Statistics II **Problem Set 3**
Professor: Marcelo J. Moreira **Solutions**
TA: Luan Borelli **August 30, 2024**

---

1. [16.25, LNs] *Consider the population model $y_t = X_t\beta + u_t$, where the product $X_t u_t$ is iid, with $\mathbb{E}(u_t|X_t) = 0$, $V(u_t|X_t) = \sigma^2$. Let $X_{T\times k} = \left[X_1', \dots, X_T'\right]'$ and suppose plim $\left(\frac{X'X}{n}\right) = \mathbb{E}(X_t'X_t) = \Sigma_{X'X}$, a positive definite matrix.*

(a) *Find the mistakes in the derivation of the asymptotic variance of $\sqrt{n}\left(\hat{\beta} - \beta\right)$ below:*

$$\begin{aligned} AV\left(\hat{\beta}\right) &= \text{plim}\left(\sqrt{n}\left(\hat{\beta}-\beta\right)\sqrt{n}\left(\hat{\beta}-\beta\right)'\right) \\ &= \text{plim}\left(\frac{X'X}{n}\right)^{-1} \text{plim}\left(\frac{X'uu'X}{n}\right) \text{plim}\left(\frac{X'X}{n}\right)^{-1} \\ &= \Sigma_{X'X}^{-1}\sigma^2\Sigma_{X'X}\Sigma_{X'X}^{-1} = \sigma^2\Sigma_{X'X}^{-1}. \end{aligned}$$

Solution. The derivation is wrong primarily because there is no reason for

$$\text{plim}\left(\frac{X'uu'X}{n}\right) = \sigma^2\Sigma_{X'X}$$

to be a valid probability limit. Indeed, notice that

$$\begin{aligned} X'uu'X &= (u'X)'(u'X) \\ &= \left(\sum_{i=1}^n u_i X_i\right)'\left(\sum_{j=1}^n u_j X_j\right) \\ &= \sum_{i=1}^n\sum_{j=1}^n u_i u_j X_i' X_j. \end{aligned}$$

The summands $u_i u_j X_i X_j'$ are not independent, thus the law of large number doesn't apply.

But, actually, the mistake arises from an even more fundamental conceptual misunderstanding regarding the definition of asymptotic variance. Asymptotic variance refers to the limit (as the sample size n approaches infinity) of the finite-sample variance of a random variable. The finite-sample variance of a generic random variable Z is defined as $V(Z) \equiv \mathbb{E}[(Z - \mathbb{E}[Z])(Z - \mathbb{E}[Z])']$. Therefore, given that $\mathbb{E}[\sqrt{n}(\hat{\beta} - \beta)] = 0$, the correct limit characterization of the asymptotic variance for $\sqrt{n}(\hat{\beta} - \beta)$ is

$$\text{AV}(\hat{\beta}) = \lim_{n\to\infty}\mathbb{E}[\sqrt{n}(\hat{\beta}-\beta)\sqrt{n}(\hat{\beta}-\beta)'] \quad \neq \text{plim}\left(\sqrt{n}\left(\hat{\beta}-\beta\right)\sqrt{n}\left(\hat{\beta}-\beta\right)'\right).$$

Using the fact that $\sqrt{n}(\hat{\beta} - \beta) = (X'X)^{-1}X'u$ we can write

$$\begin{aligned}
\text{AV}(\hat{\beta}) &= \lim_{n\to\infty} \mathbb{E}[n(X'X)^{-1}X'uu'X(X'X)^{-1}] \\
&= \lim_{n\to\infty} \mathbb{E}[n(X'X)^{-1}X'\mathbb{E}[uu'|X]X(X'X)^{-1}] \\
&= \lim_{n\to\infty} \mathbb{E}[n(X'X)^{-1}X'X(X'X)^{-1}]\sigma^2 \\
&= \lim_{n\to\infty} \mathbb{E}[n(X'X)^{-1}]\sigma^2 \\
&= \lim_{n\to\infty} \mathbb{E}\left[\left(n^{-1}\sum_{t=1}^{n} X_t'X_t\right)^{-1}\right]\sigma^2.
\end{aligned}$$

But at this point, we cannot proceed further without introducing additional assumptions. Typically, limits and expectations do not commute, meaning there is no simple way to eliminate the limit operator here. This is why directly computing the limit of a finite-sample variance is not the standard method for determining the asymptotic variance of a random variable. In the next section, we will derive the asymptotic variance using the conventional approach, which relies on the central limit theorem. □

(b) *Derive the asymptotic variance of $\sqrt{n}(\hat{\beta}-\beta)$ correctly.*

Solution. The OLS estimator can be written as

$$\hat{\beta} = (X'X)^{-1}X'y = (X'X)^{-1}X'(X\beta + u) = \beta + (X'X)^{-1}X'u.$$

Thus

$$\sqrt{n}(\hat{\beta}-\beta) = \sqrt{n}(X'X)^{-1}X'u = \left(n^{-1}\sum_{t=1}^{n} X_t'X_t\right)^{-1}\sqrt{n}\left(n^{-1}\sum_{t=1}^{n} X_t'u_t\right).$$

By the weak law of large numbers and the continuous mapping theorem,

$$\left(n^{-1}\sum_{t=1}^{n} X_t'X_t\right)^{-1} \xrightarrow{p} \mathbb{E}[X_t'X_t]^{-1}.$$

Further, by the central limit theorem,

$$\sqrt{n}\left(n^{-1}\sum_{t=1}^{n} X_t'u_t\right) \xrightarrow{d} N\left(0, \mathbb{E}[X_t'u_tu_t'X_t]\right).$$

Thus, by Slutsky's theorem,

$$\sqrt{n}(\hat{\beta}-\beta) \xrightarrow{d} N(0, \mathbb{E}[X_t'X_t]^{-2}\mathbb{E}[X_t'u_tu_t'X_t]).$$

Notice that by the law of iterated expectations we can write

$$\mathbb{E}[X_t'u_tu_t'X_t] = \mathbb{E}[X_t'\mathbb{E}[u_tu_t'|X_t]X_t] = \sigma^2\mathbb{E}[X_t'X_t].$$

Therefore the asymptotic variance of $\sqrt{n}(\hat{\beta}-\beta)$ is $\sigma^2\mathbb{E}[X_t'X_t]^{-1} = \sigma^2\Sigma_{X'X}^{-1}$. □

2. [16.31, LNs] *Consider the model*

$$y_i = x_i'\beta + u_i,$$

where (x_i', u_i) *are iid with* $u|x_i \sim N(0, \sigma^2)$. *We want to test* $H_0 : \beta = \beta_0$ *against* $H_1 : \beta \neq \beta_0$ *using the statistic*

$$W = \frac{(\hat{\beta} - \beta_0)'X'X(\hat{\beta} - \beta_0)}{\sigma^2}.$$

(a) *Is this statistic pivotal? Explain your answer.*

Solution. Under the null, yes. Observe that $\hat{\beta} - \beta_0 = (X'X)^{-1}X'u$, whence

$$W = (X'u)'(X'X\sigma^2)^{-1}X'u.$$

Notice that $X'u \sim N(0, X'X\sigma^2)$. Therefore $W \sim \chi^2_k$, which does not depend on $(\beta, \sigma)'$. $\square$

(d) *Show that*

$$\hat{W} = \frac{(\hat{\beta} - \beta_0)'X'X(\hat{\beta} - \beta_0)/k}{e'e/(n-k)}$$

has a $F(k, n-k)$ *distribution under the null.*

Solution. We have already verified that, under the null, when divided by σ^2, the numerator term has a χ^2_k distribution. For the denominator, notice that $e'e/\sigma^2 = u'Mu/\sigma^2$. Spectral decomposition on M gives $M = H\Lambda H'$, where $H'H = I_n$. Thus

$$\begin{aligned}\frac{e'e}{\sigma^2} &= \left(\frac{H'u}{\sigma}\right)' \Lambda \left(\frac{H'u}{\sigma}\right) \\ &= \left(\frac{H'u}{\sigma}\right)' \begin{bmatrix} I_{n-k} & 0 \\ 0 & 0 \end{bmatrix} \left(\frac{H'u}{\sigma}\right).\end{aligned}$$

Let $w \equiv H'u/\sigma$. Notice that $w \sim N(0, I_n)$. Partition $w = (w_1', w_2')'$, where $w_1 \sim N(0, I_{n-k})$. It follows that

$$\frac{e'e}{\sigma^2} = w_1'w_1 \sim \chi^2_{n-k}.$$

Write

$$\hat{W} = \frac{\overbrace{\frac{(\hat{\beta} - \beta_0)'X'X(\hat{\beta} - \beta_0)}{\sigma^2}}^{\sim\chi^2_k}/k}{\underbrace{\frac{e'e}{\sigma^2}}_{\sim\chi^2_{n-k}}/(n-k)} \sim F(k, n-k).$$

Notice that numerator and denominator terms are independent. Indeed, the numerator term can be written as $u'Nu/\sigma^2$, while the denominator term can be written as $u'Mu/\sigma^2$. As N and M are orthogonal, both terms are orthogonal and hence independent (due to normality). Therefore $\hat{W} \sim F(k, n-k)$. $\square$

(e) *Show that $\hat{W}$ asymptotically has a chi-square distribution with k degrees of freedom.*

Solution. Consider the above expression for $\hat{W}$. Notice that $e'e/(n-k) \xrightarrow{p} \sigma^2$, so the denominator asymptotically converges to 1. Furthermore, observe that the numerator asymptotically converges in distribution to $Z'[\mathbb{E}[x_i x_i']^{-1}\sigma^2]^{-1}Z/k \sim \chi^2_k/k$, where $Z \sim N(0, \mathbb{E}[x_i x_i']^{-1}\sigma^2)$.[1] Therefore, by Slutsky, it follows that $k\hat{W} \xrightarrow{d} \chi^2_k$. □

---

3. [16.21, LNs] Suppose the auto-regressive linear model applies to:

$$y_t = \rho y_{t-1} + u_t, \quad t = 1, \dots, T$$

where $|\rho| < 1$, u_t are iid with $E(u_t) = 0$ and $E(u_t^2) = \sigma^2$.

(a) *Show that*

$$\textbf{[I]}: \quad \frac{1}{\sqrt{T}} \sum_{t=1}^{T} y_{t-1} u_t \xrightarrow{d} N\left(0, \frac{\sigma^4}{1-\rho^2}\right) \quad \textit{and} \quad \textbf{[II]}: \quad \frac{1}{T} \sum_{t=1}^{T} y_{t-1}^2 \xrightarrow{p} \frac{\sigma^2}{1-\rho^2}.$$

Solution. To solve this problem, I need to introduce the concept of martingale difference sequences (henceforth MDS), along with a law of large numbers and a central limit theorem applicable to this class of processes. It's important to note that the Lindeberg-Feller CLT cannot be applied here, as $\{y_{t-1}u_t\}$ and $\{y_{t-1}\}$ are not sequences of independent random variables.

Definition 1 (Martingale Difference Sequence)**.** Let $\{x_t\}_{t=1}^{\infty}$ denote a sequence of random scalars with $\mathbb{E}[x_t] = 0$ for all t, and Ω_t the information set available at date t, including current and lagged values of $\{x_t\}$. If $\mathbb{E}[x_t|\Omega_{t-1}] = 0$ for $t = 2, 3, \dots$, then $\{x_t\}$ is said to be a *martingale difference sequence* (MDS) with respect to $\{\Omega_t\}$.

Theorem 1 (A Law of Large Numbers for a MDS)**.** *Let $\bar{x}_t \equiv T^{-1}\sum_{t=1}^{T} x_t$ be the sample mean from a martingale difference sequence with $\mathbb{E}|x_t|^r < M$ for some $r > 1$ and $M < \infty$. Then $\bar{x}_t \xrightarrow{p} 0$.*

Theorem 2 (A Central Limit Theorem for a MDS)**.** *Let $\{x_t\}_{t=1}^{\infty}$ be a scalar martingale difference sequence with $\bar{x}_t = T^{-1}\sum_{t=1}^{T} x_t$. Suppose that (a) $\mathbb{E}[x_t^2] = \sigma_t^2 > 0$ with $T^{-1}\sum_{t=1}^{T}\sigma_t^2 \to \sigma^2 > 0$, (b) $\mathbb{E}|x_t|^r < \infty$ for some $r > 2$ and all t, and (c) $T^{-1}\sum_{t=1}^{T} x_t^2 \xrightarrow{p} \sigma^2$. Then $T^{1/2}\bar{x}_t \xrightarrow{d} N(0, \sigma^2)$.*

Now we are ready to move on to the proofs.

[I]: For $\frac{1}{T}\sum_{t=1}^{T} y_{t-1}^2 \xrightarrow{p} \frac{\sigma^2}{1-\rho^2}$, let $x_t \equiv y_t^2 - \mathbb{E}[y_t^2]$ and observe that for all $t \geq 2$,

$$\begin{aligned}\mathbb{E}[x_t|\Omega_{t-1}] &= \mathbb{E}[y_t^2 - \mathbb{E}[y_t^2] \mid y_{t-1}, y_{t-2}, \dots, y_1, y_0] \\ &= \mathbb{E}[y_t^2 \mid y_{t-1}, y_{t-2}, \dots, y_1, y_0] - \mathbb{E}[y_t^2 \mid y_{t-1}, y_{t-2}, \dots, y_1, y_0] = 0,\end{aligned}$$

[1] This is a particular case of Exercise 6, item **(a)**. Here, $R = I_k$ and $r = \beta_0$.

whence it follows that $\{x_t\}_{t=1}^{\infty}$ is a MDS and hence $\bar{x}_{t-1} \equiv T^{-1}\sum_{t=1}^{T} x_{t-1}$ the sample mean from a MDS. Now, assuming $\mathbb{E}[u_t^4] < \infty$, it is possible to show that $\mathbb{E}[y_t^4] < \infty$ for all t, and thus $\mathbb{E}|x_{t-1}|^r < \infty$ for $r = 2$. Therefore, by the law of large numbers for martingale difference sequences we conclude that

$$\bar{x}_{t-1} = T^{-1}\sum_{t=1}^{T}(y_{t-1}^2 - \mathbb{E}[y_{t-1}^2]) = \sum_{t=1}^{T} y_{t-1}^2 - \mathbb{E}[y_{t-1}^2] \xrightarrow{p} 0,$$

from which

$$T^{-1}\sum_{t=1}^{T} y_{t-1}^2 \xrightarrow{p} \mathbb{E}[y_{t-1}^2].$$

Now, using the lag operator L, observe that since

$$\begin{aligned} y_{t-1} &= \rho y_{t-2} + u_{t-1} \\ &= \rho L y_{t-1} + u_{t-1} \end{aligned}$$

we can isolate y_{t-1} and write

$$y_{t-1} = (1-\rho L)^{-1}u_{t-1} = \sum_{j=0}^{\infty}(\rho L)^j u_{t-1} = \sum_{j=0}^{\infty}\rho^j L^j u_{t-1} = \sum_{j=0}^{\infty}\rho^j u_{t-1-j},$$

by standard geometric series and lag operator algebra arguments. Finally, taking the variance of the expression above we obtain

$$V[y_{t-1}] = V\left[\sum_{j=0}^{\infty}\rho^j u_{t-1-j}\right] = \sum_{j=0}^{\infty}\rho^{2j}V\left[u_{t-1-j}\right] = \sum_{j=0}^{\infty}(\rho^2)^j\sigma^2 = \frac{\sigma^2}{1-\rho^2},$$

where the second equality follows from the independence of u_t, and the fourth from the fact that $\rho^2 < 1$. Therefore we can conclude that

$$T^{-1}\sum_{t=1}^{T} y_{t-1}^2 \xrightarrow{p} \frac{\sigma^2}{1-\rho^2}, \tag{1}$$

as desired.

[II]: For $\frac{1}{\sqrt{T}}\sum_{t=1}^{T} y_{t-1}u_t \xrightarrow{d} N\left(0, \frac{\sigma^4}{1-\rho^2}\right)$, let $x_t \equiv y_{t-1}u_t$. Then, $\{x_t\}_{t=1}^{\infty}$ is trivially a MDS with variance $\mathbb{E}[x_t^2] = \sigma^2\mathbb{E}[y_{t-1}^2]$ and with fourth moment $\mathbb{E}[u_t^4]\mathbb{E}[y_{t-1}^4] < \infty$ by the assumption that $\mathbb{E}[u_t^4] < \infty$ (which, as argued before, also implies $\mathbb{E}[y_t^4] < \infty$ for all t). Hence, if we can show that

$$T^{-1}\sum_{t=1}^{T} x_t^2 \xrightarrow{p} \mathbb{E}[x_t^2], \tag{2}$$

then the CLT for martingale difference sequences can be applied to show that

$$T^{1/2}\bar{x}_t \xrightarrow{d} N(0, \mathbb{E}[x_t^2]),$$

or, equivalently,

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T} y_{t-1}u_t \xrightarrow{d} N(0, \sigma^2\mathbb{E}[y_{t-1}^2]) = N\left(0, \frac{\sigma^4}{1-\rho^2}\right).$$

To verify (2), notice that

$$\begin{aligned} T^{-1}\sum_{t=1}^{T} x_t^2 &= T^{-1}\sum_{t=1}^{T} u_t^2 y_{t-1}^2 \\ &= T^{-1}\sum_{t=1}^{T}(u_t^2-\sigma^2)y_{t-1}^2 + T^{-1}\sum_{t=1}^{T}\sigma^2 y_{t-1}^2. \qquad (3) \end{aligned}$$

But $(u_t^2-\sigma^2)y_{t-1}^2$ is a MDS with finite second moment, so by the law of large numbers for martingale difference sequences,

$$T^{-1}\sum_{t=1}^{T}(u_t^2-\sigma^2)y_{t-1}^2 \xrightarrow{p} 0.$$

It further follows from (1) that

$$T^{-1}\sum_{t=1}^{T}\sigma^2 y_{t-1}^2 = \sigma^2 T^{-1}\sum_{t=1}^{T} y_{t-1}^2 \xrightarrow{p} \sigma^2\frac{\sigma^2}{1-\rho} = \frac{\sigma^4}{1-\rho}.$$

Thus (3) implies

$$T^{-1}\sum_{t=1}^{T} x_t^2 \xrightarrow{p} \frac{\sigma^4}{1-\rho^2},$$

as desired. □

(b) *Is the OLS estimator consistent for ρ? Show your calculations.*

Solution. Yes. Write

$$\begin{aligned} \hat{\rho} &= \left(\sum_{t=1}^{T} y_{t-1}^2\right)^{-1}\sum_{t=1}^{T} y_{t-1}y_t \\ &= \left(\sum_{t=1}^{T} y_{t-1}^2\right)^{-1}\sum_{t=1}^{T} y_{t-1}(\rho y_{t-1}+u_t) \\ &= \rho\left(\sum_{t=1}^{T} y_{t-1}^2\right)^{-1}\sum_{t=1}^{T} y_{t-1}^2 + \left(\sum_{t=1}^{T} y_{t-1}^2\right)^{-1}\sum_{t=1}^{T} y_{t-1}u_t \\ &= \rho + \left(\sum_{t=1}^{T} y_{t-1}^2\right)^{-1}\sum_{t=1}^{T} y_{t-1}u_t. \qquad (4) \end{aligned}$$

As verified before, $y_{t-1}u_t$ is a MDS with finite fourth moments. Thus by the law of large numbers for martingale difference sequences it follows that

$$\left(\sum_{t=1}^{T} y_{t-1}^2\right)^{-1} \sum_{t=1}^{T} y_{t-1}u_t \xrightarrow{p} 0,$$

whence

$$\hat{\rho} \xrightarrow{p} \rho,$$

as desired. □

(c) *Find the limiting distribution of* $\sqrt{T}(\hat{\rho} - \rho)$.

Solution. Using (4), write

$$\begin{aligned}
\sqrt{T}(\hat{\rho} - \rho) &= \sqrt{T}\left(\sum_{t=1}^{T} y_{t-1}^2\right)^{-1} \sum_{t=1}^{T} y_{t-1}u_t \\
&= \sqrt{T}\left(\sum_{t=1}^{T} y_{t-1}^2\right)^{-1} \sqrt{T}\frac{1}{\sqrt{T}} \sum_{t=1}^{T} y_{t-1}u_t \\
&= T\left(\sum_{t=1}^{T} y_{t-1}^2\right)^{-1} \frac{1}{\sqrt{T}} \sum_{t=1}^{T} y_{t-1}u_t \\
&= \left(T^{-1}\sum_{t=1}^{T} y_{t-1}^2\right)^{-1} \frac{1}{\sqrt{T}} \sum_{t=1}^{T} y_{t-1}u_t.
\end{aligned}$$

From the results obtained in **(a)** we have

$$\frac{1}{T}\sum_{t=1}^{T} y_{t-1}^2 \xrightarrow{p} \frac{\sigma^2}{1-\rho^2} \quad \text{and} \quad \frac{1}{\sqrt{T}}\sum_{t=1}^{T} y_{t-1}u_t \xrightarrow{d} Z \sim N\left(0, \frac{\sigma^4}{1-\rho^2}\right),$$

whence, by the continuous mapping and Slutsky's theorems,

$$\sqrt{T}(\hat{\rho} - \rho) \xrightarrow{d} \left(\frac{\sigma^2}{1-\rho^2}\right)^{-1} Z = N\left(0, \frac{(1-\rho^2)^2}{\sigma^4}\frac{\sigma^4}{1-\rho^2}\right) = N(0, 1-\rho^2).$$

□

4. [14.7, LNs] *Assume that* $X_i \overset{\text{iid}}{\sim} N(\theta, 1)$.

(a) *Show that the joint pdf is a special case of an exponential family.*

Solution. The single-parameter exponential family is the class of probability distributions whose probability density function (or probability mass function, in case of discrete distributions) can be expressed as

$$f(X;\theta) = C(\theta)\exp(A(\theta)T(X))h(X),$$

where $C(\theta)$, $A(\theta)$ and $h(X)$ are known functions and $C(\theta)$ is non-negative. Since X_i, for $i = 1, \ldots, n$, are i.i.d., the joint pdf of $X \equiv (X_1, \ldots, X_n)$ is given by

$$\begin{aligned} f(X;\theta) &= \prod_{i=1}^{n}(2\pi)^{-1/2}\exp\left(-\frac{(X_i-\theta)^2}{2}\right) \\ &= (2\pi)^{-n/2}\exp\left(-\frac{1}{2}\sum_{i=1}^{n}(X_i^2 - 2\theta X_i + \theta^2)\right) \\ &= (2\pi)^{-n/2}\exp\left(-\frac{n}{2}\theta\right)\exp\left(\theta\sum_{i=1}^{n}X_i\right)\exp\left(-\frac{1}{2}\sum_{i=1}^{n}X_i^2\right). \end{aligned}$$

By setting

$$C(\theta) \equiv \exp\left(\theta\sum_{i=1}^{n}X_i\right), \quad A(\theta) \equiv \theta, \quad T(X) \equiv \sum_{i=1}^{n}X_i, \quad h(X) \equiv (2\pi)^{-n/2}\exp\left(-\frac{1}{2}\sum_{i=1}^{n}X_i^2\right)$$

it becomes clear that this joint pdf is a special case of the exponential family. □

(b) *Show that the UMP test for* $H_0 : \theta \leq \theta_0$ *against* $H_1 : \theta > \theta_0$ *reject the null when* $\sqrt{n}(\bar{X}_n - \theta_0) > c_{1-\alpha}$, *where* $c_{1-\alpha}$ *is the* $1-\alpha$ *quantile of a standard normal.*

Solution. For distributions of the exponential family with $A(\theta)$ monotone increasing, there exists a UMP for testing $H_0 : \theta \leq \theta_0$ against $H_1 : \theta > \theta_0$ characterized by the critical region

$$C_X = \{X \mid T(X) > k\},$$

where k is determined by $\alpha = \int_{C_X} f(X;\theta_0)\,dx$.[2] In particular, from **(a)** we have $A(\theta) = \theta$, which is clearly monotone increasing, and $T(X) = \sum_{i=1}^{n} X_i$. Therefore the UMP test is characterized by the critical region

$$\begin{aligned} C_X &= \left\{X \;\middle|\; \sum_{i=1}^{n}X_i > k\right\} \\ &= \left\{X \;\middle|\; \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}X_i - \theta_0\right) > k'\right\}. \end{aligned}$$

[2]For further details I refer to Lehmann and Romano's book "Testing Statistical Hypothesis". See Theorem 3.4.1 and, more specifically, Corollary 3.4.1. A rigorous reader will note that the test of Corollary 3.4.1 is actually a randomized test that also involves rejecting the null hypothesis with some probability γ when $T(x) = k$. Note, however, that for the particular case of this question, $T(X)$ is continuous when viewed as a random variable. Therefore, $T(X) = k$ implies a set of measure zero.

Under the null, $\bar{X} \equiv \frac{1}{n}\sum_{i=1}^{n} X_i \sim N(\theta_0, \frac{1}{n})$ and hence $\sqrt{n}(\bar{X}-\theta_0) \sim N(0,1)$. It follows that

$$\int_{C_X} f(X;\theta_0)\, dx = 1 - \Phi(k') = \alpha \iff k' = \Phi^{-1}(1-\alpha),$$

which is the $1-\alpha$ quantile of a standard normal distribution. □

(c) *Suppose someone discards the even observations and constructs a one-sided test using averages of the odd observations. Compare the asymptotic power (using Pitman's drift) of this test with the test using averages of all observations found in part **(b)**.*

Solution. Define the sequence $\theta_n = \theta_0 + \frac{h}{\sqrt{n}}$ of local alternatives. For the test using averages of all observations we have

$$\begin{aligned}\sqrt{n}(\bar{X}-\theta_0) &= \sqrt{n}(\bar{X}-\theta_n+\theta_n-\theta_0)\\ &= \sqrt{n}(\bar{X}-\theta_n)+\sqrt{n}(\theta_n-\theta_0)\\ &= \sqrt{n}(\bar{X}-\theta_n)+h.\end{aligned}$$

Notice that as $n \to \infty$, $\theta_n \to \theta_0$ and hence $\sqrt{n}(\bar{X}-\theta_n) \xrightarrow{d} N(0,1)$ by the central limit theorem. Therefore $\sqrt{n}(\bar{X}-\theta_0) \xrightarrow{d} Z+h \sim N(h,1)$, where $Z \sim N(0,1)$. The asymptotic local power is given by

$$\begin{aligned}\lim_{n\to\infty} P(\sqrt{n}(\bar{X}-\theta_0) \geq c) &= P(Z+h \geq c) = P(Z \geq c-h)\\ &= 1 - P(Z < c-h) = 1-\Phi(c-h)\\ &= \Phi(h-c).\end{aligned}$$

For the one-sided test using averages of the "odd" observations only, suppose, without loss of generality, that n is even. By discarding the "even" observations we obtain the new statistic

$$\bar{X}' = \frac{1}{(n/2)}\sum_{i=1}^{n/2} X_i.$$

By proceeding in the same way as before, we obtain

$$\sqrt{n}(\bar{X}'-\theta_0) = \sqrt{n}(\bar{X}'-\theta_n)+h.$$

Now $\mathbb{E}[\bar{X}'] = \theta$ and $V[\bar{X}'] = 2/n$. Thus, as $n \to \infty$, $\sqrt{n}(\bar{X}-\theta_n) \xrightarrow{d} N(0,2)$. Therefore $\sqrt{n}(\bar{X}'-\theta_0) \xrightarrow{d} \sqrt{2}Z+h \sim N(h,2)$. The asymptotic local power then becomes

$$\begin{aligned}\lim_{n\to\infty} P(\sqrt{n}(\bar{X}'-\theta_0) \geq c) &= P(\sqrt{2}Z+h \geq c) = P\left(Z \geq \frac{c-h}{\sqrt{2}}\right)\\ &= 1 - P\left(Z < \frac{c-h}{\sqrt{2}}\right) = 1-\Phi\left(\frac{c-h}{\sqrt{2}}\right)\\ &= \Phi\left(\frac{h-c}{\sqrt{2}}\right).\end{aligned}$$

Denote $\delta \equiv h - c$. Since $\sqrt{2} > 1$ and Φ is strictly increasing, it follows that

$$\Phi(\delta) > \Phi\left(\frac{\delta}{\sqrt{2}}\right) \quad \forall \delta \in \mathbb{R}.$$

That is, the test based on the full sample statistic $\bar{X}$ is (asymptotically) uniformly more powerful than the test based on the half-sample statistic $\bar{X}'$. □

(d) *Show that the UMP test for $H_0 : \theta \geq \theta_0$ against $H_1 : \theta < \theta_0$ reject the null when $\sqrt{n}(\bar{X}_n - \theta_0) < c_\alpha$. How is c_α related to $c_{1-\alpha}$.*

Solution. In this case, the UMP is characterized by the critical region

$$\begin{aligned} C_X &= \left\{ X \;\middle|\; \sum_{i=1}^n X_i < k \right\} \\ &= \left\{ X \;\middle|\; \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n X_i - \theta_0\right) < k' \right\}. \end{aligned}$$

Therefore

$$\int_{C_X} f(X;\theta_0)\, dx = \Phi(k') = \alpha \iff k' = \Phi^{-1}(\alpha),$$

which is the α quantile of a standard normal distribution.

Observe that if c_α is the α quantile and $c_{1-\alpha}$ the $1-\alpha$ quantile of a standard normal distribution, then $\Phi(c_\alpha) = \alpha$, $\Phi(c_{1-\alpha}) = 1 - \alpha$ and hence $\Phi(c_{1-\alpha}) = 1 - \Phi(c_\alpha) = \Phi(-c_\alpha)$. The last equality follows from the symmetry of the standard normal distribution. Therefore $c_{1-\alpha} = -c_\alpha$. □

---

5. [16.22, LNs] *Consider the population model $y = X\beta + u$, where u has iid components with $E(u_i) = 0$ and $E(u_i^2) = \sigma^2$. Note that $(X'X)^{1/2}(\hat{\beta} - \beta) = \sum_{i=1}^n a_{ni}u_i$ for $\hat{\beta} = (X'X)^{-1}X'y$. If the Lindeberg-Feller condition holds, then $(X'X)^{1/2}(\hat{\beta} - \beta) \xrightarrow{d} N(0, \sigma^2 I_k)$.*

(a) *Show that the Lindeberg-Feller condition holds if* $\max_{1\leq i\leq n}\|a_{ni}\| \to 0$.

Solution. Observe that

$$\begin{aligned} (X'X)^{1/2}(\hat{\beta} - \beta) &= (X'X)^{1/2}(X'X)^{-1}X'u \\ &= (X'X)^{-1/2}X'u \\ &= \left(\sum_{i=1}^n x_i x_i'\right)^{-1/2} \sum_{i=1}^n x_i u_i = \sum_{i=1}^n \left(\sum_{j=1}^n x_j x_j'\right)^{-1/2} x_i u_i. \end{aligned}$$

Thus $a_{ni} \equiv \left(\sum_{j=1}^n x_j x_j'\right)^{-1/2} x_i$. Let $z_{ni} \equiv a_{ni} u_i$. For any ε,

$$\begin{aligned}
\left| \sum_{i=1}^n \mathbb{E}[\|z_{ni}\|^2 \mathbf{1}(\|z_{ni}\| > \varepsilon)] \right| &= \sum_{i=1}^n \mathbb{E}[\|a_{ni}u_i\|^2 \mathbf{1}(\|a_{ni}u_i\| > \varepsilon)] \\
&= \sum_{i=1}^n \mathbb{E}[\|a_{ni}\|^2 u_i^2 \mathbf{1}(\|a_{ni}\| \cdot |u_i| > \varepsilon)] \\
&\leq \sum_{i=1}^n \mathbb{E}\left[\|a_{ni}\|^2 u_i^2 \mathbf{1}\left(|u_i| \cdot \max_{1\leq i\leq n} \|a_{ni}\| > \varepsilon\right)\right] < \infty.
\end{aligned}$$

Therefore the dominated convergence theorem holds and

$$\begin{aligned}
0 \leq \lim_{n\to\infty} \sum_{i=1}^n \mathbb{E}[\|z_{ni}\|^2 \mathbf{1}(\|z_{ni}\| > \varepsilon)] &\leq \lim_{n\to\infty} \sum_{i=1}^n \mathbb{E}\left[\|a_{ni}\|^2 u_i^2 \mathbf{1}\left(|u_i| \cdot \max_{1\leq i\leq n} \|a_{ni}\| > \varepsilon\right)\right] \\
&= \sum_{i=1}^n \mathbb{E}\left[\lim_{n\to\infty} \|a_{ni}\|^2 u_i^2 \mathbf{1}\left(|u_i| \cdot \max_{1\leq i\leq n} \|a_{ni}\| > \varepsilon\right)\right] \to 0,
\end{aligned}$$

provided $\max_{1\leq i\leq n} \|a_{ni}\| \to 0$, whence it follows that

$$\lim_{n\to\infty} \sum_{i=1}^n \mathbb{E}[\|z_{ni}\|^2 \mathbf{1}(\|z_{ni}\| > \varepsilon)] = 0,$$

as desired. □

(b) *For the special case $y_i = \beta_0 + x_{1i}\beta_1 + u_i$, assume that $\overline{x_{1n}}$ and $\overline{x_{1n}^2}$ are bounded. Show that the Lindeberg-Feller condition holds if $\max_{1\leq i\leq n} \|x_{1i}\| = o(n^{1/2})$.*

Solution. In this case we have $x_i = (1, x_{1i})$ and

$$\sum_{j=1}^n x_j x_j' = \begin{bmatrix} \boldsymbol{\iota}' \\ X_1' \end{bmatrix} \begin{bmatrix} \boldsymbol{\iota} & X_1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{\iota}'\boldsymbol{\iota} & \boldsymbol{\iota}' X_1 \\ X_1'\boldsymbol{\iota} & X_1' X_1 \end{bmatrix} = \begin{bmatrix} n & \sum_{i=1}^n x_{1i} \\ \sum_{i=1}^n x_{1i} & \sum_{i=1}^n x_{1i}^2 \end{bmatrix},$$

where ι is a n-dimensional vector of ones. Thus

$$\begin{aligned}
\left(\sum_{j=1}^n x_j x_j'\right)^{-1} &= \frac{1}{n\sum_{i=1}^n x_{1i}^2 - \left(\sum_{i=1}^n x_{1i}\right)^2} \begin{bmatrix} \sum_{i=1}^n x_{1i}^2 & -\sum_{i=1}^n x_{1i} \\ -\sum_{i=1}^n x_{1i} & n \end{bmatrix} \\
&= \frac{n}{n\sum_{i=1}^n x_{1i}^2 - \left(\sum_{i=1}^n x_{1i}\right)^2} \begin{bmatrix} n^{-1}\sum_{i=1}^n x_{1i}^2 & -n^{-1}\sum_{i=1}^n x_{1i} \\ -n^{-1}\sum_{i=1}^n x_{1i} & 1 \end{bmatrix} \\
&= \frac{1}{\sum_{i=1}^n x_{1i}^2 - n^{-1}\left(\sum_{i=1}^n x_{1i}\right)^2} \begin{bmatrix} n^{-1}\sum_{i=1}^n x_{1i}^2 & -n^{-1}\sum_{i=1}^n x_{1i} \\ -n^{-1}\sum_{i=1}^n x_{1i} & 1 \end{bmatrix} \\
&= \frac{1}{n\overline{x_{1n}^2} - n\overline{x}_{1n}^2} \begin{bmatrix} \overline{x_{1n}^2} & -\overline{x}_{1n} \\ -\overline{x}_{1n} & 1 \end{bmatrix}.
\end{aligned}$$

Therefore

$$\begin{aligned}
\|a_{ni}\| &= \left[x_i'\left(\sum_{j=1}^n x_jx_j'\right)^{-1/2}\left(\sum_{j=1}^n x_jx_j'\right)^{-1/2} x_i\right]^{1/2} = \left[x_i'\left(\sum_{j=1}^n x_jx_j'\right)^{-1} x_i\right]^{1/2} \\
&= \left[\frac{1}{n\overline{x_{1n}^2} - n\overline{x}_{1n}^2}\cdot\begin{bmatrix}1 & x_{1i}\end{bmatrix}\begin{bmatrix}\overline{x_{1n}^2} & -\overline{x}_{1n} \\ -\overline{x}_{1n} & 1\end{bmatrix}\begin{bmatrix}1 \\ x_{1i}\end{bmatrix}\right]^{1/2} \\
&= \left[\frac{1}{n\overline{x_{1n}^2} - n\overline{x}_{1n}^2}\cdot\left(\overline{x_{1n}^2} - 2\overline{x}_{1n}x_{1i} + x_{1i}^2\right)\right]^{1/2} \\
&= \left[\frac{1}{\overline{x_{1n}^2} - \overline{x}_{1n}^2}\cdot\left(\frac{\overline{x_{1n}^2}}{n} - 2\overline{x}_{1n}\frac{x_{1i}}{n} + \frac{x_{1i}^2}{n}\right)\right]^{1/2} \\
&= \left\{\frac{1}{\overline{x_{1n}^2} - \overline{x}_{1n}^2}\cdot\left[\frac{\overline{x_{1n}^2}}{n} - 2\overline{x}_{1n}\frac{x_{1i}}{n} + \left(\frac{x_{1i}}{\sqrt{n}}\right)^2\right]\right\}^{1/2} \\
&\leq \left\{\frac{1}{\overline{x_{1n}^2} - \overline{x}_{1n}^2}\cdot\left[\frac{\overline{x_{1n}^2}}{n} + 2\overline{x}_{1n}\frac{\max_{1\leq i\leq n}\|x_{1i}\|}{\sqrt{n}} + \left(\frac{\max_{1\leq i\leq n}\|x_{1i}\|}{\sqrt{n}}\right)^2\right]\right\}^{1/2}.
\end{aligned}$$

But observe that since $\overline{x_{1n}^2}$ is bounded, $\overline{x_{1n}^2}/n \to 0$; and, since $\max_{1\leq i\leq n}\|x_{1i}\| = o(n^{1/2})$, $\max_{1\leq i\leq n}\|x_{1i}\|/\sqrt{n} \to 0$. Then

$$\begin{aligned}
0 &\leq \lim_{n\to\infty}\|a_{ni}\| \\
&\leq \lim_{n\to\infty}\left\{\frac{1}{\overline{x_{1n}^2} - \overline{x}_{1n}^2}\cdot\left[\frac{\overline{x_{1n}^2}}{n} + 2\overline{x}_{1n}\frac{\max_{1\leq i\leq n}\|x_{1i}\|}{\sqrt{n}} + \left(\frac{\max_{1\leq i\leq n}\|x_{1i}\|}{\sqrt{n}}\right)^2\right]\right\}^{1/2} \to 0,
\end{aligned}$$

whence it follows that $\|a_{ni}\| \to 0$ and, hence, $\max_{1\leq i\leq n}\|a_{ni}\| \to 0$. Thus, from the result proved in **(a)**, the Lindeberg-Feller condition holds. □

6. [16.33, LNs] *Consider the model:*

$$y_t = x_t'\beta + u_t,$$

where $x_t = (1, t)$, $\beta = (\beta_1, \beta_2)$, u_i *are i.i.d. with* $E(u_t) = 0$, $V(u_t) = \sigma^2$, *and* $E|u_t|^3 = C$. *Hint: It may be helpful to know* $\sum_{t=1}^{N} t^2 = \frac{1}{6}N(N+1)(2N+1)$.

(a) *Show that the OLS estimator* $\hat{\beta}_N$ *is unbiased for* $\beta = (\beta_1, \beta_2)'$.

Solution. The OLS estimator can be written as

$$\begin{aligned}
\hat{\beta}_N &= \left(\sum_{t=1}^{N} x_t x_t'\right)^{-1} \left(\sum_{t=1}^{N} x_t y_t\right) \\
&= \left(\sum_{t=1}^{N} x_t x_t'\right)^{-1} \left(\sum_{t=1}^{N} x_t (x_t'\beta + u_t)\right) \\
&= \left(\sum_{t=1}^{N} x_t x_t'\right)^{-1} \left(\sum_{t=1}^{N} x_t x_t'\right)\beta + \left(\sum_{t=1}^{N} x_t x_t'\right)^{-1} \left(\sum_{t=1}^{N} x_t u_t\right) \\
&= \beta + \left(\sum_{t=1}^{N} x_t x_t'\right)^{-1} \left(\sum_{t=1}^{N} x_t u_t\right).
\end{aligned}$$

Thus, taking the unconditional expectation of $\hat{\beta}_N$ we obtain

$$\begin{aligned}
\hat{\beta}_N &= \mathbb{E}\left[\beta + \left(\sum_{t=1}^{N} x_t x_t'\right)^{-1} \left(\sum_{t=1}^{N} x_t u_t\right)\right] \\
&= \beta + \left(\sum_{t=1}^{N} x_t x_t'\right)^{-1} \left(\sum_{t=1}^{N} x_t \mathbb{E}[u_t]\right) \\
&= \beta,
\end{aligned}$$

where the second equality follows from the linearity of the expectation operator and the fact that $x_t = (1, t)$ is nonstochastic. □

(b) *Find the limiting distribution of* $N^{1/2}(\hat{\beta}_{2,N} - \beta_2)$. *Explain your answer.*

Solution. The limiting distribution will be a constant, degenerated at zero, distribution. To see why, recall that usual asymptotic arguments rely on the convergence in probability of $N^{-1}\sum_{t=1}^{N} x_t x_t'$ to a nonsingular matrix Q and in distribution of $(1/\sqrt{N})\sum_{t=1}^{N} x_t u_t$ to a $N(0, \sigma^2 Q)$ random variable, implying that $\sqrt{N}(\hat{\beta} - \beta) \xrightarrow{d} N(0, \sigma^2 Q^{-1})$. To see why this same argument cannot be used for a deterministic time trend, note that

$$\hat{\beta} - \beta = \begin{bmatrix} \sum_{t=1}^{N} 1 & \sum_{t=1}^{N} t \\ \sum_{t=1} t & \sum_{t=1}^{N} t^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=1} u_t \\ \sum_{t=1} t u_t \end{bmatrix}.$$

It is straightforward to show by induction that

$$\sum_{t=1}^{N} t = N(N+1)/2 \quad \text{and} \quad \sum_{t=1}^{N} t^2 = T(T+1)(2T+1)/6.$$

Thus

$$\begin{aligned}
N^{1/2}(\hat{\beta}-\beta) &= \sqrt{N}\begin{bmatrix} N & N(N+1)/2 \\ N(N+1)/2 & N(N+1)(2N+1)/6 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=1}^{N} u_t \\ \sum_{t=1}^{N} t u_t \end{bmatrix} \\
&= \sqrt{N}\begin{bmatrix} N & N(N+1)/2 \\ N(N+1)/2 & N(N+1)(2N+1)/6 \end{bmatrix}^{-1} \sqrt{N}\frac{1}{\sqrt{N}} \begin{bmatrix} \sum_{t=1}^{N} u_t \\ \sum_{t=1}^{N} t u_t \end{bmatrix} \\
&= N\begin{bmatrix} N & N(N+1)/2 \\ N(N+1)/2 & N(N+1)(2N+1)/6 \end{bmatrix}^{-1} \begin{bmatrix} \frac{1}{\sqrt{N}}\sum_{t=1}^{N} u_t \\ \frac{1}{\sqrt{N}}\sum_{t=1}^{N} t u_t \end{bmatrix} \\
&= \begin{bmatrix} 1 & (N+1)/2 \\ (N+1)/2 & (N+1)(2N+1)/6 \end{bmatrix}^{-1} \begin{bmatrix} \frac{1}{\sqrt{N}}\sum_{t=1}^{N} u_t \\ \frac{1}{\sqrt{N}}\sum_{t=1}^{N} t u_t \end{bmatrix}.
\end{aligned}$$

Taking the inverse we obtain

$$\left(N^{-1}\sum_{t=1}^{N} x_t x_t'\right)^{-1} = \begin{bmatrix} 1 & (N+1)/2 \\ (N+1)/2 & (N+1)(2N+1)/6 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{4N+2}{N-1} & \frac{6}{1-N} \\ \frac{6}{1-N} & \frac{12}{N^2-1} \end{bmatrix}.$$

Thus

$$\begin{aligned}
N^{1/2}(\hat{\beta}-\beta) &= \left(N^{-1}\sum_{t=1}^{N} x_t x_t'\right)^{-1} \frac{1}{\sqrt{N}}\sum_{t=1}^{N} x_t u_t \\
&= \begin{bmatrix} \frac{4N+2}{N-1} & \frac{6}{1-N} \\ \frac{6}{1-N} & \frac{12}{N^2-1} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{N}}\sum_{t=1}^{N} u_t \\ \frac{1}{\sqrt{N}}\sum_{t=1}^{N} t u_t \end{bmatrix} \\
&= \begin{bmatrix} \frac{4N+2}{N-1}\frac{1}{\sqrt{N}}\sum_{t=1}^{N} u_t + \frac{6}{1-N}\frac{1}{\sqrt{N}}\sum_{t=1}^{N} t u_t \\ \frac{6}{1-N}\frac{1}{\sqrt{N}}\sum_{t=1}^{N} u_t + \frac{12}{N^2-1}\frac{1}{\sqrt{N}}\sum_{t=1}^{N} t u_t \end{bmatrix}.
\end{aligned}$$

Therefore,

$$N^{1/2}(\hat{\beta}_{2,N}-\beta_2) = \frac{6}{1-N}\frac{1}{\sqrt{N}}\sum_{t=1}^{N} u_t + \frac{12}{N^2-1}\frac{1}{\sqrt{N}}\sum_{t=1}^{N} t u_t.$$

We will see in the next item that both $\frac{1}{\sqrt{N}}\sum_{t=1}^{N} u_t$ and $\frac{1}{\sqrt{N}}\sum_{t=1}^{N} t u_t$ have well-defined, nondegenerate, limiting distributions. From this it becomes easy to see that whatever the asymptotic distribution of $\frac{1}{\sqrt{N}}\sum_{t=1}^{N} u_t$ and $\frac{1}{\sqrt{N}}\sum_{t=1}^{N} t u_t$ are, as $N \to \infty$ both $6/(1-N) \to 0$ and $12/(N^2-1) \to 0$; therefore, so will $N^{1/2}(\hat{\beta}_{2,N}-\beta_2) \to 0$. □

(c) *Find the limiting joint distribution of $A_T(\hat{\beta}-\beta)$, where*

$$A_N = \begin{bmatrix} N^{1/2} & 0 \\ 0 & N^{3/2} \end{bmatrix},$$

where $\hat{\beta}$ is the OLS estimator of β.

Solution.

$$\begin{aligned} A_N(\hat{\beta}-\beta) &= A_N \begin{bmatrix} N & N(N+1)/2 \\ N(N+1)/2 & N(N+1)(2N+1)/6 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=1}^N u_t \\ \sum_{t=1}^N tu_t \end{bmatrix} \\ &= A_N \begin{bmatrix} N & N(N+1)/2 \\ N(N+1)/2 & N(N+1)(2N+1)/6 \end{bmatrix}^{-1} A_N A_N^{-1} \begin{bmatrix} \sum_{t=1}^N u_t \\ \sum_{t=1}^N tu_t \end{bmatrix} \\ &= \left(A_N^{-1} \begin{bmatrix} N & N(N+1)/2 \\ N(N+1)/2 & N(N+1)(2N+1)/6 \end{bmatrix} A_N^{-1} \right)^{-1} A_N^{-1} \begin{bmatrix} \sum_{t=1}^N u_t \\ \sum_{t=1}^N tu_t \end{bmatrix}. \end{aligned}$$

Notice that

$$A_N^{-1} = \begin{bmatrix} N^{-1/2} & 0 \\ 0 & N^{-3/2} \end{bmatrix}.$$

Thus, for the first term, we have

$$\begin{aligned} A_N^{-1} \begin{bmatrix} N & N(N+1)/2 \\ N(N+1)/2 & N(N+1)(2N+1)/6 \end{bmatrix} A_N^{-1} &= \begin{bmatrix} N^{-1}N & N^{-2}N(N+1)/2 \\ N^{-2}N(N+1)/2 & N^{-3}N(N+1)(2N+1)/6 \end{bmatrix} \\ &= \begin{bmatrix} 1 & (N+1)/2N \\ (N+1)/2N & (N+1)(2N+1)/6N^2 \end{bmatrix} \\ &\xrightarrow{N\to\infty} \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{bmatrix} =: Q. \end{aligned} \tag{5}$$

Turning to the second term, notice that

$$A_N^{-1} \begin{bmatrix} \sum_{t=1}^N u_t \\ \sum_{t=1}^N tu_t \end{bmatrix} = \begin{bmatrix} N^{-1/2} & 0 \\ 0 & N^{-3/2} \end{bmatrix} \begin{bmatrix} \sum_{t=1}^N u_t \\ \sum_{t=1}^N tu_t \end{bmatrix} = \begin{bmatrix} N^{-1/2}\sum_{t=1}^N u_t \\ N^{-3/2}\sum_{t=1}^N tu_t \end{bmatrix} = \begin{bmatrix} N^{-1/2}\sum_{t=1}^N u_t \\ N^{-1/2}\sum_{t=1}^N (t/N)u_t \end{bmatrix}.$$

For the first element, notice that since u_t are i.i.d. with mean zero and variance σ^2, the central limit theorem ensures that

$$\left(1/\sqrt{N}\right)\sum_{t=1}^N u_t \xrightarrow{d} N(0,\sigma^2).$$

For the second element, observe that $\{(t/N)u_t\}$ is a martingale difference sequence that satisfies the conditions for the central limit theorem for matingale difference sequences. Specifically, its variance is

$$\sigma_t^2 = \mathbb{E}[(t/N)^2 u_t^2] = \sigma^2 \cdot (t^2/N^2),$$

where

$$(1/N)\sum_{t=1}^N \sigma_t^2 = \sigma^2(1/N^3)\sum_{t=1}^N t^2 \xrightarrow{N\to\infty} \sigma^2/3.$$

Furthermore, $(1/T)\sum_{t=1}^{T}[(t/N)u_t]^2 \xrightarrow{p} \sigma^2/3$. To verify this claim, notice that

$$\begin{aligned}
\mathbb{E}\left[\left((1/N)\sum_{t=1}^{N}[(t/N)u_t]^2 - (1/N)\sum_{t=1}^{N}\sigma_t^2\right)^2\right] &= \mathbb{E}\left[\left((1/N)\sum_{t=1}^{N}[(t/N)u_t]^2 - (1/N)\sum_{t=1}^{N}(t^2/N^2)\sigma^2\right)^2\right] \\
&= \mathbb{E}\left[\left((1/N)\sum_{t=1}^{N}(t/N)^2(u_t^2-\sigma^2)\right)^2\right] \\
&= \mathbb{E}\left[(1/N)^6\sum_{t=1}^{N}t^4(u_t^2-\sigma^2)^2\right] \\
&= (1/N)^6\sum_{t=1}^{N}t^4\mathbb{E}\left[(u_t^2-\sigma^2)^2\right] \\
&= \left[(1/N)^6\sum_{t=1}^{N}t^4\right]\mathbb{E}\left[(u_t^2-\sigma^2)^2\right].
\end{aligned}$$

One can show by induction that

$$(1/N^{\nu+1})\sum_{t=1}^{N}t^{\nu} \xrightarrow{N\to\infty} 1/(\nu+1).$$

Therefore

$$N \times \left[(1/N)^6\sum_{t=1}^{N}t^4\right]\mathbb{E}\left[(u_t^2-\sigma^2)^2\right] \xrightarrow{N\to\infty} \frac{1}{5}\mathbb{E}\left[(u_t^2-\sigma^2)^2\right],$$

which implies

$$\left[(1/N)^6\sum_{t=1}^{N}t^4\right]\mathbb{E}\left[(u_t^2-\sigma^2)^2\right] \xrightarrow{N\to\infty} 0.$$

Therefore

$$(1/N)\sum_{t=1}^{N}[(t/N)u_t]^2 - (1/N)\sum_{t=1}^{N}\sigma_t^2 \xrightarrow{\text{m.s.}} 0,$$

whence it follows that

$$(1/N)\sum_{t=1}^{N}[(t/N)u_t]^2 \xrightarrow{p} \sigma^2/3.$$

Hence, $(1/\sqrt{N})\sum_{t=1}^{N}(t/N)u_t$ satisfies the central limit theorem for martingale difference sequences and thus

$$(1/\sqrt{N})\sum_{t=1}^{N}(t/N)u_t \xrightarrow{d} N(0,\sigma^2/3).$$

Finally, consider the joint distribution of $\left(N^{-1/2}\sum_{t=1}^N u_t, N^{-1/2}\sum_{t=1}^N (t/N)u_t\right)'$. Any linear combination of these elements takes the form

$$\alpha_1\left(N^{-1/2}\sum_{t=1}^N u_t\right) + \alpha_2\left(N^{-1/2}\sum_{t=1}^N (t/N)u_t\right) = N^{-1/2}\sum_{t=1}^N [\alpha_1 + \alpha_2(t/N)]u_t.$$

Observe that $[\alpha_1 + \alpha_2(t/N)]u_t$ is also a martingale difference sequence with positive variance given by $\sigma^2[\alpha_1^2 + 2\alpha_1\alpha_2(t/N) + \alpha_2^2(t/T)^2]$ satisfying

$$N^{-1}\sum_{t=1}^N \sigma^2[\alpha_1^2 + 2\alpha_1\alpha_2(t/N) + \alpha_2^2(t/N)^2] \to \sigma^2[\alpha_1 + 2\alpha_1\alpha_2(1/2) + \alpha_2^2(1/3)] = \alpha'(\sigma^2 Q)\alpha$$

for $\alpha = (\alpha_1, \alpha_2)'$. Furthermore, it is easy to show that

$$N^{-1}\sum_{t=1}^N [\alpha_1 + \alpha_2(t/N)]^2 u_t^2 \xrightarrow{p} \alpha'(\sigma^2 Q)\alpha.$$

Therefore, the central limit theorem for martingale difference sequences implies that any linear combination of two elements in the vector $\left(N^{-1/2}\sum_{t=1}^N u_t, N^{-1/2}\sum_{t=1}^N (t/N)u_t\right)'$ converges in distribution to $N(0, \alpha'(\sigma^2 Q)\alpha) = \alpha' N(0, \sigma^2 Q)$. Thus, by the Cramér-Wold device, we must have

$$\left(N^{-1/2}\sum_{t=1}^N u_t, N^{-1/2}\sum_{t=1}^N (t/N)u_t\right)' \xrightarrow{d} N(0, \sigma^2 Q).$$

From this, by employing the continuous mapping and Slutsky's theorem and recalling the limit obtained in (5), we can finally conclude that

$$\begin{aligned} A_N(\hat\beta - \beta) = \begin{bmatrix} N^{1/2}(\hat\beta_1 - \beta) \\ N^{3/2}(\hat\beta_2 - \beta) \end{bmatrix} &= \begin{bmatrix} 1 & (N+1)/2N \\ (N+1)/2N & (N+1)(2N+1)/6N^2 \end{bmatrix}^{-1} \begin{bmatrix} N^{-1/2}\sum_{t=1}^N u_t \\ N^{-1/2}\sum_{t=1}^N (t/N)u_t \end{bmatrix} \\ &\xrightarrow{d} Q^{-1}Z \sim N(0, Q^{-1}\sigma^2 Q Q^{-1}) = N(0, Q^{-1}\sigma^2), \end{aligned}$$

where $Z \sim N(0, \sigma^2 Q)$ and Q is as defined in (5). □