POINT & INTERVAL ESTIMATES

OLS ESTIMATOR $\hat{\beta} = (X'X)^{-1}X'y$ USES SAMPLE INFORMATION TO PROVIDE SPECIFIC NUMERICAL ESTIMATES OF THE COEFFICIENTS β_j. AN INTERVAL ESTIMATOR USES THE SAMPLING VARIANCE OF THE ESTIMATOR TO CONSTRUCT AN INTERVAL AROUND THE SPECIFIC VALUES OF THE COEFFICIENTS, WHICH ARE THE POINT ESTIMATES.

TYPE I & TYPE II ERRORS

A HYPOTHESIS H_0 IS EITHER TRUE OR FALSE. BUT BASED ON SAMPLE INFORMATION ONLY ONE CANNOT KNOW THE TRUE STATE OF THE WORLD & MUST MAKE A DECISION ABOUT TRUTH/FALSEHOOD OF H_0 BASED ON LIMITED INFORMATION ONLY.

	DECISIONS	
STATE OF THE WORLD	ACCEPT	REJECT
H_0 IS TRUE	CORRECT	TYPE I ERROR
H_0 IS FALSE	TYPE II ERROR	CORRECT

GENERALLY TYPE I ERROR IS TREATED AS MORE SERIOUS. THE DECISION RULE IS BASED ON A CONFIDENCE INTERVAL SO CONSTRUCTED THAT THE PROBABILITY THAT THE ~~SAMPLE~~ ESTIMATOR WOULD GENERATED AN ESTIMATE OUTSIDE OF THE INTERVAL WHEN H_0 IS TRUE IS ONLY α. HENCE IF H_0 IS REJECTED WHENEVER THE ESTIMATE FALLS OUTSIDE THE INTERVAL, THE PROBABILITY OF TYPE I ERROR IS α.

FROM A t-DISTRIBUTION WITH $p = n-k$ DEGREES OF FREEDOM, WE CAN FIND THE VALUE OF THE t-STATISTIC $t^*_{\alpha/2}$ SUCH THAT

$$PR\left\{ t \geq t^*_{\alpha/2} \right\} = \alpha/2$$

BY SYMMETRY,

$$PR\left\{ t \leq -t^*_{\alpha/2} \right\} = \alpha/2$$

HENCE $PR\left\{ -t^*_{\alpha/2} \leq t \leq t^*_{\alpha/2} \right\} = 1-\alpha$.

HERE $(1-\alpha)$ IS THE CONFIDENCE LEVEL & α IS THE SIGNIFICANCE LEVEL.

FOR THE 2-VARIABLE REGRESSION

$$t = \frac{\hat{\beta}_2 - \beta_2}{se(\hat{\beta}_2)} = \frac{\hat{\beta}_2 - \beta_2}{\sqrt{\Sigma (x_i - \bar{x})^2}}\hat{\sigma} \sim t_{n-2}$$

SUPPOSE THAT FOR $n-2$ d.f. $Pr\left\{ t \geq t^* \right\} = \alpha/2$

THEN

$$Pr\left\{ -t^* \leq \frac{\hat{\beta}_2 - \beta_2}{se(\hat{\beta}_2)} \leq t^* \right\} = 1-\alpha$$

$$\Rightarrow Pr\left\{ -t^* se(\hat{\beta}_2) \leq \hat{\beta}_2 - \beta_2 \leq t^* \cdot se(\hat{\beta}_2) \right\} = 1-\alpha.$$

$$\Rightarrow Pr\left\{ \hat{\beta}_2 - t^* \cdot se(\hat{\beta}_2) \leq \beta_2 \leq \hat{\beta}_2 + t^* \cdot se(\hat{\beta}_2) \right\} = 1-\alpha.$$

THE INTERVAL $\left(\hat{\beta}_2 \pm t^* \cdot se(\hat{\beta}_2) \right)$ IS AN INTERVAL ESTIMATOR OF β_2. $\hat{\beta}_2$ IS A POINT ESTIMATOR OF β_2.

TEST OF HYPOTHESIS $H_0: \beta_2 = \beta_2^0$
AGAINST $H_1: \beta_2 \neq \beta_2^0$

$$t = \frac{\hat{\beta}_2 - \beta_2^0}{se(\hat{\beta}_2)}.$$

REJECT H_0 IF $|t| > t^*$.

DO NOT REJECT IF $|t| < t^*$.

EXAMPLE

$$Y_i = \beta_1 + \beta_2 X_i + u_i$$

$$\bar{Y} = 20; \qquad \sum(X_i - \bar{X})^2 = 60;$$

$$\bar{X} = 10; \qquad \sum(Y_i - \bar{Y})^2 = 100;$$

$$\sum(X_i - \bar{X})(Y_i - \bar{Y}) = 30; \quad n = 22$$

$$\hat{\beta}_2 = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sum(X_i - \bar{X})^2} = \frac{30}{60} = 0.5$$

$$\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{X} = 20 - (0.5)(10) = 15$$

$$SST = \sum(Y_i - \bar{Y})^2 = 10.0$$

$$SSR = (\hat{\beta}_2)^2 \sum(X_i - \bar{X})^2 = 15$$

$$SSE = \sum e_i^2 = SST - SSR = 85$$

$$\hat{\sigma}^2 = \frac{\sum e_i^2}{n-2} = \frac{85}{22-2} = 4.25$$

$$VAR(\hat{\beta}_2) = \frac{\hat{\sigma}^2}{\sum(X_i - \bar{X})^2} = \frac{4.25}{60} = 0.0708$$

$$se(\hat{\beta}_2) = \sqrt{VAR(\hat{\beta}_2)} = 0.2661$$

SIGNIFICANCE LEVEL $= \alpha = 0.05$

$$t_{\alpha/2} = 2.086 = (t^*) \qquad (n-2 = 20)$$

HYPOTHESIS TESTING:

$$H_0: \quad \beta_2 = 1.0 \quad (= \beta_2^0)$$

$$t = \frac{\hat{\beta}_2 - \beta_2^0}{se(\hat{\beta}_2)} = \frac{0.05 - 1.0}{0.2261}$$

$$= -1.879$$

$$|t| = 1.879 < 2.089 \ (= t^*)$$

TEST OF SIGNIFICANCE OF OVERALL REGRESSION

CONSIDER THE K VARIABLE REGRESSION

$$y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + \cdots + \beta_k x_{ki} + u_i$$

AND $H_0: \quad \beta_2 = \beta_3 = \cdots = \beta_k = 0$

IF H_0 IS TRUE

$$y_i = \beta_1 + u_i \quad , \quad \hat{\beta}_1 = \bar{y}$$

$$e_i = y_i - \hat{\beta}_1 = y_i - \bar{y}$$

$$\sum e_i^2 = \sum (y_i - \bar{y})^2 \Rightarrow \quad SSE = SST.$$

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} = 0.$$

WE WANT TO TEST IF R^2 IS SIGNIFICANTLY DIFFERENT FROM 0.

THE RATIO

$$\frac{R^2}{1-R^2} = \frac{SSR/SST}{SSE/SST} = \frac{SSR}{SSE} = \frac{\hat{\beta}' X' M_0 X \hat{\beta}}{e'e}.$$

CONSIDER THE MATRIX

$$X = \begin{bmatrix} 1 & x_{12} & \cdots & x_{1k} \\ 1 & x_{22} & & x_{2k} \\ 1 & & & \\ \vdots & & & x_{nk} \\ 1 & x_{n2} & & \end{bmatrix}$$

& A MATRIX

$$J = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ 0 & 0 & & 1 \end{bmatrix}$$

X IS $n \times k$ &

J IS A $k \times (k-1)$ MATRIX.

THEN

$$XJ = \begin{bmatrix} 1 & x_{12} & \cdots & x_{1k} \\ 1 & x_{22} & & x_{2k} \\ & & & \\ 1 & x_{n2} & & x_{nk} \end{bmatrix} \begin{bmatrix} 0 & 0 & \cdot & 0 \\ 1 & 0 & \cdot & 0 \\ 0 & 1 & \cdot & 0 \\ 0 & 0 & \cdot & 1 \end{bmatrix} = \begin{bmatrix} x_{12} & \cdots & x_{1k} \\ x_{22} & \cdots & x_{2k} \\ & & \\ x_{n2} & \cdots & x_{nk} \end{bmatrix}$$

PARTITION β AS $\beta = \begin{pmatrix} \beta_1 \\ \delta \end{pmatrix}$ WHERE $\delta = \begin{pmatrix} \beta_2 \\ \beta_3 \\ \vdots \\ \beta_K \end{pmatrix}$

THEN $y = \cancel{\imath}\beta_1 + XJ\delta + u$

$M_0 y = M_0 XJ\delta + M_0 u$

DEFINE $M_0 XJ \equiv Z$

THEN $M_0 y = Z\delta + M_0 u$

$$\hat{\delta} = (Z'Z)^{-1} Z' M_0 y$$

$$= (Z'Z)^{-1} Z' M_0 (Z\delta + M_0 u)$$

$$= (Z'Z)^{-1} (Z' M_0 Z)\,\delta + (Z'Z)^{-1} Z' M_0 u$$

NOW $Z = M_0 X J$

$Z' = J'X'M_0$

$Z'M_0 = J'X'M_0M_0 = J'X'M_0 = Z'$

THUS $Z'M_0Z = Z'Z$ & $Z'M_0u = Z'u$

HENCE $\hat{\delta} = \delta + (Z'Z)^{-1}Z'u$

IF $\delta = 0$, $\hat{\delta} = (Z'Z)^{-1}Z'u$

HENCE $\hat{\delta}'Z'Z\hat{\delta} = u'Z(Z'Z)^{-1}Z'Z(Z'Z)^{-1}Z'u$

$= u'Z(Z'Z)^{-1}Z'u$

DEFINE $P_Z = Z(Z'Z)^{-1}Z'$ & $M_Z = I - P_Z$

CLEARLY BOTH MATRICES ARE SYMMETRIC & IDEMPOTENT.

ALSO $e'e = u'Mu$

HENCE,

$$\frac{R^2}{1-R^2} = \frac{\frac{\hat{\delta}' Z'Z \hat{\delta}}{\sigma^2}}{\frac{e'e}{\sigma^2}} = \frac{\left(\frac{u}{\sigma}\right)' P_z \left(\frac{u}{\sigma}\right)}{\left(\frac{u}{\sigma}\right)' M \left(\frac{u}{\sigma}\right)}.$$

FINALLY

$$P_Z M = (I - M_Z) M$$

$$= Z(Z'Z)^{-1} Z'M$$

NOW

$$Z = ~~M_0 X~~ \; M_0 X J$$

$$Z' = J'X'M_0$$

HENCE

$$Z'M = J'X'M_0 M$$

$$= J'X'\left(I - \frac{ZZ'}{n}\right) M$$

$$= J'X'M - J'X'\left(\frac{ZZ'}{n}\right) M$$

BUT

$$X'M = 0 \Rightarrow Z'M = 0$$

HENCE $Z'M = 0 \Rightarrow$

$$P_Z M = 0$$

HENCE $\left(\frac{u}{\sigma}\right)' P_Z \left(\frac{u}{\sigma}\right)$ & $\left(\frac{u}{\sigma}\right)' M \left(\frac{u}{\sigma}\right)$ ARE INDEPENDENTLY DISTRIBUTED

THUS

$$\frac{\hat{\delta}' Z'Z \hat{\delta}}{\sigma^2} \sim \chi^2_{k-1} \quad \&$$

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

$\frac{R^2/k-1}{1-R^2/n-k} = F_{k-1,n-k}$ IS A RATIO OF INDEPENDENT χ^2 VARIATES EACH DIVIDED BY THE RELEVANT DEGREES OF FREEDOM AND HAS THE χ^2 DISTRIBUTION.