

# [Example8-4,5] Heteroskedasticity Test

Kei Sakamoto

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

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

```
##
## Call:
## lm(formula = price ~ lotsize + sqrft + bdrms, data = hprice1)
##
## Coefficients:
## (Intercept)  lotsize    sqrft    bdrms
## -21.770308   0.002068   0.122778   13.852522
```

## Example 8-4 Breusch Pagan test (BP test)

$u$ はunobservableなので $\hat{u}$ (residual)で置き換え、residualの2乗は $E(\hat{u})=0$ から $\text{Var}(\hat{u})$ なので、要するにregressorが $\hat{u}$ のvarianceに対して説明力を持つかどうかtestすればいい。具体的にはauxiliary regressionのF-testをする。

Manual regression of squared residuals(auxiliary regression のsummaryで勝手にF-testまでしてくれる。)

```
summary(lm( resid(reg)^2 ~ lotsize+sqrft+bdrms, data=hprice1))
```

```
##
## Call:
## lm(formula = resid(reg)^2 ~ lotsize + sqrft + bdrms, data = hprice1)
##
## Residuals:
##   Min    1Q  Median    3Q   Max
## -9044 -2212 -1256  -97  42582
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -5.523e+03  3.259e+03  -1.694  0.09390 .
## lotsize      2.015e-01  7.101e-02   2.838  0.00569 **
## sqrft       1.691e+00  1.464e+00   1.155  0.25128
## bdrms       1.042e+03  9.964e+02   1.046  0.29877
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6617 on 84 degrees of freedom
## Multiple R-squared:  0.1601, Adjusted R-squared:  0.1301
## F-statistic: 5.339 on 3 and 84 DF, p-value: 0.002048
```

# Automatic BP test

```
library(lmtest) #lmtest割と万能だけど、LM versionなので若干結果違う
```

```
## Loading required package: zoo
```

```
##  
## Attaching package: 'zoo'
```

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

```
bptest(reg)
```

```
##  
## studentized Breusch-Pagan test  
##  
## data: reg  
## BP = 14.092, df = 3, p-value = 0.002782
```

manualでもautomaticでも1% significantなことは変わらない。  
つまりregressorが $\hat{u}$ のvarianceに対して説明力を持つので  
heteroskedasticだと言えそう。

## Example8-5 White test

modelをlog transformしてもう一回BPtestして、今度は $\hat{u}$ のvarianceはregressorsに対してnonlinearを許しているWhite testと比較する。df(自由度)下がってpower(検出力)が下がるのを避けるためにyのfittedをregressorとして入れてmodelを定義する。nonlinear terms はyのfittedに含まれていることを期待。(cross product とか馬鹿正直に全部入れると自由度鬼下がってしまう。)

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

```
##  
## Call:  
## lm(formula = log(price) ~ log(lotsize) + log(sqrft) + bdrms,  
##   data = hprice1)  
##  
## Coefficients:  
## (Intercept) log(lotsize) log(sqrft) bdrms  
## -1.29704 0.16797 0.70023 0.03696
```

# BP test

```
bptest(reg)
```

```
##  
## studentized Breusch-Pagan test  
##  
## data: reg  
## BP = 4.2232, df = 3, p-value = 0.2383
```

log transformすると説明力下がっちゃう。(xが大きいとlog transformがxの良い近似になっていない可能性。)

# White test

```
bptest(reg, ~ fitted(reg) + I(fitted(reg)^2) )
```

```
##  
## studentized Breusch-Pagan test  
##  
## data: reg  
## BP = 3.4473, df = 2, p-value = 0.1784
```

同じくlog transformのせいなのかもしれないが **nonlinear terms** 含めて**coef all 0** 説は**significance leve 10%**でも**reject** できない。 **heteroskedastic**だとは言い切れなくなってしまう。