

第七章 参数估计

参数: 反映总体某方面特征的量

参数的点估计

设总体 X 有未知参数 θ , $X_1, X_2, \dots, X_n$ 是 X 的简单随机样本。

构造合适的统计量 $\hat{\theta} = \hat{\theta}(X_1, \dots, X_n)$ 用来估计未知参数 θ , 称 $\hat{\theta}$ 为参数 θ 的点估计量

当给定样本观察值 $x_1, \dots, x_n$ 时, 称 $\hat{\theta}(x_1, \dots, x_n)$ 为参数 θ 的点估计值

矩估计法

统计思想: 以样本矩估计总体矩, 以样本矩的函数估计总体矩的函数
理论依据: 辛钦大数定律和依概率收敛的性质

总体 X 的 k 阶矩: $\mu_k = E(X^k)$

则样本 $(X_1, X_2, \dots, X_n)$ 的 k 阶矩为 $A_k = \frac{1}{n} \sum_{i=1}^n X_i^k$

由大数定理: $A_k = \frac{1}{n} \sum_{i=1}^n X_i^k \xrightarrow{P} \mu_k = E(X^k)$

$$\mu_k = A_k$$

极大似然估计法

一般地, 设离散型总体 $X \sim P(X; \theta)$, $\theta \in \Theta$, θ 未知

事件 $\{X_1 = x_1, \dots, X_n = x_n\}$ 发生的概率为 $L(\theta) = P\{X_1 = x_1, \dots, X_n = x_n\} = P(X_1; \theta) \cdots P(X_n; \theta)$
$$= \prod_{i=1}^n P(X_i; \theta)$$

极大似然原理: $L(\hat{\theta}(x_1, \dots, x_n)) = \max_{\theta \in \Theta} L(\theta)$

称 $\hat{\theta}(x_1, \dots, x_n)$ 为 θ 的极大似然估计值

相应统计量 $\hat{\theta}(x_1, \dots, x_n)$ 为 θ 的极大似然估计量 (MLE)

[说明] 1. 未知参数可能不是一个, 一般设为 $\theta = (\theta_1, \theta_2, \dots, \theta_k)$;

2. 在求 $L(\theta)$ 的最大值时, 通常转换为求: $\ln L(\theta)$ 的最大值, $\ln L(\theta)$ 称为对数似然函数.

利用 $\frac{\partial \ln L(\theta)}{\partial \theta_i} = 0, i = 1, 2, \dots, k$ 解得 $\hat{\theta}_i, i = 1, 2, \dots, k$.

3. 若 $L(\theta)$ 关于某个 θ_i 是单调增(减)函数, 此时 θ_i 的极大似然估计在其边界取得;

4. 若 $\hat{\theta}$ 是 θ 的极大似然估计, 则 $g(\hat{\theta})$ 的极大似然估计为 $g(\hat{\theta})$.

极大似然估计的不变性: $\hat{\theta}$ 是 θ 的极大似然估计, $g(x)$ 是连续函数,

则 $g(\hat{\theta})$ 是 $g(\theta)$ 的极大似然估计

估计量的评选准则

① 无偏性准则

若参数 θ 的估计量 $\hat{\theta} = \hat{\theta}(x_1, x_2, \dots, x_n)$, 满足 $E(\hat{\theta}) = \theta$, 则称 $\hat{\theta}$ 是 θ 的一个无偏估计量

若 $E(\hat{\theta}) \neq \theta$, 则 $|E(\hat{\theta}) - \theta|$ 为估计量 $\hat{\theta}$ 的偏差

若 $\lim_{n \rightarrow \infty} E(\hat{\theta}) = \theta$, 则 $\hat{\theta}$ 是 θ 的渐近无偏估计量

② 有效性准则

设 $\hat{\theta}_1, \hat{\theta}_2$ 是 θ 的两个无偏估计,

若 $D(\hat{\theta}_1) \leq D(\hat{\theta}_2)$, 对一切 $\theta \in \Theta$ 成立, 且不等号至少对某一 $\theta \in \Theta$ 成立, 则称 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 有效

③ 均方误差准则

设 $\hat{\theta}$ 是参数 θ 的点估计, 方差存在, 则称 $E(\hat{\theta} - \theta)^2$ 是估计量的均方误差, 记为 $Mse(\hat{\theta})$

$$Mse(\hat{\theta}) = D(\hat{\theta}) + [E(\hat{\theta} - \theta)]^2$$

若 $\hat{\theta}$ 是 θ 的无偏估计, 则有 $Mse(\hat{\theta}) = D(\hat{\theta})$

④ 相合性准则

设 $\hat{\theta}(x_1, \dots, x_n)$ 是参数 θ 的估计量,

若对于 $\forall \theta \in \Theta$, 当 $n \rightarrow +\infty$ 时, $\hat{\theta}_n$ 依概率收敛于 θ ,

即 $\forall \varepsilon > 0$, 有 $\lim_{n \rightarrow +\infty} P\{|\hat{\theta}_n - \theta| \geq \varepsilon\} = 0$ 成立

则称 $\hat{\theta}_n$ 为 θ 的相合估计量或一致估计量

区间估计

假设 $(x_1, \dots, x_n)$ 是总体 X 的一个样本, 区间估计的方法是给出两个统计量

$\hat{\theta}_1 = \hat{\theta}_1(x_1, \dots, x_n)$, $\hat{\theta}_2 = \hat{\theta}_2(x_1, \dots, x_n)$ 使区间 $[\hat{\theta}_1, \hat{\theta}_2]$ 以一定可靠程度盖住 θ

置信区间的定义

定义 7.3.1: 设总体 X 的分布函数 $F(x; \theta)$ 含有一个未知参数 θ , $(X_1, \dots, X_n)$ 是总体 X 的一个样本, 对给定的值 α ($0 < \alpha < 1$),

如果有两个统计量 $\hat{\theta}_L = \hat{\theta}_L(X_1, \dots, X_n)$, $\hat{\theta}_U = \hat{\theta}_U(X_1, \dots, X_n)$, 使得:

$$P\{\hat{\theta}_L(X_1, \dots, X_n) < \theta < \hat{\theta}_U(X_1, \dots, X_n)\} \geq 1 - \alpha \quad \forall \theta \in \Theta \quad (7-1)$$

则称随机区间 $(\hat{\theta}_L, \hat{\theta}_U)$ 是 θ 的 双侧置信区间 称 $1 - \alpha$ 为 置信度

$\hat{\theta}_L$ 和 $\hat{\theta}_U$ 分别称为 双侧置信下限 和 双侧置信上限.

如果 θ 的置信区间 $(\hat{\theta}_L, \hat{\theta}_U)$, 满足:

$$P\{\hat{\theta}_L(X_1, \dots, X_n) < \theta < \hat{\theta}_U(X_1, \dots, X_n)\} = 1 - \alpha \quad \forall \theta \in \Theta$$

则置信区间 $(\hat{\theta}_L, \hat{\theta}_U)$ 的含义为

若反复抽样多次 (各次得到的样本容量相等, 都为 n),

每个样本值确定一个区间 $(\hat{\theta}_L, \hat{\theta}_U)$, 每个这样的区间

或者包含 θ 的真值, 或者不包含 θ 的真值.

按伯努里大数定律, 在这些区间中, 包含 θ 真值的

约占 $100(1 - \alpha)\%$.

单侧置信上下限

定义 7.3.2 在以上定义中, 若将 (7-1) 式改为:

$$P\{\hat{\theta}_L(X_1, \dots, X_n) < \theta\} \geq 1 - \alpha, \quad \forall \theta \in \Theta \quad (7-2)$$

则称 $\hat{\theta}_L(X_1, \dots, X_n)$ 为 θ 的 单侧置信下限.

又若将 (7-2) 式改为:

$$P\{\theta < \hat{\theta}_U(X_1, \dots, X_n)\} \geq 1 - \alpha, \quad \forall \theta \in \Theta \quad (7-3)$$

则称 $\hat{\theta}_U(X_1, \dots, X_n)$ 为 θ 的 单侧置信上限.

单侧置信限和 双侧置信区间的关系

设 $\hat{\theta}_L = \hat{\theta}_L(X_1, \dots, X_n)$ 是参数 θ 的置信度为 $1 - \alpha_1$ 的单侧置信下限,

$\hat{\theta}_U = \hat{\theta}_U(X_1, \dots, X_n)$ 是参数 θ 的置信度为 $1 - \alpha_2$ 的单侧置信上限,

则 $(\hat{\theta}_L, \hat{\theta}_U)$ 是参数 θ 的置信度为 $1 - \alpha_1 - \alpha_2$ 的双侧置信区间。

Neyman 原则

在置信度达到一定的前提下, 选取精确度尽可能高的区间

枢轴量法

定义 7.3.3: 设总体 X 有概率密度 (或概率分布律) $f(x; \theta)$, 其中 θ 是待估的未知参数, 并

设 $X_1, \dots, X_n$ 是来自该总体的样本,

称样本和未知参数 θ 的函数 $G(X_1, \dots, X_n; \theta)$ 为 **枢轴量**,

如果它的分布不依赖于未知参数 θ , 且完全已知。

枢轴量和统计量的区别

- (1) 枢轴量是样本和待估参数的函数, 其分布不依赖于任何未知参数;
- (2) 统计量只是样本的函数, 其分布常常依赖于未知参数。

构造置信区间具体步骤

(1) 构造一个分布已知的枢轴量 $G(X_1, \dots, X_n; \theta)$;

(2) 对连续型总体和给定的置信度 $1 - \alpha$, 设常数 $a < b$ 满足

$$P\{a < G(X_1, \dots, X_n; \theta) < b\} = 1 - \alpha;$$

(3) 若能从 $a < G(X_1, \dots, X_n; \theta) < b$ 得到等价的不等式

$$\hat{\theta}_L(X_1, \dots, X_n) < \theta < \hat{\theta}_U(X_1, \dots, X_n)$$

那么 $(\hat{\theta}_L, \hat{\theta}_U)$ 就是 θ 的置信度为 $1 - \alpha$ 的置信区间。

注: 枢轴量 $G(X_1, \dots, X_n; \theta)$ 的构造, 通常从参数 θ

的点估计 $\hat{\theta}$

(如极大似然估计, 无偏估计, 矩估计等)

出发, 根据 $\hat{\theta}$ 的分布进行改造而得。

问: 若步骤 (3) 中 a 和 b 的解不唯一, 该如何处理?

1. 根据 Neyman 原则: 求 a 和 b 使得区间长度最短;

2. 如果最优解不存在或比较复杂, 为应用的方便, 常取 a 和 b 满足:

$$P(G(X_1, \dots, X_n; \theta) \leq a) = P(G(X_1, \dots, X_n; \theta) \geq b) = \alpha / 2$$

正态总体下常见枢轴量

(1) 单个正态总体 $N(\mu, \sigma^2)$ 情形

$$\mu \text{ 的枢轴量 } \begin{cases} \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \sim N(0, 1) & \sigma^2 \text{ 已知} \\ \frac{\sqrt{n}(\bar{X} - \mu)}{S} \sim t(n-1) & \sigma^2 \text{ 未知} \end{cases}$$

$$\sigma^2 \text{ 的枢轴量: } \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1) \quad \mu \text{ 未知}$$

(2) 二个正态总体 $N(\mu_1, \sigma_1^2), N(\mu_2, \sigma_2^2)$ 情形

$\mu_1 - \mu_2$ 的枢轴量

$$\begin{cases} \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0, 1) & \sigma_1^2, \sigma_2^2 \text{ 已知} \\ \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2) & \sigma_1^2 = \sigma_2^2 \text{ 未知} \end{cases}$$

$\frac{\sigma_1^2}{\sigma_2^2}$ 的枢轴量

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

(μ_1, μ_2 未知)

$$\text{其中 } S_w^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}, \quad S_w = \sqrt{S_w^2}$$

正态总体参数的区间估计

单个正态总体情形

设总体 $X \sim N(\mu, \sigma^2)$, $x_1, x_2, \dots, x_n$ 为来自 X 的样本, $\bar{X}$: 样本均值; S^2 : 方差; $1-\alpha$: 置信度
均值 μ 的置信区间

(1) σ^2 已知时

$\bar{X}$ 是 μ 的无偏估计, 且有 $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$, 分布完全已知,

因此可取枢轴量为 $G(X_1, \dots, X_n) = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$.

设常数 $a < b$ 且满足: $P\left\{a < \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < b\right\} = 1-\alpha$

即等价于 $P\left\{\bar{X} - \frac{\sigma}{\sqrt{n}}b < \mu < \bar{X} - \frac{\sigma}{\sqrt{n}}a\right\} = 1-\alpha$

此时区间的长度为 $L = (b-a)\sigma/\sqrt{n}$

根据正态分布的对称性知, 取

$$a = -b = -z_{\alpha/2}$$

时, 区间的长度达到最短.

从而 μ 的同等置信区间为:

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

均值 μ 的置信度 $1-\alpha$ 的置信下限是 $\bar{X} - \frac{\sigma}{\sqrt{n}}z_{\alpha/2}$

(2) σ^2 未知时

取枢轴量为 $\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$,

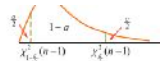
置信区间为 $\left(\bar{X} - \frac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1), \bar{X} + \frac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1)\right)$

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

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

方差 σ^2 的置信区间 (设 μ 未知)

取枢轴量为 $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$



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

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

置信区间为: $\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}, \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right)$

注: 上述所求的区间估计不是最优解.

方差 σ^2 的置信 $1-\alpha$ 的置信上限为

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

两正态总体情形

$X_1, X_2, \dots, X_{n_1}$ 来自 $N(\mu_1, \sigma_1^2)$, $Y_1, Y_2, \dots, Y_{n_2}$ 来自 $N(\mu_2, \sigma_2^2)$,

$\bar{X} = \frac{1}{n_1} \sum_{i=1}^{n_1} X_i, \bar{Y} = \frac{1}{n_2} \sum_{j=1}^{n_2} Y_j$, S_1^2 和 S_2^2 分别为第一、二个

总体的样本方差, 置信度为 $1-\alpha$.

$\mu_1 - \mu_2$ 的置信区间

(1) σ_1^2, σ_2^2 已知时

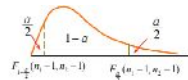
由 $\bar{X} - \bar{Y} \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$ 知,

可取枢轴量为 $\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$

置信区间为: $\left((\bar{X} - \bar{Y}) \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right)$

$\frac{\sigma_1^2}{\sigma_2^2}$ 的置信区间 (设 μ_1, μ_2 未知)

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



置信区间为:

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

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

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

(2) $\sigma_1^2 = \sigma_2^2 = \sigma^2$, σ^2 未知

此时由定理 6.3.4 知, $\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2)$

置信区间为: $\left((\bar{X} - \bar{Y}) \pm t_{\frac{\alpha}{2}}(n_1 + n_2 - 2) S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right)$

其中 $S_w^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1 + n_2 - 2}$, $S_w = \sqrt{S_w^2}$