

# 第 10 章 参数估计

设总体  $X$  的分布函数为  $F(X, \theta)$ , 其中  $\theta$  为未知参数(也可为向量). 现从总体中抽取一样本  $X_1, X_2, \dots, X_n$ , 如何依据样本估计参数  $\theta$ , 或  $\theta$  的函数  $g(\theta)$ , 此类问题称为参数估计问题. 内容包括: 点估计, 估计量标准, 区间估计.

## 10.1 点估计

### 10.1.1 矩估计法

总体  $X$  的  $k$  阶矩:  $a_k = E[X^k]$

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

用相应的样本矩去估计总体矩, 求解参数  $\theta$  的方法称为 **矩估计法**. 矩估计法的理论基础是大数定理:  $X_1, X_2, \dots, X_n$  为 i.i.d. 的随机变量, 若  $E(X) = \mu$ , 则当  $n \rightarrow \infty$  时有

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{P} \mu.$$

推论: 若  $E[X^k] = a_k$  存在, 则当  $n \rightarrow \infty$  时有

$$\frac{1}{n} \sum_{i=1}^n X_i^k \xrightarrow{P} a_k = E[X^k].$$

还可利用中心矩进行估计:

总体  $X$  的  $k$  阶中心矩:  $b_k = E[(X - E(X))^k]$  样本  $k$  阶中心矩:  $B_k = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^k$

矩估计方法: 总体  $X$  的分布函数  $F$  包含  $m$  个未知参数  $\theta_1, \theta_2, \dots, \theta_m$ ,

- 1) 求总体  $X$  的  $k$  阶矩:  $a_k = a_k(\theta_1, \theta_2, \dots, \theta_m) = E[X^k], k \in [m]$  ( $a_k$  一般为  $\theta_1, \theta_2, \dots, \theta_m$  的函数).
- 2) 计算样本的  $k$  阶矩:  $A_k = \frac{1}{n} \sum_{i=1}^n X_i^k$ .
- 3) 令样本矩等于总体矩  $A_k = a_k = a_k(\theta_1, \theta_2, \dots, \theta_m)$  ( $k = 1, 2, \dots, m$ ), 得到  $m$  个关于  $\theta_1, \theta_2, \dots, \theta_m$  的方程组.
- 4) 求解方程组得到估计量  $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_m$ .

例 10.1 设总体  $X$  的概率密度函数

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

设  $X_1, X_2, \dots, X_n$  是来自总体  $X$  的样本, 求参数  $\alpha$  的矩估计.

解 首先计算总体  $X$  的期望

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

样本  $X$  的均值  $\bar{X} = \sum_{i=1}^n X_i/n$ . 样本矩等于总体矩有

$$E(X) = \frac{\alpha+1}{\alpha+2} = \bar{X},$$

求解可得  $\alpha = (2\bar{X} - 1)/(1 - \bar{X})$ .

**例 10.2** 设  $X_1, X_2, \dots, X_n$  是来自总体  $X$  的样本, 以及总体  $X$  的密度函数为

$$f(x) = \begin{cases} \frac{1}{\theta} e^{-\frac{x-\mu}{\theta}} & x \geq \mu \\ 0 & \text{其它,} \end{cases}$$

其中  $\theta > 0$ , 求  $\mu$  和  $\theta$  的矩估计.

解 设随机变量  $Y = X - \mu$ , 则  $Y$  服从参数为  $1/\theta$  的指数分布, 有

$$E(Y) = \theta \quad \text{和} \quad \text{Var}(Y) = \theta^2.$$

由此可得  $E(X) = \mu + \theta$  和  $\text{Var}(X) = \theta^2$ . 计算对应的样本矩

$$A_1 = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i, \quad B_2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2.$$

求解方程组

$$\mu + \theta = A_1 \quad \text{和} \quad \theta^2 = B_2,$$

解得  $\mu = \bar{X} - \sqrt{\sum_{i=1}^n (X_i - \bar{X})^2/n}$  和  $\theta = \sqrt{\sum_{i=1}^n (X_i - \bar{X})^2/n}$ .

课堂习题:

- 求正态总体  $\mathcal{N}(\mu, \sigma^2)$  的  $\mu, \sigma^2$  的矩估计法.
- 求总体  $X \sim \mathcal{U}(a, b)$  中  $a, b$  的矩估计法.

### 10.1.2 最大似然估计法

设  $X_1, X_2, \dots, X_n$  是来自总体  $X$  的一个样本. 若总体  $X$  为离散型随机变量, 其分布列为  $\Pr(X = x) = \Pr(X = x; \theta)$ , 则样本  $X_1, X_2, \dots, X_n$  的分布列为

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

这里  $L(\theta)$  表示样本  $X_1 = x_1, X_2 = x_2, \dots, X_n = x_n$  发生的概率.

若总体  $X$  为连续型随机变量, 其概率密度为  $f(x; \theta)$ , 则  $X_1 = x_1, X_2 = x_2, \dots, X_n = x_n$  的联合概率密度为

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

根据概率密度定义可知  $L(\theta)$  越大, 样本  $(X_1, X_2, \dots, X_n)$  落入  $(x_1, x_2, \dots, x_n)$  的邻域内概率越大.

综合上述离散和连续两种随机变量, 统称  $L(\theta)$  为样本  $X_1, X_2, \dots, X_n$  的似然函数, 可以发现  $L(\theta)$  是  $\theta$  的函数, 若

$$\hat{\theta} = \arg \max_{\theta} L(x_1, x_2, \dots, x_n; \theta),$$

则称  $\hat{\theta}$  为  $\theta$  的最大似然估计量. 直觉而言: 最大似然估计量  $\hat{\theta}$  是使观测值  $X_1 = x_1, X_2 = x_2, \dots, X_n = x_n$  出现的概率最大.

求解最大似然估计量的步骤如下:

- i) 计算对数似然函数  $\log(L(x_1, x_2, \dots, x_n; \theta))$ ;
- ii) 求对数似然函数中参数  $\theta$  的一阶偏导, 令其等于零;
- iii) 求解方程组得到最大似然估计量  $\hat{\theta}$ .

**例 10.3** 设  $X_1, X_2, \dots, X_n$  是取自总体  $X \sim B(1, p)$  的样本, 求参数  $p$  的最大似然估计.

**解** 首先计算似然函数

$$L(p) = \prod_{i=1}^n p^{X_i} (1-p)^{1-X_i} = p^{\sum_{i=1}^n X_i} (1-p)^{n-\sum_{i=1}^n X_i},$$

从而得到对数似然函数

$$\ln L(p) = \sum_{i=1}^n X_i \ln p + \left( n - \sum_{i=1}^n X_i \right) \ln(1-p),$$

求一阶偏导并令其为零可得

$$\frac{\partial \ln L(p)}{\partial p} = \frac{1}{p} \sum_{i=1}^n X_i - \frac{1}{1-p} \left( n - \sum_{i=1}^n X_i \right) = 0.$$

由此求解  $p = \sum_{i=1}^n X_i/n = \bar{X}$ . [验证矩估计法]

下面讨论 最大似然估计不可变性

**性质 10.1** 设  $\mu(\theta)$  为  $\theta$  的函数, 且存在反函数  $\theta = \theta(\mu)$ . 若  $\hat{\theta}$  是  $\theta$  的最大似然估计, 则  $\hat{\mu} = \mu(\hat{\theta})$  是  $\mu$  的最大似然估计.

**例 10.4** 设  $X_1, X_2, \dots, X_n$  为总体  $X \sim \mathcal{N}(\mu, \sigma^2)$  的样本, 求  $\mu$  和  $\sigma > 0$  的最大似然估计.

**解** 根据高斯分布知  $X$  的概率密度为  $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ . 样本  $X_1, X_2, \dots, X_n$  的似然函数为

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

其对数似然函数为  $\ln L(\mu, \sigma) = -n \ln(2\pi)^{1/2} - n \ln \sigma - \sum_{i=1}^n (X_i - \mu)^2 / 2\sigma^2$ . 对参数  $\mu$  求导计算可得

$$\frac{\partial \ln L(\mu, \sigma)}{\partial \mu} = \sum_{i=1}^n (X_i - \mu) = 0 \implies \mu = \frac{1}{n} \sum_{i=1}^n X_i = \bar{X},$$

对  $\sigma$  求导计算可得

$$\frac{\partial \ln L(\mu, \sigma)}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^n (X_i - \mu)^2 = 0 \implies \sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2} = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2}.$$

根据最大似然估计的不变性可知方差  $\sigma^2$  的最大似然估计为  $\sigma^2 = \sum_{i=1}^n (X_i - \bar{X})^2/n$ . 下面进行验证最大似然估计的不变性: 设  $X_1, X_2, \dots, X_n$  为总体  $X \sim \mathcal{N}(\mu, \nu)$  的样本, 求  $\mu$  和  $\nu$  的最大似然估计. 根据题意可知样本  $X_1, X_2, \dots, X_n$  的对数似然函数为

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

对参数  $\mu$  求偏导计算其最大似然估计  $\mu = \sum_{i=1}^n X_i/n = \bar{X}$ , 对  $\nu$  求偏导计算可得

$$\frac{\partial \ln L(\mu, \nu)}{\partial \nu} = -\frac{n}{2\nu} + \frac{1}{2\nu^2} \sum_{i=1}^n (X_i - \mu)^2 = 0 \implies \nu = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2,$$

从而完成验证.

**例 10.5** 设总体  $X$  的密度函数为

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

设  $X_1, X_2, \dots, X_n$  是总体  $X$  的样本, 求  $\alpha$  的最大似然估计.

解 首先得到似然函数为

$$L(\alpha) = (\alpha + 1)^n \prod_{i=1}^n X_i^\alpha = (\alpha + 1)^n (X_1 X_2 \cdots X_n)^\alpha,$$

以及其对数似然函数  $\ln L(\alpha) = n \ln(\alpha + 1) + \alpha \ln(X_1 X_2 \cdots X_n)$ . 求导并令偏导为零有

$$\frac{\partial \ln L(\alpha)}{\partial \alpha} = \frac{n}{\alpha + 1} + \ln(X_1 X_2 \cdots X_n) = 0,$$

求解得

$$\alpha = \frac{-n}{\sum_{i=1}^n \ln(X_i)} - 1 = \frac{-1}{\frac{1}{n} \sum_{i=1}^n \ln(X_i)} - 1.$$

对上例, 矩估计值为  $\alpha = (2\bar{X} - 1)/(1 - \bar{X})$ , 因此矩估计值与最大似然估计值可能不同.

**例 10.6** 设  $X_1, X_2, \dots, X_n$  是总体  $X \sim U(a, b)$  的样本, 求  $a$  和  $b$  的最大似然估计.

解 当  $x \in [a, b]$  时, 总体  $X$  的概率密度为  $f(x) = 1/(b - a)$ , 其它情况为零, 因此似然函数为

$$L(a, b) = \begin{cases} \frac{1}{(b-a)^n} & a \leq X_1, X_2, \dots, X_n \leq b \\ 0 & \text{其它} \end{cases}$$

直接求偏导无法解出  $a$  和  $b$ , 此时可以从最大似然的定义出发, 应使得  $b$  尽可能小且  $a$  尽可能大, 但需满足  $a \leq X_1, X_2, \dots, X_n \leq b$ , 因此最大似然估计量为:

$$b = \max\{X_1, X_2, \dots, X_n\} \quad \text{和} \quad a = \min\{X_1, X_2, \dots, X_n\}.$$

**例 10.7** 设  $X_1, X_2, \dots, X_n$  是总体  $X$  的样本, 以及总体  $X$  的概率密度为

$$f(x) = \begin{cases} \theta e^{-(x-\mu)\theta} & x \geq \mu \\ 0 & \text{其它}, \end{cases}$$

求  $\mu$  和  $\theta$  的最大似然估计.

解 首先计算似然函数为

$$L(\theta, \mu) = \begin{cases} \theta^n e^{-\theta \sum_{i=1}^n (X_i - \mu)} & X_i \geq \mu \\ 0 & \text{其它} \end{cases}$$

进一步得到对数似然函数为

$$\ln L(\theta, \mu) = n \ln \theta - \theta \sum_{i=1}^n (X_i - \mu).$$

求偏导、并令偏导等于零有

$$\frac{\partial \ln L(\theta, \mu)}{\partial \theta} = \frac{n}{\theta} - \sum_{i=1}^n (X_i - \mu) = 0 \Rightarrow \theta = \frac{1}{\frac{1}{n} \sum_{i=1}^n (X_i - \mu)},$$

另一方面有

$$\frac{\partial \ln L(\theta, \mu)}{\partial \mu} = n\theta = 0 \Rightarrow \theta = 0,$$

此时无法求解  $\theta$  和  $\mu$  的最大似然估计. 回到似然函数的定义

$$L(\theta, \mu) = \begin{cases} \theta^n e^{-\theta \sum_{i=1}^n (X_i - \mu)} & X_1, X_2, \dots, X_n \geq \mu \\ 0 & \text{其它} \end{cases}$$

可以发现  $\mu$  越大似然函数  $L(\theta, \mu)$  越大, 但须满足  $X_i \geq \mu$  ( $i \in [n]$ ). 由此可得最大似然估计

$$\hat{\mu} = \min\{X_1, X_2, \dots, X_n\},$$

进一步求解可得

$$\hat{\theta} = \frac{1}{\frac{1}{n} \sum_{i=1}^n (X_i - \hat{\mu})}.$$

## 10.2 估计量的评价标准

前一节已经讲过不同的点估计方法, 不同的估计方法可能得到不同的估计值, 自然涉及到一个问题: 采用哪一种估计量更好, 或更好的标准是什么呢? 估计量的常用标准: 无偏性, 有效性, 一致性.

### 10.2.1 无偏性

**定义 10.1** 设  $X_1, X_2, \dots, X_n$  是来自总体  $X$  的样本, 令  $\hat{\theta} = \hat{\theta}(X_1, X_2, \dots, X_n)$  是  $\theta$  的一个估计量, 若

$$E_{X_1, X_2, \dots, X_n} [\hat{\theta}] = E_{X_1, X_2, \dots, X_n} [\hat{\theta}(X_1, X_2, \dots, X_n)] = \theta$$

则称  $\hat{\theta}$  为  $\theta$  的无偏估计.

无偏估计不要求估计值  $\hat{\theta}$  在任意情况下都等于  $\theta$ , 但在期望的情形下有  $E(\hat{\theta}) = \theta$  成立. 其意义在于无系统性偏差, 无偏性是一种对估计量常见而且重要的标准.

首先看看如下例子:

**例 10.8 (样本  $k$  阶原点矩为总体  $k$  阶原点矩的无偏估计)** 设  $X_1, X_2, \dots, X_n$  是总体  $X$  的样本, 若  $E[X^k]$  存在, 则  $A_k = \frac{1}{n} \sum_{i=1}^n X_i^k$  是总体  $a_k = E[X^k]$  的无偏估计.

**例 10.9** 设  $X_1, X_2, \dots, X_n$  是来自总体  $X$  的样本, 其期望为  $\mu$ , 方差为  $\sigma^2$ , 则: 1)  $S_0^2 = \sum_{i=1}^n (X_i - \bar{X})^2 / n$  是  $\sigma^2$  的有偏估计; 2)  $S^2 = \sum_{i=1}^n (X_i - \bar{X})^2 / (n - 1)$  是  $\sigma^2$  的无偏估计.

注意  $\hat{\theta} = \hat{\theta}(X_1, X_2, \dots, X_n)$  是  $\theta$  的无偏估计, 但并不一定有  $g(\hat{\theta})$  是  $g(\theta)$  的无偏估计, 这是因为  $E[\hat{\theta}] = \theta$  并不能推导出  $E[g(\hat{\theta})] = g(\theta)$ . 例如

$$E[\bar{X}] = E[X] = \mu \quad \text{但} \quad E[(\bar{X})^2] \neq \mu^2.$$

**例 10.10** 设  $X_1, X_2, \dots, X_n$  是总体  $X$  的样本, 以及总体  $X$  的概率密度为

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

证明:  $\bar{X} = \sum_{i=1}^n X_i/n$  和  $n \min\{X_1, X_2, \dots, X_n\}$  均是  $\theta$  的无偏估计.

证明 根据期望和指数分布的性质有

$$E[\bar{X}] = E[X] = \theta,$$

由此可知  $\bar{X}$  是  $E[X]$  的无偏估计. 设随机变量  $Z = \min\{X_1, X_2, \dots, X_n\}$ , 则有

$$\begin{aligned} F_Z(z) &= \Pr[Z \leq z] = 1 - \Pr[Z > z] \\ &= 1 - \Pr[X_1 > z] \Pr[X_2 > z] \cdots \Pr[X_n > z] \\ &= 1 - \prod_{i=1}^n (1 - \Pr[X_i \leq z]) = \begin{cases} 0 & z < 0 \\ 1 - e^{-nz/\theta} & z \geq 0. \end{cases} \end{aligned}$$

于是当  $z \geq 0$  时有

$$\Pr[Z > z] = 1 - F_Z(z) = e^{-nz/\theta}.$$

根据期望的性质有

$$E[Z] = \int_0^{+\infty} \Pr[Z > z] dz = \int_0^{+\infty} e^{-nz/\theta} dz = \frac{\theta}{n}.$$

于是有  $\theta = E[nZ]$  成立.

### 10.2.1.1 有效性

参数可能存在多个无偏估计, 若  $\hat{\theta}_1$  和  $\hat{\theta}_2$  都是  $\theta$  的无偏估计, 则可以比较方差

$$\text{Var}(\hat{\theta}_1) = E[(\hat{\theta}_1 - \theta)^2] \quad \text{和} \quad \text{Var}(\hat{\theta}_2) = E[(\hat{\theta}_2 - \theta)^2].$$

一般而言: 方差越小, 无偏估计越好.

**定义 10.2** 设  $\hat{\theta}_1 = \hat{\theta}_1(X_1, X_2, \dots, X_n)$  和  $\hat{\theta}_2 = \hat{\theta}_2(X_1, X_2, \dots, X_n)$  是  $\theta$  的两个无偏估计, 若

$$\text{Var}(\hat{\theta}_1) \leq \text{Var}(\hat{\theta}_2),$$

则称  $\theta_1$  比  $\theta_2$  有效.

**例 10.11** 设  $X_1, X_2, \dots, X_n$  是来自总体  $X$  的样本, 且  $X$  的概率密度为

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

令  $Z = \min\{X_1, X_2, \dots, X_n\}$ , 证明: 当  $n > 1$  时  $\bar{X} = \sum_{i=1}^n X_i/n$  比  $nZ$  有效.

**证明** 根据独立性有

$$\text{Var}(\bar{X}) = \frac{1}{n^2} \sum_{i=1}^n \text{Var}(X_i) = \frac{\theta^2}{n}.$$

根据例 10.10 可知随机变量  $Z$  的概率密度为

$$f(z) = \begin{cases} 0 & z < 0 \\ \frac{n}{\theta} e^{-\frac{nz}{\theta}} & z \geq 0 \end{cases}$$

从而得到

$$\text{Var}(nZ) = n^2 \text{Var}(Z) = n^2 \frac{\theta^2}{n^2} = \theta^2,$$

因此当  $n \geq 1$  时有  $\text{Var}(\bar{X}) \leq \text{Var}(nZ)$  成立, 故  $\bar{X}$  比  $nZ$  有效.

**例 10.12** 设  $X_1, X_2, \dots, X_n$  是总体  $X$  的样本, 且  $E(X) = \mu$  和  $\text{Var}(X) = \sigma^2$ . 设常数  $c_1, c_2, \dots, c_n \geq 0$  满足  $\sum_{i=1}^n c_i = 1, c_i \neq 1/n$ , 求证:  $\bar{X}$  比  $\sum_{i=1}^n c_i X_i$  有效.

**证明** 根据样本的独立同分布条件有

$$E[\bar{X}] = \mu \quad \text{和} \quad \text{Var}(\bar{X}) = \sigma^2/n.$$

根据期望的性质有  $E[\sum_{i=1}^n c_i X_i] = \mu$ , 进一步有

$$\text{Var}\left(\sum_{i=1}^n c_i X_i\right) = \sum_{i=1}^n c_i^2 \text{Var}(X_i) = \sigma^2 \sum_{i=1}^n c_i^2 \geq \frac{\sigma^2}{n}$$

这里利用不等式  $\sum_{i=1}^n c_i^2/n \geq (\sum_{i=1}^n c_i/n)^2 = 1/n^2$ , 所以有  $\text{Var}(\sum_{i=1}^n c_i X_i) \geq \text{Var}(\bar{X})$ .

下面定义有效统计量:

**定理 10.1 (Rao-Crammer 不等式)** 设随机变量  $X$  的概率密度为  $f(x; \theta)$  或分布函数为  $F(x; \theta)$ , 令

$$\text{Var}_0(\theta) = \frac{1}{n E\left[\left(\frac{\partial \ln f(X; \theta)}{\partial \theta}\right)^2\right]} \quad \text{或} \quad \text{Var}_0(\theta) = \frac{1}{n E\left[\left(\frac{\partial \ln F(X; \theta)}{\partial \theta}\right)^2\right]},$$

对任意的无偏估计量  $\hat{\theta}$  有

$$\text{Var}(\hat{\theta}) \geq \text{Var}_0(\theta),$$

称  $\text{Var}_0(\theta)$  为估计量  $\hat{\theta}$  方差的下界. 当  $\text{Var}(\hat{\theta}) = \text{Var}_0(\theta)$  时称  $\hat{\theta}$  为达到方差下界的无偏估计量, 此时  $\hat{\theta}$  为最有效估计量, 简称有效估计量.

**例 10.13** 设  $X_1, X_2, \dots, X_n$  为总体  $X$  的样本, 令总体  $X$  的密度函数为

$$f(x; \theta) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}} & x > 0 \\ 0 & x \leq 0, \end{cases}$$

证明:  $\theta$  的最大似然估计为有效估计量.

**解** 首先计算对数似然函数

$$\ln L(\theta) = -n \ln \theta - \frac{1}{\theta} \sum_{i=1}^n X_i \Rightarrow \hat{\theta} = \frac{1}{n} \sum_{i=1}^n X_i,$$

进一步得到统计量的方差

$$\text{Var}(\hat{\theta}) = \text{Var}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{\theta^2}{n}.$$

同时考察

$$\ln f(X; \theta) = -\ln \theta - \frac{X}{\theta}, \quad \frac{\partial \ln f(X; \theta)}{\partial \theta} = -\frac{1}{\theta} + \frac{X}{\theta^2}$$

所以

$$E\left[\frac{\partial \ln f(X; \theta)}{\partial \theta}\right]^2 = E\left[\left(-\frac{1}{\theta} + \frac{X}{\theta^2}\right)^2\right] = \frac{1}{\theta^4} E[(X - E[X])^2] = \frac{1}{\theta^2},$$

从而得到  $\text{Var}_0(X) = \theta^2/n = \text{Var}(\hat{\theta})$ , 因此  $\theta$  的最大似然估计是有效估计量.

### 10.2.1.2 一致性

**定义 10.3** 设  $\hat{\theta}_n = \hat{\theta}_n(X_1, X_2, \dots, X_n)$  是  $\theta$  的一个估计量, 若当  $n \rightarrow \infty$  时有  $\hat{\theta}_n \xrightarrow{P} \theta$  成立, 即对任意  $\epsilon > 0$  有

$$\lim_{n \rightarrow \infty} \Pr[|\hat{\theta}_n - \theta| > \epsilon] = 0,$$

则称  $\hat{\theta}_n$  为  $\theta$  的一致估计量.

估计量的一致性刻画了在足够多样本情形下估计量  $\hat{\theta}$  能有效逼近真实值  $\theta$ , 一致性是对估计的基本要求, 不满足一致性的估计量一般不予考虑. 下面给出满足一致性的充分条件:

**定理 10.2** 设  $\hat{\theta}_n = \hat{\theta}_n(X_1, X_2, \dots, X_n)$  是  $\theta$  的一个估计量, 若满足以下两个条件:

$$\lim_{n \rightarrow \infty} E[\hat{\theta}_n] = \theta \quad \text{和} \quad \lim_{n \rightarrow \infty} \text{Var}(\hat{\theta}_n) = 0,$$

则  $\hat{\theta}_n$  为  $\theta$  的一致估计量.

**证明** 根据  $\lim_{n \rightarrow \infty} E[\hat{\theta}_n] = \theta$  知道对任意  $\epsilon > 0$ , 存在一个  $N_0$ , 当  $n \geq N_0$  有  $|E[\hat{\theta}_n] - \theta| \leq \theta/2$ , 于是有

$$\lim_{n \rightarrow \infty} \Pr \left[ |E[\hat{\theta}_n] - \theta| > \epsilon/2 \right] = 0.$$

根据 Chebyshev 不等式有

$$\lim_{n \rightarrow \infty} \Pr \left[ |\hat{\theta}_n - E[\hat{\theta}_n]| > \epsilon/2 \right] \leq \lim_{n \rightarrow \infty} \frac{4}{\epsilon} \text{Var}(\hat{\theta}_n) = 0$$

再根据

$$\Pr \left[ |\hat{\theta}_n - \theta| > \epsilon \right] \leq \Pr \left[ |\hat{\theta}_n - E[\hat{\theta}_n]| > \epsilon/2 \right] + \Pr \left[ |E[\hat{\theta}_n] - \theta| > \epsilon/2 \right]$$

完成证明.

**定理 10.3** 设  $\hat{\theta}_{n_1}, \hat{\theta}_{n_2}, \dots, \hat{\theta}_{n_k}$  分别为  $\theta_1, \theta_2, \dots, \theta_k$  满足一致性的估计量, 对连续函数  $g: \mathbb{R}^n \rightarrow \mathbb{R}$ , 有函数  $\hat{\eta}_n = g(\hat{\theta}_{n_1}, \hat{\theta}_{n_2}, \dots, \hat{\theta}_{n_k})$  是  $\eta = g(\theta_1, \theta_2, \dots, \theta_k)$  满足一致性的估计量.

根据大数定理可知样本的  $k$  阶矩是总体  $k$  阶矩的一致估计量. 矩估计法得到的估计量一般是一致估计量. 最大似然估计量在一定条件下是一致性估计量.

**例 10.14** 设  $X_1, X_2, \dots, X_n$  是来自总体  $X$  的样本, 以及总体  $X$  的密度函数为

$$f(x; \theta) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}} & x > 0 \\ 0 & x \leq 0 \end{cases},$$

则样本均值  $\bar{X} = \sum_{i=1}^n X_i/n$  为  $\theta$  的无偏、有效、一致估计量.

由前面的例子可知估计的无偏性和有效性, 一致性可根据  $E[\bar{X}] = \theta$  以及

$$\lim_{n \rightarrow \infty} \text{Var}(\bar{X}) = \lim_{n \rightarrow \infty} \frac{\theta^2}{n} = 0.$$

**例 10.15** 设  $X_1, X_2, \dots, X_n$  是来自总体  $X \sim U(0, \theta)$  的样本, 证明:  $\theta$  的最大似然估计量是一致估计量.

**证明** 根据前面的例题可知  $\theta$  的最大似然估计为  $\hat{\theta}_n = \max(X_1, X_2, \dots, X_n)$ . 设随机变量  $Z = \max(X_1, X_2, \dots, X_n)$ , 则由  $Z$  的分布函数

$$F_Z(z) = \Pr[Z \leq z] = \Pr[\max(X_1, X_2, \dots, X_n) \leq z] = \prod_{i=1}^n \Pr[X_i \leq z] = \begin{cases} 1 & z > \theta \\ (\frac{z}{\theta})^n & z \in [0, \theta] \\ 0 & z < 0. \end{cases}$$

由此得到当  $z \in [0, \theta]$  时随机变量  $Z$  的密度函数  $f_Z(z) = nz^{n-1}/\theta^n$ , 进一步有

$$E[\hat{\theta}_n] = E[Z] = \int_0^\theta \frac{nz^n}{\theta^n} dz = \frac{n}{n+1}\theta,$$

因此  $\hat{\theta}$  是  $\theta$  的有偏估计. 另一方面有

$$E[Z^2] = \int_0^\theta \frac{nz^{n+1}}{\theta^n} dz = \frac{n}{n+2}\theta^2,$$

从而得到

$$\text{Var}(\hat{\theta}_n) = \text{Var}(Z) = E[Z^2] - (E[Z])^2 = \frac{n}{n+2}\theta^2 - \left(\frac{n\theta}{n+1}\right)^2 = \frac{n}{(n+1)^2(n+2)}\theta^2,$$

于是有

$$\lim_{n \rightarrow \infty} E[\hat{\theta}_n] = \theta \quad \text{和} \quad \lim_{n \rightarrow \infty} \text{Var}(\hat{\theta}_n) = 0,$$

由此可得  $\hat{\theta}$  是  $\theta$  的有偏、但一致估计量.

### 10.3 区间估计

**区间估计问题:** 设  $X_1, X_2, \dots, X_n$  是来自总体  $X$  的样本,  $\theta$  为总体  $X$  的分布函数  $F(x, \theta)$  的未知参数, 根据样本估计  $\theta$  的范围  $(\hat{\theta}_1, \hat{\theta}_2)$ , 其中  $\hat{\theta}_1 = \hat{\theta}_1(X_1, X_2, \dots, X_n)$  和  $\hat{\theta}_2 = \hat{\theta}_2(X_1, X_2, \dots, X_n)$ , 使得以较大的概率保证有  $\theta \in (\hat{\theta}_1, \hat{\theta}_2)$  成立. 具体而言, 对任意给定  $\alpha \in (0, 1)$ , 有

$$\Pr\left[\hat{\theta}_1(X_1, X_2, \dots, X_n) < \theta < \hat{\theta}_2(X_1, X_2, \dots, X_n)\right] \geq 1 - \alpha.$$

**定义 10.4 (置信区间与置信度)** 设  $X_1, X_2, \dots, X_n$  是来自总体  $X$  的样本, 总体  $X$  的分布函数含未知参数  $\theta$ , 找出统计量  $\hat{\theta}_1 = \hat{\theta}_1(X_1, X_2, \dots, X_n)$  和  $\hat{\theta}_2 = \hat{\theta}_2(X_1, X_2, \dots, X_n)$  ( $\hat{\theta}_1 < \hat{\theta}_2$ ), 使得

$$\Pr\left[\hat{\theta}_1 < \theta < \hat{\theta}_2\right] \geq 1 - \alpha$$

成立, 则称  $1 - \alpha$  为置信度,  $[\hat{\theta}_1, \hat{\theta}_2]$  为  $\theta$  的置信度为  $1 - \alpha$  的置信区间.

注意: 置信区间  $[\hat{\theta}_1, \hat{\theta}_2]$  是随机区间,  $1 - \alpha$  为该区间包含  $\theta$  的概率/可靠程度. 若  $\alpha = 0.05$ , 则置信度为 95%. 通常采用 95% 的置信度, 有时也可 99% 或 90% 等. 说明:

- i)  $\hat{\theta}_2 - \hat{\theta}_1$  反映了估计精度, 长度越小精度越大.
- ii)  $\alpha$  反映了估计的可靠度,  $\alpha$  越小可靠度越高.
- iii) 给定  $\alpha$ , 区间  $[\hat{\theta}_1, \hat{\theta}_2]$  的选取并不唯一确定, 通常选长度最小的一个区间.

置信区间的求解方法: 枢轴变量法.

- 1) 先找一样本函数  $W(X_1, X_2, \dots, X_n; \theta)$  包含待估参数  $\theta$ , 但不含其它参数, 函数  $W$  的分布已知, 称  $W$  为枢轴变量.
- 2) 给定置信度  $1 - \alpha$ , 根据  $W$  的分布找出临界值  $a$  和  $b$ , 使得  $\Pr[a < W < b] = 1 - \alpha$  成立.
- 3) 根据  $a < W < b$  解出  $\hat{\theta}_1 < \theta < \hat{\theta}_2$ , 则  $(\hat{\theta}_1, \hat{\theta}_2)$  为  $\theta$  的置信度为  $1 - \alpha$  的置信区间.

### 10.3.1 正态总体, 方差已知, 求期望的区间估计

设  $X_1, X_2, \dots, X_n$  是来自总体  $X \sim \mathcal{N}(\mu, \sigma^2)$  的样本, 若方差  $\sigma^2$  已知. 给定  $\alpha \in (0, 1)$ , 确定置信度为  $1 - \alpha$  下  $\mu$  的置信区间  $[\hat{\theta}_1, \hat{\theta}_2]$ . 令样本均值为  $\bar{X} = \sum_{i=1}^n X_i/n$ , 根据正太分布的性质找出枢轴变量:

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

给定置信度  $1 - \alpha$ , 找出临界值  $a$  和  $b$  使得

$$\Pr[a < W < b] = 1 - \alpha.$$

根据正态分布的性质、对称性和上分位点可知

$$\Pr[W \geq \mu_{\alpha/2}] = 1 - \alpha/2 \quad \text{和} \quad \Pr[W \leq -\mu_{\alpha/2}] = 1 - \alpha/2.$$

求解可得  $a = -\mu_{\alpha/2}$  和  $b = \mu_{\alpha/2}$ . 于是有

$$\Pr[-\mu_{\alpha/2} < W < \mu_{\alpha/2}] = 1 - \alpha.$$

根据  $W = (\bar{X} - \mu)/(\sigma/\sqrt{n})$  可得

$$\Pr\left[\bar{X} - \frac{\sigma}{\sqrt{n}}\mu_{\alpha/2} < \mu < \bar{X} + \frac{\sigma}{\sqrt{n}}\mu_{\alpha/2}\right] = 1 - \alpha.$$

**例 10.16** 某地区儿童身高服从正态分布, 现随机抽查 9 人, 高度分别为 115, 120, 131, 115, 109, 115, 115, 105, 110, 已知  $\sigma^2 = 7$  和置信度为 95%, 求期望  $\mu$  的置信区间 ( $\mu_{0.025} = 1.96$ ).

### 10.3.2 正态总体, 方差未知, 求期望的区间估计

设  $X_1, X_2, \dots, X_n$  是来自总体  $X \sim \mathcal{N}(\mu, \sigma^2)$  的样本, 若方差  $\sigma^2$  未知, 考虑期望  $\mu$  的置信度为  $1 - \alpha$  的置信区间. 设  $\bar{X} = \sum_{i=1}^n X_i/n$  和  $S^2 = \sum_{i=1}^n (X_i - \bar{X})^2/(n - 1)$ , 根据正态总体抽样定理可知:

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

由此设枢轴变量

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

给定置信度  $1 - \alpha$ , 设临界值  $a$  和  $b$  满足

$$\Pr[a \leq W \leq b] = 1 - \alpha \Rightarrow b = t_{\alpha/2}(n-1), a = -t_{\alpha/2}(n-1).$$

整理可得

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

### 10.3.3 正态总体, 求方差 $\sigma^2$ 的置信区间

设  $X_1, X_2, \dots, X_n$  是来自总体  $X \sim \mathcal{N}(\mu, \sigma^2)$  的样本, 考虑方差  $\sigma^2$  的置信度为  $1 - \alpha$  的置信区间. 设修正样本方差  $S^2 = \sum_{i=1}^n (X_i - \bar{X})^2 / (n-1)$ , 根据正态总体抽样定理有

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

由此设枢轴变量  $W = (n-1)S^2/\sigma^2$ , 设临界值  $a$  和  $b$  满足

$$\Pr[a \leq W \leq b] = 1 - \alpha.$$

根据  $\chi^2$  分布的不对称性, 采用概率对称的区间

$$\Pr[W \leq a] = \Pr[W \geq b] = \alpha/2 \Rightarrow b = \chi_{\alpha/2}^2(n-1), a = \chi_{1-\alpha/2}^2(n-1).$$

根据枢轴变量  $W = (n-1)S^2/\sigma^2$  可得

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

### 10.3.4 双正态总体情形

设  $X_1, X_2, \dots, X_n$  是来自总体  $X \sim \mathcal{N}(\mu_1, \sigma_1^2)$  的样本, 设  $Y_1, Y_2, \dots, Y_m$  是总体  $Y \sim \mathcal{N}(\mu_2, \sigma_2^2)$  的样本, 令

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i, \quad \bar{Y} = \frac{1}{m} \sum_{i=1}^m Y_i, \quad S_1^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2, \quad S_2^2 = \frac{1}{m-1} \sum_{i=1}^m (Y_i - \bar{Y})^2.$$

考虑  $\mu_1 - \mu_2$  和  $\sigma_1^2/\sigma_2^2$  的置信度为  $1 - \alpha$  的区间估计.

1) 已知方差  $\sigma_1^2$  和  $\sigma_2^2$ , 求  $\mu_1 - \mu_2$  的置信区间. 根据正态分布的性质有

$$\bar{X} \sim \mathcal{N}\left(\mu_1, \frac{\sigma_1^2}{n}\right), \quad \bar{Y} \sim \mathcal{N}\left(\mu_2, \frac{\sigma_2^2}{m}\right) \quad \bar{X} - \bar{Y} \sim \mathcal{N}\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}\right),$$

进一步有

$$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}} \sim \mathcal{N}(0, 1).$$

于是设枢轴变量

$$W = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}} \sim \mathcal{N}(0, 1),$$

求解置信区间

$$\Pr \left[ \bar{X} - \bar{Y} - \mu_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} < \mu_1 - \mu_2 < \bar{X} - \bar{Y} + \mu_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} \right] = 1 - \alpha.$$

2) 若  $\sigma_1^2$  和  $\sigma_2^2$  未知, 但已知  $\sigma_1^2 = \sigma_2^2$ , 设

$$S_W = \frac{(n-1)S_1^2 + (m-1)S_2^2}{n+m-2},$$

则考虑枢轴变量

$$W = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_W \sqrt{\frac{1}{n} + \frac{1}{m}}} \sim t(n+m-2).$$

于是有

$$\Pr \left[ -t_{\alpha/2}(n+m-2) < \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_W \sqrt{\frac{1}{n} + \frac{1}{m}}} < t_{\alpha/2}(n+m-2) \right] = 1 - \alpha.$$

3) 求方差比  $\sigma_1^2/\sigma_2^2$  的置信度为  $1 - \alpha$  的置信区间. 设枢轴变量

$$W = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F(n-1, m-1),$$

根据  $F$  分布的不对称性, 采用概率对称的区间

$$\Pr[W \leq a] = \Pr[W \geq b] = \alpha/2 \Rightarrow b = F_{\frac{\alpha}{2}}(n-1, m-1), \quad a = F_{1-\alpha/2}(n-1, m-1).$$

由此可得置信区间

$$\Pr \left[ \frac{S_1^2}{S_2^2} \frac{1}{F_{\alpha/2}(n-1, m-1)} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{S_1^2}{S_2^2} \frac{1}{F_{1-\alpha/2}(n-1, m-1)} \right] = 1 - \alpha.$$

### 10.3.5 单侧置信区间

对某些实际问题, 我们往往只关心置信区间的上限或下限, 例如, 次品率只关心上限, 产品的寿命只关心下限, 由此引入单侧置信区间及其估计.

**定义 10.5 (单侧置信区间)** 给定  $\alpha \in (0, 1)$ , 若样本  $X_1, \dots, X_n$  的统计量  $\hat{\theta}_1 = \hat{\theta}_1(X_1, X_2, \dots, X_n)$  满足

$$\Pr[\theta > \hat{\theta}_1] \geq 1 - \alpha,$$

则称  $(\hat{\theta}_1, +\infty)$  为  $\theta$  的置信度为  $1 - \alpha$  的单侧置信区间,  $\hat{\theta}_1$  称为单侧置信下限.

同理定义单侧置信上限. 对正态总体, 可以将相关置信区间的估计都扩展到单侧置信估计, 枢轴变量的定理类似, 我们将不再重复讨论, 下面仅举两个实例:

**例 10.17** 设  $X_1, X_2, \dots, X_n$  是来自总体  $X \sim \mathcal{N}(\mu, \sigma^2)$  的样本, 若方差  $\sigma^2$  已知, 求  $\mu$  的置信度为  $1 - \alpha$  的单侧置信下限和上限.

解 设样本均值  $\bar{X} = \sum_{i=1}^n X_i/n$ , 根据正态分布的性质考虑枢轴变量

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

于是有

$$\Pr\left[\frac{\bar{X} - \mu}{S/\sqrt{n}} < \mu_\alpha\right] = 1 - \alpha, \quad \Pr\left[\frac{\bar{X} - \mu}{S/\sqrt{n}} > -\mu_\alpha\right] = 1 - \alpha,$$

整理计算完成估计.

**例 10.18** 从一批出厂的灯泡中随机抽取 10 盏灯泡, 测试其寿命分别为: 1000, 1500, 1250, 1050, 950, 1000, 1150, 1050, 950, 1000, (单位: 小时). 假设这批灯泡的寿命服从正态分布, 求这批灯泡平均寿命的置信度为 95% 的单侧置信下限.

解 首先计算样本均值和样本修正方差分别为

$$\bar{X} = \sum_{i=1}^{10} X_i/10 = 1090 \quad \text{和} \quad S^2 = \sum_{i=1}^{10} (X_i - \bar{X})^2/9 = 8800/3.$$

根据正态分布的性质考虑枢轴变量

$$W = \frac{\bar{X} - \mu}{S/3} \sim t(9),$$

于是有

$$\Pr\left[\frac{\bar{X} - \mu}{S/3} < t_{0.05}(9)\right] = 0.95,$$

查表  $t_{0.05}(9) = 1.833$  可得

$$\mu > \bar{X} - t_{0.05}(9)S/3 = 1090 - \sqrt{8800/3} \times 1.833/3 > 1056.$$

### 10.3.6 非正态分布的区间估计

设  $X_1, X_2, \dots, X_n$  是来自总体  $X$  的样本, 若总体  $X$  的分布未知或非正态分布, 我们可以给出总体期望  $\mu = E[X]$  的区间估计, 方法分为两种: 利用 Concentration 不等式和中心极限定理.

(1) 首先考虑 Concentration 不等式, 若总体  $X \in [a, b]$ , 设  $\bar{X} = \sum_{i=1}^n X_i/n$ , 根据 Concentration 不等式有

$$\Pr[|\mu - \bar{X}| \geq \epsilon] \leq 2 \exp(-2n\epsilon^2/(b-a)^2).$$

令  $\alpha = 2 \exp(-2n\epsilon^2/(b-a)^2)$  求解  $\epsilon = \sqrt{(b-a)^2 \ln(2/\alpha)/n}$ , 于是有

$$\Pr[\bar{X} - \sqrt{(b-a)^2 \ln(2/\alpha)/n} < \mu < \bar{X} + \sqrt{(b-a)^2 \ln(2/\alpha)/n}] > 1 - \alpha.$$

可基于其它 Concentration 不等式给出类似的置信区间估计, 以及其它 sub-Gaussian 型随机变量的期望的置信区间估计.

(2) 利用中心极限定理, 求枢轴变量的近似分布, 再给出置信区间估计. 设总体  $X$  的期望  $E(X) = \mu$ , 方差  $\text{Var}(X) = \sigma^2$ , 利用中心极限定理设枢轴变量

$$W = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}.$$

枢轴变量  $W$  的分布近似于标准正态分布  $\mathcal{N}(0, 1)$ . 当方差  $\sigma^2$  已知时有

$$\Pr[-\mu_{\alpha/2} < \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < \mu_{\alpha/2}] \approx 1 - \alpha.$$

当方差  $\sigma^2$  未时, 用修正样本方差  $S^2 = \sum_{i=1}^n (X_i - \bar{X})^2/(n-1)$  代替方差  $\sigma^2$ , 于是有

$$\Pr[-\mu_{\alpha/2} < \frac{\bar{X} - \mu}{S^2/\sqrt{n}} < \mu_{\alpha/2}] \approx 1 - \alpha.$$

**例 10.19** 设  $X_1, X_2, \dots, X_n$  是来自总体  $X \sim \text{Ber}(p)$  的样本, 求  $p$  的置信度为  $1 - \alpha$  的区间估计.

**解** 根据 Bernoulli 分布的性质有  $X_i \in \{0, 1\}$  以及  $p = E[X]$ , 根据 Chernoff 不等式有

$$\Pr[|\bar{X} - p| > \epsilon p] \leq 2 \exp(-np\epsilon^2/3),$$

设  $\alpha = 2 \exp(-np\epsilon^2/3)$ , 于是有

$$\Pr[\bar{X} - \sqrt{3p \ln(2/\alpha)/n} < p < \bar{X} + \sqrt{3p \ln(2/\alpha)/n}] \geq 1 - \alpha,$$

最后求解  $p$  的置信区间.

方法二: 根据 Bernoulli 分布的性质有  $E[X] = p$  和  $\text{Var}(X) = p(1-p)$ , 设枢轴变量

$$W = \frac{n\bar{X} - np}{\sqrt{np(1-p)}}$$

根据中心极限定理可知  $W$  近似于标准正态分布  $\mathcal{N}(0, 1)$ . 于是有

$$\Pr \left[ -\mu_{\alpha/2} < \frac{n\bar{X} - np}{\sqrt{np(1-p)}} < \mu_{\alpha/2} \right] \approx 1 - \alpha.$$

最后求解  $p$  的近似置信区间.