

第七章 参数估计

在实际问题中,我们已知所研究的总体分布类型,但是分布中包涵 1 个或多个未知参数.那么,如何根据样本来估计未知参数,这就是所谓的参数估计.参数估计是数理统计研究的主要问题之一.

例如,假设总体 $X \sim N(\mu, \sigma^2)$, μ, σ^2 是未知参数, X_1, X_2, \dots, X_n 是来自 X 的样本,样本值是 x_1, x_2, \dots, x_n ,我们要由样本值来确定 μ 和 σ^2 的估计值,这就是参数估计问题,参数估计分为点估计 (Point estimation) 和区间估计 (Interval estimation).

第一节 点估计

所谓点估计是指把总体的未知参数估计为某个确定的值或在某个确定的点上,故点估计又称为定值估计.

定义 7.1 设总体 X 的分布函数为 $F(x, \theta)$, θ 是未知参数, X_1, X_2, \dots, X_n 是 X 的一样本,样本值为 x_1, x_2, \dots, x_n ,构造一个统计量 $\theta(X_1, X_2, \dots, X_n)$,用它的观察值 $\hat{\theta}(x_1, x_2, \dots, x_n)$ 作为 θ 的近似值,这种问题称为点估计问题.习惯上称随机变量 $\hat{\theta}(X_1, X_2, \dots, X_n)$ 为 θ 的估计量,称 $\hat{\theta}(x_1, x_2, \dots, x_n)$ 为 θ 的估计值.

构造估计量 $\hat{\theta}(X_1, X_2, \dots, X_n)$ 的方法很多,下面仅介绍矩法和最(极)大似然估计法.

1. 矩法

矩法 (Moment method of estimation) 是一种古老的估计方法.它是由英国统计学家皮尔逊 (K. Pearson) 于 1894 年首创的.它虽然古老,但目前仍常用.

矩法估计是用样本矩作为总体矩的估计的方法,用矩估计确定的估计量称为矩估计量,相应的估计值称为矩估计值.矩估计量与矩估计值统称为矩估计.

矩估计的一般作法: 设总体 $X \sim F(X; \theta_1, \theta_2, \dots, \theta_l)$ 其中 $\theta_1, \theta_2, \dots, \theta_l$ 均未知.

(1) 如果总体 X 的 k 阶矩 $\mu_k = E(X^k)$, $1 \leq k \leq l$ 均存在,则

$$\mu_k = \mu_k(\theta_1, \theta_2, \dots, \theta_l), (1 \leq k \leq l).$$

(2) 令

$$\begin{cases} \mu_1(\theta_1, \theta_2, \dots, \theta_l) A_1, \\ \mu_2(\theta_1, \theta_2, \dots, \theta_l) A_2, \\ \dots\dots\dots \\ \mu_l(\theta_1, \theta_2, \dots, \theta_l) A_l. \end{cases}$$

其中 A_k , $1 \leq k \leq l$ 为样本 k 阶矩.

求出方程组的解 $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_l$, 我们称 $\hat{\theta}_k = \hat{\theta}_k(X_1, X_2, \dots, X_n)$ 为参数 θ_k , $1 \leq k \leq l$ 的矩估计量, $\hat{\theta}_k = \hat{\theta}_k(x_1, x_2, \dots, x_n)$ 为参数 θ_k 的矩估计值.

例 7.1 设总体 X 的密度函数为:

$$f(x) = \begin{cases} (\alpha+1)x^\alpha, & 0 < x < 1, \\ 0, & \text{其他.} \end{cases}$$

其中 $\alpha > -1$, 样本为 (X_1, X_2, \dots, X_n) , 求参数 α 的矩估计量.

解 $A_1 = \bar{X}$. 由 $\mu_1 = A_1$ 及

$$\mu_1 = E(X) = \int_{-\infty}^{+\infty} xf(x)dx = \int_0^1 x(\alpha+1)x^\alpha dx = \frac{\alpha+1}{\alpha+2},$$

$$\text{有 } \bar{X} = \frac{\alpha+1}{\alpha+2}, \text{ 得 } \hat{\alpha} = \frac{1-2\bar{X}}{\bar{X}-1}.$$

例 7.2 设 $X \sim N(\mu, \sigma^2)$, 其中 μ, σ^2 未知, 试求参数 μ, σ^2 的矩估计量.

解

$$\mu_1 = E(X) = A_1 = \frac{1}{n} \sum_{i=1}^n X_i,$$

$$\mu_2 = E(X^2) = A_2 = \frac{1}{n} \sum_{i=1}^n X_i^2.$$

$$\text{又 } E(X) = \mu, \quad E(X^2) = D(X) + (E(X))^2 = \sigma^2 + \mu^2,$$

$$\text{那么 } \hat{\mu} = \bar{X}, \hat{\sigma}^2 = A_2 - (\hat{\mu})^2 = \frac{n-1}{n} S^2.$$

例 7.3 在某班期末数学考试成绩中随机抽取 9 人的成绩. 结果如表 7-1 所示, 试求该班数学成绩的平均分数、标准差的矩估计值.

表 7-1

序号	1	2	3	4	5	6	7	8	9
分数	94	89	85	78	75	71	65	63	55

解 设 X 为该班数学成绩, $\mu = E(X), \sigma^2 = D(X)$, 则

$$\bar{x} = \frac{1}{9} \sum_{i=1}^9 x_i = \frac{1}{9} (94 + 89 + \dots + 55) = 75;$$

$$\sqrt{\frac{8}{9} s^2} = \left[\frac{8}{9} \cdot \frac{1}{8} \sum_{i=1}^9 (x_i - \bar{x})^2 \right]^{1/2} \approx 12.14.$$

$$\begin{cases} \mu_1 = E(X) = A_1 = \frac{1}{9} \sum_{i=1}^9 X_i, \\ \mu_2 = E(X^2) = A_2 = \frac{1}{9} \sum_{i=1}^9 X_i^2. \end{cases}$$

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

$$\hat{\mu} = \bar{X}, \hat{\sigma}^2 = A_2 - (\hat{\mu})^2 = A_2 - (\bar{x})^2 = \frac{8}{9} S^2.$$

所以, 该班数学成绩的平均分数的矩估计值 $\hat{\mu} = \bar{x} = 75$ 分, 标准差的矩估计值

$$\hat{\sigma} = \sqrt{\frac{8}{9}s^2} = 12.14.$$

作矩法估计时无需知道总体的概率分布，只要知道总体矩即可.但矩法估计量有时不惟一，如总体 X 服从参数为 λ 的泊松分布时， \bar{X} 和 B_2 都是参数 λ 的矩法估计.

2. 极（最）大似然估计法

最大似然估计法(Maximum likelihood estimation)只能在已知总体分布的前提下进行，为了对它的思想有所了解，我们先看一个例子.

例 7.4 假定一个盒子里装有许多大小相同的黑球和白球，并且假定它们的数目之比为 3:1，但不知是白球多还是黑球多，现在有放回地从盒中抽了 3 个球，试根据所抽 3 个球中黑球的数目确定是白球多还是黑球多.

解 设所抽 3 个球中黑球数为 X ，摸到黑球的概率为 p ，则 X 服从二项分布

$$P\{X=k\} = C_3^k p^k (1-p)^{3-k}, \quad k=0, 1, 2, 3.$$

问题是 $p=1/4$ 还是 $p=3/4$? 现根据样本中黑球数，对未知参数 p 进行估计.抽样后，共有 4 种可能结果，其概率如表 7-2 所示.

表 7-2

X	0	1	2	3
$p=1/4$ 时, $P\{X=k\}$	27/64	27/64	9/64	1/64
$p=3/4$ 时, $P\{X=k\}$	1/64	9/64	27/64	27/64

假如某次抽样中，只出现一个黑球，即 $X=1$ ， $p=1/4$ 时， $P\{X=1\}=27/64$ ； $p=3/4$ 时， $P\{X=1\}=9/64$ ，这时我们会选择 $p=1/4$ ，即黑球数比白球数为 1:3.因为在一次试验中，事件“1 个黑球”发生了.我们认为它应有较大的概率 27/64 ($27/64 > 9/64$)，而 27/64 对应着参数 $p=1/4$ ，同样可以考虑 $X=0, 2, 3$ 的情形，最后可得

$$p = \begin{cases} \frac{1}{4}, & \text{当 } x=0,1 \text{ 时,} \\ \frac{3}{4}, & \text{当 } x=2,3 \text{ 时.} \end{cases}$$

(1) 似然函数.在最大似然估计法中，最关键的问题是如何求得似然函数（定义下文给出），有了似然函数，问题就简单了，下面分两种情形来介绍似然函数.

(a) 离散型总体.设总体 X 为离散型， $P\{X=x\}=p(x, \theta)$ ，其中 θ 为待估计的未知参数，假定 x_1, x_2, \dots, x_n 为样本 X_1, X_2, \dots, X_n 的一组观测值，则

$$\begin{aligned} P\{X_1=x_1, X_2=x_2, \dots, X_n=x_n\} &= P\{X_1=x_1\}P\{X_2=x_2\}\dots P\{X_n=x_n\} \\ &= p(x_1, \theta) p(x_2, \theta) \dots p(x_n, \theta) \\ &= \prod_{i=1}^n p(x_i, \theta). \end{aligned}$$

将 $\prod_{i=1}^n p(x_i, \theta)$ 看作是参数 θ 的函数，记为 $L(\theta)$ ，即

$$L(\theta) = \prod_{i=1}^n p(x_i, \theta). \quad (7-1)$$

(b) 连续型总体. 设总体 X 为连续型, 已知其分布密度函数为 $f(x, \theta)$, 其中 θ 为待估计的未知参数, 则样本 (X_1, X_2, \dots, X_n) 的联合密度为:

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

将它也看作是参数 θ 的函数, 记为 $L(\theta)$, 即

$$L(\theta) = \prod_{i=1}^n f(x_i, \theta). \quad (7-2)$$

由此可见: 不管是离散型总体, 还是连续型总体, 只要知道它的概率分布或密度函数, 我们总可以得到一个关于参数 θ 的函数 $L(\theta)$, 称 $L(\theta)$ 为似然函数.

(2) 最大似然估计

最大似然估计法的主要思想是: 如果随机抽样得到的样本观测值为 x_1, x_2, \dots, x_n , 则我们应当这样来选取未知参数 θ 的值, 使得出现该样本值的可能性最大, 即使得似然函数 $L(\theta)$ 取最大值, 从而求参数 θ 的最大似然估计 $\hat{\theta}$ 的问题, 就转化为求似然函数 $L(\theta)$ 的极值点的问题, 一般来说, 这个问题可以通过求解下面的方程来解决

$$\frac{dL(\theta)}{d\theta} = 0. \quad (7-3)$$

然而, $L(\theta)$ 是 n 个函数的连乘积, 求导数比较复杂, 由于 $\ln L(\theta)$ 是 $L(\theta)$ 的单调增函数, 所以 $L(\theta)$ 与 $\ln L(\theta)$ 在 θ 的同一点处取得极大值. 于是求解 (7-3) 可转化为求解

$$\frac{d \ln L(\theta)}{d\theta} = 0. \quad (7-4)$$

称 $\ln L(\theta)$ 为对数似然函数, 方程 (7.4) 为对数似然方程, 求解此方程就可得到参数 θ 的估计值.

如果总体 X 的分布中含有 k 个未知参数: $\theta_1, \theta_2, \dots, \theta_k$, 则最大似然估计法也适用. 此时, 所得的似然函数是关于 $\theta_1, \theta_2, \dots, \theta_k$ 的多元函数 $L(\theta_1, \theta_2, \dots, \theta_k)$, 解下列方程组, 就可得到 $\theta_1, \theta_2, \dots, \theta_k$ 的估计值,

$$\begin{cases} \frac{\partial \ln L(\theta_1, \theta_2, \dots, \theta_k)}{\partial \theta_1} = 0, \\ \frac{\partial \ln L(\theta_1, \theta_2, \dots, \theta_k)}{\partial \theta_2} = 0, \\ \dots\dots\dots \\ \frac{\partial \ln L(\theta_1, \theta_2, \dots, \theta_k)}{\partial \theta_k} = 0. \end{cases} \quad (7-5)$$

例 7.5 在泊松总体中抽取样本, 其样本值为: x_1, x_2, \dots, x_n , 试对泊松分布的未知参数 λ 作最大似然估计.

解 因泊松总体是离散型的, 其概率分布为:

$$P\{X=x\} = \frac{\lambda^x}{x!} e^{-\lambda},$$

故似然函数为:

$$L(\lambda) = \prod_{i=1}^n \frac{\lambda^{x_i}}{x_i!} e^{-\lambda} = e^{-\lambda n} \cdot \lambda^{\sum_{i=1}^n x_i} \cdot \prod_{i=1}^n \frac{1}{x_i!}.$$

$$\ln L(\lambda) = -n\lambda + \sum_{i=1}^n x_i \ln \lambda - \ln \prod_{i=1}^n (x_i!),$$

$$\frac{d \ln(\lambda)}{d \lambda} = -n + \frac{1}{\lambda} \sum_{i=1}^n x_i.$$

令 $\frac{d \ln(\lambda)}{d \lambda} = 0$, 得: $-n + \frac{1}{\lambda} \sum_{i=1}^n x_i = 0.$

所以 $\hat{\lambda}_L = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}$, 即 λ 的最大似然估计量为 $\hat{\lambda}_L = \bar{X}$ (为了和 λ 的矩法估计区别

起见, 我们将 λ 的最大似然估计记为 $\hat{\lambda}_L$).

例 7.6 设一批产品含有次品, 今从中随机抽出 100 件, 发现其中有 8 件次品, 试求次品率 θ 的最大似然估计值.

解 用最大似然法时必须明确总体的分布, 现在题目没有说明这一点, 故应先来确定总体的分布. 设

$$X_i = \begin{cases} 1, & \text{第 } i \text{ 次取次品,} \\ 0, & \text{第 } i \text{ 次取正品,} \end{cases} \quad i = 1, 2, \dots, 100,$$

则 X_i 服从两点分布, 如表 7-3 所示.

表 7-3

X_i	1	0
P	θ	$1-\theta$

设 x_1, x_2, \dots, x_{100} 为样本观测值, 则:

$$p(x_i; \theta) = P\{X_i = x_i\} = \theta^{x_i} (1-\theta)^{1-x_i},$$

$$x_i = 0, 1,$$

故似然函数为:

$$L(\theta) = \prod_{i=1}^{100} \theta^{x_i} (1-\theta)^{1-x_i} = \theta^{\sum_{i=1}^{100} x_i} (1-\theta)^{100 - \sum_{i=1}^{100} x_i}$$

由题设知 $\sum_{i=1}^{100} x_i = 8$, 所以

$$L(\theta) = \theta^8 (1-\theta)^{92}.$$

两边取对数, 得:

$$\ln L(\theta) = 8 \ln \theta + 92 \ln (1-\theta).$$

对数似然方程为:

$$\frac{d \ln L(\theta)}{d \theta} = \frac{8}{\theta} - \frac{92}{1-\theta} = 0.$$

解之得 $\hat{\theta}_L = 0.08$.

例 7.7 设 x_1, x_2, \dots, x_n 为来自正态总体 $N(\mu, \sigma^2)$ 的观测值, 试求总体未知参数 μ, σ^2 的最大似然估计.

解 因正态总体为连续型, 其密度函数为

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

所以似然函数为:

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

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

故对数似然方程组为:

$$\begin{cases} \frac{\partial \ln L(\mu, \sigma^2)}{\partial \mu} = \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu) = 0, \\ \frac{\partial \ln L(\mu, \sigma^2)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^n (x_i - \mu)^2 = 0. \end{cases}$$

解以上方程组得:

$$\begin{cases} \hat{\mu}_L = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}, \\ \hat{\sigma}_L^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = B_2. \end{cases}$$

所以
$$\begin{cases} \hat{\mu}_L = \bar{X}, \\ \hat{\sigma}_L^2 = B_2. \end{cases}$$

例 7.8 设总体 X 服从 $[0, \theta]$ 上的均匀分布, x_1, x_2, \dots, x_n 是来自 X 的样本, 求 θ 的矩估计和最大似然估计.

解 因为 $E(X) = \theta/2$, 令 $\bar{X} = E(X)$, 得

$$\hat{\theta}_{\text{矩}} = 2\bar{X}.$$

又 X 的概率密度为

$$f(x) = \begin{cases} \frac{1}{\theta}, & 0 \leq x \leq \theta, \\ 0, & \text{其他.} \end{cases}$$

所以

$$L(\theta) = \frac{1}{\theta^n}, \quad 0 \leq x_i \leq \theta.$$

要 $L(\theta)$ 最大, θ 必须尽可能小, 又 $\theta \geq x_i, i=1, 2, \dots, n$, 所以 $\hat{\theta}_L = \max_{1 \leq i \leq n} \{X_i\}$.

第二节 估计量的评价标准

设总体 X 服从 $[0, \theta]$ 上的均匀分布, 由上节例 7.8 可知 $\hat{\theta}_{\text{矩}} = 2\bar{X}$, $\hat{\theta}_L = \max_{1 \leq i \leq n} \{X_i\}$ 都是 θ 的估计, 这两个估计哪个好? 下面我们首先讨论衡量估计量好坏的标准问题.

1. 无偏性

定义 7.2 若估计量 $\hat{\theta} = \hat{\theta}(X_1, X_2, \dots, X_n)$ 的数学期望等于未知参数 θ , 即

$$E(\hat{\theta}) = \theta, \quad (7-6)$$

则称 $\hat{\theta}$ 为 θ 的无偏估计量 (Non-deviation estimator).

估计量 $\hat{\theta}$ 的值不一定就是 θ 的真值, 因为它是一个随机变量, 若 $\hat{\theta}$ 是 θ 的无偏估计, 则尽管 $\hat{\theta}$ 的值随样本值的不同而变化, 但平均来说它会等于 θ 的真值.

例 7.9 设 X_1, X_2, \dots, X_n 为总体 X 的一个样本, $E(X) = \mu$, 则样本平均数 $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ 是 μ 的无偏估计量.

证 因为 $E(X) = \mu$, 所以 $E(X_i) = \mu, i=1, 2, \dots, n$, 于是

$$E(\bar{X}) = E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n E(X_i) = \mu.$$

所以 \bar{X} 是 μ 的无偏估计量.

例 7.10 设有总体 $X, E(X) = \mu, D(X) = \sigma^2, X_1, X_2, \dots, X_n$ 为来自总体 X 的样本, 样本方差 S^2 及二阶样本中心矩 $B_2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$ 是否为总体方差 σ^2 的无偏估计?

解 因为 $E(S^2) = \sigma^2$, 所以 S^2 是 σ^2 的一个无偏估计, 这也是我们称 S^2 为样本方差的理由. 由于

$$B_2 = \frac{n-1}{n} S^2,$$

那么

$$E(B_2) = \frac{n-1}{n} E(S^2) = \frac{n-1}{n} \sigma^2,$$

所以 B_2 不是 σ^2 的一个无偏估计.

还需指出: 一般说来无偏估计量的函数并不是未知参数相应函数的无偏估计量. 例如,

当 $X \sim N(\mu, \sigma^2)$ 时, \bar{X} 是 μ 的无偏估计量, 但 \bar{X}^2 不是 μ^2 的无偏估计量, 事实上:

$$E(\bar{X})^2 = D(\bar{X}) + [E(\bar{X})]^2 = \frac{\sigma^2}{n} + \mu^2 \neq \mu^2.$$

2. 有效性

对于未知参数 θ , 如果有两个无偏估计量 $\hat{\theta}_1$ 与 $\hat{\theta}_2$, 即 $E(\hat{\theta}_1) = E(\hat{\theta}_2) = \theta$, 那么在 $\hat{\theta}_1$, $\hat{\theta}_2$ 中哪一个更好呢? 此时我们自然希望 θ 对 θ 的平均偏差 $E(\hat{\theta} - \theta)^2$ 越小越好, 即一个好的估计量应该有尽可能小的方差, 这就是有效性.

定义 7.3 设 $\hat{\theta}_1$ 和 $\hat{\theta}_2$ 都是未知参数 θ 的无偏估计, 若对任意的参数 θ , 有

$$D(\hat{\theta}_1) \leq D(\hat{\theta}_2), \quad (7-7)$$

则称 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 有效.

尽管 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 有效, 但 $\hat{\theta}_1$ 仍然还不是 θ 的真值, 只是 $\hat{\theta}_1$ 在 θ 附近取值的密集程度较 $\hat{\theta}_2$ 高, 即用 $\hat{\theta}_1$ 估计 θ 精度要高些.

例如, 对正态总体 $N(\mu, \sigma^2)$, $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$, X_i 和 \bar{X} 都是 $E(X) = \mu$ 的无偏估计量, 但

$$D(\bar{X}) = \frac{\sigma^2}{n} \leq D(X_i) = \sigma^2,$$

故 \bar{X} 较个别观测值 X_i 有效. 实际当中也是如此, 比如要估计某个班学生的平均成绩, 可用两种方法进行估计, 一种是在该班任意抽一个同学, 就以该同学的成绩作为全班的平均成绩; 另一种方法是在该班抽取 n 位同学, 以这 n 个同学的平均成绩作为全班的平均成绩, 显然第二种方法比第一种方法好.

3. 一致性

无偏性、有效性都是在样本容量 n 一定的条件下进行讨论的, 然而 $\theta(X_1, X_2, \dots, X_n)$ 不仅与样本值有关, 而且与样本容量 n 有关, 不妨记为 $\hat{\theta}_n$, 很自然地, 我们希望 n 越大时, $\hat{\theta}_n$ 对 θ 的估计应该越精确.

定义 7.4 如果 $\hat{\theta}_n$ 依概率收敛于 θ , 即 $\forall \varepsilon > 0$, 有

$$\lim_{n \rightarrow \infty} P\left\{|\hat{\theta}_n - \theta| < \varepsilon\right\} = 1, \quad (7-8)$$

则称 $\hat{\theta}_n$ 是 θ 的一致估计量 (Uniform estimator).

由辛钦大数定律可以证明：样本平均数 \bar{X} 是总体均值 μ 的一致估计量，样本方差 S^2 及二阶样本中心矩 B_2 都是总体方差 σ^2 的一致估计量。

第三节 区间估计

1. 区间估计的概念

上节我们介绍了参数的点估计，假设总体 $X \sim N(\mu, \sigma^2)$ ，对于样本 (X_1, X_2, \dots, X_n) ， $\hat{\mu} = \bar{X}$ 是参数 μ 的矩法估计和最大似然估计，并且满足无偏性和一致性。但实际上 $\bar{X} = \mu$ 的可能性有多大呢？由于 \bar{X} 是一连续型随机变量， $P\{X = \mu\} = 0$ ，即 $\hat{\mu} = \mu$ 的可能性为 0，为此，我们希望给出 μ 的一个大致范围，使得 μ 有较高的概率在这个范围内，这就是区间估计问题。

定义 7.5 设 $\hat{\theta}_1(X_1, X_2, \dots, X_n)$ 及 $\hat{\theta}_2(X_1, X_2, \dots, X_n)$ 是两个统计量，如果对于给定的概率 $1 - \alpha$ ， $0 < \alpha < 1$ ，有

$$P\{\hat{\theta}_1 < \theta < \hat{\theta}_2\} = 1 - \alpha, \quad (7-9)$$

则称随机区间 $(\hat{\theta}_1, \hat{\theta}_2)$ 为参数 θ 的**置信区间** (Confidence interval)， $\hat{\theta}_1$ 称为**置信下限**， $\hat{\theta}_2$ 称为**置信上限**， $1 - \alpha$ 叫**置信概率**或**置信度** (Confidence level)。

定义中的随机区间 $(\hat{\theta}_1, \hat{\theta}_2)$ 的大小依赖于随机抽取的样本观测值，它可能包含 θ ，也可能不包含 θ ，(7.9) 式的意义是指随机区间 $(\hat{\theta}_1, \hat{\theta}_2)$ 以 $1 - \alpha$ 的概率包含 θ 。例如，若取 $\alpha = 0.05$ ，那么置信概率为 $1 - \alpha = 0.95$ ，这时，置信区间 $(\hat{\theta}_1, \hat{\theta}_2)$ 的意义是指：在 100 次重复抽样中所得到的 100 个置信区间中，大约有 95 个区间包含参数真值 θ ，有 5 个区间不包含真值 θ ，亦即随机区间 $(\hat{\theta}_1, \hat{\theta}_2)$ 包含参数 θ 真值的频率近似为 0.95。

例 7.11 设 $X \sim N(\mu, \sigma^2)$ ， μ 未知， σ^2 已知，样本 X_1, X_2, \dots, X_n 来自总体 X ，求 μ 的置信区间，置信概率为 $1 - \alpha$ 。

解 因为 X_1, X_2, \dots, X_n 为来自 X 的样本，而 $X \sim N(\mu, \sigma^2)$ ，所以

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

对于给定的 α ，查附表 2 可得上分位点 $z_{\frac{\alpha}{2}}$ ，使得

$$P\left\{\left|\frac{\bar{X} - \mu}{\sigma / \sqrt{n}}\right| < z_{\frac{\alpha}{2}}\right\} = 1 - \alpha,$$

即

$$P\left\{\bar{X}-z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}<\mu<\bar{X}+z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right\}=1-\alpha.$$

所以 μ 的置信概率为 $1-\alpha$ 的置信区间为

$$\left(\bar{X}-z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \bar{X}+z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right). \quad (7-10)$$

由 (7.10) 式可知,置信区间的长度为 $2z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$, 若 n 越大, 置信区间就越短; 若置信概率 $1-\alpha$ 越大, α 就越小, $z_{\frac{\alpha}{2}}$ 就越大, 从而置信区间就越长.

2. 正态总体参数的区间估计

由于在大多数情况下, 我们所遇到的总体是服从正态分布的 (有的是近似正态分布), 故我们现在来重点讨论正态总体参数的区间估计问题.

在下面的讨论中, 总假定 $X \sim N(\mu, \sigma^2)$, X_1, X_2, \dots, X_n 为来自总体 X 的样本.

(1) 对 μ 的估计. 分两种情况进行讨论.

(a) σ^2 已知. 此时就是例 7.11 的情形, 结论是: μ 的置信区间为

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

置信概率为 $1-\alpha$.

(b) σ^2 未知

当 σ^2 未知时, 不能使用 (7-10) 式作为置信区间, 因为 (7-10) 式中区间的端点与 σ

有关, 考虑到 $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ 是 σ^2 的无偏估计, 将 $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ 中的 σ 换成 S 得

$$T = \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1).$$

对于给定的 α , 查附录 4 可得上分位点 $t_{\alpha/2}(n-1)$, 使得

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

即

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

所以 μ 的置信概率为 $1-\alpha$ 的置信区间为

$$\left(\bar{X}-\frac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1), \bar{X}+\frac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1)\right). \quad (7-11)$$

又因为 $\frac{S}{\sqrt{n}} = \frac{S_0}{\sqrt{n-1}}$, $S_0 = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2}$, 所以 μ 的置信区间也可写成

$$\left(\bar{X} - \frac{S_0}{\sqrt{n-1}} t_{\frac{\alpha}{2}}(n-1), \bar{X} + \frac{S_0}{\sqrt{n-1}} t_{\frac{\alpha}{2}}(n-1) \right). \quad (7-12)$$

例 7.12 某车间生产滚珠, 已知其直径 $X \sim N(\mu, \sigma^2)$, 现从某一天生产的产品中随机地抽出 6 个, 测得直径如下 (单位: mm)

14.6 15.1 14.9 14.8 15.2 15.1

试求滚珠直径 X 的均值 μ 的置信概率为 95% 的置信区间.

解 $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{6} (14.6 + 15.1 + 14.9 + 14.8 + 15.2 + 15.1) = 14.95,$

$$s_0 = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} = 0.2062,$$

$$t_{\alpha/2}(n-1) = t_{0.025}(5) = 2.571,$$

所以

$$t_{\frac{\alpha}{2}}(n-1) \frac{S_0}{\sqrt{n-1}} = 2.571 \times \frac{0.2062}{\sqrt{6-1}} = 0.24,$$

置信区间为 $(14.95 - 0.24, 14.95 + 0.24)$, 即 $(14.71, 15.19)$, 置信概率为 95%.

(2) σ^2 的置信区间, 我们只考虑 μ 未知的情形.

此时由于 $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ 是 σ^2 的无偏估计, 我们考虑 $\frac{(n-1)S^2}{\sigma^2}$, 由于

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1),$$

所以, 对于给定的 α , 有

$$P \left\{ \chi_{1-\frac{\alpha}{2}}^2(n-1) < \frac{(n-1)S^2}{\sigma^2} < \chi_{\frac{\alpha}{2}}^2(n-1) \right\} = 1-\alpha.$$

即

$$P \left\{ \frac{(n-1)S^2}{\chi_{\frac{\alpha}{2}}^2(n-1)} < \sigma^2 < \frac{(n-1)S^2}{\chi_{1-\frac{\alpha}{2}}^2(n-1)} \right\} = 1-\alpha.$$

所以, σ^2 的置信概率为 $1-\alpha$ 的置信区间为

$$\left(\frac{(n-1)S^2}{\chi_{\frac{\alpha}{2}}^2(n-1)}, \frac{(n-1)S^2}{\chi_{1-\frac{\alpha}{2}}^2(n-1)} \right) \quad (7-13)$$

或

$$\left(\frac{nS_0^2}{\chi_{\frac{\alpha}{2}}^2(n-1)}, \frac{nS_0^2}{\chi_{1-\frac{\alpha}{2}}^2(n-1)} \right),$$

其中

$$S_0^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2.$$

例 7.13 某种钢丝的折断力服从正态分布, 今从一批钢丝中任取 10 根, 试验其折断力, 得数据如下:

572	570	578	568	596
576	584	572	580	566

试求方差 σ^2 的置信概率为 0.9 的置信区间.

解 因为 $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{10} (572 + 570 + \cdots + 566) = 576.2,$

$$s_0^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2 = 71.56,$$

$\alpha = 0.10, n-1=9$, 查附表 5 得:

$$\chi_{\frac{\alpha}{2}}^2(n-1) = \chi_{0.05}^2(9) = 16.919,$$

$$\chi_{1-\frac{\alpha}{2}}^2(n-1) = \chi_{0.95}^2(9) = 3.325,$$

$$\frac{ns_0^2}{\chi_{\frac{\alpha}{2}}^2(n-1)} = \frac{10 \times 71.56}{16.919} = 42.30,$$

$$\frac{ns_0^2}{\chi_{1-\frac{\alpha}{2}}^2(n-1)} = \frac{10 \times 71.56}{3.325} = 215.22.$$

所以, σ^2 的置信概率为 0.9 的置信区间为 (42.30, 215.22).

以上仅介绍了正态总体的均值和方差两个参数的区间估计方法.

在有些问题中并不知道总体 X 服从什么分布, 要对 $E(X) = \mu$ 作区间估计, 在这种情

况下只要 X 的方差 σ^2 已知, 并且样本容量 n 很大, 由中心极限定理, $\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$ 近似地服从标

准正态分布 $N(0, 1)$, 因而 μ 的置信概率为 $1-\alpha$ 的近似置信区间为

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

小 结

参数估计问题分为点估计和区间估计.

设 θ 是总体 X 的待估计参数. 用统计量 $\hat{\theta} = \hat{\theta}(X_1, X_2, \cdots, X_n)$ 来估计 θ 称 $\hat{\theta}$ 是 θ 的估计量, 点估计只给出未知参数 θ 的单一估计.

矩估计法的做法：设总体 $X \sim F(x; \theta_1, \theta_2, \dots, \theta_l)$ 其中 $\theta_k, 1 \leq k \leq l$ 为未知参数.

- [illegible]

$\hat{\theta}_k(x_1, x_2, \dots, x_n)$ 为 θ_k 的矩估计值.

(1) 写出似然函数 $L=L(\theta_1, \theta_2, \dots, \theta_l)$

$$L = \prod_{i=1}^n P(x_i; \theta_1, \theta_2, \dots, \theta_l),$$
$$L = \prod_{i=1}^n f(x_i; \theta_1, \theta_2, \dots, \theta_l),$$

- $$\ln L = \sum_{i=1}^n \ln f(x_i; \theta_1, \theta_2, \dots, \theta_l),$$

$$\ln L = \sum_{i=1}^n \ln P(x_i; \theta_1, \theta_2, \dots, \theta_l)$$

- $$\frac{\partial \ln L}{\partial \theta_k} = 0, \quad k=1, 2, \dots, l.$$

$\hat{\theta}_k=(X_1,X_2,\cdots,X_n)$ 为 θ_k 的最大似然估计量.

在统计问题中往往先使用最大似然估计法,在此法使用不方便时,再用矩估计法进行未知参数的点估计.

对于一个未知参数可以提出不同的估计量,那么就需要给出评定估计量好坏的标准.本章介绍了三个标准:无偏性、有效性、一致性.重点是无偏性.

点估计不能反映估计的精度,我们就引入区间估计.

设 θ 是总体 X 的未知参数, $\hat{\theta}_1, \hat{\theta}_2$ 均是样本 X_1, X_2, \dots, X_n 的统计量, 若对给定值 α , $0 < \alpha < 1$ 满足 $P(\hat{\theta}_1 < \theta < \hat{\theta}_2) = 1 - \alpha$, 称 $1 - \alpha$ 为置信度或置信概率, $(\hat{\theta}_1, \hat{\theta}_2)$ 为 θ 的置信度为 $1 - \alpha$ 的置信区间.

参数的区间估计中一个典型、重要的问题是正态总体 $X (X \sim N(\mu, \sigma^2))$ 中 μ 或 σ^2 的区间估计, 其置信区间如表 7-4 所示.

表 7-4 正态总体的均值、方差的置信度为 $(1 - \alpha)$ 的置信区间

待估参数	其他参数	统计量	置信区间
μ	σ^2 已知	$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim N(0, 1)$	$\left(\bar{X} \pm \frac{\sigma}{\sqrt{n}} Z_{\frac{\alpha}{2}} \right)$
μ	σ^2 未知	$T = \frac{\bar{X} - \mu}{S / \sqrt{n}} \sim t(n-1)$	$\left(\bar{X} \pm \frac{S}{\sqrt{n}} t_{\frac{\alpha}{2}}(n-1) \right)$
σ^2	μ 未知	$\chi^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$	$\left(\frac{(n-1)S^2}{\chi_{\frac{\alpha}{2}}^2(n-1)}, \frac{(n-1)S^2}{\chi_{1-\frac{\alpha}{2}}^2(n-1)} \right)$

区间估计给出了估计的精度与可靠度 $(1 - \alpha)$, 其精度与可靠度是相互制约的, 即精度越高 (置信区间长度越小), 可靠度越低; 反之亦然. 在实际中, 应先固定可靠度, 再估计精度.

重要术语及主题

矩估计量 最大似然估计量
估计量的评选标准: 无偏性、有效性、一致性,
参数 θ 的置信度为 $(1 - \alpha)$ 的置信区间,
单个正态总体均值、方差的置信区间.

习 题 七

1. 设总体 X 服从二项分布 $b(n, p)$, n 已知, X_1, X_2, \dots, X_n 为来自 X 的样本, 求参数 p 的矩估计.
2. 设总体 X 的密度函数

$$f(x, \theta) = \begin{cases} \frac{2}{\theta^2}(\theta - x), & 0 < x < \theta, \\ 0, & \text{其他.} \end{cases}$$

X_1, X_2, \dots, X_n 为其样本, 试求参数 θ 的矩法估计.

3. 设总体 X 的密度函数为 $f(x, \theta)$, X_1, X_2, \dots, X_n 为其样本, 求 θ 的最大似然估计.

$$(1) f(x, \theta) = \begin{cases} \theta e^{-\theta x}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

$$(2) f(x, \theta) = \begin{cases} \theta x^{\theta-1}, & 0 < x < 1, \\ 0, & \text{其他.} \end{cases}$$

4. 从一批炒股票的股民一年收益率的数据中随机抽取 10 人的收益率数据, 结果如表 7-5 所示, 求这批股民的收益率的平均收益率及标准差的矩估计值.

表 7-5

序号	1	2	3	4	5	6	7	8	9	10
收益率	0.01	-0.11	-0.12	-0.09	-0.13	-0.3	0.1	-0.09	-0.1	-0.11

5. 随机变量 X 服从 $[0, \theta]$ 上的均匀分布, 今得 X 的样本观测值: 0.9, 0.8, 0.2, 0.8, 0.4, 0.4, 0.7, 0.6, 求 θ 的矩法估计和最大似然估计, 它们是否为 θ 的无偏估计.

6. 设 X_1, X_2, \dots, X_n 是取自总体 X 的样本, $E(X) = \mu$, $D(X) = \sigma^2$, $\hat{\sigma}^2 = k \sum_{i=1}^{n-1} (X_{i+1} - X_i)^2$,

问 k 为何值时, $\hat{\sigma}^2$ 为 σ^2 的无偏估计.

7. 设 X_1, X_2 是从正态总体 $N(\mu, \sigma^2)$ 中抽取的样本

$$\hat{\mu}_1 = \frac{2}{3}X_1 + \frac{1}{3}X_2; \quad \hat{\mu}_2 = \frac{1}{4}X_1 + \frac{3}{4}X_2; \quad \hat{\mu}_3 = \frac{1}{2}X_1 + \frac{1}{2}X_2;$$

试证 $\hat{\mu}_1, \hat{\mu}_2, \hat{\mu}_3$ 都是 μ 的无偏估计量, 并求出每一估计量的方差.

8. 某车间生产的螺钉, 其直径 $X \sim N(\mu, \sigma^2)$, 由过去的经验知道 $\sigma^2 = 0.06$, 今随机抽取 6 枚, 测得其长度 (单位 mm) 如下:

14.7 15.0 14.8 14.9 15.1 15.2

试求 μ 的置信概率为 0.95 的置信区间.

9. 总体 $X \sim N(\mu, \sigma^2)$, σ^2 已知, 问需抽取容量 n 多大的样本, 才能使 μ 的置信概率为 $1 - \alpha$, 且置信区间的长度不大于 L ?

10. 设某种砖头的抗压强度 $X \sim N(\mu, \sigma^2)$, 今随机抽取 20 块砖头, 测得数据如下 ($\text{kg} \cdot \text{cm}^{-2}$):

64 69 49 92 55 97 41 84 88 99
84 66 100 98 72 74 87 84 48 81

(1) 求 μ 的置信概率为 0.95 的置信区间.

(2) 求 σ^2 的置信概率为 0.95 的置信区间.

11. 设总体

$$X \sim f(x) = \begin{cases} (\theta + 1)x^\theta, & 0 < x < 1; \\ 0, & \text{其他.} \end{cases} \quad \text{其中 } \theta > -1$$

X_1, X_2, \dots, X_n 是 X 的一个样本, 求 θ 的矩估计量及最大似然估计量.

12. 设总体

$$X \sim f(x) = \begin{cases} \frac{6x}{\theta^3}(\theta - x), & 0 < x < \theta; \\ 0, & \text{其他.} \end{cases}$$

X_1, X_2, \dots, X_n 为总体 X 的一个样本, 求:

(1) θ 的矩估计量 $\hat{\theta}$;

(2) $D(\hat{\theta})$.

13. 设某种电子元件的使用寿命 X 的概率密度为

$$f(x; \theta) = \begin{cases} 2e^{-2(x-\theta)}, & x > \theta; \\ 0, & x \leq \theta. \end{cases}$$

其中 θ ($\theta > 0$) 为未知参数, 又设 x_1, x_2, \dots, x_n 是总体 X 的一组样本观察值, 求 θ 的最大似然估计值.

14. 设总体 X 的概率分布如表 7-6 所示其中 θ , $0 < \theta < 12$ 是未知参数, 利用总体的如下样本值 3, 1, 3, 0, 3, 1, 2, 3, 求 θ 的矩估计值和最大似然估计值.

表 7-6

X	0	1	2	3
P	θ^2	$2\theta(1-\theta)$	θ^2	$1-2\theta$

(2002 研考)

15. 设总体 X 的分布函数为

$$F(x, \beta) = \begin{cases} 1 - \frac{\alpha^\beta}{x^\beta}, & x > \alpha, \\ 0, & x \leq \alpha. \end{cases}$$

其中未知参数 $\beta > 1$, $\alpha > 0$, 设 X_1, X_2, \dots, X_n 为来自总体 X 的样本

(1) 当 $\alpha=1$ 时, 求 β 的矩估计量;

(2) 当 $\alpha=1$ 时, 求 β 的最大似然估计量;

(3) 当 $\beta=2$ 时, 求 α 的最大似然估计量.

(2004 研考)

16. 从正态总体 $X \sim N(3.4, 6^2)$ 中抽取容量为 n 的样本, 如果其样本均值位于区间 (1.4, 5.4) 内的概率不小于 0.95, 问 n 至少应取多大? (其中表 7-7 为部分标准正态分布表数值)

$$\varphi(z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt \quad (\text{表 7-7})$$

表 7-7

z	1.28	1.645	1.96	2.33
$\varphi(z)$	0.9	0.95	0.975	0.99

(1998 研考)

17. 设总体 X 的概率密度为

$$f(x, \theta) = \begin{cases} \theta, & 0 < x < 1, \\ 1 - \theta, & 1 \leq x < 2, \\ 0, & \text{其他.} \end{cases}$$

其中 θ 是未知参数 ($0 < \theta < 1$), X_1, X_2, \dots, X_n 为来自总体 X 的简单随机样本, 记 N 的样本值 x_1, x_2, \dots, x_n 中小于 1 的个数. 求:

(1) θ 的矩估计;

(2) θ 的最大似然估计.

(2006 研考)

18. 设总体 X 的概率密度为

$$f(x; \theta) = \begin{cases} \frac{1}{2\theta}, & 0 < x < \theta, \\ \frac{1}{2(1-\theta)}, & 0 \leq x < 1, \\ 0, & \text{其他} \end{cases}$$

其中参数 θ , $0 < \theta < 1$ 未知, X_1, X_2, \dots, X_n 是来自总体 X 的简单随机样本, \bar{X} 是样本均值。

(1) 求参数 θ 的矩估计量 $\hat{\theta}$;

(2) 判断 $4\bar{X}^2$ 是否为 θ^2 的无偏估计量, 并说明理由.

(2007 研考)

19. 设 X_1, X_2, \dots, X_n 是来自总体 $N(\mu, \sigma^2)$ 的简单随机样本, 记

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i, S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2, T = \bar{X}^2 - \frac{1}{n} S^2.$$

(1) 证明: T 是 μ^2 的无偏估计量;

(2) 当 $\mu=0$, $\sigma=1$ 时, 求 $D(T)$.

20. 设总体 X 的概率密度为

$$f(x) = \begin{cases} \lambda^2 x e^{-\lambda x}, & \lambda > 0, \\ 0, & \text{其他}. \end{cases}$$

其中参数 λ , $\lambda > 0$ 未知, X_1, X_2, \dots, X_n 是来自总体 X 的简单随机样本. 求:

(1) 参数 λ 的矩估计量;

(2) 参数 λ 的最大似然估计量.

(2009 研考)

21. 设总体 X 的概率分布如表 7-8 所示,

表 7-8

X	1	2	3
p_k	$1-\theta$	$\theta-\theta^2$	θ^2

其中参数 $\theta \in (0, 1)$ 未知. 以 N_i 表示来自总体 X 的简单随机样本 (样本容量为 n) 中等于 i 的

个数, $i=1, 2, 3$, 试求常数 a_1, a_2, a_3 , 使 $T = \sum_{i=1}^3 a_i N_i$ 为 θ 的无偏估计量, 并求 T 的方差.

(2010 研考)

22. 设 X_1, X_2, \dots, X_n 是来自总体 $N(\mu_0, \sigma^2)$ 的简单随机样本, 其中 μ_0 已知, $\sigma^2 > 0$ 未知. \bar{X} 和 S^2 分别表示样本均值和样本方差.

(1) 求参数 σ^2 的最大似然估计 $\hat{\sigma}^2$;

(2) 计算 $E(\hat{\sigma}^2)$ 和 $D(\hat{\sigma}^2)$. (2011 研考)

23. 设随机变量 X 与 Y 相互独立且分别服从正态分布 $N(\mu, \sigma^2)$ 和 $N(\mu, 2\sigma^2)$, 其中 σ 是未知参数且 $\sigma > 0$, 记 $Z = X - Y$.

(1) 求 Z 的概率密度 $f(x; \sigma^2)$;

(2) 设 Z_1, Z_2, \dots, Z_n 为来自总体 Z 的简单随机样本, 求 σ^2 的最大似然估计量 $\hat{\sigma}^2$.

(3) 证明: $\hat{\sigma}^2$ 为 σ^2 的无偏估计量. (2012 研考)

24. 设总体 X 的概率密度为

$$f(x) = \begin{cases} \frac{\theta^2}{x^3} e^{-\frac{\theta}{x}}, & x > 0, \\ 0, & \text{其他}, \end{cases}$$

其中 θ 为未知参数且大于零, X_1, X_2, \dots, X_n 是来自总体 X 的简单随机样本. 求:

(1) θ 的矩估计量;

(2) θ 的最大似然估计量. (2013 研考)

25. 设总体 X 的分布函数为

$$F(x; \theta) = \begin{cases} 1 - e^{-\frac{x^2}{\theta}}, & x \geq 0, \\ \theta, & x < 0, \end{cases}$$

其中 θ 是未知参数且大于零, X_1, X_2, \dots, X_n 是来自总体 X 的简单随机样本.

(1) 求 $E(X)$ 与 $E(X^2)$;

(2) 求 θ 的最大似然估计量 $\hat{\theta}_n$;

(3) 是否存在实数 a , 使得对任何 $\varepsilon > 0$, 都有 $\lim_{n \rightarrow \infty} P\{|\hat{\theta}_n - a| \geq \varepsilon\} = 0$? (2014 研考)