ISA 444: Business Forecasting

15: ACF and PACF

Fadel M. Megahed, PhD

Endres Associate Professor
Farmer School of Business
Miami University

@FadelMegahed
fmegahed
fmegahed@miamioh.edu
Automated Scheduler for Office Hours

Spring 2023

Quick Refresher from Last Week

- ✅ Explain the difference between fixed window and rolling origin forecasting.
- ✅ Apply several forecasting methods to the fixed forecasting window strategy.
- ✅ Apply several forecasting methods to the rolling origin forecasting window strategy.

Assignment #11: Demo Based on Class 14

We will write a function that takes ticker/symbol, and returns our `summary_df`. Then, we will:

- find the model that has the lowest MAPE for each ticker/symbol.
- count the number of times a given model has won.

Learning Objectives for Today's Class

- Explain what do we mean by population/sample mean, variance, covariance and correlation (**review**).
- Explain the population autocovariance and autocorrelation.
- Compute sample estimates of the autcovariance and autocorrelation.
- Describe the large sample distribution of the autocorrelation function.
- Explain how sample (partial) autocorrelation is calculated.
- Use R to compute both the ACF and PACF.

Review of Population Mean, Variance, Covariance & Correlation

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.

```
btc =
  tidyquant::tq_get('BTC-USD', from = "2023-01-01", to = Sys.Date() -1) |>
  dplyr::select(date, adjusted)
```

Y_2 represents the adjusted **but not observed** closing price for BTC on 2023-01-02. When we observe a value for this we have, $y_2 = 16688.47$.

Basic Population Parameter Functions

Mean Function:

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

Variance Function:

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

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

$$E[(Y - \mu_Y)(Z - \mu_Z)].$$

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

Basic Population Parameter Functions

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 $[-1, 1]$.

Population Autocovariance and Autocorrelation

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:

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

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):

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

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.

Non-graded Class Activity

05:00

Consider a white noise, centered moving average model, where w_t is distributed $iid\ N(0,1)$ and $Y_t = \frac{1}{3}(w_{t-1} + w_t + w_{t+1})$. **Please use the next 5 minutes to solve this either logically or programatically.**

- Population Mean: $E(Y_t) = ...$
- Population Variance: $\sigma^2(Y_t) = ...$
- Population Autocorrelation between times t and $t+1$: $\rho(t+1, 1) = ...$
- Population Autocorrelation between times t and $t+2$: $\rho(t+2, 1) = ...$
- Population Autocorrelation between times t and $t+3$: $\rho(t+3, 1) = ...$
- Population Autocorrelation between times t and $t+4$: $\rho(t+3, 1) = ...$

Sample Estimates of Population Parameters

and

The Large Sample Distribution of the ACF

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.

Large Sample Distribution of the ACF

- 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.
- 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 $> +1.96 \frac{1}{\sqrt{n}}$, then there is significant positive autocorrelation at a particular lag.
    - If the sample ACF is $< -1.96 \frac{1}{\sqrt{n}}$, then there is significant negative autocorrelation at a particular lag.

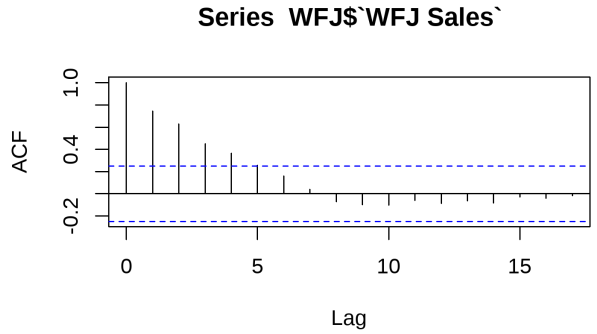
Example: The WFJ Sales Dataset

We will use R to: (a) plot the ACF for the WFJ Sales Data; (b) extract the acf values; and (c) fit a linear model where we attempt to predict sales as a function of lag1. Note that the acf plot corresponds to Figure 6.2 in your reference book; however R uses constant significance limits.

```
WFJ = readxl::read_excel(
  "../../data/WFJ_sales.xlsx") |>
  dplyr::select(1,2)

acf_results = acf(x = WFJ$`WFJ Sales`)

acf_results$acf
```

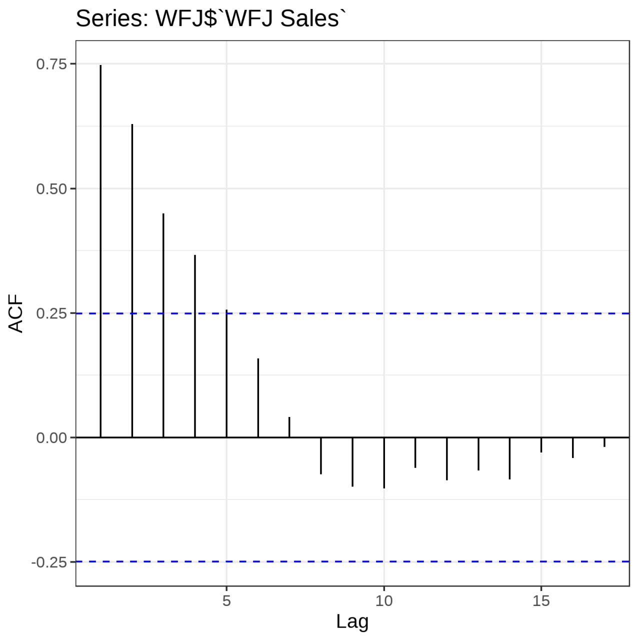



```
## , , 1
##
##              [,1]
##  [1,]  1.00000000
##  [2,]  0.74736606
##  [3,]  0.62952873
##  [4,]  0.44959855
##  [5,]  0.36613401
##  [6,]  0.25620022
##  [7,]  0.15915319
##  [8,]  0.04077136
##  [9,] -0.07380193
```

Example: The WFJ Sales Dataset

We will use R to: (a) plot the ACF for the WFJ Sales Data; (b) extract the acf values; and (c) fit a linear model where we attempt to predict sales as a function of lag1. Note that the acf plot corresponds to Figure 6.2 in your reference book; however R uses constant significance limits.

```
## for a ggplot acf plot (with no lag0)
forecast::autoplot(acf_results) +
  ggplot2::theme_bw()
```



Example: The WFJ Sales Dataset

We will use R to: (a) plot the ACF for the WFJ Sales Data; (b) extract the acf values; and (c) fit a linear model where we attempt to predict sales as a function of lag1. Note that the acf plot corresponds to Figure 6.2 in your reference book; however R uses constant significance limits.

```r
WFJ$Lag1 = dplyr::lag(WFJ$`WFJ Sales`, n =1)

# traditional lm model
model = lm(data = WFJ, formula = `WFJ Sales` ~ Lag1

# print model nicely
stargazer::stargazer(
  model, type = 'html', header = F, single.row = T)
```

	Dependent variable:
	$WFJSa \leq s$
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

Partial Autocorrelation

General Definition

Statistical Definition: Let us say that we have three variables, X, Y, and Z, all correlated, and we want to know how X and Y are correlated after we remove the effects of Z on each.

Computation Approach:

$$\hat{X} = a_1 + b_1 Z; \qquad X^* = X - \hat{X}$$

$$\hat{Y} = a_2 + b_2 Z; \qquad Y^* = Y - \hat{Y}$$

$Corr(X^*, Y^*)$ is the partial correlation between X and Y. It is the correlation that remains after we remove the effect of Z.

PACF in the Context of Time-Series Analysis

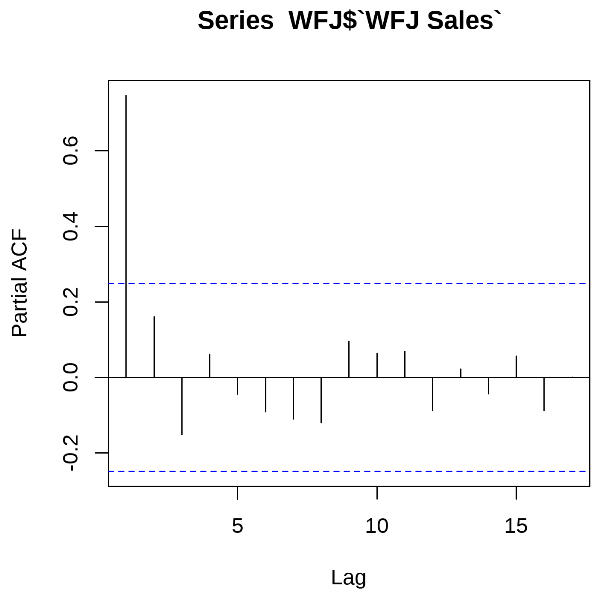
The Partial Autocorrelation between Y_t and Y_{t+k} is the correlation between Y_t and Y_{t+k} after removing the effects of $Y_{t+1}, Y_{t+2}, Y_{t+3}, \ldots, Y_{t+k-1}$.

- We plot the partial autocorrelation over multiple lags just like the autocorrelation function (ACF).
- We refer to the plotted partial autocorrelations as the PACF.

Computing the PACF in R

```
pacf_results = pacf(x = WFJ$`WFJ Sales`)

pacf_results$acf
```




```
## , , 1
##
##             [,1]
##  [1,]  0.74736606
##  [2,]  0.16077397
##  [3,] -0.15209185
##  [4,]  0.06082456
```

Recap

Summary of Main Points

By now, you should be able to do the following:

- Explain what do we mean by population/sample mean, variance, covariance and correlation (**review**).
- Explain the population autocovariance and autocorrelation.
- Compute sample estimates of the autcovariance and autocorrelation.
- Describe the large sample distribution of the autocorrelation function.
- Explain how sample (partial) autocorrelation is calculated.
- Use R to compute both the ACF and PACF.

Things to Do to Prepare for Our Next Class

- **Required:** Complete assignment12.