R\_SDERR

这是关于金融类文章如何计算不同回归标准误的R程序示例，主要参考： <http://www.r-bloggers.com/fama-macbeth-and-cluster-robust-by-firm-and-time-standard-errors-in-r/> 需要提前加载的R包：plm,lmtest，sandwich 参考文献： Thompson (2011, JFE) and Petersen (2008, WP)

首先导入Petersen的测试数据： 原始数据可从"<http://www.kellogg.northwestern.edu/faculty/petersen/htm/papers/se/test_data.dta>" 下载，为Stata的dta格式，可以用foreign包里的read.dta直接读取。

test <- read.csv("d:/R\_STDErr.csv") %>%  
 tbl\_df %>%  
 print

## Source: local data frame [5,000 x 5]  
##   
## X firmid year x y  
## 1 1 1 1 -1.113973 2.2515  
## 2 2 1 2 -0.080854 1.2423  
## 3 3 1 3 -0.237607 -1.4264  
## 4 4 1 4 -0.152486 -1.1094  
## 5 5 1 5 -0.001426 0.9147  
## 6 6 1 6 -1.212737 -1.4247  
## 7 7 1 7 -0.127273 0.7589  
## 8 8 1 8 -1.433539 0.9297  
## 9 9 1 9 -0.242196 1.0565  
## 10 10 1 10 0.460922 3.3084  
## .. .. ... ... ... ...

进行Pooling OLS回归以及Fama-Macbeth回归，因为调整标准误不涉及beta估计量，所以只需调整标准误即可

pols <- plm(y~x,test,model="pooling",index=c("firmid","year"))  
fmr <- pmg(y~x,test,index=c("year","firmid"))

定义函数来估计White异方差稳健标准误和双向聚类标准误的函数，,两个方向聚类之和减去White

##Double-clustering formula (Thompson, 2011)  
vcovDC <- function(x, ...){  
 vcovHC(x, cluster="group", ...) + vcovHC(x, cluster="time", ...) -  
 vcovHC(x, method="white1", ...)  
}

估计普通OLS标准误

coeftest(pols)

##   
## t test of coefficients:  
##   
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 0.0297 0.0284 1.05 0.3  
## x 1.0348 0.0286 36.20 <2e-16

估计White标准误

coeftest(pols, vcov=function(x) vcovHC(x, method="white1", type="HC1"))

##   
## t test of coefficients:  
##   
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 0.0297 0.0284 1.05 0.3  
## x 1.0348 0.0284 36.44 <2e-16

估计个体聚类标准误

coeftest(pols, vcov=function(x) vcovHC(x, cluster="group", type="HC1"))

##   
## t test of coefficients:  
##   
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 0.0297 0.0670 0.44 0.66  
## x 1.0348 0.0506 20.47 <2e-16

估计时间聚类标准误

coeftest(pols, vcov=function(x) vcovHC(x, cluster="time", type="HC1"))

##   
## t test of coefficients:  
##   
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 0.0297 0.0222 1.34 0.18  
## x 1.0348 0.0317 32.67 <2e-16

估计双向（时间和个体同时）聚类标准误

coeftest(pols, vcov=function(x) vcovDC(x, type="HC1"))

##   
## t test of coefficients:  
##   
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 0.0297 0.0646 0.46 0.65  
## x 1.0348 0.0525 19.72 <2e-16

注释： As Giovanni interestingly pointed out to me (in a privately circulated draft paper),it seems that the Fama-MacBeth estimator is nothing more than what econometricians call the Mean Groups estimator, and 'plm' can readily estimate this. You only need to swap the 'group' and 'time' indices.

coeftest(fmr)

##   
## t test of coefficients:  
##   
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 0.0313 0.0234 1.34 0.18  
## x 1.0356 0.0333 31.06 <2e-16