Homework 6

Problem 1. Consider the following simple linear regression model

$$Y_i = \beta X_i + U_i,$$

where $\beta \in \mathbb{R}$ is the unknown parameter. The econometrician is interested in constructing a $1-\alpha$ asymptotic confidence interval for β, where $0 < \alpha < 1/2$. Assume that the data are i.i.d. and the following assumptions hold, A-i. $\mathbb{E}(X_i U_i) = 0$. A-ii. $0 < \mathbb{E}X_i^2 < \infty$, $j = 1, \ldots, k$. A-iii. $\mathbb{E}\left(U_i^2 | X_i\right) = \sigma^2$.

Define

$$\hat{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{\beta}_n X_i\right)^2,$$

where $\hat{\beta}_n$ is the OLS estimator of β. For each confidence interval listed below indicate if it is asymptotically valid. Carefully justify your answers.

(i) $\left[\hat{\beta}_n - z_{1-\alpha/2}\hat{\sigma}_n, \hat{\beta}_n + z_{1-\alpha/2}\hat{\sigma}_n\right]$.

(ii) $\left[\hat{\beta}_n - z_{1-\alpha/2}\hat{\sigma}_n/\sqrt{n}, \hat{\beta}_n + z_{1-\alpha/2}\hat{\sigma}_n/\sqrt{n}\right]$.

(iii) $\left(-\infty, \hat{\beta}_n - z_\alpha\sqrt{\hat{\sigma}_n^2/\sum_{i=1}^n X_i^2}\right]$.

(iv) $\left\{b \in \mathbb{R} : \left(\hat{\beta}_n - b\right)^2 \leq \chi^2_{1,1-\alpha} v_n\right\}$, where

$$v_n = \frac{\sum_{i=1}^n \left(Y_i - \hat{\beta}_n X_i\right)^2 X_i^2}{\left(\sum_{i=1}^n X_i^2\right)^2}.$$

Solution. First write the asymptotic distribution of $\hat{\beta}_n$:

$$\sqrt{n}\left(\hat{\beta}_n - \beta\right) \to_d N\left(0, \frac{\sigma^2}{\mathbb{E}\left(X_i^2\right)}\right).$$

Since $\hat{\sigma}_n^2 \to_p \sigma^2$, by WLLN and continuous mapping theorem (CMT),

$$\frac{\hat{\sigma}_n^2}{\frac{1}{n}\sum_{i=1}^n X_i^2} \to_p \frac{\sigma^2}{\mathbb{E}\left(X_i^2\right)}.$$

Then by CMT and Slutsky's theorem,

$$\frac{\sqrt{n}\left(\hat{\beta}_n - \beta\right)}{\sqrt{\frac{\hat{\sigma}_n^2}{\frac{1}{n}\sum_{i=1}^n X_i^2}}} \to_d \left(\frac{\sigma^2}{\mathbb{E}\left(X_i^2\right)}\right)^{-1/2} \cdot N\left(0, \frac{\sigma^2}{\mathbb{E}\left(X_i^2\right)}\right) \sim N(0,1).$$

Or,

$$\frac{\hat{\beta}_n - \beta}{\sqrt{\frac{\hat{\sigma}_n^2}{\sum_{i=1}^n X_i^2}}} \to_d N(0,1)$$

and by CMT,

$$\left|\frac{\hat{\beta}_n - \beta}{\sqrt{\frac{\hat{\sigma}_n^2}{\sum_{i=1}^n X_i^2}}}\right| \to_d |Z|,$$

where $Z \sim N(0,1)$.

(i)

$$\begin{aligned}
\Pr\left(\beta \in \left[\hat{\beta}_n - z_{1-\alpha/2}\hat{\sigma}_n, \hat{\beta}_n + z_{1-\alpha/2}\hat{\sigma}_n\right]\right) &= \Pr\left(\left|\frac{\hat{\beta}_n - \beta}{\hat{\sigma}_n}\right| \leq z_{1-\alpha/2}\right) \\
&= \Pr\left(\left|\frac{\sqrt{n}\left(\hat{\beta}_n - \beta\right)}{\sqrt{n}\cdot\hat{\sigma}_n}\right| \leq z_{1-\alpha/2}\right).
\end{aligned}$$

By CMT,

$$\frac{1}{\sqrt{n}\cdot\hat{\sigma}_n} \to_p 0 \cdot \frac{1}{\sigma} = 0.$$

By Slutsky's theorem,

$$\frac{\sqrt{n}\left(\hat{\beta}_n - \beta\right)}{\sqrt{n}\cdot\hat{\sigma}_n} \to_d 0 \cdot N\left(0, \frac{\sigma^2}{\mathbb{E}\left(X_i^2\right)}\right) = 0.$$

So

$$\frac{\sqrt{n}\left(\hat{\beta}_n - \beta\right)}{\sqrt{n}\cdot\hat{\sigma}_n} \to_p 0$$

and

$$\Pr\left(\left|\frac{\sqrt{n}\left(\hat{\beta}_n - \beta\right)}{\sqrt{n}\cdot\hat{\sigma}_n}\right| \leq z_{1-\alpha/2}\right) \to 1.$$

It does not converge to $1-\alpha$, therefore the confidence interval is not valid.

(ii)

$$\Pr\left(\beta \in \left[\hat{\beta}_n - z_{1-\alpha/2}\hat{\sigma}_n/\sqrt{n}, \hat{\beta}_n + z_{1-\alpha/2}\hat{\sigma}_n/\sqrt{n}\right]\right) = \Pr\left(\left|\frac{\sqrt{n}\left(\hat{\beta}_n - \beta\right)}{\hat{\sigma}_n}\right| \leq z_{1-\alpha/2}\right).$$

Note

$$\frac{\sqrt{n}\left(\hat{\beta}_n - \beta\right)}{\hat{\sigma}_n} \to_d \sigma^{-1}\cdot N\left(0, \frac{\sigma^2}{\mathbb{E}\left(X_i^2\right)}\right) \sim N\left(0, \frac{1}{\mathbb{E}\left(X_i^2\right)}\right).$$

So $\Pr\left(\left|\frac{\sqrt{n}\left(\hat{\beta}_n - \beta\right)}{\hat{\sigma}_n}\right| \leq z_{1-\alpha/2}\right)$ does not converge to $1-\alpha$, unless $\mathbb{E}X_i^2 = 1$.

(iii) By CMT,

$$-\frac{\hat{\beta}_n - \beta}{\sqrt{\frac{\hat{\sigma}_n^2}{\sum_{i=1}^n X_i^2}}} \to_d (-1)\cdot N(0,1) \sim N(0,1).$$

Then,

$$\Pr\left(\beta \in \left(-\infty, \hat{\beta}_n - z_\alpha\sqrt{\hat{\sigma}_n^2/\sum_{i=1}^n X_i^2}\right]\right) = \Pr\left(-\frac{\hat{\beta}_n - \beta}{\sqrt{\frac{\hat{\sigma}_n^2}{\sum_{i=1}^n X_i^2}}} \leq -z_\alpha\right) \to \Pr\left(Z \leq -z_\alpha\right) = 1-\alpha,$$

where $Z \sim N(0,1)$.

(iv) We have

$$\frac{\frac{1}{n}\sum_{i=1}^n\left(Y_i - \hat{\beta}_n X_i\right)^2 X_i^2}{\left(\frac{1}{n}\sum_{i=1}^n X_i^2\right)^2} \to_p \frac{\mathbb{E}\left(U_i^2 X_i^2\right)}{\mathbb{E}\left(X_i^2\right)^2} = \frac{\sigma^2\mathbb{E}\left(X_i^2\right)}{\mathbb{E}\left(X_i^2\right)^2} = \frac{\sigma^2}{\mathbb{E}\left(X_i^2\right)}.$$

(Why?) Therefore, by CMT and Slutsky's theorem,

$$\sqrt{n}\left(\hat{\beta}_n-\beta\right)\left(\frac{\frac{1}{n}\sum_{i=1}^n\left(Y_i-\hat{\beta}_nX_i\right)^2X_i^2}{\left(\frac{1}{n}\sum_{i=1}^nX_i^2\right)^2}\right)^{-1}\sqrt{n}\left(\hat{\beta}_n-\beta\right)=\left(\frac{\sqrt{n}\left(\hat{\beta}_n-\beta\right)}{\sqrt{\frac{\frac{1}{n}\sum_{i=1}^n\left(Y_i-\hat{\beta}_nX_i\right)^2X_i^2}{\left(\frac{1}{n}\sum_{i=1}^nX_i^2\right)^2}}}\right)^2\to_d Z^2\sim\chi_1^2,$$

where $Z\sim N(0,1)$. So

$$\frac{\left(\hat{\beta}_n-\beta\right)^2}{v_n}\to_d\chi_1^2$$

and

$$\Pr\left(\frac{\left(\hat{\beta}_n-\beta\right)^2}{v_n}\leq\chi_{1,1-\alpha}^2\right)\to 1-\alpha.$$

The confidence interval is valid.

Problem 2. Consider the following model:

$$Y_i=\beta X_i+U_i,$$
$$X_i=\pi Z_i+V_i,$$

where Y_i is the dependent variable, X_i is a single regressor, and Z_i is a single instrument. Assume that the data are iid and: A-i $\mathbb{E}(Z_iU_i)=0$ and $\mathbb{E}(Z_iV_i)=0$; A-ii $\pi\neq 0$; A-iii $0<\mathbb{E}Z_i^2<\infty$; A-iv $0<\mathbb{E}\left(U_i^2Z_i^2\right)<\infty$.

(i) Show consistency of the IV estimator of β.

(ii) Show asymptotic normality of the IV estimator of β.

(iii) How does the asymptotic variance of the IV estimator of β depend on π?

Solution.

(i) The IV estimator of β is

$$\begin{aligned}\hat{\beta}_{IV} &= \left(\frac{1}{n}\sum_{i=1}^n Z_iX_i\right)^{-1}\frac{1}{n}\sum_{i=1}^n Z_iY_i\\ &= \left(\frac{1}{n}\sum_{i=1}^n Z_iX_i\right)^{-1}\frac{1}{n}\sum_{i=1}^n Z_i\left(\beta X_i+U_i\right)\\ &= \beta+\left(\frac{1}{n}\sum_{i=1}^n Z_iX_i\right)^{-1}\frac{1}{n}\sum_{i=1}^n Z_iU_i.\end{aligned}$$

By WLLN and CMT,

$$\left(\frac{1}{n}\sum_{i=1}^n Z_iX_i\right)^{-1}\to_p\left(\mathbb{E}\left(Z_iX_i\right)\right)^{-1}$$
$$\frac{1}{n}\sum_{i=1}^n Z_iU_i\to_p\mathbb{E}\left(Z_iU_i\right)=0.$$

Therefore, $\hat{\beta}_{IV}\to_p\beta$ is consistent.

(ii)

$$\sqrt{n}\left(\hat{\beta}_{IV}-\beta\right)=\left(\frac{1}{n}\sum_{i=1}^{n}Z_iX_i\right)^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}Z_iU_i.$$

By CLT,

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

By Slutsky's theorem,

$$\sqrt{n}\left(\hat{\beta}_{IV}-\beta\right)\to_d\left(\mathbb{E}\left(Z_iX_i\right)\right)^{-1}N\left(0,\mathbb{E}\left(U_i^2Z_i^2\right)\right)\sim N\left(0,\left(\mathbb{E}\left(Z_iX_i\right)\right)^{-2}\mathbb{E}\left(U_i^2Z_i^2\right)\right).$$

(iii) The asymptotic variance is

$$\left(\mathbb{E}\left(Z_iX_i\right)\right)^{-2}\mathbb{E}\left(U_i^2Z_i^2\right)=\frac{\mathbb{E}\left(U_i^2Z_i^2\right)}{\left(\mathbb{E}\left(Z_i\left(\pi Z_i+V_i\right)\right)\right)^2}=\frac{\mathbb{E}\left(U_i^2Z_i^2\right)}{\left(\pi\cdot\mathbb{E}\left(Z_i^2\right)\right)^2}=\frac{\mathbb{E}\left(U_i^2Z_i^2\right)}{\pi^2\mathbb{E}\left(Z_i^2\right)^2}.$$

The bigger π is, the smaller the variance is. If π is small, the variance would be large.

Problem 3. Consider the partitioned linear regression model $Y_i=X_{1i}'\beta_1+X_{2i}'\beta_2+U_i$, where β_1 is $k_1\times 1$ and β_2 is $k_2\times 1$. Let $\tilde{\beta}_{1n}$ denote the OLS estimator of β_1 from the regression of Y_i against X_{1i} alone. Assume that: **A1.** Data are iid. **A2.** $\mathbb{E}X_iX_i'$ is finite and positive definite, where $X_i=\begin{pmatrix}X_{1i}' & X_{2i}'\end{pmatrix}'$. **A3.** $\mathbb{E}X_{1i}U_i=0$. **A4.** $\mathbb{E}X_{1i}X_{2i}'=0$. **A5.** $\mathbb{E}(U_i^2|X_{1i})=\sigma_1^2$, $\mathbb{E}((X_{2i}'\beta_2)^2|X_{1i})=\sigma_2^2$, and $\mathbb{E}(U_iX_{2i}|X_{1i})=0$.

(i) Show that $\tilde{\beta}_{1n}$ is consistent.

(ii) Show that $\sqrt{n}(\tilde{\beta}_{1n}-\beta_1)\to_d N(0,V)$, where $V=(\sigma_1^2+\sigma_2^2)(\mathbb{E}X_{1i}X_{1i}')^{-1}$. Hint: Define $\varepsilon_i=X_{2i}'\beta_2+U_i$. Can the CLT be applied to $n^{-1/2}\sum_{i=1}^{n}X_{1i}\varepsilon_i$?

(iii) Let $\hat{\beta}_{1n}$ denote the OLS estimator of β_1 from the regression of Y_i against X_{1i} and X_{2i}. Explain, why the econometrician should use $\hat{\beta}_{1n}$ and not $\tilde{\beta}_{1n}$ despite the fact that $\tilde{\beta}_{1n}$ is consistent. Hints: Compare V with the asymptotic variance of $\hat{\beta}_{1n}$. Use the fact that the asymptotic variance of $\hat{\beta}_{1n}$ is $\sigma_1^2\left(p\lim\frac{X_1'M_2X_1}{n}\right)^{-1}$, where X_1 is a $n\times k_1$ matrix of observations on X_{1i}, $M_2=I_n-X_2(X_2'X_2)^{-1}X_2'$, and X_2 is a $n\times k_2$ matrix of observations on X_{2i}.

Solution.

(i)

$$\begin{aligned}\tilde{\beta}_{1n}-\beta_1 &= \left(\frac{X_1'X_1}{n}\right)^{-1}\left(\frac{X_1'X_2}{n}\beta_2+\frac{X_1'U}{n}\right)\\ &\to_p \left(\mathbb{E}X_{1i}X_{1i}'\right)^{-1}\left(\mathbb{E}X_{1i}X_{2i}'\beta_2+\mathbb{E}X_{1i}U_i\right)\\ &= \left(\mathbb{E}X_{1i}X_{1i}'\right)^{-1}\times 0=0.\end{aligned}$$

(ii) Define $\varepsilon_i=X_{2i}'\beta_2+U_i$. By A3 and A4, $\mathbb{E}X_{1i}\varepsilon_i=0$, and therefore the CLT can be applied to $n^{-1/2}\sum_{i=1}^{n}X_{1i}\varepsilon_i$, yielding $V=\left(\mathbb{E}X_{1i}X_{1i}'\right)^{-1}\mathbb{E}\varepsilon_i^2X_{1i}X_{1i}'\left(\mathbb{E}X_{1i}X_{1i}'\right)^{-1}$. By the law of iterated expectation,

$$\mathbb{E}\varepsilon_i^2X_{1i}X_{1i}'=\mathbb{E}\left(\mathbb{E}(\varepsilon_i^2|X_{1i})X_{1i}X_{1i}'\right),$$

and by Assumption A5,

$$\mathbb{E}(\varepsilon_i^2|X_{1i})=\sigma_1^2+\sigma_2^2.$$

Hence, $\mathbb{E}\varepsilon_i^2X_{1i}X_{1i}'=\left(\sigma_1^2+\sigma_2^2\right)\left(\mathbb{E}X_{1i}X_{1i}'\right)$, and the result for V follows.

(iii) The asymptotic variance of $\hat{\beta}_{1n}$ is

$$\begin{aligned}
\sigma_1^2\left(p\lim\frac{X_1'M_2X_1}{n}\right)^{-1} &= \sigma_1^2\left(\frac{X_1'X_1}{n}-\frac{X_1'X_2}{n}\left(\frac{X_2'X_2}{n}\right)\frac{X_2'X_1}{n}\right)^{-1}\\
&\to_p \sigma_1^2\left(\mathbb{E}X_{1i}X_{1i}'-\mathbb{E}X_{1i}X_{2i}'(\mathbb{E}X_{2i}X_{2i}')^{-1}\mathbb{E}X_{2i}X_{1i}'\right)^{-1}\\
&= \sigma_1^2\left(\mathbb{E}X_{1i}X_{1i}'\right)^{-1}.
\end{aligned}$$

Since $\sigma_2^2 \geq 0$, $\hat{\beta}_{1n}$ has a smaller asymptotic variance and therefore should be preferred.

Problem 4. In the linear regression model $Y_i = X_i'\beta + U_i$, where β is unknown k-vector ($k > 2$), describe how to test whether $\beta_1^2 + \beta_2^2 = 1$ (the sum of squares of the first two elements of β equals one) by answering the following questions:

(i) Describe how to construct the Wald statistic.

(ii) What is the distribution of the Wald statistic under the null?

(iii) Define the decision rule.

(iv) Describe the asymptotic behavior of the Wald statistic under the alternative, i.e. when $\beta_1^2 + \beta_2^2 \neq 1$.

Assume that $\left\{(Y_i, X_i')' : i \geq 1\right\}$ are iid, U_i and the elements of X_i have finite fourth moments, the matrix $\mathbb{E}X_iX_i'$ is positive definite, and $\mathbb{E}\left(U_iX_i\right) = 0$.

Solution. Let $\hat{\beta}_n = (\hat{\beta}_{1n}, \hat{\beta}_{2n}, \ldots, \hat{\beta}_{kn})'$ denote the OLS estimators of the coefficients.

(i) By the assumptions,

$$\sqrt{n}(\hat{\beta}_n - \beta) \to_d N\left(0, V\right),$$

where

$$V = \left(\mathbb{E}X_iX_i'\right)^{-1}\mathbb{E}U_i^2X_iX_i'\left(\mathbb{E}X_iX_i'\right)^{-1}.$$

A consistent estimator of V is

$$\hat{V}_n = \left(n^{-1}\sum_{i=1}^n X_iX_i'\right)^{-1} n^{-1}\sum_{i=1}^n \hat{U}_i^2X_iX_i'\left(n^{-1}\sum_{i=1}^n X_iX_i'\right)^{-1},$$

where

$$\hat{U}_i = Y_i - X_i'\hat{\beta}_n.$$

Define

$$h(\beta) = \beta_1^2 + \beta_2^2 - 1.$$

Note that $h : \mathbb{R}^k \to \mathbb{R}$. Next, we have

$$\frac{\partial h(\beta)}{\partial \beta'} = \begin{pmatrix} 2\beta_1 & 2\beta_2 & 0 & \ldots & 0 \end{pmatrix}.$$

The Wald statistic is defined as

$$W_n = nh(\hat{\beta}_n)'\left(\frac{\partial h(\hat{\beta}_n)}{\partial \beta'}\hat{V}_n\frac{\partial h(\hat{\beta}_n)'}{\partial \beta}\right)^{-1}h(\hat{\beta}_n).$$

(ii) By Delta method,

$$\sqrt{n}\left(h(\hat{\beta}_n) - h(\beta)\right) \to_d N\left(0, \frac{\partial h(\beta)}{\partial \beta'} V \frac{\partial h(\beta)'}{\partial \beta}\right).$$

By CMT,

$$\left(\frac{\partial h(\hat{\beta}_n)}{\partial \beta'} \hat{V}_n \frac{\partial h(\hat{\beta}_n)'}{\partial \beta}\right)^{-1} \to_p \left(\frac{\partial h(\beta)}{\partial \beta'} V \frac{\partial h(\beta)'}{\partial \beta}\right)^{-1}$$

and under H_0, by CMT and Slutsky's theorem,

$$nh(\hat{\beta}_n)' \left(\frac{\partial h(\hat{\beta}_n)}{\partial \beta'} \hat{V}_n \frac{\partial h(\hat{\beta}_n)'}{\partial \beta}\right)^{-1} h(\hat{\beta}_n) \to_d Z \cdot \left(\frac{\partial h(\beta)}{\partial \beta'} V \frac{\partial h(\beta)'}{\partial \beta}\right)^{-1} \cdot Z \sim \chi_1^2,$$

where $Z \sim N\left(0, \frac{\partial h(\beta)}{\partial \beta'} V \frac{\partial h(\beta)'}{\partial \beta}\right)$. The asymptotic distribution of W_n under the null hypothesis is χ_1^2.

(iii) The econometrician should reject H_0 when $W_n > \chi^2_{1,1-\alpha}$.

(iv) Under H_1,

$$\frac{W_n}{n} = h(\hat{\beta}_n)' \left(\frac{\partial h(\hat{\beta}_n)}{\partial \beta'} \hat{V}_n \frac{\partial h(\hat{\beta}_n)'}{\partial \beta}\right)^{-1} h(\hat{\beta}_n) \to_p \left(\beta_1^2 + \beta_2^2 - 1\right)^2 \left(\frac{\partial h(\beta)}{\partial \beta'} V \frac{\partial h(\beta)'}{\partial \beta}\right)^{-1} > 0.$$

Problem 5. Consider the following IV regression model:

$$\begin{aligned} Y_i &= \beta X_i + U_i, \\ \mathbb{E}Z_i U_i &= c, \end{aligned}$$

where β is the unknown scalar parameter, Y_i, X_i, Z_i are the dependent variable, single endogenous regressor, and single instrumental variable respectively, and c some constant (may be different from zero). Assume that c is *known*.

(i) Show that β is identified as

$$\beta = \frac{\mathbb{E}Z_i Y_i - c}{\mathbb{E}Z_i X_i}.$$

What do you have to assume about the relationship between the endogenous regressor and instrument for this identification result to be valid?

(ii) Consider the following estimator of β:

$$\hat{\beta}_{n,c} = \frac{n^{-1}\sum_{i=1}^n Z_i Y_i - c}{n^{-1}\sum_{i=1}^n Z_i X_i}.$$

Assuming that data are iid and that the assumption imposed in (i) holds, show consistency of $\hat{\beta}_{n,c}$.

(iii) Establish the asymptotic normality of $\hat{\beta}_{n,c}$ and show that its asymptotic variance is given by:

$$V_c = \frac{\mathbb{E}Z_i^2 U_i^2 - c^2}{(\mathbb{E}Z_i X_i)^2}$$

What additional assumptions about Z_i and U_i do you have to make to justify the result? Hint: Define $W_i = Z_i U_i - c$. Can the CLT be applied to $n^{-1/2}\sum_{i=1}^n W_i$?

(iv) Define $\hat{U}_i = Y_i - \hat{\beta}_{n,c} X_i$, and consider the following estimator of V_c:

$$\hat{V}_{n,c} = \frac{n^{-1}\sum_{i=1}^{n} Z_i^2 \hat{U}_i^2 - c^2}{(n^{-1}\sum_{i=1}^{n} Z_i X_i)^2}.$$

Show that $\hat{V}_{n,c}$ is a consistent estimator of V_c. Assume that the following moments exist: $\mathbb{E}Z_i^2 X_i^2$, $\mathbb{E}Z_i^2 X_i U_i$.

Solution.

(i) $c = \mathbb{E}Z_i U_i = \mathbb{E}Z_i(Y_i - \beta X_i) = \mathbb{E}Z_i Y_i - \beta \mathbb{E}Z_i X_i$. Assuming that $\mathbb{E}Z_i X_i \neq 0$, we obtain $\beta = \frac{\mathbb{E}Z_i Y_i - c}{\mathbb{E}Z_i X_i}$.

(ii) Substituting the model for Y_i in the formula for the estimator, we obtain:

$$\hat{\beta}_n = \beta + \frac{n^{-1}\sum_{i=1}^{n} Z_i U_i - c}{n^{-1}\sum_{i=1}^{n} Z_i X_i}. \tag{1}$$

By the WLLN,

$$\begin{aligned} n^{-1}\sum_{i=1}^{n} Z_i U_i \quad &\to_p \quad \mathbb{E}Z_i U_i = c, \\ n^{-1}\sum_{i=1}^{n} Z_i X_i \quad &\to_p \quad \mathbb{E}Z_i X_i \neq 0. \end{aligned}$$

Hence, by Continuous Mapping Theorem,

$$\hat{\beta}_n \to_p \beta + \frac{c-c}{\mathbb{E}Z_i X_i} = \beta \quad .$$

(iii) Re-write (1) as

$$n^{1/2}(\hat{\beta}_n - \beta) = \frac{n^{-1/2}\sum_{i=1}^{n}(Z_i U_i - c)}{n^{-1}\sum_{i=1}^{n} Z_i X_i}. \tag{2}$$

Note that $\mathbb{E}(Z_i U_i - c) = 0$. Thus, if $\mathbb{E}(Z_i U_i - c)^2 < \infty$, the CLT can be applied to obtain:

$$n^{-1/2}\sum_{i=1}^{n}(Z_i U_i - c) \to_d N(0, \mathbb{E}(Z_i U_i - c)^2).$$

Since $\mathbb{E}(Z_i U_i - c)^2 = \mathbb{E}Z_i^2 U_i^2 - c^2$, it follows from (2):

$$n^{1/2}(\hat{\beta}_n - \beta) \to_d \frac{N(0, \mathbb{E}Z_i^2 U_i^2 - c^2)}{\mathbb{E}Z_i X_i} =^d N\left(0, \frac{\mathbb{E}Z_i^2 U_i^2 - c^2}{(\mathbb{E}Z_i X_i)^2}\right).$$

(iv) Since $(n^{-1}\sum_{i=1}^{n} Z_i X_i)^2 \to_p (\mathbb{E}Z_i X_i)^2$, it remains to show that $n^{-1}\sum_{i=1}^{n} Z_i^2 \hat{U}_i^2 \to_p EZ_i^2 U_i^2$. Write $\hat{U}_i^2 = U_i^2 + (\hat{\beta}_{n,c} - \beta)^2 X_i^2 - 2(\hat{\beta}_{n,c} - \beta) X_i U_i$, so that

$$n^{-1}\sum_{i=1}^{n} Z_i^2 \hat{U}_i^2 = R_{1,n} + R_{2,n} - 2R_{3,n},$$

where

$$R_{1,n} \quad = \quad n^{-1}\sum_{i=1}^{n} Z_i^2 U_i^2 \to_p \mathbb{E}Z_i^2 U_i^2,$$

$$
\begin{aligned}
R_{2,n} &= (\hat{\beta}_{n,c}-\beta)^2 n^{-1}\sum_{i=1}^{n} Z_i^2 X_i^2,\\
R_{3,n} &= (\hat{\beta}_{n,c}-\beta) n^{-1}\sum_{i=1}^{n} Z_i^2 X_i U_i.
\end{aligned}
$$

By the WLLN and since $\mathbb{E}Z_i^2X_i^2$ and $\mathbb{E}Z_i^2X_iU_i$ exist, we have $n^{-1}\sum_{i=1}^n Z_i^2X_i^2 \to_p \mathbb{E}Z_i^2X_i^2$, and $n^{-1}\sum_{i=1}^n Z_i^2X_iU_i \to_p \mathbb{E}Z_i^2X_iU_i$. It follows that $R_{2,n}\to_p 0$ and $R_{3,n}\to_p 0$ because $\hat{\beta}_{n,c}-\beta \to_p 0$.

Problem 6. (a) Let $X_m \sim t_m$, i.e. X_m is a t-distributed random variable with m degrees of freedom. Show that $X_m \to_d N(0,1)$ as $m\to\infty$. Hints: Use the definition of the t-distribution and the WLLN. (b) Let $X_{q,m}\sim F_{q,m}$, i.e. $X_{q,m}$ is an F-distributed random variable with q and m degrees of freedom. Find the limiting distribution of $X_{q,m}$ as $m\to\infty$. Hints: Use the definition of the F-distribution and the WLLN.

Solution. By the definition of a t-distribution, $X_m = Z_0/\sqrt{(Z_1^2+\ldots+Z_m^2)/m}$, where $Z_0, Z_1,\ldots,Z_m$ are iid $N(0,1)$. As $m\to\infty$,

$$\frac{1}{m}(Z_1^2+\ldots+Z_m^2)\to_p \mathbb{E}Z_1^2 = 1,$$

and therefore by Slutsky's theorem,

$$X_m \to_p Z_0 \sim N(0,1).$$

Using the definition of $F_{q,m}$ distribution, we can write:

$$X_{q,m} = \frac{Y/q}{m^{-1}\sum_{j=1}^m Z_j^2},$$

where $Z_1,\ldots,Z_m$ are iid $N(0,1)$, and

$$Y\sim\chi_q^2$$

and independent from Z's. Since $\mathbb{E}Z_j^2 = 1$, by WLLN,

$$m^{-1}\sum_{j=1}^m Z_j^2 \to_p 1.$$

It follows that

$$X_{q,m}\to_d Y/q.$$

The result can be alternatively stated as

$$qF_{q,m}\to_d \chi_q^2.$$

Problem 7. Suppose that the econometrician has data on random variables X_i and Y_i generated from the following model:

$$
\begin{aligned}
Y_i &= X_i^3+\varepsilon_i,\\
\mathbb{E}(\varepsilon_i|X_i) &= 0.
\end{aligned}
$$

The true model is unknown to the econometrician, and he estimates a linear regression of Y_i against a constant and X_i:

$$Y_i = \hat{\beta}_{0,n}+\hat{\beta}_{1,n}X_i+\hat{U}_i,$$

where $\hat{\beta}_{0,n}$ and $\hat{\beta}_{1,n}$ are the OLS estimators, and $\hat{U}_i$ denotes the OLS residuals. Suppose that data are iid and X_i's are $N(0,1)$. Find the probability limits of $\hat{\beta}_{0,n}$ and $\hat{\beta}_{1,n}$. Does the linear regression model correctly capture the sign of the true marginal effect of X_i on Y_i? Hints: Since $X_i\sim N(0,1)$, $\mathbb{E}X_i^3=0$ and $\mathbb{E}X_i^4=3$.

Solution. Let β_0 and β_1 denote the probability limits of the corresponding OLS estimators. Recall that β_0 and β_1 are the coefficients in the best linear approximation of $\mathbb{E}(Y_i|X_i) = X_i^3$ by linear functions, i.e. β_0 and β_1 are the minimizers of $\mathbb{E}\left(X_i^3 - b_0 - b_1 X_i\right)^2$ and therefore solve the following first-order conditions:

$$
\begin{aligned}
\mathbb{E}\left(X_i^3 - \beta_0 - \beta_1 X_i\right) &= 0, \\
\mathbb{E}\left(X_i^3 - \beta_0 - \beta_1 X_i\right) X_i &= 0.
\end{aligned}
$$

Since $X_i \sim N(0,1)$, $\mathbb{E}X_i^3 = \mathbb{E}X_i = 0$, and therefore the first equation implies that $\beta_0 = 0$. From the second equation, we have

$$
0 = \mathbb{E}X_i^4 - \beta_1 \mathbb{E}X_i^2 = 3 - \beta_1,
$$

and therefore $\beta_1 = 3$. The marginal effect of X_i is $3X_i^2 > 0$. Hence, linear regression correctly captures the marginal effect in the true model.

Problem 8. In the linear regression model $Y_i = \boldsymbol{X}_i'\boldsymbol{\beta} + e_i$, where $\boldsymbol{\beta}$ is unknown k-vector ($k > 5$), describe how to test a joint null hypothesis $\beta_1\beta_4 = 1$ and $\beta_2\beta_3 = \beta_5$ (agains the alternative that at least one of the two restrictions is false) by answering the following questions:

(i) Describe how to construct the Wald statistic.

(ii) What is the distribution of the Wald statistic under the null?

(iii) Define the decision rule.

Solution. Let $\hat{\beta}_n = (\hat{\beta}_{1n}, \hat{\beta}_{2n}, \ldots, \hat{\beta}_{5n}, \ldots, \hat{\beta}_{kn})'$ denote the OLS estimators of the coefficients.

(i) By the assumptions,

$$
\sqrt{n}(\hat{\beta}_n - \beta) \to_d N(0, V),
$$

where

$$
V = \left(\mathbb{E}X_iX_i'\right)^{-1} \mathbb{E}U_i^2 X_i X_i' \left(\mathbb{E}X_iX_i'\right)^{-1}.
$$

A consistent estimator of V is

$$
\hat{V}_n = \left(n^{-1}\sum_{i=1}^n X_iX_i'\right)^{-1} n^{-1}\sum_{i=1}^n \hat{U}_i^2 X_iX_i' \left(n^{-1}\sum_{i=1}^n X_iX_i'\right)^{-1},
$$

where

$$
\hat{U}_i = Y_i - X_i'\hat{\beta}_n.
$$

Define

$$
h(\beta) = \begin{pmatrix} h_1(\beta) \\ h_2(\beta) \end{pmatrix} = \begin{pmatrix} \beta_1\beta_4 - 1 \\ \beta_2\beta_3 - \beta_5 \end{pmatrix}.
$$

Note that $h : \mathbb{R}^k \to \mathbb{R}^2$. Next, we have

$$
\begin{aligned}
\frac{\partial h(\beta)}{\partial \beta'} &= \begin{pmatrix} \frac{\partial h_1(\beta)}{\partial \beta_1} & \frac{\partial h_1(\beta)}{\partial \beta_2} & \frac{\partial h_1(\beta)}{\partial \beta_3} & \frac{\partial h_1(\beta)}{\partial \beta_4} & \frac{\partial h_1(\beta)}{\partial \beta_5} & 0 & \ldots & 0 \\ \frac{\partial h_2(\beta)}{\partial \beta_1} & \frac{\partial h_2(\beta)}{\partial \beta_2} & \frac{\partial h_2(\beta)}{\partial \beta_3} & \frac{\partial h_2(\beta)}{\partial \beta_4} & \frac{\partial h_2(\beta)}{\partial \beta_5} & 0 & \ldots & 0 \end{pmatrix} \\
&= \begin{pmatrix} \beta_4 & 0 & 0 & \beta_1 & 0 & 0 & \ldots & 0 \\ 0 & \beta_3 & \beta_2 & 0 & -1 & 0 & \ldots & 0 \end{pmatrix}.
\end{aligned}
$$

The Wald statistic is defined as

$$
W_n = nh(\hat{\beta}_n)' \left(\frac{\partial h(\hat{\beta}_n)}{\partial \beta'} \hat{V}_n \frac{\partial h(\hat{\beta}_n)'}{\partial \beta}\right)^{-1} h(\hat{\beta}_n).
$$

(ii) Similarly to Problem 4(ii), the asymptotic distribution of W_n under the null hypothesis is χ^2_2.

(iii) The econometrician should reject H_0 when $W_n > \chi^2_{2,1-\alpha}$.

Problem 9. Consider the linear regression model $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta}+\boldsymbol{e}$, where $\boldsymbol{X}$ is the $n\times k$ matrix of regressors, $\boldsymbol{Y}$ is the n-vector of observations on the dependent variable, and $\boldsymbol{\beta} \in \mathbb{R}^k$ is the vector of unknown parameters. Let $\boldsymbol{Z}$ be the $n \times k$ matrix of instruments. Assume that:

- $\boldsymbol{X}$ and $\boldsymbol{Z}$ are strongly exogenous: $\mathbb{E}\left(\boldsymbol{e}|\boldsymbol{X},\boldsymbol{Z}\right) = \mathbf{0}$.
- $\boldsymbol{e}$ is homoskedastic: $\mathbb{E}\left(\boldsymbol{e}\boldsymbol{e}'|\boldsymbol{X},\boldsymbol{Z}\right) = \sigma^2\boldsymbol{I}_n$.
- $\boldsymbol{X}$ and $\boldsymbol{Z}'\boldsymbol{X}$ have rank k.

Let $\hat{\boldsymbol{\beta}} = \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{Y}$ and $\tilde{\boldsymbol{\beta}} = \left(\boldsymbol{Z}'\boldsymbol{X}\right)^{-1}\boldsymbol{Z}'\boldsymbol{Y}$ be the OLS and IV estimators of $\boldsymbol{\beta}$ respectively.

(i) Show that $\mathbb{E}\left(\boldsymbol{e}|\boldsymbol{X}\right) = \mathbf{0}$ and $\mathbb{E}\left(\boldsymbol{e}\boldsymbol{e}'|\boldsymbol{X}\right) = \sigma^2\boldsymbol{I}_n$.

(ii) Show that the OLS and IV estimators are unbiased.

(iii) Find the exact finite sample conditional variances of $\hat{\boldsymbol{\beta}}$ and $\tilde{\boldsymbol{\beta}}$: $\mathrm{Var}\left(\hat{\boldsymbol{\beta}}|\boldsymbol{X},\boldsymbol{Z}\right)$ and $\mathrm{Var}\left(\tilde{\boldsymbol{\beta}}|\boldsymbol{X},\boldsymbol{Z}\right)$. Show that

$$\begin{aligned}&\mathrm{Var}\left(\tilde{\boldsymbol{\beta}}|\boldsymbol{X},\boldsymbol{Z}\right) - \mathrm{Var}\left(\hat{\boldsymbol{\beta}}|\boldsymbol{X},\boldsymbol{Z}\right)\\ =&\sigma^2\left(\boldsymbol{Z}'\boldsymbol{X}\right)^{-1}\boldsymbol{Z}'\left(\boldsymbol{I}_n - \boldsymbol{X}\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\right)\boldsymbol{Z}\left(\boldsymbol{X}'\boldsymbol{Z}\right)^{-1}.\end{aligned}$$

(iv) When regressors are exogenous, should the econometrician use IV or OLS ? Explain why using the result in in part (iii).

Solution.

(i) The results follow by the law of iterated expectation:

$$\begin{aligned}\mathbb{E}(\boldsymbol{e}|\boldsymbol{X}) &= \mathbb{E}\left(\mathbb{E}(\boldsymbol{e}|\boldsymbol{X},\boldsymbol{Z})\,|\boldsymbol{X}\right)\\ &= \mathbb{E}\left(\mathbf{0}|\boldsymbol{X}\right)\\ &= \mathbf{0},\\ \mathbb{E}\left(\boldsymbol{e}\boldsymbol{e}'|\boldsymbol{X}\right) &= \mathbb{E}\left(\mathbb{E}\left(\boldsymbol{e}\boldsymbol{e}'|\boldsymbol{X},\boldsymbol{Z}\right)|\boldsymbol{X}\right)\\ &= \mathbb{E}\left(\sigma^2\boldsymbol{I}_n|\boldsymbol{X}\right)\\ &= \sigma^2\boldsymbol{I}_n.\end{aligned}$$

(ii) Write

$$\begin{aligned}\hat{\boldsymbol{\beta}} &= \boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{e},\\ \tilde{\boldsymbol{\beta}} &= \boldsymbol{\beta} + (\boldsymbol{Z}'\boldsymbol{X})^{-1}\boldsymbol{Z}'\boldsymbol{e}.\end{aligned}$$

The results follow since

$$\begin{aligned}\mathbb{E}(\boldsymbol{X}'\boldsymbol{e}|\boldsymbol{X},\boldsymbol{Z}) &= \boldsymbol{X}'\mathbb{E}(\boldsymbol{e}|\boldsymbol{X},\boldsymbol{Z}) = \mathbf{0},\\ \mathbb{E}(\boldsymbol{Z}'\boldsymbol{e}|\boldsymbol{X},\boldsymbol{Z}) &= \boldsymbol{Z}'\mathbb{E}(\boldsymbol{e}|\boldsymbol{X},\boldsymbol{Z}) = \mathbf{0}.\end{aligned}$$

(iii) For the IV estimator,

$$\begin{aligned}\mathrm{Var}(\tilde{\boldsymbol{\beta}}|\boldsymbol{X},\boldsymbol{Z}) &= \mathrm{Var}\left((\boldsymbol{Z}'\boldsymbol{X})^{-1}\boldsymbol{Z}'\boldsymbol{e}|\boldsymbol{X},\boldsymbol{Z}\right) \\ &= (\boldsymbol{Z}'\boldsymbol{X})^{-1}\boldsymbol{Z}'\mathrm{Var}(\boldsymbol{e}|\boldsymbol{X},\boldsymbol{Z})\boldsymbol{Z}(\boldsymbol{X}'\boldsymbol{Z})^{-1} \\ &= \sigma^2(\boldsymbol{Z}'\boldsymbol{X})^{-1}\boldsymbol{Z}'\boldsymbol{Z}(\boldsymbol{X}'\boldsymbol{Z})^{-1}.\end{aligned}$$

For the OLS estimator, we have the usual expression:

$$\mathrm{Var}(\hat{\boldsymbol{\beta}}|\boldsymbol{X},\boldsymbol{Z}) = \sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}.$$

Lastly,

$$\begin{aligned}&(\boldsymbol{Z}'\boldsymbol{X})^{-1}\boldsymbol{Z}'\boldsymbol{Z}(\boldsymbol{X}'\boldsymbol{Z})^{-1} - (\boldsymbol{X}'\boldsymbol{X})^{-1} \\ =\ &(\boldsymbol{Z}'\boldsymbol{X})^{-1}\boldsymbol{Z}'\boldsymbol{Z}(\boldsymbol{X}'\boldsymbol{Z})^{-1} - (\boldsymbol{Z}'\boldsymbol{X})^{-1}(\boldsymbol{Z}'\boldsymbol{X})(\boldsymbol{X}'\boldsymbol{X})^{-1}(\boldsymbol{X}'\boldsymbol{Z})(\boldsymbol{X}'\boldsymbol{Z})^{-1} \\ =\ &\left(\boldsymbol{Z}'\boldsymbol{X}\right)^{-1}\boldsymbol{Z}'\left(\boldsymbol{I}_n - \boldsymbol{X}\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\right)\boldsymbol{Z}\left(\boldsymbol{X}'\boldsymbol{Z}\right)^{-1}.\end{aligned}$$

(iv) We showed that

$$\mathrm{Var}(\tilde{\boldsymbol{\beta}}|\boldsymbol{X},\boldsymbol{Z}) - \mathrm{Var}(\hat{\boldsymbol{\beta}}|\boldsymbol{X},\boldsymbol{Z}) = \sigma^2\boldsymbol{A}'\boldsymbol{M}_{\boldsymbol{X}}\boldsymbol{A},$$

where $\boldsymbol{M}_{\boldsymbol{X}} = \boldsymbol{I}_n - \boldsymbol{X}\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'$ is symmetric and idempotent and therefore positive semi-definite. Consequently, $\boldsymbol{A}'\boldsymbol{M}_{\boldsymbol{X}}\boldsymbol{A}$ is also positive semi-definite, and it follows that the OLS estimator has a smaller variance than the IV estimator. Since the OLS estimator is also unbiased with exogenous regressors, one should use OLS in this case. Note that the conclusion also follows by Gauss-Markov Theorem.

Problem 10. The econometrician is interested in estimating the effect of education on wages using individual data. The data set includes a wage variable, an education variable, and other individual characteristics. Suppose that the education variable only reports the highest level of education attained by individual i:

$$\mathrm{Education}_i = \begin{cases} 0, & \text{incomplete high school;} \\ 1, & \text{high school diploma;} \\ 2, & \text{some college;} \\ 3, & \text{college degree;} \\ 4, & \text{advanced degree.}\end{cases}$$

Consider the following equation:

$$\log Wage_i = \beta_0 + \beta_1 D_{1i} + \beta_2 D_{2i} + \beta_3 D_{3i} + \beta_4 D_{4i} + X_i'\gamma + U_i,$$

where $D_{1i} - D_{4i}$ are some dummy variables corresponding to education levels $1-4$ respectively.

(i) Show how to define the dummy variables so that their coefficients measure the expected change in log-wages (the expected percentage change) relatively to education level zero. Justify your answer.

(ii) Suppose that the dummy variables are defined as in (i). Explain how to test whether the individuals with some college education (but without a college degree) are expected to have the same earnings as the high school graduates.

(iii) Show how to define the dummy variables so that their coefficients measure the expected change in log-wages relatively to the preceding education level (the marginal effect of an education level). Justify your answer.

(iv) Suppose that the dummy variables are defined as in (iii). Explain how to tests the hypothesis described in (ii).

Solution.

(i) For $j = 1, \ldots, 4$, define
$$D_{ji} = 1\left(\text{Education}_i = j\right).$$
When $\text{Education}_i = 0$, we have the following equation:
$$\ln Wage_i = \beta_0 + X_i'\gamma + U_i.$$
When $\text{Education}_i = j$ for $j > 0$, we have
$$\ln Wage_i = \beta_0 + \beta_j + X_i'\gamma + U_i.$$
Assuming that
$$\mathbb{E}\left(U_i | \text{Education}_i\right) = 0,$$
we have
$$\mathbb{E}\left(\ln Wage_i | \text{Education}_i = j\right) - \mathbb{E}\left(\ln Wage_i | \text{Education}_i = 0\right) = \beta_j.$$

(ii) This hypothesis can be stated as
$$H_0 : \beta_1 - \beta_2 = 0.$$
It can be tested using a t-test with the following statistic:
$$T = \frac{\hat{\beta}_1 - \hat{\beta}_2}{\text{std.err}(\hat{\beta}_1 - \hat{\beta}_2)},$$
where $\text{std.err}(\hat{\beta}_1 - \hat{\beta}_2)$ stands for the standard error of the difference $\hat{\beta}_1 - \hat{\beta}_2$. To compute the standard error of $\hat{\beta}_1 - \hat{\beta}_2$, one would need the variances of the estimators for β_1 and β_2 as well as their covariance:
$$\text{Var}(\hat{\beta}_1 - \hat{\beta}_2) = \text{Var}(\hat{\beta}_1) + \text{Var}(\hat{\beta}_2) - 2\text{Cov}(\hat{\beta}_1, \hat{\beta}_2).$$

(iii) Define now
$$D_{ji} = 1\left(\text{Education}_i \geq j\right).$$
Using the same exogeneity condition as in (i), we still have
$$\begin{aligned}\mathbb{E}\left(\ln Wage_i | \text{Education}_i = 0, X_i\right) &= \beta_0 + X_i'\gamma,\\ \mathbb{E}\left(\ln Wage_i | \text{Education}_i = 1, X_i\right) &= \beta_0 + \beta_1 + X_i'\gamma.\end{aligned}$$
However, now
$$\begin{aligned}\mathbb{E}\left(\ln Wage_i | \text{Education}_i = 2, X_i\right) &= \beta_0 + \beta_1 + \beta_2 + X_i'\gamma,\\ \mathbb{E}\left(\ln Wage_i | \text{Education}_i = 3, X_i\right) &= \beta_0 + \beta_1 + \beta_2 + \beta_3 + X_i'\gamma,\\ &\text{etc}\ldots\end{aligned}$$
Therefore, for $j = 1, \ldots, 4$,
$$\mathbb{E}\left(\ln Wage_i | \text{Education}_i = j\right) - \mathbb{E}\left(\ln Wage_i | \text{Education}_i = j - 1\right) = \beta_j.$$

(iv) The hypothesis now can be stated as
$$H_0 : \beta_2 = 0.$$
It can be tested with a t-test using
$$T = \frac{\hat{\beta}_2}{\text{std.err}(\hat{\beta}_2)}.$$

Problem 11. In this question, you will derive the asymptotic distribution of the OLS estimator under endogeneity. Consider the usual linear regression model $Y_i = X_i'\beta + U_i$, where β is a $k \times 1$ vector. Assume, however, that X_i's are endogenous:

$$\mathbb{E}X_iU_i = \mu \neq 0,$$

where μ is an unknown $k \times 1$ vector. Let $\hat{\beta}_n$ denote the OLS estimator of β. Make the following additional assumptions:

A1. Data are iid.

A2. $Q = \mathbb{E}X_iX_i'$ is finite and positive definite.

A3. $\mathbb{E}\left(U_i - X_i'\delta\right)^2 X_iX_i'$ is finite and positive definite, where $\delta = Q^{-1}\mu$.

(i) Find the probability limit of $\hat{\beta}_n$.

(ii) Re-write the model as $Y_i = X_i'(\beta + \delta) + (U_i - X_i'\delta)$ and find $\mathbb{E}X_i(U_i - X_i'\delta)$.

(iii) Using the result in (ii), derive the asymptotic distribution of $\hat{\beta}_n$ and find its asymptotic variance. Explain how this result differs from the asymptotic normality of OLS with exogenous regressors. Hint: To establish asymptotic normality, $\hat{\beta}_n$ must be properly re-centered based on the result in (i).

(iv) Can $\hat{\beta}_n$ and its asymptotic distribution be used for inference about β? Explain why or why not.

(v) Suppose that the errors U_i's are homoskedastic:

$$\mathbb{E}\left(U_i^2|X_i\right) = \sigma^2 = const.$$

Consider the usual estimator of the asymptotic variance of OLS designed for a model with homoskedastic errors and exogenous regressors:

$$n^{-1}\sum_{i=1}^{n}\left(Y_i - X_i'\hat{\beta}_n\right)^2\left(n^{-1}\sum_{i=1}^{n}X_iX_i'\right)^{-1}.$$

Is it consistent for the asymptotic variance of the OLS estimator if X_i's are in fact endogenous? Explain why or why not.

(vi) Continue to assume that U_i's are homoskedastic as in (v). Consider the usual heteroskedasticity-robust asymptotic variance estimator designed for a model with exogenous regressors:

$$\left(n^{-1}\sum_{i=1}^{n}X_iX_i'\right)^{-1}\left(n^{-1}\sum_{i=1}^{n}\left(Y_i - X_i'\hat{\beta}_n\right)^2 X_iX_i'\right)\left(n^{-1}\sum_{i=1}^{n}X_iX_i'\right)^{-1}.$$

Is it consistent for the asymptotic variance of the OLS estimator if X_i's are in fact endogenous? Explain why or why not.

Solution.

(i) Write

$$\begin{aligned}
\hat{\beta}_n &= \beta + \left(n^{-1}\sum_{i=1}^{n}X_iX_i'\right)^{-1} n^{-1}\sum_{i=1}^{n}X_iU_i \\
&\to_p \beta + Q^{-1}\mu \\
&= \beta + \delta,
\end{aligned}$$

where convergence of $n^{-1}\sum_{i=1}^{n}X_iX_i' \to_p Q$ and $n^{-1}\sum_{i=1}^{n}X_iU_i \to_p \mathbb{E}X_iU_i = \mu$ hold by the WLLN, and the result in the second line holds by CMT.

(ii)

$$
\begin{aligned}
\mathbb{E}X_i(U_i - X_i'\delta) &= \mathbb{E}X_iU_i - \mathbb{E}X_iX_i'Q^{-1}\mu \\
&= \mu - QQ^{-1}\mu \\
&= 0.
\end{aligned}
$$

(iii) Write

$$
\hat{\beta}_n - (\beta + \delta) = \left(n^{-1} \sum_{i=1}^{n} X_i X_i' \right)^{-1} n^{-1} \sum_{i=1}^{n} X_i \epsilon_i,
$$

where

$$
\epsilon_i = U_i - X_i'\delta
$$

and uncorrelated with X_i by the result in (ii). Furthermore, $X_i\epsilon_i$ satisfies the assumptions of the CLT. Hence, this is a regression with all the usual assumptions, however, it has a new regression coefficient $\beta + \delta$ and new errors ϵ_i's. We have:

$$
\sqrt{n} \left(\hat{\beta}_n - (\beta + \delta) \right) \to_d N \left(0, Q^{-1} \left(\mathbb{E} \left(U_i - X_i'\delta \right)^2 X_i X_i' \right) Q^{-1} \right).
$$

Comparing to the case with exogenous regressors, the center of the asymptotic distribution is shifted by δ. Also, the asymptotic variance depends on $X_i'\delta$ through $\mathbb{E} \left(U_i - X_i'\delta \right)^2 X_i X_i'$.

(iv) Asymptotic inference about β based on the OLS estimator will be invalid since the asymptotic distribution of the OLS estimator is centered at $\beta + \delta$. The OLS estimator can be only used for testing hypotheses about $\beta + \delta$.

(v) First, we need to describe the probability limit of the estimator proposed. Write:

$$
\begin{aligned}
n^{-1} \sum_{i=1}^{n} \left(Y_i - X_i'\hat{\beta}_n \right)^2 &= n^{-1} \sum_{i=1}^{n} \left((U_i - X_i'\delta) + X_i' \left(\beta + \delta - \hat{\beta}_n \right) \right)^2 \\
&= n^{-1} \sum_{i=1}^{n} \left(\epsilon_i + X_i' \left(\beta + \delta - \hat{\beta}_n \right) \right)^2,
\end{aligned}
$$

where

$$
\epsilon_i = U_i - X_i'\delta.
$$

In view of the result in (i), $\beta + \delta - \hat{\beta}_n \to_p 0$, and therefore

$$
n^{-1} \sum_{i=1}^{n} \left(Y_i - X_i'\hat{\beta}_n \right)^2 \to_p \mathbb{E}\epsilon_i^2.
$$

Hence, the proposed estimator converges in probability to $\mathbb{E} \left(U_i - X_i'\delta \right)^2 Q^{-1}$. This would be the same as the asymptotic variance in (iii) if the errors $\epsilon_i = U_i - X_i'\delta$ were homoskedastic.
It is given that U_i's are homoskedastic. However, even if U_i's are homoskedastic, $\epsilon_i = U_i - X_i'\delta$ would be heteroskedastic:

$$
\mathbb{E}(\epsilon_i^2 | X_i) = \sigma^2 + \left(X_i'\delta \right)^2 - 2\mathbb{E} \left(U_i | X_i \right) X_i'\delta \neq const,
$$

unless $\mathbb{E} \left(U_i | X_i \right) = 0.5 X_i'\delta$. Since $\delta = Q^{-1}\mu$, $Q = \mathbb{E}X_iX_i'$, and $\mu = \mathbb{E}X_iU_i$, the law of iterated expectation implies that if $\mathbb{E} \left(U_i | X_i \right) = 0.5 X_i'\delta$, then

$$
\begin{aligned}
\mu &= \mathbb{E}X_iU_i \\
&= \mathbb{E} \left(X_i \mathbb{E}(U_i | X_i) \right) \\
&= \mathbb{E} \left(X_i \times 0.5 X_i'\delta \right)
\end{aligned}
$$

$$
\begin{aligned}
&= 0.5Q\delta \\
&= 0.5Q \times Q^{-1}\mu \\
&= 0.5\mu.
\end{aligned}
$$

However, the only solution to $\mu = 0.5\mu$ is $\mu = 0$, which contradicts the assumption that $\mathbb{E}X_iU_i \neq 0$. It follows therefore that $\epsilon_i = U_i - X_i'\delta$ are heteroskedastic. Hence, the estimator would be inconsistent for the asymptotic variance of the OLS estimator.

(vi) The model $Y_i = X_i'(\beta+\delta) + (U_i - X_i'\delta)$ is the usual linear regression with weakly exogenous regressors. The OLS estimator consistently estimates $\beta+\delta$. Its asymptotic variance has the usual "sandwich" form. Hence, with additional technical assumptions such as finite fourth moments for X_i's and $U_i - X_i'\delta$, the estimator will be consistent.

Problem 12. Consider the following model:

$$SAT_i = \beta_0 + \beta_1 Size_i + \beta_2 Size_i^2 + X_i'\gamma + U_i,$$

where SAT_i is the SAT score for individual i, $Size_i$ is the size of the graduating class, and X_i is a vector of individual covariates. Assume that $Size_i$ and X_i are exogenous, and that our usual assumptions needed for consistency and asymptotic normality of OLS are satisfied. Assume further that $\beta_1 > 0$ and $\beta_2 < 0$. You may also suppose that unrestricted OLS estimates of β_1 are strictly positive, and that unrestricted estimates of β_2 are strictly negative.

(i) According to this model, what is the optimal high school size? Justify your answer. Propose a consistent estimator for the optimal high school size, and explain why it is consistent.

(ii) Find the asymptotic distribution of the estimator proposed in (i).

(iii) Explain step by step how to construct a confidence interval for the optimal high school size. The proposed confidence interval should have asymptotic coverage probability of $1-\alpha$ for some predetermined small value of α.

Solution.

(i) Maximizing SAT with respect to $Size$, we obtain that the optimal $Size$ is given by

$$S = -\frac{\beta_1}{2\beta_2}.$$

Note that $\beta_2 < 0$ and therefore different from zero by assumption. The optimal size can be estimated using the plug-in approach:

$$\hat{S}_n = -\frac{\hat{\beta}_{1,n}}{2\hat{\beta}_{2,n}},$$

where $\hat{\beta}_{1,n}$ and $\hat{\beta}_{2,n}$ are the OLS estimators of β_1 and β_2 respectively. Since the OLS estimators of β's are consistent, $\hat{S}_n \to_p S$ by Slutsky's.

(ii) Let

$$\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \gamma \end{pmatrix}$$

be the $3 + \dim(\gamma)$ vector of parameters, and let $\hat{\beta}_n$ denote its OLS estimator. We can assume that

$$\sqrt{n}\left(\hat{\beta}_n - \beta\right) \to_d N(0, V),$$

where V is the $(3+\dim(\gamma)) \times (3+\dim(\gamma))$ asymptotic variance matrix. We can write

$$S = h(\beta) = -\frac{\beta_1}{2\beta_2}.$$

The vector of partial derivatives is given by

$$\frac{\partial h(\beta)}{\partial \beta} = \begin{pmatrix} 0 \\ -\frac{1}{2\beta_2} \\ \frac{\beta_1}{2\beta_2^2} \\ 0_{\dim(\gamma)\times 1} \end{pmatrix}.$$

By delta method, we obtain the following result:

$$\sqrt{n}(\hat{S}_n - S) \to_d N(0, v),$$

where

$$v = \begin{pmatrix} 0 \\ -\frac{1}{2\beta_2} \\ \frac{\beta_1}{2\beta_2^2} \\ 0_{\dim(\gamma)\times 1} \end{pmatrix}' V \begin{pmatrix} 0 \\ -\frac{1}{2\beta_2} \\ \frac{\beta_1}{2\beta_2^2} \\ 0_{\dim(\gamma)\times 1} \end{pmatrix}.$$

(iii) Let $\hat{V}_n$ denote some consistent estimator of the asymptotic variance of OLS. In general, V is of "sandwich" form, so $\hat{V}_n$ must be heteroskedasticity robust. The asymptotic variance of $\hat{S}_n$ can be estimated consistently by

$$\hat{v}_n = \begin{pmatrix} 0 \\ -\frac{1}{2\hat{\beta}_{2,n}} \\ -\frac{\hat{\beta}_{1,n}}{2\hat{\beta}_{2,n}^2} \\ 0_{\dim(\gamma)\times 1} \end{pmatrix}' \hat{V}_n \begin{pmatrix} 0 \\ -\frac{1}{2\hat{\beta}_{2,n}} \\ -\frac{\hat{\beta}_{1,n}}{2\hat{\beta}_{2,n}^2} \\ 0_{\dim(\gamma)\times 1} \end{pmatrix}.$$

Using the result in (ii), a confidence interval for S with asymptotic coverage $1-\alpha$ can be constructed as

$$[\hat{S}_n - z_{1-\alpha/2}\sqrt{\hat{v}_n/n}, \hat{S}_n + z_{1-\alpha/2}\sqrt{\hat{v}_n/n}],$$

where $z_{1-\alpha/2}$ is the $(1-\alpha/2)$ quantile of the standard normal distribution.

Problem 13. Consider the following model:

$$Y_i = \beta + U_i,$$

where U_i are iid $N(0,1)$ random variables, $i = 1, \ldots, n$.

(i) Perform a t test. Test at 5% significance level $H_0 : \beta = 0$ against $H_1 : \beta \neq 0$.

(ii) Test at 5% significance level $H_0 : \beta \leq 0$ against $H_1 : \beta > 0$.

(iii) Find the p-values for the tests in (i) and (ii).

(iv) Compute the power of the tests in (i) and (ii) for the true value of $\beta = -0.15$, -0.10, -0.05, 0.05, 0.10, 0.15.

Solution.

(i) The 95% two-tailed confidence interval for β is equal to $CI_{95\%} = [-0.029, 0.363]$. To test the hull hypothesis at the 5% significance level, we have to check if $\beta = 0$ falls into this confidence interval. Since it is part of the confidence interval, we cannot reject H_0. Another way to perform this test is to compute the test statistic $\hat{Z} = \frac{\hat{\beta}-0}{\sqrt{\frac{\sigma^2}{n}}} = \frac{0.167-0}{0.1}$, which is equal to 1.67 and compare it to the critical value $z_{1-\frac{\alpha}{2}} = z_{0.975} = 1.96$ (where $\alpha = 0.05$). To reject the null, we need $|\hat{Z}| > 1.96$. This is not the case.

(ii) Because the null is $H_0 = \beta \leq 0$ and the alternative is $H_1 > 0$, the test is now a one-tailed test. Hence, we need to look at the $\Phi^{-1}(0.95)$ critical value, which is equal to 1.645. The test statistic $\hat{Z} = \frac{\hat{\beta}-0}{\sqrt{\frac{\sigma^2}{n}}} = \frac{0.167-0}{0.1}$ is equal to 1.67, which is larger than 1.645. We therefore reject the null in favour of the alternative that $\beta > 0$.

(iii) In both (i) and (ii), the test statistic is equal to 1.67. The p-value in (i) is $\Pr(|Z| \geq 1.67) = 2(1 - \Phi(1.67)) = 0.095$ and the p-value in (ii) is $\Pr(Z \geq 1.67) = 1 - \Phi(1.67) = 0.047$. Z is a standard normal random variable.

(iv) The power function for the test in (i) can be constructed as follows:

$$\begin{aligned}
\pi(\beta) &= \Pr\left(\left|\frac{\hat{\beta} - \beta + \beta - 0}{0.1}\right| > 1.96\right) \\
&= \Pr\left(\left|\frac{\hat{\beta} - \beta}{0.1} + \frac{\beta}{0.1}\right| > 1.96\right) \\
&= \Pr(|Z + 10\beta| > 1.96) \\
&= \Pr(Z + 10\beta > 1.96) + Pr(Z + 10\beta < -1.96)
\end{aligned}$$

Plugging in the values for β we obtain:

$$\pi(\beta) = \begin{cases} 0.3230 & \text{if } \beta = -0.15 \\ 0.1701 & \text{if } \beta = -0.10 \\ 0.0791 & \text{if } \beta = -0.05 \\ 0.0791 & \text{if } \beta = 0.05 \\ 0.1701 & \text{if } \beta = 0.10 \\ 0.3230 & \text{if } \beta = 0.15 \end{cases}$$

For (ii), the power function becomes $\pi(\beta) = \Pr(Z + 10\beta > 1.645)$. Hence,

$$\pi(\beta) = \begin{cases} 0.0008 & \text{if } \beta = -0.15 \\ 0.0041 & \text{if } \beta = -0.10 \\ 0.0160 & \text{if } \beta = -0.05 \\ 0.1261 & \text{if } \beta = 0.05 \\ 0.2595 & \text{if } \beta = 0.10 \\ 0.4424 & \text{if } \beta = 0.15 \end{cases}$$

Problem 14. Consider the following model:

$$\begin{aligned}
Y_i &= g(X_i) + V_i, \qquad (3) \\
\mathbb{E}(V_i | X_i) &= 0, \\
\mathbb{E}\left(V_i^2 | X_i\right) &= \sigma^2,
\end{aligned}$$

where $g : \mathbb{R}^k \to \mathbb{R}$ is some unknown nonlinear function.

(i) Assume that $\mathbb{E}X_i X_i'$ is finite and positive definite, and that $\mathbb{E}(g(X_i))^2 < \infty$. Define

$$\beta = \arg\min_{b \in \mathbb{R}^k} \mathbb{E}\left(g(X_i) - X_i' b\right)^2$$

. Show that $\beta = \left(\mathbb{E}X_i X_i'\right)^{-1} \mathbb{E}X_i g(X_i)$. (We say that $X_i'\beta$ is the best linear approximation of $g(X_i)$.)

(ii) Define $U_i = Y_i - X_i'\beta$. Using the results from part (i), show that $\mathbb{E}X_iU_i = 0$.

(iii) Is it true that $\mathbb{E}(U_i|X_i) = 0$?

(iv) Find $\mathbb{E}\left(U^2|X_i\right)$. Is U_i homoskedastic?

(v) Suppose that the econometrician observes iid data $\{(Y_i, X_i) : i = 1, \dots, n\}$ generated from model (3) and let $\hat{\beta}_n = \left(\sum_{i=1}^n X_iX_i'\right)^{-1} \sum_{i=1}^n X_iY_i$. Show that $\hat{\beta}_n$ is a consistent and asymptotically normal estimator of β defined in part (i). Justify any additional assumptions you have to make. What is the asymptotic variance of $\hat{\beta}_n$?

(vi) Suggest a consistent estimator of the asymptotic variance of $\hat{\beta}_n$.

Solution.

(i) Define the function $Q(b)$ as follows:

$$\begin{aligned} Q(b) &= \mathbb{E}\left[(g(X_i) - X_i'b)^2\right] \\ &= \mathbb{E}\left[(g(X_i) - X_i'b)'(g(X_i) - X_i'b)\right] \\ &= \mathbb{E}\left[g(X_i)^2 - 2g(X_i)X_i'b - b'X_iX_i'b\right] \end{aligned}$$

Now, minimize $Q(b)$ w.r.t b. To obtain a solution (candidate), compute the first order condition for this expression:

$$\frac{\partial Q(b)}{\partial b} = \mathbb{E}\left[-2X_ig(X_i) + 2X_iX_i'b\right] = 0.$$

From here it is easy to see that the solution to this problem will be given by $\beta = \mathbb{E}\left[X_iX_i'\right]^{-1}\mathbb{E}\left[X_ig(X_i)\right]$ provided that $\mathbb{E}(X_iX_i')$ is non–singular (which follows from the positive definiteness assumption given in the exercise).

(ii) Let $U_i = Y_i - X_i^{'}\beta$. Then,

$$\begin{aligned} \mathbb{E}[X_iU_i] &= \mathbb{E}[X_i(Y_i - X_i^{'}\beta)] \\ &= \mathbb{E}[X_ig(X_i) + X_iV_i - X_iX_i^{'}\beta] \\ &= \mathbb{E}[X_ig(X_i)] + \mathbb{E}[X_iV_i] - \mathbb{E}[X_iX_i^{'}]\beta] \\ &= \mathbb{E}[X_ig(X_i)] + \mathbb{E}[X_iV_i] - \mathbb{E}[X_iX_i^{'}]\mathbb{E}[X_iX_i^{'}]^{-1}\mathbb{E}[X_ig(X_i)] \\ &= \mathbb{E}[X_ig(X_i)] + \mathbb{E}[X_iV_i] - \mathbb{E}[X_ig(X_i)] \\ &= \mathbb{E}[X_iV_i] \\ &= 0 \end{aligned}$$

where the last equality follows from the LIE, $\mathbb{E}[X_iV_i] = \mathbb{E}[\mathbb{E}[X_iV_i|X_i]] = \mathbb{E}[X_i\mathbb{E}[V_i|X_i]] = 0$, since $\mathbb{E}[V_i|X_i] = 0$ by assumption.

(iii) For $\mathbb{E}[U_i|X_i]$, observe the following:

$$\begin{aligned} \mathbb{E}[U_i|X_i] &= \mathbb{E}[(Y_i - X_i^{'}\beta)|X_i] \\ &= \mathbb{E}[(g(X_i) + V_i - X_i^{'}\beta)|X_i] \\ &= g(X_i) + \mathbb{E}[V_i|X_i] - X_i^{'}\beta \\ &= g(X_i) - X_i^{'}\beta. \end{aligned}$$

(iv) For $\mathbb{E}[U_i^2|X_i]$,

$$\begin{aligned}
\mathbb{E}[U_i^2|X_i] &= \mathbb{E}[(Y_i - X_i^{'}\beta)^2|X_i] \\
&= \mathbb{E}[(Y_i^2 - 2Y_iX_i^{'}\beta + (X_i^{'}\beta)^2)|X_i] \\
&= \mathbb{E}[(g(X_i)^2 + 2g(X_i)V_i + V_i^2 - 2g(X_i)X_i^{'}\beta - 2V_iX_i^{'}\beta + (X_i^{'}\beta)^2)|X_i] \\
&= g(X_i)^2 + 2g(X_i)\mathbb{E}[V_i|X_i] + \mathbb{E}[V_i^2|X_i] - 2g(X_i)X_i^{'}\beta - 2X_i^{'}\beta\mathbb{E}[V_i|X_i] + (X_i^{'}\beta)^2 \\
&= g(X_i)^2 + \sigma^2 - 2g(X_i)X_i^{'}\beta + (X_i^{'}\beta)^2 \\
&= [g(X_i) - X_i^{'}\beta]^2 + \sigma^2
\end{aligned}$$

so U_i is not homoskedastic.

(v) Let $\hat{\beta}_n = \left(\sum_{i=1}^n X_iX_i^{'}\right)^{-1}\left(\sum_{i=1}^n X_iY_i\right)$ and recall that $U_i = Y_i - X_i'\beta$ or $Y_i = X_i'\beta + U_i$. Then,

$$\hat{\beta}_n = \beta + \left(\frac{1}{n}\sum_{i=1}^n X_iX_i'\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^n X_iU_i\right)$$

By assumption, $E[X_iX_i']$ is finite and p.d. Similarly, from part (ii) we have that $\mathbb{E}[X_iU_i] = 0$. Moreover, the fact that $\{(Y_i, X_i)\}_{i=1}^n$ is i.i.d. implies that so are $\{(X_iX_i^{'})\}_{i=1}^n$, $\{(X_iU_i^{'})\}_{i=1}^n$. Hence, by WLLN and the CMT:

$$\begin{aligned}
\frac{1}{n}\left(\sum_{i=1}^n X_iX_i'\right)^{-1} &\longrightarrow^p \mathbb{E}[X_iX_i^{'}]^{-1} \\
\frac{1}{n}\sum_{i=1}^n X_iU_i &\longrightarrow^p \mathbb{E}[X_iU_i] = 0
\end{aligned}$$

Putting all these results together and using CMT once more, we have that:

$$\begin{aligned}
\hat{\beta}_n - \beta &= \left(\frac{1}{n}\sum_{i=1}^n X_iX_i'\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^n X_iU_i\right) \\
&\longrightarrow^p \left(\mathbb{E}[X_iX_i']\right)^{-1}\mathbb{E}[U_iX_i] \\
&= 0
\end{aligned}$$

where $\beta = \left(\mathbb{E}[X_iX_i']\right)^{-1}\mathbb{E}[X_ig(X_i)]$ is the solution to the problem in (i). Therefore $\hat{\beta}_n$ is a consistent estimator of β.

To show that $\hat{\beta}_n$ is asymptotically normal, use $U_i = Y_i - X_i\beta$ (since $\{(Y_i, X_i)\}_{i=1}^n$ is iid, so is U_i) and assume the following: (i) $\mathbb{E}|X_{i,k}|^4 < \infty$; (ii) $\mathbb{E}|U_i|^4 < \infty$. Now, use Cauchy–Schwartz inequality to show that $\mathbb{E}|X_iX_i'U_i^2| < \infty$. As all conditions required to apply CLT are met, we have the following result:

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n X_iU_i\right) \longrightarrow^d N(0, \Omega)$$

where $\Omega = \mathbb{E}[U_i^2X_iX_i^{'}]$. Notice that, since $\mathbb{E}[U_i^2|X_i] \neq \sigma^2$, Ω can not be simplified to $\sigma^2\mathbb{E}[X_iX_i']$. Using Slutsky's Theorem, we are ready to prove the following:

$$\begin{aligned}
n^{1/2}(\hat{\beta}_n - \beta) &= \left(\sum_{i=1}^n X_iX_i'\right)^{-1} n^{1/2}\left(\sum_{i=1}^n X_iU_i\right) \\
&\longrightarrow^d N(0, M_{XX}^{-1}\Omega M_{XX}^{-1})
\end{aligned}$$

where $M_{XX} = \mathbb{E}(X_iX_i')$, which shows that $\hat{\beta}_n$ is asymptotically normal distributed.

(vi) Denote by $\hat{M}_{XX} = \frac{1}{n}\sum_{i=1}^n X_i X_i'$. Given the assumptions we have made and using CMT, we know that $\hat{M}_{XX}^{-1} \to_p M_{XX}^{-1}$. Let $\hat{\Omega}_n = \frac{1}{n}\sum_{i=1}^n \hat{U}_i^2 X_i X_i'$, where $\hat{U}_i = Y_i - X_i'\hat{\beta}_n$ and $\hat{\beta}_n$ is given as above.

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^n \hat{U}_i^2 X_i X_i' &= \frac{1}{n}\sum_{i=1}^n [U_i - X_i'(\hat{\beta}_n - \beta)]^2 X_i X_i' \\ &= \frac{1}{n}\sum_{i=1}^n U_i^2 X_i X_i' - 2\frac{1}{n}\sum_{i=1}^n \left(U_i X_i'(\hat{\beta}_n - \beta)\right) X_i X_i' \\ &+ \frac{1}{n}\sum_{i=1}^n \left((\hat{\beta}_n - \beta)' X_i X_i'(\hat{\beta}_n - \beta)\right) X_i X_i'\end{aligned}$$

Under the assumptions we have made so far, we know that the first term in the right hand side of this last expression will converge in probability to $\mathbb{E}[U_i^2 X_i X_i'] = \Omega$. It can also be shown that the last two terms will converge to zero in probability. It then follows that

$$\hat{\Omega}_n = \frac{1}{n}\sum_{i=1}^n \hat{U}_i^2 X_i X_i' \longrightarrow^p \mathbb{E}[U_i^2 X_i X_i'] = \Omega$$

Finally, using CMT we have that,

$$\hat{M}_{XX}^{-1}\hat{\Omega}_n \hat{M}_{XX}^{-1} \longrightarrow^p M_{XX}^{-1}\Omega M_{XX}^{-1}$$

and so, $\hat{M}_{XX}^{-1}\hat{\Omega}_n \hat{M}_{XX}^{-1}$ is a consistent estimator for $M_{XX}^{-1}\Omega M_{XX}^{-1}$.

Problem 15. Let $\hat{\theta}$ be an estimator of a scalar parameter θ such that $\hat{\theta} \sim N(\theta, V_\theta)$, and assume that V_θ is known. Consider testing $H_0 : \theta = \theta_0$ against $H_1 : \theta \neq \theta_0$ using the following two tests:
Test 1: Reject H_0 when $\left|(\hat{\theta} - \theta_0)/\sqrt{V_\theta}\right| > z_{1-\alpha/2}$.
Test 2: Reject H_0 when $(\hat{\theta} - \theta_0)/\sqrt{V_\theta} > z_{1-\alpha}$.

(i) Is Test 1 a valid level α test (the probability of type I error is α)?

(ii) Is Test 2 a valid level α test?

(iii) Suppose that $\alpha = 0.10$, $\theta_0 = -1$, and $V_\theta = 0.25$. Compute the power functions of the two tests for the following values $\theta = -2, -1.5, -1, -0.5, 0$.

(iv) Which of the two tests (if any) is preferred? Explain.

Solution.

(i) Let Z denote a standard normal random variable. When $\theta = \theta_0$,

$$\begin{aligned}\Pr\left(\left|(\hat{\theta} - \theta_0)/\sqrt{V_\theta}\right| > z_{1-\alpha/2}\right) &= \Pr\left(\left|(\hat{\theta} - \theta)/\sqrt{V_\theta}\right| > z_{1-\alpha/2}\right) \\ &= \Pr\left(|Z| > z_{1-\alpha/2}\right) \\ &= \alpha.\end{aligned}$$

Hence, Test 1 is a valid level α test.

(ii)

$$\begin{aligned}\Pr\left((\hat{\theta} - \theta_0)/\sqrt{V_\theta} > z_{1-\alpha}\right) &= \Pr\left((\hat{\theta} - \theta)/\sqrt{V_\theta} > z_{1-\alpha}\right) \\ &= \Pr\left(Z > z_{1-\alpha}\right) \\ &= \alpha.\end{aligned}$$

Hence, Test 2 is also a valid level α test.

(iii) Let $\pi_1(\theta)$ and $\pi_2(\theta)$ denote the power functions of the two tests. Note that in view of the results in (i) and (ii), $\pi_1(-1) = \pi_2(-1) = 0.10$.

The power function of Test 1 is given by

$$\begin{aligned}
\pi_1(\theta) &= \Pr\left(\left|Z + \frac{\theta - \theta_0}{\sqrt{V_\theta}}\right| > z_{1-\alpha/2}\right) \\
&= \Pr\left(\left|Z + \frac{\theta + 1}{0.5}\right| > 1.645\right) \\
&= \Pr\left(Z > 1.645 - \frac{\theta + 1}{0.5}\right) + \Pr\left(Z < -1.645 - \frac{\theta + 1}{0.5}\right) \\
&= \Pr\left(Z > -0.355 - 2\theta\right) + \Pr\left(Z < -3.645 - 2\theta\right).
\end{aligned}$$

Hence:

$$\begin{aligned}
\pi_1(-1.5) &= \pi_1(-0.5) = \Pr(Z > 0.645) + \Pr(Z < -2.645) = 0.2578 + 0.004 = 0.2618. \\
\pi_1(-2) &= \pi_1(0) = \Pr(Z > -0.355) + \Pr(Z < -3.645) = 0.6406.
\end{aligned}$$

The power function of Test 2 is given by

$$\begin{aligned}
\pi_2(\theta) &= \Pr\left(Z + \frac{\theta - \theta_0}{\sqrt{V_\theta}} > z_{1-\alpha}\right) \\
&= \Pr\left(Z + \frac{\theta + 1}{0.5} > 1.28\right) \\
&= \Pr\left(Z > -0.72 - 2\theta\right).
\end{aligned}$$

Hence:

$$\begin{aligned}
\pi_2(-2) &= \Pr(Z > 3.28) = 0.0005. \\
\pi_2(-1.5) &= \Pr(Z > 2.28) = 0.0113. \\
\pi_2(-0.5) &= \Pr(Z > 0.28) = 0.3807. \\
\pi_2(0) &= \Pr(Z > -0.72) = 0.7642.
\end{aligned}$$

(iv) Test 2 is more powerful for $\theta > -1$, while Test 1 is more powerful for $\theta < -1$. Therefore, neither test is uniformly more powerful. Note however, that Test 2 can detect $\theta < -1$ only with probabilities smaller than 0.10, i.e. its power for such values of θ is below the probability of rejecting H_0 when it is true. Thus, if we want good all around power, Test 1 is preferred.