ISA 444: Business Forecasting

12 - Autocorrelation

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

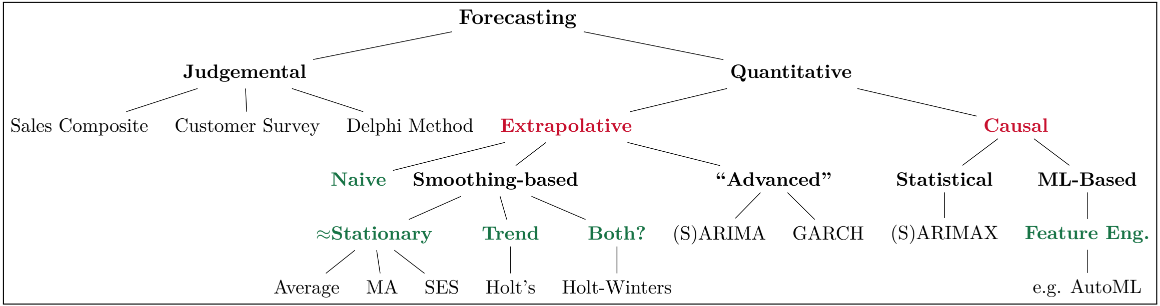
Outline

What we Covered Last Week

Main Learning Outcomes

- ☑ Recognize time series that are appropriate for linear exponential smoothing (LES).
- ☑ Use LES to forecast future observations of a time series.
- ☑ Explain when to use an additive vs. multiplicative model for a time series.
- ☑ Use classic decomposition methods to detrend and deseasonalize a time series.
- ☑ Recognize time series that are appropriate for triple exponential smoothing (HW).
- ☑ Use HW to forecast future observations of a time series.

Recap: A 10,000 Foot View of Forecasting Methods

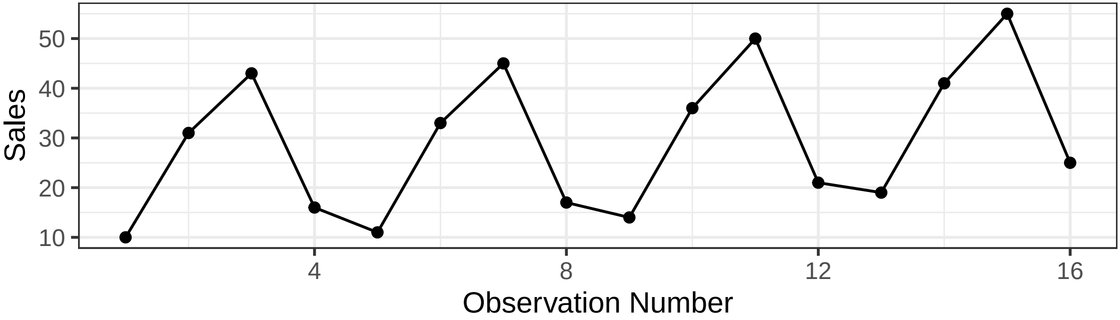


A 10,000 foot view of forecasting techniques[1]

[1] 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.

Recap: Accuracy of Holt-Winters (Additive & Multiplicative)

Last class, we discussed how to fit an additive HW method on the `BikeSalesR.xlsx` dataset; I noted that the process of fitting an additive model is similar. To ensure that you have a good grasp of this idea, let us compare the accuracy of the additive and multiplicative fits of the `hw()` for that dataset. We will use $\alpha = 0.2$, $\beta = 0.1$, and $\gamma = 0.1$.



Outline

Definition and Notation

A random variable, Y , is the outcome of a random experiment. The random nature of Y can occur through a variety of mechanisms including sampling, natural variation, etc . In time series, we write Y_t to represent the random variable at time t , where $t = 1, 2, 3, \ldots$.

Specific observed values of a random variable are written as lower case letters, y_t.

Example to demonstrate the notation

```
pacman::p_load(tidyquant, timetk)
aapl = tq_get('AAPL', from = "2021-01-07", to = Sys.Date() -1) %>%
  select(date, adjusted)
aapl_ts = timetk::tk_ts(aapl) # allows for non-equally spaced ts
```

Y_2 represents the adjusted **but not observed** closing price for the $AAPL stock on 2021-01-08. When we observe a value for this we have, $y_2 = 131.85$.

Basic Population Parameter Functions

Mean Function:

$$\mu_{Y\,t} = \mu_t = E(Y_t). \tag{1}$$

Variance Function:

$$\sigma_t^2 = E[(Y_t - \mu_t)^2]. \tag{2}$$

Covariance Function: The covariance of two random variables, Y and Z is given by

$$E[(Y - \mu_Y)(Z - \mu_Z)]. \tag{3}$$

The covariance measures the *linear dependence* between two random variables.

The Correlation Coefficient between two random variables, Y and Z is given by $\rho = \frac{E[(Y-\mu_Y)(Z-\mu_Z)]}{\sigma_Y \sigma_z}$. It measures the scaled linear dependence between two random variables, and is in the interval $[-1, 1]$.

Outline

Autocovariance Function

In time series applications, often, our best predictor of a future observation is the past values of the series. Thus, we measure the linear dependence of the series over time using the autocovariance (autocorrelation) functions. For the random variable Y observed at two different times, Y_s and Y_t , the autocovariance function is defined as:

$$\gamma(s,t) = cov(Y_s, Y_t) = E[(Y_s - \mu_s)(Y_t - \mu_t)]. \tag{4}$$

Notes:

- $\gamma(s,t) = \gamma(t,s)$.
- If $\gamma(s,t) = 0$, then Y_s and Y_t are **NOT linearly related**.
- $\gamma(t,t) = \sigma_t^2$.

Autocorrelation Function

In applications, we generally use the Autocorrelation Function (ACF):

$$\rho(s,t) = \frac{\gamma(s,t)}{\sqrt{\gamma(s,s)\gamma(t,t)}} = \frac{\gamma(s,t)}{\sqrt{\sigma_s^2 \sigma_t^2}} \tag{5}$$

Notes:

- The ACF is in the interval $[-1, 1]$.
- The ACF measures the linear predictability of the series at time t using only information from time Y_s.

An Example

Consider a white noise, centered moving average model, where w_t is distributed *iid* $N(0, 1)$ and $Y_t = \frac{1}{3}(w_{t-1} + w_t + w_{t+1})$.

- Population Mean: $E(Y_t) = 0$
- Population Variance: $\sigma^2(Y_t) = \frac{1}{3}$
- Population Autocovariance between times t and $t+1$: $\gamma(t+1, 1) = \frac{2}{9}$
- Population Autocorrelation between times t and $t+1$: $\rho(t+1, 1) = \frac{2}{3}$
- Population Autocorrelation between times t and $t+2$: $\rho(t+2, 1) = \frac{1}{3}$
- Population Autocorrelation between times t and $t+3$: $\rho(t+3, 1) = 0$
- Population Autocorrelation between times t and $t+4$: $\rho(t+3, 1) = 0$

Outline

Definitions

Sample mean:

$$\bar{y} = \frac{1}{n}\sum_{t=1}^{n} y_t$$

Sample variance:

$$\hat{\sigma}_y^2 = \frac{1}{n-1}\sum_{t=1}^{n}(y_t - \bar{y})^2$$

Standard error of the mean:

$$\hat{\sigma}_{\bar{y}}^2 = \sqrt{\frac{\hat{\sigma}_y^2}{n}} = \frac{\hat{\sigma}_y}{\sqrt{n}}$$

Lag k Sample Autocorrelation:

$$r_k = \frac{\sum_{t=k+1}^{n}(y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^{n}(y_t - \bar{y})^2}$$

Comments on the Sample ACF

- The sample ACF is very useful in helping us to determine the degree of autocorrelation in our time series.
- However, the sample ACF is subject to random sampling variability. Like the sample mean, the sample ACF has a sampling distribution.

Outline

Properties [1]

- A common heuristic is that at least 50 observations are needed to give a reliable estimate of the population ACF, and that the sample ACF should be computed up to lag $K = \frac{n}{4}$, where n is the length of the series available for training.
- Under general conditions, for large n, and $k = 1, 2, \ldots$, the ACF follows an approximate normal distribution with zero mean and standard deviation given by $\frac{1}{\sqrt{n}}$.
- This result can be used to give us a cutoff to determine if there is a statistically significant amount of autocorrelation for a given lag in a series.

Properties [2]

- `R` uses a cutoff of $\pm 1.96\frac{1}{\sqrt{n}}$ to determine statistical significance of the sample ACF.
  - That is if the sample ACF is **within** $\pm 1.96\frac{1}{\sqrt{n}}$, it is considered **NOT** significant.
  - If the sample ACF is greater than $+1.96\frac{1}{\sqrt{n}}$, then there is significant positive autocorrelation at a particular lag.
  - If the sample ACF is less than $-1.96\frac{1}{\sqrt{n}}$, then there is significant negative autocorrelation at a particular lag.

Example 1: White Noise

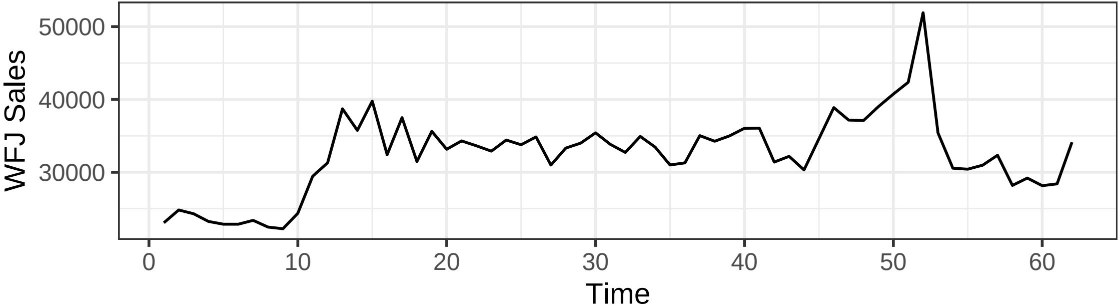
In this live coding session, we will generate the following time-series:

- White Noise
- Centered Moving Average of the White Noise Data

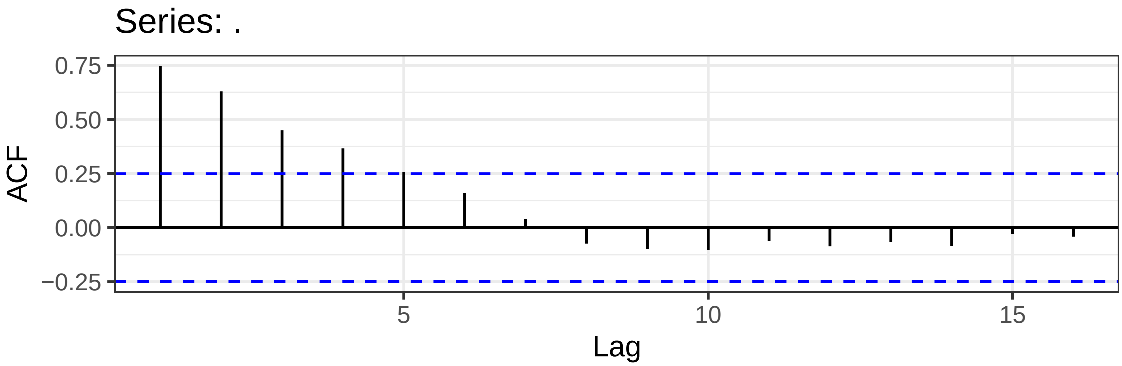
We will plot both time-series as well as their corresponding ACFs. We will also comment on the obtained results.

Example 2: The WFJ Sales Data [1]

In this live-demo example, we will use R to plot the ACF for the WFJ Sales Data. Note that this corresponds to Figure 6.2 in your textbook; however `R` uses constant significance limits.



Example 2: The WFJ Sales Data [2]



Example 2: The WFJ Sales Data [3]

	Dependent variable:
	'WFJ Sales'
Lag1	0.749*** (0.082)
Constant	8,337.702*** (2,682.195)
Observations	61
R^2	0.588
Adjusted R^2	0.581
Residual Std. Error	3,492.255 (df = 59)
F Statistic	84.367*** (df = 1; 59)
Note:	*p<0.1; **p<0.05; ***p<0.01

Outline

Summary of Main Points

Main Learning Outcomes

- Define the population mean, and variance of a random variable.
- Define the population covariance, and correlation between two random variables.
- Define the population autocovariance and autocorrelation of a random variable.
- Use sample estimates of the population mean, variance, covariance, and correlation.
- Explain the properties of the large sample distribution of the sample ACF.
- Use the large sample distribution of the sample ACF to identify significant autocorrelation in a time series.

Things to Do

- Thoroughly read Chapter 6.1 of our textbook.
- Go through the slides, examples and make sure you have a good understanding of what we have covered.
- Complete the assignment (see details in next slide).

Graded Assignment 12: Evaluating your Understanding

Please go to Canvas (click here) and answer the questions. **The assignment is due March 11, 2021 [11:59 PM, Ohio local time].**

What/Why/Prep? The purpose of this assignment is to evaluate your understanding and retention of autocorrelation. To reinforce your understanding of the covered material, I also suggest reading Chapter 6.1 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

12 - Autocorrelation

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