ISA 444: Business Forecasting

07 - Nonseasonal Smoothing (Cont.)

Fadel M. Megahed

Associate Professor
Department of Information Systems and Analytics
Farmer School of Business
Miami University
Email: fmegahed@miamioh.edu
Office Hours: Click here to schedule an appointment

Spring 2021

Outline

What we Covered Last Week

Main Learning Outcomes

- ☑ Apply transformations to a time series.
- ☑ Apply and interpret measures of forecast accuracy.
- ☑ Interpret prediction intervals for a simple forecast.
- ☑ Describe the benefits and drawbacks of judgmental and quantitative forecasing methods.
- ☑ Explain the difference between causal and extrapolative forecasting.
- ☑ Describe and apply smoothing/forecasting with a cumulative average.
- ☑ Describe and apply forecasting with a moving average.

Recap: Guidelines for Transforming Time-Series Data

Transformation Methods for Time Series

- **Stabilize the Mean**
  - **Differencing**
    - Eliminate/Reduce
      - Seasonality
      - Trend
    - Compute New Cases[1]
- **Stabilize the Variance**
  - **Power Transformations**
    - Log
    - Square Root
- **∼Stabilize both**
  - **Growth Rates**
  - **First diff of Log**[2]
    - Exponential Growth TS
      - ln to get a TS with a linear trend
      - Then diff to get a stationary series
- **Rescale**
  - **$(0-1)$ Scaling**[3]
  - **$z-$transform**[4]

A classification of common transformation approaches for time series data.[5]

[1] The COVID19 package returns cumulative cases, i.e. a first difference $\longrightarrow$ new confirmed cases.

[2] First difference of LOG $\approxeq$ percentage change. This is almost exact if the percentage change is small, but for larger percentage changes, it may differ greatly (see here for more details).

[3] Rescaling of the data from the original range so that all values are within the range of 0 and 1. Mathematically, speaking this can be achieved by calculating $y_t = \frac{x_t - \min}{\max - \min}$.

[4] One can normalize a time-series by $z_t = \frac{x_t - \mu}{\sigma}$.

[5] My (incomplete) attempt to provide you with a taxonomy for time series data transformations.

Recap: A Note on Interpreting HW 05 Q2 Result

Last week, I had two emails asking me about making sense of the Q2 result in Assignment 05. Recall the **question:** *Let us assume that we wanted to have a fairer comparison between the three countries. Therefore, you will scale the number of new cases by population. Report the scaled value for the US.*

```
pacman::p_load(tidyverse, magrittr, COVID19) # needed packages
covid = covid19(country = c('EGY', 'IND', 'USA'), start = '2020-03-01', end =
covid %>% select(id, date, confirmed, population) %>%
  mutate(newCases = confirmed - lag(confirmed),
         newCasesByPop = newCases/population) %>%
  filter(date == '2021-01-21')
```

id	date	confirmed	population	newCases	newCasesByPop
EGY	2021-01-21	159,715	98,423,595	752	0.0000
IND	2021-01-21	10,626,147	1,352,617,328	14,495	0.0000
USA	2021-01-21	24,656,646	326,687,501	193,055	0.0006

Recap: Interpreting Measures of Forecast Accuracy

In class, we have categorized measures of forecast accuracy into measures reflecting:

a. "average" forecast performance (e.g., mean error and mean percent error);
b. "variability" in forecast performance (e.g., AE, SE, MAE, and RMSE); and
c. "relative" forecast error (e.g., MAPE).

symbol	date	adjusted	naiveFC	forecastError
DOGE-USD	2021-02-08	0.079		
DOGE-USD	2021-02-09	0.070	0.079	-0.009
DOGE-USD	2021-02-10	0.073	0.070	0.003
DOGE-USD	2021-02-11	0.070	0.073	-0.003
DOGE-USD	2021-02-12	0.070	0.070	0.000
DOGE-USD	2021-02-13	0.066	0.070	-0.004
DOGE-USD	2021-02-14	0.063	0.066	-0.004

Based on the naive forecast, we can compute: (How would you interpret these results?)

ME	RMSE	MAE	MPE	MAPE
-0.0027	0.0045	0.0038	-4.06	5.54

Recap: Prediction Intervals

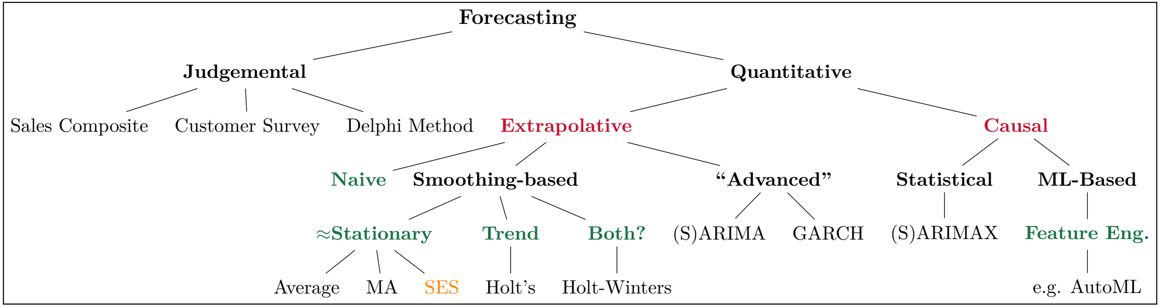
- **Point Forecasts:** future observations for which we report a single forecast observation.

- **Interval Forecast:** a range of values that are reported to forecast an outcome.

If we assume the forecast errors follow a Normal Distribution, an approximate $100(1-\alpha)$ prediction interval can be computed as follows: $\hat{F}_t \pm Z * RMSE$, where:

- $\hat{F}_t$ forecast at time t.
- The RMSE can be used as an estimate of the standard deviation of the forecast errors.
- Z is the quantile corresponding to $100(1-\frac{\alpha}{2})$ (see Section 4 of our interactive guide for more details)

Recap: A 10,000 Foot View of Forecasting Methods



A 10,000 foot view of forecasting techniques[6]

[6] An (incomplete) classification of forecasting techniques. Note that these focus on univariate time-series. Hence, they exclude popular approaches used in multivariate time series forecasting.

Learning Objectives for Today's Class

Main Learning Outcomes

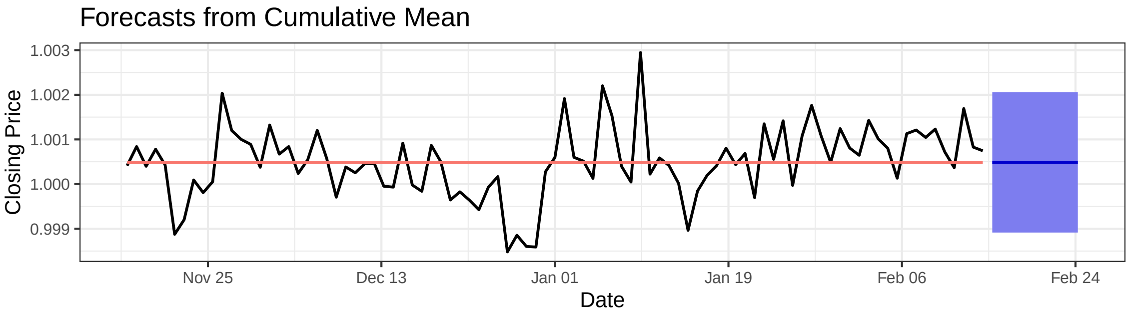
- Recognize time series that are appropriate for simple exponential smoothing (SES).
- Use SES to smooth past observations of a time series.
- Use SES to forecast future observations of a time series.
- Compare the forecasting performance of SES to other suitable techniques (i.e., methods that require similar assumptions).

Outline

Smoothing with the Overall Average

In this live coding session, let us recreate this chart building on the following code.

```
usdt = tidyquant::tq_get('USDT-USD', from = '2020-11-16', to = '2021-02-14') %
  dplyr::select(date, adjusted)
# Custom function that will be used in "hacking" the x-axis tick labels
properDates = function(x) {format(lubridate::date_decimal(x), "%b %d")}
```



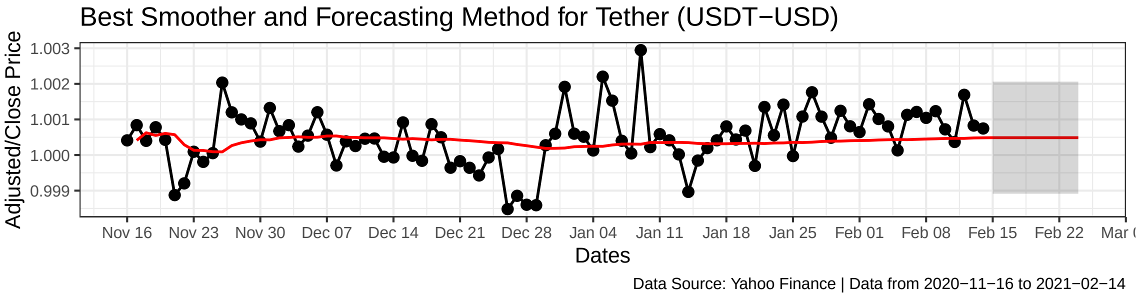
Forecasting Using Cumulative Avg, MA3 and MA7 [1]

```
usdt = tq_get('USDT-USD', from = '2020-11-16', to = '2021-02-14') %>%
  select(date, adjusted)

usdt_comp = usdt %>%
  mutate(cAVG = cummean(adjusted), # fn from dplyr
         ma3 = rollmeanr(adjusted, k = 3, na.pad = TRUE), # fn from zoo
         ma7 = rollmeanr(adjusted, k =7, na.pad = TRUE), # fn from zoo
         fcAVG = lag(cAVG), fma3 = lag(ma3), fma7 = lag(ma7) )

results = rbind(accuracy(object = usdt_comp$fcAVG, x = usdt_comp$adjusted), #
                accuracy(object = usdt_comp$fma3, x = usdt_comp$adjusted),
                accuracy(object = usdt_comp$fma7, x = usdt_comp$adjusted) )

row.names(results) = c('Cumulative AVG', 'MA3', 'MA7')
```

Forecasting Using Cumulative Avg, MA3 and MA7 [2]

	ME	RMSE	MAE	MPE	MAPE
Cumulative AVG	0.00011	0.00080	0.00060	0.01082	0.05948
MA3	0.00001	0.00078	0.00060	0.00083	0.06012
MA7	0.00005	0.00078	0.00060	0.00498	0.05974

Based on the results above, let us pick the approach with the smallest MAPE.



Outline

Definition and Basic Principles

Simple Exponential Smoothing (SES) is a method used for one-step-ahead forecasting of a time series when there is no trend or seasonal pattern, but the mean may drift slowly over time. The mean is said to have a "local level."

Similar to the idea behind a moving average, SES averages the values of the times series, but applies weights of decreasing importance to values that are farther away from the forecast. The weights of the observations "exponentially decay" as we move away from them in time.

The SES one-step-ahead forecast is given by:

$$l_{t+1} = l_t + \alpha(y_t - l_t) = \alpha y_t + (1 - \alpha) l_t, \tag{1}$$

where $0 < \alpha < 1$ is the smoothing parameter, and l_t is the level of the series at time t. Note that l_{1+1} is often denoted as f_{t+1} since it represents our one step-ahead forecast for $t + 1$.

Impact of the Smoothing Parameter: The Math

Let us examine what goes into the computations for l_4:

$$
\begin{aligned}
l_4 &= \alpha y_3 + (1-\alpha) l_3 \\
&= \alpha y_3 + (1-\alpha)[\alpha y_2 + (1-\alpha) l_2] \\
&= \alpha y_3 + \alpha(1-\alpha) y_2 + (1-\alpha)^2 l_2 \\
&= \alpha y_3 + \alpha(1-\alpha) y_2 + (1-\alpha)^2[\alpha y_1 + (1-\alpha) l_1] \\
&= \alpha y_3 + \alpha(1-\alpha) y_2 + \alpha(1-\alpha)^2 y_1 + (1-\alpha)^3 l_1 \\
&= \alpha y_3 + \alpha(1-\alpha) y_2 + \alpha(1-\alpha)^2 y_1 + (1-\alpha)^3 l_0
\end{aligned} \tag{2}
$$

Impact of the Smoothing Parameter: The Math

Let us examine what goes into the computations for l_4:

$$
\begin{aligned}
l_4 &= \alpha y_3 + (1-\alpha) l_3 \\
&= \alpha y_3 + (1-\alpha)[\alpha y_2 + (1-\alpha) l_2] \\
&= \alpha y_3 + \alpha(1-\alpha) y_2 + (1-\alpha)^2 l_2 \\
&= \alpha y_3 + \alpha(1-\alpha) y_2 + (1-\alpha)^2 [\alpha y_1 + (1-\alpha) l_1] \\
&= \alpha y_3 + \alpha(1-\alpha) y_2 + \alpha(1-\alpha)^2 y_1 + (1-\alpha)^3 l_1 \\
&= \alpha y_3 + \alpha(1-\alpha) y_2 + \alpha(1-\alpha)^2 y_1 + (1-\alpha)^3 l_0
\end{aligned} \tag{2}
$$

Impact of the Smoothing Parameter: The Math

Let us examine what goes into the computations for l_4:

$$
\begin{aligned}
l_4 &= \alpha y_3 + (1-\alpha) l_3 \\
&= \alpha y_3 + (1-\alpha)[\alpha y_2 + (1-\alpha) l_2] \\
&= \alpha y_3 + \alpha(1-\alpha) y_2 + (1-\alpha)^2 l_2 \\
&= \alpha y_3 + \alpha(1-\alpha) y_2 + (1-\alpha)^2 [\alpha y_1 + (1-\alpha) l_1] \\
&= \alpha y_3 + \alpha(1-\alpha) y_2 + \alpha(1-\alpha)^2 y_1 + (1-\alpha)^3 l_1 \\
&= \alpha y_3 + \alpha(1-\alpha) y_2 + \alpha(1-\alpha)^2 y_1 + (1-\alpha)^3 l_0
\end{aligned}
\tag{2}
$$

Impact of the Smoothing Parameter: The Math

Let us examine what goes into the computations for l_4:

$$\begin{aligned}
l_4 &= \alpha y_3 + (1-\alpha) l_3 \\
&= \alpha y_3 + (1-\alpha)[\alpha y_2 + (1-\alpha) l_2] \\
&= \alpha y_3 + \alpha(1-\alpha) y_2 + (1-\alpha)^2 l_2 \\
&= \alpha y_3 + \alpha(1-\alpha) y_2 + (1-\alpha)^2 [\alpha y_1 + (1-\alpha) l_1] \\
&= \alpha y_3 + \alpha(1-\alpha) y_2 + \alpha(1-\alpha)^2 y_1 + (1-\alpha)^3 l_1 \\
&= \alpha y_3 + \alpha(1-\alpha) y_2 + \alpha(1-\alpha)^2 y_1 + (1-\alpha)^3 l_0
\end{aligned} \tag{2}$$

Impact of the Smoothing Parameter: The Math

Let us examine what goes into the computations for l_4:

$$\begin{aligned} l_4 &= \alpha y_3 + (1-\alpha) l_3 \\ &= \alpha y_3 + (1-\alpha)[\alpha y_2 + (1-\alpha) l_2] \\ &= \alpha y_3 + \alpha(1-\alpha) y_2 + (1-\alpha)^2 l_2 \\ &= \alpha y_3 + \alpha(1-\alpha) y_2 + (1-\alpha)^2 [\alpha y_1 + (1-\alpha) l_1] \\ &= \alpha y_3 + \alpha(1-\alpha) y_2 + \alpha(1-\alpha)^2 y_1 + (1-\alpha)^3 l_1 \\ &= \alpha y_3 + \alpha(1-\alpha) y_2 + \alpha(1-\alpha)^2 y_1 + (1-\alpha)^3 l_0 \end{aligned} \tag{2}$$

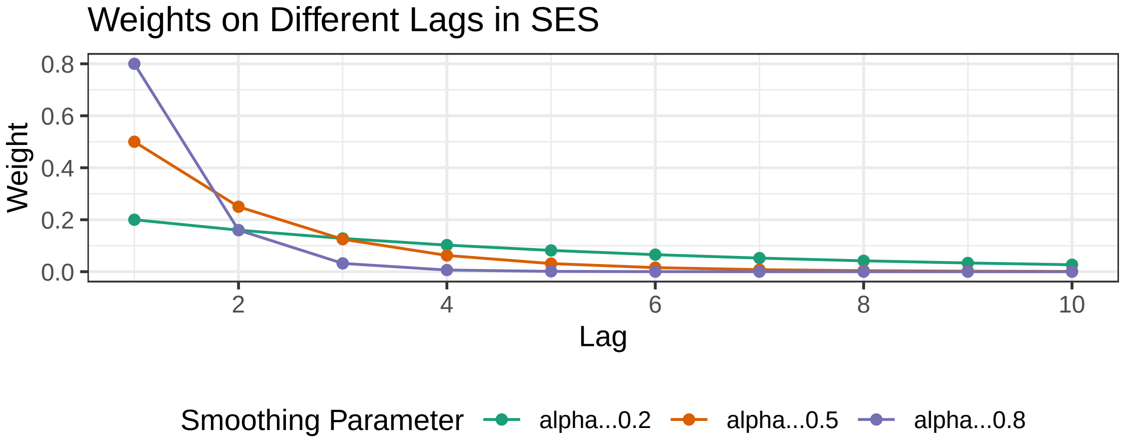
Note that SES needs two parameters: (a) the smoothing paramater α, and (b) the initial value for the level (i.e., l_0). Note that we will use $l_0 = l_1 = y_1$.

Impact of the Smoothing Parameter: The Math (Cont.)

For l_{10}, the weights of the observed values at t are distributed as follows:

t	alpha...0.2	alpha...0.5	alpha...0.8
9	0.20000	0.50000	0.80000
8	0.16000	0.25000	0.16000
7	0.12800	0.12500	0.03200
6	0.10240	0.06250	0.00640
5	0.08192	0.03125	0.00128
4	0.06554	0.01562	0.00026
3	0.05243	0.00781	0.00005
2	0.04194	0.00391	0.00001
1	0.03355	0.00195	0.00000
0	0.02684	0.00098	0.00000

Impact of the Smoothing Parameter: Visually

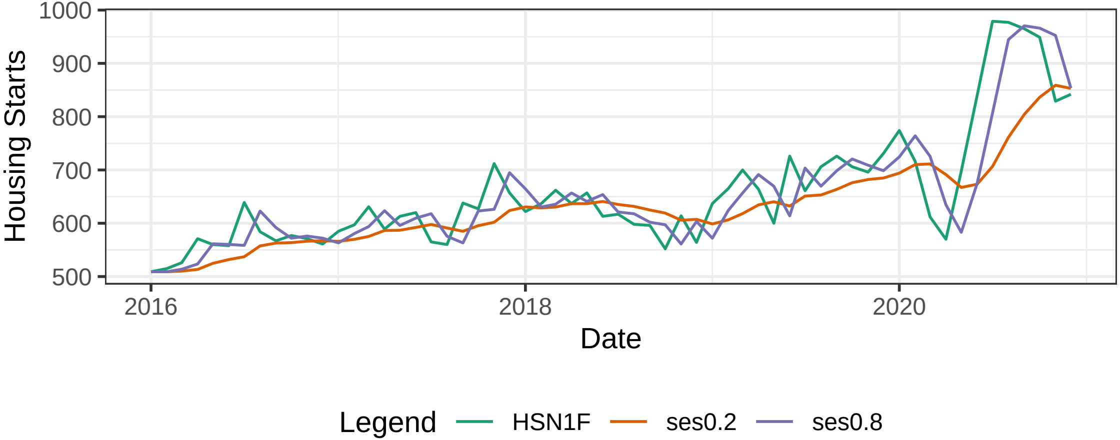


Example 1: Hand Calculations

In the example below, I have read the `HSN1F` macroeconomic variable from FRED between 2020-01-01 and 2020-07-01. If you were to select $\alpha = 0.2$ and initialize $l_0 = 774$, please fill the last three columns in the table (by performing "manual computations").

DATE	HSN1F	Forecast	Forecast Error
2020-01-01	774	...	...
2020-02-01	716	...	...
2020-03-01	612	...	...
2020-04-01	570	...	...
2020-05-01	698	...	...
2020-06-01	840	...	...
2020-07-01	979	...	...

Example 1: Charting HSN1F & its SES (2016-01 to 2020-12)

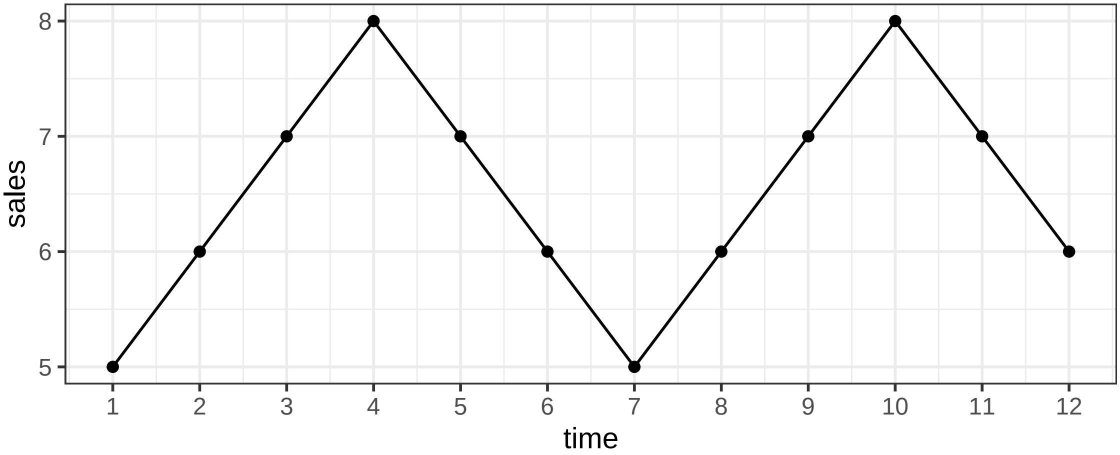


Example 2: Problem Definition

In this example, we are using R to forecast sales mimicking Table 3.3 in our textbook. See P. 72. Note the difference between the generated output here and that in the Table. If you wanted to exactly replicate that result, what minor edit to the R forecast is needed?

```
if(require(pacman)==FALSE) install.packages("pacman") # install pacman if not
pacman::p_load(magrittr, tidyverse, fpp2, scales) # load (and install if need

df = data.frame(time = 1:12,
                sales = c(5, 6, 7, 8, 7, 6, 5, 6, 7, 8, 7, 6))

df %>%
  ggplot(aes(x= time, y = sales)) + # setting the canvas
  geom_line() + geom_point() + # lines with dots highlighted
  scale_x_continuous(breaks = pretty_breaks(12)) + # making x_axis pretty (fr
  theme_bw() # using our typical black and white theme
```

Example 2: Charting the Original Data



Example 2: Using R to Compute the SES

```
sales_ts = ts(df$sales, start = 1, frequency = 1)
sesRes = ses(sales_ts, initial = "simple", alpha = 0.3, h =3, level = 95)
summary(sesRes) %>% xtable() %>% print()
```

Forecast method: Simple exponential smoothing

Model Information: Simple exponential smoothing

Call: ses(y = sales_ts, h = 3, level = 95, initial = "simple," alpha = 0.3)

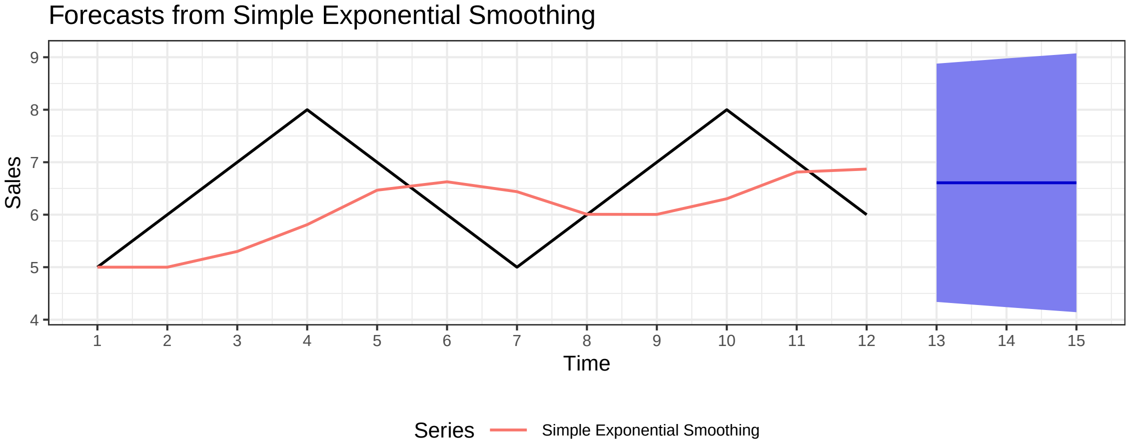
Smoothing parameters: alpha = 0.3

Initial states: l = 5

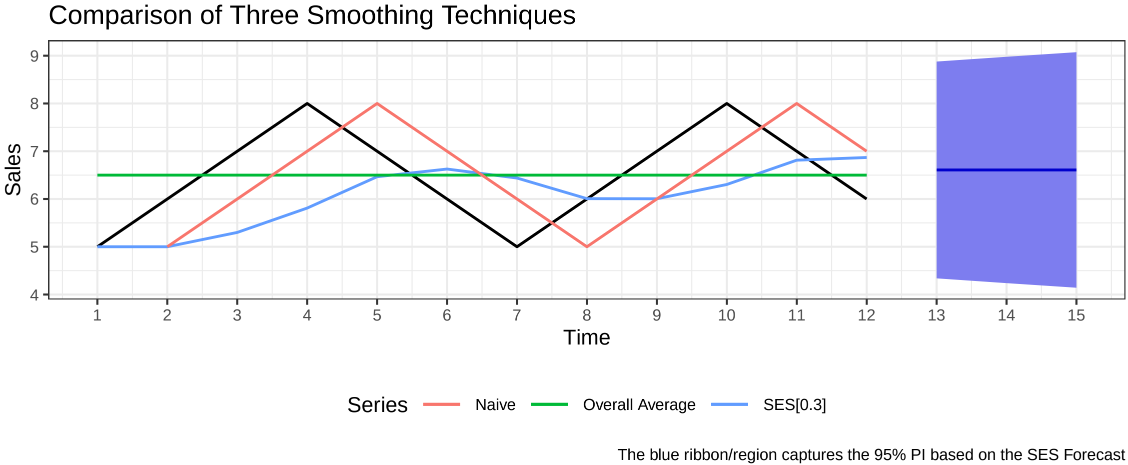
sigma: 1.158 Error measures: ME RMSE MAE MPE MAPE MASE ACF1 Training set 0.4466964 1.157992 0.9369689 5.018111 13.98867 0.9369689 0.4396956

Forecasts: Point Forecast Lo 95 Hi 95 13 6.608107 4.338484 8.877730 14 6.608107 4.238551

Example 2: Using R to Chart the Data, SES & Forecast



Example 2: Comparing SES with Other Smoothing Methods



Discussion Question

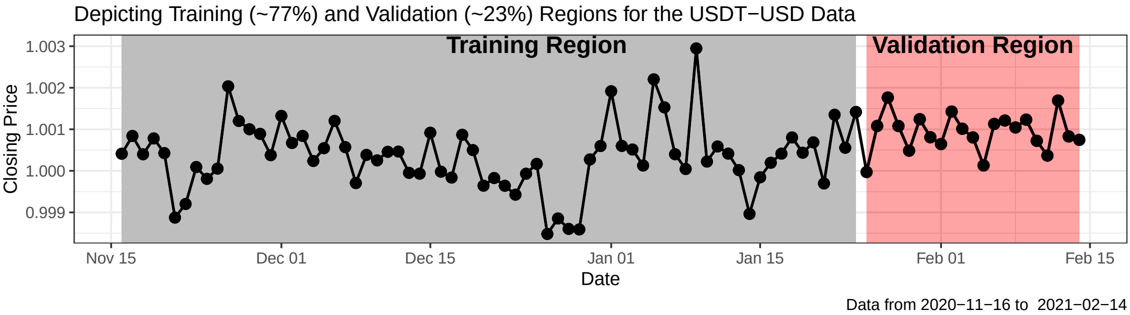
If you had to make a subjective choice for the value of the smoothing constant, what value would you choose for:

a. **a product with long-term steady sales** and

b. **a stock/cryptocurrency (e.g., AAPL, AMZN, BTC-USD, ADA-USD)?**

Training and Validation Samples [1]

Often you determine your smoothing parameter based on a training or baseline sample of observations, not the entire series. Then you apply the model using the smoothing parameter to the new observations and evaluate the fit on the out-of-sample observations.



Training and Validation Samples [2]

1. Determine the size of the training, or baseline sample.[7]
   a. Training sample size is usually 70-80% of the total available data.
   b. Training sample should maintain time order. With time series, the training sample usually consists of observations at the beginning of the sample, while validation sample consists of observations at the end of the available data.
2. Select the smoothing parameter based on the observations in the training sample only.
3. Evaluate the "in-sample" performance of the forecast using RMSE and graphs using the training sample.
4. Apply the model chosen in #2 to the validation sample.
5. Evaluate the "out-of-sample" performance of the forecast using RMSE and graphs.

[7] Slide is based on Dr. Allison Jones-Farmer's Handouts for ISA 444, Spring 2020.

Optimizing the Smoothing Parameter: WFJ Sales Series

To illustrate the aforementioned concepts, let us examine the data for the WFJ Sales Example (i.e., Example 3.2 in our textbook). Per the textbook example, we will use the first the 26 observations as the estimation sample.

```
pacman::p_load(readxl)
download.file("https://github.com/fmegahed/businessForecasting/blob/master/ass
              destfile = "Data/WFJ_sales.xlsx", mode = "wb")
WFJ = read_excel("Data/WFJ_sales.xlsx") %>% select(c(1,2))
```

This example will be coded live in class to obtain the results below.

The optimal alpha obtained using R is equal to 0.727.

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	587.398	2970.310	2058.263	1.521	6.423	0.862	-0.199

The Validation Results

A continuation of the live coding session, where we: (a) examine the validation results; and (b) print the combined training and validation results in one table.

	ME	RMSE	MAE	MPE	MAPE
Training Set	587.398	2970.310	2058.263	1.521	6.423
Validation Set	-76.797	3915.053	2562.870	-0.931	7.336

Outline

Summary of Main Points

Main Learning Outcomes

- Recognize time series that are appropriate for simple exponential smoothing (SES).
- Use SES to smooth past observations of a time series.
- Use SES to forecast future observations of a time series.
- Compare the forecasting performance of SES to other suitable techniques (i.e., methods that require similar assumptions).

Things to Do

- **Recommended:** Thoroughly read Chapter 3.1-3.3 of our book.
- Go through the slides, examples and make sure you have a good understanding of what we have covered.
- **Required:** Complete the graded assignment (see details in next slide).

Graded Assignment 07: Evaluating your Understanding

Please go to Canvas (click here) and answer the questions. **Due February 18, 2021 [11:59 AM, Ohio local time] | Will be available starting from 5PM (Feb 15, 2021)**

What/Why/Prep? The purpose of this assignment is to evaluate your understanding and retention of the material covered up to the end of Class 07. To reinforce your understanding of the covered material, I also suggest reading Chapter 3.1-3.3 of the book.

General Guidelines:

- Individual assignment.
- This is **NOT** a timed assignment.
- Proctorio is NOT required for this assignment.
- You will need to have R installed (or accessible through the Remote Desktop)

ISA 444: Business Forecasting

07 - Nonseasonal Smoothing (Cont.)

Fadel M. Megahed

Associate Professor
Department of Information Systems and Analytics
Farmer School of Business
Miami University
Email: fmegahed@miamioh.edu
Office Hours: Click here to schedule an appointment

Spring 2021