Economics 620, Lecture 5:

The K-Variable Linear Model II

Third assumption (Normality):

$$
\begin{aligned}
& y; q(X\beta, \sigma^2 I_N) \\
\Rightarrow \quad & p(y) = \frac{1}{(2\pi\sigma^2)^{(N/2)}} \exp\left(-\frac{1}{2\sigma^2}(y - X\beta)'(y - X\beta)\right)
\end{aligned}
$$

where N is the sample size.

The log likelihood function is

$$\ell(\beta, \sigma^2) = c - \tfrac{N}{2} \ln \sigma^2 - \tfrac{1}{2\sigma^2}(y - X\beta)'(y - X\beta).$$

Proposition: The LS estimator $\hat{\beta}$ is the ML estimator.

Proposition: The ML estimator for σ^2 is

$$\sigma^2_{ML} = e'e/N.$$

Proof: To find the ML estimator for σ^2, we solve the FOC:

$$\begin{aligned}\frac{\partial \ell}{\partial \sigma^2} &= -\frac{N}{2\sigma^2} + \frac{1}{2\sigma^4}(y - X\beta)'(y - X\beta) = 0 \\ &\Rightarrow \sigma^2 = (y - X\beta)'(y - X\beta)/N\end{aligned}$$

Plugging in the MLE for β gives the MLE for σ^2

Proposition: The distribution of $\hat{\beta}$ given a value of σ^2 is $q(\beta, \sigma^2(X'X)^{-1})$.

Proof: Since $\hat{\beta}$ is a linear combination of jointly normal variables, it is normal. ■

Fact: If A is an $N \times N$ idempotent matrix with rank r, then there exists an $N \times N$ matrix C with

$C'C = I = CC'$ (orthogonal)

$C'AC = \Lambda$,

where:

$$\Lambda = \begin{bmatrix} 1 \ldots 0 \ldots 0 \\ 0 \ldots 1 \ldots 0 \\ \ldots\ldots\ldots\ldots \\ 0 \ldots\ldots\ldots 0 \end{bmatrix} = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}.$$

C is the matrix whose columns are orthornormal eigenvectors of A.

Lemma: Let $z \sim q(0, I_N)$ and A be an $N \times N$ idempotent matrix with rank r. Then

$$z'Az \sim \chi^2(r).$$

Proof:

$z'Az = z'CC'ACC'z = \tilde{z}C'AC\tilde{z} = \tilde{z}'\Lambda\tilde{z}$, where $\tilde{z}' = z'C$.

But $\tilde{z}$ is normal with mean zero and variance:

$$E\tilde{z}\tilde{z}' = EC'zz'C = C'(Ezz')C = C'C = I.$$

So, $z'Az = \tilde{z}'\Lambda\tilde{z}$ is the sum of squares of r standard normal variables, i.e., $z'Az \sim \chi^2(r)$. ■

Proposition:

$$\frac{N\sigma^2_{ML}}{\sigma^2} \sim \chi^2(N-K)$$

Proof: Note that $\sigma^2_{ML} = e'e/N = \varepsilon' M\varepsilon/N$.

$$\Rightarrow \frac{N\sigma^2_{ML}}{\sigma^2} = \frac{\varepsilon' M\varepsilon}{\sigma^2} \sim \chi^2(N-K),$$

using the previous lemma with $z = \varepsilon/\sigma$. ■

Proposition: $cov(\sigma^2_{ML}, \hat{\beta}) = 0$

Proof: $Ee(\hat{\beta} - \beta)' = EM\varepsilon((X'X)^{-1}X'\varepsilon)'$

$$= EM\varepsilon\varepsilon' X(X'X)^{-1}$$

$$= \sigma^2 MX(X'X)^{-1} = 0$$

$\Rightarrow e$ and $\hat{\beta}$ are *independent*.

(This depends on normality: zero covariance $\Rightarrow$ independence)

$\Rightarrow e'e$ and $\hat{\beta}$ are *independent*.■

So, we have the *complete* sampling distribution of $\hat{\beta}$ and σ^2_{ML}.

Note on t-testing:

We now that $\frac{\hat{\beta}-\beta_k}{\sigma_{\beta_k}} \sim q(0,1)$ where σ^2_{β}, is the k^{th} diagonal element of $\sigma^2(X'X)^{-1}$.

Estimating σ^2 by s^2 gives a statistic which is $t(N-K)$, using the same argument as in simple regression.

Simultaneous Restrictions

In multiple regression we can test several restrictions simultaneously. Why is this useful?

Recall our expenditure system:

$$\begin{aligned}\ln z_j &= \ln \frac{a_j}{\sum a_\ell} + \ln m - \ln p_j \\ \text{or } y &= \beta_0 + \beta_1 \ln m + \beta_2 \ln p_j + \varepsilon\end{aligned}$$

We are interested in the hypothesis $\beta_1 = 1$ *and* $\beta_2 = -1$. A composite hypothesis like this cannot be tested with the tools we have developed so far.

Lemma: Let $z \sim q(0, I)$, and A and B be symmetric idempotent matrices such that $AB = 0$.

Thus A and B are projections to orthogonal spaces. Then $a = z'Az$ and $b = z'Bz$ are independent.

Proof:

$a = z'A'Az =$ sum of squares of Az

$b = z'B'Bz =$ sum of squares of Bz.

Note that both Az and Bz are normal with mean zero.

$cov(Az, Bz) = EAzz'B' = AEzz'B' = AB' = 0$

We are done. (why?) ■

Note: A similar argument shows that $z'Az$ and Lz are independent if $AL' = 0$.

Testing

Definition: Suppose $v \sim \chi^2(k)$ and $u \sim \chi^2(p)$ are independent. Then

$$F = \frac{v/k}{u/p} \sim F(k,p).$$

Lemma: Let M and M^* be idempotent with $MM^* = M^*$, $e = M\varepsilon$, $e^* = M^*\varepsilon$, $\varepsilon \sim q(0, \sigma^2 I)$.

Then

$$F = \frac{(e'e - e^{*\prime}e^*)/(trM - trM^*)}{e^{*\prime}e^*/trM^*} \sim F(trM - trM^*, trM^*).$$

Proof: $\sigma^{-2} tr M^*$ times the denominator is $\chi^2(tr M^*)$

As for the numerator:

$$e'e - e^{*\prime}e^* = \varepsilon' M' M \varepsilon - \varepsilon' M^{*\prime} M^* \varepsilon = \varepsilon'(M - M^*)\varepsilon.$$

Note that: $(M - M^*)(M - M^*) = M^2 - M^*M - MM^* + M^{*2} = M - M^*$ (*idempotent*).

So $e'e - e^{*\prime}e^* = \varepsilon'(M - M^*)\varepsilon$.

Thus, the numerator upon multiplication by $\sigma^{-2} tr(M - M^*)$ is distributed as

$$\chi^2(tr(M - M^*)).$$

It only remains to show that the numerator and the denominator are independent.

But $(M - M^*)M^* = 0$, so we are done. ■

Interpretation:

$R[M^*] \subset R[M]$, i.e.

$e'e$ is a restricted sum of squares

$e^{*\prime}e^*$ is an unrestricted sum of squares.

F looks at the normalized reduction in "fit" caused by the restriction.

What sort of restrictions meet the conditions of the lemma?

Proposition: Let X be $N \times H$ and X^* be $N \times K$ where $H < K$. $(R[X] \subset R[X^*])$.

Suppose $X = X^*A$ $\qquad$ (A is $K \times H$).

Let $M = I - X(X'X)^{-1}X'$ and $M^* = I - X^*(X^{*\prime}X^*)^{-1}X^{*\prime}$.

Then M and M^* are idempotent and $MM^* = M^*$.

Example 1: Leaving out variables

Consider $y = X_1\beta_1 + X_2\beta_2 + \varepsilon$ where X_1 is $N \times K_1$ and X_2 is $N \times K_2$.

Hypothesis: $\beta_2 = 0$, i.e., X_2 is *not* in the model.

Using the notation from the previous proposition, $X = X_1$ and $X^* = [X_1 X_2]$

$$X = X^* A,\ A = \begin{bmatrix} I \\ 0 \end{bmatrix}$$

Note that: $trM = N - K_1$, $trM^* = N - K_1 - K_2$.

$$F = \frac{(e'e - e^{*\prime}e^*)/K_2}{e^{*\prime}e^*/(N - K_1 - K_2)}.$$

Thus:

e is from the regression of y on $X = X_1$, and
e^* is from the regression of y on $X^* = [X_1 X_2]$.

The degrees of freedom in the numerator is the number of restrictions.

Example 2: Testing the equality of regression coefficients in two samples.

Consider

$y_1 = X_1\beta_1 + \varepsilon_1$ where y_1 is $N_1 \times 1$ and X_1 is $N_1 \times K$, and
$y_2 = X_2\beta_2 + \varepsilon_2$ where y_2 is $N_2 \times 1$ and X_2 is $N_2 \times K$.

Hypothesis: $\beta_1 = \beta_2$

Combine the observations from the samples:

$$y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \; X^* = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix}, \; \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$$

The unrestricted model is

$$y = X^*\beta + \varepsilon = X_1\beta_1 + X_2\beta_2 + \varepsilon.$$

$$X = X^*A,\ A = \begin{bmatrix} I \\ I \end{bmatrix}$$

Note that $trM^* = N_1 + N_2 - 2K$ and

$$trM = N_1 + N_2 - K.$$

Run the restricted and unrestricted regressions, and calculate

$$F = \frac{(e'e - e^{*\prime}e^*)/K}{e^{*\prime}e^*/(N_1 + N_2 - 2K)}.$$

Example 3: Testing the equality of a subset of coefficients

Consider

$$y_1 = X_1\beta_1 + X_2\beta_2 + \varepsilon_1$$

where X_1 is $N_1 \times K_1$ and X_2 is $N_1 \times K_2$

and

$$y_2 = X_3\beta_3 + X_4\beta_4 + \varepsilon_2$$

where X_3 is $N_2 \times K_1$ and X_4 is $N_2 \times K_4$

Hypothesis: $\beta_1 = \beta_3$

The unrestricted regression is

$$y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & X_2 & 0 & 0 \\ 0 & 0 & X_3 & X_4 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} + \varepsilon$$

$$= X^*\beta + \varepsilon.$$

With the restriction, we have

$$\begin{aligned}
y &= \begin{bmatrix} X_1 & X_2 & 0 \\ X_3 & 0 & X_4 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \varepsilon \\
&= X\tilde{\beta} + \varepsilon. \\
X &= X^* A,\, A = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ I & 0 & 0 \\ 0 & 0 & I \end{bmatrix}.
\end{aligned}$$

Thus,

the test statistics is:

$$F = \frac{(e'e - e^{*\prime}e^*)/K_1}{e^{*\prime}e^*/(N_1 + N_2 - 2K_1 - K_2 - K_4)}.$$

Another way to look at the condition of the lemma:

Let β^* be the unrestricted coefficient vector and β be the restricted coefficient vector.

The lemma requires that there exist a matrix A such that $\beta^* = A\beta$.

What kinds or restrictions cannot be brought into this framework??

Consider $Ey = X_1\beta_1$ versus

$Ey = X_2\beta_2$.

The combined model is not in consideration.

