ISA 444: Business Forecasting

06 - Nonseasonal Smoothing

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 Class

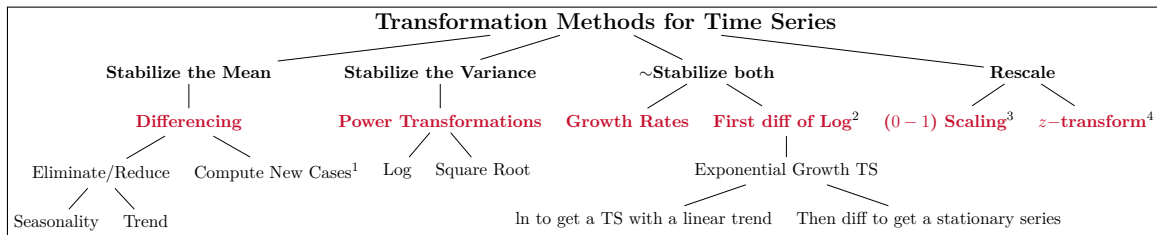**Main Learning Outcomes**

☑ **Apply transformations to a time series.**
☑ **Apply and interpret measures of forecast accuracy.**
☑ **Interpret prediction intervals for a simple forecast.**

Recap: Guidelines for Transforming Time-Series Data

Transformation Methods for Time Series

Stabilize the Mean   **Stabilize the Variance**   **~Stabilize both**   **Rescale**

Differencing   **Power Transformations**   **Growth Rates**   **First diff of Log**[2]   **(0 − 1) Scaling**[3]   **z−transform**[4]

Eliminate/Reduce   Compute New Cases[1]   Log   Square Root   Exponential Growth TS

Seasonality   Trend   ln to get a TS with a linear trend   Then diff to get a stationary series

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

---

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

[2] First difference of LOG ≅ 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: 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
ADA-USD	2021-02-01	0.402	
ADA-USD	2021-02-02	0.426	0.402
ADA-USD	2021-02-03	0.441	0.426
ADA-USD	2021-02-04	0.441	0.441
ADA-USD	2021-02-05	0.538	0.441
ADA-USD	2021-02-06	0.628	0.538
ADA-USD	2021-02-07	0.663	0.628
ADA-USD	2021-02-08	0.683	0.663

Based on the naive forecast, we can compute: (note the results are concerning!! Why?)

ME	RMSE	MAE	MPE	MAPE
0.0401	0.0533	0.0402	7.08	7.10

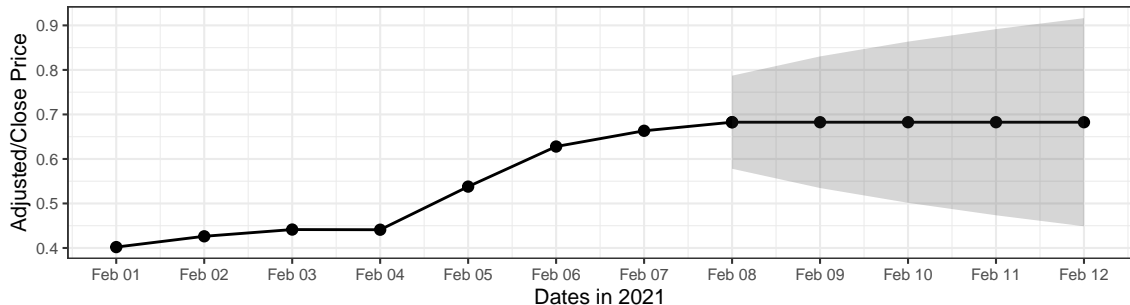# Recap: Prediction Intervals [1]

- **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: Prediction Intervals [2]

Naive Forecast for ADA−USD

Data Source: Yahoo Finance | Data from 2021−02−01 to 2021−02−08

There are two important observations to be made: (a) future predictions do not change, and (b) prediction interval is more advanced than the approach detailed last class/slide (**Why? Why is that better?**) **You will NOT be required to compute this PI by hand.**

Learning Objectives for Today's Class

Main Learning Outcomes

- 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.

Outline

Key Terms

Smoothing *is usually done to help us better see patterns/trends, e.g., in time series.*[6]

- *Generally smooth out the irregular roughness to see a clearer signal.*

- *For seasonal data, we can smooth out the seasonality so that we can identify the trend.*

- *Smoothing does not provide us with a model, but it can be a good first step in describing various components of the series.*

A **filter** is sometimes used to describe a smoothing procedure. For example, we have applied a median filter of window size 21 to smooth wearable sensors' data.
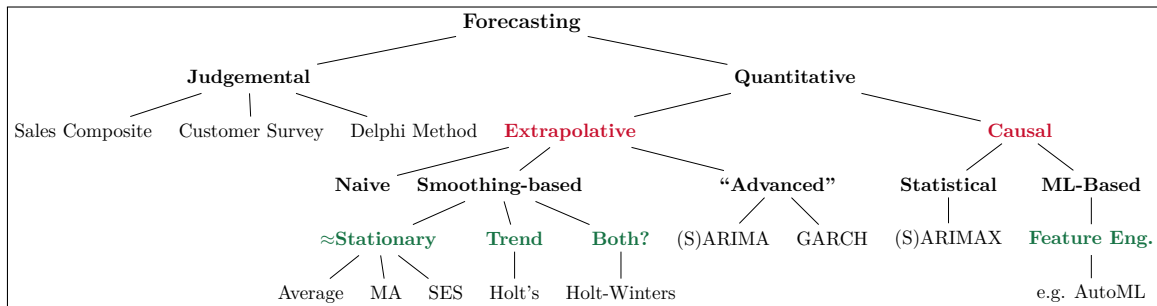
Forecast *is a prediction or estimate of an actual outcome expected in a future time period or for another situation.*[7]

---

[6]Definition and bullets are based on STAT 510: Applied Time Series Analysis | Penn State.
[7]From: Ord, K., Fildes, R., & Kourentzes, N. (2017). Principles of Business Forecasting (2nd ed., p. 3).

A 10,000 Foot View of Forecasting Methods

A 10,000 foot view of forecasting techniques[8]

---

[8]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.

Definitions

Judgmental Forecasting: The process of producing forecasts based on purely subjective information. The integration of subjective information may be made informally or through a structured process. The forecasts may also be obtained by aggregating the subjective forecasts of a number of individuals.

Quantitative Forecasting: Forecasting based on the application of an explicit analysis of numerical data. This kind of forecasting may be extrapolative, causal, or a blend of both.[9]

- *Causal Forecast:* a dependent variable is forecast using explanatory variables.
- *Extrapolative:* a dependent variable is forecast using only the past values of the dependent variables. The future is "extrapolated" from the past.

---

[9]Definitions are based on Dr. Allison Jones-Farmer's lecture notes, Miami University, Spring 2020.

Some Potential Drawbacks for Judgmental Forecasts

When using judgmental forecasts, forecasters may succumb to:

- An **availability bias**, when the forecaster relies too heavily on easily available and memorable information.
- The **representativeness heuristic**, when the forecaster matches a situation to a similar earlier event without taking into account its frequency of occurrence.
- The **anchoring** and **adjustment heuristic**, when the forecaster uses (anchors onto) an initial value such as the last observation and then adjusts the value to give a revised forecast.
- Over-optimism or **motivational bias** when the forecaster is motivated to bias the forecast towards a preferred state.

These biases can lead to invalid forecasts; they may lead to poor decision making, particularly when combined with overconfidence in their beliefs as to the accuracy of their forecasts.[10]

---

[10] Slide adapted from Dr. Allison Jones-Farmer's lecture notes, Miami University, Spring 2020.

Big Picture for Our Course

The remainder of this course focuses on Quantitative Forecasting

- Chapters 3, 4, 6 focus on Extrapolative Forecasting techniques.

- Chapter 7-8 Focus on Causal (regression-based) Forecasting techniques.

- At the end of the semester we will briefly learn how to blend Extrapolative and Causal Forecasting techniques.

Outline

Causal Models

- Causal methods use data from **sources other than the series being predicted**.

- If Y is the phenomenon to forecast and $X_1, X2, \ldots, X_n$ are the n variables we believe to be related to Y, then a causal model is one in which the forecast for Y is some function of these variables: $Y = f(X_1, X2, \ldots, X_n)$.

- Econometric models are causal models in which the relationship between Y and $X_1, X2, \ldots, X_n$ is linear. That is:

$$Y = a_o + a_1 X_1 + a_2 X_2 + \cdots + a_n X_n \tag{1}$$

for some constants $a_1, a_2, \ldots, a_n$

# Causal Models: Does Racism Affect Voting?

- Go to https://ed.ted.com/on/HTDYPiuQ
- Watch the 9 minute video
- Answer the five multiple choice questions under the think tab
- Be ready to discuss the open-ended question under the discuss tab

Does racism affect how you vote? – Nate Silver's TED Talk

Extrapolative Methods: A General Framework

Some **forecasting methods** we will discuss use the values of the series to extrapolate into the future. These **extrapolative methods** often work well for short term forecasts.
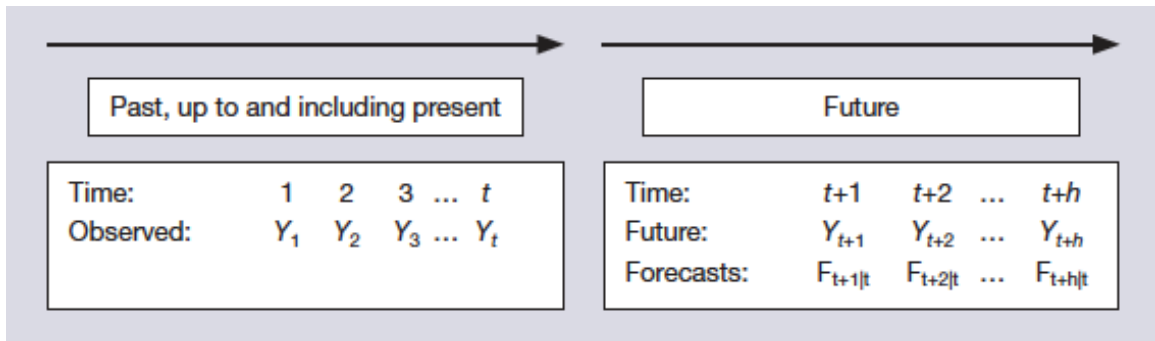
General Framework for Forecasting with a Single Series

Extrapolative Methods: Rolling Origin

Rolling Origin Forecasts: The closer you are to the forecast horizon, the better the forecast. Often, as you gain more information, you update the forecast.

Example:

- On Sunday, you might forecast the weather for the entire week ahead: Monday, Tuesday, Wednesday, Thursday, Friday, Saturday.

- Monday is pretty accurate, but by the time you get to Saturday's forecast, it is not so accurate.

- Once Monday comes, you update the forecast, now forecasting Tuesday-Sunday.

- When Tuesday comes, you will update, and forecast Wednesday-Monday.

- This is a rolling origin forecast.

Outline

Key Point: "Weak" Stationarity

Today, you are being introduced to methodologies where the time series exhibits **NO** trends and no seasonal patterns.[11]

- *A stationary time series is one whose properties do not depend on the time at which the series is observed.*

- *Some cases can be confusing — a time series with cyclic behavior (but with no trend or seasonality) is stationary. This is because the cycles are not of a fixed length, so before we observe the series we cannot be sure where the peaks and troughs of the cycles will be.*

- *In general, a stationary time series will have no predictable patterns in the long-term. Time plots will show the series to be roughly horizontal (although some cyclic behavior is possible), with constant variance.*

---

[11]The text presented here is from Hyndman, R.J., & Athanasopoulos, G. (2018) *Forecasting: Principles and Practice*, 2nd Edition, OTexts.com/fpp2.
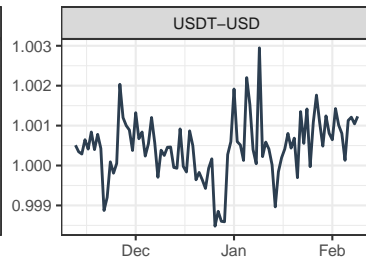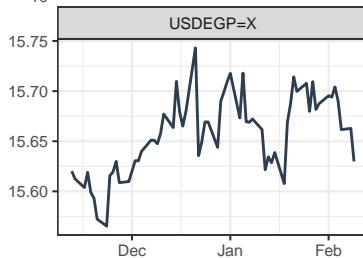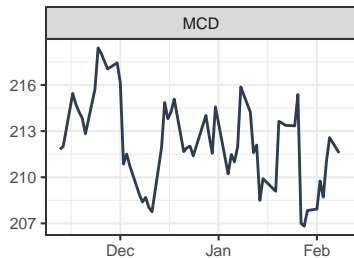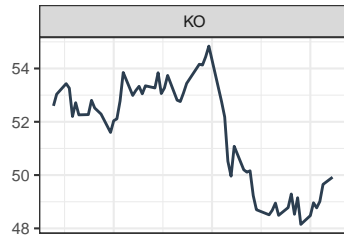
Weak Stationarity: A Visual
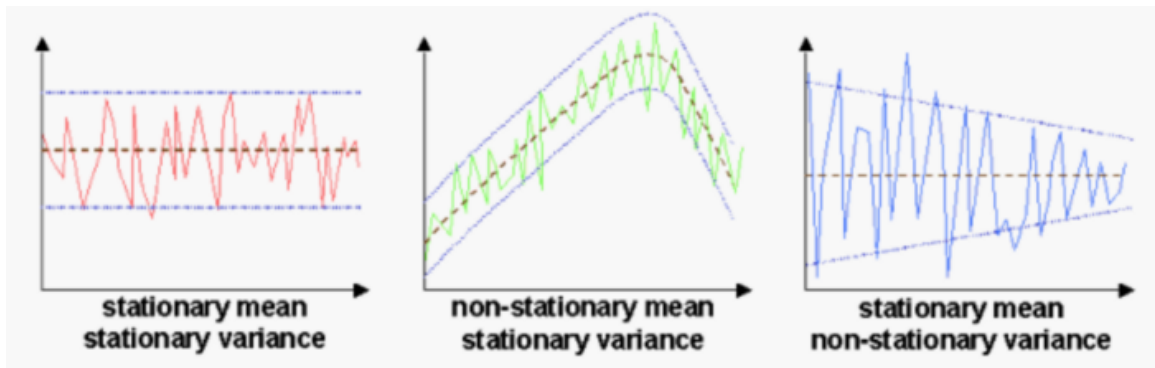
Constancy in mean and variance[12]

---
[12]Image Source: Palachy, S. (2019). Stationarity in time series analysis. towardsdatascience.com.

Weak Stationarity: Which asset(s) is somewhat stationary?

Overall Average [1]

If you have a series that stays pretty constant over time, you could just use the overall average for smoothing/forecasting. In R, we can use the `meanf()` from fpp2 to smooth and forecast using the overall mean.
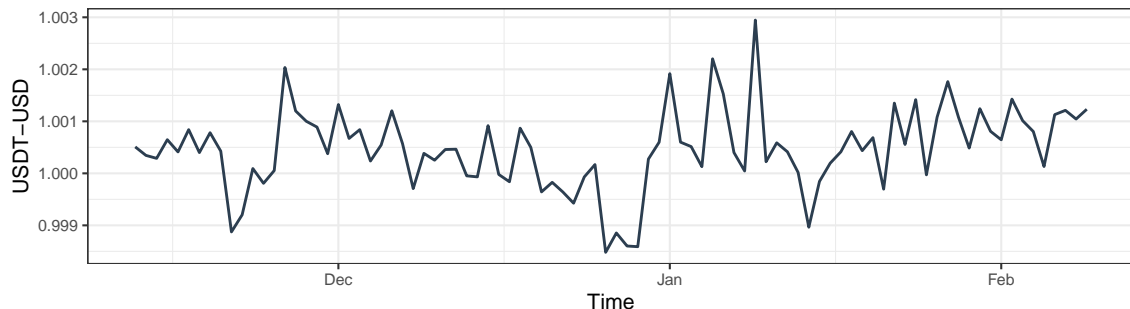
Overall Average [2]

Forecasts from Mean

Cumulative Average [1]

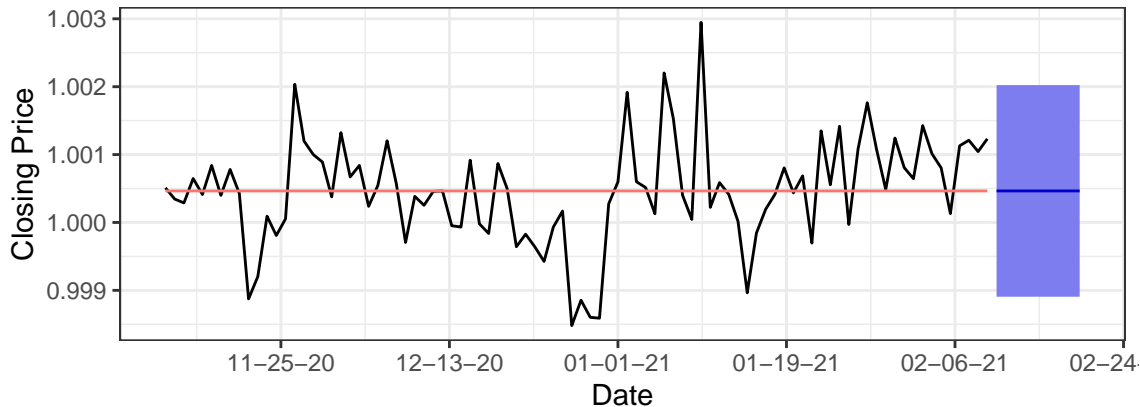**If you have a series that stays pretty constant over time, you could just constantly update the mean as you gain more information.** This can be computed using the cummean() from tidyverse.

date	adjusted	cuMean
2021-01-11	1.0006	1.0006
2021-01-12	1.0004	1.0005
2021-01-13	1.0000	1.0003
2021-01-14	0.9990	1.0000
2021-01-15	0.9998	1.0000
2021-01-16	1.0002	1.0000
2021-01-17	1.0004	1.0001
2021-01-18	1.0008	1.0002
2021-01-19	1.0004	1.0002
2021-01-20	1.0007	1.0002
2021-01-21	0.9997	1.0002
2021-01-22	1.0013	1.0003
2021-01-23	1.0006	1.0003
2021-01-24	1.0014	1.0004
2021-01-25	1.0000	1.0004

Cumulative Average [2]

Smoothing using Cumulative Mean
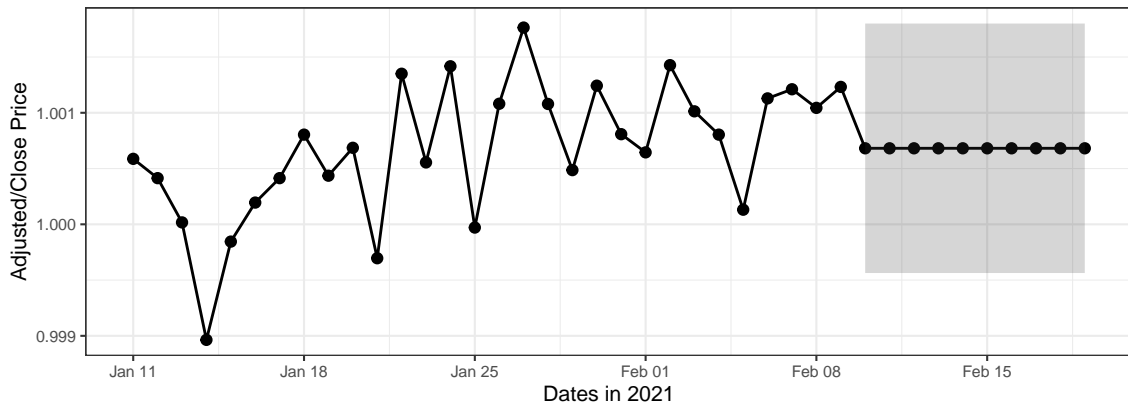
Data Source: Yahoo Finance | Data from 2021−01−11 to 2021−02−09

Moving Average [1]

In this live coding demo, we will smooth the USDT series using: (a) the overall/global average, (b) cumulative mean, (c) 3-day moving average, and (d) 7-day moving average for USDT. Then, we will plot the data (similar to next slide).

date	adjusted	globalAVG	cumMean	ma3	ma7
2021-01-11	1.0006	1.0007	1.0006		
2021-01-12	1.0004	1.0007	1.0005		
2021-01-13	1.0000	1.0007	1.0003	1.0003	
2021-01-14	0.9990	1.0007	1.0000	0.9998	
2021-01-15	0.9998	1.0007	1.0000	0.9996	
2021-01-16	1.0002	1.0007	1.0000	0.9997	
2021-01-17	1.0004	1.0007	1.0001	1.0002	1.0001
2021-01-18	1.0008	1.0007	1.0002	1.0005	1.0001
2021-01-19	1.0004	1.0007	1.0002	1.0006	1.0001
2021-01-20	1.0007	1.0007	1.0002	1.0006	1.0002
2021-01-21	0.9997	1.0007	1.0002	1.0003	1.0003
2021-01-22	1.0013	1.0007	1.0003	1.0006	1.0005

Moving Average [2]
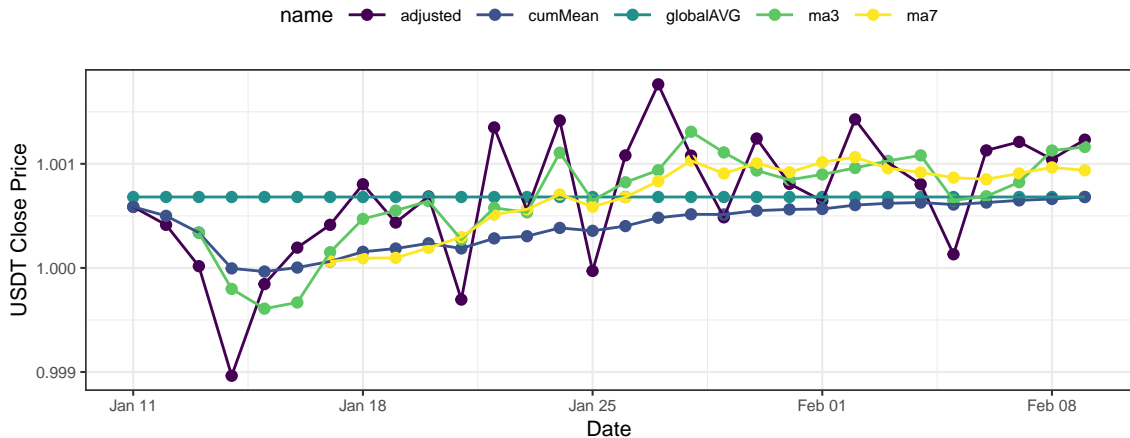
Difference between a Smoothing and Forecasting Problem

So how do you differentiate between smoothing and forecasting if we were performing the calculations manually?

date	adjusted	globalAVG	cumMean	ma3	ma7	fAVG	fCM	fma3	fma7
2021-01-11	1.001	1.001	1.001						
2021-01-12	1.000	1.001	1.001			1.001	1.001		
2021-01-13	1.000	1.001	1.000	1.000		1.001	1.001		
2021-01-14	0.999	1.001	1.000	1.000		1.001	1.000	1.000	
2021-01-15	1.000	1.001	1.000	1.000		1.001	1.000	1.000	
2021-01-16	1.000	1.001	1.000	1.000		1.001	1.000	1.000	
2021-01-17	1.000	1.001	1.000	1.000	1.000	1.001	1.000	1.000	
2021-01-18	1.001	1.001	1.000	1.000	1.000	1.001	1.000	1.000	1.000
2021-01-19	1.000	1.001	1.000	1.001	1.000	1.001	1.000	1.000	1.000
2021-01-20	1.001	1.001	1.000	1.001	1.000	1.001	1.000	1.001	1.000

	ME	RMSE	MAE	MPE	MAPE
Overall Mean	0.0000	0.0006	0.0005	0.0003	0.0492
Cumulative Mean	0.0003	0.0006	0.0005	0.0286	0.0539
MA3	0.0001	0.0006	0.0005	0.0068	0.0501
MA7	0.0002	0.0005	0.0005	0.0173	0.0454

Outline

1. Preface

2. Forecasting Terminology

3. Causal vs Explarotive Forecasts

4. Smoothing and/or Forecasting with Means

5. **Recap**

Summary of Main Points

Main Learning Outcomes

- 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.

Things to Do

- **Recommended:** Thoroughly read Chapter 3.1-3.2 of our book.

- Go through the slides, examples and make sure you have a good understanding of what we have covered.

- If you are interested in additional practice problems, please consider the following problems from your textbook. To access these datasets, please click here.

    - **Practice Problem:** Exercise 3.1 (RMSE is 2.54 and 2.28 for the MA 3 and MA7, respectively).

- **Highly Recommended:** Go through the Week 03 Self-Paced Study Guide.

- **Required:** Complete the graded assignment (see details in next slide).

Graded Assignment 06: Evaluating your Understanding

Please go to Canvas (click here) and answer the questions. **Due February 15, 2021 [11:40 AM, Ohio local time] | Will be available starting from 5PM (Feb 10, 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 06. To reinforce your understanding of the covered material, I also suggest reading Chapter 3.1-3.2 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

06 - Nonseasonal Smoothing

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