AQ061-3-M-ODL-TSF Time Series Analysis and Forecasting

Topic 1 – Characteristic of Time Series (Part II)

TOPIC LEARNING OUTCOMES

At the end of this topic, you should be able to:

1. have a broad understanding of time series characteristics, correlogram, autocorrelation, and stationary time series.
2. explain the characteristics of time series.
3. solve the model using computer software and interpret the results.

Contents & Structure

- Introduction
- Features of Time Series Data
- Time Series Plots
- Stationary Time Series

Recap From Last Lesson

- Questions to ask to trigger last week's key learning points

Time Series Plots - correlogram

- Correlogram /Autocorrelation Function (ACF)
  - It is a plot of graph of the autocorrelations for various lags of time series
  - Patterns in the correlogram is used to analyze the key feature of the data
    - Are the data random ?
    - Do the data have a trend ?
    - Are the data stationary ?
    - Are the data seasonal ?

Autocorrelation

- When a variable is measured over time, observations in different time periods are frequently related or correlated.
- This correlation is measured using the autocorrelation coefficient, ρ_k

Example

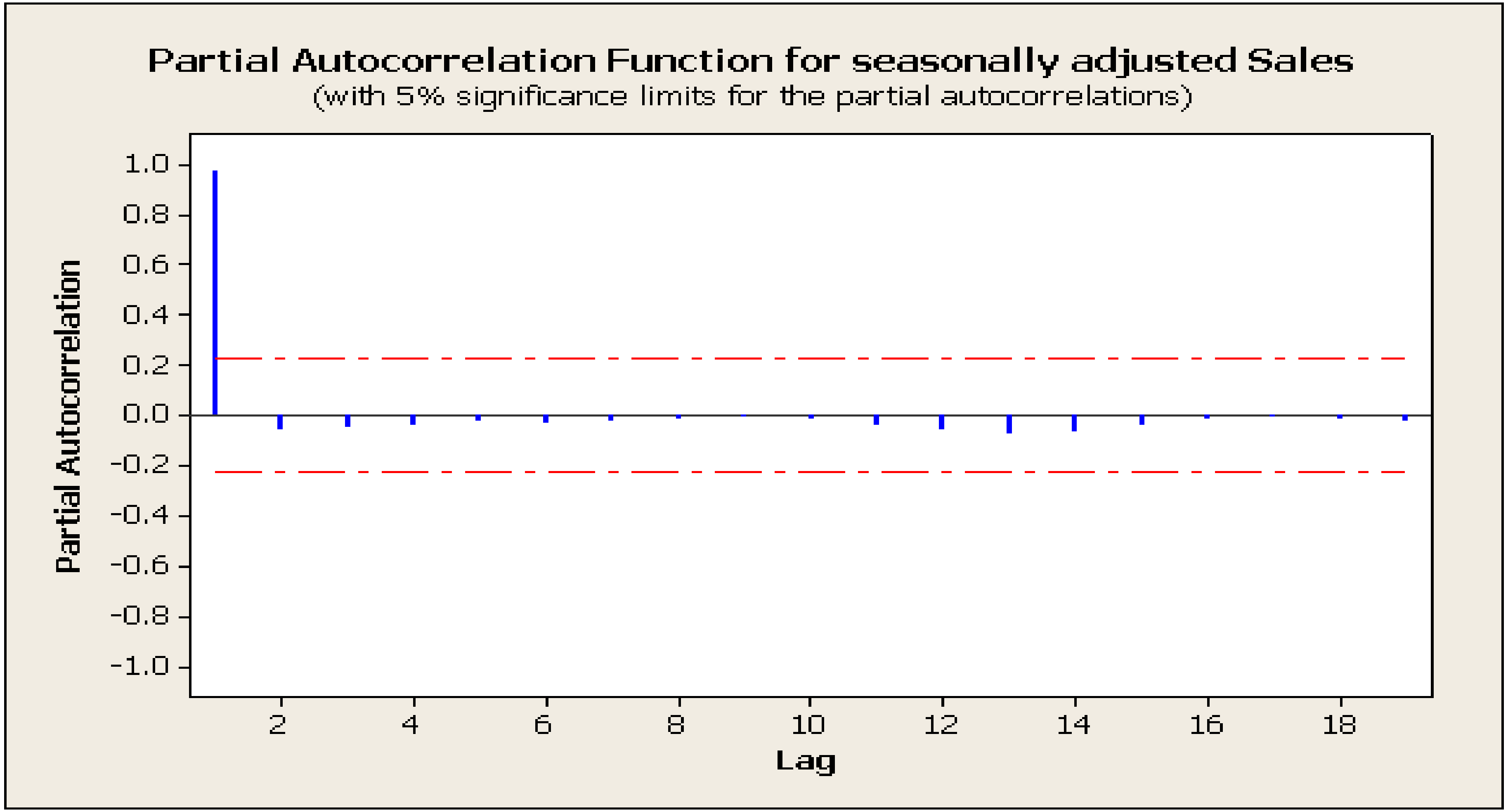
Time, t	Observed, Y_t	Lag 1, Y_{t-1}	Lag 2, Y_{t-2}
1	5		
2	8	5	
3	2	8	5
4	8	2	8
5	5	8	2
6	10	5	8
7	2	10	5
8	4	2	10
9	7	4	2

Autocorrelation, ρ_1

Autocorrelation, ρ_2

Stationary

- If a series is stationary, the value of ρ_k should diminish rapidly toward zero as k increases, generally after the second or third time lag.
- A stationary time series is one whose basic statistical properties such as the mean, variance, remain constant over time

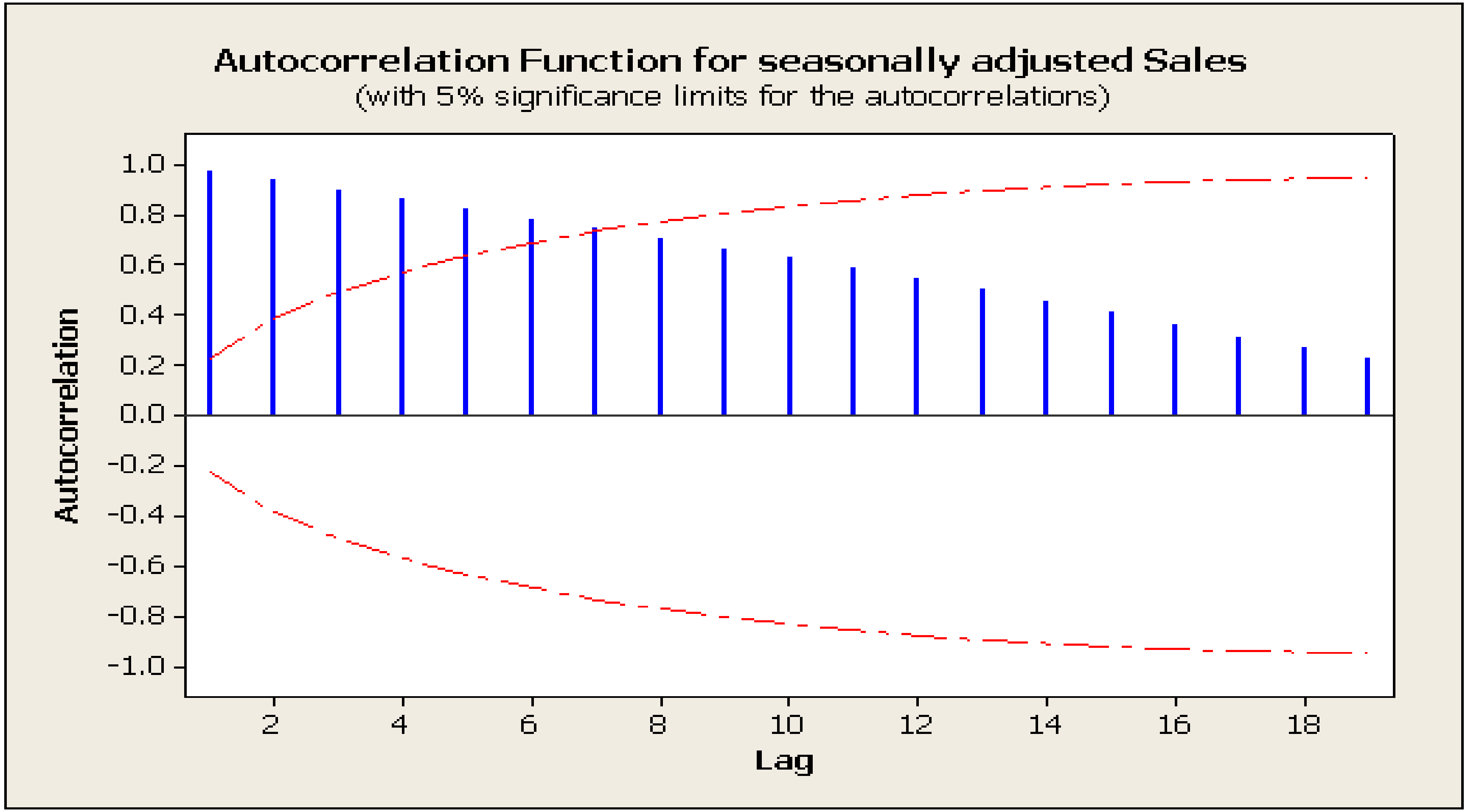


This is a typical of stationary series:

- Autocorrelation at time lag 1 is close to one and the autocorrelation for the time lag 2 through 18 are close to zero.

Trend

- If a series has a trend, Y_t and Y_{t-k} are highly correlated and the autocorrelation coefficients are typically significantly different from zero for the first several time lags and then gradually drop towards zero as the number of periods increases.
- The autocorrelation coefficient for time lag1 is often large (close to 1). The autocorrelation coefficient for time lag 2 will also be large but smaller than lag 1.

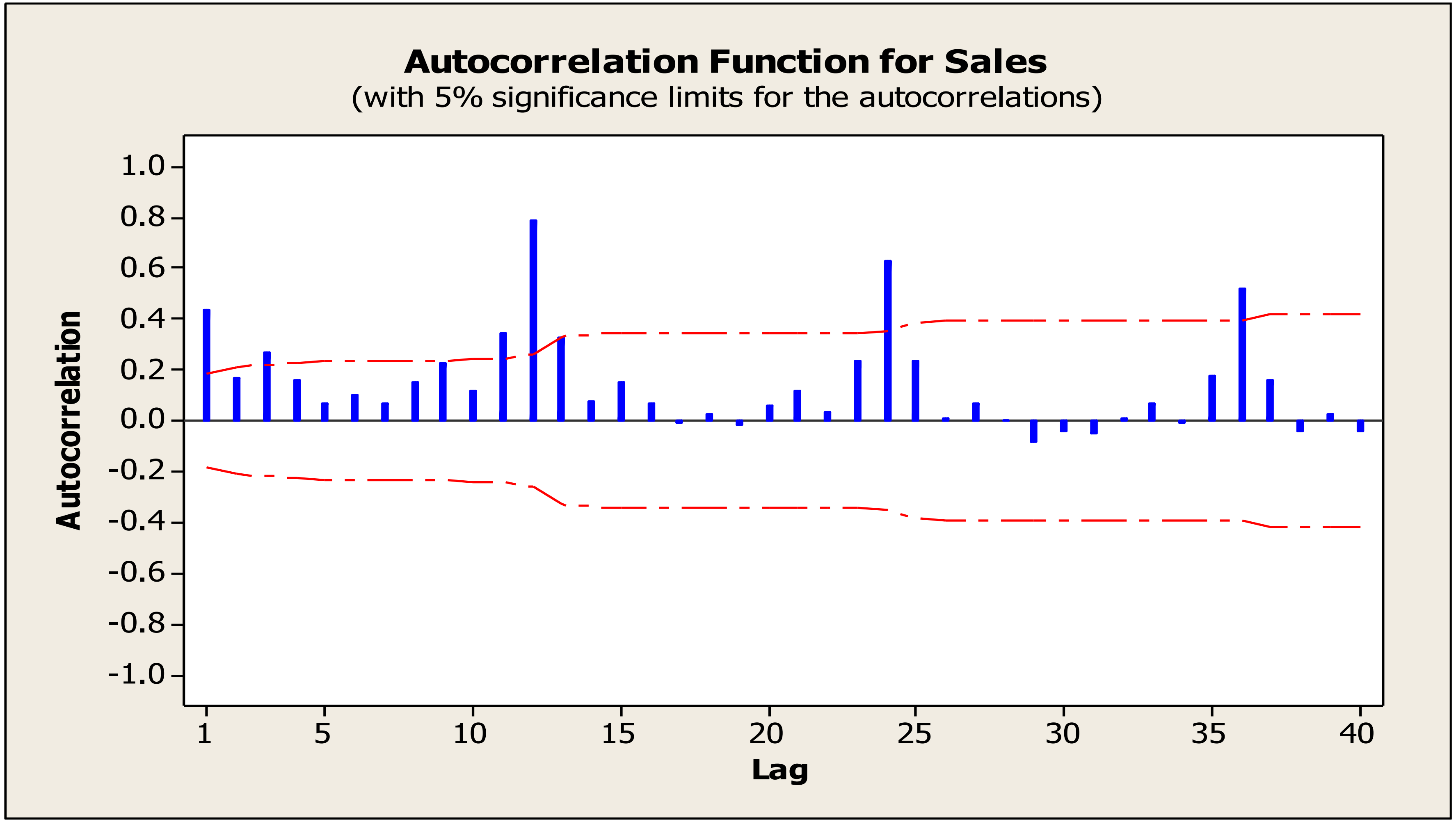


This is a typical for a non-stationary series:

- Large significant ACF for the first 7 time lag
- Slow decrease in the size of the autocorrelations.

Seasonal

- If a series has a seasonal pattern, a significant autocorrelation coefficient will occur at the seasonal time lag or multiples of the seasonal lag.
- The seasonal lag is 4 for quarterly data and 12 for monthly data.

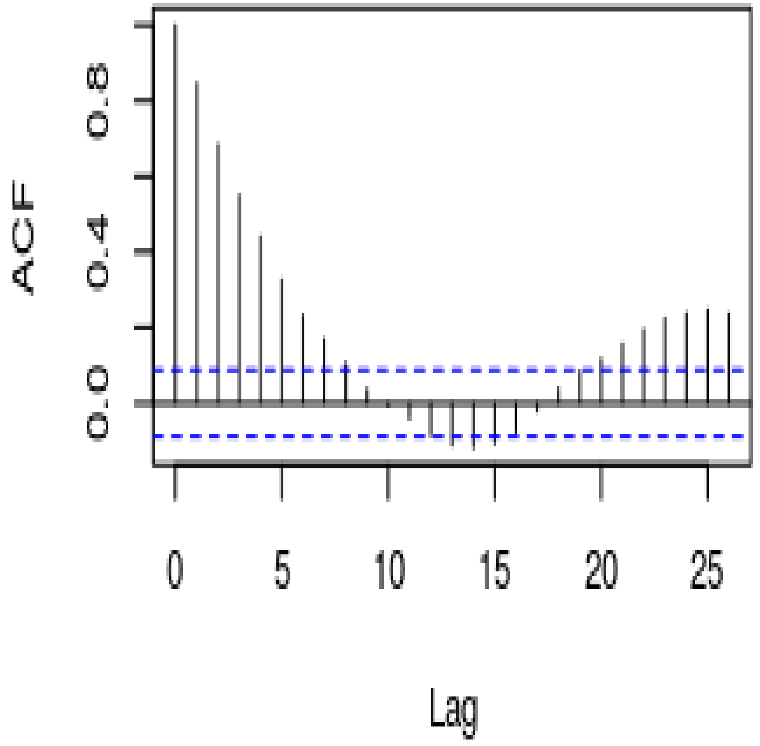


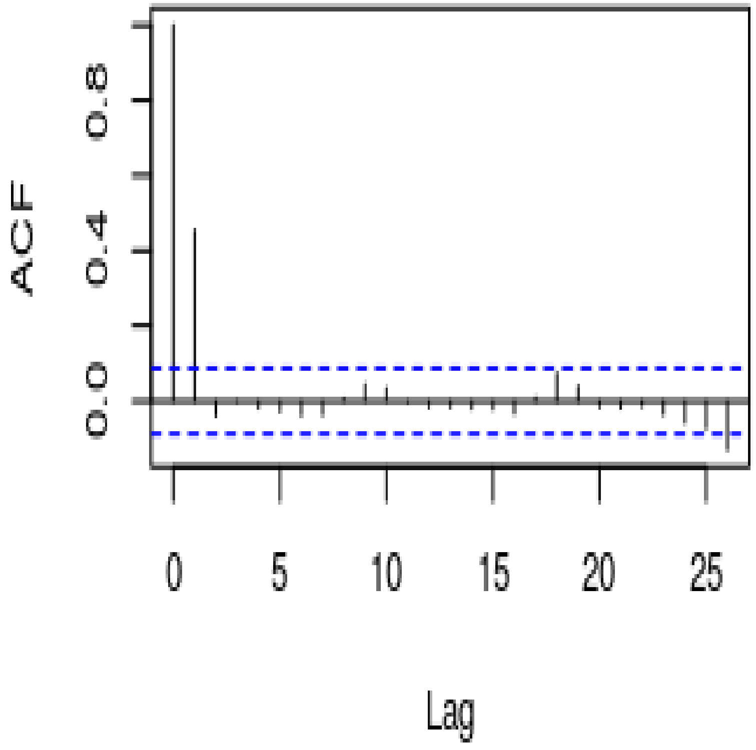
This is a typical for a non-stationary series:

- Significant ACF at the seasonality time lag (lag 12, 24, 36 etc) and slow decline in the autocorrelations.

Example

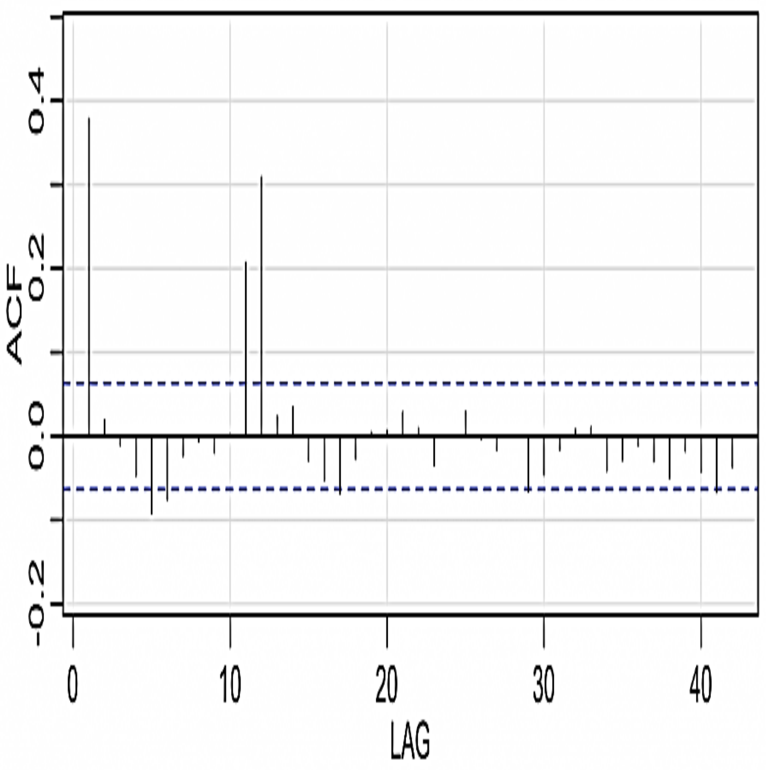
Check whether the following data is stationary.

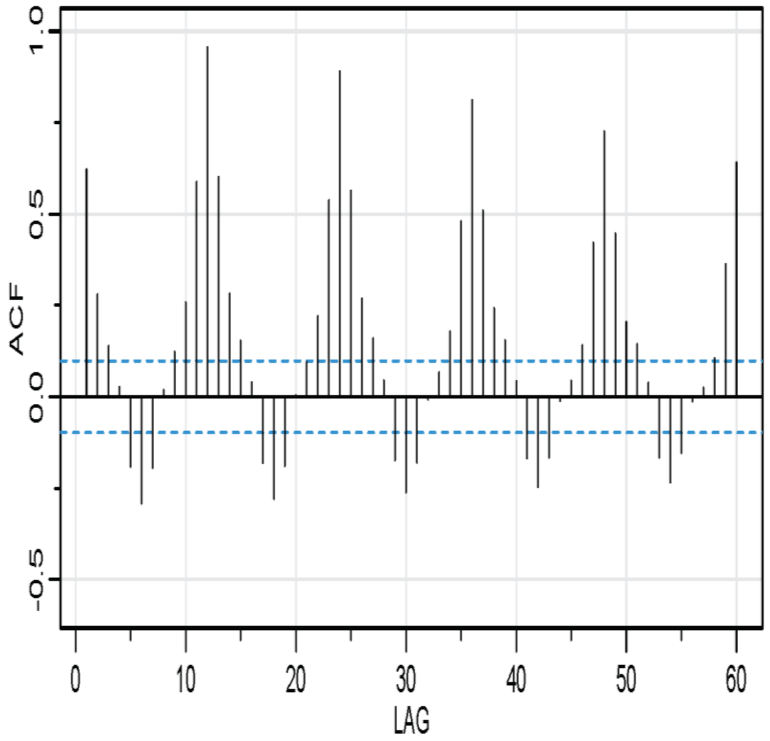




Example

Check whether the following data is stationary.





Coding Exercise

Construct ACF plot for the following data and check whether the series is stationary.

1. quakes.dat
2. sales.dat

Use **library(forecast)** and **acf(x, ylim=c(-1,1))**

Differencing

- Often non-stationary series can be made stationary through differencing.
- Differencing continues until stationary is achieved.
- First differencing:

$$\nabla y_t = y_t - y_{t-1}$$

- Second differencing:

$$\nabla^2 y_t = y_t - 2y_{t-1} + y_{t-2}$$

- In practice, it is almost never necessary to go beyond second difference, because real data generally involve only first or second level non-stationarity.

Seasonal Differencing

- Defined as a difference between a value and a value with lag that is a multiple of S
  - With S = 12, which occur with monthly data, a seasonal difference is

$$\nabla^{12} y_t = y_t - y_{t-12}$$

  - With S = 4, which occur with quarterly data, a seasonal difference is

$$\nabla^{4} y_t = y_t - y_{t-4}$$

- Differencing for trend and seasonality
  - When both trend and seasonality are present, we may need to apply both a non-seasonal first difference and a seasonal difference.

$$(y_t - y_{t-1}) - (y_{t-12} - y_{t-13})$$

Unit Root Test

- Also known as Augmented Dickey-Fuller (ADF) test.
- If a time series has a unit root, it shows a systematic pattern that is unpredictable.
- The null hypothesis is $\phi = 0$ (it contains a unit root, unpredictable systematic pattern) in which states that it is non-stationary against the alternative hypothesis ($\phi < 0$) which states that it does not contain a unit root (it is stationary)

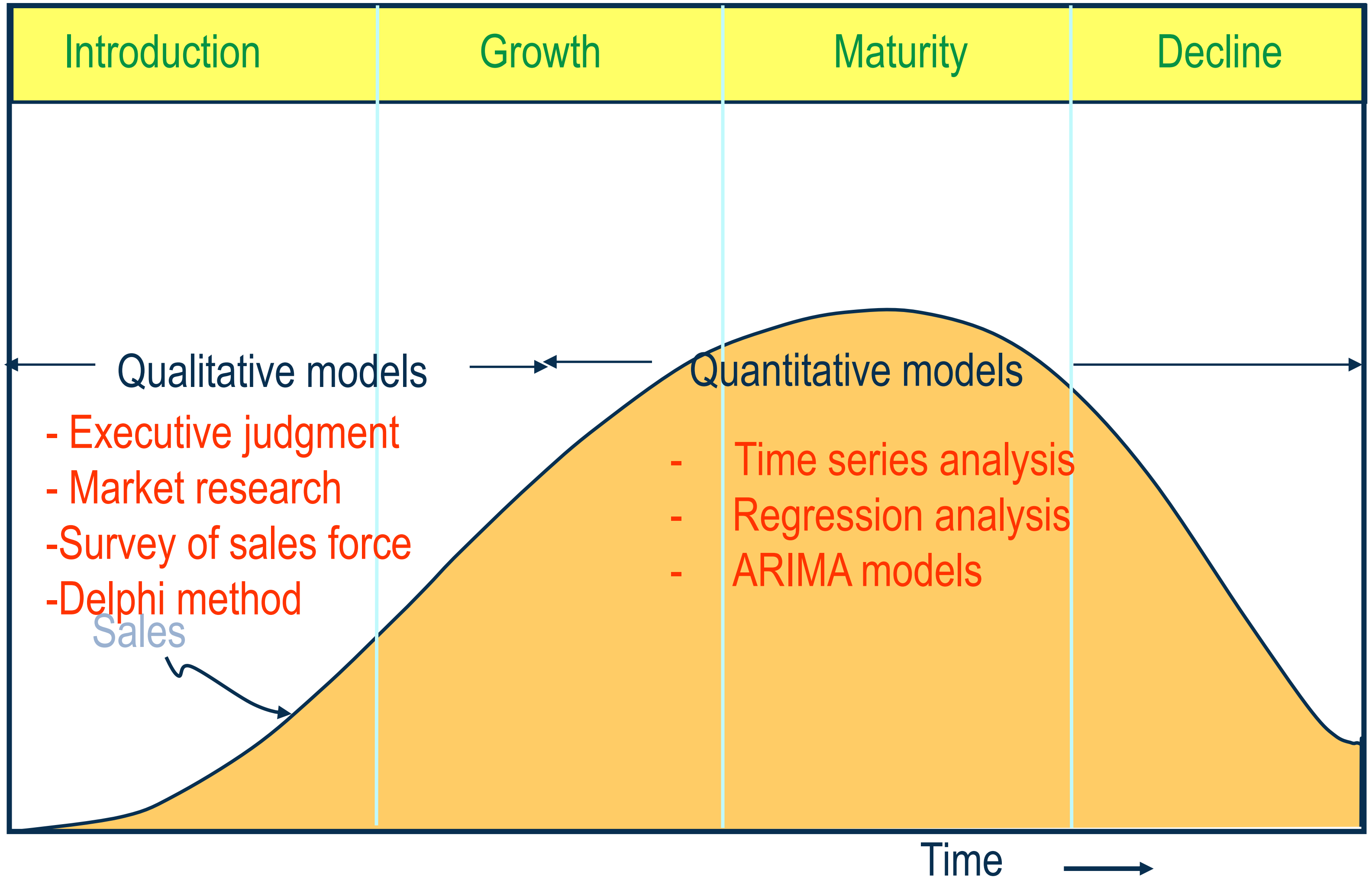
Coding Exercise

Apply Unit Root Test to test for stationary. Apply differencing if the data is not stationary.

1. quakes.dat
2. sales.dat

Use **library(tseries)**, **adf.test(y)** and **diff(x, difference)**

Forecasting During the Life Cycle



Types of Forecasting Methods

- Forecasting methods are classified into two groups:

	Qualitative Methods	Quantitative Methods
Characteristics	Based on human judgment, opinions; subjective and nonmathematical.	Based on mathematics, quantitative in nature
Strengths	Can incorporate latest changes in the environment and “inside information”	Consistent and objective; able to consider much information and data at one time.
Weakness	Can bias the forecast and reduce forecast accuracy.	Often quantitative data are not available. Only as good as the data on which they are based.

Qualitative Methods

Type	Characteristics	Strengths	Weaknesses
Executive opinion	A group of managers meet & come up with a forecast	Good for strategic or new-product forecasting	One person's opinion can dominate the forecast
Market research	Uses surveys & interviews to identify customer preferences	Good determinant of customer preferences	It can be difficult to develop a good questionnaire
Delphi method	Seeks to develop a consensus among a group of experts	Excellent for forecasting long-term product demand, technological changes, and scientific advances	Time consuming to develop

Delphi Method

Time Series Forecasting Process

Look at the data (Time Series Plot) → Forecast using one or more techniques → Evaluate the technique and pick the best one.

Observations from the Plot	Techniques to try	Ways to evaluate
Data is reasonably **stationary** (no trend or seasonality)	• Moving Averages • Simple Exponential Smoothing	• MAD • MAPE • MSE
Data shows a consistent **trend**	**Regression** • Linear, Quadratic, Exponential trend • Holt's method	• MAD • MAPE • MSE
Data shows both a **trend** and a **seasonal pattern**	**Decomposition** • Additive • Multiplicative **Holt's Winter method** **ARIMA** models	• MAD • MAPE • MSE • AIC & BIC
Data shows gradually **increasing variance** over time (**short periods of increased variation**)	• ARCH Models • GARCH Models	• MAD • MAPE • MSE • AIC & BIC

Consequences of incorrect Forecasting

- Overforecasting
  - (forecasting too much demand) results in overproduction, which has consequences of excessive inventory, which is demand
- Underforecasting
  - (not forecasting enough demand) results in underproduction, which has consequences of not meeting customer demand, and subsequently, the possibility of losing their business to competitors.

Seven Steps in the Forecasting System

- Determine the use of the forecast
- Select the items to be forecasted
- Determine the time horizon of the forecast
- Select the forecasting model(s)
- Gather the data needed to make the forecast
- Make the forecast
- Validate and implement the results

Review Questions

Summary / Recap of Main Points

- understanding of time series characteristics, correlogram, autocorrelation, and stationary time series.
- the characteristics of time series.
- solve the model using computer software and interpret the results.

What To Expect Next Week

In Class

Preparation for Class

Smoothing Techniques