Chapter 1

M-Estimator and MLE - Exercises

Shu Shen

1.1 M-Estimator

Exercise 22.1 Take the model $Y = X'\theta + e$ where e is independent of X and has known density function $f(e)$ which is continuously differentiable.

(a) Show that the conditional density of Y given $X = x$ is $f(y - x'\theta)$.

(b) Find the functions $\rho(Y, X, \theta)$ and $\psi(Y, X, \theta)$.

(c) Calculate the asymptotic covariance matrix.

Solution. 22.1

(a) Since $e = Y - X'\theta$ is independent of X,

$$f\left(Y - X'\theta \mid X\right) = f(e \mid X) = f(e)$$

then the conditional density of Y given $X = x$ is $f(y - x'\theta)$.

(b)

$$\rho(Y, X, \theta) = \left(Y - X'\theta\right)^2$$

$$\begin{aligned}
\psi(Y, X, \theta) &= \frac{\partial}{\partial\theta}\rho(Y, X, \theta) \\
&= \frac{\partial}{\partial\theta}\left(Y - X'\theta\right)^2 \\
&= -2\left(Y - X'\theta\right)X \\
&= -2eX
\end{aligned}$$

(c)

$$\frac{\partial}{\partial\theta}\psi\left(Y,X,\theta\right)=\frac{\partial}{\partial\theta}\left(-2\left(Y-X'\theta\right)X\right)=2XX'$$

the asymptotic covariance matrix is

$$\begin{aligned}&\left(E\left[\frac{\partial}{\partial\theta}\psi_i\left(\theta_0\right)\right]\right)^{-1}E\left[\psi_i\left(\theta_0\right)\psi_i'\left(\theta_0\right)\right]\left(E\left[\frac{\partial}{\partial\theta}\psi_i\left(\theta_0\right)\right]\right)^{-1}\\&=\left(E\left[2XX'\right]\right)^{-1}E\left[\left(-2eX\right)\left(-2eX\right)'\right]\left(E\left[2XX'\right]\right)^{-1}\\&=\left(E\left[XX'\right]\right)^{-1}E\left[e^2XX'\right]\left(E\left[XX'\right]\right)^{-1}\end{aligned}$$

Exercise 22.4 For the estimator described in Exercise 22.2 (Take the model $Y=X'\theta+e$. Consider the m-estimator of θ with $\rho\left(Y,X,\theta\right)=g\left(Y-X'\theta\right)$.) set $g\left(u\right)=1-\cos\left(u\right)$.

(a) Sketch $g\left(u\right)$. Is $g\left(u\right)$ continuous? Differentiable? Second differentiable?

(b) Find the functions $\rho\left(Y,X,\theta\right)$ and $\psi\left(Y,X,\theta\right)$.

(c) Calculate the asymptotic covariance matrix.

Solution. 22.4

(a)

$g\left(u\right)$ is continuous.

Since $\frac{\partial}{\partial u}g\left(u\right)=\sin u$ and $\frac{\partial}{\partial u}g\left(u\right)$ is continuous, $g\left(u\right)$ is differentiable.

Since $\frac{\partial^2}{\partial u^2}g\left(u\right)=\cos u$ and $\frac{\partial^2}{\partial u^2}g\left(u\right)$ is continuous, $g\left(u\right)$ is second differentiable.

(b)

$$\rho\left(Y,X,\theta\right)=g\left(Y-X'\theta\right)=1-\cos\left(Y-X'\theta\right)$$

$$\begin{aligned}\psi\left(Y,X,\theta\right)&=\frac{\partial}{\partial\theta}\rho\left(Y,X,\theta\right)\\&=\frac{\partial}{\partial\theta}\left(1-\cos\left(Y-X'\theta\right)\right)\\&=\sin\left(e\right)X\end{aligned}$$

(c)

$$\frac{\partial}{\partial\theta}\psi\left(Y,X,\theta\right)=\frac{\partial}{\partial\theta}\left(\sin\left(Y-X'\theta\right)X\right)=\cos\left(e\right)XX'$$

the asymptotic covariance matrix is

$$\begin{aligned}&\left(E\left[\frac{\partial}{\partial\theta}\psi_i\left(\theta_0\right)\right]\right)^{-1}E\left[\psi_i\left(\theta_0\right)\psi_i'\left(\theta_0\right)\right]\left(E\left[\frac{\partial}{\partial\theta}\psi_i\left(\theta_0\right)\right]\right)^{-1}\\&=\left(E\left[\cos\left(e\right)XX'\right]\right)^{-1}E\left[\left(\sin\left(e\right)X\right)\left(\sin\left(e\right)X\right)'\right]\left(E\left[\cos\left(e\right)XX'\right]\right)^{-1}\\&=\left(E\left[\cos\left(e\right)XX'\right]\right)^{-1}E\left[\sin^2\left(e\right)XX'\right]\left(E\left[\cos\left(e\right)XX'\right]\right)^{-1}\end{aligned}$$

1.2 MLE

Exercise 10.1 Let X be distributed Poisson: $\pi(k) = \frac{\exp(-\theta)\theta^k}{k!}$ for nonnegative integer k and $\theta > 0$.

(a) Find the log-likelihood function $\ell_n(\theta)$.

(b) Find the MLE $\hat{\theta}$ for θ.

Solution. 10.1

(a) The mass function is

$$\pi(k \mid \theta) = \frac{\exp(-\theta)\,\theta^k}{k!}$$

The log mass function is

$$\log \pi(k \mid \theta) = -\theta + k \log \theta - \log k!$$

The log-likelihood function is

$$\ell_n(\theta) = \sum_{i=1}^{n} \left(-\theta + k_i \log \theta - \log k_i!\right) = -n\theta + n\bar{k}_n \log \theta - \sum_{i=1}^{n} \log k_i!$$

(b) The F.O.C. for $\hat{\theta}$ is

$$\frac{\partial}{\partial \theta} \ell_n\left(\hat{\theta}\right) = -n + \frac{n\bar{k}_n}{\hat{\theta}} = 0$$

The solution is

$$\hat{\theta} = \bar{k}_n$$

The S.O.C. is

$$\frac{\partial^2}{\partial \theta^2} \ell_n\left(\hat{\theta}\right) = -\frac{n\bar{k}_n}{\hat{\theta}^2} < 0$$

as required.

Exercise 10.6 Let X be Bernoulli $\pi(X \mid p) = p^x (1-p)^{1-x}$.

(a) Calculate the information for p by taking the variance of the score.

(b) Calculate the information for p by taking the expectation of (minus) the second derivative. Did you obtain the same answer?

Solution. 10.6

(a) $\pi(X \mid p) = p^x (1-p)^{1-x}$. We know that $E[X] = E[X^2] = p_0$ and $var[X] = p_0(1-p_0)$. The log mass function is

$$\log \pi(x \mid p) = x \log p + (1-x)\log(1-p)$$

, with expectation

$$\ell(p) = E[\log \pi(x \mid p)] = p_0 \log p + (1-p_0)\log(1-p)$$

The derivative of the log mass function is

$$\frac{\partial}{\partial p}\log \pi(x \mid p) = \frac{x}{p} - \frac{1-x}{1-p}$$

The efficient score is the first derivative evaluated at X and p_0

$$S = \frac{\partial}{\partial p}\log \pi(X \mid p_0) = \frac{X}{p_0} - \frac{1-X}{1-p_0}$$

It has expectation

$$E[S] = \frac{E[X]}{p_0} - \frac{1-E[X]}{1-p_0} = 0$$

and variance of the score

$$\begin{aligned} var[S] &= var\left[\frac{X}{p_0} - \frac{1-X}{1-p_0}\right] \\ &= var\left[\frac{X-p_0}{p_0(1-p_0)}\right] \\ &= var\left[\frac{X}{p_0(1-p_0)}\right] \\ &= \frac{p_0(1-p_0)}{(p_0(1-p_0))^2} \\ &= \frac{1}{p_0(1-p_0)} \end{aligned}$$

the Fisher information is

$$\mathscr{I}_0 = E[SS'] = var[S] = \frac{1}{p_0(1-p_0)}$$

(b) The second derivative of the log mass function is

$$\frac{\partial^2}{\partial p^2} \log \pi(x \mid p) = -\frac{x}{p^2} - \frac{1-x}{(1-p)^2}$$

The expectation of (minus) the second derivative

$$\begin{aligned}
\mathscr{H}_0 &= -E\left[\frac{\partial^2}{\partial p^2} \log \pi(X \mid p_0)\right] \\
&= E\left[\frac{X}{p_0^2} + \frac{1-X}{(1-p_0)^2}\right] \\
&= \frac{E[X]}{p_0^2} + \frac{1-E[X]}{(1-p_0)^2} \\
&= \frac{p_0}{p_0^2} + \frac{1-p_0}{(1-p_0)^2} = \frac{1}{p_0(1-p_0)}
\end{aligned}$$

$\mathscr{I}_0 = \mathscr{H}_0$ and the information equality holds.

Exercise 10.10 Take the model $f(x) = \theta \exp(-\theta x)$, $x \geq 0$, $\theta > 0$.

(a) Find the Cramér-Rao lower bound for θ.

(b) Recall the MLE $\hat{\theta}$ for θ from above. Notice that this is a function of the sample mean. Use this formula and the delta method to find the asymptotic distribution for $\hat{\theta}$.

(c) Find the asymptotic distribution for $\hat{\theta}$ using the general formula for the asymptotic distribution of MLE. Do you find the same answer as in part (b)?

Solution. 10.10

(a) The log density is

$$\log f(x \mid \theta) = \log \theta - \theta x$$

The second derivative of the log density is

$$\frac{\partial^2}{\partial \theta^2} \log f(x \mid \theta) = -\frac{1}{\theta^2}$$

The expectation of (minus) the second derivative

$$\mathscr{H}_0 \quad = -E\left[\tfrac{\partial^2}{\partial \theta^2} \log f(X \mid \theta_0)\right] = \tfrac{1}{\theta_0^2}$$

The Fisher information is

$$\mathscr{I}_0 = \mathscr{H}_0 = \frac{1}{\theta_0^2}$$

The Cramér-Rao lower bound is

$$CRLB = (n\mathscr{I}_0)^{-1} = \frac{1}{n}\theta_0^2$$

(b) The log-likelihood function is

$$\ell_n(\theta) = \sum_{i=1}^{n} (\log\theta - \theta x) = n\log\theta - n\theta\bar{X}_n$$

The F.O.C. for $\hat{\theta}$ is

$$\frac{\partial}{\partial\theta}\ell_n\left(\hat{\theta}\right) = \frac{n}{\hat{\theta}} - n\bar{X}_n = 0$$

The solution is

$$\hat{\theta} = \frac{1}{\bar{X}_n}$$

The S.O.C. is

$$\frac{\partial^2}{\partial\theta^2}\ell_n\left(\hat{\theta}\right) = -\frac{n}{\hat{\theta}^2} < 0$$

as required.

Therefore, the MLE is

$$\hat{\theta} = \frac{1}{\bar{X}_n}$$

The derivative of the log density is

$$\frac{\partial}{\partial\theta}\log f(x \mid \theta) = \frac{1}{\theta} - x$$

The efficient score is the first derivative evaluated at X and θ_0

$$S = \frac{\partial}{\partial\theta}\log f(x \mid \theta_0) = \frac{1}{\theta_0} - X$$

It has expectation

$$E[S] = \frac{1}{\theta_0} - E[X] = 0$$

and variance of the score

$$var[S] \quad = var\left[\tfrac{1}{\theta_0} - X\right] = var[X]$$

Then we can get

$$E\left[X\right] = \frac{1}{\theta_0}$$

$$var\left[X\right] = var\left[S\right] = \mathscr{I}_0 = \frac{1}{\theta^2}$$

So,

$$\sqrt{n}\left(\bar{X}_n - \frac{1}{\theta_0}\right) \xrightarrow{d} N\left(0, \frac{1}{\theta^2}\right)$$

Take $h(u) = \frac{1}{u}$, so that by the Delta Method

$$\sqrt{n}\left(\hat{\theta} - \theta_0\right) = \sqrt{n}\left(\frac{1}{\bar{X}_n} - \frac{1}{\frac{1}{\theta_0}}\right) \xrightarrow{d} N\left(0, \left(-\frac{1}{\left(\frac{1}{\theta_0}\right)^2}\right)^2 var\left[X\right]\right) = N\left(0, \theta_0^2\right)$$

(c)

$$\sqrt{n}\left(\widehat{\theta} - \theta_0\right) \xrightarrow{d} N\left(0, \mathscr{I}_0^{-1}\right) = N\left(0, \theta_0^2\right)$$