N(0,1)$. 但此处σ 未知, 仍不能以$\sqrt{n}(\bar{X}-\theta)/\sigma$ 作为枢轴变量. 因为n 充分大, 样本标准差S 是σ 的一个相合估计, 故可近似地用S 代替σ, 用类似获得极限分布(3.12)的方法, 可证

$$\sqrt{n}(\bar{X}-\theta)/S \xrightarrow{\mathscr{L}} N(0,1).$$

此时可将$T=\sqrt{n}(\bar{X}-\theta)/S$ 作为枢轴变量, 它的极限分布与θ 无关, 令

$$P(|\sqrt{n}(\bar{X}-\theta)/S| \leq u_{\alpha/2}) \approx 1-\alpha,$$

解上式括号中的不等式, 得到θ 的置信系数近似为$1-\alpha$ 的置信区间为

$$\left[\bar{X}-\frac{S}{\sqrt{n}}u_{\alpha/2}, \bar{X}+\frac{S}{\sqrt{n}}u_{\alpha/2}\right].$$