Heteroskedasticity

Statistics and Econometrics

Jiahua Wu

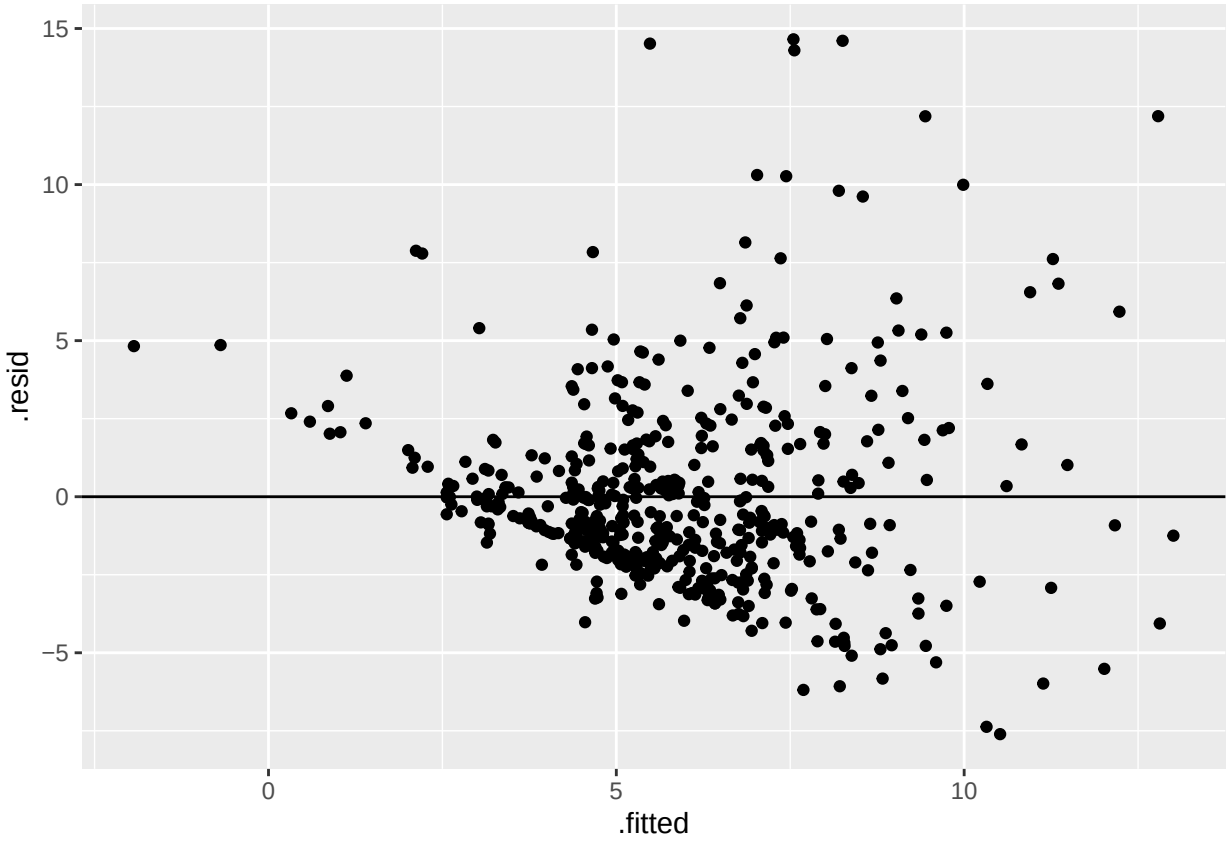
Tests for heteroskedasticity

In the class, we discussed two formal tests for heteroskedasticity, i.e., Breusch-Pagan test and White test. We use the example on page 13 to discuss how to run the tests in R.

Before we run formal tests, it is proven to be useful to check out the residual plot, and get some intuitive idea whether the variance in the residual is constant or not.

```
load("wage1.RData")
wage.m1 <- lm(wage ~ educ + exper + tenure, data)

# residual plot
ggplot(wage.m1, aes(.fitted, .resid)) + geom_point() + geom_hline(yintercept = 0)
```



We plot residuals against fitted values using *ggplot*. Judging from the plot, there exhibits some evidence of heteroskedasticity. The variation in residuals is small with small fitted values, and tends to get larger as fitted value increases.

Next we run the two formal tests. The R command for the tests is *bptest*() in the *lmtest* package. By default, *bptest* regresses $\hat{u}^2$ on all the independent variables, which is based on Breusch-Pagan test. If we want to use White test, we need to specify the independent variables to be the fitted value and squared of the fitted value

in the second step. The implementation of *bptest*() is based on an alternative test, rather than the F test as shown on the slides. So we only need to care about the p-value in the output. A smaller p-value provides stronger evidence for heteroskedasticity.

```
library(lmtest)
# BP test
bptest(wage.m1)
```

```
##
##  studentized Breusch-Pagan test
##
## data:  wage.m1
## BP = 43.096, df = 3, p-value = 2.349e-09
```

```
# White test
fitted.wage <- wage.m1$fitted.values
bptest(wage.m1, ~ fitted.wage + I(fitted.wage^2))
```

```
##
##  studentized Breusch-Pagan test
##
## data:  wage.m1
## BP = 51.777, df = 2, p-value = 5.713e-12
```

Heteroskedasticity-robust inference

We can derive robust standard errors using *sandwich* package in R. There are multiple ways to adjust standard errors to make it robust against heteroskedasticity. For details, please check out "sandwich.pdf" posted on the Hub (btw, it is not examinable). We can run robust t test and robust F test as follows.

```
library(sandwich)

# calculate robust variance and covariance matrix
vcov.robust <- vcovHC(wage.m1, "HC1")

# t test
coeftest(wage.m1, vcov = vcov.robust)
```

```
##
## t test of coefficients:
##
##              Estimate Std. Error t value  Pr(>|t|)
## (Intercept) -2.872735   0.807415 -3.5579 0.0004078 ***
## educ         0.598965   0.061014  9.8169 < 2.2e-16 ***
## exper        0.022340   0.010555  2.1165 0.0347731 *
## tenure       0.169269   0.029278  5.7814 1.277e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# F test
linearHypothesis(wage.m1, "educ - exper = 0", white.adjust = "hc1")
```

```
## Linear hypothesis test
##
## Hypothesis:
## educ - exper = 0
```

```
##
## Model 1: restricted model
## Model 2: wage ~ educ + exper + tenure
##
## Note: Coefficient covariance matrix supplied.
##
##   Res.Df Df      F    Pr(>F)
## 1    523
## 2    522  1 98.394 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```