Homework 4

Joe Diaz

9/17/2019

Question 1

Describe a situation or problem from your job, everyday life, current events, etc., for which exponential smoothing would be appropriate. What data would you need? Would you expect the value of alpha (the first smoothing parameter) to be closer to 0 or 1, and why?

An e-commerce business might be interested in doing an exponential smoothing model on its weekly sales. Depending on the items being sold, there could be seasonal effects and trend that would show how certain effects influence sales. Alpha would depend in this case. If the e-commerce business sells only one thing, I'd expect the value to be closer to 0 as there is probably high variability on sales for that 1 item. If the business is more diversified, the alpha should be set closer to 1 as you'd expect variability of sales to decrease with a more general offering of goods.

Question 2

Using the 20 years of daily high temperature data for Atlanta (July through October) from Homework 2 Question 5, build and use an exponential smoothing model to help make a judgment of whether the unofficial end of summer has gotten later over the 20 years. (Part of the point of this assignment is for you to think about how you might use exponential smoothing to answer this question.) Note: in R, you can use either HoltWinters (simpler to use) or the smooth package's es function (harder to use, but more general). If you use es, the Holt-Winters model uses model="AAM" in the function call (the first and second constants are used "A"dditively, and the third (seasonality) is used "M"ultiplicatively; the documentation doesn't make that clear

```
temps <- read.table("temps.txt", header = TRUE)


df <- ts( as.vector(unlist(temps[,2:21])),start=1996,end=2015,frequency=123)


model1 <- HoltWinters(df,seasonal="multiplicative")
```

To explain what is happening with the code below…

first we turn the output of the HoltWinters function to a matrix (specifically, the seasonal component)

Then, we run cumsum on each of the periods to detect at which point of the year is the change detected.

Then we iterate over each item of the output of our cumsum to determine if it falls under each threshold.

We then inspect that to determine if the change is occuring at a different time each year. If it is occuring later, then we can say that summer is lasting longer and longer.

```
exp_smth_df <- matrix(model1$fitted[,4],ncol = 123)
```

```
## Warning in matrix(model1$fitted[, 4], ncol = 123): data length [2215] is
## not a sub-multiple or multiple of the number of rows [19]
```

```
cusum <- function(xt,prev_st,mean,C) {
```

```r
  return(max(c(0,prev_st+(mean-xt-C))))

}

df_mean <- mean(exp_smth_df[1:31])

C <- sd(exp_smth_df[1:31])/2

thresh <- sd(exp_smth_df[1:31])*5

temps <- matrix(nrow = 19,ncol = 123)
temps[1,] <- 0


for( c in 1:ncol(exp_smth_df)){

  for(i in 2:nrow(exp_smth_df)){
  temps[i,c] <- cusum(xt = exp_smth_df[i,c]
                       , prev_st = temps[i-1,c]
                       , mean = df_mean
                       , C = C
                       )
}


}



is_above_thresh <- function(x,threshold=thresh){

  if(x<threshold){
    return(1)
  } else{
    return(0)
  }

}

is_thresh <-  matrix(nrow = 19,ncol = 123)

for(c in 1:ncol(is_thresh)){
  for(r in 1:nrow(is_thresh)){
    is_thresh[r,c]<-is_above_thresh(temps[r,c],thresh)


  }

}

colSums(is_thresh)
```

```
##   [1] 19 19 19 19  5  2  2 19 11 18 12  4  4 19 19 19 19  4  2  6 19 13 19
##  [24] 11  3  3 19 19 19 16  4  1 13 19 15 19 10  3  3 19 18 19 18  4  2 15
```

```
##  [47] 19 19 18 10  2  3 19 18 19 16  4  1 18 19 17 16  9  3  2 19 18 19 16
##  [70]  4  2 19 19 16 14  9  3  2 19 17 19 15  4  2 19 19 14 16  7  2  2 19
##  [93] 18 19 14  4  2 19 19 15 13  7  3  2 19 17 19 13  4  2 19 19 14 13  6
## [116]  3  2 19 12 19 14  6  4
```

To explain the output above, I've set the instances of no change to be 1 and change to be 0. As a result, the sum of each column would tell us at which point was there change detected.

Judging from the output of the script, it seems there is no discernable pattern and more importantly, summer doesn't seem to be lasting longer and longer.