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

估计量评价 $\hat{\theta}: X_1 \times \dots \times X_n \longrightarrow \mathbb{R}$.

- 无偏性 $\mathbb{E}[\hat{\theta}] = \theta$. $\mathbb{E}[\hat{\theta}] = \mathbb{E}[\theta] = \theta$.
  - 渐近无偏 $\mathbb{E}[\hat{\theta}] \to \theta_{(n\to+\infty)}$ $Var[\hat{\theta}] \to Var[\theta] = 0$
  - 有效性: $Var[\hat{\theta}]$
- 相合性 (一致性).

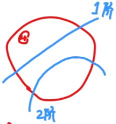
  $\hat{\theta} \xrightarrow{P} \theta$.

  (Tod · WLNL)

$X \sim F_X(\cdot;\theta)$, $\theta \in \Theta$

样本 $(X_1, \dots, X_n)$.

Q: How to estimate $\theta \in \Theta$ from $(X_1, \dots, X_n)$?



- 点估计: $(x_1, \dots, x_n) \longmapsto \hat{\theta}$
  - 矩法: 通过 k阶矩 构造方程组.

    解方程组得 $\hat{\theta}_M$ ⇒ 问题: 存在性?

  - 最大似然估计:

    似然: $X \sim f_X(\cdot;\theta)$.

$$P(X_1 = x_1, \dots, X_n = x_n; \theta) = \prod_{i=1}^{n} P(X_i = x_i; \theta)$$

$$\Rightarrow -\frac{1}{n}\ln P(\dots) = -\sum_{i=1}^{n} \frac{1}{n} \ln P(X_i = x_i;\theta)$$

Let $(x_1, \dots, x_n) \sim Q$

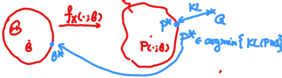
$Q(X = x_i) = \frac{1}{n}$

$$= -\sum_{i=1}^{n} Q(X = x_i) \cdot \ln P(X = x_i;\theta)$$

KL divergence $= \mathbb{E}_{X\sim Q}[-\ln P(\cdot;\theta)]$

$KL(P \| Q)$

$\{f_X(\cdot;\theta) \mid \theta \in \Theta\}$



$KL(P\|Q)$ 在 $P = Q$ 时

取值最小 $KL(P\|P) = \mathbb{E}_{X\sim P}[-\ln P(\dots)]$ = Entropy of P

熵

参数的区间估计

区间估计: 参数的 区间估计 表示以

$$T_1(X_1, X_2, \cdots, X_n) \text{ 和 } T_2(X_1, X_2, \cdots, X_n)$$

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

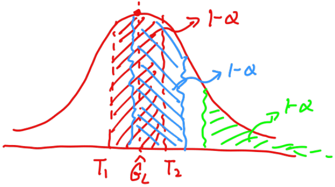
$$\boldsymbol{P}(\boldsymbol{T}_1(\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n) \le \boldsymbol{g}(\boldsymbol{\theta}) \le \boldsymbol{T}_2(\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n)) \overset{1-\alpha}{=} 1$$

则称 $[T_1(X_1, X_2, \cdots, X_n), T_2(X_1, X_2, \cdots, X_n)]$ 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。

考察 $\alpha \to 1$

置信度 $1-\alpha \to 0$

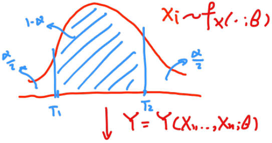
$\Rightarrow \hat{\theta}_L$: 最大似然



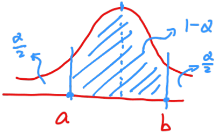
区间估计的一般方法

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 $Y(X_1, X_2, \cdots, X_n, \theta)$，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

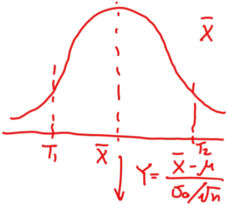
$$T_1(X_1, X_2, \cdots, X_n) \leq \theta \leq T_2(X_1, X_2, \cdots, X_n)$$

即有了 θ 的 $1-\alpha$ 置信区间。

正态分布总体均值的区间估计

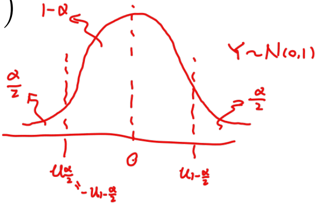
对于正态分布总体 $X \sim N(\mu, \sigma_0^2)$，其中 σ_0 已知，求 μ 的置信度为 $\underline{1-\alpha}$ 的区间估计。
假设 $(X_1, X_2, \cdots, X_n)$ 为其样本，根据推论 6.3.1 知　$\hat{\mu} = \overline{X}$

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



所以

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



要使 $P[T_1 \le \overline{X} \le T_2] = 1-\alpha.$

$\qquad ||$

$P[u_{\frac{\alpha}{2}} \le Y \le u_{1-\frac{\alpha}{2}}].$

代入 $Y = \frac{\overline{X}-\mu}{\sigma_0/\sqrt{n}}$ 得

$u_{\frac{\alpha}{2}} \le \frac{\overline{X}-\mu}{\sigma_0/\sqrt{n}} \le u_{1-\frac{\alpha}{2}} \Rightarrow \overline{X} - \frac{\sigma_0}{\sqrt{n}} u_{\frac{\alpha}{2}} \le \mu \le \overline{X} - \frac{\sigma_0}{\sqrt{n}} u_{\frac{\alpha}{2}}$

$\Uparrow$

$\overline{X} - \frac{\sigma_0}{\sqrt{n}} \cdot u_{1-\frac{\alpha}{2}} \le \mu \le \overline{X} + \frac{\sigma_0}{\sqrt{n}} \cdot u_{1-\frac{\alpha}{2}}$

正态分布总体均值的区间估计

由

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

知

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

变化为

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

所以我们有 μ 的置信度为 $1-\alpha$ 的区间估计为

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

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

$\bar{X} \sim N(\mu, \frac{1}{n}\sigma^2)$.

用样本方差均值.

$S_n^2 = \frac{1}{n}\sum(X_i-\bar{X})^2 \to \sigma^2$.

根据推论 6.3.2 有

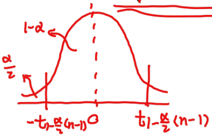
$$\frac{\overline{X}-\mu}{S_n}\sqrt{n-1} \sim t(n-1)$$

$\bar{X} \sim N(\mu, \frac{1}{n}\sigma^2)$

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

$\frac{n\cdot S_n^2}{\sigma^2} = \sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{\sigma}\right)^2 \sim \chi^2(n-1)$

$\Rightarrow \frac{\frac{\bar{X}-\mu}{\cancel{\sigma}/\sqrt{n}}}{\left(\frac{\cancel{n}\cdot S_n^2}{\cancel{\sigma^2}}/n-1\right)^{\frac{1}{2}}} = \frac{\bar{X}-\mu}{\sqrt{S_n^2}}\sqrt{n-1} \sim t(n-1)$



由

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

知 μ 的置信度为 $1-\alpha$ 的区间估计为

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

σ^2 的区间估计

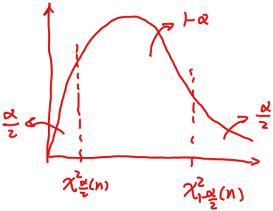
$\mu = \mu_0$，其中 μ_0 已知，求 σ^2 的区间估计。

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

$$P(\chi^2_{\frac{\alpha}{2}}(n) \le Z \le \chi^2_{1-\frac{\alpha}{2}}(n)) = 1 - \alpha$$

σ^2 的区间估计

即

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

得到 σ^2 置信度为 $1-\alpha$ 的区间估计为

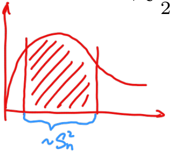
$$\left[\frac{\sum_{i=1}^n (X_i-\mu_0)^2}{\chi^2_{1-\frac{\alpha}{2}}(n)},\quad \frac{\sum_{i=1}^n (X_i-\mu_0)^2}{\chi^2_{\frac{\alpha}{2}}(n)}\right]$$

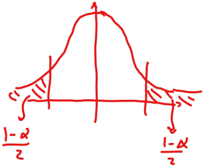
Q: Why not use $\bar{X} \sim N(\mu_0, \frac{1}{n}\sigma^2)$

$\sqrt{n}\,\frac{\bar{X}-\mu_0}{\sigma} \sim N(0,1)$.

$\Rightarrow P[-u_{1-\frac{\alpha}{2}} \leq \sqrt{n}(\bar{X}-\mu_0)\cdot\frac{1}{\sigma} \leq u_{1-\frac{\alpha}{2}}] = 1-\alpha$

$\Rightarrow P[\frac{n\cdot(\bar{X}-\mu_0)^2}{\sigma^2} \leq (u_{1-\frac{\alpha}{2}})^2] = 1-\alpha \Rightarrow \sigma^2 \geq \frac{n\cdot(\bar{X}-\mu)^2}{(u_{1-\frac{\alpha}{2}})^2}$





σ^2 区间估计

$\sum\left(\frac{x_i-\bar{x}}{\sigma}\right)^2$

$\bar{x}$ 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

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

所以 σ^2 的置信度 $1-\alpha$ 的区间估计为

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

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 $(X_1, X_2, \cdots, X_m)$， $(Y_1, Y_2, \cdots, Y_n)$， 其中总体 X 与 Y 的样本均值与方差分别为

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

$\mu_1 - \mu_2$ 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

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

所以

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

$\mathrm{Var}[\bar{X} - \bar{Y}]$
$= \mathrm{Var}\,\bar{X} + \mathrm{Var}\,\bar{Y}$
$= \frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}$

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

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

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

$\sum_{i=1}^{m}\left(\frac{X_i-\bar{X}}{\sigma}\right)^2 \sim \chi^2(m-1)$， $\sum_{j=1}^{n}\left(\frac{Y_j-\bar{Y}}{\sigma}\right)^2 \sim \chi^2(n-1)$

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

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

$\sqrt{\underbrace{m\cdot S_m^2(X) + n\cdot S_n^2(Y) \,/\, (m+n-2)}_{S_w^2}\cdot\boxed{\sigma^2}}$

$\bar{X} \sim N(\mu_1, \frac{1}{m}\sigma^2)$, $\bar{Y} \sim N(\mu_2, \frac{1}{n}\sigma^2)$

$\sim \sqrt{\chi^2(m+n-2)/(m+n-2)}$

以及 t 分布关于 y 轴对称的性质，得到

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

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

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

$\frac{\sigma_1^2}{\sigma_2^2}$ 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

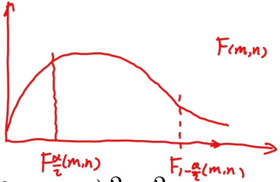
$$\frac{1}{\sigma_1^2}\sum_{i=1}^{m}(X_i-\mu_1)^2 \sim \chi^2(m), \quad \frac{1}{\sigma_2^2}\sum_{i=1}^{n}(Y_i-\mu_2)^2 \sim \chi^2(n)$$

根据 F 分布的定义有

$$\frac{\sum_{i=1}^{m} X_i/m}{\sum_{j=1}^{n} Y_j/n} \sim F(m,n)$$

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

$\frac{\sigma_1^2}{\sigma_2^2}$ 的区间估计



特别的

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

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

$\bar{X}$ $\bar{Y}$

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

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

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

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