Homework 8: Partial Solutions

Chapter 10, Problem 1. $\mathbb{E}(X) = \alpha\beta$ and $\mathbb{E}(X^2) = \mathbb{V}(X) + (\mathbb{E}(X))^2 = \alpha\beta^2 + (\alpha\beta)^2$ and hence we solve

$$\begin{aligned} \widehat{\alpha}\widehat{\beta} &= \frac{1}{n}\sum_{i=1}^{n} X_i \\ \widehat{\alpha}\widehat{\beta}^2 + \widehat{\alpha}^2\widehat{\beta}^2 &= \frac{1}{n}\sum_{i=1}^{n} X_i^2 \end{aligned}$$

which gives

$$\widehat{\alpha} = \frac{\overline{X}^2}{S^2} \quad \text{and} \quad \widehat{\beta} = \frac{S^2}{\overline{X}}$$

where $S^2 = \sum_{i=1}^{n}(X_i - \overline{X})^2/n$.

Chapter 10, Problem 2. $\mathbb{E}(X) = (a+b)/2$ and $\mathbb{E}(X^2) = (b-a)^2/12 + ((a+b)/2)^2 = (b^2 + ba + a^2)/3$ and hence we solve

$$\begin{aligned} \frac{\widehat{a} + \widehat{b}}{2} &= \frac{1}{n}\sum_{i=1}^{n} X_i \\ \frac{\widehat{b}^2 + \widehat{b}\widehat{a} + \widehat{a}^2}{3} &= \frac{1}{n}\sum_{i=1}^{n} X_i^2. \end{aligned}$$

Let $X_{(1)} = \min\{X_1, \ldots, X_n\}$ and $X_{(n)} = \max\{X_1, \ldots, X_n\}$. The likelihood is

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

This is maximized by

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

Now $\tau = \int x dF(x) = \mathbb{E}(X) = (a+b)/2$. The mle is $\widehat{\tau} = (\widehat{a} + \widehat{b})/2 = (X_{(1)} + X_{(2)})/2$.

The nonparametric plug-in estimator is $\widetilde{\tau} = n^{-1}\sum_i X_i$. The MSE is $\mathsf{MSE}(\widetilde{\tau}) = \mathsf{bias}^2 + \mathbb{V} = 0 + (b-a)^2/(12n) = (b-a)^2/(12n)$. 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.

Chapter 10, Problem 3. $\mathbb{P}(X < \tau) = \mathbb{P}(Z < (\tau - \mu)/\sigma) = \Phi((\tau - \mu)/\sigma) = .95$. Solving for τ we get $\tau = \mu + \sigma\Phi^{-1}(.05) \equiv g(\mu, \tau)$. The mle is $\widehat{\tau} = \widehat{\mu} + \widehat{\sigma}\Phi^{-1}(.05)$ where $\widehat{\mu} = n^{-1}\sum_{i=1}^n X_i$ and $\widehat{\sigma^2} = n^{-1}\sum_{i=1}^n (X_i - \widehat{\mu})^2$. The gradient of g

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

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

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

The estimates standard error is

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

An approximate $1 - \alpha$ confidence interval is

$$\widehat{\tau} \pm z_{\alpha/2}\widehat{\mathsf{se}}\,.$$

Chapter 10, Problem 4. The mle is $\widehat{\theta} = X_{(n)}$, the maximum data point. Note that $\widehat{\theta} \le \theta$. Hence,

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

$$
\begin{aligned}
&= \left(1 - \frac{\epsilon}{\theta}\right)^n \\
&\to 0
\end{aligned}
$$

as $n \to \infty$.

Chapter 10, Problem 5. $\mathbb{E}(X) = \lambda$ so the method of moments estimator is $\widehat{\lambda} = n^{-1}\sum_{i=1}^n X_i$. The likelihood is $\lambda^{\sum_i X_i} e^{-n\lambda}$, the log-likelihood is $\ell(\lambda) = \sum_i X_i \log\lambda - n\lambda$. The mle is obtained by setting $\ell'(\lambda) = 0$ yielding $\widehat{\lambda} = n^{-1}\sum_{i=1}^n X_i$. Now, $f(x;\lambda) = \lambda^x e^{-\lambda}$ so

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

and

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

Thus,

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

Chapter 10, Problem 6. (a) $\psi = \mathbb{P}(Y_1 = 1) = \mathbb{P}(X_1 > 0) = \mathbb{P}(X_1 - \theta > -\theta) = \mathbb{P}(Z > -\theta) = 1 - \mathbb{P}(Z < -\theta) = 1 - \Phi(-\theta)$. The mle is $\widehat{\psi} = 1 - \Phi(-\widehat{\theta}) = 1 - \Phi(-\overline{X})$.

(b) Let $g(\theta) = 1 - \Phi(-\theta) = \Phi(\theta)$. Then, $g'(\theta) = \phi(\theta)$. The estimated standard error of ψ is $\widehat{\mathsf{se}} = \widehat{\mathsf{se}}\,(\widehat{\theta})|g'(\widehat{\theta})| = \phi(\widehat{\theta})/\sqrt{n} = \phi(\overline{X})/\sqrt{n}$. 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 $\mathbb{E}(Y_1) = \psi$. Consistency follows from the weak law of large numbers.

(d) Note that $Y_1 \sim \text{Bernoulli}(\psi)$ so $\mathbb{V}(Y_1) = \psi(1-\psi)$ and $\mathbb{V}(\widetilde{\psi}) = \mathbb{V}(Y_1)/n = \psi(1-\psi)/n$. 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 $\mathbb{E}(X_1) \equiv \mu$. So $\widehat{\psi} = 1 - \Phi(-\overline{X})$ 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.

Chapter 10, Problem 7. (a) $\widehat{\psi} = \widehat{p}_1 - \widehat{p}_2$.

(b) The likelihood is

$$L(p_1, p_2) = p_1^{X_1}(1-p_1)^{n_1 - X_1} p_2^{X_2}(1-p_2)^{n_2 - X_2}.$$

The matrix H of second derivatives is

$$H = \begin{bmatrix} -\frac{X_1}{p_1^2} - \frac{1-X_1}{(1-p_1)^2} & 0 \\ 0 & -\frac{X_2}{p_2^2} - \frac{1-X_2}{(1-p_2)^2} \end{bmatrix}.$$

Since $E(X_1) = n_1 p_1$ and $E(X_2) = n_2 p_2$, the Fisher information matrix is

$$I(p_1, p_2) = E(-H) = \begin{bmatrix} \frac{n_1}{p_1(1-p_1)} & 0 \\ 0 & \frac{n_2}{p_2(1-p_2)} \end{bmatrix}.$$

(c) $\psi = g(p_1, p_2) = p_1 - p_2$ and the gradient of g is

$$\nabla g = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

By the delta method, the estimated standard error of $\widehat{\psi}$ is

$$\widehat{se} = \sqrt{(\nabla g)^T I^{-1}(\widehat{p}_1, \widehat{p}_2)(\nabla g)} = \left\{ \frac{\widehat{p}_1(1-\widehat{p}_1)}{n_1} + \frac{\widehat{p}_2(1-\widehat{p}_2)}{n_2} \right\}^{1/2}.$$

(d) The bootstrap code is:

```
B <- 10000
tau.boot <- rep(0,B)
for(i in 1:B){
     xx1 <- rbinom(1,n1,p1.hat)
     xx2 <- rbinom(1,n2,p2.hat)
     tau.boot[i] <- (xx1/n1)-(xx2/n2)
     }
```