LECTURE 5: THE K-VARIABLE
LINEAR MODEL II

Third assumption (Normality):

\[ y ; q (X\beta, \sigma^2 I_N) \]

\[ p(y) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left(-\frac{1}{2\sigma^2} (y - X\beta)'(y - X\beta)\right) \]

where \( N \) is the sample size.

The log likelihood function is

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

**Proposition:** The LS estimator \( \hat{\beta} \) is the ML estimator.
**Proposition:** The ML estimator for $\sigma^2$ is
\[ \sigma_{ML}^2 = \frac{e'e}{N}. \]

**Proof:** To find the ML estimator for $\sigma^2$, we solve the FOC:
\[
\frac{\partial \ell}{\partial \sigma^2} = -\frac{N}{2\sigma^2} + \frac{1}{2\sigma^4} (y - X\beta)'(y - X\beta) = 2E \left[ \sum \frac{(x_i - \bar{x})\varepsilon_i}{\sum (x_i - \bar{x})^2} \sum (x_i - \bar{x})(\varepsilon_i - \bar{\varepsilon}) \right]
\]

**Proposition:** The distribution of $\hat{\beta}$ given a value of $\sigma^2$ 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:

$$
\Lambda = \begin{bmatrix}
1 & \cdots & 0 & \cdots & 0 \\
0 & \cdots & 1 & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & \cdots & 0
\end{bmatrix} = \begin{bmatrix}
I_r & 0 \\
0 & 0
\end{bmatrix}.
$$

$C$ is the matrix whose columns are orthonormal 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:

$$z'Az = z'CC'ACz = \tilde{z}'C'AC\tilde{z} = \tilde{z}'\Lambda\tilde{z},$$

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: \[
\frac{N\sigma_{ML}^2}{\sigma^2} \sim \chi^2(N - K)
\]

Proof: Note that \(\sigma_{ML}^2 = e'e/N = \varepsilon'M\varepsilon/N\).

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

using the previous lemma with \(z = \varepsilon/\sigma\). \(\blacksquare\)

Proposition: \(\text{cov}(\sigma_{ML}^2, \hat{\beta}) = 0\)

Proof: \[
Ee(\hat{\beta} - \beta)' = E M\varepsilon((X'X)^{-1}X'\varepsilon)' = E M\varepsilon\varepsilon'X(X'X)^{-1} = \sigma^2MX(X'X)^{-1} = 0
\]
⇒ e and $\hat{\beta}$ are independent.
(This depends on normality: zero covariance ⇒ independence)
⇒ e'e and $\hat{\beta}$ are independent.

So, we have the complete sampling distribution of $\hat{\beta}$ and $\sigma^{ML}_2$.

*Note on t-testing:*

$$\frac{\hat{\beta}_k - \beta_k}{\sigma_{\beta_k}} \sim q(0,1)$$

We know that $\sigma_{\beta_k}$ is the kth diagonal element of $\sigma^2 (X'X)^{-1}$.

Estimating $\sigma^2$ 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:

\[ \ln z_j = \ln \frac{a_j}{\sum a_{\ell}} + \ln m - \ln p_j \]

or \[ y = \beta_0 + \beta_1 \ln m + \beta_2 \ln p_j + \varepsilon \]

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:

\[ a = z'A'Az = \text{sum of squares of } Az \]
\[ b = z'B'Bz = \text{sum of squares of } Bz. \]

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

\[ \text{cov}(Az, Bz) = EAzz'B' = A Ezz' B' = AB' = 0. \]
We are done. (why?) ■

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

N.M. Kiefer, Cornell University, Economics 620, Lecture 5
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 $MM^* = M^*$, $e = Me$, $e^* = M^*e$, $\varepsilon \sim q(0, \sigma^2I)$. Then

$$F = \frac{(e'e - e^*e^*) / (\text{tr } M - \text{tr } M^*)}{e^*e^* / \text{tr } M^*} \sim F(\text{tr } M - \text{tr } M^*, \text{ tr } M).$$
Proof: \( \sigma^{-2} \text{tr } M^* \) times the denominator is \( \chi^2(\text{tr } M^*) \).

As for the numerator:

\[
e'e - e'*e* = \varepsilon'M'M\varepsilon - \varepsilon'M'*M\varepsilon = \varepsilon'(M - M^*)\varepsilon.
\]

Note that: 
\[
(M - M^*)(M - M^*) = M^2 - M*M - MM^* + M^*2 = M - M^* \quad \text{(idempotent)}.
\]

So \( e'e - e'*e* = \varepsilon'(M - M^*)\varepsilon \).
Thus, the numerator upon multiplication by \( \sigma^{-2} \text{tr } (M - M^*) \) is distributed as \( \chi^2(\text{tr } (M - M^*)) \).

It only remains to show that the numerator and the denominator are independent. 
But \( (M - M^*)M^* = 0 \), so we are done. ■
Interpretation:

\[ R[M^*] \subset R[M], \text{i.e.} \]

e'\epsilon e \text{ is a restricted sum of squares}
\[ e^{*'} e^* \text{ is an unrestricted sum of squares.} \]

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

Let \[ M = I - X(X'X)^{-1}X' \]
\[ M^* = I - X^*(X'^*X^*)^{-1}X'^*. \]

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 \( X^* = [X_1 \ X_2] \)

\[
X = X^* A, \quad A = \begin{bmatrix} I \\ 0 \end{bmatrix}
\]

Note that: \( \text{tr} \ M = N - K_1 \), \( \text{tr} \ M^* = N - K_1 - K_2 \).

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

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

\[ y_1 = X_1 \beta_1 + \epsilon_1 \]  where \( y_1 \) is \( N_{1 \times 1} \) and \( X_1 \) is \( N_{1 \times K} \), and

\[ y_2 = X_2 \beta_2 + \epsilon_2 \]  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:

\[
\begin{bmatrix}
y_1 \\
y_2
\end{bmatrix},
\begin{bmatrix}
X^* = X_1 & 0 \\
0 & X_2
\end{bmatrix},
\beta = \begin{bmatrix}
\beta_1 \\
\beta_2
\end{bmatrix}
\]
The unrestricted model is
\[ y = X^* \beta + \varepsilon = X_1 \beta_1 + X_2 \beta_2 + \varepsilon. \]

\[ X = X^* A, \quad A = \begin{bmatrix} I \\ I \end{bmatrix} \]

Note that \( \text{tr} \ M^* = N_1 + N_2 - 2K \) and
\[ \text{tr} \ M = N_1 + N_2 - K. \]

Run the restricted and unrestricted regressions, and calculate

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

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

\[
= X^* \beta + \varepsilon.
\]
With the restriction, we have

\[ y = \begin{bmatrix} X_1 & X_2 & 0 \\ X_3 & 0 & X_4 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_4 \end{bmatrix} + \varepsilon \]

\[ = X\tilde{\beta} + \varepsilon. \]

Thus,

\[ X = X^*A, \quad A = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 1 & 0 & 0 \\ 0 & 0 & I \end{bmatrix}. \]

The test statistic is:

\[ F = \frac{(e'e - e'^*e^*) / K_1}{e'^*e^* / (N_1 + N_2 - 2K_1 - K_2 - K_4)}. \]
Another way to look at the condition of the lemma:

Let $\beta^*$ be the unrestricted coefficient vector and $\beta$ be the restricted coefficient vector.

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