Using RStudio Cloud for an Introductory Statistics Class

Lori Lewis

Santa Rosa Junior College

What is R?

R is a free, open source software program for statistical analysis.

What is Rstudio?

- Is a free, open source integrated development environment for R.
- Its interface is organized so that the user can clearly view graphs, tables, code and output at the same time.

Why use R cloud?

- Free and open source.

- You can share, make projects, and teach from it.

- There is no installation

- You can access it with a link

How do I use it in my class?

- I use it as a substitute for a graphing calculator
- I created scripts that serve as the basis for small projects that I assign throughout the semester
  - The projects reinforce different topics such as the Central Limit Theorem
  - This gives me an opportunity to introduce analysis of real world data

How to get started?

- Go to https://posit.cloud/
- Create an account

Examples Today

- Barplots
- Histograms
- Mean, median, and standard deviation
- Five number summary and boxplot
- Least-squares regression
- Area under the normal curve

Frequency Distributions and Bar plots

- Let's look at a problem from Statistics: Informed Decisions Using Data by Michael Sullivan III

  Problem: A physical therapist wants to determine types of rehabilitation required by her patients. To do so, she obtains a simple random sample of 30 of her patients and records the body part requiring rehabilitation. See Table 1. Construct a frequency distribution of location of injury.

Table 1		
Back	Back	Hand
Wrist	Back	Groin
Elbow	Back	Back
Back	Shoulder	Shoulder
Hip	Knee	Hip
Neck	Knee	Knee
Shoulder	Shoulder	Back
Back	Back	Back
Knee	Knee	Back
Hand	Back	Wrist

Source: Krystal Catton, student at Joliet Junior College.

Students will make the Frequency Distribution

Body Part	Tally	Frequency
Back	||||| ||||| ||	12
Wrist	||	2
Elbow			1
Hip	||	2
Shoulder	||||	4
Knee	|||||	5
Hand	||	2
Groin			1
Neck			1

How to draw a Histogram

- The following Example is from mathisfun.com

Example: How much is that puppy growing?

Each month you measure how much weight your pup has gained and get these results:

0.5, 0.5, 0.3, −0.2, 1.6, 0, 0.1, 0.1, 0.6, 0.4

Mean and standard deviation

- The following Example is from mathisfun.com

Example: How much is that puppy growing?

Each month you measure how much weight your pup has gained and get these results:

0.5, 0.5, 0.3, −0.2, 1.6, 0, 0.1, 0.1, 0.6, 0.4

Five-number summary and boxplot

- The following Example is from mathisfun.com

Example: How much is that puppy growing?

Each month you measure how much weight your pup has gained and get these results:

0.5, 0.5, 0.3, −0.2, 1.6, 0, 0.1, 0.1, 0.6, 0.4

Scatter diagrams and correlation coefficient

Let's consider the following data

x	2	4	6	6	7
y	4	8	10	13	20

Least-squares regression

Let's consider the following data

x	2	4	6	6	7
y	4	8	10	13	20

Area under the Normal Curve

- Let's look at a problem from Statistics: Informed Decisions Using Data by Michael Sullivan III

Problem: A pediatrician obtains the heights of her three-year-old female patients. The heights are approximately normally distributed, with mean 38.72 inches and standard deviation 3.17 inches. Use the normal model to determine the proportion of the three-year-old females that have a height less than 35 inches.

Area = proportion

35.55 38.72 41.89   X

35
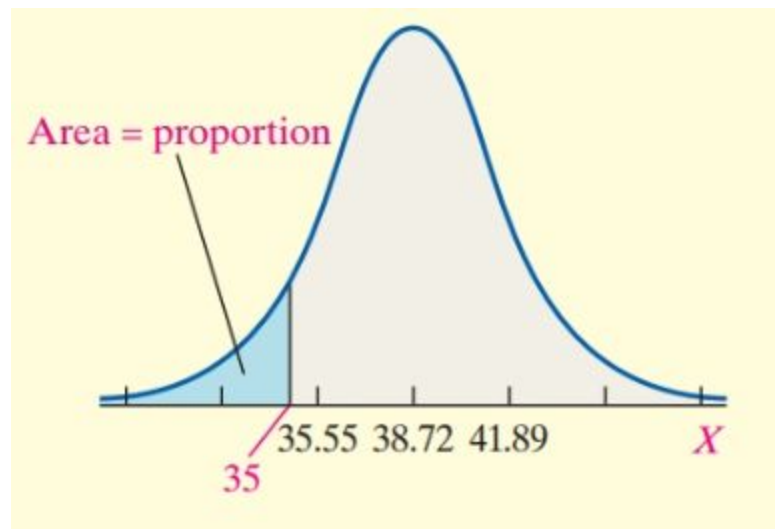
- **Normal distribution/ same code 8.1 just be careful with the standard deviation**

Finding the area under the curve

Area to the left

$P(x < a)$ or $P(x \leq a)$ pnorm(a, μ, σ)

Area to the right

$P(x > b)$ or $P(x \geq b)$ pnorm(b, μ, σ, lower.tail=FALSE)

Area in between

$P(a < x < b)$ or

$P(a \leq x \leq b)$ pnorm(b, μ, σ) - pnorm(a, μ, σ)

Questions?

For more information, here are some resources

- https://media.pearsoncmg.com/ph/esm/esm_sullivan_sst6e_21/r/sst6e_r.html
- https://www.statmethods.net/r-tutorial/index.html
- https://cengel.github.io/R-intro/backgroud.html