

Ch 7. 参数估计

钟友良

zhongyl@scut.edu.cn

Outline

7.0 前言

7.1 参数的点估计

7.2 估计量优劣势的评价

7.3 参数的区间估计

Outline

7.0 前言

7.1 参数的点估计

7.2 估计量优劣势的评价

7.3 参数的区间估计

Outline

7.0 前言

7.1 参数的点估计

7.2 估计量优劣势的评价

7.3 参数的区间估计

什么是参数

总体参数基本分为下面几类

- ▶ 总体分布中所含的参数: 比如正态总体 $N(\mu, \sigma^2)$ 中的 μ 和 σ^2 。
- ▶ 总体参数的函数: 比如服从正态分布 $N(\mu, \sigma^2)$ 的总体, 其取值不超过某个给定常数 a 的概率, 即 $P(X \leq a) = \Phi(\frac{a-\mu}{\sigma})$ 。
- ▶ 分布的数字特征: 比如均值 $E[X]$, 方差 $Var[X]$ 等。

参数估计的形式

参数估计的形式有以下两种：

- ▶ 点估计。
- ▶ 区间估计。

参数的点估计

设 (X_1, X_2, \dots, X_n) 是来自总体 X 的一个容量为 n 的样本，称用于估计未知参数 θ 的统计量

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

为 θ 的 **点估计量**。

点估计的分类

本课程要求掌握的点估计方法分为以下几类

- ▶ 矩法估计
- ▶ 最大似然估计
- ▶ 顺序统计量估计

矩法估计

所谓矩法估计，就是用样本的矩来估计总体的矩。
比如对于总体 X ，其概率密度函数 $f_X(x, \theta_1, \theta_2, \dots, \theta_k)$ 含有 k 个未知参数 $\theta_1, \theta_2, \dots, \theta_k$ 。假若 $\theta_1, \theta_2, \dots, \theta_k$ 是总体前 k 阶原点矩 $E[X^i]$ ， $i = 1, 2, \dots, k$ 的函数，即

$$\begin{cases} \theta_1 = \theta_1(E[X], E[X^2], \dots, E[X^k]) \\ \theta_2 = \theta_2(E[X], E[X^2], \dots, E[X^k]) \\ \dots \\ \theta_k = \theta_k(E[X], E[X^2], \dots, E[X^k]) \end{cases}$$

Q: 如何估计 $\theta_1, \theta_2, \dots, \theta_k$?

矩法估计的具体操作

- ▶ 矩法估计的基本条件：把总体的参数表示成总体各阶 **原点矩** 的函数。
- ▶ 矩法估计的基本方法：用前 k 阶样本原点矩 $\frac{1}{n} \sum_{i=1}^n X_i^j$ 代替 X 的前 k 阶原点矩 $E[X^j]$, $j = 1, 2, \dots, k$, 得到 θ_j 的 **矩法估计量** $\hat{\theta}_j$,

$$\begin{cases} \hat{\theta}_1 = \theta_1(\frac{1}{n} \sum_{i=1}^n X_i, \frac{1}{n} \sum_{i=1}^n X_i^2, \dots, \frac{1}{n} \sum_{i=1}^n X_i^k) \\ \dots \\ \hat{\theta}_k = \theta_k(\frac{1}{n} \sum_{i=1}^n X_i, \frac{1}{n} \sum_{i=1}^n X_i^2, \dots, \frac{1}{n} \sum_{i=1}^n X_i^k) \end{cases}$$

- ▶ 对样本进行一次观测，假设 (x_1, x_2, \dots, x_n) 是样本的一个观测值，那么将其带入 $\hat{\theta}_j$ 得到的数值 $\hat{\theta}_j(x_1, x_2, \dots, x_n)$ 就是参数 θ_j 的 **矩法估计值**。

矩法估计应用

- ▶ 总体均值 $E[X]$ 的矩法估计: $\hat{\mu}_M = \bar{X}$ 。
- ▶ 一批产品合格率 (不合格率) 的估计:
由于总体 $X \sim B(1, p)$, $p = E[X]$ 。将 $E[X]$ 换成 \bar{X} , 即可得到合格率的矩法估计量

$$\hat{p}_M = \bar{X}$$

假如 $n = 100$, 其中一次抽取中有 3 件不合格品, 那么

$$\hat{p}_M = \frac{1}{100} \times 3 = 0.03.$$

即 0.03 为矩法估计值。

矩法估计应用

- ▶ 总体方差 $\sigma^2 = \text{Var}[X]$ 以及标准差 σ 的估计。
因为

$$\sigma^2 = E[X^2] - (E[X])^2, \quad \sigma = \sqrt{E[X^2] - (E[X])^2}$$

作矩法估计

$$\widehat{E[X]} = \overline{X}, \quad \widehat{E[X^2]} = \overline{X^2}$$

$\hat{=} \frac{1}{n} \sum X_i$ $\hat{=} \frac{1}{n} \sum X_i^2$

有

$$\widehat{\sigma^2} = \overline{X^2} - \overline{X}^2 = S_n^2, \quad \widehat{\sigma} = \sqrt{\overline{X^2} - \overline{X}^2} = S_n$$

总体的各阶中心矩的矩法估计量就是样本的相应阶中心矩。

例 7.1.3

设总体在某一区间上均匀取值，试用矩法估计该区间的左右端点。

解：根据题意，设总体 $X \sim U[a, b]$ ，此题要估计的就是 a 和 b 。均匀分布的数字特征我们知道，

$$E[X] = \frac{a+b}{2}, \quad \text{Var}[X] = \frac{(b-a)^2}{12}$$

例 7.1.3

反解得

$$\begin{cases} a = E[X] - \sqrt{3\text{Var}[X]} = \mu - \sqrt{3}\sigma \\ b = E[X] + \sqrt{3\text{Var}[X]} = \mu + \sqrt{3}\sigma \end{cases}$$

依据将 μ, σ 换成其相应的矩法估计量, 有

$$\hat{a}_M = \bar{X} - \sqrt{3}S_n$$

$$\hat{b}_M = \bar{X} + \sqrt{3}S_n$$

矩法估计量不唯一 (例 7.1.4)

设总体的分布密度函数为

$$f_X(x, \theta) = \frac{\theta}{2} e^{-\theta|x|}, \quad -\infty < x < \infty, \theta > 0$$

求 θ 的矩法估计量。
考虑总体的偶数阶矩，

$E[X], E[X^2], \dots$
 $E[X], E[X^2], \dots$
 $E[X^2], E[X^4], \dots$

$$E[X^{2k}] = \theta \int_0^{\infty} x^{2k} e^{-\theta x} dx = \frac{(2k)!}{\theta^{2k}}$$

即

$$\theta = \left(\frac{(2k)!}{E[X^{2k}]} \right)^{\frac{1}{2k}}$$

矩法估计量不唯一 (例 7.1.4)

再由

$$\widehat{E[X^{2k}]}_M = \frac{1}{n} \sum_{i=1}^n X_i^{2k}$$

所以

$$\hat{\theta}_M = \left(\frac{(2k)!n}{\sum_{i=1}^n X_i^{2k}} \right)^{\frac{1}{2k}}$$

最大似然估计

最大似然估计的基本思想是 某一次 抽样的观测结果是 概率最大 的。

问题分类：

- ▶ 离散型总体：联合概率分布在观测点处概率最大。
- ▶ 连续型总体：联合概率密度函数在观测点处取值最大。

似然函数

- ▶ 总体 X 服从某种离散型分布, 含有参数 θ , 某次观察值为 (x_1, x_2, \dots, x_n) , 其 似然函数 为

↪ Likelihood

$$L(x_1, \dots, x_n; \theta) = P(X_1 = x_1, \dots, X_n = x_n; \theta)$$

- ▶ 总体 X 服从某种连续型分布, 含有参数 θ , 某次观察值为 (x_1, x_2, \dots, x_n) , 其 似然函数 为

$$L(x_1, \dots, x_n; \theta) = f(X_1 = x_1, \dots, X_n = x_n; \theta)$$

最大似然估计

如上定义 **似然函数** $L(x_1, x_2, \dots, x_n, \theta)$, 如果某统计量 $\hat{\theta} = \hat{\theta}(X_1, X_2, \dots, X_n)$, 使得

$$\begin{aligned} & L(x_1, x_2, \dots, x_n; \hat{\theta}(x_1, x_2, \dots, x_n)) \\ &= \max_{\theta} L(x_1, x_2, \dots, x_n; \theta) \end{aligned}$$

就称 $\hat{\theta}(x_1, x_2, \dots, x_n)$ 为 θ 的 **最大似然估计量**

。

在数学优化问题中, 一个函数 (一元或者是多元) 的极值点等价于使这个函数的导数 (或者梯度) 为零的点。而极大值点不仅满足上述条件还会使得目标函数的 Hessian 矩阵在该点负定。

最大似然估计量

对服从参数为 λ 的指数分布的总体 X , 使用最大似然估计。

由于 $X \sim \text{Exp}(\lambda)$, 假设从中抽取容量为 n 的样本, 那么当 $x_i > 0, i = 1, 2, \dots, n$ 的时候, 并假设样本的一次观测值为 (x_1, x_2, \dots, x_n) , 那么

$$L(x_1, x_2, \dots, x_n, \lambda) = \lambda^n \exp \left(-\lambda \sum_{i=1}^n x_i \right)$$

最大似然估计量

为了方便讨论, 我们运用 $\ln x$ 的单调性, 来研究

$$\max_{\lambda} \ln L(x_1, x_2, \dots, x_n, \lambda)$$

计算一下

$$G(x_1, x_2, \dots, x_n, \lambda) = \ln L(x_1, x_2, \dots, x_n, \lambda) = n \ln(\lambda) - \lambda$$

对上式求其驻点, 即求使得

$$\frac{dG}{d\lambda} = 0$$

的 λ , 得到

$$\lambda^* = \frac{n}{\sum_{i=1}^n x_i}$$

最大似然估计量

验证下

$$\frac{d^2 G}{d\lambda^2} \Big|_{\lambda^*} = -\frac{n}{\lambda^2} \Big|_{\lambda^*} = -\frac{(\sum_{i=1}^n x_i)^2}{n} < 0$$

所以 λ^* 确实为 G 从而是 L 的极大值点, 即

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

最大似然估计量

对于服从 $[0, \theta]$ 上均匀分布的总体 X , 估计区间的右端点 θ 。

假设样本容量为 n , 那么 X 似然函数为

$$L(x_1, x_2, \dots, x_n, \theta) = \begin{cases} \frac{1}{\theta^n}, & x_1, x_2, \dots, x_n \in [0, \theta] \\ 0, & \text{else} \end{cases}$$

最大似然估计量

对于 **确定** 的观测值 (x_1, x_2, \dots, x_n) , 要使得 L 取到最大值, 那么

$$\theta = \max\{x_1, x_2, \dots, x_n\}$$

即

$$\theta = x_{(n)} \Rightarrow \hat{\theta}_L = X_{(n)}$$

也就是说区间右端点的最大似然估计量为顺序统计量 $X_{(n)}$.

最大似然估计量

我们也可以用矩法估计 θ ，因为

$$E[X] = \frac{\theta}{2},$$

所以 $\theta = 2E[X]$ ，说明

$$\hat{\theta}_M = 2\bar{X}$$

即区间右端点的矩法估计量为 $2\bar{X}$.

Outline

7.0 前言

7.1 参数的点估计

7.2 估计量优劣势的评价

7.3 参数的区间估计

无偏估计量

定义 7.2.1: 设总体 $X \sim F_X(\cdot, \theta)$, $\theta \in \Theta$, g 为 θ 的函数, $T(X_1, X_2, \dots, X_n)$ 是 $g(\theta)$ 的一个估计量, 如果

$$E[T(X_1, X_2, \dots, X_n)] = g(\theta),$$

称 $T(X_1, X_2, \dots, X_n)$ 为 $g(\theta)$ 的 **无偏估计量**。如果

$$\lim_{n \rightarrow \infty} E[T(X_1, X_2, \dots, X_n)] = g(\theta)$$

则称 $T(X_1, X_2, \dots, X_n)$ 为 $g(\theta)$ 的 **渐进无偏估计量**。

无偏与渐进无偏

分析当 $X \sim U[0, \theta]$, 其区间右端点的矩估计量 $\hat{\theta}_M$ 与最大似然估计量 $\hat{\theta}_L$ 是否为 θ 的无偏估计量?

► 对于矩法估计量

$$\hat{\theta}_M = 2\bar{X}$$

计算

$$\begin{aligned} E_{\theta}[\hat{\theta}_M] &= 2E_{\theta}[\bar{X}] = \frac{2}{n} \sum_{i=1}^n E_{\theta}[X_i] \\ &= \frac{2}{n} \cdot n \cdot \frac{\theta}{2} = \theta \end{aligned}$$

所以 $\hat{\theta}_M$ 是 θ 的无偏估计量。

无偏与渐进无偏

- ▶ 对于最大似然估计量

$$\hat{\theta}_L = X_{(n)}$$

根据 (6.3.19), 知

$$f_{X_{(n)}}(x) = n[F_X(x)]^{n-1}f_X(x)$$

而

$$f_X(x) = \begin{cases} \frac{1}{\theta}, & x \in [0, \theta] \\ 0, & \text{else} \end{cases}$$

并且

$$F_X(x) = \begin{cases} 0, & x < 0 \\ \frac{x}{\theta}, & x \in [0, \theta] \\ 1, & \text{else} \end{cases}$$

无偏与渐进无偏

所以

$$f_{X_{(n)}}(x) = \begin{cases} n \left(\frac{x}{\theta}\right)^{n-1} \frac{1}{\theta}, & x \in [0, \theta] \\ 0, & \text{else} \end{cases}$$

所以

$$E_{\theta}[\hat{\theta}_L] = E_{\theta}[X_{(n)}] = \int_0^{\theta} x n \frac{x^{n-1}}{\theta^n} dx = \frac{n}{n+1} \theta$$

所以 $\hat{\theta}_L$ 不是 θ 的无偏估计量，而是 θ 的渐进无偏估计量。

无偏与渐进无偏

我们令

$$\hat{\theta} = \frac{n+1}{n}\hat{\theta}_L$$

那么 $\hat{\theta}$ 就是 θ 的无偏估计量。若 $\hat{\theta}$ 是 θ 的无偏估计量，那么 $g(\hat{\theta})$ 一般不是 $g(\theta)$ 的无偏估计量，除非 $g(x)$ 是线性函数。例如 S_n^{*2} 是 σ^2 的无偏估计量，但是 S_n^* 却不是 σ 的无偏估计量。

有效估计量

对于总体 X 的参数 $g(\theta)$ 的两个 无偏估计量 $T_1(X_1, X_2, \dots, X_n)$, $T_2(X_1, X_2, \dots, X_n)$, 若有

$$\mathbf{Var}[T_1] < \mathbf{Var}[T_2]$$

则称 T_1 比 T_2 有效。

由于方差反映随机变量取值的波动程度, 认为波动较小的统计量更有效是合理的。

有效估计

判断总体 $X \sim U[0, \theta]$ 的区间右端点的矩估计量和最大似然估计量哪一个更有效?

$$\begin{aligned} \text{Var}[\hat{\theta}_M] &= \text{Var}[2\bar{X}] = \text{Var}\left[\frac{2}{n} \sum_{i=1}^n X_i\right] \\ &= \frac{4}{n^2} \sum_{i=1}^n \text{Var}[X_i] = \frac{4}{n^2} n \frac{\theta^2}{12} = \frac{\theta^2}{3n} \end{aligned}$$

无偏估计量

$$\begin{aligned} \text{Var}[\hat{\theta}_L] &= \text{Var}\left[\frac{n+1}{n}X_{(n)}\right] \\ &= \frac{(n+1)^2}{n^2} [E[X_{(n)}^2] - (E[X_{(n)}])^2] \\ &= \frac{(n+1)^2}{n^2} \left\{ \int_0^\theta x^2 n \frac{x^{n-1}}{\theta^n} dx - \frac{n^2}{(n+1)^2} \theta^2 \right\} \\ &= \frac{(n+1)^2}{n^2} \left\{ \frac{n}{n+2} \theta^2 - \frac{n^2}{(n+1)^2} \theta^2 \right\} \\ &= \frac{\theta^2}{n(n+2)} \end{aligned}$$

无偏估计量

显然的, 当 $n > 1$, 就有

$$\text{Var}[\hat{\theta}_L] < \text{Var}[\hat{\theta}_M]$$

即 $\hat{\theta}_L$ 比 $\hat{\theta}_M$ 要好。

一致最小方差无偏估计量

定义 7.2.2: 设总体 $X \sim F_X(\cdot, \theta)$, $\theta \in \Theta$, 若 $T_0(X_1, X_2, \dots, X_n)$ 是 $g(\theta)$ 的无偏估计量, 且对 $g(\theta)$ 的任意无偏估计量 $T(X_1, X_2, \dots, X_n)$ 都有

$$\text{Var}[T_0] \leq \text{Var}[T], \forall \theta \in \Theta$$

则称 T_0 是 $g(\theta)$ 的 **一致最小方差无偏估计量**。

一致最小方差无偏估计量一般不易求得, 除非事先知道 $g(\theta)$ 无偏估计量的方差的下界, 再去寻找达到该方差的无偏估计量。

相合估计量

定义 7.2.3: 设总体 $X \sim F_X(\cdot, \theta)$, $\theta \in \Theta$, 并且设 $T(X_1, X_2, \dots, X_n)$ 是 $g(\theta)$ 的估计量, 如果对于任意的 $\epsilon > 0$, 有

$$\lim_{n \rightarrow \infty} P(|T(X_1, X_2, \dots, X_n) - g(\theta)| \geq \epsilon) = 0, \forall \theta \in \Theta,$$

则称 $T(X_1, X_2, \dots, X_n)$ 为 $g(\theta)$ 的 **相合估计量**。

由弱大数定律, $\overline{X^k}$ 是总体 k 阶原点矩的相合估计量。

相合估计量的判定依据

若 $\hat{\theta}$ 为 θ 的相合估计, 且 $\eta = g(\theta)$ 为 θ 的连续函数, 那么 $\hat{\eta} = g(\hat{\theta})$ 为 η 的相合估计量。

- ▶ $\overline{X^k}$ 是 $E[X^k]$ 的相合估计量。
- ▶ S_n^2 和 S_n^{*2} 都是 σ^2 的相合估计量。
- ▶ S_n 和 S_n^* 都是 σ 的相合估计量。

Outline

7.0 前言

7.1 参数的点估计

7.2 估计量优劣势的评价

7.3 参数的区间估计

参数的区间估计

区间估计: 参数的 **区间估计** 表示以

$$T_1(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n) \text{ 和 } T_2(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n)$$

为一个 **随机区间** 的端点, 并且使得该随机区间包含参数的 **概率** 满足给定的条件。

参数的区间估计

定义 7.3.1: 设总体 $X \sim F_X(\cdot, \theta)$, $\theta \in \Theta$,
(X_1, X_2, \dots, X_n) 为样本。若有统计量
 $T_1(X_1, X_2, \dots, X_n)$ 和 $T_2(X_1, X_2, \dots, X_n)$, 使得对
给定的 $\alpha (0 < \alpha < 1)$ 有

$$P(T_1(\mathbf{X}_1, \dots, \mathbf{X}_n) \leq g(\theta) \leq T_2(\mathbf{X}_1, \dots, \mathbf{X}_n)) = 1$$

则称 $[T_1(X_1, X_2, \dots, X_n), T_2(X_1, X_2, \dots, X_n)]$ 是
 $g(\theta)$ 的 置信度为 $1 - \alpha$ 的 区间估计, 或叫做 估计区间, 置信区间。

区间估计的一般方法

置信度为 $1 - \alpha$ 的区间估计的一般方法可以分为下面三步：

- ▶ 构造一个样本和未知参数 θ 的函数，记为 $Y(X_1, X_2, \dots, X_n, \theta)$ ，使得 Y 服从的分布 **不依赖** 于 θ 。
- ▶ **适当选择** 两个常数 a, b ，使得

$$P(a \leq Y \leq b) = 1 - \alpha$$

- ▶ 根据 $a \leq Y \leq b$ **反解出** θ 的范围

$$T_1(X_1, X_2, \dots, X_n) \leq \theta \leq T_2(X_1, X_2, \dots, X_n)$$

即有了 θ 的 $1 - \alpha$ 置信区间。

正态分布总体均值的区间估计

对于正态分布总体 $X \sim N(\mu, \sigma_0^2)$, 其中 σ_0 已知, 求 μ 的置信度为 $1 - \alpha$ 的区间估计。

假设 (X_1, X_2, \dots, X_n) 为其样本, 根据推论 6.3.1 知

$$\bar{X} \sim N\left(\mu, \frac{\sigma_0^2}{n}\right)$$

所以

$$Y = \frac{\bar{X} - \mu}{\sqrt{\frac{\sigma_0^2}{n}}} \sim N(0, 1)$$

正态分布总体均值的区间估计

由

$$P(|Y| \leq u_{1-\frac{\alpha}{2}}) = 1 - \alpha$$

知

$$P\left(\left|\frac{\bar{X} - \mu}{\sqrt{\frac{\sigma_0^2}{n}}}\right| \leq u_{1-\frac{\alpha}{2}}\right) = 1 - \alpha$$

变化为

$$P\left(\bar{X} - u_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_0^2}{n}} \leq \mu \leq \bar{X} + u_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_0^2}{n}}\right) = 1 - \alpha.$$

所以我们有 μ 的置信度为 $1 - \alpha$ 的区间估计为

$$\left[\bar{X} - u_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_0^2}{n}}, \bar{X} + u_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_0^2}{n}}\right]$$

正态总体参数的区间估计

总体方差 σ^2 未知, 给出总体均值 μ 的区间估计量

根据推论 6.3.2 有

$$\frac{\bar{X} - \mu}{S_n} \sqrt{n-1} \sim t(n-1)$$

由

$$P\left(\left|\frac{\bar{X} - \mu}{S_n} \sqrt{n-1}\right| \leq t_{1-\frac{\alpha}{2}}(n-1)\right) = 1 - \alpha$$

知 μ 的置信度为 $1 - \alpha$ 的区间估计为

$$\left[\bar{X} - t_{1-\frac{\alpha}{2}} \frac{S_n}{\sqrt{n-1}}, \bar{X} + t_{1-\frac{\alpha}{2}} \frac{S_n}{\sqrt{n-1}}\right]$$

σ^2 的区间估计

$\mu = \mu_0$, 其中 μ_0 已知, 求 σ^2 的区间估计。

由于总体与样本独立同分布, 所以

$$Y_i = \frac{X_i - \mu_0}{\sigma} \sim N(0, 1)$$

根据命题 6.3.2 知

$$Z = \sum_{i=1}^n Y_i^2 \sim \chi^2(n)$$

所以

$$P(\chi_{\frac{\alpha}{2}}^2(n) \leq Z \leq \chi_{1-\frac{\alpha}{2}}^2(n)) = 1 - \alpha$$

σ^2 的区间估计

即

$$P\left(\frac{1}{\chi_{\frac{\alpha}{2}}^2(n)} \geq \frac{\sigma^2}{\sum_{i=1}^n (X_i - \mu_0)^2} \geq \frac{1}{\chi_{1-\frac{\alpha}{2}}^2(n)}\right) = 1 - \alpha$$

得到 σ^2 置信度为 $1 - \alpha$ 的区间估计为

$$\left[\frac{\sum_{i=1}^n (X_i - \mu_0)^2}{\chi_{1-\frac{\alpha}{2}}^2(n)}, \quad \frac{\sum_{i=1}^n (X_i - \mu_0)^2}{\chi_{\frac{\alpha}{2}}^2(n)} \right]$$

σ^2 区间估计

μ 未知, 估计 σ^2 的置信度 $1 - \alpha$ 的区间。
由抽样分布基本定理,

$$\frac{nS_n^2}{\sigma^2} \sim \chi^2(n-1)$$

令

$$P\left(\chi_{\frac{\alpha}{2}}^2(n-1) \leq \frac{nS_n^2}{\sigma^2} \leq \chi_{1-\frac{\alpha}{2}}^2(n-1)\right) = 1 - \alpha$$

所以 σ^2 的置信度 $1 - \alpha$ 的区间估计为

$$\left[\frac{nS_n^2}{\chi_{1-\frac{\alpha}{2}}^2(n-1)}, \quad \frac{nS_n^2}{\chi_{\frac{\alpha}{2}}^2(n-1)} \right]$$

两个正态总体参数的区间估计

设 $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$, 且 X, Y 相互独立, 样本分别为 (X_1, X_2, \dots, X_m) , (Y_1, Y_2, \dots, Y_n) , 其中总体 X 与 Y 的样本均值与方差分别为

$$\bar{X}, \bar{Y}, S_m^2, S_n^2$$

$\mu_1 - \mu_2$ 和 $\frac{\sigma_1}{\sigma_2}$ 的区间估计量分别是什么?

$\mu_1 - \mu_2$ 的区间估计

σ_1^2 和 σ_2^2 已知, 求 $\mu_1 - \mu_2$ 的区间估计。由推论 6.3.1

$$\bar{X} \sim N(\mu_1, \frac{\sigma_1^2}{m}), \bar{Y} \sim N(\mu_2, \frac{\sigma_1^2}{n}),$$

所以

$$\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}} \sim N(0, 1)$$

然后令

$$P\left(\left|\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}}\right| \leq u_{1-\frac{\alpha}{2}}\right) = 1 - \alpha$$

$\mu_1 - \mu_2$ 的区间估计

求得 $u_1 - u_2$ 的置信度为 $1 - \alpha$ 区间估计为

$$\left[(\bar{X} - \bar{Y}) - u_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}, \right. \\ \left. (\bar{X} - \bar{Y}) + u_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}} \right]$$

两个正态总体参数的区间估计

$\sigma_1 = \sigma_2 = \sigma$ 未知, 求 $\mu_1 - \mu_2$ 的区间估计。
由推论 6.3.4 知

$$\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{S_w \sqrt{\frac{1}{m} + \frac{1}{n}}} \sim t(m + n - 2)$$

以及 t 分布关于 y 轴对称的性质, 得到

$$P\left(\left|\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{S_w \sqrt{\frac{1}{m} + \frac{1}{n}}}\right| \leq t_{1-\frac{\alpha}{2}}(m + n - 2)\right) = 1 - \alpha$$

两个正态总体参数的区间估计

知 $\mu_1 - \mu_2$ 的置信度为 $1 - \alpha$ 的区间估计量为

$$\begin{aligned} & [(\bar{X} - \bar{Y}) - t_{1-\frac{\alpha}{2}}(m+n-2)S_w\sqrt{\frac{1}{m} + \frac{1}{n}}, \\ & (\bar{X} - \bar{Y}) + t_{1-\frac{\alpha}{2}}(m+n-2)S_w\sqrt{\frac{1}{m} + \frac{1}{n}}] \end{aligned}$$

$\frac{\sigma_1^2}{\sigma_2^2}$ 的区间估计

μ_1, μ_2 已知, 求 $\frac{\sigma_1^2}{\sigma_2^2}$ 的区间估计。

首先, 下面的结论是显然的

$$\frac{1}{\sigma_1^2} \sum_{i=1}^m (X_i - \mu_1)^2 \sim \chi^2(m), \quad \frac{1}{\sigma_2^2} \sum_{i=1}^n (Y_i - \mu_2)^2 \sim \chi^2(n)$$

根据 F 分布的定义有

$$\frac{n \sum_{i=1}^m (X_i - \mu_1)^2 \sigma_2^2}{m \sum_{i=1}^n (Y_i - \mu_2)^2 \sigma_1^2} \sim F(m, n)$$

$\frac{\sigma_1^2}{\sigma_2^2}$ 的区间估计

特别的

$$P\left(F_{\frac{\alpha}{2}}(m, n) \leq \frac{n \sum_{i=1}^m (X_i - \mu_1)^2 \sigma_2^2}{m \sum_{i=1}^n (Y_i - \mu_2)^2 \sigma_1^2} \leq F_{1-\frac{\alpha}{2}}(m, n)\right) = 1$$

由此可得 $\frac{\sigma_1^2}{\sigma_2^2}$ 的置信度为 $1 - \alpha$ 的区间估计

$$\left[\frac{1}{F_{1-\frac{\alpha}{2}}(m, n)} \frac{n \sum_{i=1}^m (X_i - \mu_1)^2}{m \sum_{i=1}^n (Y_i - \mu_2)^2}, \frac{1}{F_{\frac{\alpha}{2}}(m, n)} \frac{n \sum_{i=1}^m (X_i - \mu_1)^2}{m \sum_{i=1}^n (Y_i - \mu_2)^2} \right]$$

$\frac{\sigma_1^2}{\sigma_2^2}$ 的区间估计

μ_1, μ_2 未知, 求 $\frac{\sigma_1^2}{\sigma_2^2}$ 的区间估计。

根据推论 6.3.3 有

$$\frac{mS_m^2}{nS_n^2} \frac{\sigma_2^2}{\sigma_1^2} \frac{n-1}{m-1} \sim F(m-1, n-1)$$

所以

$$P\left(F_{\frac{\alpha}{2}}(m-1, n-1) \leq \frac{mS_m^2}{nS_n^2} \frac{\sigma_2^2}{\sigma_1^2} \frac{n-1}{m-1} \leq F_{1-\frac{\alpha}{2}}(m-1, n-1)\right) = 1 - \alpha$$

可得 $\frac{\sigma_1^2}{\sigma_2^2}$ 的置信度为 $1 - \alpha$ 的区间估计量为

$$\left[\frac{1}{F_{1-\frac{\alpha}{2}}(m-1, n-1)} \frac{m(n-1)S_m^2}{n(m-1)S_n^2}, \frac{1}{F_{\frac{\alpha}{2}}(m-1, n-1)} \frac{m(n-1)S_m^2}{n(m-1)S_n^2} \right]$$