

Statement and Minutes Forecast

William D. MacMillan

11/8/2016

This code produces point estimates and residual scores for Fed statements and minutes. This first chunk is used to set the cutoff date for the forecast (which predicts 7 days out, i.e. in accordance with the blackout), and set whether you are predicting statements or minutes. You must set your own username and password to utilize the `pRattle` package.

```
library(pRattle)
library(lubridate)
library(forecast)
library(xts)

your_username = "your-username"
your_password = "your-password"
cutoffdate<-"2016-10-24"
statement<-T

df<-get_scores('frc', agg.level='daily', email=your_username, pwd=your_password)
```

The following chunk performs a series of data munging tasks in order to construct the data in the proper time series format for estimation.

```
dates<-as.Date(as.Date("1990-01-01"):as.Date(cutoffdate), origin="1970-01-01")

empty<-xts(,order.by=dates)

df<-merge(empty,df,join="left")
df<-na.locf(df)

df2<-get_scores('frc', type='raw', email=your_username, pwd=your_password)
df_orig<-df2
df2<-subset(df2, speaker %in% c("FOMC Statement", "Minutes FRC"))
df2$statement[df2$speaker=="FOMC Statement"]<-1
df2$minutes[df2$speaker=="Minutes FRC"]<-1

df2$date<-as.Date(df2$date)

df2<-xts(df2[,c('statement', 'minutes')], order.by=df2$date)

ff<-merge(df, df2, join='left')

names(ff)[1]<-'score'
ff<-na.fill(ff, 0)
ff<-ff["19980101/"]
```

After having assembeled the data correctly, we now will fit the model, and produce the forecast.

```

fit<-auto.arima(ff$score, xreg=ff[,c("statement", "minutes")])

if(statement==T){
  xdata<-data.frame(statement=c(0,0,0,0,0,0,1), minutes=c(0,0,0,0,0,0,0))
} else {
  xdata<-data.frame(statement=c(0,0,0,0,0,0,0), minutes=c(0,0,0,0,0,0,1))
}

sfcore<-forecast(fit, h=7, xreg=xdata)
sfcore

```

```

##      Point Forecast      Lo 80      Hi 80      Lo 95      Hi 95
## 6873      0.43127543 -0.3289633  1.191514 -0.7314094  1.593960
## 6874      0.37077403 -0.5356624  1.277210 -1.0155008  1.757049
## 6875      0.32227621 -0.6554525  1.300005 -1.1730309  1.817583
## 6876      0.28963586 -0.7209980  1.300270 -1.2559953  1.835267
## 6877      0.26766803 -0.7590940  1.294430 -1.3026289  1.837965
## 6878      0.25288308 -0.7822172  1.287983 -1.3301662  1.835932
## 6879     -0.02870813 -1.0683777  1.010961 -1.6187456  1.561329

```

The last row is the point estimate of the statement or minute. We'll continue with the calculation of the residual score, which can be highly useful for a more contextualized interpretation of the score.

```

formean<-subset(df_orig, date<=ymd(cutoffdate)&date>=ymd(cutoffdate)-months(12))
if(statement==T){
  formean<-subset(formean, speaker=='FOMC Statement')
} else {
  formean<-subset(formean, speaker=='Minutes FRC')
}

baseline<-mean(formean$score)

print("Point estimate")

```

```
## [1] "Point estimate"
```

```
sfc$mean[7]
```

```
## [1] -0.02870813
```

```
print("Residual estimate")
```

```
## [1] "Residual estimate"
```

```
sfc$mean[7] - baseline
```

```
## [1] 0.5841684
```