第七章 参数估计

统计推断

两大基本问题

- 估计
  - 点估计
  - 区间估计
- 假设检验

点估计

估计量的评选标准

区间估计

问题的提出

参数估计是统计推断的基本问题之一，实际工作中碰到的总体，它的分布类型往往是知道的，只是不知道总体分布中的某些参数或全部参数。

例如：考察某灯泡厂生产的一批灯泡的寿命分布。由经验可知灯泡的寿命服从指数分布，但指数分布的具体参数并不知道。我们希望能从该批灯泡中抽取若干灯泡进行试验，利用试验所得寿命数据来推断指数分布中的未知参数。

因此，参数估计问题就是要通过样本来估计总体分布中所包含的未知参数。

参数估计方法

点估计法

以样本的某一函数值作为总体分布中未知参数的估计值。

区间估计法

根据样本给出总体分布中未知参数的具有一定可信度的区间范围。

下面将分别介绍这两种方法。

§1 点估计

一、点估计问题的提法

设总体X的分布函数形式$F(x;\theta)$已知, θ为待估参数, $X_1, X_2, \cdots X_n$是来自X的一个样本, $x_1, x_2, \cdots x_n$为相应的一个样本值, 点估计问题就要构造一个适当的统计量$\theta(X_1, X_2, \cdots X_n)$, 用其观察值$\hat{\theta}(x_1, x_2, \cdots x_n)$来估计未知参数, 称

$\hat{\theta}(X_1, X_2, \cdots X_n)$为$\theta$的估计量(Estimator);

$\hat{\theta}(x_1, x_2, \cdots x_n)$为$\theta$的估计值(Estimated Value)。

说 明

① $F(x;\theta)$ 中待估参数θ可以是1个或多个参数；

② 估计量和估计值统称为估计，都简记为 $\hat{\theta}$ ；

③ 由于估计量是样本的函数，故对不同的样本值，得到待估参数的估计值往往也是不同的。

点估计问题的关键在于构造合适的统计量，其方法很多，我们仅介绍两种常用的方法：

点估计方法
- 矩估计法
- 最(极)大似然估计法

二、矩估计法

设RV X，其分布律或概率密度含有**k个待估参数** $\theta_1, \theta_2, \ldots, \theta_k$，

离散总体 $P\{X = x_i\} = p(x_i; \theta_1, \ldots, \theta_k),\ i = 1, 2, \ldots$

连续总体 PDF为 $f(x; \theta_1, \ldots, \theta_k)$

假设总体X 的前k 阶矩存在，即

$$\mu_l = E\left(X^l\right) = \begin{cases} \sum_{i=1}^{n} x_i^l\, p(x_i; \theta_1, \theta_2, \cdots, \theta_k), & X \text{ 为离散型RV} \\ \int_{-\infty}^{+\infty} x^l f(x; \theta_1, \theta_2, \cdots, \theta_k)\,\mathrm{d}x, & X \text{ 为连续型RV} \end{cases}$$

$(l = 1, 2, \cdots k)$

$$\mu_l = E\left(X^l\right) = \begin{cases} \sum_{i=1}^{n} x_i^l p(x_i; \theta_1, \ldots, \theta_k) \\ \int_{-\infty}^{+\infty} x^l f(x; \theta_1, \ldots, \theta_k) \mathrm{d}x \end{cases}$$

可以发现总体矩 μ_l 一般是θ的函数，则

$$\theta_l = \theta_l(\mu_1, \mu_2, \ldots, \mu_k), \quad l = 1, 2, \ldots, k$$

若能估计出总体矩 μ_l，便可估计出 θ_l。

设 $X_1, X_2, \ldots, X_n$是来自X 的样本，其样本矩为

$$A_l = \frac{1}{n} \sum_{i=1}^{n} X_i^l, \quad l = 1, 2, \cdots k$$

在第六章中，我们知道样本矩与总体矩之间的关系：

$$A_l \xrightarrow{P} \mu_l$$

$$g(A_1, A_2, \cdots A_k) \xrightarrow{P} g(\mu_1, \mu_2, \cdots \mu_k)$$

因此，要估计$\theta_l = \theta_l(\mu_1, \mu_2, ..., \mu_k)$，$l=1,2, \ldots, k$,

用样本矩作为相应总体矩的估计量，而用样本矩的连续函数作为相应总体矩的连续函数的估计量，这种得到θ_l估计的方法就称为矩估计法。

矩估计法的基本步骤

① 设总体 X 的分布中含有k 个待估参数 $\theta_1,\theta_2,\cdots,\theta_k$ 分别写出前k阶总体矩和样本矩；

$$\mu_l = E\left(X^l\right) = \mu_l(\theta_1,\theta_2,\cdots,\theta_k),\quad l=1,2,\cdots k$$

$$A_l = \frac{1}{n}\sum_{i=1}^{n} X_i^l,\quad l=1,2,\cdots k$$

② 令相应的总体矩和样本矩相等，建立方程/组：

$$\mu_l(\theta_1,\theta_2,\cdots,\theta_k) = A_l,\ l=1,2,\cdots k$$

③ 解方程，得到未知参数的矩估计量：

$$\hat{\theta}_l = \hat{\theta}_l(A_1,A_2,\cdots,A_k),\ l=1,2,\cdots k$$

以$\hat{\theta}_l = \hat{\theta}_l(A_1, A_2, \cdots, A_k)$作为参数$\theta_l$的估计量，称为**矩估计量**，其观察值称为**矩估计值**。

说明：

☞ 矩估计法的**基本思想**是：基于未知参数与总体矩之间存在的函数关系，用样本矩来代替相应总体矩，从此来求出未知参数的估计量。

☞ 利用依概率收敛的性质知：一般

$$\hat{\theta}_l(A_1, A_2, \cdots, A_k) \xrightarrow{P} \theta_l(\mu_1, \mu_2, \cdots, \mu_k)$$ 真值

例1　设总体X 的均值 $E(X)=\mu$，方差 $D(X)=\sigma^2$ 都存在，但$\mu,\sigma^2(>0)$未知，$X_1,X_2,\ldots,X_n$是来自X 的样本，试求μ、σ^2的矩估计量。

解：(1)求总体矩(前2阶)

$$\mu_1=E(X)=\mu$$

$$\mu_2=E(X^2)=D(X)+E^2(X)=\mu^2+\sigma^2$$

样本矩(前2阶)：$A_1=\frac{1}{n}\sum_{i=1}^{n}X_i=\bar{X}$,　　$A_2=\frac{1}{n}\sum_{i=1}^{n}X_i^2$

(2)令$\begin{cases}\mu_1=A_1\\ \mu_2=A_2\end{cases}$　　即$\begin{cases}\mu=\bar{X}\\ \mu^2+\sigma^2=\frac{1}{n}\sum_{i=1}^{n}X_i^2\end{cases}$

(3) 解方程组 $\begin{cases} \mu = \bar{X} \\ \mu^2 + \sigma^2 = \dfrac{1}{n}\sum_{i=1}^{n} X_i^2 \end{cases}$

得 $\begin{cases} \hat{\mu} = \bar{X} \\ \hat{\sigma}^2 = \dfrac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2 = \dfrac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 \end{cases}$

结 论

对任意分布的总体，若总体均值与方差都存在，

则**总体均值的矩估计量是样本均值**；

总体方差的矩估计量是样本的二阶中心矩。

如：① $X \sim N(\mu,\sigma^2)$，试求μ、σ^2的矩估计量？

② $X \sim \pi(\lambda)$，求λ的矩估计量？

由于$E(X)=\lambda$，λ的一个矩估计为$\hat{\lambda}=\bar{X}$

又由于$D(X)=\lambda$，λ的另一个矩估计为

$$\hat{\lambda}=\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2$$

由此可见：**一个参数的矩估计量可能不惟一。**

究竟采用哪一个作为参数λ的矩估计量？

阶数低的矩估计、估计量的评选标准

矩估计量的性质（不变性）

若$\hat{\theta}$为θ的矩估计量，$g(\theta)$为θ的已知连续函数，

则$g\left(\hat{\theta}\right)$也是$g(\theta)$的矩估计量。

矩估计法小结

优点：无论总体为何种分布，只要总体矩存在，就可以得到总体分布中未知参数的估计量.

缺点：总体矩不存在，就不能用矩估计法；
当总体分布类型已知时，矩估计法没有充分利用总体分布提供的信息。

三、最(极)大似然估计法

1. 基本思想

引例1　甲乙两个射手，甲的命中率为0.9，乙的命中率为0.1，现在他们中的一个人向目标射击了一发，结果命中，估计是谁射击的？

一般认为，只发了一枪便打中，且甲的命中率比乙的高，看来这一枪是甲射中的。

三、最(极)大似然估计法

1. 基本思想

引例2　一车上装有大小不同的两筐水果。第一筐中有85%是香蕉，15%是桔子；第二筐中15%是苹果，85%是桔子。忽然从车上掉下一个桔子，此时让你估计它是从那一筐里掉出来的？

一般认为是从第二筐里调出来的可能性大。

在已得试验结果情况下，我们应该寻找使这个结果出现可能性最大的事件作为估计结果。

这个引例所作的推断就体现了最大似然估计的基本思想 ：概率最大的事件最有可能发生。

针对点估计问题，最大似然估计的基本思想：在参数θ的取值范围内挑选使样本取其观察值的概率最大的参数。

在具体介绍最大似然估计法之前，先介绍一下似然函数的定义。

2. 似然函数

设$X_1, X_2, \ldots, X_n$是来自总体X的一个样本，

$x_1, x_2, \ldots, x_n$是其样本值。

➢ 若X为离散型RV, 其分布律为$P\{X = x_i\} = p(x_i;\theta)$, θ为待估参数, $\theta \in \Theta$ (Θ为θ可能取值的范围), 则样本的似然函数为

$$L(\theta) = L(x_1, x_2, \cdots, x_n;\theta) = \prod_{i=1}^{n} p(x_i;\theta)$$

意义 **样本$X_1, X_2, \ldots, X_n$取其观察值$x_1, x_2, \ldots, x_n$的概率**

➢ 若X为连续型RV, 其PDF为$f(x;\theta)$, θ为待估参数, $\theta \in \Theta$, 则样本的似然函数为

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

意 义　**样本点$(X_1, X_2, \dots, X_n)$落在点$(x_1, x_2, \dots, x_n)$的邻域(边长分别为$dx_1, dx_2, \dots, dx_n$的n维立方体)内的近似概率** $\prod_{i=1}^{n} f(x_i;\theta)\,\mathrm{d}x_i \left(\text{其中因子}\prod_{i=1}^{n}\mathrm{d}x_i\text{不随}\theta\text{变化}\right)$ **仅随θ变化的部分。**

3. 最大似然估计法

引入似然函数后，最大似然估计法的**基本思想**就变成为：

选取$\hat{\theta}\in\Theta$使得似然函数$L(\theta)$在$\hat{\theta}$处取最大值

定 义

若$L\left(x_1,x_2,\cdots,x_n;\hat{\theta}\right)=\max\limits_{\theta\in\Theta}L\left(x_1,x_2,\cdots,x_n;\theta\right)$

则称$\hat{\theta}(x_1,x_2,\cdots,x_n)$为参数$\theta$的**最大似然估计值**；

$\hat{\theta}(X_1,X_2,\cdots,X_n)$为参数$\theta$的**最大似然估计量**。

$$\left.\begin{array}{l}\text{最大似然估计值}\\ \text{最大似然估计量}\end{array}\right\}\text{最大似然估计 MLE}$$

（Maximum Likehood Estimator）

多数情况下，总体的分布关于待估参数可微，因此，可以利用微分学中求最值的方法来获得参数的MLE。于是给出如果下求解过程。

最大似然估计法的一般步骤

① 写出样本的似然函数 $L(\theta)$；

② 求解对数似然函数 $\ln L(\theta)$；

③ 求对数似然函数的导数 $\dfrac{d\ln L(\theta)}{d\theta}$；

④ 建立对数似然方程 $\dfrac{d\ln L(\theta)}{d\theta}=0$；

⑤ 解对数似然方程，得到θ的MLE $\hat{\theta}_{\text{MLE}}$。

说 明

☞ 考虑到$L(\theta)$是连乘积的形式，为便于求驻点，可取对数$\ln L(\theta)$，化为累加和形成后再求导比较方便。

$L(\theta)>0$，由于$\ln x$是x的单调增函数，则$\ln L(\theta)$与$L(\theta)$在θ的同一值处达到各自的最大值，故可以利用对数似然方程来获得似然函数取最大值的解。

说 明

☞ 驻点可能是最大值点，也可能是最小值点，因此，必须避免使用最小值点。

☞ 当待估参数有多个时，可以令对数似然函数对每个待估参数分别求偏导，来建立对数似然方程组，以此获得待估参数的最大似然估计。

☞ 如果似然函数没有驻点、或其偏导数不存在、甚至本身不连续时，上述利用求导方法获得待估参数的MLE是行不通的，这时则直接用最大似然估计法的基本思想来求解。

例2　设总体$X \sim b(1,p)$，$X_1, X_2, \dots, X_n$是来自X的样本，试求p的MLE？

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

似然函数 $L(p)=\prod_{i=1}^{n}P\{X=x_i\}=p^{\sum_{i=1}^{n}x_i}(1-p)^{n-\sum_{i=1}^{n}x_i}$

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

求偏导数 $\dfrac{d\ln L(p)}{dp}=\dfrac{\sum_{i=1}^{n}x_i}{p}-\dfrac{n-\sum_{i=1}^{n}x_i}{1-p}$　令其$=0$解得

$\hat{p}_{MLE}=\frac{1}{n}\sum_{i=1}^{n}x_i=\bar{x}$　故p的最大似然估计为$\bar{X}$

对比p的估计量$\hat{p}_M=\bar{X}$，发现p **的MLE与矩估计相同**。

例3　设总体$X \sim N(\mu,\sigma^2)$，μ、σ^2未知，若 $X_1, X_2, \ldots, X_n$是来自X 的样本，试求μ、σ^2的MLE？

解：X 的PDF为

$$f\left(x;\mu,\sigma^2\right)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\},\quad -\infty<x<+\infty$$

似然函数

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

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

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

求偏导数 $\begin{cases} \dfrac{\partial \ln L(\mu,\sigma^2)}{\partial \mu} = \dfrac{1}{\sigma^2}\displaystyle\sum_{i=1}^{n}(x_i-\mu) \\ \dfrac{\partial \ln L(\mu,\sigma^2)}{\partial \sigma^2} = -\dfrac{n}{2\sigma^2} + \dfrac{1}{2(\sigma^2)^2}\displaystyle\sum_{i=1}^{n}(x_i-\mu)^2 \end{cases}$

解对数似然方程 $\begin{cases} \dfrac{\partial \ln L(\mu,\sigma^2)}{2\mu} \overset{令}{=} 0 \\ \dfrac{\partial \ln L(\mu,\sigma^2)}{2\sigma^2} \overset{令}{=} 0 \end{cases} \Rightarrow \begin{cases} \hat{\mu} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}x_i = \bar{x} \\ \hat{\sigma}^2 = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}(x_i-\bar{x})^2 \end{cases}$

故μ、σ^2的MLE分别为 $\hat{\mu}_{\text{MLE}} = \bar{X}$, $\hat{\sigma}^2_{\text{MLE}} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}(X_i-\bar{X})^2$

说明 **μ、σ^2的MLE与其矩估计相同**

例4　设总体$X\sim U(a,b)$，a、b未知，若 $X_1, X_2, \dots,$ X_n是来自X 的样本，试求a、b 的矩估计量和MLE？

解：(1) 总体矩

$$\mu_1 = E(X) = \frac{a+b}{2}$$

$$\mu_2 = E(X^2) = D(X) + E^2(X) = \frac{(b-a)^2}{12} + \left(\frac{a+b}{2}\right)^2$$

样本矩：$A_1 = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X},\qquad A_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$

令$\begin{cases}\mu_1 = A_1\\ \mu_2 = A_2\end{cases}$　解得$\begin{cases}\hat{a} = \bar{X} - \sqrt{\frac{3}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2}\\ \hat{b} = \bar{X} + \sqrt{\frac{3}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2}\end{cases}$

(2) 求MLE　　PDF为 $f(x;\theta)=\begin{cases}\dfrac{1}{b-a}, & a\le x\le b\\ 0\ , & else\end{cases}$

似然函数

$$L(a,b)=\prod_{i=1}^{n} f(x_i;\theta)=\begin{cases}\dfrac{1}{(b-a)^n}, & a\le x_i\le b, i=1,2,\cdots n\\ 0 & else\end{cases}$$

求偏导数　$\dfrac{\partial L(\theta)}{\partial a}=\dfrac{n}{(b-a)^{n+1}},\quad \dfrac{\partial L(\theta)}{\partial b}=\dfrac{-n}{(b-a)^{n+1}}$

令它们为0，则 *a*、*b* 至少有一个为∞，此结果无意义，这是由于似然函数在最大值处的导数不为0。

因此，不能利用求导方法来获得a、b的MLE，只能利用**最大化似然函数原则**来求解。

$$L(a,b)=\begin{cases}\dfrac{1}{(b-a)^n}, & a\le x_i\le b, i=1,2,\cdots n\\ 0 & else\end{cases}$$

为使 $L(a, b)$ 达到max，则要使 $(b-a)$ 达到min，

即 $b\downarrow$和 $a\uparrow$；又考虑到 $a\le x_1, x_2, \ldots, x_n \le b$，

但 b不能小于$\max\{x_1, x_2, \ldots, x_n\}$，

a不能大于$\min\{x_1, x_2, \ldots, x_n\}$，否则$L(a, b)=0$。

因此，***a***、***b***的最大似然估计值为

$$\hat{a} = \min\{x_1, \ldots, x_n\} = x_{(1)}$$

$$\hat{b} = \max\{x_1, \ldots, x_n\} = x_{(n)}$$

a、***b***的最大似然估计量为 $\hat{a}_{\mathrm{MLE}} = X_{(1)}$，$\hat{b}_{\mathrm{MLE}} = X_{(n)}$

对比***a***, ***b***的矩估计量 $\hat{a}_M = \bar{X} - \sqrt{\frac{3}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2}$

$$\hat{b}_M = \bar{X} + \sqrt{\frac{3}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

显然，***a***、***b***的MLE与矩估计不同。

思考

① 待估参数的最大似然估计是否一定存在？

② 若存在，是否唯一？

下面通过一个例题来回答这两个问题。

例 设总体$X\sim U\left(\theta-\frac{1}{2},\theta+\frac{1}{2}\right)$, 样本$X_1,X_2,\cdots,X_n$, 求$\hat{\theta}_{\mathrm{MLE}}$?

解：似然函数 $L(\theta)=\begin{cases}1, & \theta-\frac{1}{2}\le x_i\le\theta+\frac{1}{2}, i=1,2,\cdots n\\ 0 & else\end{cases}$

$L(\theta)$仅取0和1两个值，故凡使$L(\theta)$=1的θ都可作为θ的MLE。此时，

$$x_{(n)}-\frac{1}{2}\le\theta\le x_{(1)}+\frac{1}{2}$$

因此，θ的MLE可能不存在、也可能唯一存在、也可能有很多个。

4. 最大似然估计的性质

设$\hat{\theta}$是总体分布中参数θ的最大似然估计，函数 $u=u(\theta),\theta\in\Theta$具单值反函数$\theta=\theta(u),u\in U$，则 $\hat{u}=u(\hat{\theta})$是$u(\theta)$的最大似然估计。

这一性质称为最大似然估计的不变性。

此性质说明：当待估参数θ的函数$u(\theta)$存在单值反函数时，则待估参数函数的MLE在待估参数取MLE处取得。

此外，最大似然估计的不变性可推广至总体分布中含有多个未知参数的情况。

证：当$\hat{\theta}$为θ的MLE时，下式成立，

$$L\left(x_1, x_2, \cdots, x_n; \hat{\theta}\right) = \max_{\theta \in \Theta} L\left(x_1, x_2, \cdots, x_n; \theta\right) \quad (*)$$

当$\theta = \hat{\theta}$，$\hat{u} = u(\hat{\theta})$，

由于$u = u(\theta)$有单值反函数$\theta = \theta(u)$

所以$\hat{\theta} = \theta(\hat{u})$　　　　则(*)式可写成

$$L\left(x_1, x_2, \cdots, x_n; \theta(\hat{u})\right) = \max_{u \in U} L\left(x_1, x_2, \cdots, x_n; \theta(u)\right)$$

$\therefore \hat{u} = u(\hat{\theta})$是$u(\theta)$的MLE。　得证

举例说明

正态分布中方差σ^2的MLE为$\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2$

而$\sigma=u=\sqrt{\sigma^2}$有单值反函数$\sigma^2=u^2\left(u>0\right)$,

故σ的MLE为$\hat{\sigma}=\sqrt{\hat{\sigma}^2}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2}$

最大似然估计法小结

最大似然估计法的**特点**：

- ☞ 需要已知总体的分布，才能获得估计量。
- ☞ 最大似然估计法中，需求解对数似然方程(组)，然而除了一些简单情况，多数情况下(例如二参数威布尔分布)求解对数似然方程(组)比较困难，往往需要在计算机上通过一些**数值迭代算法**才能求出近似解，如牛顿—拉弗森算法，拟牛顿法等。

小结

- 本节介绍了点估计问题的两种求解方法：

  矩估计法、最大似然估计法

作业

Pages 173, 174：

第1, 2, 3, 5, 7题

练习：设$X_1,\ldots,X_n$为来自总体$X\sim U(\theta,1)$,的一个样本, 其中θ为未知参数, 若θ的MLE为$\hat{\theta}_{MLE}=X_{(1)}$, 试求$\mathrm{P}(\theta<X\le 1/2)$的MLE。

$$解:\mathrm{P}\left(\theta<X\le\frac{1}{2}\right)=\int_{\theta}^{\frac{1}{2}}f(x)\mathrm{d}x=\int_{\theta}^{\frac{1}{2}}\frac{1}{1-\theta}\mathrm{d}x=\frac{1-2\theta}{2-2\theta}$$

由于函数$u=\dfrac{1-2\theta}{2-2\theta}$具有单值反函数$\theta=\dfrac{1-2u}{2-2u}$,由MLE的不变性知

$$\mathrm{P}\left(\theta<X\le\frac{1}{2}\right)=\frac{1-2\theta}{2-2\theta}的MLE为$$

$$\hat{\mathrm{P}}\left(\theta<X\le\frac{1}{2}\right)=\frac{1-2\hat{\theta}_{\mathrm{MLE}}}{2-2\hat{\theta}_{\mathrm{MLE}}}=\frac{1-2X_{(1)}}{2-2X_{(1)}}$$