BST169: Course Work Project answer

sn0wfree

11/10/2016

Contents

1 BST169: Course Work Project

there are 5 questions

1.1 topic 1:

1. Consider the model:

$$y_i = \beta_0 + \beta_1 * x_{1,i} + \beta_2 * x_{2,i} + e_i \ (1)$$

What is the requirement for e_i such that the following test statstics will be valid to test H0: $\beta_1 + \beta_2 = 1$?

- $W = N * (SSR_R - SSR_U)/SSR_U$ (Wald).
- $LM = N * (SSR_R - SSR_U)/SSR_R$ (Lagrange Multiplier),
- $LR = N * ln(SSR_R/SSR_U)$ (Likelihood Ratio)

where SSR_R is the sum of squared residuals obtained from the restricted model, while SSR_R is from the unrestricted model.

1.1.1 ansewer

chi-sq distribution

$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon_i$ $y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon_i$ with $\beta_1 + \beta_2 = 1 => y_i = \beta_0 + \beta_1 x_1 + (1 - \beta_1) x_2 + \epsilon_i$ $y_i - x_2 = \beta_0 + \beta_1 (x_1 - x_2) + \epsilon_i$ in lecture 3

homoscedasity

1.2 topic 2

2. For the data set **pbp.csv**, can we use the **three test statistics** mentioned in the previous question to test H0 : $\beta_1 + \beta_2 = 1$? Why? If W and LM are not valid, how can one modify them for the test? What is your conclusion from the valid test?

bootstrap lecture:h-test or white test

```
pbp=read.csv("/Users/sn0wfree/Dropbox/PhD_1st_study/BST169_Econometrics/Crousework_Project/pbp.csv")
head(pbp)
```

```
##   X        y       x1        x2
## 1 1 4.238746 6.271882 2.8827052
## 2 2 5.807991 9.890907 0.6961033
## 3 3 3.796173 5.868733 3.4438338
## 4 4 2.692767 2.356197 0.5413317
## 5 5 2.429828 6.077795 2.8432213
## 6 6 4.967426 5.528258 2.0317105
```

```
str(pbp)
```

```
## 'data.frame':    1000 obs. of  4 variables:
##  $ X : int  1 2 3 4 5 6 7 8 9 10 ...
##  $ y : num  4.24 5.81 3.8 2.69 2.43 ...
##  $ x1: num  6.27 9.89 5.87 2.36 6.08 ...
##  $ x2: num  2.883 0.696 3.444 0.541 2.843 ...
```

1.3 topic 3

3. Generate y_i from the following model,

$$y_i = \beta_0 + \beta_1 * x_{1,i} + (1 - \beta_1) * x_{2,i} + \sqrt{x_{1,i}} * \epsilon_1 \quad (2)$$

where $x_{1,i}$ follows chi-squared distribution with **2** degrees of freedom. Generate ϵ_1 from student t distribution with 6 degrees of freedom and $x_{2,i}$~$U(0, 10)$. Check whether Wald, LR and LM in Question 1 follow chi-squared distribution by Monte Carlo.(The R command: ks.test(,'pchisq',2) can be used.) If W and LM are not valid, calculate the correct test statistics and also verify them by Monte Carlo. Please consider different sample sizes.

1.3.1 lecture Monte Carlo

from mc1.r mc2.r

lecture Monte Carlo

sample size; estimation:power of test

1.4 topic 4

Compare the size of different test statistics (frequencies of making Type 1 error) from Monte Carlo using 5% level of significance for different sample sizes. Explain the results.

1.5 topic 5

For the data set pbp.csv, suppose Equation (2) is the true model. Use proper bootstrapped errors from the true model to study whether different test statistics for H0 : $\beta_1 + \beta_2 = 1$ in the previous questions follow chi-squared distribution. Explain your results.

---

reject null, make type I error