Ministry of Higher Education and Scientific Research
University Center of El Bayadh
Institute of Economic Sciences, Commercial Sciences,
and Management Sciences

A Pedagogical Handbook on the Subject of :

Time Series Analysis 1

Intended for Third-Year Undergraduate Students – Specialization in Quantitative economics and Statistics

Prepared by:
Dr :Bourdache Chahrazed
Associate Professor-A-

Academic Year: 2025 – 2026

Table of contents

Table of contents

Introduction

Introduction

Time series analysis is considered one of the fundamental tools in the quantitative analysis of economic and statistical phenomena, due to its capacity to understand the behavior of variables over time, reveal their structural characteristics, and support forecasting and decision-making. This field has gained increasing importance in contemporary economic studies, owing to its wide application in analyzing both macroeconomic and microeconomic data, such as gross domestic product, prices, unemployment rates, demand, and other variables that exhibit a temporal dimension.

This pedagogical handbook aims to provide third-year Bachelor's students in Quantitative Economics and Statistics with the essential knowledge for the course Time Series Analysis 1. It offers a theoretical and methodological framework that allows students to understand the nature of time series and their components, as well as to identify the statistical properties governing their behavior. The course focuses on the essential aspects that enable students to analyze time-dependent data scientifically and accurately, linking theoretical concepts with practical applications relevant to the economic field.

The handbook begins by presenting the general concepts related to time series and their basic components, along with methods for estimating and analyzing these components. It then addresses stationarity, a fundamental prerequisite for time series analysis, and studies autocorrelation and partial autocorrelation, and their role in diagnosing a series' structure. Finally, exponential smoothing models are introduced as one of the most important methods used for forecasting time series, given their simplicity and effectiveness in many practical applications.

In preparing this handbook, a clear and progressive pedagogical approach has been adopted, taking into account the students' knowledge level, while emphasizing scientific accuracy and simplification of concepts as much as possible. The aim is to enable students to master the fundamentals of time series analysis and to apply them effectively in their academic and practical studies.

Chapter One: Fundamentals of Time Series and Their Components

The study and analysis of time series represent a fundamental topic in statistics, focusing on understanding and explaining the behavior of phenomena over specific time intervals (e.g., annual, semi-annual, quarterly). The objectives of time series analysis can be summarized as follows: first, to obtain a comprehensive description of the process that generates the time series; second, to construct a model that explains the behavior of the series and to use the results for forecasting its future evolution; and third, to exert control over the generating process by examining potential outcomes when certain model parameters are modified. Achieving these objectives requires a thorough analytical study of time series models, relying on both statistical and mathematical methods

I - Definition of a Time Series:

A time series is defined as a set of numerical observations indexed in time order, where the data points are inherently linked to the temporal dimension and thus vary over time. In time series analysis, time itself is treated as the independent variable used to explain the behavior of the phenomenon under study (the dependent variable), especially when explanatory variables are absent or when no data are available for these variables[1].

In contrast, in traditional linear regression models, the dependent variable is explained and its values are estimated based on one or more independent variables, under the assumption that all other influencing conditions remain constant[2]. However, when sufficient information on the explanatory variable(s) is unavailable, analysts must rely on alternative approaches to understand and model the behavior of the dependent variable in a time series context.

Two principal approaches are commonly adopted in such situations, which are discussed in the sections that follow:

---

[1] Chatfield C, Xing H, **The Analysis of Time Series: An Introduction with R**, 7th ed, Chapman & Hall/CRC, New York, 2025, p 7.

[2] Wooldridge J. M, **Introductory Econometrics: A Modern Approach**, 6th ed, Cengage Learning, 2016, p 120.

1- Using Time as an Independent Variable:

In this approach, time itself is treated as the independent variable to model and explain the phenomenon, usually through the **trend component**. The relationship can be expressed as:

$$y_i = f(t, \varepsilon_t)$$

where y_i represents the observed value at time t, and ε_t denotes the random error term, capturing unobserved influences.

2- Using Past Values of the Dependent Variable:

This approach relies on the historical behavior of the dependent variable to forecast its future values. It is typically implemented using regression-based models or moving average models, expressed as[1]:

$$y_t = f(y_{t-1}, y_{t-2}, y_{t-3}, \dots \dots .., y_{t-k}, \varepsilon_t)$$

Where:

$y_{t-1}, y_{t-2}, y_{t-3}, \dots \dots .., y_{t-k}$ are past values used to predict the current observation y_i.

Time series models are employed in several situations, including:

-When there is no causal relationship between variables:

Time series analysis allows for modeling and forecasting without relying on explicit explanatory variables [2] .

- When sufficient data on independent variables are unavailable:

In such cases, time itself or past values of the dependent variable can be used to predict future outcomes.

---

[1] Box G, Jenkins M, Reinsel C, Ljung G. M, **Time Series Analysis: Forecasting and Control** , 5th ed,Wiley, 2015, p 20.

[2] Chatfield C, Xing H, op.cit, p 15.

- When regression models are statistically, economically, or predictively weak:

Time series models can provide better insights in situations where classical regression models perform poorly, as measured by model indicators such as correlation and determination coefficients, statistical tests, and economic significance[1].

In certain forecasting situations, such as estimating the future volume of imports of a specific commodity, analysts may face the absence of a clear causal relationship among explanatory variables or insufficient data on independent variables traditionally used in economic theory (e.g., consumption, prices, consumer preferences). In such cases, forecasting relies primarily on the internal structure of the time series itself, and two main approaches are typically employed.

First Approach: Using Time as an Independent Variable

This approach treats **time** as an independent or explanatory variable, focusing on the **series' trend component**. The future values of the import volume can be modelled as:

$$y_t = \beta_0 + \beta_1 t + \varepsilon_t$$

where y_t represents the import volume at time t, β_0 and β_1 are parameters representing the intercept and trend slope, respectively, and ε_t is the error term capturing random fluctuations.

After estimating the parameters, the predicted future values are obtained as follows:

$$\hat{y}_{t+1} = \hat{\beta}_0 + \hat{\beta}_1 (t+1)$$
$$\hat{y}_{t+2} = \hat{\beta}_0 + \hat{\beta}_1 (t+2)$$
$$\vdots$$
$$\hat{y}_{t+k} = \hat{\beta}_0 + \hat{\beta}_1 (t+k)$$

[1] Wooldridge J. M, op.cit, p 125.

This approach is particularly effective when the primary driver of change is a long-term trend over time, and when external explanatory variables are unavailable or unreliable[1].

Second approach: Using Past Values of the Dependent Variable

Alternatively, future import volumes can be estimated from historical values of the dependent variable. This approach assumes that past values contain information about the series' dynamics and can be used to predict future outcomes. The model is expressed as:

$$y_t = f(y_{t-1}, y_{t-2}, y_{t-3}, \dots\dots\dots, y_{t-k}, \varepsilon_t)$$

II- Components of a Time Series:

A time series is generally composed of four main components, namely: the trend component, the seasonal component, the cyclical component, and the irregular (or random) component[2]:

1- Trend Component "T_t":

The trend component represents the long-term general direction of the phenomenon under study. It reflects the overall movement of the time series over an extended period and captures persistent changes, such as economic growth or structural transformations. In most practical applications, the trend component is approximated by a straight line, especially when the series' long-term evolution follows a roughly linear pattern. Statistically, the trend component is commonly expressed as a function of time:

$$X_t = \hat{\beta}_0 + \hat{\beta}_1 t$$

2- Seasonal Component "S_t ":

The seasonal component represents regular and systematic fluctuations that occur at fixed intervals within a year or over shorter periods. These variations arise from seasonal or periodic factors

[1] Chatfield C, Xing H, op.cit, p 12.
[2] Kendall M. G, Ord J. K, **Time Series** , 3rd ed, London, 1990, p p 11–13.

influenced by external conditions and tend to follow a recurring and predictable pattern. The seasonal component reflects short-term changes in the phenomenon under study, typically within a one-year horizon, and captures variations attributable to climatic, institutional, or habitual factors.

Common examples include household electricity consumption over a 24-hour period, agricultural production across seasons, and the consumption of certain beverages that varies according to seasonal patterns.

3- Cyclical Component "C_t":

The cyclical component reflects the effects of economic activity over the medium and long term. It captures fluctuations in the time series that are associated with the different phases of the business cycle, namely recession, recovery, expansion, and contraction. These cyclical movements represent oscillations around the trend component and recur over time, although they do not follow a fixed or strictly regular periodic pattern.

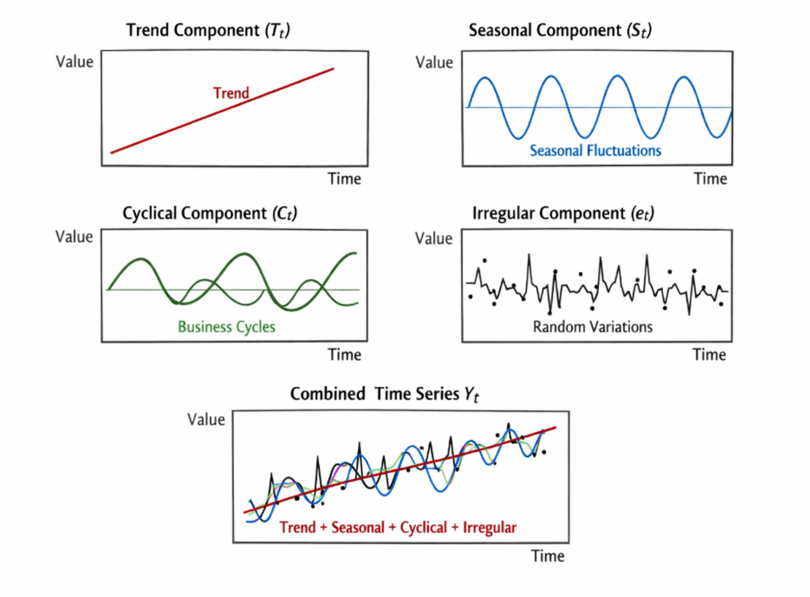
4- Irregular (Stochastic) Component "e_t":

The irregular, or stochastic, component represents variations that are difficult to control or regulate, as they arise from non-systematic and unpredictable factors. These fluctuations are characterized by their sudden occurrence and short duration, and they cannot be explained by the trend, seasonal, or cyclical components. Due to their random nature and limited analytical significance, the effect of this component can often be removed from the time series data in order to obtain a series free from irregular variations.

Typical examples include sudden declines in production resulting from failures in production facilities, unexpected disruptions, or other unforeseen events

The following figure summarizes the graphical pattern of these components:

Figure (I-1): Graphical representation of the components of a time series.



Accordingly, the observed time series Y_t can be expressed as a function of its components as follows:

$$Y_t = f(T_t, C_t, S_t, e_t)$$

However, if we exclude the cyclical component—since it generally appears only in very long time series—the series can be written as:

$$Y_t = f(T_t, C_t, S_t)$$

The functional relationship of the series can be represented using the following classical models:

- ✓ **Additive Model:**

$$Y_t = T_t + C_t + S_t$$

- ✓ **Multiplicative Model:**

$$Y_t = T_t \times C_t \times S_t$$

- ✓ **Mixed Model:**

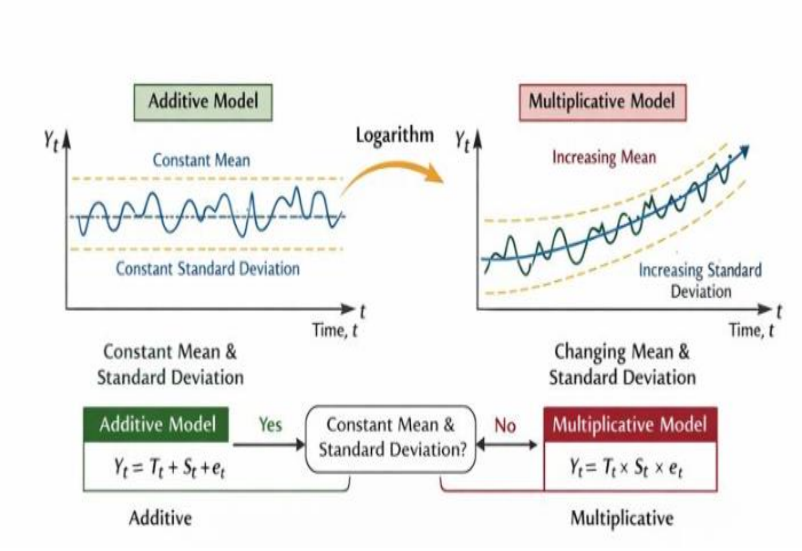
$$Y_t = (T_t \times S_t) + e_t$$

These models provide flexible ways to decompose and analyze the time series depending on the nature of the fluctuations and the relative magnitude of the components[1].

The nature of the time series model can be determined by calculating the arithmetic mean and the standard deviation of the series. If both of these measures remain approximately constant over time, the series is considered to follow an **additive model**. Conversely, if they vary over time, the series corresponds to a **multiplicative model**.

Furthermore, by applying a logarithmic transformation to a multiplicative or mixed model, the series can be converted into a standard additive model, which facilitates estimation, interpretation, and forecasting.

Figure (I-2): Determining the Nature of a Time Series Model.



[1] Brockwell P. J, Davis R. A**, Introduction to Time Series and Forecasting** , 3rd ed, Springer, New York**,** 2016, p p 3–5.

III- Methods for Identifying the Structure and Components of a Time Series:

There are several statistical tests used to identify the components of a time series. It is well known that these components can initially be detected through graphical analysis, by examining and interpreting the conditions under which the series was generated. However, graphical analysis alone is not sufficient to accurately identify the fundamental components of a time series. Therefore, it becomes necessary to rely on statistical tests in order to achieve a more rigorous and reliable identification of these components.

1- Graphical Method for Identifying and Detecting Time Series Components:

The graphical method is widely used as an initial approach for identifying and detecting the components of a time series[1]. Its application requires a high level of accuracy in the graphical presentation of the data, due to the considerable difficulty often faced by researchers in distinguishing the underlying components of the series[2]. In many situations, visual inspection alone does not provide sufficient evidence to clearly isolate the structural characteristics of a time series.

In general, when a time series exhibits an upward or downward movement over time, accompanied by regular and relatively constant fluctuations, it can be assumed that the series follows an **additive structure**, either increasing or decreasing[3]. In this case, the appropriate model can be expressed as:

$$Y_t = X_t + S_t + \varepsilon_t$$

[1] Chatfield. C**, The Analysis of Time Series: An Introduction**, 6th ed, Chapman & Hall/CRC, United Kingdom, 2004, p p 7–10.

[2] Wei W. S**, Time Series Analysis: Univariate and Multivariate Methods,** 2nd ed, Pearson Addison Wesley, United States, 2006, p p 4–6.

[3] Hyndman .R. J, Athanasopoulos . G**, Forecasting: Principles and Practice,** 2nd ed, OTexts, Australia, 2018, p p 21–24.

or equivalently as a linear trend model:

$$Y_t = \beta_0 + \beta_1 t + \varepsilon_t$$

Where:

Y_t :represents the dependent variable or the phenomenon under study.

$X_t = \beta_0 + \beta_1 t$: trend component.

S_t: represents the seasonal component.

ε_t: corresponds to the random or irregular component.

This specification is referred to as an **additive model**, since the observed value of the time series is obtained as the sum of its different components.

Conversely, when the magnitude of the fluctuations increases as time progresses, the time series is said to follow a **multiplicative structure**. In such cases, the model may be written as:

$$Y_t = X_t.S_t.\varepsilon_t$$

or alternatively:

$$Y_t = X_t.S_t.(1 + \varepsilon_t)$$

Although the graphical approach is useful for preliminary analysis, it is generally insufficient for accurately identifying all fundamental components of a time series. With the exception of the seasonal component—which is often evident through visual inspection—most components require confirmation through statistical and econometric tests[1].

2- The Analytical Method for Identifying and Detecting Time Series Components:

Given the limited effectiveness of the graphical method in clearly identifying the fundamental components of a time series, the analytical

[1] Hyndman .R. J, Athanasopoulos . G, op.cit, p 23.

method is adopted as a more rigorous and reliable approach. This method relies on statistical procedures and tests that allow for an objective assessment of the presence or absence of the main components of the time series, particularly the trend component[1].

2-1- Identification and Detection of the Trend Component:

The detection of the trend component constitutes a crucial step in time series analysis, as it reflects the long-term behavior of the studied phenomenon. Several statistical tests are employed for this purpose, among which nonparametric tests occupy a prominent place due to their flexibility and limited distributional assumptions[2].

2-1-1- Nonparametric Tests:

Nonparametric tests are used to detect the existence of a trend component without imposing any specific probability distribution on the random error term ε_t . These tests are particularly useful when the assumptions of classical parametric methods—such as normality of errors—are not satisfied. This is why they are called *distribution-free tests*[3].

- ✓ **Runs Test (Test of Randomness):**

The Runs Test is one of the most widely used nonparametric tests for assessing the randomness of a time series. It is based on the analysis of the sequence of signs of deviations from a central value, usually the median. The test aims to determine whether the observed series is generated by a purely random process or exhibits a systematic structure, such as a trend.

If the time series is random, this indicates the absence of a trend component. Conversely, a departure from randomness suggests the presence of a trend component.

---

[1] Kendall M. G, Ord J. K, op. cit, p p 22–27.
[2] Chatfield. C, op.cit, p 36.
[3] Conover .W, **Practical Nonparametric Statistics**, 3rd ed, Wiley, United States, 1999, p p 3–6.

The hypotheses of the test are formulated as follows:

$$\begin{cases} H_0: \text{The time series is random (no trend component)} \\ H_1: \text{The time series is not random (presence of a trend component)} \end{cases}$$

The implementation of the test proceeds through the following steps:

- The observations are ordered according to time;

- The median (Md) of the series is computed. If the sample size T is odd, the median corresponds to the observation of rank:

$$m = {}^{(T+1)}/_{2}$$

whereas if T is even, the median is given by:

$$m = {}^{T}/_{2} \rightarrow \ Md = \frac{Y_m + Y_{m+1}}{2}$$

- Observations greater than the median are assigned a positive sign, while those smaller than the median are assigned a negative sign;

- A *run* is defined as a sequence of identical signs. The total number of runs R is then computed.

For small samples ($m \leq 20$), critical values R_l and R_u are used directly. For larger samples ($m > 20$), the test statistic is approximated by the standard normal distribution:

$$|Z| = \frac{R - U_R}{\sigma_R}$$

$$\sigma_R = \sqrt{\frac{m(m+1)}{2m-1}} \qquad\qquad U_R = m + 1$$

If the observed number of runs is significantly smaller than the upper critical value, the null hypothesis of randomness is rejected, indicating the presence of a trend component.

Remark: In the presence of a trend component, the number of runs is expected to be relatively small, whereas a large number of runs suggests randomness and the absence of a trend.

Example:

To verify the existence of a **trend component** in the following data related to the sales of a firm from 2006 to 2020, the **Runs Test** is applied. (All values are expressed in **millions)**

T	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
observation	**14.12**	**14.42**	**16.12**	**16.78**	**17.34**	**17.86**	**18.37**	**18.96**	**19.56**	**20.19**	**20.84**	**21.52**	**22.19**	**22.81**	**23.45**

Solution:

It is observed that the chronological order coincides with the ascending order of the data.

Since T is odd and equal to 15, we have:

$$m = {(T+1)}/{2} = {(15+1)}/{2} = 8$$
$$Md = 18.96 \quad \leftarrow \quad m = 8$$

Moreover, the number of runs is:

$$R = 2$$

From the statistical table, the lower and upper critical values are obtained as follows:

$$R_l = 14.12 \qquad\qquad R_U = 23.45$$

Decision:

The null hypothesis H_0 is rejected since:

$$R < R_l < R_U$$

Therefore, the time series contains a **trend component**.

- ✓ **Turning Point Test :**

The designation *Turning Point Test* is not entirely accurate, since the test does not focus on the turning points themselves. Rather, it is concerned with the number of upward and downward movements of the time series. In other words, the test examines the number of sign changes from positive to negative or vice versa, based on the first-order differences of the series[1].

The first-order differences are computed as follows:

$$\Delta Y_{\mathrm{t}} = \mathrm{Y}_{\mathrm{t}} - \mathrm{Y}_{\mathrm{t}-1}$$

Formulation of the Test :

$$\begin{cases} H_0\text{: The time series is random (no trend component)} \\ H_1\text{: The time series is not random (presence of a trend component)} \end{cases}$$

Construction of the Test :

- The first-order differences of the time series are calculated. A positive sign is assigned to positive differences, and a negative sign to negative differences. Let u denote the number of sign changes in the sequence of ΔY_{t}.

---

[1] Kendall M. G, Ord J. K, op.cit, p 28.

- The Turning Point Test is generally applied when the number of observations t exceeds 10.

- The null hypothesis H_0 is rejected if the following condition is satisfied:

$$|Z| > Z_{\propto/2}$$

where the standardized test statistic Z is given by:

$$|Z| = \frac{u - u_u}{\sigma_u}$$

$$\sigma_u = \sqrt{\frac{16T - 29}{90}} \qquad u_u = \frac{2(t-2)}{3}$$

A statistically significant deviation of Z from zero indicates non-randomness in the series and, consequently, the existence of a trend component.

Example:

The following table contains 12 observations related to the consumption of a group of households.

T	1	2	3	4	5	6	7	8	9	10	11	12
C_t	155	158	163	171	153	156	162	172	162	164	173	181
Diff	−−	**3**	**5**	**8**	**−18**	**3**	**6**	**10**	**−10**	**2**	**9**	**8**
Sign	−−−	+	+	+	−	+	+	+	−	+	+	+

Solution :

Since the number of observations $T > 10$, this test can be applied.

In this example, we have:

$$u = 5$$

$$|Z| = \frac{\mathrm{u} - \mathrm{u_u}}{\sigma_\mathrm{u}}$$

$$u_u = \frac{2(t-2)}{3} = \frac{2(12-2)}{3} = \frac{20}{3} = 6.66$$

$$\sigma_u = \sqrt{\frac{16T-29}{90}} = \sqrt{\frac{16(12)-29}{90}} = \sqrt{\frac{163}{90}} = 1.34$$

$$|Z| = \frac{5-6.66}{1.34} = 1.23$$

We also have:

$$\mathrm{Z}_{\propto/2} = \mathrm{Z}_{2,5} = 1{,}96$$

Since:

$$|Z| < \mathrm{Z}_{\propto/2}$$

we reject the alternative hypothesis H_1 and accept the null hypothesis H_0, that is, we conclude that the time series is random. Consequently, the time series does **not** contain a trend component.

- ✓ **Rank Correlation Test for Detecting Trend Components:**

The Rank Correlation Test, also known as Daniel's Test, relies on Spearman's rank correlation coefficient. It is considered one of the most effective nonparametric tests for detecting the presence of a trend component in a time series. To apply this test, the following steps are followed[1] :

---

[1] Chatfield. C, op.cit, p p 40-46.

- **Assign ranks** to the values of the time series from the smallest to the largest, denoted R_t.

- **Compute the rank correlation coefficient** between the time element t and the ranks of the series values R_t:

$$r_s = \frac{COV(R_t, t)}{\delta_t . \delta_{Rt}}$$

Where :

$$V(Rt) = V(t) = \frac{T^2 - 1}{12}$$

δ_t and δ_{Rt} represent the standard deviations of t and Rt, respectively.

For practical computation, the coefficient can also be expressed as:

$$r_s = 1 - \frac{6 \sum_{t=1}^{n} d_t^2}{n(n^2 - 1)} \qquad , \qquad d_{\mathrm{t}} = \mathrm{Rt} - \mathrm{t}$$

where n is the number of observations. The coefficient r_s satisfies:

$$-1 \leq r_s \leq 1$$

$r_s > 0$ indicates a positive trend;

$r_s < 0$ indicates a negative trend;

$r_s \approx 0$ indicates no trend.

The hypothesis test is formulated as follows:

$$\begin{cases} H_0\text{: The time series is random (no trend component)} \\ H_1\text{: The time series contains a trend component} \end{cases}$$

The calculated coefficient is compared with its **tabulated critical value** to determine the presence or absence of a trend.

Application by Sample Size[1] :

- **Small samples ($n \leq 30$)**:

If the absolute value of the calculated coefficient exceeds the tabulated critical value:

$$|r_s| > r_{\alpha/2}$$

the time series is considered to contain a trend component; otherwise, no trend is detected. Here, $r_{\alpha/2}$ is the tabulated value of Spearman's statistic.

- **Large samples ($n > 30$)**:

In this case, the test statistic follows an approximate normal distribution, and the decision rule is:

$$|Z| > Z_{\alpha/2}$$

where:

$$Z = \frac{r - \mu_r}{\delta_r}$$

Where:

$$\mu_r = 0$$

Hence :

[1] Conover .W, op.cit, p p 256–258.

$$Z = \frac{r}{SD_r} = r\sqrt{n-1}$$

given that:

$$\delta_r = \frac{1}{\sqrt{n-1}}$$

Therefore :

$$Z = r\sqrt{n-1}$$

The Daniel/Spearman test is widely used in economic and statistical applications because it is nonparametric and robust to violations of classical assumptions such as normality and linearity. It is particularly suitable for detecting trends in[1]:

- Aggregate consumption ;

- Industrial production indices ;

- Price series of commodities or financial assets;

- Import/export volumes.

It also complements other **nonparametric tests** such as the **Runs Test** and **Turning Point Test**, providing a robust framework for identifying trend components in both short and long time series.

Example :

Consider the following time series. The objective is to determine whether a trend component exists or not by using the Spearman rank correlation test at a 5% significance level.

[1] Conover .W**,** op.cit, p p 256–258.

12	11	10	9	8	7	6	5	4	3	2	1	t
181	173	164	162	172	162	156	153	171	163	158	155	Y_t
12	**11**	**8**	**5.5**	**10**	**5.5**	**3**	**1**	**9**	**7**	**4**	**2**	**Rt**
0	**0**	**2**	**3.5**	**−2**	**1.5**	**3**	**4**	**−5**	**−4**	**−2**	**−1**	$\mathbf{d_t}$
0	**0**	**4**	**12.25**	**4**	**2.25**	**9**	**16**	**25**	**16**	**4**	**1**	$\mathbf{d_t^2}$

$$\sum_{t=1}^{12} \mathbf{d_t^2} = \mathbf{1+4+16+25+16+9+2.25+4}$$

$$\mathbf{+\,12.25+4+0+0=93.5}$$

$$r_s = 1 - \frac{6(93.5)}{12(144-1)} = 0.67$$

$$T = 12,\ \ \alpha = 5\% \qquad \rightarrow r_{\alpha/2} = r_{2.5} = 0.5804$$

this value is obtained from the Spearman rank correlation table**.**

We observe that the calculated value is greater than the tabulated value, that is:

$$|r_s| > r_{\alpha/2} \rightarrow 0.67 > 0.5804$$

Therefore, the null hypothesis H_0 is rejected, which implies that the time series **contains a trend component**.

2-1-2- parametric Tests:

Parametric tests are among the most important methods used to detect the presence of a trend component in a time series. Their role is not limited to identifying whether a trend exists; rather, they also help determine the appropriate transformation or modeling strategy that allows the studied series to converge toward stationarity. These tests are based on the assumption that the time series contains a deterministic trend component in addition to a random component, while explicitly assuming that the error term follows a known probability distribution, most commonly the normal distribution[1].

Within this framework, parametric tests provide a formal statistical basis for distinguishing between trend-stationary and non-stationary processes. This distinction is of fundamental importance in time series analysis, as many econometric models and forecasting techniques require stationarity as a necessary condition for valid inference[2].

- ✓ **Unit Root Tests:**

Issues related to testing for the presence of a unit root in time series representing various economic variables attracted considerable attention among researchers during the 1980s. This growing interest stemmed from the recognition that many macroeconomic and financial time series exhibit persistent behavior that cannot be adequately captured by traditional trend-stationary models[3].

In response to this concern, several unit root testing procedures were developed, beginning with the seminal contributions of Fuller (1976) and Dickey and Fuller (1979, 1981). These foundational works led to the formulation of one of the most influential and widely used tests in time series analysis, namely the Dickey–Fuller (DF) unit root test. The test was designed to examine whether a stochastic trend dominates the behavior of a time series, implying non-stationarity, or whether the series can be rendered stationary through appropriate transformations such as differencing[4].

---

[1] Kendall M. G, Ord J. K, op.cit, p p 25-27.

[2] Gujarati. D, Porter. C, **Basic Econometrics**, 5th ed, McGraw-Hill, New York, United States, 2009, p p 741–743.

[3] Hamilton. J. D, **Time Series Analysis**, Princeton University Press, Princeton, United States, 1994, p p 47–49

[4] Dickey. D. A, Fuller, W. A, **Distribution of the Estimators for Autoregressive Time Series with a Unit Root**, Journal of the American Statistical Association, 74(366), United States, 1979, p p 427–431.

Unit root tests are particularly important because the presence of a unit root implies that shocks to the series have permanent effects, thereby affecting both the interpretation of economic relationships and the reliability of forecasts. Consequently, these tests play a central role in the preliminary analysis of economic time series prior to model specification and estimation[1] .

- ✓ **The Simple Dickey–Fuller Test (DF, 1979):**

The starting point of unit root testing is the following autoregressive model of order one:

$$X_t = \rho X_{t-1} + \varepsilon_t \quad \dots\dots\dots(1)$$

where:

$$\varepsilon_t \sim i.i.d.(0.\sigma_\varepsilon^2)$$

Depending on the value of the parameter "ρ", three possible cases can be distinguished[2]:

- If $|\rho| < 1$: the time series X_t is stationary, since current observations receive more weight than past observations.

- If $\rho = 1$: the time series X_t is non-stationary, and current and past observations play the same role; this case corresponds to a random walk process.

- If $|\rho| > 1$: the time series X_t is non-stationary, with its variance increasing exponentially over time, and past observations receiving more weight than current observations.

Within the framework of unit root testing, particular attention is devoted to the case where $\rho = 1$. In this context, Dickey and Fuller

[1] Enders. W, **Applied Econometric Time Series**, 4th ed, Wiley, Hoboken, United States, 2015, p p 208–210.

[2] Hamilton. J. D, op.cit, p p 48-49.

proposed a **left-sided unit root test** based on the following hypotheses[1]:

$$\begin{cases} H_0: \rho = 1 \quad \text{random walk, non} - \text{stationary series} \\ \qquad H_1: |\rho| < 1 \quad \text{stationary series} \end{cases}$$

By subtracting X_{t-1} from both sides of equation (1), the model can be rewritten in first-difference form as one of the following specifications:

$$\Delta X_t = \varphi X_{t-1} + \varepsilon_t \dots\dots\dots (\mathbf{I})$$

$$\Delta X_t = \varphi X_{t-1} + \beta_0 + \varepsilon_t \dots\dots\dots (\mathbf{II})$$

$$\Delta X_t = \varphi X_{t-1} + \beta_0 + \beta_1 t + \varepsilon_t \dots\dots\dots (\mathbf{III})$$

where:

$$\varphi = \rho - 1 \qquad , \qquad \varepsilon_t \sim \text{i.i.d.}(0.\sigma_\epsilon^2)$$

Testing the hypothesis $H_0: \rho = 1$ is therefore equivalent to testing:

$$\begin{cases} H_\mathbf{0}: \varphi = 0 \\ H_\mathbf{1}: |\varphi| < 0 \end{cases}$$

using one of the models (I), (II), or (III), depending on whether the series contains a constant and/or a deterministic time trend[2].

To test this hypothesis, Dickey and Fuller proposed two complementary approaches. The first is based on the non-standard (restricted) distribution of the Ordinary Least Squares (OLS) estimators of the parameters in the above models. The second relies on the computation of a *t*-statistic associated with the estimated coefficient φ

---

[1] Dickey. D. A, Fuller, W. A, op.cit, p p 427-428.

[2] Enders. W, op.cit, p p 214-216.

. However, under the null hypothesis $H_0: \varphi = 0$, the distribution of this statistic is neither standard normal nor asymptotically Student-*t*[1].

For this reason, Dickey and Fuller derived specific critical values for each of the three model specifications and for different sample sizes. These critical values were later tabulated by Fuller (1976) and Dickey and Fuller (1979). The decision rule is therefore based on comparing the calculated value of the test statistic $t_{\hat{\rho}}$ (or equivalently $t_{\hat{\varphi}}$) with the corresponding Dickey–Fuller critical value. If the calculated statistic is greater (in absolute value) than the tabulated critical value, the null hypothesis is rejected and the series is concluded to be stationary; otherwise, the null hypothesis is accepted, indicating that the time series is non-stationary and contains a unit root[2].

- ✓ **Augmented Dickey-Fuller Test (ADF, 1981) :**

Although the simple DF test is theoretically important, it assumes no serial correlation in the error term, which is rarely satisfied in practice, especially in economic time series that often exhibit high dynamics and delayed effects[3].

Since the DF models include only X_{t-1} as the explanatory variable, omitting relevant lagged terms often results in autocorrelated errors. To overcome this limitation, Dickey and Fuller (1981) proposed the Augmented Dickey-Fuller (ADF) test by adding lagged differences of the dependent variable, yielding the following models:

$$\Delta X_t = \varphi X_{t-1} + \sum_{j=2}^{p} \gamma_j \, \Delta X_{t-j} + \varepsilon_t \; \dots \dots \dots . (\,\mathbf{IV}\,)$$

$$\Delta X_t = \varphi X_{t-1} + \beta_0 + \sum_{j=2}^{p} \gamma_j \, \Delta X_{t-j} + \varepsilon_t \; \dots \dots \dots . (\,\mathbf{V}\,)$$

$$\Delta X_t = \varphi X_{t-1} + \beta_0 + \beta_1 t + \sum_{j=2}^{p} \gamma_j \, \Delta X_{t-j} + \varepsilon_t \; \dots \dots \dots . (\,\mathbf{VI}\,)$$

---

[1] Dickey. D. A, Fuller, W. A, op.cit, p p 429-430.
[2] Hamilton. J. D, op.cit, p p 50-52.
[3] Enders. W, op.cit, p 135.

Where :

$$\varphi = (\rho - 1)(1 - \theta_1 - \cdots - \theta_{p-1})$$

ρ represents the number of lagged differences included to absorb potential autocorrelation in the residuals. The hypotheses remain:

$$\begin{cases} H_0: \varphi = 0 & \textit{presence of unit root} \\ H_1: |\varphi| < 0 & \textit{stationary series} \end{cases}$$

Dickey and Fuller (1981) demonstrated that the distributions of the estimators in the ADF models are similar to those in the original DF test, allowing the use of the same critical value tables.

Selecting the appropriate number of lags p is crucial. Over- or underestimating p can lead to misleading results. Common approaches include using information criteria such as Akaike Information Criterion (AIC) or Schwarz Bayesian Criterion (BIC), or employing residual autocorrelation tests such as Box-Pierce or Ljung-Box, stopping at the first lag that satisfies the null hypothesis of no residual autocorrelation[1].

2-2- Detecting and Identifying the Seasonal Component:

The seasonal (or periodic) component represents the systematic and recurrent variations observed in a time series over short and regular time intervals within a given year. These variations arise from factors that repeat themselves according to a fixed calendar pattern, such as climatic conditions, institutional arrangements, social habits, or production cycles. Typical examples include domestic or industrial electricity consumption over a 24-hour cycle, monthly agricultural output, or seasonal demand for certain goods and services[2].

The presence of seasonal fluctuations allows the researcher to gain an initial understanding of how the studied phenomenon evolves

[1] Gujarati. D, Porter. C, op.cit, 664.

[2] Gujarati. D, Porter. C, op.cit. p 743.

within short time horizons. A common preliminary approach to identifying this component relies on graphical representation, where the seasonal pattern often appears clearly through regular and repeated oscillations around the general trend of the series[1].

Nevertheless, although graphical analysis is a useful exploratory tool, it requires a high degree of accuracy in data presentation and interpretation. In many empirical situations, visual inspection alone may be insufficient or even misleading, particularly when the seasonal component interacts with other components of the time series, such as the trend or irregular fluctuations[2].

For this reason, researchers generally complement graphical methods with analytical procedures to formally verify the existence and significance of the seasonal component. These analytical methods aim to isolate seasonal effects and assess their statistical relevance, thereby providing more reliable conclusions regarding the structure of the time series[3]. Among the most widely used analytical approaches are those based on seasonal indices, decomposition techniques, and regression models incorporating seasonal dummy variables.

Among the most widely used analytical methods for detecting seasonality is the **Kruskal–Wallis test (KW)**.

Kruskal–Wallis Test (KW):

The Kruskal–Wallis test is one of the most commonly applied non-parametric tests for determining the presence or absence of a seasonal component in a time series. It is denoted by **KW** and is based on the following empirical statistic[4]:

$$KW = \frac{12}{n(n+1)} \times \sum \frac{R_i^2}{m_i} - 3(n+1)$$

[1] Hyndman. R. J., Athanasopoulos. G, op.cit, p 29.

[2] Chatfield. C, op.cit, p 18.

[3] Wei W. S, op.cit, p 12.

[4] Conover .W, op.cit, p p 288.

This statistic follows a chi-square distribution with (P−1) degrees of freedom:

$$KW \to \chi^2(\mathrm{P}-1)$$

Where:

R_i^2: sum of the ranks corresponding to season i;
m_i: number of observations associated with each season, which in most cases corresponds to the number of years;
P: periodicity of the seasonal component. For instance, $\mathrm{P} = 4$ for quarterly data, $\mathrm{P} = 12$ for monthly data, and so forth.

- ✓ **Steps for Applying the Kruskal–Wallis Test:**

1- Removing the Trend Component :

1-1- Estimation of the Trend Component:

The trend component is first identified and estimated by fitting a linear trend model using the Ordinary Least Squares (OLS) method[1]:

$$X_t = \beta_0 + \beta_1 t + \varepsilon_t$$

The estimation consists of minimizing the sum of squared residuals:

$$Min \sum_{i=1}^{n} e_i^2 = \underset{\widehat{\beta_0},\widehat{\beta_1}}{\mathrm{Min}} \sum_{i=1}^{n} (Y_i - \hat{\beta}_0 - \hat{\beta}_1 t)^2$$

This leads to the following normal equations:

[1] Gujarati. D, Porter. C, op.cit. p p 65-67.

$$\begin{cases} \dfrac{\partial \sum_{i=1}^{n}(Y_i - \hat{\beta}_0 - \hat{\beta}_1 t)^2}{\partial \hat{\beta}_0} = 0 \\ \dfrac{\partial \sum_{i=1}^{n}(Y_i - \hat{\beta}_0 - \hat{\beta}_1 t)^2}{\partial \hat{\beta}_1} = 0 \end{cases}$$

Solving these equations yields:

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{t}$$

$$\hat{\beta}_1 = \frac{cov(Y_t, t)}{v(t)} = \frac{\sum tY - n\bar{t}\bar{Y}}{\sum t^2 - n\bar{t}^2}$$

Given that:

$$t = 1,2,3, \dots \dots \dots, n$$

$$v(t) = \frac{n^2 - 1}{12} \qquad , \qquad \bar{t} = \frac{n + 1}{2}$$

The estimators can be rewritten as:

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \left(\frac{n + 1}{2}\right)$$

$$\hat{\beta}_1 = \frac{\sum \frac{tY}{n} - \left(\frac{n+1}{2}\right)\bar{Y}}{\frac{n^2 - 1}{12}}$$

1-2- Removal of the Trend Component:

The trend component is removed by computing the residual series w_t:

$$w_t = Y_t - \hat{X}_t = Y_t - \left(\hat{\beta}_0 - \hat{\beta}_1 t\right)$$

2- Detecting the Seasonal Component :

- Rank the residuals w_t and adjust the ranks in the case of ties;

- Compute the calculated value of the KW statistic;

- Determine the critical value $\chi^2(\alpha, p-1)$ at significance level α;

- Compare the calculated and critical values:

  - ✓ If $KW > \chi^2(\alpha, p-1)$, the time series contains a seasonal component;
  - ✓ If $KW < \chi^2(\alpha, p-1)$, the time series does not contain a seasonal component.

Note:

If the time series is found not to contain a trend component, the third step can be directly applied, since in this case:

	1	2	3	4
2016	14	20	44	21
2017	10	19	64	32
2018	12	12	68	29
2019	7	18	60	36
2020	6	11	64	50

$$w_t = Y_t$$

Example:
Consider the following time series:

Verify the presence or absence of the seasonal component.

Solution:
To detect the seasonal component, the following steps are followed:

1- Verification of the presence or absence of the trend component:

To determine whether a trend component exists, the **Spearman rank correlation coefficient** is used. It is computed as follows:

$$r_s = 1 - \frac{6\sum_{t=1}^{n} d_t^2}{n(n^2-1)}$$

where d_t denotes the difference between the rank of the observation and the rank of time, and n is the number of observations.

a. Computation of the ranks R_t:

	1	**2**	**3**	**4**
1	7	10	15	11
2	3	9	18.5	13
3	5.5	5.5	20	12
4	2	8	17	14
5	1	4	18.5	16

b- Calculation of the squared ranks d_t^2:

	1	**2**	**3**	**4**
1	36	64	144	49
2	4	9	132.25	25
3	12.25	20.25	81	0
4	121	36	4	4
5	256	196	0.25	16

c- Determination of the calculated value of the Spearman rank correlation coefficient:

$$r_s = 1 - \frac{6(1210)}{7980} = 0.9022$$

d- Determination of the tabulated or critical value for the Spearman statistic:

$$t = 20 \;, \alpha = 5\% \quad \rightarrow r_{\alpha/2} = r_{2.5} = 0.4451$$

We observe that the calculated value is greater than the tabulated value; therefore, the time series does **not** contain a general trend component.

2- Determination of the Presence or Absence of the Seasonal Component:

In this case, $\mathrm{w_t} = \mathrm{Y_t}$; therefore, we calculate the ranks of the original table values $\mathrm{Y_t}$.

a- Calculation of the ranks $\boldsymbol{R_t}$:

We notice the presence of tied values; therefore, we adjust them accordingly. Then, we calculate the R_j values, which represent the sum of ranks for each season, segment of the year, or the studied periods.

	1	**2**	**3**	**4**
1	7	10	15	11
2	3	9	18.5	13
3	5.5	5.5	20	12
4	2	8	17	14
5	1	4	18.5	16
$\boldsymbol{R_j}$	18.5	36.5	89	66

Total number of values or sample size: $n = 20$

Number of values or observations corresponding to each season:

$$m_1 = m_2 = m_3 = m_4 = 5$$

b- Calculating the computed value of KW:

$$KW = \frac{12}{20(20+1)} \times \left[\frac{(18.5)^2}{5} + \frac{(36.5)^2}{5} + \frac{(89)^2}{5} + \frac{(66)^2}{5}\right] - 3(20+1) = 16.72$$

c- Determining the critical value:

$$\chi^2(5\%\%, 3ddl) = 7.815$$

d- Comparison between the computed and critical values: We observe that the computed value is greater than the critical value, i.e.,

$$\text{KW} > \chi^2(\alpha, p-1)$$

Thus, we conclude that the time series contains a seasonal component.

IV – Methods for Determining the Form of a Time Series:

Once the existence of a seasonal component in a time series has been confirmed, the subsequent step is to determine the structural form of the series. In practice, three primary forms are distinguished: **additive (aggregated)**, **multiplicative**, and **mixed models**. Accurately identifying the form of a time series is essential, as it directly affects forecasting accuracy and the selection of appropriate smoothing and modeling techniques[1].

1- Annual Mean Method:

The annual mean method is particularly useful when a year is divided into regular intervals, such as months, quarters, or semesters. The procedure involves the following steps:

[1] Chatfield, op.cit, p p 65-68.

- Calculate the annual arithmetic mean for each year;

- Compute the differences between the original values of each year and their corresponding annual mean.

- ✓ If these differences form an arithmetic sequence or are approximately equal, the series follows an additive model.
- ✓ If these differences form a geometric sequence, meaning that the differences increase or multiply from year to year, the series follows a multiplicative model.

This method provides an initial and intuitive understanding of the series structure, but it is sensitive to extreme values and irregular fluctuations[1].

2- Annual Standard Deviation Method:

The annual standard deviation method complements the mean method. Here, the standard deviation of the observations within each year is calculated[2]:

- ✓ If the annual standard deviations are approximately constant, this supports an additive model, suggesting that the variability of the seasonal and random components remains stable across time.
- ✓ If the annual standard deviations vary significantly, it indicates a multiplicative model, reflecting that the variability grows or diminishes proportionally with the level of the series.

This approach provides an important diagnostic check for heteroscedasticity, which is critical for selecting appropriate forecasting models.

---

[1] Harvey. A. C, **Forecasting, Structural Time Series Models and the Kalman Filter**, Cambridge University Press, Cambridge, UK, 1990, p p 120-125.

[2] Hyndman. R. J, Athanasopoulos. G, **Forecasting: Principles and Practice**, 3rd Edition; OTexts, Melbourne, Australia, 2021, p p 107-110.

3- Regression Equation Method:

The regression-based method offers a more formal analytical approach to classify the time series type. The method relies on the regression of the standard deviation of each period (σ_i) against the corresponding mean ($\overline{Y}_i$) of the period:

$$\sigma_i = \hat{\beta}_0 + \hat{\beta}_1 \overline{Y}_i$$

where the slope coefficient $\hat{\beta}_1$ is computed as:

$$\hat{\beta}_1 = \frac{\frac{\sum \sigma \overline{Y}}{n} - \overline{\sigma}\overline{\overline{Y}}}{\frac{\sum \overline{Y}^2}{n} - \overline{\overline{Y}}^2}$$

Interpretation of $\hat{\beta}_1$ provides a classification of the time series form:

- $\hat{\beta}_1 < 0.05$→ indicates an **additive (linear) model**,
- $\hat{\beta}_1 > 0.1$ → indicates a **multiplicative model**,
- $0.05 \leq \hat{\beta}_1 \leq 0.1$→ indicates a **mixed model**, where both additive and multiplicative effects coexist.

The regression method is highly recommended because it combines trend and variability information in a single quantitative measure, allowing researchers to make informed decisions before applying smoothing, decomposition, or forecasting techniques[1].

Example:

The following time series illustrates the evolution of sales for a particular product over a period of five years, where each year is divided into four quarters.

[1] Chatfield, op.cit, p p 70-73.

	1	2	3	4
1	20	28	22	34
2	19	39	25	44
3	21	49	33	55
4	23	60	37	66
5	24	71	42	76

- Use the annual mean method to determine the form of the time series.

- Use the annual standard deviation method to determine the form of the time series.

- Use the regression equation method to determine the form of the time series.

Solution:

1- Annual Mean Method:

Step 1: Calculate the annual mean for each year, $\overline{Y}$.
Step 2: Compute the differences between the original values and the corresponding annual mean for each year.

The following table presents the various calculations involved in this method:

	1	2	3	4	$\overline{Y}$
1	−6	2	−4	8	26
2	−12.75	7.25	−6.75	12.25	13.75
3	−18.5	9.5	−6.5	15.5	39.5
4	−23.5	13.5	−9.5	19.5	46.51
5	−29.25	17.75	−11.25	22.75	53.25

We observe that the seasonal differences (or seasonal variations) for the first quarter, for example, increase from one year to another (6, 12.75, 18.5, 23.5, 29.25). This indicates that the time series follows a **multiplicative model**, which can be expressed as follows:

$$Y_t = T_t \times C_t \times S_t$$

2- Annual Standard Deviation Method:

This method consists of a single step, which is the calculation of the annual standard deviation for each year. The following table presents the various calculations involved in this method.

Year	$\overline{Y}$	σ_i
1	26	5.447
2	31.75	10.13
3	39.5	13.37
4	46.5	17.36
5	53.25	21.29

We observe that the annual standard deviations are not constant from one year to another; therefore, the model corresponding to this time series is the **multiplicative model**.

- ✓ **Regression Equation Method:**

The following steps are applied:

- Calculation of the annual means for each year, $\overline{Y}$;

- Calculation of the annual standard deviations for each year, σ_i;

- Estimation of the regression coefficient $\hat{\beta}_1$.

$$\hat{\beta}_1 = \frac{\frac{\sum \sigma \bar{Y}}{n} - \bar{\sigma}\bar{\bar{Y}}}{\frac{\sum \bar{Y}^2}{n} - \bar{\bar{Y}}^2}$$

$$\bar{\bar{Y}}^2 = 1552.517 \text{ ؛ } \bar{\bar{Y}} = \frac{197.01}{5} = 39.601 \text{ ؛ } \bar{\sigma} = \frac{67.627}{5} = 13.5254$$

$$\sum \sigma \bar{Y} = 2985.32 \text{ ؛ } \sum \bar{Y}^2 = 8243.05$$

$$\hat{\beta}_1 = \frac{\frac{2985.32}{5} - 13.5254 \times 39.601}{\frac{8243.05}{5} - 1552.517} = \frac{51.3463}{96.094} = 0.53$$

Since the value of the regression coefficient is greater than 0.1:

$$\left(\hat{\beta}_1 > 0.1\right)$$

we conclude that the time series follows a **multiplicative model**.

V – Analysis of Time Series Subject to Seasonal Variations :

In many economic, demographic, and social phenomena, the time series under study is not only influenced by a random component and a long-term trend, but also by seasonal or periodic fluctuations. These fluctuations reflect systematic variations that recur regularly over fixed periods of time, such as months, quarters, or semesters. Accounting for the seasonal component is therefore essential in order to correctly model, interpret, and forecast the behavior of the series[1].

Seasonal fluctuations are typically associated with climatic conditions, institutional arrangements, or behavioral patterns, and their identification constitutes a fundamental step in time series analysis[2].

---

[1] Kendall. M. G, Stuart. A, **The Advanced Theory of Statistics**, *Vol. 3,* Griffin, 1976, p p 310–317.
[2] Chatfield, op.cit, p p 13-16.

When seasonality is present, the series may be represented using one of three standard forms: the **additive model**, the **multiplicative model**, or the **mixed model**. The selection of the appropriate form depends on whether seasonal variations remain constant in absolute value or vary proportionally with the level of the series[1].

In the present analysis, attention is restricted to the **additive model**, for which parameter estimation can be efficiently conducted using the **Buys-Ballot table** combined with the ordinary least squares (OLS) method. It is important to note that the Buys-Ballot procedure is applicable **exclusively** to additive seasonal structures.

- ✓ **Estimation of the Additive Time Series Model Using the Buys-Ballot Table**

The additive model is written as[2]:

$$Y_t = T_t + S_t + e_t$$

where:

Y_t: denotes the observed value of the series at time ttt;

T_t: represents the trend component;

S_t: is the seasonal component;

e_t: is the random component, assumed to be white noise with zero mean and constant variance.

A fundamental condition of the additive model is that seasonal fluctuations are **constant in magnitude over time**. Consequently, the

[1] Makridakis. S, Wheelwright. S. C, Hyndman. R. J, **Forecasting: Methods and Applications,** 3rd ed, Wiley, 1998, p p 19–23.
[2] Box. G. E. P, Jenkins. G. M, Reinsel. G. C, op.cit, p p 8–10.

sum of the seasonal components over a complete cycle must be equal to zero.

$$\sum S_{\mathrm{j}} = 0$$

This restriction ensures parameter identifiability and prevents confounding between the trend and seasonal components[1].

- ✓ **Temporal Indexing and Model Reformulation:**

To simplify the estimation procedure, the time index t is decomposed as:

$$t = \big(j + m(i - 1)\big)$$

where:

j: denotes the sub-period within the year (month, quarter, etc.);

m: is the number of sub-periods per year;

i: denotes the year index.

Let the seasonal-adjusted intercept be defined as:

$$a_j = a + S_j$$

Summing over one complete seasonal cycle yields:

$$\sum a_j = \sum a + \sum S_j$$

Since $\sum S_{\mathrm{j}} = 0$, it follows that:

$$\sum a_j = ma + 0 \rightarrow a = \frac{\sum a_j}{m}$$

[1] Kendall. M. G, Stuart. A, op.cit, p 312.

Accordingly, the additive model can be rewritten as:

$$Y_{ij} = b\big(j + m(i-1)\big) + a_j + \varepsilon_{ij}$$

This formulation separates the trend and seasonal effects in a manner consistent with classical decomposition techniques.

- ✓ **Estimation by Ordinary Least Squares :**

The parameters b and a_j are estimated by minimizing the sum of squared residuals:

$$\sum_{i}\sum_{j}\varepsilon_{ij}^2 = \sum_{i}\sum_{j}\left[Y_{ij} - b(j + m(i-1)) - a_j\right]^2 = \min$$

Applying the OLS method leads to the following normal equations:

$$\frac{\partial \sum_i \sum_j \varepsilon_{ij}^2}{\partial a_j} = 0 \ldots\ldots\ldots (1)$$

$$\frac{\partial \sum_i \sum_j \varepsilon_{ij}^2}{\partial b} = 0 \ldots\ldots\ldots (2)$$

These equations allow the separate estimation of the trend and seasonal parameters[1].

- ✓ **Estimation of the Model Parameters:**

Estimation of the Intercept Parameter a:

Solving the first normal equation yields:

$$a = \bar{\bar{Y}} - b\left[\frac{m+1}{2} + \frac{m(n-1)}{2}\right] = \bar{\bar{Y}} - b\left(\frac{m.n+1}{2}\right)$$

where $\bar{\bar{Y}}$ denotes the overall mean of the series.

Estimation of the Trend Coefficient b:

The slope of the trend component is given by:

[1] Chatfield, op.cit, p p 33-35.

$$b = \frac{12\left[\sum i\bar{Y}_i - n\frac{(n+1)}{2}\bar{\bar{Y}}\right]}{mn(n^2-1)}$$

This expression is derived under the assumption of equally spaced observations and linear trend behavior.

Estimation of the Seasonal Component S_j:

Since $a_j = a + S_j$, the seasonal component is obtained as:

$$S_j = a_j - a$$

which can be written as:

$$S_{\mathrm{j}} = \bar{\mathrm{Y}}_{\mathrm{j}} - \bar{\bar{\mathrm{Y}}} - \mathrm{b}\left(\mathrm{j} - \frac{(m+1)}{2}\right)$$

This formula enables the isolation of seasonal effects once the trend has been removed.

The parameters a and b correspond respectively to the intercept $\hat{\beta}_0$ and the slope $\hat{\beta}_1$ of the linear trend model. Combined with the seasonal indices S_{j}, they form the basis of the **Buys-Ballot decomposition**, which allows the joint estimation of trend, seasonal, and irregular components in additive time series models[1].

[1] Kendall. M. G, Stuart. A, op.cit, p p 315-317.

Chapter Two: Stationarity, Autocorrelation, and Partial Autocorrelation

The concepts of stationarity, autocorrelation, and partial autocorrelation constitute fundamental pillars in time series analysis due to their critical role in understanding the dynamic structure of a series and identifying the statistical properties governing its temporal behavior. Most statistical and econometric models used in time series analysis, particularly regression-based models and Box–Jenkins methodologies, implicitly or explicitly assume that the series under study is stationary, or can be transformed into a stationary series.

Stationarity refers to the constancy of the core statistical properties of a time series—such as the mean, variance, and autocovariance—over time. It is a necessary condition to ensure the validity of statistical inference and the accuracy of future forecasts. In the absence of stationarity, results derived from standard models can be misleading and unreliable, necessitating rigorous stationarity tests and appropriate corrective procedures.

Autocorrelation provides a vital analytical tool for examining the relationship between current and past values of a series, reflecting the extent to which the series depends on its own historical behavior. Partial autocorrelation, on the other hand, measures the direct relationship between values of the series at a specific lag after removing the effects of intermediate lags, making it an essential instrument for determining the appropriate order of a time series model.

I - Definition of stationarity in time series:

The concept of stationarity is pivotal in time series analysis, particularly due to the inherent sensitivity of time series data and its critical role in modeling for future value predictions. To ensure the reliability of any forecasting model, it is essential to rigorously assess the stochastic properties of the time series.

A time series is considered stationary if its statistical properties, specifically the mathematical expectation (mean) and variance, remain constant over time. This stability implies that the series' distribution does not change; thus, it does not exhibit any systematic trends or seasonal patterns. In practical terms, a stationary time series may fluctuate around a constant mean with constant variance, indicating the absence of a long-term upward or downward trend.

To formally test for stationarity, various statistical tests can be used, such as the Augmented Dickey-Fuller (ADF) and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) tests. These tests help determine whether a unit root is present in the series, indicating non-stationarity. If the series is found to be non-stationary, transformations such as differencing, detrending, or applying logarithmic scaling may be necessary to achieve stationarity, thereby making the time series suitable for further analysis and forecasting.

So, the stationarity of a time series is a fundamental criterion that underpins the validity of predictive models. A stable time series, characterized by constant mean and variance, is essential for accurate forecasting and is a prerequisite for many time series modeling techniques, including ARIMA and other autoregressive models.

II - Linear Models for Time Series :

Time series analysis involves modeling temporal data to uncover patterns, forecast future values, and understand underlying dynamics. Among the most widely used tools are linear models, which include Autoregressive (AR), Moving Average (MA), and Autoregressive Moving Average (ARMA) models. we will analyse these models as follows[1] :

2-1- moving average process Ma(q) :

A series Y_t is said to follow a moving average process of order q if its current value can be explained and interpreted as a weighted average of past random error values and we write:

$$MA(q): y_t = \theta_0 + \theta_1 \zeta_{t-1} + \theta_2 \zeta_{t-2} + \cdots + \theta_q \zeta_{t-q} + \zeta_t$$

[1] Sharif Hossain, **Econometric Analysis an applied approach to business and economics**, Cambridge scholars publishing, United kingdom, 2024, p p 364-381.

Where :

ζ_t: is white noise that follows a normal distribution ;
θ_0 : represents the constant term;
$\theta_1, \theta_2, , , \theta_q$ represent the model coefficients, which determine the effect of past residuals (the impact of random shocks) on the current value of the series.

Example 1:

$$MA(1): y_t = \theta_0 + \theta_1 \zeta_{t-1} + \zeta_t$$

$$MA(2): y_t = \theta_0 + \theta_1 \zeta_{t-1} + \theta_2 \zeta_{t-2} + \zeta_t$$

model is typically expressed using the lag operator (L) where:

$$L^d \zeta_t = \zeta_{t-d}$$

And we write :

$$Y_t = \theta_0 + \theta_1 L \zeta_t - \theta_2 L^2 \zeta_t + \cdots + \theta_q L^q \zeta_t + \zeta_t$$

$$Y_t = \theta_0 + (1 - \theta_1 L - \theta_2 L^2 - \cdots - \theta_q L^q)\zeta_t = \theta_0 + \zeta_t$$

$$Y_t = \theta_0 + \theta_q(L)\zeta_t$$

Where :

$$\theta_q(L) = 1 - \theta_1 L - \theta_2 L^2 - \cdots - \theta_q L^q$$

Stationarity of mooving averege process ma(q) :

To ensure that the MA(q) model is stable, it must meet the conditions of stationarity, which are the constancy of the mean, variance, and covariance over time[1].

---

[1] Marc S.Paolella, **Linear Models and Time series Analysis regression (Anova, Arma and Garch)**, wily series in probability and statistics,university of Zurich, Switzerland, 2019, P 295.

We can write the MA(q) model as follows:

$$y_t = \mu - \theta_1 \zeta_{t-1} - \theta_2 \zeta_{t-2} - \cdots - \theta_q \zeta_{t-q} + \zeta_t$$

Where :

ζ_t: is white noise that follows a normal distribution.

The expected value of the series $y_t$is:

$$E(y_{t)} = E(\mu - \theta_1 \zeta_{t-1} - \theta_2 \zeta_{t-2} - \cdots - \theta_q \zeta_{t-q} + \zeta_t)$$

$$E(y_t) = \mu - \theta_1 E(\zeta_{t-1}) - \theta_2 E(\zeta_{t-2}) - \cdots - \theta_q E(\zeta_{t-q}) + E(\zeta_{t-1}) = \mu$$

This indicates that the mean is constant and independent of time. The variance of the series y_t is:

$$Var(Y_t) = E(Y_t - \mu)^2$$

$$Var(Y_t) = E(Y_t - \mu)^2 = E\left(\mu - \theta_1 \zeta_{t-1} - \theta_2 \zeta_{t-2} - \cdots - \theta_q \zeta_{t-q} + \zeta_t\right)^2$$

$$Var(Y_t) = E\left({\zeta_t}^2 + \theta_1 {\zeta_{t-1}}^2 + \theta_2 {\zeta_{t-2}}^2 + \cdots + \theta_q {\zeta_{t-q}}^2 + \theta_i \theta_j \zeta_{t-i} \zeta_{t-j} + \cdots - \theta_s \zeta_t \zeta_{t-s} - \cdots\right)$$

$$Var(Y_t) = E\left({\zeta_t}^2\right) + {\theta_1}^2 E({\zeta_{t-1}}^2) + {\theta_2}^2 E({\zeta_{t-2}}^2) + \cdots + {\theta_q}^2 E({\zeta_{t-q}}^2) + \theta_i \theta_j E(\varepsilon_{t-i} \varepsilon_{t-j}) + \cdots - \theta_s E(\varepsilon_t \varepsilon_{t-s}\Big) - \cdots$$

$$Var(Y_t) = \sigma_\zeta^2 (1 + \sum_{j=1}^{q} {\theta_j}^2)$$

So the variance is stable and independent of time.

For the covariance :

$$\boldsymbol{cov}(Y_t, Y_{t-k}) = E((Y_t - \mu)(Y_{t-k} - \mu))$$

$$= \mathrm{E}(\zeta_t - \theta_1\zeta_{t-1} - \theta_2\zeta_{t-2} - \cdots - \theta_q\zeta_{t-q})(\zeta_{t-k} - \theta_1\zeta_{t-1-k} - \theta_2\zeta_{t-2-k} - \cdots - \theta_q\zeta_{t-q-k})$$

$$= E\left(-\theta_k{\zeta_{t-k}}^2 + \theta_1\theta_{k+1}{\zeta_{t-1-k}}^2 + \theta_2\theta_{k+2}{\zeta_{t-2-k}}^2 + \cdots + \theta_q\theta_{q-k}{\zeta_{t-q}}^2 + \theta_i\theta_j\zeta_{t-i}\zeta_{t-j}\right) i \neq j$$

$$if \;\; K = 1,2,3,,,,q \qquad cov(Y_t, Y_{t-k}) = {\zeta_t}^2(-\theta_k + \theta_1\theta_{k+1} + \theta_2\theta_{k+2} + \cdots + \theta_q\theta_{q-k})$$

$$if \; K > q \qquad \boldsymbol{cov}(Y_t, Y_{t-k}) = 0$$

This shows that the covariance between values of the series at different time points is also constant and does not depend on the specific time at which the observations are made.

Thus, we can conclude that moving average models are stable by definition since they are sums of random shocks.

2-2- autoregressive process AR(p):

An autoregressive process of order *p* describes a time series Y_t where the current observation is modeled as a linear combination of its *p* preceding values:

$$AR(p): y_t = \varphi_0 + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \cdots + \varphi_p y_{t-p} + \zeta_t$$

Where :

ζ_t: is white noise that follows a normal distribution.

$\varphi_1, \varphi_2, , ,$: are autoregressive coefficients

Example 2:

$$AR(1): y_t = \varphi_0 + \varphi_1 y_{t-1} + \zeta_t$$

$$AR(2): y_t = \varphi_0 + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \zeta_t$$

The autoregressive model is typically expressed using the lag operator (L) where:

$$L^p y_t = y_{t-p}$$

And we write :

$$Y_t = \varphi_0 + \varphi_1 L y_t - \varphi_2 L^2 y_t + \cdots + \varphi_p L^p y_t + \zeta_t$$

$$Y_t - \varphi_1 L y_t - \varphi_2 L^2 y_t - \cdots - \varphi_p L^p y_t = \varphi_0 + \zeta_t$$

$$(1 - \varphi_1 L - \varphi_2 L^2 - \cdots - \varphi_p L^p) y_t = \varphi_0 + \zeta_t$$

$$\varphi_p(L) y_t = \varphi_0 + \zeta_t$$

Where :

$$\varphi_p(L) = 1 - \varphi_1 L - \varphi_2 L^2 - \cdots - \varphi_p L^p$$

Stationarity of autoregressive process AR(p) :

For the AR(p) model to be stable, it must meet the conditions of stationarity, which include the constancy of the mean, variance, and covariance over time[1] :

[1] Marc S.Paolella, op.cit**,** P188.

The expected value of the series $y_t$is:

$$E(y_t) = \varphi_0 + \varphi_1 E(y_{t-1}) + \varphi_2 E(y_{t-2}) + \cdots + \varphi_p E(y_{t-p}) + E(\zeta_t)$$

$$\mu = \varphi_0 + \varphi_1 \mu + \varphi_2 \mu + \cdots + \varphi_p \mu$$

$$\mu = \frac{\varphi_0}{1 - \sum_{i=1}^{p} \varphi_i}$$

For the mean to be constant and independent of time, it must be finite. Additionally, the condition must hold that :

$$\sum_{i=1}^{p} \varphi_i < 1$$

This condition is necessary but not sufficient to ensure stationarity, as there are other conditions that must also be met.

As for the variance of the series y_t:

$$Var(\, Y_t \,) = E(\, Y_t - \mu)^2$$
$$= E\left(\varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \cdots + \varphi_p y_{t-p} + \zeta_t\right)^2$$

Therefore, we can say that for the autoregressive model AR(p) to be stable, it must be expressed in a form where the sum of the coefficients is greater than one. The roots of the characteristic polynomial must lie outside the unit circle.

2-3- Autoregressive Moving Average (ARMA (p, q)) Models:

A series Y_t is said to follow a mixed process combining moving averages and autoregression if:

$$ARMA(p,q): y_t = \gamma + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \cdots + \varphi_p y_{t-p} + \theta_1 \zeta_{t-1} + \theta_2 \zeta_{t-2} + \cdots + \theta_q \zeta_{t-q} + \zeta_t$$

Where :

ζ_t: is white noise that follows a normal distribution.

$\varphi_1, \varphi_2, , \varphi_p, \theta_1, \theta_2, , , \theta_q$: Are the model coefficients

Example 3:

$$ARMA(1,1): y_t = \gamma + \varphi_1 y_{t-1} + \theta_1 \zeta_{t-1} + \zeta_t$$

$$ARMA(2,2): y_t = \gamma + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \theta_1 \zeta_{t-1} + \theta_2 \zeta_{t-2} + \zeta_t$$

And we can write the Autoregressive Moving Average model using the lag operator as follows :

$$Y_t - \varphi_1 L y_t - \varphi_2 L^2 y_t - \cdots - \varphi_p L^p y_t = \gamma + \zeta_t - \theta_1 \zeta_{t-1} - \theta_2 \zeta_{t-2} \ldots - \theta_q \zeta_{t-q}$$

$$\left(1 - \varphi_1 L - \varphi_2 L^2 - \cdots - \varphi_p L^p\right) y_t = \gamma + \left(1 - \theta_1 L - \theta_2 L^2 - \cdots - \theta_q L^q\right) \zeta_t$$

$$\varphi_p(L) y_t = \gamma + \theta_q(L) \zeta_t$$

III- Statistical Properties of Stationary Series:

A time series Y_t is deemed stable if it adheres to the following essential conditions[1] :

1- Time-Averaged Stability:

[1] Dimitrios Asteriou, Stephen G. Hall, **Applied Econometrics**, Second Edition, PALGRAVE MACMILLAN, the United Kingdom, 2011, p 267.

The expected value of the series remains constant over time, which can be expressed mathematically as:

$$E\ (Y_t) = E\ (Y_{t+h}) = \mu$$

This condition indicates that the mean of the series does not change, regardless of the time period considered, thus ensuring that the series does not exhibit a long-term upward or downward trend.

2- Stability of Variance Over Time:

The variance of the series must also be constant over time, represented as :

$$Var(Y_t) = Var\ (Y_{t+h}) = \sigma^2$$

This implies that the degree of variability in the series remains uniform across different time periods, which is crucial for valid statistical inference and modeling.

3- Independence of Covariance from Time:

The covariance between values of the series at different time points should not depend on the specific time at which the observations are made. This can be expressed as:

$$\boldsymbol{cov}(Y_t , Y_{t+h}) = E[(Y_t - \mu\ )(Y_{t+h} - \mu\ )]$$

This condition ensures that the relationship between observations is consistent over time, allowing for reliable predictions based on past values.

While graphical representations of the time series can provide initial insights into its stability, such as identifying trends or seasonal patterns, they are not sufficient for a comprehensive analysis. Therefore, it is imperative to employ specialized statistical tests, which will be discussed in subsequent sections, to rigorously assess the stability of the time series. These tests provide a more definitive evaluation, ensuring that any conclusions drawn from the data are robust and reliable.

4- White noise series (Blanc Bruit Un) :

The white noise series or white noise is characterized by a sequence of observations that are both independent and identically distributed (i.i.d.). This sequence is fundamental in time series analysis and is denoted by the symbol BB.

The defining properties of a white noise series Y_t are as follows[1] :

- ✓ **Zero Mean** :

The expected value of the series is constant and equal to zero for all time points :

$$E\ (Y_t) = 0 \qquad \forall t$$

This indicates that there is no systematic bias in the observations over time.

- ✓ **Constant Variance:**

The variance of the series remains constant across all time points, represented as:

$$Var(Y_t) = \sigma^2 \quad \forall t$$

This property ensures that the variability of the observations does not change over time, maintaining a consistent level of uncertainty.

- ✓ **Zero Covariance**:

[1] Eric Dor, **Econometrie**, pearson Education, France, 2009, p151.

The covariance between observations at different time points is zero when the time indices are different, expressed as:

$$\boldsymbol{cov}(Y_t, Y_{t+s}) = 0 \quad \forall t \neq s$$

This condition signifies that there is no linear relationship between the values of the series at different times, reinforcing the independence of the observations.

In general a white noise series can be formally expressed as:

$$Y_t \rightsquigarrow BB(0, \sigma^2)$$

This notation indicates that the series Y_t follows a white noise distribution with a mean of 0 and a constant variance σ^2. White noise serves as a crucial building block in time series modeling, often used as a benchmark to assess the randomness of a series and to formulate more complex models, such as autoregressive integrated moving average (ARIMA) models.

IV- Autocorrelation function:

The autocorrelation function (ACF), often denoted as ρ_h ,quantifies the relationship between a time series Y_t and its lagged values at different time intervals, or lags, denoted by h. Specifically, it measures how the current value of the series correlates with its past values[1].

Formally, the autocorrelation function is defined as:

$$\rho_h = \frac{cov(Y_t, Y_{t-h})}{\sigma_{Y_t} * \sigma_{Y_{t-h}}}$$

[1] Box. G. E. P, Jenkins. G. M, Reinse,op.cit, p 33.

This equation illustrates that the autocorrelation at lag h is the covariance between Y_t and Y_{t-h} normalized by the product of their standard deviations[1].

An alternative representation of the ACF can be expressed as :

$$r_h = \widehat{\rho_h} = \frac{\sum_{t=h+1}^{n}(Y_t - \overline{Y})\,(Y_{t-h} - \overline{Y})}{\sum_{t=1}^{n}(Y_t - \overline{Y})^2}$$

$\overline{Y}$: represents the mean of the series, and this formulation emphasizes the empirical calculation of the ACF based on observed data[2].

4-1 - Properties of the autocorrelation function :

The autocorrelation function possesses several important properties that are crucial for understanding its behavior[3]:

- Symmetry: The ACF is symmetrical around zero, meaning that :

$$\rho(h) = \rho(-h)$$

This indicates that the correlation between Y_t and Y_{t-h} is the same as that between Y_t and Y_{t+h}.

- Bounded Range: The values of the autocorrelation function are confined between -1 and 1, expressed as:

$$-1 \leq \rho_h \leq 1$$

A value of 1 indicates perfect positive correlation, while -1 indicates perfect negative correlation.

- Lag Zero: When the lag *h* is zero, the autocorrelation function equals 1, reflecting that any series is perfectly correlated with itself. Mathematically, this can be expressed as:

---

[1] Brockwell. P. J, Davis. R. A, op. cit, p 19.
[2] Enders. W, op.cit, p 67.
[3] Wei, op.cit, p 48.

$$\rho_0 = \frac{cov(Y_t, Y_t)}{\sigma_{Y_t} * \sigma_{Y_t}} = \frac{Var(\ Y_t\ )}{Var(\ Y_t\ )} = 1$$

The selection of lag h is critical for the analysis of time series data. When dealing with series containing fewer than 150 observations, it is recommended to limit the lag to $\frac{n}{6}$. Conversely, for series with more than 150 observations, a lag of $h = \frac{n}{5}$ is suggested.

In practice, the choice of lag also depends on the frequency of the data:

- For monthly or quarterly data, a common lag value is h=24;

- For daily data, the lag is typically set between 30 and 36;

- For annual data, a lag of 15 to 20 is appropriate.

The autocorrelation function is a fundamental tool in time series analysis, enabling researchers and analysts to explore the relationships within data over time. By understanding the properties and appropriate applications of the ACF, one can gain insights into the stability and patterns of time series components, ultimately enhancing the quality of forecasts and analyses.

4- 2- Autocorrelation function (ACF) confidence interval:

The autocorrelation function (ACF) is a fundamental tool in time series analysis, providing insights into the relationships between a series and its lagged values. Understanding the confidence intervals associated with the ACF is crucial for evaluating the statistical significance of these correlations, particularly in the context of large sample sizes[1].

4- 3- Distribution of the Autocorrelation Coefficient :

---

[1] Brockwell. P. J, Davis. R. A, op.cit, p 27.

For large samples, specifically when the sample size exceeds 30 observations, the distribution of the sample autocorrelation coefficient r_h can be approximated by a normal distribution centered around a mean of zero. This approximation is based on the central limit theorem, which asserts that, given a sufficiently large sample size, the sampling distribution of the mean will tend to be normally distributed, regardless of the shape of the population distribution[1].

The variance of the autocorrelation coefficient is expressed as:

$$var(r_h) = \frac{1}{n}\left(1 + 2\sum_{i=1}^{h} r^2{}_i\right)$$

Where :

n : is the total number of observations and

ri : represents the autocorrelation coefficients at various lags.

This formula indicates that the variance of rh not only depends on the sample size but also incorporates the cumulative effects of autocorrelations at lower lags, reflecting the interconnectedness of the time series data.

And we write :

$$r_h \rightsquigarrow N\left(0, var(r_h) = \frac{1}{n}\left(1 + 2\sum_{i=1}^{h} r_i^2\right)\right)$$

- ✓ **Constructing the Confidence Interval:**

At a 5% significance level, the confidence interval for the autocorrelation coefficient r_h can be constructed using the following formula:

[1] Hamilton, op.cit, p 111.

$$r_h \in \left[-1.96\sqrt{var(r_h)}, 1.96\sqrt{var(r_h)}\ \right]$$

Here, the critical value of 1.96 corresponds to the z-score for a two-tailed test at the 5% significance level. This interval provides a range of plausible values for the autocorrelation coefficient, allowing analysts to assess whether the observed autocorrelation is statistically significant.

- ✓ **Interpretation of the Confidence Interval :**

If the calculated values of r_h fall within this confidence interval, we fail to reject the null hypothesis, which posits that $\rho_h = 0$ at the 5% significance level. This outcome implies that there is no significant autocorrelation at lag (h), suggesting that past values do not have a meaningful influence on current values.

Moreover, when all correlation coefficients across various lags are contained within this confidence interval, it indicates that they are statistically insignificant. In practical terms, this suggests that the time series Y_t lacks memory, general trends, or seasonal components. Therefore, the series can be considered stable, with its values confined within a narrower range:

$$r_h \in \left[\frac{-1.96}{\sqrt{n}}, \frac{1.96}{\sqrt{n}}\ \right]$$

Where:

$\frac{-1.96}{\sqrt{n}}, \frac{1.96}{\sqrt{n}}$ **:** are the lower and upper bounds of the confidence interval, respectively. the value of 1.96 corresponds to the critical value from the standard normal distribution for a two-tailed test at the 5% significance level.

This range further emphasizes the stability of the series, as it indicates that the autocorrelation coefficients are not only insignificant but also close to zero, reinforcing the notion that the series behaves in a random manner without systematic patterns.

4- 4- Practical Implications for Time Series Analysis :

Understanding the confidence intervals for the autocorrelation function is vital for several reasons[1]:

1- Model Selection :

Analysts can use the significance of autocorrelation coefficients to determine appropriate models for forecasting. If significant autocorrelations are present, models such as ARIMA (Autoregressive Integrated Moving Average) may be warranted.

2- Data Stability:

The identification of statistically insignificant autocorrelations suggests that the time series is stable, which is a desirable property in many applications, including economic forecasting and quality control.

3- Hypothesis Testing:

The framework provided by confidence intervals allows researchers to conduct hypothesis testing regarding the presence of autocorrelation, facilitating more rigorous statistical analyses.

The confidence interval for the autocorrelation function is a critical aspect of time series analysis, allowing researchers to draw meaningful conclusions about the relationships within their data. By comprehensively understanding how to calculate and interpret these intervals, analysts can enhance the robustness of their analyses, ensure the validity of their models, and make informed decisions based on the underlying dynamics of the time series. This analytical rigor is essential for advancing knowledge and practices in fields ranging from economics to environmental science, where time-dependent data play a pivotal role.

[1] Box. G. E. P, Jenkins. G. M, op.cit, p p 93-96.

Example 4:

if we want to calculate the autocorrelation function values for the following series we will follow this steps:

T	1	2	3	4	5	6
Y_t	9	10	14	11	14	8

when h=1 :

t	Y_t	Y_{t-1}	$(Y_t - \overline{Y})$	$(Y_{t-1} - \overline{Y})$	$(Y_t - \overline{Y})(Y_{t-1} - \overline{Y})$	$(Y_t - \overline{Y})^2$
1	9		-2			4
2	10	9	-1	-2	2	1
3	14	10	3	-1	-3	9
4	11	14	0	3	0	0
5	14	11	3	0	0	3
6	8	14	-3	3	-9	9
$\overline{Y}$	11				-10	2

$$r_1 = \widehat{\rho_1} = \frac{\sum_{t=h+1}^{n}(Y_t - \overline{Y})\,(Y_{t-1} - \overline{Y})}{\sum_{t=1}^{n}(Y_t - \overline{Y})^2} = \frac{-10}{26} = -0.3$$

when h=2 :

t	Y_t	Y_{t-2}	$(Y_t - \overline{Y})$	$(Y_{t-2} - \overline{Y})$	$(Y_t - \overline{Y})(Y_{t-2} - \overline{Y})$
1	9		-2		
2	10		-1		
3	14	9	3	-2	-6
4	11	10	0	-1	0
5	14	14	3	3	9
6	8	11	-3	0	0
$\overline{Y}$	11				3

$$r_2 = \widehat{\rho_2} = \frac{\sum_{t=h+1}^{n}(Y_t - \overline{Y})\,(Y_{t-2} - \overline{Y})}{\sum_{t=1}^{n}(Y_t - \overline{Y})^2} = \frac{3}{26} = 0.09$$

when h=3 :

t	Y_t	Y_{t-3}	$(Y_t - \overline{Y})$	$(Y_{t-3} - \overline{Y})$	$(Y_t - \overline{Y})(Y_{t-3} - \overline{Y})$
1	9		-2		
2	10		-1		
3	14		3		
4	11	9	0	-2	0
5	14	10	3	-1	-3
6	8	14	-3	3	-9
$\overline{Y}$	11				-12

$$r_3 = \widehat{\rho_3} = \frac{\sum_{t=h+1}^{n}(Y_t - \overline{Y})\,(Y_{t-3} - \overline{Y})}{\sum_{t=1}^{n}(Y_t - \overline{Y})^2} = \frac{-12}{26} = -0.3$$

We follow the same steps to calculate r_4 and r_5

h	1	2	3	4	5
r_h	-0.3	0.09	-0.3	-0.09	0.1

V- Partial Autocorrelation Function (FACP):

The Partial Autocorrelation Function (PACF) is a fundamental statistical tool in time series analysis that measures the correlation between a current observation and its lagged values after eliminating the linear influence of the intermediate lags. Unlike the autocorrelation function (ACF), which reflects both direct and indirect dependencies, the PACF isolates the *direct* relationship between observations separated by a given lag.

The partial autocorrelation at lag h is denoted by $\emptyset_{hh}$ and is defined as follows:

When : h = 0

$$\widehat{\emptyset_{hh}} = \widehat{\emptyset_{00}} = 1$$

This indicates that the correlation of a variable with itself at lag 0 is always 1.

When : h=1

$$\widehat{\emptyset_{hh}} = \widehat{\emptyset_{11}} = r_1$$

Here r_1 represents the autocorrelation at lag 1, which measures the correlation between the current value and the immediately preceding value.

And when : h= 2,3,4,….

$$\widehat{\emptyset_{hh}} = r_{hh} = \left|\frac{P_h^*}{P_h}\right|$$

In this expression, P_h and P_h^* represent the determinants of specific matrices that encapsulate the autocorrelations of the time series.

Whereas:

The matrix *Ph* is constructed as follows:

$$|P_h| = \begin{vmatrix} 1 & r_1 & r_2 & . & . & . & r_{h-2} & r_{h-1} \\ r_1 & 1 & r_1 & r_2 & . & . & . & r_{h-2} \\ r_2 & r_1 & 1 & r_1 & r_2 & . & . & . \\ . & . & r_1 & . & r_1 & r_2 & . & . \\ . & . & . & r_1 & . & r_1 & r_2 & . \\ . & . & . & . & r_1 & 1 & r_1 & r_2 \\ r_{h-2} & r_{h-3} & . & . & . & r_1 & 1 & r_1 \\ r_{h-1} & r_{h-2} & . & . & . & . & r_1 & 1 \end{vmatrix}$$

This matrix encompasses the autocorrelations *ri* for lags up to $h-1$, reflecting the relationships between the time series values at different lags.

And the matrix *Ph∗* is defined similarly but includes modifications to account for the partial correlations:

$$|P_h^*| = \begin{vmatrix} 1 & r_1 & \mathrm{r}_2 & . & . & . & r_{h-2} & r_1 \\ r_1 & 1 & r_1 & \mathrm{r}_2 & . & . & . & r_2 \\ \mathrm{r}_2 & r_1 & 1 & r_1 & . & . & . & . \\ . & . & r_1 & . & r_1 & . & . & . \\ . & . & . & r_1 & . & r_1 & . & . \\ . & . & . & . & r_1 & 1 & r_1 & . \\ r_{h-2} & r_{h-3} & . & . & . & r_1 & 1 & r_{h-1} \\ r_{h-1} & r_{h-2} & . & . & . & . & r_1 & r_h \end{vmatrix}$$

This matrix is structured to isolate the direct effects of the lags, allowing for the calculation of the PACF.

The PACF is particularly useful in identifying the order of autoregressive (AR) models in time series analysis[1]:

- Significant PACF Values: If the PACF shows significant values up to a certain lag *p* and then tails off, this suggests that an autoregressive model of order *p* may be appropriate.

- Insignificant PACF Values: If the PACF values drop to insignificance after a few lags, it indicates that higher-order autoregressive terms may not be necessary.

The Partial Autocorrelation Function (PACF) is an essential component of time series analysis, providing insights into the direct relationships between current and past values while controlling for

[1] Enders. w, op.cit, p65.

intermediate influences. By understanding how to compute and interpret the PACF, analysts can make informed decisions regarding model selection and improve the accuracy of their forecasts. This tool enhances the robustness of time series modeling, particularly in identifying appropriate structures for autoregressive models.

5-1- Properties of the Partial Autocorrelation Function:

In the case of large samples, specifically when the sample size exceeds 30 observations, the distribution of the estimated PACF values r_{hh} can be approximated by a normal distribution with a mean of zero. This property allows for the construction of confidence intervals, which can be used to assess the significance of the PACF values and The PACF at lag h can be expressed as[1]:

$$r_{hh} \sim N(0, var(r_{hh}) = \frac{1}{n})$$

where *n* is the sample size. This indicates that as the sample size increases, the estimated PACF values will become more concentrated around the mean of zero.

The confidence interval for the PACF at a 5% significance level can be constructed using the following formula:

$$r_{hh} \in \left[\frac{-1.96}{\sqrt{n}}, \frac{1.96}{\sqrt{n}}\right]$$

Where:

$\frac{-1.96}{\sqrt{n}}, \frac{1.96}{\sqrt{n}}$ are the lower and upper bounds of the confidence interval, respectively. The value of 1.96 corresponds to the critical value from the standard normal distribution for a two-tailed test at the 5% significance level.

[1] Box. G. E. P, Jenkins. G. M, op.cit, p p 67-69.

Accepting the Null Hypothesis: If the calculated PACF values $\emptyset_{hh}$fall within this confidence interval, we fail to reject the null hypothesis, which states that $\emptyset_{hh} = 0$ at the 5% significance level. This implies that there is no significant partial autocorrelation at lag h.

Rejecting the Null Hypothesis: Conversely, if the PACF values fall outside of this interval, we reject the null hypothesis, suggesting that there is a significant partial autocorrelation at lag hh. This indicates that past values have a direct influence on the current value, warranting further investigation into the structure of the time series.

5-2- Practical Implications for Time Series Analysis:

The confidence interval for the PACF is essential for several reasons[1]:

1-Model Selection:

By identifying significant PACF values, analysts can determine the appropriate order of autoregressive terms in ARIMA models, enhancing the accuracy of forecasts.

2-Understanding Data Dynamics:

The PACF helps in understanding the underlying dynamics of the time series, particularly in distinguishing between direct and indirect relationships among observations.

3-Statistical Testing:

The framework provided by confidence intervals allows for rigorous hypothesis testing regarding the presence of partial autocorrelation, facilitating more robust statistical analyses.

The properties of the Partial Autocorrelation Function, particularly its confidence intervals, play a crucial role in time series analysis. By understanding how to calculate and interpret these

[1] Box. G. E. P, Jenkins. G. M, op.cit, p p 68-72.

intervals, analysts can draw meaningful conclusions about the relationships within their data, guiding model selection and enhancing the robustness of their analyses. This understanding is vital for effective forecasting and decision-making in various fields that rely on time-dependent data.

Example 5:

Based on the previous example, we will calculate the partial autocorrelation function as follows:

when h=1 we know that:

$$\widehat{\emptyset_{11}} = r_1 \text{ so } \widehat{\emptyset_{11}} = -0.31$$

When : h=2

$$\widehat{\emptyset_{22}} = r_{22} = \frac{\begin{vmatrix} 1 & r_1 \\ r_1 & r_2 \end{vmatrix}}{\begin{vmatrix} 1 & r_1 \\ r_1 & 1 \end{vmatrix}} = \frac{r_2 - {r_1}^2}{1 - {r_1}^2}$$

$$\frac{r_2 - {r_1}^2}{1 - {r_1}^2} = \frac{0.09 - (-0.31)^2}{1 - (-0.31)^2} = -0.004$$

When : h=3

$$\widehat{\emptyset_{33}} = r_{33} = \frac{\begin{vmatrix} 1 & r_1 & r_1 \\ r_1 & 1 & r_2 \\ r_2 & r_1 & r_3 \end{vmatrix}}{\begin{vmatrix} 1 & r_1 & r_2 \\ r_1 & 1 & r_1 \\ r_2 & r_1 & 1 \end{vmatrix}}$$

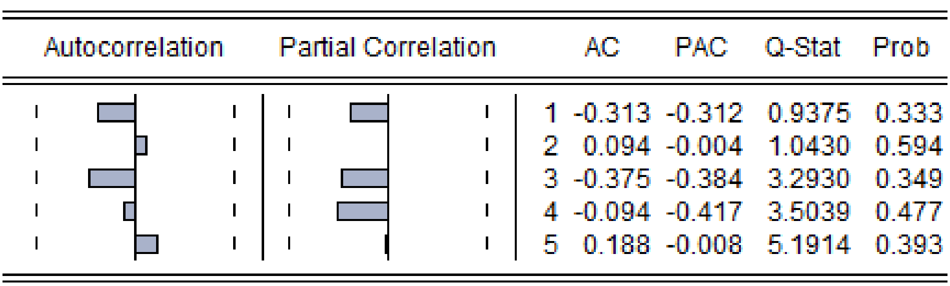
$$= \frac{1 * \begin{vmatrix} 1 & r_2 \\ r_1 & r_3 \end{vmatrix} - r_1 * \begin{vmatrix} r_1 & r_2 \\ r_2 & r_3 \end{vmatrix} + r_1 * \begin{vmatrix} r_1 & 1 \\ r_2 & r_1 \end{vmatrix}}{1 * \begin{vmatrix} 1 & r_1 \\ r_1 & 1 \end{vmatrix} - r_1 * \begin{vmatrix} r_1 & r_1 \\ r_2 & 1 \end{vmatrix} + r_2 * \begin{vmatrix} r_1 & 1 \\ r_2 & r_1 \end{vmatrix}}$$

$$= \frac{1*(r_3 - r_1 r_2) - r_1*(r_1 r_3 - r_2{}^2) + r_1*(r_1{}^2 - r_2)}{1*(1 - r_1{}^2) - r_1*(r_1 - r_1 r_2) + r_2*(r_1{}^2 - r_2)}$$

$$= \frac{r_1{}^3 - r_1 r_2(2 - r_2) + r_3(1 - r_1{}^2)}{1 - r_2{}^2 - 2r_1{}^2*(1 - r_2)}$$

$$r_{33} = -\frac{48}{125} = -0.38$$

To compute the Partial Autocorrelation Function (PACF) for lags to h=5, we follow the same procedure as previously described.and we also got the same results using the EViews 10 programme as it shows:



Date: 11/15/25 Time: 22:54
Sample: 2001 2006
Included observations: 6

	AC	PAC	Q-Stat	Prob
1	-0.313	-0.312	0.9375	0.333
2	0.094	-0.004	1.0430	0.594
3	-0.375	-0.384	3.2930	0.349
4	-0.094	-0.417	3.5039	0.477
5	0.188	-0.008	5.1914	0.393

5-3- Yule -Walker method (for calculating partial autocorrelation function):

The Yule-Walker equations provide a systematic approach to calculating the Partial Autocorrelation Function (PACF) for time series data. This method is particularly useful when dealing with larger lags, as it simplifies the calculations compared to the determinant method previously discussed. The Yule-Walker equations relate the autocorrelations of a time series to its parameters, enabling efficient computation of PACF values[1].

[1] Box. G. E. P, Jenkins. G. M, op.cit, p p 64-70.

1- Yule-Walker Equation for PACF :

The Yule-Walker equation for calculating the PACF at lag *i* can be expressed as follows:

$$\widehat{\emptyset_{ii}} = r_{ii} = \frac{r_i - \sum_{j=1}^{i-1}(r_{i-1}, j * r_{i-j})}{1 - \sum_{j=1}^{i-1}(r_{i-1}, j * r_{i-j})}$$

Where :

$i = 1,2,3,,,,h$;

j =1,2,3,,,, i-1;

$\widehat{\emptyset_{ii}}$:The estimated partial autocorrelation at lag *I*;

r_i: The autocorrelation at lag *I*;

r_{i-1}, j: The autocorrelation at lag *i*−1 and lag *j*.

$$\boldsymbol{r_{ij}} = r_{i-1}, j - r_{ii} r_{i-1,i-j}$$

2- Steps to Calculate PACF Using the Yule-Walker Method :

- ✓ **Calculate Autocorrelations:**

Compute the autocorrelation coefficients $r_1, r_{2,} r_{3,...,}$, r_h for the time series up to the desired maximum lag *h*.

- ✓ **Set Up the Yule-Walker Equations:**

For each lag *i* from 1 to *h*, apply the Yule-Walker equation:

For *i*=1:

$$\widehat{\emptyset_{11}} = r_1$$

For *i*=2:

$$\widehat{\emptyset_{22}} = \frac{r_2 - r_{11} * r_1}{1 - {r_{11}}^2}$$

For *i*=3:

$$\widehat{\emptyset_{33}} = \frac{r_3 - (r_{21} * r_2 + r_{22} * r_1)}{1 - ({r_{21}}^2 + {r_{22}}^2)}$$

And Continue this process for higher lags *i*.

3-Iterate for Higher Lags:

Repeat the calculation for each *i* until you reach the desired maximum lag *h*.

Example 6:

To illustrate the application of the Yule-Walker method, let's consider a hypothetical set of autocorrelation values:

Assume the following autocorrelation values:

*r*1=0.5, *r*2=0.3 and *r*3=0.2

Calculating PACF Values:

For *i*=1:

$$\widehat{\emptyset_{11}} = r_1 = 0.5$$

For *i*=2:

$$\widehat{\emptyset_{22}} = \frac{r_2 - r_{11} * r_1}{1 - {r_{11}}^2} = \frac{0.3 - 0.5 * 0.5}{1 - 0.5^2} = \frac{0.05}{0.75} = 0.067$$

For *i*=3:

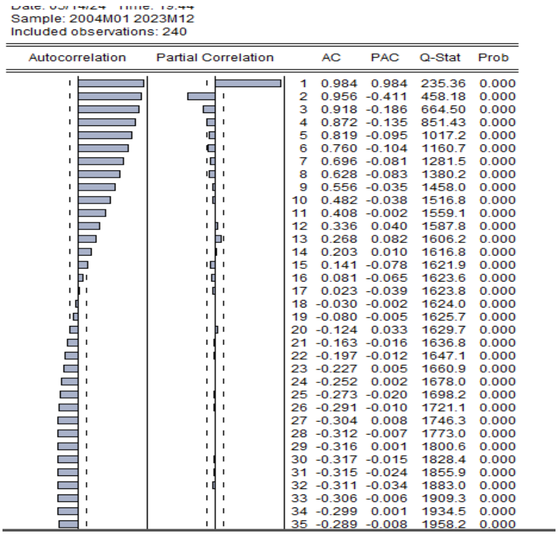
$$\widehat{\emptyset_{33}} = \frac{r_3 - (r_{21} * r_2 + r_{22} * r_1)}{1 - ({r_{21}}^2 + {r_{22}}^2)}$$

$$= \frac{0.2 - (0.5 * 0.3 + 0.067 * 0.5)}{1 - (0.5^2 + 0.067^2)} = 0.0221$$

The Yule-Walker method provides a systematic and efficient way to calculate the Partial Autocorrelation Function (PACF) for larger lags in time series data. By utilizing the autocorrelation coefficients and applying the Yule-Walker equations, analysts can derive PACF values that are essential for model selection and understanding the underlying dynamics of the time series. This method enhances the robustness of time series analysis, particularly in autoregressive modeling contexts.

Exapmle 7:

In this example we will represent the Graphical representation of the autocorrelation and partial autocorrelation function (Correlogramme) and to make it we will use a Series of Algerian monthly inflation rates, set at 240 views from January 2004 to December 2023 (2001=100), obtained from the Bank of Algeria and available on its official website (https://www.bank-of-algeria.dz/, s.d.)



	AC	PAC	Q-Stat	Prob
1	0.984	0.984	235.36	0.000
2	0.956	-0.411	458.18	0.000
3	0.918	-0.186	664.50	0.000
4	0.872	-0.135	851.43	0.000
5	0.819	-0.095	1017.2	0.000
6	0.760	-0.104	1160.7	0.000
7	0.696	-0.081	1281.5	0.000
8	0.628	-0.083	1380.2	0.000
9	0.556	-0.035	1458.0	0.000
10	0.482	-0.038	1516.8	0.000
11	0.408	-0.002	1559.1	0.000
12	0.336	0.040	1587.8	0.000
13	0.268	0.082	1606.2	0.000
14	0.203	0.010	1616.8	0.000
15	0.141	-0.078	1621.9	0.000
16	0.081	-0.065	1623.6	0.000
17	0.023	-0.039	1623.8	0.000
18	-0.030	-0.002	1624.0	0.000
19	-0.080	-0.005	1625.7	0.000
20	-0.124	0.033	1629.7	0.000
21	-0.163	-0.016	1636.8	0.000
22	-0.197	-0.012	1647.1	0.000
23	-0.227	0.005	1660.9	0.000
24	-0.252	0.002	1678.0	0.000
25	-0.273	-0.020	1698.2	0.000
26	-0.291	-0.010	1721.1	0.000
27	-0.304	0.008	1746.3	0.000
28	-0.312	-0.007	1773.0	0.000
29	-0.316	0.001	1800.6	0.000
30	-0.317	-0.015	1828.4	0.000
31	-0.315	-0.024	1855.9	0.000
32	-0.311	-0.034	1883.0	0.000
33	-0.306	-0.006	1909.3	0.000
34	-0.299	0.001	1934.5	0.000
35	-0.289	-0.008	1958.2	0.000

Chapter Three: Exponential Smoothing Models for Time Series Forecasting

Time series smoothing serves as a critical preprocessing step in temporal data analysis, aimed at mitigating transient or stochastic fluctuations to facilitate robust analysis of underlying structural patterns. Within this paradigm, exponential smoothing emerges as a foundational methodology, employing a weighted moving average in which historical observations are assigned geometrically decaying weights. This approach systematically attenuates noise while preserving essential components such as trends, seasonality, and non-systematic variation, thereby enhancing the interpretability and predictive accuracy of time-series models.

I- Simple Exponential Smoothing:

Time series smoothing is performed to eliminate temporary or incidental fluctuations and study the series effectively. In this context, we refer to the exponential smoothing method. This method is used when the time series contains only the random component and shows an approximately constant behavior. and the model for this method is given by the following form[1]:

$$\widehat{y_t} = \alpha y_t + \alpha(1-\alpha)yt - 1 + \alpha(1-\alpha)^2 yt - 2 + \ldots + \alpha(1-\alpha)^n y_{t-n} \ldots (1)$$

This model can be summarized as follows :

$$\widehat{y_t} = \alpha \sum_{k=0}^{n} (1-\alpha)^k y_{t-k}$$

Where :

α : is the smoothing constant

[1] Hyndman. R. J , Athanasopoulos. G, op. cit, p145.

$(1-\alpha)^k$: is the weighting factor.

As K increases, the weight becomes smaller and less significant.

To simplify this equation :

1-Lag the obtained form by one period and multiply it by $(\mathbf{1}-\boldsymbol{\alpha})$:

$$(1-\alpha)\hat{y}_{t-1} = \alpha(1-\alpha)y_t + \alpha(1-\alpha)yt-1 + \alpha(1-\alpha)^2yt - 2 + \ldots + \alpha(1-\alpha)^n y_{t-n} \ldots (2)$$

2-when we subtract equation (1) from equation (2) we obtain :

$$\widehat{y_t} = \alpha y_t + (1-\alpha)\hat{y}_{t-1}$$

Notes :

1-The selection of the initial forecast value $\widehat{y_t}$ is a critical step in exponential smoothing, as it influences the trajectory of subsequent predictions. Two primary approaches are employed:

-Naïve Initialization: $\widehat{y_t} = y_1$

-Averaging Initialization: $\frac{\mathbf{1}}{\boldsymbol{k}}\sum_{i=1}^{k} y_t$

2-Averaging the first k observations mitigates the impact of potential outliers or measurement errors in the initial data points. This approach is favored in stable environments where early observations are assumed to reflect the underlying process more reliably ;

3-The weights form a geometric series, ensuring the forecast remains a convex combination of past observations. The decay $(1-\alpha)$

rate dictates the memory of the model, with smaller α values extending the influence of older data;

4-The parameter α is pivotal in balancing responsiveness and stability so The larger the value of α, the more sensitive the forecast is to recent fluctuations, making the forecast more responsive but potentially less stable. the smaller the value of α, the more weight is given to older observations, making the forecast smoother but less responsive to recent changes.

if **(α → 1)**: Emphasizes recent observations, ideal for volatile series with rapid shifts. But it may Overfit to noise, leading to erratic forecasts.

if (α → 0): Prioritizes historical patterns, suitable for stable systems with gradual evolution. but it may Lag behind genuine structural changes.

Therefore, suitable values for α are those lying between the two extremes $\alpha \in]0,1[$. and in this case, the predicted values are less affected by random fluctuations.

Simple Exponential Smoothing is a versatile and intuitive tool for forecasting stationary time series. While its simplicity is a strength, practitioners must carefully initialize forecasts and calibrate α to balance responsiveness and smoothness

Example 1:

The table below shows a set of recorded observations for a time series y_t.

-Forecast the next three periods using the Exponential Smoothing method. The smoothing constant α is equal to 0.3.

t	1	2	3	4	5	6	7	8
y_t	30	40	40	30	20	20	30	30

-To Forecast the next three periods using the Exponential Smoothing method we must follow these steps :

$$\widehat{y_t} = y_1 = 30$$

$$\widehat{y_2} = 0.3 * y_1 + 0.7 * \widehat{y_1} = 0.3 * 30 + 0.7 * 30 = 30$$

$$\hat{y}_3 = 0.3 * x_2 + 0.7 * \widehat{y_2} = 0.3 * 40 + 0.7 * 30 = 33$$

...

$$\widehat{x_9} = 0.3 * y_8 + 0.7 * \widehat{y_8} = 0.3 * 30 + 0.7 * 27.65 = 28.36$$

$$\widehat{x_{10}} = 0.3 * y_8 + 0.7 * \widehat{y_8} = 0.3 * 30 + 0.7 * 27.65 = 28.36$$

$$\widehat{x_{10}} = \widehat{x_{11}} = \widehat{x_{12}}$$

t	y_t	$\widehat{y_t}$	$e_t = y_t - \widehat{y_t}$
1	30	30	0
2	40	30,00	10,00
3	40	33,00	7,00
4	30	35,10	-5,10
5	20	33,57	-13,57
6	20	29,50	-9,50
7	30	26,65	3,35
8	30	27,65	2,35
9		28,36	
10		28,36	
11		28,36	

Example 2:

The following time series shows the evolution of the prices of a particular type of fuel.

t	01/2020	02/2020	03/2020	04/2020	05/2020	06/2020	07/2020	08/2020
Price	30	40	40	30	20	20	30	30

Using a short-term forecasting model, estimate fuel prices for the next three months, given that: $\alpha = 0.3$

Solution:

Initial value:

$$\hat{Y}_1 = Y_1 = 30$$

Then, the remaining values are estimated as follows:

$$\hat{Y}_2 = 0.3(40) + (1 - 0.3)(30) = 33$$

$$\hat{Y}_3 = 0.3(40) + (1 - 0.3)(33) = 35.1$$

$$\hat{Y}_4 = 0.3(30) + (1 - 0.3)(35.1) = 33.57$$

$$\hat{Y}_5 = 0.3(20) + (1 - 0.3)(33.57) = 29.499$$

$$\hat{Y}_6 = 0.3(20) + (1 - 0.3)(29.499) = 26.65$$

$$\hat{Y}_7 = 0.3(30) + (1 - 0.3)(26.65) = 27.65$$

$$\hat{Y}_8 = 0.3(30) + (1 - 0.3)(27.65) = 28.36$$

Forecast of fuel prices for 09/2020:

$$\hat{Y}_{9/2020} = 0.3\left(Y_{8/2020}\right) + 0.7\left(\hat{Y}_{8/2020}\right)$$
$$= 0.3(30) + 0.7(28.36) = 28.852$$

Forecast of fuel prices for 10/2020:

$$\hat{Y}_{10/2020} = 0.3(30) + 0.7(28.852) = 29.1964$$

Forecast of fuel prices for 11/2020:

$$\hat{Y}_{11/2020} = 0.3(30) + 0.7(29.1964) = 29.4374$$

II- Double Exponential Smoothing:

Double Exponential Smoothing (DES), formally recognized as Brown's linear exponential smoothing, constitutes a pivotal advancement in time series forecasting for datasets exhibiting linear trends. This method refines traditional exponential smoothing by incorporating a dual-stage smoothing mechanism to concurrently estimate level and trend components, while maintaining parsimony through a single smoothing parameter α. Below is a rigorous academic exposition of its theoretical foundations, operational mechanics, and empirical implications.

The DES framework posits a linear structural model with additive disturbances :

$$Y_t = a_t + b_t + \varepsilon_t$$

Where :

$$\varepsilon_t \sim BB(0, \sigma^2)$$

a_t: Latent level component at time , representing the baseline value;

b_t :Latent trend component, capturing systematic linear growth;

ε_t: White noise error term with constant variance σ^2.

This formulation assumes local linearity in the trend, making DES particularly suited for short-to-medium-term forecasting horizons where nonlinearities are negligible.

To apply this method, we proceed in **two stages**[1]:

Stage 1: Primary Smoothing (Level Adjustment)

The first smoothing operation generates an intermediate series , attenuating noise while preserving trend information :

$$\hat{y}_t = \alpha y_t + (1-\alpha)\hat{y}_{t-1}$$

And : $\hat{y}_t = y_1$

$\hat{y}_t$: serves as a noise-reduced proxy for the latent level a_t, with α modulating the trade-off between responsiveness and smoothness.

Stage 2: Secondary Smoothing (Trend Extraction)

A subsequent smoothing pass on $\hat{y}_t$ isolates the trend component:

$$\hat{\hat{y}}_t = \alpha y_t + (1-\alpha)\hat{\hat{y}}_{t-1}$$

And: $\hat{\hat{y}}_t = \hat{y}_t = y_1$

This stage mitigates **phase lag** inherent in single smoothing, enabling unbiased trend estimation.

[1]Roustant. O, **introduction aux séries chnologiques,axe méthodes statistique et application**, Ecole nationale supérieure des mines , saint-etienne, France, 2008, p 20.

- ✓ **Parameter Derivation and Trend Decomposition :**

The level a_t and trend b_t are analytically derived from the smoothed series:

$$a_t = 2\hat{y}_t - \hat{\hat{y}}_t$$

$$b_t = \frac{\alpha}{1-\alpha}(\hat{y}_t - \hat{\hat{y}}_t)$$

$$\lambda = \frac{\alpha}{1-\alpha}$$

where:

a_t: corrects for the smoothing-induced lag by extrapolating the difference between single and double-smoothed values.

b_t: scales the inter-smoothing discrepancy to quantify the instantaneous trend velocity .

- ✓ **Forecasting :**

The estimated values for the series are obtained by:

$$\widehat{y_{t+k}} = a_t + b_t.\mathrm{k}$$

The forecasted values are calculated as follows :

$$y^p_{t-1,t} = a_t + b_t; [t = 2, n+1]$$

$$y^p_{n,n+h} = a_{n+1} + b_{n+1} * h; [h = 2,3,4,,,]$$

Notes:

1-Forecasting with Double Exponential Smoothing is based on a straight line. Its coordinates are the latest estimated values of a_t and b_t

This forecast can be extended to future horizons k, but k should not be too large, as large values would lead to predictions far from the true values.

2-The appropriate value for the smoothing coefficient α is the one that minimizes the sum of squared errors of past forecasts:

$$Min\sum_{t=1}^{n}(y_t - \widehat{y_t})^2 = Min\sum_{t=1}^{n}(\varepsilon_t)^2$$

The linearity assumption renders DES unsuitable for long-term forecasts as cumulative errors in trend extrapolation amplify exponentially .

Example 1:

The table below shows a set of recorded observations for a time series y_t.

-Forecast the next three periods using the Double Exponential Smoothing method. The smoothing constant α is equal to 0.5.

t	1	2	3	4	5	6	7	8
xt	10	20	20	30	40	40	50	50

-To Forecast the next three periods using the Double Exponential Smoothing method we must follow these steps :

1-

$$\widehat{x_1} = x_1 = 10$$

$$\widehat{x_2} = 0.5 * x_1 + 0.5 * \widehat{x_1} = 0.5 * 10 + 0.5 * 10 = 10$$

$\widehat{x_3} = 0.5 * x_2 + 0.5 * \widehat{x_2} = 0.5 * 20 + 0.5 * 10 = 15$

…

$\widehat{x_9} = 0.5 * x_8 + 0.5 * \widehat{x_8} = 0.5 * 50 + 0.5 * 42.97 = 46.49$

2-

$\widehat{\widehat{x_1}} = \widehat{x_1} = x_1 = 10$

$\widehat{\widehat{x_1}} = 0.5 * \widehat{x_2} + 0.5 * \widehat{\widehat{x_1}} = 0.5 * 10 + 0.5 * 10 = 10$

$\widehat{\widehat{x_1}} = 0.5 * \widehat{x_3} + 0.5 * \widehat{\widehat{x_2}} = 0.5 * 15 + 0.5 * 10 = 12.5$

…

$$\widehat{\widehat{x_9}} = 0.5 * \widehat{x_8} + 0.5 * \widehat{\widehat{x_8}} = 0.5 * 46.69 + 0.5 * 36.88 = 41.69$$

3- we have :

$$\lambda = \frac{\alpha}{1-\alpha} = \frac{0.5}{1-0.5} = 1$$

so

$$b_t = (\widehat{y_t} - \widehat{\widehat{y_t}})$$

4-

$x_{8,ç}^{p} = a_9 + b_9 = 51.29 + 4.80 = 56.09$

$x_{8,10}^{p} = a_9 + b_9 * 2 = 51.29 + 4.80 * 2 = 60.89$

$x_{8,11}^{p} = a_9 + b_9 = 51.29 + 4.80 * 3 = 65.69$

$x_{8,12}^{p} = a_9 + b_9 = 51.29 + 4.80 * 4 = 70.49$

t	x_t	$\widehat{x_t}$	$\widehat{\widehat{x_t}}$	a_t	b_t	$x^p_{t-1,t}$
1	10	10	10			
2	20	10,00	10,00	10,00	0,00	10,00
3	20	15,00	12,	17,50	2,50	20,00
4	30	17,50	15,00	20,00	2,50	22,50
5	40	23,75	19,38	28,12	4,37	32,49
6	40	31,88	25,63	38,13	6,25	44,38
7	50	35,94	30,79	41,09	5,15	46,24
8	50	42,97	36,88	49,06	6,09	55,15
9		46,49	41,69	51,29	4,80	56,09
10						60,89
11						65,69
12						70,49

Example 2:

Consider the following table:

t	1	2	3	4	5	6	7	8	9	10	11	12
Y_t	10	20	20	30	40	40	50	50				

Estimate the forecast for the remaining periods in the table using the double exponential smoothing method, given that α = 0.5.

Solution:

Step 1 – Calculation of $\widehat{Y}_t$: The initial value is $\hat{Y}_1 = Y_1 = 10$

$$\hat{Y}_2 = 0.5(20) + 0.5(10) = 15$$

$$\hat{Y}_3 = 0.5(20) + 0.5(15) = 17.5$$

$$\hat{Y}_4 = 0.5(30) + 0.5(17.5) = 23.75$$

$$\hat{Y}_5 = 0.5(40) + 0.5(23.75) = 31.875$$

$$\hat{Y}_6 = 0.5(40) + 0.5(31.875) = 35.93$$

$$\hat{Y}_7 = 0.5(50) + 0.5(35.93) = 42.96$$

$$\hat{Y}_8 = 0.5(50) + 0.5(42.96) = 46.48$$

Step 2 – Calculation of $\hat{\hat{Y}}_t$: The initial value is given by:

$$\hat{\hat{Y}}_1 = \hat{Y}_1 = Y_1 = 10$$

$$\hat{Y}_2 = 0.5(15) + 0.5(10) = 12.5$$

$$\hat{Y}_3 = 0.5(17.5) + 0.5(12.5) = 15$$

$$\hat{Y}_4 = 0.5(23.75) + 0.5(15) = 19.375$$

$$\hat{Y}_5 = 0.5(31.875) + 0.5(19.375) = 25.625$$

$$\hat{Y}_6 = 0.5(35.93) + 0.5(25.625) = 30.77$$

$$\hat{Y}_7 = 0.5(42.96) + 0.5(30.77) = 36.86$$

$$\hat{Y}_8 = 0.5(46.48) + 0.5(36.86) = 41.67$$

Third – Calculation of a_t:

$$a_t = 2\hat{Y}_t - \hat{\hat{Y}}_t$$

$$a_1 = 2(10) - (10) = 10$$

$$a_2 = 2(15) - (12.5) = 17.5$$

$$a_3 = 2(17.5) - (15) = 20$$

$$a_4 = 2(23.75) - (19.375) = 28.125$$

$$a_5 = 2(31.875) - (25.625) = 38.125$$

$$a_6 = 2(35.93) - (30.77) = 41.09$$

$$a_7 = 2(42.96) - (36.86) = 49.05$$

$$a_8 = 2(46.48) - (41.67) = 51.29$$

Fourth – Calculation of b_{t}:

$$b_{\mathrm{t}} = \frac{\alpha}{1-\alpha}\left(\widehat{\mathrm{Y}}_{\mathrm{t}} - \hat{\hat{Y}}_{\mathrm{t}}\right)$$

$$b_1 = \frac{\alpha}{1-\alpha}\left(\widehat{\mathrm{Y}}_1 + \hat{\hat{Y}}_1\right) = \frac{0.5}{1-0.5}(10-10) = 0$$

$$b_2 = (15) - (12.5) = 2.5$$

$$b_3 = (17.5) - (15) = 2.5$$

$$b_4 = (23.75) - (19.375) = 4.375$$

$$b_5 = (31.875) - (25.625) = 6.25$$

$$b_6 = (35.93) - (30.77) = 5.16$$

$$b_7 = (42.96) - (36.86) = 6.1$$

$$b_8 = (46.48) - (41.67) = 4.81$$

Fifth – Calculation of the estimated series:

$$\mathrm{Y}_1^{\mathrm{P}} = \mathrm{a}_1 + \mathrm{b}_1 = 10 + 0 = 10$$

$$\mathrm{Y}_2^{\mathrm{P}} = 17.5 + 2.5 = 20$$

$$\mathrm{Y}_3^{\mathrm{P}} = 20 + 2.5 = 22.5$$

$$\mathrm{Y}_4^{\mathrm{P}} = 28.125 + 4.375 = 32.495$$

$$\mathrm{Y}_5^{\mathrm{P}} = 38.125 + 6.25 = 44.375$$

$$\mathrm{Y}_6^{\mathrm{P}} = 41.09 + 5.16 = 46.25$$

$$\mathrm{Y}_7^{\mathrm{P}} = 49.05 + 6.1 = 55.16$$

$$\mathrm{Y}_8^{\mathrm{P}} = 51.29 + 4.81 = 56.1$$

Sixth – Calculation of the forecasted values (completion of the table):

$$Y_{8,9}^{\mathrm{p}} = 51.29 + 4.81 \times 2 = 60.91$$

$$Y_{8,10}^{\mathrm{p}} = 51.29 + 4.81 \times 3 = 65.72$$

$$Y_{8,11}^{\mathrm{p}} = 51.29 + 4.81 \times 4 = 70.53$$

$$Y_{8,12}^{\mathrm{p}} = 51.29 + 4.81 \times 5 = 75.34$$

We summarize the previous calculations in the following table:

t	Y_t	$\hat{Y}_t$	$\hat{\hat{Y}}_t$	a_t	b_t	Y_t^p
1	10	10	10	10	0	10
2	20	15	12.5	17.5	2.5	20
3	20	17.5	15	20	2.5	22.5
4	30	23.75	19.375	28.125	4.375	32. 5
5	40	31.875	25.625	38.125	6. 25	44.375
6	40	35.93	30.77	41.09	5.16	46.25
7	50	42.96	36.86	49.06	6.1	55.16
8	50	46.48	41.67	51.29	4.81	56. 1
9						60.91
10						65.72
11						70.53
12						75.34

III- Comparative Analysis between Exponential Smoothing and Double Exponential Smoothing :

Exponential Smoothing and Double Exponential Smoothing differ in several key operational and methodological aspects, particularly in their treatment of trends, parameter requirements, and forecasting capabilities. Exponential Smoothing operates under the assumption of stationarity, making it suitable for datasets without inherent trends by relying solely on a single smoothing parameter α to balance recent observations against historical averages. This simplicity grants it low computational complexity and effectiveness in short-term forecasting scenarios where trend dynamics are negligible. In contrast, Double Exponential Smoothing extends this framework by explicitly modeling linear trends through the introduction of a second parameter b, which governs trend adaptation alongside the level-smoothing α. This dual-parameter system allows it to capture sustained directional movements in data, thereby supporting medium-term forecasts through linear extrapolation. However, this enhanced capability comes at the cost of moderate computational complexity, requiring joint optimization of α and b to avoid overfitting while maintaining predictive accuracy. The methods' divergent approaches reflect their specialized applications: the former excels in stable environments like inventory management for routine products, while the latter proves valuable in contexts with persistent trends, such as sales projections for growing markets[1].

Aspect	Exponential Smoothing	Double Exponential Smoothing
trend	No trend (stationary series)	Linear trend
parameters	α	α, b
Forcast horizon	Short-term	Medium-term
Model complexity	Low	Moderate

Double Exponential Smoothing remains a cornerstone in linear trend forecasting due to its **computational efficiency** and **interpretability**. However, its reliance on linear extrapolation necessitates caution in long-horizon applications. Contemporary

[1] Hyndman. R. J , Athanasopoulos. G, op. cit, p p 28-41.

adaptations, such as damped trends or machine learning integrations, address these constraints while preserving DES's foundational principles. For practitioners, rigorous parameter calibration and initialization remain imperative to harness DES's full potential in dynamic environments.

IV- Holt's Model:

Holt's Exponential Smoothing Model, is a foundational technique in time series forecasting designed to handle data exhibiting a linear trend. Unlike Simple Exponential Smoothing, which assumes stationarity and focuses solely on smoothing the level of the series, Holt's model incorporates both the level (baseline value) and the trend (slope or rate of change), making it more robust for datasets with consistent upward or downward movements. This model was developed by Charles C. Holt in the 1950s and remains widely used in fields like economics, supply chain management, and sales forecasting. Below, it provide a comprehensive explanation, including its components, mathematical underpinnings, advantages, limitations, and practical considerations.

This model incorporates two parameters : the first for smoothing the level a_{0t} and the second for smoothing the slope $a_{1t}$with the smoothing coefficients constrained between zero and one.

Where :

Smoothing the level $a_{0t}$with smoothing coefficient α, it represents the smoothed estimate of the current value of the series at time;

Smoothing the slope $a_{1t}$with smoothing coefficient β, it represents the smoothed estimate of the slope or rate of change in the series at time t. This component allows the model to extrapolate future values based on the observed trend direction and magnitude;

The model uses two smoothing parameters to control how much weight is given to recent observations versus historical estimates :

α **(Level Smoothing Coefficient)**: Ranges from 0 to 1. A higher α (closer to 1) makes the level more responsive to recent data, while a lower α (closer to 0) emphasizes historical values for stability.

β **(Trend Smoothing Coefficient)**: Also ranges from 0 to 1. A higher β allows the trend to adapt quickly to changes, whereas a lower β stabilizes the trend against noise.

These parameters are constrained to $[0, 1]$ to ensure the smoothing process remains weighted and interpretable, preventing overemphasis on any single observation.

Note :

When $\alpha = \beta$ the Holt model is equivalent to the double exponential smoothing model by Brown.

The formulation of this model is as follows :

level update :

$$a_{0t} = \alpha y_t + (1 - \alpha)(a_{0t-1} + a_{1t-1})$$

Where:

$a_{0t-1} + a_{1t-1}$: Represents the one-step-ahead forecast from the previous period, adjusted by the trend.

Trend update :

$$a_{1t} = \beta(a_{0t} - a_{0t-1}) + (1 - \beta)a_{1t-1}$$

This is smoothed with the previous trend estimate, controlled by β.

The forecast for horizon h based is generated linearly as the following:

$$\hat{y}_{t+1} = a_{0t} + ha_{1t}$$

For longer horizons, the forecast extends the current level by adding the trend multiplied by the number of steps, assuming a constant linear trend.

This linear extrapolation constitutes a key feature of the method; however, it implicitly assumes that the underlying trend remains stable over time. Such an assumption may be violated in the presence of nonlinear dynamics or pronounced seasonal fluctuations. Consequently, the forecasting accuracy of the model may deteriorate when the data exhibit structural changes or complex patterns.

Initialization also plays a critical role in the overall performance of the model. In particular, the choice of initial values for the level and trend components significantly influences the quality of early forecasts, especially when the available historical data are limited. Poor initialization may lead to biased estimates and slow convergence toward the true underlying components.

The initial level is set to the first observation:

$$a_{01} = x_1$$

The initial trend is set to zero, assuming no prior knowledge of directional movement. (In practice, this can be adjusted based on domain knowledge

$$a_{11} = 0$$

Brown's Double Exponential Smoothing (DES) and Holt's Model both address trends in time series forecasting but differ in key areas. Brown's DES uses a single parameter $\alpha$for uniform smoothing of level and trend, implicitly handling trends via repeated operations and offering lower computational load with limited flexibility for linear trends only. In contrast, Holt's Method employs dual parameters (α and β) for independent level and trend control, explicitly modeling trends with a dedicated equation, requiring more complex joint optimization but allowing adaptability to damped or nonlinear trends through an additional damping parameter, enhancing versatility for varied data patterns[1].

Aspect	Brown's DES	Holt's Method
Parameterization	Single parameter (α)	Dual parameters (α,β)
Trend Dynamics	Implicitly derived via double smoothing	Explicitly modeled with separate trend equation
Computational Load	Lower (fewer parameters to optimize)	Higher (requires joint optimization of α,β)
Flexibility	Limited to linear trends	Adaptable to damped or nonlinear trends via ϕ

V- Holt-Winters Model:

The Holt-Winters Model, often referred to as Triple Exponential Smoothing or Winters' Method, stands as a foundational approach in time series forecasting, particularly for datasets exhibiting seasonal patterns. Pioneered by Charles C. Holt and Peter Winters in the 1960s, this method builds upon Holt's Double Exponential Smoothing by

[1] Hyndman. R. J , Athanasopoulos. G, op. cit, p 157.

introducing a third smoothing component dedicated to seasonality, thereby overcoming a key drawback of its predecessor: the inability to account for recurring periodic fluctuations, such as monthly sales cycles or quarterly demand spikes. This enhancement has made the Holt-Winters Model a staple in sales forecasting software, where it employs three distinct smoothing processes to refine predictions with remarkable precision and adaptability.

The model decomposes a time series into three additive components (assuming additive seasonality; multiplicative variants exist for proportional effects)[1]:

Level $\boldsymbol{a_{0t}}$: The smoothed baseline value of the series at time t, representing the underlying trend-adjusted average.

Trend $\boldsymbol{a_{1t}}$: The smoothed slope, capturing the linear growth or decline over time.

Seasonal Component S_t: The periodic variation, smoothed to reflect recurring patterns (higher sales in December). The parameter p denotes the seasonal period (12 for monthly data, 4 for quarterly).

Three smoothing parameters control the model's responsiveness:

Level smoothing coefficient $\boldsymbol{\alpha}$: weights recent observations in level updates

$$0 \leq \alpha \leq 1$$

Trend smoothing coefficient $\boldsymbol{\beta}$: adjusts trend adaptation

$$0 \leq \beta \leq 1$$

[1] Roustant. O, op.cit, p 20.

Seasonal smoothing coefficient γ : governs seasonal updates. Higher values make the model more reactive to recent seasonal changes.

$$0 \leq \gamma \leq 1$$

These parameters are optimized to minimize forecast errors, often via algorithms like nonlinear least squares or grid search.

The model formulation is as follows:

$$a_{0t} = \alpha\left(\frac{y_t}{S_{t-p}}\right) + (1-\alpha)(a_{0t-1} + a_{1t-1})$$

$$a_{1t} = \beta(a_{0t} + a_{0t-1}) + (1-\beta)a_{1t-1}$$

$$S_t = \gamma\left(\frac{y_t}{a_{0t}}\right) + (1-\gamma)S_{t-p}$$

Where:

The seasonal component is estimated by dividing the raw observation by the current level and smoothing it with past seasonal values.

The use of S_{t-p} in the level equation ensures the model uses the most recent available seasonal estimate, avoiding circularity.

The parameter (p) represents the period, equal to 12 for monthly data and 4 for quarterly data.

The forecast for horizon (h) is given as follows:

If $1 \leq h \leq p$ than :

$$\hat{y}_{t+1} = (a_{0t} + ha_{1t})S_{t-p+h}$$

If $1 \leq h \leq 2p$ than :

$$\hat{y}_{t+1} = (a_{0t} + ha_{1t})S_{t-p+2h} \quad 1 \leq h \leq p$$

With the knowledge that:

$$a_{0p} = \bar{y}$$

And

$$a_{1p} = 0$$

This technique was introduced by Holt and Winters. It consists of Holt's **double exponential smoothing** for the trend component, using two smoothing parameters, combined with **single exponential smoothing** for the seasonal component as proposed by Winters. Consequently, this method involves **three smoothing parameters**. In this context, it is necessary to distinguish between two cases: the **multiplicative form** and the **additive form**.

First: The Multiplicative Form

In this case, the time series is expressed as follows:

$$Y_t = (a_t + b_t t).S_t.e_t$$

Three different smoothing operations are applied to this model:

- ✓ Smoothing of the level a_t with smoothing parameter α_1, where $\alpha_1 \in (0,1)$;
- ✓ Smoothing of the trend coefficient b_t with smoothing parameter α_2, where $\alpha_2 \in (0,1)$;
- ✓ Smoothing of the seasonal component S_t with smoothing parameter α_3, where $\alpha_3 \in (0,1)$.

Model specification:

- ✓ ***Level smoothing*:**

$$a_t = \alpha_1\left(\frac{Y_t}{S_{t-p}}\right) + (1-\alpha_1)(a_{t-1}+b_{t-1})$$

- ✓ ***Trend smoothing*:**

$$b_t = \alpha_2(a_t - a_{t-1}) + (1-\alpha_2)b_{t-1}$$

- ✓ ***Seasonal smoothing*:**

$$S_t = \alpha_3\left(\frac{Y_t}{a_t}\right) + (1-\alpha_3)S_{t-p}$$

Forecasting:

$$Y_{t+k} = (a_t + Kb_t).S_{t-p+k} \quad For \quad 1 \le K \le P$$

Second: The Additive Form

In this case, the time series is written as:

$$Y_t = (a_t + b_t t) + S_t + e_t$$

Model specification:

- ✓ ***Level smoothing*:**

$$a_t = \alpha_1(Y_t - S_{t-p}) + (1-\alpha_1)(a_{t-1}+b_{t-1})$$

- ✓ ***Trend smoothing*:**

$$b_t = \alpha_2(a_t - a_{t-1}) + (1-\alpha_2)b_{t-1}$$

- ✓ ***Seasonal smoothing*:**

$$S_t = \alpha_3(Y_t - a_t) + (1-\alpha_3)S_{t-p}$$

The values α_1, α_2, and α_3 are chosen so as to **minimize the sum of squared residuals**, $\sum \varepsilon_t^2$.

Forecasting:

$$Y_{t+k} = a_t + b_t K + S_{t-p+k}, \qquad for\ \ 1 \le K \le P$$

$$Y_{t+k} = a_t + b_t K + S_{t-2p+k}, \qquad for\ \ P+1 \le K \le 2P$$

✓ **Starting point of the Holt–Winters model:**

There is a difficulty in determining the starting point of this technique, as is the case with other smoothing methods. The initial values can be estimated using the **ordinary least squares (OLS)** method, or alternatively, for the first year, one may proceed as follows:

1. **Initialization of seasonality:**

The seasonal component (or seasonal indices) for the first year can be estimated after distinguishing between the **multiplicative form** and the **additive form**, as follows:

- **First: Initial seasonal component for the multiplicative form:**

$$S_t = \frac{Y_t}{\overline{Y}}$$

where $\overline{Y}$ represents the average of the first P values (or the first P observations) of the first year.

- **Second: Initial seasonal component for the additive form:**

$$S_t = Y_t - \overline{Y}$$

2. Initialization of the level a:

$$a_{\mathrm{p}} = \overline{\mathrm{Y}}$$

3. Initialization of the trend coefficient b:
We assume that

$$\mathrm{b}_P = 0$$

that is, the trend values for the first year are set equal to zero.

Example 1:

Forecast the sales for the next 6 months of the fourth year using the Holt-Winters model with trend and seasonality knowing that :

$$\alpha = 0.3, \beta = 0.1, \gamma = 0.2$$

Year 1 – January	401,60	Year 2 – January	263,90	Year 3 – January	393,40
February	395,70	February	289,90	February	316,20
March	451,00	March	337,00	March	428,60
April	427,60	April	374,00	April	467,60
May	496,80	May	292,70	May	501,00
June	467,70	June	398,60	June	487,40
July	352,30	July	421,70	July	463,30
August	182,10	August	173,80	August	165,90
September	522,20	September	522,10	September	595,10
October	687,20	October	642,40	October	698,10
November	1080,3	November	984,20	November	1012,10
December	1391,6	December	1307,6	December	1380,00

We have :

$$a_{av-year\ 1} = \bar{y} = 571.34$$

$$S_t = \frac{427.6}{571.34} = 0.75$$

$$a_{0déc-year1} = 571.34$$

$$a_{1déc-year1} = 0$$

$$a_{1sep-year2} = 0.3 * \frac{522.1}{0.91} + 0.7 * (512.9 - 0.3) = 530.2$$

$$a_{1sep-year2} = 0.1 * \frac{530.2}{512.9} + 0.9 * (-0.3) = 1.5$$

$$S_{sep-year2} = 0.2 * \frac{522.1}{530.2} + 0.8 * (0.91) = 0.93$$

$$\hat{y}_{sep-year\ 4} = (512.9 + (-0.3) * 1) * 0.91 = 468.51$$

$$\hat{y}_{sep-year4} = (583.6 + 1.9 * 9) * 0.94 = 563.23$$

The results are shown in the following table :

t	y_t	a_{0t}	a_{1t}	S_t	$\hat{y}_t$
Year 1 – January	401,6			0,7	
F	395,7			0,69	
M	451			0,79	
A	427,6			0,75	
M	496,8			0,87	
J	467,7			0,82	
J	352,3			0,62	
A	182,1			0,32	
S	522,2			0,91	
O	687,2			1,2	
N	1080,3			1,89	
D	1391,6	571,3	0	2,44	
Year 2 – January	263,9	512,6	-5,9	0,67	

F	289,9	480,3	-8,5	0,67	351
M	337	458,3	-9,9	0,78	372
A	374	463,8	-8,3	0,76	336
M	292,7	419,8	-11,9	0,84	396
J	398,6	431,6	-9,5	0,84	334
J	421,7	500,6	-1,7	0,66	260
A	173,8	512,9	-0,3	0,32	159
S	522,1	530,2	1,5	0,93	469
o	642,4	532,4	1,6	1,2	639
N	984,2	529,9	1,1	1,88	1010
D	1307,6	532,8	1,3	2,44	1294
Year 3 – January	393,4	551,3	3	0,67	355
F	316,2	528,6	0,5	0,66	374
M	428,6	535,5	1,1	0,78	412
A	467,6	560,2	3,5	0,77	408
M	501	574,6	4,6	0,84	471
J	487,4	579,5	4,6	0,84	486
J	463,3	618,9	8,1	0,68	387
A	165,9	593,1	4,7	0,31	202
S	595,1	610,8	6	0,94	555
O	698,1	605,8	4,9	1,19	742
N	1012,1	588,6	2,7	1,85	1151
D	1380	583,6	1,9	2,42	1442
Year 4 – January					395
F					387
M					461
A					458
M					500
J					500

Example 2:

Use the additive form of the Holt–Winters model to forecast the quarterly sales of a given product, given that:

$$\alpha_3 = 0.3 \text{ ،} \alpha_2 = 0.1 \text{ ،} \alpha_1 = 0.4$$

	1	2	3	4	$\overline{Y}_i$	σ_i
2018	1248	1392	1057	3159	1714	842.69
2019	891	1065	1118	2934	1502	831.02
2020	1138	1456	1224	3090	1727	795.48
$\overline{Y}_j$	1092	1304	1133	3061	$\overline{\overline{Y}} = 1647.7$	
σ_i	149	171	69	94		$\bar{\sigma}_i = 829.74$

Forecast the sales for the four quarters of 2021 and the first quarter of 2022.

Solution:
Use the Holt–Winters exponential smoothing method.

Applied formulas:

Level smoothing a:

$$a_t = \alpha_1\left(Y_t - S_{t-p}\right) + (1 - \alpha_1)(a_{t-1} + b_{t-1})$$

For example:

$$a_7 = 0.4\left(1118 - (-657)\right) + 0.6[1488.95 + (-21.08)] = 1590.73$$

Trend smoothing b:

$$b_t = \alpha_2(a_t - a_{t-1}) + (1 - \alpha_2)b_{t-1}$$

For example:

$$b_7 = 0.1(1590.73 - 1488.95) + 0.9(-21.08) = -8.79$$

Seasonal smoothing S:

$$S_t = \alpha_3(Y_t - a_t) + (1 - \alpha_3)S_{t-p}$$

For example:

$$S_7 = 0.3(1118 - 1590.73) + 0.7(-657) = -601.72$$

After computing the seasonal components, it was observed that their sum does not equal zero; therefore, an adjustment is performed:

$$S_7^* = -601.72 - \left(\frac{-56.3}{4}\right) = -587.64$$

Sales estimation using the estimated model:

$$Y_{t+K} = a_t + b_t K + S^*_{t-p+k}$$

$$Y_7 = a_6 + b_6 K + S^*_{6-4+1}$$

$$Y_7 = 1488.95 + (-21.08) + (-657) = 810.88$$

Forecasting starting from the first quarter of 2021:

$$Y_{t+k} = a_t + b_t K + S_{t-p+k}, \qquad if \quad 1 \le K \le P$$
$$Y_{t+k} = a_t + b_t K + S_{t-2p+k}, \qquad if \quad P + 1 \le K \le 2P$$

$$\hat{Y}_{13} = 1696.85 + 1 \times 3.88 + (-512.67) = 1188.05$$
$$\hat{Y}_{14} = 1696.85 + 2 \times 3.88 + (-317.15) = 1387.46$$
$$\hat{Y}_{15} = 1696.85 + 3 \times 3.88 + (-579.33) = 1129.15$$
$$\hat{Y}_{16} = 1696.85 + 4 \times 3.88 + (1409.15) = 3121.51$$
$$\hat{Y}_{17} = 1696.85 + 5 \times 3.88 + (-512.67) = 1203.56$$

The following table presents the remaining results:

T	Year–Quarter	Y_t	a_t	b_t	S_t	S_t^*	$\hat{Y}_t$
1	2018−1	1248	−	−	−466	−466	−
2	2	1392	−	−	−322	−322	−
3	3	1057	−	−	−657	−657	−
4	4	3159	1714	0	1445	1445	−
5	2019−1	891	1571.2	−14.28	−530.26	−516.19	1248
6	2	1065	1488.95	−21.08	−352.59	−338.51	1234.92
7	3	1118	1590.73	−8.79	−601.72	−587.64	810.88
8	4	2934	1544.76	−12.51	1428.27	1442.34	3026.93
9	2020−1	1138	1581.03	−7.63	−504.09	−512.67	1016.06
10	2	1456	1661.84	1.21	−308.56	−317.15	1234.88
11	3	1224	1712.49	7.16	−570.75	−579.33	1075.41
12	4	3090	1696.85	3.88	1417.74	1409.15	3161.99
13	2021−1	−	−	−	−	−	1188.05
14	2	−	−	−	−	−	1387.46
15	3	−	−	−	−	−	1129.15
16	4	−	−	−	−	−	3121.51
17	2022−1	−	−	−	−	−	1203.56

References

References

1- Box G, Jenkins M, Reinsel C, Ljung G. M, ***Time Series Analysis: Forecasting and Control*** , 5th ed,Wiley, 2015.

2- Brockwell P. J, Davis R. A**, Introduction to Time Series and Forecasting** , 3rd ed, Springer, New York**,** 2016.

3- Chatfield C, Xing H, **The Analysis of Time Series: An Introduction with R**, (7th ed.), Chapman & Hall/CRC, New York, 2025.

4- Chatfield. C**, The Analysis of Time Series: An Introduction**, 6th ed, Chapman & Hall/CRC, United Kingdom, 2004.

5- Conover .W**, Practical Nonparametric Statistics**, 3rd ed, Wiley, United States, 1999.

6- Dickey. D. A, Fuller, W. A, **Distribution of the Estimators for Autoregressive Time Series with a Unit Root**, Journal of the American Statistical Association, 74(366), United States, 1979.

7- Dimitrios Asteriou, Stephen G. Hall, **Applied Econometrics**, Second Edition, PALGRAVE MACMILLAN, the United Kingdom, 2011.

8- Eric Dor, **Econometrie**, pearson Education, France, 2009.

9- Enders. W, **Applied Econometric Time Series**, 4th ed, Wiley, Hoboken, United States, 2015.

10- Gujarati. D, Porter. C, **Basic Econometrics**, 5th ed, McGraw-Hill, New York, United States, 2009.

11- Hyndman .R. J, Athanasopoulos . G**, Forecasting: Principles and Practice,** 2nd ed, OTexts, Australia, 2018.

12- Hyndman. R. J, Athanasopoulos. G, **Forecasting: Principles and Practice**, 3rd Edition; OTexts, Melbourne, Australia, 2021.

13- Hamilton. J. D, **Time Series Analysis**, Princeton University Press, Princeton, , United States, 1994.

References

14- Harvey. A. C, **Forecasting, Structural Time Series Models and the Kalman Filter**, Cambridge University Press, Cambridge, UK, 1990.

15- Kendall M. G, Ord J. K**, Time Series**, 3rd ed, Edward Arnold, United Kingdom, 1990.

16- Kendall. M. G, Stuart. A, **The Advanced Theory of Statistics**, *Vol. 3*, Griffin, 1976.

17- Marc S.Paolella, **Linear Models and Time series Analysis regression (Anova, Arma and Garch)**, wily series in probability and statistics,university of Zurich, Switzerland, 2019.

18- Makridakis. S, Wheelwright. S. C, Hyndman. R. J, **Forecasting: Methods and Applications,** 3rd ed, Wiley, 1998.

19- Roustant. O, **introduction aux séries chnologiques,axe méthodes statistique et application**, Ecole nationale supérieure des mines , saint-etienne, France, 2008.

20- Sharif Hossain,**Econometric Analysis an applied approach to business and economics**, Cambridge scholars publishing, United kingdom, 2024.

21- Wei W. S**, Time Series Analysis: Univariate and Multivariate Methods,** 2nd ed, Pearson Addison Wesley, United States, 2006.

22- Wooldridge J. M, **Introductory Econometrics: A Modern Approach**, 6th ed, Cengage Learning, 2016.