Stat 513

Homework key 1

January 13, 2015

REGULAR PROBLEMS

1. Let $X_1, \ldots, X_n$ be a sequence of i.i.d. random variables having mean μ, variance $\sigma^2 > 0$, and finite third and fourth central moments and γ_4, respectively. Find an asymptotic joint distribution for the sample mean $\bar{X}$ and sample variance S_n^2.

Ans: We will show that

$$\sqrt{n}\begin{pmatrix}\bar{X}-\mu \\ S_n^2-\sigma^2\end{pmatrix} \xrightarrow{\mathcal{D}} N_2\left[0, \begin{pmatrix}\sigma^2 & \gamma_3 \\ \gamma_3 & \gamma_4-\sigma^4\end{pmatrix}\right]$$

We will show that using Cramer-Wald theorem by proving the following: For any $l_1, l_2 \in \mathbb{R}$,

$$\sqrt{n}(l_1(\bar{X}-\mu)+l_2(S_n^2-\sigma^2)) \xrightarrow{\mathcal{D}} N\left(0, \begin{pmatrix}l_1 & l_2\end{pmatrix}\begin{pmatrix}\sigma^2 & \gamma_3 \\ \gamma_3 & \gamma_4-\sigma^4\end{pmatrix}\begin{pmatrix}l_1 \\ l_2\end{pmatrix}.\right)$$

Now

$$\sqrt{n}(l_1(\bar{X}-\mu)+l_2(S_n^2-\sigma^2)) \quad (1)$$

$$= \sqrt{n}(l_1(\bar{X}-\mu)+l_2(\sum_{i=1}^{n}\frac{X_i^2}{n-1}-\bar{X}^2-\sigma^2)) \quad (2)$$

$$(3)$$

For large the above will be distributed same as

$$\sqrt{n}(l_1(\bar{X}-\mu)+l_2(\sum_{i=1}^{n}\frac{X_i^2}{n}-\bar{X}^2-\sigma^2))$$

$$= \sqrt{n}(l_1(\bar{X}-\mu)+l_2(\sum_{i=1}^{n}\frac{(X_i-\mu)^2}{n}-\sigma^2)) - l_2\sqrt{n}(\mu^2-2\mu\bar{X}+\bar{X}^2)$$

Now

$$\sqrt{n}(\mu^2-2\mu\bar{X}+\bar{X}^2) = [\sqrt{n}(\bar{X}-\mu)].(\bar{X}-\mu)$$

$$\sqrt{n}(\bar{X}-\mu) \xrightarrow{\mathcal{D}} N(0,\sigma^2)$$

$$\bar{X}-\mu \xrightarrow{P} 0$$

by CLT and WLLN respectively. Hence by Slutsky's theorem, $l_2\sqrt{n}(\mu^2-2\mu\bar{X}+\bar{X}^2) \xrightarrow{\mathcal{D}} 0$, or $\xrightarrow{P} 0$. Hence, the expression in (1) becomes

$$\sqrt{n}(l_1(\bar{X}-\mu)+l_2(\sum_{i=1}^{n}\frac{(X_i-\mu)^2}{n}-\sigma^2)) + O_P(1).$$

Hence it is enough to show that

$$\sqrt{n}(l_1(\bar{X}-\mu)+l_2(\sum_{i=1}^{n}\frac{(X_i-\mu)^2}{n}-\sigma^2)) \xrightarrow{\mathcal{D}} N\left(0, \begin{pmatrix}l_1 & l_2\end{pmatrix}\begin{pmatrix}\sigma^2 & \gamma_3 \\ \gamma_3 & \gamma_4-\sigma^4\end{pmatrix}\begin{pmatrix}l_1 \\ l_2\end{pmatrix}\right). \quad (4)$$

since the rest follows by Slutsky's Theorem. The LHS can be re-written as

$$\sqrt{n}\left(\frac{\sum_{i=1}^{n}(l_1X_i-l_1\mu+l_2(X_i-\mu)^2-l_2\sigma^2)}{n}\right) = \sqrt{n}\frac{\sum_{i=1}^{n}T_i}{n}$$

where $T_i = l_1 X_i - l_1\mu + l_2(X_i-\mu)^2 - l_2\sigma^2) = \begin{pmatrix} l_1 & l_2 \end{pmatrix}\begin{pmatrix} X_i-\mu \\ (X_i-\mu)^2-\sigma^2 \end{pmatrix}$. By CLT,

$$\frac{\sum_{i=1}^n T_i}{\sqrt{n}} \xrightarrow{\mathcal{D}} N(0, V(T_1))$$

where

$$\begin{aligned} Var(T_1) =& Var\left(\begin{pmatrix} l_1 & l_2 \end{pmatrix}\begin{pmatrix} X_1-\mu \\ (X_1-\mu)^2-\sigma^2 \end{pmatrix}\right) \\ =& \begin{pmatrix} l_1 & l_2 \end{pmatrix}\begin{pmatrix} \sigma^2 & \gamma_3 \\ \gamma_3 & \gamma_4-\sigma^4 \end{pmatrix}\begin{pmatrix} l_1 \\ l_2 \end{pmatrix} \end{aligned}$$

This settles (4) which completes the proof.
An alternative solution could have first considered the random vector

$$\vec{Z}_i = \begin{pmatrix} X_i \\ (X_i-\mu)^2 \end{pmatrix} \sim \left(\begin{pmatrix} \mu \\ \sigma^2 \end{pmatrix}, \mathbf{V} = \begin{pmatrix} \sigma^2 & \gamma_3 \\ \gamma_3 & \gamma_4-\sigma^4 \end{pmatrix} \right),$$

where the covariance of X_i and $(X_i-\mu)^2$ is easily found by

$$\begin{aligned} Cov(X_i,(X_i-\mu)^2) &= E[X_i(X_i-\mu)^2) - E[X_i]\,E[(X_i-\mu)^2] \\ &= E[(X_i-\mu)^3 + \mu X_i - \mu)^2) - E[X_i]\,E[(X_i-\mu)^2] = \gamma_3 + \mu\sigma^2 - \mu\sigma^2 = \gamma_3 \end{aligned}$$

Then because the $\vec{Z}_i$'s are i.i.d., the multivariate CLT tells us that

$$\sqrt{n}\left(\begin{pmatrix} \overline{X}_n \\ \frac{1}{n}\sum_{i=1}^n (X_i-\mu)^2 \end{pmatrix} - \begin{pmatrix} \mu \\ \sigma^2 \end{pmatrix}\right) \rightarrow_d \mathcal{N}_2\left(\vec{0}, \mathbf{V}\right)$$

Now, because

$$\sum_{i=1}^n (X_i-\mu)^2 = \sum_{i=1}^n (X_i-\overline{X}_n)^2 + n(\overline{X}_n-\mu)^2,$$

we have

$$\sqrt{n}\left(\begin{pmatrix} \overline{X}_n \\ \frac{1}{n}\sum_{i=1}^n (X_i-\mu)^2 \end{pmatrix} - \begin{pmatrix} \mu \\ \sigma^2 \end{pmatrix}\right) = \sqrt{n}\left(\begin{pmatrix} \overline{X}_n \\ s_n^2 \end{pmatrix} - \begin{pmatrix} \mu \\ \sigma^2 \end{pmatrix}\right) + \begin{pmatrix} 0 \\ \frac{1}{\sqrt{n}}s_n^2 + \sqrt{n}(\overline{X}_n-\mu)^2 \end{pmatrix}$$

Now we also know by the CLT that $\sqrt{n}(\overline{X}_n-\mu) \rightarrow_d \mathcal{N}(0,\sigma^2)$, by the WLLN that $(\overline{X}_n-\mu) \rightarrow_p 0$, and by previous homeworks that $s_n^2 \rightarrow_p \sigma^2$, so

$$\frac{1}{\sqrt{n}}s_n^2 + \sqrt{n}(\overline{X}_n-\mu)^2 \rightarrow_p 0$$

and Slutsky's theorem then gives us our result.

2. Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables having mean μ and variance $\sigma^2 > 0$, and let $Y_1, Y_2, \ldots$ be a sequence of i.i.d. random variables having mean ν and variance $\tau^2 > 0$. Further suppose that X_i's and Y_j's are totally independent. For notational convenience, denote $\vec{X}_n = (X_1, \ldots, X_n)$ and $\vec{Y}_n = (Y_1, \ldots, Y_n)$.

   (a) Find a method of moments estimator $\hat{\theta}_{mn} = \hat{\theta}_{mn}(\vec{X}_m, \vec{Y}_n)$ for $\theta = \mu - \nu$. Show that your estimator is unbiased.

   (b) Let $N = m+n$. Suppose further that $m/N \to \lambda \in (0,1)$ as $N \to \infty$. Find an asymptotic distribution for $\hat{\theta}_{mn}$. That is, show that there is some $a_N \to \infty$ such that $a_N(\hat{\theta}_{mn} - \theta)$ converges in distribution, and identify the distribution.

Ans:

(a) It is useful to redefine the sample in terms of a new random variable T_i so that $(T_1, \ldots, T_N) = (X_1, \ldots, X_m, Y_1, \ldots Y_n)$ as follows:

$$T_i = \delta_i X_i + (1-\delta_i)Y_{i-m},$$

$$\delta_i = \mathbf{1}[i \leq m].$$

Clearly $E[T_i|\delta_i = d_i] = d_i\mu + (1-d_i)\nu = \nu + d_i\theta$. It follows that a method of moments estimator for theta is

$$\hat{\theta}_{mn} = \frac{1}{m}\sum_{i=1}^{m} T_i + \frac{1}{n}\sum_{j=m+1}^{m+n} T_j = \bar{X}_m - \bar{Y}_n.$$

As the difference in empirical means, the estimator is seen to be unbiased:

$$E[\hat{\theta}_{mn}] = E[\bar{X}_m - \bar{Y}_n] = \mu - \nu = \theta.$$

(b) *Note: You were asked to solve this problem by considering the Lindeberg-Feller CLT. The problem would have been quite easy to solve using the following theorem:*
Theorem Suppose the sequence of random variables $X_1, X_2, \ldots$ has asymptotic distribution $X_n \to_d X$, and the sequence of random variables $Y_1, Y_2, \ldots$ has asymptotic distribution $Y_n \to_d Y$. Further suppose that each X_i is totally independent of all the Y_j's, and each Y_j is totally independent of all the X_i's. Let X^* and Y^* be independent random variables such that X^* and X are identically distributed and Y^* and Y are identically distributed. Then as $min(m,n) \to \infty$,

$$X_n + Y_m \to_d X^* + Y^*.$$

As should be obvious with a little reflection, if there is dependence between the X_i's and Y_j's, we would need to consider the limiting distribution of the sums in a different way.
In problem 6, we find an asymptotic distribution under the weaker condition that $\min(m,n) \to \infty$ as $N \to \infty$. Since $m/N \to \lambda$, we have $m \to \infty$, so that the result from problem 6 applies:

$$\frac{\hat{\theta}_{mn} - \theta}{\sqrt{\frac{\sigma^2}{m} + \frac{\tau^2}{n}}} \to_d \mathcal{N}(0,1).$$

However, we desire an asymptotic distribution for a quantity of the form

$$a_N(\hat{\theta}_{mn} - \theta),$$

which requires that we re-express the sequence as

$$\begin{aligned}
\frac{\hat{\theta}_{mn} - \theta}{\sqrt{\frac{\sigma^2}{m} + \frac{\tau^2}{n}}} &= \frac{\sqrt{mn/N}}{\sqrt{mn/N}} \frac{\hat{\theta}_{mn} - \theta}{\sqrt{\frac{\sigma^2}{m} + \frac{\tau^2}{n}}} \\
&= \frac{\sqrt{\frac{mn}{N}}(\hat{\theta}_{mn} - \theta)}{\sqrt{\frac{n}{N}\sigma^2 + \frac{m}{N}\tau^2}} \\
&= \frac{\sqrt{\frac{mn}{N}}(\hat{\theta}_{mn} - \theta)}{\sqrt{(1-\lambda)\sigma^2 + \lambda\tau^2}} \frac{\sqrt{(1-\lambda)\sigma^2 + \lambda\tau^2}}{\sqrt{\frac{n}{N}\sigma^2 + \frac{m}{N}\tau^2}}.
\end{aligned}$$

We recognize that as $N \to \infty$, the quantity

$$b_n^{-1} \equiv \sqrt{(1-\lambda)\sigma^2 + \lambda\tau^2} / \sqrt{\frac{n}{N}\sigma^2 + \frac{m}{N}\tau^2} \to_p \sqrt{(1-\lambda)\sigma^2 + \lambda\tau^2} / \sqrt{(1-\lambda)\sigma^2 + \lambda\tau^2} = 1.$$

By Slutsky's theorem, we obtain that

$$\frac{\hat{\theta}_{mn} - \theta}{\sqrt{\frac{\sigma^2}{m} + \frac{\tau^2}{n}}} \cdot b_n = \frac{\sqrt{\frac{mn}{N}}(\hat{\theta}_{mn} - \theta)}{\sqrt{(1-\lambda)\sigma^2 + \lambda\tau^2}} \to_d \mathcal{N}(0,1).$$

It is now clear that

$$a_N(\hat{\theta}_{mn} - \theta) = \sqrt{\frac{mn}{N}}(\hat{\theta}_{mn} - \theta) \to_d \mathcal{N}(0, (1-\lambda)\sigma^2 + \lambda\tau^2) \equiv \mathcal{N}(0, V(\sigma^2, \tau^2, \lambda)).$$

(Note: To determine the correct form of a_N, we recognize that the sequence involved a function of m^{-1} and n^{-1}. Since we assume here that m/N and n/N converge to constants as N increases, it is natural to choose $a_N = mn/N$ to make m/N and n/N appear explicitly in our sequence.)

3. In the setting of problem 2, we are often faced with the problem that σ^2 and τ^2 are "nuisance" parameters that are unknown, but are not central to our primary statistical question. They are necessary, however, when trying to compute confidence intervals or perform hypothesis tests. In settings where

$$a_N(\hat{\theta}_N - \theta) \rightarrow_d \mathcal{N}(0, V(\sigma^2, \tau^2, \lambda))$$

we most often proceed by finding a consistent estimator $\hat{V} = V(\hat{\sigma}_m^2, \hat{\tau}_n^2, \hat{\lambda})$ such that

$$a_N \left(\frac{\hat{\theta}_N - \theta}{\sqrt{\hat{V}}} \right) \rightarrow_d \mathcal{N}(0, 1).$$

Using the notation of problem 1, find the asymptotic consistency properties of $\hat{V}/V$ for each of the following commonly used estimators of V. Explicitly consider the two possibilities that $\sigma^2 = \tau^2$ and $\sigma^2 \neq \tau^2$ and explicitly consider possible values λ.

(a) *Use of a pooled variance estimate*: $\hat{\sigma}_m^2 = \hat{\tau}_n^2 = s_{P,mn}^2$ where

$$s_{P,mn}^2 = \frac{\sum_{i=1}^m (X_i - \bar{X}_m)^2 + \sum_{j=1}^n (Y_j - \bar{Y}_n)^2}{m + n - 2}.$$

(b) *Use of separate sample variance estimates*:

$$\hat{\sigma}_m^2 = \frac{1}{m-1} \sum_{i=1}^m (X_i - \bar{X}_m)^2 \qquad \hat{\tau}_n^2 = \frac{1}{n-1} \sum_{j=1}^n (Y_j - \bar{Y}_n)^2.$$

(c) *Use of an estimate derived from a combined population*: Combine the two samples into a single sample, letting $\bar{W}_{mn}$ be the sample mean

$$\bar{W}_{mn} = \frac{m\bar{X}_m + n\bar{Y}_n}{m + n}$$

$$\hat{\sigma}_m^2 = \hat{\tau}_n^2 = \frac{\sum_{i=1}^m (X_i - \bar{W}_{mn})^2 + \sum_{j=1}^n (Y_j - \bar{W}_{mn})^2}{m + n - 1}.$$

Ans:

(a) We will become more familiar with the inferential concepts of hypothesis testing and confidence intervals later in the course, though this problem highlights important considerations associated with these topics. When the limit of the ratio $\hat{V}/V$ is > 1, confidence intervals based on $\hat{V}$ to be too wide. The associated inference will be conservative, or more likely to *not* reject the null hypothesis. For a limit of < 1, confidence intervals will be too narrow, inference will be anti-conservative and we inflate the probability of a false rejection (the Type I error rate will be greater than the nominal level).

We can express $s_{P,mn}^2$ in terms of the unbiased variance estimates from a single sample,

$$s_X^2 = \frac{1}{m-1} \sum_{i=1}^m (X_i - \bar{X}_m)^2 \qquad s_Y^2 = \frac{1}{n-1} \sum_{j=1}^n (Y_j - \bar{Y}_n)^2.$$

Doing so, we see that

$$s_{P,mn}^2 = \frac{m-1}{m+n-2} s_X^2 + \frac{n-1}{m+n-2} s_Y^2 \rightarrow_p \lambda\sigma^2 + (1-\lambda)\tau^2.$$

Now, we consider our estimate of the asymptotic variance

$$\hat{V} = (1 - \hat{\lambda})\hat{\sigma}_m^2 + \hat{\lambda}\hat{\tau}_n^2 = s_{P,mn}^2.$$

We are now prepared to study the quantity $\hat{V}/V$. We find that

$$\frac{\hat{V}}{V} = \frac{\frac{m-1}{m+n-2} s_X^2 + \frac{n-1}{m+n-2} s_Y^2}{(1-\lambda)\sigma^2 + \lambda\tau^2} \rightarrow_p \frac{\lambda\sigma^2 + (1-\lambda)\tau^2}{(1-\lambda)\sigma^2 + \lambda\tau^2},$$

which is not equal to 1 in general (which means that $\hat{V}$ is not always consistent for V). However, if $\sigma^2 = \tau^2$ or $\lambda = \frac{1}{2}$, then we have the desired consistency. For values of λ close to 1, the limit of $\hat{V}/V$ will behave like σ^2/τ^2. Similarly, for λ near 0, $\hat{V}/V$ will converge towards a value nearly equal to τ^2/σ^2.

(b) In this case, $\hat{\sigma}_m^2 = s_X^2$ and $\hat{\tau}_n^2 = s_Y^2$. We have an estimate of the asymptotic variance that is

$$\hat{V} = (1-\hat{\lambda})s_X^2 + \hat{\lambda}s_Y^2.$$

Invoking the consistency of the sample variances, we have that, for any $\sigma^2, \tau^2, \lambda$,

$$\frac{\hat{V}}{V} \to_p \frac{(1-\lambda)\sigma^2 + \lambda\tau^2}{(1-\lambda)\sigma^2 + \lambda\tau^2} = 1.$$

This means that $\hat{V}$ based on separate variance estimates will always be consistent for the quantity V, as desired.

(c) Since the statistic again assumes equal variances, we see that

$$\begin{aligned}
\hat{V} &= \hat{\sigma}_m^2 = \hat{\tau}_n^2 \\
&= \frac{\sum_{i=1}^m (X_i - \bar{W}_{mn})^2 + \sum_{j=1}^n (Y_j - \bar{W}_{mn})^2}{m+n-1} \\
&= \frac{\sum_{i=1}^m (X_i - \bar{X}_m + \bar{X}_m - \bar{W}_{mn})^2 + \sum_{j=1}^n (Y_j - \bar{Y}_n + \bar{Y}_n - \bar{W}_{mn})^2}{m+n-1} \\
&= \frac{\sum_{i=1}^m (X_i - \bar{X}_m)^2 + \left(\frac{n}{m+n}\right)^2 m(\bar{X}_m - \bar{Y}_n)^2 + \sum_{j=1}^n (Y_j - \bar{Y}_n)^2 + \left(\frac{m}{m+n}\right)^2 n(\bar{Y}_n - \bar{X}_m)^2}{m+n-1} \\
&= \frac{m-1}{m+n-1}s_X^2 + \frac{n-1}{m+n-1}s_Y^2 + \frac{m}{m+n}\frac{n}{m+n-1}(\bar{X}_m - \bar{Y}_n)^2 \\
&\to_p \lambda\sigma^2 + (1-\lambda)\tau^2 + \lambda(1-\lambda)(\mu-\nu)^2.
\end{aligned}$$

Studying the limiting ratio of the variances, we find that

$$\frac{\hat{V}}{V} \to_p \frac{\lambda\sigma^2 + (1-\lambda)\tau^2 + \lambda(1-\lambda)(\mu-\nu)^2}{(1-\lambda)\sigma^2 + \lambda\tau^2}.$$

In order for this limit to equal 1, it is necessary for $\mu = \nu$ in addition to at least one of the conditions in (a). When $\mu \neq \nu$, the limiting ratio can become arbitrarily greater than 1, even if the variances are equal. This means that the larger our effect size is in terms of $|\theta| = |\mu - \nu|$, the more conservative our inference will be!

4. Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables having mean $\mu > 0$ and variance $\sigma^2 > 0$, and $Y_1, \ldots, Y_n$ be a sequence of i.i.d. random variables having mean $v > 0$ and variance $\tau^2 > 0$. Further suppose that X_i and Y_j are independent if $i \neq j$, but that the correlation between X_i and Yj is ρ if $i = j$. For notational convenience, denote $\overrightarrow{X} \sim (X_1, \ldots X_n)$ and $\overrightarrow{Y} \sim (Y_1, \ldots, Y_n)$. Define sample covariance

$$\begin{aligned}
\hat{\rho_n} = \hat{\rho_n}(\overrightarrow{X}, \overrightarrow{Y}) &= \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}) \\
&= \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})Y_i \\
&= \frac{1}{n}\sum_{i=1}^n X_iY_i - \bar{X}\bar{Y}.
\end{aligned}$$

Find an asymptotic distribution for $\hat{\rho}_n$.

Ans: Notice that If we replace X_i by $X_i - \mu$ and Y_i by $Y_i - \mu$ then $\hat{\rho}_n$ remains the same. So from now on by X_i' we will denote $X_i - \mu$ and by Y_i',we will denote $Y_i - v$.

$$\begin{aligned}
&\sqrt{n}\hat{\rho}_n \\
&= \frac{\sum_{i=1}^n X_i'Y_i'}{\sqrt{n}} - \sqrt{n}\bar{X}'\bar{Y}'
\end{aligned}$$

By CLT,

$$\frac{\sum_{i=1}^n X_i'Y_i'}{\sqrt{n}} \xrightarrow{\mathcal{D}} N(0, V(X_i'Y_i'))$$

$$\sqrt{n}\bar{X}' \xrightarrow{\mathcal{D}} N(0, \sigma^2),$$

and by WLLN

$$\bar{Y}' \xrightarrow{P} 0.$$

Hence, by Slutsky's theorem we get that

$$\sqrt{n}\hat{\rho}_n = \frac{\sum_{i=1}^n X_i'Y_i'}{\sqrt{n}} - \sqrt{n}\bar{X}'\bar{Y}' \xrightarrow{\mathcal{D}} N(0, V(X_i'Y_i')).$$

Now in our case, $V(X_i'Y_i') = V((X_i - \mu)(Y_i - v))$.

5. Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables having $X_i \sim N(\mu, \sigma^2)$. Show that sample mean $\bar{X}$ and sample variance S_n^2 are independent.
Ans: Since $\bar{X}$ and S_n^2 are multiples of $\sum_{i=1}^n X_i$ and $\sum_{i=1}^n (X_i - \bar{X})^2$, it is enough to show that $\sum_{i=1}^n X_i$ and $\sum_{i=1}^n (X_i - \bar{X})^2$ are independent. Let $\overrightarrow{X} = (X_1, X_2, \ldots X_n)^T$. Then $\bar{X} = \overrightarrow{1}^T\overrightarrow{X}$ where $\overrightarrow{1} = (1, 1, \ldots)^T$. Now we know from Stat 512 that $\sum_{i=1}^n (X_i - \bar{X})^2 = \overrightarrow{X}^T\left(I - \frac{J}{n}\right)\overrightarrow{X}$ where $J = \overrightarrow{1}\,\overrightarrow{1}^T$. Also we learned that $\left(I - \frac{J}{n}\right)^2 = \left(I - \frac{J}{n}\right)$. Let us define

$$\overrightarrow{v} = \left(I - \frac{J}{n}\right)\overrightarrow{X}.$$

Then

$$\sum_{i=1}^n (X_i - \bar{X})^2 = \overrightarrow{v}^T\overrightarrow{v}.$$

Hence it is enough to show that $\overrightarrow{v}$ and $\bar{X}$ are independent. The rest will follow since S_n^2 is a function solely depending on $\overrightarrow{v}\dot{\text{W}}$e will show $cov(\overrightarrow{v}, \bar{X}) = 0$. This implies that they are independent noting the fact that $\overrightarrow{v}, \bar{X}$ are linear combinations of $\overrightarrow{X}$. So $\overrightarrow{v}$ as well as $\bar{X}$ are normal random variables and recall that uncorrelated normal random variables are independent. Now,

$$\begin{aligned}
Cov(\overrightarrow{v}, \bar{X}) &= cov\left(\left(I - \frac{J}{n}\right)\overrightarrow{X}, \overrightarrow{X}^T\overrightarrow{1}\right) \\
&= \left(I - \frac{J}{n}\right)cov(\overrightarrow{X}, \overrightarrow{X})\overrightarrow{1} \\
&= \sigma^2\left(I - \frac{J}{n}\right)\overrightarrow{1} \\
&= \overrightarrow{1} - \frac{n}{n}\overrightarrow{1} = 0
\end{aligned}$$

which completes the proof.

MORE INVOLVED PROBLEMS

6. Consider again the setting of problem 1, but merely presume that as $N \to \infty$, we only know that $\min(m,n) \to \infty$. Prove an asymptotic distribution for $\hat{\theta}_{mn}$ in this setting.
Ans:
We can not appeal directly to the standard Central Limit Theorem, instead appealing to the Lindeberg-Feller CLT. In order to show the desired convergence, a strategy is to express $\hat{\theta}_{mn} - \theta$ as a sum of mean zero random variables. A similar definition to the T_i we saw in problem 2 will work here:

$$(Z_1, \dots, Z_N) \equiv \Big((X_1 - \mu)/m, \dots, (X_m - \mu)/m, -(Y_1 - \nu)/n, \dots, -(Y_n - \nu)/n\Big),$$

$$Z_i = \frac{X_i - \mu}{m}\delta_i + \frac{Y_{i-m} - \nu}{n}(1 - \delta_i),$$
$$\delta_i = \mathbf{1}[i \le m].$$

Clearly we have $E[Z_i] = 0$ for any $1 \le i \le N$. We also have

$$Var[Z_i] = \delta_i \frac{\sigma^2}{m^2} + (1 - \delta_i)\frac{\tau^2}{n^2}.$$

The Lindeberg-Feller CLT states that if the Lindeberg Condition holds, then

$$\frac{S_N}{\sqrt{V_N}} \to_d \mathcal{N}(0,1),$$

where we define

$$S_N = \sum_{i=1}^{N} Z_i = \hat{\theta}_{mn} - \theta, \qquad V_N = \sum_{i=1}^{N} Var[Z_i] = \frac{\sigma^2}{m} + \frac{\tau^2}{n}.$$

The Lindeberg Condition requires that we show

$$\lim_{N \to \infty} L_N(\epsilon) = 0$$

for any $\epsilon > 0$, where

$$\begin{aligned}
L_N(\epsilon) &= \frac{1}{V_N}\sum_{i=1}^{N} E\Big[Z_i^2 \mathbf{1}[|Z_i| > \epsilon\sqrt{V_N}]\Big] \\
&= \frac{1}{V_N}\sum_{i=1}^{m} E\left[\left(\frac{X_i - \mu}{m}\right)^2 \mathbf{1}[|X_i - \mu| > m\epsilon\sqrt{V_N}]\right] + \frac{1}{V_N}\sum_{i=m+1}^{N} E\left[\left(\frac{Y_{i-m} - \nu}{n}\right)^2 \mathbf{1}[|Y_{i-m} - \nu| > n\epsilon\sqrt{V_N}]\right] \\
&\equiv L_N^X(\epsilon) + L_N^Y(\epsilon).
\end{aligned}$$

We now show that $L_N^X(\epsilon) \to 0$ as $N \to \infty$. Noting that the X_i are i.i.d. random variables, we can simplify the summation as

$$\begin{aligned}
L_N^X &= \frac{1}{V_N}\sum_{i=1}^{m} E\left[\left(\frac{X_i - \mu}{m}\right)^2 \mathbf{1}[|X_i - \mu| > m\epsilon\sqrt{V_N}]\right] \\
&= \frac{1}{V_N} mE\left[\left(\frac{X_1 - \mu}{m}\right)^2 \mathbf{1}[|X_1 - \mu| > m\epsilon\sqrt{V_N}]\right] \\
&= \frac{1}{mV_N} E\left[(X_1 - \mu)^2 \mathbf{1}[|X_1 - \mu| > m\epsilon\sqrt{V_N}]\right] \\
&= \frac{1}{\sigma^2 + (m/n)\tau^2} E\left[(X_1 - \mu)^2 \mathbf{1}[|X_1 - \mu| > m\epsilon\sqrt{V_N}]\right] \\
&< \frac{1}{\sigma^2} E\left[(X_1 - \mu)^2 \mathbf{1}[|X_1 - \mu| > m\epsilon\sqrt{V_N}]\right] \\
&= \frac{1}{\sigma^2} E\left[(X_1 - \mu)^2 \mathbf{1}[|X_1 - \mu| > \epsilon\sqrt{m\sigma^2 + (m^2/n)\tau^2}]\right] \\
&\to \frac{1}{\sigma^2} E\left[(X_1 - \mu)^2 \cdot 0\right] \\
&= 0,
\end{aligned}$$

since $\min(m, n) \to \infty$. By the squeeze theorem for sequences, we have that

$$\lim_{N \to \infty} L_N^X(\epsilon) = 0, \quad \forall \epsilon > 0.$$

An entirely similar argument shows that

$$\lim_{N \to \infty} L_N^Y(\epsilon) = 0, \quad \forall \epsilon > 0.$$

This shows that the Lindeberg condition is satisfied. Hence, we have that

$$\frac{S_N}{\sqrt{V_N}} = \frac{\hat{\theta}_{mn} - \theta}{\sqrt{\frac{\sigma^2}{m} + \frac{\tau^2}{n}}} \to_d \mathcal{N}(0, 1)$$

by the Lindeberg-Feller CLT.