STA 522/Solutions to Homework 6

Problem 7.1

Given an x, the MLE is simply the value of θ that maximizes the likelihood. The MLEs corresponding to $x = 0, 1, 2, 3, 4$ are $\hat{\theta} = 1, 1, (2 \text{ or } 3), 3, 3$ respectively.

Problem 7.2

Part (a): Since α is known, the likelihood function for β is given by:

$$L(\beta \mid \underline{x}) = \prod_{i=1}^{n} \frac{1}{\Gamma(\alpha)\beta^{\alpha}} \; x_i^{\alpha-1} \; e^{-x_i/\beta} = \frac{1}{\Gamma(\alpha)^n \beta^{n\alpha}} \; \left(\prod_{i=1}^{n} x_i^{\alpha-1} \right) \; \exp\left[-\frac{1}{\beta} \sum_{i=1}^{n} x_i \right]$$

The log-likelihood function is given by:

$$\log L(\beta \mid \underline{x}) = -n log \Gamma(\alpha) - n\alpha \log \beta + (\alpha - 1) \sum_{i=1}^{n} \log x_i - \frac{1}{\beta} \sum_{i=1}^{n} x_i$$

To maximize $\log L(\beta \mid \underline{x})$ we consider the first derivative test:

$$\frac{\partial \; \log L(\beta \mid \underline{x})}{\partial \beta} = -\frac{n\alpha}{\beta} + \frac{1}{\beta^2} \sum_{i=1}^{n} x_i \gtreqless 0 \iff \frac{1}{\beta^2} \sum_{i=1}^{n} x_i \gtreqless \frac{n\alpha}{\beta} \iff \beta \lesseqgtr \frac{1}{n\alpha} \sum_{i=1}^{n} x_i = \frac{1}{\alpha} \overline{x}$$

This shows that $\hat{\beta} = \frac{1}{\alpha} \overline{X}$ is maximum likelihood estimator for α.

Part (b): We shall consider successive optimization. The log-likelihood function for (α, β) is given by (same as in part (a); only α is also unknown here)

$$\log L(\alpha, \beta \mid \underline{x}) = -n \log \Gamma(\alpha) - n\alpha \log \beta + (\alpha - 1) \sum_{i=1}^{n} \log x_i - \frac{1}{\beta} \sum_{i=1}^{n} x_i$$

From part (a), for any α, the log-likelihood is maximized when $\beta = \hat{\beta} = \overline{x}/\alpha$. Plugging $\hat{\beta}$ into $\log L(\alpha, \beta \mid \underline{x})$ we get the following profile log-likelihood for α:

$$\log \tilde{L}(\alpha \mid \underline{x}) = \log L(\alpha, \hat{\beta} \mid \underline{x}) = -n \log \Gamma(\alpha) - n\alpha \log(\overline{x}/\alpha) + (\alpha - 1) \sum_{i=1}^{n} \log x_i - n\alpha$$

The MLE $\hat{\alpha}$ of α is obtained by numerically maximizing the above log-likelihood. The corresponding MLE of β is $\tilde{\beta} = \overline{x}/\hat{\alpha}$.

Problem 7.3

Given the data $\underline{x}$, let $\hat{\theta}$ be the MLE of $\theta \in \Theta$, where Θ denotes the parameter space. Then

$$\begin{aligned} & L(\hat{\theta} \mid \underline{x}) \geq L(\theta^* \mid \underline{x}) \text{ for all } \theta^* \in \Theta && (\hat{\theta} \text{ is MLE}) \\ \iff & \log L(\hat{\theta} \mid \underline{x}) \geq \log L(\theta^* \mid \underline{x}) \text{ for all } \theta^* \in \Theta && (\log \text{ is an increasing function}) \end{aligned}$$

This completes the proof.

Problem 7.7

First find the likelihood function of θ. Here $\theta \in \Theta = \{0, 1\}$, with

$$L(\theta = 0 \mid \underline{x}) = \prod_{i=1}^{n} I(0 < x_i < 1) = I(0 < x_{(1)} < x_{(n)} < 1)$$

and

$$L(\theta = 1 \mid \underline{x}) = \prod_{i=1}^{n} \frac{1}{2\sqrt{x_i}} \; I(0 < x_i < 1) = \frac{1}{2^n \prod_{i=1}^{n} \sqrt{x_i}} \; I(0 < x_{(1)} < x_{(n)} < 1)$$

Therefore

$$\frac{L(\theta = 1 \mid \underline{x})}{L(\theta = 0 \mid \underline{x})} = \frac{1}{2^n \prod_{i=1}^{n} \sqrt{x_i}} \gtreqless 1 \iff 1 \gtreqless 2^n \prod_{i=1}^{n} \sqrt{x_i}$$

Thus, the MLE of θ is

$$\hat{\theta} = \begin{cases} 1 & \text{if } \; 2^n \prod_{i=1}^{n} \sqrt{x_i} < 1 \\ 0 & \text{if } \; 2^n \prod_{i=1}^{n} \sqrt{x_i} > 1 \\ 0 \text{ or } 1 & \text{if } \; 2^n \prod_{i=1}^{n} \sqrt{x_i} = 1 \end{cases}$$

Problem 7.9

Here $X_1, X_2, \ldots, X_n \sim$ iid Uniform(θ), $\theta > 0$. We have seen in class (lecture 5 & 6) that the method of moments and the method of maximum likelihood estimators of θ are $\hat{\theta}_{MM} = 2\overline{X}$ and $\hat{\theta} = X_{(n)}$ respectively. We shall compare the two estimators using their mean squared errors.

For $\hat{\theta}_{MM}$ we have

$$\mathrm{E}_\theta(\hat{\theta}_{MM}) = 2\,\mathrm{E}_\theta(\overline{X}) \overset{\text{iid}}{=} 2\,\mathrm{E}(X_1) = 2\frac{\theta}{2} = \theta \quad \text{for all } \theta$$

i.e., $\hat{\theta}_{MM}$ is unbiased for θ. Hence,

$$\mathrm{MSE}_\theta(\hat{\theta}_{MM}) = \mathrm{Var}_\theta(\hat{\theta}_{MM}) = 4\,\mathrm{Var}_\theta(\overline{X}) \overset{\text{iid}}{=} \frac{4}{n}\,\mathrm{Var}_\theta(X_1) = \frac{4}{n} \cdot \frac{\theta^2}{12} = \frac{\theta^2}{3n}.$$

For $\hat{\theta}$, we have (from lecture 7)

$$\mathrm{E}_\theta(\hat{\theta}) = \frac{n}{n+1}\theta \quad \text{and} \quad \mathrm{Var}_\theta(\hat{\theta}) = \left(\frac{n}{n+1}\right)^2 \mathrm{Var}_\theta\left(\frac{n}{n+1}\; X_{(n)}\right) = \frac{n}{(n+1)^2(n+2)}\theta^2$$

Therefore,

$$\begin{aligned} \mathrm{MSE}_\theta(\hat{\theta}) &= \left(\mathrm{Bias}_\theta(\hat{\theta})\right)^2 + \mathrm{Var}_\theta(\hat{\theta}) \\ &= \left(\frac{n}{n+1}\theta - \theta\right)^2 + \frac{n}{(n+1)^2(n+2)}\theta^2 \\ &= \frac{1}{(n+1)^2}\theta^2 + \frac{n}{(n+1)^2(n+2)}\theta^2 \\ &= \frac{2n+2}{(n+1)^2(n+2)}\theta^2 = \frac{2\theta^2}{(n+1)(n+2)} \end{aligned}$$

Thus,

$$\begin{aligned} \mathrm{MSE}_\theta(\hat{\theta}_{MM}) - \mathrm{MSE}_\theta(\hat{\theta}) &= \frac{\theta^2}{3n} - \frac{2\theta^2}{(n+1)(n+2)} \\ &= \frac{n^2 - 3n + 2}{3n(n+1)(n+2)}\;\theta^2 = \frac{(n-1)(n-2)}{3n(n+1)(n+2)}\;\theta^2 \end{aligned}$$

Hence $\mathrm{MSE}_\theta(\hat{\theta}_{MM}) = \mathrm{MSE}_\theta(\hat{\theta})$ for all θ when $n = 1, 2$ and $\mathrm{MSE}_\theta(\hat{\theta}_{MM}) > \mathrm{MSE}_\theta(\hat{\theta})$ for all θ for $n \geq 3$. Hence, in terms of having a smaller MSE, $\hat{\theta}$ is preferred.