Time Series Analysis

Module 1

A time series represents a sequence of values observed over time, such as stock prices or temperature measurements. Modeling and predicting are distinct: **predicting** refers to guessing future events, while **modeling** involves constructing mathematical representations of a series to gain insights and make predictions.

In forecasting, common mistakes, such as predicting only the next value rather than multiple future values, are often found in misleading models. Many models, like those using LSTMs in stock price prediction, fail due to improper validation techniques and overfitting.

The **Random Walk Hypothesis** suggests that stock prices follow a random path, making it impossible to predict future values precisely. However, **volatility clustering** indicates that when prices are highly volatile, they tend to remain volatile in the near future.

Instead of predicting stock prices, it is preferable to model **returns**, particularly **log returns**, as they are more stable and stationary. This allows for consistent predictions because returns tend to come from the same distribution over time.

Finally, the **Efficient Market Hypothesis (EMH)** posits that all available information is already reflected in stock prices, making it difficult to consistently outperform the market. EMH exists in three forms: **weak**, **semi-strong**, and **strong**, each describing different levels of market efficiency.

In conclusion, effective time series analysis for financial data focuses on modeling returns, addressing stationarity, and acknowledging market efficiency.

Module 2

Exponential smoothing is a time series technique based on moving averages (MA). It reduces fluctuations in the data to reveal trends. There are two primary types: **Simple Moving Average (SMA)** and **Exponential Moving Average (EMA)**. While SMA assigns equal importance to all past data points, EMA gives greater weight to more recent data, making it more responsive to short-term changes.

A simple trading strategy using moving averages involves calculating two averages with different time windows (e.g., 50 and 100 steps). When the short-term average crosses the long-term average from below, it signals a buy opportunity, and crossing from above signals a sell.

Holt’s Linear Trend Model extends simple exponential smoothing by incorporating a trend. The model includes three components: the **level equation** (capturing the average smoothed value), the **trend equation** (capturing the slope), and the **forecast equation** (which predicts future values based on the level and trend).

Holt-Winter’s Model further adds a **seasonality component** to account for repeating patterns in the data. The model includes equations for level, trend, and seasonality, providing more accurate forecasts for time series with seasonal variations.

These models offer significant improvements in forecasting trends in time series, especially for stock prices and other financial data. While exponential smoothing focuses on smoothing data points, the Holt and Holt-Winter models address the need for both trend detection and seasonality adjustments. The **Holt-Winter’s Model** is especially effective when a time series exhibits regular seasonal fluctuations.

For implement these models, detailed code for both Holt and Holt-Winter models was referred online, alongside helpful tutorials and videos to understand practical applications

Module 3

Auto-Regressive (AR) models predict the current value of a time series using a linear combination of past values. The parameter ‘p’ represents the number of lagged values included. The model can be expressed as:

$y_t = b + a_1 y_{t-1} + a_2 y_{t-2} + \ldots + a_p y_{t-p} + \epsilon_t$.

Moving Average (MA) models, in contrast, are based on past error terms. The parameter ‘q’ represents how many past errors contribute to the current value:

$y_t = c + \epsilon_t + b_1 \epsilon_{t-1} + \ldots + b_q \epsilon_{t-q}$.

ARMA (p, q) combines both AR and MA components, but requires a stationary time series.

ARIMA (p, d, q) adds the concept of differencing (d) to handle non-stationary time series. ARIMA applies differencing to the data before using ARMA modeling, effectively making the series stationary. The model's parameters are learned through optimization, typically by minimizing RMSE.

Stationarity is crucial for these models. The **Augmented Dickey-Fuller (ADF)** test helps determine if a series is stationary. A low p-value (< 0.05) indicates stationarity, guiding how much differencing (d) is needed.

Auto-correlation Function (ACF) helps select the MA parameter 'q' by analyzing correlations at different lags. The number of points crossing a significance margin in an ACF plot suggests the value of 'q'.

Partial Auto-correlation Function (PACF) aids in selecting the AR parameter 'p' by isolating the direct correlation between the current point and past lags.

ARIMA models are commonly used in time series forecasting, but for many stock price data, ARIMA(0,1,0)—a random walk—often emerges as the best model, highlighting the challenge of predicting stock prices. Variants like **seasonal ARIMA** and **Auto ARIMA** extend this approach for more complex datasets.

Module 4

In this module, we explore machine learning methods to predict values in a time series, focusing on models like Linear Regression, Multi-Layered Perceptrons (MLPs), Convolutional Neural Networks (CNNs), and Recurrent Neural Networks (RNNs), including Long Short-Term Memory (LSTM) networks.

Linear regression, a fundamental statistical method, forms the basis for predicting a value by learning from past data points. However, more advanced machine learning models, such as MLPs, CNNs, and RNNs, offer significant improvements by learning complex patterns in time series data.

- **Multi-Layered Perceptrons (MLPs)** consist of multiple layers of neurons and can capture non-linear relationships in time series data.
- **Convolutional Neural Networks (CNNs)**, typically used in image processing, are adapted for time series by recognizing patterns and trends across temporal data, using convolutional filters to capture local dependencies.
- **Recurrent Neural Networks (RNNs)**, designed to handle sequential data, leverage information from previous time steps. RNNs retain memory of past inputs, which is crucial for time series prediction.

- **Long Short-Term Memory (LSTM)**, a variant of RNN, overcomes the limitation of vanishing gradients in traditional RNNs by maintaining long-term dependencies. This is particularly useful for time series data that require memory of distant past events.

To understand how these models can be applied to time series forecasting, several useful resources are available. The YouTube videos provided were used to learn the theory and application of CNNs, RNNs, and LSTMs in predicting time series data, which is crucial for gaining hands-on experience with these models.

These machine learning techniques are powerful tools for modeling time series data, especially when traditional methods like ARIMA are insufficient for capturing complex patterns.

Module 5

In this module, we explore **GARCH** (Generalized Auto-Regressive Conditional Heteroskedasticity), a model used to predict volatility in time series data, especially in financial markets. **Heteroskedasticity** refers to changing variances over time, where the variance of residuals is not constant. GARCH models volatility clustering, the phenomenon where high-volatility periods are followed by more high-volatility periods, and low-volatility periods by more low-volatility periods.

1) ARCH (Auto-Regressive Conditional Heteroskedasticity):

ARCH models focus on capturing variance from past error terms. In the ARCH(1) model, the time-dependent variance σ_t^2 is a function of past squared residuals (ϵ_{t-1}^2). By extending to ARCH(p), multiple past error terms are considered. This helps model volatility in financial data, where squared residuals often show significant lags.

2) GARCH:

GARCH extends ARCH by including both past error terms and past variances, allowing for greater persistence in volatility. The **GARCH(1,1)** model, for example, is given by:

$\sigma_t^2 = \omega + a_1\epsilon_{t-1}^2 + b_1\sigma_{t-1}^2$, where past errors and variances both affect the current variance, capturing volatility clustering.

Applications:

GARCH is commonly used in financial modeling, particularly in pricing financial instruments like options. The **Black-Scholes option pricing model** relies on accurate estimates of a stock's volatility, which can be modeled using GARCH. Additionally, **ARIMA-GARCH** models are popular, where ARIMA models the mean and GARCH handles the volatility of residuals.

For further learning, several YouTube videos and Engle's original paper on GARCH were used.

GARCH is crucial for understanding financial market behavior and modeling time series data with variable volatility.

Module 6

In this module, we expand on time series prediction models by exploring **VARMA**, **SARIMA**, and the **Heston Model**.

1) VARMA (Vector AutoRegressive Moving Average):

The **VARMA** model is an extension of the ARIMA model. While ARIMA predicts future values based on past data from a single time series, VARMA uses multiple time series to predict future values. By incorporating the historical values of related variables, VARMA can provide better predictions, particularly in multivariate time series, like predicting stock returns based on multiple financial indicators.

2) SARIMA (Seasonal ARIMA):

SARIMA introduces a seasonality component to the ARIMA model. This model accounts for regular, repeating patterns (seasonality) in the data in addition to trend and noise. For example, it can model quarterly or annual trends in economic or sales data. SARIMA is particularly useful for data where the effect of seasonality cannot be ignored. It allows ARIMA models to capture not just past values but also recurring seasonal influences in a time series.

3) Facebook Prophet:

Facebook Prophet is an advanced model designed for time series forecasting with an emphasis on capturing trends, seasonality, and holiday effects. It is user-friendly, robust to missing data, and capable of handling outliers, making it popular for forecasting tasks. It automatically detects the seasonality in data and is widely used in business applications like sales forecasting.

4) Heston Model:

The **Heston Model** is a **stochastic volatility model** used in finance to predict the volatility of assets. Unlike traditional models where volatility is constant, the Heston model assumes volatility varies over time, better capturing real-world asset price movements. This model is essential for pricing financial derivatives like options.

Each of these models helps handle complex time series data, enhancing prediction capabilities in finance and business

Financial Derivatives

Financial derivatives are financial instruments whose value is derived from an underlying asset, such as a stock price. Common types of financial derivatives include **Futures**, **Options**, **Forwards**, **Swaps**, and **Swaptions**. Just as cheese is derived from milk or ketchup from tomatoes, financial derivatives derive their value from another financial entity.

Futures and **Options** are the most widely known derivatives. In Futures contracts, two parties agree to buy or sell an asset at a future date for a specified price. In contrast, Options provide the buyer with the right, but not the obligation, to buy or sell an asset at a predetermined price within a specific time frame.

Derivatives are crucial in finance for managing risk, speculating on price movements, and improving market efficiency. They are widely used in trading and investment strategies.

In quantitative finance, arbitrage refers to the practice of profiting from price differences in two or more markets. For example, if a product costs $5 in one market and $6 in another, one could buy it in the cheaper market and sell it in the more expensive one, profiting from the price difference. Arbitrage portfolios are constructed in such a way that their current value is less than or equal to zero, while their future value is non-negative with a possibility of being strictly positive. The **Monotonicity Theorem** states that if one portfolio's future value is always greater than or equal to another's, then the present value follows the same relationship. This leads to the replication theorem, which is used extensively in derivative pricing.

Zero Coupon Bonds (ZCB) are basic instruments that have a value of 1 at maturity and are discounted back using either continuous or annually compounded interest rates.

A simple approach to price options involves computing expected payouts and discounting them using a risk-free rate. For example, a call option with three possible future outcomes can be priced by calculating the expected payout and discounting it. However, using such naive models can lead to arbitrage opportunities, which are unsustainable in a realistic financial environment.

In a risk-neutral world, option pricing can be handled by constructing a portfolio that replicates the option's payoff. By adjusting the amount of stock held in the portfolio, the value of the portfolio at different future outcomes can be made equal, and by the replication theorem, its present value is equal to the option's present value. This is known as **risk-neutral pricing**, which discounts the expected future payoff using the risk-free rate.

Extending this approach to multiple time-steps, the **binomial model** divides time into discrete intervals and assumes that at each step, the stock price either rises by a certain percentage or falls by another. Replication portfolios can be built at each node, and risk-neutral probabilities are used to calculate the option price at each time step. The general rule for an arbitrage-free binomial model is that the down movement must be less than the risk-free rate, which in turn must be less than the up movement.

In conclusion, these fundamental concepts, including arbitrage, replication, and risk-neutral pricing, are essential for valuing derivatives like options and ensuring that no arbitrage opportunities exist in financial markets.

The binomial pricing model provides a way to replicate the payout of a derivative contract using stocks and zero-coupon bonds (ZCBs). For a contract with a payout of γ in the down state and $\gamma + \beta$ in the up state, the replicating portfolio consists of λ stocks and μ ZCBs. The replication price can be derived and expressed as a function of these variables, leading to the generalized formula for pricing derivatives using risk-neutral probabilities.

This concept can be extended to multiple timesteps, where the price of a contract with a payout dependent on the stock price is the expected value under the risk-neutral measure. As the number of timesteps increases and approaches zero, this leads to the continuous-time Black-Scholes option pricing formula.

Martingales are central to pricing in finance. A process is a martingale if its expected future value, given its current value, equals its present value. This property holds in the context of financial markets under the risk-neutral probability measure. Stock prices, when discounted by the growth factor (money market account), form a martingale.

The Fundamental Theorem of Asset Pricing states that there are no arbitrage opportunities if and only if there exists a risk-neutral probability distribution. Under this measure, asset prices (relative to the money market or another chosen numeraire) are martingales.

A numeraire is any positive asset used as a reference for pricing other assets. The theorem generalizes pricing by stating that for any numeraire, the ratio of asset prices to the numeraire is a martingale under the risk-neutral measure. This leads to the risk-neutral pricing formula when the numeraire is a zero-coupon bond.

The Black-Scholes model is a foundational concept in quantitative finance, structured around two primary domains: discrete time and continuous time. The derivation of the Black-Scholes formula begins in the discrete time framework, employing specific assumptions:

1. Interest rates remain constant throughout the option's expiry.
2. The stock price can only take two values in the next time step.
3. There is no bid-ask spread for the stock.
4. Arbitrage opportunities do not exist.

While the second assumption—limiting stock price movements to two outcomes—may seem restrictive, it serves as a basis for modeling stock prices as Geometric Brownian Motion in the continuous case. Despite its limitations, the model remains widely used due to its effectiveness.

In the binomial tree model, let the actual probability of upward or downward price movement be set at 1/2 for simplicity. The logarithmic return, defined as $\lambda_n = \log(S_n/S_{n-1})$, leads to the formulation $S_n = S_{n-1}e^{\lambda_n}$. If S_T represents the stock price at maturity T and there are N steps on the binomial tree of size ΔT (where $T = N\Delta T$), we define $Y_T = \log S_T - \log S_0 = \lambda_N + \lambda_{N-1} + \dots + \lambda_1$.

Assuming $E(Y_T) = \mu T$ and $Var(Y_T) = \sigma^2 T$, we can observe that as the number of time steps N increases, both the mean and variance of Y_T remain constant. The expected value of each return $E(\lambda_i)$ equates to $\mu\Delta T$, and the variance $Var(\lambda_i)$ equals $\sigma^2\Delta T$. Given that λ_i is a random variable with two possible outcomes, we define its probabilities in terms of μ and σ.

The risk-neutral probability $p*p^*$p∗ can be expressed as $p*=r(\Delta T-d)u-dp^* = \backslash frac\{r(\Delta T - d)\}\{u - d\}$p∗=u−dr(ΔT−d) , where r is the constant interest rate. By relating this back to the logarithmic returns, we can express $log(1+u)log(1 + u)$log(1+u) in terms of the mean and volatility, leading to an approximation for small ΔT. This establishes relationships between stock price movements, their logarithmic returns, and the parameters of the binomial tree.

As the analysis continues, we derive expressions for the stock price behavior in terms of its expected value and volatility, laying the groundwork for transitioning from discrete time to the continuous case, ultimately leading to the derivation of the Black-Scholes option pricing formula. This transition is facilitated by leveraging the Central Limit Theorem, which provides a framework for understanding the distribution of stock prices over continuous time intervals.

Probability Primer

Stochastic Processes and Markov Chains

Stochastic processes are mathematical models that represent a sequence of random variables ordered in time or space, encapsulating the evolution of systems in a probabilistic manner. Key concepts include:

- **State Space:** The complete set of possible states that a system can inhabit.
- **Trajectory/Path:** A specific realization of the stochastic process over time, capturing the evolution of the system.
- **Types of Stochastic Processes:**
    - **Discrete-time vs. Continuous-time:** Discrete processes involve observations at distinct intervals, while continuous processes allow for observations at any moment.
    - **Discrete-state vs. Continuous-state:** Discrete-state processes consist of countable states, whereas continuous-state processes feature a continuum of possible states.

Markov Chains are a specific type of stochastic process characterized by the memoryless property, where future states depend only on the current state. Key elements include:

- **Transition Matrix:** Represents probabilities of transitioning between states.
- **State Classification:**
    - **Recurrent State:** A state that will definitely be revisited.
    - **Transient State:** A state that might not be revisited once left.
    - **Absorbing State:** A state that, once entered, cannot be exited.
- **Stationary Distribution:** A probability distribution that remains unchanged over time.

The **Chapman-Kolmogorov equations** define n-step transition probabilities in Markov chains.

Queueing Systems study the behavior of queues where entities wait for service. Key components include:

- **Arrival Process:** Describes how entities arrive (e.g., Poisson process).
- **Service Process:** Outlines how entities are serviced (e.g., exponential service times).
- **Queue Discipline:** Rules for service order (e.g., FIFO).
- **Kendall's Notation:** A notation format (A/B/C) to describe queueing systems.

Applications in Finance

Stochastic processes are instrumental in finance, modeling phenomena like asset price movements, interest rate changes, and risk dynamics.

- **Asset Price Modeling:** Used to capture the random fluctuations in asset prices for applications like option pricing and risk management.
- **Risk Management:** Quantifies market, credit, and operational risks using measures like Value-at-Risk (VaR) and Expected Shortfall (ES).
- **Portfolio Optimization:** Involves selecting asset weights to maximize expected returns relative to risk.

Geometric Brownian Motion (GBM) is a key continuous-time stochastic process modeling stock prices, underlying the Black-Scholes model for option pricing. The **Black-Scholes Model** provides a formula for pricing European options based on assumptions about market efficiency and asset return distributions.

Interest Rate Models describe the evolution of interest rates, vital for pricing bonds and managing interest rate risk. Finally, **Value at Risk (VaR)** estimates potential portfolio losses over a specific period at a given confidence level, essential for risk management and regulatory compliance in financial institutions.