

Nonnested-test: Encomprehensive-test, Davidson-MacKinnon-test

Kei Sakamoto

misspecification の test であることは変わらないが、前回の nested model の test でなく、log transform したほうがいいかなどの test なので、nonnested model の test。

```
load("~/計量経済学演習/R data sets for 5e/hprice1.RData")
hprice1<-data
```

two alternative models

```
model1 <- lm(price ~ lotsize + sqrft + bdrms, data=hprice1)
model2 <- lm(price ~ log(lotsize) + log(sqrft) + bdrms, data=hprice1)
```

Encomprehensive-test(manual way)

```
#modelE : comprehensive model
modelE<-lm(price ~lotsize + sqrft + bdrms + log(lotsize) + log(sqrft), data=hprice1)
library(car)

## Loading required package: carData

linearHypothesis(modelE,matchCoefs(modelE,"log"))

## Linear hypothesis test
##
## Hypothesis:
## log(lotsize) = 0
## log(sqrft) = 0
##
## Model 1: restricted model
## Model 2: price ~ lotsize + sqrft + bdrms + log(lotsize) + log(sqrft)
##
##   Res.Df    RSS Df Sum of Sq    F    Pr(>F)
## 1      84 300724
## 2      82 252340  2    48383 7.8613 0.0007526 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Encomprehensive-test(automatic way)

```
library(lmtest)
```

```
## Loading required package: zoo

##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##      as.Date, as.Date.numeric

encomptest(model1,model2, data=hprice1)

## Encompassing test
##
## Model 1: price ~ lotsize + sqrft + bdrms
## Model 2: price ~ log(lotsize) + log(sqrft) + bdrms
## Model E: price ~ lotsize + sqrft + bdrms + log(lotsize) + log(sqrft)
##      Res.Df Df       F      Pr(>F)
## M1 vs. ME      82 -2 7.8613 0.0007526 ***
## M2 vs. ME      82 -2 7.0508 0.0014943 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

M2 vs ME までやってくれるのが automatic way のいいところで、M1 vs ME を見れば当然結果は同じになる。だが一番起こってほしくないことが起きている。M1 と M2 が両方 reject されてる。結局どっちがいいのかわからん。これが Encomprehensive-model の欠陥だと言える。observation の数が 88 で若干少ないからかも。

Davidson-MacKinnon test

M2 の fitted を regressor として含むのがこの test だが、自動でやってくれる関数が見つからない(time series の J-test しかない)ので manual way のみ。

```
modelDM<-lm(price ~lotsize + sqrft + bdrms + fitted(model2), data=hprice1)
summary(modelDM)

##
## Call:
## lm(formula = price ~ lotsize + sqrft + bdrms + fitted(model2),
##      data = hprice1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -108.216  -40.458   -5.462    25.343   215.995
```

```
##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -1.194e+01  2.944e+01  -0.405    0.686
## lotsize      5.690e-04  9.986e-04   0.570    0.570
## sqrft        5.354e-02  3.803e-02   1.408    0.163
## bdrms         3.592e+00  1.033e+01   0.348    0.729
## fitted(model2) 6.123e-01  3.159e-01   1.938    0.056 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 58.88 on 83 degrees of freedom
## Multiple R-squared:  0.6865, Adjusted R-squared:  0.6714
## F-statistic: 45.45 on 4 and 83 DF,  p-value: < 2.2e-16
```

p-value が 0.056 なので微妙……。自信を持って model1 が正しいとは言いきれないので色々全部推定結果提出するのが吉。