Module 00: Preliminaries

Econometrics II

Max Heinze (mheinze@wu.ac.at)
Department of Economics, WU Vienna

Sannah Tijani (stijani@wu.ac.at)
Department of Economics, WU Vienna

October 16, 2025

Introduction

Installing R

Introduction to R with RStudio

Matrices and Vectors

Introduction

Introduction

Welcome to the course **Econometrics II**!

Introduction

Welcome to the course **Econometrics II**!

In this course, we will build on our knowledge from Econometrics I to explore how we can use data to provide evidence for hypotheses and find answers to the questions we pose.

Introduction

Welcome to the course **Econometrics II**!

In this course, we will build on our knowledge from Econometrics I to explore how we can use data to provide evidence for hypotheses and find answers to the questions we pose.

To do this, we need solid **mathematical** and **statistical basics**. Broadly speaking, these are things already covered in school and repeated in the statistics and mathematics lectures in the CBK. Above all, you should remember this from your Econometrics I course – but in case you do not, you can use this **self-study slide set** to refresh your knowledge.

Introduction

Welcome to the course **Econometrics II**!

In this course, we will build on our knowledge from Econometrics I to explore how we can use data to provide evidence for hypotheses and find answers to the questions we pose.

To do this, we need solid **mathematical** and **statistical basics**. Broadly speaking, these are things already covered in school and repeated in the statistics and mathematics lectures in the CBK. Above all, you should remember this from your Econometrics I course – but in case you do not, you can use this **self-study slide set** to refresh your knowledge.

If something is unclear, there will be enough time to ask questions in the course, but understanding this content is a prerequisite for this course.

Introduction

Welcome to the course **Econometrics II**!

In this course, we will build on our knowledge from Econometrics I to explore how we can use data to provide evidence for hypotheses and find answers to the questions we pose.

To do this, we need solid **mathematical** and **statistical basics**. Broadly speaking, these are things already covered in school and repeated in the statistics and mathematics lectures in the CBK. Above all, you should remember this from your Econometrics I course – but in case you do not, you can use this **self-study slide set** to refresh your knowledge.

If something is unclear, there will be enough time to ask questions in the course, but understanding this content is a prerequisite for this course.

The slide set also contains an **introduction to R** with interactive code examples.

Introduction

Welcome to the course **Econometrics II**!

In this course, we will build on our knowledge from Econometrics I to explore how we can use data to provide evidence for hypotheses and find answers to the questions we pose.

To do this, we need solid **mathematical** and **statistical basics**. Broadly speaking, these are things already covered in school and repeated in the statistics and mathematics lectures in the CBK. Above all, you should remember this from your Econometrics I course – but in case you do not, you can use this **self-study slide set** to refresh your knowledge.

If something is unclear, there will be enough time to ask questions in the course, but understanding this content is a prerequisite for this course.

The slide set also contains an **introduction to R** with interactive code examples.

You also have access to **Econometrics I slides** via Canvas. **Review these** if you do not feel confident about your knowledge of what we learned in Econometrics I.

Introduction

Installing R

Introduction to R with RStudio

Matrices and Vectors

Random Variables

Installing RStudio

Installing RStudio

At various points in this course, not least in **assignments**, we will need software to perform statistical calculations. Which software you use is up to you. Our **recommendation** is **R**, as it is what we usually discuss example code in.

Installing RStudio

At various points in this course, not least in **assignments**, we will need software to perform statistical calculations. Which software you use is up to you. Our **recommendation** is **R**, as it is what we usually discuss example code in.

A comfortable way to use R is with the *integrated development environment* **RStudio**. RStudio is the interface we use to write code in R; R itself is a separate program that executes our code and provides results.

Installing RStudio

At various points in this course, not least in **assignments**, we will need software to perform statistical calculations. Which software you use is up to you. Our **recommendation** is **R**, as it is what we usually discuss example code in.

A comfortable way to use R is with the *integrated development environment* **RStudio**. RStudio is the interface we use to write code in R; R itself is a separate program that executes our code and provides results.

Installing R and RStudio

An installation guide and the **download for R** can be found at cran.r-project.org.

An installation guide and the **download for RStudio** can be found at posit.co/download/rstudio-desktop/

Installing RStudio

At various points in this course, not least in **assignments**, we will need software to perform statistical calculations. Which software you use is up to you. Our **recommendation** is **R**, as it is what we usually discuss example code in.

A comfortable way to use R is with the *integrated development environment* **RStudio**. RStudio is the interface we use to write code in R; R itself is a separate program that executes our code and provides results.

Installing R and RStudio

An installation guide and the **download for R** can be found at cran.r-project.org.

An installation guide and the **download for RStudio** can be found at posit.co/download/rstudio-desktop/

Having statistical software (e.g., R) installed is a prerequisite for the course.

Introduction

Installing R

Introduction to R with RStudio

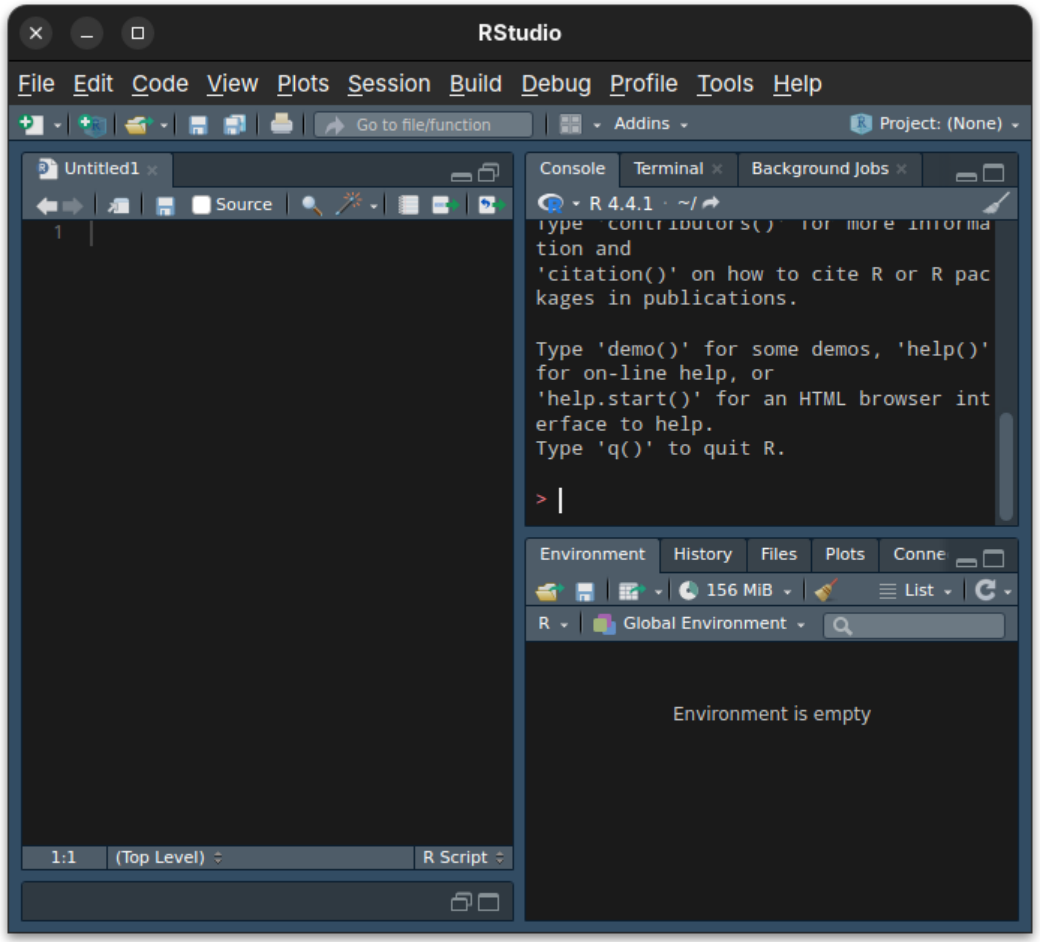
Matrices and Vectors

Random Variables

Analysis of a Random Variable

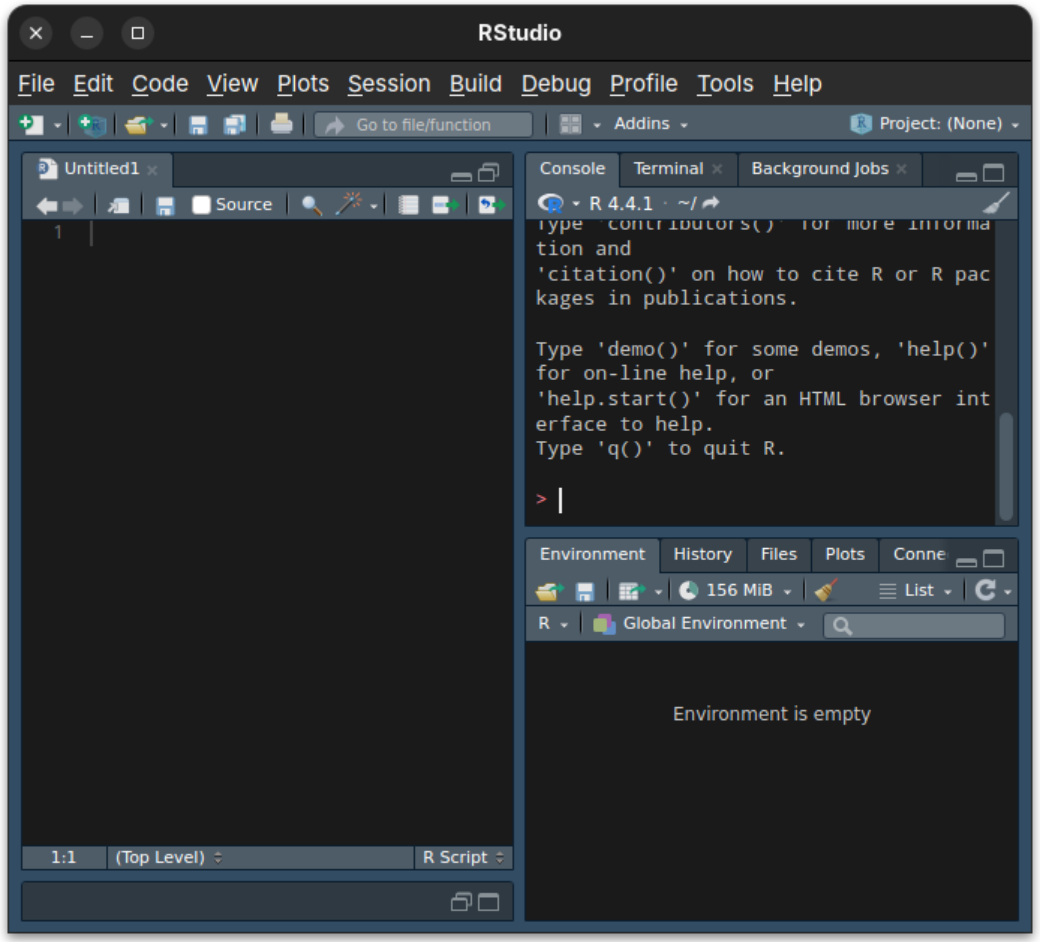
RStudio

A typical **RStudio** window looks like this:



The default layout is slightly different, and the default theme is light. Both can be changed in the settings.

RStudio

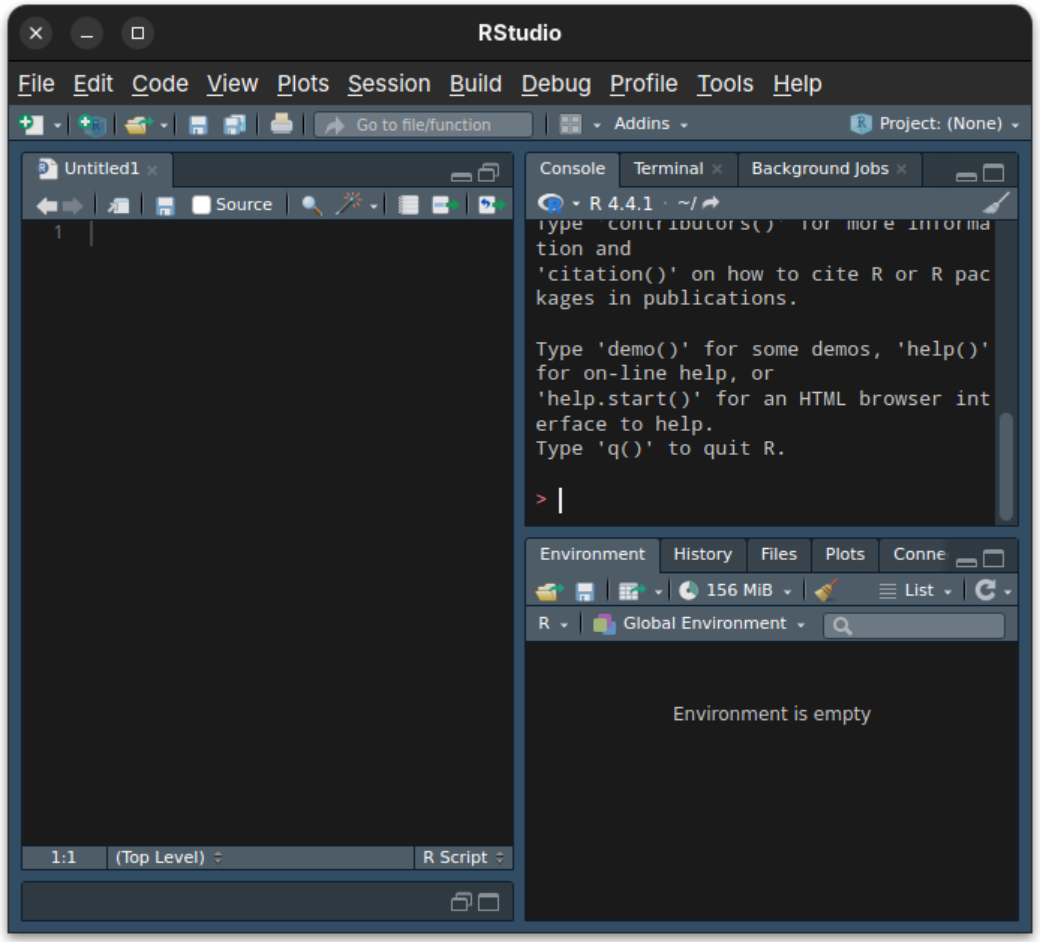


The default layout is slightly different, and the default theme is light. Both can be changed in the settings.

A typical **RStudio** window looks like this:

- In the **Console**, we can execute **R** commands. For example, we can type `1+1` and press `Enter` to confirm.

RStudio

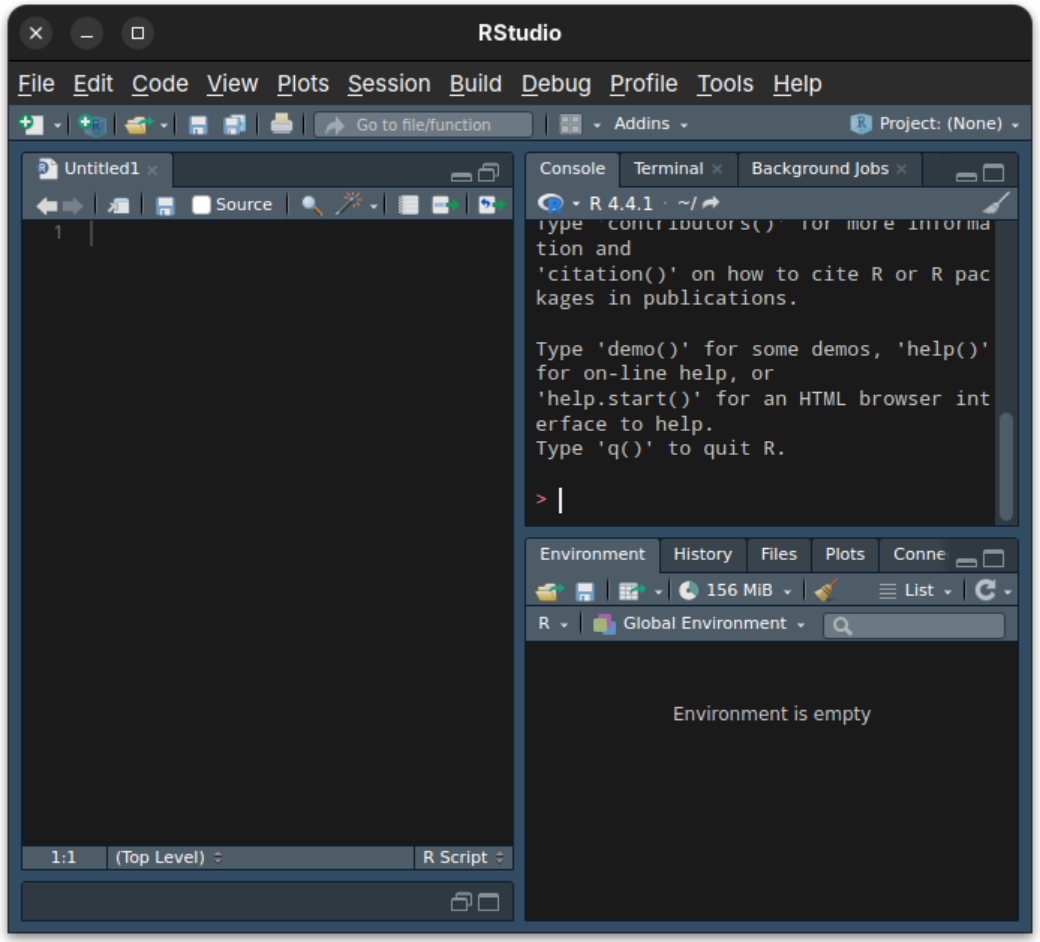


The default layout is slightly different, and the default theme is light. Both can be changed in the settings.

A typical **RStudio** window looks like this:

- In the **Console**, we can execute **R** commands. For example, we can type `1+1` and press `Enter` to confirm.
- The **Environment** shows all currently defined variables and datasets.

RStudio

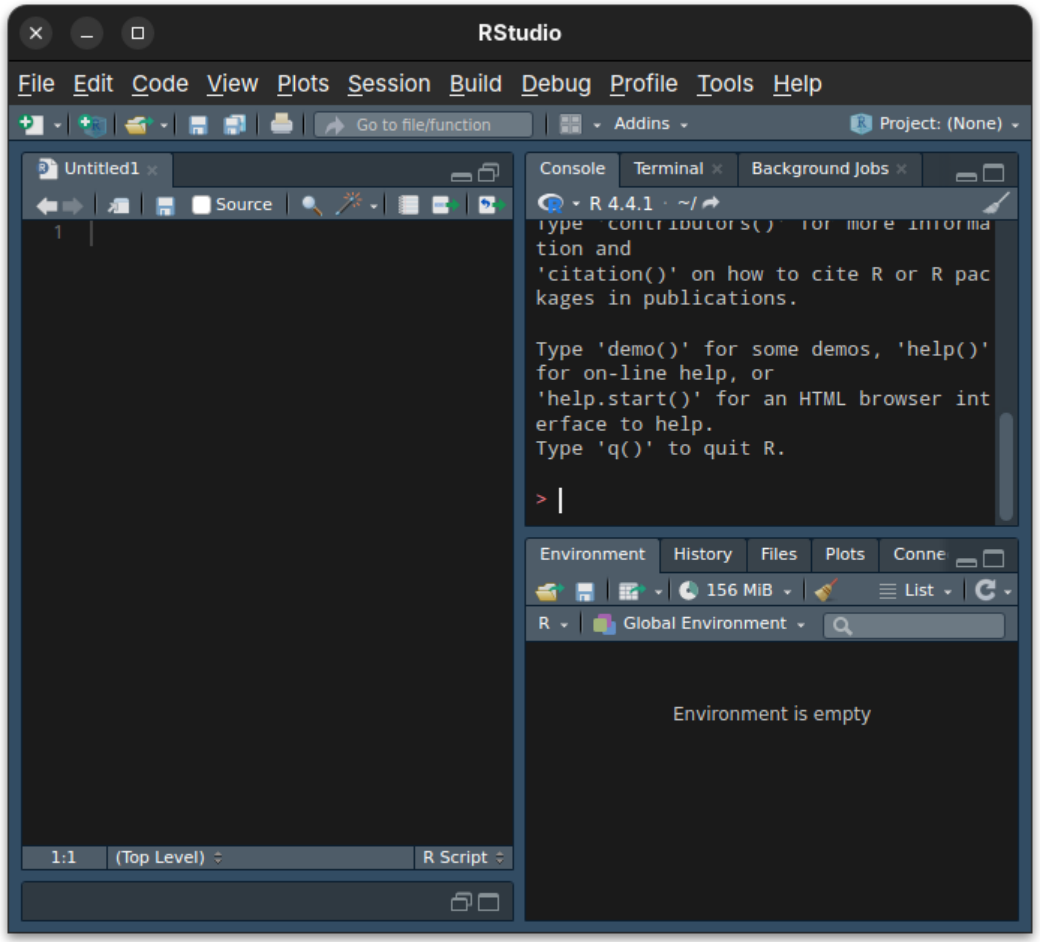


The default layout is slightly different, and the default theme is light. Both can be changed in the settings.

A typical **RStudio** window looks like this:

- In the **Console**, we can execute **R** commands. For example, we can type `1+1` and press `Enter` to confirm.
- The **Environment** shows all currently defined variables and datasets.
- In the **Plots** pane, generated graphics are displayed.

RStudio

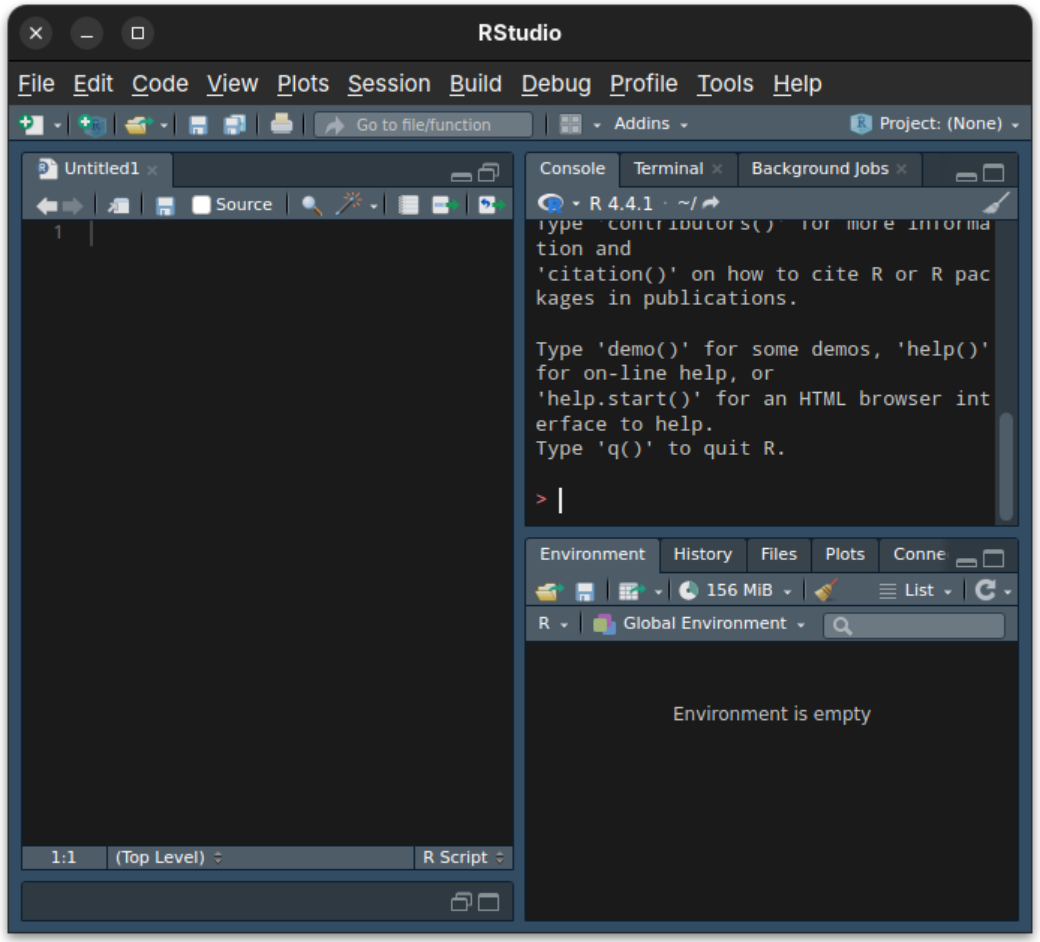


The default layout is slightly different, and the default theme is light. Both can be changed in the settings.

A typical **RStudio** window looks like this:

- In the **Console**, we can execute **R** commands. For example, we can type `1+1` and press `Enter` to confirm.
- The **Environment** shows all currently defined variables and datasets.
- In the **Plots** pane, generated graphics are displayed.
- In the **Help** pane, we can access the documentation.

RStudio



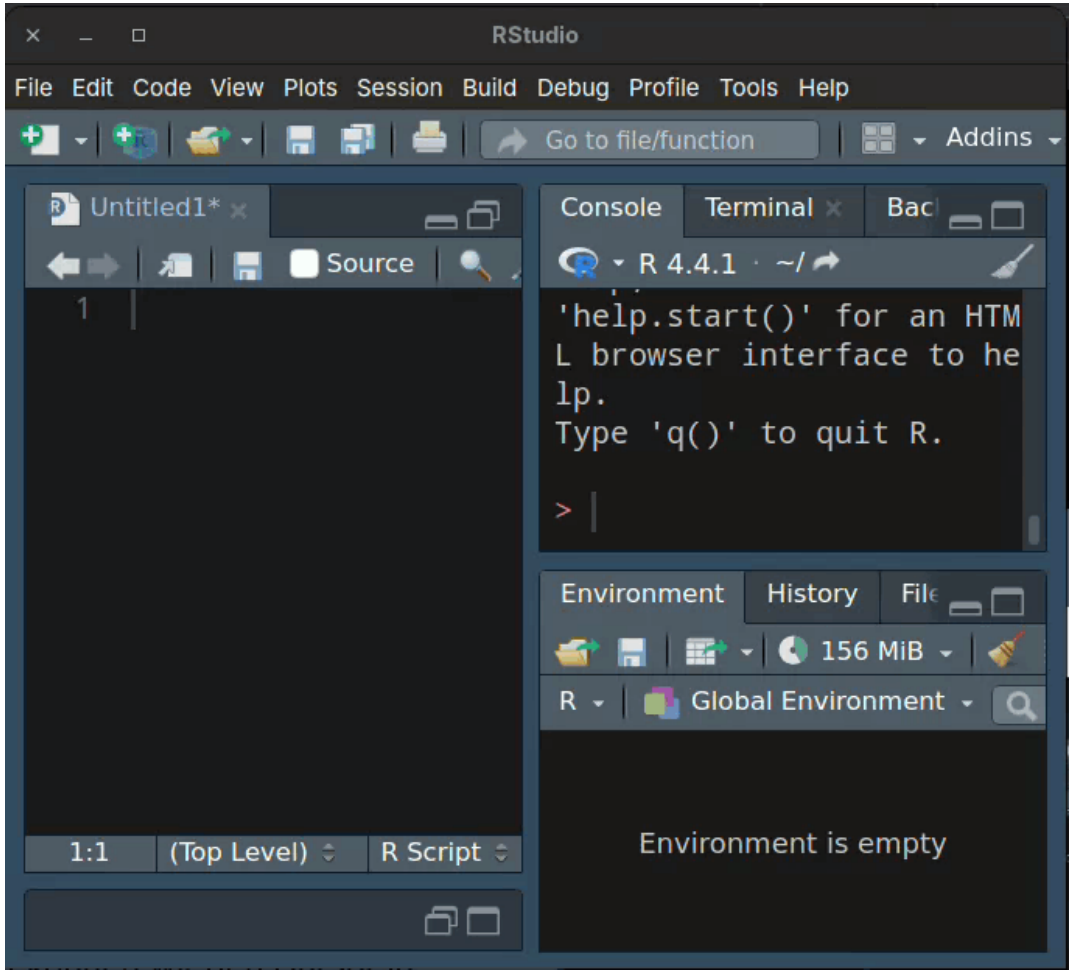
The default layout is slightly different, and the default theme is light. Both can be changed in the settings.

A typical **RStudio** window looks like this:

- In the **Console**, we can execute **R** commands. For example, we can type `1+1` and press `Enter` to confirm.
- The **Environment** shows all currently defined variables and datasets.
- In the **Plots** pane, generated graphics are displayed.
- In the **Help** pane, we can access the documentation.
- In the large pane on the left, we can open a **script**.

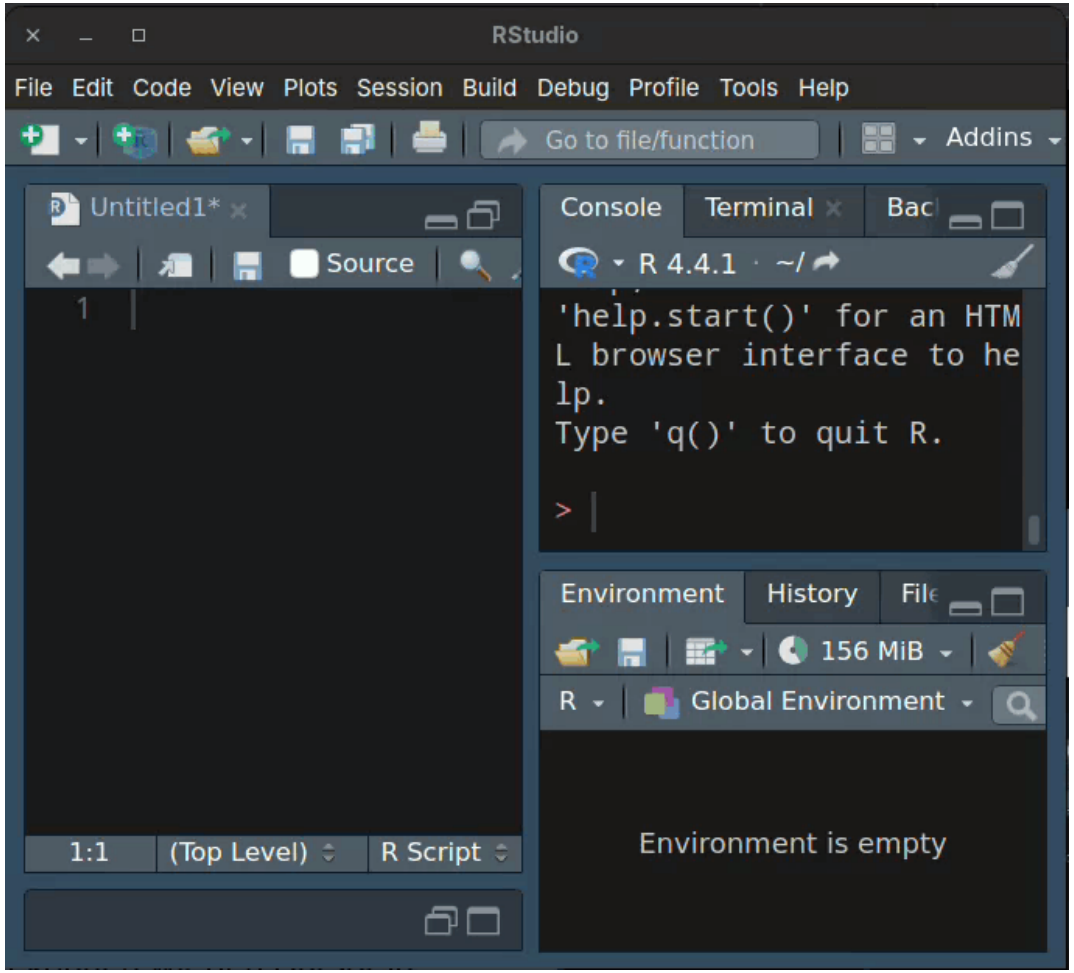
Script vs. Console

Script vs. Console



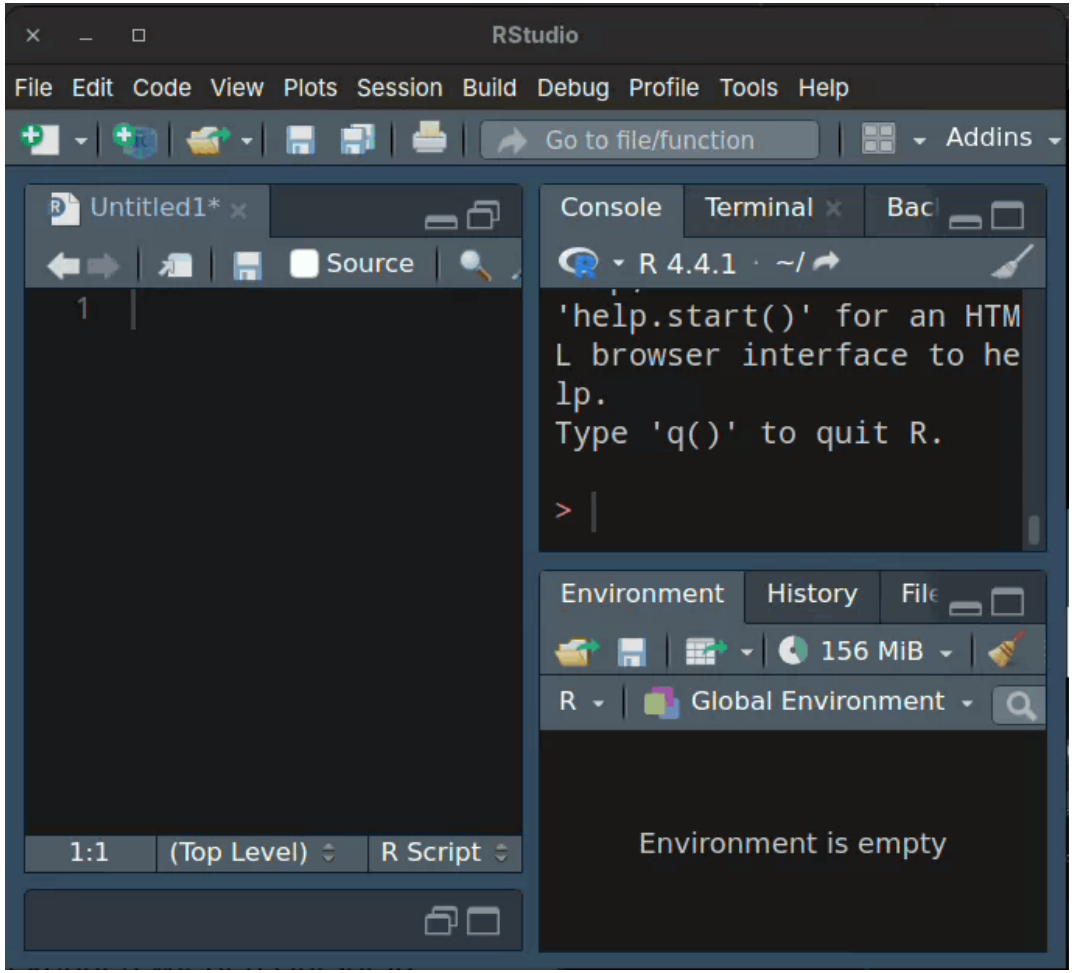
Script vs. Console

- With a **script**, we can save different lines of code and execute them whenever needed.



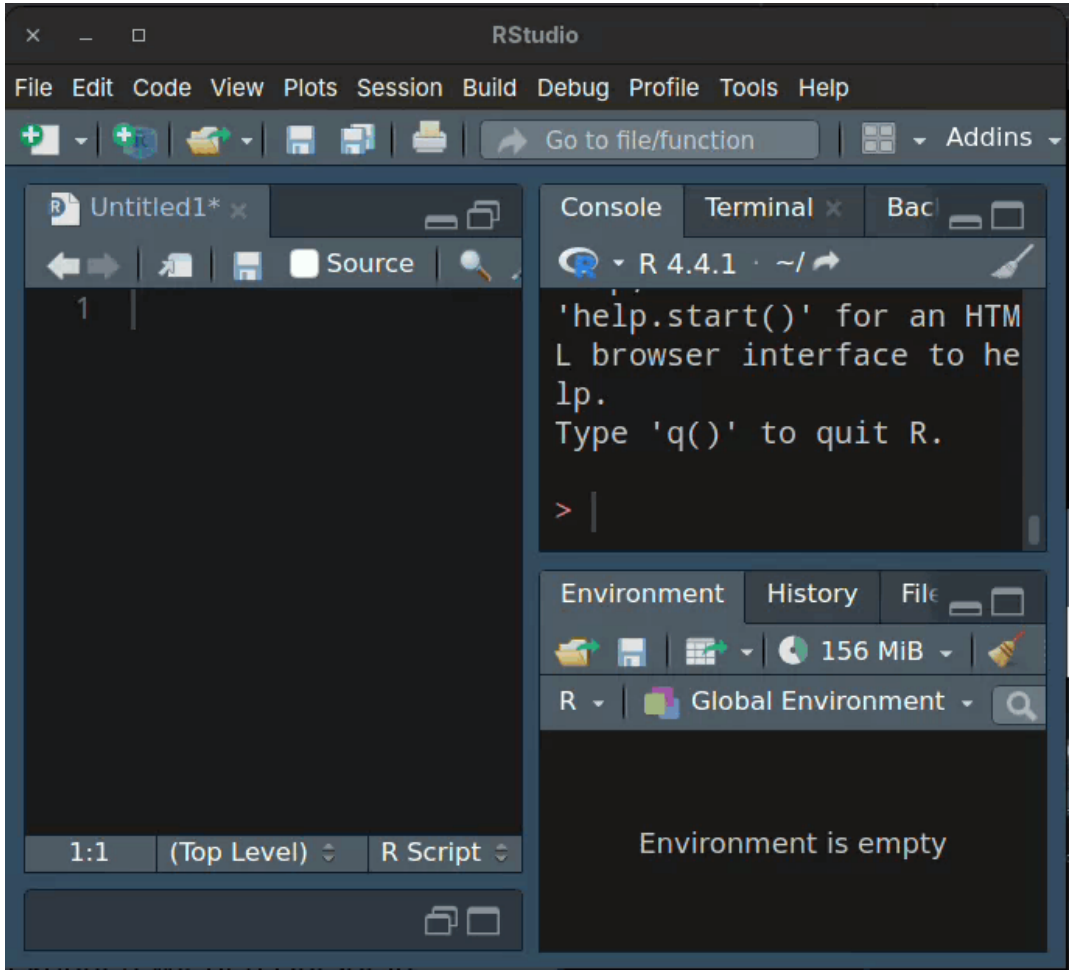
Script vs. Console

- With a **script**, we can save different lines of code and execute them whenever needed.
- We can also execute code directly in the **console**, but we cannot save it there.



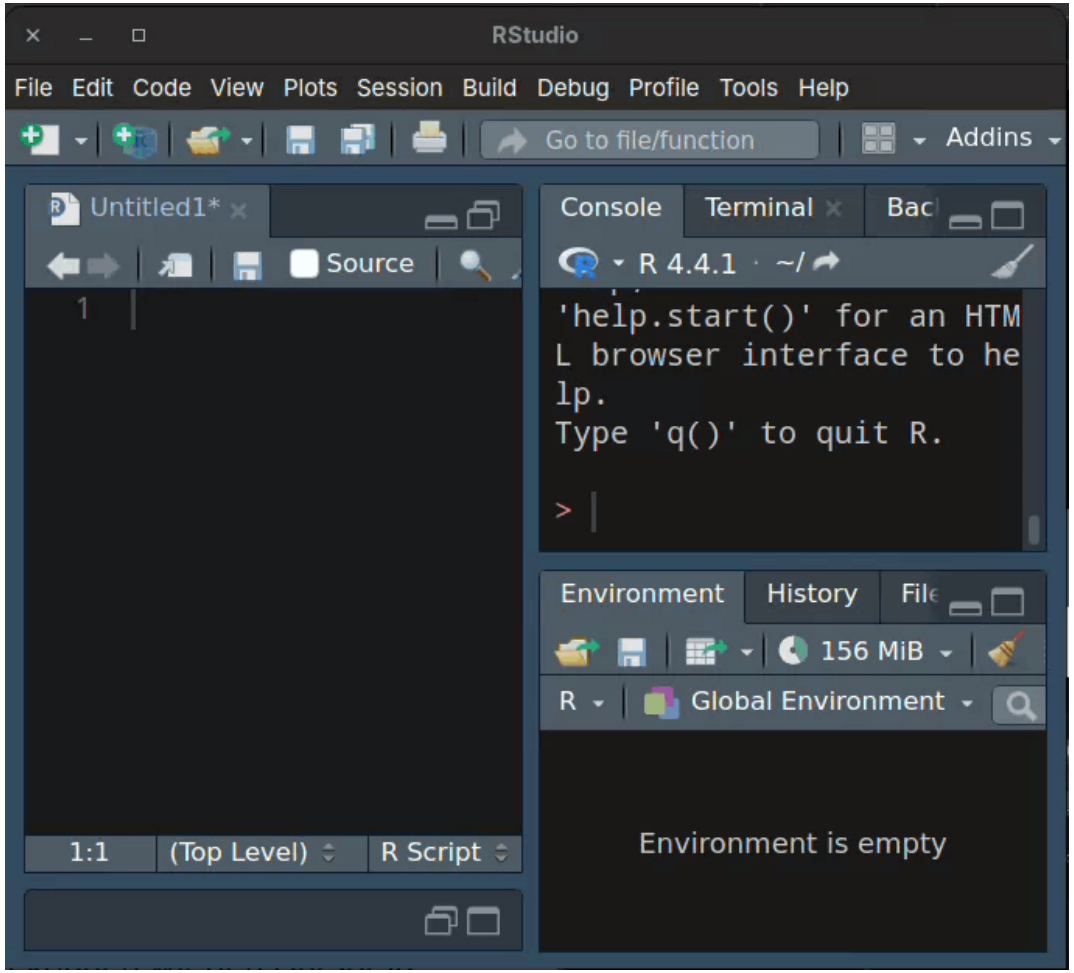
Script vs. Console

- With a **script**, we can save different lines of code and execute them whenever needed.
- We can also execute code directly in the **console**, but we cannot save it there.
- In the script, we can place the cursor on a line and execute it with `Ctrl` + `Enter` (Linux, Windows) or `Cmd` + `Enter` (macOS). It will then be "inserted" into the console.



Script vs. Console

- With a **script**, we can save different lines of code and execute them whenever needed.
- We can also execute code directly in the **console**, but we cannot save it there.
- In the script, we can place the cursor on a line and execute it with `Ctrl` + `Enter` (Linux, Windows) or `Cmd` + `Enter` (macOS). It will then be "inserted" into the console.
- Alternatively, we can highlight multiple lines and execute them, or execute the entire script at once.



Basics in R

Basics in R

Base R comes with many useful **functions**, but sometimes we will need functions for specific econometric purposes that are not included in Base R. These are often included in packages created by developers and then (ideally) published in the Comprehensive R Archive Network (CRAN). We can install these additional packages as follows:

```r
install.packages("tidyverse")

library(tidyverse)

update.packages()
```

Basics in R

Base R comes with many useful **functions**, but sometimes we will need functions for specific econometric purposes that are not included in Base R. These are often included in packages created by developers and then (ideally) published in the Comprehensive R Archive Network (CRAN). We can install these additional packages as follows:

```
install.packages("tidyverse")

library(tidyverse)

update.packages()
```

The above functions **install**, **load**, and **update** packages (Note: `install.packages()` requires quotation marks). We can access a function's documentation with a `?`:

```
?install.packages
```

Basics in R

Base R comes with many useful **functions**, but sometimes we will need functions for specific econometric purposes that are not included in Base R. These are often included in packages created by developers and then (ideally) published in the Comprehensive R Archive Network (CRAN). We can install these additional packages as follows:

```
install.packages("tidyverse")

library(tidyverse)

update.packages()
```

The above functions **install**, **load**, and **update** packages (Note: `install.packages()` requires quotation marks). We can access a function's documentation with a `?`:

```
?install.packages
```

Functions can be recognized by the parentheses. They may or may not contain arguments.

Variables and Vectors

Variables and Vectors

With `<-` (or `=`), we can **assign** a value to a variable name.

This can either be a scalar, a vector, or something else (more on that later).

We can print a variable with `print()`. It is also sufficient to just write the variable name (without `print()`).

The code on the right is **interactive** and can be modified and executed.

R Code | Start Over | Run Code

```r
# Lines starting with # are comments
# and are ignored.

my_scalar <- 12

my_vector <- c(5, 2, 13, 15)

my_character_vector <- c("Vienna", "Linz", "Graz")

# Works without print()
print(my_vector)

print(my_character_vector)

```

Mathematics

Mathematics

We can also use R as a calculator. The interactive code on the right demonstrates various **mathematical operations**.

Mathematics

We can also use R as a calculator. The interactive code on the right demonstrates various **mathematical operations**.

R Code | Start Over | Run Code

```r
1 + 3

# We can also save the result:
my_sum <- 4 + 12

c(5, 2) + c(7, 3)

2^2
sqrt(2) # Square root
log(100) # Natural logarithm
log(100, base = 10)

```

Mathematics

We can also use R as a calculator. The interactive code on the right demonstrates various **mathematical operations**.

We can also define **matrices** and perform calculations with them.

R Code | Start Over | Run Code

```r
my_matrix <- cbind(c(1, 2, 3), c(
2 * my_matrix
```

R Code | Start Over | Run Code

```r
1 + 3

# We can also save the result:
my_sum <- 4 + 12

c(5, 2) + c(7, 3)

2^2
sqrt(2) # Square root
log(100) # Natural logarithm
log(100, base = 10)

```

Descriptive Statistics

Descriptive Statistics

Of course, R is primarily a **statistical programming language**. So let us simulate 100 dice rolls:

R Code | Start Over | Run Code

```r
# We draw 100 times
#from the values 1 to 6
throws <- sample(1:6,
       size = 100,
       replace = TRUE)
```

Execute this code so we can proceed.

Descriptive Statistics

Of course, R is primarily a **statistical programming language**. So let us simulate 100 dice rolls:

R Code | Start Over | Run Code

```r
# We draw 100 times
#from the values 1 to 6
throws <- sample(1:6,
        size = 100,
        replace = TRUE)

```

Execute this code so we can proceed.

What is the **mean** of the rolls?

R Code | Start Over | Run Code

```r
mean(throws)
```

Descriptive Statistics

Of course, R is primarily a **statistical programming language**. So let us simulate 100 dice rolls:

R Code | Start Over | Run Code

```r
# We draw 100 times
#from the values 1 to 6
throws <- sample(1:6,
        size = 100,
        replace = TRUE)

```

Execute this code so we can proceed.

What is the **mean** of the rolls?

R Code | Start Over | Run Code

```r
mean(throws)
```

We can also calculate other measures:

R Code | Start Over | Run Code

```r
summary(throws)
```

Other functions include `median()`, `min()`, `max()`, `length()`, `var()`, `sd()`, `sum()`, ...

Data

Data

We use different functions depending on the file format to read in data: `read.csv()` for CSV, `readRDS()` for RDS, ...

Data

We use different functions depending on the file format to read in data: `read.csv()` for CSV, `readRDS()` for RDS, ...

Some datasets are already available in R, which is especially convenient for practice purposes. One example is `mtcars`. We can view the first rows using `head()`.

R Code | Start Over | Run Code

```
1 head(mtcars)
```

Data

We use different functions depending on the file format to read in data: `read.csv()` for CSV, `readRDS()` for RDS, ...

Some datasets are already available in R, which is especially convenient for practice purposes. One example is `mtcars`. We can view the first rows using `head()`.

R Code | Start Over | Run Code

```
1 head(mtcars)
```

When reading data, e.g., as CSV, we must assign it to a name, e.g., through `my_data <- read.csv("data.csv")`. Incidentally, we can export data in a similar way: `write.csv(my_data, "my_data.csv")`.

Dataframes

Dataframes

The structure in which data is stored is called a **dataframe**. The rows of a dataframe correspond to individual observations, and the columns correspond to variables. With `View()`, we can view the dataset in a separate window. We can also find out, e.g., the number of columns and rows:

R Code | Start Over | Run Code

```
1  ncol(mtcars); nrow(mtcars)
```

Dataframes

The structure in which data is stored is called a **dataframe**. The rows of a dataframe correspond to individual observations, and the columns correspond to variables. With `View()`, we can view the dataset in a separate window. We can also find out, e.g., the number of columns and rows:

R Code | Start Over | Run Code

```
1  ncol(mtcars); nrow(mtcars)
```

Using square brackets, we can access specific rows and columns. `mtcars[1,]` is the first row of `mtcars`, `mtcars[,1]` is the first column. We can also access individual variables using the following notation: `mtcars$mpg`.

Logic

Logic

What happens when we execute this code?

R Code | Start Over | Run Code

```r
# Double == means "is equal to," because single = is used for assignments
# Also: <= for ≤, >= for ≥, != for ≠.
1 == 2
```

Logic

What happens when we execute this code?

R Code | Start Over | Run Code

```r
# Double == means "is equal to," because single = is used for assignments
# Also: <= for ≤, >= for ≥, != for ≠.
1 == 2
```

We can also use `TRUE` and `FALSE` to filter values:

R Code | Start Over | Run Code

```r
test_vector <- c(1, 2, 3)     # Define vector
test_vector > 1               # Check: What is >1?
test_vector[test_vector > 1] # Filter: Only elements >1 (so only TRUE)
```

Further Operations

Further Operations

We can also combine multiple functions:

R Code | Start Over | Run Code

```
1 log(mean(mtcars$mpg))
2 # How must this code be changed to calculate
```

Further Operations

We can also combine multiple functions:

R Code | Start Over | Run Code

```
1 log(mean(mtcars$mpg))
2 # How must this code be changed to calculate
```

Or draw graphs:

R Code | Start Over | Run Code

```
1 hist(mtcars$mpg, breaks = 20)  # Histogram wi
2 # Try creating a boxplot with boxplot()!
```

Scatterplots

Scatterplots

Often, we want to draw a **scatterplot** to show the relationship between two variables.

R Code | Start Over | Run Code

```
1 plot(mtcars$hp, mtcars$mpg)
```

Scatterplots

Often, we want to draw a **scatterplot** to show the relationship between two variables.

Try adding the line `abline(lm(mpg~hp, data=mtcars), col="red")` to draw a red **regression line** on the plot!

R Code | Start Over | Run Code

```
1 plot(mtcars$hp, mtcars$mpg)
```

Scatterplots

Often, we want to draw a **scatterplot** to show the relationship between two variables.

Try adding the line `abline(lm(mpg~hp, data=mtcars), col="red")` to draw a red **regression line** on the plot!

How strongly are `mpg` and `hp` correlated?

R Code | Start Over | Run Code

```
1 cor(mtcars$hp, mtcars$mpg)
```

R Code | Start Over | Run Code

```
1 plot(mtcars$hp, mtcars$mpg)
```

Introduction

Installing R

Introduction to R with RStudio

Matrices and Vectors

Random Variables

Analysis of a Random Variable

Analysis of Two Random Variables

Matrices and Vectors

Matrices and Vectors

We denote matrices with bold uppercase letters ($\boldsymbol{X}$) and (column) vectors with bold lowercase letters ($\boldsymbol{x}$):

$$\boldsymbol{X} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & \cdots & \cdots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} \end{pmatrix}, \quad \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Matrices and Vectors

We denote matrices with bold uppercase letters ($\boldsymbol{X}$) and (column) vectors with bold lowercase letters ($\boldsymbol{x}$):

$$\boldsymbol{X} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & \cdots & \cdots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} \end{pmatrix}, \quad \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Bold is not strictly necessary. If we want to clarify in handwriting that something is a matrix or vector, we can instead underline them: $\underline{X}$ or $\underline{x}$.

Dimensions and Transposition

Dimensions and Transposition

On the previous slide, $\boldsymbol{X}$ had **dimensions** $n \times k$, and $\boldsymbol{x}$ had dimension n. $\boldsymbol{x}$ was written as a column vector, but we can also write $\boldsymbol{x}$ as a row vector. To do this, we must **transpose** the vector: Simply put, rows become columns, and columns become rows. We denote the transposition with a small apostrophe (or optionally a superscript T):

$$\boldsymbol{x}' = (x_1, x_2, \ldots, x_n)$$

Dimensions and Transposition

On the previous slide, $\boldsymbol{X}$ had **dimensions** $n \times k$, and $\boldsymbol{x}$ had dimension n. $\boldsymbol{x}$ was written as a column vector, but we can also write $\boldsymbol{x}$ as a row vector. To do this, we must **transpose** the vector: Simply put, rows become columns, and columns become rows. We denote the transposition with a small apostrophe (or optionally a superscript T):

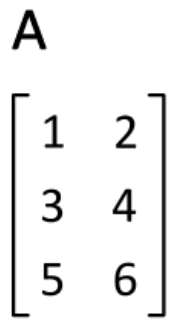
$$\boldsymbol{x}' = (x_1, x_2, \ldots, x_n)$$

We can also **transpose matrices**. A matrix that previously had dimensions $n \times k$ will have dimensions $k \times n$ after transposition.

If we transpose a transposed matrix, we get:

$$(\boldsymbol{X}')' = \boldsymbol{X}$$

Furthermore, $(\boldsymbol{XZ})' = \boldsymbol{Z}'\boldsymbol{X}'$.



Animation from Wikipedia on matrix transposition.

Special Matrices

Special Matrices

In econometrics, we encounter various **special matrices**:

Special Matrices

In econometrics, we encounter various **special matrices**:

- An $n \times n$ matrix is called a **square matrix**.

Special Matrices

In econometrics, we encounter various **special matrices**:

- An $n \times n$ matrix is called a **square matrix**.
- If for this square matrix $\boldsymbol{X}' = \boldsymbol{X}$ holds, the matrix is **symmetric**.

Special Matrices

In econometrics, we encounter various **special matrices**:

- An $n \times n$ matrix is called a **square matrix**.
- If for this square matrix $\boldsymbol{X}' = \boldsymbol{X}$ holds, the matrix is **symmetric**.
- A square ($n \times n$) matrix whose elements outside the main diagonal are all 0 is called a **diagonal matrix** and can be written as $\text{diag}(x_1, x_2, \ldots, x_n)$.

Special Matrices

In econometrics, we encounter various **special matrices**:

- An $n \times n$ matrix is called a **square matrix**.
- If for this square matrix $\boldsymbol{X}' = \boldsymbol{X}$ holds, the matrix is **symmetric**.
- A square ($n \times n$) matrix whose elements outside the main diagonal are all 0 is called a **diagonal matrix** and can be written as $\operatorname{diag}(x_1, x_2, \ldots, x_n)$.
- A diagonal matrix whose main diagonal elements are all 1 is called an **identity matrix** and is denoted by $\boldsymbol{I}$.

Special Matrices

In econometrics, we encounter various **special matrices**:

- An $n \times n$ matrix is called a **square matrix**.
- If for this square matrix $\boldsymbol{X}' = \boldsymbol{X}$ holds, the matrix is **symmetric**.
- A square ($n \times n$) matrix whose elements outside the main diagonal are all 0 is called a **diagonal matrix** and can be written as $\operatorname{diag}(x_1, x_2, \ldots, x_n)$.
- A diagonal matrix whose main diagonal elements are all 1 is called an **identity matrix** and is denoted by $\boldsymbol{I}$.
- An $n \times k$ matrix whose all elements are 0 is called a **zero matrix** and is denoted by $\mathbf{0}_{n \times k}$.

Rank

Rank

The **rank** of a matrix is defined as the *dimension of the vector space spanned by the columns of a matrix* and is denoted by $\text{rank}(\boldsymbol{X})$. Simply put, the rank of a matrix corresponds to the **number of its linearly independent columns**. A column is linearly independent of the others if it cannot be expressed as a linear combination of them (i.e., as a sum of multiples of the other columns). Consider the following matrix:

Rank

The **rank** of a matrix is defined as the *dimension of the vector space spanned by the columns of a matrix* and is denoted by $\text{rank}(\boldsymbol{X})$. Simply put, the rank of a matrix corresponds to the **number of its linearly independent columns**. A column is linearly independent of the others if it cannot be expressed as a linear combination of them (i.e., as a sum of multiples of the other columns). Consider the following matrix:

$$\boldsymbol{X} = \begin{pmatrix} 12 & 2 & 10 \\ 3 & 1 & 2 \\ 7 & 4 & 3 \\ 8 & 6 & 2 \end{pmatrix}$$

This matrix has rank 2. It has three columns, but the third column is a linear combination of the first two: $x_{i3} = x_{i1} + (-1) \cdot x_{i2}$.

If a matrix has the maximum possible rank for its dimensions, it is said to have **full rank**.

Matrix Addition and Scalar Multiplication

Matrix Addition and Scalar Multiplication

Matrix addition occurs element by element:

$$\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} \end{pmatrix} + \begin{pmatrix} z_{11} & z_{12} & \cdots & z_{1k} \\ z_{21} & z_{22} & \cdots & z_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ z_{n1} & z_{n2} & \cdots & z_{nk} \end{pmatrix} = \begin{pmatrix} x_{11} + z_{11} & x_{12} + z_{12} & \cdots & x_{1k} + z_{1k} \\ x_{21} + z_{21} & x_{22} + z_{22} & \cdots & x_{2k} + z_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} + z_{n1} & x_{n2} + z_{n2} & \cdots & x_{nk} + z_{nk} \end{pmatrix}$$

Matrix Addition and Scalar Multiplication

Matrix addition occurs element by element:

$$\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} \end{pmatrix} + \begin{pmatrix} z_{11} & z_{12} & \cdots & z_{1k} \\ z_{21} & z_{22} & \cdots & z_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ z_{n1} & z_{n2} & \cdots & z_{nk} \end{pmatrix} = \begin{pmatrix} x_{11} + z_{11} & x_{12} + z_{12} & \cdots & x_{1k} + z_{1k} \\ x_{21} + z_{21} & x_{22} + z_{22} & \cdots & x_{2k} + z_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} + z_{n1} & x_{n2} + z_{n2} & \cdots & x_{nk} + z_{nk} \end{pmatrix}$$

Multiplication of a matrix by a scalar also occurs element by element:

$$\alpha \cdot \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} \end{pmatrix} = \begin{pmatrix} \alpha x_{11} & \alpha x_{12} & \cdots & \alpha x_{1k} \\ \alpha x_{21} & \alpha x_{22} & \cdots & \alpha x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha x_{n1} & \alpha x_{n2} & \cdots & \alpha x_{nk} \end{pmatrix}$$

Multiplication of Two Matrices

Multiplication of Two Matrices

The **multiplication of two matrices** is somewhat more complex. Let $\boldsymbol{X}$ be a 2×3 matrix and $\boldsymbol{Z}$ be a 3×2 matrix. Then we can multiply the matrices as follows:

$$\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \end{pmatrix} \cdot \begin{pmatrix} z_{11} & z_{12} & z_{13} \\ z_{21} & z_{22} & z_{23} \end{pmatrix} = \begin{pmatrix} x_{11}z_{11} + x_{12}z_{21} & x_{11}z_{12} + x_{12}z_{22} & x_{11}z_{13} + x_{12}z_{23} \\ x_{21}z_{11} + x_{22}z_{21} & x_{21}z_{12} + x_{22}z_{22} & x_{21}z_{13} + x_{22}z_{23} \\ x_{31}z_{11} + x_{32}z_{21} & x_{31}z_{12} + x_{32}z_{22} & x_{31}z_{13} + x_{32}z_{23} \end{pmatrix}$$

Multiplication of Two Matrices

The **multiplication of two matrices** is somewhat more complex. Let $\boldsymbol{X}$ be a 2×3 matrix and $\boldsymbol{Z}$ be a 3×2 matrix. Then we can multiply the matrices as follows:

$$\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \end{pmatrix} \cdot \begin{pmatrix} z_{11} & z_{12} & z_{13} \\ z_{21} & z_{22} & z_{23} \end{pmatrix} = \begin{pmatrix} x_{11}z_{11} + x_{12}z_{21} & x_{11}z_{12} + x_{12}z_{22} & x_{11}z_{13} + x_{12}z_{23} \\ x_{21}z_{11} + x_{22}z_{21} & x_{21}z_{12} + x_{22}z_{22} & x_{21}z_{13} + x_{22}z_{23} \\ x_{31}z_{11} + x_{32}z_{21} & x_{31}z_{12} + x_{32}z_{22} & x_{31}z_{13} + x_{32}z_{23} \end{pmatrix}$$

The following table helps visualize the process:

	z_{11}, z_{21}	z_{12}, z_{22}	z_{13}, z_{23}
x_{11}, x_{12}	$x_{11}z_{11} + x_{12}z_{21}$	$x_{11}z_{12} + x_{12}z_{22}$	$x_{11}z_{13} + x_{12}z_{23}$
x_{21}, x_{22}	$x_{21}z_{11} + x_{22}z_{21}$	$x_{21}z_{12} + x_{22}z_{22}$	$x_{21}z_{13} + x_{22}z_{23}$
x_{31}, x_{32}	$x_{31}z_{11} + x_{32}z_{21}$	$x_{31}z_{12} + x_{32}z_{22}$	$x_{31}z_{13} + x_{32}z_{23}$

Multiplication of Two Matrices

The **multiplication of two matrices** is somewhat more complex. Let $\boldsymbol{X}$ be a 2×3 matrix and $\boldsymbol{Z}$ be a 3×2 matrix. Then we can multiply the matrices as follows:

$$\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \end{pmatrix} \cdot \begin{pmatrix} z_{11} & z_{12} & z_{13} \\ z_{21} & z_{22} & z_{23} \end{pmatrix} = \begin{pmatrix} x_{11}z_{11} + x_{12}z_{21} & x_{11}z_{12} + x_{12}z_{22} & x_{11}z_{13} + x_{12}z_{23} \\ x_{21}z_{11} + x_{22}z_{21} & x_{21}z_{12} + x_{22}z_{22} & x_{21}z_{13} + x_{22}z_{23} \\ x_{31}z_{11} + x_{32}z_{21} & x_{31}z_{12} + x_{32}z_{22} & x_{31}z_{13} + x_{32}z_{23} \end{pmatrix}$$

The following table helps visualize the process:

	z_{11}, z_{21}	z_{12}, z_{22}	z_{13}, z_{23}
x_{11}, x_{12}	$x_{11}z_{11} + x_{12}z_{21}$	$x_{11}z_{12} + x_{12}z_{22}$	$x_{11}z_{13} + x_{12}z_{23}$
x_{21}, x_{22}	$x_{21}z_{11} + x_{22}z_{21}$	$x_{21}z_{12} + x_{22}z_{22}$	$x_{21}z_{13} + x_{22}z_{23}$
x_{31}, x_{32}	$x_{31}z_{11} + x_{32}z_{21}$	$x_{31}z_{12} + x_{32}z_{22}$	$x_{31}z_{13} + x_{32}z_{23}$

It is easy to see that matrices can **only** be multiplied **if the number of columns in the left matrix** equals the number of **rows in the right matrix**.

Properties of Operations

Properties of Operations

- Both matrix addition and scalar multiplication are commutative: $\boldsymbol{X} + \boldsymbol{Z} = \boldsymbol{Z} + \boldsymbol{X}$ and $\alpha \boldsymbol{X} = \boldsymbol{X} \alpha$.

Properties of Operations

- Both matrix addition and scalar multiplication are commutative: $\boldsymbol{X} + \boldsymbol{Z} = \boldsymbol{Z} + \boldsymbol{X}$ and $\alpha \boldsymbol{X} = \boldsymbol{X} \alpha$.
- Unlike scalar multiplication, matrix multiplication is **not commutative**, meaning $\boldsymbol{X}\boldsymbol{Z}$ is generally not the same as $\boldsymbol{Z}\boldsymbol{X}$.

Properties of Operations

- Both matrix addition and scalar multiplication are commutative: $\boldsymbol{X} + \boldsymbol{Z} = \boldsymbol{Z} + \boldsymbol{X}$ and $\alpha \boldsymbol{X} = \boldsymbol{X} \alpha$.
- Unlike scalar multiplication, matrix multiplication is **not commutative**, meaning $\boldsymbol{X}\boldsymbol{Z}$ is generally not the same as $\boldsymbol{Z}\boldsymbol{X}$.
- Due to the condition for the dimensions of the two matrices, it may even be the case that $\boldsymbol{X}\boldsymbol{Z}$ exists, but $\boldsymbol{Z}\boldsymbol{X}$ does not.

Properties of Operations

- Both matrix addition and scalar multiplication are commutative: $\boldsymbol{X} + \boldsymbol{Z} = \boldsymbol{Z} + \boldsymbol{X}$ and $\alpha\boldsymbol{X} = \boldsymbol{X}\alpha$.
- Unlike scalar multiplication, matrix multiplication is **not commutative**, meaning $\boldsymbol{X}\boldsymbol{Z}$ is generally not the same as $\boldsymbol{Z}\boldsymbol{X}$.
- Due to the condition for the dimensions of the two matrices, it may even be the case that $\boldsymbol{X}\boldsymbol{Z}$ exists, but $\boldsymbol{Z}\boldsymbol{X}$ does not.
- A matrix $\boldsymbol{X}$ is called **idempotent** if $\boldsymbol{X}\boldsymbol{X} = \boldsymbol{X}$.

Properties of Operations

- Both matrix addition and scalar multiplication are commutative: $\boldsymbol{X} + \boldsymbol{Z} = \boldsymbol{Z} + \boldsymbol{X}$ and $\alpha\boldsymbol{X} = \boldsymbol{X}\alpha$.
- Unlike scalar multiplication, matrix multiplication is **not commutative**, meaning $\boldsymbol{XZ}$ is generally not the same as $\boldsymbol{ZX}$.
- Due to the condition for the dimensions of the two matrices, it may even be the case that $\boldsymbol{XZ}$ exists, but $\boldsymbol{ZX}$ does not.
- A matrix $\boldsymbol{X}$ is called **idempotent** if $\boldsymbol{XX} = \boldsymbol{X}$.
- Additionally, for matrices with the appropriate dimensions:

Properties of Operations

- Both matrix addition and scalar multiplication are commutative: $\boldsymbol{X} + \boldsymbol{Z} = \boldsymbol{Z} + \boldsymbol{X}$ and $\alpha \boldsymbol{X} = \boldsymbol{X} \alpha$.
- Unlike scalar multiplication, matrix multiplication is **not commutative**, meaning $\boldsymbol{X}\boldsymbol{Z}$ is generally not the same as $\boldsymbol{Z}\boldsymbol{X}$.
- Due to the condition for the dimensions of the two matrices, it may even be the case that $\boldsymbol{X}\boldsymbol{Z}$ exists, but $\boldsymbol{Z}\boldsymbol{X}$ does not.
- A matrix $\boldsymbol{X}$ is called **idempotent** if $\boldsymbol{X}\boldsymbol{X} = \boldsymbol{X}$.
- Additionally, for matrices with the appropriate dimensions:
  - $(\boldsymbol{X} + \boldsymbol{Z}) + \boldsymbol{A} = \boldsymbol{X} + (\boldsymbol{Z} + \boldsymbol{A})$

Properties of Operations

- Both matrix addition and scalar multiplication are commutative: $\boldsymbol{X} + \boldsymbol{Z} = \boldsymbol{Z} + \boldsymbol{X}$ and $\alpha \boldsymbol{X} = \boldsymbol{X} \alpha$.
- Unlike scalar multiplication, matrix multiplication is **not commutative**, meaning $\boldsymbol{X}\boldsymbol{Z}$ is generally not the same as $\boldsymbol{Z}\boldsymbol{X}$.
- Due to the condition for the dimensions of the two matrices, it may even be the case that $\boldsymbol{X}\boldsymbol{Z}$ exists, but $\boldsymbol{Z}\boldsymbol{X}$ does not.
- A matrix $\boldsymbol{X}$ is called **idempotent** if $\boldsymbol{X}\boldsymbol{X} = \boldsymbol{X}$.
- Additionally, for matrices with the appropriate dimensions:
  - $(\boldsymbol{X} + \boldsymbol{Z}) + \boldsymbol{A} = \boldsymbol{X} + (\boldsymbol{Z} + \boldsymbol{A})$
  - $\boldsymbol{X} + \mathbf{0} = \boldsymbol{X}$

Properties of Operations

- Both matrix addition and scalar multiplication are commutative: $\boldsymbol{X} + \boldsymbol{Z} = \boldsymbol{Z} + \boldsymbol{X}$ and $\alpha \boldsymbol{X} = \boldsymbol{X} \alpha$.
- Unlike scalar multiplication, matrix multiplication is **not commutative**, meaning $\boldsymbol{X}\boldsymbol{Z}$ is generally not the same as $\boldsymbol{Z}\boldsymbol{X}$.
- Due to the condition for the dimensions of the two matrices, it may even be the case that $\boldsymbol{X}\boldsymbol{Z}$ exists, but $\boldsymbol{Z}\boldsymbol{X}$ does not.
- A matrix $\boldsymbol{X}$ is called **idempotent** if $\boldsymbol{X}\boldsymbol{X} = \boldsymbol{X}$.
- Additionally, for matrices with the appropriate dimensions:
  - $(\boldsymbol{X} + \boldsymbol{Z}) + \boldsymbol{A} = \boldsymbol{X} + (\boldsymbol{Z} + \boldsymbol{A})$
  - $\boldsymbol{X} + \mathbf{0} = \boldsymbol{X}$
  - $(\boldsymbol{X}\boldsymbol{Z})\boldsymbol{A} = \boldsymbol{X}(\boldsymbol{Z}\boldsymbol{A})$

Properties of Operations

- Both matrix addition and scalar multiplication are commutative: $\boldsymbol{X} + \boldsymbol{Z} = \boldsymbol{Z} + \boldsymbol{X}$ and $\alpha \boldsymbol{X} = \boldsymbol{X} \alpha$.
- Unlike scalar multiplication, matrix multiplication is **not commutative**, meaning $\boldsymbol{X}\boldsymbol{Z}$ is generally not the same as $\boldsymbol{Z}\boldsymbol{X}$.
- Due to the condition for the dimensions of the two matrices, it may even be the case that $\boldsymbol{X}\boldsymbol{Z}$ exists, but $\boldsymbol{Z}\boldsymbol{X}$ does not.
- A matrix $\boldsymbol{X}$ is called **idempotent** if $\boldsymbol{X}\boldsymbol{X} = \boldsymbol{X}$.
- Additionally, for matrices with the appropriate dimensions:
  - $(\boldsymbol{X} + \boldsymbol{Z}) + \boldsymbol{A} = \boldsymbol{X} + (\boldsymbol{Z} + \boldsymbol{A})$
  - $\boldsymbol{X} + \boldsymbol{0} = \boldsymbol{X}$
  - $(\boldsymbol{X}\boldsymbol{Z})\boldsymbol{A} = \boldsymbol{X}(\boldsymbol{Z}\boldsymbol{A})$
  - $\boldsymbol{A}(\boldsymbol{X} + \boldsymbol{Z}) = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{A}\boldsymbol{Z}$

Properties of Operations

- Both matrix addition and scalar multiplication are commutative: $\boldsymbol{X} + \boldsymbol{Z} = \boldsymbol{Z} + \boldsymbol{X}$ and $\alpha \boldsymbol{X} = \boldsymbol{X} \alpha$.
- Unlike scalar multiplication, matrix multiplication is **not commutative**, meaning $\boldsymbol{X}\boldsymbol{Z}$ is generally not the same as $\boldsymbol{Z}\boldsymbol{X}$.
- Due to the condition for the dimensions of the two matrices, it may even be the case that $\boldsymbol{X}\boldsymbol{Z}$ exists, but $\boldsymbol{Z}\boldsymbol{X}$ does not.
- A matrix $\boldsymbol{X}$ is called **idempotent** if $\boldsymbol{X}\boldsymbol{X} = \boldsymbol{X}$.
- Additionally, for matrices with the appropriate dimensions:
  - $(\boldsymbol{X} + \boldsymbol{Z}) + \boldsymbol{A} = \boldsymbol{X} + (\boldsymbol{Z} + \boldsymbol{A})$
  - $\boldsymbol{X} + \mathbf{0} = \boldsymbol{X}$
  - $(\boldsymbol{X}\boldsymbol{Z})\boldsymbol{A} = \boldsymbol{X}(\boldsymbol{Z}\boldsymbol{A})$
  - $\boldsymbol{A}(\boldsymbol{X} + \boldsymbol{Z}) = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{A}\boldsymbol{Z}$
  - $(\boldsymbol{X} + \boldsymbol{Z})\boldsymbol{A} = \boldsymbol{X}\boldsymbol{A} + \boldsymbol{Z}\boldsymbol{A}$

Properties of Operations

- Both matrix addition and scalar multiplication are commutative: $\boldsymbol{X} + \boldsymbol{Z} = \boldsymbol{Z} + \boldsymbol{X}$ and $\alpha \boldsymbol{X} = \boldsymbol{X} \alpha$.
- Unlike scalar multiplication, matrix multiplication is **not commutative**, meaning $\boldsymbol{X}\boldsymbol{Z}$ is generally not the same as $\boldsymbol{Z}\boldsymbol{X}$.
- Due to the condition for the dimensions of the two matrices, it may even be the case that $\boldsymbol{X}\boldsymbol{Z}$ exists, but $\boldsymbol{Z}\boldsymbol{X}$ does not.
- A matrix $\boldsymbol{X}$ is called **idempotent** if $\boldsymbol{X}\boldsymbol{X} = \boldsymbol{X}$.
- Additionally, for matrices with the appropriate dimensions:
  - $(\boldsymbol{X} + \boldsymbol{Z}) + \boldsymbol{A} = \boldsymbol{X} + (\boldsymbol{Z} + \boldsymbol{A})$
  - $\boldsymbol{X} + \boldsymbol{0} = \boldsymbol{X}$
  - $(\boldsymbol{X}\boldsymbol{Z})\boldsymbol{A} = \boldsymbol{X}(\boldsymbol{Z}\boldsymbol{A})$
  - $\boldsymbol{A}(\boldsymbol{X} + \boldsymbol{Z}) = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{A}\boldsymbol{Z}$
  - $(\boldsymbol{X} + \boldsymbol{Z})\boldsymbol{A} = \boldsymbol{X}\boldsymbol{A} + \boldsymbol{Z}\boldsymbol{A}$
  - $\boldsymbol{I}\boldsymbol{X} = \boldsymbol{X}\boldsymbol{I} = \boldsymbol{X}$

Properties of Operations

- Both matrix addition and scalar multiplication are commutative: $\boldsymbol{X} + \boldsymbol{Z} = \boldsymbol{Z} + \boldsymbol{X}$ and $\alpha \boldsymbol{X} = \boldsymbol{X} \alpha$.
- Unlike scalar multiplication, matrix multiplication is **not commutative**, meaning $\boldsymbol{X}\boldsymbol{Z}$ is generally not the same as $\boldsymbol{Z}\boldsymbol{X}$.
- Due to the condition for the dimensions of the two matrices, it may even be the case that $\boldsymbol{X}\boldsymbol{Z}$ exists, but $\boldsymbol{Z}\boldsymbol{X}$ does not.
- A matrix $\boldsymbol{X}$ is called **idempotent** if $\boldsymbol{X}\boldsymbol{X} = \boldsymbol{X}$.
- Additionally, for matrices with the appropriate dimensions:
  - $(\boldsymbol{X} + \boldsymbol{Z}) + \boldsymbol{A} = \boldsymbol{X} + (\boldsymbol{Z} + \boldsymbol{A})$
  - $\boldsymbol{X} + \mathbf{0} = \boldsymbol{X}$
  - $(\boldsymbol{X}\boldsymbol{Z})\boldsymbol{A} = \boldsymbol{X}(\boldsymbol{Z}\boldsymbol{A})$
  - $\boldsymbol{A}(\boldsymbol{X} + \boldsymbol{Z}) = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{A}\boldsymbol{Z}$
  - $(\boldsymbol{X} + \boldsymbol{Z})\boldsymbol{A} = \boldsymbol{X}\boldsymbol{A} + \boldsymbol{Z}\boldsymbol{A}$
  - $\boldsymbol{I}\boldsymbol{X} = \boldsymbol{X}\boldsymbol{I} = \boldsymbol{X}$
  - $(\alpha \boldsymbol{X})\boldsymbol{Z} = \boldsymbol{X}(\alpha \boldsymbol{Z}) = \alpha(\boldsymbol{X}\boldsymbol{Z})$

Inverse

Inverse

A **square** matrix $\boldsymbol{X}$ is called **invertible** if there exists a matrix $\boldsymbol{X}^{-1}$ such that:

$$\boldsymbol{X}\boldsymbol{X}^{-1} = \boldsymbol{X}^{-1}\boldsymbol{X} = \boldsymbol{I}.$$

In this case, $\boldsymbol{X}^{-1}$ is called the **inverse** of $\boldsymbol{X}$. If no such matrix exists, $\boldsymbol{X}$ is called **singular** or **non-invertible**. If an inverse exists, it is **unique**.

Inverse

A **square** matrix $\boldsymbol{X}$ is called **invertible** if there exists a matrix $\boldsymbol{X}^{-1}$ such that:

$$\boldsymbol{X}\boldsymbol{X}^{-1} = \boldsymbol{X}^{-1}\boldsymbol{X} = \boldsymbol{I}.$$

In this case, $\boldsymbol{X}^{-1}$ is called the **inverse** of $\boldsymbol{X}$. If no such matrix exists, $\boldsymbol{X}$ is called **singular** or **non-invertible**. If an inverse exists, it is **unique**.

- A matrix $\boldsymbol{X}$ is **invertible** if and only if it has **full rank**.

Inverse

A **square** matrix $\boldsymbol{X}$ is called **invertible** if there exists a matrix $\boldsymbol{X}^{-1}$ such that:

$$\boldsymbol{X}\boldsymbol{X}^{-1} = \boldsymbol{X}^{-1}\boldsymbol{X} = \boldsymbol{I}.$$

In this case, $\boldsymbol{X}^{-1}$ is called the **inverse** of $\boldsymbol{X}$. If no such matrix exists, $\boldsymbol{X}$ is called **singular** or **non-invertible**. If an inverse exists, it is **unique**.

- A matrix $\boldsymbol{X}$ is **invertible** if and only if it has **full rank**.
- For invertible matrices, the following holds:

Inverse

A **square** matrix $\boldsymbol{X}$ is called **invertible** if there exists a matrix $\boldsymbol{X}^{-1}$ such that:

$$\boldsymbol{X}\boldsymbol{X}^{-1} = \boldsymbol{X}^{-1}\boldsymbol{X} = \boldsymbol{I}.$$

In this case, $\boldsymbol{X}^{-1}$ is called the **inverse** of $\boldsymbol{X}$. If no such matrix exists, $\boldsymbol{X}$ is called **singular** or **non-invertible**. If an inverse exists, it is **unique**.

- A matrix $\boldsymbol{X}$ is **invertible** if and only if it has **full rank**.
- For invertible matrices, the following holds:
  - $(\boldsymbol{X}^{-1})^{-1} = \boldsymbol{X}$

Inverse

A **square** matrix $\boldsymbol{X}$ is called **invertible** if there exists a matrix $\boldsymbol{X}^{-1}$ such that:

$$\boldsymbol{X}\boldsymbol{X}^{-1} = \boldsymbol{X}^{-1}\boldsymbol{X} = \boldsymbol{I}.$$

In this case, $\boldsymbol{X}^{-1}$ is called the **inverse** of $\boldsymbol{X}$. If no such matrix exists, $\boldsymbol{X}$ is called **singular** or **non-invertible**. If an inverse exists, it is **unique**.

- A matrix $\boldsymbol{X}$ is **invertible** if and only if it has **full rank**.
- For invertible matrices, the following holds:
  - $(\boldsymbol{X}^{-1})^{-1} = \boldsymbol{X}$
  - $(\boldsymbol{X}')^{-1} = (\boldsymbol{X}^{-1})'$

Inverse

A **square** matrix $\boldsymbol{X}$ is called **invertible** if there exists a matrix $\boldsymbol{X}^{-1}$ such that:

$$\boldsymbol{X}\boldsymbol{X}^{-1} = \boldsymbol{X}^{-1}\boldsymbol{X} = \boldsymbol{I}.$$

In this case, $\boldsymbol{X}^{-1}$ is called the **inverse** of $\boldsymbol{X}$. If no such matrix exists, $\boldsymbol{X}$ is called **singular** or **non-invertible**. If an inverse exists, it is **unique**.

- A matrix $\boldsymbol{X}$ is **invertible** if and only if it has **full rank**.
- For invertible matrices, the following holds:
  - $(\boldsymbol{X}^{-1})^{-1} = \boldsymbol{X}$
  - $(\boldsymbol{X}')^{-1} = (\boldsymbol{X}^{-1})'$
  - $(\boldsymbol{X}\boldsymbol{Z})^{-1} = \boldsymbol{Z}^{-1}\boldsymbol{X}^{-1}$

Installing R

Introduction to R with RStudio

Matrices and Vectors

Random Variables

Analysis of a Random Variable

Analysis of Two Random Variables

Random Variables

Random Variables

Suppose we observe a random event, such as a coin toss or the roll of a die. A **random variable** is a variable that takes on a value depending on the observed event. We denote it with a capital letter:

$$X$$

Random Variables

Suppose we observe a random event, such as a coin toss or the roll of a die. A **random variable** is a variable that takes on a value depending on the observed event. We denote it with a capital letter:

$$X$$

We denote all possible outcomes with the corresponding lowercase letter:

$$x_i$$

Discrete Random Variables

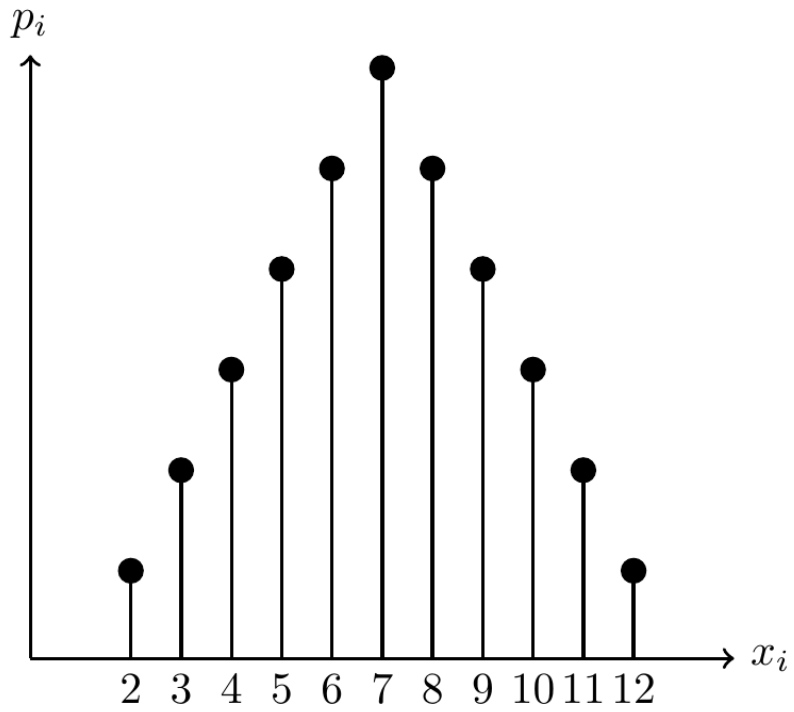
Discrete Random Variables

A **discrete random variable** is a random variable that can have only a finite or countably infinite number of possible outcomes. If the variable is called X, we denote the outcomes as x_i and the corresponding probabilities as p_i. Note that the sum of all probabilities $\sum_i p_i$ must equal 1.

Discrete Random Variables

A **discrete random variable** is a random variable that can have only a finite or countably infinite number of possible outcomes. If the variable is called X, we denote the outcomes as x_i and the corresponding probabilities as p_i. Note that the sum of all probabilities $\sum_i p_i$ must equal 1.

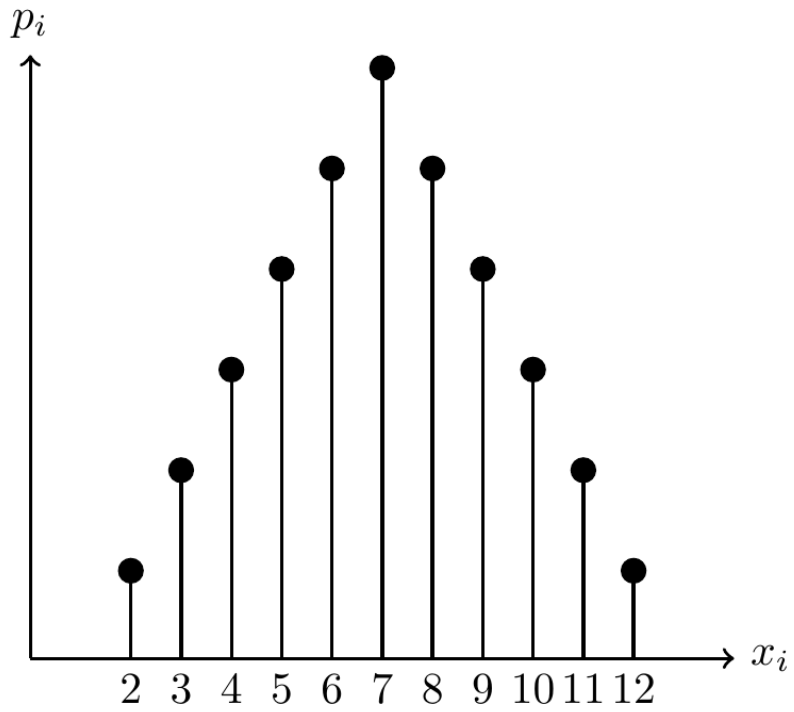
An example of a discrete random variable would be the roll of two dice. The possible outcomes are $\{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$, and the corresponding probabilities are $\{\frac{1}{36}, \frac{2}{36}, \frac{3}{36}, \frac{4}{36}, \frac{5}{36}, \frac{6}{36}, \frac{5}{36}, \frac{4}{36}, \frac{3}{36}, \frac{2}{36}, \frac{1}{36}\}$. To the right, the **probability mass function** (PMF) is illustrated:



Discrete Random Variables

A **discrete random variable** is a random variable that can have only a finite or countably infinite number of possible outcomes. If the variable is called X, we denote the outcomes as x_i and the corresponding probabilities as p_i. Note that the sum of all probabilities $\sum_i p_i$ must equal 1.

An example of a discrete random variable would be the roll of two dice. The possible outcomes are $\{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$, and the corresponding probabilities are $\{\frac{1}{36}, \frac{2}{36}, \frac{3}{36}, \frac{4}{36}, \frac{5}{36}, \frac{6}{36}, \frac{5}{36}, \frac{4}{36}, \frac{3}{36}, \frac{2}{36}, \frac{1}{36}\}$. To the right, the **probability mass function** (PMF) is illustrated:



A **Bernoulli variable** is a discrete random variable that can only take on two outcomes, such as a coin toss.

Continuous Random Variables (1)

Continuous Random Variables (1)

A **continuous random variable** is a random variable that can take on an **uncountably infinite** number of different outcomes.

Continuous Random Variables (1)

A **continuous random variable** is a random variable that can take on an **uncountably infinite** number of different outcomes.

We know that there are infinitely many outcomes and that the sum of all of these equals 1. It follows that the probability of any single outcome is zero. Therefore, there is no probability mass function as with discrete random variables.

Continuous Random Variables (1)

A **continuous random variable** is a random variable that can take on an **uncountably infinite** number of different outcomes.

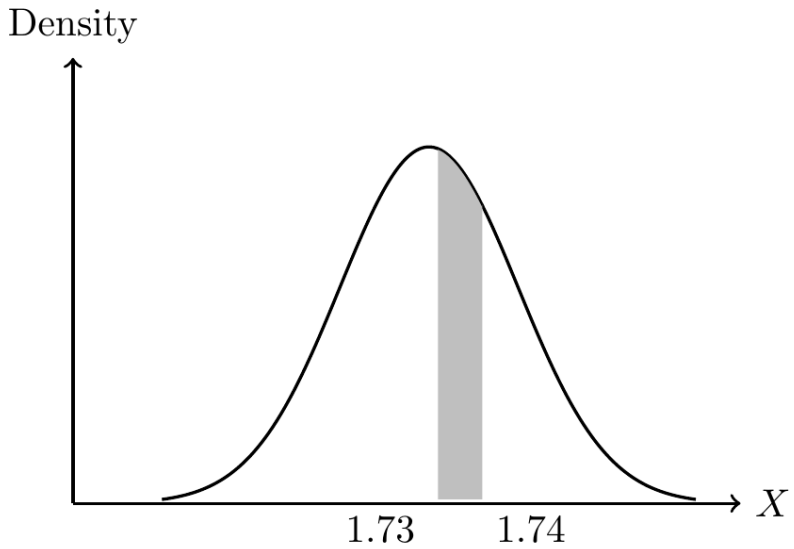
We know that there are infinitely many outcomes and that the sum of all of these equals 1. It follows that the probability of any single outcome is zero. Therefore, there is no probability mass function as with discrete random variables.

What we can do, however, is draw a **probability density function** (PDF). It tells us the probability that the outcome falls within a certain interval. The total area under the PDF is equal to 1.

Continuous Random Variables (2)

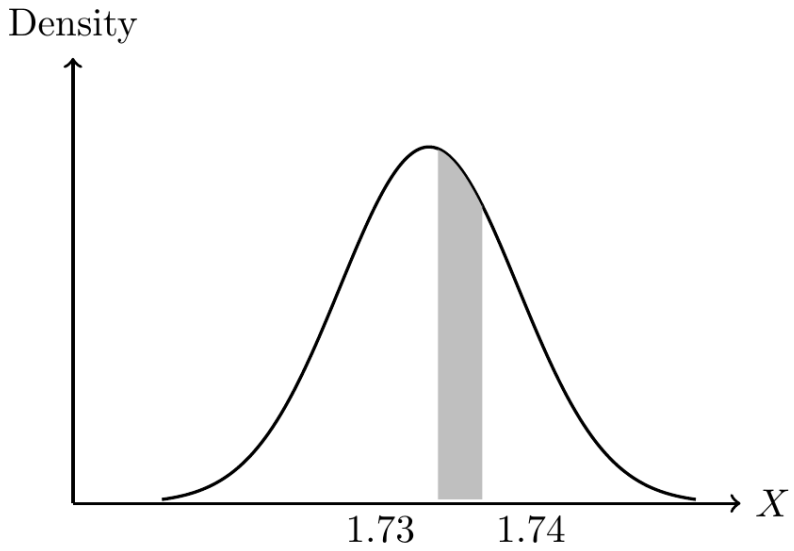
Continuous Random Variables (2)

An example of such a variable would be the **height of a person**. It would make no sense to ask for the probability that a person is exactly 1.734681092536 meters tall. This probability is zero. But we can look at the PDF and determine how likely it is that the person's height lies between 1.73 and 1.74 meters:



Continuous Random Variables (2)

An example of such a variable would be the **height of a person**. It would make no sense to ask for the probability that a person is exactly 1.734681092536 meters tall. This probability is zero. But we can look at the PDF and determine how likely it is that the person's height lies between 1.73 and 1.74 meters:

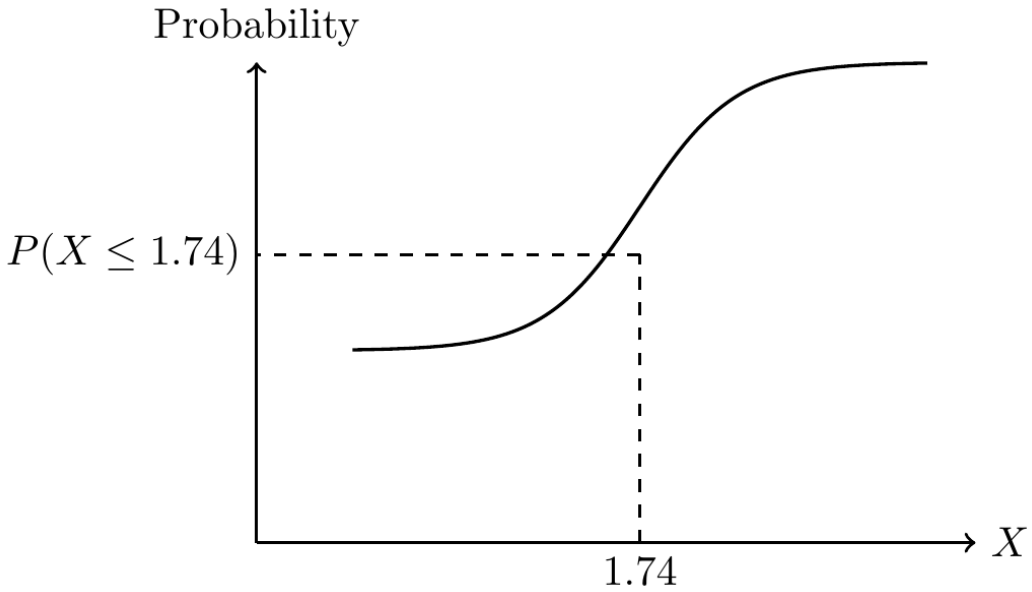


Whether an infinite set of numbers is **countable** or **uncountable** can be intuitively answered. All **natural** numbers $\mathbb{N}$ are countably infinite. We can clearly specify a way to count them (start at 0, then 1, then 2, then 3, ...), we just don't know where and when the path *ends*. After all, it is still infinitely long. All **real** numbers $\mathbb{R}$, however, are uncountably infinite. We cannot define a unique way to include all numbers. Suppose we start at 0, then 0.001 what about all the numbers in between? And all the numbers between those numbers? There is no way to count them all.

Cumulative Distribution Function

Cumulative Distribution Function

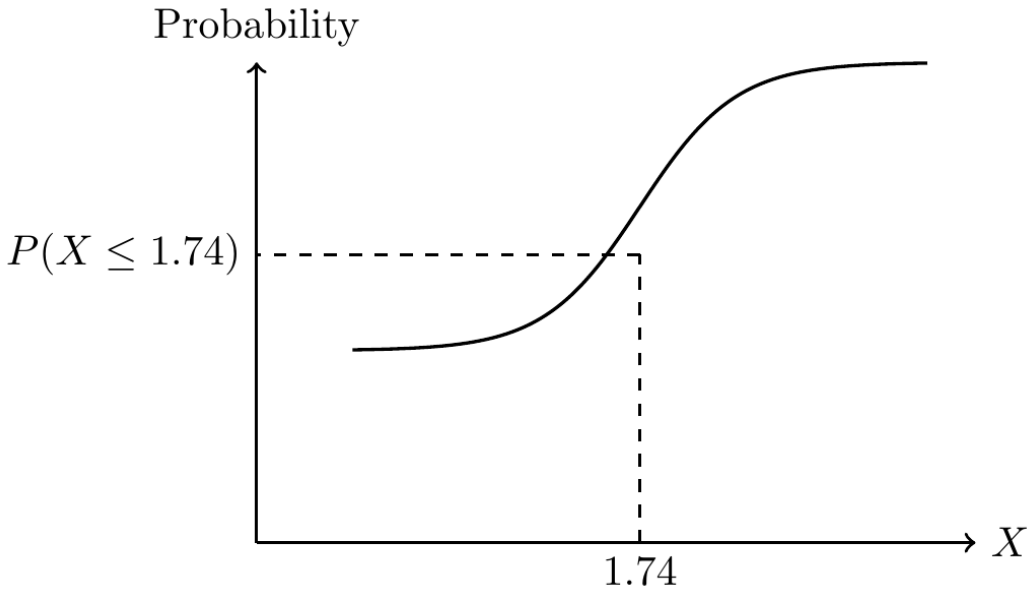
In addition to the **probability density function**, we can draw the **cumulative distribution function** (CDF). It represents the probability that the outcome is less than or equal to a certain value. The function is strictly monotonically increasing:



Cumulative Distribution Function

In addition to the **probability density function**, we can draw the **cumulative distribution function** (CDF). It represents the probability that the outcome is less than or equal to a certain value. The function is strictly monotonically increasing:

The dashed line shows how to read the plot: The value of the density function at $X = 1.74$ represents the probability that a randomly selected person is shorter than or exactly 1.74 meters tall.



Introduction to R with RStudio

Matrices and Vectors

Random Variables

Analysis of a Random Variable

Analysis of Two Random Variables

Expected Value (1)

Expected Value (1)

Let's return to our example of rolling a die. The outcome is a discrete random variable with the following outcomes and associated probabilities:

Expected Value (1)

Let's return to our example of rolling a die. The outcome is a discrete random variable with the following outcomes and associated probabilities:

Outcome	Probability
1	$\frac{1}{6}$
2	$\frac{1}{6}$
3	$\frac{1}{6}$
4	$\frac{1}{6}$
5	$\frac{1}{6}$
6	$\frac{1}{6}$

Expected Value (2)

Expected Value (2)

The **expected value** (or **expectation**) is a concept that allows us to simply analyze what value we can *expect* when rolling the die. We calculate it as the weighted arithmetic mean of the outcomes, with their respective probabilities as weights. We denote the expected value with a capital E:

Expected Value (2)

The **expected value** (or **expectation**) is a concept that allows us to simply analyze what value we can *expect* when rolling the die. We calculate it as the weighted arithmetic mean of the outcomes, with their respective probabilities as weights. We denote the expected value with a capital E:

$$\mathrm{E} \equiv \sum_{i=1}^{n} x_i p_i$$

Expected Value (2)

The **expected value** (or **expectation**) is a concept that allows us to simply analyze what value we can *expect* when rolling the die. We calculate it as the weighted arithmetic mean of the outcomes, with their respective probabilities as weights. We denote the expected value with a capital E:

$$\mathrm{E} \equiv \sum_{i=1}^{n} x_i p_i$$

The expected value of a fair die is 3.5. If we draw more and more outcomes from this distribution, i.e., roll the die many times, the average of all rolls will approach the expected value more and more. As long as we work with discrete variables, all of this is relatively simple to interpret. For continuous variables, it becomes more challenging, but the general intuition remains the same.

Rules for Expected Values

Rules for Expected Values

In econometrics, we work a lot with expected values, so it is useful to know some rules for handling them.

Rules for Expected Values

In econometrics, we work a lot with expected values, so it is useful to know some rules for handling them.

- For a constant c: $\mathrm{E}(c) = c$

Rules for Expected Values

In econometrics, we work a lot with expected values, so it is useful to know some rules for handling them.

- For a constant c: $\mathrm{E}(c) = c$
- For random variables X, Y and constants c, d:
  $\mathrm{E}(c \cdot X + d \cdot Y) = c \cdot \mathrm{E}(X) + d \cdot \mathrm{E}(Y)$

Rules for Expected Values

In econometrics, we work a lot with expected values, so it is useful to know some rules for handling them.

- For a constant c: $\mathrm{E}(c) = c$
- For random variables X, Y and constants c, d:
  $\mathrm{E}(c \cdot X + d \cdot Y) = c \cdot \mathrm{E}(X) + d \cdot \mathrm{E}(Y)$
- For constants $c_1, \ldots, c_n$ and random variables $X_1, \ldots, X_n$:
  $\mathrm{E}\left(\sum_{i=1}^{n} c_i X_i\right) = \sum_{i=1}^{n} c_i \mathrm{E}(X_i)$

Rules for Expected Values

In econometrics, we work a lot with expected values, so it is useful to know some rules for handling them.

- For a constant c: $\mathrm{E}(c) = c$
- For random variables X, Y and constants c, d:
  $\mathrm{E}(c \cdot X + d \cdot Y) = c \cdot \mathrm{E}(X) + d \cdot \mathrm{E}(Y)$
- For constants $c_1, \ldots, c_n$ and random variables $X_1, \ldots, X_n$:
  $\mathrm{E}\left(\sum_{i=1}^{n} c_i X_i\right) = \sum_{i=1}^{n} c_i \mathrm{E}(X_i)$
- For two *independent* random variables X, Y: $\mathrm{E}(XY) = \mathrm{E}(X)\mathrm{E}(Y)$

Variance (1)

Variance (1)

Often, the expected value alone is not sufficient to analyze a distribution. Imagine you own a company that manufactures screws. You have two machines producing them. You advertise that your screws are all 35 millimeters long, but in reality, the length of the screws is randomly distributed: The expected value of the screw length is 35 mm for both machines. However, machine A mostly produces screws very close to the desired length, while machine B sometimes produces screws as short as 33 mm or as long as 37 mm. What is the difference between these two machines with identical expected values?

Variance (1)

Often, the expected value alone is not sufficient to analyze a distribution. Imagine you own a company that manufactures screws. You have two machines producing them. You advertise that your screws are all 35 millimeters long, but in reality, the length of the screws is randomly distributed: The expected value of the screw length is 35 mm for both machines. However, machine A mostly produces screws very close to the desired length, while machine B sometimes produces screws as short as 33 mm or as long as 37 mm. What is the difference between these two machines with identical expected values?

The answer is **variance**. Simply put: The expected value shows us the "center" of a distribution. Variance, on the other hand, indicates how far the outcomes tend to deviate from this expectation. We denote it as $\mathrm{Var}(X)$ and calculate it as follows:

Variance (1)

Often, the expected value alone is not sufficient to analyze a distribution. Imagine you own a company that manufactures screws. You have two machines producing them. You advertise that your screws are all 35 millimeters long, but in reality, the length of the screws is randomly distributed: The expected value of the screw length is 35 mm for both machines. However, machine A mostly produces screws very close to the desired length, while machine B sometimes produces screws as short as 33 mm or as long as 37 mm. What is the difference between these two machines with identical expected values?

The answer is **variance**. Simply put: The expected value shows us the "center" of a distribution. Variance, on the other hand, indicates how far the outcomes tend to deviate from this expectation. We denote it as $\mathrm{Var}(X)$ and calculate it as follows:

$$\mathrm{Var}(X) \equiv \mathrm{E}\left((X-\mu)^2\right),$$

where $\mu = \mathrm{E}(X)$.

Variance (2)

Variance (2)

It is evident that the variance of any constant is zero. Additionally, the following rule applies for a random variable X and constants a, b:

$$\mathrm{Var}(aX + b) = a^2\mathrm{Var}(X) + \mathrm{Var}(b) = a^2\mathrm{Var}(X)$$

The standard deviation, denoted as $\mathrm{sd}(X)$, is simply the square root of the variance.

Analysis of Two Random Variables

Joint Probability Distribution

Joint Probability Distribution

Suppose X and Y are two discrete random variables. In addition to their individual distributions, we can describe their **joint distribution** using a joint probability function:

$$f_{X,Y}(x, y) = P(X = x, Y = y)$$

Joint Probability Distribution

Suppose X and Y are two discrete random variables. In addition to their individual distributions, we can describe their **joint distribution** using a joint probability function:

$$f_{X,Y}(x, y) = P(X = x, Y = y)$$

This function simply specifies the probability for each combination of X and Y. If X and Y are independent, then:

$$f_{X,Y}(x, y) = f_X(x) f_Y(y),$$

where $f(x)$ and $f(y)$ are the probability functions for X and Y, respectively. Two random variables are **independent** if the outcome of X does not affect the probabilities of the possible outcomes of Y.

Conditional Distribution

Conditional Distribution

Another important concept is the **conditional distribution**. The conditional probability density function describes how the outcome of X affects that of Y:

$$f_{Y|X}(y|x) = P(Y = y|X = x) = \frac{f_{X,Y}(x, y)}{f_X(x)}, \text{ for all } f_X(x) > 0$$

Conditional Distribution

Another important concept is the **conditional distribution**. The conditional probability density function describes how the outcome of X affects that of Y:

$$f_{Y|X}(y|x) = P(Y = y|X = x) = \frac{f_{X,Y}(x, y)}{f_X(x)}, \text{ for all } f_X(x) > 0$$

If X and Y are independent, the outcome of X does not affect Y, and thus $f_{Y|X}(y|x) = f_Y(y)$.

Covariance

Covariance

Covariance is similar to a "two-variable version" of variance. It allows us to analyze two distributions together. It is defined as follows and denoted as $\operatorname{Cov}(X, Y)$:

$$\operatorname{Cov}(X, Y) \equiv \operatorname{E}\left((X - \mu_X)(Y - \mu_Y)\right),$$

where $\mu_X = \operatorname{E}(X)$ and $\mu_Y = \operatorname{E}(Y)$.

Covariance

Covariance is similar to a "two-variable version" of variance. It allows us to analyze two distributions together. It is defined as follows and denoted as $\operatorname{Cov}(X, Y)$:

$$\operatorname{Cov}(X, Y) \equiv \operatorname{E}\left((X - \mu_X)(Y - \mu_Y)\right),$$

where $\mu_X = \operatorname{E}(X)$ and $\mu_Y = \operatorname{E}(Y)$.

The sign of the covariance can be intuitively interpreted. If the covariance is positive, we expect Y to be above its mean when X is as well. If the covariance is negative, we expect Y to be below its mean when X is above its mean. Simply put, a positive covariance indicates that two variables are positively associated, and vice versa. A covariance of 0 means that there is no relationship. If X and Y are independent, the covariance is always 0.

Covariance

Covariance is similar to a "two-variable version" of variance. It allows us to analyze two distributions together. It is defined as follows and denoted as $\operatorname{Cov}(X, Y)$:

$$\operatorname{Cov}(X, Y) \equiv \operatorname{E}\left((X - \mu_X)(Y - \mu_Y)\right),$$

where $\mu_X = \operatorname{E}(X)$ and $\mu_Y = \operatorname{E}(Y)$.

The sign of the covariance can be intuitively interpreted. If the covariance is positive, we expect Y to be above its mean when X is as well. If the covariance is negative, we expect Y to be below its mean when X is above its mean. Simply put, a positive covariance indicates that two variables are positively associated, and vice versa. A covariance of 0 means that there is no relationship. If X and Y are independent, the covariance is always 0.

An association in this sense does not necessarily imply causation, but more on that in the course :)

Rules for Covariance

Rules for Covariance

The following rules apply for covariance:

Rules for Covariance

The following rules apply for covariance:

$$\operatorname{Cov}(X, Y) = \operatorname{E}(XY) - \operatorname{E}(X)\operatorname{E}(Y)$$

Rules for Covariance

The following rules apply for covariance:

$$\operatorname{Cov}(X, Y) = \operatorname{E}(XY) - \operatorname{E}(X)\operatorname{E}(Y)$$

For constants a, b, c, d:

$$\operatorname{Cov}(aX + b, cY + d) = a \cdot c \cdot \operatorname{Cov}(X, Y)$$

Conditional Expectation

Conditional Expectation

Suppose we have two random variables X and Y that are somehow connected. We want to know the expectation of Y, given that X takes on a specific value. This is called the **conditional expectation** and is denoted as $\mathrm{E}(Y|X=x)$. The following rules apply for conditional expectations:

Conditional Expectation

Suppose we have two random variables X and Y that are somehow connected. We want to know the expectation of Y, given that X takes on a specific value. This is called the **conditional expectation** and is denoted as $\mathrm{E}(Y|X=x)$. The following rules apply for conditional expectations:

- If $f(x)$ is any function: $\mathrm{E}\left(f(X)|X\right) = f(X)$

Conditional Expectation

Suppose we have two random variables X and Y that are somehow connected. We want to know the expectation of Y, given that X takes on a specific value. This is called the **conditional expectation** and is denoted as $\mathrm{E}(Y|X=x)$. The following rules apply for conditional expectations:

- If $f(x)$ is any function: $\mathrm{E}\left(f(X)|X\right) = f(X)$
- For two functions $f(X)$ and $g(x)$: $\mathrm{E}\left(f(X)Y + g(X) \mid X\right) = f(X)\mathrm{E}(Y|X) + g(X)$

Conditional Expectation

Suppose we have two random variables X and Y that are somehow connected. We want to know the expectation of Y, given that X takes on a specific value. This is called the **conditional expectation** and is denoted as $\mathrm{E}(Y|X=x)$. The following rules apply for conditional expectations:

- If $f(x)$ is any function: $\mathrm{E}\left(f(X)|X\right) = f(X)$
- For two functions $f(X)$ and $g(x)$: $\mathrm{E}\left(f(X)Y + g(X) \mid X\right) = f(X)\mathrm{E}(Y|X) + g(X)$
- If X and Y are independent: $\mathrm{E}(Y|X) = \mathrm{E}(Y)$

Conditional Expectation

Suppose we have two random variables X and Y that are somehow connected. We want to know the expectation of Y, given that X takes on a specific value. This is called the **conditional expectation** and is denoted as $\mathrm{E}(Y|X=x)$. The following rules apply for conditional expectations:

- If $f(x)$ is any function: $\mathrm{E}\left(f(X)|X\right) = f(X)$
- For two functions $f(X)$ and $g(x)$: $\mathrm{E}\left(f(X)Y + g(X) \mid X\right) = f(X)\mathrm{E}(Y|X) + g(X)$
- If X and Y are independent: $\mathrm{E}(Y|X) = \mathrm{E}(Y)$
- The **law of iterated expectations**: $\mathrm{E}\left(\mathrm{E}(Y|X)\right) = \mathrm{E}(Y)$