STATS 2107
Statistical Modelling and Inference II
Tutorial 5
Solutions

Sharon Lee, Matt Ryan

Semester 2 2022

1. Suppose $y_1, y_2, \ldots, y_n$ are independent $Po(\lambda_i)$ observations with

$$\lambda_i = \theta x_i$$

where $\theta > 0$ is the unknown parameter and $x_1,\ x_2,\ \ldots,\ x_n$ are given positive constants.

(a) Find the log-likelihood, $\ell(\theta; \boldsymbol{y})$, and the score function, $S(\theta; \boldsymbol{y})$.

Solutions:

$$L(\lambda_i, \boldsymbol{y}) = \prod_{i=1}^{n} \frac{e^{-\lambda_i} \lambda_i^{y_i}}{y_i!}$$

Substitute in $\lambda_i = \theta x_i$ leads to

$$L(\theta, \boldsymbol{y}) = \prod_{i=1}^{n} \frac{e^{-\theta x_i} (\theta x_i)^{y_i}}{y_i!} = e^{-\theta \sum\limits_{i=1}^{n} x_i} \theta^{\sum\limits_{i=1}^{n} y_i} \prod_{i=1}^{n} \frac{x_i^{y_i}}{y_i!}$$

$$\Rightarrow \ell(\theta, \boldsymbol{y}) = -\theta \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i \log(\theta) + \log\left(\prod_{i=1}^{n} \frac{x_i^{y_i}}{y_i!}\right).$$

$$S(\theta; \boldsymbol{y_i}) = \frac{\partial \ell}{\partial \theta} = -\sum_{i=1}^{n} x_i + \frac{\sum\limits_{i=1}^{n} y_i}{\theta}.$$

(b) Find the maximum likelihood estimate, $\hat{\theta}$, and the Fisher information, I_θ.

Solutions:

To get the MLE, set the score function to zero and solve:

$$-\sum_{i=1}^{n} x_i + \frac{\sum_{i=1}^{n} y_i}{\hat{\theta}} = 0$$

$$\Rightarrow \hat{\theta} = \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i} = \frac{\bar{y}}{\bar{x}}.$$

Fisher information

$$\begin{aligned}
I_\theta &= -E\left[\frac{\partial^2 \ell}{\partial \theta^2}\right] \\
&= -E\left[\frac{\partial}{\partial \theta}\left(-\sum_{i=1}^{n} x_i + \frac{\sum_{i=1}^{n} Y_i}{\theta}\right)\right] \\
&= E\left[\frac{\sum_{i=1}^{n} Y_i}{\theta^2}\right] && \text{by linearity} \\
&= \frac{\sum_{i=1}^{n} E[Y_i]}{\theta^2} \\
&= \frac{\sum_{i=1}^{n} \theta x_i}{\theta^2} \\
&= \frac{\sum_{i=1}^{n} x_i}{\theta}.
\end{aligned}$$

---

2. Consider a **single** binomial observation y from $\text{Bin}(n, \theta)$ where the number of trials is n and the probability of success is θ. Assume n is known.
   (a) Give the log-likelihood $\ell(\theta; y)$.

Solutions:

$$\begin{aligned}
\ell(\theta; y) &= \log\left[\binom{n}{y}\theta^y(1-\theta)^{n-y}\right] \\
&= \log\binom{n}{y} + \log\theta^y + \log(1-\theta)^{n-y} \\
&= y\log\theta + (n-y)\log(1-\theta) + \log\binom{n}{y}
\end{aligned}$$

(b) Find the Score function and the Fisher information about θ.

Solutions:

The score function is

$$\begin{aligned} S(\theta; y) &= \frac{\partial \ell}{\partial \theta} \\ &= \frac{y}{\theta} - \frac{n-y}{1-\theta}. \end{aligned}$$

The second derivative of the log likelihood:

$$\frac{\partial^2 \ell}{\partial \theta^2} = -\frac{y}{\theta^2} - \frac{n-y}{(1-\theta)^2}.$$

Hence, the Fisher information is

$$\begin{aligned} I_\theta &= -E\left(\frac{\partial^2 \ell}{\partial \theta^2}\right) \\ &= -E\left(-\frac{Y}{\theta^2} - \frac{n-Y}{(1-\theta)^2}\right) \\ &= \frac{1}{\theta^2}E(Y) + \frac{1}{(1-\theta)^2}(n - E(Y)) \\ &= \frac{n\theta}{\theta^2} + \frac{n-n\theta}{(1-\theta)^2} && \text{since } E(Y) = n\theta \\ &= \frac{n}{\theta} + \frac{n}{1-\theta} = \frac{n}{\theta(1-\theta)}. \end{aligned}$$

(c) Find the MLE $\hat{\theta}$.

Solutions:
Set the Score function equal to zero and solve to find the MLE:

$$\begin{aligned} S(\theta; x) &= \frac{\partial \ell}{\partial \theta} = 0 \\ &\Rightarrow \frac{y}{\theta} - \frac{n-y}{1-\theta} = 0 \\ &\Rightarrow \frac{y - n\theta}{\theta(1-\theta)} = 0 \\ &\Rightarrow \theta = y/n \\ &\Rightarrow \hat{\theta} = y/n. \end{aligned}$$

(d) Find expressions for the score test statistic, U, and the log-likelihood ratio test statistic, G^2, for testing the null hypothesis $H_0 : \theta = \theta_0$ versus $H_a : \theta \neq \theta_0$.

Solutions:
The score statistic is

$$U = \frac{S(\theta_0; Y)}{\sqrt{I_{\theta_0}}} = \frac{\frac{Y-n\theta_0}{\theta_0(1-\theta_0)}}{\sqrt{\frac{n}{\theta_0(1-\theta_0)}}} = \frac{Y - n\theta_0}{\sqrt{n\theta_0(1-\theta_0)}}$$

$$\left[= \frac{Y/n - \theta_0}{\sqrt{\theta_0(1-\theta_0)/n}} = \frac{\hat{\theta} - \theta_0}{\sqrt{\theta_0(1-\theta_0)/n}}\right].$$

The LRT statistic is

$$\begin{aligned}
G^2 &= -2\left(\ell(\theta_0; Y) - \ell(\hat{\theta}; Y)\right) \\
&= -2\left(Y\log\theta_0 + (n-Y)\log(1-\theta_0) - Y\log\hat{\theta} - (n-Y)\log(1-\hat{\theta})\right) \\
&= -2\left(Y\log\theta_0 + (n-Y)\log(1-\theta_0) - Y\log\frac{Y}{n} - (n-Y)\log\left(1-\frac{Y}{n}\right)\right) \\
&= -2\left(Y\log\frac{n\theta_0}{Y} + (n-Y)\log\frac{n(1-\theta_0)}{n-Y}\right) \\
&\left[= -2\left(Y\log\frac{\theta_0}{\hat{\theta}} + (n-Y)\log\frac{1-\theta_0}{1-\hat{\theta}}\right)\right].
\end{aligned}$$

(e) State the asymptotic distributions of U and G^2, respectively, under H_0.

Solutions:
Asymptotically, the LRT test statistic $G^2 = Z^2 \sim \chi_1^2$; the score test statistic is the familiar $Z \sim N(0,1)$.

---

3. Consider the simple linear regression model with no intercept, that is,

$$Y_i = \beta x_i + \epsilon_i, \quad \epsilon_i \overset{i.i.d.}{\sim} N(0, \sigma^2), \quad i = 1, 2, \ldots, n,$$

where $\boldsymbol{\theta} = (\beta, \sigma^2)$ are the unknown parameters.
(a) Write down the log-likelihood $\ell(\boldsymbol{\theta}; \boldsymbol{y})$

Solutions:
We have $Y_i \sim N(\beta x_i, \sigma^2)$ independently.

$$\begin{aligned}
L(\beta, \sigma^2; \boldsymbol{y}) &= \prod_{i=1}^{n} (2\pi)^{-\frac{1}{2}} \left(\sigma^2\right)^{-\frac{1}{2}} e^{-\dfrac{1}{2\sigma^2}(y_i - \beta x_i)^2} \\
&= (2\pi)^{-\frac{n}{2}} \left(\sigma^2\right)^{-\frac{n}{2}} e^{-\dfrac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \beta x_i)^2} \\
\ell(\beta, \sigma^2; \boldsymbol{y}) &= -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\left(\sigma^2\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \beta x_i)^2
\end{aligned}$$

(b) Find the score vector $S(\boldsymbol{\theta}; \boldsymbol{y})$.

Solutions:

$$S(\boldsymbol{\theta}; \boldsymbol{y}) = \begin{bmatrix} S(\beta; \boldsymbol{y}) \\ S(\sigma^2; \boldsymbol{y}) \end{bmatrix} = \begin{bmatrix} \dfrac{\partial \ell}{\partial \beta} \\ \dfrac{\partial \ell}{\partial \sigma^2} \end{bmatrix} = \begin{bmatrix} \dfrac{1}{\sigma^2}\displaystyle\sum_{i=1}^{n}(y_i - \beta x_i)x_i \\ -\dfrac{n}{2\sigma^2} + \dfrac{1}{2\sigma^4}\displaystyle\sum_{i=1}^{n}(y_i - \beta x_i)^2 \end{bmatrix}$$

(c) Find the Fisher information matrix $\boldsymbol{I_\theta}$.

Solutions:

$$\boldsymbol{I_\theta} = \begin{bmatrix} E\left[-\dfrac{\partial \ell^2}{\partial \beta^2}\right] & E\left[-\dfrac{\partial \ell^2}{\partial \beta \partial \sigma^2}\right] \\ E\left[-\dfrac{\partial \ell^2}{\partial \sigma^2 \partial \beta}\right] & E\left[-\dfrac{\partial \ell^2}{\partial \sigma^4}\right] \end{bmatrix} = \begin{bmatrix} \dfrac{1}{\sigma^2}\displaystyle\sum_{i=1}^n x_i^2 & 0 \\ 0 & \dfrac{n}{2\sigma^4} \end{bmatrix}.$$

The elements of $\boldsymbol{I_\theta}$ can be calculated as follows.

$$\begin{aligned} \frac{\partial \ell^2}{\partial \beta^2} &= \frac{\partial}{\partial \beta} S(\beta; \boldsymbol{y}) = -\frac{1}{\sigma^2}\sum_{i=1}^n x_i^2 \\ \frac{\partial \ell^2}{\partial \beta \partial \sigma^2} &= \frac{\partial}{\partial \sigma^2} S(\beta; \boldsymbol{y}) = -\frac{1}{\sigma^4}\sum_{i=1}^n (y_i - \beta x_i) x_i \\ \frac{\partial \ell^2}{\partial \sigma^2 \partial \beta} &= \frac{\partial}{\partial \beta} S(\sigma^2; \boldsymbol{y}) = -\frac{1}{\sigma^4}\sum_{i=1}^n (y_i - \beta x_i) x_i \\ \frac{\partial \ell^2}{\partial \sigma^4} &= \frac{\partial}{\partial \sigma^2} S(\sigma^2; \boldsymbol{y}) = \frac{n}{2\sigma^4} - \frac{1}{\sigma^6}\sum_{i=1}^n (y_i - \beta x_i)^2 \end{aligned}$$

It follows from the above that

$$\begin{aligned} E\left[-\frac{\partial \ell^2}{\partial \beta^2}\right] &= E\left[\frac{1}{\sigma^2}\sum_{i=1}^n x_i^2\right] = \frac{1}{\sigma^2}\sum_{i=1}^n x_i^2 \\ E\left[-\frac{\partial \ell^2}{\partial \beta \partial \sigma^2}\right] &= E\left[\frac{1}{\sigma^4}\sum_{i=1}^n (y_i - \beta x_i) x_i\right] = \frac{1}{\sigma^4}\sum_{i=1}^n (E[Y_i] - \beta x_i) x_i = 0 \\ E\left[-\frac{\partial \ell^2}{\partial \sigma^4}\right] &= E\left[-\frac{n}{2\sigma^4} + \frac{1}{\sigma^6}\sum_{i=1}^n (y_i - \beta x_i)^2\right] = -\frac{n}{2\sigma^4} - \frac{1}{\sigma^6} E\left[\sum_{i=1}^n (y_i - \beta x_i)^2\right] \\ &= -\frac{n}{2\sigma^4} + \frac{\sigma^2}{\sigma^6} E\left[\sum_{i=1}^n \left(\frac{y_i - \beta x_i}{\sigma}\right)^2\right] = -\frac{n}{2\sigma^4} + \frac{n}{\sigma^4} = \frac{n}{2\sigma^2}, \end{aligned}$$

and noting that:

- $E[Y_i] = \beta x_i$ and x_i is not a random variable; and
- $Z_i = \dfrac{y_i - \beta x_i}{\sigma} \sim N(0,1)$, and hence $\displaystyle\sum_{i=1}^n Z_i^2 \sim \chi_n^2$.

(d) Find the MLEs $\hat{\beta}$ and $\hat{\sigma}^2$.

Solutions:
We solve $S(\boldsymbol{\theta}; \boldsymbol{y}) = \boldsymbol{0}$ for $\boldsymbol{\theta} = (\beta, \sigma^2)$.

The MLE for β:

$$
\begin{aligned}
0 &= S(\beta; \boldsymbol{y}) \\
0 &= \sum_{i=1}^{n} (y_i - \beta x_i) x_i \\
0 &= \sum_{i=1}^{n} x_i y_i - \beta \sum_{i=1}^{n} x_i^2 \\
\beta \sum_{i=1}^{n} x_i^2 &= \sum_{i=1}^{n} x_i y_i \\
\hat{\beta} &= \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}
\end{aligned}
$$

The MLE for σ^2:

$$
\begin{aligned}
0 &= S(\sigma^2; \boldsymbol{y}) \\
0 &= -\frac{n}{\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^{n} (y_i - \beta x_i)^2 \\
\frac{n}{2\sigma^2} &= \frac{1}{2\sigma^4} \sum_{i=1}^{n} (y_i - \beta x_i)^2 \\
n\sigma^2 &= \sum_{i=1}^{n} (y_i - \beta x_i)^2 \\
\hat{\sigma}^2 &= \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{\beta} x_i)^2
\end{aligned}
$$

---

4. Suppose that $Y_1, Y_2, \ldots, Y_n$ are independent and identically distributed with density function

$$f(y; \theta) = e^{-(y-\theta)}, \quad y \geq \theta$$

and $f(y; \theta) = 0$ otherwise.
Find the MLE of θ. ***Hint: Take note of the region*** $y \geq \theta$ ***where the density is positive.***

Solutions:
Using standard steps (i.e. solving $S(\theta; \boldsymbol{y}) = 0$) will lead to problems as $f(y; \theta)$ is an increasing function of θ:

$$
\begin{aligned}
L(\theta; \boldsymbol{y}) &= \prod_{i=1}^{n} e^{-(y_i - \theta)} = e^{n\theta - \sum_{i=1}^{n} y_i} \\
\ell(\theta; \boldsymbol{y}) &= n\theta - \sum_{i=1}^{n} y_i \\
S(\theta; \boldsymbol{y}) &= n
\end{aligned}
$$

We want to maximize $\ell(\theta; \boldsymbol{y})$ with respect to θ. Given that $\ell(\theta; \boldsymbol{y}) = n\theta - \sum_{i=1}^{n} y_i$, we would choose the largest θ possible. But recall the constraint that $f(y; \theta)$ is positive only if $y_i \geq \theta$. Hence, θ must be at most as large as the smallest y_i. It follows that $\hat{\theta} = y_{(1)}$, where $y_{(1)} = \min(y_1, y_2, \ldots, y_n)$.

---

5. Suppose $Y_1, Y_2, \ldots, Y_n \overset{i.i.d.}{\sim} \text{Exp}(\lambda)$ so that the density function is

$$f_{Y_i}(y;\lambda) = \lambda e^{-\lambda y_i}\,.$$

Consider the equivalent parameterisation in terms of $\theta = \frac{1}{\lambda}$ where

$$f_{Y_i}(y;\theta) = \frac{1}{\theta} e^{-\frac{1}{\theta} y_i}\,.$$

By considering the transformation $\Phi(\lambda) = \frac{1}{\lambda} = \theta$, do the following:

(a) Calculate the log-likelihoods $\ell_\lambda(\lambda; \boldsymbol{y})$ and $\ell_\theta(\theta; \boldsymbol{y})$. Verify directly that $\ell_\lambda(\lambda; \boldsymbol{y}) = \ell_\theta(\Phi(\lambda); \boldsymbol{y})$ and $\ell_\theta(\theta; \boldsymbol{y}) = \ell_\lambda(\Phi^{-1}(\theta); \boldsymbol{y})$.

Solutions:
For λ:

The likelihood is:

$$\begin{aligned} L(\lambda; \boldsymbol{y}) &= \prod_{i=1}^{n} f_{Y_i}(y;\lambda) \\ &= \prod_{i=1}^{n} \lambda e^{-\lambda y_i} \\ &= \lambda^n e^{-\lambda \sum_{i=1}^{n} y_i} \\ &= \lambda^n e^{-\lambda n \bar{y}}\,. \end{aligned}$$

Hence the log-likelihood is

$$\begin{aligned} \ell(\lambda; \boldsymbol{y}) &= \log\left(\lambda^n e^{-\lambda n \bar{y}}\right) \\ &= n\log(\lambda) - \lambda n \bar{y}\,. \end{aligned}$$

For θ:

The likelihood is:

$$\begin{aligned} L(\theta; \boldsymbol{y}) &= \prod_{i=1}^{n} f_{Y_i}(y;\theta) \\ &= \prod_{i=1}^{n} \frac{1}{\theta} e^{-\frac{1}{\theta} y_i} \\ &= \left(\frac{1}{\theta}\right)^n e^{-\frac{1}{\theta} \sum_{i=1}^{n} y_i} \\ &= \frac{1}{\theta^n} e^{-\frac{1}{\theta} n \bar{y}}\,. \end{aligned}$$

Hence the log-likelihood is

$$\begin{aligned} \ell(\theta; \boldsymbol{y}) &= \log\left(\frac{1}{\theta^n} e^{-\frac{1}{\theta} n \bar{y}}\right) \\ &= -n\log(\theta) - \frac{1}{\theta} n \bar{y}\,. \end{aligned}$$

Now, we have

$\ell_\lambda(\lambda; \boldsymbol{y}) = \ell_\theta(\Phi(\lambda); \boldsymbol{y})$ and $\ell_\theta(\theta; \boldsymbol{y}) = \ell_\lambda(\Phi^{-1}(\theta); \boldsymbol{y})$

$$\begin{aligned}
\ell_\theta(\Phi(\lambda); \boldsymbol{y}) &= \ell_\theta\left(\frac{1}{\lambda}; \boldsymbol{y}\right) \\
&= -n\log\left(\frac{1}{\lambda}\right) - \frac{1}{\frac{1}{\lambda}} n\bar{y} \\
&= n\log(\lambda) - \lambda n\bar{y} \\
&= \ell_\lambda(\lambda; \boldsymbol{y}) \,,
\end{aligned}$$

and

$$\begin{aligned}
\ell_\lambda(\Phi^{-1}(\theta); \boldsymbol{y}) &= \ell_\lambda\left(\frac{1}{\theta}; \boldsymbol{y}\right) \\
&= n\log\left(\frac{1}{\theta}\right) - \frac{1}{\theta} n\bar{y} \\
&= -n\log(\theta) - \frac{1}{\theta} n\bar{y} \\
&= \ell_\theta(\theta; \boldsymbol{y}) \,,
\end{aligned}$$

as required.

(b) Calculate $\hat{\lambda}$, the maximum likelihood estimate of λ. Hence, calculate the the maximum likelihood estimate of θ.

Solutions:
The score function is:

$$\begin{aligned}
S(\lambda; \boldsymbol{y}) &= \frac{\partial}{\partial\lambda} \ell_\lambda(\lambda; \boldsymbol{y}) \\
&= \frac{\partial}{\partial\lambda}\left[n\log(\lambda) - \lambda n\bar{y}\right] \\
&= \frac{n}{\lambda} - n\bar{y} \,.
\end{aligned}$$

Evaluating this equal to zero will give the MLE $\hat{\lambda}$ such that $\frac{n}{\hat{\lambda}} - n\bar{y} = 0$. Hence

$$\hat{\lambda} = \frac{1}{\bar{y}} \,.$$

By Theorem 15 we know that $\hat{\theta} = \Phi(\hat{\lambda})$, so

$$\hat{\theta} = \bar{y} \,.$$
