



UNIVERSITY
OF OSLO



OSLO UNIVERSITY HOSPITAL



How can R and Quarto help?

A one year recap on teaching statistics to medical students

Chi Zhang, PhD

OCBE (Oslo Center for Biostatistics and Epidemiology)
@ University of Oslo, Oslo University Hospital

2024-06-13 R/Medicine 2024

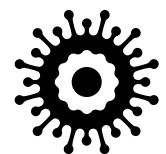
About me



Lead developer of MF9130 course website with R+quarto+webR



50% lecturer in biostatistics @ Faculty of Medicine, University of Oslo, Norway



Statistician turned R developer - RWD, EHR and large public health registries

Website: <https://andreaczhang.github.io/>

LinkedIn: [Chi Zhang](#)

GitHub: [@andreaczhang](#)

Email: chi.zhang@medisin.uio.no



Carpentry@UiO



This is a recap on ...

A story of

Zero to Hero

(sort of)



quarto®



GitHub



MF9130: introduction to statistics

Faculty of Medicine, University of Oslo

PhD students at UiO (and/or hospital) + other Norwegian universities. **Many have their own data - motivated!**

Background: **clinicians, biologists, lab researcher, psychologists, nutritionists, ...**

Statistical competency: very basic - basic: over 75%

Software: none, some SPSS / Stata.

Few know R, basic IT competency

MF9130: 8 days intensive course, offered 3 times per year (8/30 credits required for PhD)

Topics: probability and distributions, sensitivity/specificity, commonly used hypothesis tests, regression, survival analysis

How it has always been:

Lectures and exercise sessions

Focus on theory and hand-calculation, no emphasis on software / data analysis

Since 2023 Spring class: use **R**, teach more **data skills**; 2024: added **WebR**

Change I: course website

R + Quarto + GitHub + WebR

MF9130E - V24

HomeGet StartedScheduleR Lab and CodeAbout

MF9130E - Introductory course in Statistics

Welcome!

You are on the course website for **MF9130E - Introductory Course in Statistics** at the Faculty of Medicine, University of Oslo.

On this page

Welcome!

Preparation

Schedule and course material

ocbe-uio / teaching_mf9130e

Type to search

>_+⌚🔗📧👤

<> Code⌚ Issues🔗 Pull requests⌚ Actions📁 Projects📖 Wiki🛡 Security📈 Insights⚙ Settings

🔴 teaching_mf9130e (Public)

📌 Edit Pins👁 Watch 0🔗 Fork 9⭐ Starred 13

🔗 main🔗 4 Branches🔗 0 Tags🔍 Go to file🔗 +<> Code⚙

👤 andreaczhang added link ✓7983eff · last month🕒 139 Commits

_extensions

add demo of webr

7 months ago

docs

added link

last month

About

Course material for MF9130E

🔗 ocbe-uio.github.io/teaching_mf9130e/

📖 Readme

📄 MIT license

📈 Activity

Open access to students and instructors

Lecture notes, data and code more **organized** than university portal

Collaboration: can be **reused** with minimal effort

Change I: course website

R + Quarto + GitHub + **WebR**

💡 Formula: mean

The mean of data $X = (x_1, x_2, \dots, x_n)$, $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$

```
# enter the data
heart <- c(11.5, 14.75, 13.75, 10.5, 14.75,
          13.5, 10.75, 9.5, 11.75, 12,
          10.5, 11.75, 10, 14.5, 12,
          11, 14, 15, 11.5, 10.25)
```

You can either compute the sample mean by summing each data point, and divide by command `mean()`.

```
# compute the mean
sum_heart <- 11.5 + 14.75 + 13.75 + 10.5 + 14.75 +
  13.5 + 10.75 + 9.5 + 11.75 + 12 +
  10.5 + 11.75 + 10 + 14.5 + 12 +
  11 + 14 + 15 + 11.5 + 10.25
```

```
# this is the sum
sum_heart
```

```
[1] 243.25
```

```
# sample size: 20
n <- 20
# if we do not know the size, can find out with length(heart)
sum_heart/n
```

```
[1] 12.1625
```

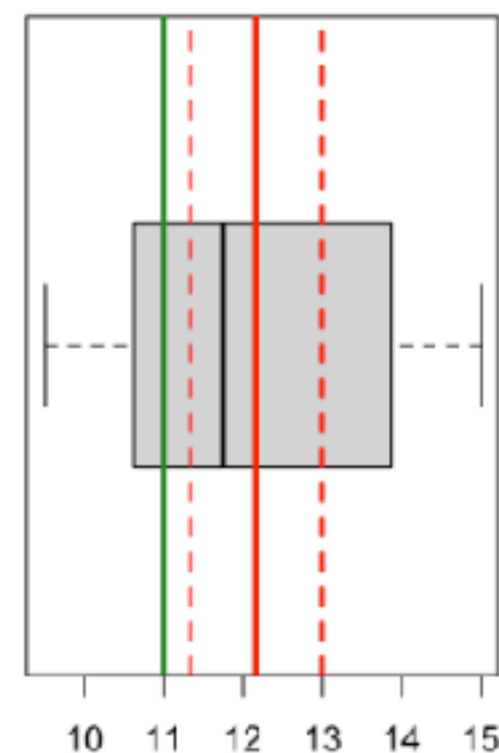
```
# formula: sum(heart)/length(heart)
mean(heart) # should be the same as above
```

```
[1] 12.1625
```

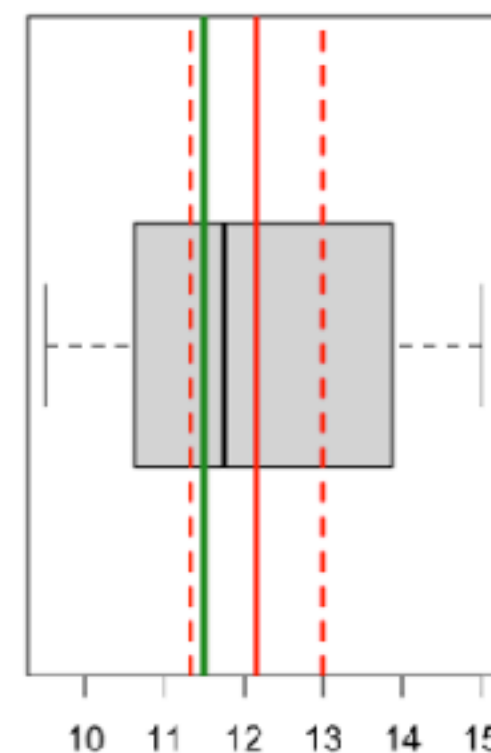
```
# boxplot
par(mfrow = c(1, 2))
# compare with 11
boxplot(heart, horizontal = T, main = 'Compare with mean = 11')
abline(v = mean(heart), col = 'red', lwd = 3)
abline(v = c(ci_lower, ci_upper), col = 'red', lwd = 2,
      lty = 'dashed')
abline(v = 11, col = 'forestgreen', lwd = 3)

# compare with 11.5
boxplot(heart, horizontal = T, main = 'Compare with mean = 11.5')
abline(v = mean(heart), col = 'red', lwd = 3)
abline(v = c(ci_lower, ci_upper), col = 'red', lwd = 2,
      lty = 'dashed')
abline(v = 11.5, col = 'forestgreen', lwd = 3)
```

Compare with mean = 11



Compare with mean = 11.5



Learn by **verification**
formula vs preset function/pkg

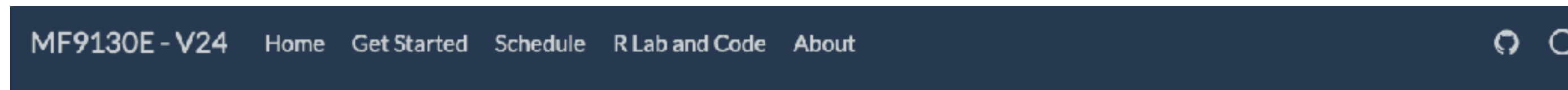
Learn by **visualization**

Extremely useful for teaching
statistical concepts:

summary statistics, p-values, t-
tests etc

Change I: course website (**NEW 2024**)

R + Quarto + GitHub + **WebR**



Binomial and Normal distribution

</> Code

WEBR STATUS

● Ready!

On this page you will find some examples and code for probability distributions.

① Note

In your take-home exam, you will NOT be tested on programming (e.g. simulation) or theoretical results. Content on this page is to help you understand the math behind topics in the next days.

Randomness and simulation (optional)

A random sample is **random**. For example, if you throw a 6 sided dice twice, you will possibly have different outcomes. This process can be **simulated** with the line below.

You can copy and paste this line to your R console, and run it a few times. Alternatively, click on **Run Code** button below a few times. You should see different results every time.

▶ Run Code

```
1 # take two numbers from 1,2,3,4,5,6
2 # replace = T means the same number can appear in both first and second
3 sample(1:6, size = 2, replace = T)
```

On this page

Randomness and simulation (optional)

Binomial distribution

Normal distribution

Bar plot

Example: birth weight

Binomial vs Normal distribution

WebR extension has made it more interactive to demonstrate certain concepts such as **randomness and distribution**

Students can try the code in the browser at home

Try it out [here](#)

Change II: Guided R lab sessions

How we improved

We kept the good things:

Live coding demo - the Carpentry way

Sticky notes system (depends on helper capacity)

An complete example of data analysis (**data import, EDA, tests, interpret**)

Give sufficient time to practice and troubleshoot

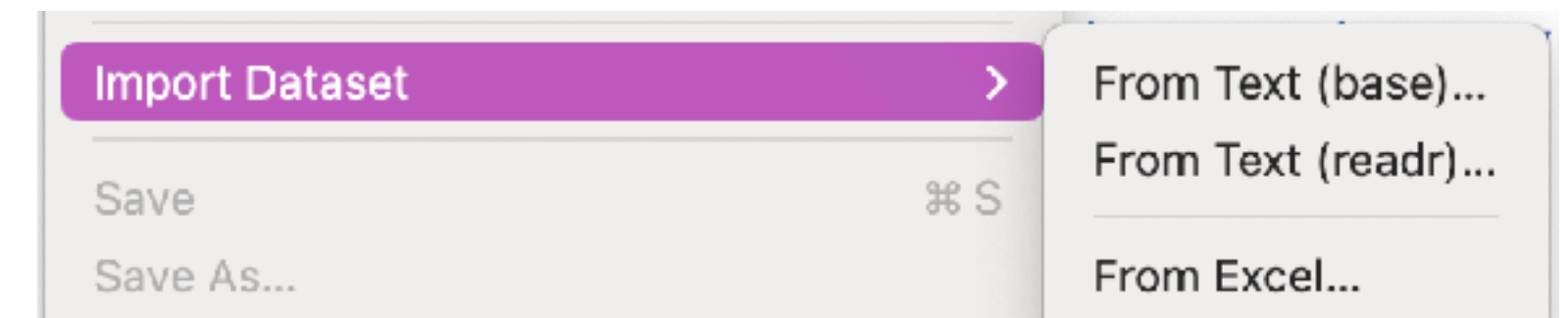
Some **NEW practices**:

Use **R project**, avoid working directory config

Use RStudio **point-and-click for data import**, rather than 'read_csv(path)'

Less ambitious, use **few but good datasets**

Repeat the basic workflow many times



How did it go?

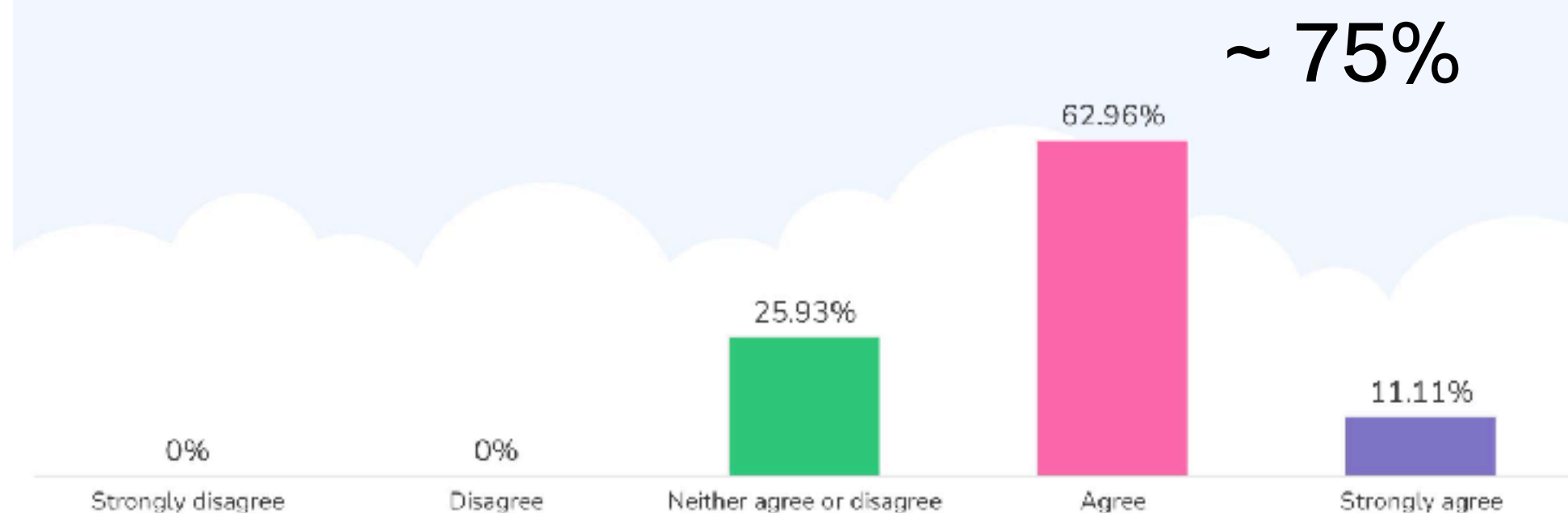
Based on student survey

By the end of day 4 (first week), **every one** can

- load a dataset, extract the target variable
- make a histogram, carry out a test

More **confidence** and interests in statistics

You know how to apply the statistical concepts discussed in the course to your own research



Generally **positive** feedback:

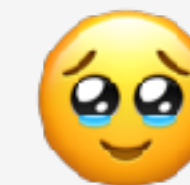


More organized course material,

Skills taught are very **useful for their own work**

Interactive (live demo) lab sessions particularly helpful

Negative: “R is not user friendly”



Less negative in 2024, possibly because students who do not like coding choose the alternative course (in stata)

Find out more

Zero to Hero



MF9130E course website

https://ocbe-uio.github.io/teaching_mf9130e/

Zero to Quarto Workshop (R Ladies Zurich 2023)

<https://andreaczhang.github.io/workshop-02quarto/>

qtwAcademic

an R package to get you started building a quarto website, tailored for academics



Let's chat!

Website: <https://andreaczhang.github.io/>

LinkedIn: [Chi Zhang](#)

GitHub: [@andreaczhang](#)

Email: chi.zhang@medisin.uio.no