Lecture 9

Sampling

agenda

Announcements	HW revisions; “Homework” this week; Midterm 1 on Tuesday, October 3rd! Rmd recommendations.
Mini-lecture (will not be on exam)	Creating your own simulation
To R	Simulation activity in groups
Midterm 1	Discuss topics covered on exam, format, revisions

homework 1 comments

- READ THE DIRECTIONS!!!
- When asked to explain in a few sentences, the sentences should be complete
  - You may need to check spelling/grammar (just use Google docs)
-

I need help :)

I need help catching up on hw feedback and I thought of a cool way to do it (well, I think it's cool).

I am going to post the solutions to homework 3 tonight. I want you to go through your own solutions, and make comments/notes/provide helpful feedback as if you were looking at somebody else's work.

Use the EMRN rubric to assess your assignment.

Rmd recommendations

- Do not run help() or ? in Rmd
- Make sure the line is not getting cut off in pdf (sometimes removing the tab helps)
- Keep it clean and organized
  - do not run unnecessary code
  - load all libraries ONCE at the beginning
- Print the entire dataset (nobody did this, just saying though)
- PLEASE please please set `message = FALSE` and `warning = FALSE` in setup chunk
- When creating tables, use the same number of significant digits for all statistics
-

R notes

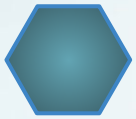
- Do not delete rows of a dataset just because there is an NA, instead:
  - Use na.rm = FALSE in mean/sd functions
  - “temporarily” delete inside a function
- Almost always, the id column is not necessary to plot/take summary statistics about
- You do not need to create a factor, then recreate a factor but add labels
- Utilize comments!
-

Simulations

Some useful functions:

- sample functions
- r____ functions
- for loops

sample function (base)



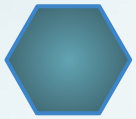
sample()

takes a sample of the specified size from the elements of x

```
sample(x, size, replace = FALSE,
       prob = NULL,...)
```

- x = vector of elements to choose from
- size = number of items to choose
- replace = sample with replacement?
- prob = vector of probability weights for each element being sampled (x)

sample function (tidyverse)



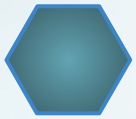
dplyr::sample_n()

takes a sample of the specified size from the rows of a tbl

```
sample_n(tbl, size, replace = FALSE,
        prob = NULL,...)
```

- tbl = data frame
- size = number of rows to select
- replace = sample with replacement?
- weight = vector of probability weights for each element being sampled (weights are standardized to sum to 1)

probability distributions



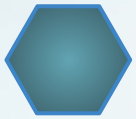
r____()

creates a random sample from a distribution with specific parameters

```
Beta, Binomial, Chi-squared,
Exponential, F, Gamma, Geometric,
Hypergeometric, Normal, ...
```

- arguments change depending on the distribution

for loop



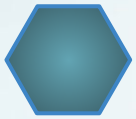
for()

iterating over a sequence (e.g., do *something* for each row in a dataset)

```
for(var in seq){
    do stuff
}
```

- var = typically a letter
  - i, j, k, etc.
- seq = sequence of var
  - e.g., 1:100 or 1:nrow(data)
- do stuff
  - what do you need to iterate

if ... else



if()... else()

if a condition is true,
“do this”, otherwise,
“do that”

```
if(condition1){
    do stuff
}
elseif(condition2) {
    do other stuff
}
else {
    do other other stuff
}
```

- condition: should evaluate to TRUE or FALSE

logical conditions:

- == (equal to)
- != (not equal to)
- <, >, <=, >= less/greater than (or equal to)

To R!

Open sims.R in

About the midterm

01

Format

.Rmd file (like hw)

02

How

Make your own copy of my project on posit.cloud, upload .Rmd and .pdf/.doc

03

When

During class

04

~~Cheating~~

Will not get to include midterm in final portfolio (max you could earn is a C), plus Academic Dishonesty Incident Report filed

05

Revisions

Feedback on exams provided within one week. You will have until the next class to submit rewrites (see next slide)

06

What to study

See midterm 1 topics list in Week 6 module

How “(un)grading" works

01

Take Exam

Tues, Oct 26 during class

Give yourself an E, M, R, N

02

Feedback

Wendy provides feedback, plus

✅: E, M

✘: R, N

03

Revisions

All students can revise midterm if they wish

04

Reflection

First mid-semester reflection (will go in final portfolio)