Homework 3

Problem 1. (a) Prove the “Squeeze Rule”: If $0 \leq X_n \leq Y_n$ and $Y_n \to_p 0$, then $X_n \to_p 0$; (b) Prove: $X_n \to_p 0$ if and only if $|X_n| \to_p 0$.

Solution. (a) For any $\epsilon > 0$,

$$\Pr(Y_n \leq \epsilon) \leq \Pr(X_n \leq \epsilon) \leq 1.$$

Then,

$$\Pr(Y_n \leq \epsilon) = \Pr(|Y_n - 0| \leq \epsilon) \to 1 \Longrightarrow \Pr(X_n \leq \epsilon) = \Pr(|X_n - 0| \leq \epsilon) \to 1.$$

(b) “$\Longrightarrow$ part”: by continuous mapping theorem, since the mapping $x \mapsto |x|$ is continuous. “$\Longleftarrow$ part”: straightforward.

Problem 2. Provide a counter example to show that $X_n \to_d X$ and $Y_n \to_d Y$ does not imply $X_n + Y_n \to_d X + Y$. Hint: Consider an iid random sample $X_1, ..., X_n$ with $\mathbb{E}X_1 = 0$ and $n^{1/2}\overline{X}_n$ and $-n^{1/2}\overline{X}_n$.

Solution. Let Z be a random variable such that $Z \sim N(0, \sigma^2)$, where $\sigma^2 = \text{Var}(X_1)$. Then by CLT,

$$n^{1/2}\overline{X}_n \to_d Z$$

and

$$-n^{1/2}\overline{X}_n = (-1) \times \left(n^{1/2}\overline{X}_n\right) \to_d (-1) \times Z \sim N\left(0, \sigma^2\right).$$

Therefore, it is also true that $-n^{1/2}\overline{X}_n \to_d Z \sim N\left(0, \sigma^2\right)$. Note

$$0 = \left(n^{1/2}\overline{X}_n\right) + \left(-n^{1/2}\overline{X}_n\right) \nrightarrow_d Z + Z \sim N\left(0, 4\sigma^2\right).$$

Problem 3. Let $\widehat{\boldsymbol{\theta}}_n = \left(\widehat{\theta}_{n,1}, \ldots, \widehat{\theta}_{n,k}\right)'$ be an estimator of the k-vector of parameters $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)'$. Suppose that $\widehat{\boldsymbol{\theta}}_n \to_p \boldsymbol{\theta}$, and $n^{1/2}\left(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right) \to_d W \sim N(0, \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma}$ is a positive definite $k \times k$ matrix. Use the delta method or CMT to find the (non-degenerate, i.e., not a constant) asymptotic distributions of the following quantities after a suitable normalization. "Suitable normalization" means subtraction of a constant and/or multiplication by a constant (could be dependent on n).

1. $n^{1/2}\left(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right)' \boldsymbol{c}$ where $\boldsymbol{c} \in \mathbb{R}^k$ is a vector of constants.
2. $\widehat{\theta}_{n,1}$.
3. $n\left(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right)'\left(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right)$.
4. $\widehat{\theta}_{n,1} - \widehat{\theta}_{n,2}$.
5. $\widehat{\theta}_{n,1}\widehat{\theta}_{n,2}/\widehat{\theta}_{n,3}$, provided that $\theta_3 \neq 0$.

Solution.

1. Define $\boldsymbol{X}_n = n^{1/2}\left(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right)$ and $h(\boldsymbol{X}_n) = \boldsymbol{X}_n'\boldsymbol{c}$. By the Continuous Mapping Theorem we have

$$n^{1/2}\left(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right)' \boldsymbol{c} = h(\boldsymbol{X}_n) \to_d h(\boldsymbol{W}) = \boldsymbol{W}'\boldsymbol{c}$$

By the property of normal distribution we have

$$n^{1/2}\left(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right)' \boldsymbol{c} \to_d \boldsymbol{W}'\boldsymbol{c} \sim N(\boldsymbol{0}, \boldsymbol{c}'\boldsymbol{\Sigma}\boldsymbol{c}).$$

2. Set $\boldsymbol{c} = (1, 0, ..., 0)'$. Then, it follows from Part (i) that

$$n^{1/2}(\widehat{\theta}_{n,1} - \theta_1) \to_d N(0, \sigma_{11}^2),$$

where σ_{11}^2 is the first diagonal element of $\boldsymbol{\Sigma}$.

3. Since $n^{1/2}\left(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right) \to_d \boldsymbol{W}$, by the Continuous Mapping Theorem,

$$n(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta})'(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) = \left[n^{1/2}\left(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right)\right]'\left[n^{1/2}\left(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right)\right] \to_d \boldsymbol{W}'\boldsymbol{W}.$$

4. Set $\boldsymbol{c} = (1, -1, 0, ..., 0)'$.It follows from Part (i) that

$$n^{1/2}(\widehat{\theta}_{n,1} - \widehat{\theta}_{n,2} - \theta_1 + \theta_2) \to_d N(0, \sigma_{11}^2 - 2\sigma_{12} + \sigma_{22}^2),$$

where σ_{11}^2 and σ_{22}^2 are the first and second diagonal element of $\boldsymbol{\Sigma}$, and σ_{12} is the element on the first row and second column of $\boldsymbol{\Sigma}$.

5. Put $h(\boldsymbol{\theta}) = \frac{\theta_1\theta_2}{\theta_3}$, apply the Delta method

$$n^{1/2}(\frac{\widehat{\theta}_{n,1}\widehat{\theta}_{n,2}}{\widehat{\theta}_{n,3}} - \frac{\theta_1\theta_2}{\theta_3}) = n^{1/2}(h(\widehat{\boldsymbol{\theta}}_n) - h(\boldsymbol{\theta})) \to_d \frac{\partial h(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}\boldsymbol{W}$$

where

$$\frac{\partial h(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} = \left(\frac{\theta_2}{\theta_3}, \frac{\theta_1}{\theta_3}, \frac{-\theta_1\theta_2}{\theta_3^2}, 0, ..., 0\right)'.$$

Then by the property of Normal density

$$n^{1/2}(\frac{\widehat{\theta}_{n,1}\widehat{\theta}_{n,2}}{\widehat{\theta}_{n,3}} - \frac{\theta_1\theta_2}{\theta_3}) \to_d N(0, \frac{\partial h(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}\boldsymbol{\Sigma}\frac{\partial h(\boldsymbol{\theta})'}{\partial \boldsymbol{\theta}}).$$

Problem 4. Suppose that $\hat{\theta}_n \to_p \theta$ and $\hat{\beta}_n \to \beta$, where θ and β are two scalar parameters. Without relying on Slutsky's Theorem, show:

1. $c\hat{\theta}_n \to_p c\theta$, where c is a constant.
2. $\hat{\theta}_n\hat{\beta}_n \to_p \theta\beta$.

Solution. (i) Suppose $c \neq 0$. Then $\Pr\left(|c\hat{\theta}_n - c\theta| > \varepsilon\right) = \Pr\left(|\hat{\theta}_n - \theta| > \frac{\varepsilon}{|c|}\right) \to 0$ as $n \to \infty$. If $c = 0$, then $c\hat{\theta}_n = 0 \to_p c\theta = 0$.

(ii) First, note that $\hat{\theta}_n\hat{\beta}_n - \theta\beta = (\hat{\theta}_n - \theta + \theta)(\hat{\beta}_n - \beta + \beta) - \theta\beta = \left(\hat{\theta}_n - \theta\right)\left(\hat{\beta}_n - \beta\right) + \left(\hat{\theta}_n - \theta\right)\beta + \left(\hat{\beta}_n - \beta\right)\theta$. Then, $\left(\hat{\theta}_n - \theta\right)\beta + \left(\hat{\beta}_n - \beta\right)\theta \to_p 0$ by Part (i). Then, for any $\epsilon > 0$,

$$\begin{aligned}\Pr\left(\left|\left(\hat{\theta}_n - \theta\right)\left(\hat{\beta}_n - \beta\right)\right| > \varepsilon\right) &\leq \Pr\left(\left|\hat{\theta}_n - \theta\right| > \sqrt{\varepsilon} \text{ or } \left|\hat{\beta}_n - \beta\right| > \sqrt{\varepsilon}\right) \\ &\leq \Pr\left(\left|\hat{\theta}_n - \theta\right| > \sqrt{\varepsilon}\right) + \Pr\left(\left|\hat{\beta}_n - \beta\right| > \sqrt{\varepsilon}\right) \to 0 \text{ as } n \to \infty.\end{aligned}$$

Thus, $\hat{\theta}_n\hat{\beta}_n - \theta\beta \to_p 0$.

Problem 5. Suppose that $\mathbb{E}\left(\hat{\theta}_n\right) \to \theta$ and $\mathrm{Var}(\hat{\theta}_n) \to 0$ as $n \to \infty$. Show that $\hat{\theta}_n \to_p \theta$.

Solution. $\hat{\theta}_n$ converges in probability to θ if for all $\varepsilon > 0$, $\Pr\left(\left|\hat{\theta}_n - \theta\right| \geq \varepsilon\right) \to 0$ as $n \to \infty$. First, decompose the Mean Squared Error (MSE) into

$$\begin{aligned} MSE\left(\hat{\theta}_n\right) &= \mathbb{E}\left(\hat{\theta}_n - \theta\right)^2 = \mathbb{E}\left(\hat{\theta}_n - \mathbb{E}\hat{\theta}_n + \mathbb{E}\hat{\theta}_n - \theta\right)^2 \\ &= \mathbb{E}\left(\hat{\theta}_n - \mathbb{E}\hat{\theta}_n\right)^2 + \left(\mathbb{E}\hat{\theta}_n - \theta\right)^2 + 2\mathbb{E}\left(\hat{\theta}_n - \mathbb{E}\hat{\theta}_n\right)\left(\mathbb{E}\hat{\theta}_n - \theta\right) \\ &= \mathbb{E}\left(\hat{\theta}_n - \mathbb{E}\hat{\theta}_n\right)^2 + \left(\mathbb{E}\hat{\theta}_n - \theta\right)^2 = \text{Var}\left(\hat{\theta}_n\right) + \text{Bias}\left(\hat{\theta}_n\right)^2, \end{aligned}$$

where the last line follows by the fact that $\mathbb{E}\left(\hat{\theta}_n - \mathbb{E}\hat{\theta}_n\right) = 0$.

Then, using Markov's Inequality,

$$\Pr\left(\left|\hat{\theta}_n - \theta\right| \geq \varepsilon\right) \leq \frac{\mathbb{E}\left|\hat{\theta}_n - \theta\right|^2}{\varepsilon^2} = \frac{\mathbb{E}\left(\hat{\theta}_n - \theta\right)^2}{\varepsilon^2} = \frac{\text{Var}\left(\hat{\theta}_n\right) + \text{Bias}\left(\hat{\theta}_n\right)^2}{\varepsilon^2} \to 0 \text{ as } n \to \infty,$$

since by assumption, $\text{Var}\left(\hat{\theta}_n\right) \to 0$ and $\mathbb{E}\hat{\theta}_n - \theta \to 0$ as $n \to \infty$.

Problem 6. Consider the linear model (with independently and identically distributed (i.i.d.) observations):

$$Y_i = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + U_i$$

with $\mathbb{E}U_i = \mathbb{E}U_i X_{1,i} = \mathbb{E}U_i X_{2,i} = 0$. Suppose we know that $\beta_2 = \beta_1$ and conduct a constrained LS estimation of β_1:

$$\min_{b_0, b_1} \sum_{i=1}^{n} \left(Y_i - b_0 - b_1 X_{1,i} - b_1 X_{2,i}\right)^2.$$

1. Find the expression for the constrained LS estimator $\left(\widehat{\beta}_0, \widehat{\beta}_1\right)$ that solve the above minimization problem.
2. Assume that the restriction $\beta_2 = \beta_1$ is true. Derive the large-sample (asymptotic) distribution of $\widehat{\beta}_1$.

Solution. Denote $\overline{X}_1 = n^{-1}\sum_{i=1}^{n} X_{1,i}$ and $\overline{X}_2 = n^{-1}\sum_{i=1}^{n} X_{2,i}$. The constrained LS:

$$\widehat{\beta}_1 = \frac{\sum_{i=1}^{n}\left(X_{1,i} + X_{2,i} - \overline{X}_1 - \overline{X}_2\right) Y_i}{\sum_{i=1}^{n}\left(X_{1,i} + X_{2,i} - \overline{X}_1 - \overline{X}_2\right)^2}.$$

And

$$\widehat{\beta}_0 = \overline{Y} - \widehat{\beta}_1\left(\overline{X}_1 + \overline{X}_2\right).$$

For (ii),

$$\begin{aligned} \widehat{\beta}_1 &= \frac{\sum_{i=1}^{n}\left(X_{1,i} + X_{2,i} - \overline{X}_1 - \overline{X}_2\right) Y_i}{\sum_{i=1}^{n}\left(X_{1,i} + X_{2,i} - \overline{X}_1 - \overline{X}_2\right)^2} \\ &= \frac{\sum_{i=1}^{n}\left(X_{1,i} + X_{2,i} - \overline{X}_1 - \overline{X}_2\right)\left(\beta_0 + \beta_1 X_{1,i} + \beta_1 X_{2,i} + U_i\right)}{\sum_{i=1}^{n}\left(X_{1,i} + X_{2,i} - \overline{X}_1 - \overline{X}_2\right)^2} \\ &= \beta_1 + \frac{\frac{1}{n}\sum_{i=1}^{n}\left(X_{1,i} + X_{2,i} - \overline{X}_1 - \overline{X}_2\right) U_i}{\frac{1}{n}\sum_{i=1}^{n}\left(X_{1,i} + X_{2,i} - \overline{X}_1 - \overline{X}_2\right)^2}. \end{aligned}$$

By WLLN and Continuous Mapping Theorem,

$$\frac{1}{n}\sum_{i=1}^{n}\left(X_{1,i} + X_{2,i} - \overline{X}_1 - \overline{X}_2\right)^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_{1,i} + X_{2,i}\right)^2 - \left(\overline{X}_1 + \overline{X}_2\right)^2$$

$$\to_p \quad \operatorname{Var}(X_{1,i} + X_{2,i}).$$

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\left(X_{1,i}+X_{2,i}-\overline{X}_1-\overline{X}_2\right)U_i &= \frac{1}{n}\sum_{i=1}^{n}\left(X_{1,i}+X_{2,i}-\mathbb{E}\left(X_{1,i}+X_{2,i}\right)\right)U_i \\ &\quad + \left(\overline{X}_1-\mathbb{E}\left(X_{1,i}\right)\right)\frac{1}{n}\sum_{i=1}^{n}U_i + \left(\overline{X}_2-\mathbb{E}\left(X_{2,i}\right)\right)\frac{1}{n}\sum_{i=1}^{n}U_i.\end{aligned}$$

Since $n^{-1/2}\sum_{i=1}^{n}U_i \to_d N\left(0,\mathbb{E}\left(U_i^2\right)\right)$, $\overline{X}_1-\mathbb{E}\left(X_{1,i}\right)\to_p 0$ and $\overline{X}_2-\mathbb{E}\left(X_{2,i}\right)\to_p 0$, by Slutsky's theorem,

$$\begin{aligned}\left(\overline{X}_1-\mathbb{E}\left(X_{1,i}\right)\right)\frac{1}{\sqrt{n}}\sum_{i=1}^{n}U_i &\to_p 0 \\ \left(\overline{X}_2-\mathbb{E}\left(X_{2,i}\right)\right)\frac{1}{\sqrt{n}}\sum_{i=1}^{n}U_i &\to_p 0.\end{aligned}$$

By CLT,

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(X_{1,i}+X_{2,i}-\mathbb{E}\left(X_{1,i}+X_{2,i}\right)\right)U_i \to_d N\left(0,\mathbb{E}\left(U_i^2\left(X_{1,i}+X_{2,i}-\mathbb{E}\left(X_{1,i}+X_{2,i}\right)\right)^2\right)\right).$$

By Slutsky's theorem,

$$\begin{aligned}\sqrt{n}\left(\widehat{\beta}_1-\beta_1\right) &= \frac{\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(X_{1,i}+X_{2,i}-\overline{X}_1-\overline{X}_2\right)U_i}{\frac{1}{n}\sum_{i=1}^{n}\left(X_{1,i}+X_{2,i}-\overline{X}_1-\overline{X}_2\right)^2} \\ &\to_d \operatorname{Var}\left(X_{1,i}+X_{2,i}\right)^{-1}N\left(0,\mathbb{E}\left(U_i^2\left(X_{1,i}+X_{2,i}-\mathbb{E}\left(X_{1,i}+X_{2,i}\right)\right)^2\right)\right).\end{aligned}$$

Problem 7. Suppose we observe the i.i.d. random sample $\{(Y_i, X_i)\}_{i=1}^n$ with X_i being a scalar. Take the linear model

$$\begin{aligned}Y_i &= X_i\beta + e_i \\ \mathbb{E}\left(e_i|X_i\right) &= 0.\end{aligned}$$

Consider the estimator

$$\widehat{\beta} = \frac{\sum_{i=1}^{n}X_i^3Y_i}{\sum_{i=1}^{n}X_i^4}.$$

Find the asymptotic distribution of $\sqrt{n}\left(\widehat{\beta}-\beta\right)$.

Problem 8. Let $\{\theta_n : n \geq 1\}$ be a random sequence such that $\Pr\left(\theta_n = 0\right) = (n-1)/n$, and $\Pr\left(\theta_n = n^2\right) = 1/n$. Note that the only possible values for θ_n are zero and n^2.

1. Show that $\lim_{n\to\infty}\mathbb{E}\theta_n = \infty$.

2. Does θ_n converge in probability to some limit? If yes, prove. If not, explain why.

Solution. (i) $\mathbb{E}\theta_n = 0\cdot(n-1)/n + n^2\cdot 1/n = n \to \infty$.
(ii) $\theta_n \to_p 0$, since for any $\epsilon > 0$, $\Pr\left(|\theta_n| > \epsilon\right) = \Pr\left(\theta_n = n^2\right) = 1/n \to 0$.