

Testing Multiple Hypotheses

- (I) Testing exclusion Restrictions
- (II) Overall F Test
- III Overall Linear Hypotheses Restrictions.

MLB

$$\log(\text{salary}) = \beta_1 + \beta_2 \text{years} + \beta_3 \text{gamesyr} + \beta_4 \text{bAVG} + \beta_5 \text{HRunsyr} + \beta_6 \text{rbisyr} + \epsilon.$$

1993 salary

years # years in league

games average games per year.

Performance	{	bAVG	career BA
		HR	HR/season
		RBI	

H₀: performance has no effect

$$\beta_4 = \beta_5 = \beta_6 = 0$$

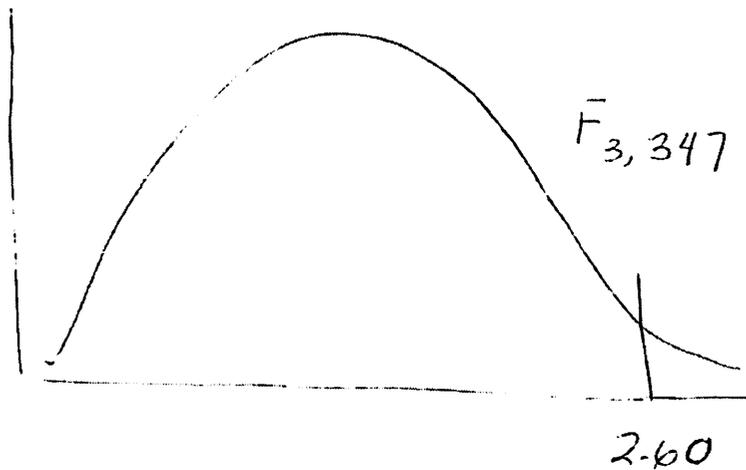
H_A: H₀ NOT TRUE

test stat

$$\frac{(SSR_R - SSR_u) / \# \text{ rehs.}}{SSR_u / (n - k)}$$

$k = \#$ params in unref model.

$$F = \frac{(198.311 - 183.186) / 3}{(183.186) / (353 - 6)} = 9.55$$



$9.55 > 2.60$ Reject.

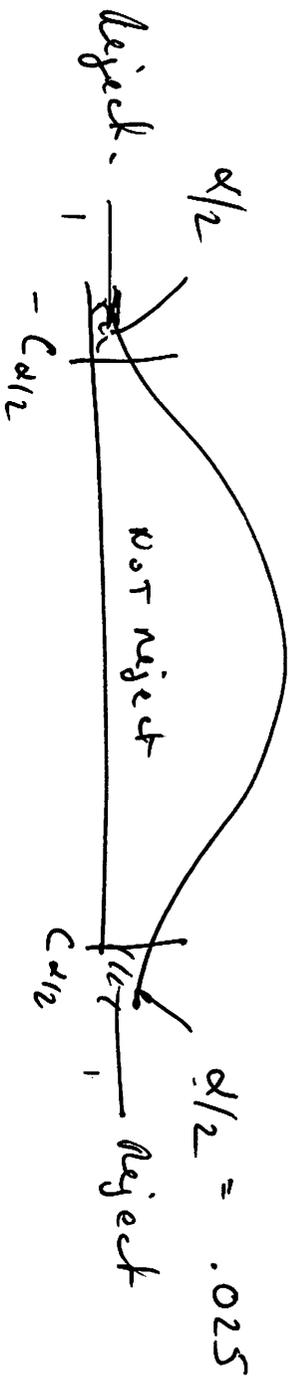
Performance Matters.

Testing a single hypothesis about β_i .

$H_0: \beta_i = c$ $H_A: \beta_i \neq c$

Pick α (Prob of Type I error)

$$t = \frac{\hat{\beta}_i - c}{SE(\hat{\beta}_i)} \quad \text{a } t\text{-stat if } H_0 \text{ True.}$$



if $|t| > c_{\alpha/2}$ then Reject H_0 .

or if t value $< -c_{\alpha/2}$ then Reject H_0 .

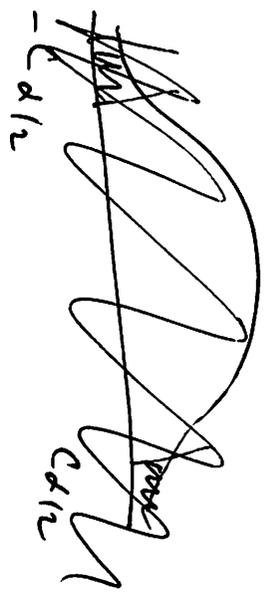
Single Hypothesis, Multiple Parameters

$H_0: \beta_2 + \beta_3 = 1$ $H_A: \beta_2 + \beta_3 > 1$

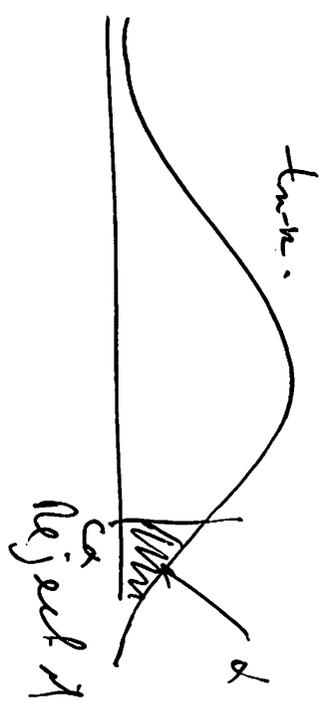
$H_0: \beta_2 + \beta_3 - 1 = 0$ $\beta_2 + \beta_3 - 1 > 0$

$$t = \frac{\beta_2 + \beta_3 - 1}{\sqrt{\text{SE}(\beta_2 + \beta_3 - 1)}} \approx 0 \quad \text{mark it}$$

$\text{SE}(\beta_2 + \beta_3 - 1)$ H_0 True.



Reject if



Reject if $t > c_{\alpha}$.

Testing Multiple Hypotheses.

$$y_i = \beta_1 + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + e_i$$

$$E(e_i | x_i) = 0$$

$$\text{Var}(e_i | x_i) = \sigma^2$$

$\text{Cov}(e_i, e_j | x_i, x_j) = 0 \quad i \neq j$
no perfect collinearity

OCS is normally dist (approximately, at least)

$$H_0: \beta_2 = 0, \beta_3 = 0, \beta_4 = 0 \quad \text{AND} \quad \text{AND}$$

$$H_A: \beta_2 \neq 0 \text{ or } \beta_3 \neq 0 \text{ or } \beta_4 \neq 0$$

hypotheses $\equiv J$ (example $J=3$)

$$y_i = \beta_1 + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + e_i$$

$$E(e_i | x_{i2}, \dots, x_{ik}) = 0$$

$$\text{Var}(e_i | x_{i2}, \dots, x_{ik}) = \sigma^2$$

$$\text{Cov}(e_i, e_j | x_i) = 0 \quad i \neq j$$

No perfect collinearity.

Under this DGP. Then LS is at least approximately normally distributed.

$$\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k \overset{\sim}{\sim} N(\quad)$$

If $e_i \sim N(0, \sigma^2)$ Then LS is normally distributed.

① Test - Statistic. (NOTE: Model is Homoskedastic)

$$\text{Var}(e_i | y) = \sigma^2$$

variance.

i) Estimate the model using LS.

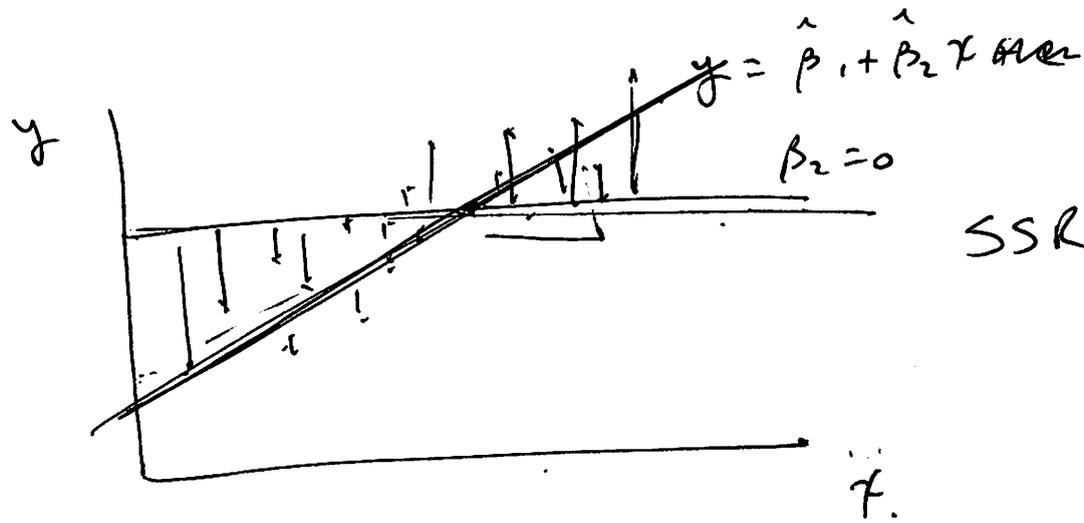
~~est~~

ii) ~~Est~~ Estimate a Restricted Model

where the hypothesis under H_0 are imposed on the model are Restrictions

F test statistic

$$F = \frac{(SSR_R - SSR_u) / J}{SSR_u / (n - k)}$$



SSR is minimized
By OLS

$SSR_R >>$ $H_0: \beta_2 = 0$ $y = \beta_1 + e$ Restricted.
 $SSR_u <<$ $H_A: \beta_2 \neq 0$ $y = \beta_1 + \beta_2 x + e$ Unrest.

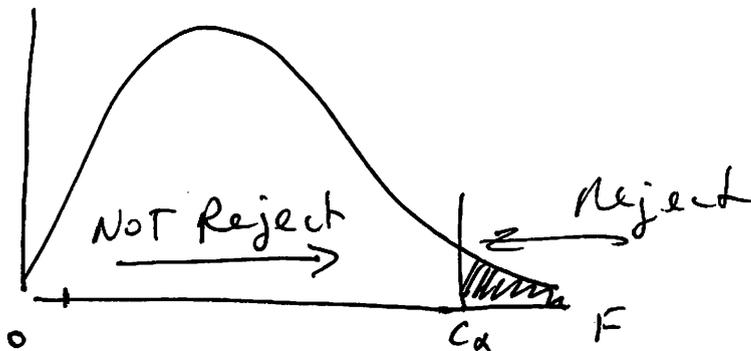
$$SSR_R - SSR_u \geq 0$$

If n is "large" enough

Then $F \sim F_{5, n-k}$ if H_0 is true.

as $n \rightarrow \infty$ $F_{5, n-k} \rightarrow \chi^2_{5/5}$.

$F_{5, n-k}$



if $F > C_\alpha$ then reject H_0 .

IG Model / DATA are

Heteroskedastic you need to compute a different Statistic.

Wald Stat use the LS coeff estimates as well as a consistent estimator of LS Precision (Variance-covariance).

Example $H_0: \beta_2 = 0$ $H_a: \beta_2 \neq 0$

$$\frac{\hat{\beta}_2 - 0}{\text{SE}(\hat{\beta}_2)} \sim t_{n-k} \text{ if } H_0 \text{ True}$$

And if

$\text{SE}(\hat{\beta}_2) \xrightarrow{P} \text{SE}(\hat{\beta}_2)$
consistently est STS errors.

If the model is heteroskedastic
you will use a consistent
estimator of LS & covariance
matrix (i.e., 1. Theil is ROBUST
to the existence of heteroskedastic
errors.

ROBUST = even if model is misspecified
(heteroskedastic) then cov is estimated
consistently.

HCCME

Heteroskedasticity Consistent Covariance Matrix Estimator.

Huber (1962), White (1978) - White's covariance estimator.
White's standard errors.

Matrix representation of Linear Model.

$$y_i = \beta_1 + \beta_2 x_i + e_i \quad i = 1, 2, \dots, n$$

$$y_1 = \beta_1 + \beta_2 x_1 + e_1$$

$$y_2 = \beta_1 + \beta_2 x_2 + e_2$$

$$y_3 = \beta_1 + \beta_2 x_3 + e_3$$

⋮

$$y_n = \beta_1 + \beta_2 x_n + e_n$$

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \quad e = \begin{pmatrix} e_1 \\ e_2 \\ e_3 \\ \vdots \\ e_n \end{pmatrix} \quad X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \quad \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$$

$$y = X\beta + e$$

$$E(e_i | x_1, x_2, \dots, x_n) = 0 \text{ strong exogeneity.}$$

$$E(e_n | X) = 0$$

$$\{ \text{Var}(e_i) = \sigma^2 \quad \text{Cov}(e_i, e_j) = 0 \quad i \neq j \}$$

$$\text{Cov}(e) = E \left[(e - E(e)) (e - E(e))^T \right]$$

$$\text{and } E(e) = 0$$

$$= E \left[e e^T \right] = \text{Cov}(e)$$

$$3 \cdot 3^{-1} = 1$$

$$3 \cdot \frac{1}{3} = 1$$

$$A \cdot A^{-1} = I_n$$

$$\text{Cov}(e) = \begin{pmatrix} \sigma^2 & 0 & \dots & 0 \\ 0 & \sigma^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma^2 \end{pmatrix} = \sigma^2 I_n$$

$$I_m = \begin{bmatrix} 1 & 0 & & 0 \\ 0 & 1 & & \\ & & \ddots & \\ 0 & & & 1 \end{bmatrix}$$

$$A \cdot I = A$$

$$I \cdot A = A$$

~~TS~~ ~~also~~ Method of Moments definable.

$$E(e|X) = 0 \quad \text{strong exog.}$$

$\Rightarrow e$ and X cannot be correlated if strong exog.

$$E \begin{bmatrix} X^T e \\ 2 \times m \quad n \times 1 \\ 2 \times 1 \end{bmatrix} = 0$$

$$E \begin{bmatrix} X^T (y - X\beta) \\ 2 \times m \quad n \times 1 - n \times 1 \\ 2 \times 1 \end{bmatrix} = 0 \quad \text{Pop Moment}$$

$$X \quad 2 \times m$$

$$X^T \quad 2 \times m$$

$$\text{Sample } \frac{1}{n} \begin{bmatrix} X^T (y - X\hat{\beta}) \\ 2 \times m \quad n \times 1 \\ 2 \times 1 \end{bmatrix} = 0$$

Solve for $\hat{\beta}$ is now. Since it we sample to solve this.

$$\text{Cov}(\hat{\beta})_{2 \times 2} = \begin{bmatrix} \text{Var}(\hat{\beta}_1) & \text{Cov}(\hat{\beta}_1, \hat{\beta}_2) \\ \text{Cov}(\hat{\beta}_1, \hat{\beta}_2) & \text{Var}(\hat{\beta}_2) \end{bmatrix} = E \left[\left(\hat{\beta} - E(\hat{\beta}) \right) \left(\hat{\beta} - E(\hat{\beta}) \right)^T \right]_{2 \times 2}$$

↳ or $k \times k$ in multiple regression
if OLS is unbiased $E(\hat{\beta}) = \beta$

$$\text{Cov}(\hat{\beta}) = E \left[\left(\hat{\beta} - \beta \right) \left(\hat{\beta} - \beta \right)^T \right] =$$

$$\begin{aligned} \hat{\beta} &= (X^T X)^{-1} X^T y = (X^T X)^{-1} X^T (\beta + \epsilon) \\ &= \beta + (X^T X)^{-1} X^T \epsilon \end{aligned}$$

$$\hat{\beta} - \beta = \underset{n \times 1}{(X^T X)^{-1}} \underset{2 \times 2}{X^T} \underset{2 \times n}{\epsilon}$$

$$E \left[\left((X^T X)^{-1} X^T \epsilon \right) \left((X^T X)^{-1} X^T \epsilon \right)^T \right]$$

$$(A B C)^T = C^T B^T A^T$$

$$(X^T X)^{-1} = \left[(X^T X)^{-1} \right]^T$$

Symmetric!

$$\begin{pmatrix} \underline{y} = X\underline{\beta} + \underline{e} & \underline{e} \sim (0, \sigma^2 I_n) \\ \underline{\hat{\beta}} = (X^T X)^{-1} X^T \underline{y} & \text{Rank}(X) = 2 \text{ (or } k) \end{pmatrix}$$

$(X^T X)$ 2×2 — Square. # Col's = # Rows
 Symmetric
 $n \times n$ $n \times 2$

$X^T X$ can be inverted as long as you have no perfect collinearity.

Properties of LS:

$$\begin{aligned} E(\hat{\beta} | X) &= E \left[(X^T X)^{-1} X^T y \mid X \right] \\ &= E \left[(X^T X)^{-1} X^T (X\beta + e) \mid X \right] \\ &= E \left[(X^T X)^{-1} X^T X \beta + (X^T X)^{-1} X^T e \mid X \right] \\ &= E(\beta) + E \left[(X^T X)^{-1} X^T e \mid X \right] \\ &= \beta + (X^T X)^{-1} X^T 0 = \beta \end{aligned}$$

$$X^T y - X^T X \hat{\beta} = 0$$

$$(X^T X) \hat{\beta} = X^T y$$

$$(X^T X)^{-1} X^T X \hat{\beta} = (X^T X)^{-1} X^T y = \hat{\beta} \quad \begin{matrix} \text{matrix} \\ \text{matrix} \\ \text{matrix} \end{matrix} \quad \begin{matrix} \text{matrix} \\ \text{matrix} \\ \text{matrix} \end{matrix} = \begin{matrix} \beta_1 \\ \beta_2 \end{matrix}$$

The MOM is same as LS for the model

$$S = \sum_{i=1}^n e_i^2 = e_1^T e_2^T \dots e_n^T$$

$$= (e_1 \ e_2 \ \dots \ e_n) \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix} = \begin{matrix} 1 \times n \\ n \times 1 \end{matrix}$$

$$e_1^2 + e_2^2 + \dots + e_n^2 = S$$

$$S = e^T e = (y - X\beta)^T (y - X\beta)$$

$$\frac{dS}{d\beta} = 0 \quad \text{solve for } \beta$$

$$E \left[\begin{matrix} e \\ e \\ \vdots \\ e \end{matrix} \begin{matrix} 1 \\ x \end{matrix} \right] = \sigma^2 I_n \text{ from the Model.}$$

$$\text{Cov}(\beta) = E \left[(X^T X)^{-1} X^T e e^T X (X^T X)^{-1} \right]$$

$$\text{Cov}(\beta) = \underbrace{E \left[(X^T X)^{-1} X^T \sigma^2 I_n X (X^T X)^{-1} \right]}_{E(e e^T | X)}$$

$$\text{Cov}(\hat{\beta}) = \sigma^2 (X^T X)^{-1} X^T X (X^T X)^{-1} = \sigma^2 (X^T X)^{-1}$$

just estimate σ^2 using usual estimator.

Suppose $e \sim (0, \sigma^2 \psi)$

$$E(e e^T) = \text{Cov}(e) = \sigma^2 \psi = \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_n^2 \end{bmatrix} \neq \sigma^2 I_n \text{ but heteroskedastic}$$

No autocorrelation

$$\text{Cov}(\hat{\beta}) = E \left((X^T X)^{-1} X^T e e^T X (X^T X)^{-1} \right)$$

$$(X^T X)^{-1} X^T \sigma^2 \psi X (X^T X)^{-1} = \sigma^2 (X^T X)^{-1} X^T \psi X (X^T X)^{-1}$$

LS covariance if ψ errors are heteroskedastic.

Write (and Huber) figured out a way

to estimate

$$(2) \quad \underbrace{\sigma^2 (X^T X)^{-1} X^T y X (X^T X)^{-1}}_{\text{using LS estimate of } \beta} \text{ consistently}$$

HCCME is an estimator of $\text{Var}(\hat{\beta})$

"Sandwich" covariance.

$$e_i = y - \beta_1 x_i - \beta_2 x_i^2 \quad e_i^2 \quad i=1, 2, \dots, n$$

$$\underbrace{X^T y X}_{HCO} = X^T \begin{bmatrix} e_1^2 & 0 \\ e_2^2 & 0 \\ e_3^2 & 0 \\ \vdots & 0 \\ e_n^2 & 0 \end{bmatrix} X$$

$$\text{EST Cov}(\hat{\beta}) = \sigma^2 (X^T X)^{-1} X^T y X (X^T X)^{-1}$$

1. if errors of Model are Homoskedastic (and not auto correlated)

OCS is BLUE and you estimate $\sigma^2 (X^T X)^{-1}$

2. if errors of Model are heteroskedastic

LS is no longer BLUE

LS is unbiased, But not efficient.

LS now has a sandwich covariance.

$$\sigma^2 (X^T X)^{-1} X^T \Phi \Phi^T (X^T X)^{-1} \quad (2)$$

which depends on σ^2 AND Φ

Valid tests and CI require a consistent estimator of (2). And HCCME is easily computed.