概率论与数理统计第十三次作业

ZhaohengLi 2017050025

2020 年 5 月 26 日

1 5.4.11

根据题意已知:

$$c(\overline{x}-\mu_1) \sim N(0, \frac{c^2\sigma^2}{n})$$
$$d(\overline{y}-\mu_2) \sim N(0, \frac{d^2\sigma^2}{m})$$

$$\frac{(n-1)s_x^2}{\sigma^2} \sim X^2(n-1)$$
$$\frac{(m-1)s_y^2}{\sigma^2} \sim X^2(m-1)$$

因 $\overline{x}, \overline{y}, s_x^2, s_y^2$ 的独立性,可以得到:

$$c(\overline{x}-\mu_1) + d(\overline{y}-\mu_2) \sim N(0, \frac{c^2\sigma^2}{n} + \frac{d^2\sigma^2}{m})$$

$$\frac{(n+m-2)s_w^2}{\sigma^2} = \frac{(n-1)s_x^2}{\sigma^2} + \frac{(m-1)s_y^2}{\sigma^2} \sim X^2(n+m-2)$$

所以可以得到:

$$t = \frac{c(\overline{x}-\mu_1) + d(\overline{y}-\mu_2)}{s_w\sqrt{\frac{c^2}{n}+\frac{d^2}{n}}} = \frac{[c(\overline{x}-\mu_1) + d(\overline{y}-\mu_2)]/\sqrt{\frac{c^2\sigma^2}{n}+\frac{d^2\sigma^2}{n}}}{\sqrt{\frac{(n+m-2)s_w^2}{\sigma^2}/(n+m-2)}} \sim t(n+m-2)$$

2 5.4.12

根据题意可以知道:

$$x_{n+1} \sim N(\mu, \sigma^2), \overline{x}_n \sim N(\mu, \frac{\sigma^2}{n})$$

$$\frac{(n-1)s_n^2}{\sigma^2} \sim X^2(n-1)$$

$$x_{n+1} - \overline{x}_n \sim N(0, \sigma^2 + \frac{\sigma^2}{n}) = N(0, \frac{n+1}{n}\sigma^2)$$

因此有：

$$t=\frac{(x_{n+1}-\overline{x}_n)/\sqrt{\frac{n+1}{n}}}{s_n}=\frac{(x_{n+1}-\overline{x}_n)/\sqrt{\frac{n+1}{n}\sigma^2}}{\sqrt{\frac{(n-1)s_n^2}{\sigma^2}/(n-1)}}\sim t(n-1)$$

因此自由度为 $n-1$。

3 5.4.19

如果 $X\sim F(x)$，且 $F(x)$ 为连续严格增函数，那么 $Y=F(x)\sim U(0,1)$，$Y=F(x)$ 的分布函数为：

$$F_Y(y)=P(F(X)\leq y)=P(X\leq F^{-1}(y))=F(F^{-1}(y))=y$$

当 $y\leq 0, F_Y(y)=0$，当 $y\geq 1, F_Y(y)=1$，所以 $F(X)\sim U(0,1)$。

如果 $Y\sim U(0,1)$，那么 $Z=-lnY\sim X^2(2)$。当 $z\leq 0, F_X(z)=0$，当 $z>0$，有：

$$F_Z(z)=P(-lnY\leq z)=P(Y\geq e^{-z})=1-e^{-z}$$

这是参数为 1 的指数分布函数，也是自由度为 2 的 X^2 的分布函数，即 $Z=-lnY\sim X^2(2)$。

由 $X_1,X_2,\cdots,X_n$ 的相互独立性可以知道 $F(X_1),F(X_2),\cdots,F(X_n)$ 相互独立，因此 $u=-2\sum_{i=1}^n F(x_i)\sim X^2(2n)$。

4 5.5.1

由几何分布性质可知 $T\sim Nb(n,\theta)$，因此可以得到分布列为：

$$P(T=t)=\binom{n+t-1}{t}\theta^n(1-\theta)^t,\quad t=0,1,2,cdots. \tag{1}$$

在给定 $T=t$ 之后，对任意的一个样本 $x_1,\cdots,x_n(\sum_{i=1}^n=t)$ 有：

$$P(X_1=x_1,\cdots,X_n=x_n|T=t)=\frac{P(X_1=x_1,\cdots,X_{n-1}=x_{n-1},X_n=t-\sum_{i=1}^{n-1}x_i)}{P(T=t)}=\frac{1}{\binom{n+t-1}{t}} \tag{2}$$

该条件分布与 θ 无关，因而 $T=\sum_{i=1}^n x_i$ 是充分统计量。

5 5.5.4

由题意可以知道条件密度函数为：

$$p_\mu(x_1,\cdots,x_n|T=t)=\frac{p_\mu(x_1,\cdots,x_n)}{p_\mu(t)}=\frac{1}{\sqrt{n}}(2\pi)^{-(n-1)/2}exp\{-\frac{1}{2}(\sum_{i=1}^n x_i^2-\frac{t^2}{n})\}$$

可以看到它与 μ 无关，因此 $T=\sum_{i=1}^n x_i$ 是充分统计量。

6 5.5.5

样本的联合密度函数为:

$$p(x_1, x_2, \cdots, x_n; \theta) = \theta^n (x_1 x_2 \cdots x_n)^{\theta-1}$$

令

$$T = \prod_{i=1}^{n} x_i$$

$$g(t; \theta) = t^{\theta-1}\theta^n, h(x_1, \cdots, x_n) = 1$$

由因子分解定理,

$$T = \prod_{i=1}^{n} x_i$$

为 θ 的充分统计量。另外 T 的一一变换得到的统计量，如

$$x_1, \cdots, x_n$$

的几何平均

$$(x_1 \cdots x_n)^{1/n}$$

或其对数

$$\frac{1}{n}\sum_{i=1}^{n} lnx_i$$

都是 θ 的充分统计量。

7 5.5.12

总体的密度函数为:

$$p(x; \theta) = \begin{cases} \frac{1}{\theta}, & \theta < x < 2\theta \\ 0, & others \end{cases} \tag{3}$$

于是样本的联合密度为:

$$p(x_1, \cdots, x_n; \theta) = (\frac{1}{\theta})^n I_{\theta < x(1) < x(n) < 2\theta^n}$$

令 $t = x_{(1)}, t_2 = x_{(n)}$，并取

$$g(t; \theta) = (\frac{1}{\theta})^n I_{\theta < t_1 < t_2 < 2\theta}, h(x) = 1$$

由因子分解定理，$T(t_1, t_2) = (x_{(1)}, x_{(n)})$ 为 θ 的充分统计量。

8 5.5.15

联合密度函数为:

$$p(x_1,\cdots,x_n;\theta)=(C(\theta))^n exp\{\sum_{j=i}^{n}\sum_{i=1}^{k}Q_i(\theta)T_i(x_j)\}\prod_{j=1}^{n}h(x_j)$$

由分子分解定理可以知道，$T(x)=(\sum_{j=1}^n T_i(x_j),\cdots,\sum_{j=1}^n T_k(x_j))$ 为充分统计量。

9 6.1.5

先求得三个统计量的数学期望:

$$E(\hat{\mu}_1)=\frac{1}{2}\mu+\frac{1}{3}\mu+\frac{1}{6}\mu=\mu$$

$$E(\hat{\mu}_2)=\frac{1}{3}\mu+\frac{1}{3}\mu+\frac{1}{3}\mu=\mu$$

$$E(\hat{\mu}_3)=\frac{1}{6}\mu+\frac{1}{6}\mu+\frac{2}{3}\mu=\mu$$

这说明他们都是总体均值 μ 的无偏估计, 下面求他们的方差，设总体的方差为 σ^2，则:

$$Var(\hat{\mu}_1)=\frac{1}{4}\sigma^2+\frac{1}{9}\sigma^2+\frac{1}{36}\sigma^2=\frac{7}{18}\sigma^2$$

$$Var(\hat{\mu}_2)=\frac{1}{9}\sigma^2+\frac{1}{9}\sigma^2+\frac{1}{9}\sigma^2=\frac{1}{3}\sigma^2$$

$$Var(\hat{\mu}_3)=\frac{1}{36}\sigma^2+\frac{1}{36}\sigma^2+\frac{4}{9}\sigma^2=\frac{1}{2}\sigma^2$$

可以看到 $\hat{\mu}_2$ 的有效性最好，$\hat{\mu}_3$ 最差。

10 6.1.10

使用反证法，假设 $T(x_1,\cdots,x_n)$ 为 $g(\theta)$ 的无偏估计，那么:

$$(\frac{1}{\sqrt{2\pi}})^n\int_{-\infty}^{\infty}T(x_1,\cdots,x_n)exp\{-\sum_{i=1}^{n}\frac{(x_i-\theta)^2}{2}\}dx_1\cdots dx_n=|\theta|$$

由上式可以知道，等式左边关于 θ 处处可导，而等式右边在 $\theta=0$ 的时候不存在导数，因此假设不成立。

因此没有无偏估计。

11 6.2.4

(1)

$$E(X)=\frac{1}{3}\theta$$

因此参数 θ 的矩估计为 $\hat{\theta}=3\overline{x}$

(2)

$$E(X)=\frac{\theta+1}{\theta+2}$$

因此参数 θ 的矩估计为 $\hat{\theta}=\frac{1-2\overline{x}}{\overline{x}-1}$

(3)

$$E(X)=\frac{\sqrt{\theta}}{\sqrt{\theta}+1}$$

因此参数 θ 的矩估计为 $\hat{\theta}=(\frac{\overline{x}}{1-\overline{x}})^2$

(4)

总体均值和方差计算结果为:

$$E(X)=\theta+\mu, E(X^2)=2\theta^2+2\mu\theta+\mu^2, Var(X)=\theta^2$$

因此参数估计为 $\hat{\theta}=3, \hat{\mu}=\overline{x}-s$

12 6.3.3

(1) 似然函数:

$$L(\theta)=(\frac{1}{2\theta})^n e^{-\frac{\sum_{i=1}^n |x_i|}{\theta}}$$

对数似然函数为:

$$lnL(\theta)=-nln2\theta-\frac{\sum_{i=1}^n |x_i|}{\theta}$$

对 θ 求导并令其得零得到似然方程，解得:

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

因为

$$\frac{\partial lnL(\theta)}{\partial\theta^2}|_{\hat{\theta}}=\frac{n^3}{(\sum|x_i|)^2}<0$$

因此 $\hat{\theta}$ 是 θ 的最大似然估计。

(2)

此处的似然函数为:

$$L(\theta) = I_{\{\theta-\frac{1}{2}<x_{(1)}<x_{(n)}<\theta+\frac{1}{2}\}}$$

他只有两个取值，0,1，为了使似然函数取 1，θ 的取值范围是 $x_{(n)}-\frac{1}{2}<\theta<x_{(1)}+\frac{1}{2}$，因此 θ 的最大似然估计 $\hat{\theta}$ 可以取范围中的任意值，这说明 MLE 可能不止一个。

(3)

似然函数为：

$$L(\theta) = \frac{1}{(\theta_2-\theta_1)^n} I_{\{\theta_1<x_{(1)}<x_{(n)}<\theta_2\}}$$

要使 $L(\theta)$ 尽可能大，那么先示性函数应为 1，这说明 $\theta_1<x_{(1)}<x_{(n)}<\theta_2$；其次 $\theta_2-\theta_1$ 要尽量小，综上可知，θ_1 的最大的然估计应为 $x_{(1)}$，θ_2 的最大似然估计为 $x_{(n)}$。

13 6.3.7

(1) 根据题意可以求得

$$E(\overline{x}) = \frac{3\theta}{2}, Var(\overline{x}) = \frac{\theta^2}{12n}$$

因为 $E(\hat{\theta}) = \frac{2}{3}E(\overline{x}) = \theta$，这说明 $\hat{\theta} = \frac{2}{3}\overline{x}$ 是参数 θ 的无偏估计，因此可以得到：

$$Var(\hat{\theta}) = \frac{\theta^2}{27n} \to 0$$

因此相合估计得证。

(2) 根据题意，最大似然估计为：

$$\hat{\theta}_{mle} = \frac{x_{(n)}}{2}$$

进一步求得均值和方差分别为：

$$E(x_{(n)}) = \frac{2n+1}{n+1}\theta$$

$$Var(x_{(n)}) = \frac{n\theta^2}{(n+2)(n+1)^2}$$

可以看到：

$$E(\hat{\theta}) = \frac{2n+1}{2(n+1)}\theta \to \theta(n \to +\infty)$$

$$Var(\hat{\theta}) = \frac{n\theta^2}{4(n+1)^2(n+2)} \to 0(n \to +\infty)$$

因此不是无偏估计，是相合估计。

14 6.3.9

根据题意求得均方误差为:

$$MSE(\hat{\theta}_a) = Var(a\overline{x}) + (E(a\overline{x}) - \theta)^2 = \frac{a^2\theta^2}{n} + (a-1)^2\theta^2$$

对 a 求导并令其为 0，可以得到当 $a_0 = \frac{n}{n+1}$ 时，上式取得最小值，并且有:

$$MSE(\hat{\theta}_{a_0}) = \frac{1}{n+1}\theta^2 < \frac{1}{n}\theta^2 = MSE(\overline{x})$$

这证明了在均方误差准则下存在一个优于 $\overline{x}$ 的估计，这也说明，有偏估计并不是总比无偏估计差。