Homework 10 Solutions

(1) Let $\theta = (a, b)$. The density is

$$f(x;\,\theta) = \begin{cases} \frac{1}{b-a} & \text{if } a < x < b \\ 0 & \text{otherwise.} \end{cases}$$

The likelihood function is

$$\mathcal{L}(\theta) = \prod_i f(X_i;\,\theta).$$

If $a < X_i < b$ then the i^{th} observation contributes a factor of $1/(a-b)$ otherwise it makes the product 0. Hence,

$$\mathcal{L}(\theta) = \begin{cases} \left(\frac{1}{b-a}\right)^n & \text{if } a < X_{(1)},\ X_{(n)} < b \\ 0 & \text{otherwise} \end{cases}$$

where $X_{(1)}$ is the smallest data point and $X_{(n)}$ is the largest data point. The likelihood is maximized by taking

$$\hat{a} = X_{(1)},\ \hat{b} = X_{(n)}.$$

(1b) τ is the mean, which, for a $\text{Uniform}(a, b)$ is $\tau = (a+b)/2$. The mle is $\hat{\tau} = (X_{(1)} + X_{(n)})/2$.

(1c) The nonparametric plug-in estimator is $\tilde{\tau} = n^{-1}\sum_i X_i$. The MSE is $MSE(\tilde{\tau}) = \text{bias}^2 + \text{Var} = 0 + (b-a)^2/(12n) = (b-a)^2/(12n)$. With $b = 3, a = 1, n = 10$ this is is .0333. By simulation, The mle has MSE about .015, substantially smaller than the nonparametric plug-in.

(2) We need the Fisher information matrix for μ and σ. Now

$$f(x;\,\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\}.$$

And

$$\begin{bmatrix} \frac{\partial^2 \log f}{\partial \mu^2} & \frac{\partial^2 \log f}{\partial \mu \partial \sigma} \\ \frac{\partial^2 \log f}{\partial \sigma \partial \mu} & \frac{\partial^2 \log f}{\partial \sigma^2} \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sigma^2} & -\frac{2(X-\mu)}{\sigma^3} \\ -\frac{2(X-\mu)}{\sigma^3} & \frac{1}{\sigma^2} - \frac{3(X-\mu)^2}{\sigma^4} \end{bmatrix}.$$

Since $E(X-\mu)=0$ and $E(X-\mu)^2=\sigma^2$, we have

$$I(\mu,\sigma)=\begin{bmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{2}{\sigma^2} \end{bmatrix}.$$

The inverse Fisher information is

$$I^{-1}(\mu,\sigma)=\begin{bmatrix} \sigma^2 & 0 \\ 0 & \frac{\sigma^2}{2} \end{bmatrix}.$$

(3) $P(X<\tau)=P(Z<(\tau-\mu)/\sigma)=\Phi((\tau-\mu)/\sigma)=.95$. Solving for τ we get $\tau=\mu+\sigma\Phi^{-1}(.05)\equiv g(\mu,\tau)$. The mle is $\hat{\tau}=\hat{\mu}+\hat{\sigma}\Phi^{-1}(.05)$ where $\hat{\mu}=n^{-1}\sum_i X_i$ and $\hat{\sigma^2}=n^{-1}\sum_i(X_i-\hat{\mu})^2$. The gradient of g

$$\nabla g=\begin{pmatrix} 1 \\ \Phi^{-1}(.05) \end{pmatrix}.$$

The asymptotic standard error of $\hat{\tau}$ is

$$se=\sqrt{\frac{(\nabla g)^T I^{-1}(\nabla g)}{n}}=\frac{\sigma}{\sqrt{n}}\sqrt{1+\frac{\Phi^{-1}(.05)^2}{2}}.$$

The estimates standard error is

$$\hat{se}=\frac{\hat{\sigma}}{\sqrt{n}}\sqrt{1+\frac{\Phi^{-1}(.05)^2}{2}}.$$

An approximate $1-\alpha$ confidence interval is

$$\hat{\tau}\pm z_{\alpha/2}\hat{se}.$$

(4) The mle is $\hat{\theta}=X_{(n)}$, the maximum data point. Note that $\hat{\theta}\le\theta$. Hence,

$$\begin{aligned} P(|\hat{\theta}-\theta|>\epsilon) &= P(\hat{\theta}<\theta-\epsilon) \\ &= P(X_1<\theta-\epsilon)^n \\ &= \left(\frac{\theta-\epsilon}{\theta}\right)^n \\ &= \left(1-\frac{\epsilon}{\theta}\right)^n \\ &\to 0 \end{aligned}$$

as $n \to \infty$.

(5) $f(x;\, \lambda) = \lambda^x e^{-\lambda}$ so

$$\frac{\partial \log f}{\partial \lambda} = \frac{x}{\lambda} - 1$$

and

$$\frac{\partial^2 f}{\partial \lambda^2} = -\frac{x}{\lambda^2}.$$

Thus,

$$I(\lambda) = E\left(\frac{x}{\lambda^2}\right) = \frac{1}{\lambda}.$$

(6a) $\psi = P(Y_1 = 1) = P(X_1 > 0) = P(X_1 - \theta > -\theta) = P(Z > -\theta) = 1 - P(Z < -\theta) = 1 - \Phi(-\theta)$. The mle is $\hat{\psi} = 1 - \Phi(-\hat{\theta}) = 1 - \Phi(-\overline{X})$.

(b) Let $g(\theta) = 1 - \Phi(-\theta) = \Phi(\theta)$. Then, $g'(\theta) = \phi(\theta)$. The estimated standard error of ψ is $\hat{se} = \hat{se}(\hat{\theta})|g'(\hat{\theta})| = \phi(\hat{\theta})/\sqrt{n} = \phi(\overline{X})/\sqrt{n}$. An approximate 95 per cent confidence interval is

$$1 - \Phi(-\overline{X}) \pm 2\frac{\phi(\overline{X})}{\sqrt{n}}.$$

(c) $\widetilde{\psi}$ has mean $E(Y_1) = \psi$. Consistency follows from the weak law of large numbers.

(d) Note that $Y_1 \sim \text{Bernoulli}(\psi)$ so $\text{Var}(Y_1) = \psi(1-\psi)$ and $\text{Var}(\widetilde{\psi}) = \text{Var}(Y_1)/n = \psi(1-\psi)/n$. The ARE is

$$\frac{\psi(1-\psi)}{\phi(\theta)} = \frac{\Phi(\theta)(1-\Phi(\theta))}{\phi(\theta)}.$$

(e) By the law of large numbers, $\overline{X}$ converges in probability to $E(X_1) \equiv \mu$. So $\hat{\psi} = 1 - \Phi(-\overline{X})$ converges in probability to $1 - \Phi(-\mu) = \Phi(\mu)$. The true value of ψ is $P(X > 0) = 1 - P(X < 0) = 1 - F_X(0)$. For an arbitrary distribution F_X, we have $1 - F_X(0) \neq \Phi(\mu)$ so the mle is inconsistent. On the other hand, $\widetilde{\psi}$ is still consistent.