

第七章 参数估计

1. 随机地取 8 只活塞环,测得它们的直径为(以 mm 计)

$$\begin{array}{cccc} 74.001 & 74.005 & 74.003 & 74.001 \\ 74.000 & 73.998 & 74.006 & 74.002 \end{array}$$

试求总体均值 μ 及方差 σ^2 的矩估计值,并求样本方差 s^2 .

解 总体均值 μ 的矩估计值,总体方差 σ^2 的矩估计值分别为

$$\hat{\mu} = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2.$$

由给出的观察值得

$$\begin{aligned} \hat{\mu} &= 74 + \frac{1}{8} [0.001 + 0.005 + 0.003 + 0.001 \\ &\quad + 0 + (-0.002) + 0.006 + 0.002] \\ &= 74.002, \\ \hat{\sigma}^2 &= \frac{1}{8} \sum_{i=1}^8 (x_i - \bar{x})^2 \\ &= \frac{1}{8} [(-0.001)^2 + 0.003^2 + 0.001^2 \\ &\quad + (-0.001)^2 + (-0.002)^2 + (-0.004)^2 + 0.004^2 + 0^2] \\ &= 6 \times 10^{-6}, \\ s^2 &= \frac{1}{7} \sum_{i=1}^8 (x_i - \bar{x})^2 = \frac{8}{7} \hat{\sigma}^2 = \frac{8}{7} \times 6 \times 10^{-6} = 6.86 \times 10^{-6}. \end{aligned}$$

事实上,只需将观察值输入具有统计功能的计算器,就能直接读出 \bar{x} , $\sqrt{\hat{\sigma}^2}$ 和 s . 读者应掌握计算器的统计功能的用法.

2. 设 X_1, X_2, \dots, X_n 为总体的一个样本, x_1, x_2, \dots, x_n 为一相应的样本值. 求下列各总体的概率密度或分布律中的未知参数的矩估计量和矩估计值.

$$(1) f(x) = \begin{cases} \theta c^\theta x^{-(\theta+1)}, & x > c, \\ 0, & \text{其他,} \end{cases}$$

其中 $c > 0$ 为已知, $\theta > 1$, θ 为未知参数.

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

其中 $\theta > 0$, θ 为未知参数.

$$(3) P\{X=x\} = \binom{m}{x} p^x (1-p)^{m-x}, x=0,1,2,\dots,m,$$

其中 $0 < p < 1$, p 为未知参数.

$$\begin{aligned} \text{解 (1)} \mu_1 &= \int_{-\infty}^{\infty} x f(x) dx = \int_c^{\infty} x \theta c^{\theta} x^{-(\theta+1)} dx \\ &= \theta c^{\theta} \int_c^{\infty} x^{-\theta} dx = \frac{\theta c^{\theta} x^{-\theta+1}}{-\theta+1} \Big|_c^{\infty} = \frac{c\theta}{\theta-1}, \end{aligned}$$

由此得

$$\theta = \frac{\mu_1}{\mu_1 - c}.$$

在上式中以 \bar{X} 代替 μ_1 , 得到 θ 的矩估计量和矩估计值分别为

$$\hat{\theta} = \frac{\bar{X}}{\bar{X} - c}, \quad \hat{\theta} = \frac{\bar{x}}{\bar{x} - c}.$$

$$(2) \mu_1 = \int_0^1 x \sqrt{\theta} x^{\sqrt{\theta}-1} dx = \int_0^1 \sqrt{\theta} x^{\sqrt{\theta}} dx = \frac{\sqrt{\theta}}{\sqrt{\theta}+1}.$$

由此得

$$\theta = \left(\frac{\mu_1}{1 - \mu_1} \right)^2.$$

在上式中以 \bar{X} 代替 μ_1 , 得 θ 的矩估计量和矩估计值分别为

$$\hat{\theta} = \left(\frac{\bar{X}}{1 - \bar{X}} \right)^2, \quad \hat{\theta} = \left(\frac{\bar{x}}{1 - \bar{x}} \right)^2.$$

(3) 因 $\mu_1 = E(X) = mp$, 得 $p = \frac{\mu_1}{m}$, 以 \bar{X} 代替 μ_1 , 得到 p 的矩估计量和矩估计值分别为

$$\hat{p} = \frac{\bar{X}}{m}, \quad \hat{p} = \frac{\bar{x}}{m}.$$

3. 求上题中各未知参数的最大似然估计值和估计量.

解 (1) 设 x_1, x_2, \dots, x_n 是一个样本值. 似然函数为

$$\begin{aligned} L &= L(x_1, x_2, \dots, x_n; \theta) = \prod_{i=1}^n \theta c^{\theta} x_i^{-(\theta+1)} \\ &= (\theta c^{\theta})^n \prod_{i=1}^n x_i^{-(\theta+1)} = (\theta c^{\theta})^n \left(\prod_{i=1}^n x_i \right)^{-(\theta+1)}. \end{aligned}$$

$$\ln L = n(\ln \theta + \theta \ln c) - (\theta + 1) \ln \prod_{i=1}^n x_i.$$

$$\text{令} \quad \frac{d}{d\theta} \ln L = n \left(\frac{1}{\theta} + \ln c \right) - \sum_{i=1}^n \ln x_i = 0,$$

得 θ 的最大似然估计值为

$$\hat{\theta} = 1 / \left(\frac{1}{n} \sum_{i=1}^n \ln x_i - \ln c \right),$$

θ 的最大似然估计量为

$$\hat{\theta} = 1 / \left(\frac{1}{n} \sum_{i=1}^n \ln X_i - \ln c \right).$$

(2) 设 x_1, x_2, \dots, x_n 是一个样本值. 似然函数为

$$L = \prod_{i=1}^n (\sqrt{\theta} x_i^{\sqrt{\theta}-1}) = \theta^{n/2} \left(\prod_{i=1}^n x_i \right)^{\sqrt{\theta}-1},$$

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

令

$$\frac{d}{d\theta} \ln L = \frac{n}{2\theta} + \frac{1}{2\sqrt{\theta}} \sum_{i=1}^n \ln x_i = 0,$$

得 θ 的最大似然估计值为

$$\hat{\theta} = \frac{n^2}{\left(\sum_{i=1}^n \ln x_i \right)^2}.$$

θ 的最大似然估计量为

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

(3) 设 x_1, x_2, \dots, x_n 是一个样本值. 似然函数为

$$L = \prod_{i=1}^n P\{X_i = x_i\} = \prod_{i=1}^n \binom{m}{x_i} p^{x_i} (1-p)^{m-x_i}$$

$$= \left[\prod_{i=1}^n \binom{m}{x_i} \right] p^{\sum_{i=1}^n x_i} (1-p)^{nm - \sum_{i=1}^n x_i},$$

$$\ln L = \ln \prod_{i=1}^n \binom{m}{x_i} + \left(\sum_{i=1}^n x_i \right) \ln p + (nm - \sum_{i=1}^n x_i) \ln (1-p),$$

令

$$\frac{d}{dp} \ln L = \left(\sum_{i=1}^n x_i \right) \frac{1}{p} - (nm - \sum_{i=1}^n x_i) \frac{1}{1-p} = 0,$$

得 p 的最大似然估计值为

$$\hat{p} = \frac{\sum_{i=1}^n x_i}{nm} = \frac{\bar{x}}{m}, \quad \text{其中 } \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i,$$

p 的最大似然估计量为

$$\hat{p} = \frac{\bar{X}}{m}.$$

4. (1) 设总体 X 具有分布律

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

其中 $\theta(0 < \theta < 1)$ 为未知参数. 已知取得了样本值 $x_1 = 1, x_2 = 2, x_3 = 1$. 试求 θ 的矩估计值和最大似然估计值.

(2) 设 X_1, X_2, \dots, X_n 是来自参数为 λ 的泊松分布总体的一个样本, 试求 λ 的最大似然估计量及矩估计量.

(3) 设随机变量 X 服从以 r, p 为参数的负二项分布, 其分布律为

$$P\{X = x_k\} = \binom{x_k - 1}{r - 1} p^r (1 - p)^{x_k - r}, \quad x_k = r, r + 1, \dots,$$

其中 r 已知, p 未知. 设有样本值 x_1, x_2, \dots, x_n , 试求 p 的最大似然估计值.

解 (1) (i) $\mu_1 = \theta^2 + 2 \times 2\theta(1 - \theta) + 3 \times (1 - \theta)^2 = 3 - 2\theta$.

解得 $\theta = \frac{1}{2}(3 - \mu_1)$, 故得 θ 的矩估计值为

$$\hat{\theta} = \frac{1}{2}(3 - \bar{x}).$$

今 $\bar{x} = \frac{1}{3}(x_1 + x_2 + x_3) = \frac{1}{3}(1 + 2 + 1) = \frac{4}{3}$, 故 θ 的矩估计值为 $\hat{\theta} = \frac{5}{6}$.

(ii) 由给定的样本值, 得似然函数为

$$\begin{aligned} L &= \prod_{i=1}^3 P\{X_i = x_i\} = P\{X_1 = 1\}P\{X_2 = 2\}P\{X_3 = 1\} \\ &= \theta^2 \cdot 2\theta(1 - \theta) \cdot \theta^2 = 2\theta^5(1 - \theta), \end{aligned}$$

$$\ln L = \ln 2 + 5 \ln \theta + \ln(1 - \theta).$$

$$\text{令} \quad \frac{d}{d\theta} \ln L = \frac{5}{\theta} - \frac{1}{1 - \theta} = 0,$$

得 θ 的最大似然估计值为 $\hat{\theta} = \frac{5}{6}$.

(2) (i) 设 x_1, x_2, \dots, x_n 是相应于样本 X_1, X_2, \dots, X_n 的样本值, 则似然函数为

$$L = \prod_{i=1}^n \left(\frac{\lambda^{x_i} e^{-\lambda}}{x_i!} \right) = e^{-n\lambda} \lambda^{\sum_{i=1}^n x_i} / \prod_{i=1}^n x_i!,$$

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

$$\text{令} \quad \frac{d}{d\lambda} \ln L = -n + \left(\sum_{i=1}^n x_i \right) / \lambda = 0,$$

得 λ 的最大似然估计值为 $\hat{\lambda} = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}$, 最大似然估计量为

$$\hat{\lambda} = \bar{X}.$$

(ii) 因 $\mu_1 = E(X) = \lambda$, 故 λ 的矩估计量也是 $\hat{\lambda} = \bar{X}$.

(3) 似然函数为

$$\begin{aligned} L(p) &= \prod_{k=1}^n P\{X = x_k\} = \prod_{k=1}^n \binom{x_k-1}{r-1} p^r (1-p)^{x_k-r} \\ &= \left[\prod_{k=1}^n \binom{x_k-1}{r-1} \right] p^{nr} (1-p)^{\sum_{k=1}^n x_k - nr}, \end{aligned}$$

$$\ln L = \ln C + nr \ln p + \left(\sum_{k=1}^n x_k - nr \right) \ln(1-p), C \text{ 为常数.}$$

令
$$\frac{d}{dp}(\ln L) = \frac{nr}{p} - \frac{1}{1-p} \left(\sum_{k=1}^n x_k - nr \right) = 0,$$

得 p 的最大似然估计值为

$$\hat{p} = \frac{r}{\bar{x}}.$$

5. 设某种电子器件的寿命(以 h 计) T 服从双参数的指数分布, 其概率密度为

$$f(t) = \begin{cases} \frac{1}{\theta} e^{-(t-c)/\theta}, & t \geq c, \\ 0, & \text{其他,} \end{cases}$$

其中 c, θ ($c, \theta > 0$) 为未知参数. 自一批这种器件中随机地取 n 件进行寿命试验. 设它们的失效时间依次为 $x_1 \leq x_2 \leq \cdots \leq x_n$.

(1) 求 θ 与 c 的最大似然估计值.

(2) 求 θ 与 c 的矩估计量.

解 (1) 似然函数为

$$L(\theta, c) = L(x_1, x_2, \cdots, x_n; \theta, c)$$

$$= \begin{cases} \prod_{i=1}^n \frac{1}{\theta} e^{-(x_i-c)/\theta}, & x_1, x_2, \cdots, x_n \geq c, \\ 0, & \text{其他.} \end{cases}$$

由题设 $x_1 \leq x_2 \leq \cdots \leq x_n$, 故 $x_1, x_2, \cdots, x_n \geq c$ 相当于 $x_1 \geq c$, 因而上式相当于

$$L(\theta, c) = \begin{cases} \frac{1}{\theta^n} e^{-(\sum_{i=1}^n x_i - nc)/\theta} & c \leq x_1, \\ 0, & c > x_1. \end{cases}$$

可知当 $c \leq x_1$ 时, $L(\theta, c)$ 随 c 的增加而递增, 而当 $c > x_1$ 时 $L(\theta, c) = 0$, 因而对于固定的 θ , $L(\theta, c)$ 在 $c = x_1$ 取到最大值, 从而知应取 $\hat{c} = x_1$.

另外,当 $c \leq x_1$ 时,将 $L(\theta, c)$ 取自然对数得

$$\ln L(\theta, c) = -n \ln \theta - \frac{1}{\theta} \left(\sum_{i=1}^n x_i - nc \right).$$

令 $\frac{\partial}{\partial \theta} \ln L(\theta, c) = 0$, 得

$$\frac{\partial}{\partial \theta} \ln L(\theta, c) = -\frac{n}{\theta} + \frac{\sum_{i=1}^n x_i - nc}{\theta^2} = 0,$$

于是

$$\theta = \bar{x} - c.$$

由此可知 c, θ 的最大似然估计值为

$$\begin{cases} \hat{c} = x_1, \\ \hat{\theta} = \bar{x} - x_1. \end{cases}$$

$$(2) \mu_1 = \int_{-\infty}^{\infty} t f(t) dt = \int_c^{\infty} \frac{t}{\theta} e^{-(t-c)/\theta} dt,$$

令 $u = \frac{t-c}{\theta}$, 得

$$\mu_1 = \int_0^{\infty} (\theta u + c) e^{-u} du = c + \theta \Gamma(2) = c + \theta,$$

$$\mu_2 = \int_{-\infty}^{\infty} t^2 f(t) dt = \int_c^{\infty} \frac{t^2}{\theta} e^{-(t-c)/\theta} dt,$$

令 $u = \frac{t-c}{\theta}$, 得

$$\begin{aligned} \mu_2 &= \int_0^{\infty} (\theta u + c)^2 e^{-u} du = \theta^2 \Gamma(3) + 2c\theta \Gamma(2) + c^2 = 2\theta^2 + 2c\theta + c^2 \\ &= (c + \theta)^2 + \theta^2, \end{aligned}$$

由此得

$$\theta = \sqrt{\mu_2 - \mu_1^2}, \quad c = \mu_1 - \sqrt{\mu_2 - \mu_1^2}.$$

将上两式中的 μ_1, μ_2 分别换成 $A_1 = \bar{X}, A_2 = \frac{1}{n} \sum_{i=1}^n X_i^2$, 并注意到 $A_2 - A_1^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$, 就得到 θ 及 c 的矩估计量为

$$\hat{\theta} = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2}, \quad \hat{c} = \bar{X} - \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2}.$$

6. 一地质学家为研究密歇根湖湖滩地区的岩石成分,随机地自该地区取 100 个样品,每个样品有 10 块石子,记录了每个样品中属石灰石的石子数. 假设这 100 次观察相互独立,并且由过去经验知,它们都服从参数为 $m=10, p$ 的二

项分布, p 是这地区一块石子是石灰石的概率. 求 p 的最大似然估计值. 该地质学家所得的数据如下:

样品中属石灰石的石子数 i	0	1	2	3	4	5	6	7	8	9	10
观察到 i 块石灰石的样品个数	0	1	6	7	23	26	21	12	3	1	0

解 设 X 为一个样品中属石灰石的石子数, 则 $X \sim b(10, p)$, 由本章习题第 3 题知 p 的最大似然估计值为

$$\hat{p} = \frac{\bar{x}}{10}.$$

由给出的数据得

$$\begin{aligned}\bar{x} &= \frac{1}{100} [0 \times 0 + 1 \times 1 + 6 \times 2 + 7 \times 3 + \cdots + 3 \times 8 + 1 \times 9 + 0 \times 10] \\ &= 4.99,\end{aligned}$$

于是 $\hat{p} = \frac{4.99}{10} = 0.499.$

7. (1) 设 X_1, X_2, \dots, X_n 是来自总体 X 的一个样本, 且 $X \sim \pi(\lambda)$, 求 $P\{X=0\}$ 的最大似然估计值.

(2) 某铁路局证实一个扳道员在五年内所引起的严重事故的次数服从泊松分布. 求一个扳道员在五年内未引起严重事故的概率 p 的最大似然估计. 使用下面 122 个观察值. 下表中, r 表示一扳道员五年中引起严重事故的次数, s 表示观察到的扳道员人数.

r	0	1	2	3	4	5
s	44	42	21	9	4	2

解 (1) 设 x_1, x_2, \dots, x_n 是相应于样本 X_1, X_2, \dots, X_n 的样本值. 本题需求

$$P\{X=0\} = \frac{\lambda^0 e^{-\lambda}}{0!} = e^{-\lambda}$$

的最大似然估计.

由第 4 题知 λ 的最大似然估计值 $\hat{\lambda} = \bar{x}$, 又由于函数 $u = e^{-\lambda}$ 具有单值反函数: $\lambda = -\ln u$, 由最大似然估计的不变性知 $P\{X=0\} = e^{-\lambda}$ 的最大似然估计值为

$$\hat{P}\{X=0\} = e^{-\bar{x}}.$$

(2) 由所给数据, 得

$$\begin{aligned}\bar{x} &= \frac{1}{122} \sum_{i=1}^{122} x_i \\ &= \frac{1}{122} (44 \times 0 + 42 \times 1 + 21 \times 2 + 9 \times 3 + 4 \times 4 + 2 \times 5) \\ &= \frac{137}{122}.\end{aligned}$$

由(1)知,扳道员五年内未引起严重事故的概率 $p = P\{X=0\} = e^{-\lambda}$ 的最大似然估计值为

$$\hat{p} = \hat{P}\{X=0\} = e^{-\bar{x}} = e^{-137/122} = 0.3253.$$

8. (1) 设 X_1, X_2, \dots, X_n 是来自概率密度为

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

的总体的样本, θ 未知, 求 $U = e^{-1/\theta}$ 的最大似然估计值.

(2) 设 X_1, X_2, \dots, X_n 是来自正态总体 $N(\mu, 1)$ 的样本. μ 未知, 求 $\theta = P\{X > 2\}$ 的最大似然估计值.

(3) 设 x_1, x_2, \dots, x_n 是来自总体 $b(m, \theta)$ 的样本值, 又 $\theta = \frac{1}{3}(1 + \beta)$, 求 β 的最大似然估计值.

解 (1) 先求 θ 的最大似然估计. 似然函数为

$$L(\theta) = \prod_{i=1}^n \theta x_i^{\theta-1} = \theta^n \left(\prod_{i=1}^n x_i \right)^{\theta-1},$$

$$\ln L(\theta) = n \ln \theta + (\theta - 1) \ln \left(\prod_{i=1}^n x_i \right).$$

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

得 θ 的最大似然估计值为

$$\hat{\theta} = \frac{-n}{\sum_{i=1}^n \ln x_i}. \quad (*_1)$$

$U = e^{-1/\theta}$ 具有单调反函数, 故由最大似然估计的不变性知 U 的最大似然估计值为

$$\hat{U} = e^{-1/\hat{\theta}}, \text{ 其中 } \hat{\theta} \text{ 由 } (*_1) \text{ 所确定.}$$

(2) 已知 μ 的最大似然估计为 $\hat{\mu} = \bar{x}$. 而 $\theta = P\{X > 2\} = 1 - P\{X \leq 2\} = 1 - \Phi(2 - \mu)$ 具有单调反函数. 由最大似然估计的不变性得 $\theta = P\{X > 2\}$ 的最大似然估计值为

$$\hat{\theta} = 1 - \Phi(2 - \hat{\mu}) = 1 - \Phi(2 - \bar{x}).$$

(3) 由本章习题第 3 题知二项分布 $X \sim b(m, \theta)$ 的参数 θ 的最大似然估计为 $\hat{\theta} = \bar{x}/m$. 由最大似然估计的不变性得 $\beta = 3\theta - 1$ 的最大似然估计值为

$$\hat{\beta} = 3\hat{\theta} - 1 = \frac{3\bar{x}}{m} - 1.$$

9. (1) 验证教材第六章 § 3 定理四中的统计量

$$S_w^2 = \frac{n_1 - 1}{n_1 + n_2 - 2} S_1^2 + \frac{n_2 - 1}{n_1 + n_2 - 2} S_2^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

是两总体公共方差 σ^2 的无偏估计量 (S_w^2 称为 σ^2 的合并估计).

(2) 设总体 X 的数学期望为 μ , X_1, X_2, \dots, X_n 是来自 X 的样本, a_1, a_2, \dots, a_n 是任意常数, 验证 $(\sum_{i=1}^n a_i X_i) / \sum_{i=1}^n a_i$ (其中 $\sum_{i=1}^n a_i \neq 0$) 是 μ 的无偏估计量.

证 (1) 注意到 $E(S_1^2) = E(S_2^2) = \sigma^2$, 即有

$$\begin{aligned} E(S_w^2) &= E\left[\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}\right] \\ &= \frac{1}{n_1 + n_2 - 2} [(n_1 - 1)E(S_1^2) + (n_2 - 1)E(S_2^2)] \\ &= \sigma^2, \end{aligned}$$

故 S_w^2 是 σ^2 的无偏估计量.

$$\begin{aligned} (2) E\left[\left(\sum_{i=1}^n a_i X_i\right) / \sum_{i=1}^n a_i\right] &= \sum_{i=1}^n E(a_i X_i) / \sum_{i=1}^n a_i \\ &= \sum_{i=1}^n a_i E(X_i) / \sum_{i=1}^n a_i = \mu \sum_{i=1}^n a_i / \sum_{i=1}^n a_i = \mu, \end{aligned}$$

这就证明了 $(\sum_{i=1}^n a_i X_i) / \sum_{i=1}^n a_i$ 是 μ 的无偏估计量.

10. 设 X_1, X_2, \dots, X_n 是来自总体 X 的一个样本, 设 $E(X) = \mu, D(X) = \sigma^2$.

(1) 确定常数 c , 使 $c \sum_{i=1}^{n-1} (X_{i+1} - X_i)^2$ 为 σ^2 的无偏估计.

(2) 确定常数 c 使 $(\bar{X})^2 - cS^2$ 是 μ^2 的无偏估计 (\bar{X}, S^2 是样本均值和样本方差).

$$\begin{aligned} \text{解 } (1) E\left[c \sum_{i=1}^{n-1} (X_{i+1} - X_i)^2\right] &= c \sum_{i=1}^{n-1} E[(X_{i+1} - X_i)^2] \\ &= c \sum_{i=1}^{n-1} \{D(X_{i+1} - X_i) + [E(X_{i+1} - X_i)]^2\} \\ &= c \sum_{i=1}^{n-1} [D(X_{i+1}) + D(X_i)] = 2\sigma^2(n-1)c \end{aligned}$$

(因 X_{i+1}, X_i 相互独立且 $E(X_{i+1} - X_i) = 0$).

$$\text{要使 } E\left[c \sum_{i=1}^{n-1} (X_{i+1} - X_i)^2\right] = 2\sigma^2(n-1)c = \sigma^2,$$

应取 $c = \frac{1}{2(n-1)}$.

(2) 要使

$$E[(\bar{X})^2 - cS^2] = E(\bar{X}^2) - cE(S^2)$$

$$= \left(\frac{\sigma^2}{n} + \mu^2 \right) - c\sigma^2 = \mu^2,$$

应取 $c = \frac{1}{n}$.

11. 设总体 X 的概率密度为

$$f(x; \theta) = \begin{cases} \frac{1}{\theta} x^{(1-\theta)/\theta}, & 0 < x < 1, \\ 0, & \text{其他.} \end{cases} \quad 0 < \theta < \infty,$$

X_1, X_2, \dots, X_n 是来自总体 X 的样本.

(1) 验证 θ 的最大似然估计量是 $\hat{\theta} = \frac{-1}{n} \sum_{i=1}^n \ln X_i$.

(2) 证明 $\hat{\theta}$ 是 θ 的无偏估计量.

解 (1) 似然函数为

$$L(\theta) = \frac{1}{\theta^n} \prod_{i=1}^n x_i^{(1-\theta)/\theta},$$

$$\ln L(\theta) = -n \ln \theta + \frac{1-\theta}{\theta} \ln \prod_{i=1}^n x_i.$$

$$\text{令 } \frac{d}{d\theta} \ln L(\theta) = -\frac{n}{\theta} + \left(\sum_{i=1}^n \ln x_i \right) \left(\frac{-1}{\theta^2} \right) = 0,$$

$$\text{得 } -n\theta = \sum_{i=1}^n \ln x_i.$$

于是得到 θ 的最大似然估计量为

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

$$(2) \text{ 因 } E[-\ln X] = \int_0^1 (-\ln x) \cdot \frac{1}{\theta} x^{\frac{1}{\theta}-1} dx$$

$$= -x^{\frac{1}{\theta}} \ln x \Big|_0^1 + \int_0^1 \frac{1}{x} x^{\frac{1}{\theta}} dx = \theta.$$

从而知

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

故 $\hat{\theta}$ 为 θ 的无偏估计.

12. 设 X_1, X_2, X_3, X_4 是来自均值为 θ 的指数分布总体的样本, 其中 θ 未知. 设有估计量

$$T_1 = \frac{1}{6}(X_1 + X_2) + \frac{1}{3}(X_3 + X_4),$$

$$T_2 = \frac{1}{5}(X_1 + 2X_2 + 3X_3 + 4X_4),$$

$$T_3 = \frac{1}{4}(X_1 + X_2 + X_3 + X_4).$$

(1) 指出 T_1, T_2, T_3 中哪几个是 θ 的无偏估计量.

(2) 在上述 θ 的无偏估计中指出哪一个较为有效.

解 已知对于均值为 θ 的指数分布总体 X , 有 $E(X) = \theta, D(X) = \theta^2$, 于是 $E(X_i) = \theta, D(X_i) = \theta^2, i=1, 2, 3, 4$. 所以

$$\begin{aligned} E(T_1) &= \frac{1}{6}[E(X_1) + E(X_2)] + \frac{1}{3}[E(X_3) + E(X_4)] \\ &= \frac{\theta}{3} + \frac{2\theta}{3} = \theta, \end{aligned}$$

$$\begin{aligned} E(T_2) &= \frac{1}{5}[E(X_1) + 2E(X_2) + 3E(X_3) + 4E(X_4)] \\ &= \frac{1}{5}(\theta + 2\theta + 3\theta + 4\theta) = 2\theta, \end{aligned}$$

$$\begin{aligned} E(T_3) &= \frac{1}{4}[E(X_1) + E(X_2) + E(X_3) + E(X_4)] \\ &= \frac{1}{4}(\theta + \theta + \theta + \theta) = \theta. \end{aligned}$$

以上结果表明 T_1, T_3 都是 θ 的无偏估计量, 但 T_2 不是 θ 的无偏估计量.

$$\begin{aligned} \text{又 } D(T_1) &= D\left[\frac{1}{6}(X_1 + X_2) + \frac{1}{3}(X_3 + X_4)\right] \\ &= \frac{1}{36}D(X_1 + X_2) + \frac{1}{9}D(X_3 + X_4) \\ &= \frac{1}{36}[D(X_1) + D(X_2)] + \frac{1}{9}[D(X_3) + D(X_4)] \\ &= \frac{2\theta^2}{36} + \frac{2\theta^2}{9} = \frac{5}{18}\theta^2, \end{aligned}$$

$$\begin{aligned} \text{而 } D(T_3) &= D\left[\frac{1}{4}(X_1 + X_2 + X_3 + X_4)\right] = \frac{1}{16}\sum_{i=1}^4 D(X_i) \\ &= \frac{4}{16}\theta^2 = \frac{1}{4}\theta^2 < D(T_1), \end{aligned}$$

故统计量 T_3 较 T_1 有效.

13. (1) 设 $\hat{\theta}$ 是参数 θ 的无偏估计, 且有 $D(\hat{\theta}) > 0$, 试证 $\hat{\theta}^2 = (\hat{\theta})^2$ 不是 θ^2 的无偏估计.

(2) 试证明均匀分布

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

中未知参数 θ 的最大似然估计量不是无偏的.

证 (1) 由 $D(\hat{\theta}) > 0$ 及 $E(\hat{\theta}) = \theta$, 得知 $\hat{\theta}^2 = (\hat{\theta})^2$ 的数学期望为

$$E(\hat{\theta}^2) = E[(\hat{\theta})^2] = D(\hat{\theta}) + [E(\hat{\theta})]^2 = D(\hat{\theta}) + \theta^2 > \theta^2,$$

故 $\hat{\theta}^2 = (\hat{\theta})^2$ 不是 θ^2 的无偏估计量.

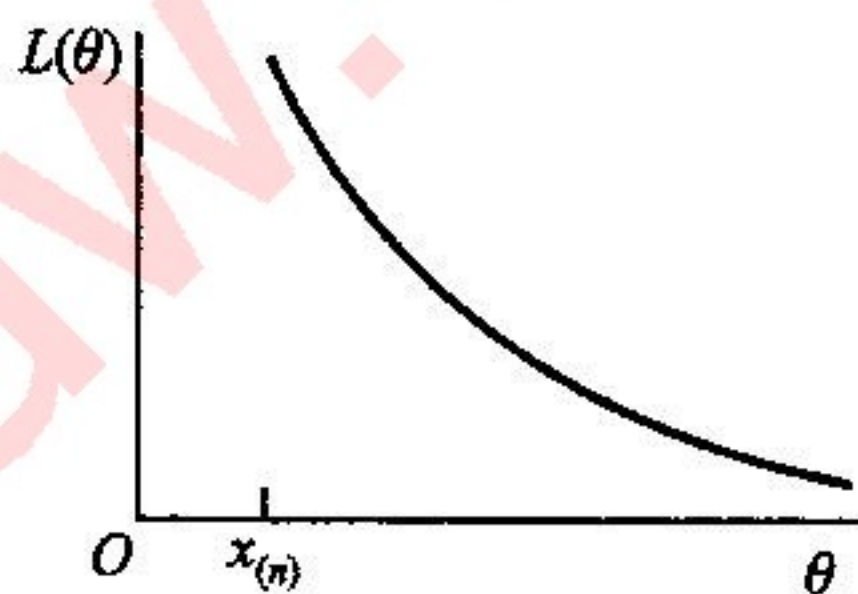
(2) 似然函数为

$$L(\theta) = \begin{cases} \frac{1}{\theta^n}, & 0 < x_1, x_2, \dots, x_n \leq \theta, \\ 0, & \text{其他.} \end{cases}$$

记 $x_{(n)} = \max\{x_1, x_2, \dots, x_n\}$, 由于 $x_1, x_2, \dots, x_n \leq \theta$ 相当于 $x_{(n)} \leq \theta$, 因而上式相当于

$$L(\theta) = \begin{cases} \frac{1}{\theta^n}, & \theta \geq x_{(n)}, \\ 0, & \theta < x_{(n)}. \end{cases}$$

可知当 $\theta < x_{(n)}$ 时, $L(\theta) = 0$; 而当 $\theta \geq x_{(n)}$ 时, $L(\theta)$ 随 θ 的增加而减少; 故 $L(\theta)$ 在 $x_{(n)}$ 处取到最大值 (题 7.13 图), 即得 θ 的最大似然估计值为



题 7.13 图

$$\hat{\theta} = x_{(n)} = \max\{x_1, x_2, \dots, x_n\},$$

于是 θ 的最大似然估计量为

$$\hat{\theta} = \max\{X_1, X_2, \dots, X_n\}.$$

本题总体 X 的分布函数为

$$F(x) = \begin{cases} 0, & x < 0, \\ \frac{x}{\theta}, & 0 \leq x \leq \theta, \\ 1, & x > \theta. \end{cases}$$

得 $\hat{\theta} = \max\{X_1, X_2, \dots, X_n\}$ 的分布函数为

$$F_{\hat{\theta}}(z) = F_{\max}(z) = [F(z)]^n = \begin{cases} 0, & z < 0, \\ \left(\frac{z}{\theta}\right)^n, & 0 \leq z \leq \theta, \\ 1, & z > \theta. \end{cases}$$

由此推得 $\hat{\theta}$ 的概率密度为

$$f_{\hat{\theta}}(z) = f_{\max}(z) = \begin{cases} \frac{n}{\theta} \left(\frac{z}{\theta}\right)^{n-1}, & 0 \leq z \leq \theta, \\ 0, & \text{其他,} \end{cases}$$

$$\begin{aligned}\text{于是 } E(\hat{\theta}) &= \int_{-\infty}^{\infty} z f_{\hat{\theta}}(z) dz = \int_0^{\theta} n \left(\frac{z}{\theta} \right)^n dz \\ &= \frac{n\theta}{n+1} \left(\frac{z}{\theta} \right)^{n+1} \Big|_0^{\theta} = \frac{n\theta}{n+1} \neq \theta,\end{aligned}$$

故 $\hat{\theta}$ 不是 θ 的无偏估计量.

14. 设从均值为 μ , 方差为 $\sigma^2 > 0$ 的总体中分别抽取容量为 n_1, n_2 的两独立样本. \bar{X}_1 和 \bar{X}_2 分别是两样本的均值. 试证: 对于任意常数 a, b ($a+b=1$), $Y = a\bar{X}_1 + b\bar{X}_2$ 都是 μ 的无偏估计, 并确定常数 a, b 使 $D(Y)$ 达到最小.

解 由 $E(\bar{X}_1) = \mu = E(\bar{X}_2)$ 以及 $a+b=1$, 得知

$$\begin{aligned}E(Y) &= E(a\bar{X}_1 + b\bar{X}_2) = aE(\bar{X}_1) + bE(\bar{X}_2) \\ &= a\mu + b\mu = (a+b)\mu = \mu,\end{aligned}$$

即对于任意 a, b , 只要 $a+b=1$, 则 Y 都是 μ 的无偏估计量. 又

$$D(\bar{X}_1) = \frac{\sigma^2}{n_1}, \quad D(\bar{X}_2) = \frac{\sigma^2}{n_2},$$

且 \bar{X}_1, \bar{X}_2 相互独立, 由此得

$$\begin{aligned}D(Y) &= D(a\bar{X}_1 + b\bar{X}_2) = D(a\bar{X}_1) + D(b\bar{X}_2) \\ &= a^2 D(\bar{X}_1) + b^2 D(\bar{X}_2) = \left(\frac{a^2}{n_1} + \frac{b^2}{n_2} \right) \sigma^2.\end{aligned}$$

将 $b=1-a$ 代入上式, 得到

$$D(Y) = \left[\frac{a^2}{n_1} + \frac{(1-a)^2}{n_2} \right] \sigma^2.$$

$$\text{令 } \frac{d}{da} D(Y) = \left[\frac{2a}{n_1} - \frac{2(1-a)}{n_2} \right] \sigma^2 = 0,$$

得

$$a = \frac{n_1}{n_1 + n_2},$$

从而

$$b = 1 - a = \frac{n_2}{n_1 + n_2},$$

又由于

$$\frac{d^2}{da^2} D(Y) = \left(\frac{2}{n_1} + \frac{2}{n_2} \right) \sigma^2 > 0,$$

故知当 $a = \frac{n_1}{n_1 + n_2}, b = \frac{n_2}{n_1 + n_2}$ 时, $D(Y)$ 达到最小.

15. 设有 k 台仪器, 已知用第 i 台仪器测量时, 测定值总体的标准差为 σ_i ($i=1, 2, \dots, k$). 用这些仪器独立地对某一物理量 θ 各观察一次, 分别得到 X_1, X_2, \dots, X_k . 设仪器都没有系统误差, 即 $E(X_i) = \theta$ ($i=1, 2, \dots, k$). 问 a_1, a_2, \dots, a_k 取何值,

方能使使用 $\hat{\theta} = \sum_{i=1}^k a_i X_i$ 估计 θ 时, $\hat{\theta}$ 是无偏的, 并且 $D(\hat{\theta})$ 最小?

解 要使 $\hat{\theta}$ 是 θ 的无偏估计, 则必须

$$\theta = E(\hat{\theta}) = E[a_1 X_1 + a_2 X_2 + \cdots + a_k X_k] = (a_1 + a_2 + \cdots + a_k)\theta.$$

即必须
$$a_1 + a_2 + \cdots + a_k = 1. \quad (\text{A})$$

又由题设知 $D(X_i) = \sigma_i^2, i=1, 2, \cdots, k$, 且 X_1, X_2, \cdots, X_k 相互独立, 故有

$$\begin{aligned} D(\hat{\theta}) &= D\left(\sum_{i=1}^k a_i X_i\right) = \sum_{i=1}^k D(a_i X_i) = \sum_{i=1}^k a_i^2 D(X_i) \\ &= a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2 + \cdots + a_k^2 \sigma_k^2. \end{aligned}$$

为求 $D(\hat{\theta})$ 在条件(A)下的最小值, 用拉格朗日乘数法, 作函数

$$\begin{aligned} g(a_1, a_2, \cdots, a_k, \lambda) \\ = a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2 + \cdots + a_k^2 \sigma_k^2 + \lambda(a_1 + a_2 + \cdots + a_k - 1). \end{aligned}$$

令

$$\begin{aligned} \frac{\partial g}{\partial a_1} &= 2a_1 \sigma_1^2 + \lambda = 0, & \frac{\partial g}{\partial a_2} &= 2a_2 \sigma_2^2 + \lambda = 0, \cdots, \\ \frac{\partial g}{\partial a_k} &= 2a_k \sigma_k^2 + \lambda = 0, & \frac{\partial g}{\partial \lambda} &= \sum_{i=1}^k a_i - 1 = 0. \end{aligned}$$

得

$$\begin{aligned} a_1 &= -\frac{\lambda}{2\sigma_1^2}, & a_2 &= -\frac{\lambda}{2\sigma_2^2}, & \cdots, \\ a_k &= -\frac{\lambda}{2\sigma_k^2}, & \sum_{i=1}^k a_i &= 1. \end{aligned} \quad (\text{B})$$

将前 k 个式子代入末式, 得

$$-\frac{\lambda}{2} \sum_{i=1}^k \frac{1}{\sigma_i^2} = 1.$$

在上式中记 $\sum_{i=1}^k \frac{1}{\sigma_i^2} = \frac{1}{\sigma_0^2}$, 即有 $-\frac{\lambda}{2} = \sigma_0^2$, 代入(B)式, 得

$$a_1 = \frac{\sigma_0^2}{\sigma_1^2}, \quad a_2 = \frac{\sigma_0^2}{\sigma_2^2}, \quad \cdots, \quad a_k = \frac{\sigma_0^2}{\sigma_k^2},$$

即当 $a_i = \frac{\sigma_0^2}{\sigma_i^2}, i=1, 2, \cdots, k$ 时, 用 $\hat{\theta} = \sum_{i=1}^k a_i X_i$ 估计 θ 时, $\hat{\theta}$ 是无偏的且其方差为最小(相对于其他无偏估计来说).

16. 设某种清漆的 9 个样品, 其干燥时间(以 h 计)分别为

6.0 5.7 5.8 6.5 7.0 6.3 5.6 6.1 5.0

设干燥时间总体服从正态分布 $N(\mu, \sigma^2)$. 求 μ 的置信水平为 0.95 的置信区间.

(1) 若由以往经验知 $\sigma=0.6$ h, (2) 若 σ 为未知.

解 (1) 总体 $X \sim N(\mu, \sigma^2)$, μ 未知, σ^2 已知, X_1, X_2, \dots, X_n 是来自 X 的样本, 需要求 μ 的置信水平为 $1-\alpha$ 的置信区间, 即需确定随机区间 $(\underline{\theta}, \bar{\theta})$, 使得

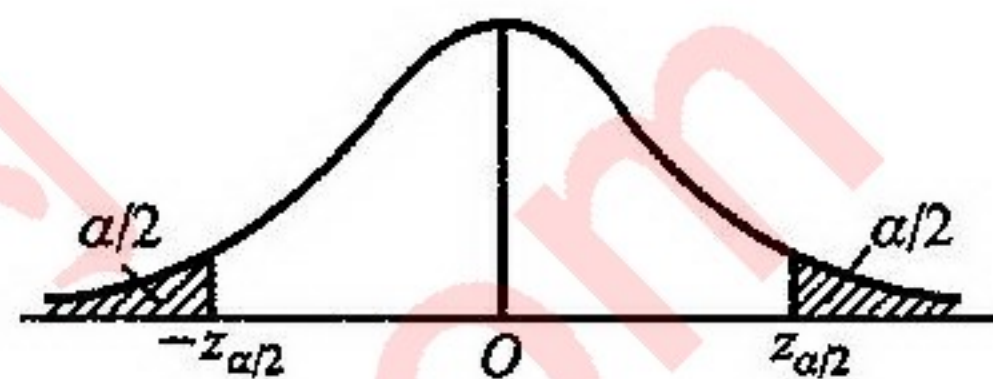
$$P\{\underline{\theta} < \mu < \bar{\theta}\} = 1 - \alpha.$$

考虑

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

因 $\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$, 如题 7.16 图有

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



题 7.16 图

在 $\{ \}$ 内的不等式中解出 μ , 得

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

即得 μ 的一个置信水平为 $1-\alpha$ 的置信区间为

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

今 $n=9, \sigma=0.6, 1-\alpha=0.95, \alpha/2=0.025, z_{0.025}=1.96$, 并算得 $\bar{x}=6$, 得到 μ 的一个置信水平为 0.95 的置信区间为

$$\left(6 \pm \frac{0.6}{3}z_{0.025}\right) = (6 \pm 0.2 \times 1.96) = (5.608, 6.392).$$

(2) σ 为未知, 由于

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

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

在 $\{ \}$ 内的不等式中解出 μ , 得

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

即得 μ 的一个置信水平为 $1-\alpha$ 的置信区间为

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

今 $n=9, 1-\alpha=0.95, \alpha/2=0.025, t_{0.025}(8)=2.306$, 并算得 $\bar{x}=6, s^2=0.33$, 得到 μ 的一个置信水平为 0.95 的置信区间为

$$\left(6 \pm \frac{\sqrt{0.33}}{3} \times 2.306\right) = (6 \pm 0.442) = (5.558, 6.442).$$

17. 分别使用金球和铂球测定引力常数(单位: $10^{-11} \text{m}^3 \cdot \text{kg}^{-1} \cdot \text{s}^{-2}$).

(1) 用金球测定观察值为

6.683 6.681 6.676 6.678 6.679 6.672

(2) 用铂球测定观察值为

6.661 6.661 6.667 6.667 6.664

设测定值总体为 $N(\mu, \sigma^2)$, μ, σ^2 均为未知. 试就(1), (2)两种情况分别求 μ 的置信水平为 0.9 的置信区间, 并求 σ^2 的置信水平为 0.9 的置信区间.

解 (1) $n=6, \bar{x}=6.678, s^2=0.00387^2, 1-\alpha=0.9, \alpha/2=0.05, t_{\alpha/2}(n-1)=t_{0.05}(5)=2.0150$, 得 μ 的一个置信水平为 0.9 的置信区间为

$$\begin{aligned} (\bar{x} \pm \frac{s}{\sqrt{n}} t_{\alpha/2}(n-1)) &= (6.678 \pm \frac{0.00387}{\sqrt{6}} \times 2.0150) \\ &= (6.678 \pm 0.003) = (6.675, 6.681). \end{aligned}$$

由于

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

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

在{ }内的不等式中解出 σ^2 , 得

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

即得 σ^2 的一个置信水平为 $1-\alpha$ 的置信区间为

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

查表知 $\chi_{\alpha/2}^2(n-1) = \chi_{0.05}^2(5) = 11.071, \chi_{1-\alpha/2}^2(5) = \chi_{0.95}^2(5) = 1.145$, 得到 σ^2 的一个置信水平为 0.9 的置信区间为

$$\left(\frac{(6-1) \times 0.00387^2}{11.071}, \frac{(6-1) \times 0.00387^2}{1.145}\right) = (6.8 \times 10^{-6}, 6.5 \times 10^{-5}).$$

(2) $n=5, \bar{x}=6.664, s^2=0.003^2, t_{0.05}(4)=2.1318, \chi_{0.05}^2(4)=9.488, \chi_{0.95}^2(4)=0.711$, 得到 μ 的一个置信水平为 0.9 的置信区间为

$$\begin{aligned} (6.664 \pm \frac{0.003}{\sqrt{5}} \times 2.1318) &= (6.664 \pm 0.003) \\ &= (6.661, 6.667), \end{aligned}$$

σ^2 的一个置信水平为 0.9 的置信区间为

$$\begin{aligned} \left(\frac{(5-1) \times 0.003^2}{9.488}, \frac{(5-1) \times 0.003^2}{0.711}\right) \\ = (3.8 \times 10^{-6}, 5.06 \times 10^{-5}). \end{aligned}$$

18. 随机地取某种炮弹 9 发做试验, 得炮口速度的样本标准差 $s=11$ m/s. 设炮口速度服从正态分布. 求这种炮弹的炮口速度的标准差 σ 的置信水平为

0.95 的置信区间.

解 今 $n=9, s=11, 1-\alpha=0.95, \alpha/2=0.025, \chi_{\alpha/2}^2(n-1)=\chi_{0.025}^2(8)=17.535$, $\chi_{1-\alpha/2}^2(n-1)=\chi_{0.975}^2(8)=2.180$, 得到标准差 σ 的一个置信水平为 0.95 的置信区间为

$$\left(\frac{\sqrt{n-1}S}{\sqrt{\chi_{\alpha/2}^2(n-1)}}, \frac{\sqrt{n-1}S}{\sqrt{\chi_{1-\alpha/2}^2(n-1)}} \right) = \left(\frac{\sqrt{8} \times 11}{\sqrt{17.535}}, \frac{\sqrt{8} \times 11}{\sqrt{2.180}} \right) = (7.4, 21.1).$$

19. 设 X_1, X_2, \dots, X_n 是来自分布 $N(\mu, \sigma^2)$ 的样本, μ 已知, σ 未知.

(1) 验证 $\sum_{i=1}^n (X_i - \mu)^2 / \sigma^2 \sim \chi^2(n)$. 利用这一结果构造 σ^2 的置信水平为 $1-\alpha$ 的置信区间.

(2) 设 $\mu=6.5$, 且有样本值 7.5, 2.0, 12.1, 8.8, 9.4, 7.3, 1.9, 2.8, 7.0, 7.3. 试求 σ 的置信水平为 0.95 的置信区间.

解 (1) 因 $X_i \sim N(\mu, \sigma^2)$, 故 $\frac{X_i - \mu}{\sigma} \sim N(0, 1), i=1, 2, \dots, n$.

由 $\frac{X_1 - \mu}{\sigma}, \frac{X_2 - \mu}{\sigma}, \dots, \frac{X_n - \mu}{\sigma}$ 相互独立, 得

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

于是有 $P\left\{ \chi_{1-\alpha/2}^2(n) < \sum_{i=1}^n \frac{(X_i - \mu)^2}{\sigma^2} < \chi_{\alpha/2}^2(n) \right\} = 1 - \alpha$,

即有 $P\left\{ \frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi_{\alpha/2}^2(n)} < \sigma^2 < \frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi_{1-\alpha/2}^2(n)} \right\} = 1 - \alpha$.

得 σ^2 的置信水平为 $1-\alpha$ 的置信区间为

$$\left\{ \frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi_{\alpha/2}^2(n)}, \frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi_{1-\alpha/2}^2(n)} \right\}$$

(2) 现在 $n=10, \mu=6.5, 1-\alpha=0.95, \alpha=0.05$, 由样本值经计算得

$\sum_{i=1}^{10} (X_i - \mu)^2 = 102.69$, 查表知, $\chi_{0.025}^2(10) = 20.483, \chi_{0.975}^2(10) = 3.247$, 于是

σ^2 的置信水平为 0.95 的置信区间为 (5.013, 31.626). σ 的置信水平为 0.95 的置信区间为 (2.239, 5.624).

20. 在 17 题中, 设用金球和用铂球测定时测定值总体的方差相等. 求两个测定值总体均值差的置信水平为 0.90 的置信区间.

解 由于 $\sigma_1^2 = \sigma_2^2 = \sigma^2$ 未知, 因而

$$\frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2),$$

其中 $S_w = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}$, 即有

$$P\left\{-t_{\alpha/2}(n_1 + n_2 - 2) \leq \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \leq t_{\alpha/2}(n_1 + n_2 - 2)\right\} = 1 - \alpha.$$

在 () 内的不等式中解出 $\mu_1 - \mu_2$, 得

$$\begin{aligned} P\left\{\bar{X}_1 - \bar{X}_2 - t_{\alpha/2}(n_1 + n_2 - 2)S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \leq \mu_1 - \mu_2 \leq \bar{X}_1 - \bar{X}_2 + t_{\alpha/2}(n_1 + n_2 - 2)S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right\} \\ = 1 - \alpha, \end{aligned}$$

即得 $\mu_1 - \mu_2$ 的一个置信水平为 $1 - \alpha$ 的置信区间为

$$(\bar{X}_1 - \bar{X}_2 \pm t_{\alpha/2}(n_1 + n_2 - 2)S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}).$$

今 $n_1 + n_2 - 2 = 9$, $t_{\alpha/2}(n_1 + n_2 - 2) = t_{0.05}(9) = 1.8331$, $\bar{x}_1 - \bar{x}_2 = 0.014$, 得到两个测定值总体均值差 $\mu_{\text{金}} - \mu_{\text{铂}}$ 的一个置信水平为 0.90 的置信区间为

$$\begin{aligned} \left(0.014 \pm 1.8331 \sqrt{\frac{5 \times 0.00387^2 + 4 \times 0.003^2}{9}} \sqrt{\frac{1}{6} + \frac{1}{5}}\right) \\ = (0.014 \pm 0.004) = (0.010, 0.018). \end{aligned}$$

21. 随机地从 A 批导线中抽 4 根, 又从 B 批导线中抽 5 根, 测得电阻 (Ω) 为

A 批导线: 0.143 0.142 0.143 0.137

B 批导线: 0.140 0.142 0.136 0.138 0.140

设测定数据分别来自分布 $N(\mu_1, \sigma^2)$, $N(\mu_2, \sigma^2)$, 且两样本相互独立. 又 μ_1, μ_2, σ^2 均为未知. 试求 $\mu_1 - \mu_2$ 的置信水平为 0.95 的置信区间.

解 将 A 批导线测定数据的均值、方差分别记为 \bar{x}_1, s_1^2 ; B 批的均值、方差分别记为 \bar{x}_2, s_2^2 . 则有,

$$n_1 = 4, \quad \bar{x}_1 = 0.14125, \quad 3s_1^2 = 0.00002475;$$

$$n_2 = 5, \quad \bar{x}_2 = 0.1392, \quad 4s_2^2 = 0.0000208,$$

$$s_w^2 = \frac{3s_1^2 + 4s_2^2}{7} = (0.00255)^2, \quad 1 - \alpha = 0.95, \quad \alpha/2 = 0.025,$$

$$t_{\alpha/2}(n_1 + n_2 - 2) = t_{0.025}(7) = 2.3646.$$

得 $\mu_1 - \mu_2$ 的一个置信水平为 0.95 的置信区间

$$\begin{aligned} & \left(\bar{x}_1 - \bar{x}_2 \pm t_{\alpha/2}(n_1 + n_2 - 2) s_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right) \\ &= \left(0.002\ 05 \pm 2.364\ 6 \times 0.002\ 55 \times \sqrt{\frac{1}{4} + \frac{1}{5}} \right) \\ &= (0.002 \pm 0.004) = (-0.002, 0.006). \end{aligned}$$

22. 研究两种固体燃料火箭推进器的燃烧率. 设两者都服从正态分布, 并且已知燃烧率的标准差均近似地为 0.05 cm/s, 取样本容量为 $n_1 = n_2 = 20$. 得燃烧率的样本均值分别为 $\bar{x}_1 = 18$ cm/s, $\bar{x}_2 = 24$ cm/s, 设两样本独立. 求两燃烧率总体均值差 $\mu_1 - \mu_2$ 的置信水平为 0.99 的置信区间.

解 分别记两种固体燃料火箭推进器的燃烧率总体为 X_1 和 X_2 , 按题意 $X_1 \sim N(\mu_1, \sigma_1^2)$, $X_2 \sim N(\mu_2, \sigma_2^2)$, 其中 μ_1, μ_2 未知, σ_1^2, σ_2^2 已知, 两样本独立, 此时有

$$\begin{aligned} \bar{X}_1 - \bar{X}_2 &\sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right), \\ \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} &\sim N(0, 1), \end{aligned}$$

因而有

$$P\left\{-z_{\alpha/2} < \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} < z_{\alpha/2}\right\} = 1 - \alpha.$$

在 $\{ \}$ 内的不等式中解出 $\mu_1 - \mu_2$, 得

$$\begin{aligned} P\left\{(\bar{X}_1 - \bar{X}_2) - z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{X}_1 - \bar{X}_2) + z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right\} \\ = 1 - \alpha, \end{aligned}$$

即得 $\mu_1 - \mu_2$ 的一个置信水平为 $1 - \alpha$ 的置信区间为

$$\left(\bar{X}_1 - \bar{X}_2 \pm z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \right).$$

今 $\sigma_1 = \sigma_2 = 0.05$, $n_1 = n_2 = 20$, $\bar{x}_1 = 18$, $\bar{x}_2 = 24$, $1 - \alpha = 0.99$, $\alpha/2 = 0.005$, $z_{0.005} = 2.57$, 得到 $\mu_1 - \mu_2$ 的一个置信水平为 0.99 的置信区间为

$$\begin{aligned} & \left(18 - 24 \pm 2.57 \sqrt{\frac{(0.05)^2}{20} + \frac{(0.05)^2}{20}} \right) \\ &= (-6 \pm 0.04) = (-6.04, -5.96). \end{aligned}$$

23. 设两位化验员 A, B 独立地对某种聚合物含氯量用相同的方法各做 10 次测定, 其测定值的样本方差依次为 $s_A^2 = 0.541\ 9$, $s_B^2 = 0.606\ 5$. 设 σ_A^2, σ_B^2 分别为 A, B 所测定的测定值总体的方差. 设总体均为正态的, 且两样本独立. 求方差比 σ_A^2/σ_B^2 的置信水平为 0.95 的置信区间.

解 本题是两个总体均值均未知时,求两总体方差比的置信区间的问题. 已知 $n_1 = n_2 = 10$, $s_A^2 = 0.5419$, $s_B^2 = 0.6065$, $1 - \alpha = 0.95$, $\frac{\alpha}{2} = 0.025$, $F_{0.025}(9, 9) = 4.03$, $F_{0.975}(9, 9) = 0.2481$, 得 σ_A^2/σ_B^2 的一个置信水平为 0.95 的置信区间为

$$\left(\frac{s_A^2}{s_B^2} \frac{1}{F_{0.025}(9, 9)}, \frac{s_A^2}{s_B^2} \frac{1}{F_{0.975}(9, 9)} \right) = (0.222, 3.601).$$

24. 在一批货物的容量为 100 的样本中,经检验发现有 16 只次品,试求这批货物次品率的置信水平为 0.95 的置信区间.

解 本题是(0-1)分布总体 X 的参数的区间估计问题,现在样本容量 $n=100$ 是一个大样本, X 的分布律为

$$f(x, p) = p^x (1-p)^{1-x}, \quad x=0, 1,$$

其中 p 为待估参数,已知 $E(X) = p$, $D(X) = p(1-p)$. 设 X_1, X_2, \dots, X_n 是一个样本,由中心极限定理知,近似地有

$$\frac{\sum_{i=1}^n X_i - np}{\sqrt{np(1-p)}} = \frac{n\bar{X} - np}{\sqrt{np(1-p)}} \stackrel{\text{近似地}}{\sim} N(0, 1),$$

于是有

$$P\left\{-z_{\alpha/2} < \frac{n\bar{X} - np}{\sqrt{np(1-p)}} < z_{\alpha/2}\right\} \approx 1 - \alpha.$$

在()的不等式中解出 p ,得

$$P\{p_1 < p < p_2\} \approx 1 - \alpha,$$

其中

$$p_1 = \frac{1}{2a}(-b - \sqrt{b^2 - 4ac}), \quad p_2 = \frac{1}{2a}(-b + \sqrt{b^2 - 4ac}),$$

$$a = n + z_{\alpha/2}^2, \quad b = -(2n\bar{X} + z_{\alpha/2}^2), \quad c = n\bar{X}^2,$$

即得 p 的一个置信水平为 $1 - \alpha$ 的近似置信区间为

$$(p_1, p_2).$$

今 $n=100$, $\bar{x}=0.16$, $1 - \alpha = 0.95$, $\alpha/2 = 0.025$, $z_{0.025} = 1.96$, 经计算得 $a=103.84$, $b=-35.84$, $c=2.56$, $p_1=0.101$, $p_2=0.244$, 得这批货物的次品率的一个置信水平为 0.95 的近似置信区间为

$$(0.101, 0.244).$$

25. (1) 求第 16 题中 μ 的置信水平为 0.95 的单侧置信上限.

(2) 求第 21 题中 $\mu_1 - \mu_2$ 的置信水平为 0.95 的单侧置信下限.

(3) 求第 23 题中方差比 σ_A^2/σ_B^2 的置信水平为 0.95 的单侧置信上限.

解 (1) 总体 $X \sim N(\mu, \sigma^2)$, μ 未知, 先设 $\sigma=0.6$, X_1, X_2, \dots, X_n 是来自 X 的样本, 需要求 μ 的置信水平为 $1 - \alpha = 0.95$ 的单侧置信上限. 即需确定统计量

$\bar{\mu}$ 使得

$$P\{\mu < \bar{\mu}\} = 1 - \alpha.$$

考虑

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

(因在这里 μ 前有“ $-$ ”号,故在 $\{ \}$ 内写上不等式“ $? < \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$ ”.)

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

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

在 $\{ \}$ 内的不等式中解出 μ , 得

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

于是得 μ 的置信水平为 $1 - \alpha$ 的单侧置信上限为

$$\bar{\mu} = \bar{X} + \frac{\sigma}{\sqrt{n}}z_\alpha.$$

在第 16 题中, $n=9, \bar{x}=6, \sigma=0.6, \alpha=0.05, z_\alpha=1.645$, 得所求的单侧置信上限为

$$\bar{\mu} = 6 + \frac{0.6}{\sqrt{9}} \times 1.645 = 6.329.$$

类似地, 当 σ 未知时, 可得 μ 的置信水平为 $1 - \alpha$ 的单侧置信上限为

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

在第 16 题中, $n=9, \bar{x}=6, s=0.5745, t_\alpha(n-1)=t_{0.05}(8)=1.8595$. 得所求的单侧置信上限为

$$\bar{\mu} = 6 + \frac{0.5745}{\sqrt{9}} \times 1.8595 = 6.356.$$

(2) 在第 21 题中, $\bar{x}_1 - \bar{x}_2 = 0.00205, n_1 = 4, n_2 = 5, 1 - \alpha = 0.95, \alpha = 0.05$, $\mu_1 - \mu_2$ 的置信水平为 0.95 的单侧置信下限为

$$\begin{aligned}\underline{\mu_1 - \mu_2} &= \bar{x}_1 - \bar{x}_2 - t_\alpha(n_1 + n_2 - 2)s_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \\ &= 0.00205 - 1.8946 \times 0.00255 \sqrt{\frac{1}{4} + \frac{1}{5}} \\ &= -0.0012.\end{aligned}$$

(3) 在第 23 题中, $s_A^2/s_B^2 = 0.5419/0.6065, n_1 = n_2 = 10, 1 - \alpha = 0.95, \alpha = 0.05$. 本题需要求的是“?”满足

$$P\left\{\frac{\sigma_A^2}{\sigma_B^2} < ?\right\} = 0.95,$$

注意到

$$\frac{S_A^2/S_B^2}{\sigma_A^2/\sigma_B^2} \sim F(9,9),$$

即有

$$P\left\{\frac{S_A^2/S_B^2}{\sigma_A^2/\sigma_B^2} > F_{0.95}(9,9)\right\} = 0.95,$$

由{ }中的不等式解出 $\frac{\sigma_A^2}{\sigma_B^2} < \frac{S_A^2/S_B^2}{F_{0.95}(9,9)}$, 于是有

$$P\left\{\frac{\sigma_A^2}{\sigma_B^2} < \frac{S_A^2}{S_B^2 F_{0.95}(9,9)}\right\} = 0.95.$$

故得 $\frac{\sigma_A^2}{\sigma_B^2}$ 的置信水平为 0.95 的单侧置信上限为

$$\begin{aligned}\left(\frac{\sigma_A^2}{\sigma_B^2}\right) &= \frac{s_A^2}{s_B^2} \frac{1}{F_{0.95}(9,9)} \\ &= \frac{0.5419}{0.6065} F_{0.95}(9,9) \\ &= \frac{0.5419}{0.6065} \times 3.18 = 2.84.\end{aligned}$$

26. 为研究某种汽车轮胎的磨损特性, 随机地选择 16 只轮胎, 每只轮胎行驶到磨坏为止, 记录所行驶的路程(以 km 计)如下:

41 250 40 187 43 175 41 010 39 265 41 872 42 654 41 287
38 970 40 200 42 550 41 095 40 680 43 500 39 775 40 400

假设这些数据来自正态总体 $N(\mu, \sigma^2)$, 其中 μ, σ^2 未知, 试求 μ 的置信水平为 0.95 的单侧置信下限.

解 由 $\frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t(n-1)$, 有

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

即

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

得 μ 的置信水平为 $1 - \alpha$ 的单侧置信下限为

$$\underline{\mu} = \bar{X} - \frac{S}{\sqrt{n}} t_\alpha(n-1).$$

今 $n=16, \bar{x}=41\,116.875, s=1\,346.842, 1-\alpha=0.95, \alpha=0.05, t_\alpha(n-1)=t_{0.05}(15)=1.753\,1$, 故得 μ 的置信水平为 0.95 的单侧置信下限为

$$\underline{\mu} = 41\,116.875 - \frac{1\,346.842}{4} \times 1.753\,1 = 40\,527.$$

27. 科学上的重大发现往往是由年轻人作出的. 下面列出了自 16 世纪中叶至 20 世纪早期的十二项重大发现的发现者和他们发现时的年龄:

发现内容	发现者	发现时间	年龄
1. 地球绕太阳运转	哥白尼(Copernicus)	1543	40
2. 望远镜、天文学的基本定律	伽利略(Galileo)	1600	36
3. 运动原理、重力、微积分	牛顿(Newton)	1665	23
4. 电的本质	富兰克林(Franklin)	1746	40
5. 燃烧是与氧气联系着的	拉瓦锡(Lavoisier)	1774	31
6. 地球是渐进过程演化成的	莱尔(Lyell)	1830	33
7. 自然选择控制演化的证据	达尔文(Darwin)	1858	49
8. 光的场方程	麦克斯韦(Maxwell)	1864	33
9. 放射性	居里(Curie)	1896	34
10. 量子论	普朗克(Plank)	1901	43
11. 狭义相对论, $E=mc^2$	爱因斯坦(Einstein)	1905	26
12. 量子论的数学基础	薛定谔(Schrödinger)	1926	39

设样本来自正态总体, 试求发现者的平均年龄 μ 的置信水平为 0.95 的单侧置信上限.

解 μ 的置信水平为 $1-\alpha$ 的单侧置信上限为

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

现在 $n=12, \bar{x}=35.58, s=7.22, 1-\alpha=0.95, \alpha=0.05, t_{\alpha}(n-1)=t_{0.05}(11)=1.7959$, 故得

$$\bar{\mu} = 35.58 + \frac{7.22}{\sqrt{12}} \times 1.7959 = 39.32,$$

约为 39 岁零 4 个月.