Lecture Notes: Introduction to Modeling Time Series

Justin Offutt, VP Quantitative Development

Fintech @ IU — April 2025

1. What is a Stochastic Process?

A **stochastic process** is a mathematical model for a collection of random outcomes that evolve according to some index set, T:

$$\{X_t\}_{t \in T}$$

Each X_t is a random variable, or a function that assigns a real value to each possible outcome in a probability space:

$$X_t : \Omega \to \mathbb{R}$$

So in order to define a stochastic process, we need a probability space:

- **Sample space** Ω: the set of all possible outcomes (I like to think of this as every possible reality in a multidimensional universe)
- **Event space (sigma-algebra)** $\mathcal{F}$: the set of all events ($\mathcal{P}(\Omega)$ we can sensibly assign probabilities to
- **Probability measure** $\mathbb{P}$: a function that assigns a probability to each event in $\mathcal{F}$, s.t:
  - $\mathbb{P}(A) \geq 0$ for all $A \in \mathcal{F}$
  - $\mathbb{P}(\Omega) = 1$ (the probability of something occurring in the sample space is certain)
  - $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ if $A \cap B = \emptyset$ (the probability of events is additive)

Each element $\omega \in \Omega$ represents one possible scenario (or one world in a multiverse), and the stochastic process is the underlying rule set which governs what the value of each X_t would be in that scenario.

Chess Analogy

Imagine tracking a chess player's rating over time:

- The rating changes after each match depending on a win/loss (no draw for simplicity sake)

- The result of each match is uncertain, affected by randomness (opponent, fatigue, luck)
- The rating is a value that evolves with some degree of uncertainty over the course of matches played

This is exactly what a stochastic process models: a variable with uncertainty (rating) evolving across an index (games played).

Simple Example: Chess Rating Process

Let's define a very simple stochastic process:

- Let $T = \{1, 2, 3\}$: representing the first three matches played.
- Start with rating 800
- In each game:
  - Win: gain 10 points
  - Loss: lose 10 points
- $\Omega = \{WWW, WWL, WLW, WLL, LWW, LWL, LLW, LLL\}$: all possible ordered win/loss sequences.

Event space: $\mathcal{F} = 2^{\Omega}$, the power set of Ω. That means it contains every subset of Ω, for example:

- $\{WWW\}$: "win all games"
- $\{WWL, WLW, LWW\}$: "win 2 out of 3 games"
- $\{LLL\}$: "lose all games"
- $\{WLW, WLL, LWL\}$: "exactly one win"
- Ω: "some outcome happens"
- $\emptyset$: "no outcome happens"

Probability measure: Suppose the player has a 60% chance of winning each game independently, derived from a win rate and other uncertain metric measured.

- $\mathbb{P}(WWW) = 0.6^3 = 0.216$
- $\mathbb{P}(WWL) = 0.6^2 \times 0.4 = 0.144$
- $\mathbb{P}(WLW) = 0.6 \times 0.4 \times 0.6 = 0.144$
- $\mathbb{P}(LLL) = 0.4^3 = 0.064$
- $\mathbb{P}(WLW \text{ or } LWW) = 0.144 + 0.144 = 0.288$

Define the stochastic process: Let $X_t(\omega)$ be the player's rating after game t if the outcome is ω.

For example, for $\omega = WLW$ (think of this as one possible world out of the entire sample space, the world in which you win, lose, and win the three chess games):

$$X_1(\omega) = 810$$
$$X_2(\omega) = 800$$
$$X_3(\omega) = 810$$

Then the realization is:

$$\{X_1(\omega), X_2(\omega), X_3(\omega)\} = \{810, 800, 810\}$$

Think of it like this: the probability space is what governs the potential outcomes for the random process X_t, the above series is simply a realization (observation) of one possible $\omega \in \Omega$, in which your rating after the three games is 800 ELO.

How the Event Space Plays Into the Stochastic Process

The event space $\mathcal{F}$ is essential because the events allow us to ask meaningful questions about the stochastic process.

- Example event: $A = \{WWW, WWL\}$ = "player wins first 2 games"
- Compute: $\mathbb{P}(A) = \mathbb{P}(WWW) + \mathbb{P}(WWL) = 0.216 + 0.144 = 0.36$
- Another event: $B = \{WLW, LWW\}$ = "player wins exactly 2 games with one loss in the middle or first"
- Compute: $\mathbb{P}(B) = 0.144 + 0.144 = 0.288$
- Suppose we define a random variable X_3: the player's rating after 3 games
- We can form events like:

$$\{\omega \in \Omega : X_3(\omega) > 820\} = \{WWW\} \in \mathcal{F}$$

- Then compute: $\mathbb{P}(X_3 > 820) = \mathbb{P}(WWW) = 0.216$

Bottom line: Without $\mathcal{F}$, we couldn't define or calculate probabilities of outcomes involving X_t. It is what makes the stochastic process probabilistically useful.

2. What is a Time Series?

A **time series** is a special type of stochastic process where the index set T is time — usually $T = \mathbb{N}$ or a finite time horizon $\{1, 2, \ldots, n\}$.

$$\{X_t\}_{t=1}^{n}$$

- Each X_t represents a quantity measured at time t
- The sequence $\{x_t\}$ you observe in the real world is a **realization**
- Time series can be discrete (daily prices) or continuous (e.g., Brownian motion)

Example: Apple stock prices over 4 days:

$$A_4 = \{198.78, 202.54, 196.23, 200.01\}$$

This is one realization of the time series process $\{A_t\}$.

Why Modeling Financial Time Series is So Hard

Let's now reflect on what it would take to define a full stochastic process for a real financial asset, building on our example of Apple stock.

To do this, we would need:

- A sample space Ω: the set of all possible price trajectories Apple stock could take. This isn't just 8 outcomes like in chess, but *infinitely many*, since prices are real numbers and can take an uncountable number of trajectories.
- An event space $\mathcal{F}$: which subsets of these price paths we want to consider "events" (e.g., "price exceeds 210 next week", "price falls 5% this month") — each of which must be mathematically measurable and sensible.
- A probability measure $\mathbb{P}$: which assigns probabilities to all these events. This is incredibly hard because market movements depend on:
  - Global macroeconomics
  - Trader psychology
  - Random news shocks

    $\vdots$

    $\vdots$
  - (this list is infinitely long)

In practice, we often have to
guess the structure of $\mathbb{P}$ from statistical data and the development of theory.

Compare this to the chess example:

- Only 8 outcomes, easily modeled.
- Probabilities can be assigned using coin-like assumptions.

Markets? Infinitely many outcomes. No obvious assumptions. Probabilities must be guessed.

Why Quants Get Paid So Much

This complexity is exactly why **quants** are so valuable. They attempt to:

- Build stochastic processes that reflect real market dynamics
- Estimate probabilities of events like crashes, spikes, or volatility shifts
- Predict expected returns, risk, and correlations

Doing this well requires:

- Deep math (measure theory, stochastic calculus, optimization)
- Coding and simulation (Python, C++, Monte Carlo)
- Intuition for economic and behavioral patterns

3. Stationarity and Non-Stationarity

Before moving into a broad overview of a few time-series models, understanding stationarity is crucial because many time series models (like AR, MA, ARIMA) rely on it to produce meaningful insight.

Stationarity

A time series is **stationary** if its statistical properties do not change over time.

- **Strict Stationarity:** The full probability distribution of the process is invariant under shifts in time (it doesn't matter where you look).
- **Weak Stationarity (Second-order):**
  - **Constant mean:**
  $$\mathbb{E}[X_t] = \mu \quad \text{for all } t \in T$$
  - **Constant variance:**
  $$\mathrm{Var}(X_t) = \mathbb{E}[(X_t - \mu)^2] = \sigma^2 \quad \text{for all } t \in T$$
  - **Autocovariance depends only on lag:**
  $$\mathrm{Cov}(X_t, X_{t+h}) = \gamma(h) \quad \text{for all } t, h \in T$$
- If a series satisfies weak stationarity, most linear models (like AR) can be, and in practice are applied.

Non-Stationarity

A time series is **non-stationary** when its mean, variance, or other properties change over time.

- Can arise from trends, seasonality, structural changes, or unit roots.
- Non-stationary data can often be made stationary by transformations:
  - **Differencing:** Subtracting the previous value
  - **De-trending:** Removing a trend
  - **Log transformations:** Reducing exponential growth behavior

Example: Stock prices often exhibit a long-term upward trend violating stationarity.

4. Overview of Common Time Series Models

Each of the models we study in time series analysis is an attempt to represent certain types of stochastic processes. They are just *attempts*.

White Noise

- A white noise process consists of i.i.d. random variables with constant mean and variance, and no autocorrelation.
- Models pure randomness. 0 signal and structure.
- Often used as a baseline or residual correction term in more complex models. If a model is effective, its residuals should look like white noise.
- **Example:** The day-to-day random variation in stock returns due to uninformative retail trades.

Autoregressive (AR) Models

- AR models assume the current value depends on a finite number of past values.
- Captures momentum, inertia, or short-term memory.
- Used when a variable shows correlation with its own lagged values (remember auto-covariance).
- **Example:** Stock prices with short-term momentum trends (think momentum or swing trading).

ARIMA Models

- ARIMA stands for AutoRegressive Integrated Moving Average.
- It combines autoregression, differencing to remove trends, and moving averages.

- Designed for non-stationary time series with trends, cycles, and/or seasonality.
- **Example 1:** Electricity demand over time, which varies seasonally.
- **Example 2:** Monthly airline passenger counts (very seasonal).
- **Example 3:** GDP or inflation, where long-term upward trends exist but short-term fluctuations matter.

GARCH Models

- GARCH stands for Generalized AutoRegressive Conditional Heteroskedasticity.
- Models time varying volatility (think volatility clustering).
- Useful when the variance is not constant and tends to cluster over time.
- **Example 1:** Financial returns during crises or news releases.
- **Example 2:** Exchange rates which may experience calm and volatile periods.
- **Example 3:** Cryptocurrency price movements.

Takeaway: These models try to capture specific characteristics of real-world time series. No model fully captures the complexity of financial markets, they are just tools. Each model is useful when its assumptions *roughly* match the behavior of the data.