回归分析

Qiufei

1/21/2017

本次分析任务是对CPI和央行资产负债表增速之间的关系做回归分析。

## 首先是下载数据

library(Quandl)  
# 央行资产负债表增速的quandl code代码是PBCHINA/REP\_04.11  
# 中国cpi增速数据的quandl code代码是NBSC/A01030101\_M  
codes = c("PBCHINA/REP\_04.11","NBSC/A01030101\_M")  
  
reg\_data = Quandl(codes)

## 其次修改下数据集的名字

我们下载的数据已经保存在数据集reg\_data中，该数据集有三列，第一列是时间数据，第二列是央行资产负债表规模数据，第三列是cpi数据。

第一列的列名Date，比较好认，但是第二列和第三列的列名太长，也不好认。所以我们修改一下列名。

将第一列的列名修改为date，将第二列的列名修改为pbc\_asset，将第三列的列名修改为cpi。

colnames(reg\_data) = c('date','pbc\_asset','cpi')

## 计算央行资产负债表规模的增速

因为15年有些数据没有，所以去掉2015年以前的数据。

又因为我们要计算资产负债表的增速百分比，和CPI的增速百分比，所以要做些转换。

# 为了使用filter函数  
library(dplyr)  
  
# 为了使用ROC函数  
library(TTR)  
  
## drop those NA observations  
reg\_data = filter(reg\_data,date>'2005-12-31')  
## get growth rate  
reg\_data[,2] = ROC(reg\_data[,2],type = 'discrete')\*100  
reg\_data[,3] = reg\_data[,3] - 100

## 最后是对数据进行回归分析

在回归分析中，将cpi作为因变量（Y），将pbc\_asset作为自变量（X）。

my\_model = lm(cpi~pbc\_asset,reg\_data)  
  
summary(my\_model)

##   
## Call:  
## lm(formula = cpi ~ pbc\_asset, data = reg\_data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.31918 -0.35927 0.01544 0.28228 2.19536   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.13665 0.06151 2.222 0.02826 \*   
## pbc\_asset 0.12991 0.03865 3.361 0.00105 \*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.55 on 116 degrees of freedom  
## (14 observations deleted due to missingness)  
## Multiple R-squared: 0.08875, Adjusted R-squared: 0.0809   
## F-statistic: 11.3 on 1 and 116 DF, p-value: 0.001051