BU.610.740: Forecasting Models for Business Intelligence
Chapter 0: Prelude

Ali Eshragh

Spring II, 2024

Topics Covered in These Slides

1. **Introduction**
   - Time Series
   - Case Study
2. **Essentials of Probability**
   - Probability Distribution Function
   - Mathematical Expectation
   - Normal Distribution

Time Series

Definition (Time Series Data)

A **time series** is a collection of **observations** indexed according to the order they are obtained in **time**.

Definition (Time Series Analysis)

Time series analysis encompasses a range of **statistical models** and **machine learning techniques** aimed at **analyzing** time series data.

Objective

The **primary objective** of time series analysis is to **develop** appropriate models that provide plausible descriptions for time series data to accurately **forecast** the **future** behavior of the system.

Examples

- Apple **daily** stock price
- Walmart **weekly** retail sales
- Amazon **monthly** demand for electronics
- Microsoft **annual** revenue

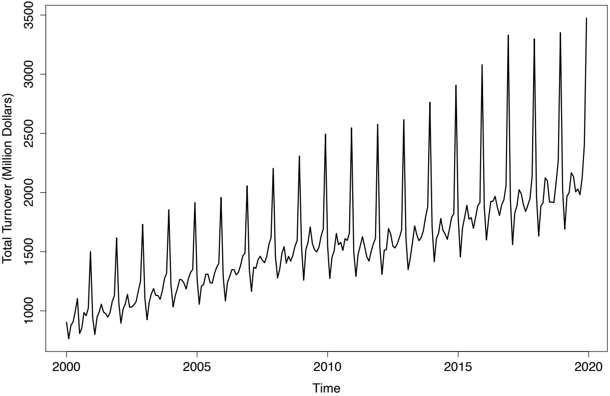
Applications

- In **business analytics**, time series analysis is key for **forecasting** sales trends, **analyzing** consumer behavior over time, and **optimizing** inventory levels, aiding in **strategic** decision-making and operational efficiency.

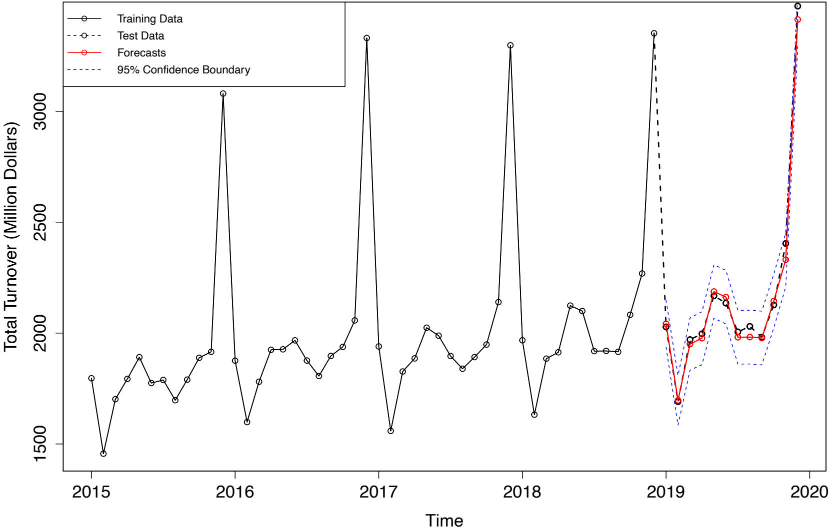
- In **health analytics**, time series analysis enables the **prediction** of disease spread, **tracks** patient admissions, and **assesses** intervention effectiveness, guiding resource allocation and **public health** strategies.

- In **financial analytics**, time series analysis is essential for **forecasting** market movements, **assessing** risk, and **developing** trading strategies, enabling investors and institutions to make **data-driven** investment decisions.

Monthly Turnover of Clothing in Australia



Monthly Turnover of Clothing in Australia Forecasts



Quote

Chatfield (2004)

"Anyone who tries to analyze a time series **without plotting** it first is asking for **trouble**."

Outline

1. Introduction
   - Time Series
   - Case Study
2. Essentials of Probability
   - Probability Distribution Function
   - Mathematical Expectation
   - Normal Distribution

Probability Density Function

Definition (Probability Density Function)

The function $f_X(x)$ is defined as the **probability distribution function** or **probability density function** (**pdf**) for the **continuous** random variable X, over the set of real numbers, if the following conditions are met:

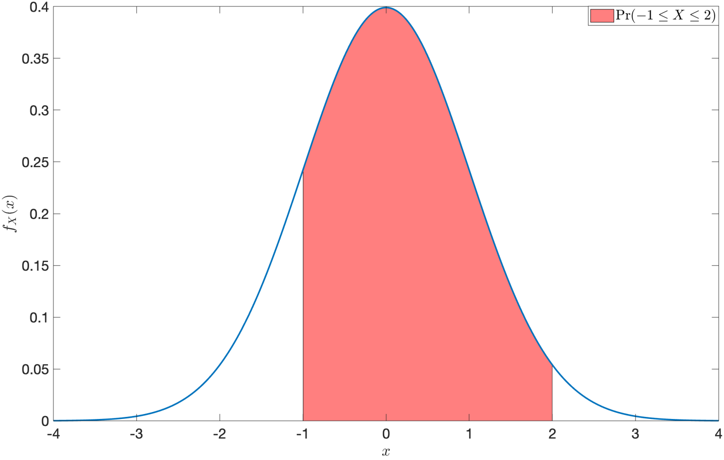
1. $f_X(x) \geq 0 \quad$ for all $x \in \mathbb{R}$
2. $\int_{-\infty}^{\infty} f_X(x)dx = 1$

Probability and PDF

For a **continuous** random variable X with pdf $f_X(x)$,

$$\Pr(a < X < b) := \int_a^b f_X(x)dx$$

Probability Distribution with Shaded Area



Joint Probability Density Function

Definition (Joint PDF)

The function $f_{X_1,\ldots,X_n}(x_1,\ldots,x_n)$ is defined as the **joint probability density function** for the **continuous** random variables $X_1,\ldots,X_n$, over the set of real numbers, if the following conditions are met:

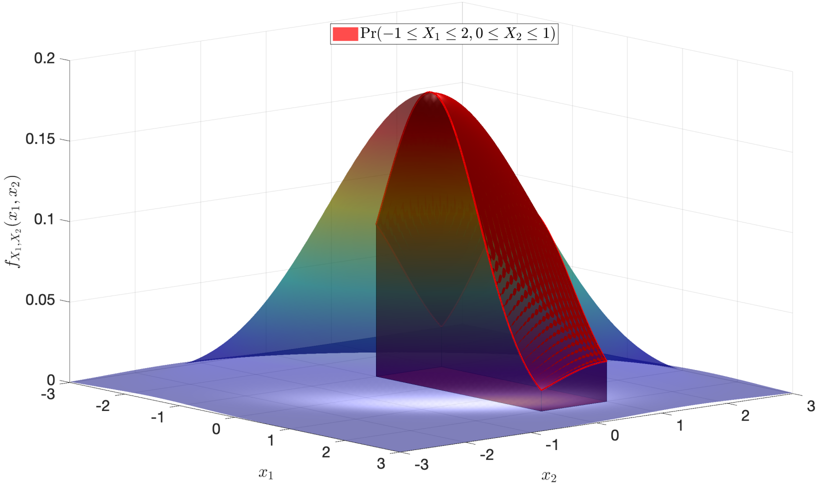
1. $f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) \geq 0 \quad$ for all $(x_1,\ldots,x_n) \in \mathbb{R}^n$
2. $\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) dx_1 \ldots dx_n = 1$

Probability and Joint PDF

For continuous random variables $X_1,\ldots,X_n$ with joint pdf $f_{X_1,\ldots,X_n}(x_1,\ldots,x_n)$,

$$\Pr(a_1 < X_1 < b_1, \ldots, a_n < X_n < b_n) := \int_{a_n}^{b_n} \cdots \int_{a_1}^{b_1} f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) dx_1 \ldots dx_n$$

Joint PDF with Highlighted Volume



Independent Random Variables

Definition (Independent Random Variables)

Let $X_1, \ldots, X_n$ be random variables with a joint pdf $f_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$ and marginal distributions $f_{X_1}(x_1), \ldots, f_{X_n}(x_n)$, respectively. The random variables $X_1, \ldots, X_n$ are said to be **statistically independent** if and only if

$$f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = f_{X_1}(x_1) \ldots f_{X_n}(x_n)$$

for all values of $(x_1, \ldots, x_n) \in \mathbb{R}^n$.

Mean

Definition (Mean)

Let X be a random variable with pdf $f_X(x)$. The **expected value**, or **mean**, of the random variables X is

$$\mathbb{E}[X] := \int_{-\infty}^{\infty} x f_X(x) dx$$

and is **denoted** by μ.

Variance

Definition (Variance)

Let X be a random variable with pdf $f_X(x)$ and mean μ. The **variance** of the random variables X is

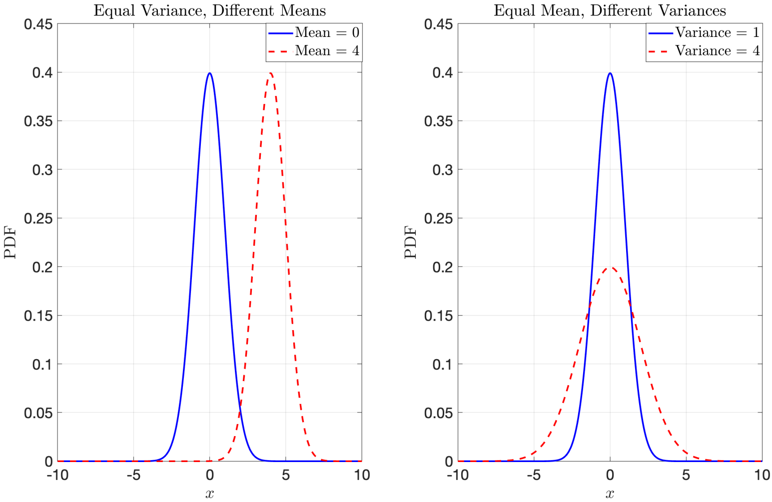
$$\begin{aligned} \mathrm{Var}(X) &:= \mathbb{E}[(X-\mu)^2] \\ &= \mathbb{E}[X^2] - \mu^2 \end{aligned}$$

and **denoted** by σ^2.

Definition (Standard Deviation)

The positive **square root** of the variance, σ, is known as the **standard deviation** of the random variable X.

Mean vs. Variance: Location vs. Scale



Covariance

Definition (Covariance)

Let X and Y be two random variables with means μ_X and μ_Y, respectively. The **covariance** of the two random variables X and Y is defined as

$$\begin{aligned}\mathrm{Cov}(X,Y) &:= \mathbb{E}[(X-\mu_X)(Y-\mu_Y)] \\ &= \mathbb{E}[XY]-\mu_X\mu_Y\end{aligned}$$

- It can be **shown** that $\mathrm{Cov}(X,Y) \in \mathbb{R}$
- It is readily **observed** that $\mathrm{Cov}(X,X) = \mathrm{Var}(X)$

Correlation

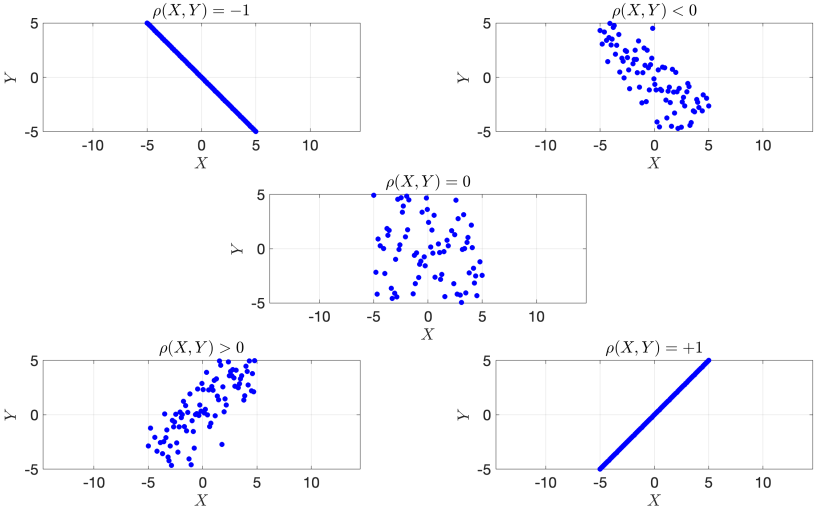
Definition (Correlation)

Let X and Y be two random variables with standard deviations σ_X and σ_Y, respectively. The **correlation** between the two random variables X and Y, **denoted** by $\rho(X, Y)$, is given by

$$\rho(X, Y) := \frac{\mathrm{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

- It can be **shown** that $-1 \leq \rho(X, Y) \leq 1$
- If $\rho(X, Y) = 0$, then we say X and Y are **uncorrelated** random variables; otherwise, they are considered **correlated**.
- In some contexts, $\rho(X, Y)$ is also referred to as the **correlation coefficient** or the **linear correlation**.

The Spectrum of Correlation



Independence vs. Uncorrelated

- If X and Y are two **independent** random variables, then they are **uncorrelated**, that is, $\rho(X, Y) = 0$

- However, if X and Y are two **uncorrelated** random variables, then X and Y **may or may not** be **independent**.

Example

Example

A company employing **dynamic pricing** for online courses examines user engagement, denoted by X, **uniformly** distributed over the interval $[-1, 1]$. Therefore, the **probability density** function of X is given by:

$$f(x) = \begin{cases} \frac{1}{2} & \text{for } -1 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Revenue, modeled as $Y = 10X^2$, shows the magnitude of engagement impacts revenue more than its positive or negative direction. Given this, are X and Y:

1. **dependent** or **independent**?
2. **correlated** or **uncorrelated**?

Normal Distribution

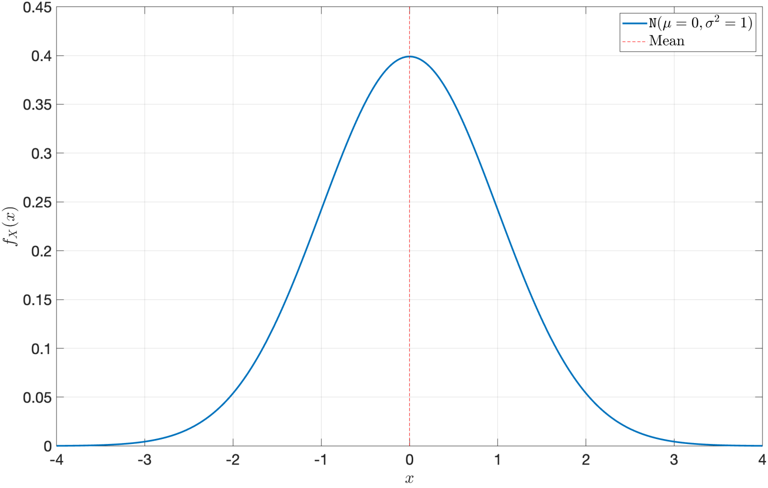
Definition (Normal Distribution)

The continuous random variable X has the **Normal** (or **Gaussian**) distribution with **parameters** μ and σ^2, denoted by $X \sim \mathtt{N}(\mu, \sigma^2)$, if and only if the **pdf** of the random variable X is

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \text{for } x \in \mathbb{R}$$

- It can easily be **shown** that $\mathbb{E}[X] = \mu$ and $\mathsf{Var}(X) = \sigma^2$

Key Characteristics of the Normal Distribution



Bivariate Normal Distribution

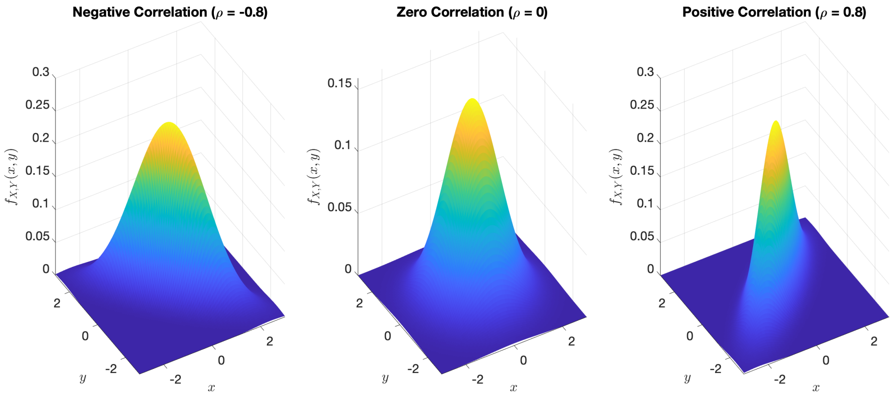
Definition (Bivariate Normal Distribution)

Let the two random variables X and Y have **Normal** distributions $\mathtt{N}(\mu_X, \sigma_X^2)$ and $\mathtt{N}(\mu_Y, \sigma_Y^2)$, respectively, with a **correlation** coefficient of ρ. The joint distribution of (X, Y) is called the **Bivariate Normal** distribution with the **joint pdf**

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \times \exp\left(\frac{-1}{2(1-\rho^2)}\left(\frac{(x-\mu_X)^2}{\sigma_X^2} - \frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}\right)\right)$$

for $(x,y) \in \mathbb{R}^2$.

Bivariate Normal Distributions: Varying Correlations



Assignment 1

Question 1

Show that if two random variables $X \sim \mathtt{N}[\mu_X, \sigma_X^2]$ and $Y \sim \mathtt{N}[\mu_Y, \sigma_Y^2]$ are **uncorrelated**, then they are **independent** as well.

In general, if two random variables are **uncorrelated**, they **may or may not** be **independent**; however, for two **normal** random variables, being uncorrelated **guarantees** they are independent.

References

1. R.H. Shumway and D.S. Stoffer, *Time Series Analysis and Its Applications With R Examples*, Springer, New York, 2017.
2. R.E. Walpole, R.H. Myers, S.L. Myers and K.E. Ye, *Probability and Statistics for Engineers and Scientists*, Pearson, New Jersey, 2017.