第四章　参数估计

作者: 徐建军　向一波

中国科学技术大学

统计与金融系

完成日期: 2017年11月24日　　E-mail: xjj1994@mail.ustc.edu.cn

2.

$EX = p_1 + 2p_2 + 3(1 - p_1 - p_2) = 3 - 2p_1 - p_2$

$EX^2 = p_1 + 4p_2 + 9(1 - p_1 - p_2) = 9 - 8p_1 - 5p_2$

令$\overline{X} = \dfrac{n_1 + 2n_2 + 3n_3}{n} = 3 - 2p_1 - p_2$，$\overline{X^2} = \dfrac{n_1 + 4n_2 + 9n_3}{n} = 9 - 8p_1 - 5p_2$

可得矩估计为$\hat{p}_1 = \dfrac{n_1}{n}$，$\hat{p}_2 = \dfrac{n_2}{n}$

5.

$$EX = \int_0^{+\infty} \frac{4x^3}{\theta^3\sqrt{\pi}} e^{-\frac{x^2}{\theta^2}} dx \xlongequal{t=\frac{x^2}{\theta^2}} \frac{2\theta}{\sqrt{\pi}} \int_0^{+\infty} te^{-t} dt = \frac{2\theta}{\sqrt{\pi}}$$

$\therefore \theta$的矩估计$\hat{\theta} = \dfrac{\sqrt{\pi}}{2}\overline{X}$

$$EX^2 = \int_0^{+\infty} \frac{4x^4}{\theta^3\sqrt{\pi}} e^{-\frac{x^2}{\theta^2}} dx \xlongequal{t=\frac{x^2}{\theta^2}} \frac{2\theta^2}{\sqrt{\pi}} \int_0^{+\infty} t^{\frac{3}{2}} e^{-t} dt = \frac{2\theta^2}{\sqrt{\pi}} \Gamma\left(\frac{5}{2}\right) = \frac{3}{2}\theta^2$$

$$\therefore Var(X) = \left(\frac{3}{2} - \frac{4}{\pi}\right)\theta^2$$

$$\therefore Var(\hat{\theta}) = \frac{\pi}{4} Var(\overline{X}) = \frac{\pi}{4n} Var(X) = \frac{\pi}{4n}\left(\frac{3}{2} - \frac{4}{\pi}\right)\theta^2$$

6.

(1)

$P(X = 0) = e^{-\lambda}$

$\because \lambda$的MLE为$\overline{X}$

$\therefore e^{-\lambda}$的MLE为$e^{-\overline{X}}$

(2)

$p = P(X = 0) = e^{-\lambda}$，$MLE$为$e^{-\overline{X}}$

$\overline{X} = \dfrac{44 \times 0 + 42 \times 1 + 21 \times 2 + 9 \times 3 + 4 \times 4 + 2 \times 5}{122} = \dfrac{137}{122}$

9.

似然函数$L(x;\theta) = (\theta^2)^{n_1}\,(2\theta(1-\theta))^{n_2}\,(1-\theta)^{2n_3} = \theta^{2n_1+n_2}(1-\theta)^{n_2+2n_3} \cdot 2^{n_2}$

对数似然$l(x;\theta) = (2n_1 + n_2)\ln\theta + (n_2 + 2n_3)\ln(1-\theta) + n_2 \ln 2$

似然方程$\dfrac{\partial l(x;\theta)}{\partial\theta} = \dfrac{2n_1 + n_2}{\theta} - \dfrac{n_2 + 2n_3}{1-\theta} = 0$

MLE为$\hat{\theta} = \dfrac{2n_1 + n_2}{2n}$

10.

(1)

$L(x;\theta) = \dfrac{1}{(2\theta)^n}\exp\left\{-\dfrac{\sum_{i=1}^n |x_i|}{\theta}\right\}$

$l(x;\theta) = -n\log(2\theta) - \dfrac{\sum_{i=1}^n |x_i|}{\theta}$

令$\dfrac{\partial l(x;\theta)}{\partial\theta} = -\dfrac{n}{\theta} + \dfrac{\sum_{i=1}^n |x_i|}{\theta^2} = 0$

可得MLE为$\hat{\theta} = \dfrac{\sum_{i=1}^n |x_i|}{n}$

(2)

$L(x;\theta) = \mathbf{I}_{(\theta-1/2<x_{(1)}\le x_{(n)}<\theta+1/2)} = \mathbf{I}_{(x_{(n)}-1/2<\theta<x_{(1)}+1/2)}$

$\therefore (X_{(n)} - 1/2\,,\ X_{(1)} + 1/2)$中任意值都是$MLE$

(3)

$L(x;\theta) = \dfrac{1}{(\theta_2-\theta_1)^n}\mathbf{I}_{(\theta_1<x_{(1)}\le x_{(n)}<\theta_2)}$

$\therefore MLE$为$\hat{\theta}_1 = X_{(1)}\,,\ \hat{\theta}_2 = X_{(n)}$

11.

$\theta = P(X \ge 2) = P\left(\dfrac{X-\mu}{\sigma} \ge \dfrac{2-\mu}{\sigma}\right) = 1 - \Phi\left(\dfrac{2-\mu}{\sigma}\right)$

$\therefore \theta$的MLE为$1 - \Phi\left(\dfrac{2-\overline{X}}{\sqrt{m_2}}\right)$，其中$m_2 = \dfrac{1}{n}\sum_{i=1}^n (X_i - \overline{X})^2$

16.

$$L(x;\theta)=\frac{1}{(c-1)^n\theta^n}\mathbf{I}_{(\theta\le x_{(1)}\le x_{(n)}\le c\theta)}=\frac{1}{(c-1)^n\theta^n}\mathbf{I}_{(\frac{x_{(n)}}{c}\le\theta\le x_{(1)})}$$

$\therefore MLE$为$\hat{\theta}=\dfrac{X_{(n)}}{c}$

$\because EX=\dfrac{c+1}{2}\theta$

$\therefore$矩估计为$\hat{\theta}=\dfrac{2}{c+1}\overline{X}$，且$E\hat{\theta}=\theta$，无偏

18.

$n_1\sim B(n,p_1)$，$n_2\sim B(n,p_2)$，$n_3\sim B(n,1-p_1-p_2)$

又$\because p_2=2p_1=2p$

$\therefore E(\hat{p}_1)=\dfrac{1}{n}E(n_1)=\dfrac{1}{n}np=p$

$E(\hat{p}_2)=\dfrac{1}{2n}E(n_2)=\dfrac{1}{2n}n\cdot 2p=p$

$E(\hat{p}_3)=\dfrac{1}{3}-\dfrac{1}{3n}E(n_3)=\dfrac{1}{3}-\dfrac{1}{3n}n\cdot(1-3p)=p$

$\therefore$都是无偏的

$Var(\hat{p}_1)=\dfrac{1}{n^2}np(1-p)=\dfrac{p(1-p)}{n}$

$Var(\hat{p}_2)=\dfrac{1}{4n^2}n\cdot 2p(1-2p)=\dfrac{p(1-2p)}{2n}$

$Var(\hat{p}_3)=\dfrac{1}{9n^2}n\cdot 3p(1-3p)=\dfrac{p(1-3p)}{3n}$

$\therefore Var(\hat{p}_3)$最小

19.

总体X :随机摸出一个硬币连掷两次，掷出正面的次数

$P(X=0)=\dfrac{\theta}{4N}$

$P(X=1)=\dfrac{\theta}{2N}$

$P(X=2)=\dfrac{\theta}{4N}+\dfrac{N-\theta}{N}$

$\therefore EX=2-\dfrac{\theta}{N}$

θ的矩估计为$\hat{\theta}_M=(2-\overline{X})\cdot N$

又$\because \overline{X} = \dfrac{n_1 + 2n_2}{n}$

$\therefore$矩估计$\hat{\theta}_M = \dfrac{(2n_0 + n_1)N}{n}$

$L(x;\theta) = p_0^{n_0} p_1^{n_1} p_2^{n_2}$

$l(x;\theta) = n_0 \ln \dfrac{\theta}{4N} + n_1 \ln \dfrac{\theta}{2N} + n_2 \ln\left(\dfrac{\theta}{4N} + \dfrac{N-\theta}{N}\right)$

令$\dfrac{\partial l(x;\theta)}{\partial \theta} = \dfrac{n_0}{\theta} + \dfrac{n_1}{\theta} - \dfrac{3n_2}{4N - 3\theta} = 0$

可得θ的MLE为$\hat{\theta}_{MLE} = \dfrac{4N(n_0 + n_1)}{3n}$

20.

(1)

$EX = \sigma + \theta$

$\therefore$矩估计$\hat{\theta}_1 = \overline{X} - \sigma$

$L(x;\theta) = \dfrac{1}{\sigma^n} \exp\left\{-\dfrac{\sum_{i=1}^n (x_i - \theta)}{\sigma}\right\} \mathbf{I}_{(x_{(1)} > \theta)}$

$\therefore \theta$的MLE为$\hat{\theta}_2 = X_{(1)}$

(2)

$E\hat{\theta}_1 = \theta$，$\therefore \hat{\theta}_1$ 无偏

下面先求$X_{(1)}$的密度函数

$P(X_{(1)} \le t) = 1 - P(X_{(1)} \ge t) = 1 - P(X_1 \ge t)P(X_2 \ge t)\cdots P(X_n \ge t) \qquad (t > \theta)$

$= 1 - [1 - F_X(t)]^n$

$\therefore X_{(1)}$的密度函数为$f_{X_{(1)}}(t) = n\left[1 - F_X(t)\right]^{n-1} f_X(t) = \dfrac{n}{\sigma} \exp\left\{-\dfrac{n(t-\theta)}{\sigma}\right\} \quad (t > \theta)$

$\therefore E\hat{\theta}_2 = EX_{(1)} = \dfrac{\sigma}{n} + \theta$，不是无偏的

修正$\tilde{\theta}_2 = X_{(1)} - \dfrac{\sigma}{n}$

(3)

$Var(\tilde{\theta}_1) = \dfrac{Var(X)}{n} = \dfrac{\sigma^2}{n}$

$Var(\tilde{\theta}_2) = Var\left(X_{(1)}\right) = \dfrac{\sigma^2}{n^2}$

$\therefore n = 1$时一样，$n > 1$时，$\tilde{\theta}_2$较优

21.

(1) σ 已知

$1-\alpha$置信区间为$\left[\overline{X}-\frac{\sigma u_{\alpha/2}}{\sqrt{n}}, \overline{X}+\frac{\sigma u_{\alpha/2}}{\sqrt{n}}\right]$

其中$n=9$, $\overline{X}=6$, $\sigma=0.6$, $\alpha=0.05$, $u_{\alpha/2}=1.96$

代入计算得95%置信区间为[5.608, 6.392]

(2) σ 未知

$1-\alpha$置信区间为$\left[\overline{X}-\frac{S\,t_{n-1}(\alpha/2)}{\sqrt{n}}, \overline{X}+\frac{S\,t_{n-1}(\alpha/2)}{\sqrt{n}}\right]$

其中$n=9$, $\overline{X}=6$, $S=0.574$, $\alpha=0.05$, $t_{n-1}(\alpha/2)=2.306$

代入计算得95%置信区间为[5.558, 6.442]

23.

(1)

[0.163 , 0.717]

(2)

[−0.146 , 1.026]

25.

方差未知的情形

[1783.85 , 2116.15]

27.

(1) μ已知

记$S_n^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2$，则$\frac{nS_n^2}{\sigma^2}\sim\chi_n^2$

$\because P\left(\chi_n^2(1-\alpha/2)\le\frac{nS_n^2}{\sigma^2}\le\chi_n^2(\alpha/2)\right)=1-\alpha$

$\therefore \sigma^2$的$1-\alpha$置信区间为$\left[\frac{nS_n^2}{\chi_n^2(\alpha/2)}, \frac{nS_n^2}{\chi_n^2(1-\alpha/2)}\right]$

带入数据得[0.142 , 0.893]

(2) μ未知

置信区间为$\left[\frac{(n-1)S^2}{\chi^2_{n-1}(\alpha/2)}, \frac{(n-1)S^2}{\chi^2_{n-1}(1-\alpha/2)}\right]$

带入数据得[0.152, 1.074]

29.

(1)

σ_1和σ_2已知

$\mu_1-\mu_2$的$1-\alpha$置信区间为$\left[(\bar{X}-\bar{Y})-u_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}, (\bar{X}-\bar{Y})+u_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right]$

带入数据[−0.712, 0.412]

(2)

$\sigma_1^2=\sigma_2^2$未知

$\mu_1-\mu_2$的$1-\alpha$置信区间为$\left[(\bar{X}-\bar{Y})-S\,t_{n_1+n_2-2}(\alpha/2)\sqrt{\frac{n_1+n_2}{n_1n_2}}, (\bar{X}-\bar{Y})+S\,t_{n_1+n_2-2}(\alpha/2)\sqrt{\frac{n_1+n_2}{n_1n_2}}\right]$

其中$S^2=\left[\sum_{i=1}^{n_1}(X_i-\bar{X})^2+\sum_{j=1}^{n_2}(Y_j-\bar{Y})^2\right]\Big/(n_1+n_2-2)$

带入数据[−0.648, 0.348]

32.

μ_1和μ_2未知时

σ_1^2/σ_2^2的$1-\alpha$置信区间为$\left[(S_1^2/S_2^2)F_{n_2-1,n_1-1}(1-\alpha/2), (S_1^2/S_2^2)F_{n_2-1,n_1-1}(\alpha/2)\right]$

带入数据[0.22, 3.61]

34.

设n个人中有Y_n个人支持，则$Y_n\sim B(n,p)$

当n足够大时，由中心极限定理

$$\frac{Y_n-np}{\sqrt{np(1-p)}}\sim N(0,1)$$

$$\therefore P\left(-u_{\alpha/2}\le\frac{Y_n-np}{\sqrt{np(1-p)}}\le u_{\alpha/2}\right)\approx 1-\alpha$$

可改写为

$$P(A \le p \le B) \approx 1 - \alpha$$

其中A, B是二次方程

$$\frac{(Y_n - np)^2}{np(1-p)} = u_{\alpha/2}^2$$

的两个根，即

$$A, B = \frac{n}{n + u_{\alpha/2}^2}\left(\hat{p} + \frac{u_{\alpha/2}^2}{2n} \pm u_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n} + \frac{u_{\alpha/2}^2}{4n^2}}\right)$$

A取负号，B取正号，$\hat{p} = Y_n/n$，
带入数据得区间估计为$[0.689, 0.850]$

(2)
$\because P\left(\frac{Y_n - np}{\sqrt{np(1-p)}} \le u_\alpha\right) \approx 1 - \alpha$，可得$p$的$1-\alpha$置信下限为

$$\frac{n}{n + u_\alpha^2}\left(\hat{p} + \frac{u_\alpha^2}{2n} - u_\alpha\sqrt{\frac{\hat{p}(1-\hat{p})}{n} + \frac{u_\alpha^2}{4n^2}}\right)$$

其中u_α为上α分位数，查表得$u_{0.05} = 1.6449$，带入数据得p的95%置信下界为0.705