第七章 参数估计

什么是参数估计?

参数是刻画总体某方面概率特性的数量.

当此数量未知时, 从总体抽出一个样本, 用某种方法对这个未知参数进行估计就是参数估计.

例如, $X \sim N(\mu, \sigma^2)$,

若μ, σ^2未知, 通过构造样本的函数, 给出它们的估计值或取值范围就是参数估计的内容.

点估计

区间估计

第七章 参数估计

第一节 矩估计法

第二节 极大似然估计法

- 估计量的定义
- 构造估计量的方法
  - 矩估计法
  - 极大似然法

一. 估计量的定义

设： 总体X的分布中含有未知参数θ

$X_1,X_2,\cdots,X_n$是总体X的一个样本

$x_1,x_2,\cdots,x_n$为样本值

点估计： 构造一个适当的统计量$\hat{\theta}(X_1,X_2,\cdots,X_n)$

用这个统计量的观察值$\hat{\theta}(x_1,x_2,\cdots,x_n)$

来估计未知参数θ

称： $\hat{\theta}(X_1,X_2,\cdots,X_n)$为$\theta$**估计量**

$\hat{\theta}(x_1,x_2,\cdots,x_n)$ 为θ**估计值**

估计量与估计值统称为**估计**

二. 构造统计量的方法

用样本的各阶矩来估计总体的各阶矩

1. 矩估计法

总体X的分布中含一个未知参数θ

设 总体X一阶矩存在 $E(X)=\mu(\theta)$ 数学期望

样本一阶矩为 A_1

$$\mu(\theta)=A_1$$

解 $\hat{\theta}=\hat{\theta}(X_1,X_2,\cdots,X_n)$矩估计**量**

$\hat{\theta}(x_1,x_2,\cdots,x_n)$矩估计**值**

例1 设总体 $X \sim \begin{pmatrix} 1 & 2 & 3 \\ 1-\theta^2 & \frac{\theta^2}{2} & \frac{\theta^2}{2} \end{pmatrix}$，$\theta(0<\theta<1)$未知.

现得一样本值1, 3, 2, 3, 求 θ 矩估计值.

解 设 X_1, X_2, X_3, X_4 是总体X 的一个样本

$$(x_1, x_2, x_3, x_4) = (1,3,2,3), \quad \bar{X} = \frac{1}{4}\sum_{i=1}^{4} X_i, \quad \bar{x} = \frac{9}{4}$$

总体一阶矩 $E(X) = 1 + \frac{3}{2}\theta^2$

样本一阶矩 $A_1 = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}$

令 $\boxed{1 + \frac{3}{2}\theta^2 = A_1} \longrightarrow \hat{\theta} = \sqrt{\frac{2}{3}(A_1 - 1)} = \sqrt{\frac{2}{3}(\bar{X} - 1)}$

矩估计值 $\hat{\theta} = \sqrt{\frac{2}{3}(\bar{x} - 1)} = \sqrt{\frac{5}{6}}$

例2 设总体X密度 $f(x)=\begin{cases}(\theta+1)x^{\theta}, & 0<x<1,\\ 0, & \text{其他}\end{cases}$

$\theta>-1$未知. 求 θ 的矩估计.

解 设 $X_1,X_2,\cdots,X_n$ 是总体X的一个样本

$x_1,x_2,\cdots,x_n$ 为样本值

$$E(X)=\int_{-\infty}^{\infty}xf(x)dx=\int_0^1(\theta+1)x^{\theta+1}dx=\frac{\theta+1}{\theta+2}$$

样本一阶矩 $A_1=\frac{1}{n}\sum_{i=1}^{n}X_i=\overline{X}$

令 $\boxed{\frac{\theta+1}{\theta+2}=A_1}$ $\longrightarrow$ $\hat{\theta}=\frac{2A_1-1}{1-A_1}=\frac{2\overline{X}-1}{1-\overline{X}}$ 矩估计量

$\hat{\theta}=\frac{2\overline{x}-1}{1-\overline{x}}$ 矩估计值

矩估计法

2. 总体X的分布中含两个未知参数θ_1,θ_2.

设 总体一,二阶矩均存在 $E(X)=\mu_1(\theta_1,\theta_2)$

$$E(X^2)=\mu_2(\theta_1,\theta_2)$$

样本一,二阶矩为 A_1,A_2 $\quad A_1=\frac{1}{n}\sum_{i=1}^{n}X_i=\overline{X}\quad A_2=\frac{1}{n}\sum_{i=1}^{n}X_i^2$

令 $\begin{cases}\mu_1(\theta_1,\theta_2)=A_1\\\mu_2(\theta_1,\theta_2)=A_2\end{cases}\Rightarrow\begin{cases}\hat\theta_1(X_1,X_2,\cdots,X_n)\\\hat\theta_2(X_1,X_2,\cdots,X_n)\end{cases}$ 矩估计**量**

$$\begin{cases}\hat\theta_1(x_1,x_2,\cdots,x_n)\\\hat\theta_2(x_1,x_2,\cdots,x_n)\end{cases}$$ 矩估计**值**

例3 设总体X服从均匀分布$U(a,b)$，a,b未知.

$X_1,X_2,\cdots,X_n$是来自总体X的一个样本.

求a,b的矩估计量.

解 $$E(X)=\frac{a+b}{2}$$

$$E(X^2)=D(X)+E^2(X)$$

$$E(X^2)=D(X)+\left(E(X)\right)^2=\frac{(b-a)^2}{12}+\frac{(a+b)^2}{4}$$

令 $$\begin{cases}\dfrac{a+b}{2}=A_1\\[2ex]\dfrac{(b-a)^2}{12}+\dfrac{(a+b)^2}{4}=A_2\end{cases}$$

$$A_1=\frac{1}{n}\sum_{i=1}^{n}X_i=\overline{X}$$

$$A_2=\frac{1}{n}\sum_{i=1}^{n}X_i^2$$

例3 设总体X服从均匀分布$U(a,b)$, a,b未知. $X_1,X_2,\cdots,X_n$是来自总体X的一个样本. 求a,b的矩估计量.

解

$$\begin{cases} \dfrac{a+b}{2}=A_1 \\ \dfrac{(b-a)^2}{12}+\dfrac{(a+b)^2}{4}=A_2 \end{cases}$$

$$A_2-A_1^2=\frac{(b-a)^2}{12}$$

$$\sqrt{3(A_2-A_1^2)}=\frac{b-a}{2}$$

$$\begin{cases} \hat{a}=A_1-\sqrt{3(A_2-A_1^2)}=\bar{X}-\sqrt{3\left(\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i^2-\bar{X}^2\right)} \\ \hat{b}=A_1+\sqrt{3(A_2-A_1^2)}=\bar{X}+\sqrt{3\left(\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i^2-\bar{X}^2\right)} \end{cases}$$

矩估计量

例4. 设总体X的均值为μ，方差为σ^2都存在，且

$\sigma^2>0$, $X_1, X_2, \cdots X_n$是总体 X 的一个样本

(1). μ，σ^2均未知, 求：μ, σ^2的矩估计量

(2)当总体(某种灯泡寿命) $X \sim N(\mu, \sigma^2)$ μ, σ^2未知，

今取 4 只灯泡， 测得其寿命(小 时), 如下：

1502, 1453, 1367, 1650 （小时）

求：μ, σ^2的矩估计值

例5. 设总体X的均值为 μ,方差为 σ^2 都存在，且 $\sigma^2>0$, $X_1, X_2, \cdots X_n$ 是总体 X 的一个样本

(1). μ，σ^2 均未知, 求： μ,σ^2 的矩估计量

$$E(X^i)=A_i$$

$$i=1,2,\cdots,k$$

$$A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k$$

解： $E(X)=\mu$

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

现令： $\begin{cases} \mu = A_1 \\ \sigma^2+\mu^2 = A_2 \end{cases}$

解得: $\begin{cases} \hat{\mu}=A_1=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i=\bar{X} \\ \hat{\sigma}^2=A_2-A_1^2=\dfrac{1}{n}\sum\limits_{i=1}^{n}(X_i-\bar{X})^2 \end{cases}$

解得: $\begin{cases} \hat{\mu} = A_1 = \overline{X} \\ \hat{\sigma}^2 = A_2 - A_1^2 = \dfrac{1}{n}\sum\limits_{i=1}^{n}\left(X_i - \overline{X}\right)^2 \end{cases}$

从而得 μ, σ^2 的矩估计量为: $\hat{\mu} = \overline{X}$

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

结论: 不论总体X服从什么分布，总体均值$E(X) = \mu$

与方差$D(X) = \sigma^2$的矩估计量的表达式是相同的。

$$\hat{\mu} = \overline{X}, \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

例4 设总体X的均值为 μ, 方

$\sigma^2 > 0$, $X_1, X_2, \cdots X_n$ 是总体 X 的一个样本

(2).当总体(某种灯泡寿命) $X \sim N(\mu, \sigma^2)$ μ, σ^2未知,

今取 4 只灯泡, 测得其寿命(小 时), 如下:

1502, 1453, 1367, 1650 （小时）求: μ, σ^2的矩估计值

解: $\overline{X} = \frac{1}{4}(1502 + 1453 + 1367 + 1650) = 1493$

$$\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2 = \frac{1}{4}[(1502-1493)^2 + (1453-1493)^2 + (1367-1493)^2 + (1650-1493)^2] = 10551$$

$\therefore$ 灯泡寿命的均值与方差的矩估计值为:

$\hat{\mu} = 1493$, $\hat{\sigma}^2 = 10551$

矩估计法

3. 总体X的分布中含 k 个未知参数 $\theta_1,\theta_2,\cdots,\theta_k$

设 总体前k阶矩均存在, $E\left(X^l\right)=\mu_l\left(\theta_1,\theta_2,\cdots,\theta_k\right)$

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

样本前 k 阶矩为 $A_1,A_2,\cdots,A_k$

令 $\begin{cases}\mu_1\left(\theta_1,\theta_2,\cdots,\theta_k\right)=A_1\\ \mu_2\left(\theta_1,\theta_2,\cdots,\theta_k\right)=A_2\\ \quad\vdots\\ \mu_k\left(\theta_1,\theta_2,\cdots,\theta_k\right)=A_k\end{cases}$ $\Longrightarrow$ $\hat{\theta}_i\left(X_1,X_2,\cdots,X_n\right)$

矩估计量

$i=1,2,\cdots,k$

矩估计法的优点：

简单易行，使用时并不需要事先知道总体的分布。

矩估计法的缺点：

在总体分布已知的场合,没有充分利用分布的信息。

2. 极大似然法

极大似然法是在总体类型已知条件下使用的一种参数估计方法 。

这个方法是英国统计学家费歇尔（ **Fisher** ）在 **1922** 年提出来的， 是一种应用非常广泛的统计方法。

想法

选择参数的估计值， 使观测到的样本出现的概率最大

最大可能性估计

Fisher

极大似然法的基本思想：

引例

已知袋中有100个大小、形状完全一样的球，分黑白两种颜色。不知道是黑球多还是白球多，但知道两种颜色球的数量比为95:5。现从中随机地抽取一个球，发现是黑球。

问：袋中黑球多，还是白球多？

设 $X = \begin{cases} 1, & \text{取到黑球} \\ 0, & \text{取到白球} \end{cases}$

$X \sim (0,1)$

X	0	1
p_k	$1-p$	p

已知样本 X_1的值 $x_1 = 1$

问题：$\hat{p} = 0.95 \; or \; 0.05$?

应选择参数的估计值，使样本值出现的概率达到最大。

1. 离散型

设总体X的分布律: $P\{X=x\}=p(x;\theta)$, $\theta\in\Theta$, θ 未知

$X_1,X_2,\cdots,X_n$是总体X的一个样本

$x_1,x_2,\cdots,x_n$为样本值

$$P\{(X_1,X_2,\cdots,X_n)=(x_1,x_2,\cdots,x_n)\}$$

$$=P\{X_1=x_1,\ X_2=x_2,\ \cdots,\ X_n=x_n\}$$

$$=P\{X_1=x_1\}P\{X_2=x_2\}\cdots P\{X_n=x_n\}$$

$$=P\{X=x_1\}P\{X=x_2\}\cdots P\{X=x_n\}$$

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

1. 离散型

设总体X的分布律: $P\{X=x\}=p(x;\theta)$, $\theta\in\Theta$, θ 未知

$X_1,X_2,\cdots,X_n$是总体X的一个样本

$x_1,x_2,\cdots,x_n$ 为样本值

称 $L(\theta)=\prod_{i=1}^{n}p(x_i;\theta)$为样本的似然函数

$$\frac{d}{d\theta}L(\theta)=0$$

$L(\theta)$的意义——样本值出现的概率

令 $\frac{d}{d\theta}\ln L(\theta)=0$ $\quad\theta\in\Theta$

解 $\hat{\theta}(x_1,x_2,\cdots,x_n)$ 最大似然估计值

$\hat{\theta}(X_1,X_2,\cdots,X_n)$ 最大似然估计量

例5 设总体 $X \sim \begin{pmatrix} 1 & 2 & 3 \\ 1-\theta^2 & \dfrac{\theta^2}{2} & \dfrac{\theta^2}{2} \end{pmatrix}$, $\theta(0<\theta<1)$未知.

现得一样本值1, 3, 2, 3, 求 θ 最大似然估计值.

解 设 X_1, X_2, X_3, X_4 是总体X的一个样本

而 $(x_1, x_2, x_3, x_4) = (1,3,2,3)$ 是一个样本值

$$P\{(X_1, X_2, X_3, X_4) = (1,3,2,3)\}$$

选择参数的估计值，使观测到的样本出现的概率最大

例5 设总体 $X \sim \begin{pmatrix} 1 & 2 & 3 \\ 1-\theta^2 & \dfrac{\theta^2}{2} & \dfrac{\theta^2}{2} \end{pmatrix}$，$\theta(0<\theta<1)$未知.

现得一样本值1, 3, 2, 3，求 θ 最大似然估计值.

解 $P\{(X_1, X_2, X_3, X_4) = (1,3,2,3)\}$

$$= P\{X_1 = 1,\ X_2 = 3,\ X_3 = 2,\ X_4 = 3\}$$

$$= P\{X_1 = 1\} P\{X_2 = 3\} P\{X_3 = 2\} P\{X_4 = 3\}$$

$$= P\{X = 1\} P\{X = 3\} P\{X = 2\} P\{X = 3\}$$

$$= \frac{1}{8}\theta^6\left(1-\theta^2\right) \triangleq L(\theta)$$

例5 设总体 $X \sim \begin{pmatrix} 1 & 2 & 3 \\ 1-\theta^2 & \dfrac{\theta^2}{2} & \dfrac{\theta^2}{2} \end{pmatrix}$，$\theta(0<\theta<1)$未知.

现得一样本值1, 3, 2, 3, 求 θ 最大似然估计值.

解 $$L(\theta)=\frac{1}{8}\theta^6(1-\theta^2),\quad 0<\theta<1$$

$$\ln L(\theta)=\ln\frac{1}{8}+6\ln\theta+\ln(1-\theta^2)$$

令 $$\frac{d}{d\theta}\ln L(\theta)=\frac{6}{\theta}-\frac{2\theta}{1-\theta^2}=0$$

解出 $\hat{\theta}=\dfrac{\sqrt{3}}{2}$ 最大似然估计值

矩估计值

$$\hat{\theta}=\sqrt{\frac{5}{6}}$$

2. 连续型

设总体X的概率密度：$f(x;\theta)$，$\theta\in\Theta$. θ未知

$X_1,X_2,\cdots,X_n$ 是总体X的一个样本

$x_1,x_2,\cdots,x_n$ 为样本值

记 $L(\theta)=\prod_{i=1}^{n} f(x_i;\theta)$ 称为样本的似然函数

令 $\boxed{\dfrac{d}{d\theta}\ln L(\theta)=0}$ $\theta\in\Theta$

解 $\hat{\theta}(x_1,x_2,\cdots,x_n)$ 最大似然估计**值**

$\hat{\theta}(X_1,X_2,\cdots,X_n)$ 最大似然估计**量**

例6 设总体X密度 $f(x)=\begin{cases}(\theta+1)x^{\theta}, & 0<x<1,\\ 0, & 其他\end{cases}$

$\theta>-1$未知. 求 θ 的最大似然估计.

解 设 $X_1, X_2, \cdots, X_n$ 是总体X的一个样本

$x_1, x_2, \cdots, x_n$ 为样本值

$$L(\theta)=\begin{cases}\prod_{i=1}^{n}(\theta+1)x_i^{\theta}=(\theta+1)^n(x_1x_2\cdots x_n)^{\theta}, & 0<x_i<1,\ i=1,2,\cdots,n.\\ 0, & 其他\end{cases}$$

$0<x_i<1,\ i=1,2,\cdots,n$ 时

$$\ln L(\theta)=n\ln(\theta+1)+\theta\sum_{i=1}^{n}\ln x_i$$

例6 设总体X密度 $f(x)=\begin{cases}(\theta+1)x^{\theta}, & 0<x<1,\\ 0, & 其他\end{cases}$

$\theta>-1$未知. 求 θ 的最大似然估计.

解 $\ln L(\theta)=n\ln(\theta+1)+\theta\sum_{i=1}^{n}\ln x_i$

令 $\frac{d}{d\theta}\ln L(\theta)=\boxed{\frac{n}{\theta+1}+\sum_{i=1}^{n}\ln x_i=0}$

$\longrightarrow \hat{\theta}=-\frac{n}{\sum_{i=1}^{n}\ln x_i}-1$ 最大似然估计值

$\hat{\theta}=-\frac{n}{\sum_{i=1}^{n}\ln X_i}-1$ 最大似然估计量

矩估计值

$$\boxed{\hat{\theta}=\frac{2\bar{X}-1}{1-\bar{X}}}$$

例7 设总体X密度 $f(x)=\begin{cases}(\theta+1)x^{\theta}, & 0<x<1,\\ 0, & \text{其他}\end{cases}$

$\theta>-1$ 未知. 试求 $P(X<1/2)$ 的极大似然估计值。

解 $\ln L(\theta)=n\ln(\theta+1)+\theta\sum_{i=1}^{n}\ln x_i$

令 $\dfrac{d}{d\theta}\ln L(\theta)=\dfrac{n}{\theta+1}+\sum_{i=1}^{n}\ln x_i=0$

$$\hat{\theta}=-\frac{n}{\sum_{i=1}^{n}\ln x_i}-1$$

最大似然估计值

$$P(X<\frac{1}{2})=\int_0^{1/2}(\theta+1)x^{\theta}dx=(\frac{1}{2})^{\theta+1}$$

故 $P(X<1/2)$ 的极大似然估计值为

$$(\frac{1}{2})^{\hat{\theta}+1}=(\frac{1}{2})^{-\frac{n}{\sum_{i=1}^{n}\ln x_i}}$$

二. 极大似然估计法

3. 总体X的分布含两个未知参数 θ_1,θ_2.

似然函数 $L(\theta_1,\theta_2)$

令 $$\begin{cases} \dfrac{\partial}{\partial\theta_1}\ln L(\theta_1,\theta_2)=0 \\ \dfrac{\partial}{\partial\theta_2}\ln L(\theta_1,\theta_2)=0 \end{cases} \longrightarrow \begin{cases} \hat{\theta}_1(x_1,x_2,\cdots,x_n) \\ \hat{\theta}_2(x_1,x_2,\cdots,x_n) \end{cases}$$

最大似然估计值

例8. 设某机床加工的轴的直径与图纸规定的中心尺寸的偏差服从 $N(\mu,\sigma^2)$，其中 μ,σ^2 未知.

从中随机抽取n 根轴，测得其偏差为 $x_1,x_2,\cdots,x_n$

试求：μ,σ^2的极大似然估计.

解：$\because$ X 的密度函数为：$f(x;\mu,\sigma^2)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

作似然函数：

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

$$=\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

$$L(\theta_1,\theta_2,\cdots,\theta_k)=\prod_{i=1}^{n}f(x_i;\theta_1,\theta_2,\cdots,\theta_k)$$

对 L 两边取对数得:

$$L=\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n\cdot e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

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

令:
$$\begin{cases}\dfrac{\partial\ln L}{\partial\mu}=\dfrac{1}{2\sigma^2}\sum_{i=1}^{n}2(x_i-\mu)=\boxed{\dfrac{1}{2\sigma^2}\sum_{i=1}^{n}2x_i-n\mu=0}\\[2ex] \dfrac{\partial\ln L}{\partial\sigma^2}=\boxed{-\dfrac{n}{2\sigma^2}+\dfrac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(x_i-\mu)^2=0}\end{cases}$$

与矩估计法所得的结论是一致的

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

解得:
$$\begin{cases}\hat{\mu}=\dfrac{1}{n}\sum_{i=1}^{n}x_i=\bar{x}\\[2ex] \hat{\sigma}^2=\dfrac{1}{n}\sum_{i=1}^{n}(x_i-\hat{\mu})^2=\dfrac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2\end{cases}$$

$\hat{\mu},\hat{\sigma}^2$是μ,σ^2的极大似然估计值(量)

例9 设总体 $X \sim U(a,b)$， a, b未知，$X_1, X_2,\ldots,X_n$是来自X的样本，$x_1, x_2,\ldots, x_n$是样本值。

求a, b的极大似然估计量 $\hat{a}, \hat{b}$

解：
$$f(x,a,b)=\begin{cases}1/(b-a), & a \le x \le b \\ 0, & \text{其它}\end{cases}$$

$$L(a,b)=\prod_{i=1}^{n} f(x_i,a,b)=\frac{1}{(b-a)^n},$$

当似然函数不可微或方程组无解时，则应根据定义直接寻求能使 $L(\theta_1,\theta_2,\cdots,\theta_k)$ 达到最大值的解 $\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_k$ 作为极大似然估计值(量)。

例9. $X \sim U(a,b)$, 求a, b的极大似然估计量 $\hat{a}, \hat{b}$

解: $$f(x,a,b)=\begin{cases} 1/(b-a), & a \le x \le b \\ 0, & \text{其它} \end{cases}$$

$$L(a,b)=\frac{1}{(b-a)^n},$$

用谁估计a?用谁估计b?

$$x_{(1)}=\min\{x_1,x_2,\cdots,x_n\} \qquad x_{(n)}=\max\{x_1,x_2,\cdots,x_n\}$$

a,b的极大似然估计值: $\hat{a}=x_{(1)}=\min\limits_{1\le i\le n} x_i$, $\hat{b}=x_{(n)}=\max\limits_{1\le i\le n} x_i$

a,b的极大似然估计量: $\hat{a}=\min\limits_{1\le i\le n} X_i$, $\hat{b}=\max\limits_{1\le i\le n} X_i$

小结

矩估计法的步骤：

(1) 计算总体矩：

离散型： $\mu_i = E(X^i) = \sum_{x_l \in R_X} x_l^i \cdot P(x_l, \theta_1, \theta_2, \cdots \theta_k) \quad i = 1, 2, \cdots k$

连续型： $\mu_i = E(X^i) = \int_{-\infty}^{+\infty} x^i \cdot f(x; \theta_1, \theta_2, \cdots, \theta_k)\, dx \quad i = 1, 2, \cdots, k$

(2) 用样本矩估计总体矩： $\mu_i = E(X^i) = A_i$

$i = 1, 2, \cdots, k$

(3) 解方程组，得矩估计量：

$\hat{\theta}_i = \theta_i(A_1, A_2, \cdots, A_k)$是$\theta_i$的矩估计量，$i = 1, 2, \cdots, k$

极大似然估计法的步骤：

1) 计算似然函数： $L(\theta_1,\theta_2,\cdots,\theta_k)=\prod_{i=1}^{n} p(x_i;\theta_1,\theta_2,\cdots,\theta_k)$

$$L(\theta_1,\theta_2,\cdots,\theta_k)=\prod_{i=1}^{n} f(x_i;\theta_1,\theta_2,\cdots,\theta_k)$$

2) 计算似然函数的最大值点：

取$(\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_k)\in\Theta$，使$L(\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_k)=\max\limits_{\hat{\theta}\in\Theta} L(\theta_1,\theta_2,\cdots,\theta_k)$

求$\max\limits_{\hat{\theta}\in\Theta} L \xrightarrow{\text{可微}}$ 求$L(\theta_1,\theta_2,\cdots,\theta_k)$的驻点：

$$\frac{\partial L(\theta_1,\theta_2,\cdots,\theta_k)}{\partial\theta_i}=0,\ i=1,2,\cdots,k$$

$\xrightarrow{\ln x\text{单调}}$ 求$\ln L(\theta_1,\theta_2,\cdots,\theta_k)$的驻点：

$$\frac{\partial \ln L(\theta_1,\theta_2,\cdots,\theta_k)}{\partial\theta_i}=0,\ i=1,2,\cdots,k$$

- 求解方程组得：$\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_k$为极大似然估计值(量).

1) 计算似然函数：$L(\theta_1,\theta_2,\cdots,\theta_k)=\prod_{i=1}^{n} p(x_i;\theta_1,\theta_2,\cdots,\theta_k)$

$$L(\theta_1,\theta_2,\cdots,\theta_k)=\prod_{i=1}^{n} f(x_i;\theta_1,\theta_2,\cdots,\theta_k)$$

2) 计算似然函数的最大值点：

当似然函数不可微或方程组无解时，则应根据定义直接寻求能使 $L(\theta_1,\theta_2,\cdots,\theta_k)$ 达到最大值的解 $\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_k$ 作为极大似然估计值(量)。