Ch-07 估计

7. 1 最大似然估计

使用 $X=(X_1,\cdots,X_n)\sim P_\theta,\theta\in\Theta$ 表示统计模型。

引入两类求 $\hat{\theta}(x_1,\cdots,x_n)$ 的方法**——最大似然法**的模型。第一类模型是**离散统计模型**：设 $(X_1,\cdots,X_n)$ 为样本，其中 $X_i(i=1,\cdots,n)$ 为离散型随机变量，样本**分布列**具有下列一般形式

$$P_\theta((X_1,\cdots,X_n)=(x_1,\cdots,x_n))=\prod_{i=1}^n P_\theta(X_i=x_i),\quad \theta\in\Theta$$

此处 θ 为参数。对于固定的样本值 $(x_1,\cdots,x_n)$，作为参数 θ 的函数

$$L(\theta)=\prod_{i=1}^n P_\theta(X_i=x_i)=P_\theta(X_i=x_i,i=1,\cdots,n)$$

称为似然函数。

第二种是**连续统计模型**：此时 $X_i(i=1,\cdots,n)$ 为连续型随机变量，样本 $(X_1,\cdots,X_n)$ 具有联合密度

$$\prod_{i=1}^n p(x_i,\theta),\quad \theta\in\Theta$$

对于固定的样本值 $(x_1,\cdots,x_n)$，θ 的函数

$$L(\theta)=\prod_{i=1}^n p(x_i,\theta)=p_{\theta,\vec{X}}(x_1,\cdots,x_n)$$

也称为**似然函数**。

似然函数 $L(\theta)$ 就是当总体参数为 θ 的情况下，事件 $\{(X_1,\cdots,X_n)=(x_1,\cdots,x_n)\}$ 的概率。最大似然估计就是挑选使 $P_\theta((X_1,\cdots,X_n)=(x_1,\cdots,x_n))$ 达到最大的 θ 值作为真值的估计。

定义 1.1 设 $\theta \in \Theta$ 为统计模型 $(X_1, \cdots, X_n) \sim P_\theta$ 的参数。统计模型可为连续型，也可为离散型。设 $x_1, \cdots, x_n$ 为总体的样本值。若存在 $\hat{\theta}(x_1, \cdots, x_n)$ 使得

$$L(\hat{\theta}(x_1, \cdots, x_n)) = \max_{\theta \in \Theta} L(\theta)$$

则称 $\hat{\theta}(x_1 \cdots, x_n)$ 为 θ 的**最大似然估计**（简称 ML 估计）。

若 $\hat{\theta}$ 为参数 θ 的 ML 估计，则 θ 的函数 $g(\theta)$ 的 ML 估计定义为 $g(\hat{\theta})$.

7.2 矩估计

定义 2.1 设 $X_1, \cdots, X_n$ 为来自总体 $X \sim F_\theta(\theta \in \Theta)$ 的一个样本。若所涉及的矩存在，则

(1) l 阶样本矩 $a_l = \dfrac{1}{n}\sum_{i=1}^{n} X_i^l$ 为相应的总体矩 $\alpha_l = E_\theta(X^l)$ 的矩估计，$l \in N^+$

(2) 若存在连续函数 ϕ 使 $g(\theta) = \phi(\alpha_1, \cdots, \alpha_k)$ 成立，则 $g(\theta)$ 的矩估计定义为

$$\widehat{g(\theta)} = \phi(a_1, \cdots, a_k)$$

其中 a_l 为相应的总体矩 $a_l(l = 1, \cdots, k)$ 的样本矩。

7.3 估计的无偏性

定义 3.1 设 $X_1, \cdots, X_n \sim \mathrm{iid} F(x, \theta)$，$\theta \in \Theta$ 为一个统计模型，$g(\theta)$ 为待估量，统计量 $T(X_1, \cdots, X_n)$ 称为 $g(\theta)$ 的无偏估计，如果 T 满足

$$E_\theta(T(X_1, \cdots, X_n)) = g(\theta), \quad \forall\, \theta \in \Theta$$

无偏估计在平均意义下是准确的。

定理 3.1 设总体 X 的方差 $var(X)$ 存在且为有限，$X_1, \cdots, X_n$ 为 X 的一个样本，则

$$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

不能片面追求无偏性

7.4 无偏估计的优良性

无偏性是对估计的最基本的要求，通常用均方误差作为刻画一个估计偏离目标值的波动的度量。

设 $X_1, \cdots, X_n \sim \mathrm{iid} F_\theta(x) (\theta \in \Theta)$ 为统计模型，$g(\theta)$ 为待估量，$g(\theta)$ 的估计量 $T(X_1, \cdots, X_n)$ 的**均方误差**定义为

$$R(\theta, T) = E_\theta[T(X_1, \cdots, X_n) - g(\theta)]^2$$

均方误差 $R(\theta, T)$ 依赖于未知参数 θ，是刻画估计的性能的一个数量指标。寻找使均方误差越小越好的估计，并希望找到具有一致最小均方误差的估计：对任何参数的真值 θ，该估计相应的均方误差均达到最小。但一致最小均方误差不一定存在，因此所估计的均方误差也被称为**风险函数**。

定义 4.1 设 $X_1, \cdots, X_n \sim \mathrm{iid} F_\theta(x)$，$\theta \in \Theta$，其中 $F_\theta(x)$ 为 $X_i (i = 1, \cdots, n)$ 的共同分布函数。又设 $g(\theta)$ 为一维的待估量，统计量 $T(X_1, \cdots, X_n)$ 为 $g(\theta)$ 的一个估计量。T 称为 $g(\theta)$ 的**最小方差无偏估计**（UMVU 估计），如果满足

1. T 是 $g(\theta)$ 的无偏估计
2. 对于 $g(\theta)$ 的任何其他无偏估计 $\widetilde{T}$，其方差不比 T 的小，即

$$var_\theta(T) \le var_\theta(\widetilde{T}), \quad \forall\, \theta \in \Theta$$

定义 4.2 设 $(X_1, \cdots, X_n) \sim F_\theta(z_1, \cdots, z_k), \theta \in \Theta.$ 统计量 $T(X_1, \cdot, X_n)$ 称为**充分统计量**，如果对于任何其他的统计量 $\widetilde{T}(X_1, \cdots, X_n)$，$\widetilde{T}$ 在 T 的值已知的条件下的条件分布与参数 θ 无关。

定理 4.1 因子分解定理 设 $X_1,\cdots,X_n\sim \mathrm{iid}p(x,\theta),\theta\in\Theta$，其中 $p(x,\theta)$ 在连续情况为分布密度，在离散情况为分布列。若 $(X_1,\cdots,X_n)$ 的联合密度具有下列形式

$$\prod_{i=1}^{n}p(x_i,\theta)=q_\theta(T(x_1,\cdots,x_n))\cdot h(x_1,\cdots,x_n)$$

式中 $h(x_1,\cdots,x_n)$ 与参数 θ 无关，$q_\theta(T)$ 表示它可以写成与 θ 有关，但作为 $x_1,\cdots,x_n$ 的函数具有复合函数 $q_\theta(T(x_1,\cdots,x_n))$ 的形式，则 $T(x_1,\cdots,x_n)$ 是充分统计量。

充分数据量可在不丢失数据所含信息的条件下将数据进行简化。充分统计量不一定唯一。

定义 4.3 设 $X_1,\cdots,X_n\sim \mathrm{iid}p(x,\theta)$，$\theta\in\Theta$，其中 θ 为参数，也可以为向量函数。设 T 为充分统计量。若对任何满足下列条件的统计量 $\phi(T(X_1,\cdots,X_n))$：

$$E_\theta[\phi(T(X_1,\cdots,X_n))]\equiv 0,\quad \forall\,\theta\in\Theta$$

可推知 $P_\theta(\phi(T(X_1,\cdots,X_n))=0)=1,\forall\,\theta\in\Theta$，则称 T 为**完全充分统计量**。

定理 4.2 设 $X_1,\cdots,X_n\sim \mathrm{iid}p(x,\theta),\theta\in\Theta$，$T$ 为完全充分统计量。设 $\phi(T)$ 满足

$$E_\theta[\phi(T)]=g(\theta),\quad \forall\,\theta\in\Theta$$

则 $\phi(T)$ 为 $g(\theta)$ 的 UMVU 估计。

对于待估量 $g(\theta)$，只要找到依赖于完全充分统计量的函数 $\phi(T)$，使得 $\phi(T)$ 是 $g(\theta)$ 的无偏估计，则 $\phi(T)$ 就是 $g(\theta)$ 的 UMVU 估计。因此，寻找 $g(\theta)$ 的 UMVU 估计时，只需要在完全充分统计量的函数中寻找即可。并且，在完全充分统计量 T 存在的前提下，$g(\theta)$ 的 UMVU 估计必定是 T 的函数。

定义 4.4 设 $p(x,\theta),\theta\in\Theta$ 为密度函数或分布列。$\{p(x,\theta),\theta\in\Theta\}$ 称为构成**指数族分布**（或指数型分布），若 $p(x,\theta)$ 能分解成如下几个因子的形式：

$$p(x,\theta)=S(\theta)h(x)\exp\{\sum_{k=1}^{m}C_k(\theta)T_k(\theta)\},\quad \theta\in\Theta$$

引理 4.1 设 $X \sim p(x,\theta), \theta \in \Theta$，其分布族为指数族，又设 $X_1,\cdots,X_n$ 为 X 的一个样本。将 $(X_1,\cdots,X_n)$ 看成随机向量，则 $(X_1,\cdots,X_n)$ 的联合分布密度为

$$\prod_{i=1}^{n} p(x_i,\theta) = S^n(\theta)\prod_{i=1}^{n} h(x_i)\exp\{\sum_{k=1}^{m} C_k(\theta)\cdot\sum_{i=1}^{n} T_k(x_i)\}$$

联合分布也是指数族分布。

定理 4.3 设 X 具有指数族分布，其密度函数为

$$p(x,\theta) = S(\theta)h(x)\exp\{\sum_{k=1}^{m} C_k(\theta)T_k(\theta)\}, \quad \theta \in \Theta$$

若

1. Θ 是 $\mathbb{R}^m$ 中有内点的集合
2. $(C_1,\cdots,C_m)$：$\Theta \to \mathbb{R}^m$ 一一对应的连续函数
3. $C_k(\theta), 1 \le k \le m$，线性无关；$T_k(x), 1 \le k \le m$，线性无关

则 $(T_1(X),\cdots,T_m(X))$ 是该分布族的完全充分统计量。

7.5 估计的相合性

定义 5.1 设 $X_1,\cdots,X_n$ 为来自某总体 $X \sim F_\theta(x)$ 的一个样本，待估参数为 $g(\theta)$，$T_n(X_1,\cdots,X_n)$ 为 $g(\theta)$ 的一个估计。若对任何 $\epsilon > 0$，

$$P_\theta(\mid T_n(X_1,\cdots,X_n) - g(\theta) \mid \ge \epsilon) \to 0, \quad n \to \infty$$

则称估计序列 $T_n(X_1,\cdots,X_n)$ 为 $g(\theta)$ 的**相合估计**，或称 $g(\theta)$ 的估计 $T_n(X_1,\cdots,X_n)$ 具有**相合性**。

估计的相合性是对估计的最基本的要求。

定理 5.1 设 $X_1,\cdots,X_n \sim \mathrm{iid} F_\theta(x), \theta \in \Theta, E_\theta(X_i)\ (i=1,2,\cdots,n)$ 存在且有限，则

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} E_\theta(X_1) \quad (n \to \infty)$$

或等价地，对任何 $\epsilon > 0$，有

$$P_\theta(\mid \overline{X}_n - E_\theta(X_1) \mid \geq \epsilon) \to 0 \quad (n \to \infty)$$

推论 5.1 设 $X_1, \cdots, X_n \sim \text{iid} F_\theta(x), \theta \in \Theta$，则 $\alpha_l = E_\theta(X_1^l)$ 的矩估计 $a_l = \dfrac{1}{n}\sum_{i=1}^n X_i^l$ 为 α_l 的相合估计。

定理 5.2 设 $X_1, \cdots, X_n \sim \text{iid} F_\theta(x), \theta \in \Theta$. 若 θ 的函数 $g(\theta)$ 的矩估计 $\hat{g}_\theta(X_1, \cdots, X_n)$ 存在，则 $\hat{g}_n(X_1, \cdots, X_n)$ 必为 $g(\theta)$ 的相合估计。

只要存在矩估计，就可以保证矩估计的相合性。

设 $X_1, \cdots, X_n \sim \text{iid} N(\mu, \sigma^2)$. 参数 σ/μ 称为正态分布的**变异系数**。

7.6 估计的渐进分布

定理 6.2 中心极限定理 设 $X_i (i = 1, \cdots, n)$ 是独立同分布的，$E(X_i) = \mu$，$var(X_i) = \sigma^2 < \infty$，那么 $\sqrt{n}(\overline{X} - \mu) \xrightarrow{d} N(0, \sigma^2)$.

定理 6.3 Δ 方法 设 T_n 为 θ 的估计。若

$$\sqrt{n}(T_n - \theta) \xrightarrow{d} N(0, \tau^2) \quad (n \to \infty)$$

则对于函数 $h(\theta)$，当 $h'(\theta)$ 存在且不为 0 时，有

$$\sqrt{n}[h(T_n) - h(\theta)] \xrightarrow{d} N(0, \tau^2[h'(\theta)]^2) \quad (n \to \infty)$$

定义 6.2 设 $X_1, \cdots, X_n \sim \text{iid} F_\theta, \theta \in \Theta$，$T(X_1, \cdots, X_n)$ 是 $g(\theta)$ 的估计。若对每一个 $\theta \in \Theta$ 下式成立

$$\sqrt{n}[T(X_1, \cdots, X_n) - g(\theta)] \xrightarrow{d} N(0, \sigma^2) \quad (n \to \infty)$$

则称估计 $T(X_1, \cdots, X_n)$ 是**渐近正态**的，其渐近分布为 $N(0, \sigma^2)$. 渐近分布 $N(0, \sigma^2)$ 的方差 σ^2 称为估计的**渐近方差**（σ^2 也可依赖于 θ）。

7.7 置信区间与置信限

用一个“点” $T(X_1,\cdots,X_n)$ 估计 $g(\theta)$ 的值称为**点估计**，用**置信区间**或**置信限**估计 $g(\theta)$ 的方式称为**区间估计**。

如果给出两个统计量 $\underline{T}$ 和 $\overline{T}$，对很小的正数 α 保证满足 $P(\underline{T}\le a\le\overline{T})\ge 1-\alpha$ 即可。通常称 $1-\alpha$ 为**置信度**或**置信水平**，称 $[\underline{T},\overline{T}]$ 为置信度为 $1-\alpha$ 的**置信区间**。

有的问题只需要考虑**置信上限**或**置信下限**，置信上限和置信下限统称**置信限**。置信区间和置信限都是对参数的一种估计，称为**区间估计**。

定义 7.1 设 $X_1,\cdots,X_n\sim \text{iid}F_\theta\ (\theta\in\Theta)$ 为某统计模型，其中 θ 可为向量参数。又设 $g(\theta)$ 为 θ 的实值函数（在统计中 $g(\theta)$ 也称为参数或一维参数）。

1. 设 $\underline{T}$ 和 $\overline{T}$ 为满足条件 $\underline{T}<\overline{T}$ 的两个统计量，$\alpha\in(0,1)$ 为某常数。若对任意 $\theta\in\Theta$，有
$$P_\theta(\underline{T}\le g(\theta)\le\overline{T})\ge 1-\alpha$$
则称 $[\underline{T},\overline{T}]$ 为 $g(\theta)$ 的置信度是 $1-\alpha$ 的**置信区间**。
2. 设 $\underline{T}$ 为某统计量，$\alpha\in(0,1)$ 为某常数。若对任意 $\theta\in\Theta$，有
$$P_\theta(\underline{T}\le g(\theta))\ge 1-\alpha$$
则称 $\underline{T}$ 为 $g(\theta)$ 的置信度为 $1-\alpha$ 的**置信下限**。
3. 设 $\overline{T}$ 为某统计量，$\alpha\in(0,1)$ 为某常数。若对任意 $\theta\in\Theta$，有
$$P_\theta(g(\theta)\le\overline{T})\ge 1-\alpha$$
则称 $\overline{T}$ 为 $g(\theta)$ 的置信度为 $1-\alpha$ 的**置信上限**。

一、枢轴量法

枢轴量是与参数有关而其分布与参数无关的随机变量。

定义 7.2 设 $X_1,\cdots,X_n\sim \text{iid}F_\theta\ (\theta\in\Theta)$ 为其统计模型，其中 θ 可为向量参数。又设 $g(\theta)$ 为 θ 的实值函数。$g(\theta)$ 的样本 $X_1,\cdots,X_n$ 的函数 $h(X_1,\cdots,X_n;g(\theta))$ 称为**枢轴量**，如果它的分布和 θ 无关。

二、正态分布中参数的置信区间

枢轴量为 $\dfrac{\sqrt{n}(\overline{X}-\mu)}{\sigma}$，由 $P(|\dfrac{\sqrt{n}(\overline{X}-\mu)}{\sigma}| \leq z_{1-\alpha/2}) = 1-\alpha$，得置信区间为 $[\overline{X} - \dfrac{\sigma z_{1-\alpha/2}}{\sqrt{n}}, \overline{X} + \dfrac{\sigma z_{1-\alpha/2}}{\sqrt{n}}]$.

总体模型为 $X \sim N(\mu, \sigma_0^2)$，$\sigma_0^2$ 已知，则 X 的均值 μ 的置信度为 $1-\alpha$ 的置信区间为 $[\overline{X} - \dfrac{\sigma_0 z_{1-\alpha/2}}{\sqrt{n}}, \overline{X} + \dfrac{\sigma_0 z_{1-\alpha/2}}{\sqrt{n}}]$.

设 $X_1, \cdots, X_n \sim \mathrm{iid} N(\mu, \sigma^2), \mu \in (-\infty, +\infty), \sigma^2 > 0$

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

2.
$$\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \overline{X})^2 \sim \chi^2(n-1)$$

3. $\overline{X}$ 与 $\sum_{i=1}^{n}(X_i - \overline{X})^2$ 相互独立，则

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

的分布为自由度 $n-1$ 的 t 分布。

三、参数的近似置信区间

定理 7.2 设 $X_1, \cdots, X_n \sim \mathrm{iid} F_\theta, \theta \in \Theta$，又设 $T(X_1, \cdots, X_n)$ 是 $g(\theta)$ 的渐近正态估计，当 $n \to \infty$ 时，

$$\sqrt{n}[T(X_1, \cdots, X_n) - g(\theta)] \xrightarrow{d} N(0, \sigma^2)$$

1. 若 σ^2 已知，则 $g(\theta)$ 的置信度为 $1-\alpha$ 的近似置信区间为

$$[T(X_1, \cdots, X_n) - \frac{\sigma}{\sqrt{n}} z_{1-\alpha/2}, T(X_1, \cdots, X_n) + \frac{\sigma}{\sqrt{n}} z_{1-\alpha/2}]$$

2. 若 σ^2 未知，则 $g(\theta)$ 的置信度为 $1-\alpha$ 的近似置信区间为

$$[T(X_1,\cdots,X_n)-\frac{\hat{\sigma}}{\sqrt{n}}z_{1-\alpha/2},T(X_1,\cdots,X_n)+\frac{\hat{\sigma}}{\sqrt{n}}z_{1-\alpha/2}]$$

其中 $\hat{\sigma}$ 是 σ 的相合估计，$\hat{\sigma}=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})}$

总结：

设 $X_1,\cdots,X_n\sim \mathrm{iid}N(\mu,\sigma^2),\mu\in(-\infty,+\infty),\sigma^2>0$，求 μ 的置信度为 $1-\alpha$ 的置信区间

1. σ^2 已知：

   设 $(x_1,\cdots,x_n)$ 是总体的样本 $(X_1,\cdots,X_n)$ 的观察值，枢轴量 $h(X_1,\cdots,X_n;\mu)=\frac{\sqrt{n}(\overline{X}-\mu)}{\sigma}\sim N(0,1)$，令 $z_{1-\alpha/2}$ 为标准正态分布的 $1-\alpha/2$ 分位数，则

   $$P(\mid\frac{\sqrt{n}(\overline{X}-\mu)}{\sigma}\mid\leq z_{1-\alpha/2})=1-\alpha$$

   由此可得

   $$P(\overline{X}-\frac{\sigma}{\sqrt{n}}z_{1-\alpha/2}\leq\mu\leq\overline{X}+\frac{\sigma}{\sqrt{n}}z_{1-\alpha/2})=1-\alpha$$

   则参数 μ 的置信度为 $1-\alpha$ 的置信区间为

   $$[\overline{X}-\frac{\sigma}{\sqrt{n}}z_{1-\alpha/2},\overline{X}+\frac{\sigma}{\sqrt{n}}z_{1-\alpha/2}]$$

2. σ^2 未知：

   设 $(x_1,\cdots,x_n)$ 是总体的样本 $(X_1,\cdots,X_n)$ 的观察值，令 $\hat{\sigma}=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2}$，枢轴量 $h(X_1,\cdots,X_n;\mu)=\frac{\sqrt{n}(\overline{X}-\mu)}{\hat{\sigma}}\sim N(0,1)$，令 $t_{1-\alpha/2}(n-1)$ 为自由度为 $n-1$ 的 t 分布的 $1-\alpha/2$ 分位数，则

   $$P(\mid\frac{\sqrt{n}(\overline{X}-\mu)}{\hat{\sigma}}\mid\leq t_{1-\alpha/2}(n-1))=1-\alpha$$

   由此可得

   $$P(\overline{X}-\frac{\hat{\sigma}}{\sqrt{n}}t_{1-\alpha/2}(n-1)\leq\mu\leq\overline{X}+\frac{\hat{\sigma}}{\sqrt{n}}t_{1-\alpha/2}(n-1))=1-\alpha$$

   则参数 μ 的置信度为 $1-\alpha$ 的置信区间为

   $$[\overline{X}-\frac{\hat{\sigma}}{\sqrt{n}}t_{1-\alpha/2}(n-1),\overline{X}+\frac{\hat{\sigma}}{\sqrt{n}}t_{1-\alpha/2}(n-1)],\qquad \text{其中 }\hat{\sigma}=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2}$$

设 $X_1,\cdots,X_n\sim \text{iid}N(\mu,\sigma^2),\mu\in(-\infty,+\infty),\sigma^2>0$，求 μ 的置信度为 $1-\alpha$ 的置信上限

1. σ^2 已知：

   设 $(x_1,\cdots,x_n)$ 是总体的样本 $(X_1,\cdots,X_n)$ 的观察值，枢轴量 $h(X_1,\cdots,X_n;\mu)=\dfrac{\sqrt{n}(\overline{X}-\mu)}{\sigma}\sim N(0,1)$，令 $z_{1-\alpha}$ 为标准正态分布的 $1-\alpha$ 分位数，则

   $$P(\frac{\sqrt{n}(\overline{X}-\mu)}{\sigma}\geq -z_{1-\alpha})=1-\alpha$$

   由此可得

   $$P(\mu\leq \overline{X}+\frac{\sigma}{\sqrt{n}}z_{1-\alpha})=1-\alpha$$

   则参数 μ 的置信度为 $1-\alpha$ 的置信上限为

   $$\overline{X}+\frac{\sigma}{\sqrt{n}}z_{1-\alpha}$$

2. σ^2 未知：

   设 $(x_1,\cdots,x_n)$ 是总体的样本 $(X_1,\cdots,X_n)$ 的观察值，令 $\hat{\sigma}=\sqrt{\dfrac{1}{n-1}\sum\limits_{i=1}^{n}(X_i-\overline{X})^2}$，枢轴量 $h(X_1,\cdots,X_n;\mu)=\dfrac{\sqrt{n}(\overline{X}-\mu)}{\hat{\sigma}}\sim N(0,1)$，令 $t_{1-\alpha}(n-1)$ 为自由度为 $n-1$ 的 t 分布的 $1-\alpha$ 分位数，则

   $$P(\frac{\sqrt{n}(\overline{X}-\mu)}{\hat{\sigma}}\geq -t_{1-\alpha}(n-1))=1-\alpha$$

   由此可得

   $$P(\mu\leq \overline{X}+\frac{\hat{\sigma}}{\sqrt{n}}t_{1-\alpha}(n-1))=1-\alpha$$

   则参数 μ 的置信度为 $1-\alpha$ 的置信上限为

   $$\overline{X}+\frac{\hat{\sigma}}{\sqrt{n}}t_{1-\alpha}(n-1),\qquad \text{其中 }\hat{\sigma}=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2}$$