Lecture 7
Time Series Analysis

Dr. Lydia Wahid

Agenda

- Time Series Analysis Introduction
- Time Series Analysis Techniques
- Exponential Smoothing
- Auto Regression
- Moving Average Model
- ACF
- PACF
- ARMA
- ARIMA

Time Series Analysis Introduction

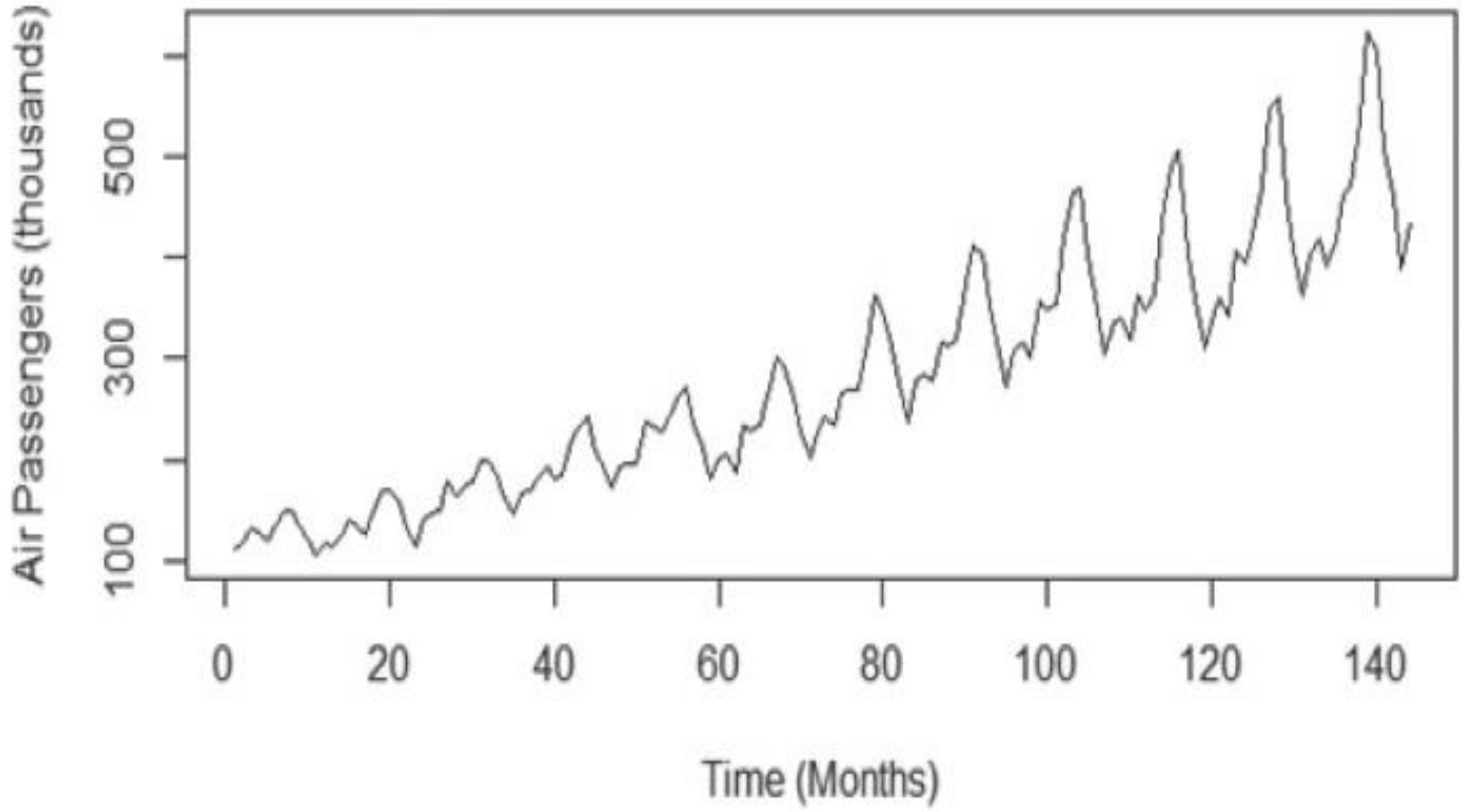
Time Series Analysis Introduction

- **Time series analysis** attempts to model the underlying structure of observations taken over time.

- A ***time series*** is an ordered sequence of equally spaced values over time.

- For example, the following figure provides a plot of the monthly number of international airline passengers over a 12-year period:

Time Series Analysis Introduction



In this example, the time series consists of an ordered sequence of 144 values.

Time Series Analysis Introduction

- Following are the goals of Time Series Analysis:
  - Identify and model the structure of the time series.
  - Forecast future values in the time series.

- Time series analysis has many applications in finance, economics, biology, engineering, retail, and manufacturing.

Time Series Analysis Introduction

- **Univariate Time Series** consists of the values taken by a *single* variable at periodic time instances over a period.
- **Multivariate Time Series** consists of the values taken by *multiple* variables at the same periodic time instances over a period.
- A time series can consist of the following components:
  - Trend
  - Seasonality
  - Cyclic
  - Random

Time Series Analysis Introduction

- ***Trend*** refers to the long-term movement in a time series.
  - It indicates whether the observation values are increasing or decreasing over time.
  - Examples of trends are a steady increase in sales month over month.
- ***Seasonality*** component describes the fixed, periodic fluctuation in the observations over time.
  - As the name suggests, the seasonality component is often related to the calendar.
  - For example, monthly retail sales can fluctuate over the year due to the weather and holidays.

Time Series Analysis Introduction

- ***Cyclic*** component also refers to a periodic fluctuation, but one that is not as fixed as in the case of a seasonality component.
  - For example, retails sales are influenced by the general state of the economy.
  - Thus, a retail sales time series can often follow the lengthy boom-bust cycles of the economy.
- ***Random*** component refers to random fluctuations.
  - These fluctuations are unforeseen, uncontrollable and unpredictable.
  - For example: earthquakes, wars, floods.

Time Series Analysis Techniques

Time Series Analysis Techniques

- **Naïve Methods:** These are simple estimation techniques, such as the predicted value is given the value equal to **mean of preceding values** of the time dependent variable, or **previous actual value.**

- **Exponential Smoothing:** Exponential smoothing uses a **weighted average of past time series values** as a forecast.

- **Auto Regression:** Auto regression predicts the values of future time periods as a **function of values at previous time periods**.

Time Series Analysis Techniques

- **Moving Average Model:** The value of a time series is a linear function of the **errors at previous time steps** of a stationary timeseries.

- **ARMA Model:** **A**uto-**R**egressive **M**oving-**A**verage models the value of a time series as a linear function of previous values and errors at previous time steps of a stationary timeseries.

- **ARIMA Model:** **A**uto-**R**egressive **I**ntegrated **M**oving-**A**verage is similar to ARMA with *differencing* applied to use it for nonstationary time series.

Time Series Analysis: Exponential Smoothing

- Exponential smoothing uses a weighted average of past time series values as a forecast.
- The weights become smaller as the observations move farther into the past. The exponential smoothing equation follows:

$$F_{t+1} = \alpha Y_t + (1 - \alpha)F_t$$

where

F_{t+1} = forecast of the time series for period $t + 1$
Y_t = actual value of the time series in period t
F_t = forecast of the time series for period t
α = smoothing constant ($0 \leq \alpha \leq 1$)

Time Series Analysis: Exponential Smoothing

- The forecast for period t + 1 is a weighted average of the **actual value in period t** and the **forecast for period t**.

- The weight given to the actual value in period t is the **smoothing constant α** and the weight given to the forecast in period t is **$1 - \alpha$**.

- **The exponential smoothing forecast for any period is actually a weighted average of all the previous actual values of the time series.**

Time Series Analysis: Exponential Smoothing

- The general equation will be:

$$\hat{y}_{T+1|T} = \sum_{j=0}^{T-1} \alpha(1-\alpha)^j y_{T-j} + (1-\alpha)^T \ell_0.$$

The process has to start somewhere, so we let the first fitted value at time 1 be denoted by ℓ_0

- The last term becomes tiny for large T. So, the weighted average form leads to the forecast Equation:

$$\hat{y}_{T+1|T} = \alpha y_T + \alpha(1-\alpha) y_{T-1} + \alpha(1-\alpha)^2 y_{T-2} + \cdots,$$

Time Series Analysis: Exponential Smoothing

- Let us illustrate by working with a time series involving only three periods of data: Y1, Y2, and Y3.
- To initiate the calculations, we let F1 equal the actual value of the time series in period 1; that is, **F1 = Y1**. Hence, the forecast for period 2 is:

$$\begin{aligned} F_2 &= \alpha Y_1 + (1 - \alpha)F_1 \\ &= \alpha Y_1 + (1 - \alpha)Y_1 \\ &= Y_1 \end{aligned}$$

Time Series Analysis: Exponential Smoothing

- The forecast for period 3 is:

$$F_3 = \alpha Y_2 + (1-\alpha)F_2 = \alpha Y_2 + (1-\alpha)Y_1$$

- Finally, substituting this expression for F3 in the expression for F4, we obtain:

$$\begin{aligned} F_4 &= \alpha Y_3 + (1-\alpha)F_3 \\ &= \alpha Y_3 + (1-\alpha)[\alpha Y_2 + (1-\alpha)Y_1] \\ &= \alpha Y_3 + \alpha(1-\alpha)Y_2 + (1-\alpha)^2 Y_1 \end{aligned}$$

Time Series Analysis: Exponential Smoothing

➢ **Example:** Consider the following data:

Week	Time Series Value
1	17
2	21
3	19
4	23
5	18
6	16
7	20

- To start the calculations we set the exponential smoothing forecast for period 2 equal to the actual value of the time series in period 1.
- Thus, with **Y1 = 17**, we set **F2 = 17** to initiate the computations.
- Referring to the time series data, we find an actual time series value in period 2 of **Y2 = 21**. Thus, period 2 has a forecast error of 21 – 17 = 4.

Time Series Analysis: Exponential Smoothing

- **Example:** Continuing with the exponential smoothing computations using a smoothing constant of α = 0.2, we obtain the following forecast for period 3:

$$F_3 = .2Y_2 + .8F_2 = .2(21) + .8(17) = 17.8$$

- Once the actual time series value in period 3, Y3 = 19, is known, we can generate a forecast for period 4 as follows:

$$F_4 = .2Y_3 + .8F_3 = .2(19) + .8(17.8) = 18.04$$

- …and so on.

Time Series Analysis: Exponential Smoothing

- **How to choose the value of the smoothing constant α?**
  - Although any value of a between 0 and 1 is acceptable, some values will yield better forecasts than others.
  - The new forecast F_{t+1} is equal to the previous forecast F_t plus an adjustment, which is the smoothing constant α times the most recent forecast error, $Y_t – F_t$.
  - That is, the forecast in period t + 1 is obtained by adjusting the forecast in period t by a fraction of the forecast error.

Time Series Analysis: Exponential Smoothing

- **How to choose the value of the smoothing constant α?**
  - If the time series contains large random variability, a small value of the smoothing constant is preferred.
  - For a time series with relatively little random variability, forecast errors are more likely to represent a change in the level of the series. Thus, larger values of the smoothing constant provide the advantage of quickly adjusting the forecasts; this allows the forecasts to react more quickly to changing conditions.
  - We will choose the value of α based on the **Mean Squared Error (MSE)**.

Time Series Analysis: Exponential Smoothing

- The table shows the MSE for α=0.2

Sales Time Series with Smoothing Constant $\alpha = .2$

Week	Time Series Value	Forecast	Forecast Error	Squared Forecast Error
1	17			
2	21	17.00	4.00	16.00
3	19	17.80	1.20	1.44
4	23	18.04	4.96	24.60
5	18	19.03	−1.03	1.06
6	16	18.83	−2.83	8.01
7	20	18.26	1.74	3.03
8	18	18.61	−.61	.37
9	22	18.49	3.51	12.32
10	20	19.19	.81	.66
11	15	19.35	−4.35	18.92
12	22	18.48	3.52	12.39
		Totals	10.92	98.80

Time Series Analysis: Exponential Smoothing

- **How to choose the value of the smoothing constant α?**
  - Would a different value of α provide better results in terms of a lower MSE value?
  - The most straightforward way to answer this question is simply to try another value for α

Time Series Analysis: Exponential Smoothing

- The table shows the MSE for α=0.3
- Thus, after trying α=0.2 and α=0.3, α=0.2 is preferred since it yields lower MSE.

Sales Time Series with Smoothing Constant $\alpha = .3$

Week	Time Series Value	Forecast	Forecast Error	Squared Forecast Error
1	17			
2	21	17.00	4.00	16.00
3	19	18.20	.80	.64
4	23	18.44	4.56	20.79
5	18	19.81	−1.81	3.28
6	16	19.27	−3.27	10.69
7	20	18.29	1.71	2.92
8	18	18.80	−.80	.64
9	22	18.56	3.44	11.83
10	20	19.59	.41	.17
11	15	19.71	−4.71	22.18
12	22	18.30	3.70	13.69
		Totals	8.03	102.83

Time Series Analysis: Auto Regression

- For a stationary time series, *yt*, an ***autoregressive model of order p***, denoted AR(p), is expressed as shown in the equation:

$$y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + \varepsilon_t$$

where δ is a constant for a nonzero-centered time series:

ϕ_j is a constant for j = 1, 2, . . ., p

y_{t-j} is the value of the time series at time $t - j$

$\phi_p \neq 0$

$\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ for all t

A particular point in the time series can be expressed as a **linear combination of the prior p values** of the time series plus a **random error term**.

Time Series Analysis: Auto Regression

- $\mathcal{E}_t$ is often called a ***white noise process*** and is used to represent random, independent fluctuations that are part of the time series.

- **Stationary Time Series** means that behavior of a system is characterized by **non**-changing statistical properties over **time** such as the mean, variance and autocorrelation.

- **Time series** are **stationary if** they do not have trend or seasonal effects.

Time Series Analysis: **Auto Regression**

- Different methods can be used for parameters estimations such as:
  - Maximum-Likelihood estimation
  - Expectation Maximization (EM) algorithm
  - Kalman Filter

Time Series Analysis: Auto Regression

- Once the parameters of the autoregression have been estimated, the autoregression can be used to forecast an arbitrary number of periods into the future.

- First use t to refer to the first period for which data is not yet available; substitute the known preceding values into the autoregressive equation.

- Next, use t to refer to the next period for which data is not yet available; again the autoregressive equation is used to make the forecast.

Time Series Analysis: Moving Average Model

- For a time series, *yt*, a ***moving average model of order q***, denoted MA(q), is expressed as shown in the following equation:

$$y_t = c + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q},$$

where C is a constant

ε_{t-j} is the value of error at time series y_{t-j}

θ_j *for* $j=1,2,..q$ are constants

The value of a time series **is a linear combination** of the **current white noise** term and the **prior q white noise** terms.

Stationary Time Series

- It is necessary to remove any trends or seasonality in the time series. This step is necessary to achieve a time series with certain properties to which autoregressive and moving average models can be applied.
- Such a time series is known as a **stationary time series**. A time series, y_t for t=1,2,3,.. is a stationary time series if the following three conditions are met:
  a. The expected value (mean) of y_t is a constant for all values of t.
  b. The variance of y_t is finite. $\sigma^2 = \frac{\Sigma(x_i - \mu)^2}{n}$
  c. The covariance of y_t and y_{t+h} depends only on the value of h= 0,1,2, . . .for all t (not t).

Stationary Time Series

- **Covariance:**
  - The covariance of y_t and y_{t+h} is a measure of how the two variables, vary together.
  - If two variables are independent of each other, their **covariance is zero.**
  - If the variables change together in the same direction, the variables have a **positive covariance**.
  - Conversely, if the variables change together in the opposite direction, the variables have a **negative covariance**.

Stationary Time Series

- **Covariance:**
  - For a stationary time series, by condition (a), the mean is a constant, the covariance notation is:

$$cov(h) = E[(y_t - \mu)(y_{t+h} - \mu)]$$

  - By part (c), the covariance between two points in the time series can be nonzero, as long as the value of the covariance is only a function of h. An example for h= 3:

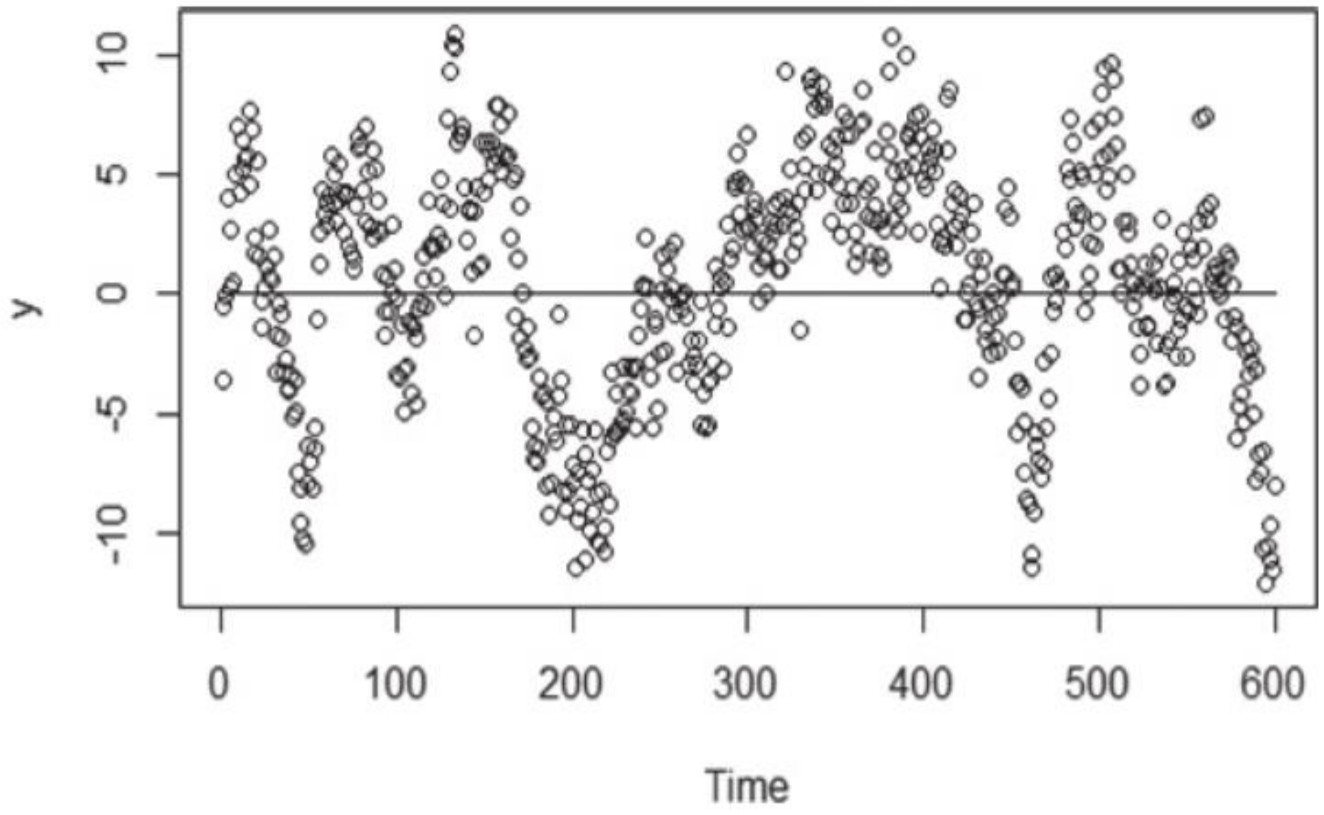
$$cov(3) = cov(y_1, y_4) = cov(y_2, y_5) = \ldots$$

Autocorrelation Function (ACF)

- **Covariance VS Correlation**
  - Covariance is a quantitative measure of the degree to which the deviation of one variable from its mean is related to the deviation of another variable from its mean.
  - Correlation quantifies the **strength** and **direction** of relationship between two variables (normalized covariance).
  - Correlation can only take values between +1 and -1. A correlation of +1 indicates that random variables have a direct and strong relationship.

Autocorrelation Function (ACF)

- It appears that each point is somewhat dependent on the past points.
- The difficulty is that the plot does not provide insight into the covariance of the variables in the time series and its underlying structure.



Autocorrelation Function (ACF)

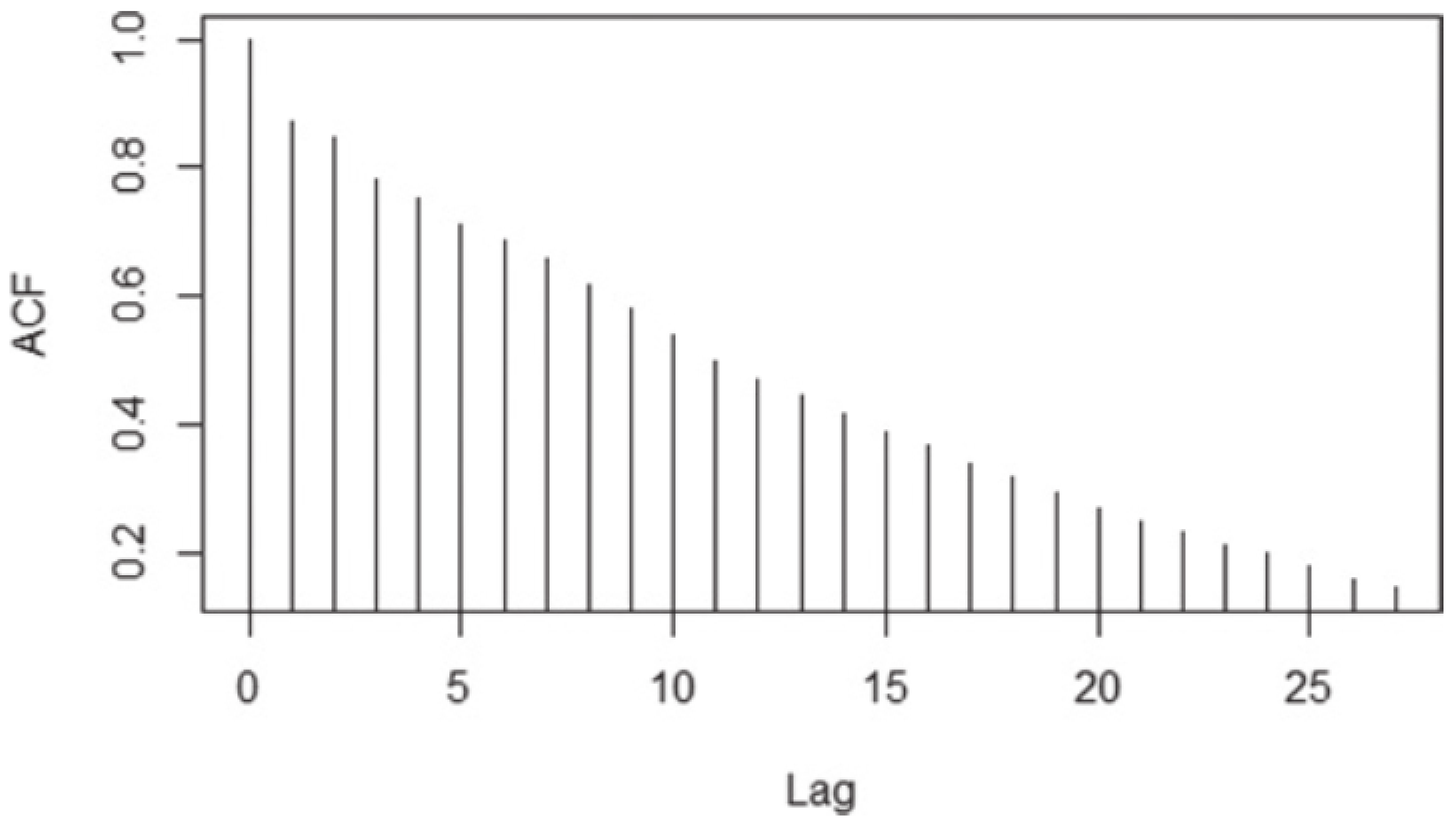
- The plot of autocorrelation function (ACF) provides this insight. For a stationary time series, the ACF is defined as:

$$ACF(h) = \frac{cov(y_t, y_{t+h})}{\sqrt{cov(y_t, y_t)\,cov(y_{t+h}, y_{t+h})}} = \frac{cov(h)}{cov(0)}$$

- Because the cov(0) is the variance, the ACF is analogous to the correlation function of two variables and the value of the ACF falls between –1 and 1.
- Thus, the closer the absolute value of ACF(h) is to 1, the more useful y_t can be a predictor of y_{t+h}

Autocorrelation Function (ACF)

- Using the same dataset, the plot of the ACF is as shown in the figure.
- By convention, the quantity h in the ACF is referred to as the lag, the difference between the time points t and t + h.



Autocorrelation Function (ACF)

- At lag 0, the ACF provides the correlation of every point with itself. So ACF(0) always equals 1.
- According to the ACF plot, at lag 1 the correlation between y_t and y_{t+h} is approximately 0.9, which is very close to 1. So y_{t-1} appears to be a good predictor of the value of y_t..
- Because ACF(2) is around 0.8, y_{t-2} also appears to be a good predictor of the value of y_t.
- A similar argument could be made for lag 3 to lag 8. (All the autocorrelations are greater than 0.6.)
- In other words, a model can be considered that would express yt as a linear sum of its previous 8 terms. Such a model is known as an autoregressive model of order 8.

Partial autocorrelation function (PACF)

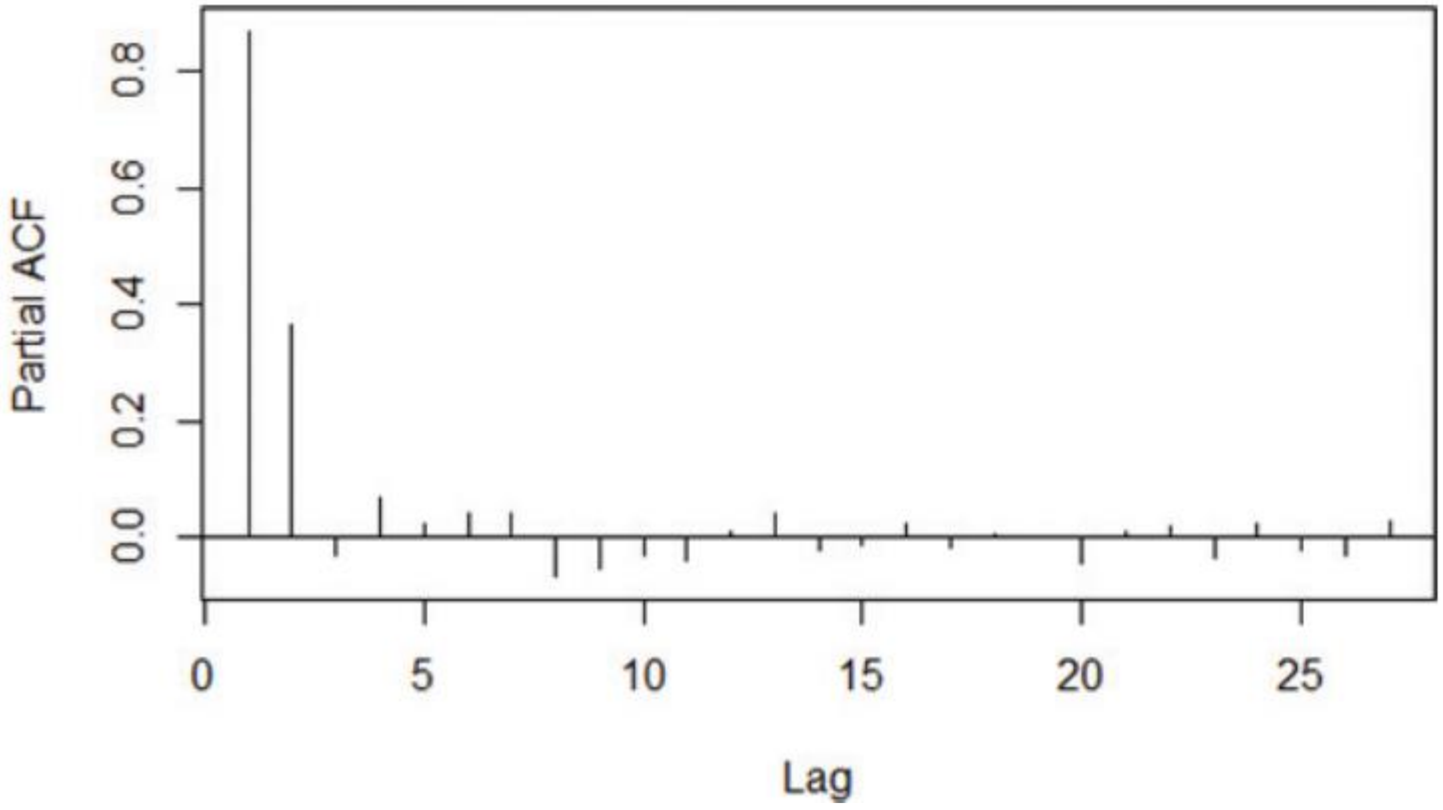
- If y_t and y_{t-1} are correlated, then y_{t-1} and y_{t-2} are also correlated. y_t and y_{t-2} might be correlated because they are both connected to y_{t-1} not because y_{t-2} contains information that can be used to forecast yt

- What is needed is a measure of the autocorrelation between yt and yt+h for h = 1, 2, 3... with the effect of y_{t+1} to y_{t+h-1} values excluded from the measure.

- The partial autocorrelation function (PACF) provides such a measure.

Partial autocorrelation function (PACF)

- For the earlier example, the PACF plot illustrates that after lag 2, the value of the PACF is sharply reduced.
- Such a plot indicates that an **AR(2)** is a good candidate model for the time series given.



Time Series Analysis: **ARMA**

- In fact, it is often useful to combine the two representations AR(p) and an MA(q) into one model.

- The combination of these two models for a stationary time series results in an ***Autoregressive Moving Average model, ARMA(p,q)***, which is expressed as shown in the following equation:

Time Series Analysis: **ARMA**

$$y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p}$$
$$+ \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q}$$

where δ is a constant for a nonzero-centered time series

ϕ_j is a constant for j = 1, 2, ..., p

$\phi_p \neq 0$

θ_k is a constant for k = 1, 2, ..., q

$\theta_q \neq 0$

$\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ for all t

Time Series Analysis: **ARIMA**

- Time series may not be stationary.
- Since such a time series does not meet the requirement of a constant expected value (mean), the data needs to be adjusted.
- An option is to compute the difference between successive y-values. This is known as ***differencing***.
- In other words, for the n values in a given time series compute the differences as shown in the following equation:

$$d_t = y_t - y_{t-1} \quad \text{for } t = 2, 3, \ldots, n$$

Time Series Analysis: **ARIMA**

- It is often necessary to account for seasonal patterns in time series. The seasonal pattern could be determined, and the time series appropriately adjusted.
- An alternative is to use a seasonal autoregressive integrated moving average model, denoted ARIMA(p,d,q) × (P,D,Q)s where:
  - p, d, and q are the same as defined previously.
  - s denotes the seasonal period.
  - P is the number of terms in the AR model across the s periods.
  - D is the number of differences applied across the s periods.
  - Q is the number of terms in the MA model across the s periods.

Time Series Analysis: **ARIMA**

- For a time series with a seasonal pattern, following are typical values of s:
  - 52 for weekly data
  - 12 for monthly data
  - 7 for daily data

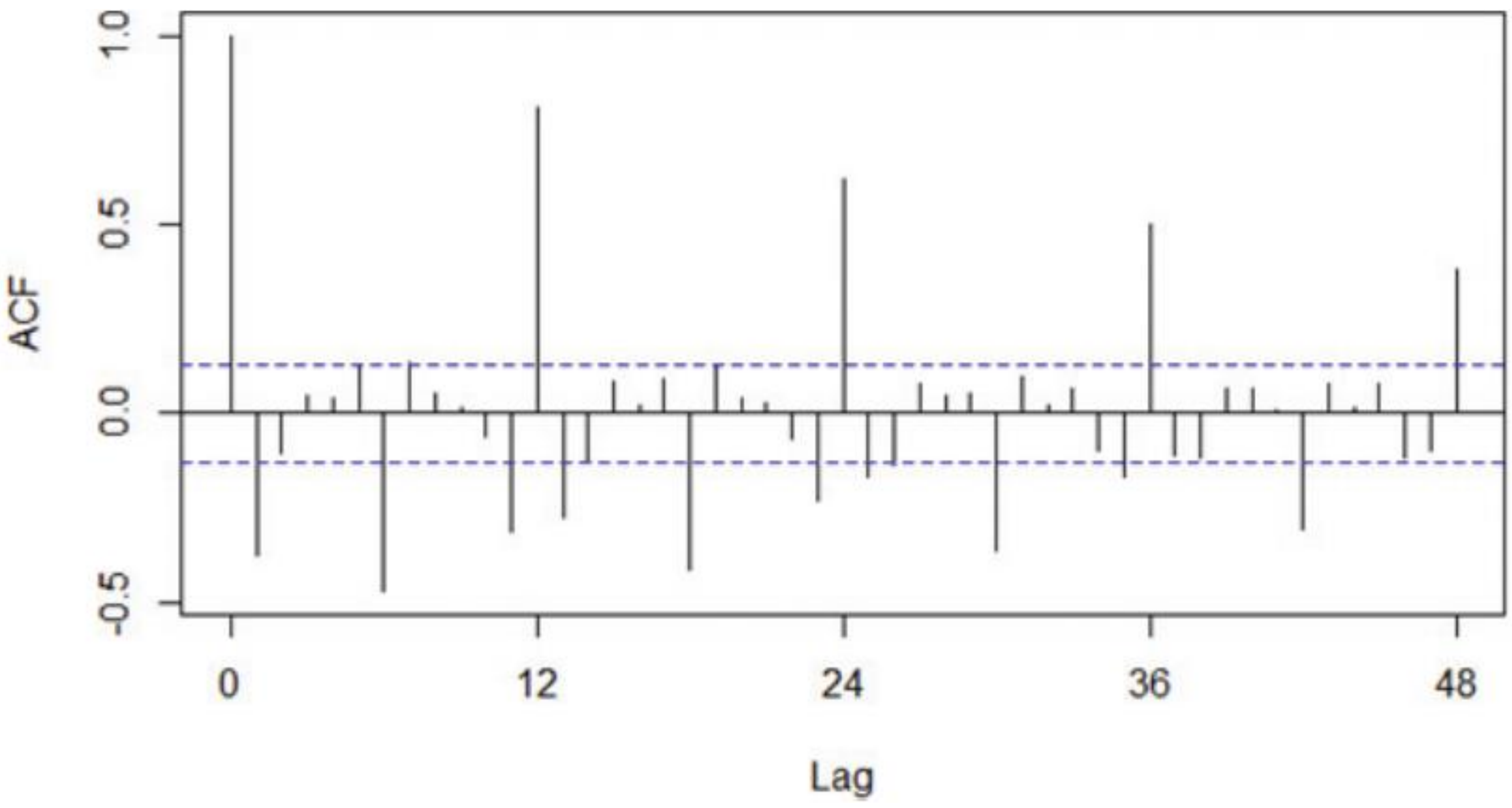
Time Series Analysis: **ARIMA**

- For a large country, the monthly gasoline production measured in millions of barrels has been obtained for the past 240 months (20 years).

- A market research firm requires some short-term gasoline production forecasts to assess the petroleum industry's ability to deliver future gasoline supplies and the effect on gasoline prices.

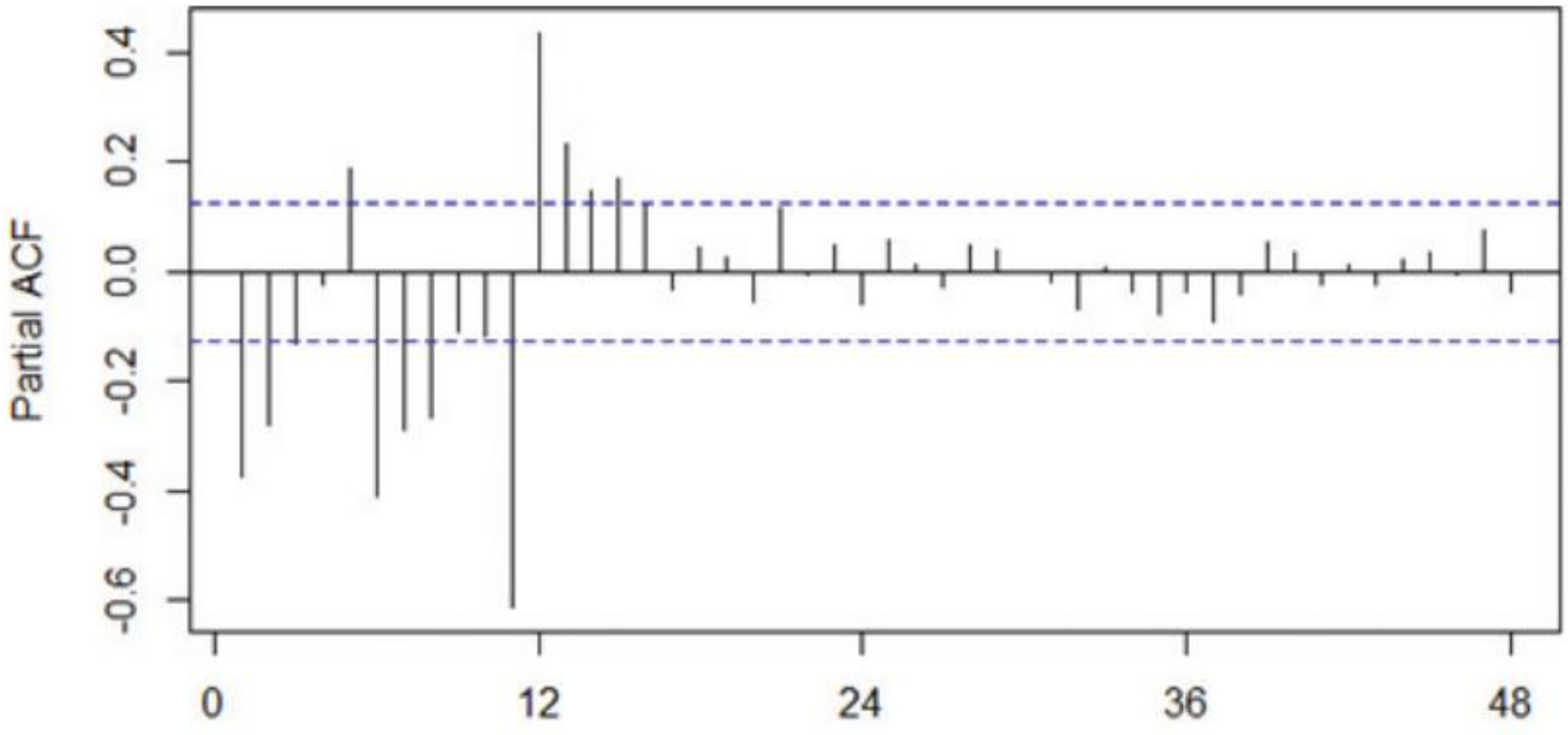
Time Series Analysis: **ARIMA**

- The figure shows several significant ACF values. The slowly decaying ACF values at lags 12, 24, 36, and 48 are of particular interest.
- The figure indicates a seasonal autoregressive pattern every 12 months.



Time Series Analysis: **ARIMA**

- In the PACF plot, value at lag 12 is quite large, but the PACF values are close to zero at lags 24, 36, and 48.
- Thus, a seasonal AR(1) model with period = 12 will be considered.



Time Series Analysis: **ARIMA**

- The arima() function in R uses Maximum Likelihood Estimation (MLE) to estimate the model coefficients.
- The values of the model coefficients are determined such that the value of the log likelihood function is maximized.
- The R output provides several measures that are useful for comparing the appropriateness of one fitted model against another fitted model. These measures are:
  - AIC (Akaike Information Criterion)
  - AICc (Akaike Information Criterion, corrected)
  - BIC (Bayesian Information Criterion)

Time Series Analysis: **ARIMA**

- Because these criteria impose a penalty based on the number of parameters included in the models, the preferred model is the fitted model with the smallest AIC, AICc, or BIC value.

- The following table provides the information criteria measures for the ARIMA fitted models of the gasoline problem.

Time Series Analysis: ARIMA

ARIMA Model $(p,d,q) \times (P,Q,D)_S$	AIC	AICc	BIC
$(0,1,0) \times (1,0,0)_{12}$	1561.38	1561.43	1568.33
$(0,1,1) \times (1,0,0)_{12}$	1472.43	1472.53	1482.86
$(0,1,2) \times (1,0,0)_{12}$	1474.25	1474.42	1488.16
$(1,1,0) \times (1,0,0)_{12}$	1504.29	1504.39	1514.72
$(1,1,1) \times (1,0,0)_{12}$	1474.22	1474.39	1488.12

In this dataset, the $(0,1,1) \times (1,0,0)_{12}$ model does have the lowest AIC, AICc, and BIC values compared to the same criterion measures for the other ARIMA models.

Time Series Analysis: Box-Jenkins

- Developed by George Box and Gwilym Jenkins, is the **Box-Jenkins** methodology for time series analysis.
- Box-Jenkins methodology is used to apply an **ARIMA** (or **ARMA**) model to a given time series.
- The Box-Jenkins method applies iteratively the following three steps:
  1. Identification
  2. Estimation
  3. Diagnostic Checking

Time Series Analysis: **Box-Jenkins**

1. **Identification:**
   - Assess whether the time series is stationary, and if not, make it stationary.
   - Identify the parameters of an ARIMA model for the data (i.e. **p** and **q**)
     - Two diagnostic plots can be used to help choose the p and q: **Autocorrelation Function** and **Partial Autocorrelation Function**
2. **Estimation:**
   - Calculate the coefficients of the model.
3. **Diagnostic Checking:**
   - Evaluate the fitted model in the context of the available data and check for areas where the model may be improved.

Thank You