MATH 392 Problem Set 3

Sam D. Olson

Note on PS3

Time Taken

Total Time on Exercises from the Book: 9 1/2 hours Solving Exercises from the Book: 5 hours LaTeX writeup of Exercises from the Book: 4 1/2 hours

Additional Notes

- This Problem Set does not contain the Case Study. It will be turned in on Monday, 2/17/2020.
- Answers to All 8 problems are included.

Exercises from the book

7.5: 5

Let $y = \sum_{i}^{n} x_i$

Additionally, denote: $\bar{x} = x_1, ..., x_n$

Then, for $y > 0$ the likelihood function is given by:

$$f_n\left(\bar{x} \mid \theta\right) = \frac{e^{-n\theta}\theta^y}{\prod\limits_{i=1}^{n} x_i!}$$

Let $L\left(\theta\right) = log\left(f_n\left(\bar{x} \mid \theta\right)\right)$

Note that $\prod\limits_{i=1}^{n} x_i!$ does not depend on θ, as such, let $\psi = \prod\limits_{i=1}^{n} x_i!$

Then we may write:

$L\left(\theta\right) = -n\theta + ylog(\theta) - \psi$

Then we may take the derivative of $L\left(\theta\right)$ with respect to θ, giving us:

$$\frac{\partial L\left(\theta\right)}{\partial \theta} = -n + \frac{y}{\theta}$$

To determine the M.L.E. of θ, we maximize this equation and set it equal to zero, such that:

$$0 = -n + \frac{y}{\theta} \rightarrow n = \frac{y}{\theta} \rightarrow \hat{\theta} = \frac{y}{n}$$

(A): Thus, the M.L.E. of θ is $\hat{\theta} = \frac{y}{n} = \bar{x}_n$.

(B): If $y = 0$, and if $\theta \neq 0$ ($\theta > 0$), then $\theta = 0$ is not in the parameter space, Ω, then the M.L.E. of θ does not exist.

For additional commentary, we may note that if $y = 0$, then the likelihood function $f_n(\bar{x} \mid \theta)$ is a decreasing function of θ. As such, the maximum of the derivative of $L(\theta)$ occurs when $\theta = 0$. However, $\theta > 0$ is a condition of θ, and as such, the M.L.E. of θ does not exist.

7.5: 8

Let $y = \sum\limits_{i}^{n} x_i$

Additionally, denote: $\bar{x} = x_1, ..., x_n$

We note the likelihood function, $f_n(\bar{x} \mid \theta)$, as:

$$f_n(\bar{x} \mid \theta) = \left\{ \begin{matrix} e^{\theta - y} \ for \ x > \theta \\ 0 \ for \ x \leq \theta \end{matrix} \right\}$$

Let $L(\theta) = log(f_n(\bar{x} \mid \theta))$

Then, for $x > \theta$, we may write:

$L(\theta) = \theta - y$

Then we may take the derivative of $L(\theta)$ with respect to θ, setting it equal to 0 to maximize the likelihood function, giving us:

$$\frac{\partial L(\theta)}{\partial \theta} = 0 \rightarrow 1 = 0$$

(A): As this is markedly false, we conclude the M.L.E. of θ does not exist.

Alternatively, we may not that $\forall x_i \in \bar{x}, f_n(\bar{x} \mid \theta)$ will be maximized when θ is maximized (while holding the relation $\theta < x$). As such, $\theta < min(x_1, ..., x_n)$. However, due to the **strict inequality**, the maximum $\theta = min(x_1, ..., x_n)$ cannot be used as an estimate of θ. As such, the M.L.E. of θ does not exist.

We may take advantage of the example provided in the book to create another version of the p.d.f. for which the M.L.E. of θ will exist. To do so, let the **strict inequality** be swapped with the **weak inequality**, such that we may write:

$$f(x \mid \theta) = \left\{ \begin{matrix} e^{\theta - y} \ for \ x \geq \theta \\ 0 \ for \ x < \theta \end{matrix} \right\}$$

We may then note the likelihood function of this p.d.f as:

$$f_n(\bar{x} \mid \theta) = \left\{ \begin{matrix} e^{\theta - y} \ for \ x \geq \theta \\ 0 \ for \ x < \theta \end{matrix} \right\}$$

We may then note: The likelihood function will be non-zero for $\theta \leq min(x_1, ..., x_n)$.

(B): Thus, the M.L.E. of θ, $\hat{\theta} = min(x_1, ..., x_n)$

7.5: 11

Note, as we are dealing with the interval $[\theta_1, \theta_2]$, we know $\theta_1 \leq \theta_2$

We may then write the p.d.f. of each x_i as:

$$f(x_i \mid \theta_1, \theta_2) = \left\{ \begin{matrix} \frac{1}{\theta_2 - \theta_1} \ for \ \theta_1 \leq x \leq \theta_2 \\ 0 \ otherwise \end{matrix} \right\}$$

Letting $\bar{x} = x_1, ..., x_n$, we may then write the likelihood function as:

$$f_n\left(\bar{x} \mid \theta_1, \theta_2\right) = \left\{ \begin{array}{c} \frac{1}{(\theta_2 - \theta_1)}^n \; for \; \theta_1 \leq min(x_1, ..., x_n) \leq max(x_1, ..., x_n) \leq \theta_2 \\ 0 \;\; otherwise \end{array} \right\}$$

We may then note that $f_n\left(\bar{x} \mid \theta_1, \theta_2\right)$ is maximized when $\theta_2 - \theta_1$ is minimized.

Thus, we may note:

$min(\theta_2) = max(x_1, ..., x_n)$

And

$max(\theta_1) = min(x_1, ..., x_n)$

Thus, the M.L.E. of θ may be written: $\theta = (\hat{\theta}_1 = min(x_1, ..., x_n), \hat{\theta}_2 = max(x_1, ..., x_n))$.

7.5: 12

Let $\theta = \theta_1, ..., \theta_n$

Note the following relations for use later: $k \geq 2$, $0 \leq \theta_i \leq 1$, $\theta_1 + ... + \theta_k = 1$, and $n_1 + ... + n_k = n$

Then, the likelihood function is given by:

$$f_n\left(\bar{x} \mid \theta_1, ..., \theta_k\right) = \prod_{i=1}^{k} \theta_i{}^{n_i} = \theta_1{}^{n_1} ... \theta_k{}^{n_k}$$

For ease of computation, let $L\left(\theta\right) = log\left(f_n\left(\bar{x} \mid \theta_1, ..., \theta_k\right)\right)$

Note that: $\theta_k = 1 - \sum_{i=1}^{k-1} \theta_i$

Then we may write:

$L\left(\theta\right) = \sum_{i=1}^{k} n_i log(\theta_i)$

Then we may take the derivative with respect to θ_i, giving us:

$$\frac{\partial L\left(\theta\right)}{\partial \theta_i} = \frac{\partial \left(n_k \left(1 - \sum_{i=1}^{k-1} \theta_i \right) + \sum_{i=1}^{k-1} n_i log(\theta_i) \right)}{\partial \theta_i}$$

This evaluates to:

$$\frac{\partial L\left(\theta\right)}{\partial \theta_i} = -\frac{n_k}{\left(1 - \sum_{i=1}^{k-1} \theta_i \right)} + \frac{n_i}{\theta_i} = -\frac{n_k}{\theta_k} + \frac{n_i}{\theta_i}$$

To determine the M.L.E. of θ_i, we maximize this equation and set it equal to zero, such that:

$0 = -\frac{n_k}{\theta_k} + \frac{n_i}{\theta_i} \rightarrow \frac{n_k}{\theta_k} = \frac{n_i}{\theta_i}$ for $i = 1, ..., k-1$

For $i = 1, ..., k$, if we set $\theta_i = \alpha n_i$, then it follows:

$1 = \sum_{i=1}^{k} \theta_i = \sum_{i=1}^{k} \alpha n_i = \alpha \sum_{i=1}^{k} n_i$

Noting that $\sum_{i=1}^{n} n_i = n$, we may then say:

$1 = \alpha n \rightarrow \alpha = \frac{1}{n}$

We may then write, for $i = 1, ..., k$, the M.L.E. of θ_i is:

$\hat{\theta}_i = \alpha n_i = n_i/n$

Intuitively, and similar to other applications of the M.L.E., this results says that the M.L.E. of the individuals of the type i is the proportion of individuals of type i.

7.6: 3

Let m denote the median of the Exponential distribution.

Note the median of the Exponential distribution may be derived from the following equation:

$\int\limits_{0}^{m} \beta e^{-\beta x} dx = \frac{1}{2}$

Evaluating this integral gives us:

$\frac{1}{2} = 1 - e^{-\beta m} \rightarrow 1 = 2 - e^{-\beta m}$

Solving for m, we take the log of the above equation, giving us:

$0 = log(2) - \beta m \rightarrow \beta m = log(2) \rightarrow m = \frac{log(2)}{\beta}$

Thus, for $\hat{\beta}$, the sample median, $\hat{m} = \frac{log(2)}{\hat{\beta}}$

All praise the glorious tables at the back of the book For our purposes, note the mean of the Exponential distribution, $\mu = \frac{1}{\beta}$.

Solving for β, we then have $\beta = \frac{1}{\mu}$

For the sample mean, $\bar{x}_n$, note: $\hat{\beta} = \frac{1}{\bar{x}_n}$

Thus, for the sample median, we have:

$\hat{m} = \frac{log(2)}{\frac{1}{\bar{x}_n}} = \bar{x}_n log(2)$

7.6: 12

Note from the prior exercise, the parameter $\hat{\beta} = \frac{1}{\bar{x}_n}$

Additionally, we may once again **praise the glorious tables at the back of the book** and note the mean of the Exponential distribution is given by: $\mu = \frac{1}{\beta}$

Via application of the Law of Large Numbers, we may then write:

$\lim\limits_{n \to \infty} Pr \left(| \bar{x}_n - \frac{1}{\beta} | \right) = 1 \equiv \bar{x}_n \xrightarrow{p} \frac{1}{\beta} = \mu$

And

$\lim\limits_{n \to \infty} Pr \left(| \bar{\beta} - \beta | \right) = 1 \equiv \bar{\beta} \xrightarrow{p} \beta$

It then follows that the sequence of M.L.E.s of β, $\hat{\beta}$, is a consistent sequence of M.L.E.s of β

7.6: 14

The M.L.E., $\hat{p}$, is the proportion of butterflies in the sample with the special type of marking on their wings. Of note, this holds regardless of the sampling plan, as noted in the section *Sampling Plans* (pp.439-441) of the book. Therefore, using this method, we have:

(A): $\hat{p} = \frac{5}{43}$

And

(B): $\hat{p} = \frac{3}{58}$

Alternatively, and with greater difficulty, we can take the likelihood function using the Binomial distribution, written:

$$f_n\left(\bar{x}\mid p\right) = \binom{n}{x}p^x\left(1-p\right)^{n-x}$$

Where n denotes the total number of butterflies and x denotes the number of butterflies with the special type of marking on their wings.

Let $L\left(p\right) = log\left(f_n\left(\bar{x}\mid p\right)\right)$

Taking the derivative with respect to p gives us:

$$\frac{\partial L\left(p\right)}{\partial p} = \frac{\partial\left(log\binom{n}{x} + xlog(p) + (n-x)log(1-p)\right)}{\partial p} = \frac{x}{p} - \frac{n-x}{1-p}$$

Setting this equation equal to zero, we then solve for p:

$\frac{n-x}{1-p} = \frac{x}{p} \to p\frac{n-x}{1-p} = x \to \frac{1}{p} = \frac{n-x+x}{x} = \frac{n}{x}$

Solving for p gives us:

$p = \frac{x}{n}$, which confirms our above use of the sample proportion for the M.L.E. of p.

7.6: 23

Note, the mean is the first moment and the variance is the second moment. Their respective relations are given as follows:

(1): 1st Moment, $X_i = \frac{\alpha}{\alpha+\beta}$ (2): 2nd Moment, $X_i^2 = \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}$

We set these, X_i and X_i^2 respectively, equal to the sample moments m_1 and m_2, and then use relations (1) and (2) to solve for α and β.

Using equation (1), we have:

$(\alpha+\beta)\,m_1 = \alpha \to \beta m_1 = \alpha - \alpha m_1 \to \beta = \frac{\alpha}{m_1} - \alpha$

Substituting this into equation (2), and solving for α we have:

$\alpha\beta = (\alpha+\beta)^2\,(\alpha+\beta+1)m_2$

It then follows:

$$\frac{\alpha^2}{m_1} - \alpha^2 = \left(\alpha + \frac{\alpha}{m_1} - \alpha\right)^2\left(\alpha + \frac{\alpha}{m_1} - \alpha + 1\right)m_2$$

$$= \left(\frac{\alpha}{m_1}\right)^2\left(\frac{\alpha}{m_1} + 1\right)m_2$$

$\alpha^2\left(\frac{1}{m_1} - 1\right) = \alpha^2\left(\frac{1}{m_1^2}\right)\left(\frac{\alpha}{m_1} + 1\right)m_2$

$\to \frac{1}{m_1} - 1 = \frac{1}{m_1^2}\left(\frac{\alpha}{m_1}\right)\left(\frac{\alpha}{m_1} + 1\right)m_2$

$\to \left(\frac{1}{m_1} - 1\right)\left(\frac{m_1^2}{m_2}\right) = \frac{\alpha}{m_1} + 1$

$\to \left(\frac{m_1(1-m_1)}{m_2}\right) = \frac{\alpha}{m_1} + 1$

Thus

$\alpha = m_1\left(\frac{m_1(1-m_1)}{m_2}\right) - 1$

Using this simplification of α, we may now express β as follows:

$$\beta = \frac{\alpha}{m_1} - \alpha$$

$$= \alpha \left(\frac{1 - m_1}{m_1} \right)$$

$$= \left(\frac{\alpha}{m_1} - \alpha \right) m_1 \left(\left(\frac{m_1 (1 - m_1)}{m_2} \right) - 1 \right)$$

$$= (1 - m_1) \left(\frac{m_1 (1 - m_1)}{m_2} \right) - 1$$

(A): We thus have:

$\hat{\alpha} = m_1 \left(\frac{m_1(1-m_1)}{m_2} \right) - 1$

And

$\hat{\beta} = (1 - m_1) \left(\frac{m_1(1-m_1)}{m_2} \right) - 1$

Where our method of moments estimator for θ, $\hat{\theta} = \left(\hat{\alpha}, \hat{\beta} \right)$.

The likelihood function for the Beta distribution is:

$f_n \left(\alpha, \beta, | \bar{x} \right) = \prod\limits_{i=1}^{n} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x_i^{\alpha-1} (1 - x_i)^{\beta-1}$

Which we may rewrite as:

$f_n \left(\alpha, \beta \mid \bar{x} \right) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \prod\limits_{i=1}^{n} x_i^{\alpha-1} \prod\limits_{i=1}^{n} (1 - x_i)^{\beta-1}$

Letting $L \left(\alpha, \beta \mid \bar{x} \right) = log \left(f_n \left(\alpha, \beta, \mid \bar{x} \right) \right)$, we thus have:

$L \left(\alpha, \beta \mid \bar{x} \right) = nlog(\Gamma(\alpha + \beta)) - nlog(\Gamma(\alpha)) - nlog(\Gamma(\beta)) + (\alpha - 1) \sum\limits_{i=1}^{n} log(x_i) + (\beta - 1) \sum\limits_{i=1}^{n} log(1 - x_i)$

We would then need to derive with respect to α and with respect to β. However, and I will freely admit this is a bit *hand-wavey*, we may instead note that the M.L.E. of θ involves the products $\prod\limits_{i=1}^{n} x_i$ and $\prod\limits_{i=1}^{n} (1 - x_i)$. By comparison, the method of moments estimator for θ noted in (A) involves the *sum* of these values (noting the definition of the method of moments estimator). As such, we may note that the method of moments estimator for θ and the M.L.E. of θ are not equal.

As a very fun side note, there is no closed–form solution to the system of equations noted above (for the derivatives of the Log-likelihood function with respect to α and β). This is also grounds for proof that the method of moments estimator and the M.L.E. are not equal as it would (hopefully) simplify to the equations noted as the estimators in (A) if the two were equal.