MAST20005/MAST90058: Week 12 Solutions

1. (a) Working with just a single observation to start with. The score function is,

$$\frac{\partial}{\partial p}\log p^x(1-p)^{1-x} = \frac{x}{p} - \frac{1-x}{1-p}.$$

Differentiating once more gives,

$$\frac{\partial}{\partial p}\left(\frac{x}{p} - \frac{1-x}{1-p}\right) = -\frac{x}{p^2} - \frac{1-x}{(1-p)^2}.$$

Recall that $\mathbb{E}(X) = p$. Therefore the Fisher information is,

$$-\mathbb{E}\left[-\frac{X}{p^2} - \frac{1-X}{(1-p)^2}\right] = \frac{p}{p^2} + \frac{1-p}{(1-p)^2} = \frac{1}{p(1-p)}.$$

Thus the Cramér–Rao lower bound, using the full sample, is $p(1-p)/n$.

(b) Since $\operatorname{var}(\bar{X}) = n^{-2}(\operatorname{var}(X_1) + \cdots + \operatorname{var}(X_n)) = p(1-p)/n$, the sample proportion $\bar{X}$ attains the Cramér–Rao lower bound.

2. (a) The score function is,

$$\frac{\partial}{\partial \theta}\log\left\{\left(\frac{1}{\sqrt{2\pi\theta}}\right)^n e^{-\sum_i (x_i-\mu)^2/(2\theta)}\right\} = -\frac{n}{2\theta} + \sum_i \frac{(x_i-\mu)^2}{2\theta^2}.$$

Setting this equal to 0 and solving for θ gives the result.

(b) Working with just a single observation to start with,

$$\frac{\partial^2}{\partial \theta^2}\log\left\{\left(\frac{1}{\sqrt{2\pi\theta}}\right) e^{-(x-\mu)^2/(2\theta)}\right\} = \frac{\partial}{\partial \theta}\left\{-\frac{1}{2\theta} + \frac{(x-\mu)^2}{2\theta^2}\right\} = \frac{1}{2\theta^2} - \frac{(x-\mu)^2}{\theta^3}.$$

Since $\mathbb{E}[(X-\mu)^2] = \theta$, the Fisher information is,

$$-\mathbb{E}\left[\frac{1}{2\theta^2} - \frac{(X-\mu)^2}{\theta^3}\right] = -\frac{1}{2\theta^2} + \frac{1}{\theta^2} = \frac{1}{2\theta^2},$$

and the Cramér–Rao lower bound, using the full sample, is $2\theta^2/n$.

(c) $\mathrm{N}(\theta, 2\theta^2/n)$

(d) Note that $(X_i-\mu)^2/\theta \sim \chi^2_1$. Thus, $n\hat{\theta}/\theta = \sum_i (X_i-\mu)^2/\theta \sim \chi^2_n$.

3. (a) Writing out the likelihood,

$$\begin{aligned} f(x_1,\ldots,x_n \mid \theta) &\propto \theta^{-2n}\left(\prod_{i=1}^n x_i\right) e^{-\sum_{i=1}^n x_i/\theta} \\ &= \left(\prod_{i=1}^n x_i\right) e^{-y/\theta - 2n\log\theta} \end{aligned}$$

and the factorisation theorem yields that $Y = \sum_{i=1}^n X_i$ is sufficient for θ.

(b) The log-likelihood and score functions are, respectively:

$$\begin{aligned} \ell(\theta) &= -2n\log(\theta) - \frac{y}{\theta} + \text{const.} \\ s(\theta) &= \frac{\partial \ell}{\partial \theta} = -\frac{2n}{\theta} + \frac{y}{\theta^2} \end{aligned}$$

(c) Solving $s(\theta) = 0$ gives the MLE, $\hat{\theta} = \bar{X}/2$.

(d) Differentiating the score function and treating the data as random gives,

$$-\frac{\partial^2 \ell}{\partial \theta^2} = -\frac{2n}{\theta^2} + \frac{2Y}{\theta^3},$$

which has expected value,

$$\mathbb{E}\left(-\frac{\partial^2 \ell}{\partial \theta^2}\right) = -\frac{2n}{\theta^2} + \frac{2n(2\theta)}{\theta^3} = \frac{2n}{\theta^2}.$$

Hence the Cramér–Rao lower bound is $\theta^2/(2n)$.

(e) The maximum likelihood estimate is $\hat{\theta} = 10.5/2 = 5.25$ and an approximate 95% confidence interval is $5.25 \pm 1.96 \times \sqrt{5.25^2/70} = (4.02, 6.48)$.

For questions 4–6, only the answers are given below. For full details of the derivation see the video consultation *Understanding Sufficiency* on the LMS.

4. Let the first five tosses of the coin be $X_1, \ldots, X_5$ and the second five be $Y_1, \ldots, Y_5$. Then $T = \sum_{i=1}^{5} X_i - \sum_{i=1}^{5} Y_i$ is sufficient for p.

5. Suppose we observe a random sample of size n.

   (a) $X_{(n)}$ is sufficient for θ.

   (b) $X_{(1)}$ and $X_{(n)}$ are jointly sufficient for θ.

6. Suppose we observe a random sample of size n.

   (a) $\sum_i X_i$ is sufficient for θ.

   (b) $X_{(1)}$ is sufficient for θ.

   (c) $\sum_i X_i$ and $X_{(1)}$ are jointly sufficient for θ.

7. (a) Let the sample frequencies (counts) of the three possible observations be f_0, f_1, f_2. Note that $f_0 + f_1 + f_2 = n$. The likelihood function is, $L(\theta) = (1-\theta)^{f_0} (\frac{3}{4}\theta)^{f_1} (\frac{1}{4}\theta)^{f_2} = \left[3^{f_1}\right] \left[(1-\theta)^{f_0} (\frac{1}{4}\theta)^{n-f_0}\right]$. Therefore, by the factorisation theorem we see that f_0 is sufficient for θ. (Note that we referred to this statistic as Z in the week 3 problems.)

   (b) We want to use an estimator based on this statistic in order to most efficiently capture the relevant information. This explains why the esimator based on $\bar{X}$ was not optimal: it is trying to use information that distinguishes between observations of 1 and 2, which is irrelevant information.