

MATH 557 - EXERCISES 5 SOLUTIONS

1 In this model, if $X = \sum_{i=1}^n X_i$, the MLE of θ is $\hat{\theta}_n = \bar{X} = X/n$.

(a) This is not a 1-1 mapping, so some care is needed. Inverting the transformation yields that

$$\theta = \frac{1 \pm \sqrt{1 - 4\tau}}{2}$$

so the likelihood needs to be worked out for both cases. In the first case

$$L(\tau|\underline{x}) = \left(\frac{1 + \sqrt{1 - 4\tau}}{2} \right)^x \left(\frac{1 - \sqrt{1 - 4\tau}}{2} \right)^{n-x}$$

and in the second

$$L(\tau|\underline{x}) = \left(\frac{1 - \sqrt{1 - 4\tau}}{2} \right)^x \left(\frac{1 + \sqrt{1 - 4\tau}}{2} \right)^{n-x}$$

where $0 < \tau < 1/4$. After some manipulation it follows that in both cases $x < n - x$ or $x > n - x$,

$$\hat{\tau} = \frac{1}{4} - \left(\frac{1}{2} - \frac{x}{n} \right)^2 = \bar{x}(1 - \bar{x})$$

with the same result if $x = n - x$, and so the estimator is

$$\hat{\tau}_n(\underline{X}) = \bar{X}(1 - \bar{X})$$

so in fact the non 1-1 nature of the reparameterization is not problematic.

(b) Despite the non 1-1 nature of the reparameterization, the Delta method can still be used. If $g(t) = t(1 - t)$, then $\dot{g}(t) = 1 - 2t$, which is non zero if $t \neq 1/2$. Thus for $\theta \neq 1/2$, by the CLT

$$\sqrt{n}(\bar{X} - \theta) \xrightarrow{d} Z \sim \text{Normal}(0, \theta(1 - \theta))$$

and thus by the Delta method

$$\sqrt{n}(\hat{\tau}_n(\underline{X}) - \tau) \xrightarrow{d} Z \sim \text{Normal}(0, \tau(1 - 2\theta)^2)$$

or, for large n

$$\hat{\tau}_n(\underline{X}) \sim \text{AN}(\tau, \tau(1 - 2\theta)^2/n)$$

If $\theta = 1/2$, this approximation yields a degenerate limiting distribution, so a second order Delta method must be used. We have that $\ddot{g}(t) = -2$, so from 556,

$$n(\hat{\tau}_n(\underline{X}) - 1/4) \xrightarrow{d} -\frac{1}{4}Q \quad \text{where} \quad Q \sim \chi_1^2$$

or, for large n

$$\hat{\tau}_n(\underline{X}) \dot{\sim} \frac{1}{4} - \frac{1}{4}\text{Gamma}(1/4, 2n)$$

2 Note first that

$$E_{f_{T_{1n}}}[T_{1n}] = \frac{1}{n} \sum_{i=1}^n E_{f_{\mathcal{X}}}[X_i^2] - 1 = \frac{1}{n} \sum_{i=1}^n (\mu^2 + 1) - 1 = \mu^2$$

and

$$\begin{aligned} E_{f_{T_{2n}}}[T_{2n}] &= \frac{1}{n^2} E_{f_{\mathcal{X}}} \left[\left(\sum_{i=1}^n X_i \right)^2 \right] - \frac{1}{n} \\ &= \frac{1}{n^2} \sum_{i=1}^n E_{f_{\mathcal{X}}}[X_i^2] + \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n E_{f_{\mathcal{X}}}[X_i X_j] - \frac{1}{n} \\ &= \frac{1}{n^2} n(\mu^2 + 1) + 0 - \frac{1}{n} = \mu^2. \end{aligned}$$

so both statistics are unbiased (and hence asymptotically unbiased). Thus the ARE is the ratio of the asymptotic variances. For T_{1n} , we know by the CLT that

$$\sqrt{n} \left(\frac{1}{n} \sum_{i=1}^n X_i^2 - \theta \right) \xrightarrow{d} Z \sim \text{Normal}(0, \gamma)$$

where $\theta = \mu^2 + 1$, and $\gamma = \text{Var}_{f_{\mathcal{X}}}[X_i^2]$, so the asymptotic variance of T_{1n} is γ . For T_{2n} , we have by the CLT

$$\sqrt{n}(\bar{X} - \mu) \xrightarrow{d} Z \sim \text{Normal}(0, 1)$$

so by the Delta Method, for $\mu \neq 0$

$$\sqrt{n}(\bar{X}^2 - \mu^2) \xrightarrow{d} Z \sim \text{Normal}(0, 4\mu^2)$$

and for $\mu = 0$, as $\sqrt{n}\bar{X} \xrightarrow{d} \text{Normal}(0, 1)$,

$$n\bar{X}^2 \xrightarrow{d} Q \sim \chi_1^2$$

where Q has variance 2. Hence

$$\text{ARE}_{\mu}(T_{1n}, T_{2n}) = \begin{cases} \frac{4\mu^2}{\gamma} & \mu \neq 0 \\ \frac{2}{\gamma} & \mu = 0 \end{cases}$$

Note however that there is a different rescaling in the $\mu = 0$ case. So in terms of large sample comparison,

$$T_{1n} \asymp \text{Normal}(0, \gamma/n) \quad T_{2n} \asymp \text{Gamma}(1/2, n/2)$$

yielding a large sample variance ratio of $\frac{2}{n\gamma}$.

3 In this case, $X \sim \text{Exp}(\phi)$ and $Y \sim \text{Exp}(\theta\phi)$, so if $\lambda_1 = \phi$ and $\lambda_2 = \theta\phi$, the MLEs are $\hat{\lambda}_1 = 1/\bar{X}$ and $\hat{\lambda}_2 = 1/\bar{Y}$, so that by invariance

$$\hat{\phi} = \hat{\lambda}_1 = 1/\bar{X} \quad \hat{\theta} = \hat{\lambda}_2/\hat{\lambda}_1 = \bar{X}/\bar{Y}$$

Thus

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} \mathcal{Z} \sim \text{Normal}(\underline{0}, \mathcal{I}(\theta)^{-1})$$

where $\mathcal{I}(\theta)$ is the Fisher Information. We have

$$L(\theta, \phi|x, y) = \phi^2 \theta \exp\{-[\phi x + \theta \phi y]\} \quad x, y > 0$$

so that

$$l(\theta, \phi|x, y) = \log L(\theta, \phi) = 2 \log \phi + \log \theta - (\phi x + \theta \phi y)$$

and

$$\begin{aligned} \frac{\partial l}{\partial \phi} &= \frac{2}{\phi} - x - \theta y & \frac{\partial l}{\partial \theta} &= \frac{1}{\theta} - \phi y \\ \frac{\partial^2 l}{\partial \phi^2} &= -\frac{2}{\phi^2} & \frac{\partial^2 l}{\partial \theta^2} &= -\frac{1}{\theta^2} & \frac{\partial^2 l}{\partial \phi \partial \theta} &= -y \end{aligned}$$

yielding the matrix $\Psi(\theta, \phi, X, Y)$ and Fisher information

$$\Psi(\theta, \phi, X, Y) = \begin{bmatrix} \frac{2}{\phi^2} & Y \\ Y & \frac{1}{\theta^2} \end{bmatrix} \quad \mathcal{I}(\theta, \phi) = \begin{bmatrix} \frac{1}{\theta^2} & \frac{1}{\phi\theta} \\ \frac{1}{\phi\theta} & \frac{2}{\phi^2} \end{bmatrix}$$

as $E[Y] = 1/(\phi\theta)$. Thus

$$\mathcal{I}(\theta, \phi)^{-1} = \begin{bmatrix} 2\theta^2 & -\phi\theta \\ -\phi\theta & \phi^2 \end{bmatrix}$$

Thus

$$\sqrt{n} \begin{pmatrix} \hat{\theta} - \theta \\ \hat{\phi} - \phi \end{pmatrix} \xrightarrow{d} Z \sim \text{Normal} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{bmatrix} 2\theta^2 & -\phi\theta \\ -\phi\theta & \phi^2 \end{bmatrix} \right)$$

or, for large n ,

$$\begin{pmatrix} \hat{\theta} \\ \hat{\phi} \end{pmatrix} \sim \text{AN} \left(\begin{pmatrix} \theta \\ \phi \end{pmatrix}, \begin{bmatrix} 2\theta^2 & -\phi\theta \\ -\phi\theta & \phi^2 \end{bmatrix} \right)$$

4 For the Poisson case, for $\lambda > 0$

$$f_{X|\lambda}(x|\lambda) = \frac{e^{-\lambda} \lambda^x}{x!} \quad x = 0, 1, 2, \dots$$

$$l_n(\lambda) = -n\lambda + t_n \log \lambda - \sum_{i=1}^n \log x_i! \quad \text{where } t_n = \sum_{i=1}^n x_i$$

$$\dot{l}_n(\lambda) = -n + \frac{t_n}{\lambda} \quad \ddot{l}_n(\lambda) = -\frac{t_n}{\lambda^2}$$

and hence the MLE, from $\dot{l}_n(\hat{\lambda}_n) = 0$, is $\hat{\lambda}_n = t_n/n = \bar{x}$, with estimator $T_n/n = \bar{X}$. Using the formulae from notes

$$W_n = n(\tilde{\theta}_n - \theta_0)^\top \hat{I}_n(\tilde{\theta}_n)(\tilde{\theta}_n - \theta_0) \quad (\text{WALD})$$

and if

$$Z_n \equiv Z_n(\theta_0) = \frac{1}{\sqrt{n}} \dot{l}_n(\theta_0).$$

then

$$R_n = Z_n^\top \mathcal{I}(\theta_0)^{-1} Z_n \quad (\text{SCORE})$$

we have, in the 1-d case

$$W_n = (\tilde{\theta}_n - \theta_0)^2 (n\hat{l}_n(\tilde{\theta}_n)) = -(\tilde{\theta}_n - \theta_0)^2 \ddot{l}_n(\tilde{\theta}_n)$$

and

$$R_n = \frac{\{Z_n(\theta_0)\}^2}{\mathcal{I}(\theta_0)}$$

Thus we have

- **Wald Statistic:**

$$W_n = -(\tilde{\theta}_n - \theta_0)^2 \ddot{l}_n(\tilde{\theta}_n) = -(\bar{X} - \lambda_0)^2 \left(\frac{-S_n}{(\bar{X})^2} \right) = n \frac{(\bar{X} - \lambda_0)^2}{\bar{X}} = \frac{n(\hat{\lambda}_n - \lambda_0)^2}{\hat{\lambda}_n}$$

- **Rao Statistic:** in this case, we can compute the Fisher Information $\mathcal{I}(\lambda_0)$ exactly - we have

$$\mathcal{I}(\lambda_0) = \mathbb{E}_{f_{X|\lambda_0}} [-\Psi(\lambda_0, X)] = \mathbb{E}_{f_{X|\lambda_0}} \left[\frac{X}{\lambda_0^2} \right] = \frac{1}{\lambda_0^2} \mathbb{E}_{f_{X|\lambda_0}} [X] = \frac{\lambda_0}{\lambda_0^2} = \frac{1}{\lambda_0}$$

so therefore

$$R_n = \frac{\{Z_n(\theta_0)\}^2}{\mathcal{I}(\theta_0)} = \frac{\left(\frac{1}{\sqrt{n}} \left(\frac{T_n}{\lambda_0} - n \right)^2 \right)}{\frac{1}{\lambda_0}} = \frac{\lambda_0}{n} \left(\frac{T_n}{\lambda_0} - n \right)^2 = \frac{n(\bar{X} - \lambda_0)^2}{\lambda_0}$$

If the Fisher Information can be computed exactly, then the exact version should be used for the Score statistic rather than an estimated version. If the Fisher information cannot be obtained, we might instead use the observed information. Here that would imply that

$$R_n = - \left\{ \dot{l}_n(\theta_0) \right\}^2 \left\{ \ddot{l}_n(\theta_0) \right\}^{-1} = \frac{- \left(\frac{T_n}{\lambda_0} - n \right)^2}{-\frac{T_n}{\lambda_0^2}} = \frac{(T_n - n\lambda_0)^2}{T_n} = \frac{n(\bar{X} - \lambda_0)^2}{\bar{X}}$$

that is, identical to Wald.

- **Likelihood Ratio Statistic:** by definition, using the notation Λ_n for the likelihood ratio

$$\Lambda_n = \frac{L(\lambda_0)}{L(\hat{\lambda}_n)} = \frac{e^{-n\lambda_0} \lambda_0^{T_n}}{e^{-n\hat{\lambda}_n} \hat{\lambda}_n^{T_n}} = \exp \left\{ n(\hat{\lambda}_n - \lambda_0) - T_n(\log \hat{\lambda}_n - \log \lambda_0) \right\}$$

or equivalently

$$-2 \log \Lambda_n = -2n(\hat{\lambda}_n - \lambda_0) + 2T_n(\log \hat{\lambda}_n - \log \lambda_0)$$

For a $1 - \alpha$ confidence interval, we utilize the result that each of the test statistics has an approximate χ_1^2 distribution as $n \rightarrow \infty$. For W_n and R_n , we have

$$\left\{ \lambda : n(\hat{\lambda}_n - \lambda)^2 / \hat{\lambda}_n \leq c_{1-\alpha} \right\} \quad \text{and} \quad \left\{ \lambda : n(\hat{\lambda}_n - \lambda)^2 / \lambda \leq c_{1-\alpha} \right\}$$

respectively, where $c_{1-\alpha}$ is the $1 - \alpha$ quantile of the χ_1^2 distribution. For the LRT, we have

$$\left\{ \lambda : -2n(\hat{\lambda}_n - \lambda) + 2t_n(\log \hat{\lambda}_n - \log \lambda) \leq c_{1-\alpha} \right\}$$