概率论与数理统计

概率论与数理统计教研团队[1]

[1] 数学科学学院
哈尔滨工程大学

2024 年　春

大工至善
大学至真

哈爾濱工程大學
HARBIN ENGINEERING UNIVERSITY

目 录

第七章：参数估计

- 7.1 点估计
  - 矩估计
  - 极大似然估计
  - 点估计量的评选标准
- 7.2 区间估计
  - 区间估计的含义
  - 单个正态总体 $N(\mu,\sigma^2)$ 均值 μ 和方差 σ^2 区间估计
  - 两个正态总体的参数区间估计
  - 单侧置信区间
- 参数估计习题

在统计推断的很多实际问题中，经常是总体的分布函数形式已知，但总体分布中的参数是未知的，要完全确定这些参数的真值是不可能的，只能根据分布类型及抽取的样本提供的信息来对未知参数进行估计，这就是参数估计.

参数估计是统计推断的一种，从估计形式可区分为点估计与区间估计；从构造估计量的方法可区分为矩法估计、极大似然估计、最小二乘估计、贝叶斯估计等.

点估计

点估计问题的一般提法：设总体 X 的分布函数 $F(x,\theta)$（此处仅讨论一个未知参数情形，多于一个未知参数时可类似讨论）的形式已知，θ 是待估参数，$X_1,X_2,\cdots,X_n$ 是总体 X 的一个样本，$x_1,x_2,\cdots,x_n$ 是相应的一个样本值. 点估计问题就是要构造一个适当的统计量 $T(X_1,X_2,\cdots,X_n)$，用它的观察值 $T(x_1,x_2,\cdots,x_n)$ 作为参数 θ 的近似值. 称 $\hat{\theta}=T(X_1,X_2,\cdots,X_n)$ 为 θ 的点估计量，$\hat{\theta}=T(x_1,x_2,\cdots,x_n)$ 为 θ 的点估计值. 通常估计量和估计值都简记为 $\hat{\theta}$，简称为估计.

下面介绍两种常用的点估计方法：矩估计法和极大似然估计法.

矩估计

由辛钦大数定律，样本的原点矩依概率收敛到相应的总体矩，用样本矩替换总体矩，进而求出总体中未知参数的估计，基于这种思想求估计量的方法称为矩法，用矩法求得的估计称为矩法估计，简称矩估计.

矩估计

由辛钦大数定律，样本的原点矩依概率收敛到相应的总体矩，用样本矩替换总体矩，进而求出总体中未知参数的估计，基于这种思想求估计量的方法称为矩法，用矩法求得的估计称为矩法估计，简称矩估计.

总体 X 的分布中有 k 个未知参数 $\theta_1,\theta_2,\cdots,\theta_k$，且总体的 $l(l=1,2,\cdots,k)$ 阶矩 $\mu_l=E(X^l)$ 存在，$X_1,X_2,\cdots,X_n$ 是总体 X 的一个样本，样本的 $l(l=1,2,\cdots,k)$ 阶矩为 $A_l=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i^l$，则有参数 $\theta_1,\theta_2,\cdots,\theta_k$ 的方程组

$$\begin{cases}\mu_1=E(X)=\mu_1(\theta_1,\theta_2,\cdots,\theta_k)=A_1\\ \mu_2=E(X^2)=\mu_2(\theta_1,\theta_2,\cdots,\theta_k)=A_2\\ \qquad\qquad\vdots\\ \mu_k=E(X^k)=\mu_k(\theta_1,\theta_2,\cdots,\theta_k)=A_k\end{cases}$$

方程组的解 $\hat\theta_i=\hat\theta_i(X_1,X_2,\cdots,X_n)$ 称为 $\theta_i(i=1,2,\cdots,k)$ 的矩估计量.

例 1

设总体 X 的概率密度为

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

$X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本，求参数 θ 的矩估计量.

例 1

设总体 X 的概率密度为

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

$X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本，求参数 θ 的矩估计量.

解：因为

$$\mu_1 = E(X) = \int_0^1 x\theta x^{\theta-1}\mathrm{d}x = \frac{\theta}{1+\theta}$$

例 1

设总体 X 的概率密度为

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

$X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本，求参数 θ 的矩估计量.

解：因为

$$\mu_1=E(X)=\int_0^1 x\theta x^{\theta-1}\mathrm{d}x=\frac{\theta}{1+\theta}$$

以 A_1 代替 μ_1 得方程 $\dfrac{\theta}{1+\theta}=\overline{X}$，解得参数 θ 的矩估计量为 $\hat{\theta}=\frac{\overline{X}}{1-\overline{X}}$.

例 2

设总体 X 服从参数为 λ 的泊松分布，$X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本，样本值为 $x_1, x_2, \cdots, x_n$，求参数 λ 的矩估计量.

例 2

设总体 X 服从参数为 λ 的泊松分布，$X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本，样本值为 $x_1, x_2, \cdots, x_n$，求参数 λ 的矩估计量.

解：因为 $\mu_1 = E(X) = \lambda$，以 A_1 代替 μ_1 得方程 $\lambda = \overline{X}$.

例 2

设总体 X 服从参数为 λ 的泊松分布，$X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本，样本值为 $x_1, x_2, \cdots, x_n$，求参数 λ 的矩估计量.

解：因为 $\mu_1 = E(X) = \lambda$，以 A_1 代替 μ_1 得方程 $\lambda = \overline{X}$.

参数 λ 的矩估计量为

$$\hat{\lambda} = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

例 2

设总体 X 服从参数为 λ 的泊松分布，$X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本，样本值为 $x_1, x_2, \cdots, x_n$，求参数 λ 的矩估计量.

解：因为 $\mu_1 = E(X) = \lambda$，以 A_1 代替 μ_1 得方程 $\lambda = \overline{X}$.

参数 λ 的矩估计量为

$$\hat{\lambda} = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

参数 λ 的矩估计值为

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

例 3

设总体 X 服从区间 $[a,b]$ 上的均匀分布，$X_1,X_2,\cdots,X_n$ 是来自总体 X 的样本，求参数 a,b 的矩估计.

例 3

设总体 X 服从区间 $[a,b]$ 上的均匀分布，$X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本，求参数 a,b 的矩估计.

解： 因为总体一阶矩为 $\mu_1 = E(X) = \int_a^b \frac{x}{b-a}\mathrm{d}x = \dfrac{a+b}{2}$，

总体二阶矩为 $\mu_2 = E(X^2) = D(X) + \left[E(X)\right]^2 = \dfrac{(b-a)^2}{12} + \dfrac{(a+b)^2}{4}$.

例 3

设总体 X 服从区间 $[a,b]$ 上的均匀分布，$X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本，求参数 a, b 的矩估计.

解：因为总体一阶矩为 $\mu_1 = E(X) = \int_a^b \frac{x}{b-a}\mathrm{d}x = \dfrac{a+b}{2}$，

总体二阶矩为 $\mu_2 = E(X^2) = D(X) + \left[E(X)\right]^2 = \dfrac{(b-a)^2}{12} + \dfrac{(a+b)^2}{4}$.

令 $\mu_1 = A_1, \mu_2 = A_2$，**得方程组**

$$\begin{cases} \dfrac{a+b}{2} = A_1 \\ \dfrac{(b-a)^2}{12} + \dfrac{(a+b)^2}{4} = A_2 \end{cases}$$

方程组

$$\begin{cases} \dfrac{a+b}{2} = A_1 \\ \dfrac{(b-a)^2}{12} + \dfrac{(a+b)^2}{4} = A_2 \end{cases}$$

的解为参数 a, b 的矩估计，即

$$\hat{a} = A_1 - \sqrt{3(A_2 - A_1^2)} = \overline{X} - \sqrt{\frac{3}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}$$

$$\hat{b} = A_1 + \sqrt{3(A_2 - A_1^2)} = \overline{X} + \sqrt{\frac{3}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}$$

例 4

设总体 X 的均值 μ 和方差 σ^2 都存在，且有 $\sigma > 0$，μ 和 σ^2 均未知，$X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本，求参数 μ, σ^2 的矩估计量.

例 4

设总体 X 的均值 μ 和方差 σ^2 都存在，且有 $\sigma > 0$，μ 和 σ^2 均未知，$X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本，求参数 μ, σ^2 的矩估计量.

解：因为总体一阶矩和二阶矩为

$$\mu_1 = E(X) = \mu, \mu_2 = E(X^2) = D(X) + \left[E(X)\right]^2 = \sigma^2 + \mu^2$$

例 4

设总体 X 的均值 μ 和方差 σ^2 都存在，且有 $\sigma > 0$，μ 和 σ^2 均未知，$X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本，求参数 μ, σ^2 的矩估计量.

解：因为总体一阶矩和二阶矩为

$$\mu_1 = E(X) = \mu, \mu_2 = E(X^2) = D(X) + \left[E(X)\right]^2 = \sigma^2 + \mu^2$$

令 $\mu_1 = A_1, \mu_2 = A_2$，得方程组

$$\begin{cases} \mu = A_1 \\ \sigma^2 + \mu^2 = A_2 \end{cases}$$

例 4

设总体 X 的均值 μ 和方差 σ^2 都存在，且有 $\sigma > 0$，μ 和 σ^2 均未知，$X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本，求参数 μ, σ^2 的矩估计量.

解：因为总体一阶矩和二阶矩为

$$\mu_1 = E(X) = \mu, \mu_2 = E(X^2) = D(X) + \left[E(X)\right]^2 = \sigma^2 + \mu^2$$

令 $\mu_1 = A_1, \mu_2 = A_2$，**得方程组**

$$\begin{cases} \mu = A_1 \\ \sigma^2 + \mu^2 = A_2 \end{cases}$$

所以 μ, σ^2 **的估计量为** $\hat{\mu} = A_1 = \overline{X}, \widehat{\sigma^2} = A_2 - A_1^2 = \frac{1}{n} \sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2$.

例 4

设总体 X 的均值 μ 和方差 σ^2 都存在，且有 $\sigma > 0$，μ 和 σ^2 均未知，$X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本，求参数 μ, σ^2 的矩估计量.

解：因为总体一阶矩和二阶矩为

$$\mu_1 = E(X) = \mu, \mu_2 = E(X^2) = D(X) + \left[E(X)\right]^2 = \sigma^2 + \mu^2$$

令 $\mu_1 = A_1, \mu_2 = A_2$，得方程组

$$\begin{cases} \mu = A_1 \\ \sigma^2 + \mu^2 = A_2 \end{cases}$$

所以 μ, σ^2 的估计量为 $\hat{\mu} = A_1 = \overline{X}, \hat{\sigma^2} = A_2 - A_1^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$.

注： 不同分布类型的总体，总体均值和方差的矩估计的表达式相同.

注： 矩估计的特点

(1) 矩估计原理简单、使用方便，使用时可以不知总体的分布，在实际问题中被广泛使用；

(2) 总体原点矩不存在的分布矩估计不能用（如柯西分布等）；

(3) 矩估计只涉及总体的一些数字特征，并未用到总体的分布，因此矩估计量实际上只集中了总体的部分信息，于是在体现总体分布特征上往往性质较差，只有在样本容量较大时，才能保障它的优良性，因而理论上讲，矩估计是以大样本为应用对象的.

极大似然估计

极大似然估计是建立在极大似然原理基础上的一个统计方法，是概率论在统计学中的应用.

极大似然原理：若事件 A 发生的概率与参数 θ 有关，θ 取值不同则 $P(A)$ 也不同. 记事件 A 发生的概率为 $P(A|\theta)$. 若一次试验事件 A 发生了，可认为此时 θ 的取值是在其定义域内使得 $P(A|\theta)$ 达到最大的那一个，这就是极大似然原理.

设总体 X 是离散型随机变量，其分布律为 $P\{X=x\}=p(x;\theta)(\theta\in\Theta)$ 的形式为已知，θ 为未知参数，Θ 为 θ 的可能取值范围. 设 $X_1,X_2,\cdots,X_n$ 是来自总体 X 的样本，则样本 $(X_1,X_2,\cdots,X_n)$ 的分布律为

$$\begin{aligned}&P\{X_1=x_1,X_2=x_2,\cdots,X_n=x_n\}\\&=P\{X_1=x_1\}P\{X_2=x_2\}\cdots P\{X_n=x_n\}\\&=\prod_{i=1}^{n}p(x_i;\theta)\end{aligned}$$

对于给定的样本值 $x_1,x_2,\cdots,x_n$，样本的分布律是 θ 的函数，记作

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

称 $L(\theta)$ 为似然函数.

设总体 X 是连续型随机变量，其概率密度为 $f(x;\theta)(\theta \in \Theta)$ 的形式为已知，θ 为未知参数，Θ 为 θ 的可能取值范围. 设 $X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本，则样本 $(X_1, X_2, \cdots, X_n)$ 的概率密度为

$$\begin{aligned} f(x_1, x_2, \cdots, x_n; \theta) &= f(x_1;\theta) f(x_2;\theta) \cdots f(x_n;\theta) \\ &= \prod_{i=1}^{n} f(x_i;\theta) \end{aligned}$$

对于给定的样本值 $x_1, x_2, \cdots, x_n$，样本的概率密度是 θ 的函数，记作

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

称 $L(\theta)$ 为似然函数.

注： 对于给定的样本值 $x_1, x_2, \cdots, x_n$, 当 θ 取 θ_1 和 θ_2 两个不同的值时，若似然函数值 $L(\theta_1) > L(\theta_2)$，表明当 $\theta = \theta_1$ 时，在总体中抽取这一组指定样本值的可能性更大一些，这种可能性称之为似然，换句话说，θ 取 θ_1 比取 θ_2 有更大的似然获得样本值 $x_1, x_2, \cdots, x_n$.

定义 1

设总体 X 是离散型或者是连续型随机变量，极大似然估计就是用使似然函数 $L(\theta)$ 达到最大值的 $\hat{\theta}$ 作为 θ 的估计，即若

$$\hat{\theta} = \arg\max_{\theta \in \Theta} L(\theta)$$

则称 $\hat{\theta}$ 为 θ 的极大似然估计值.

定义 1

设总体 X 是离散型或者是连续型随机变量，极大似然估计就是用使似然函数 $L(\theta)$ 达到最大值的 $\hat{\theta}$ 作为 θ 的估计，即若

$$\hat{\theta}=\arg\max_{\theta\in\Theta} L(\theta)$$

则称 $\hat{\theta}$ 为 θ 的极大似然估计值.

显然，$\hat{\theta}$ 是样本值 $x_1, x_2, \cdots, x_n$ 的函数，即 $\hat{\theta}=\hat{\theta}(x_1, x_2, \cdots, x_n)$. 如果把样本值换为样本 $X_1, X_2, \cdots, X_n$，则有 $\hat{\theta}=\hat{\theta}(X_1, X_2, \cdots, X_n)$，称其为 θ 的极大似然估计量. 这种求未知参数估计量的方法称为极大似然估计法，亦称为最大似然估计法.

总体含有一个未知参数的极大似然估计一般步骤:

1. **根据总体分布形式给出似然函数 $L(\theta)$;**
2. **求似然函数的最大值点，最大值点即为总体未知参数的极大似然估计.**
   - **若似然函数 $L(\theta)$ 单调，最大值点在区间端点取得.**
   - **若似然函数 $L(\theta)$ 可导，由 $\frac{\mathrm{d}}{\mathrm{d}\theta}L(\theta)=0$(称为似然方程) 得到驻点，分析驻点得到最大值点. 在实践中，由于求导数的需要，往往将似然函数 $L(\theta)$ 取对数，得到对数似然函数 $\ln L(\theta)$；若对数似然函数可导，分析对数似然函数的驻点得到似然函数 $L(\theta)$ 的最大值点.**
   - **似然函数不可导, 则根据似然函数定义求解.**

总体含有多个未知参数的极大似然估计步骤与上述过程类似.

例 5

设总体 $X \sim b(1,p)$，$X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本，其样本值为 $x_1, x_2, \cdots, x_n$，求参数 p 的极大似然估计.

例 5

设总体 $X \sim b(1,p)$，$X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本，其样本值为 $x_1, x_2, \cdots, x_n$，求参数 p 的极大似然估计.

解：总体 X 的分布律为：$P\{X=x\}=p^x(1-p)^{1-x}, x=0,1$

例 5

设总体 $X \sim b(1,p)$，$X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本，其样本值为 $x_1, x_2, \cdots, x_n$，求参数 p 的极大似然估计.

解：总体 X 的分布律为：$P\{X=x\} = p^x(1-p)^{1-x}, x=0,1$

似然函数为

$$L(p) = \prod_{i=1}^{n} P\{X=x_i\} = \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} = p^{\sum\limits_{i=1}^{n} x_i}(1-p)^{n-\sum\limits_{i=1}^{n} x_i}$$

例 5

设总体 $X \sim b(1,p)$，$X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本，其样本值为 $x_1, x_2, \cdots, x_n$，求参数 p 的极大似然估计.

解：总体 X 的分布律为： $P\{X=x\} = p^x(1-p)^{1-x}, x=0,1$

似然函数为

$$L(p) = \prod_{i=1}^{n} P\{X=x_i\} = \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} = p^{\sum\limits_{i=1}^{n} x_i}(1-p)^{n-\sum\limits_{i=1}^{n} x_i}$$

对数似然函数为

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

对数似然方程为

$$\frac{\mathrm{d}}{\mathrm{d}p}\ln L(p)=\left(\sum_{i=1}^{n}x_i\right)\frac{1}{p}-\left(n-\sum_{i=1}^{n}x_i\right)\frac{1}{1-p}=0$$

对数似然方程为

$$\frac{\mathrm{d}}{\mathrm{d}p}\ln L(p)=\left(\sum_{i=1}^{n}x_i\right)\frac{1}{p}-\left(n-\sum_{i=1}^{n}x_i\right)\frac{1}{1-p}=0$$

可以验证，此对数似然函数的驻点为似然函数的最大值点，故 p 的极大似然估计值为

对数似然方程为

$$\frac{\mathrm{d}}{\mathrm{d}p}\ln L(p)=\left(\sum_{i=1}^{n}x_i\right)\frac{1}{p}-\left(n-\sum_{i=1}^{n}x_i\right)\frac{1}{1-p}=0$$

可以验证，此对数似然函数的驻点为似然函数的最大值点，故 p 的极大似然估计值为

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

对数似然方程为

$$\frac{\mathrm{d}}{\mathrm{d}p}\ln L(p)=\left(\sum_{i=1}^{n}x_i\right)\frac{1}{p}-\left(n-\sum_{i=1}^{n}x_i\right)\frac{1}{1-p}=0$$

可以验证，此对数似然函数的驻点为似然函数的最大值点，故 p 的极大似然估计值为

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

极大似然估计量为

对数似然方程为

$$\frac{\mathrm{d}}{\mathrm{d}p}\ln L(p)=\left(\sum_{i=1}^{n}x_i\right)\frac{1}{p}-\left(n-\sum_{i=1}^{n}x_i\right)\frac{1}{1-p}=0$$

可以验证，此对数似然函数的驻点为似然函数的最大值点，故 p 的极大似然估计值为

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

极大似然估计量为

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

例 6

已知总体 X 的概率密度为 $f(x;\lambda)=\begin{cases}\lambda e^{-\lambda x}, & x>0, \\ 0, & \text{其他.}\end{cases}$，其中参数 $\lambda>0$，$X_1,X_2,\cdots,X_n$ 是来自总体 X 样本，求参数 λ 的极大似然估计.

例 6

已知总体 X 的概率密度为 $f(x;\lambda)=\begin{cases}\lambda e^{-\lambda x}, & x>0,\\ 0, & \text{其他}.\end{cases}$，其中参数 $\lambda>0$，$X_1,X_2,\cdots,X_n$ 是来自总体 X 样本，求参数 λ 的极大似然估计.

解：设 $x_1,x_2,\cdots,x_n$ 是样本 $X_1,X_2,\cdots,X_n$ 的一个样本值，似然函数为

$$L(\lambda)=\prod_{i=1}^{n} f(x_i;\lambda)=\lambda^n e^{-\lambda\sum\limits_{i=1}^{n} x_i}$$

例 6

已知总体 X 的概率密度为 $f(x;\lambda)=\begin{cases}\lambda e^{-\lambda x}, & x>0,\\ 0, & \text{其他}.\end{cases}$，其中参数 $\lambda>0$，$X_1,X_2,\cdots,X_n$ 是来自总体 X 样本，求参数 λ 的极大似然估计.

解：设 $x_1,x_2,\cdots,x_n$ 是样本 $X_1,X_2,\cdots,X_n$ 的一个样本值，似然函数为

$$L(\lambda)=\prod_{i=1}^{n} f(x_i;\lambda)=\lambda^n e^{-\lambda\sum\limits_{i=1}^{n} x_i}$$

对数似然函数为

$$\ln L(\lambda)=-nln\lambda-\lambda n\overline{x}$$

对数似然方程为

$$\frac{\mathrm{d}\ln L(\lambda)}{\mathrm{d}\lambda} = -\frac{n}{\lambda} + n\overline{x} = 0$$

对数似然方程为

$$\frac{\mathrm{d}\ln L(\lambda)}{\mathrm{d}\lambda} = -\frac{n}{\lambda} + n\overline{x} = 0$$

可以验证，此对数似然函数的驻点为似然函数的最大值点，故 λ 的极大似然估计值为

对数似然方程为

$$\frac{\mathrm{d}\ln L(\lambda)}{\mathrm{d}\lambda}=-\frac{n}{\lambda}+n\overline{x}=0$$

可以验证，此对数似然函数的驻点为似然函数的最大值点，故 λ 的极大似然估计值为

$$\hat{\lambda}=\frac{1}{\overline{x}}$$

对数似然方程为

$$\frac{\mathrm{d}\ln L(\lambda)}{\mathrm{d}\lambda} = -\frac{n}{\lambda} + n\overline{x} = 0$$

可以验证，此对数似然函数的驻点为似然函数的最大值点，故 λ 的极大似然估计值为

$$\hat{\lambda} = \frac{1}{\overline{x}}$$

极大似然估计量为

对数似然方程为

$$\frac{\mathrm{d}\ln L(\lambda)}{\mathrm{d}\lambda}=-\frac{n}{\lambda}+n\overline{x}=0$$

可以验证，此对数似然函数的驻点为似然函数的最大值点，故 λ 的极大似然估计值为

$$\hat{\lambda}=\frac{1}{\overline{x}}$$

极大似然估计量为

$$\hat{\lambda}=\frac{1}{\overline{X}}$$

例 7

设总体 $X \sim N(\mu, \sigma^2)$，μ, σ^2 是未知参数，$X_1, X_2, \cdots, X_n$ 是来自总体 X 的一个样本，求参数 μ, σ^2 的极大似然估计.

例 7

设总体 $X \sim N(\mu, \sigma^2)$，μ, σ^2 是未知参数，$X_1, X_2, \cdots, X_n$ 是来自总体 X 的一个样本，求参数 μ, σ^2 的极大似然估计.

解：设 $x_1, x_2, \cdots, x_n$ 为样本值，总体 X 的概率密度为

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, -\infty < x < +\infty$$

例 7

设总体 $X \sim N(\mu,\sigma^2)$，μ,σ^2 是未知参数，$X_1,X_2,\cdots,X_n$ 是来自总体 X 的一个样本，求参数 μ,σ^2 的极大似然估计.

解：设 $x_1,x_2,\cdots,x_n$ 为样本值，总体 X 的概率密度为

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}},-\infty<x<+\infty$$

似然函数为

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

对数似然函数为

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

对数似然函数为

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

对数似然方程组为

$$\begin{cases} \dfrac{\partial \ln L(\mu,\sigma^2)}{\partial \mu} = \dfrac{1}{\sigma^2}\left(\sum\limits_{i=1}^{n} x_i - n\mu\right) = 0 \\ \dfrac{\partial \ln L(\mu,\sigma^2)}{\partial \sigma^2} = -\dfrac{n}{2\sigma^2} + \dfrac{1}{2(\sigma^2)^2}\sum\limits_{i=1}^{n}(x_i-\mu)^2 = 0 \end{cases}$$

对数似然函数为

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

对数似然方程组为

$$\begin{cases} \dfrac{\partial \ln L(\mu,\sigma^2)}{\partial \mu} = \dfrac{1}{\sigma^2}\left(\sum\limits_{i=1}^{n} x_i - n\mu\right) = 0 \\ \dfrac{\partial \ln L(\mu,\sigma^2)}{\partial \sigma^2} = -\dfrac{n}{2\sigma^2} + \dfrac{1}{2(\sigma^2)^2}\sum\limits_{i=1}^{n}(x_i-\mu)^2 = 0 \end{cases}$$

因此，μ,σ^2 的极大似然估计值分别为 $\hat{\mu} = \overline{x}, \hat{\sigma^2} = \dfrac{1}{n}\sum\limits_{i=1}^{n}(x_i-\overline{x})^2$

对数似然函数为

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

对数似然方程组为

$$\begin{cases} \dfrac{\partial \ln L(\mu,\sigma^2)}{\partial \mu} = \dfrac{1}{\sigma^2}\left(\sum\limits_{i=1}^{n} x_i - n\mu\right) = 0 \\ \dfrac{\partial \ln L(\mu,\sigma^2)}{\partial \sigma^2} = -\dfrac{n}{2\sigma^2} + \dfrac{1}{2(\sigma^2)^2}\sum\limits_{i=1}^{n}(x_i-\mu)^2 = 0 \end{cases}$$

因此，μ,σ^2 的极大似然估计值分别为 $\hat{\mu}=\overline{x}, \hat{\sigma^2} = \frac{1}{n}\sum\limits_{i=1}^{n}(x_i-\overline{x})^2$

μ,σ^2 的极大似然估计量分别为 $\hat{\mu}=\overline{X}, \hat{\sigma^2} = \frac{1}{n}\sum\limits_{i=1}^{n}(X_i-\overline{X})^2$

例 8

设总体 X **服从区间** $[a, b]$ **上的均匀分布，** $X_1, X_2, \cdots, X_n$ **是来自总体** X **的样本，求参数** a, b **的极大似然估计.**

例 8

设总体 X 服从区间 $[a,b]$ 上的均匀分布，$X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本，求参数 a, b 的极大似然估计.

解：总体 X 的概率密度为

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

例 8

设总体 X 服从区间 $[a,b]$ 上的均匀分布，$X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本，求参数 a, b 的极大似然估计.

解：总体 X 的概率密度为

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

对于给定的样本值 $x_1, x_2, \cdots, x_n$，似然函数为

$$L(a,b) = \prod_{i=1}^{n} f(x_i;a,b) = \frac{1}{(b-a)^n}$$

似然函数

$$L(a,b)=\prod_{i=1}^{n}f(x_i;a,b)=\frac{1}{(b-a)^n}$$

显然，$b-a$ 越小，似然函数函数值越大，即 b 应尽量小，而 a 应尽量大. 记 $x_{min}=\min(x_1,x_2,\cdots,x_n),x_{max}=\max(x_1,x_2,\cdots,x_n)$，结合条件 $a\le x_{min}\le x_{max}\le b$，当 $a=x_{min},b=x_{max}$ 时似然函数取得最大值，故参数 a,b 的极大似然估计值为

$$\hat{a}=x_{min}=\min(x_1,x_2,\cdots,x_n),\hat{b}=x_{max}=\max(x_1,x_2,\cdots,x_n)$$

参数 a,b 的极大似然估计量为

$$\hat{a}=\min(X_1,X_2,\cdots,X_n),\hat{b}=\max(X_1,X_2,\cdots,X_n)$$

极大似然估计的不变性

定理 2
若 $\hat{\theta}$ 是 θ 的极大似然估计，则对于 θ 的函数 $u=u(\theta)$，$u(\hat{\theta})$ 是 $u(\theta)$ 的极大似然估计.

例如，正态总体的参数 σ^2 的极大似然估计为

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

函数 $u=u(\sigma^2)=\sqrt{\sigma^2}$，根据极大似然估计的不变性有 $u(\hat{\sigma^2})$ 是 $u(\sigma^2)$ 的极大似然估计，即

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

是 σ 的极大似然估计.

点估计量的评选标准 — 无偏性

对于同一个参数，用不同的方法得到的估计量可能不相同，采用哪一个估计量好呢？最常用的有三个评选标准：无偏性、有效性、一致性.

点估计量的评选标准 — 无偏性

对于同一个参数，用不同的方法得到的估计量可能不相同，采用哪一个估计量好呢？最常用的有三个评选标准：无偏性、有效性、一致性.

定义 3

若估计量 $\hat{\theta}=\hat{\theta}(X_1, X_2, \cdots, X_n)$ 的数学期望存在，且对于任意的 $\theta \in \Theta$，有

$$E\left(\hat{\theta}\right)=\theta$$

则称 $\hat{\theta}$ 是 θ 的无偏估计量，简称无偏估计，否则称为是有偏估计量.

例 9

设总体 X 的 k 阶矩 $\mu_k = E\left(X^k\right)(k \geq 1)$ 存在，$X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本，证明不论总体服从什么分布，k 阶样本矩 $A_k = \dfrac{1}{n}\sum_{i=1}^{n} X_i^k$ 是 k 阶总体矩 μ_k 的无偏估计.

例 9

设总体 X 的 k 阶矩 $\mu_k = E\left(X^k\right)(k \geq 1)$ 存在，$X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本，证明不论总体服从什么分布，k 阶样本矩 $A_k = \dfrac{1}{n}\sum_{i=1}^{n} X_i^k$ 是 k 阶总体矩 μ_k 的无偏估计.

证明：令 $\hat{\mu}_k = A_k$ 是总体 k 阶矩 μ_k 的估计量，因为

例 9

设总体 X 的 k 阶矩 $\mu_k = E\left(X^k\right)(k \geq 1)$ 存在，$X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本，证明不论总体服从什么分布，k 阶样本矩 $A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$ 是 k 阶总体矩 μ_k 的无偏估计.

证明：令 $\hat{\mu}_k = A_k$ 是总体 k 阶矩 μ_k 的估计量，因为

$$E(\hat{\mu}_k) = E\left(A_k\right) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i^k\right) = \frac{1}{n}\sum_{i=1}^{n} E\left(X_i^k\right) = \mu_k$$

例 9

设总体 X 的 k 阶矩 $\mu_k = E\left(X^k\right)(k \geq 1)$ 存在，$X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本，证明不论总体服从什么分布，k 阶样本矩 $A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$ 是 k 阶总体矩 μ_k 的无偏估计.

证明：令 $\hat{\mu}_k = A_k$ 是总体 k 阶矩 μ_k 的估计量，因为

$$E(\hat{\mu}_k) = E\left(A_k\right) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i^k\right) = \frac{1}{n}\sum_{i=1}^{n} E\left(X_i^k\right) = \mu_k$$

所以 k 阶样本矩 A_k 是 k 阶总体矩 μ_k 的无偏估计.

例 9

设总体 X 的 k 阶矩 $\mu_k = E\left(X^k\right)(k \geq 1)$ 存在，$X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本，证明不论总体服从什么分布，k 阶样本矩 $A_k = \dfrac{1}{n}\sum\limits_{i=1}^{n} X_i^k$ 是 k 阶总体矩 μ_k 的无偏估计.

证明：令 $\hat{\mu}_k = A_k$ 是总体 k 阶矩 μ_k 的估计量，因为

$$E(\hat{\mu}_k) = E\left(A_k\right) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i^k\right) = \frac{1}{n}\sum_{i=1}^{n} E\left(X_i^k\right) = \mu_k$$

所以 k 阶样本矩 A_k 是 k 阶总体矩 μ_k 的无偏估计.

特别地，不论总体 X 服从什么分布，只要数学期望存在，样本均值 $\overline{X}$ 总是总体 X 的数学期望 $\mu_1 = E(X)$ 的无偏估计.

例 10

设总体 X 的均值 μ 和方差 $\sigma^2 > 0$ 存在但未知，$X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本，则 σ^2 的估计量 $\hat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$ 是有偏估计.

例 10

设总体 X 的均值 μ 和方差 $\sigma^2 > 0$ 存在但未知，$X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本，则 σ^2 的估计量 $\hat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$ 是有偏估计.

证明：因为样本方差 $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$，从而有 $\hat{\sigma^2} = \frac{n-1}{n}S^2$

例 10

设总体 X 的均值 μ 和方差 $\sigma^2>0$ 存在但未知，$X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本，则 σ^2 的估计量 $\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2$ 是有偏估计.

证明：因为样本方差 $S^2=\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2$**，从而有** $\hat{\sigma}^2=\frac{n-1}{n}S^2$

$$E(\hat{\sigma}^2)=E\left(\frac{n-1}{n}S^2\right)=\frac{n-1}{n}E(S^2)=\frac{n-1}{n}\sigma^2\neq\sigma^2$$

例 10

设总体 X 的均值 μ 和方差 $\sigma^2>0$ 存在但未知，$X_1,X_2,\cdots,X_n$ 是来自总体 X 的样本，则 σ^2 的估计量 $\hat{\sigma^2}=\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2$ 是有偏估计.

证明：因为样本方差 $S^2=\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2$，从而有 $\hat{\sigma^2}=\frac{n-1}{n}S^2$

$$E(\hat{\sigma^2})=E\left(\frac{n-1}{n}S^2\right)=\frac{n-1}{n}E(S^2)=\frac{n-1}{n}\sigma^2\neq\sigma^2$$

所以 $\hat{\sigma^2}$ 是有偏估计.

例 10

设总体 X 的均值 μ 和方差 $\sigma^2>0$ 存在但未知，$X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本，则 σ^2 的估计量 $\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2$ 是有偏估计.

证明：因为样本方差 $S^2=\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2$，从而有 $\hat{\sigma}^2=\frac{n-1}{n}S^2$

$$E(\hat{\sigma}^2)=E\left(\frac{n-1}{n}S^2\right)=\frac{n-1}{n}E(S^2)=\frac{n-1}{n}\sigma^2\neq\sigma^2$$

所以 $\hat{\sigma}^2$ 是有偏估计.

注： 二阶以上的样本中心距，一般不再是总体中心距的无偏估计. 由于 $E(S^2)=\sigma^2$，因此一般选择样本方差 S^2 作为总体方差 σ^2 的估计量.

点估计量的评选标准 — 有效性

一个未知参数的估计量仅用无偏性的标准评选是不够的. 一方面，无偏估计量可能不止一个；另一方面，无偏性仅反映估计量在参数真值周围波动，而没有反映出分散程度.

因为方差是随机变量的取值与其数学期望的偏离程度的度量. 参数 θ 的两个无偏估计量为 θ_1, θ_2，若 $D(\theta_1) < D(\theta_2)$，由 θ_1 得到的估计值会更有效地反映待估参数的真值.

点估计量的评选标准 — 有效性

一个未知参数的估计量仅用无偏性的标准评选是不够的. 一方面，无偏估计量可能不止一个；另一方面，无偏性仅反映估计量在参数真值周围波动，而没有反映出分散程度.

因为方差是随机变量的取值与其数学期望的偏离程度的度量. 参数 θ 的两个无偏估计量为 θ_1, θ_2，若 $D(\theta_1) < D(\theta_2)$，由 θ_1 得到的估计值会更有效地反映待估参数的真值.

定义 4

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 θ 得无偏估计，若有

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

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

例 11

设 X_1, X_2 是来自正态总体 $N(\mu, 1)$ 的一个样本，下列三个无偏估计量 $\hat{\mu}_1 = \frac{2}{3}X_1 + \frac{1}{3}X_2, \hat{\mu}_2 = \frac{1}{4}X_1 + \frac{3}{4}X_2, \hat{\mu}_3 = \frac{1}{2}X_1 + \frac{1}{2}X_2$ 哪一个更有效?

例 11

设 X_1, X_2 是来自正态总体 $N(\mu, 1)$ 的一个样本，下列三个无偏估计量 $\hat{\mu}_1 = \frac{2}{3}X_1 + \frac{1}{3}X_2, \hat{\mu}_2 = \frac{1}{4}X_1 + \frac{3}{4}X_2, \hat{\mu}_3 = \frac{1}{2}X_1 + \frac{1}{2}X_2$ 哪一个更有效?

解：$E(\hat{\mu}_1) = E(\hat{\mu}_2) = E(\hat{\mu}_3) = \mu$，故 $\hat{\mu}_1, \hat{\mu}_2, \hat{\mu}_3$ 都是参数 μ 的无偏估计.

例 11

设 X_1, X_2 是来自正态总体 $N(\mu, 1)$ 的一个样本，下列三个无偏估计量 $\hat{\mu}_1 = \frac{2}{3}X_1 + \frac{1}{3}X_2, \hat{\mu}_2 = \frac{1}{4}X_1 + \frac{3}{4}X_2, \hat{\mu}_3 = \frac{1}{2}X_1 + \frac{1}{2}X_2$ 哪一个更有效?

解：$E(\hat{\mu}_1) = E(\hat{\mu}_2) = E(\hat{\mu}_3) = \mu$，故 $\hat{\mu}_1, \hat{\mu}_2, \hat{\mu}_3$ 都是参数 μ 的无偏估计.

$$D(\hat{\mu}_1) = D\left(\frac{2}{3}X_1 + \frac{1}{3}X_2\right) = \frac{4}{9}D(X_1) + \frac{1}{9}D(X_2) = \frac{5}{9}$$

$$D(\hat{\mu}_2) = D\left(\frac{1}{4}X_1 + \frac{3}{4}X_2\right) = \frac{1}{16}D(X_1) + \frac{9}{16}D(X_2) = \frac{5}{8}$$

$$D(\hat{\mu}_3) = D\left(\frac{1}{2}X_1 + \frac{1}{2}X_2\right) = \frac{1}{4}D(X_1) + \frac{1}{4}D(X_2) = \frac{1}{2}$$

例 11

设 X_1, X_2 是来自正态总体 $N(\mu, 1)$ 的一个样本，下列三个无偏估计量 $\hat{\mu}_1 = \frac{2}{3}X_1 + \frac{1}{3}X_2, \hat{\mu}_2 = \frac{1}{4}X_1 + \frac{3}{4}X_2, \hat{\mu}_3 = \frac{1}{2}X_1 + \frac{1}{2}X_2$ 哪一个更有效?

解：$E(\hat{\mu}_1) = E(\hat{\mu}_2) = E(\hat{\mu}_3) = \mu$，故 $\hat{\mu}_1, \hat{\mu}_2, \hat{\mu}_3$ 都是参数 μ 的无偏估计.

$$D(\hat{\mu}_1) = D\left(\frac{2}{3}X_1 + \frac{1}{3}X_2\right) = \frac{4}{9}D(X_1) + \frac{1}{9}D(X_2) = \frac{5}{9}$$

$$D(\hat{\mu}_2) = D\left(\frac{1}{4}X_1 + \frac{3}{4}X_2\right) = \frac{1}{16}D(X_1) + \frac{9}{16}D(X_2) = \frac{5}{8}$$

$$D(\hat{\mu}_3) = D\left(\frac{1}{2}X_1 + \frac{1}{2}X_2\right) = \frac{1}{4}D(X_1) + \frac{1}{4}D(X_2) = \frac{1}{2}$$

因为 $D(\hat{\mu}_3)$ 最小，所以 $\hat{\mu}_3$ 在三个无偏估计中是最有效的.

定义 5

设 $\hat{\theta} = \hat{\theta}(X_1, X_2, \cdots, X_n)$ 为参数 θ 的估计量，若对于任意 $\theta \in \Theta$，当 $n \to +\infty$ 时，$\hat{\theta}$ 依概率收敛于 θ，即对于任意的 $\varepsilon > 0$，有

$$\lim_{n \to +\infty} P\left\{|\hat{\theta} - \theta| < \varepsilon\right\} = 1$$

则称 $\hat{\theta}$ 是 θ 的一致估计量.

定义 5
设 $\hat{\theta}=\hat{\theta}(X_1,X_2,\cdots,X_n)$ 为参数 θ 的估计量，若对于任意 $\theta\in\Theta$，当 $n\to+\infty$ 时，$\hat{\theta}$ 依概率收敛于 θ，即对于任意的 $\varepsilon>0$，有

$$\lim_{n\to+\infty} P\left\{|\hat{\theta}-\theta|<\varepsilon\right\}=1$$

则称 $\hat{\theta}$ 是 θ 的一致估计量.

一致性估计量仅在样本容量 n 足够大时，才显示其优越性. 矩法得到的估计量一般为一致估计量; 在一定条件下，极大似然估计具有一致性.

区间估计

区间估计（interval estimate）是在点估计的基础上，给出总体参数估计的一个区间范围，该区间通常由样本统计量加减估计误差得到. 与点估计不同，进行区间估计时，根据样本统计量的抽样分布可以对样本统计量与总体参数的接近程度给出一个概率度量.

区间估计的定义

定义 6

设总体 X 的分布函数 $F(x;\theta)$ 含有一个未知参数 θ，对于给定值 $\alpha(0<\alpha<1)$，由来自总体 X 的样本 $X_1,X_2,\cdots,X_n$ 确定的两个统计量 $\underline{\theta}=\underline{\theta}(X_1,X_2,\cdots,X_n)$ 和 $\overline{\theta}=\overline{\theta}(X_1,X_2,\cdots,X_n)(\underline{\theta}<\overline{\theta})$，使得

$$P\{\underline{\theta}<\theta<\overline{\theta}\}=1-\alpha$$

则称随机区间 $(\underline{\theta},\overline{\theta})$ 是参数 θ 的置信度为 $1-\alpha$ 的置信区间. $\underline{\theta}$ 称为置信下限，$\overline{\theta}$ 称为置信上限，$1-\alpha$ 称为置信度.

对参数 θ 进行区间估计的一般步骤：

(1) **寻找一个只包含样本 $X_1, X_2, \cdots, X_n$ 和待估参数 θ，而不包含其他未知参数的函数 $W = W(X_1, X_2, \cdots, X_n; \theta)$；而且 W 的分布完全已知，并且不依赖于参数 θ 及其他未知参数，称具有这样性质的函数为枢轴量.**

对参数 θ 进行区间估计的一般步骤：

(1) 寻找一个只包含样本 $X_1, X_2, \cdots, X_n$ 和待估参数 θ，而不包含其他未知参数的函数 $W = W(X_1, X_2, \cdots, X_n; \theta)$；而且 W 的分布完全已知，并且不依赖于参数 θ 及其他未知参数，称具有这样性质的函数为枢轴量.

(2) 对于给定的置信度 $1-\alpha$, 确定两个实数 a, b, 使得

$$P\{a < W(X_1, X_2, \cdots, X_n; \theta) < b\} = 1-\alpha$$

其中 $0 < \alpha < 1, a, b$ 由确定性分布的分位点确定.

对参数 θ 进行区间估计的一般步骤:

(1) 寻找一个只包含样本 $X_1, X_2, \cdots, X_n$ 和待估参数 θ，而不包含其他未知参数的函数 $W = W(X_1, X_2, \cdots, X_n; \theta)$；而且 W 的分布完全已知，并且不依赖于参数 θ 及其他未知参数，称具有这样性质的函数为枢轴量.

(2) 对于给定的置信度 $1-\alpha$, 确定两个实数 a, b, 使得

$$P\{a < W(X_1, X_2, \cdots, X_n; \theta) < b\} = 1 - \alpha$$

其中 $0 < \alpha < 1, a, b$ 由确定性分布的分位点确定.

(3) 由 $a < W(X_1, X_2, \cdots, X_n; \theta) < b$ 得到等价形式 $\underline{\theta} < \theta < \overline{\theta}$, 得到置信度为 $1-\alpha$ 的置信区间 $(\underline{\theta}, \overline{\theta})$.

单个正态总体 $N(\mu,\sigma^2)$ 均值 μ 和方差 σ^2 区间估计

给定置信度 $1-\alpha$, $X_1,X_2,\cdots,X_n$ 为来自正态总体 $X\sim N(\mu,\sigma^2)$ 的样本, $\overline{X},S^2$ 分别为样本均值和样本方差.

1. **σ^2 已知，均值 μ 的区间估计**

单个正态总体 $N(\mu,\sigma^2)$ 均值 μ 和方差 σ^2 区间估计

给定置信度 $1-\alpha$, $X_1,X_2,\cdots,X_n$ 为来自正态总体 $X\sim N(\mu,\sigma^2)$ 的样本, $\overline{X},S^2$ 分别为样本均值和样本方差.

1. **σ^2 已知，均值 μ 的区间估计**

   (1) **因为 $\overline{X}$ 是 μ 的无偏估计，且 $\overline{X}\sim N\left(\mu,\dfrac{\sigma^2}{n}\right)$，因此选取枢轴量**

$$W=\frac{\overline{X}-\mu}{\dfrac{\sigma}{\sqrt{n}}}$$

单个正态总体 $N(\mu,\sigma^2)$ 均值 μ 和方差 σ^2 区间估计

给定置信度 $1-\alpha$, $X_1, X_2, \cdots, X_n$ 为来自正态总体 $X \sim N(\mu,\sigma^2)$ 的样本, $\overline{X}, S^2$ 分别为样本均值和样本方差.

1. σ^2 已知，均值 μ 的区间估计

(1) 因为 $\overline{X}$ 是 μ 的无偏估计，且 $\overline{X} \sim N\left(\mu, \dfrac{\sigma^2}{n}\right)$，因此选取枢轴量

$$W = \frac{\overline{X}-\mu}{\dfrac{\sigma}{\sqrt{n}}}$$

(2) 由标准正态分布的 α 分位点, 对于给定的置信度 $1-\alpha(0<\alpha<1)$，有

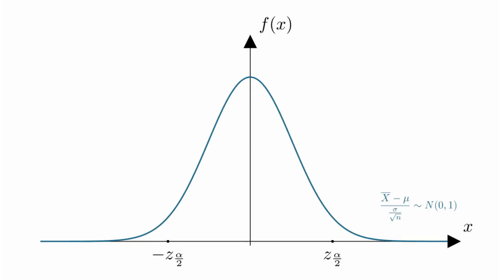
$$P\left\{-z_{\frac{\alpha}{2}} < \frac{\overline{X}-\mu}{\dfrac{\sigma}{\sqrt{n}}} < z_{\frac{\alpha}{2}}\right\} = 1-\alpha$$

(3) 等价变形得到

$$P\left\{\overline{X}-\frac{\sigma}{\sqrt{n}}z_{\frac{\alpha}{2}}<\mu<\overline{X}+\frac{\sigma}{\sqrt{n}}z_{\frac{\alpha}{2}}\right\}$$

因此，参数 μ 的置信度为 $1-\alpha$ 的置信区间为

$$\left(\overline{X}-\frac{\sigma}{\sqrt{n}}z_{\frac{\alpha}{2}},\overline{X}+\frac{\sigma}{\sqrt{n}}z_{\frac{\alpha}{2}}\right)$$



2. σ^2 未知，均值 μ 的区间估计

2. σ^2 未知，均值 μ 的区间估计

(1) 因为 S^2 是 σ^2 的无偏估计，且有 $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$，因此，进一步可构造枢轴量

$$W = \frac{\overline{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t(n-1)$$

2. σ^2 未知，均值 μ 的区间估计

(1) 因为 S^2 是 σ^2 的无偏估计，且有 $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$，因此，进一步可构造枢轴量

$$W = \frac{\overline{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t(n-1)$$

(2) 由 t 分布的 α 分位点, 对于给定的置信度 $1-\alpha(0<\alpha<1)$，有

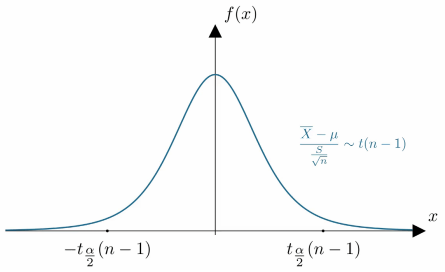
$$P\left\{-t_{\frac{\alpha}{2}}(n-1) < \frac{\overline{X} - \mu}{\frac{S}{\sqrt{n}}} < t_{\frac{\alpha}{2}}(n-1)\right\} = 1-\alpha$$

(3) 等价变形得到

$$P\left\{\overline{X}-\frac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1)<\mu<\overline{X}+\frac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1)\right\}=1-\alpha$$

因此，参数 μ 的置信度为 $1-\alpha$ 的置信区间为

$$\left(\overline{X}-\frac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1),\overline{X}+\frac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1)\right)$$



3. 方差 σ^2 的置信区间

3. 方差 σ^2 的置信区间

（1）因为 S^2 是 σ^2 的无偏估计，所以选取枢轴量

$$W=\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$$

3. 方差 σ^2 的置信区间

(1) **因为 S^2 是 σ^2 的无偏估计，所以选取枢轴量**

$$W = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

(2) **由 χ^2 分布分位点，对于给定的置信度 $1-\alpha(0<\alpha<1)$，有**

$$P\left\{\chi^2_{1-\frac{\alpha}{2}}(n-1) < \frac{(n-1)S^2}{\sigma^2} < \chi^2_{\frac{\alpha}{2}}(n-1)\right\} = 1-\alpha$$

3. 方差 σ^2 的置信区间

(1) **因为 S^2 是 σ^2 的无偏估计，所以选取枢轴量**

$$W=\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$$

(2) **由 χ^2 分布分位点，对于给定的置信度 $1-\alpha(0<\alpha<1)$，有**

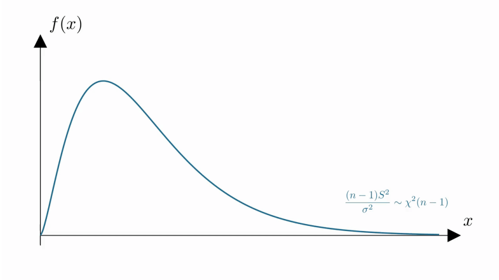
$$P\left\{\chi^2_{1-\frac{\alpha}{2}}(n-1)<\frac{(n-1)S^2}{\sigma^2}<\chi^2_{\frac{\alpha}{2}}(n-1)\right\}=1-\alpha$$

(3) **等价变形得到**

$$P\left\{\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}<\sigma^2<\frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}\right\}=1-\alpha$$

因此，参数 σ^2 的置信度为 $1-\alpha$ 的置信区间为

$$\left(\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}, \frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}\right)$$



例 1

有一大批袋装大米，现随机抽取 16 袋，称得质量如下（单位：克)：

$$
\begin{aligned}
&497, 502, 499, 496, 514, 505, 493, 496,\\
&504, 498, 497, 489, 506, 506, 499, 503
\end{aligned}
$$

设袋装大米的质量近似服从正态分布，试求总体均值 μ 和总体方差 σ^2 的置信度为 0.95 的置信区间.

例 1

有一大批袋装大米，现随机抽取 16 袋，称得质量如下（单位：克）：

$$\begin{aligned}&497, 502, 499, 496, 514, 505, 493, 496,\\&504, 498, 497, 489, 506, 506, 499, 503\end{aligned}$$

设袋装大米的质量近似服从正态分布，试求总体均值 μ 和总体方差 σ^2 的置信度为 0.95 的置信区间.

解：因为 $1-\alpha=0.95$，可知 $\alpha=0.05, \frac{\alpha}{2}=0.025, n-1=15$,
$t_{0.025}(15)=3.1315, \chi^2_{0.025}(15)=27.488, \chi^2_{0.975}(15)=6.262.$

例 1

有一大批袋装大米，现随机抽取 16 袋，称得质量如下（单位：克）：

$$\begin{aligned}&497, 502, 499, 496, 514, 505, 493, 496,\\&504, 498, 497, 489, 506, 506, 499, 503\end{aligned}$$

设袋装大米的质量近似服从正态分布，试求总体均值 μ 和总体方差 σ^2 的置信度为 0.95 的置信区间.

解：因为 $1-\alpha=0.95$，**可知** $\alpha=0.05, \frac{\alpha}{2}=0.025, n-1=15,$
$t_{0.025}(15)=3.1315, \chi^2_{0.025}(15)=27.488, \chi^2_{0.975}(15)=6.262.$

由已给数据计算可得 $\overline{x}=500.25, S=6.0388.$

均值 μ 的置信度为 0.95 的置信区间为

$$\left(\overline{X}-\frac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1),\overline{X}+\frac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1)\right)$$

代入样本值及样本标准差得 $(497.0, 503.5)$.

均值 μ 的置信度为 0.95 的置信区间为

$$\left(\overline{X}-\frac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1),\overline{X}+\frac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1)\right)$$

代入样本值及样本标准差得 $(497.0, 503.5)$.

方差 σ^2 的置信度为 0.95 的置信区间为

$$\left(\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)},\frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}\right)$$

代入样本方差得 $(19.90, 87.35)$.

两个正态总体均值差 $\mu_1 - \mu_2$ 和方差比 $\dfrac{\sigma_1^2}{\sigma_2^2}$ 的区间估计

在实际工作中，经常需要比较两个数据之间是否存在差异，例如，已知某产品的某一质量指标服从正态分布，但由于原料、设备条件、操作人员不同，或改变工艺过程等因素，引起总体均值、总体方差有所改变，这需要考虑两个正态总体均值差或方差比的估计问题. 若均值差 $\mu_1 - \mu_2$ 的置信区间包含零，则在实际应用中, 我们就认为两总体的均值没有显著差别；若方差比 $\dfrac{\sigma_1^2}{\sigma_2^2}$ 的置信区间包含 1，则在实际应用中，我们就认为两总体的方差没有显著差别. 否则，会认为具有显著差别.

两个正态总体均值差 $\mu_1-\mu_2$ 和方差比 $\dfrac{\sigma_1^2}{\sigma_2^2}$ 的区间估计

在实际工作中，经常需要比较两个数据之间是否存在差异，例如，已知某产品的某一质量指标服从正态分布，但由于原料、设备条件、操作人员不同，或改变工艺过程等因素，引起总体均值、总体方差有所改变，这需要考虑两个正态总体均值差或方差比的估计问题. 若均值差 $\mu_1-\mu_2$ 的置信区间包含零，则在实际应用中, 我们就认为两总体的均值没有显著差别；若方差比 $\dfrac{\sigma_1^2}{\sigma_2^2}$ 的置信区间包含 1，则在实际应用中，我们就认为两总体的方差没有显著差别. 否则，会认为具有显著差别.

设给定置信度为 $1-\alpha(0<\alpha<1)$，$X_1,X_2,\cdots,X_{n_1}$ 是来自总体 $N(\mu_1,\sigma_1^2)$ 的样本，$\overline{X},S_1^2$ 分别是样本均值与样本方差；$Y_1,Y_2,\cdots,Y_{n_2}$ 是来自总体 $N(\mu_2,\sigma_2^2)$ 的样本，$\overline{Y},S_2^2$ 分别是样本均值与样本方差，且这两个样本互相独立.

1. **方差 σ_1^2, σ_2^2 均为已知，求均值差 $\mu_1 - \mu_2$ 的置信区间**

1. 方差 σ_1^2, σ_2^2 均为已知，求均值差 $\mu_1 - \mu_2$ 的置信区间

（1）选取枢轴量

$$Z = \frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

1. 方差 σ_1^2, σ_2^2 均为已知，求均值差 $\mu_1-\mu_2$ 的置信区间

(1) 选取枢轴量

$$Z=\frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}\sim N(0,1)$$

(2) 对于给定的置信度 $1-\alpha(0<\alpha<1)$，有

$$P\left\{-z_{\frac{\alpha}{2}}<\frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}<z_{\frac{\alpha}{2}}\right\}=1-\alpha$$

(3) 等价变形得到

$$P\left\{\overline{X}-\overline{Y}-z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}<\mu_1-\mu_2<\overline{X}-\overline{Y}+z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right\}=1-\alpha$$

因此，均值差 $\mu_1-\mu_2$ 的置信度为 $1-\alpha$ 的置信区间为

$$\left(\overline{X}-\overline{Y}-z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}},\overline{X}-\overline{Y}+z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right)$$

(3) **等价变形得到**

$$P\left\{\overline{X}-\overline{Y}-z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}<\mu_1-\mu_2<\overline{X}-\overline{Y}+z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right\}=1-\alpha$$

因此，均值差 $\mu_1-\mu_2$ 的置信度为 $1-\alpha$ 的置信区间为

$$\left(\overline{X}-\overline{Y}-z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}},\overline{X}-\overline{Y}+z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right)$$

(3) **等价变形得到**

$$P\left\{\overline{X}-\overline{Y}-z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}<\mu_1-\mu_2<\overline{X}-\overline{Y}+z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right\}=1-\alpha$$

因此，均值差 $\mu_1-\mu_2$ 的置信度为 $1-\alpha$ 的置信区间为

$$\left(\overline{X}-\overline{Y}-z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}},\overline{X}-\overline{Y}+z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right)$$

注： **若方差 $\sigma_1^2\neq\sigma_2^2$ 未知，当样本容量 n_1,n_2 都大于 30 时，可以用 S_1,S_2 分别代替 σ_1,σ_2，得均值差 $\mu_1-\mu_2$ 的置信度为 $1-\alpha$ 的近似的置信区间近似为**

$$\left(\overline{X}-\overline{Y}-z_{\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}},\overline{X}-\overline{Y}+z_{\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}\right)$$

2. **方差 $\sigma_1^2=\sigma_2^2=\sigma^2$，$\sigma^2$ 未知，均值差 $\mu_1-\mu_2$ 的置信区间**

2. 方差 $\sigma_1^2=\sigma_2^2=\sigma^2$，$\sigma^2$ 未知，均值差 $\mu_1-\mu_2$ 的置信区间

(1) 选取枢轴量

$$T=\frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{S_W\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\sim t(n_1+n_2-2)$$

其中 $S_W=\sqrt{\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}$

2. 方差 $\sigma_1^2=\sigma_2^2=\sigma^2$，$\sigma^2$ 未知，均值差 $\mu_1-\mu_2$ 的置信区间

(1) 选取枢轴量

$$T=\frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{S_W\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\sim t(n_1+n_2-2)$$

其中 $S_W=\sqrt{\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}$

(2) 对于给定的置信度 $1-\alpha(0<\alpha<1)$，有

$$P\left\{-t_{\frac{\alpha}{2}}(n_1+n_2-2)<\frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{S_W\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}<t_{\frac{\alpha}{2}}(n_1+n_2-2)\right\}=1-\alpha$$

(3) **等价变形得到**

$$P\left\{\overline{X}-\overline{Y}-t_{\frac{\alpha}{2}}(n_1+n_2-2)S_W\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}<\mu_1-\mu_2<\right.$$

$$\left.\overline{X}-\overline{Y}+t_{\frac{\alpha}{2}}(n_1+n_2-2)S_W\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right\}=1-\alpha$$

因此，均值差 $\mu_1-\mu_2$ **的置信度为** $1-\alpha$ **的置信区间为**

$$\left(\overline{X}-\overline{Y}-t_{\frac{\alpha}{2}}(n_1+n_2-2)S_W\sqrt{\frac{1}{n_1}+\frac{1}{n_2}},\overline{X}-\overline{Y}+t_{\frac{\alpha}{2}}(n_1+n_2-2)S_W\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right)$$

3. 方差比 $\dfrac{\sigma_1^2}{\sigma_2^2}$ 的区间估计

3. 方差比 $\dfrac{\sigma_1^2}{\sigma_2^2}$ 的区间估计

（1）**选取枢轴量**

$$\frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}} \sim F(n_1-1, n_2-1)$$

3. 方差比 $\dfrac{\sigma_1^2}{\sigma_2^2}$ 的区间估计

(1) 选取枢轴量

$$\frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}} \sim F(n_1-1, n_2-1)$$

(2) 由 F 分布的分位点, 对于给定的置信度 $1-\alpha(0<\alpha<1)$，有

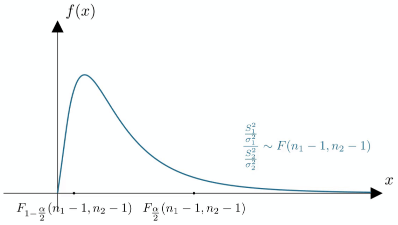
$$P\left\{F_{1-\frac{\alpha}{2}}(n_1-1, n_2-1) < \frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}} < F_{\frac{\alpha}{2}}(n_1-1, n_2-1)\right\} = 1-\alpha$$

(3) **等价变形得到**

$$P\left\{\frac{S_1^2}{S_2^2}\frac{1}{F_{\frac{\alpha}{2}}(n_1-1,n_2-1)}<\frac{\sigma_1^2}{\sigma_2^2}<\frac{S_1^2}{S_2^2}\frac{1}{F_{1-\frac{\alpha}{2}}(n_1-1,n_2-1)}\right\}=1-\alpha$$

因此，方差比 $\frac{\sigma_1^2}{\sigma_2^2}$ **的置信度为** $1-\alpha$ **的置信区间为**

$$\left(\frac{S_1^2}{S_2^2}\frac{1}{F_{\frac{\alpha}{2}}(n_1-1,n_2-1)},\frac{S_1^2}{S_2^2}\frac{1}{F_{1-\frac{\alpha}{2}}(n_1-1,n_2-1)}\right)$$



例 2

有 I，II 两种型号的子弹，为比较它们的枪口速度，设 X 表示 I 型子弹枪口速度，Y 表示 II 型子弹枪口速度. 假定 $X \sim N(\mu_1, \sigma^2), Y \sim N(\mu_2, \sigma^2)$，随机抽取 30 发子弹，其中 I 型 10 发，II 型 20 发，由实验数据计算得到

$$\overline{x} = 500(m/s), s_1 = 1.10(m/s), \overline{y} = 496(m/s), s_2 = 1.20(m/s)$$

求 $\mu_1 - \mu_2$ 的置信度为 0.95 的置信区间.

例 2

有 I，II 两种型号的子弹，为比较它们的枪口速度，设 X 表示 I 型子弹枪口速度，Y 表示 II 型子弹枪口速度. 假定 $X \sim N(\mu_1, \sigma^2), Y \sim N(\mu_2, \sigma^2)$，随机抽取 30 发子弹，其中 I 型 10 发，II 型 20 发，由实验数据计算得到

$$\overline{x} = 500(m/s), s_1 = 1.10(m/s), \overline{y} = 496(m/s), s_2 = 1.20(m/s)$$

求 $\mu_1 - \mu_2$ 的置信度为 0.95 的置信区间.

解：由已知条件 $n_1 = 10, n_2 = 20, \alpha = 0.05$ 可知 $t_{0.025}(28) = 2.0484$ 及

$$S_W = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}} = 1.1688$$

故所求 $\mu_1 - \mu_2$ 的置信度为 0.95 的置信区间为

$$\left(\overline{x}-\overline{y}-S_W\times t_{0.025}(28)\sqrt{\frac{1}{10}+\frac{1}{20}},\overline{x}-\overline{y}+S_W\times t_{0.025}(28)\sqrt{\frac{1}{10}+\frac{1}{20}}\right)$$

即 $(3.07, 4.93)$. 由于置信下限大于零，在实际应用中就认为 μ_1 比 μ_2 大.

故所求 $\mu_1-\mu_2$ 的置信度为 0.95 的置信区间为

$$\left(\overline{x}-\overline{y}-S_W\times t_{0.025}(28)\sqrt{\frac{1}{10}+\frac{1}{20}},\overline{x}-\overline{y}+S_W\times t_{0.025}(28)\sqrt{\frac{1}{10}+\frac{1}{20}}\right)$$

即 $(3.07,4.93)$. 由于置信下限大于零，在实际应用中就认为 μ_1 比 μ_2 大.

注:

(1) 若 $\mu_1-\mu_2$ 的置信区间包含零，则在实际应用中认为两个总体的均值没有显著差别.

(2) 若 $\frac{\sigma_1^2}{\sigma_2^2}$ 的置信区间包含 1，则在实际应用中认为两个总体的方差没有显著差别.

单侧置信区间

对于总体中的未知参数 θ，通过两个统计量 $\underline{\theta}$ 和 $\overline{\theta}$，给出参数 θ 的双侧置信区间. 但是在一些实际问题中，我们只关心未知参数取值的上限或下限. 例如对于设备、元件的寿命来说，寿命越长越好，所以只关心寿命的下限不少于多少；而产品的废品率当然越低越好，所以只关心废品率的上限不超过多少．这样对参数 θ 进行估计的时候，就要用到单侧置信区间.

定义 7

对于给定值 $\alpha(0 < \alpha < 1)$，若由样本$X_1, X_2, \cdots, X_n$ 确定的统计量 $\underline{\theta} = \underline{\theta}(X_1, X_2, \cdots, X_n)$ 满足

$$P\{\underline{\theta} < \theta\} = 1 - \alpha$$

则称随机区间 $(\underline{\theta}, +\infty)$ 是参数 θ 的置信度为 $1-\alpha$ 的单侧置信区间，$\underline{\theta}$ 称为单侧置信下限.

又若统计量 $\overline{\theta} = \overline{\theta}(X_1, X_2, \cdots, X_n)$，满足

$$P\{\theta < \overline{\theta}\} = 1 - \alpha$$

则称随机区间 $(-\infty, \overline{\theta})$ 为置信度为 $1-\alpha$ 的单侧置信区间，$\overline{\theta}$ 称为单侧置信上限.

正态总体 $N(\mu, \sigma^2)$ 均值 μ 的单侧置信区间

设已给定置信度为 $1-\alpha(0<\alpha<1)$，$X_1, X_2, \cdots, X_n$ 为总体 $X \sim N(\mu, \sigma^2)$ 的样本，$\overline{X}, S^2$ 分别是样本均值与样本方差.

1. **方差 σ^2 为已知，均值 μ 的单侧置信区间**

正态总体 $N(\mu, \sigma^2)$ 均值 μ 的单侧置信区间

设已给定置信度为 $1-\alpha(0<\alpha<1)$，$X_1, X_2, \cdots, X_n$ 为总体 $X \sim N(\mu, \sigma^2)$ 的样本，$\overline{X}, S^2$ 分别是样本均值与样本方差.

1. **方差 σ^2 为已知，均值 μ 的单侧置信区间**

   (1) **因为 $\overline{X}$ 是 μ 的无偏估计，且 $\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$，因此选取枢轴量**

$$W = \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$$

正态总体 $N(\mu,\sigma^2)$ 均值 μ 的单侧置信区间

设已给定置信度为 $1-\alpha(0<\alpha<1)$，$X_1, X_2, \cdots, X_n$ 为总体 $X \sim N(\mu,\sigma^2)$ 的样本，$\overline{X}, S^2$ 分别是样本均值与样本方差.

1. **方差 σ^2 为已知，均值 μ 的单侧置信区间**

   (1) **因为 $\overline{X}$ 是 μ 的无偏估计，且 $\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$，因此选取枢轴量**

$$W = \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$$

   (2) **由标准正态分布的 α 分位点, 对于给定的置信度 $1-\alpha(0<\alpha<1)$，有**

$$P\left\{\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} > -z_\alpha\right\} = 1-\alpha \quad \text{或} \quad P\left\{\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} < z_\alpha\right\} = 1-\alpha$$

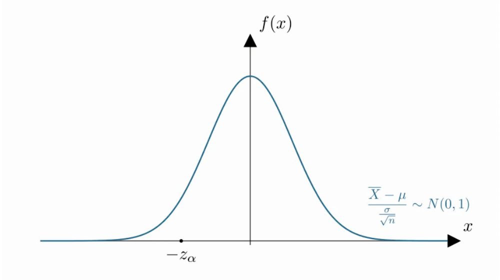
(3) 等价变形得到

$$P\left\{\mu < \overline{X} + \frac{\sigma}{\sqrt{n}} z_\alpha\right\} = 1 - \alpha \quad \text{或} \quad P\left\{\mu > \overline{X} - \frac{\sigma}{\sqrt{n}} z_\alpha\right\} = 1 - \alpha$$

因此，参数 μ 的置信度为 $1-\alpha$ 的置信区间为

$$\left(-\infty, \overline{X} + \frac{\sigma}{\sqrt{n}} z_\alpha\right) \quad \text{或} \quad \left(\overline{X} - \frac{\sigma}{\sqrt{n}} z_\alpha, +\infty\right)$$

单侧置信上限为 $\overline{X} + \frac{\sigma}{\sqrt{n}} z_\alpha$，单侧置信下限为 $\overline{X} - \frac{\sigma}{\sqrt{n}} z_\alpha$.



2. σ^2 未知，均值 μ 的单侧置信区间

2. σ^2 未知，均值 μ 的单侧置信区间

(1) 因为 S^2 是 σ^2 的无偏估计，且有 $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$，因此，进一步可构造枢轴量

$$W = \frac{\overline{X} - \mu}{\dfrac{S}{\sqrt{n}}} \sim t(n-1)$$

2. σ^2 未知，均值 μ 的单侧置信区间

(1) 因为 S^2 是 σ^2 的无偏估计，且有 $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$，因此，进一步可构造枢轴量

$$W = \frac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}} \sim t(n-1)$$

(2) 由 t 分布的 α 分位点, 对于给定的置信度 $1-\alpha(0<\alpha<1)$，有

$$P\left\{\frac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}} > -t_\alpha(n-1)\right\} = 1-\alpha \quad \text{或} \quad P\left\{\frac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}} < t_\alpha(n-1)\right\} = 1-\alpha$$

2. σ^2 未知，均值 μ 的单侧置信区间

(1) 因为 S^2 是 σ^2 的无偏估计，且有 $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$，因此，进一步可构造枢轴量

$$W = \frac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}} \sim t(n-1)$$

(2) 由 t 分布的 α 分位点, 对于给定的置信度 $1-\alpha(0<\alpha<1)$，有

$$P\left\{\frac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}} > -t_\alpha(n-1)\right\} = 1-\alpha \quad \text{或} \quad P\left\{\frac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}} < t_\alpha(n-1)\right\} = 1-\alpha$$

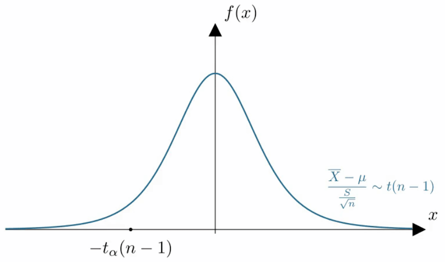
(3) 等价变形得到

$$P\left\{\mu < \overline{X} + \frac{S}{\sqrt{n}} t_\alpha(n-1)\right\} = 1-\alpha \quad \text{或} \quad P\left\{\mu > \overline{X} - \frac{S}{\sqrt{n}} t_\alpha(n-1)\right\} = 1-\alpha$$

因此，参数 μ 的置信度为 $1-\alpha$ 的单侧置信区间为

$$\left(-\infty, \overline{X}+\frac{S}{\sqrt{n}}t_{\alpha}(n-1)\right) \quad \text{或} \quad \left(\overline{X}-\frac{S}{\sqrt{n}}t_{\alpha}(n-1), +\infty\right)$$

单侧置信上限为 $\overline{X}+\dfrac{S}{\sqrt{n}}t_{\alpha}(n-1)$，单侧置信下限为 $\overline{X}-\dfrac{S}{\sqrt{n}}t_{\alpha}(n-1)$.



3. 方差 σ^2 的单侧置信区间

3. 方差 σ^2 的单侧置信区间

(1) **因为 S^2 是 σ^2 的无偏估计，所以选取枢轴量**

$$W=\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$$

3. 方差 σ^2 的单侧置信区间

(1) 因为 S^2 是 σ^2 的无偏估计，所以选取枢轴量

$$W = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

(2) 由 χ^2 分布分位点，对于给定的置信度 $1-\alpha(0<\alpha<1)$，有

$$P\left\{\frac{(n-1)S^2}{\sigma^2} < \chi^2_{\alpha}(n-1)\right\} = 1-\alpha \quad 或 \quad P\left\{\frac{(n-1)S^2}{\sigma^2} > \chi^2_{1-\alpha}(n-1)\right\} = 1-\alpha$$

3. 方差 σ^2 的单侧置信区间

(1) 因为 S^2 是 σ^2 的无偏估计，所以选取枢轴量

$$W=\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$$

(2) 由 χ^2 分布分位点，对于给定的置信度 $1-\alpha(0<\alpha<1)$，有

$$P\left\{\frac{(n-1)S^2}{\sigma^2}<\chi^2_{\alpha}(n-1)\right\}=1-\alpha \quad 或 \quad P\left\{\frac{(n-1)S^2}{\sigma^2}>\chi^2_{1-\alpha}(n-1)\right\}=1-\alpha$$

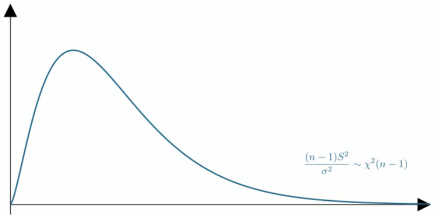
(3) 等价变形得到

$$P\left\{\sigma^2>\frac{(n-1)S^2}{\chi^2_{\alpha}(n-1)}\right\}=1-\alpha \quad 或 \quad P\left\{\sigma^2<\frac{(n-1)S^2}{\chi^2_{1-\alpha}(n-1)}\right\}=1-\alpha$$

因此，参数 σ^2 的置信度为 $1-\alpha$ 的单侧置信区间为

$$\left(\frac{(n-1)S^2}{\chi^2_{\alpha}(n-1)}, +\infty\right) \quad \text{或} \quad \left(0, \frac{(n-1)S^2}{\chi^2_{1-\alpha}(n-1)}\right)$$

单侧置信下限为 $\dfrac{(n-1)S^2}{\chi^2_{\alpha}(n-1)}$，**单侧置信上限为** $\dfrac{(n-1)S^2}{\chi^2_{1-\alpha}(n-1)}$.



例 3

从一批电子元件中随机地抽取 9 只做寿命测试，测得数据 (单位：10^3 小时）为

$$10.5, 11.2, 9.7, 12.1, 10.9, 9.6, 11.5, 9.9, 10.6$$

设电子元件寿命服从正态分布 $N(\mu, \sigma^2)$，以 95% 的把握估计这批电子元件的平均寿命不少于多少?

例 3

从一批电子元件中随机地抽取 9 只做寿命测试，测得数据 (单位：10^3 小时）为

$$10.5, 11.2, 9.7, 12.1, 10.9, 9.6, 11.5, 9.9, 10.6$$

设电子元件寿命服从正态分布 $N(\mu, \sigma^2)$，以 95% 的把握估计这批电子元件的平均寿命不少于多少?

解：因为 σ^2 未知，所以选取枢轴 $\frac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}} \sim t(n-1)$，由题意可知

$1-\alpha = 0.95, \alpha = 0.05, n = 9$，查表可得 $t_{0.05}(8) = 1.8595$，计算样本均值和样本标准差为 $\overline{x} = 10.7, S = 0.85$，所以 μ 的置信度为 0.95 的单侧置信区间为

例 3

从一批电子元件中随机地抽取 9 只做寿命测试，测得数据 (单位：10^3 小时）为

$$10.5, 11.2, 9.7, 12.1, 10.9, 9.6, 11.5, 9.9, 10.6$$

设电子元件寿命服从正态分布 $N(\mu, \sigma^2)$，以 95% 的把握估计这批电子元件的平均寿命不少于多少?

解：因为 σ^2 未知，所以选取枢轴 $\frac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}} \sim t(n-1)$，由题意可知 $1-\alpha=0.95, \alpha=0.05, n=9$，查表可得 $t_{0.05}(8)=1.8595$，计算样本均值和样本标准差为 $\overline{x}=10.7, S=0.85$，所以 μ 的置信度为 0.95 的单侧置信区间为

$$\left(\overline{x}-\frac{S}{\sqrt{n}}t_{\alpha}(n-1), +\infty\right)=(10.17\times 10^3 h, +\infty)$$

例 3

从一批电子元件中随机地抽取 9 只做寿命测试，测得数据 (单位：10^3 小时）为

$$10.5, 11.2, 9.7, 12.1, 10.9, 9.6, 11.5, 9.9, 10.6$$

设电子元件寿命服从正态分布 $N(\mu, \sigma^2)$，以 95% 的把握估计这批电子元件的平均寿命不少于多少?

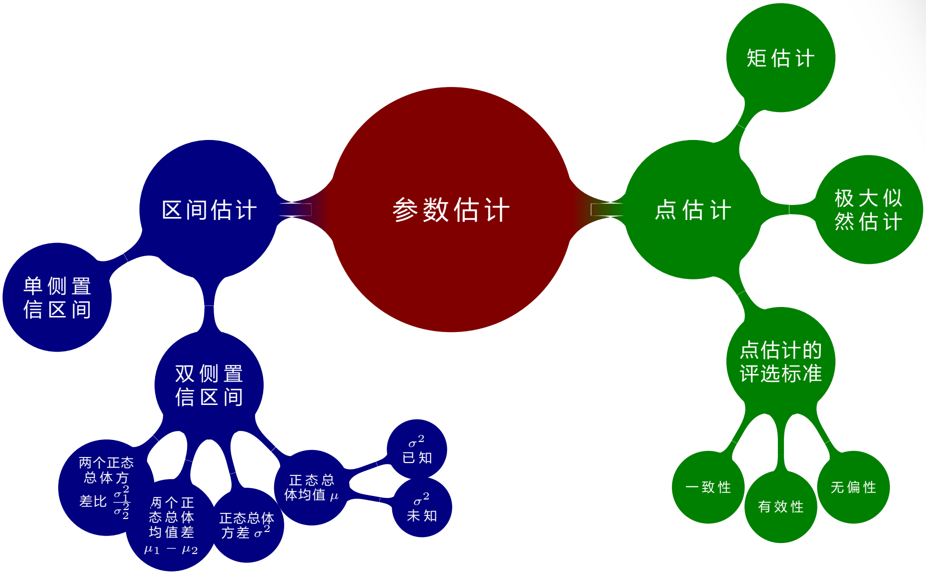
解：因为 σ^2 未知，所以选取枢轴 $\frac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}} \sim t(n-1)$，由题意可知 $1-\alpha = 0.95, \alpha = 0.05, n = 9$，查表可得 $t_{0.05}(8) = 1.8595$，计算样本均值和样本标准差为 $\overline{x} = 10.7, S = 0.85$，所以 μ 的置信度为 0.95 的单侧置信区间为

$$\left(\overline{x} - \frac{S}{\sqrt{n}} t_\alpha(n-1), +\infty\right) = (10.17 \times 10^3 h, +\infty)$$

所以以 95% 的把握估计这批电子元件的平均寿命不少于 $10.17 \times 10^3 h$

参数估计习题

知识点总结


参数估计
区间估计
单侧置信区间
双侧置信区间
两个正态总体方差比 $\frac{\sigma_1^2}{\sigma_2^2}$
两个正态总体均值差 $\mu_1-\mu_2$
正态总体方差 σ^2
正态总体均值 μ
σ^2 已知
σ^2 未知
点估计
矩估计
极大似然估计
点估计的评选标准
一致性
有效性
无偏性

例 1

设总体 X 服从参数为 λ 的泊松分布, 其中参数 λ 未知, $X_1, X_2, \cdots, X_n$ 为来自总体的样本，求 λ 的矩估计量和极大似然估计量.

例 1

设总体 X 服从参数为 λ 的泊松分布，其中参数 λ 未知，$X_1, X_2, \cdots, X_n$ 为来自总体的样本，求 λ 的矩估计量和极大似然估计量.

解：总体 X 的分布律为

$$P\{X = x\} = \frac{\lambda^x e^{-\lambda}}{x!}, x = 0, 1, \cdots$$

例 1

设总体 X 服从参数为 λ 的泊松分布，其中参数 λ 未知，$X_1, X_2, \cdots, X_n$ 为来自总体的样本，求 λ 的矩估计量和极大似然估计量.

解：总体 X 的分布律为

$$P\{X = x\} = \frac{\lambda^x e^{-\lambda}}{x!}, x = 0, 1, \cdots$$

总体一阶矩 $E(X) = \lambda$，样本一阶矩 $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$，用样本矩代替总体矩得到参数 λ 的矩估计量

例 1

设总体 X 服从参数为 λ 的泊松分布，其中参数 λ 未知，$X_1, X_2, \cdots, X_n$ 为来自总体的样本，求 λ 的矩估计量和极大似然估计量.

解：总体 X 的分布律为

$$P\{X = x\} = \frac{\lambda^x e^{-\lambda}}{x!}, x = 0, 1, \cdots$$

总体一阶矩 $E(X) = \lambda$，样本一阶矩 $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$，用样本矩代替总体矩得到参数 λ 的矩估计量

$$\hat{\lambda} = \overline{X}$$

设 $x_1, x_2, \cdots, x_n$ 为样本值，似然函数为

$$L(\lambda) = \prod_{i=1}^{n} P\{X = x_i\} = \prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!} = \frac{\lambda^{\sum_{i=1}^{n} x_i} e^{-n\lambda}}{x_1! x_2! \cdots x_n!}$$

设 $x_1, x_2, \cdots, x_n$ 为样本值，似然函数为

$$L(\lambda)=\prod_{i=1}^{n} P\{X=x_i\}=\prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!}=\frac{\lambda^{\sum_{i=1}^{n} x_i} e^{-n\lambda}}{x_1! x_2! \cdots x_n!}$$

对数似然函数为

$$\ln L(\lambda)=\ln \lambda \sum_{i=1}^{n} x_i-\ln(x_1! x_2! \cdots x_n!)-n\lambda$$

设 $x_1, x_2, \cdots, x_n$ 为样本值，似然函数为

$$L(\lambda)=\prod_{i=1}^{n} P\{X=x_i\}=\prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!}=\frac{\lambda^{\sum_{i=1}^{n} x_i} e^{-n\lambda}}{x_1! x_2! \cdots x_n!}$$

对数似然函数为

$$\ln L(\lambda)=\ln \lambda \sum_{i=1}^{n} x_i-\ln(x_1! x_2! \cdots x_n!)-n\lambda$$

对数似然方程为

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

设 $x_1, x_2, \cdots, x_n$ 为样本值，似然函数为

$$L(\lambda)=\prod_{i=1}^{n} P\{X=x_i\}=\prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!}=\frac{\lambda^{\sum_{i=1}^{n} x_i} e^{-n\lambda}}{x_1!x_2!\cdots x_n!}$$

对数似然函数为

$$\ln L(\lambda)=\ln \lambda \sum_{i=1}^{n} x_i-\ln(x_1!x_2!\cdots x_n!)-n\lambda$$

对数似然方程为

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

解得 $\lambda=\frac{1}{n} \sum_{i=1}^{n} x_i=\overline{x}$，即 λ 的极大似然估计值为 $\hat{\lambda}=\frac{1}{n} \sum_{i=1}^{n} x_i=\overline{x}$，
故 λ 的极大似然估计量 $\hat{\lambda}=\frac{1}{n} \sum_{i=1}^{n} X_i=\overline{X}$.

例 2

设总体 X 的概率密度为，

$$f(x)=\frac{1}{2\sigma}e^{-\frac{|x|}{\sigma}},-\infty<x<+\infty$$

其中 σ 是未知参数，$X_1,X_2,\cdots,X_n$ 为来自总体的样本，求 σ 的矩估计量和极大似然估计量.

例 2

设总体 X 的概率密度为,

$$f(x)=\frac{1}{2\sigma}e^{-\frac{|x|}{\sigma}},-\infty<x<+\infty$$

其中 σ 是未知参数，$X_1,X_2,\cdots,X_n$ 为来自总体的样本，求 σ 的矩估计量和极大似然估计量.

解：总体一阶矩为

$$\mu_1=E(X)=\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x=\int_{-\infty}^{+\infty}x\frac{1}{2\sigma}e^{-\frac{|x|}{\sigma}}\mathrm{d}x=0$$

总体二阶矩为

$$\mu_2 = E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x)\mathrm{d}x = \int_{-\infty}^{+\infty} x^2 \frac{1}{2\sigma} e^{-\frac{|x|}{\sigma}} \mathrm{d}x = 2\sigma^2$$

总体二阶矩为

$$\mu_2 = E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x)\mathrm{d}x = \int_{-\infty}^{+\infty} x^2 \frac{1}{2\sigma} e^{-\frac{|x|}{\sigma}} \mathrm{d}x = 2\sigma^2$$

令 $A_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$ 代替 μ_2，得到 σ 的矩估计量为

$$\hat{\sigma} = \sqrt{\frac{1}{2n}\sum_{i=1}^{n} X_i^2}$$

总体二阶矩为

$$\mu_2 = E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x)\mathrm{d}x = \int_{-\infty}^{+\infty} x^2 \frac{1}{2\sigma} e^{-\frac{|x|}{\sigma}} \mathrm{d}x = 2\sigma^2$$

令 $A_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$ **代替** μ_2**，得到** σ **的矩估计量为**

$$\hat{\sigma} = \sqrt{\frac{1}{2n}\sum_{i=1}^{n} X_i^2}$$

设 $x_1, x_2, \cdots, x_n$ **为样本值，似然函数为**

$$L(\sigma) = \prod_{i=1}^{n} f(x_i) = \prod_{i=1}^{n} \frac{1}{2\sigma} e^{-\frac{|x_i|}{\sigma}} = \frac{1}{(2\sigma)^n} e^{-\frac{1}{\sigma}\sum_{i=1}^{n}|x_i|}$$

对数似然函数为

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

对数似然函数为

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

对数似然方程为

$$\frac{\mathrm{d}\ln L(\sigma)}{\mathrm{d}\sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^2}\sum_{i=1}^{n}|x_i| = 0$$

对数似然函数为

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

对数似然方程为

$$\frac{\mathrm{d}\ln L(\sigma)}{\mathrm{d}\sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^2}\sum_{i=1}^{n}|x_i| = 0$$

则 σ 的极大似然估计值为

$$\hat{\sigma} = \frac{1}{n}\sum_{i=1}^{n}|x_i|$$

故 σ 的极大似然估计量为

$$\hat{\sigma} = \frac{1}{n}\sum_{i=1}^{n}|X_i|$$

例 3

设 $X_1, X_2, \cdots, X_n$ 为来自正态总体 $X \sim N(\mu, 1)$ 的样本，参数 μ 未知，求 $\theta = P\{X > 1\}$ 的极大似然估计值.

例 3

设 $X_1, X_2, \cdots, X_n$ 为来自正态总体 $X \sim N(\mu, 1)$ 的样本，参数 μ 未知，求 $\theta = P\{X > 1\}$ 的极大似然估计值.

解：因为总体 $X \sim N(\mu, 1)$，所以有

$$\theta = P\{X > 1\} = 1 - P\{X \le 1\} = 1 - P\left\{\frac{X-\mu}{1} \le \frac{1-\mu}{1}\right\}$$

例 3

设 $X_1, X_2, \cdots, X_n$ 为来自正态总体 $X \sim N(\mu, 1)$ 的样本，参数 μ 未知，求 $\theta = P\{X > 1\}$ 的极大似然估计值.

解：因为总体 $X \sim N(\mu, 1)$，所以有

$$\theta = P\{X > 1\} = 1 - P\{X \le 1\} = 1 - P\left\{\frac{X-\mu}{1} \le \frac{1-\mu}{1}\right\}$$

由此可知 $\theta = 1 - \Phi(1-\mu)$.

例 3

设 $X_1, X_2, \cdots, X_n$ 为来自正态总体 $X \sim N(\mu, 1)$ 的样本，参数 μ 未知，求 $\theta = P\{X > 1\}$ 的极大似然估计值.

解：因为总体 $X \sim N(\mu, 1)$，所以有

$$\theta = P\{X > 1\} = 1 - P\{X \leq 1\} = 1 - P\left\{\frac{X-\mu}{1} \leq \frac{1-\mu}{1}\right\}$$

由此可知 $\theta = 1 - \Phi(1-\mu)$.

因为正态总体均值 μ 的极大似然估计值为 $\hat{\mu} = \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$，利用极大似然估计的不变性，可知 μ 的函数的极大似然估计值，故有

例 3

设 $X_1, X_2, \cdots, X_n$ 为来自正态总体 $X \sim N(\mu, 1)$ 的样本，参数 μ 未知，求 $\theta = P\{X > 1\}$ 的极大似然估计值.

解：因为总体 $X \sim N(\mu, 1)$，所以有

$$\theta = P\{X > 1\} = 1 - P\{X \le 1\} = 1 - P\left\{\frac{X - \mu}{1} \le \frac{1 - \mu}{1}\right\}$$

由此可知 $\theta = 1 - \Phi(1 - \mu)$.

因为正态总体均值 μ 的极大似然估计值为 $\hat{\mu} = \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$，利用极大似然估计的不变性，可知 μ 的函数的极大似然估计值，故有

$$\hat{\theta} = 1 - \Phi(1 - \hat{\mu}) = 1 - \Phi(1 - \overline{x})$$

例 4

设总体 X 的概率密度为

$$f(x)=\begin{cases}\dfrac{1}{\theta}e^{-\frac{x-\mu}{\theta}}, & x\geq\mu,\\ 0, & \text{其他}.\end{cases}$$

其中 θ,μ 均是未知参数且 $\theta>0$，$X_1,X_2,\cdots,X_n$ 为来自总体 X 的样本，求 θ,μ 的矩估计量和极大似然估计量.

例 4

设总体 X 的概率密度为

$$f(x)=\begin{cases}\dfrac{1}{\theta}e^{-\frac{x-\mu}{\theta}}, & x\geq\mu,\\ 0, & \text{其他}.\end{cases}$$

其中 θ,μ 均是未知参数且 $\theta>0$，$X_1,X_2,\cdots,X_n$ 为来自总体 X 的样本，求 θ,μ 的矩估计量和极大似然估计量.

解：总体 X 的一阶与二阶矩为

$$\mu_1=E(X)=\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x=\int_{\mu}^{+\infty}x\frac{1}{\theta}e^{-\frac{x-\mu}{\theta}}\mathrm{d}x=\theta+\mu$$

$$\mu_2 = E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x) \mathrm{d}x = \int_{\mu}^{+\infty} x^2 \frac{1}{\theta} e^{-\frac{x-\mu}{\theta}} \mathrm{d}x = \theta^2 + (\theta + \mu)^2$$

$$\mu_2 = E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x)\mathrm{d}x = \int_{\mu}^{+\infty} x^2 \frac{1}{\theta} e^{-\frac{x-\mu}{\theta}} \mathrm{d}x = \theta^2 + (\theta+\mu)^2$$

用样本一阶与二阶矩替代总体一阶与二阶矩，有

$$\begin{cases} \theta + \mu = \overline{X} = \dfrac{1}{n}\sum\limits_{i=1}^{n} X_i \\ \theta^2 + (\theta+\mu)^2 = A_2 = \dfrac{1}{n}\sum\limits_{i=1}^{n} X_i^2 \end{cases}$$

$$\mu_2 = E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x)\mathrm{d}x = \int_{\mu}^{+\infty} x^2 \frac{1}{\theta} e^{-\frac{x-\mu}{\theta}} \mathrm{d}x = \theta^2 + (\theta+\mu)^2$$

用样本一阶与二阶矩替代总体一阶与二阶矩，有

$$\begin{cases} \theta + \mu = \overline{X} = \dfrac{1}{n}\sum\limits_{i=1}^{n} X_i \\ \theta^2 + (\theta+\mu)^2 = A_2 = \dfrac{1}{n}\sum\limits_{i=1}^{n} X_i^2 \end{cases}$$

解得

$$\begin{cases} \theta = \sqrt{\dfrac{1}{n}\sum\limits_{i=1}^{n} X_i^2 - \overline{X}^2} = \sqrt{\dfrac{1}{n}\sum\limits_{i=1}^{n} \left(X_i - \overline{X}\right)^2} \\ \mu = \overline{X} - \theta = \overline{X} - \sqrt{\dfrac{1}{n}\sum\limits_{i=1}^{n} \left(X_i - \overline{X}\right)^2} \end{cases}$$

故 θ, μ 的矩估计量为

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

故 θ,μ 的矩估计量为

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

设 $x_1,x_2,\cdots,x_n$ 为样本值，由总体密度可知，有 $x_i\geq\mu,i=1,2,\cdots,n$，似然函数为

$$L(\theta,\mu)=\prod_{i=1}^{n}f(x_i)=\prod_{i=1}^{n}\frac{1}{\theta}e^{-\frac{x_i-\mu}{\theta}}=\frac{1}{\theta^n}e^{-\frac{\sum\limits_{i=1}^{n}(x_i-\mu)}{\theta}}$$

故 θ, μ 的矩估计量为

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

设 $x_1, x_2, \cdots, x_n$ 为样本值，由总体密度可知，有 $x_i \geq \mu, i=1,2,\cdots,n$，似然函数为

$$L(\theta,\mu)=\prod_{i=1}^{n} f(x_i)=\prod_{i=1}^{n}\frac{1}{\theta}e^{-\frac{x_i-\mu}{\theta}}=\frac{1}{\theta^n}e^{-\frac{\sum_{i=1}^{n}(x_i-\mu)}{\theta}}$$

对数似然函数为

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

似然方程组为

$$\begin{cases} \dfrac{\partial \ln L(\theta,\mu)}{\partial \theta} = -\dfrac{n}{\theta} + \dfrac{1}{\theta^2}\left(\sum\limits_{i=1}^{n} x_i - n\mu\right) = 0 \\ \dfrac{\partial \ln L(\theta,\mu)}{\partial \mu} = \dfrac{n}{\theta} = 0 \end{cases}$$

似然方程组为

$$\begin{cases} \dfrac{\partial \ln L(\theta, \mu)}{\partial \theta} = -\dfrac{n}{\theta} + \dfrac{1}{\theta^2}\left(\sum_{i=1}^{n} x_i - n\mu\right) = 0 \\ \dfrac{\partial \ln L(\theta, \mu)}{\partial \mu} = \dfrac{n}{\theta} = 0 \end{cases}$$

此方程组无解，对数似然函数 $\ln L(\theta, \mu)$ 无驻点，由极值存在的必要条件可知，在 $\theta > 0, \mu \leq \min(x_1, x_2, \cdots, x_n)$ 区域内无极值点，$\ln L(\theta, \mu)$ 的最大值必在区域的边界上取到.

似然方程组为

$$\begin{cases} \dfrac{\partial \ln L(\theta,\mu)}{\partial \theta} = -\dfrac{n}{\theta} + \dfrac{1}{\theta^2}\left(\sum\limits_{i=1}^{n} x_i - n\mu\right) = 0 \\ \dfrac{\partial \ln L(\theta,\mu)}{\partial \mu} = \dfrac{n}{\theta} = 0 \end{cases}$$

此方程组无解，对数似然函数 $\ln L(\theta,\mu)$ 无驻点，由极值存在的必要条件可知，在 $\theta > 0, \mu \le \min(x_1, x_2, \cdots, x_n)$ 区域内无极值点，$\ln L(\theta,\mu)$ 的最大值必在区域的边界上取到.

又因为 $\frac{\partial \ln L(\theta,\mu)}{\mu} = \frac{\theta}{n} > 0$，说明 $\ln L(\theta,\mu)$ 关于 μ 单调递增，所以当 $L(\theta,\mu)$ 取最大值时必有 $\mu = \min(x_1, x_2, \cdots, x_n)$.

似然方程组为

$$\begin{cases} \dfrac{\partial \ln L(\theta,\mu)}{\partial \theta} = -\dfrac{n}{\theta} + \dfrac{1}{\theta^2}\left(\sum\limits_{i=1}^{n} x_i - n\mu\right) = 0 \\ \dfrac{\partial \ln L(\theta,\mu)}{\partial \mu} = \dfrac{n}{\theta} = 0 \end{cases}$$

此方程组无解，对数似然函数 $\ln L(\theta,\mu)$ 无驻点，由极值存在的必要条件可知，在 $\theta > 0, \mu \leq \min(x_1, x_2, \cdots, x_n)$ 区域内无极值点，$\ln L(\theta,\mu)$ 的最大值必在区域的边界上取到.

又因为 $\frac{\partial \ln L(\theta,\mu)}{\mu} = \frac{\theta}{n} > 0$，说明 $\ln L(\theta,\mu)$ 关于 μ 单调递增，所以当 $L(\theta,\mu)$ 取最大值时必有 $\mu = \min(x_1, x_2, \cdots, x_n)$.

当 $\mu = \min(x_1, x_2, \cdots, x_n)$ 时，对数似然函数为 θ 的函数

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

当 $\theta > 0$ 时，得到函数的驻点方程

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

当 $\theta > 0$ 时，得到函数的驻点方程

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

解得 $\theta = \frac{1}{n}\left(\sum_{i=1}^{n} x_i - n\mu\right) = \overline{x} - \mu$，**因此对数似然函数在**

$$\theta = \overline{x} - \mu, \mu = \min(x_1, x_2, \cdots, x_n)$$

时取最大值. 故参数 θ, μ **的极大似然估计量为**

当 $\theta > 0$ 时，得到函数的驻点方程

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

解得 $\theta = \frac{1}{n}\left(\sum_{i=1}^{n} x_i - n\mu\right) = \overline{x} - \mu$，因此对数似然函数在

$$\theta = \overline{x} - \mu, \mu = \min(x_1, x_2, \cdots, x_n)$$

时取最大值. 故参数 θ, μ 的极大似然估计量为

$$\hat{\mu} = \min(X_1, X_2, \cdots, X_n)$$

$$\hat{\theta} = \overline{X} - \hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i - \min(X_1, X_2, \cdots, X_n)$$

例 5

设总体的概率密度为

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

$X_1, X_2, \cdots, X_n$ 是来自此总体的样本, 求未知参数 θ 的极大似然估计量 $\hat{\theta}$，判断 $\hat{\theta}$ 是否为无偏估计量.

例 5

设总体的概率密度为

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

$X_1,X_2,\cdots,X_n$ 是来自此总体的样本, 求未知参数 θ 的极大似然估计量 $\hat{\theta}$，判断 $\hat{\theta}$ 是否为无偏估计量.

解：对于给定的样本值 $x_1,x_2,\cdots,x_n$，似然函数为

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

似然函数 $L(\theta)$ 为 θ 的单调递减函数，当 θ 取最小值时，$L(\theta)$ 取到最大值.

故 θ 的极大似然估计量为

故 θ 的极大似然估计量为

$$\hat{\theta} = \max(X_1, X_2, \cdots, X_n)$$

故 θ 的极大似然估计量为

$$\hat{\theta} = \max(X_1, X_2, \cdots, X_n)$$

由题意可知总体 X 的分布函数为

$$F_X(x) = \begin{cases} 0, & x < 0, \\ \dfrac{x}{\theta}, & 0 \le x \le \theta, \\ 1, & x > \theta. \end{cases}$$

故 θ 的极大似然估计量为

$$\hat{\theta} = \max(X_1, X_2, \cdots, X_n)$$

由题意可知总体 X 的分布函数为

$$F_X(x) = \begin{cases} 0, & x < 0, \\ \dfrac{x}{\theta}, & 0 \le x \le \theta, \\ 1, & x > \theta. \end{cases}$$

从而极大似然估计量的分布函数为

$$F_{\hat{\theta}}(z) = \left[F_X(z)\right]^n = \begin{cases} 0, & z < 0, \\ \left(\dfrac{z}{\theta}\right)^n, & 0 \le z \le \theta, \\ 1, & z > \theta. \end{cases}$$

因此 $\hat{\theta}$ 的概率密度为

$$f_{\hat{\theta}}(z) = F'_{\hat{\theta}}(z) = \begin{cases} \dfrac{n}{\theta}\left(\dfrac{z}{\theta}\right)^{n-1}, & 0 \leq z \leq \theta, \\ 0, & \text{其他}. \end{cases}$$

因此 $\hat{\theta}$ 的概率密度为

$$f_{\hat{\theta}}(z)=F_{\hat{\theta}}'(z)=\begin{cases}\dfrac{n}{\theta}\left(\dfrac{z}{\theta}\right)^{n-1}, & 0\le z\le\theta,\\ 0, & \text{其他.}\end{cases}$$

因为

$$E\left(\hat{\theta}\right)=\int_{-\infty}^{+\infty}zf_{\hat{\theta}}(z)\mathrm{d}z=\int_0^{\theta}n\frac{z^n}{\theta^n}\mathrm{d}z=\frac{n\theta}{n+1}\neq\theta$$

因此 $\hat{\theta}$ 的概率密度为

$$f_{\hat{\theta}}(z) = F'_{\hat{\theta}}(z) = \begin{cases} \dfrac{n}{\theta}\left(\dfrac{z}{\theta}\right)^{n-1}, & 0 \le z \le \theta, \\ 0, & \text{其他}. \end{cases}$$

因为

$$E\left(\hat{\theta}\right) = \int_{-\infty}^{+\infty} z f_{\hat{\theta}}(z)\mathrm{d}z = \int_0^{\theta} n\frac{z^n}{\theta^n}\mathrm{d}z = \frac{n\theta}{n+1} \neq \theta$$

所以参数 θ 的极大似然估计量 $\hat{\theta} = \max(X_1, X_2, \cdots, X_n)$ 是有偏的估计量.

例 6

设总体 X 服从正态分布 $X \sim N(\mu, \sigma^2)$，其中 σ^2 已知，若已知样本容量和置信度 $1-\alpha$ 均不变，则对于不同的样本观测值，总体均值 μ 的置信区间的长度 ______.

(A) 变长　　(B) 变短

(C) 不变　　(D) 不能确定

例 6

设总体 X 服从正态分布 $X \sim N(\mu, \sigma^2)$，其中 σ^2 已知，若已知样本容量和置信度 $1-\alpha$ 均不变，则对于不同的样本观测值，总体均值 μ 的置信区间的长度 ______.

(A) 变长 (B) 变短

(C) 不变 (D) 不能确定

解：当 σ^2 已知时，均值 μ 置信度为 $1-\alpha$ 的置信区间为

$$\left(\overline{X} - \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}}, \overline{X} + \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}}\right)$$

置信区间长度为 $2\frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}}$，所以与样本值无关.

例 6

设总体 X 服从正态分布 $X \sim N(\mu, \sigma^2)$，其中 σ^2 已知，若已知样本容量和置信度 $1-\alpha$ 均不变，则对于不同的样本观测值，总体均值 μ 的置信区间的长度 $\underline{\quad C \quad}$.

(A) 变长　　(B) 变短

(C) 不变　　(D) 不能确定

解：当 σ^2 已知时，均值 μ 置信度为 $1-\alpha$ 的置信区间为

$$\left(\overline{X} - \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}}, \overline{X} + \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}}\right)$$

置信区间长度为 $2\frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}}$，所以与样本值无关.

例 7

设总体 $X \sim N(\mu, \sigma^2)$，其中 μ, σ^2 均为未知参数，$X_1, X_2, \cdots, X_n$ 为来自总体的样本，若关于 μ 的置信度为 $1-\alpha$ 的置信区间长度为 L，求 $E(L^2)$

例 7

设总体 $X \sim N(\mu, \sigma^2)$，其中 μ, σ^2 均为未知参数，$X_1, X_2, \cdots, X_n$ 为来自总体的样本，若关于 μ 的置信度为 $1-\alpha$ 的置信区间长度为 L，求 $E(L^2)$

解：当 σ^2 未知时，均值 μ 置信度为 $1-\alpha$ 的置信区间为

$$\left(\overline{X} - \frac{S}{\sqrt{n}} t_{\frac{\alpha}{2}}(n-1), \overline{X} + \frac{S}{\sqrt{n}} t_{\frac{\alpha}{2}}(n-1)\right)$$

例 7

设总体 $X \sim N(\mu, \sigma^2)$，其中 μ, σ^2 均为未知参数，$X_1, X_2, \cdots, X_n$ 为来自总体的样本，若关于 μ 的置信度为 $1-\alpha$ 的置信区间长度为 L，求 $E(L^2)$

解：当 σ^2 未知时，均值 μ 置信度为 $1-\alpha$ 的置信区间为

$$\left(\overline{X} - \frac{S}{\sqrt{n}} t_{\frac{\alpha}{2}}(n-1), \overline{X} + \frac{S}{\sqrt{n}} t_{\frac{\alpha}{2}}(n-1)\right)$$

置信区间长度为 $L = \dfrac{2S}{\sqrt{n}} t_{\frac{\alpha}{2}}(n-1)$.

例 7

设总体 $X \sim N(\mu, \sigma^2)$，其中 μ, σ^2 均为未知参数，$X_1, X_2, \cdots, X_n$ 为来自总体的样本，若关于 μ 的置信度为 $1-\alpha$ 的置信区间长度为 L，求 $E(L^2)$

解：当 σ^2 未知时，均值 μ 置信度为 $1-\alpha$ 的置信区间为

$$\left(\overline{X} - \frac{S}{\sqrt{n}} t_{\frac{\alpha}{2}}(n-1), \overline{X} + \frac{S}{\sqrt{n}} t_{\frac{\alpha}{2}}(n-1)\right)$$

置信区间长度为 $L = \dfrac{2S}{\sqrt{n}} t_{\frac{\alpha}{2}}(n-1)$.

$$E\left(L^2\right) = E\left(\frac{4S^2}{n} t_{\frac{\alpha}{2}}^2(n-1)\right) = \frac{4E\left(S^2\right)}{n} t_{\frac{\alpha}{2}}^2(n-1) = \frac{4\sigma^2}{n} t_{\frac{\alpha}{2}}^2(n-1)$$

例 8

从总体 $X \sim N(\mu, 6^2)$，抽取容量为 n 的样本，要求总体均值位于区间 $(1.4, 5.4)$ 内的概率不小于 0.95，问样本容量 n 至少应取多大? $(\Phi(1.96) = 0.975)$

例 8

从总体 $X \sim N(\mu, 6^2)$，抽取容量为 n 的样本，要求总体均值位于区间 $(1.4, 5.4)$ 内的概率不小于 0.95，问样本容量 n 至少应取多大? $(\Phi(1.96) = 0.975)$

解：当 $\sigma^2 = 6^2$ 时，均值 μ 置信度为 $1 - \alpha = 0.95$ 的置信区间为

$$\left(\overline{x} - \frac{6}{\sqrt{n}} z_{\frac{\alpha}{2}}, \overline{x} + \frac{6}{\sqrt{n}} z_{\frac{\alpha}{2}}\right) = (1.4, 5.4)$$

解得 $\overline{x} = 3.4, \frac{6}{\sqrt{n} z_{\frac{\alpha}{2}}} = 2$，于是 $n = 9\left(z_{\frac{\alpha}{2}}\right)^2$.

例 8

从总体 $X \sim N(\mu, 6^2)$，抽取容量为 n 的样本，要求总体均值位于区间 $(1.4, 5.4)$ 内的概率不小于 0.95，问样本容量 n 至少应取多大? $(\Phi(1.96) = 0.975)$

解：当 $\sigma^2 = 6^2$ **时，均值** μ **置信度为** $1-\alpha = 0.95$ **的置信区间为**

$$\left(\overline{x} - \frac{6}{\sqrt{n}} z_{\frac{\alpha}{2}}, \overline{x} + \frac{6}{\sqrt{n}} z_{\frac{\alpha}{2}}\right) = (1.4, 5.4)$$

解得 $\overline{x} = 3.4, \frac{6}{\sqrt{n} z_{\frac{\alpha}{2}}} = 2$**，于是** $n = 9\left(z_{\frac{\alpha}{2}}\right)^2$.

因为总体均值 μ **位于区间** $(1.4, 5.4)$ **内的概率不小于 0.95，即** $1-\alpha \geq 0.95$**，从而** $\alpha \leq 0.05$**，进一步有** $z_{\frac{\alpha}{2}} \geq z_{\frac{0.05}{2}}$**，所以有**

$$n = 9\left(z_{\frac{\alpha}{2}}\right)^2 \geq 9\left(z_{\frac{0.05}{2}}\right)^2 \approx 34.57 \qquad n\text{至少应取}35$$

概率论与数理统计

概率论与数理统计教研团队[1]

[1] 数学科学学院
哈尔滨工程大学

2024 年 春

大工至善
大学至真

HARBIN ENGINEERING UNIVERSITY