Chapter # 01

⇒ **Forecasting :**

Forecasting is defined as the process of making predictions about the future events based on past and present event data. It involves using historical data to estimate future outcomes or trends.

The "Forecasting : Principles and Practice," book emphasizes that forecasting is essential for decision-making in various fields such as economics, finance, marketing, and operation management.

Furthermore, the book highlights that forecasting is not about making precise predictions, but rather about providing estimates of future uncertainity. It acknowledges that forecasting errors are unavoidable due to the inherent randomness and complexity of many real world phenomena.

Overall, the definition of forecasting in this book emphasizes its practical importance in helping organizations and individuals anticipate future trends, making informed decisions, and plan effectively.

⇒ Getting Started :

1) Forecasts throughout History :
People have always been curious about what the future holds. Sometimes, predicting the future was seen as special and divine, while at other times, it was.

seen as illegal. Long ago, prophets or fortune tellers would try to guess what would happen based on signs like the appearance of a sheep's liver or by inhaling certain fumes.

2) Famous Mistakes in predictions:
Even Smart people can make big mistakes when guessing what the future will be like, e.g., some experts in the past said thinks like, "There will only be a need for a few computers in the world" or "Nobody would want a computer at home". These predictions turned out to be very wrong, as we now see computers everywhere -

i) Why good forecasting matters? the ability to predict the future is ~~so~~ really important for businesses. If a company can forecast well, it can make better decisions and have an advantage over its competitors. This book will teach reliable methods for making forecasts that have been tested and proven to work.

So, basically predicting the future has always been important, but it's also really hard to do. Businesses that can do it well have a big advantage.

1.18 What can be forecast?

Forecasting is necessary in various situations, such as predicting future demand for energy, scheduling staff, or managing inventory. It helps in effective planning and decision making.

*) Factors effecting Forecast Accuracy:
The predictability of an event depend on factors like our understanding of its causes, the availability of data, how similar the future is to the past, and whether the forecast can influence the event itself. eg: Short-term forecasts of electricity demand are usually accurate because we understand the factors driving demand, have plenty of data and can assume that past behaviors will continue.

*) Limitations of forecasting:
Some events, like currency

exchange rates, are harder to forecast because they may not fully understand the factors influencing them, and forecasts themselves can effect the outcomes. In such cases, forecasts may not be better than random guesses.

*) Adapting to changing environments:
Good forecasting models adapt to changing environments by capturing the way they are moving. They don't assume that the environment is static but rather that the changes observed will continue into the future.

1.2

*) Variety of forecasting methods:
Forcasting methods can range from simple to

complex, depending on factors like the availability of data and the predictability of the quantity being forecasted. Sometimes, when there is no data available, judgemental forecasting methods, where experts make educated guesses, are used.

Overall, forecasting is about understanding what can be predicted accurately and choosing the appropriate method to do so based on the available data and circumstances.

1.2 Forecasting, Goals and planning- It explains the importance of forecasting in business decision making and clarifies the distinction b/w forecasting, goals, and planning.

*) Forecasting ⇒ Predicting the future accurately based on available information, including historical data and upcoming events that could effect forecasts.

*) Goals ⇒ What a business aims to achieve, ideally based on forecasts and plans. However, goals are often set without clear plans or consideration of their feasibility.

*) Planning ⇒ Responding to forecasts and goals by determining necessary actions to align forecasts with goals, bridging the gap between aspiration and likely outcomes.

⇒ Different types of forecasts are needed fore various timeframes.

i) Short-term : Crucial for scheduling personnel, production, and transportation, and predicting demand.

ii) Medium-term : Necessary for planning resource requirements like raw materials, personnel and equipment.

iii) Long-term : Used for strategic planning, considering market opportunities, environmental factors and internal resources.

To develop an effective forecasting system, organizations must employ various approaches, including identifying forecasting problems, applying different methods, and continuously evaluting and improving forecasting methods. Strong organizational support is vital for succesful forecasting implementation.

1.3⇒ Determining what to forecast:
In the initial phases of a forecasting project, decisions must be made regarding the scope of the forecasts. This involves considerations such as:

i) Determining what to forecast: At the outset, decision must be made about what to forecast. For instance, in a manufacturing setting, one must decide if forecasts are needed for every product line or groups of products, every sales outlet or regions, and whether to use weekly, monthly, or annual data. The forecasting horizon (timeframe) and frequency of forecasts are also important considerations.

ii) Considerations for forecasting horizon and frequency: Different models are required based on the forecast horizon and frequency. ie. Short-term forecasts may require automated systems due to their frequent productions.

iii) Understanding Users Needs: It is crucial to communicate with stake-holders who will use the forecasts to understand their requirements and how they intend to use the forecasts. This ensures that the forecasts meet their needs effectively.

iv) Data Collection: Once the forecast requirements are determined, the next step is to gather or locate the necessary data. This data may already exist within the organization, such as sales records or historical demand data. However, locating ~~the~~ and collecting the data can be time consuming and may require effort before developing suitable forecasting methods.

In summary, the early stages of a forecasting projects involve making decisions about what to forecasts, considering factors like forecast horizon and frequency, understanding user needs and collecting the necessary data for analysis.

1.4 Forecasting data and methods:
The appropriate forecasting methods depend largely on what data are available.

*1. Data availability & Forecasting Methods:
The choice of forecasting methods depends largely on the availability of data. If no relevent data exist, qualitative forecasting methods must be used. These methods are not purely guesswork - there are well - developed structured approaches to obtain good forecasts without using historical data.

→ Quantitative forecasting: Quantitative forecasting method can be applied when numerical information about the past is available, and it is reasonable to assume that some aspects of past patterns will continue into the future. There is a wide range of quantitative methods, each with its own properties, accuracies, and costs.

↳ Time Series forecasting: This book primarily focuses on time series data, which are observed sequentially over time at regular intervals, e.g, includes sales results, rainfall and stock prices. Time series forecasting involves estimating how the sequence of observations will continue into the future. Popular ~~smoothing~~ methods include exponential smoothing and ARIMA models.

↳ Predicted variables: Predicted variables can be useful in time series forecasting, allowing for more explanatory variables. These variables like temperature, or economic indicators, help explain variations in the data. Mixed models, combine time series and predicted variables, offering a hybrid approach.

↳ Choosing forcasting models:
The choice b/w time series, explanatory or mixed models depend on factors such as the understanding of the system, the availability of future predictor values, the need for explanation vs prediction, and the accuracy of competing models.

In summary, the selection of forecasting method depends on the availability of data, the nature of the data patterns, and the specific requirements and constraints of the forecasting projects.

1.5 Some Case Studies :-

The following four cases highlights different challenges in forecasting.

↳ Case 1 : The clint, a large manufacture of disposable tableware, struggled with inaccurate forecasts generated by their in house software. The software used basic methods like averaging and simple regression, but failed to capture the complex patterns in their data. The challange was to improve the forecasting accuracy while working within limitations of the existing COBOL software.

↳ Case 2 : The Austrialian federal govt needed to forecast the annual budget for the Pharmaceutical benefit Scheme (PBS), which heavily relies on the sales volumes of various Pharmaceutical products. The challenge here was to develop a forecasting method that could handle trends, ~~seasonalality~~ seasonality, and sudden changes in sales volumes due to factors like changes in subsidized drugs or competitor prices.

→ Case 38 A car fleet company sought assistance in forecasting vehical resale values to better control profits. However their specialists were reluctant to adopt statistical models, viewing them as a threat to their jobs. The challenge was to develop accurate forecasts despite the lack of cooperation and to utilize the large amount of available data effectively.

4 Case 4 : A leading airline in Australia needed weekly forecasts of air passenger traffic on major domestic routes, considering factors like school holidays, sporting events and competition behavior. The historical data was complicated by factors such as a pilots strike, and launch and failure of a new airline, and changes in seat classification. The challenge was to develop a robust forecasting model that could handle the complexities of air travel demand.

In each case, the challenge was to develop forecasting methods that could accurately capture the underlying patterns in the data while addressing specific constraints and complexities unique to each situation.

16 ⇒ The basic steps in forecasting:

In a forecasting task, there are typically 5 fundamental steps.

1) Problem definition:

The initial step involves precisely defining the forecasting problem, understanding its purpose, stakeholders, and its integration within the organization's operation.

2) Gathering information:

This Phase entails collecting both statistical data and insights from domain experts who are involved in data collection and utilization of forecasts.

3) Preliminary Analysis :
Before collecting / selecting a forecasting model, it's essential to conduct exploratory analysis, including graphing data to identify pattern, trends, seasonality, and outliers.

4) Choosing and fitting models :
After understanding data characteristics, select a suitable forecasting model based on historical data availability, relationships b/w variables, and forecast requirements.

Compare multiple potential models and choose the most appropriate one.

5) Using and Evaluating a forecasting Models

Once a model is selected and parameters are estimated, utilize it to generate forecasts. Evaluate the model's performance by comparing forecasted values with actual values/data and assess forecast accuracy using various models methods. Consider practical issues like handling missing values, outliers, and organizational aspects related to implementing and acting on forecasts.

These steps provide a structured approach to conducting forecasting tasks, ensuring thorough analysis, appropriate model selecting and effective utilization of forecasts.

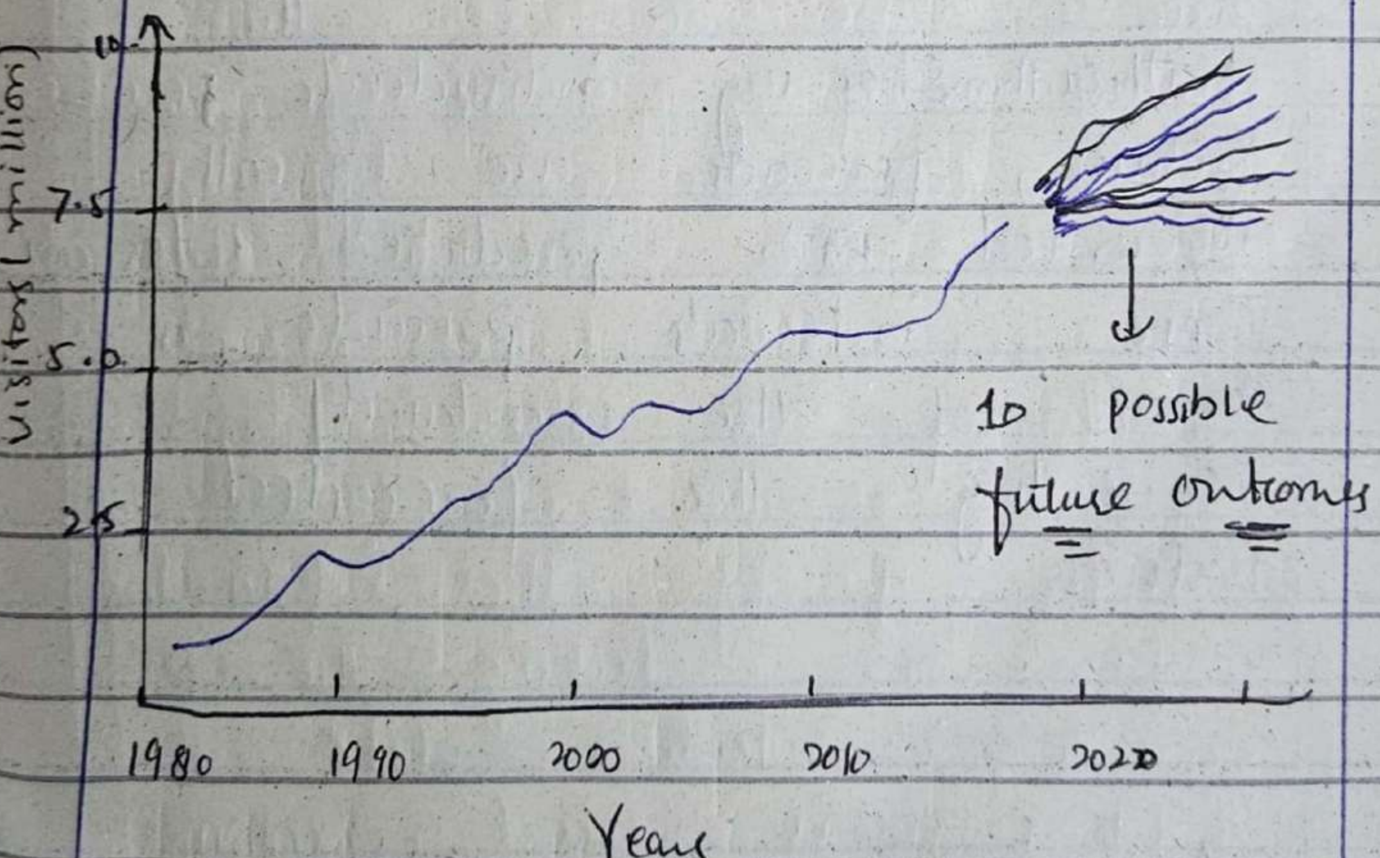
1.7 The statistical forecasting perspective

In forecasting, the think are are trying to predict is unknown and treated as a random variable, For instance, next months total sales could vary within a range of possible values until the actual sales are known at month-end.

As the forecast horizon (timeframe) extents further into the future, the uncertainity surrounding the forecasts increases. This is because there are more potential scenarios or outcomes.

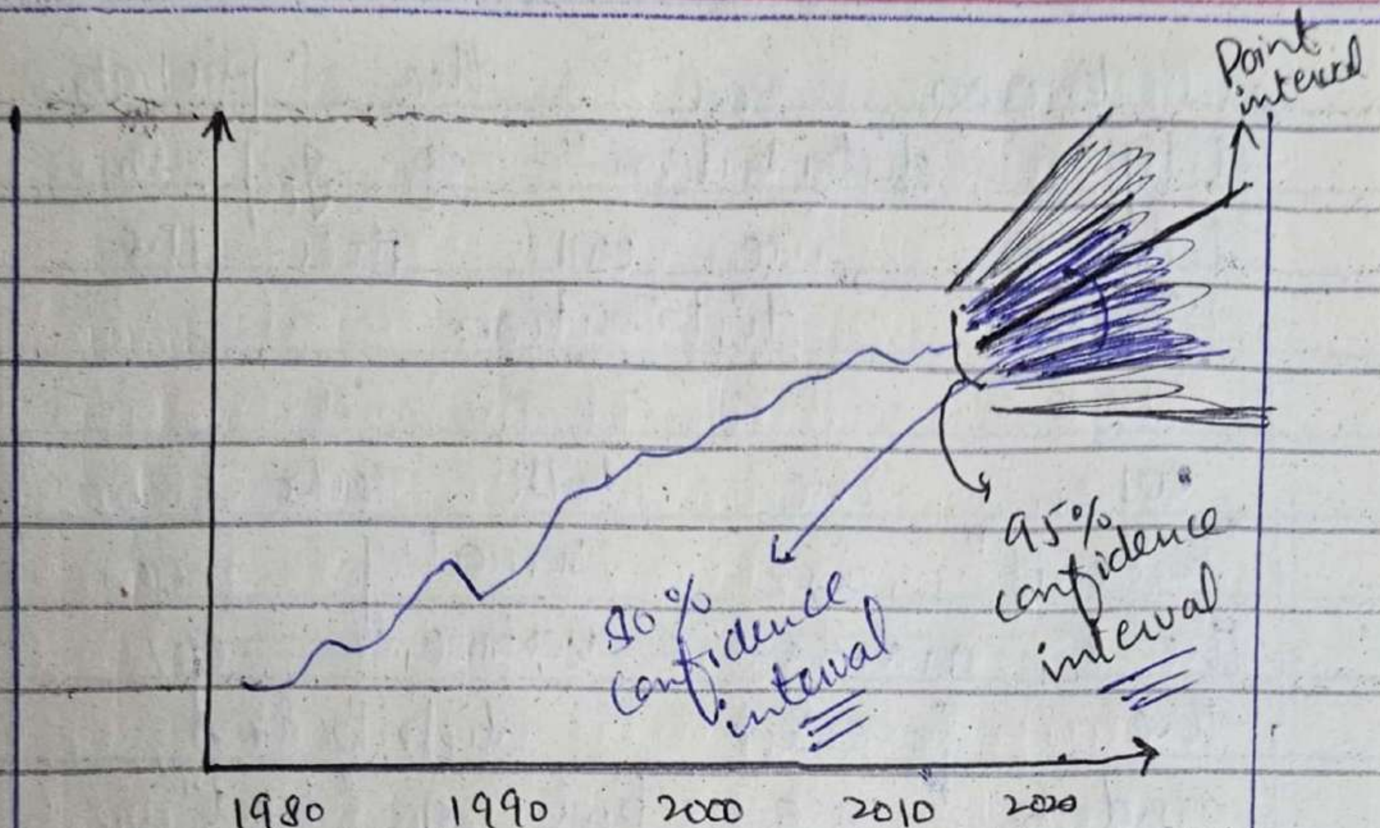
- we can imagine many possible futures, each yielding a different value for the thing we wish to forecast. Plotted in future are the total international visitors from 1980 - 2015. Also shown are ten possible future from 2016 - 2025.

Total international visitors to Australia.



Forecasting often involves estimating the middle of the range of possible values for a random variable. Prediction intervals are used to indicate the range of values the variable could like take within a certain level of probability, such as a 95% prediction interval.

Rather than showing individuals possible future, forecasts are typically presented with prediction intervals. These intervals provide a sense of the uncertainity surronding the forecasted value



we will use the transcript "t" for time i.e y_t will denote the observation at time t.

Suppose we denote all the information we have observed as and we want to forecast y_t. We then write $y_t|$ meaning "the random variable y_t given what we know in". The set of values that this random variable could take, along with their relative probabilities is known as the "probability distribution" of y_t |. In forecasting, we call this the forecast distribution

When we talk about "forecast", we usually mean the avarage value of the forecast distribution, and we put a "hat" over y to show this. Thus we write the forecast of y_t as $\hat{y}_t$, meaning the average of the possible values that y_t could take given everything we know

To specify the information used in making the forecasts, notation like $\hat{y}_{t|t-1}$ is used, meaning the forecast of y_t considering all previous observations up to time $t-1$. Similarly, $\hat{y}_{T+h|T}$ represents the forecasts of y_{T+h} considering all observations up to time T.

(Lecture 1) Macroeconomic forecasting

Macroeconomic forecasting is the process of predicting future economic conditions and trends on a national or global scale. It involves analyzing indicators like GDP growth, inflation, unemployment, and interest rates to anticipate economic changes. Methods include statistical models, historical data analysis, leading indicators, surveys, scenario analysis, and machine learning. This forecasting is vital for decision-making by policymakers, businesses, investors and individuals, but its subject to uncertainties and revisions due to economic fluctuations and unforeseen events.

⇒ Properties of time Series:

Properties of time series data include.

i) Trend : long term increase or decrease movement or direction.

ii) Seasonality : Predictable patterns at fixed intervals.

iii) Cyclical Variation : Fluctuations related to business cycle.

iv) Randomness / Noise : Irregular, unpredictable components.

v) Autocorrelation : Correlation b/w current and past values.

vi) Stationarity : Stable statistical properties over time. ie

mean = constant

variance = constant

covariance doesn't change with time / means it doesn't depend on time.

Understanding these properties is essential for analyzing and forecasting time series data effectively.

→ Univariate Analysis :

Univariate analysis is the simplest form of analyzing data. The term "univariate" signifies that we are dealing with one variable in the dataset. Unlike more complex analysis that explores relationships or cause, univariate analysis focus solely on describing patterns within a single variable.

→ Univariate analysis aims to understand the distribution of values for a single variable

⇒ Stochastic Process vs time Series :

i) Stochastic process :

A stochastic process is a mathematical concept defined as a family of random variables. These variables evolve over time, and their behavior is inherently random.

- Stochastic processes can be observed in both continuous time (where the random variables change smoothly) and discrete time (where the random variables change at specific intervals).
- The term "stochastic" emphasizes the uncertainity associated with the process.
- Example of stochastic processes include Poisson processes.

Where events happen continuously and independently at a constant rate within a given interval.

(ii) Time Series:

A time series is a specific

type of stochastic process.

- It consists of a sequence of observations collected at equally spaced time intervals (e.g., daily, monthly, yearly)

- Unlike stochastic processes, which can be continuous, time series are discrete-time by nature.

- Time series data often arise from real world observations, such as stock prices, temperature measurements, or economic indicators.

- Time series analysis involve studying patterns, trends and seasonality within these observed data points.

Summary:
A stochastic process is a collection of

random variables, while a time Series is a collection of actual numerical values (observations) obtained from a stochastic process.

⇒ Probability Density function:
The PDF represents the likelihood of a continuous random variable falling within a specific range of values. Here are some key points

i) Normal distribution: (Bell Curve)
The most common example of a PDF is the normal distribution. The Graph of Normal distribution is a symmetric bell-shaped curve.
The area under the curve b/w any two specified values represents the probability of a random variable falling

within that range.

ii) PDF for a specific variables

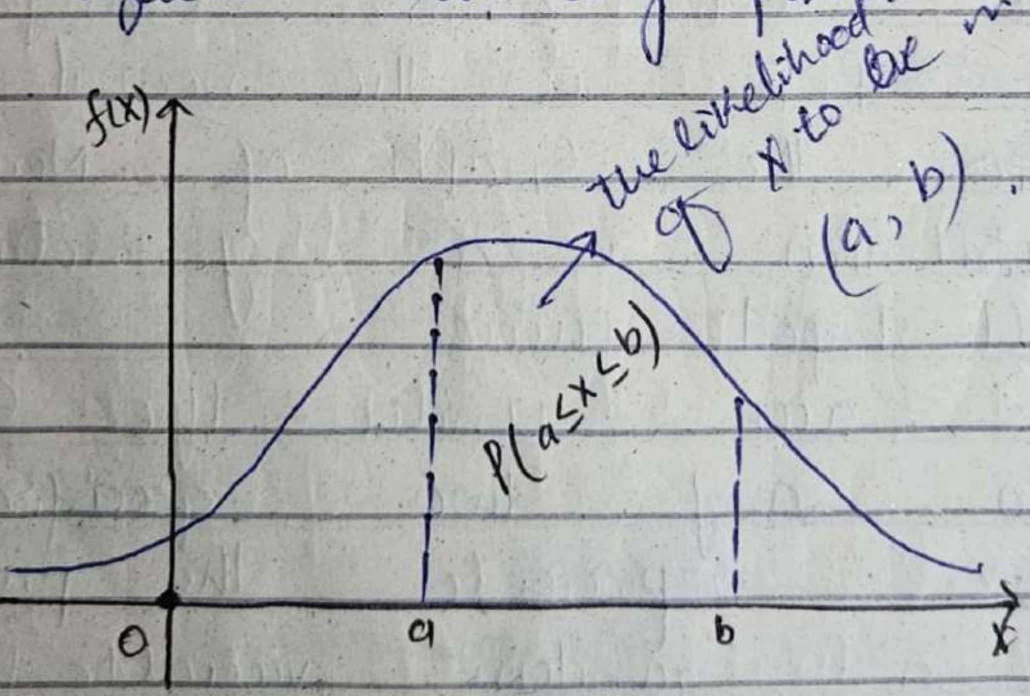
↳ Suppose we have a continuous random variable X.

↳ The PDF, denoted as f(X), provides the likelihood of X lying in an interval (a, b).

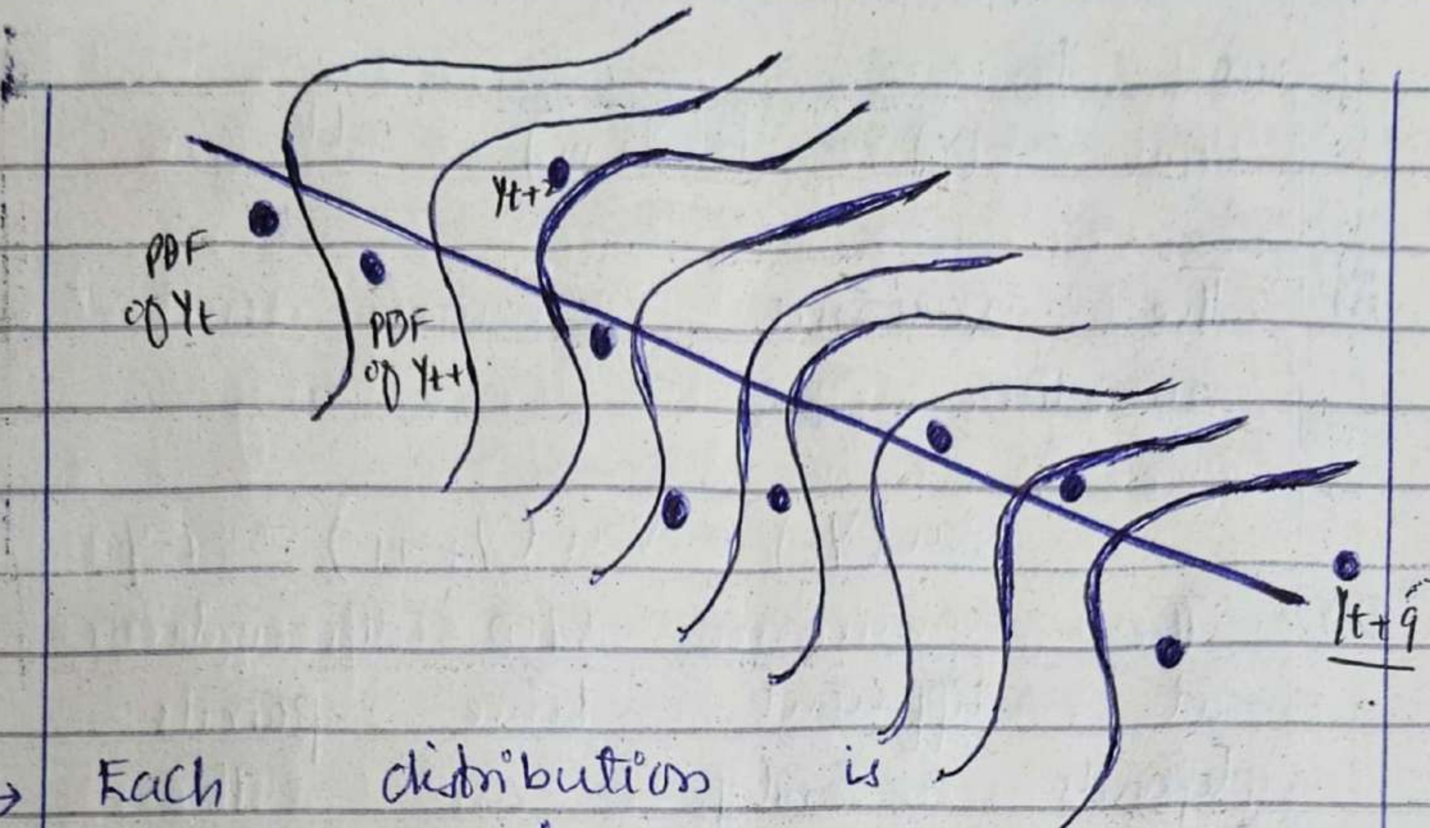
↳ The formula for calculating this probability is:

$$P(a < X < b) = \int_a^b f(x)\,dx$$

↳ The graph of f(x) is +ive or non-negative at any point.



Graph ⇒ 2



↳ Each distribution is a draw from a random process.

⇒ Two General "classes" of processes.

1) Stationary process :

A process is considered stationary if its statistical properties remain constant over time.

→ Key properties of a stationary process

i) The mean remains the same regardless of the time point. Mathematically we have

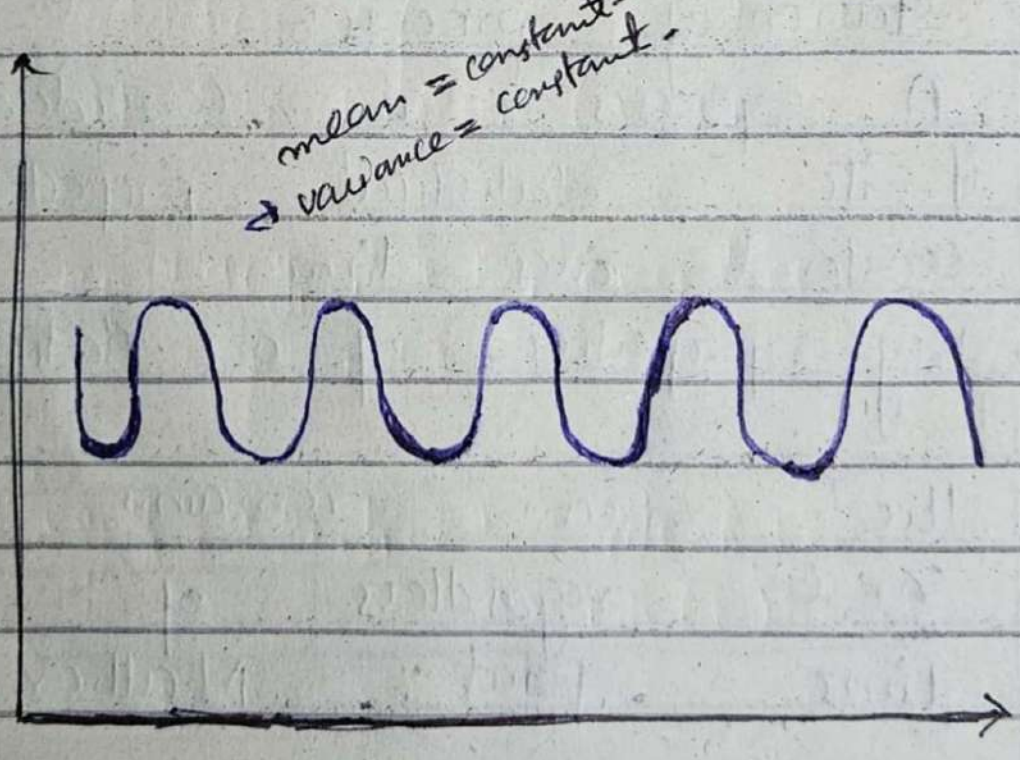
$$E(Y_t) = E(Y_{t+1}) = \mu$$

ii) The variance remains constant over time.

$$Var(Y_t) = Var(Y_{t+1}) = \sigma_Y^2$$

iii) The covariance b/w observations at different time points depends only on the time (j) between them, not on any specific reference period

$$Cov(Y_t + Y_{t+j}) = Cov(Y_s, Y_{s+j}) = \gamma_j$$



2). Non-stationary process :

- A process is considered non-stationary if its statistical properties changed over time. Common examples of non-stationary behaviors include trends, seasonality, and structural breaks.
- Non-stationary processes often exhibits changing means, variances or covariances over time.

- Autoregressive moving average model

Autoregressive model predicts a variable's future values based on its past values alone, without considering other variables.

- Autoregressive → selfregressive.

① $y_t = a_0 + a_1 y_{t-1} + \varepsilon_t$ (AR(1))
↳ autoregressive of order 1.

② $y_t = \varepsilon_t + \beta_1 \varepsilon_{t-1} \Rightarrow$ MA(1)

The addition of ① and ② makes it ARMA (1, 1)

$$y_t = a_0 + a_1 y_{t-1} + \beta_1 \varepsilon_{t-1} + \varepsilon_t$$
↳ ARMA(1, 1)

$$y = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + \cdots a_p y_{t-p} + \varepsilon_t$$
↳ AR(P)

$$y_t = \varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \cdots + \beta_p \varepsilon_{t-p} + \varepsilon_t$$

$\Rightarrow MA(p)$.

The addition of both makes it ARMA(P, q).

$$y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + \cdots + a_p y_{t-p} + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \cdots + \beta_q \varepsilon_{t-q} + \varepsilon_t$$

$\hookrightarrow$ ARMA(P, q)

$\varepsilon \Rightarrow$ unobserved components, we give weightage to it, e.g.: any positive or negative shock.

If any process is shock observes, it is called stationary process or short memory process, if a process is not an observer of shock, then it is called long memory process.

$$-1 < B_1 < 1, \; -1 < \alpha_1 < 1$$

we have to put conditions/restrictions on coefficients i.e α and β.

Box - Jinkins Methodology:

The Box-Jenkins methodology is a systematic approach for time series analysis and forecasting developed by George Box and Gwilym Jenkins.

It involves three stages:

(i) Identifying pattern in data.
(ii) Estimating model parameters.
(iii) Checking model adequacy.

I) Stage 1:

How to identify what numbers of lag should be incorporated? once it is identified, then?

II) Stage 2:

Is our model Correct?
Does, it pass statistical tests?

III forecasting 3:

"True model is never known."

Therefore, we call candidate models, then select one of the best.

↳ IDentification of true models: Through measure of linear correlation.

degree of linearity b/w two variables:

Partial correlation: (Fish waugh Theorem). Measuring the relationship b/w two variables holding other variables constant (eg Blood pressure and age example).

$y_t = \varepsilon_t + \beta_1 \varepsilon_{t-1}$ (MA(1))

measuring with y_t and applying expectations:

$$E(y_t y_t) = E\left[(\varepsilon_t + \beta_1 \varepsilon_{t-1})(\varepsilon_t + \beta_1 \varepsilon_{t-1})\right]$$

ε_t is white noise means.

↳ $E(\varepsilon_t) = \text{mean} = 0.$

$V(\varepsilon_t) = \text{variance} = \delta^2$

$\text{cov}(\varepsilon_t, \varepsilon_j) = 0.$

$$V(y_t) = E[y_t - E(y_t)]^2 \Rightarrow E\, y(t)$$

$$\underbrace{E(y_t)}_{=0}$$

$$V(y_t y_t) = E[(\varepsilon_t + B_1 \varepsilon_{t-1})(\varepsilon_t + B_1 \varepsilon_{t-1})]$$

$$V(y_t) = E(\varepsilon_t)^2 + 2B_1 E\underbrace{(\varepsilon_t \varepsilon_{t-1})}_{=0} + B_1^2 E(\varepsilon_{t-1}^2)$$

$= 0$ (because of no autocorrelation)

$$V(y_t) = \sigma^2 + \beta_1^2 \sigma^2 = \sigma^2 (1 + B_1^2)$$

Now multiplying y_t with y_{t-1}

$$\underbrace{E(y_t, y_{t-1})}_{\text{covariance}} = E[(\varepsilon_t + B_1 \varepsilon_{t-1})(\varepsilon_t + \beta_1 \varepsilon_{t-2})]$$

covariance

$$Cov(y_t, y_{t-1}) = E(\varepsilon_t \varepsilon_{t-1} + \beta_1 \varepsilon_{t-1}^2 + B_1 \varepsilon_t \varepsilon_{t-2} + B_2 \varepsilon_{t-1} \varepsilon_{t-2}]$$

$$\underbrace{Cov(y_t, y_{t-1})}_{\gamma_1} = \beta_1 E(\varepsilon_{t-1}^2) = \beta_1 \sigma^2$$

$\therefore \gamma_0 = Var(y_t) = \sigma_1^2 + (1 + \beta_1^2)$

Multiplying with two lags:

$E(y_t y_{t-2}) = E[(\varepsilon_t + \beta_1 \varepsilon_{t-1})(\varepsilon_{t-2} + \beta_1 \varepsilon_{t-3})]$

$\gamma_2 = 0$

$Var(y_t) = \delta^2 (1 + \beta_1^2) = \gamma_0$

$Cov(y_t, y_{t-1}) = \beta_1 \delta^2 = \gamma_1$

$Cov(y_t, y_{t-2}) = 0 = \gamma_2$

$Cov(y_t, y_{t-3}) = 0 = \gamma_3$

$$\boxed{\text{But } Corr(y_t, y_{t-1}) = \frac{Cov(y_t, y_{t-1})}{var(y_t)}}$$

$Corr(y_t, y_t) = \frac{\gamma_0}{\gamma_0} = 1$

$\delta_1 = Corr(y_t, y_{t-1}) = \frac{\gamma_1}{\gamma_0} = \frac{\beta_1 \cancel{\delta^2}}{(1+\beta_1^2)\cancel{\delta^2}} \Rightarrow \frac{\beta_1}{(1+\beta_1^2)}$

$\delta_2 = \frac{\gamma_2}{\gamma_0} = 0, \quad \delta_3 = 0$

$y_t = \varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2}$ } ↳ (homework)

Now, AR(1) process,

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \varepsilon_t$$

multiplying with y_t and, applying expectations

$$E(y_t y_t) = E[a_1 y_{t-1} + \varepsilon_t]^2$$

$$= a_1^2 E(y_{t-1}^2) + E(E_t^2) + 2\alpha_1 E(y_{t-1} \varepsilon_t)$$

$$Var(y_t) = \alpha_1^2 \sigma^2 = \sigma^2$$

now,

$$E(y_t y_{t-1}) = E[(a_1 y_{t-1} + \varepsilon_t)(a_1 y_{t-2} + \varepsilon_{t-1})]$$

$$= a_1^2 E(y_{t-1} y_{t-2}) + 0.$$

$$\gamma_1 = a_1 \gamma_0$$

$$\gamma_2 = a_1 \gamma_1 = a_1 \cdot a_1 \gamma_0 = a_1^2 \gamma_0$$

$$\gamma_3 = a_1^3 \gamma_0$$

$$\gamma_4 = a_1^4 \gamma_0$$

MA(1)

$\delta_0 = 1$

$\delta_1 = ---$

$\delta_2 = 0$

AR(1)

$\delta_0 = 1$

$\delta_1 = a_1$

$\delta_2 = a_1^2$

$\delta_3 = a_1^3$

$\delta_4 = a_1^4$

$\vdots$

$\delta_k = a_1^k$

Find $\delta_0, \delta_1, \delta_2, \delta_3, \delta_4$ --- for MA(2)

$$y_t = \beta_0 + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2}.$$

Autoregressive : A statistical data is autoregressive if it predicts future values based on past values.
ie: an autoregressive model might seek to predict a stock's price based on its past performance.

The observations at previous time series/steps are useful to predict the value at the next time step (correlation)
(Positive or negative correlation)
Because the correlation is calculated b/w the variables and itself at previous time steps, it is called autocorrelation, also called a serial correlation because of the sequenced structure of time series data.

$$\hat{y} = b_0 + b_1 x_1$$

$$x_{(t+1)} = b_0 + b_1 x_{t-1} + b_2 x_{t-2}$$

4. Partial autocorrelation function:
PACF gives the direct partial correlation of a stationary time series with its own lagged values, while controlling for the influence of intermediate lags.

: Fundamentals: The ACF and PACF plots are used to figure out the order of AR, MA, and ARMA models.

1)→ Autoregressive model (AR):
An AR model in time series analysis predicts future values of a variable based on its past values. It assumes that the current value of the variable depends linearly on its past values, with the addition of random noise. The order of AR model

model determines how many past values are used for prediction.

$$Y_t = \alpha_1 Y_{t-1} + \text{----} + \alpha_p Y_{t-p}$$

To figure out the model of (AR) we need to look at PACF

2) Moving average model (MA):

A moving average (MA) model in time series analysis predicts future values by considering the average of past error in predictions, rather than looking directly at past values of the variable

$$\hat{Y}_t = \epsilon_t + \beta_1 \epsilon_{t-1} + \beta_2 \epsilon_{t-2} + \text{---} + \beta_q \epsilon_{t-q}$$

To figure out the order of MA model, we need to look at ACF.

→ White Noise:

A time series may be white noise if.

(i) mean = μ = 0
(ii) variance = δ^2 (constant)
(iii) Covariance = 0.

A white noise can be stationary but a stationary process can't be a white noise.

	ACF	PACF
AR (P)	Geometric decay towards zero (may oscillate if $b<0$)	cutoff after lag P.
MA (Q)	Cutoff after lag q.	Geometric decay (oscillating if $\phi<0$)
ARMA	Decay (direct or oscillatory) after lag q	decay (direct or oscillatory) after lag P.

$y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + \varepsilon_t \rightarrow AR(2)$

$y_t = a_0 + a_1 y_{t-1} + \varepsilon_t \rightarrow AR(1)$

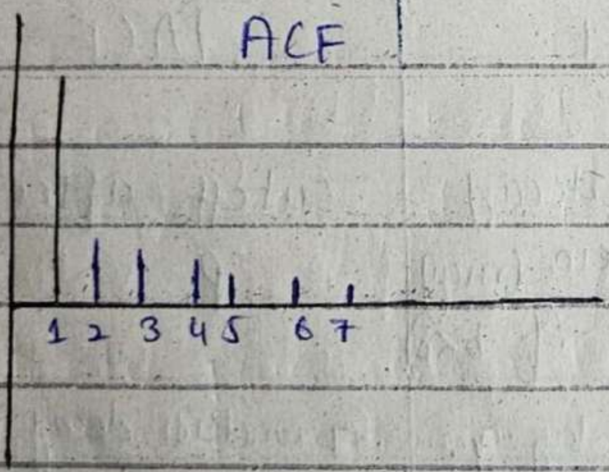
$y_t = \varepsilon_t + B_1 \varepsilon_{t-1} \rightarrow MA(1)$

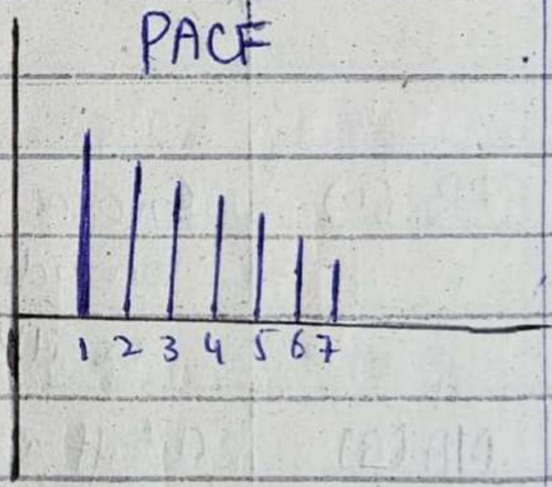
$y_t = \varepsilon_t + \beta_1 \varepsilon_{t-1} + B_2 \varepsilon_{t-2} \rightarrow MA(2)$

$y_t = a_0 + a_1 y_{t-1} + B_1 \varepsilon_{t-1} + \varepsilon_t$ (ARMA $^{(1,1)}$ model)

↳ mixture of both AR and MA

⇒ MA model:

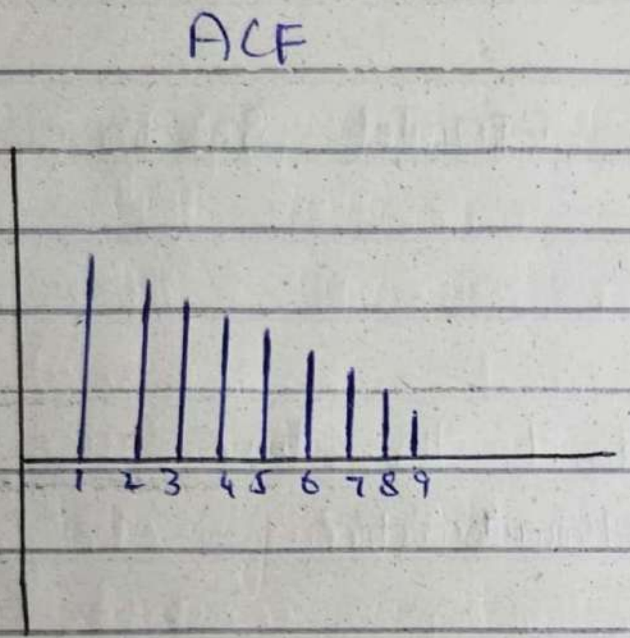




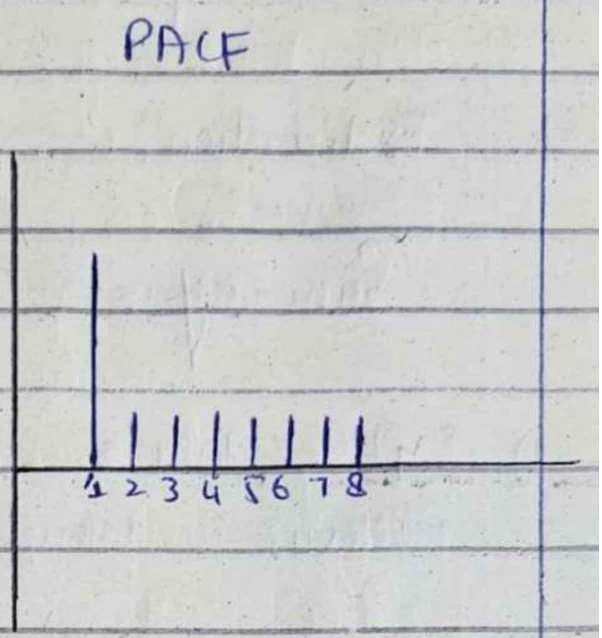
cutoff after lag 1

Geometric decay after lag 1

⇒ AR model :



Geometric decay after lag 1



Cutoff after lag 1.

⇒ Tips :

i) ACF's that do not go to zero could be a sign of non-stationarity.

ii) ACF of both AR, ARMA decay gradually, drops to zero for MA.

iii) PACF decays gradually for ARMA, MA, drops to 0 for AR.

Possible approach : begin with Parsimonious low order AR, Check residuals to decide on possible MA terms.

⇒ Estimation and model Selections:

Summary:

1) Explore data: Understand the time series patterns and characteristics

2) Autocorrelation analysis: Examine ACF and PACF plots for potential ARMA terms.

3) Propose models: Based on autocorrelation patterns, suggest ARMA model specification.

4) Estimate parameters: Use methods like OLS to estimate model parameters.

5) Compare models: Use criteria like AIC or BIC to compare model fit.

6) Diagnostic Checks : Verify model validity through residual analysis.

7) Refinement and Selection : Adjust models based on diagnostics and select the best - performing one

s) Compare models :
Fit vs parsimonous (information criteria)

Additional parameters (lags) automatically improve fitness of the model by reduce forecast quality.

"Fit vs parsimony" means balancing how well a model explains data (fit) with its simplicity (Parsimony). Fit aims for accuracy with more complexity, while parsimony prefers simplicity with fewer parameters. It's about finding the right balance between explaining data well and keeping the model manageable.

⇒ Non Stationary:

Nonstationary refers to the property of time series data where the statistical properties such as mean, variance, and autocorrelation change overtime. In other words, the data's characteristics change rather than remaining constant.

⇒ Why is it important?

It's important to recognize nonstationarity because many time series analysis techniques assume stationarity, meaning the statistical properties remain constant overtime. Nonstationarity violates these assumptions, leading to unreliable analysis and misleading conclusions.

⇒ How do we determine whether a time series is nonstationary?

Determining whether a time series is nonstationary often involves visual inspection and statistical tests.

⇒ White Noise :

While noise is defined by a zero mean, constant variance, and zero correlation.

White noise can be stationary but the stationary data can't be white noise.

White Noise error :

The error term represents that portion of the dependent variable that the independent variables in the model donot explain. When the error term is white noise, it indicates that the model's residuals are random and contain no autocorrelation, mean past values do not influence the future values.

⇒ Eliminating autocorrelation:

↳ Aim to eliminate autocorrelation in residuals, because it suggests that the model doesn't capture the lag values / Structure adequately. To address it your aim is to ensure that residuals exhibits no significant autocorrelation.

② Ploting Standardized residuals:
Standardized residuals ~~plotted as~~ denoted as (ε_{it}) are the residuals divided by their standard deviation. By plotting them, you can easily inspect whether they meet certain criteria. In this case, you're looking for no more

then 5% of the standardized residuals to fall outside the range [+2, -2] across all periods. This range corresponds to approximately 95% confidence intervals if the residuals are normally distributed.

3) Assessing Autocorrelation at different lags:

This involves examining the Autocorrelation function (ACF) and partial autocorrelation function (PACF) at different lag values. The Box-Pierce Statistic (or Ljung-Box test) is a formal statistical test used to assess the joint significance of autocorrelations up to a certain lags. s.

4) Box-Pierce Statistics:

This statistic denoted as Q, is calculated by summing the square of autocorrelations up to lag - s. Under the null hypothesis (HO) that all autocorrelations are zero, Q follows a chi-square distribution with s degree of freedom.

(*) A significant Q value suggests that at least one autocorrelation is non-zero, indicating that the residuals are not white noise

- By conducting these diagnostics, you can evalute whether time series model adequately Capture the underlying data patterns and whether the residuals exhibits the characteristics of white Noise. This helps ensure the reliability of the model forecasts and inferences.

⇒ Forecastibility :

Forecastibility is the ability of the model to predict future values accurately. TO access it:

i) Develop a model such as ARMA or ARIMA etc, using historical data.

ii) Use part of data to estimate parameters and the rest to evaluate predictions.

iii) Evalute forecast errors using metrics like (MSPE) mean squared predicted errors.

iv) Check the structural breaks to ensure the stability of data over time.

⇒ Account for possible structural breaks :

Does the same model apply equally

well to entire sample, or do parameters change (significantly) within the sample?
means is there any structural break?
Structural changes in data are shifts in the underlying data generating process overtime.
They can result from.

i) Economic conditions.
ii) Regime shifts.
iii) Policy changes.
iv) Technological advancements.
v) Natural disasters.
vi) Measurement Errors.

Detecting and addressing these changes is crucial for accurate modeling and forecasting. Techniques like the chow test or CUSUM test help to identify and manage structural breaks in the data.

(Non Stationary) (Slides)

Deterministic trend & (Non stationary Slides)

A deterministic trend refers to a Systematic and predicted pattern or directionality in a time Series that is driven by non random factors. In other words, it is a trend that follows a specific, non random function over time.

for examples

$y_t = \mu + \beta t + u_t$, where u_t is an independently and identically distributed (iid) random error term, and β is parameter representing the slope, or the Rate of Change over time.

This equation represents a deterministic trend model where the trend is a linear function of time. The parameter β controls the slope of the trend. $\beta = 0.45$, the trend increases by 0.45 units for each unit increase in time.

$$y_t = u + \overset{\beta}{0.45t} + u_t$$

This model is deterministic because the trend component, $u + \beta t$ is completely determined by the values of u and β, and there is no random variations in the trend itself. The randomness in the series comes solely from the random error term u_t.

→ Stochastic trend :
Random trend varies over time.
Stochastic trends are characterized by randomness and lack of predictability. In the

simplest case, we have a random walk process

$$y_t = y_{t-1} + \varepsilon_t$$

ε_t = (iid) Random shock at time t

A random walk is a special case of stochastic process, specifically a discrete-Time stochastic process. It's characterized by its tendency to exhibit movements that appears random and unpredictable. In a random walk, each step is determined by adding a random variable to the previous step.

*) Random walk with drift:
This adds a constant term (drift) to the basic Random walk equation. Mathematically it is represented as

$$y_t = \mu + y_{t-1} + \varepsilon_t$$

where μ represents a drift

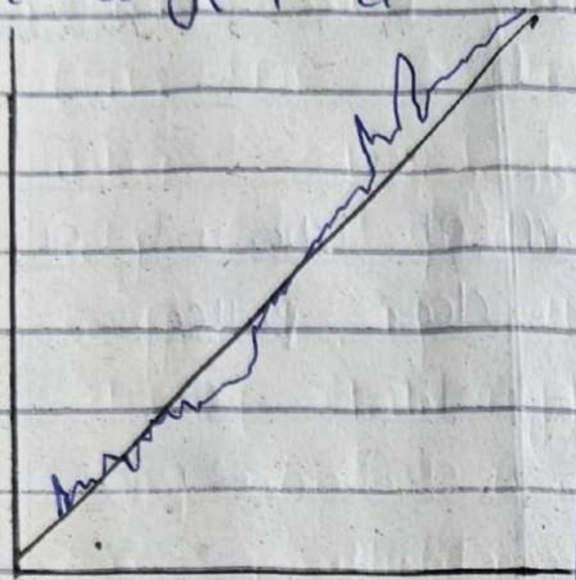
term, or the deterministic trend component, if $\mu > 0$, it indicates a positive drift, meaning that the series tends to increase over time, conversely, if $\mu < 0$, it indicates a negative drift, implying that the series tends to decrease over time. If $\mu = 0$, then there is no drift, and a series fluctuates around a constant level.

Random walk with drift	Deterministic Trend
→ Combines a random component ε_t with a deterministic trend.	Follows a specific, non random function over time.
↳ While it includes a trend component, it also involves random fluctuations, making	It remains constant over time.

it less predictable than a purely deterministic trend	
The series may show overall directionality due to ~~later~~ trend and short-term randomness, but individual movements can be unpredictable.	It shows a clear pattern.
$y_t = \mu + y_{t-1} + \varepsilon_t$	$y_t = \mu + \beta t + u_t$
μ = drift or trend component, where y_{t-1} is the previous observation, and ε_t is the Random Shock at time t.	μ, β are known parameters representing intercept and slope of the trend. here u_t = error term which is not accounted by the deterministic trend component $(\mu + \beta t)$ visually

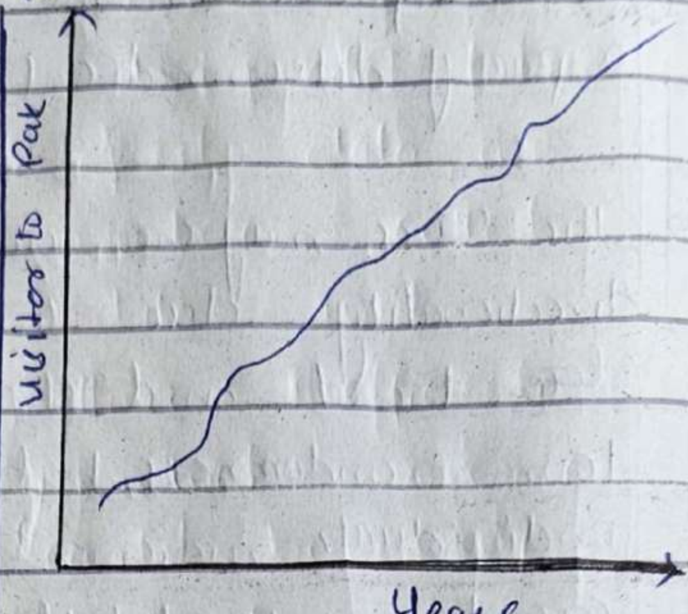
drift component is positive/above

$Y_t = \mu + Y_{t-1} + \epsilon_t$



So here in the Random walk with drift graph the trend is because the drift term μ and the Randomness is because of the Error term

$Y_t = \mu + \beta t + u_t$



Here the trend is determined by $\mu + \beta t$ and the small fluctuation in the trend is because of u_t error term, which is small and negligible relative to trend component.

⇒ Consequences of Non-stationarity :

1) "Shocks Do not die out" : In non stationary series, shocks to the series donot dissipate over time. Instead they have a persistent effects, leading to long lasting impact on the series. This make it challenging to interpret the effects of shocks and can complicate forecasting.

2) Statistical Consequences :

*) Non normal distribution of test statistics :
when we analyze data, we often assume it behaves in a certain predictable way. But if the data doesn't stay stable over time (non-stationary), our tests might not work as expected. This can lead us to misunderstand what the data is telling us.

*) Bias in AR Coefficients; Poor forecastibility.

When we use certain methods (like AR models) to predict future values based on past data, we rely on specific numbers called coefficients. If the data isn't stable (stationary), these coefficients might be wrong, which means our predictions could be way off. So we can't trust our forcasts as much.

↳ Explanations :

① Shocks doesn't die out :

In an AR(1) model like

$$y_t = by_{t-1} + \epsilon_t ,$$

the impact of shocks (disturbance) depends on the value of b.

⇒ Three cases :

if $b < 0$ ↳ each shock effect diminishes over time. As time passes, the influence of each shock gets smaller until it almost disappears. This is because b^t approches to 0 as time (t) goes to infinity.

If $b = 1$ ⇒ The effect of each shock persists indefinitely. Each shock has a lasting impact on the series, and the variance of the series keeps growing over time.

If $b > 1$ ⇒ Each shock's influence becomes stronger over time. As time goes on, the impact of each shock increases, leading to larger and larger fluctuations in the series.

(iii) Statistical consequences of non-stationary for multivariate regressions:

Non-stationary series in multivariate regressions can lead to:

i) High R^2 :- This might falsely suggest strong explanations for variation, even if the relationship isn't real

ii) Significant t-statistics :- These may appear, but they could be misleading due to coincidental correlations.

iii) True test : if residuals are stationary to ensure the reliability of the regression.

Unit Root testing : Methods and Problems :

⇒ Unit Root testing :

Unit Root testing is a statistical technique used to assess whether a time series variable has a unit root, indicating non stationarity.

If the unit root test results in rejecting the null hypothesis, it indicates that the series is stationary, and there is no unit root.
Conversely, if the null hypothesis cannot be rejected, it suggests that the series is non stationary and likely have a unit root.

There are several types of unit Root tests commonly used in time Series analysis -

Such as DF, ADF PP (Phi-llips - Perron) unit root test.

↳ How do you find out if a series is stationary or not?

Understanding whether a series is stationary involves considering its order of integration. A series which becomes stationary after being differenced once is said to be integrated of order 1 and is being denoted by I(1). In general a series which is stationary after being differenced d times is said to be integrated of order d, denoted by I(d). A series which is stationary without differencing, is said to be I(0).

↳ Examples

Consider a model

$$Y_t = b_0 + Y_{t-} + \varepsilon_t$$

↳ This model is $I(1)$ because differencing once $(\Delta Y_t = Y_t - Y_{t-1})$ ~~mean~~ makes it stationary.

↳ $\Delta Y = b_0 + \varepsilon_t$, which is $I(0)$ (i.e.) stationary without any differencing.

↳ If differencing a time series Y_t three times results in stationary, then it is integrated of order 3, denoted by $I(3)$. Then equation of such series would be.

$$\Delta^3 Y_t = Y_t - Y_{t-1} - (Y_{t-1} - Y_{t-2}) - (Y_{t-2} - Y_{t-3}) = \varepsilon_t$$

Breakdown of $I(3)$ difference equations

i) First difference : $(Y_t - Y_{t-1})$ removes the autoregressive term Y_{t-1}.

$((Y_t - Y_{t-1}) -$

ii) Second difference ↳ $(Y_{t-1} - Y_{t-2})$ removes.

the autoregressive term Y_{t-2}

(iii) Third difference : $((Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}) - (Y_{t-2} - Y_{t-3}))$ removes the autoregressive term Y_{t-3}.

After 3 differencing, the autoregressive terms are completely eliminated, and the resulting equation only includes the random error term ε_t, leading to a stationary series. There the coefficient b does not appear in the I(3) differenced equation.

Identifying non-stationary processes

Eye ball the data:
(informal procedure for identifying non-stationary process)

Informal procedure for identifying non-stationary processes involve visually examining the data for trends, changes in mean, and other patterns that suggests a lack of statistical equilibrium. If visual inspection indicates potential non-stationarity, further analysis using formal tests can provide additional confirmation.

Diagnostic test - Correlogram &
Indeed, for a stationary process, the ACF or correlogram decays rapidly as the lag (k) increases. This means that there is little to no correlation b/w observations at different time plots, indicating that the series has no memory beyond short term fluctuations-

eg: if the correlogram for a stationary process dies out rapidly, it implies that observations from 1980 are not significantly correlated with observations from 1985 or any other time point. Each observation in the series behaves independently of others, supporting the notion of stationarity.

Therefore, if you observe a rapidly decaying correlogram, it provides evidence in favor of the series being stationary, as there is no persistent relationship b/w observations over time.

1) Statistical tests for stationarity:

Simple t-test:

lets suppose we have setup

AR(1) process with drift (b_0)

$$Y_t = b_0 + b Y_{t-1} + \varepsilon_t \rightarrow (i)$$

$\varepsilon_t \Rightarrow$ iid

Simple approach is to estimate eq (i) using OLS and examine b.

if $b > 1$ or $= 1$

then it is non-stationary

but if $b < 1$ (stationary)

because the influence of the lag die's out when time approaches to ∞.

↳ The simple t-test checks if the coefficient ~~be~~ "b_0" estimated from an AR(1) model with drift is significantly different from 1. If b_0 is less than 1, it indicates stationarity; otherwise it suggests non-stationarity. The test rejects the ($H_0 = \cancel{b_0} = 1$) when the t-statistics (TS) is sufficiently negative (below -1.65 at the 5% significance level)

problem with simple test approach:

The simple test we use to check if an AR(1) process with drift is stationary has a problem in small samples. It's called Dynamic bias. This happens when we include lagged dependent variables in the test. When the coefficient $b = 1$, it means the process isn't stationary, and our usual analysis methods don't work correctly. Dynamic bias means that our estimates of b tends to be lower than it should be, especially in small samples. This can make us think the process is stationary when it is not.

2) Dickey - Fuller (DF) Test :

The Dickey Fuller test checks if a Time series has a unit root, which indicates non Stationarity.

→ The test is based on an (AR) model :

$$\Delta Y_t = \alpha + \beta Y_{t-1} + \varepsilon_t$$

where

① → $\Delta Y_t = Y_t - Y_{t-1}$

$\beta = bY_{t-1} - Y_{t-1} \Rightarrow (b-1) Y_{t-1}$

② → $\beta \Rightarrow (b-1)$ (coefficient of lagged value)

③ → α = intercept

④ → ε_t = error term.

①→ The Null-hypothesis (H_0), series has a unit root, meaning it's not Stationary : $H_0 = \beta \stackrel{?}{=} 1$

• If $\beta = 1$ ⇒ unit root ⇒ nonstationary

2) Calculate the ~~t-test~~ test statistics using a t. test.

$$TS = \frac{B_1 - B_0}{SE(B_1)}$$

3) Reject the null hypothesis (H_0) in favor of the alternative hypothesis (H_1) if the t. statistics is < 1

large negative values of the test statistics indicate rejection of the H_0, suggesting that the series is stationary.

In summary, the Dickey- Fuller test assesses stationarity by examining whether the coefficient of the lagged value in an autoregressive model is significantly different from 1. If it is significantly less than 1, it indicates stationarity, while a value close to or equal to 1 suggests non-stationarity.

↳ Variants of Dickey - Fuller test :
The different regression can be used to test the presence of a unit root.

i) $\Delta Y = \beta Y_{t-1} + \varepsilon_t$
For testing if Y is pure Random Walk

ii) $\Delta Y = b_0 + \beta Y_{t-1} + \varepsilon_t$
For testing if Y is RW with drift.

iii) $\Delta Y = b_0 + \beta Y_{t-1} + \underbrace{b_2 t}_{\text{deterministic}} + \varepsilon_t$

For testing if Y is RW with drift and deterministic Trend.

⇒ Incorporating Time trends in DF test for unit roots
Sometimes series clearly displays

an upword or downword trend (non stationary mean). Should therefor incorporate trend in the regression used for the DF Test.

$$\Delta Y = b_0 + \beta Y_{t-1} + b_2 \, trend + \varepsilon_t$$

It may be the case that Y_t will be stationary around a trend, although if a trend is not included series is non-stationary.

3) Augmented Dickey Fuller (ADF) test for unit root :

The augmented Dickey fuller test is a statistical test used to determine whether a time series is stationary or nonstationary. It is the extension of Dickey-Fuller test, designed to handle autocorrelation in the residuals of the model.

↳ ADF test models the differenced time series (ΔY_t) as a function of lagged values of the time series and its differences

$$\Delta Y_t = b_0 + BY_{t-1} + a_1 \Delta Y_{t-1} + a_2 \Delta Y_{t-2} + a_3 \Delta Y_{t-3} + a_4 \Delta Y_{t-4} + \varepsilon$$

The lagged difference (ΔY_{t-1}) captures the autocorrelation in the residuals.

↳ A General to specific approach is used to select the most appropriate model by eliminating insignificant variables.

*) Initially, all the lagged values are included in the model.

*) Insignificance of variables is tested, and non significant terms are removed iteratively to arrive at a parsimonious model.

↳ Once the final model is

selected, it's checked for auto-correlation in the residuals. Autocorrelation indicates that the model may not adequately capture all the relevent impfostmation in the data.

↳ The significance of the variables in the final model is assessed using an F test. Significant variables contribute to the explanatory power of the model, while insignificant ones may be removed to simplify the model.

By incorporating lagged differences and iteratively refining the model, the ADF test addresses the issue of auto-correlation in the residuals, providing a more reliable assessment of the presence of the unit root in the time series data.

⇒ Choose between alternative models - The model - progress result &

When choosing the best model for the ADF test, we com use criteria lik AIC, BIC. These criteria help us balance how well the model fits the data with its complexity. Lower values are better.

So, we compare different model based on these values and choose the one with the lowest AIC, SBC, HQC. This ensures we pick a model that fits the data well without being too complicated.

So, AIC, SBC, HQC, or some

other criteria may be used beside having judgement for appropriate lag selection.

4) Phillips Perron test:

The Phillips Perron test checks if a time series is stable over time or if it's changing a lot.

The PP test statistics measures the significance of the coefficient of the lagged series in the regression. If this coefficient is statistically significant, it suggest the presence of a unit root and non-stationarity.

5) ADF vs PP tests:

	Feature	ADF	PP. Test
1	underlying assumption	No autocorrelation in error term	No assumption regarding AC.

(2)	Applicability	Data with no AC (significant one)	Data with potential AC.
(3)	Implementation	Require lag selection	No lag selection required
(4)	Advantages :	Common - well established	more robust to AC in data.

In practice, it's often recommended to run both tests and compare the results. If they agree on the presence or absence of a unit root, you can be more confident in the conclusion.

General to Specific or Specific to General Technique :

General to specific :
Start the model including all potential variables, then eliminate non-significant ones step by

step until you have the most relevent ones left.

↳ Specific to General :
Begin with a simpler model and add variables one by one, testing their significance until no additional variables significantly improve the model.

Both approaches aim to reach a parsimonious model, but they differ in how they progress from initial model to the final one.

A Parsimonious model is one that includes only the essential variables or components necessary to explain the data, avoiding unnecessary complexity.

↳ Multiple roots:-

$$\Delta^2 y_t = b_0 + B_1 \Delta y_{t-1} + \epsilon_t$$

If the coefficient B_1 in the first difference of time series is significant, it suggests that the series might have a 2nd ~~year~~ order autoregressive process. In This case, the second-order autoregressive model is estimated, and if B_2 coefficient is also significant, it implies that original series is stationary.

if $B_1 \Rightarrow$ significant then model becomes

$$\Delta^2 y_t = b_0 + B_1 \Delta y_{t-1} + B_2 \Delta y_{t-2} + \epsilon_t$$

↳ then test B_2 significance

if it is then conclude

$y_t \Rightarrow$ stationary.

→ Panel unit Root tests

A panel unit root test is a statistical method used to assess whether a set of time series data, observed ~~across~~ across different individuals or entities over time, exhibits unit root behavior over time. Unit root behavior suggests that the data series is non-stationary, meaning it has a stochastic trend and does not trend to return to a stable mean time.

OR:

A panel unit root test checks if multiple time series from different entities share a common non-stationary pattern, like a random trend, it is useful for analyzing data with both cross-sectional (different entities) and time-series dimensions.

For this panel unit root test we use Im-Pesaran-Shin test.

→ Im-Pesaran Unit root test:
It is a panel unit root test. It's designed to determine whether a panel dataset exhibits unit root properties, which implies non-stationarity. Unit root test are important in time series analysis as they help assess the stationarity of the variable over time. Non-stationarity can

effect. these reliability analyses and forecasting models :-

Limitations of IPSE tests:

① Null-Hypothesis ⇒ It tests if all coefficients are zero, indicating disagreement if any coefficient differs.

② There's uncertainity about the test's asymptotic theory, particularly with varied sample size.

③ Critical values :- They depend on coefficient magnitudes, affecting test interpretations.

④ Error terms :- Serial and contemporaneous uncorrelated error terms are assumed, requiring autocorrelation checks.

(5) Correlation correction & Adjustments may not fully eliminate correlation across equations, especially if variables are non-stationary.

These highlights challenges in using the IPSE effectively.

(5) The Kwiatkowski - ~~Ph~~ Phillips - Schmidt - Shin) Test (~~KP~~ (KPSS)

The KPSS test checks if a time series has a trend or not. It looks whether the variance of the series is stable over time. If the variance is constant, it suggests the series is trend stationary. Critical values are used to decide if the test rejects or accepts the idea of trend stationary.

$y_t = \beta_0 D_t + \mu_t + u_t$.

D_t = deterministic components.

μ_t = random walk

u_t = error term.

↳ Null hypothesis :

variance of error term is ~~unit root~~ ⇒ zero, indicating that the series is trend stationary.

H_0 = variance = ~~unit~~ root.

↳ H_1 = not a ~~unit root~~ non Stationary

↳ Critical value :

These value are used to determine whether to reject or accept the null hypothesis.

① if T. statistics > critical value H_0 rejected.

② if T-statistics < Critical value

H_0 = accepted. indicating that the series is trend stationary.

Important:

DF, ADF and PP are called unit root test; the null hypothesis is that y_t has a unit root; is $I(1)$ or higher.

KPSS is a stationary test; null hypothesis is
$$y_t = I(0).$$

ARIMA Slides

AR =?
MA =?
ARMA =?

↳ Autoregressive model (AR) :
The AR model represents a regression of the variable against its own lagged values. It assumes that that the current value of a variable depends linearly on its past values i,e its influenced by its own values.

AR models are used when the current value of the series can be predicted based on its past values. They are useful in modeling phenomena like stock prices, weather patterns, or any other time dependent data where the current value depends on past values.

↳ Moving averages:
The MA models represents the relationship between an observation and a residual error term based on a moving average of the past errors. It assumes that that the current value of the variable depends linearly on the errors made in the past.

↳ Autoregressive moving average models:
The ARMA model combines both AR and MA models. It captures both the linear relationship with the past values (AR) and the dependency on

the error terms (MA)
ARMA models are beneficial when the time series data exhibit both auto-correlation and a moving average structure, they are widely used in economics for forecasting.

↳ Autoregressive Integrated moving Average (ARIMA) model :

When the ARMA model is non stationary, ~~they~~ then the model is extended to ARIMA model by adding differencing to make the series stationary

ARIMA models are suitable for non stationary time series data where trends and seasonality exists

So the AR, MA, ARMA or ARIMA models are used in time series analysis to understand and forecast sequential data points. The accuracy of the forecast can be improved by selecting appropriate model parameters, such as the lag orders or the degree of differencing, and by incorporating additional variables or variables with lags.

⇒ Important :

How can ~~we~~ we tell that AR, MA, ARMA or ARIMA is the one that best explains the time series data :

Selecting the appropriate model AR, MA, ARMA or ARIMA to explain your time series data involves assessing the data's characteristics and comparing the goodness of fit of the different models

Here's the general approach that determine that which models best explains the data.

↳ Visual Inspection of
Plot the Autocorreletion (ACF) and partial autocorrelation function (PACF) of the time series data

Look for patterns in the ACF or PACF plots that suggests the presence of AR or MA behavior i.e

(i) If the ACF decreases exponentially and cuts-off after a certain lag, it suggests an AR process.

ii) If the PACF cuts off after a certain lag, it suggests an MA process.

iii) If both ACF and PACF shows significant correlation at multiple lags, it suggests an ARMA process.

↳ Model ~~End~~ identification :
Based on visual inspection, identify candidate model (AR, MA, ARMA) that seem appropriate for explaining the data

Use (AIC) or (BIS) : to compare the canditate models quantitatively. Lower values of AIC or BIC indicate a better trade-off b/w model complexity and goodness of fit.

↳ Model Estimation:
Estimate the Parameters of each candidate model using methods like least square estimation.
Assess the goodness of fit of each model by examining diagnostic plots, such as residuals plots and normally tests.

↳ Model Selection:
Choose the model with the lowest AIC or BIC values as the preferred model

↳ Model validation:

Validate the selected model by comparing its forecasts with actual data on a validation dataset.

↳ Iterative Refinement:

If the selected model doesn't perform well, consider refining the model by adding or removing components (eg. adding seasonal terms or adjusting parameter values.

↳ Repeat the model identification, estimation, selection and validation steps until a satisfactory model is found.

By following these steps, you can systematically compare AR, MA, ARMA models to determine which one best explains the data and provide accurate forecasts.

Box-Jenkins Methodology:

The Box-Jenkins Methodology is a Approach for the time series analysis and forecasting. It primarily focuses on ARIMA model, although it can incorporate elements of AR, MA models as well. It typically involves the following steps.

1) Identification:

Identify patterns or structurals in the time series data

*) Visual inspection of the time series plot to identify trends, seasonality or irregularity.

*) Analysis of ACF and

PACF functions to identify (AR) and (MA) pattern.

*) Application of differencing to achieve stationarity if the data exhabits trends or seasonality.

*) The objective is to determine the appropriate order (P, d, q) for the ARIMA model, representing the AR, I and MA components, respectively.

2) Estimation:
Estimate the Parameters of ARIMA model based on the identified patterns. use the Methods such as least square estimation.

3) Diagnostic checking:
Evaluate the adequacy

and goodness of fit of the estimated ARIMA model.

*) Conduct diagnostic checks on the residuals to access model adequacy.
*) Check the autocorrelation, heteroscedasticity and normality in the residual using statistical tests and graphical methods.

4) Forecasting:
Once the ARIMA model is validated, it can be used to generate forecasts for future time period.

5) Model Refinement and updating:
Continuously monitor and update the ARIMA model to maintain forecast accuracy overtime.

By following these steps, analysts can systematically apply the Box-Jenkins methodology to analyze the time series data, develop ARIMA models, and generate accurate forecasts for decision making and planning purposes.

{ Spurious Regression }

Spurious regression occurs when two or more time series variables appears to be correlated in a regression analysis, but there is no genuine causal relationship b/w them. This misleading correlation arises due to the presence of the common trend or other factors that effect both the variables simultaneously, leading to erroneous conclusions about their relationship.

The spurious regressions arises from the presence of common cause between two variables. A common cause occurs when both variables are influenced by the 3rd factor leading to a correlation b/w them. However, despite this correlation, neither variable cause directly the other. When a regression analysis is conducted on such variables, it may incorrectly suggests a strong relationship b/w them, attribbuting causality where none exist. This misleading regression result is termed a spurious regression or a nonsence regression

↳ A Causal Relationship : A consumption function :

Analysis Conducted on the data regarding annual GDP and consumption in Singapore. The regression aims to explore the relationship b/w GDP and consumption, following the Keynesian consumption function hypothesis the income (GDP) determine consumption.

The results shows a high R^2 of 99.6% suggesting that most of the variation in Singaporean consumption can be explained by Singaporean GDP. the coefficient of GDP is significant, indicating a strong influence on consumption.

However, the issue arises with the regression. Graphs of residual reveals pattern, suggesting the model may be wrong. Econometricians often try fixing this by adding more variables or changing the model, but this can make things worse. Also estimated effect of GDP on consumption is unstable - it changes alot with different variables in the model. So, we should not rely too much on these results, suggesting they might not be accurate for making policies.

4 A correllation without Causation or The spurious regression occurs when two variables are correlated but are not causing each other directly. We can illustrate this Phenomena by running a regression of (Sgpcon) on (SAF GDP), which logically should not have a significant relationship. However

The regression yields a high coefficient of determination R^2 of 97%, with a highly significant coefficient for (SAF GDP). This result is nonsensical when interpreted causally, as it suggests that an increase in (SAF GDP) would lead to a substantial increase in Singaporean consumption

The distinction b/w nominal and real econometrics emphasizing the importance of understanding the underlying causal relationship b/w variables. Nominal econometrics focuses on correlation without necessarily considering causation, while real econometrics seeks to uncover the causal mechanisms behind observed relationships.

There are limitations of traditional approaches to addressing spurious regressions, such as adding missing variables or transforming variables to achieve stationarity. While these ~~results~~ methods may improve the regression results to some extent, they donot address the fundamental issue of distinguishing b/w genuine causal relationship and spurious correlations.

The solution of the problem of the spurious regression lies in the deeper study of causal structures connecting variables. This involves steping outside the confines of nominal econometrics and adopting a real econometrics approach, which considers the underlying meaning of variables and their relationships.

Investigation whether a transformation in the regression analysis lead to distinguish b/w genuine causal relationships and spurious ones.

Conduct a regression of the growth rates of (Sgp con) on the growth rates of (SAF GDP)

The regression results shows that the growth rates of Sgp con are explained by a highly significant growth rates of the

(SAF GDP). The residuals from this regression appear random, supporting its validity.

However, the problem arises when interpreting such regressions that there is no economic meaning of such regressions.

To address this issue, they proposes a solution involving the inclusion of additional variables in the regression models. By adding the growth rates of both SgpGDP and SAF GDP, the regression yields more accurate results, Singapore GDP emerges as a highly significant determinant, while the SAF GDP is not significant at the 95% level.