ST 793 : Solution of Homework-3

Salil Koner

September 27, 2019

Problem 3.9

We want to test

$$H_0 : \lambda_1 = \lambda_2 = \cdots = \lambda_n = \lambda$$

The likelihood function is given by

$$L(\lambda_1, \ldots, \lambda_n\,; Y_1, Y_2, \ldots, Y_n) = e^{-\sum_{i=1}^{n} \lambda_i} \frac{\Pi_{i=1}^{n} \lambda^{Y_i}}{\Pi_{i=1}^{n} Y_i!} I(Y_i = 0, 1, 2, \ldots \quad \forall\, i)$$

Which means the log-likelihood is given by

$$l(\lambda_1, \ldots, \lambda_n\,; Y_1, Y_2, \ldots, Y_n) = -\sum_{i=1}^{n} \lambda_i + \sum_{i=1}^{n} Y_i \log_e \lambda_i - \sum_{i=1}^{n} \log_e Y_i! \tag{1}$$

The score function is given by

$$\mathbf{S}(\lambda_1, \ldots, \lambda_n) = \left(\frac{Y_1}{\lambda_1} - 1, \frac{Y_2}{\lambda_2} - 1, \ldots, \frac{Y_n}{\lambda_n} - 1 \right)^{\top} \tag{2}$$

And the Information matrix is given by

$$\begin{aligned} \mathbf{I}(\lambda_1, \ldots, \lambda_n) &= \operatorname{diag}\left(E\left(\frac{Y_1}{\lambda_1^2}\right), \ldots, E\left(\frac{Y_n}{\lambda_n^2}\right) \right) \\ &= \operatorname{diag}\left(\left(\frac{1}{\lambda_1}\right), \ldots, \left(\frac{1}{\lambda_n}\right) \right) \end{aligned} \tag{3}$$

Under H_0, the likelihood function in (1) simplifies to

$$l(\lambda\,; Y_1, Y_2, \ldots, Y_n) = -n\lambda + \log_e \lambda \sum_{i=1}^{n} Y_i - \sum_{i=1}^{n} \log_e Y_i! \tag{4}$$

So, we get the MLE under H_0 by setting the first derivative (with respect to *lambda*) of log-likelihood in (4), which is

$$\tilde{\lambda} = \bar{Y} = \frac{1}{n} \sum_{i=1}^{n} Y_i$$

The score test-statistic is given by

$$\begin{aligned} T_s &= \mathbf{S}\left(\tilde{\lambda}, \ldots, \tilde{\lambda}\right)^{\top} \mathbf{I}\left(\tilde{\lambda}, \ldots, \tilde{\lambda}\right)^{-1} \mathbf{S}\left(\tilde{\lambda}, \ldots, \tilde{\lambda}\right) \\ &= \left(\frac{Y_1}{\bar{Y}} - 1, \frac{Y_2}{\bar{Y}} - 1, \ldots, \frac{Y_n}{\bar{Y}} - 1 \right)^{\top} \bar{Y} \mathbf{I}_n \left(\frac{Y_1}{\bar{Y}} - 1, \frac{Y_2}{\bar{Y}} - 1, \ldots, \frac{Y_n}{\bar{Y}} - 1 \right) \\ &= \sum_{i=1}^{n} \left(\frac{Y_i - \bar{Y}}{\bar{Y}} \right)^2 \bar{Y} \\ &= \sum_{i=1}^{n} \frac{\left(Y_i - \bar{Y}\right)^2}{\bar{Y}} \end{aligned} \tag{5}$$

This completes the proof.

Problem 3.13

(a) The score function is given by

$$\begin{aligned}
\mathbf{S}(\boldsymbol{\beta}) &= \frac{\partial}{\partial \boldsymbol{\beta}} \log L(\boldsymbol{\beta}) \\
&= \sum_{i=1}^{k} \sum_{j=1}^{m_i} \left[Y_{ij} \frac{p_i'(\boldsymbol{\beta})}{p_i(\boldsymbol{\beta})} \frac{\partial}{\partial \boldsymbol{\beta}} \left(\mathbf{x}_i^\top \boldsymbol{\beta} \right) + (n_{ij} - Y_{ij}) \frac{-p_i'(\boldsymbol{\beta})}{1 - p_i(\boldsymbol{\beta})} \frac{\partial}{\partial \boldsymbol{\beta}} \left(\mathbf{x}_i^\top \boldsymbol{\beta} \right) \right] \\
&= \sum_{i=1}^{k} \sum_{j=1}^{m_i} \left[Y_{ij} \frac{p_i'(\boldsymbol{\beta})}{p_i(\boldsymbol{\beta})} \mathbf{x}_i - (n_{ij} - Y_{ij}) \frac{p_i'(\boldsymbol{\beta})}{1 - p_i(\boldsymbol{\beta})} \mathbf{x}_i \right] \\
&= \sum_{i=1}^{k} \sum_{j=1}^{m_i} \left[\frac{Y_{ij}}{p_i(\boldsymbol{\beta})} - \frac{(n_{ij} - Y_{ij})}{1 - p_i(\boldsymbol{\beta})} \right] p_i'(\boldsymbol{\beta}) \mathbf{x}_i \qquad (6) \\
&= \sum_{i=1}^{k} \sum_{j=1}^{m_i} \left[\frac{Y_{ij} - n_{ij} p_i(\boldsymbol{\beta})}{p_i(\boldsymbol{\beta}) \left(1 - p_i(\boldsymbol{\beta})\right)} \right] p_i'(\boldsymbol{\beta}) \mathbf{x}_i \qquad (7)
\end{aligned}$$

(b) From (6), if we take a derivative of score function we get the observed total information matrix as

$$\begin{aligned}
\mathbf{I}_T(\mathbf{Y}, \boldsymbol{\beta}) &= -\frac{\partial}{\partial \boldsymbol{\beta}^\top} \mathbf{S}(\boldsymbol{\beta}) \\
&= -\sum_{i=1}^{k} \sum_{j=1}^{m_i} \left[Y_{ij} \left(\frac{p_i(\boldsymbol{\beta}) p_i''(\boldsymbol{\beta}) - [p_i'(\boldsymbol{\beta})]^2}{[p_i(\boldsymbol{\beta})]^2} \right) - (n_{ij} - Y_{ij}) \left(\frac{(1 - p_i(\boldsymbol{\beta}))\, p_i''(\boldsymbol{\beta}) + [p_i'(\boldsymbol{\beta})]^2}{[1 - p_i(\boldsymbol{\beta})]^2} \right) \right] \mathbf{x}_i \mathbf{x}_i^\top
\end{aligned}$$

Using the fact that $E(Y_{ij}) = n_{ij} p_i(\boldsymbol{\beta})$, we get the Fisher Information matrix as

$$\begin{aligned}
\mathbf{I}_T(\boldsymbol{\beta}) &= \mathrm{E}\left(\mathbf{I}_T(\mathbf{Y}, \boldsymbol{\beta}) \right) \\
&= -\sum_{i=1}^{k} \sum_{j=1}^{m_i} \left[n_{ij} p_i(\boldsymbol{\beta}) \left(\frac{p_i(\boldsymbol{\beta}) p_i''(\boldsymbol{\beta}) - [p_i'(\boldsymbol{\beta})]^2}{[p_i(\boldsymbol{\beta})]^2} \right) - (n_{ij} - n_{ij} p_i(\boldsymbol{\beta})) \left(\frac{(1 - p_i(\boldsymbol{\beta}))\, p_i''(\boldsymbol{\beta}) + [p_i'(\boldsymbol{\beta})]^2}{[1 - p_i(\boldsymbol{\beta})]^2} \right) \right] \mathbf{x}_i \mathbf{x}_i^\top \\
&= -\sum_{i=1}^{k} \sum_{j=1}^{m_i} \left[n_{ij} \left(\frac{p_i(\boldsymbol{\beta}) p_i''(\boldsymbol{\beta}) - [p_i'(\boldsymbol{\beta})]^2}{p_i(\boldsymbol{\beta})} \right) - n_{ij} \left(\frac{(1 - p_i(\boldsymbol{\beta}))\, p_i''(\boldsymbol{\beta}) + [p_i'(\boldsymbol{\beta})]^2}{1 - p_i(\boldsymbol{\beta})} \right) \right] \mathbf{x}_i \mathbf{x}_i^\top \\
&= -\sum_{i=1}^{k} \sum_{j=1}^{m_i} n_{ij} \left[\left(p_i''(\boldsymbol{\beta}) - \frac{[p_i'(\boldsymbol{\beta})]^2}{p_i(\boldsymbol{\beta})} \right) - \left(p_i''(\boldsymbol{\beta}) + \frac{[p_i'(\boldsymbol{\beta})]^2}{1 - p_i(\boldsymbol{\beta})} \right) \right] \mathbf{x}_i \mathbf{x}_i^\top \\
&= \sum_{i=1}^{k} \sum_{j=1}^{m_i} n_{ij} \left[\frac{1}{p_i(\boldsymbol{\beta})} + \frac{1}{1 - p_i(\boldsymbol{\beta})} \right] [p_i'(\boldsymbol{\beta})]^2\, \mathbf{x}_i \mathbf{x}_i^\top \\
&= \sum_{i=1}^{k} \sum_{j=1}^{m_i} n_{ij} \left[\frac{1}{p_i(\boldsymbol{\beta}) \left(1 - p_i(\boldsymbol{\beta})\right)} \right] [p_i'(\boldsymbol{\beta})]^2\, \mathbf{x}_i \mathbf{x}_i^\top \\
&= \sum_{i=1}^{k} n_{i+} \left[\frac{[p_i'(\boldsymbol{\beta})]^2}{p_i(\boldsymbol{\beta}) \left(1 - p_i(\boldsymbol{\beta})\right)} \right] \mathbf{x}_i \mathbf{x}_i^\top \qquad \left[n_{i+} = \sum_{j=1}^{m_i} n_{ij} \right]
\end{aligned}$$

(c) When $F(x) = (1 + \exp(-x))^{-1}$, some interesting thing happens. Notice that

$$F'(x) = \frac{e^{-x}}{(1 + e^{-x})^2} = F(x)(1 - F(x))$$

Which means,

$$p_i'(\boldsymbol{\beta}) = p_i(\boldsymbol{\beta}) \left(1 - p_i(\boldsymbol{\beta})\right)$$

Then, from (7) the score function can be simplied as

$$\mathbf{S}(\boldsymbol{\beta}) = \sum_{i=1}^{k}\sum_{j=1}^{m_i}\left[Y_{ij} - n_{ij}p_i(\boldsymbol{\beta})\right]\mathbf{x}_i \tag{8}$$

Taking second derivative of (8), we get the observed information matrix as

$$\begin{aligned}\mathbf{I}_T(\mathbf{Y}, \boldsymbol{\beta}) &= -\frac{\partial}{\partial \boldsymbol{\beta}^\top}\mathbf{S}(\boldsymbol{\beta}) \\ &= \sum_{i=1}^{k}\sum_{j=1}^{m_i}\left[n_{ij}p_i'(\boldsymbol{\beta})\right]\mathbf{x}_i\mathbf{x}_i^\top \\ &= \sum_{i=1}^{k}\sum_{j=1}^{m_i} n_{ij}\left[p_i(\boldsymbol{\beta})\left(1 - p_i(\boldsymbol{\beta})\right)\right]\mathbf{x}_i\mathbf{x}_i^\top \end{aligned} \tag{9}$$

We see that the observed information matrix does not involve any Y_{ij} component, so after taking expectation the Information matrix and the observed information matrix becomes equal for that particular form of F, the form of information matrix is as in (9). Hence the proof.

Problem 3.15

The log-likelihood function is given by

$$\ell\left(p_1, \ldots, p_n\right) = \sum_{i=1}^{n}\log_e\binom{m_i}{Y_i} + \sum_{i=1}^{n} Y_i \log_e p_i + \sum_{i=1}^{n}(m_i - Y_i)\log_e(1 - p_i) \tag{10}$$

Taking partial derivatives with respect to p_i we get the score function as

$$\mathbf{S}(\mathbf{p}) = \begin{pmatrix}\frac{Y_1}{p_1} - \frac{m_1 - Y_1}{1-p_1} & \frac{Y_2}{p_2} - \frac{m_2 - Y_2}{1-p_2} & \cdots & \frac{Y_n}{p_n} - \frac{m_n - Y_n}{1-p_n}\end{pmatrix}^\top \tag{11}$$

And taking a further derivative the we get second derivative as

$$\frac{\partial}{\partial \mathbf{p}^\top}\mathbf{S}(\mathbf{p}) = \begin{pmatrix} -\frac{Y_1}{p_1^2} - \frac{m_1 - Y_1}{(1-p_1)^2} & 0 & \cdots & 0 \\ 0 & -\frac{Y_2}{p_2^2} - \frac{m_2 - Y_2}{(1-p_2)^2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & -\frac{Y_n}{p_n^2} - \frac{m_n - Y_n}{(1-p_n)^2} \end{pmatrix}$$

So, the Fisher information matrix is given by

$$\mathbf{I}_T(\mathbf{p}) = -\mathrm{E}\left(\frac{\partial}{\partial \mathbf{p}^\top}\mathbf{S}(\mathbf{p})\right) \tag{12}$$

$$= \begin{pmatrix} \frac{m_1}{p_1} + \frac{m_1}{1-p_1} & 0 & \cdots & 0 \\ 0 & \frac{m_2}{p_2} + \frac{m_2}{1-p_2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \frac{m_n}{p_n} + \frac{m_n}{1-p_n} \end{pmatrix} \tag{13}$$

$$= \begin{pmatrix} \frac{m_1}{p_1(1-p_1)} & 0 & \cdots & 0 \\ 0 & \frac{m_2}{p_2(1-p_2)} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \frac{m_n}{p_n(1-p_n)} \end{pmatrix} \tag{14}$$

Let $\tilde{\mathbf{p}}$ be the MLE of $\mathbf{p}$ under H_0. Then the score statistic is given by

$$\begin{aligned}
T_s &= \mathbf{S}(\tilde{\mathbf{p}})^\top \mathbf{I}_T^{-1}(\tilde{\mathbf{p}})\mathbf{S}(\tilde{\mathbf{p}}) && (15)\\
&= \sum_{i=1}^{n} \left(\frac{Y_i}{\tilde{p}_i} - \frac{m_i - Y_i}{1-\tilde{p}_i}\right)^2 \frac{\tilde{p}_i(1-\tilde{p}_i)}{m_i} && (16)\\
&= \sum_{i=1}^{n} \left[\frac{Y_i - m_i\tilde{p}_i}{\tilde{p}_i\,(1-\tilde{p}_i)}\right]^2 \frac{\tilde{p}_i(1-\tilde{p}_i)}{m_i} && (17)\\
&= \sum_{i=1}^{n} \frac{(Y_i - m_i\tilde{p}_i)^2}{m_i\tilde{p}_i\,(1-\tilde{p}_i)} && (18)
\end{aligned}$$

Hence the proof.

Problem 3.18

The log-likelihood function is given by

$$\ell(\mu_1,\mu_2) = -\frac{1}{2}\left[\sum_{j=1}^{n_1}(Y_{1j}-\mu_1)^2 + \sum_{j=1}^{n_2}(Y_{2j}-\mu_2)^2\right] \tag{19}$$

Under $H_0 : \mu_1 = \mu_2 = \mu$, the MLE of μ is obtained by replacing $\mu_1 = \mu_2 = \mu$ in the likelihood function and setting its derivative equal to zero.

$$0 = \frac{\partial}{\partial\mu}\ell(\mu,\mu) = \left[\sum_{j=1}^{n_1}(Y_{1j}-\mu) + \sum_{j=1}^{n_2}(Y_{2j}-\mu)\right]$$

Which implies that the MLE under H_0 is

$$\tilde{\mu}_1 = \tilde{\mu}_2 = \tilde{\mu} = \frac{n_1\bar{Y}_1 + n_2\bar{Y}_2}{n_1+n_2} \tag{20}$$

It is given that under $H_0 \cup H_1 = \{(\mu_1,\mu_2) : \mu_1 \le \mu_2\}$, the MLE is given by

$$(\hat{\mu}_1,\hat{\mu}_2) = \begin{cases} (\bar{Y}_1,\bar{Y}_2) & \text{if } \bar{Y}_1 < \bar{Y}_2 \\ (\tilde{\mu},\tilde{\mu}) & \text{otherwise} \end{cases} \tag{21}$$

Note that, twice log-likelihood can be re-written as

$$-2\ell(\mu_1,\mu_2) = \left[\sum_{j=1}^{n_1}\left(Y_{1j}-\bar{Y}_1\right)^2 + \sum_{j=1}^{n_2}\left(Y_{2j}-\bar{Y}_2\right)^2\right] + \left[n_1\left(\bar{Y}_1-\mu_1\right)^2 + n_2\left(\bar{Y}_2-\mu_2\right)^2\right] \tag{22}$$

This means,

$$-2\ell(\tilde{\mu},\tilde{\mu}) = \left[\sum_{j=1}^{n_1}\left(Y_{1j}-\bar{Y}_1\right)^2 + \sum_{j=1}^{n_2}\left(Y_{2j}-\bar{Y}_2\right)^2\right] + \left[n_1\left(\bar{Y}_1-\tilde{\mu}\right)^2 + n_2\left(\bar{Y}_2-\tilde{\mu}\right)^2\right] \tag{23}$$

and

$$\begin{aligned}
-2\ell(\hat{\mu}_1,\hat{\mu}_2) &= \left[\sum_{j=1}^{n_1}\left(Y_{1j}-\bar{Y}_1\right)^2 + \sum_{j=1}^{n_2}\left(Y_{2j}-\bar{Y}_2\right)^2\right] + \left[n_1\left(\bar{Y}_1-\hat{\mu}_1\right)^2 + n_2\left(\bar{Y}_2-\hat{\mu}_2\right)^2\right] && (24)\\
&= \begin{cases} \sum_{j=1}^{n_1}\left(Y_{1j}-\bar{Y}_1\right)^2 + \sum_{j=1}^{n_2}\left(Y_{2j}-\bar{Y}_2\right)^2 & \text{if } \bar{Y}_1 < \bar{Y}_2 \\ -2\ell(\tilde{\mu},\tilde{\mu}) & \text{otherwise} \end{cases} && (25)
\end{aligned}$$

So, the likelihood ratio test is given by

$$\begin{aligned}
T_{LR} &= -2\{\ell(\tilde{\mu},\tilde{\mu}) - \ell(\hat{\mu}_1,\hat{\mu}_2)\} && (26)\\
&= \begin{cases} n_1\left(\bar{Y}_1-\tilde{\mu}_1\right)^2 + n_2\left(\bar{Y}_2-\tilde{\mu}_2\right)^2 & \text{if } \bar{Y}_1<\bar{Y}_2 \\ 0 & \text{otherwise}\end{cases} && (27)
\end{aligned}$$

Now, we can do further simplification, such as

$$\begin{aligned}
n_1\left(\bar{Y}_1-\tilde{\mu}_1\right)^2 + n_2\left(\bar{Y}_2-\tilde{\mu}_2\right)^2 &= n_1\left(\bar{Y}_1 - \frac{n_1\bar{Y}_1+n_2\bar{Y}_2}{n_1+n_2}\right)^2 + n_2\left(\bar{Y}_2 - \frac{n_1\bar{Y}_1+n_2\bar{Y}_2}{n_1+n_2}\right)^2\\
&= n_1n_2^2\left(\frac{\bar{Y}_1-\bar{Y}_2}{n_1+n_2}\right)^2 + n_1^2n_2\left(\frac{\bar{Y}_1-\bar{Y}_2}{n_1+n_2}\right)^2\\
&= n_1n_2\left(\frac{\bar{Y}_1-\bar{Y}_2}{n_1+n_2}\right)^2(n_1+n_2)\\
&= \frac{\left(\bar{Y}_1-\bar{Y}_2\right)^2}{\frac{1}{n_1}+\frac{1}{n_2}}
\end{aligned}$$

So, the likelihood ratio test-statistic can be simplified as

$$T_{LR} = \left(\frac{\bar{Y}_2-\bar{Y}_1}{\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\right)^2 I\left(\bar{Y}_2-\bar{Y}_1>0\right) \tag{28}$$

Note that, under H_0,

$$\bar{Y}_2-\bar{Y}_1 \sim \text{Normal}\left(0,\ \frac{1}{n_1}+\frac{1}{n_2}\right) \implies \frac{\bar{Y}_2-\bar{Y}_1}{\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim \text{Normal}\,(0,1)$$

This means, the null distribution of T_{LR} follows the same distribution as $Z^2 I(Z>0)$ where Z follows a standard normal distribution. CDF and density of that particular distribution we will derive at the next problem. It turns out that the testing procedure under $\alpha = 0.05$ will be

$$\text{Reject } H_0 \quad \text{if } T_{LR} > 2.70554$$

Problem 3.19

Let $W = Z^2 I(Z>0)$, then $P(W=0) = P(Z\le 0) = 0.5$. This means W has point mass of 0.5 at 0. And W can not be negative. The CDF of W is given by

$$\begin{aligned}
F_W(w) = P(W\le w) &= \begin{cases} 0 & \text{if } w<0\\ 0.5 & \text{if } w=0\\ 0.5 + P\left(Z^2\le w, Z>0\right) & \text{if } w>0\end{cases}\\
&= \begin{cases} 0 & \text{if } w<0\\ 0.5 & \text{if } w=0\\ 0.5 + P\left(0<Z\le\sqrt{w}\right) & \text{if } w>0\end{cases}\\
&= \begin{cases} 0 & \text{if } w<0\\ 0.5 & \text{if } w=0\\ 0.5 + \Phi\left(\sqrt{w}\right) - \Phi(0) & \text{if } w>0\end{cases}\\
&= \begin{cases} 0 & \text{if } w<0\\ 0.5 & \text{if } w=0\\ \Phi\left(\sqrt{w}\right) & \text{if } w>0\end{cases}
\end{aligned}$$

This means that W follows a mixture distribution and its density is given by

$$\begin{aligned} f_W(w) &= 0.5I(w=0) + 0.5w^{-1/2}\phi\left(\sqrt{w}\right)I(w>0) \\ &= 0.5I(w=0) + 0.5\left(\frac{1}{\sqrt{2\pi}}w^{1/2-1}e^{-w/2}\right)I(w>0) \\ &= \frac{1}{2}I(w=0) + \frac{1}{2}\,\chi_1^2\,I(w>0) \end{aligned} \tag{29}$$

So, for any $\alpha > 0.5$, the formula for αth quantile is given by

$$w_q = \left(\Phi^{-1}(\alpha)\right)^2$$

or equivalently,

$$w_q = \chi^2_{1\,;\,(2\alpha-1)}$$

where $\chi^2_{1,\tau}$ is the τth quantile of chi-square distribution with 1 degrees of freedom and Φ is the CDF of standard normal distribution. This means, the 0.90, 0.95 and 0.99 quantiles of W is given by

$$w_{.9} = 1.64237 \quad w_{.95} = 2.70554 \quad w_{.99} = 5.41189$$