ISYE 6051 : Homework 4
2/17/2021

Table of Contents

7.1 Determining current situations for exponential smoothing

7.2 Using Exponential Smoothing to determine the unofficial end of summer

<u>7.1 Current Situations where exponential smoothing is relevant</u>

Scheduling the administration of vaccines to protect the population against series illness and death related to COVID-19 has proven to be a challenge for larger municipalities. The raw numbers of eligible recipients, the fluctuating availability in the allocation per states and the ability of individuals to schedule vaccines at multiple locations makes exponential smoothing a viable model to determine the percentage of shot 1 versus shot 2 recipients to schedule.

States request allocations of shots on different days of the week – Thursdays for shot 1 and Sundays for shot 2. The obvious method to scheduling would be to schedule the number of participants based on the allocation requested and approved for each phase of the shot.

However, widely publicized issues with vaccine sites either not having enough recipients (and having to beg people to come take shots who may not be in the eligibility group so that the vaccine doesn't go to waste) or having to turn away eligible recipients due to overscheduling proves that the obvious method isn't a strong model. Two of the most challenging issues are:

1) ***Eligible receipients don't show up for the shot 1 appointment for any number of reasons and***
2) ***Shot 2 eligible receipients don't show up due to the fact that in larger municipalities they can register in more than one county***

and will go to the first one that has an available appointment after having signed up in multiple locations.

The vaccine is date and temperature sensitive; waste is of paramount concern and both shots are exactly the same in formulation. This is where exponential smoothing utilizing forecasting and trends would be of benefit. These are short term models where the most recent data would be of value in determining the potentially best allocation of vaccines on a given day. In this case the α value would be closer to 1 based on the high possibility of randomness based on several factors such as location, day of the week, weather conditions and any other number of variances.

7.2 Using Exponential Smoothing

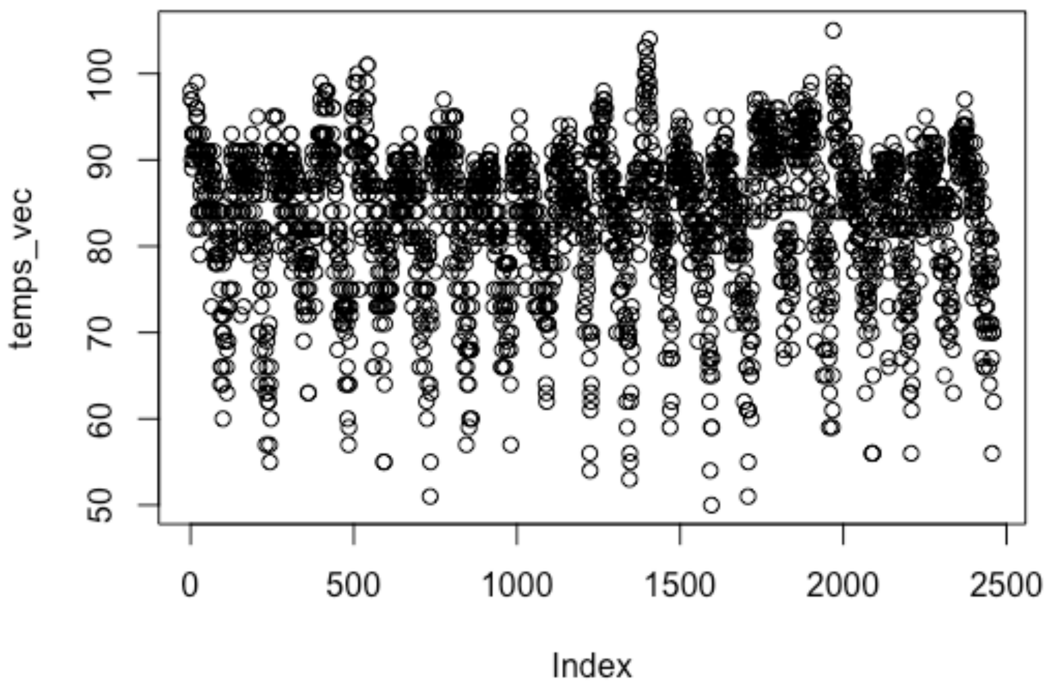
Using the temps.txt data, the assignment is to develop a model to determine the unofficial end of summer with exponential smoothing. Exponential smoothing is valuable when there is "noise" or variations in the pattern of data. In order to determine if the temps.txt contains variations, I started with creation of a data vector that was plotted for visualization.

```
# Set the working directory
rm(list = ls())
setwd("~/Documents/ISYE6501 Intro to Analytics
Modeling/FA_SP_hw4")
# Read the data in
matrix1 <- read.table("temps.txt", header = TRUE)
head(matrix1)
tail(matrix1)

temps_vec <- as.vector(unlist(matrix1[,2:21]))
temps_vec
plot(temps_vec)
```

Viewing the head and tail of ingested temps.txt data

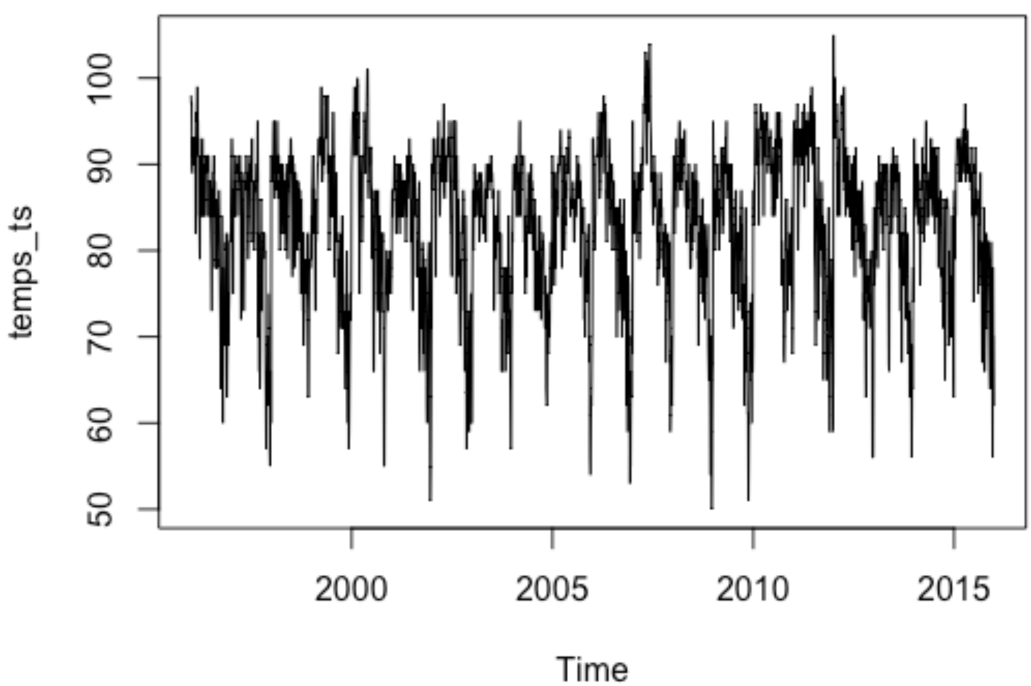
```
> head(matrix1)
   DAY X1996 X1997 X1998 X1999 X2000 X2001 X2002 X2003
X2004 X2005 X2006 X2007
1 1-Jul    98    86    91    84    89    84    90    73    82    91    93    95
2 2-Jul    97    90    88    82    91    87    90    81    81    89    93    85
3 3-Jul    97    93    91    87    93    87    87    87    86    86    93    82
4 4-Jul    90    91    91    88    95    84    89    86    88    86    91    86
5 5-Jul    89    84    91    90    96    86    93    80    90    89    90    88
6 6-Jul    93    84    89    91    96    87    93    84    90    82    81    87
  X2008 X2009 X2010 X2011 X2012 X2013 X2014 X2015
1    85    95    87    92   105    82    90    85
2    87    90    84    94    93    85    93    87
3    91    89    83    95    99    76    87    79
4    90    91    85    92    98    77    84    85
5    88    80    88    90   100    83    86    84

tail(matrix1)
      DAY X1996 X1997 X1998 X1999 X2000 X2001 X2002 X2003
X2004 X2005 X2006
118 26-Oct    75    71    79    69    75    64    68    68    79    61    62
119 27-Oct    75    57    79    75    78    51    69    64    81    63    66
120 28-Oct    81    55    79    73    80    55    75    57    78    62    63
121 29-Oct    82    64    78    72    75    63    75    70    75    64    72
122 30-Oct    82    66    82    75    77    72    68    77    78    69    73
123 31-Oct    81    60    79    75    78    71    60    75    82    70    68
   X2007 X2008 X2009 X2010 X2011 X2012 X2013 X2014 X2015
118    68    70    65    85    77    80    61    84    67
119    67    59    60    76    79    70    69    84    56
120    70    50    71    74    74    56    64    77    78
121    62    59    75    68    59    56    75    73    70
122    67    65    66    71    61    56    78    68    70
123    71    67    69    75    65    65    74    63    62
```

Plot of vector data to visualize noise.



The vector data is converted to time series data and plotted for visualization of noise. Seasonality in temperature changes (both up and down across the years) appear.

```
# Start at the year 1996 and use the frequency equal to 123 which
represents the number of days Jul 1 - Oct 31
temps_ts <- ts(temps_vec, start = 1996, frequency = 123)
temps_ts
plot(temps_ts)
```



There are 3 types of exponential smoothing models that can be used. However, each one should be used based on data complexity (weighted average, trending and seasonality).
Single exponential smoothing looks at weighted data. It assumes that there is not much variance. Data that is closer to the date of the observation is weighted heavier.

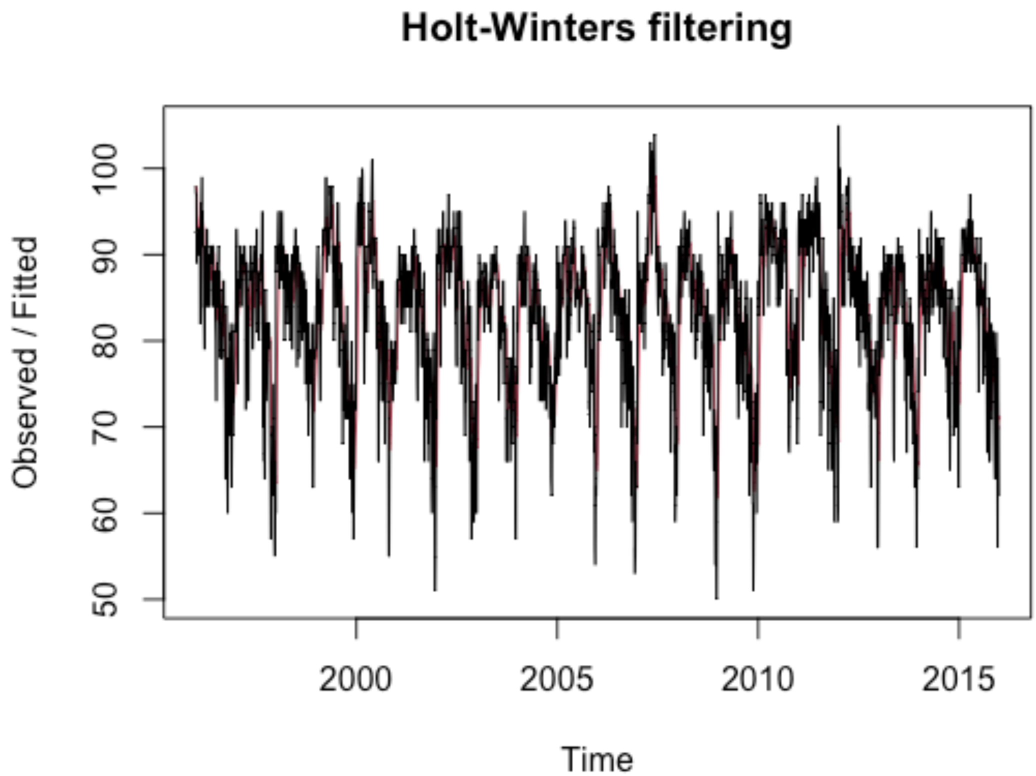
Double exponential smoothing assumes that there are trends in data. In the case of the temps.txt data a possible trend could be longer ends of summer.
Triple exponential smoothing looks at data where there is seasonality and trends can be extracted.

Single Exponential smoothing using the HoltWinters function. HoltWinters can be used for simple, trend and seasonality modeling. I used alpha = 0.2 as the common default.

```
#Single exponential smoothing view. Use of HoltWinters function with
beta and gamme = FALSE
#for weighted average
install.packages("forecast")
library(forecast)

singleexp <- HoltWinters(temps_ts,
                        alpha = 0.2, #common starting point for alpha
                        beta = FALSE,
                        gamma = FALSE)
singleexp
summary(singleexp)
plot(singleexp)
```

```
Holt-Winters exponential smoothing without trend and without seasonal
component.

Call:
HoltWinters(x = temps_ts, alpha = 0.2, beta = FALSE, gamma =
FALSE)

Smoothing parameters:
 alpha: 0.2
 beta : FALSE
```

```
 gamma: FALSE

Coefficients:
    [,1]
a 69.0748
> summary(singleexp)
             Length Class  Mode
fitted       4918   mts    numeric
x            2460   ts     numeric
alpha           1   -none- numeric
beta            1   -none- logical
gamma           1   -none- logical
coefficients    1   -none- numeric
seasonal        1   -none- character
SSE             1   -none- numeric
call            5   -none- call
```

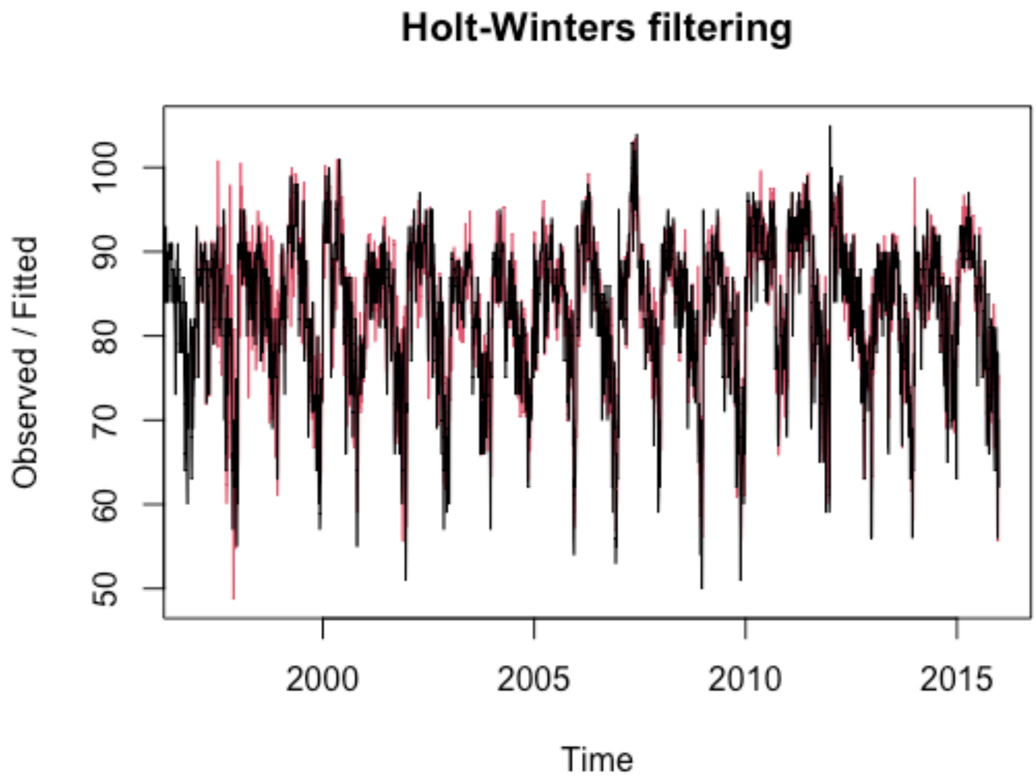


In order to determine if this visual seasonality and possibly by extracting trends that it can be demonstrated that summers are unofficially becoming longe, HoltWinters using trend and seasonality will be used (triple exponential smoothing). Multiplicative or additive modeling can be done in the triple exponential smoothing. Since seasonality is consistent across all of the years (in other words there are seasonal changes each year that is being measured), the additive form will most likely be the best to use.

```
set.seed(1)
tripleexp <- HoltWinters(temps_ts,
                         seasonal = 'additive')
tripleexp
summary(tripleexp)

plot(tripleexp)
```

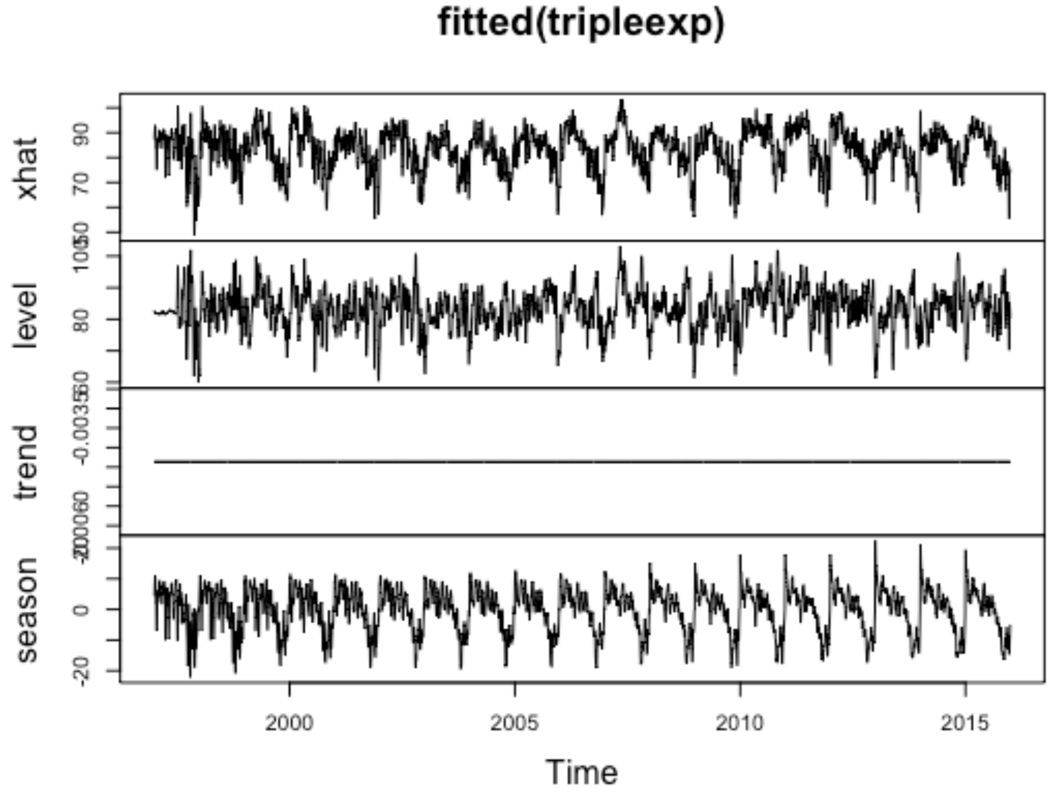
```
cat('\tThe Alpha for the single exponential method is', tripleexp$alpha, '\n')
cat('\tThe Trend for the single exponential method is', tripleexp$beta, '\n')
cat('\tThe Seasonality for the single exponential method is', tripleexp$gamma, '\n')
```

```
summary(tripleexp)
             Length Class  Mode
fitted       9348   mts    numeric
x            2460   ts     numeric
alpha           1   -none- numeric
beta            1   -none- numeric
gamma           1   -none- numeric
coefficients  125   -none- numeric
seasonal        1   -none- character
SSE             1   -none- numeric
call            3   -none- call
> plot(tripleexp)
> cat('\tThe Alpha for the triple exponential method is', tripleexp$alpha, '\n')
        The Alpha for the triple exponential method is 0.6610618
> cat('\tThe Trend for the triple exponential method is', tripleexp$beta, '\n')
        The Trend for the triple exponential method is 0
> cat('\tThe Seasonality for the triple exponential method is', tripleexp$gamma, '\n')
        The Seasonality for the triple exponential method is 0.6248076
```



The fitted data is in red. Note that the Trend metric is 0. This would indicate that there are no trends detected in this dataset. The alpha, level, trend and seasonality views are plotted below for visualization. Xhat represents the raw fitted data with the decompositions of level, trend and seasonality visualized below.

```
par(mfrow=c(1,2))
> plot(fitted(tripleexp))
```



Again, trend is visually 0 therefore the fitted data is not showing trends that can easily proven that summers are lasting longer thru the 20 years evaluated.

Row	***Xhat***	***Level***	***Trend***	***Seasonality***
[1,]	87.17619	82.87739	-0.004362918	4.303159
[2,]	90.32925	82.09550	-0.004362918	8.238119
[3,]	92.96089	81.87348	-0.004362918	11.091777
[4,]	90.93360	81.89497	-0.004362918	9.042997
[5,]	83.99752	81.93450	-0.004362918	2.067387
[6,]	84.04358	81.93177	-0.004362918	2.116168
[7,]	75.06732	81.89860	-0.004362918	-6.826922
[8,]	87.04284	81.84974	-0.004362918	5.197468
[9,]	84.01829	81.81705	-0.004362918	2.205599
[10,]	87.05875	81.80060	-0.004362918	5.262509
[11,]	84.04807	81.75740	-0.004362918	2.295029

[12,]	88.04445	81.72126	-0.004362918	6.327550

No change in the length of seasons is seen across the visual data. Using the triple exponential smoothing model there does not appear to be data to support a statement the unofficial summer is getting longer over the 20 year data sample used.