Homework 5: Partial Solutions

Chapter 8, Problem 1. $Y = n\widehat{F}_n(x) \sim \text{Binomial}(n,p)$ where $p = \mathbb{P}(X < x) = F(x)$. So $\mathbb{E}(\widehat{F}(x)) = \mathbb{E}(Y/n) = F(x)$ and $\mathbb{V}(\widehat{F}(x)) = n^{-2}\mathbb{V}(Y) = p(1-p)/n = F(x)(1-F(x))/n$. The MSE is $\mathsf{bias}^2 + \mathbb{V} = \mathbb{V} = F(x)(1-F(x))/n$. Hence $\text{mse} \to 0$ which implies that $\widehat{F}(x) \xrightarrow{\text{qm}} F(x)$ which implies that $\widehat{F}(x) \xrightarrow{\text{P}} (x)$.

Chapter 8, Problem 2. The quantity of interest is $\delta = p - q = \int x dF_1(x) - \int x dF_2(x)$ and the plug-in estimator is $\widehat{\delta} = \int x d\widehat{F}_1(x) - \int x d\widehat{F}_2(x) = n^{-1}\sum_i X_i - m^{-1}\sum_i Y_i \equiv \widehat{p} - \widehat{q}$. The variance of the estimator is $\mathbb{V}(\widehat{\delta}) = \mathbb{V}(hatp - \widehat{q}) = \mathbb{V}(hatp) + \mathbb{V}(\widehat{q}) = p(1-p)/n + q(1-q)/m$ and the estimated standard error is

$$\widehat{\mathsf{se}} = \sqrt{\frac{\widehat{p}(1-\widehat{p})}{n} + \frac{\widehat{q}(1-\widehat{q})}{m}}.$$

Since $z_{\alpha/2} = z_{.1/2} = z_{.05} = 1.64$, an approximate 90 per cent interval is

$$\widehat{p} - \widehat{q} \pm 1.64\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n} + \frac{\widehat{q}(1-\widehat{q})}{m}}.$$

Chapter 8, Problem 3. We can write

$$\widehat{F}(x) = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

where $Y_i \sim \text{Bernoulli}(p)$ with $\mathbb{E}(Y_i) = p = F(x)$ and $\mathbb{V}(Y_i) = p(1-p) = F(x)(1-F(x))$. Hence,

$$\frac{\sqrt{n}(\widehat{F}(x) - F(x))}{F(x)(1-F(x))} \rightsquigarrow N(0,1).$$

6.

$$\mathbb{E}(\widehat{\lambda}) = \mathbb{E}(X_1) = \lambda$$

and hence $\mathsf{bias} = 0$.

$$\mathbb{V}(\widehat{\lambda}) = \frac{\mathbb{V}(X_1)}{n} = \frac{\lambda}{n}$$

and se $\sqrt{\lambda/n}$

7. The CDF G of $\widehat{\theta}$ is

$$\begin{aligned} G(\widehat{\theta}) &= \mathbb{P}(\widehat{\Theta} \leq \widehat{\theta}) \\ &= \mathbb{P}(\max\{X_1, \ldots, X_n\} \leq \widehat{\theta}) \\ &= \prod_{i=1}^{n} \mathbb{P}(X_i \leq \widehat{\theta}) \\ &= F_\theta(\widehat{\theta})^n \\ &= \left(\frac{\widehat{\theta}}{\theta}\right)^n. \end{aligned}$$

The density is therefore

$$g(\widehat{\theta}) = \left(\frac{n}{\theta}\right)\left(\frac{\widehat{\theta}}{\theta}\right)^{n-1}.$$

Thus,

$$\mathbb{E}_\theta(\widehat{\theta}) = \int_0^\theta \widehat{\theta} g(\widehat{\theta}) d\widehat{\theta} = \frac{n\theta}{n+1}$$

and

$$\mathsf{bias} = \frac{n\theta}{n+1} - \theta = -\frac{\theta}{n+1}.$$

Also,

$$\mathbb{E}_\theta(\widehat{\theta}^2) = \int_0^\theta \widehat{\theta}^2 g(\widehat{\theta}) d\widehat{\theta} = \frac{n\theta^2}{n+2}$$

and so

$$\mathbb{V}_\theta(\widehat{\theta}) = \frac{n\theta^2}{n+2} - \left(\frac{n\theta}{n+1}\right)^2 = \frac{n\theta^2}{(n+2)(n+1)^2}.$$

The mse is

$$\mathsf{bias}^2 + \mathbb{V} = \left(\frac{\theta}{n+1}\right)^2 + \frac{n\theta^2}{(n+2)(n+1)^2} = \frac{2\theta^2}{(n+1)(n+2)}.$$

8. Recall that $\mathbb{E}(X_i) = \theta/2$, $\mathbb{V}(X_i) = \theta^2/12$. So

$$\mathbb{E}_\theta(2\overline{X}) = 2\mathbb{E}_\theta(\overline{X}) = 2\frac{\theta}{2} = \theta$$

and hence $\mathsf{bias} = 0$. Now

$$\mathbb{V}_\theta(2\overline{X}) = 4\mathbb{V}_\theta(\overline{X}) = \frac{4\sigma^2}{n} = \frac{4\theta^2}{12n} = \frac{\theta^2}{3n}.$$

Since this estimator is unbiased,

$$\text{mse} = \mathbb{V}_\theta(\widehat{\theta}) = \frac{\theta^2}{3n}.$$

9. $\mu = \mathbb{E}(X_i) = 1/2$ and $\sigma^2 = \mathbb{V}(X_i) = 1/12$. By the CLT,

$$\frac{\sqrt{n}(\overline{X} - \mu)}{\sigma} = \sqrt{12n}\left(\overline{X} - \frac{1}{2}\right) \rightsquigarrow N(0,1).$$

Now $Y = g(\overline{X})$ where $g(s) = s^2$. And $g'(s) = 2s$ and $g'(\mu) = g'(1/2) = 2(1/2) = 1$. From the delta method,

$$\frac{\sqrt{n}(Y - g(\mu))}{|g'(\mu)|\sigma} = \sqrt{12n}\left(\overline{X} - \frac{1}{4}\right) \rightsquigarrow N(0,1).$$

10. $Y_n = g(\overline{X}_1, \overline{X}_2)$ where $g(s_1, s_2) = s_1/s_2$. By the central limit theorem,

$$\sqrt{n}\begin{pmatrix} \overline{X}_1 - \mu_1 \\ \overline{X}_2 - \mu_2 \end{pmatrix} \rightsquigarrow N(0, \Sigma).$$

Now

$$\nabla g(s) = \begin{pmatrix} \frac{\partial g}{\partial s_1} \\ \frac{\partial g}{\partial s_2} \end{pmatrix} = \begin{pmatrix} \frac{1}{s_2} \\ -\frac{s_1}{s_2^2} \end{pmatrix}$$

and so

$$\nabla_\mu^T \Sigma \nabla_\mu = \begin{pmatrix} \dfrac{1}{\mu_2} & -\dfrac{\mu_1}{\mu_2^2} \end{pmatrix} \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix} \begin{pmatrix} \frac{1}{\mu_2} \\ -\frac{\mu_1}{\mu_2} \end{pmatrix} = \frac{\sigma_{11}\mu_2^2 - 2\mu_2\mu_1\sigma_{12} + \mu_1^2\sigma_{22}}{\mu_2^4}$$

Therefore,

$$\sqrt{n}(\overline{X}_1\overline{X}_2 - \mu_1\mu_2) \rightsquigarrow N\Big(0, \frac{\sigma_{11}\mu_2^2 - 2\mu_2\mu_1\sigma_{12} + \mu_1^2\sigma_{22}}{\mu_2^4}\Big).$$