An Introduction to Probability

With Applications to Computational Finance using R

Martin Summer

11 January, 2026

The Motto of this Course

The Motto of this Course

"What I hear, I forget; What I see, I remember; What I do, I understand"

Confucius, 551-479 BC

What is this lecture about?

Probability, R, Finance

- These lecture notes support the course **An Introduction to Probability - with Applications to Computational Finance using R**.
- The course introduces essential probability concepts that every finance practitioner must understand using R within the context of computational finance.
- **Why Probability?**
    - Because Finance is about making decisions under uncertainty.
    - Probability is the most powerful conceptual tool we have at hand to think about uncertainty and decision making under uncertainty in a systematic an rational way.
- **Why R ?**
    - Because it is a language which is well adapted to probability and hands on coding of applied probability problems.
    - The computer can make abstract problems tangible.

Course Objectives

- **Understand the fundamentals of probability**: concepts, rules, and theorems.
- **Learn R programming for probability simulations**:
    - Generating random variables.
    - Visualizing probabilities.
    - Solving real-world problems computationally.
- **Apply probability concepts to finance**: Use cases in risk management, pricing, and investment decisions.

Building Abstract Concepts

Conceptual Experiments

- How can we build abstract concepts like **probability**, **random phenomena**, and **chance** by our own hands?
- Probability gains practical value through real or conceptual experiments, such as:
  - Future changes in a **stock price index**.
  - The future value of a **portfolio of securities**.
  - The chance that a **creditor cannot repay a loan**.

Simulations with Computers

- Computers allow us to **simulate** a wide variety of random phenomena:
    - Models of random **fluctuations in asset prices**.
    - Models of **financial risks**.
    - Exploring **future scenarios** through simulation.
- The arrival of computers has been a **revolution** in the mathematics of probability:
    - Enables both theoretical reflections and practical simulations.
    - Abstract concepts can be built **hands-on** using computational tools.

R Example for Simulation

Simulating a Stock Price in R

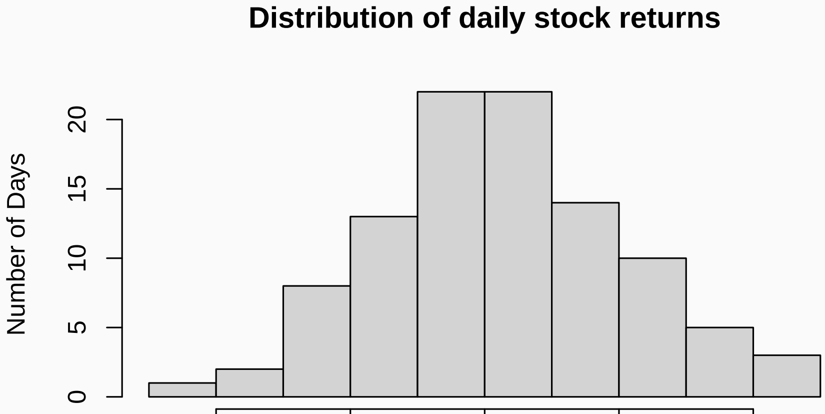
```r
# Simulate daily returns for a stock


# Ensure reproducibility
set.seed(123)
# Number of days
n <- 100
# Normal distribution
daily_returns <- rnorm(n, mean = 0.001, sd = 0.02)
# Compute cumulative returns, starting price = 100
price <- cumprod(1 + daily_returns) * 100
price[1:10]
```

```
 [1]  98.97905  98.62237 101.79547 102.04081 102.40670 106.02
 [8] 104.50237 103.17132 102.35490
```

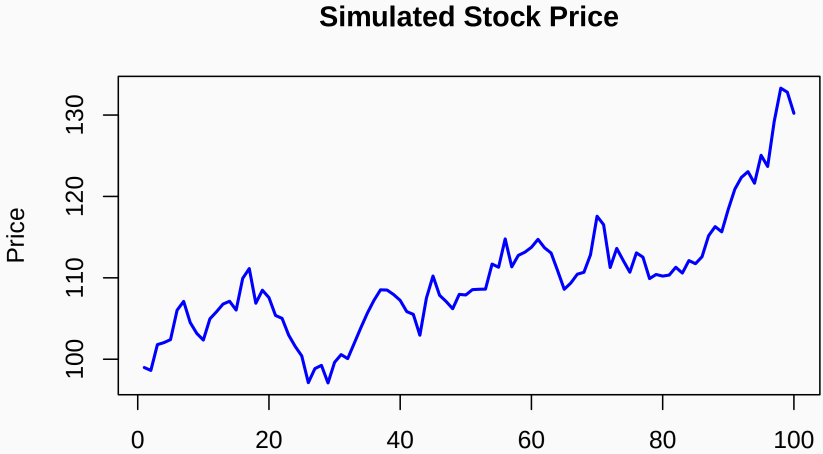
Visualize the distribution of daily returns:

```r
hist(daily_returns,
     main = "Distribution of daily stock returns",
     xlab = "Returns",
     ylab = "Number of Days")
```



Visualize the Price Dynamics

```r
# Visualize the stock price
plot(price, type = "l", col = "blue", lwd = 2,
     main = "Simulated Stock Price",
     xlab = "Days", ylab = "Price")
```



Insights from the Example

Key Takeaways

- Highlights:
  - How randomness influences stock price movements.
  - The usefulness of R for simulations and visualization.
- Critical Reflection:
  - The real world is **different** from theoretical and simulated worlds.
  - The real world is **richer and more complex** than both theory and simulation.
  - Features in the real world may be **absent** in theory or simulations but are crucial for explanations.

The Interdependence of Theory, Simulation, and Reality

- Theoretical concepts, simulations, and applications are **interdependent**.
- Developing a sense of their differences and connections is a core goal of this course.

Building Virtual Objects and Simulations in R

The Power of Simulations

- Simulations allow us to:
    - **Construct and manipulate** virtual objects.
    - Explore and analyze random phenomena.
    - Understand complex systems through hands-on experimentation.

Why R and RStudio?

- R Programming Language:
    - Well-suited for probability and hands-on coding of applied problems.
    - Widely used in data analysis, statistics, and data science.
- RStudio IDE:
    - Provides an integrated environment for coding, visualization, and simulation.
    - A popular tool in both industry and academia.

Leveraging AI Tools in Learning

The Role of Large Language Models (LLMs)

- Since this course was first taught in 2021, AI tools such as large language models (LLMs) have emerged.
- Examples include ChatGPT, Claude, and Gemini.
- These tools create new opportunities for learning and reinforcing concepts:
    - Generate examples.
    - Explain difficult topics.
    - Debug R code.
    - Translate code into familiar languages.

Integrating AI into This Course

- Throughout the course, I will show you how to:
  - Use LLMs to **enhance your learning experience**.
  - Leverage AI for **interactive and engaging studies**. # Using an LLM for Learning

Benefits of LLMs in This Course

- ChatGPT offers capabilities to support learning, such as:
  - Clarifying concepts.
  - Generating examples.
  - Debugging R code.
  - Practicing exercises.
  - Simulating discussions.
  - Learning best practices in R programming.
- **Example Use Cases**:
  - **Clarify Concepts**:

    > “Explain the concept of a probability distribution with an example.”
  - **Generate Examples**:

    > “Can you give me an example of a random variable and how it applies in finance?”

- **Example Use Cases**:
  - **Clarify Concepts**:

    > “Explain the concept of a probability distribution with an example.”
  - **Generate Examples**:

Practical Examples of using an LLM

More Use Cases

- **Debug R Code**:

  Paste your code and ask for help: > "Why does this R code not run, and how can I fix it?"

- **Practice Exercises**:

  Generate exercises to reinforce learning: > "Create three exercises to practice calculating probabilities for dice rolls."

- **Simulate Discussions**:

  Test your understanding with interactive discussions: > "I think the variance of a constant is zero. Am I correct? Explain why or why not."

- **Learn R Best Practices**:

  Ask for coding tips: > "What are the best practices for writing clean and efficient R code?"

Important Considerations

Verifying AI Outputs

- While ChatGPT is powerful, always **verify outputs**, especially for:
    - Complex calculations.
    - Detailed explanations.
    - Programming suggestions.
- Use additional tools for cross-checking:
    - **Wolfram|Alpha**:

      A computational engine available for free:

      https://www.wolframalpha.com/

Downloading and Installing R

Steps to Get R

- R is an open-source project maintained by an international team of developers.
- The software is available at:
  - Comprehensive R Archive Network (CRAN)

Download Instructions

1. Visit the “Download and install R” section at CRAN.
2. Select the link corresponding to your operating system (Windows, Mac, or Linux).
3. Choose **precompiled binaries** for an easier installation.
   - Optionally, build R from source if you have the tools and expertise.
4. Install either the 32-bit or 64-bit version:
   - **64-bit versions** handle larger files and datasets more efficiently.

Downloading and Installing RStudio

Why Use RStudio?

- RStudio: A user-friendly application that simplifies R coding.
- Provides a consistent interface across all operating systems.

Steps to Install RStudio

1. Visit: https://posit.co/download/rstudio-desktop/
2. Select **RStudio Desktop** and follow the download instructions.
    - RStudio Desktop is free to use.
3. Ensure you have a version of R installed before using RStudio.

Ready to Start

- If R and RStudio are successfully installed, you're ready to begin.
- Throughout this course:
  - I will demonstrate code using RStudio.
  - Other options (e.g., Jupyter Notebooks) are available for advanced setups.

Using R with Jupyter Notebooks

Advanced Setup

- R can also be used in Jupyter Notebooks.
- For those interested:
    - Ask ChatGPT for step-by-step instructions: > “Please give me a step-by-step instruction on how to set up R in Jupyter Notebooks.”
- Note: This setup will not be covered in this course, but you’re welcome to explore it further.

Signing Up for an LLM

Steps to Create a Free Account

1. Visit website
2. Click “Create a free account.”
    - Sign up using an email address, Google, or Microsoft account.
3. Verify your email and complete the registration process.
4. Log in to start using LLM for learning and exploration.

Free vs. Paid Versions

- The free versions of most LLMs are sufficient for the learning tasks in this course.
- Paid subscriptions offer higher usage limits and additional features.
- Explore the free version of your preferred LLM to understand its potential.

Prerequisites

An Intuitive Introduction

- This course is an **elementary introduction** to probability and R.
- **No prior knowledge** of probability or R is required.
- You can rely on these **lecture notes** and slides without needing additional textbooks.

Designed for All Levels

- For newcomers:
    - Learn essential concepts from scratch.
    - Develop curiosity and excitement about the field.
- For experienced learners:
    - Gain new perspectives and deepen understanding of familiar concepts.

Course Structure

Parallel Development of Concepts

- Probability and R concepts are developed **together**.
- Build knowledge by **constructing and experimenting** with concepts on the computer.
- Reinforces:
    - Probability understanding.
    - R programming skills.

Main Chapters and Projects

- The course is divided into **five chapters**:
    - Each corresponds to one double lecture.
- After each lecture:
    - Work on a **project** to practice and deepen understanding.
    - Projects include discussions and worked solutions.

Lecture Highlights

Lecture 1: Foundations

- Explore the historical context of probability.
- Learn basic concepts using a **coin-toss game**:
    - Application: Modelling security prices.
- Address a real-world problem:
    - Understand coincidences and cryptographic safety (e.g., hash functions in Bitcoin).
- **Project 1**: Design secure transaction identifiers for a digital payment system.

Lecture 2: Frequencies and Data Manipulation

- Discuss probability and frequency relationships.
- Explore **Benford's Law** for detecting anomalies in data.
- Enhance R skills with data manipulation and structures.
- **Project 2**: Analyze and manipulate data for probabilistic insights.

Lecture Highlights (Continued)

Lecture 3: Conditional Probability

- Understand **dependence** and its role in finance.
- Learn probability updates with new data.
- Applications:
    - Risk management.
    - Investment decisions.
- **Project 3**: Apply conditional probability to real-world financial scenarios.

Lecture 4: Random Variables and Asset Dynamics

- Explore:
    - Expected value.
    - Variance.
    - Covariance and correlation.
- Understand **binomial lattice models** for asset price dynamics.
- Develop R skills for programming control structures.
- **Project 4**: Simulate asset dynamics using probabilistic models.

Lecture 5: Continuous Random Variables

- Study the **normal distribution**:
    - Its power and limitations in finance.
- Optimize R code for performance and efficiency.
- **Project 5**: Analyze continuous random phenomena and risks in finance.

Acknowledgements

Special Thanks

- Branko Urošević:
    - For entrusting me to develop this course and supporting the project.
- Past Students:
    - For their feedback and enthusiasm, which improved these notes.

References

Key Resources

1. **Probability**:
    - William Feller: *An Introduction to Probability Theory and Its Applications.*
    - Karl Schmedder: *An Intuitive Introduction to Probability* (Coursera).
2. **R Programming**:
    - Garrett Grolemund: *Hands-On Programming with R.*
3. **Finance**:
    - David Luenberger: *Investment Science.*
4. **History and Philosophy**:
    - Persi Diaconis and Brian Skyrms: *10 Great Ideas About Chance.*

- These notes combine well-known ideas with new approaches to teaching probability and R.