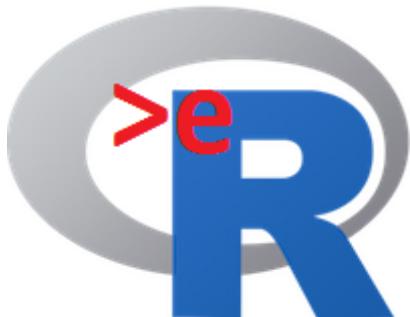




This course was developed as a part of several VLIR-UOS projects:

- Cross-cutting Statistics: 2011-2016, 2017.
- Cross-cutting Statistics: 2017.
- Statistics for development : 2018-2022.
- The >eR-BioStat platform ITP project: 2024-2026.



The >eR-Biostat initiative
Making R based education materials in
statistics accessible for all

Basic concept of statistical inference and estimation using R: content development

Ziv Shkedy (Hasselt University)

LAST UPDATE: 06/2026



ER-BioStat

 <https://github.com/eR-Biostat>

 @erbiostat

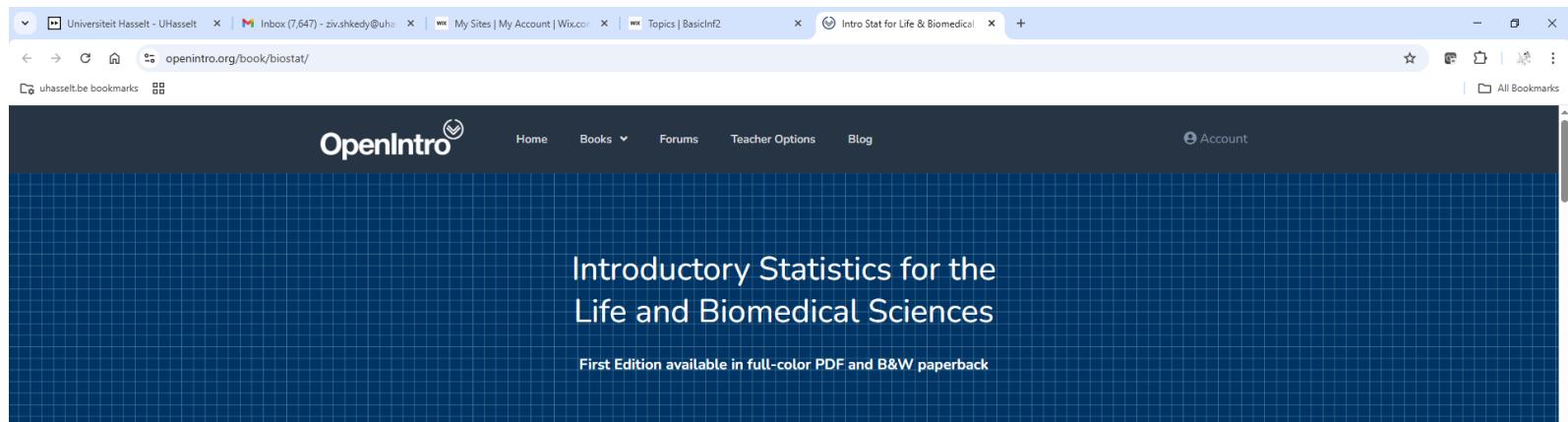
Basic concepts of Statistical inference and estimation: course's content

Information about the course level

- The course was developed at an undergraduate level:
 - Students in Statistics.
 - Students in life science.
- But:
 - Applied materials for master students.
 - Short course on basic statistics.
 - Training course for (non statisticians) PhD students.
- Data science: basic course about statistics.

Reference book

- Online book of Jullie Vu and David Harrington.



Textbook Pedagogy

Introduction to Statistics for the Life and Biomedical Sciences has been written to be used in conjunction with a set of self-paced learning labs. These labs guide students through learning how to apply statistical ideas and concepts discussed in the text with the R computing language.

The text discusses the important ideas used to support an interpretation (such as the notion of a confidence interval), rather than the process of generating such material from data (such as computing a confidence interval for a particular subset of individuals in a study). This allows students whose main focus is understanding statistical concepts to not be distracted by the details of a particular software package. In our experience, however, we have found that many students enter a research setting after only a single course in statistics. These students benefit from a practical introduction to data analysis that incorporates the use of a statistical computing language.

In a classroom setting, we have found it beneficial for students to start working through the labs after having been exposed to the corresponding material in the text, either from self-reading or through an instructor presenting the main ideas. The labs are organized by chapter, and each lab corresponds to a particular section or set of sections in the text.

There are traditional exercises at the end of each chapter that do not require the use of computing. More complicated methods, such as multiple regression, do not lend themselves to hand calculation and computing is necessary for gaining practical experience with these methods. The lab exercises for these later chapters become an increasingly important part of mastering the material.



<https://www.openintro.org/book/biostat/>

Reference book

Chapter 4

Foundations for inference

- 4.1 Variability in estimates
- 4.2 Confidence intervals
- 4.3 Hypothesis testing
- 4.4 Notes
- 4.5 Exercises

Chapter 5

Inference for numerical data

- 5.1 Single-sample inference with the *t*-distribution
- 5.2 Two-sample test for paired data
- 5.3 Two-sample test for independent data
- 5.4 Power calculations for a difference of means
- 5.5 Comparing means with ANOVA
- 5.6 Notes
- 5.7 Exercises

Chapter 8

Inference for categorical data

- 8.1 Inference for a single proportion
- 8.2 Inference for the difference of two proportions
- 8.3 Inference for two or more groups
- 8.4 Chi-square tests for the fit of a distribution
- 8.5 Outcome-based sampling: case-control studies
- 8.6 Notes
- 8.7 Exercises

- Introduction to statistic and life science.
- The course is based on three chapters from Jullie and Dave's book.
- Focus: inference and estimation.
- Example today: how to produce Chapter 4.

Online book for Chapter 4

The screenshot shows a web browser window displaying an RStudio RPubs chapter page. The title bar includes tabs for 'Universiteit Hasselt - UHasselt', 'Inbox (7,647) - ziv.shkedy@uhasselt.be', 'My Sites | My Account | Wix.com', 'Topics | BasicInfo2', and 'RPubs - Chapter 4 (Temp)'. The main content area is titled 'RPubs by RStudio' and features a sidebar with a table of contents:

1. Variability in estimates
2. Standard error of the mean
3. Confidence intervals
4. Hypothesis testing
5. Hypothesis testing and confidence intervals
6. Decision error (Type I and Type II error)

The main content area includes a timestamp '26-03-2025' and author information 'Ziv Shkedy and Thi Huyen Nguyen based on Chapter 4 in the book of Julie Vu and Dave Harrington *Introductory Statistics for the Life and Biomedical Sciences* (<https://www.openintro.org/book/biostat/>)'. A large section header 'Foundations for inference using R' is followed by a sub-section '1. Variability in estimates'.

1. Variability in estimates

1.2 A point estimate for the population parameter

A natural way to estimate features of the population, such as the population mean weight, is to use the corresponding summary statistic calculated from the sample. For example, the sample mean \bar{x} is a point parameter estimate for the population (unknown) mean μ and the sample variance s^2 is a point parameter estimate for the population variance σ^2 .

1.3 Example: the wind speed in the airquality dataset

Data and point estimates

The airquality dataset gives information about 153 daily air quality measurements in New York, May to September 1973.

```
## [1] 153   6
```

The first 6 lines of the dataset are shown below.

```
## [1] 153   6
```

At the bottom of the page, there is a footer with the title 'Chapter 4 (Temp)', the author 'by Ziv Shkedy', and the last update 'Last updated 11 days ago'. It also includes a search bar, social sharing icons, and a toolbar with options like 'Comments (-)', 'Share', and 'Hide Toolbars'. The bottom right corner shows the date and time '14:42 6/06/2025'.

This presentation

- How do we produce the files for Chapter 4.
- From the laptop to the website.

How do we produce the HTML files ?

Content development

- How can we create/change content ?
- How can we upload online ?
- How to create a link of the online file to a website ?
- Producing the content in three steps.

From Rmd to HTML (laptop)

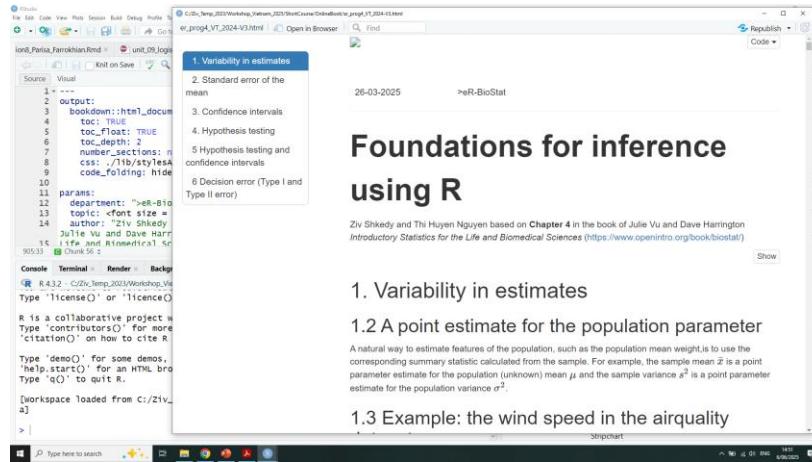
The screenshot shows the RStudio interface. On the left, the R Markdown file 'unit_09_logistic_regression.Rmd' is open, displaying R code and a YAML front matter section. The front matter includes parameters like 'output: bookdown::html_document2:', 'params:', and 'author: "Ziv Shkedy and Thi Huyen Nguyen based on Chapter 4 in the book of Julie Vu and Dave Harrington "Introductory Statistics for the Life and Biomedical Sciences" (<https://www.openintro.org/book/biostat/>)"'. The right side of the interface shows the 'Environment' tab with a list of packages installed in the user library, such as 'abind', 'airports', 'aplore3', 'gridExtra', 'knitr', 'lattice', 'lazyeval', 'magrittr', 'mnorm', 'openintro', 'pander', 'plotly', 'raster', 'readr', 'rmarkdown', 'tidyverse', 'viridis', and 'viridisLite'. Below the environment is a 'Plots' tab showing a histogram of a normal distribution.

- R Markdown program:
er_prog4_VT_2024_V1_Rmd

The screenshot shows the RStudio interface with the generated HTML output for 'er_prog4_VT_2024_V1.html'. The title of the page is 'er_prog4_VT_2024_V1.html'. The content of the page is identical to the R Markdown source code, including the YAML front matter, the R code, and the resulting HTML structure. The page is titled 'Foundations for inference using R' and contains sections 1. Variability in estimates, 1.2 A point estimate for the population parameter, and 1.3 Example: the wind speed in the airquality.

- R Markdown's output.
- HTML file.

Offline to Online



- Laptop.

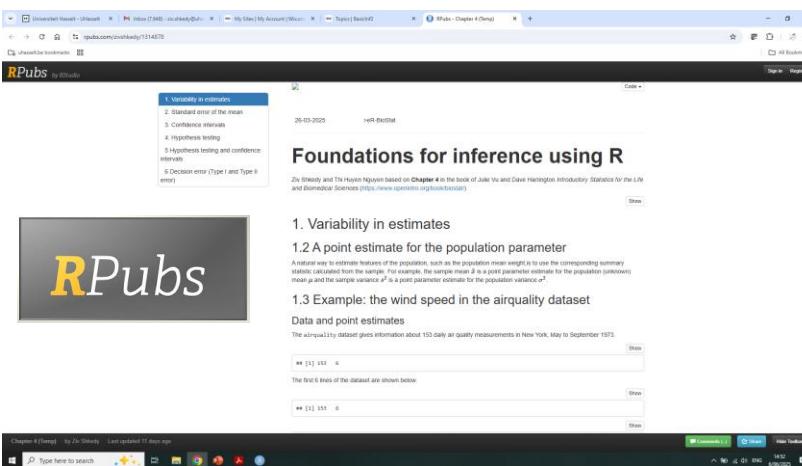
Foundations for inference using R

1. Variability in estimates

1.2 A point estimate for the population parameter

A natural way to estimate features of the population, such as the population mean weight, is to use the corresponding summary statistic calculated from the sample. For example, the sample mean \bar{x} is a point parameter estimate for the population (unknown) mean μ and the sample variance s^2 is a point parameter estimate for the population variance σ^2 .

1.3 Example: the wind speed in the airquality



- RPubs.
- Online (HTML).
 - Everybody can see and use.

<https://rpubs.com/zivshkedy/1323673>

Online at the website

<https://rpubs.com/zivshkedy/1323673>

The screenshot shows a web browser window with multiple tabs open. The active tab is titled 'Rpubs - Chapter 4 (Temp)' and displays a document titled 'Foundations for inference using R'. The document content includes a table of contents with sections like 'Variability in estimates', 'Standard error of the mean', 'Confidence Intervals', etc. Below the table of contents, there's a section titled 'Foundations for inference using R' with a sub-section '1. Variability in estimates'. It also contains code snippets and a small image of the R logo.

<https://erbiostat.wixsite.com/basicinf>

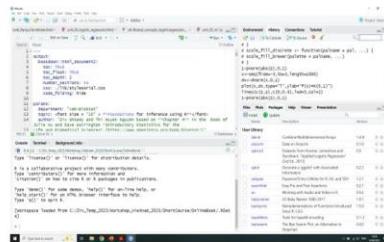
The screenshot shows a Wix website with a large 'WIX' logo. The main content area is divided into several green boxes containing links to 'R programs' for different chapters (Chapter 4, Chapter 5, Chapter 6). Below these are sections for 'Online book' and 'This part of the course is based on unit 4 in Vu & Harrison course and it covers the following topics...'. At the bottom, there are more green boxes for 'Slides (PDF) Unit 4', 'Slides (PDF) Unit 5', 'Slides (PDF) Unit 6', and 'Slides (PDF) Unit 8'. A blue arrow points from the text 'Links to the course materials.' down towards the 'WIX' website screenshot.

Links to the course materials.

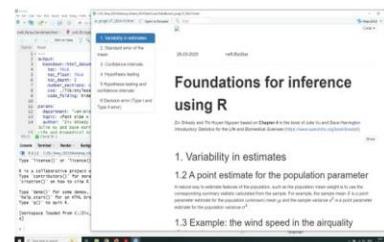
- HRML: RPUs.
- Rmd: GitHub.

Step 1: how to produce the HTML file ?

From Rmd to HTML (laptop)



- R Markdown program.



- R Markdown's output.
- HTML file.

The HTML file

A screenshot of a Microsoft Edge browser window displaying an R Markdown generated HTML file. The title bar shows the file path: C:/Ziv_Temp_2023/Workshop_Vietnam_2025/ShortCourse/OnlineBook/er_prog4_VT_2024-V3.html. The page content includes a sidebar menu, a header with a date and author information, a main title, and several sections of text and code snippets.

1. Variability in estimates

2. Standard error of the mean
3. Confidence intervals
4. Hypothesis testing
5. Hypothesis testing and confidence intervals
6. Decision error (Type I and Type II error)

26-03-2025 >eR-BioStat

Foundations for inference using R

Ziv Shkedy and Thi Huyen Nguyen based on **Chapter 4** in the book of Julie Vu and Dave Harrington *Introductory Statistics for the Life and Biomedical Sciences* (<https://www.openintro.org/book/biostat/>)

1. Variability in estimates

1.2 A point estimate for the population parameter

A natural way to estimate features of the population, such as the population mean weight, is to use the corresponding summary statistic calculated from the sample. For example, the sample mean \bar{x} is a point parameter estimate for the population (unknown) mean μ and the sample variance s^2 is a point parameter estimate for the population variance σ^2 .

1.3 Example: the wind speed in the airquality dataset

Data and point estimates

The airquality dataset gives information about 153 daily air quality measurements in New York, May to September 1973.

```
## [1] 153   6
```

The first 6 lines of the dataset are shown below.

```
## [1] 153   6
```

	Ozone	Solar.R	Wind	Temp	Month	Day
## 1	41	190	7.4	67	5	1
## 2	36	118	8.0	72	5	2
## 3	12	149	12.6	74	5	3
## 4	19	313	11.5	62	5	4

The Rmd file

- You will need:
 - Install R packages.
 - Tex/Latex...

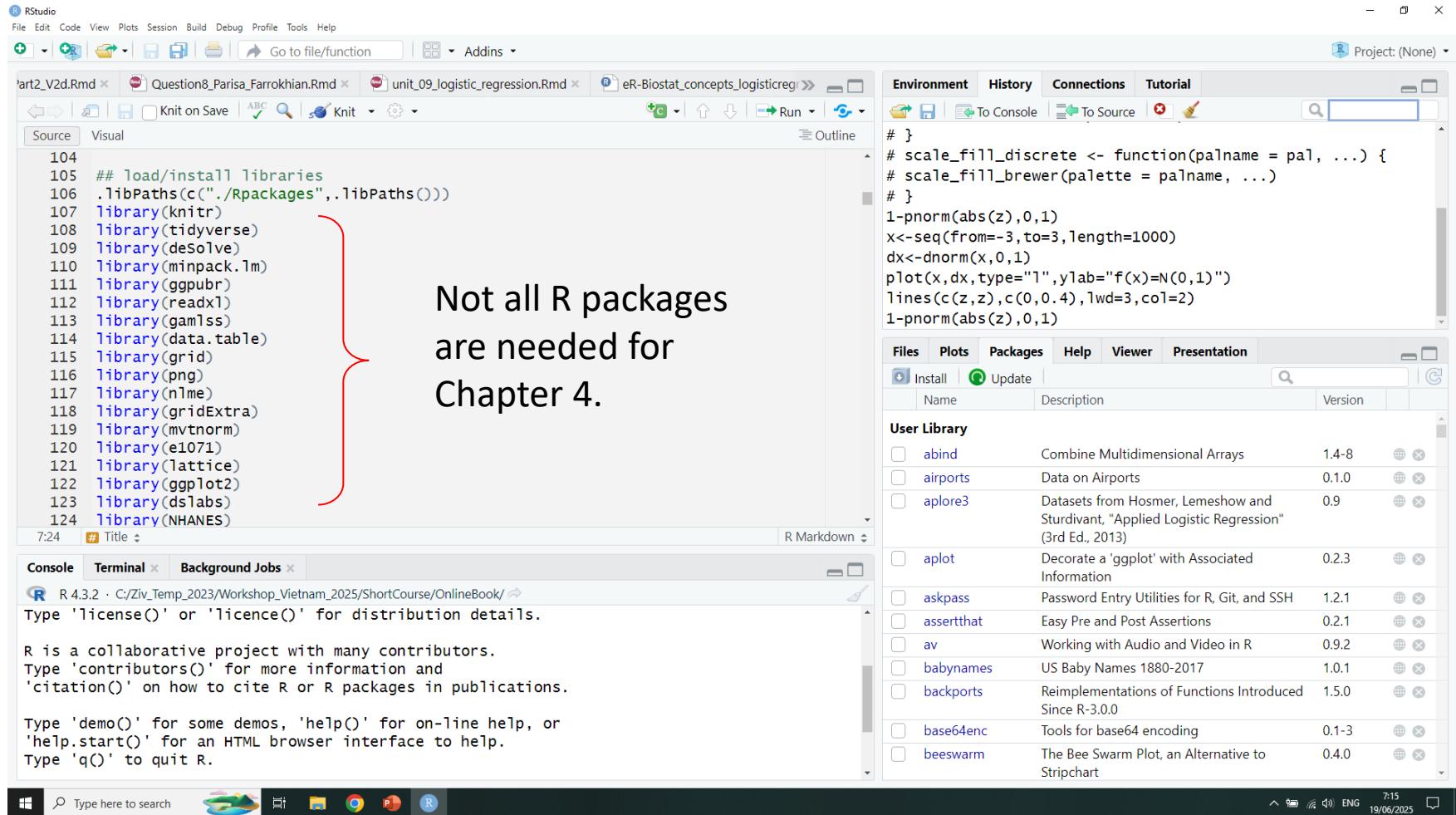
The Rmd file: YMAL

The screenshot shows the RStudio interface with the following details:

- File Menu:** File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help.
- Toolbar:** Go to file/function, Addins.
- Project:** Project: (None).
- Source Editor:** Contains an Rmd file with the following content:

```
1 ---  
2 output:  
3   bookdown::html_document2:  
4     toc: TRUE  
5     toc_float: TRUE  
6     toc_depth: 2  
7     number_sections: no  
8     css: ./lib/stylesArial.css  
9     code_folding: hide  
10  
11 params:  
12   department: ">eR-BioStat"  
13   topic: <font size = "10" > **Foundations for inference using R**</font>  
14   author: "Ziv Shkedy and Thi Huyen Nguyen based on **Chapter 4** in the book of  
Julie Vu and Dave Harrington *Introductory Statistics for the  
Life and Biomedical Sciences* (https://www.openintro.org/book/biostat/)"  
15   date: "26-03-2025"  
16   endCode: FALSE  
17   RmdLocation: ""  
18  
19 ---  
20
```
- Environment Tab:** Shows the R environment with packages like abind, airports, aplore3, aplot, askpass, assertthat, av, babynames, backports, base64enc, and beeswarm.
- Console Tab:** Displays the R startup message and basic help information.
- Taskbar:** Shows the Windows taskbar with various pinned icons.

The Rmd file: R packages



Not all R packages are needed for Chapter 4.

```
104  
105 ## load/install libraries  
106 .libPaths(c("./Rpackages", .libPaths()))  
107 library(knitr)  
108 library(tidyverse)  
109 library(deSolve)  
110 library(minpack.lm)  
111 library(ggpubr)  
112 library(readxl)  
113 library(gam1ss)  
114 library(data.table)  
115 library(grid)  
116 library(png)  
117 library(nlme)  
118 library(gridExtra)  
119 library(mvtnorm)  
120 library(e1071)  
121 library(lattice)  
122 library(ggplot2)  
123 library(ds-labs)  
124 library(NHANES)
```

R 4.3.2 · C:/Ziv_Temp_2023/Workshop_Vietnam_2025/ShortCourse/OnlineBook/ ↵
Type 'license()' or 'licence()' for distribution details.
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

Environment History Connections Tutorial

Project: (None)

Name	Description	Version
abind	Combine Multidimensional Arrays	1.4-8
airports	Data on Airports	0.1.0
aplore3	Datasets from Hosmer, Lemeshow and Sturdivant, "Applied Logistic Regression" (3rd Ed., 2013)	0.9
aplot	Decorate a 'ggplot' with Associated Information	0.2.3
askpass	Password Entry Utilities for R, Git, and SSH	1.2.1
assertthat	Easy Pre and Post Assertions	0.2.1
av	Working with Audio and Video in R	0.9.2
babynames	US Baby Names 1880-2017	1.0.1
backports	Reimplementations of Functions Introduced Since R-3.0.0	1.5.0
base64enc	Tools for base64 encoding	0.1-3
beeswarm	The Bee Swarm Plot, an Alternative to Stripchart	0.4.0

Type here to search

7:15 19/06/2025 ENG

The Rmd file: free text

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
+ Addins Go to file/function Addins Project: (None)

part2_V2d.Rmd x Question8_Parisafarrokhan.Rmd x unit_09_logistic_regression.Rmd x eR-Biostat_concepts_logisticreg...
Knit on Save ABC Knit Run Outline

Source Visual

```
139 # 1. Variability in estimates
140
141 ## 1.2 A point estimate for the population parameter
142
143 A natural way to estimate features of the population, such as the population mean weight, is to use the corresponding summary statistic calculated from the sample. For example, the sample mean  $\bar{x}$  is a point parameter estimate for the population (unknown) mean  $\mu$  and the sample variance  $s^2$  is a point parameter estimate for the population variance  $\sigma^2$ .
144
145 ## 1.3 Example: the wind speed in the airquality dataset
146
147 ### Data and point estimates
148
149 The <tt>airquality</tt> dataset gives information about 153 daily air quality measurements in New York, May to September 1973.
150
151 ``{r, echo=TRUE, message=FALSE, warning=FALSE}
152 dim(airquality)
153
154
```

7:24 Title R Markdown

Console Terminal Background Jobs

R 4.3.2 · C:/Ziv_Temp_2023/Workshop_Vietnam_2025/ShortCourse/OnlineBook/ Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors. Type 'contributors()' for more information and 'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or 'help.start()' for an HTML browser interface to help. Type 'q()' to quit R.

Environment History Connections Tutorial
To Console To Source

```
# }
# scale_fill_discrete <- function(palname = pal, ...){
#   scale_fill_brewer(palette = palname, ...)
# }
1-pnorm(abs(z),0,1)
x<-seq(from=-3,to=3,length=1000)
dx<-dnorm(x,0,1)
plot(x,dx,type="l",ylab="f(x)=N(0,1)")
lines(c(z,z),c(0,0.4),lwd=3,col=2)
1-pnorm(abs(z),0,1)
```

Files Plots Packages Help Viewer Presentation
Install Update Name Description Version

User Library

Name	Description	Version
abind	Combine Multidimensional Arrays	1.4-8
airports	Data on Airports	0.1.0
aplore3	Datasets from Hosmer, Lemeshow and Sturdivant, "Applied Logistic Regression" (3rd Ed., 2013)	0.9
aplot	Decorate a 'ggplot' with Associated Information	0.2.3
askpass	Password Entry Utilities for R, Git, and SSH	1.2.1
assertthat	Easy Pre and Post Assertions	0.2.1
av	Working with Audio and Video in R	0.9.2
babynames	US Baby Names 1880-2017	1.0.1
backports	Reimplementations of Functions Introduced Since R-3.0.0	1.5.0
base64enc	Tools for base64 encoding	0.1-3
beeswarm	The Bee Swarm Plot, an Alternative to Stripchart	0.4.0

The HTML file: free text

New Tab er_prog4_VT_2024-V3.knit C:/Ziv_Temp_2023/Workshop_Vietnam_2025/ShortCourse/OnlineBook/er_prog4_VT_2024-V3.html uhasselt.be bookmarks

1. Variability in estimates

1.2 A point estimate for the population parameter

A natural way to estimate features of the population, such as the population mean weight, is to use the corresponding summary statistic calculated from the sample. For example, the sample mean \bar{x} is a point parameter estimate for the population (unknown) mean μ and the sample variance s^2 is a point parameter estimate for the population variance σ^2 .

1.3 Example: the wind speed in the airquality dataset

Data and point estimates

The airquality dataset gives information about 153 daily air quality measurements in New York, May to September 1973.

```
## [1] 153 6
```

The first 6 lines of the dataset are shown below.

```
## [1] 153 6
```

	Ozone	Solar.R	Wind	Temp	Month	Day
## 1	41	190	7.4	67	5	1
## 2	36	118	8.0	72	5	2
## 3	12	149	12.6	74	5	3
## 4	18	313	11.5	62	5	4
## 5	NA	NA	14.3	56	5	5
## 6	28	NA	14.9	66	5	6

The variable Wind is the average wind speed in miles per hour at 0700 and 1000 hours at LaGuardia Airport. The mean wind speed is $\hat{\mu} = \bar{x} = 9.95$ and the sample standard deviation is $\hat{\sigma} = s = 3.52$. This sample mean is a point estimate of the population mean. If a different random sample of 153 days were taken the new sample mean would likely be different as a result of sampling variation. Note that while estimates generally vary from one sample to another, the population mean (μ) and variance (σ^2) are fixed values (the parameters in the population).

```
## [1] 9.957516
```

```
## [1] 3.523001
```

Type here to search

The Rmd file: sections, subsections..

The screenshot shows the RStudio interface with an Rmd file open. Red arrows point to several code blocks that represent sections and subsections:

- Line 139: `# 1. Variability in estimates`
- Line 141: `## 1.2 A point estimate for the population parameter`
- Line 147: `### Data and point estimates`

The RStudio environment includes:

- Project: (None)
- Environment tab (selected)
- History, Connections, Tutorial tabs
- To Console, To Source buttons
- Code editor with R code and red annotations
- Outline panel
- Files, Plots, Packages, Help, Viewer, Presentation tabs
- User Library table listing various R packages
- Console tab showing R startup messages
- Background Jobs tab
- Taskbar at the bottom with search bar and icons

The HTML file: sections, subsections..

A screenshot of a Microsoft Edge browser window displaying an HTML document titled "er_prog4_VT_2024-V3.html". The page content is as follows:

1. Variability in estimates

- 1.2 A point estimate for the population parameter
- 1.3 Example: the wind speed in the airquality dataset
- 1.4 The variability of the sample mean
- 1.5 The sampling distribution for the mean
- 1.6 The Central Limit Theorem
- 2. Standard error of the mean
- 3. Confidence intervals
- 4. Hypothesis testing
- 5. Hypothesis testing and confidence intervals
- 6. Decision error (Type I and Type II error)

1.2 A point estimate for the population parameter

A natural way to estimate features of the population, such as the population mean weight, is to use the corresponding summary statistic calculated from the sample. For example, the sample mean \bar{x} is a point parameter estimate for the population (unknown) mean μ and the sample variance s^2 is a point parameter estimate for the population variance σ^2 .

1.3 Example: the wind speed in the airquality dataset

Data and point estimates

The airquality dataset gives information about 153 daily air quality measurements in New York, May to September 1973.

[1] 153 6

The first 6 lines of the dataset are shown below.

[1] 153 6

	Ozone	Solar.R	Wind	Temp	Month	Day
## 1	41	190	7.4	67	5	1
## 2	36	118	8.0	72	5	2
## 3	12	149	12.6	74	5	3
## 4	18	313	11.5	62	5	4
## 5	NA	NA	14.3	56	5	5
## 6	28	NA	14.9	66	5	6

The variable Wind is the average wind speed in miles per hour at 0700 and 1000 hours at LaGuardia Airport. The mean wind speed is $\hat{\mu} = \bar{x} = 9.95$ and the sample standard deviation is $\hat{\sigma} = s = 3.52$. This sample mean is a point estimate of the population mean. If a different random sample of 153 days were taken the new sample mean would likely be different as a result of sampling variation. Note that while estimates generally vary from one sample to another, the population mean (μ) and variance (σ^2) are fixed values (the parameters in the population).

[1] 9.957516

[1] 3.523001

Red arrows point to the following sections:

- 1. Variability in estimates
- 1.2 A point estimate for the population parameter
- 1.3 Example: the wind speed in the airquality dataset

The Rmd file: R code

The screenshot shows the RStudio interface with an Rmd file open. The code editor on the left contains R code, and the right pane shows the Environment, History, Connections, and Tutorial tabs. A specific block of code in the editor is highlighted with a blue rectangle.

```
139 # 1. Variability in estimates
140
141 ## 1.2 A point estimate for the population parameter
142
143 A natural way to estimate features of the population, such as the population mean weight, is to use the corresponding summary statistic calculated from the sample. For example, the sample mean  $\bar{x}$  is a point parameter estimate for the population (unknown) mean  $\mu$  and the sample variance  $s^2$  is a point parameter estimate for the population variance  $\sigma^2$ .
144
145 ## 1.3 Example: the wind speed in the airquality dataset
146
147 ### Data and point estimates
148
149 The <tt>airquality</tt> dataset gives information about 153 daily air quality measurements in New York, May to September 1973.
150
151 ``{r, echo=TRUE, message=FALSE, warning=FALSE}
152 dim(airquality)
153
154
```

Console output:

```
R 4.3.2 · C:/Ziv_Temp_2023/Workshop_Vietnam_2025/ShortCourse/OnlineBook/ ↵
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.
```

The HTML file: R output

New Tab er_prog4_VT_2024-V3.knit

File C:/Ziv_Temp_2023/Workshop_Vietnam_2025/ShortCourse/OnlineBook/er_prog4_VT_2024-V3.html

uhasselt.be bookmarks

1. Variability in estimates

- 1.2 A point estimate for the population parameter
- 1.3 Example: the wind speed in the airquality dataset
- 1.4 The variability of the sample mean
- 1.5 The sampling distribution for the mean
- 1.6 The Central Limit Theorem
- 2. Standard error of the mean
- 3. Confidence intervals
- 4. Hypothesis testing
- 5. Hypothesis testing and confidence intervals
- 6. Decision error (Type I and Type II error)

1.2 A point estimate for the population parameter

A natural way to estimate features of the population, such as the population mean weight, is to use the corresponding summary statistic calculated from the sample. For example, the sample mean \bar{x} is a point parameter estimate for the population (unknown) mean μ and the sample variance s^2 is a point parameter estimate for the population variance σ^2 .

1.3 Example: the wind speed in the airquality dataset

Data and point estimates

The airquality dataset gives information about 153 daily air quality measurements in New York, May to September 1973.

```
## [1] 153   6
```

The first 6 lines of the dataset are shown below.

```
## [1] 153   6
```

	Ozone	Solar.R	Wind	Temp	Month	Day
## 1	41	190	7.4	67	5	1
## 2	36	118	8.0	72	5	2
## 3	12	149	12.6	74	5	3
## 4	18	313	11.5	62	5	4
## 5	NA	NA	14.3	56	5	5
## 6	28	NA	14.9	66	5	6

The variable Wind is the average wind speed in miles per hour at 0700 and 1000 hours at LaGuardia Airport. The mean wind speed is $\hat{\mu} = \bar{x} = 9.95$ and the sample standard deviation is $\hat{\sigma} = s = 3.52$. This sample mean is a point estimate of the population mean. If a different random sample of 153 days were taken the new sample mean would likely be different as a result of sampling variation. Note that while estimates generally vary from one sample to another, the population mean (μ) and variance (σ^2) are fixed values (the parameters in the population).

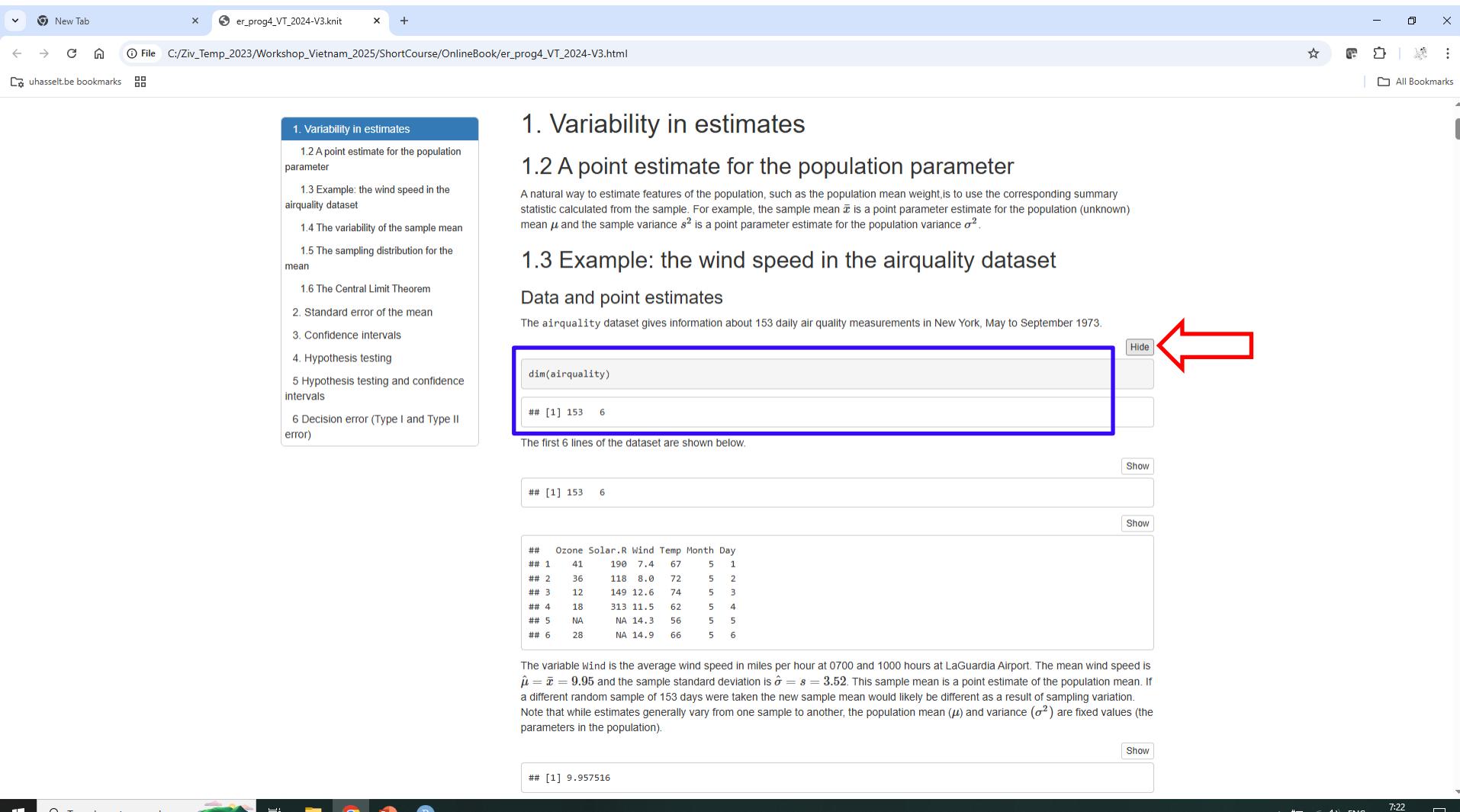
```
## [1] 9.957516
```

```
## [1] 3.523001
```

Type here to search

7:17 19/06/2025 ENG

The HTML file: R code & R output



A screenshot of a web browser displaying an R Markdown output page. The page title is "The HTML file: R code & R output". The main content is titled "1. Variability in estimates" and includes sections like "1.2 A point estimate for the population parameter" and "1.3 Example: the wind speed in the airquality dataset". A sidebar on the left lists topics such as "Variability in estimates", "A point estimate for the population parameter", "Example: the wind speed in the airquality dataset", "The variability of the sample mean", "The sampling distribution for the mean", "The Central Limit Theorem", "Standard error of the mean", "Confidence intervals", "Hypothesis testing", "Hypothesis testing and confidence intervals", and "Decision error (Type I and Type II error)".

The "1.3 Example: the wind speed in the airquality dataset" section contains R code and its output:

```
dim(airquality)
## [1] 153   6
```

The first 6 lines of the dataset are shown below.

```
## [1] 153   6
```

	Ozone	Solar.R	Wind	Temp	Month	Day
## 1	41	190	7.4	67	5	1
## 2	36	118	8.0	72	5	2
## 3	12	149	12.6	74	5	3
## 4	18	313	11.5	62	5	4
## 5	NA	NA	14.3	56	5	5
## 6	28	NA	14.9	66	5	6

The variable `Wind` is the average wind speed in miles per hour at 0700 and 1000 hours at LaGuardia Airport. The mean wind speed is $\hat{\mu} = \bar{x} = 9.95$ and the sample standard deviation is $\hat{\sigma} = s = 3.52$. This sample mean is a point estimate of the population mean. If a different random sample of 153 days were taken the new sample mean would likely be different as a result of sampling variation. Note that while estimates generally vary from one sample to another, the population mean (μ) and variance (σ^2) are fixed values (the parameters in the population).

```
## [1] 9.957516
```

The Rmd file: R code

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help

Project: (None)

art2_V2d.Rmd × Question8_Parisa_Farrokhan.Rmd × unit_09_logistic_regression.Rmd × eR-Biostat_concepts_logisticreg ×

Source Visual

```
measurements in New York, May to September 1973.
```

150
151 ````{r, echo=TRUE, message=FALSE, warning=FALSE}`
152 `dim(airquality)`
153 `...`
154
155 The first 6 lines of the dataset are shown below.
156
157 ````{r, echo=TRUE, message=FALSE, warning=FALSE}`
158 `dim(airquality)`
159 `head(airquality)`
160 `...`

161
162 The variable `<tt>Wind</tt>` is the average wind speed in miles per hour at 0700 and 1000 hours at LaGuardia Airport.
163 The mean wind speed is $\hat{\mu} = \bar{x} = 9.95$ and the sample standard deviation is $\hat{\sigma} = s = 3.52$. This sample mean is a point estimate of the population mean. If a different random sample of 153 days were taken the new sample mean would likely be different as a result of sampling variation. Note that while estimates generally vary from one sample to another, the population mean (μ) and variance (σ^2) are fixed values (the parameters in the

7:24 # Title

Console Terminal × Background Jobs ×

R 4.3.2 · C:\Ziv_Temp_2023\Workshop_Vietnam_2025\ShortCourse\OnlineBook/ ↵
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

Environment History Connections Tutorial

To Console To Source

```
# }  
# scale_fill_discrete <- function(palname = pal, ...){  
# scale_fill_brewer(palette = palname, ...)  
# }  
1-pnorm(abs(z),0,1)  
x<-seq(from=-3,to=3,length=1000)  
dx<-dnorm(x,0,1)  
plot(x,dx,type="l",ylab="f(x)=N(0,1)")  
lines(c(z,z),c(0,0.4),lwd=3,col=2)  
1-pnorm(abs(z),0,1)
```

Files Plots Packages Help Viewer Presentation

Install Update

Name	Description	Version
User Library		
abind	Combine Multidimensional Arrays	1.4-8
airports	Data on Airports	0.1.0
aplore3	Datasets from Hosmer, Lemeshow and Sturdivant, "Applied Logistic Regression" (3rd Ed., 2013)	0.9
aplot	Decorate a 'ggplot' with Associated Information	0.2.3
askpass	Password Entry Utilities for R, Git, and SSH	1.2.1
assertthat	Easy Pre and Post Assertions	0.2.1
av	Working with Audio and Video in R	0.9.2
babynames	US Baby Names 1880-2017	1.0.1
backports	Reimplementations of Functions Introduced Since R-3.0.0	1.5.0
base64enc	Tools for base64 encoding	0.1-3
beeswarm	The Bee Swarm Plot, an Alternative to Stripchart	0.4.0

The HTML file: R output

New Tab er_prog4_VT_2024-V3.knit +

File C:/Ziv_Temp_2023/Workshop_Vietnam_2025/ShortCourse/OnlineBook/er_prog4_VT_2024-V3.html

uhasselt.be bookmarks All Bookmarks

1. Variability in estimates
1.2 A point estimate for the population parameter
1.3 Example: the wind speed in the airquality dataset
1.4 The variability of the sample mean
1.5 The sampling distribution for the mean
1.6 The Central Limit Theorem
2. Standard error of the mean
3. Confidence intervals
4. Hypothesis testing
5 Hypothesis testing and confidence Intervals
6 Decision error (Type I and Type II error)

[1] 153 6

The first 6 lines of the dataset are shown below.

[1] 153 6

Ozone Solar.R Wind Temp Month Day

	Ozone	Solar.R	Wind	Temp	Month	Day
## 1	41	190	7.4	67	5	1
## 2	36	118	8.0	72	5	2
## 3	12	149	12.6	74	5	3
## 4	18	313	11.5	62	5	4
## 5	NA	NA	14.3	56	5	5
## 6	28	NA	14.9	66	5	6

The variable Wind is the average wind speed in miles per hour at 0700 and 1000 hours at LaGuardia Airport. The mean wind speed is $\hat{\mu} = \bar{x} = 9.95$ and the sample standard deviation is $\hat{\sigma} = s = 3.52$. This sample mean is a point estimate of the population mean. If a different random sample of 153 days were taken, the new sample mean would likely be different as a result of sampling variation. Note that while estimates generally vary from one sample to another, the population mean (μ) and variance (σ^2) are fixed values (the parameters in the population).

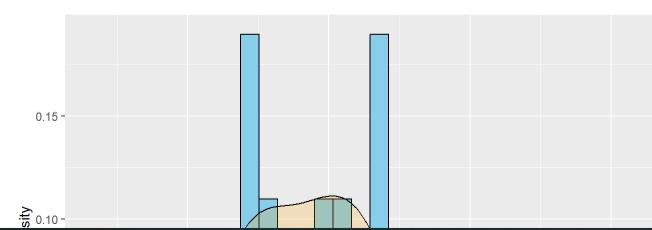
[1] 9.957516

[1] 3.523001

Visualization

Figure 1 and 2 show a histogram and boxplot of the data. Both Figures indicate a slightly skewed distribution to the right.

Show



Windows Type here to search  7:24 ENG 19/06/2025 27

The HTML file: formulas and equations...

RPubs - Chapter 4 (Temp) rpubs.com/zivshkedy/1323673

Topics | BasicInf2

uhasselt.be bookmarks

Sign in Register

RPubs by RStudio

1. Variability in estimates
2. Standard error of the mean
3. Confidence intervals
4. Hypothesis testing
4.1 The Formal Approach to Hypothesis Testing
4.2 Example: wind speed in New York 1973
4.3 Example: the NHANES data set analysis of the number of sleep hours per night
5 Hypothesis testing and confidence intervals
6 Decision error (Type I and Type II error)

Figure 16: Histogram and violin of a random sample for the BMI (n=50).

A 95% C.I for the mean BMI

The standard error of the sample mean is equal to

$$\frac{s}{\sqrt{n}} = \sqrt{\frac{87.67340}{50}} = 1.3241.$$

[1] 1.324186

For a 95% confidence level, $\alpha = 0.025$ and $Z_{\alpha} = 1.96$. The confidence interval is given by

$$\bar{x} \pm m = \bar{x} \pm 1.96 \times SE.$$

For the BMI sample we get

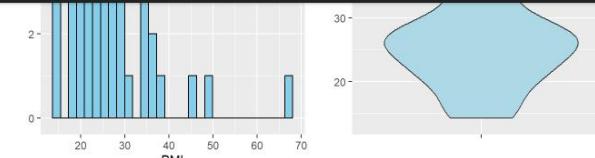
[1] 24.17259 29.36341

This implies that with a confidence of 95% we can conclude that the true mean in the population, μ , is between 24.16 to 29.36. Note that the true mean $\mu = 26.78$.

Visualization

The histogram of the sample with a 95% confidence interval is shown in Figure 17.

BMI



Comments (-) Share Hide Toolbars

Chapter 4 (Temp) by Ziv Shkedy Last updated 21 days ago

Type here to search

8:43 10/07/2025

The Rmd file: formulas and equations...

The screenshot shows the RStudio interface with an Rmd file open. The code editor window displays R code, with several lines containing mathematical formulas in TeX style being highlighted by red boxes and arrows. A callout bubble points to one of these formulas with the text "Math formula in Tex style". The R console window at the bottom shows standard R startup messages and workspace details.

Source Visual

```
632  
633 ## A 95% C.I for the mean BMI  
634 The standard error of the sample mean is equal to  
635  
636 
$$\frac{s}{\sqrt{n}} = \sqrt{\frac{87.67349}{50}} = 1.3241$$
  
637  
638 ## r, echo=TRUE, message FALSE, warning FALSE  
639 n<-length(x.bmi)  
640 SD.x<-sqrt(var(x.bmi))  
641 SE<-SD.x/sqrt(n)  
642 SE  
643  
644  
645 For a 95% confidence level,  $\alpha=0.025$  and  $Z_{\alpha/2}=1.96$ . The confidence interval is given by  
646  
647 
$$\bar{x} \pm s \times Z_{\alpha/2}$$
  
648  
649 For the BMI sample we get  
650  
651 # Title :
```

Environment History Connections Tutorial

```
# }  
# scale_fill_discrete <- function(palname = pal, ...){  
# scale_fill_brewer(palette = palname, ...)  
# }  
1-pnorm(abs(z),0,1)  
x<-seq(from=-3,to=3,length=1000)  
dx<-dnorm(x,0,1)  
",ylab="f(x)=N(0,1)")  
0.4),lwd=3,col=2)  
1)
```

Files Plots Packages Help Viewer Presentation

User Library

Name	Description	Version
abind	Combine Multidimensional Arrays	1.4-8
airports	Data on Airports	0.1.0
aplore3	Datasets from Hosmer, Lemeshow and Sturdivant, "Applied Logistic Regression" (3rd Ed., 2013)	0.9
aplot	Decorate a 'ggplot' with Associated Information	0.2.3
askpass	Password Entry Utilities for R, Git, and SSH	1.2.1
assertthat	Easy Pre and Post Assertions	0.2.1
av	Working with Audio and Video in R	0.9.2
babynames	US Baby Names 1880-2017	1.0.1
backports	Reimplementations of Functions Introduced Since R-3.0.0	1.5.0
base64enc	Tools for base64 encoding	0.1-3
beeswarm	The Bee Swarm Plot, an Alternative to Stripchart	0.4.0

R 4.3.2 · C:/Ziv_Temp_2023/Workshop_Vietnam_2025/ShortCourse/OnlineBook/

Type 'demo()' for some demos, 'help()' for on-line help, or 'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[Workspace loaded from C:/Ziv_Temp_2023/Workshop_Vietnam_2025/ShortCourse/OnlineBook/.RData]

>

Type here to search

The HTML file: formulas and equations...

RPubs - Chapter 4 (Temp) rpubs.com/zivshkedy/1323673

Topics | BasicInf2

uhasselt.be bookmarks

Sign in Register

RPubs by RStudio

1. Variability in estimates
2. Standard error of the mean
3. Confidence intervals
4. Hypothesis testing
4.1 The Formal Approach to Hypothesis Testing
4.2 Example: wind speed in New York 1973
4.3 Example: the NHANES data set analysis of the number of sleep hours per night
5 Hypothesis testing and confidence intervals
6 Decision error (Type I and Type II error)

Figure 16: Histogram and violin of a random sample for the BMI (n=50).

A 95% C.I for the mean BMI

The standard error of the sample mean is equal to

$$\frac{s}{\sqrt{n}} = \sqrt{\frac{87.67340}{50}} = 1.3241.$$

[1] 1.324186

For a 95% confidence level $\alpha = 0.025$ and $Z_\alpha = 1.96$. The confidence interval is given by

$$\bar{x} \pm m = \bar{x} \pm 1.96 \times SE.$$

For the BMI sample we get

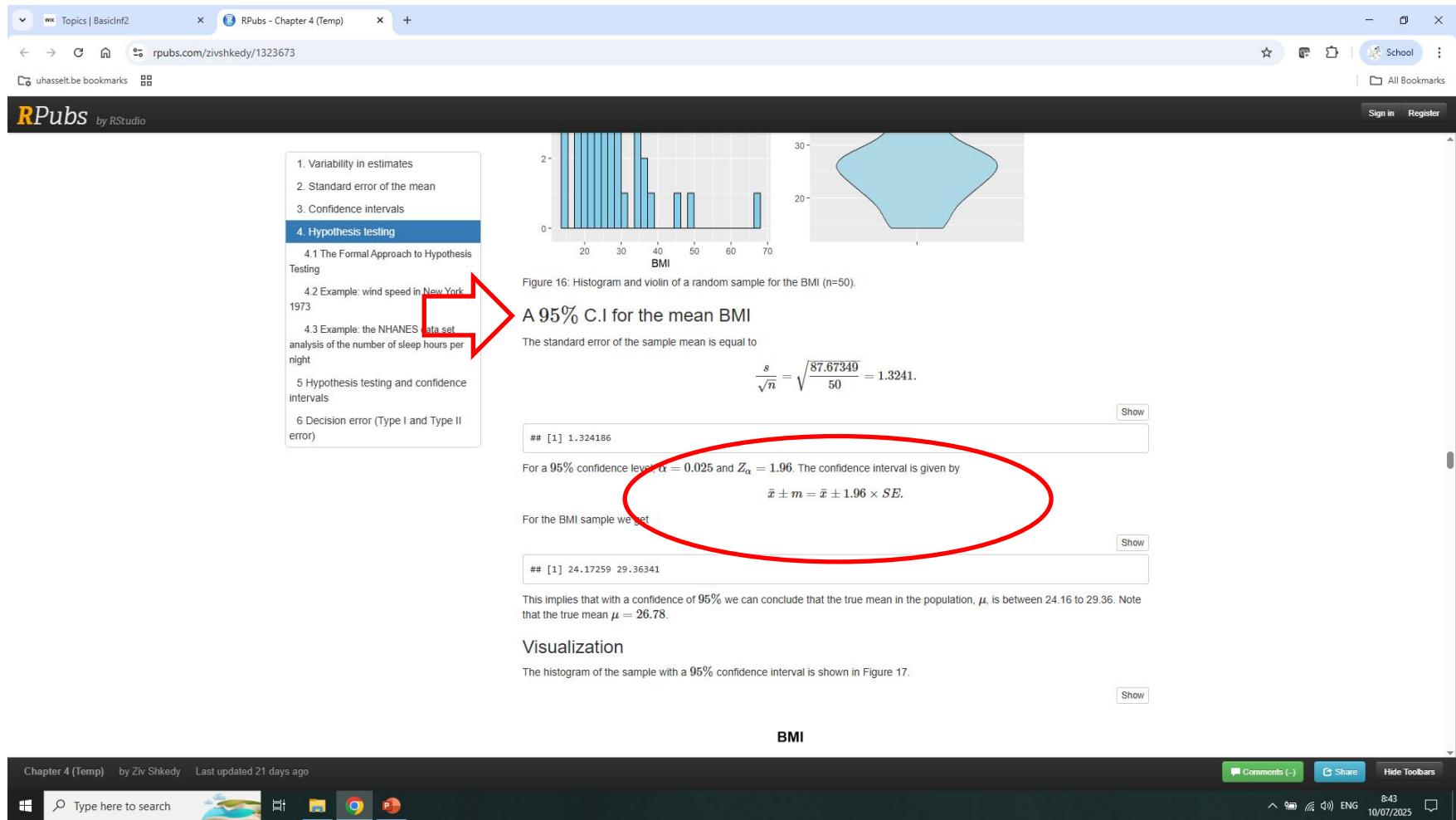
[1] 24.17259 29.36341

This implies that with a confidence of 95% we can conclude that the true mean in the population, μ , is between 24.16 to 29.36. Note that the true mean $\mu = 26.78$.

Visualization

The histogram of the sample with a 95% confidence interval is shown in Figure 17.

BMI



Comments (-) Share Hide Toolbars

Chapter 4 (Temp) by Ziv Shkedy Last updated 21 days ago

Type here to search

8:43 10/07/2025

The Rmd file: formulas and equations...

The screenshot shows the RStudio interface with several windows open:

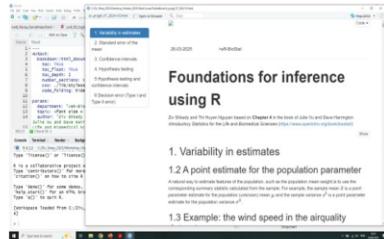
- Code Editor (Source tab):** Displays R code for calculating standard error (SE) and confidence intervals. A red circle highlights the formula $\bar{x} \pm 1.96 \times SE$ in the code.
- Environment Tab:** Shows the R environment with various objects and their values.
- Plots Tab:** Shows a normal distribution plot with a mean line at 0 and a shaded confidence interval centered around 0.
- Packages Tab:** Shows the User Library with installed packages like abind, airports, aplore3, aplot, askpass, assertthat, av, babynames, backports, base64enc, and beeswarm.
- Console Tab:** Shows the R console output, including the version (R 4.3.2), workspace loading, and help messages.
- Terminal Tab:** Shows the terminal output, which is currently empty.
- Background Jobs Tab:** Shows no active background jobs.

Math formula in Tex style

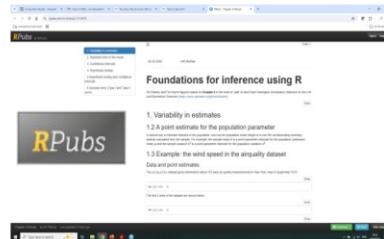
$$\bar{x} \pm 1.96 \times SE$$

Step 2: how to upload the HTML online ?

Offline to Online



- Laptop.



- Online (HTML).
 - Everybody can see and use.

Rpubs

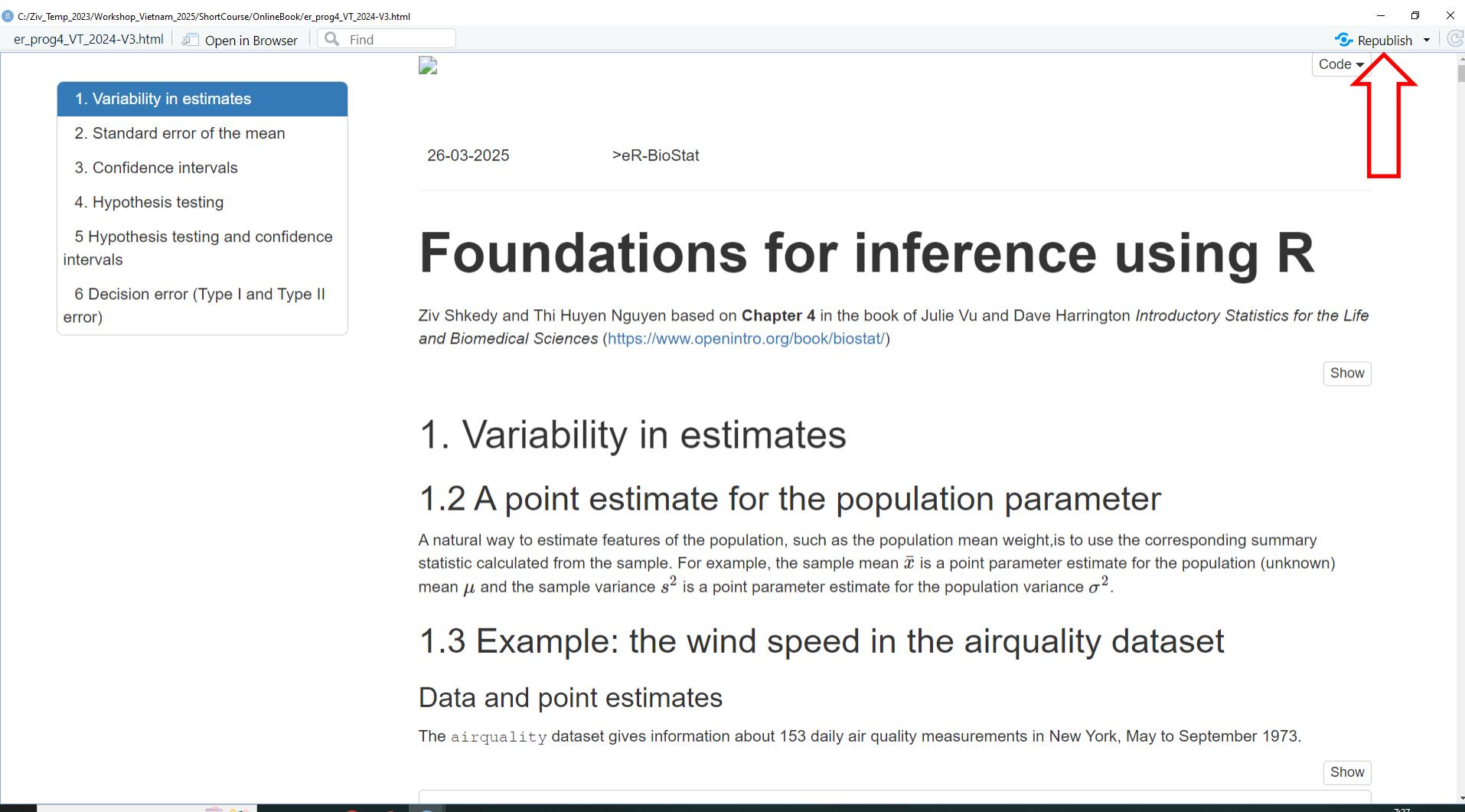
- For the next step, an account in RPubs is needed.

Knitting the file

The screenshot shows the RStudio interface with the following components:

- Top Bar:** Contains the RStudio logo, menu items (File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help), and a "Project: (None)" indicator.
- Source Editor:** Displays an R Markdown code block. A red arrow points from the "Knit" button in the toolbar to the code editor area. The code includes YAML front matter and R code for generating a plot of a normal distribution.
- Console:** Shows the R environment and a welcome message from R 4.3.2.
- Environment Tab:** Shows the user library with various R packages listed.
- Bottom Bar:** Includes a search bar, pinned application icons (File Explorer, Task View, Google Chrome, Microsoft Edge, File History, and R), and system status indicators.

Upload the HTML to RPubs



C:/Ziv_Temp_2023/Workshop_Vietnam_2025/ShortCourse/OnlineBook/er_prog4_VT_2024-V3.html
er_prog4_VT_2024-V3.html | Open in Browser | Find | Republish | Code ▾

1. Variability in estimates
2. Standard error of the mean
3. Confidence intervals
4. Hypothesis testing
5 Hypothesis testing and confidence intervals
6 Decision error (Type I and Type II error)

26-03-2025 >eR-BioStat

Foundations for inference using R

Ziv Shkedy and Thi Huyen Nguyen based on **Chapter 4** in the book of Julie Vu and Dave Harrington *Introductory Statistics for the Life and Biomedical Sciences* (<https://www.openintro.org/book/biostat/>)

Show

1. Variability in estimates

1.2 A point estimate for the population parameter

A natural way to estimate features of the population, such as the population mean weight, is to use the corresponding summary statistic calculated from the sample. For example, the sample mean \bar{x} is a point parameter estimate for the population (unknown) mean μ and the sample variance s^2 is a point parameter estimate for the population variance σ^2 .

1.3 Example: the wind speed in the airquality dataset

Data and point estimates

The `airquality` dataset gives information about 153 daily air quality measurements in New York, May to September 1973.

Show

Upload the HTML to RPubs

C:/Ziv_Temp_2023/Workshop_Vietnam_2025/ShortCourse/OnlineBook/er_prog4_VT_2024-V3.html
er_prog4_VT_2024-V3.html | Open in Browser | Find | Republish | Code ▾

1. Variability in estimates
2. Standard error of the mean
3. Confidence intervals
4. Hypothesis testing
5 Hypothesis testing and confidence intervals
6 Decision error (Type I and Type II error)

26-03-2025 >eR-BioStat

Foundations for inference using R

Ziv Shkedy and Biometris

RPubs

This document has already been published on RPubs. You can choose to either update the existing RPubs document, or create a new one.

IMPORTANT: All documents published to RPubs are publicly visible.
You should only publish documents that you wish to share publicly.

1. Variability in estimates
1.2 A point estimate
A natural way to estimate a population parameter is to use the corresponding summary statistic calculated from the sample. For example, the sample mean \bar{x} is a point parameter estimate for the population (unknown) mean μ and the sample variance s^2 is a point parameter estimate for the population variance σ^2 .

1.3 Example: the wind speed in the airquality dataset

Data and point estimates

The airquality dataset gives information about 153 daily air quality measurements in New York, May to September 1973.

Windows Taskbar: Type here to search, File, Settings, Task View, Start, Google Chrome, R, Show, ENG, 7:28, 19/06/2025

Upload the HTML to RPubs

The screenshot shows a web browser window with the URL rpubs.com/zivshkedy. The page displays a list of recent publications by the user 'zivshkedy'. The publications are arranged in a grid:

- Chapter 4 (Temp)** - Published 2 minutes ago
- VDEDA1** - Published 3 days ago
- Chapter 8 (Temp)** - Published 24 days ago
- Chapter 5 (Temp)** - Published 24 days ago

A red arrow points from the 'Published' tab to the 'Recent' tab, indicating that the 'Recent' tab should be selected.

At the bottom of the screen, the Windows taskbar is visible with various icons and the system tray showing the date and time (19/06/2025) and battery level (735).

Upload the HTML to RPubs

The screenshot shows a web browser window with the URL rpubs.com/zivshkedy/1323673. The page is titled "RPubs - Chapter 4 (Temp)" by Ziv Shkedy. The main content is a document titled "Foundations for inference using R". The sidebar on the left lists sections: 1. Variability in estimates, 2. Standard error of the mean, 3. Confidence intervals, 4. Hypothesis testing, 5 Hypothesis testing and confidence intervals, and 6 Decision error (Type I and Type II error). The document header includes the date 26-03-2025 and the author >eR-BioStat. The main title "Foundations for inference using R" is displayed prominently. Below it, a note states: "Ziv Shkedy and Thi Huyen Nguyen based on Chapter 4 in the book of Julie Vu and Dave Harrington *Introductory Statistics for the Life and Biomedical Sciences* (<https://www.openintro.org/book/biostat/>)". The document content includes sections on variability in estimates, point estimates for population parameters, an example using the airquality dataset, and data and point estimates. It also shows R code snippets and their outputs.

- Online.
- Everybody can see (and download) the file.

1. Variability in estimates

1.2 A point estimate for the population parameter

A natural way to estimate features of the population, such as the population mean weight, is to use the corresponding summary statistic calculated from the sample. For example, the sample mean \bar{x} is a point parameter estimate for the population (unknown) mean μ and the sample variance s^2 is a point parameter estimate for the population variance σ^2 .

1.3 Example: the wind speed in the airquality dataset

Data and point estimates

The airquality dataset gives information about 153 daily air quality measurements in New York, May to September 1973.

```
## [1] 153   6
```

The first 6 lines of the dataset are shown below.

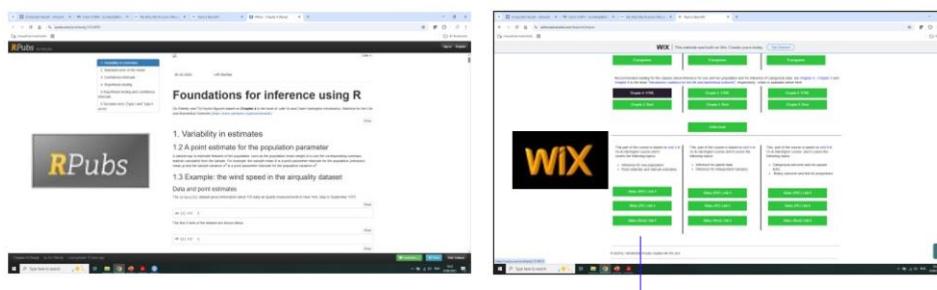
```
## [1] 153   6
```

Comments (-) Share Hide Toolbars

Step 3: how to link the content to a website ?

Online at the website

<https://erbiostat.wixsite.com/basicinf>



- HRML: RPus.
- Rmd: GitHub.

WIX

- For this step, an account in WIX is needed.

Link a content to a website

- Suppose that we want to link the content created for Chapter 4 to a website.
- The user will be able to read the chapter from the website.



- Example: a website for a course in which Chapter 4 is used.

My WIX websites

A website that uses Chapter 4.

Sites
View and manage all your websites in one place.

Create New Folder Create New Site

Filter Search... View Options

Site Name	URL
vdeda1	https://erbiostat.wixsite...
ITP6	https://erbiostat.wixsite...
BasicInf2	https://erbiostat.wixsite...
MOIEL1	https://erbiostat.wixsite...
KARDSA1	https://erbiostat.wixsite...
Erbiostat	https://erbiostat.wixsite...
Newcourse1	https://erbiostat.wixsite...
Rintro	https://erbiostat.wixsite...
ITP5	https://erbiostat.wixsite...
itpb0	https://erbiostat.wixsite...
itpa0	https://erbiostat.wixsite...
ITP4	https://erbiostat.wixsite...
BasicInf1	https://erbiostat.wixsite...
ITP3	https://erbiostat.wixsite...
ITP2	https://erbiostat.wixsite...

Final Hours to Get 50% OFF 01:14:46 Upgrade Now

Offer valid for initial subscription on select Premium Plans purchased during the sale.

The course website

Universiteit Hasselt - UHasselt × | Inbox (7,678) - ziv.shkedy@uhasselt.be × | Wix Website Editor | BasicInfo2 × + editor.wix.com/html/editor/web/renderer/edit/9d2b9308-a82c-48be-a5e5-a918253b8b12?metaSiteId=1765393e-f382-448d-bde2-8a5c007498cd

uhasselt.be bookmarks All Bookmarks

You're now in Preview mode

Upgrade your website to remove Wix ads [Upgrade Now](#)

R
Basic concepts of statistical inference using R
>eR-BioStat

Home About Topics Online tutorials Contact

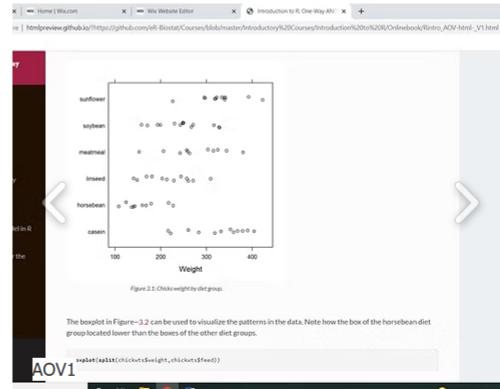
This course consists of 3-4 classes (each of three hours) which are focused on statistical inference at a basic level. The course covers the following topics:

- Inference for continuous variables: one population.
- Inference for numerical variables: two populations.
- Inference for categorical variables: one & two populations.

The three classes can be given as a complete course (at a basic level) or as a part of a specific course in basic statistical analysis (with extra modules in statistical modelling).

The course was developed at an [undergraduate](#) level courses (for both statistician and non statisticians).

Figure 2.1 Chick weight by diet group.



The boxplot in Figure 2.2 can be used to visualize the patterns in the data. Note how the box of the horsebean diet group is located lower than the boxes of the other diet groups.

```
boxplot(weight ~ diet, data = chicks)
```

AOV1

This is an open source course and all source files used to produce the slides are available online (in PP, Tex or Rmd formats).

THE COURSE IS UNDER DEVELOPMENT AND THIS IS NOT THE FINAL VERSION OF THE COURSE.



Type here to search



7:46 19/06/2025

Chapter 4

You're now in Preview mode

Upgrade your website to remove Wix ads [Upgrade Now](#)

basic concepts of inference (one population)

This chapter introduces the important concepts in drawing estimates from samples from one population. We discuss methods of inference and estimation for a population mean and cover three topics:

- Point estimates for a population mean.
- Interval estimates methods.
- Testing hypotheses about the population mean.

External datasets are available in the data repository.

[Slides \(PDF\): Chapter 4](#)

[Slides \(PP\): Chapter 4](#)

[R programm](#)

Inference for continuous data (one population & two populations)

This chapter discusses the settings where The methods discussed in the previous class are applied to multiple populations. This include:

- The analysis of paired observations
- The comparison of two or more independent groups.

Both normal and t distribution are used for both interval estimation and inference.

External datasets are available in the data repository.

[Slides \(PDF\): Chapter 5](#)

[Slides \(PP\): Chapter 5](#)

[R programm](#)

Inference for binary and categorical data

In this chapter, methods for estimation and inference are extended to categorical data, such as binomial proportions or data in two-way tables. We cover the following topics:

- Tests and estimation methods for proportion in one population.
- Tests and estimation methods for proportions in two populations.
- Chi-square tests.

External datasets are available in the data repository.

[Slides \(PDF\): Chapter 8](#)

[Slides \(PP\): Chapter 8](#)

[R programm](#)

Recommended reading for the classes about inference for one and two population and for inference of categorical data are [Chapter 4](#), [Chapter 5](#) and [Chapter 8](#) in the book "Introductory statistics for the life and biomedical sciences", respectively, which is available online here:

[Chapter 4: HTML](#)

[Chapter 4: Rmd](#)

[Chapter 5: HTML](#)

[Chapter 5: Rmd](#)

[Chapter 8: HTML](#)

[Chapter 8: Rmd](#)

7:46
19/06/2025
ENG

44

Chapter 4 in HTML

Universiteit Hasselt - UHasselt × | Inbox (7,678) - ziv.shkedy@uhasselt.be × | wx Setup | Wix.com × | wx Wix Website Editor | BasicInfo2 × +

editor.wix.com/html/editor/web/renderer/edit/9d2b9308-a82c-48be-a5e5-a918253b8b12?metaSiteId=1765393e-f382-448d-bde2-8a5c007498cd

uhasselt.be bookmarks All Bookmarks

You're now in Preview mode

Upgrade your website to remove Wix ads [Upgrade Now](#)

Recommended reading for the classes about inference for one and two population and for inference of categorical data are [Chapter 4](#), [Chapter 5](#) and [Chapter 8](#) in the book "Introductory statistics for the life and biomedical sciences", respectively, which is available online here:

[Chapter 4: HTML](#) [Chapter 5: HTML](#) [Chapter 8: HTML](#)

[Chapter 4: Rmd](#) [Chapter 5: Rmd](#) [Chapter 8: Rmd](#)

[Online book](#)

This part of the course is based on [unit 4](#) in Vu & Harrington course and it covers the following topics:

- Inference for one population.
- Point estimate and interval estimates.

[Slides \(PDF\): Unit 4](#) [Slides \(PP\): Unit 4](#) [Slides \(Rmd\): Unit 4](#)

This part of the course is based on [unit 5](#) in Vu & Harrington course and it covers the following topics:

- Inference for paired data.
- Inference for independent samples.

[Slides \(PDF\): Unit 5](#) [Slides \(PP\): Unit 5](#) [Slides \(Rmd\): Unit 5](#)

This part of the course is based on [unit 8](#) in Vu & Harrington course and it covers the following topics:

- Categorical outcome and chi-square tests.
- Binary outcome and test for proportions.

[Slides \(PDF\): Unit 8](#) [Slides \(PP\): Unit 8](#) [Slides \(Rmd\): Unit 8](#)

© 2020 by >eR-BioStat. Proudly created with Wix.com

https://rpubs.com/zivshkedy/1314878

Type here to search

7:47 19/06/2025 ENG

Chapter 4 on RPubs

Universiteit Hasselt - UHasselt | Inbox (7,679) - ziv.shkedy@uhasselt.be | wx Setup | Wix.com | wx | Wix Website Editor | BasicInfo2 | RPubs - Chapter 4 (Temp)

rpubs.com/zivshkedy/1323673

uhasselt.be bookmarks | All Bookmarks

RPubs by RStudio Sign in Register

1. Variability in estimates

2. Standard error of the mean

3. Confidence intervals

4. Hypothesis testing

5 Hypothesis testing and confidence intervals

6 Decision error (Type I and Type II error)

26-03-2025 >eR-BioStat

Foundations for inference using R

Ziv Shkedy and Thi Huyen Nguyen based on **Chapter 4** in the book of Julie Vu and Dave Harrington *Introductory Statistics for the Life and Biomedical Sciences* (<https://www.openintro.org/book/biostat/>)

Show

1. Variability in estimates

1.2 A point estimate for the population parameter

A natural way to estimate features of the population, such as the population mean weight, is to use the corresponding summary statistic calculated from the sample. For example, the sample mean \bar{x} is a point parameter estimate for the population (unknown) mean μ and the sample variance s^2 is a point parameter estimate for the population variance σ^2 .

1.3 Example: the wind speed in the airquality dataset

Data and point estimates

The airquality dataset gives information about 153 daily air quality measurements in New York, May to September 1973.

Show

```
## [1] 153   6
```

The first 6 lines of the dataset are shown below.

Show

```
## [1] 153   6
```

Show

Comments (-) Share Hide Toolbars

Course website

Universiteit Hasselt - UHasselt | Inbox (7,678) - ziv.shkedy@uhasselt.be | wx Setup | Wix.com | wx Wix Website Editor | BasicInf2

uhasselt.be bookmarks All Bookmarks

WIX Site Settings Dev Mode Hire a Professional Help 50% OFF Upgrade Save Preview Publish

Page: Topics basicinf2.com is available. Connect Your Domain

Recommended reading for the classes about inference for one and two population and for inference of categorical data are [Chapter 4](#), [Chapter 5](#) and [Chapter 8](#) in the book "Introductory statistics for the life and biomedical sciences", respectively, which is available online here:

[Chapter 4: HTML](#) [Chapter 5: HTML](#) [Chapter 8: HTML](#)

[Chapter 4: Rmd](#) [Chapter 5: Rmd](#) [Chapter 8: Rmd](#)

[Online book](#)

This part of the course is based on [unit 4](#) in Vu & Harrington course and it covers the following topics:

- Inference for one population.
- Point estimate and interval estimates.

[Slides \(PDF\): Unit 4](#) [Slides \(PP\): Unit 4](#) [Slides \(Rmd\): Unit 4](#)

This part of the course is based on [unit 5](#) in Vu & Harrington course and it covers the following topics:

- Inference for paired data.
- Inference for independent samples.

[Slides \(PDF\): Unit 5](#) [Slides \(PP\): Unit 5](#) [Slides \(Rmd\): Unit 5](#)

This part of the course is based on [unit 8](#) in Vu & Harrington course and it covers the following topics:

- Categorical outcome and chi-square tests.
- Binary outcome and test for proportions.

[Slides \(PDF\): Unit 8](#) [Slides \(PP\): Unit 8](#) [Slides \(Rmd\): Unit 8](#)

© 2020 by >eR-BioStat. Proudly created with Wix.com

Create the link to RPubs

The screenshot shows a Wix website editor interface. The top navigation bar includes tabs for 'Site', 'Settings', 'Dev Mode', 'Hire a Professional', and 'Help'. A banner at the top says 'basicinf2.com is available. Connect Your Domain'. The main content area displays a grid of green cards representing course materials. A red arrow points from a green card labeled 'R programm' to a 'Change Text' tool icon. The cards are organized into three columns:

- Column 1:** Contains cards for 'Slides (PDF): Chapter 4', 'Slides (PP): Chapter 4', 'R programm', 'Chapter 4: HTML', and 'Chapter 4: Rmd'.
- Column 2:** Contains cards for 'Slides (PDF): Chapter 5', 'Slides (PP): Chapter 5', 'R programm', 'Chapter 5: HTML', and 'Chapter 5: Rmd'.
- Column 3:** Contains cards for 'Slides (PDF): Chapter 8', 'Slides (PP): Chapter 8', 'R programm', 'Chapter 8: HTML', and 'Chapter 8: Rmd'.

Below the grid, there is a section titled 'Online book' with three paragraphs describing the course's focus on different units:

- This part of the course is based on [unit 4](#) in Vu & Harrington course and it covers the following topics:
 - Inference for one population.
 - Point estimate and interval estimates.
- This part of the course is based on [unit 5](#) in Vu & Harrington course and it covers the following topics:
 - Inference for paired data.
 - Inference for independent samples.
- This part of the course is based on [unit 8](#) in Vu & Harrington course and it covers the following topics:
 - Categorical outcome and chi-square tests.
 - Binary outcome and test for proportions.

At the bottom of the page, there are sections for 'Slides (PDF): Unit 4', 'Slides (PP): Unit 4', and 'Slides (Rmd): Unit 4' on the left; 'Slides (PDF): Unit 5', 'Slides (PP): Unit 5', and 'Slides (Rmd): Unit 5' in the middle; and 'Slides (PDF): Unit 8', 'Slides (PP): Unit 8', and 'Slides (Rmd): Unit 8' on the right. The bottom navigation bar includes icons for search, file, browser, and other system functions, along with a timestamp '7:49 19/06/2025'.

Create the link to RPubs

Screenshots showing the Wix Website Editor interface and a modal dialog for linking to RPubs.

The Wix editor interface shows a grid of green cards representing course materials. A modal dialog is open, prompting the user to "What do you want to link to?" with options like "None", "Page", "Web address", etc. The "Web address" option is selected, and the URL input field contains <https://rpubs.com/zivshkedy/1323673>.

Modal Dialog Content:

- What do you want to link to?
- None
- Page
- Web address
- Section or anchor
- Top/bottom of page
- Document
- Email
- Phone number
- Lightbox

What's the web address (URL)?

How does it open?
 New window
 Current window

Advanced SEO Settings

Red Text Overlay:

<https://rpubs.com/zivshkedy/1323673>

The Wix editor interface includes a sidebar with icons for Site, Settings, Dev Mode, Hire a Professional, Help, and a search bar at the bottom.

Which account do you need for this system ?

To go local....

	Platform	Why
1	RPubs 	To upload HTML online
2	GitHub 	To upload and store PDF, HTML, R programs etc online
3	WIX 	To develop a website
4	R studio/ R markdown 	To develop and produce content