Department of Mathematics and Statistics

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

# CSE Exercises – Week 8

① Revisiting Exercise 4(b) from last week:

(a) Find the ML estimate of $c$ analytically.

(b) Find the ML estimate of $c$ numerically using fminbnd.

(c) Explain why, in this case, the normal approximation should not be used to obtain a confidence interval for $c$? In spite of this, use the normal approximation to find a 95% confidence interval for $c$ anyway. Comment on your result.

(d) We can use Monte Carlo simulation to get an approximate confidence interval for $c$:

(i) Use the ML estimate, $\hat{c}$, in the distribution, i.e. let $c = \hat{c}$ in $f(x; c)$.

(ii) Generate $X_1, X_2 \overset{IID}{\sim} f(x; \hat{c})$. Recall from exercise 3 in week 3 that we can do this by the inverse CDF method.

(iii) Find the ML estimate of $c$ for the values of $X_1$ and $X_2$ from step (ii), i.e.

find $\hat{c}_1 = \arg\max_{c>0} f(X_1, X_2 ; c)$ .

(iv) Repeat steps (ii) and (iii) $N$ times to get $\hat{c}_1, \hat{c}_2, \ldots, \hat{c}_N$ .

(v) Sort the estimates from (iv) in increasing order to get the order statistics, $\hat{c}_{(1)}, \hat{c}_{(2)}, \ldots, \hat{c}_{(N)}$ .

(vi) An approximate 95% confidence interval is given by $(\hat{c}_{(0.025N)}, \hat{c}_{(0.975N)})$.

Use this procedure with $N = 1000$ to find an approximate 95% confidence interval for $c$.

Later on, you will learn about a method for constructing confidence intervals known as the bootstrap. The procedure that we have used here is sometimes called a "parametric bootstrap". It is worthwhile keeping in mind that the bootstrap is simulation-based.

(e) Next, we want to check that when the sample size $n$ is "large enough", the normal approximation confidence interval and the parametric bootstrap confidence interval are about the same.

Suppose that $n = 30$, $X_1, \ldots, X_n \overset{\text{IID}}{\sim} f(x ; c)$, and

$$\sum_{i=1}^{n} \log X_i = -7.7335 .$$

(i) Find the ML estimate of $c$.

(ii) Find a 95% confidence interval for $c$ using the normal approximation for $\hat{c}$.

(iii) Find a 95% confidence interval for $c$ using the "parametric bootstrap" procedure.

(2) Given that $X_1, \ldots, X_n \overset{IID}{\sim} f(x;\theta)$, where

$$f(x;\theta) = \frac{1}{\theta} \exp\left(-\frac{x}{\theta}\right) \exp\left[-\exp\left(-\frac{x}{\theta}\right)\right],$$

for $x \in \mathbb{R}$ and $\theta > 0$. Now suppose that $n = 5$ and that $X_1 = -0.15$, $X_2 = 0.27$, $X_3 = 1.33$, $X_4 = -1.71$ and $X_5 = -0.89$.

(a) Find the ML estimate of $\theta$ using fminbnd.

(b) Use the "parametric bootstrap" procedure with $N = 1000$ to find an approximate 95% confidence interval for $\theta$.

(c) Let $\sigma^2$ denote the variance of $X \sim f(x;\theta)$. Given that $\sigma^2 = \theta^2 \pi^2 / 6$, find the ML estimate of $\sigma^2$.

(d) By noting that $\sigma^2$ is a one-to-one monotone increasing function of $\theta$ for $\theta > 0$, find an approximate 95% confidence interval for $\sigma^2$.

(3) Given that $X_1, \ldots, X_n \overset{\text{IID}}{\sim} f(x;\theta)$, where

$$f(x;\theta) = \frac{1}{4\theta}\,\text{sech}^2\left(\frac{x}{2\theta}\right),$$

for $x \in \mathbb{R}$ and $\theta > 0$. Now suppose that $n = 6$ and $X_1 = 0{\cdot}32$, $X_2 = 1{\cdot}56$, $X_3 = 3{\cdot}11$, $X_4 = -0{\cdot}01$, $X_5 = -2{\cdot}27$ and $X_6 = -3{\cdot}41$.

(a) Find the ML estimate of $\theta$ using fminbnd.

(b) Use the "parametric bootstrap" procedure to find an approximate 95% confidence interval for $\theta$.

(c) Let $\lambda = 1/\sqrt{\theta}$. Find the ML estimate of $\lambda$.

(d) Find an approximate 95% confidence interval for $\lambda$.

Solutions

(1) From week 7, exercise 4(b),

$$l_2(c) = 2\log c + (c-1)(\log x_1 + \log x_2), \quad c > 0.$$

(a) $$\frac{\partial l_2}{\partial c} = \frac{2}{c} + \log x_1 + \log x_2 .$$

Equating to 0 and solving for c, the ML estimator for c is

$$\hat{c} = -2/(\log x_1 + \log x_2) .$$

When $X_1 = 0.7$ and $X_2 = 0.9$,

$$\hat{c} = -2/(\log 0.7 + \log 0.9)$$

$$= 4.3287 .$$

(b) Using fminbnd, $\hat{c} = 4.3287$ .

(c) In this case, the sample size of $n = 2$ is too small for the normal approximation to be valid for $\hat{c}$ .

Fisher information,

$$I_1(c) = -E_c\left(\frac{\partial^2 l(c)}{\partial c^2}\right)$$

$$= -E_c\left[\frac{\partial}{\partial c}\left(\frac{1}{c} + \log x\right)\right]$$

$$= -E_c\left(-\frac{1}{c^2}\right) = \frac{1}{c^2}.$$

$$I_2(c) = 2\,I_1(c) = \frac{2}{c^2}$$

$$se_2 \approx \sqrt{1/I_2(c)} = c/\sqrt{2}$$

$$\hat{se}_2 = \hat{c}/\sqrt{2}$$

$\therefore$ an approximate 95% CI for $c$ is

$$\hat{c} \pm 1.96\,\hat{se}_2 = \hat{c} \pm 1.96\,\hat{c}/\sqrt{2}$$

$$= 4.3287 \pm \frac{1.96 \times 4.3287}{\sqrt{2}}$$

$$= (-1.67,\ 10.33).$$

Clearly, there is a problem with this confidence interval because the left-hand limit is $-1.6706$, which violates the fact that $c > 0$.

(d) Using the procedure with $N = 1000$, an approximate 95% confidence interval for $c$ is

$$(\hat{c}_{(25)},\ \hat{c}_{(975)}) = (1.57,\ 38.74).$$

Since this is a simulation-based procedure, your answers will differ slightly from mine.

(e) With $n = 30$, $X_1, \ldots, X_n \overset{\text{IID}}{\sim} f(x; c)$

and $\sum_{i=1}^{n} \log X_i = -7.7335$,

(i) ML estimate of $c$ is

$$\hat{c} = -n \Big/ \sum_{i=1}^{n} \log X_i$$

$$= -30 / -7.7335 = 3.8792.$$

(ii) Normal approximation 95% CI is

$$\hat{c} \pm 1.96\, \hat{c}/\sqrt{n}$$

$$= 3.8792 \pm \frac{1.96 \times 3.8792}{\sqrt{30}}$$

$$= (2.49, \ 5.28).$$

(iii) Parametric bootstrap 95% CI with $N = 1000$ is

$$(\hat{c}_{(25)}, \ \hat{c}_{(975)}) = (2.80, \ 5.80).$$

(2) Given that $X_1, \ldots, X_n \overset{\text{IID}}{\sim} f(x;\theta)$, where

$$f(x;\theta) = \frac{1}{\theta}\exp\left(-\frac{x}{\theta}\right)\exp\left[-\exp\left(-\frac{x}{\theta}\right)\right],$$

for $x \in \mathbb{R}$ and $\theta > 0$.

(a)
$$L_n(\theta) = \prod_{i=1}^{n} f(x_i;\theta)$$

$$\begin{aligned}
\ell_n(\theta) &= \log L_n(\theta) \\
&= \sum_{i=1}^{n} \log f(x_i;\theta) \\
&= \sum_{i=1}^{n}\left[-\log\theta - \frac{x_i}{\theta} - \exp\left(-\frac{x_i}{\theta}\right)\right] \\
&= -n\log\theta - \frac{1}{\theta}\sum_{i=1}^{n} x_i - \sum_{i=1}^{n}\exp\left(-\frac{x_i}{\theta}\right)
\end{aligned}$$

$$\hat{\theta} = \underset{\theta>0}{\arg\max}\ \ell_n(\theta).$$

Using fminbnd, $\hat{\theta} = 1.2967$.

(b) Using the "parametric bootstrap" procedure with $N = 1000$, an approximate 95% confidence interval for $\theta$ is

$$(\hat{\theta}_{(25)}, \hat{\theta}_{(975)}) = (0.4836, 2.1455).$$

(c) By the equivariant property of the ML estimator, the ML estimate of $\sigma^2$ is

$$\begin{aligned}
\hat{\sigma}^2 &= \hat{\theta}^2\pi^2/6 \\
&= (1.2967)^2\pi^2/6 = 2.7658.
\end{aligned}$$

(d) Since $\sigma^2$ is a one-to-one monotone increasing function of $\theta$ for $\theta > 0$, an approximate 95% confidence interval for $\sigma^2$ can be obtained by applying the function to the end-points of the confidence interval for $\theta$. Using the result from part (b), the required confidence interval for $\sigma^2$ is

$$\left( \hat{\theta}^2_{(25)} \pi^2/6 \,,\; \hat{\theta}^2_{(975)} \pi^2/6 \right)$$

$$= \left( 0.4836^2 \pi^2/6 \,,\; 2.1455^2 \pi^2/6 \right)$$

$$= ( 0.3848 \,,\; 7.5721 ) .$$

(3) Given that $X_1, \ldots, X_n \overset{\text{IID}}{\sim} f(x;\theta)$, where

$$f(x;\theta) = \frac{1}{4\theta}\,\text{sech}^2\left(\frac{x}{2\theta}\right),$$

for $x \in \mathbb{R}$ and $\theta > 0$.

(a)
$$L_n(\theta) = \prod_{i=1}^{n} f(x_i;\theta)$$
$$\begin{aligned} \ell_n(\theta) &= \log L_n(\theta) = \sum_{i=1}^{n} \log f(x_i;\theta) \\ &= \sum_{i=1}^{n}\left[-\log(4\theta) + \log\left(\text{sech}^2\left(\frac{x_i}{2\theta}\right)\right)\right] \\ &= -n\log(4\theta) + \sum_{i=1}^{n}\log\left(\text{sech}^2\left(\frac{x_i}{2\theta}\right)\right). \end{aligned}$$

Using fminbnd, $\hat{\theta} = 1.3232$.

(b) Using the "parametric bootstrap" procedure with $N = 1000$, an approximate 95% confidence interval for $\theta$ is

$$(\hat{\theta}_{(25)}, \hat{\theta}_{(975)}) = (0.5710, 2.2490).$$

(c) By the equivariant property, ML estimate of $\lambda$ is

$$\hat{\lambda} = 1/\sqrt{\hat{\theta}} = 1/\sqrt{1.3232} = 0.8693.$$

(d) Notice that $\lambda$ is a one-to-one monotone decreasing function of $\theta$ for $\theta > 0$.

Therefore, an approximate 95% confidence interval for $\lambda$ is

$$\left( 1/\sqrt{\hat{\theta}_{(975)}} \;,\; 1/\sqrt{\hat{\theta}_{(25)}} \right)$$

note the reversal
because the function is decreasing

$$= \left( 1/\sqrt{2.2490} \;,\; 1/\sqrt{0.5710} \right)$$

$$= (0.6668,\ 1.3233).$$