

Midterm 2 Solutions to Additional Practice Questions

6.3

The natural estimator for μ_k , based on the analogy principle, is

$$\hat{\mu}_k = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^k$$

The estimator looks similar to $\hat{\sigma}^2$ for $\text{Var}(X)$ that we discussed in class, and, based on that, we would expect it to be biased.

7.7

a) To show that $\bar{X}_n^*$ is unbiased,

$$\begin{aligned} \mathbb{E}[\bar{X}_n^*] &= \mathbb{E}\left[\frac{1}{n} \sum_{i=1}^n w_i X_i\right] \\ &= \frac{1}{n} \sum_{i=1}^n w_i \mathbb{E}[X_i] \\ &= \frac{1}{n} \sum_{i=1}^n w_i \mathbb{E}[X] \\ &= \mathbb{E}[X] \frac{1}{n} \sum_{i=1}^n w_i \\ &= \mathbb{E}[X] \end{aligned}$$

where the first equality holds by the definition of $\bar{X}_n^*$, the second equality holds because $1/n$ is constant, expectations can pass through sums, and w_i 's are constant, the third equality holds because X_i 's are identically distributed, the fourth equality holds because $\mathbb{E}[X]$ doesn't vary with i , and the last equality holds because $n^{-1} \sum_{i=1}^n w_i = 1$.

b) For the variance,

$$\begin{aligned} \text{Var}(\bar{X}_n^*) &= \text{Var}\left(\frac{1}{n} \sum_{i=1}^n w_i X_i\right) \\ &= \frac{1}{n^2} \sum_{i=1}^n w_i^2 \text{Var}(X_i) \\ &= \frac{1}{n^2} \sum_{i=1}^n w_i^2 \text{Var}(X) \\ &= \text{Var}(X) \frac{1}{n^2} \sum_{i=1}^n w_i^2 \end{aligned}$$

where the first equality holds by the definition of $\bar{X}_n^*$, the second equality holds because $1/n$ is constant, the X_i 's are independent (implying that variance can pass through the sum), and

the w_i 's are constant, the third equality holds because the X_i 's are identically distributed, and the last equality holds because $\text{Var}(X)$ does not change with i .

- c) To show that $\bar{X}_n^* \xrightarrow{p} \mu$, we can show that, for any $\epsilon > 0$, $P(|\bar{X}_n^* - \mu| \geq \epsilon) \rightarrow 0$ as $n \rightarrow \infty$. From Chebyshev's inequality, we have that

$$\begin{aligned} P(|\bar{X}_n^* - \mu| \geq \epsilon) &\leq \frac{\text{Var}(\bar{X}_n^*)}{\epsilon^2} \\ &= \text{Var}(X) \frac{1}{n^2} \sum_{i=1}^n w_i^2 / \epsilon^2 \end{aligned}$$

where the first line uses the result from part (a) that $\bar{X}_n^*$ is unbiased for μ , the second line uses the result about the sampling variance of $\bar{X}_n^*$ from part (b). Since ϵ is fixed and $\text{Var}(X) < \infty$, $P(|\bar{X}_n^* - \mu| \geq \epsilon) \rightarrow 0$ if $\frac{1}{n^2} \sum_{i=1}^n w_i^2 \rightarrow 0$, which is exactly the condition in the problem.

- d) As stated in the textbook, the condition in part (d) is that $\max_{i \leq n} w_i \rightarrow 0$ as $n \rightarrow \infty$. However, this condition can never hold under the setup of the problem. The largest weight is always at least as large as the average weight, which is 1 in the problem. I think that the intended condition is that

$$\frac{\max_{i \leq n} w_i}{n} \rightarrow 0$$

Since $\sum_{j=1}^n w_j = n$, w_i/n is the share of the total weight given to observation i , so this condition says that no single observation receives a non-negligible share of the total weight as n gets large.

To show that this is a sufficient condition for the condition in part (c), notice that $w_i^2 = w_i \cdot w_i \leq (\max_{j \leq n} w_j) w_i$, so that

$$\begin{aligned} \frac{1}{n^2} \sum_{i=1}^n w_i^2 &\leq \frac{1}{n^2} \left(\max_{j \leq n} w_j \right) \sum_{i=1}^n w_i \\ &= \frac{\max_{j \leq n} w_j}{n} \cdot \frac{1}{n} \sum_{i=1}^n w_i \\ &= \frac{\max_{j \leq n} w_j}{n} \\ &\rightarrow 0 \end{aligned}$$

where the third line uses that $\frac{1}{n} \sum_{i=1}^n w_i = 1$.

7.11

For this problem, we are interested in estimating $\mu = \exp(\mathbb{E}[\log(X)])$ (the geometric mean of X) using the estimator $\hat{\mu} = \left(\prod_{i=1}^n X_i\right)^{1/n}$. Notice that

$$\hat{\theta} := \log(\hat{\mu}) = \frac{1}{n} \sum_{i=1}^n \log(X_i) \xrightarrow{p} \mathbb{E}[\log(X)] =: \theta$$

where the equality holds because the log of a product is the sum of the logs, and the convergence in probability holds by the law of large numbers under the condition that $\{X_i\}_{i=1}^n$ are iid and $\mathbb{E}|\log(X)| < \infty$.

Then, we have that $\mu = \exp(\theta)$ and $\hat{\mu} = \exp(\hat{\theta})$. Since $\exp$ is a continuous function, $\hat{\mu} \xrightarrow{p} \mu$ by the continuous mapping theorem.

7.12

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

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

If we want to pick δ such that $\mathbb{P}(|Z| > \delta) \leq 0.05$, 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 $Z \sim \mathcal{N}(0, 1)$, then the δ that solves $\mathbb{P}(|Z| > \delta) = 0.05$ 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.

8.2

Consider the following estimator of $\mu'_2 = \mathbb{E}[X^2]$:

$$\hat{\mu}'_2 = \frac{1}{n} \sum_{i=1}^n X_i^2$$

From the central limit theorem, we have that

$$\sqrt{n}(\hat{\mu}'_2 - \mu'_2) = \sqrt{n} \left(\frac{1}{n} \sum_{i=1}^n X_i^2 - \mathbb{E}[X^2] \right) \xrightarrow{d} \mathcal{N}(0, v^2)$$

where

$$v^2 = \mathbb{E}[(X^2 - \mathbb{E}[X^2])^2]$$

which is just the variance of X^2 . For the central limit theorem to apply, we need that $\{X_i\}_{i=1}^n$ is an iid sample and that $\mathbb{E}[(X^2)^2] = \mathbb{E}[X^4] < \infty$.

8.7 (a)

From the delta method, we know that $\sqrt{n}(h(\hat{\theta}) - h(\theta)) \xrightarrow{d} \mathcal{N}(0, h'(\theta)^2 v^2)$ where v^2 is the asymptotic variance of $\sqrt{n}(\hat{\theta} - \theta)$. Since $h(\theta) = \theta^2$, $h'(\theta) = 2\theta$. Thus,

$$\sqrt{n}(\hat{\theta}^2 - \theta^2) \xrightarrow{d} \mathcal{N}(0, 4\theta^2 v^2)$$

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

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

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 notation for these elements, 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,

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

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.

c) In this case (which requires $\theta_2 \neq 0$)

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

and, therefore,

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

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,

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

13.6(a,b)

a) In this case,

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

$|t| > 1.96 \implies$ we would reject $\mathbb{H}_0$ at the 5% significance level.

b) In this case,

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

$|t| < 1.96 \implies$ we would fail to reject $\mathbb{H}_0$ at the 5% significance level.

13.11

For this problem, we have two independent samples for Madison and Ann Arbor, and each of them has exactly n observations. For this problem, let X 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

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

We are interested in $\mathbb{H}_0 : \mathbb{E}[X] - \mathbb{E}[Y] = 0$. To carry out the test, it's natural to construct a test statistic based on $\bar{X} - \bar{Y}$. The key step here is to figure out the limiting distribution of

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

where $Z_1 \sim \mathcal{N}(0, \sigma_X^2)$ and where $Z_2 \sim \mathcal{N}(0, \sigma_Y^2)$ 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, $Z_1 - Z_2 \sim \mathcal{N}(0, \sigma_X^2 + \sigma_Y^2)$ (i.e., there is no additional covariance term). Thus, we can construct a test statistic

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

where

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

Under $\mathbb{H}_0$, $t \xrightarrow{d} \mathcal{N}(0, 1)$, while under the alternative t diverges. Then, we can compare the absolute value of t that we calculated above to a critical value, e.g., 1.96 at a 5% significance level, to conduct the test of whether or not the rent in the two cities is the same.

14.4 (a, b, in addition: provide a t-statistic for $\mathbb{H}_0 : \beta = 1$, do you reject $\mathbb{H}_0$?)

a) In this case,

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

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

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

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

provide a t-statistic for $\mathbb{H}_0 : \beta = 1$, do you reject $\mathbb{H}_0$?

$$t = \frac{1.57 - 1}{0.44} \approx 1.30$$

Since $|t| < 1.96$, we would not reject $\mathbb{H}_0$ at the 5% significance level.