Statistical Inference about β_k

Confidence Interval for β_k

$(1-\alpha)100\%$ CI for $\beta_k = \hat{\beta}_k \pm t_{\frac{\alpha}{2}}(n-K)se(\hat{\beta}_k)$

Hypothesis Testing for β_k

$$\mathrm{H}_0 : \beta_k = 0.6$$

$$\mathrm{H}_1 : \beta_k \neq 0.6 \qquad t_{cal} = \frac{\hat{\beta}_k - 0.6}{se(\hat{\beta}_k)}$$

$$\left|t_{cal}\right| < t_{\frac{\alpha}{2}}(n-K) \Rightarrow \text{accept } \mathrm{H}_0\text{. Otheriwse, reject } \mathrm{H}_0\text{.}$$

Testing for Effect of X_k on Y

Mean-independence of Y on X_k

$$\mathrm{H}_0 : \beta_k = 0$$

$$\mathrm{H}_1 : \beta_k \neq 0$$

$$t_{cal} = \frac{\hat{\beta}_k}{se(\hat{\beta}_k)}$$

Accept H_0 => X_k has no significant effect on Y

Overall F-test (1)

Assumption

There is a constant term in the model or X_1 is a vector of one. Why?

Test for mean-independence of Y on $[X_2, X_3, \ldots, X_K]$

$$\mathrm{H}_0 : \beta_2 = \beta_3 = \ldots = \beta_K = 0$$

$$\mathrm{H}_1 : \beta_2 \neq \beta_3 \neq \ldots \neq \beta_K \neq 0$$

Overall F-test (2)

We are choosing between

$Y = \beta_1 + \varepsilon$ ---- (H_0)

expect low R^2 when all X_k's are included

$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_K X_K + \varepsilon$ ---- (H_1)

expect higher R^2

Overall F-test (3)

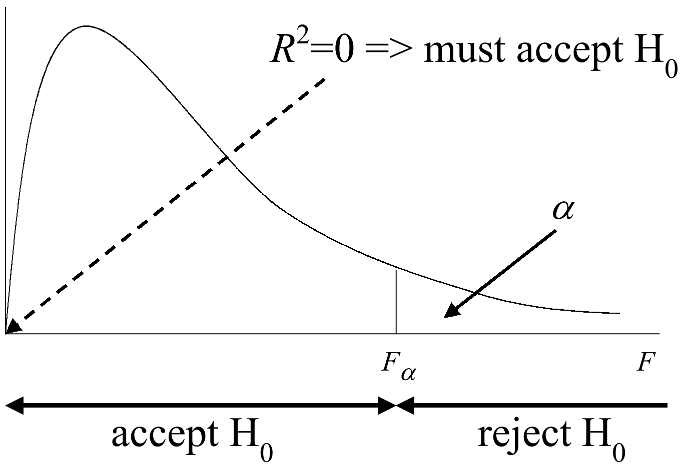
$$F_{cal} = \frac{R^2}{1-R^2}\frac{n-K}{K-1} \sim F(K-1, n-K)$$

Accept H_0 if $F_{cal} < F_{\alpha}(K\text{-}1,n\text{-}K)$. Otherwise, reject H_0. Note that

1) an F-test is always right-tailed.

2) we need a positive R^2.

Overall F-test (4)



Generalized F-test (1)

$$H_0 : \mathbf{H}(\boldsymbol{\beta}) = \mathbf{0}$$

$$H_1 : \mathbf{H}(\boldsymbol{\beta}) \neq \mathbf{0}$$

where

$\mathbf{H}(\beta)$ is a Mx1 vector function of β

Note that M must be less than K.

Generalized F-test (2)

$$H_0 : \begin{bmatrix} H_1(\boldsymbol{\beta}) \\ H_2(\boldsymbol{\beta}) \\ \vdots \\ H_M(\boldsymbol{\beta}) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad H_1 : \begin{bmatrix} H_1(\boldsymbol{\beta}) \\ H_2(\boldsymbol{\beta}) \\ \vdots \\ H_M(\boldsymbol{\beta}) \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

or

$$H_0 : H_1(\boldsymbol{\beta}) = 0, H_2(\boldsymbol{\beta}) = 0, \ldots, H_M(\boldsymbol{\beta}) = 0$$

$$H_1 : H_1(\boldsymbol{\beta}) \neq 0, H_2(\boldsymbol{\beta}) \neq 0, \ldots, H_M(\boldsymbol{\beta}) \neq 0$$

Generalized F-test (3)

Linear Restriction

$\mathbf{H}(\beta)$ is a Mx1 vector linear function of β

$$\mathbf{H}(\beta) = \mathbf{R}\beta - \mathbf{r}$$

where $\mathbf{R}$ is an MxK coefficient matrix with

Rank=M

$\mathbf{r}$ is a Mx1 constant vector

$$H_0 : \mathbf{R\beta} - \mathbf{r} = \mathbf{0} \text{ or } \mathbf{R\beta} = \mathbf{r}$$

$$H_1 : \mathbf{R\beta} - \mathbf{r} \neq \mathbf{0} \text{ or } \mathbf{R\beta} \neq \mathbf{r}$$

Generalized F-test (4)

Two approaches

- Restricted Least Square (RLS)
- Wald Test

Restricted Least Square (1)

Require two LS runs

Unrestricted run is the OLS run on the original model

==> SSR_U

where

SSR_U is the sum of squared residuals from the unrestricted run

Restricted Least Square (2)

Restricted LS run is as follows

$$\min_{\beta} [\mathbf{Y}-\mathbf{X}\beta]^T[\mathbf{Y}-\mathbf{X}\beta]$$

subject to $\mathbf{R}\beta=\mathbf{r}$

==> SSR_R

where

SSR_R is the sum of squared residuals from the restricted run

Restricted Least Square (3)

Transform RLS to OLS (Elimination Approach)

Define $\mathbf{R}$=[$\mathbf{A}$ $\mathbf{B}$] where

$\mathbf{A}$ is an MxM invertible sub-matrix of $\mathbf{R}$

$\mathbf{B}$ is the Mx(K-M) sub-matrix containing columns of $\mathbf{R}$ not in $\mathbf{A}$

Restricted Least Square (4)

Define $\mathbf{X}$=[$\mathbf{V}$ $\mathbf{W}$] where

$\mathbf{V}$ is an NxM sub-matrix of $\mathbf{X}$

$\mathbf{W}$ is the Nx(K-M) sub-matrix containing columns of $\mathbf{X}$ not in $\mathbf{V}$

Restricted Least Square (5)

Re-write the restriction as

$$[\mathbf{A}\,\mathbf{B}]\begin{bmatrix}\boldsymbol{\gamma}\\ \boldsymbol{\delta}\end{bmatrix} = \mathbf{r} \quad \text{or } \mathbf{A}\boldsymbol{\gamma} + \mathbf{B}\boldsymbol{\delta} = \mathbf{r}$$

where

γ is a Mx1 subset of β

δ is a (K-M)x1 subset of β

Restricted Least Square (6)

Re-write the model as

$$\mathbf{Y} = [\mathbf{V}\,\mathbf{W}]\begin{bmatrix}\boldsymbol{\gamma}\\ \boldsymbol{\delta}\end{bmatrix} + \boldsymbol{\varepsilon}$$

$$= \mathbf{V}\boldsymbol{\gamma} + \mathbf{W}\boldsymbol{\delta} + \boldsymbol{\varepsilon}$$

Restricted Least Square (7)

Since $\mathbf{A}$ is invertible,

$$\boldsymbol{\gamma} = \mathbf{A}^{-1}[\mathbf{r} - \mathbf{B}\boldsymbol{\delta}]$$

Substitute into the model.

$$\mathbf{Y} = \mathbf{V}\mathbf{A}^{-1}[\mathbf{r} - \mathbf{B}\boldsymbol{\delta}] + \mathbf{W}\boldsymbol{\delta} + \boldsymbol{\varepsilon}$$

$$\mathbf{Y} - \mathbf{V}\mathbf{A}^{-1}\mathbf{r} = [\mathbf{W} - \mathbf{V}\mathbf{A}^{-1}\mathbf{B}]\boldsymbol{\delta} + \boldsymbol{\varepsilon}$$

Restricted Least Square (8)

$$\mathbf{P} = \mathbf{Z}\boldsymbol{\delta} + \boldsymbol{\varepsilon}$$

where $\mathbf{P} = \mathbf{Y} - \mathbf{V}\mathbf{A}^{-1}\mathbf{r}$, $\mathbf{Z} = \mathbf{W} - \mathbf{V}\mathbf{A}^{-1}\mathbf{B}$

Apply OLS

$$\hat{\boldsymbol{\delta}} = [\mathbf{Z}^{\mathrm{T}}\mathbf{Z}]^{-1}\mathbf{Z}^{\mathrm{T}}\mathbf{P}$$

$$\hat{\boldsymbol{\gamma}} = \mathbf{A}^{-1}[\mathbf{r} - \mathbf{B}\hat{\boldsymbol{\delta}}]$$

$$\hat{\sigma}_R^2 = \frac{SSR_R}{n-(K-M)} \quad SSR_R = [\mathbf{P} - \mathbf{Z}\hat{\boldsymbol{\delta}}]^{\mathrm{T}}[\mathbf{P} - \mathbf{Z}\hat{\boldsymbol{\delta}}]$$

Restricted Least Square (9)

$$\mathrm{V}(\hat{\boldsymbol{\delta}}) = \sigma^2[\mathbf{Z}^{\mathrm{T}}\mathbf{Z}]^{-1}$$

$$\mathrm{V}(\hat{\boldsymbol{\gamma}}) = \mathbf{A}^{-1}\mathbf{B}\mathrm{V}(\hat{\boldsymbol{\delta}})\mathbf{B}^{\mathrm{T}}[\mathbf{A}^{\mathrm{T}}]^{-1}$$

$$= \sigma^2\mathbf{A}^{-1}\mathbf{B}[\mathbf{Z}^{\mathrm{T}}\mathbf{Z}]^{-1}\mathbf{B}^{\mathrm{T}}[\mathbf{A}^{\mathrm{T}}]^{-1}$$

$$\mathrm{COV}(\hat{\boldsymbol{\gamma}}, \hat{\boldsymbol{\delta}}) = \sigma^2\mathbf{A}^{-1}\mathbf{B}[\mathbf{Z}^{\mathrm{T}}\mathbf{Z}]^{-1}$$

$$\mathrm{V}(\hat{\boldsymbol{\beta}}_R) = \sigma^2\begin{bmatrix} \mathbf{A}^{-1}\mathbf{B}[\mathbf{Z}^{\mathrm{T}}\mathbf{Z}]^{-1}\mathbf{B}^{\mathrm{T}}[\mathbf{A}^{\mathrm{T}}]^{-1} & \mathbf{A}^{-1}\mathbf{B}[\mathbf{Z}^{\mathrm{T}}\mathbf{Z}]^{-1} \\ [\mathbf{Z}^{\mathrm{T}}\mathbf{Z}]^{-1}\mathbf{B}^{\mathrm{T}}[\mathbf{A}^{\mathrm{T}}]^{-1} & [\mathbf{Z}^{\mathrm{T}}\mathbf{Z}]^{-1} \end{bmatrix}$$

Restricted Least Square (10)

<u>Lagrange Method</u>

FOC $\quad -\mathbf{X}^{\mathbf{T}}[\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}] + \mathbf{R}^{\mathbf{T}}\hat{\boldsymbol{\lambda}} = \mathbf{0}$

$$\mathbf{X}^{\mathrm{T}}\mathbf{Y} - \mathbf{X}^{\mathrm{T}}\mathbf{X}\hat{\boldsymbol{\beta}}_R - \mathbf{R}^{\mathbf{T}}\hat{\boldsymbol{\lambda}} = \mathbf{0}$$

$$\hat{\boldsymbol{\beta}}_R = [\mathbf{X}^{\mathrm{T}}\mathbf{X}]^{-1}[\mathbf{X}^{\mathrm{T}}\mathbf{Y} - \mathbf{R}^{\mathbf{T}}\hat{\boldsymbol{\lambda}}]$$

$$= [\mathbf{X}^{\mathrm{T}}\mathbf{X}]^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{Y} - [\mathbf{X}^{\mathrm{T}}\mathbf{X}]^{-1}\mathbf{R}^{\mathbf{T}}\hat{\boldsymbol{\lambda}}$$

$$= \hat{\boldsymbol{\beta}}_U - [\mathbf{X}^{\mathrm{T}}\mathbf{X}]^{-1}\mathbf{R}^{\mathbf{T}}\hat{\boldsymbol{\lambda}}$$

Restricted Least Square (11)

Substitute into **Rβ=r**

$$[\mathbf{R}\hat{\boldsymbol{\beta}}_U - \mathbf{r}] - \mathbf{R}[\mathbf{X}^{\mathrm{T}}\mathbf{X}]^{-1}\mathbf{R}^{\mathbf{T}}\hat{\boldsymbol{\lambda}} = \mathbf{0}$$

$$\hat{\boldsymbol{\lambda}} = \mathbf{S}^{-1}[\mathbf{R}\hat{\boldsymbol{\beta}}_U - \mathbf{r}]$$

where $\mathbf{S} = \mathbf{R}[\mathbf{X}^{\mathrm{T}}\mathbf{X}]^{-1}\mathbf{R}^{\mathbf{T}}$

$$\hat{\boldsymbol{\beta}}_R = \hat{\boldsymbol{\beta}}_U - [\mathbf{X}^{\mathrm{T}}\mathbf{X}]^{-1}\mathbf{R}^{\mathbf{T}}\mathbf{S}^{-1}[\mathbf{R}\hat{\boldsymbol{\beta}}_U - \mathbf{r}]$$

$$= \left[\mathbf{I} - [\mathbf{X}^{\mathrm{T}}\mathbf{X}]^{-1}\mathbf{R}^{\mathbf{T}}\mathbf{S}^{-1}\mathbf{R}\right]\hat{\boldsymbol{\beta}}_U$$

$$+ [\mathbf{X}^{\mathrm{T}}\mathbf{X}]^{-1}\mathbf{R}^{\mathbf{T}}\mathbf{S}^{-1}\mathbf{r}$$

Restricted Least Square (12)

$$\mathrm{V}(\hat{\boldsymbol{\beta}}_R) = \sigma^2\mathbf{D}[\mathbf{X}^{\mathrm{T}}\mathbf{X}]^{-1}\mathbf{D}^{\mathbf{T}}$$

where $\mathbf{D} = \mathbf{I} - [\mathbf{X}^{\mathrm{T}}\mathbf{X}]^{-1}\mathbf{R}^{\mathbf{T}}\mathbf{S}^{-1}\mathbf{R}$

$$\hat{\sigma}_R^2 = \frac{SSR_R}{n-(K-M)}$$

where $SSR_R = [\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}_R]^{\mathrm{T}}[\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}_R]$

Prove that both RLS and LM yield identical result

Restricted Least Square (13)

$$F_{cal} = \frac{(SSR_R - SSR_U)/M}{SSR_U/(n-K)} \sim F(M, n-K)$$

where

M is the number of restriction equations or constraints or the number of rows in matrix **R**

Note that $\mathrm{df}_U = n\text{-}K$ and $\mathrm{df}_R = n\text{-}(K\text{-}M)$

Restricted Least Square (14)

$F_{\mathrm{cal}} < F_{\alpha}(M,n\text{-}K)$ ==>Accept H_0

or restriction holds

$F_{\mathrm{cal}} > F_{\alpha}(M,n\text{-}K)$ ==>Reject H_0 or restriction

does not holds

Wald Test (1)

Require only the Unrestricted run

==> $\hat{\boldsymbol{\beta}}, \hat{\sigma}^2$

$$F_{cal} = [\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r}]^{\mathrm{T}} \left[\mathbf{R}[\mathbf{X}^{\mathrm{T}}\mathbf{X}]^{-1}\mathbf{R}^{\mathrm{T}}\right]^{-1} [\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r}] \frac{1}{\widehat{\sigma^2} M}$$

$$\sim F(M, n-K)$$

Accept H_0 if $F_{cal} < F_{\alpha}(M,n\text{-}K)$.
Otherwise, reject H_0.

Wald Test (2)

<u>Concept</u>

Note that, given H_0 is true,

$$[\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r}] \sim \mathrm{MVN}(\mathbf{0}, \sigma^2\mathbf{R}[\mathbf{X}^{\mathrm{T}}\mathbf{X}]^{-1}\mathbf{R}^{\mathrm{T}})$$

Standardize a normal vector

$$\mathbf{Z} = \left[\sigma^2\mathbf{R}[\mathbf{X}^{\mathrm{T}}\mathbf{X}]^{-1}\mathbf{R}^{\mathrm{T}}\right]^{-\frac{1}{2}} [\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r}]$$

Wald Test (3)

Note that Z is a vector of *M* iid

standard normal RV's

$$\mathbf{Z}^{\mathrm{T}}\mathbf{Z} = [\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r}]^{\mathrm{T}} \left[\sigma^2\mathbf{R}[\mathbf{X}^{\mathrm{T}}\mathbf{X}]^{-1}\mathbf{R}^{\mathrm{T}}\right]^{-1} [\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r}]$$

$$\sim \chi^2(M)$$

Wald Test (4)

$$F_{cal} = \frac{\dfrac{\mathbf{Z}^{\mathrm{T}}\mathbf{Z}}{M}}{\dfrac{(n-K)\dfrac{\hat{\sigma}^2}{\sigma^2}}{n-K}} \sim F(M, n-K)$$

$$= [\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r}]^{\mathrm{T}} \left[\mathbf{R}[\mathbf{X}^{\mathrm{T}}\mathbf{X}]^{-1}\mathbf{R}^{\mathrm{T}}\right]^{-1} [\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r}] \frac{1}{\hat{\sigma}^2 M}$$

Example#1 (1)

Overall F-test is a simple case of

Generalized F-tests with

$$\underset{(K-1)\text{x}K}{\mathbf{R}} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}, \mathbf{r} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Example#1 (2)

<u>RLS Approach</u>

Since the restriction set is simple, the restricted model can be written as

$$Y_{\text{i}} = \beta_1 + \varepsilon_{\text{i}}$$

By OLS => $\hat{\beta}_1 = \overline{Y}$

$$SSR_R = \sum_{i=}^{n} (Y_i - \hat{\beta}_1)^2 = \sum_{i=}^{n} (Y_i - \overline{Y})^2$$

Example#1 (3)

Note that SSR_{R} = SST of the unrestricted model.

$$F_{cal} = \frac{SST_U - SSR_U}{SSR_U} \frac{n-K}{K-1}$$

$$= \frac{(SST_U - SSR_U)/SST_U}{SSR_U/SST_U} \frac{n-K}{K-1}$$

$$= \frac{R^2}{1-R^2} \frac{n-K}{K-1}$$

Example#1 (4)

<u>Wald Test</u> (single-run)

See Eviews example

Example#2 (1)

<u>Removing</u> X_2 and X_3

$$H_0 : \beta_2 = 0, \beta_3 = 0$$

$$H_1 : \beta_2 \neq 0, \beta_3 \neq 0$$

Use this **R** and **r** in the test

$$\underset{2\text{x}K}{\mathbf{R}} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \end{bmatrix}, \mathbf{r} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Example#2 (2)

<u>RLS Approach</u>

Since the restriction set is simple, the restricted model can be written as

$$Y_i = \beta_1 + \beta_4 X_{4i} + \ldots + \beta_K X_{Ki} + \varepsilon_i$$

Example#3 (1)

$$H_0 : \beta_2 = 0, \beta_3 = 0, \beta_4 + \beta_5 = 1$$

$$H_1 : \beta_2 \neq 0, \beta_3 \neq 0, \beta_4 + \beta_5 \neq 1$$

Use this **R** and **r** in the test

$$\underset{3\text{x}K}{\mathbf{R}} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & \cdots & 0 \end{bmatrix}, \mathbf{r} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Example#3 (2)

<u>RLS Approach</u>

Since the restriction set is simple, the restricted model can be written as

$$Y_i = \beta_1 + \beta_4 X_{4i} + (1-\beta_4) X_{5i} + \ldots + \beta_K X_{Ki} + \varepsilon_i$$

$$Y_i - X_{5i} = \beta_1 + \beta_4 (X_{4i} - X_{5i}) + \beta_6 X_{6i} \ldots + \beta_K X_{Ki} + \varepsilon_i$$

See EViews example

Normality Tests

- Cumulative Normal plot
- Goodness-of-fit test (a Chi-square test)
- Jarque-Bera Test

Cumulative Normal Plot (1)

If X is normal, graph of inverse CDF of cumulative relative frequency versus X will exhibit linearity

Step 1 Sort X

Step 2 Calculate Cumulative Relative frequency F for each X. Note that

$$0 \le F \le 1$$

Cumulative Normal Plot (2)

Step 3 Calculate (look for in the Z-table) the Z value for the area on left equal to F

Step 4 Plot Z against standardized X

If the graph is linear with slope of +1, ==> X~Normal

Jarque-Bera Normality Test (1)

$$\mathrm{H}_0 : S = 0, \kappa = 3$$
$$\mathrm{H}_1 : S \neq 0, \kappa \neq 3$$

where S is skewedness

κ is Kurtosis

$$\chi^2_{cal} = (n-K)\left(\frac{1}{6}\hat{S}^2 + \frac{1}{24}(\hat{\kappa}-3)^2\right) \sim \chi^2_{\alpha}(2)$$

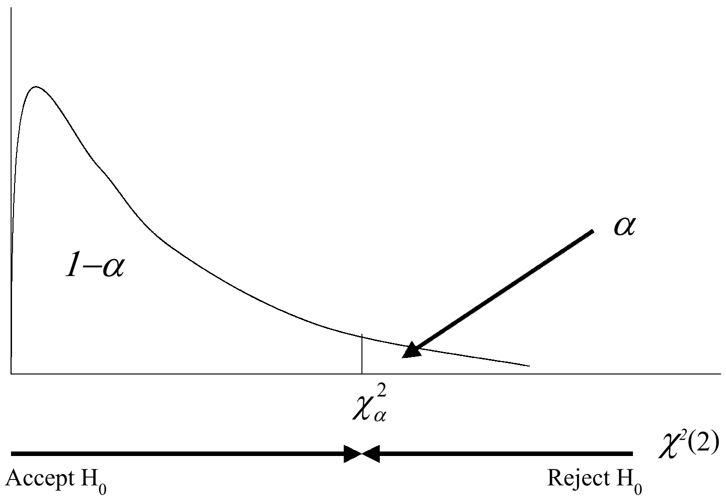
Jarque-Bera Normality Test (2)

where
$$\hat{\sigma} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}$$
$$\hat{S} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_i - \overline{X}}{\hat{\sigma}}\right)^3$$
$$\hat{\kappa} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_i - \overline{X}}{\hat{\sigma}}\right)^4$$

Perform a right-tailed χ^2-test

Note: different definition for skewedness and kurtosis

Jarque-Bera Normality Test (3)



Prediction Interval of Y (1)

$$\mathrm{E}(Y \mid \mathbf{X}_0) = \mathbf{X}_0\boldsymbol{\beta}$$
$$\widehat{\mathrm{E}(Y \mid \mathbf{X}_0)} = \mathbf{X}_0\hat{\boldsymbol{\beta}}$$

Is an <u>unbiased</u> estimator of $\mathrm{E}(Y|\mathbf{X}_0)$

where $\mathbf{X}_0 = [X_{10}, X_{20}, \ldots, X_{K0}]$

$$\mathrm{V}\left(\widehat{\mathrm{E}(Y \mid \mathbf{X}_0)}\right) = \mathbf{X}_0\mathbf{V}(\hat{\boldsymbol{\beta}})\left[\mathbf{X}_0\right]^{\mathrm{T}}$$
$$= \sigma^2\mathbf{X}_0[\mathbf{X}^{\mathrm{T}}\mathbf{X}]^{-1}\left[\mathbf{X}_0\right]^{\mathrm{T}}$$

Prediction Interval of Y (2)

$(1-\alpha)100\%$ CI for $\mathrm{E}(Y|\mathbf{X}_0) =$
$$= \mathbf{X}_0\hat{\boldsymbol{\beta}} + t_{\frac{\alpha}{2}}(n-K)\,\mathrm{se}\left(\widehat{\mathrm{E}(Y \mid \mathbf{X}_0)}\right)$$

where
$$\mathrm{se}\left(\widehat{\mathrm{E}(Y \mid \mathbf{X}_0)}\right) = \sqrt{\widehat{\sigma^2}\mathbf{X}_0[\mathbf{X}^{\mathrm{T}}\mathbf{X}]^{-1}\left[\mathbf{X}_0\right]^{\mathrm{T}}}$$

Prediction Interval of Y (3)

$(1\text{-}\alpha)100\%$ PI for $\mathrm{Y}|\mathbf{X}_0 =$

$$= \mathbf{X}_0\hat{\boldsymbol{\beta}} + t_{\frac{\alpha}{2}}(n-K)\mathrm{se}(Y \mid \mathbf{X}_0)$$

where

$$\mathrm{se}(Y \mid \mathbf{X}_0) = \sqrt{\widehat{\sigma^2}\left(1 + \mathbf{X}_0[\mathbf{X}^{\mathrm{T}}\mathbf{X}]^{-1}[\mathbf{X}_0]^{\mathrm{T}}\right)}$$