Assignment 2

Tuesday, 6 September 2022 9:34 AM

a.

$$f(x_1,\dots x_n;\theta) = b(x_1,\dots,x_n)\exp(c(\theta)T(x_1,\dots,x_n))/a(\theta)$$

i. First, take the logarithm of the above p.d.f. to obtain the log likelihood

$$\log L(\theta) = \log(b(x_1,\dots x_n)) + c(\theta)T(x_1,\dots,x_n) - \log(a(\theta))$$

Now, differentiate with respect to θ,

$$\frac{\partial \log L(\theta)}{\partial\theta} = \frac{\partial c(\theta)}{\partial\theta}T(x_1,\dots,x_n) - \frac{\partial \log(a(\theta))}{\partial\theta} \qquad (1)$$

Assuming regular conditions, the root of the above gives the ML estimate, so

$$0 = \frac{\partial c(\theta)}{\partial\theta}T(x_1,\dots,x_n) - \frac{\partial \log(a(\theta))}{\partial\theta} \qquad (2)$$

but, taking the expectation of (1) gives

$$E\left(\frac{\partial \log L(\theta)}{\partial\theta}\right) = 0$$

$$= \frac{\partial c(\theta)}{\partial\theta}E(T(X_1,\dots,X_n)) - \frac{\partial \log(a(\theta))}{\partial\theta}$$

Hence $\frac{\partial \log(a(\theta))}{\partial\theta} = \frac{\partial c(\theta)}{\partial\theta}E(T(X_1,\dots,X_n))$

and so equation (2) can be expressed as

$$0 = \frac{\partial c(\theta)}{\partial\theta}T(x_1,\dots,x_n) - \frac{\partial c(\theta)}{\partial\theta}E(T(X_1,\dots,X_n))$$

$$\Rightarrow \frac{\partial c(\theta)}{\partial\theta}\left[E(T(X_1,\dots,X_n)) - T(x_1,\dots x_n)\right] = 0$$

$$\Rightarrow E(T(X_1,\dots X_n)) - T(x_1,\dots,x_n) = 0$$

and finally

$$E(T(X_1,\dots,X_n)) = T(x_1,\dots,x_n)$$

But since we've taken this at the MLE, $\hat\theta$,

$$\left[E(T(X_1,\dots,X_n))\right]_{\theta=\hat\theta} = T(x_1,\dots,x_n)$$

ii. For a 1-parameter exponential distribution,

$$I(\theta) = -\frac{\partial^2 \log L(\theta)}{\partial\theta^2}$$

where

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

$$= f(x_1,\dots,x_n;\theta)$$

$$= b(x_1,\dots,x_n)\frac{\exp(c(\theta)T(x_1,\dots,x_n))}{a(\theta)}$$

$$\Rightarrow \log L(\theta) = \log(b(x_1,\dots,x_n)) + \log\left(\frac{\exp(c(\theta)T(x_1,\dots,x_n))}{a(\theta)}\right)$$

$$= \log(b(x_1,\dots x_n)) + c(\theta)T(x_1,\dots x_n) - \log(a(\theta))$$

$$\Rightarrow \frac{\partial \log L(\theta)}{\partial\theta} = \frac{\partial c(\theta)}{\partial\theta}T(x_1,\dots,x_n) - \frac{\partial \log(a(\theta))}{\partial\theta}$$

$$\Rightarrow \frac{\partial \log L(\theta)}{\partial \theta} = \frac{\partial c(\theta)}{\partial \theta} T(x_1, \dots, x_n) - \frac{\partial \log(a(\theta))}{\partial \theta}$$

$$\Rightarrow \frac{\partial^2 \log L(\theta)}{\partial \theta^2} = \frac{\partial^2 c(\theta)}{\partial \theta^2} T(x_1, \dots, x_n) - \frac{\partial^2 \log(a(\theta))}{\partial \theta^2}$$

Hence

$$I(\theta) = \frac{\partial^2 \log(a(\theta))}{\partial \theta^2} - \frac{\partial^2 c(\theta)}{\partial \theta^2} T(x_1, \dots, x_n)$$

Finally, the Fisher information is given by

$$\mathcal{I}(\theta) = \mathbb{E}(I(\theta))$$

Since we're evaluating this at the ML estimate, we have

$$\begin{aligned}
\mathcal{I}(\hat{\theta}) &= \mathbb{E}(I(\hat{\theta})) \\
&= \mathbb{E}\left(\frac{\partial^2 \log(a(\hat{\theta}))}{\partial \hat{\theta}^2} - \frac{\partial^2 c(\hat{\theta})}{\partial \hat{\theta}^2} T(x_1, \dots x_n)\right) \\
&= \frac{\partial^2 \log(a(\hat{\theta}))}{\partial \hat{\theta}^2} - \mathbb{E}\left(\frac{\partial^2 c(\hat{\theta})}{\partial \hat{\theta}^2} T(x_1, \dots, x_n)\right) \\
&= \frac{\partial^2 \log(a(\hat{\theta}))}{\partial \hat{\theta}^2} - \frac{\partial^2 c(\hat{\theta})}{\partial \hat{\theta}^2} \mathbb{E}(T(x_1, \dots, x_n))
\end{aligned}$$

It was shown in part i) that, at the MLE,

$$\left[\mathbb{E}(T(x_1, \dots x_n))\right]_{\theta = \hat{\theta}} = T(x_1, \dots, x_n)$$

and so,

$$\begin{aligned}
\mathcal{I}(\hat{\theta}) &= \frac{\partial^2 \log(a(\hat{\theta}))}{\partial \hat{\theta}^2} - \frac{\partial^2 c(\hat{\theta})}{\partial \hat{\theta}^2} T(x_1, \dots, x_n) \\
&= I(\hat{\theta})
\end{aligned}$$

QED

b. i. We have a truncated Poisson distribution (at 0), with

$$f(x; \theta) = \frac{e^{-\theta} \theta^x}{A(\theta) x!}$$

and $A(\theta) = 1 - e^{-\theta}$

If the sample mean is an unbiased estimator of θ, we'd expect

$$\mathbb{E}(\bar{X}) = \theta \quad \text{or} \quad \mathbb{E}(\bar{X}) - \theta = \text{bias}(\bar{X})$$

where the bias is the deviation of $\mathbb{E}(\bar{X})$ from θ.

$$\begin{aligned}
\mathbb{E}(\bar{X}) &= \mathbb{E}\left(\frac{1}{n} \sum_{i=1}^{n} X_i\right) \\
&= \frac{1}{n} \mathbb{E}\left(\sum_{i=1}^{n} X_i\right) \\
&= \frac{1}{n}\left(n \cdot \mathbb{E}(X)\right)
\end{aligned}$$

Now to compute $\mathbb{E}(X)$:

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{x=1}^{\infty} x \frac{\theta^x e^{-\theta}}{(1-e^{-\theta}) x!} \\
&= \frac{1}{(1-e^{-\theta})} \sum_{x=1}^{\infty} x \frac{\theta^x e^{-\theta}}{x!} \\
&= \frac{1}{(1-e^{-\theta})}\left(-0 \cdot \frac{\theta^0 e^{-\theta}}{0!} + \sum_{x=0}^{\infty} x \frac{\theta^x e^{-\theta}}{x!}\right) \\
&= \frac{\theta}{(1-e^{-\theta})}
\end{aligned}$$

where the last step involved the expectation of the poisson distribution: $\sum_{x=0}^{\infty} x \frac{\theta^x e^{-\theta}}{x!} = \theta$

So, $$\mathbb{E}(\bar{X}) = \frac{1}{n}\left(n \cdot \mathbb{E}(X)\right)$$
$$= \frac{\theta}{1-e^{-\theta}}$$

$$\Rightarrow \quad bias(\bar{x}) = \mathbb{E}(\bar{x}) - \theta$$
$$= \frac{\theta}{1-e^{-\theta}} - \theta$$
$$= \frac{\theta - \theta(1-e^{-\theta})}{(1-e^{-\theta})}$$
$$= \frac{\theta e^{-\theta}}{(1-e^{-\theta})} = \frac{\theta}{e^{\theta}-1}$$

Therefore $\mathbb{E}(\bar{x})$ is a biased estimator of θ, with bias $\theta/(e^{\theta}-1)$

ii. The likelihood function of the truncated Poisson dist. is

$$\mathcal{L}(\theta) = \prod_{x=1}^{n} \frac{e^{-\theta}\theta^x}{(1-e^{-\theta})x!}$$
$$= b(x_1,\dots,x_n)\cdot\left(\frac{e^{-\theta}}{1-e^{-\theta}}\right)^n \cdot e^{\log(\theta)\cdot\sum_{x=1}^{n}(x)}$$
$$= b(x_1,\dots,x_n)\cdot\frac{\exp\left(c(\theta)T(x_1,\dots,x_n)\right)}{a(\theta)}$$

which is exactly of the form of the regular exponential family, with

$$b(x_1,\dots,x_n) = \left(\prod_{x=1}^{n} x!\right)^{-1}$$
$$c(\theta) = \log(\theta)$$
$$T(x_1,\dots,x_n) = \sum_{x=1}^{n} x$$
$$a(\theta) = \left(\frac{1-e^{-\theta}}{e^{-\theta}}\right)^n = (e^{\theta}-1)^n$$

iii. To start with, find the log likelihood function and differentiate w.r.t. θ. Since $f(x;\theta)$ is in the exp. family,

$$\frac{\partial \log \mathcal{L}(\theta)}{\partial\theta} = \frac{\partial c(\theta)}{\partial\theta}T(x_1,\dots,x_n) - \frac{\partial \log(a(\theta))}{\partial\theta}$$

as found in part a). We have

$$c(\theta) = \log(\theta) \Rightarrow \frac{\partial c(\theta)}{\partial\theta} = \frac{1}{\theta}$$
$$a(\theta) = (e^{\theta}-1)^n \Rightarrow \log(a(\theta)) = n\log(e^{\theta}-1)$$
$$\Rightarrow \frac{\partial \log(a(\theta))}{\partial\theta} = \frac{ne^{\theta}}{e^{\theta}-1}$$

And so
$$\frac{\partial \log \mathcal{L}(\theta)}{\partial\theta} = \frac{\sum_{x=1}^{n} x}{\theta} - \frac{ne^{\theta}}{e^{\theta}-1} \qquad (3)$$

The expectation of the score statistic is 0, hence

$$\mathbb{E}\left(\frac{\sum_{x=1}^{n} x}{\hat{\theta}} - \frac{ne^{\hat{\theta}}}{e^{\hat{\theta}}-1}\right) = 0$$

$$\mathbb{E}\left(\sum_{i} x\right) \quad ne^{\hat{\theta}}$$

$$\mathbb{E}\left(\frac{\sum_{x=1}^{n} x}{\hat{\theta}} - \frac{ne^{\hat{\theta}}}{e^{\hat{\theta}}-1}\right) = 0$$

$$\mathbb{E}\left(\frac{\sum_{x=1}^{n} x}{\hat{\theta}}\right) = \frac{ne^{\hat{\theta}}}{e^{\hat{\theta}}-1}$$

$$\frac{1}{\hat{\theta}}\,\mathbb{E}\left(\sum_{x=1}^{n} x\right) = \frac{ne^{\hat{\theta}}}{e^{\hat{\theta}}-1}$$

And so substituting this into equation (3) for the MLE gives

$$0 = \frac{\sum_{x=1}^{n} x}{\hat{\theta}} - \frac{\mathbb{E}\left(\sum_{x=1}^{n} x\right)}{\hat{\theta}}$$

$$\Rightarrow \left[\mathbb{E}\left(\sum_{x=1}^{n} x\right)\right]_{\theta=\hat{\theta}} = \sum_{x=1}^{n} x$$

which agrees with that proposed in part a)

To show the result in part aii, we must differentiate the log likelihood once more:

$$\frac{\partial^2 \log \mathcal{L}(\theta)}{\partial \theta^2} = \frac{\partial}{\partial \theta}\left(\frac{\partial \log \mathcal{L}(\theta)}{\partial \theta}\right)$$

$$= \frac{\partial}{\partial \theta}\left(\frac{\partial c(\theta)}{\partial \theta} T(x_1, \ldots, x_n) - \frac{\partial \log(a(\theta))}{\partial \theta}\right)$$

$$= \frac{\partial^2 c(\theta)}{\partial \theta^2} T(x_1, \ldots, x_n) - \frac{\partial^2 \log(a(\theta))}{\partial \theta^2}$$

where

$$\frac{\partial^2 c(\theta)}{\partial \theta} = \frac{\partial}{\partial \theta}\left(\frac{\partial c(\theta)}{\partial \theta}\right) = \frac{\partial}{\partial \theta}\left(\frac{1}{\theta}\right) = -\frac{1}{\theta^2}$$

$$\frac{\partial^2 \log(a(\theta))}{\partial \theta^2} = \frac{\partial}{\partial \theta}\left(\frac{\partial \log(a(\theta))}{\partial \theta}\right) = \frac{\partial}{\partial \theta}\left(\frac{ne^{\theta}}{e^{\theta}-1}\right) = -\frac{ne^{\theta}}{(e^{\theta}-1)^2}$$

And so

$$\frac{\partial^2 \log \mathcal{L}(\theta)}{\partial \theta^2} = \frac{ne^{\theta}}{(e^{\theta}-1)^2} - \frac{\sum_{x=1}^{n} x}{\theta^2}$$

As in part aii:

$$I(\theta) = -\frac{\partial^2 \log \mathcal{L}(\theta)}{\partial \theta^2} = \frac{\sum_{x=1}^{n} x}{\theta^2} - \frac{ne^{\theta}}{(e^{\theta}-1)^2}$$

$$\Rightarrow \mathbb{E}(I(\theta)) = \mathbb{E}\left(\frac{\sum_{x=1}^{n} x}{\theta^2} - \frac{ne^{\theta}}{(e^{\theta}-1)^2}\right)$$

$$= \mathbb{E}\left(\frac{1}{\theta^2}\sum_{x=1}^{n} x\right) - \frac{ne^{\theta}}{(e^{\theta}-1)^2}$$

$$= \frac{1}{\theta^2}\,\mathbb{E}\left(\sum_{x=1}^{n} x\right) - \frac{ne^{\theta}}{(e^{\theta}-1)^2}$$

but at the MLE, $\left[\mathbb{E}\left(\sum_{x=1}^{n} x\right)\right]_{\theta=\hat{\theta}} = \sum_{x=1}^{n} x$, and so

$$\mathbb{E}(I(\hat{\theta})) = I(\hat{\theta})$$

but $\mathcal{I}(\theta) = \mathbb{E}(I(\theta))$, and so, at the MLE,

$$\mathcal{I}(\hat{\theta}) = I(\hat{\theta})$$

$$= \frac{1}{\hat{\theta}^2}\sum_{x=1}^{n} x - \frac{ne^{\hat{\theta}}}{(e^{\hat{\theta}}-1)^2}$$

as in part a)ii.