Econometrics Practice

Date: 01/02/2007

OUTLINE

1.NTS:$var(\hat{\beta}) = E[var(\hat{\beta}|X)] + var[E(\hat{\beta}|X)]$

2.NTS: $cov(\hat{\beta}_{OLS}, \hat{\varepsilon}|X) = 0$

3. NTS: $var(s^2|X)$

4. One tailed t-test

5. M. S. Bartlett Identities

1. We have to show that

$$var(\hat{\beta}) = E[var(\hat{\beta}|X)] + var[E(\hat{\beta}|X)]$$

Proof. We have the following hint: By definition

$$var(\hat{\beta}|X) = E[(\hat{\beta} - E(\hat{\beta}|X))(\hat{\beta} - E(\hat{\beta}|X))^T|\ X]$$

and also

$$var[E(\hat{\beta}|X)] = E\{[E(\hat{\beta}|X) - E(\hat{\beta})][E(\hat{\beta}|X) - E(\hat{\beta})]^T\}$$

Let

$$\begin{aligned} d &= \hat{\beta} - E(\hat{\beta}|X) \\ a &= \hat{\beta} - E(\hat{\beta}) \\ c &= E(\hat{\beta}|X) - E(\hat{\beta}) \end{aligned}$$

then we observe first that

$$\begin{aligned} d &= a - c \\ dd^T &= aa^T - ca^T - ac^T + cc^T \end{aligned}$$

by taking the unconditional expectation and using the linearity of the expectation we have

$$E(dd^T) = E(aa^T) - E(ca^T) - E(ac^T) + E(cc^T)$$

By Law of Total expectation

$$E(dd^T) = E(E(dd^T|X)$$

inserting the expressions and using the first hint

$$E(dd^T) = E[(\hat{\beta} - E(\hat{\beta}|X))(\hat{\beta} - E(\hat{\beta}|X))^T|\ X] = E[var(\hat{\beta}|X)]$$

By definition of variance

$$E(aa^T) = var(\hat{\beta})$$

From the second hint we know that

$$E(cc^T) = var[E(\hat{\beta}|X)]$$

Finally we need to show what $E(ca^T)$ is. We first exploit Law of Total expectations

$$E(ca^T) = E[E(ca^T|X)]$$

then inserting the expressions

$$\begin{aligned}
E[E(ca^T|X)] &= E\{E[E(\hat{\beta}|X) - E(\hat{\beta})(\hat{\beta} - E(\hat{\beta}))^T|X]\} \\
&= E\{E(\hat{\beta}|X) - E(\hat{\beta})E[(\hat{\beta} - E(\hat{\beta}))^T|X]\} \\
&= E\{E(\hat{\beta}|X) - E(\hat{\beta})(E[(\hat{\beta}|X] - E(\hat{\beta})^T)\} \\
&= E(cc^T) = var[E(\hat{\beta}|X)]
\end{aligned}$$

The same way

$$E(ac^T) = var[E(\hat{\beta}|X)]$$

So we have shown that

$$\begin{aligned}
E(dd^T) &= E(aa^T) - E(ca^T) - E(ac^T) + E(cc^T) \\
E[var(\hat{\beta}|X)] &= var(\hat{\beta}) - 2var[E(\hat{\beta}|X)] + var[E(\hat{\beta}|X)] \\
E[var(\hat{\beta}|X)] &= var(\hat{\beta}) - var[E(\hat{\beta}|X)] \\
var(\hat{\beta}) &= E[var(\hat{\beta}|X)] + var[E(\hat{\beta}|X)].
\end{aligned}$$

2. We have to show the 4th finite sample property of OLS that we have introduced in class, namely

$$cov(\hat{\beta}_{OLS}, \hat{\varepsilon}|X) = 0$$

Proof. Under the assumptions 1-4 of the classical regression model, first we exploit the definition

$$cov(\hat{\beta}_{OLS}, \hat{\varepsilon}|X) = E[(\hat{\beta}_{OLS} - E(\hat{\beta}|X))(\hat{\varepsilon} - E(\hat{\varepsilon}|X))^T|X]$$

by the unbiasedness property of OLS, i.e $E(\hat{\beta}|X) = \beta$ and strict exogeneity of error terms, i.e $E(\hat{\varepsilon}|X) = 0$

$$E[(\hat{\beta}_{OLS} - \beta)\hat{\varepsilon}^T|X]$$

recall from the lecture that the sampling error is

$$(\hat{\beta}_{OLS} - \beta) = (X^T X)^{-1} X y - \beta = (X^T X)^{-1} X (X\beta + \varepsilon) - \beta = (X^T X)^{-1} X \varepsilon$$

and

$$\begin{aligned}
\hat{\varepsilon} &= M\varepsilon \Rightarrow \hat{\varepsilon}^T = \varepsilon^T M \\
M &= (\mathbf{I}_n - X(X^T X)^{-1} X^T) \\
M &= M^T \quad M = MM
\end{aligned}$$

given these expressions we have

$$E[(X^T X)^{-1} X \varepsilon\varepsilon^T M|X] = (X^T X)^{-1} X E[\varepsilon\varepsilon^T|X] M = \sigma^2 (X^T X)^{-1} X M$$

recall that annihilator matrix

$$XM = 0 \Rightarrow \sigma^2 (X^T X)^{-1} X M \Rightarrow cov(\hat{\beta}_{OLS}, \hat{\varepsilon}|X) = 0.$$

3. We have to show that under the assumptions 1-5

$$var(s^2|X) = \frac{2\sigma^4}{n-K}$$

Proof. First we recall that

$$s^2 = \frac{\hat{\varepsilon}^T \hat{\varepsilon}}{n-K}$$

A4:

$$E(\varepsilon\varepsilon^T|X) = \sigma^2 \mathbf{I}_n$$

we can also write A4 as

$$E(\frac{\varepsilon}{\sigma}\frac{\varepsilon^T}{\sigma}|X) = \mathbf{I}_n$$

Given A5:

$$\varepsilon|X \sim N(0, \sigma^2 \mathbf{I}_n)$$

we can exploit the fact from statistics that

$$\begin{aligned}
z &\sim N(0, \mathbf{I}) \\
& \quad M \text{ idempotent} \\
rank(M) &= m \\
z^T M z &\sim \mathcal{X}_m^2
\end{aligned}$$

and apply to

$$q = \frac{\varepsilon^T}{\sigma} M \frac{\varepsilon}{\sigma} \sim \mathcal{X}^2_{rank(M)}$$

recall from class in case of idempotent matrices rank(M)=trace(M), we have shown in class tha trace(M)=n-K, so

$$q \sim \mathcal{X}^2_{n-K}$$

Since

$$s^2 = \frac{\varepsilon^T M \varepsilon}{n-K} \Rightarrow \frac{s^2}{\sigma^2} = \frac{q}{n-K}$$

so

$$var(\frac{s^2}{\sigma^2}) = \frac{1}{(n-K)^2} var(q)$$

given the distribution of q

$$var(q) = 2 \times d.o.f = 2(n-K)$$

finally we have

$$var(s^2) = \frac{\sigma^4}{(n-K)^2} 2(n-K) = \frac{2\sigma^4}{n-K}$$

4. **One tailed t-test:** Under the assumptions 1-5 of classical regression model, we have the following model in our mind

$$\begin{aligned} y &= \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i \\ \varepsilon_i | X &\sim N(O, \sigma^2) \end{aligned}$$

we can write it in a compact way

$$\begin{aligned} y &= W\beta + \varepsilon \\ W &= \begin{bmatrix} \mathbf{1} & x_1 & x_2 \end{bmatrix} \\ \beta &= \begin{bmatrix} \beta_0 & \beta_1 & \beta_2 \end{bmatrix}^T \end{aligned}$$

We have the following null hypothesis

$$H_0 : \beta_1 \geq 4\beta_2 \; or \; \gamma = \beta_1 - 4\beta_2 \geq 0$$

Under null $\gamma = 0$ we have

$$\hat{\beta}_1 - 4\widehat{\beta}_2 | X \sim N(\beta_1 - 4\beta_2, var(\hat{\beta}_1 - 4\widehat{\beta}_2 | X))$$

recall that

$$\begin{aligned} var(\widehat{\beta}|X)) &= \sigma^2 (X^T X)^{-1} \\ var(\hat{\beta}_1 - 4\widehat{\beta}_2 | X) &= var(\hat{\beta}_1 | X) + 16 var(\hat{\beta}_2 | X) - 8 cov(\hat{\beta}_1, \hat{\beta}_2 | X) \end{aligned}$$

labeling ij'th element of $(X^TX)^{-1}$ as $(X^TX)^{-1}_{ij}$ we cab express the above equation as

$$\sigma^2(X^TX)^{-1}_{22} + 16\sigma^2(X^TX)^{-1}_{33} - 8\sigma^2(X^TX)^{-1}_{23}$$

Finally we can construct our one sided test

$$t = \frac{\hat{\beta}_1 - 4\widehat{\beta}_2}{\sqrt{s^2(X^TX)^{-1}_{22} + 16s^2(X^TX)^{-1}_{33} - 8s^2(X^TX)^{-1}_{23}}} \sim_{\gamma=0} t_{n-3}$$

notice that we replaced σ^2 with its estimator since the true value is not observed. Then we can decide we will reject or not reject the null hypothesis comparing the test statitics(that we have calculated) with the critical value(given the confidence level, usually t_c=2). We reject the null hypothesis if

$$t_c \leq t_{(n-3,\alpha)}$$

where n-3=n-K and a is the maximum size of the test. If it were a two sided test, then we reject the null hypothesis whenever

$$t_c \leq t_{(n-3,\alpha/2)}$$

5. **Bartlett Identities:** We will introduce two identities and prove them.

i) Expected value of the score is 0. $\Rightarrow E(S(\theta)) = 0$

Proof. Assume we have the likelihood function

$$f(z;\ \theta), \theta \in \Omega$$

By the property of the density function

$$\int_Z f(z;\ \theta)dz = 1 \Rightarrow \frac{\partial}{\partial\theta}\int_Z f(z;\ \theta)dz = 0$$

then we take the derivative inside the integral and use divide/multiply trick

$$\begin{aligned}\int_Z \frac{\partial f(z;\ \theta)}{\partial\theta}\frac{1}{f(z;\ \theta)}f(z;\ \theta)dz &= \int_Z \frac{\partial \log f(z;\ \theta)}{\partial\theta}f(z;\ \theta)dz = \\ &= \int_Z S(\theta)f(z;\ \theta)dz = E(S(\theta)) = 0\end{aligned}$$

ii) Variance of the score is the information matrix $\Rightarrow var(S(\theta)) = I(\theta) = E[-\frac{\partial^2 \log L(\theta)}{\partial\theta\partial\theta^T}|_{\widetilde{\theta}=\theta}]$

Proof. Using the finding in the previous point

$\int_Z \frac{\partial \log f(z;\ \theta)}{\partial\theta}f(z;\theta)dz = 0$

$\frac{\partial}{\partial\theta^T}\int_Z \frac{\partial \log f(z;\,\theta)}{\partial\theta} f(z;\theta)dz = \int_Z (\frac{\partial^2 \log f(z;\,\theta)}{\partial\theta\partial\theta^T} f(z;\theta) + \frac{\partial f(z;\,\theta)}{\partial\theta^T}\frac{1}{f(z;\,\theta)} f(z;\,\theta))dz = 0$

$\Rightarrow \int_Z \frac{\partial^2 \log f(z;\,\theta)}{\partial\theta\partial\theta^T} f(z;\theta)dz + \int_Z \frac{\partial \log f(z;\,\theta)}{\partial\theta}\frac{\partial \log f(z;\,\theta)}{\partial\theta^T} f(z;\,\theta))dz = 0$

$\Rightarrow E(S(\theta)S(\theta)^T) = E(-\frac{\partial^2 \log f(z;\,\theta)}{\partial\theta\partial\theta^T}) \Rightarrow^{E(S(\theta))=0} var(S(\theta)) = E(-\frac{\partial^2 \log f(z;\,\theta)}{\partial\theta\partial\theta^T}).$