Additional Practice Questions 2 Solutions

7.6

a)

$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 \xrightarrow{p} \mathrm{E}[X^2]$$

by the WLLN, provided that $\mathrm{E}[X^2] < \infty$.

b)

$$\frac{1}{n}\sum_{i=1}^{n} X_i^3 \xrightarrow{p} \mathrm{E}[X^3]$$

by the WLLN, provided that $\mathrm{E}[|X^3|] < \infty$.

c) WLLN and CMT do not imply convergence of $\max_{i \leq n} X_i$

d)

$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)^2 \xrightarrow{p} \mathrm{E}[X^2] - \mathrm{E}[X]^2$$

where $\frac{1}{n}\sum_{i=1}^{n} X_i^2 \xrightarrow{p} \mathrm{E}[X^2]$ and $\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{p} \mathrm{E}[X]$ by the WLLN, then we can use the CMT for the squared term and the for the sum. We need that $\mathrm{E}[X^2] < \infty$ (recall from Lyapunov's inequality that if $\mathrm{E}[X^2]$ exists, then $\mathrm{E}[X]$ will also exist).

e)

$$\frac{\sum_{i=1}^{n} X_i^2}{\sum_{i=1}^{n} X_i} = \frac{\frac{1}{n}\sum_{i=1}^{n} X_i^2}{\frac{1}{n}\sum_{i=1}^{n} X_i} \xrightarrow{p} \frac{\mathrm{E}[X^2]}{\mathrm{E}[X]}$$

where the sample averages converge by the WLLN and the CMT can be used for dividing one by the other. We need that $\mathrm{E}[X^2] < \infty$ here.

f) From the WLLN, we have that $\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{p} \mathrm{E}[X]$ provided that $\mathrm{E}[|X|] < \infty$. The function $h(u) = \mathbf{1}\{u > 0\}$ is continuous at all $u \neq 0$. Therefore, by the CMT, we have that

$$\mathbf{1}\left\{\frac{1}{n}\sum_{i=1}^{n} X_i > 0\right\} \xrightarrow{p} \mathbf{1}\{\mathrm{E}[X] > 0\}$$

as long as $\mathrm{E}[X] \neq 0$.

g) $\frac{1}{n}\sum_{i=1}^{n} \xrightarrow{p} \mathrm{E}[XY]$ by the WLLN provided that $\mathrm{E}[|XY|] < \infty$. [Recall: from the Cauchy-Schwarz inequality, a sufficient condition for $\mathrm{E}[XY]$ existing is that $\mathrm{E}[X^2] < \infty$ and $\mathrm{E}[Y^2] < \infty$.]

7.12

From Chebyshev's inequality (and given that $\mathrm{E}[Z] = 0$ and $\mathrm{var}(Z) = 1$), we have that

$$\mathrm{P}(|Z| > \delta) \leq \frac{1}{\delta^2}$$

If we want to pick δ such that $\mathrm{P}(|Z| > \delta) \leq 0.05$, we can solve for δ in

$$\begin{aligned} \frac{1}{\delta^2} &= 0.05 \\ \Leftrightarrow \frac{1}{0.05} &= \delta^2 \\ \Leftrightarrow \delta &\approx 4.47 \end{aligned}$$

If, instead, we knew that $Z \sim N(0, 1)$, then the δ that solves $\mathrm{P}(|Z| > \delta) = 0.05$ is $\delta = 1.96$ (this is just the critical value for $\alpha = 0.05$).

Here we are thinking about bounding the tail probability for Z. If we know the distribution of Z, then it makes sense that this bound will be tighter. And in the case for Chebyshev's inequality, where we don't impose that we know the distribution of Z, it makes sense that the bound will be less tight.

8.1

a)

$$\begin{aligned} \mathrm{E}[X] &= 0\mathrm{P}(X = 0) + 1\mathrm{P}(X = 1) \\ &= 0(1 - p) + 1p \\ &= p \end{aligned}$$

b) Given the result from part (a), the moment estimator of p is

$$\hat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}$$

c) Given the result in part (b), we have that

$$\begin{aligned} \mathrm{var}(\hat{p}) &= \mathrm{var}(\bar{X}) \\ &= \frac{\mathrm{var}(X)}{n} \end{aligned}$$

I think that we have computed $\mathrm{var}(X)$ when X follows a Bernoulli distribution before, but for completeness, notice that

$$\mathrm{E}[X^2] = (0)^2\mathrm{P}(X = 0) + (1)^2\mathrm{P}(X = 1) = p$$

Therefore,

$$\mathrm{var}(X) = \mathrm{E}[X^2] - \mathrm{E}[X]^2 = p - p^2 = p(1 - p)$$

Thus,

$$\text{var}(\hat{p}) = \frac{p(1-p)}{n}$$

d) Since $\hat{p} = \bar{X}$ and $p = \text{E}[X]$, we immediately have from the central limit theorem that

$$\sqrt{n}(\hat{p} - p) = \sqrt{n}(\bar{X} - \text{E}[X]) \xrightarrow{d} N(0, \sigma^2)$$

where $\sigma^2 = \text{var}(X) = p(1-p)$.

8.8

For each question, if we let $\theta = (\theta_1, \theta_2)'$ and $h(\theta)$ denote the function in each particular part of the problem, then we mainly need to compute $\nabla h(\theta)$.

a) In this case

$$\nabla h(\theta) = \begin{bmatrix} \theta_2 \\ \theta_1 \end{bmatrix}$$

and we have that

$$\sqrt{n}(\hat{\theta}_1\hat{\theta}_2 - \theta_1\theta_2) \xrightarrow{d} N(0, \nabla h(\theta)'\Sigma\nabla h(\theta))$$

You could perhaps simplify the expression $\nabla h(\theta)'\Sigma\nabla h(\theta)$ a bit more; it would depend on the particular elements of the matrix Σ though and I am thinking that, since the problem didn't give these specific notation, that it is implicitly saying that the previous expression is enough.

b) In this case,

$$\nabla h(\theta) = \begin{bmatrix} \exp(\theta_1 + \theta_2) \\ \exp(\theta_1 + \theta_2) \end{bmatrix}$$

Thus,

$$\sqrt{n}(\exp(\hat{\theta}_1 + \hat{\theta}_2) - \exp(\theta_1 + \theta_2)) \xrightarrow{d} N(0, \nabla h(\theta)'\Sigma h(\theta))$$

Just to be clear here, the expression on the right hand side has the same notation as for part (a), but the particular expression for $\nabla h(\theta)$ is not the same, so they are not identical in practice (sorry for the confusing notation).

c) In this case

$$\nabla h(\theta) = \begin{bmatrix} \frac{1}{\theta_2^2} \\ -\frac{2\theta_1}{\theta_2^3} \end{bmatrix}$$

and, therefore,

$$\sqrt{n}\left(\frac{\hat{\theta}_1}{\hat{\theta}_2^2} - \frac{\theta_1}{\theta_2^2}\right) \xrightarrow{d} N(0, \nabla h(\theta)'\Sigma h(\theta))$$

d) In this case

$$\nabla h(\theta) = \begin{bmatrix} 3\theta_1^2 + \theta_2^2 \\ 2\theta_1\theta_2 \end{bmatrix}$$

and, therefore,

$$\sqrt{n}\left(\hat{\theta}_1^3 + \hat{\theta}_1\hat{\theta}_2^2 - (\theta_1^3 + \theta_1\theta_2^2)\right) \xrightarrow{d} N(0, \nabla h(\theta)'\Sigma h(\theta))$$

13.6(a,b)

a)

$$t = \frac{\bar{X}}{\text{se}(\bar{X})} = \frac{1.2}{0.4} = 3$$

$|t| > 1.96 \implies$ we would reject H_0 at the 5% significance level.

b)

$$t = \frac{\bar{X}}{\text{se}(\bar{X})} = \frac{-1.6}{0.9} \approx 1.78$$

$|t| < 1.96 \implies$ we would fail to reject H_0 at the 5% significance level.

13.11

The wording of this problem is somewhat unclear to me, but I am interpreting this as saying that you have two independent samples for Madison and Ann Arbor, and that each of them has exactly n observations. For this problem, let X denote denote the monthly rent of a person in Madison, and let Y denote the monthly rent of a person in Ann Arbor. We are able to estimate

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{and} \quad \frac{1}{n}\sum_{i=1}^{n} Y_i$$

We are interested in $H_0 : \text{E}[X] - \text{E}[Y] = 0$. The key step here is to figure out the limiting distribution of

$$\begin{aligned}\sqrt{n}(\bar{X} - \bar{Y} - (\text{E}[X] - \text{E}[Y])) &= \sqrt{n}(\bar{X} - \text{E}[X]) - \sqrt{n}(\bar{Y} - \text{E}[Y]) \\ &\xrightarrow{d} Z_1 - Z_2\end{aligned}$$

where $Z_1 \sim N(0, \sigma_X^2)$ and where $Z_2 \sim N(0, \sigma_Y^2)$ and where $\sigma_X^2 = \text{var}(X)$ and $\sigma_Y^2 = \text{var}(Y)$. Importantly, since we have independent samples from Madison and Ann Arbor, $Z_1 + Z_2 \sim N(0, \sigma_X^2 + \sigma_Y^2)$ (i.e., there is no additional covariance term). Then, we can construct a test statistic

$$\frac{\sqrt{n}(\bar{X} - \bar{Y})}{\sqrt{\hat{\sigma}_X^2 + \hat{\sigma}_Y^2}}$$

where

$$\hat{\sigma}_X^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}) \quad \text{and} \quad \hat{\sigma}_Y^2 = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \bar{Y})$$

Then, we can compare the absolute value of t that we calculated above to some critical value, e.g., 1.96, to conduct the test of whether or not the rent in the two cities is the same.

14.1

a)

$$[\hat{\theta} \pm 1.96\text{s.e.}(\hat{\theta})] = [2.45 \pm 1.96(0.14)] \approx [2.18, 2.72]$$

b) The only thing that changes for the 90% confidence interval is the critical value

$$[\hat{\theta} \pm 1.64\text{s.e.}(\hat{\theta})] = [2.45 \pm 1.64(0.14)] \approx [2.22, 2.68]$$

14.4

a)

$$\hat{\beta} = \exp(\hat{\theta}) = \exp(0.45) \approx 1.57$$

b) As a first step, notice that $h'(\theta) = \exp(\theta)$. Also, use the notation V to indicate the asymptotic variance of $\sqrt{n}(\hat{\theta} - \theta)$. The problem does not give us V, but we do know that $\text{s.e.}(\hat{\theta}) = 0.28 = \frac{\sqrt{\hat{V}}}{\sqrt{n}}$. From the delta method, we have that

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} N(0, \underbrace{h'(\theta)^2 V}_{=:V_\beta})$$

This suggests that $\text{s.e.}(\hat{\beta}) = \frac{\sqrt{\hat{V}_\beta}}{\sqrt{n}} = h'(\hat{\theta})\frac{\sqrt{\hat{V}}}{\sqrt{n}} = h'(\hat{\theta})\text{s.e.}(\hat{\theta}) \approx (1.57)(0.28) \approx 0.44$

c)

$$[\hat{\beta} \pm 1.96\text{s.e.}(\hat{\beta})] = [1.57 \pm 1.96(0.44)] = [0.71, 2.43]$$

d)

$$[\hat{\theta} \pm 1.96\text{s.e.}(\hat{\theta})] = [0.45 \pm 1.96(0.28)] = [-0.10, 1.00]$$

If we take exp() of the lower and upper confidence interval immediately above, we get

$$[0.90, 2.71]$$

$$\mathrm{P}\Big(\exp(\hat{\theta} - 1.96\text{s.e.}(\hat{\theta})) < \exp(\theta) < \exp(\hat{\theta} + 1.96\text{s.e.}(\hat{\theta}))\Big)$$

Notice that this is, at least, a different confidence interval that the normal confidence interval from part (c).

The last part is about why this is a valid way to form a confidence interval. This is a challenging question and one that I had to think about for a while. I'm 90% sure the answer below is correct, but if you think there are issues, let me know.

As a first step, notice that, from the setup of the problem, we have that

$$\frac{(\hat{\theta} - \theta)}{\text{se}(\hat{\theta})} \xrightarrow{d} Z \sim N(0, 1)$$

and from the extended continuous mapping theorem, we have that

$$\exp\left(\frac{(\hat{\theta} - \theta)}{\text{se}(\hat{\theta})}\right) \xrightarrow{d} \exp(Z)$$

A useful thing here (that we have not mentioned in class before) is that, if $Z \sim N(0, 1)$, the $\exp(Z)$ follows a log-normal distribution with parameters $\mu = 0$ and $\sigma^2 = 1$. Let z_p denote the pth quantile of a standard normal random variable, and let q_p denote the p quantile of a log-normal random variable with $\mu = 0$ and $\sigma^2 = 1$. A useful property of log-normally random variables is that $\log(q_p) = z_p$ (we will use this below). Now because $\exp\left(\frac{(\hat{\theta}-\theta)}{\text{se}(\hat{\theta})}\right) \xrightarrow{d} \exp(Z)$, we have that, asymptotically,

$$\begin{aligned}
0.95 &= \mathrm{P}\left(q_{.025} \leq \exp\left(\frac{(\hat{\theta} - \theta)}{\text{se}(\hat{\theta})}\right) \leq q_{.975}\right) \\
&= \mathrm{P}\left(\log(q_{.025})\text{se}(\hat{\theta}) \leq (\hat{\theta} - \theta) \leq \log(q_{.975})\text{se}(\hat{\theta})\right) \\
&= \mathrm{P}\Big(-\hat{\theta} + z_{.025}\,\text{se}(\hat{\theta}) \leq -\theta \leq -\hat{\theta} + z_{.975}\,\text{se}(\hat{\theta})\Big) \\
&= \mathrm{P}\Big(\hat{\theta} - z_{.975}\,\text{se}(\hat{\theta}) \leq \theta \leq \hat{\theta} - z_{.025}\,\text{se}(\hat{\theta})\Big) \\
&= \mathrm{P}\Big(\hat{\theta} - 1.96\,\text{se}(\hat{\theta}) \leq \theta \leq \hat{\theta} + 1.96\,\text{se}(\hat{\theta})\Big) \\
&= \mathrm{P}\Big(\exp\left(\hat{\theta} - 1.96\,\text{se}(\hat{\theta})\right) \leq \beta \leq \exp\left(\hat{\theta} + 1.96\,\text{se}(\hat{\theta})\right)\Big)
\end{aligned}$$

where the first line holds because (at least asymptotically) there is a 95% probability that a log-normally distributed random variable is between $q_{.025}$ and $q_{.975}$, the second equality holds by taking the logarithm of all terms and then multiplying all terms by $\text{se}(\hat{\theta})$, the third equality subtracts $\hat{\theta}$ from all terms and by the connection between quantiles of the log-normal and normal distributions mentioned in the previous paragraph, the fourth equality multiplies each term by -1 (which causes the inequalities to flip directions) and then re-arranges the inequalities, the fifth equality holds because $z_{.975} = 1.96$ and $z_{.025} = -1.96$, the sixth equality takes the exponential of each term (and uses that $\beta = \exp(\theta)$). Finally, notice on the last line that this is exactly the same interval as we computed earlier in this problem, and that it covers β with 95% probability (which is what we've just shown) completes the result.