

Hypothesis Testing

$$Y_L = \beta_0 + \beta_1 X_L + U_L$$

formulate two opposing hypotheses

H_0 : Null hypothesis

H_A : Alternative hypothesis

two sided test

$$H_0: \beta_i = \beta_i^*$$

$$H_A: \beta_i \neq \beta_i^*$$

$i=2 \rightarrow$ two sided test

one sided tests

$$H_0: \beta_i > \beta_i^*$$

$$H_A: \beta_i \leq \beta_i^*$$

or

$$H_0: \beta_i < \beta_i^*$$

$$H_A: \beta_i \geq \beta_i^*$$

$i=1$

testing

1st set α (often 0.05, significance level; probability of TYPE I error)

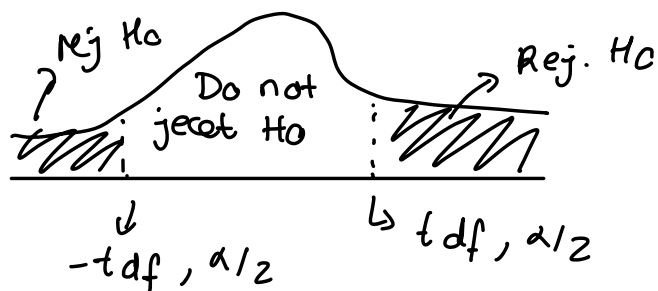
the distribution of the test static under the maintained null hypothesis

$$H_0: \beta_i = \beta_i^*$$

$$H_A: \beta_i \neq \beta_i^*$$

$$t_c = \frac{\hat{\beta}_i - \beta_i^*}{\hat{\sigma}_{\hat{\beta}_i}} \sim t_{df}, \alpha/2$$

decision rule



decision rule:

i) compute t_c

ii) Rej H_0 if $|t_c| \geq t_{df, \alpha/2}$

Restricted & Unrestricted Models

restricted model: the model obtained under the maintained null hypothesis

consider

$$(1) Y_t = \beta_0 + \beta_1 X_t + u_t$$

examples:

$$H_0: \beta_0 = 0$$

$$H_A: \beta_0 \neq 0$$

unrestricted model: (1)

$$\text{restricted model: } Y_t = \beta_1 X_t + u_t$$

examples

$$\text{URM } Y_t = \beta_0 + \beta_1 X_t + u_t$$

$$H_0: \beta_1 = 0 \quad H_A: \beta_1 \neq 0$$

$$\text{RM } Y_t = \beta_0 + u_t$$

empirical examples

$$Y_t = 0.08 - 1.50 X_t \quad R^2 = 0.90$$

$$(4.00) \quad (-4.50)$$

parenthesis could either be t -ratios under $H_0: \beta_i = 0$ or estimated standard error

* standard error is always positive. So in this example numbers are t -ratios

$$\beta_0 = 0.08 : t_0 = 4.00$$

$$\beta_1 = -1.5 : t_1 = -4.5$$

Hypothesis for slope (β_1)

$$H_0 : \beta_1 = 0$$

$$H_A : \beta_1 \neq 0$$

Hypothesis for intercept (β_0)

$$H_0 : \beta_0 = 0$$

$$H_A : \beta_0 \neq 0$$

$$\alpha = 0.05 \quad t_0 = 4.0 \quad t_1 = -4.5$$

for β_1 , since $t_1 = -4.5$, and it's below -1.96 , we reject the null hypothesis $H_0: \beta_1 = 0$, so there is statistically significant evidence that X_t has a negative impact on Y_t

for β_0 $t = 4$ and it is significantly above 1.96

We reject the null hypothesis $H_0: \beta_0 = 0$. Therefore, there is statistically significant evidence that the expected value Y_t when $X_t = 0$ is not zero and is indeed positive

consider

$$Y_t = \beta_0 + \beta_1 X_t + U_t$$

$$H_0: \beta_0 = \beta_0^*$$

$$H_A: \beta_0 \neq \beta_0^*$$

$$\beta_1 = \beta_1^*$$

$$\beta_1 \neq \beta_1^*$$

for instance

$$H_0: \begin{matrix} \beta_0 = 0 \\ \beta_1 = 0 \end{matrix}$$

$$H_A: \begin{matrix} \beta_0 \neq 0 \\ \beta_1 \neq 0 \end{matrix}$$

More than one linear restriction. therefore the t-test cannot be used. we need to consider F-test

3- The following regression model is given

$$Y_t = \beta \ln X_t + U_t$$

$$E(U_t) = 0 \text{ where } t = 1, \dots, T$$

$$E(U^2) = \sigma^2$$

$$E(U_t U_s) = 0 \quad t \neq s$$

a) derive the least squares estimator of β

$$\min SSR = \sum \hat{U}_t^2$$

$$= \sum (Y_t - \hat{Y}_t)^2$$

$$= \sum (Y_t - \hat{\beta} \ln X_t)^2$$

$$\frac{\partial SSR}{\partial \hat{\beta}} = -2 \sum (Y_t - \hat{\beta} \ln X_t) \ln X_t = 0$$

$$\sum Y_t \ln X_t - \hat{\beta} (\ln X_t)^2 = 0$$

$$\sum Y_t \ln X_t = \hat{\beta} (\ln X_t)^2$$

$$\hat{\beta} = \frac{\sum Y_t \ln X_t}{\sum (\ln X_t)^2}$$

b) derive the variance of $\hat{\beta}$

$$\hat{\beta} = \sum \underbrace{\frac{\ln X_t}{\sum (\ln X_t)^2}}_{w_t} Y_t \Rightarrow \hat{\beta} = \sum w_t Y_t$$

$$\hat{\beta} = \sum w_t (\beta \ln X_t + U_t) = \sum \beta w_t \ln X_t + \sum w_t U_t$$

$$\hat{\beta} = \beta \sum w_t \ln X_t + \sum w_t U_t$$

$$\hat{\beta} = \beta \sum \frac{\ln X_t}{\sum (\ln X_t)^2} \ln X_t + \sum w_t U_t$$

$$\hat{\beta} = \beta \underbrace{\sum \frac{\ln X_t^2}{\sum (\ln X_t)^2}}_{=1} + \sum w_t u_t$$

$$\hat{\beta} = \beta + \sum w_t u_t \rightarrow E(\hat{\beta}) = E(\beta + \sum w_t u_t)$$

$$E(\hat{\beta}) = E(\beta) + E(\sum w_t u_t)$$

$$= \beta + \sum w_t \underbrace{E(u_t)}_{=0}$$

$$\underline{E(\hat{\beta}) = \beta}$$

$$\text{Var}(\hat{\beta}) = E(\hat{\beta} - E(\hat{\beta}))^2 = E(\hat{\beta} - \beta)^2 = E(\sum w_t u_t)^2$$

$$E(w_1 u_1 + w_2 u_2 + w_3 u_3 + \dots + w_T u_T)^2$$

$$= E(w_1^2 u_1^2 + w_2^2 u_2^2 + w_3^2 u_3^2 + \dots + w_T^2 u_T^2 + w_1 w_2 u_1 u_2 + w_1 w_3 u_1 u_3 + \dots + w_{T-1} w_T u_{T-1} u_T)$$

$$= w_1^2 \underbrace{E(u_1^2)}_{=\sigma^2} + w_2^2 \underbrace{E(u_2^2)}_{=\sigma^2} + \dots + w_T^2 \underbrace{E(u_T^2)}_{=\sigma^2} + \dots +$$

$$w_1 w_2 \underbrace{E(u_1 u_2)}_{=0} + \dots + w_{T-1} w_T \underbrace{E(u_{T-1} u_T)}_{=0}$$

$$w_1^2 \sigma^2 + w_2^2 \sigma^2 + \dots + w_T^2 \sigma^2$$

$$= \sigma^2 \sum w_t^2$$

$$= \sigma^2 \sum \left(\frac{\ln X_t}{\sum (\ln X_t)^2} \right)^2 = \sigma^2 \sum \frac{(\ln X_t)^2}{(\sum (\ln X_t)^2)^2}$$

$$\text{Var}(\hat{\beta}) = \sigma^2 \frac{\sum (\ln X_t)^2}{(\sum (\ln X_t)^2)^2} = \boxed{\frac{\sigma^2}{\sum (\ln X_t)^2}}$$

c) another estimator $\hat{\beta} = \frac{\bar{Y}}{\bar{\ln X}}$ is suggested. Compare $\tilde{\beta}$ & $\hat{\beta}$ in terms of efficiency

$$\tilde{\beta} = \frac{\bar{Y}}{\bar{\ln X}} = \frac{\frac{\sum Y_t}{T}}{\frac{\sum \ln X_t}{T}} = \frac{\sum Y_t}{\sum \ln X_t}$$

$$\tilde{\beta} = \frac{\sum (\beta \ln X_t + u_t)}{\sum \ln X_t} = \frac{\beta \sum \ln X_t + \sum u_t}{\sum \ln X_t}$$

$$\tilde{\beta} = \beta + \frac{\sum u_t}{\sum \ln X_t}$$

$$\text{var}(\tilde{\beta}) = E(\tilde{\beta} - \beta)^2 = E\left(\frac{\sum u_t}{\sum \ln X_t}\right)^2$$

$$\text{var}(\tilde{\beta}) = E\left(\frac{(\sum u_t)^2}{(\sum \ln X_t)^2}\right) = \frac{1}{(\sum \ln X_t)^2} \cdot E(u_1 + u_2 + \dots + u_T)^2$$

$$\text{var}(\tilde{\beta}) = \frac{1}{(\sum \ln X_t)^2} E(u_1^2 + u_2^2 + \dots + u_T^2 + u_1 u_2 + u_1 u_3 + \dots + u_{T-1} u_T)$$

$$\text{var}(\tilde{\beta}) = \frac{1}{(\sum \ln X_t)^2} \underbrace{E(u_1^2)}_{\sigma^2} + \underbrace{E(u_2^2)}_{\sigma^2} + \dots + \underbrace{E(u_T^2)}_{\sigma^2} + \underbrace{E(u_1 u_2)}_0 + \dots + \underbrace{E(u_{T-1} u_T)}_0$$

$$\boxed{\text{var}(\tilde{\beta}) = \frac{T\sigma^2}{(\sum \ln X_t)^2}} \quad \text{var}(\hat{\beta}) = \frac{\sigma^2}{(\sum \ln X_t)^2}$$

$$\text{var}(\tilde{\beta}) > \text{var}(\hat{\beta})$$

$\hat{\beta}$ is more efficient than $\tilde{\beta}$

d) It is claimed that sum of the residuals of the model given above will be equal to zero. Do you agree?

$$\sum \hat{u}_t = 0 ?$$

if this is true, then

$$\sum \hat{u}_t = \sum (Y_t - \hat{Y}_t) = \sum (Y_t - \hat{\beta} \ln X_t) = 0$$

$$\hat{\beta} = \frac{\sum Y_t}{\sum \ln X_t} = \frac{\bar{Y}}{\ln \bar{X}} \quad \text{which is not true}$$

$$\hat{\beta} = \frac{\sum Y_t \ln X_t}{\sum (\ln X_t)^2} \rightarrow \underline{\sum \hat{u}_t \neq 0}$$