- Final Project - Poster Presentation Information is available on the course website. https://ntaback.github.io/UofT_STA130/Fall2018/Finalproject/sta130posterpresentation_student_Fall2018.html
- Form a group of 3 to 4 today.

Motivation of confidence interval

- We estimate a characteristic of a population from incomplete, imperfect observed data.

- What is a range of plausible values for what it could actually be?
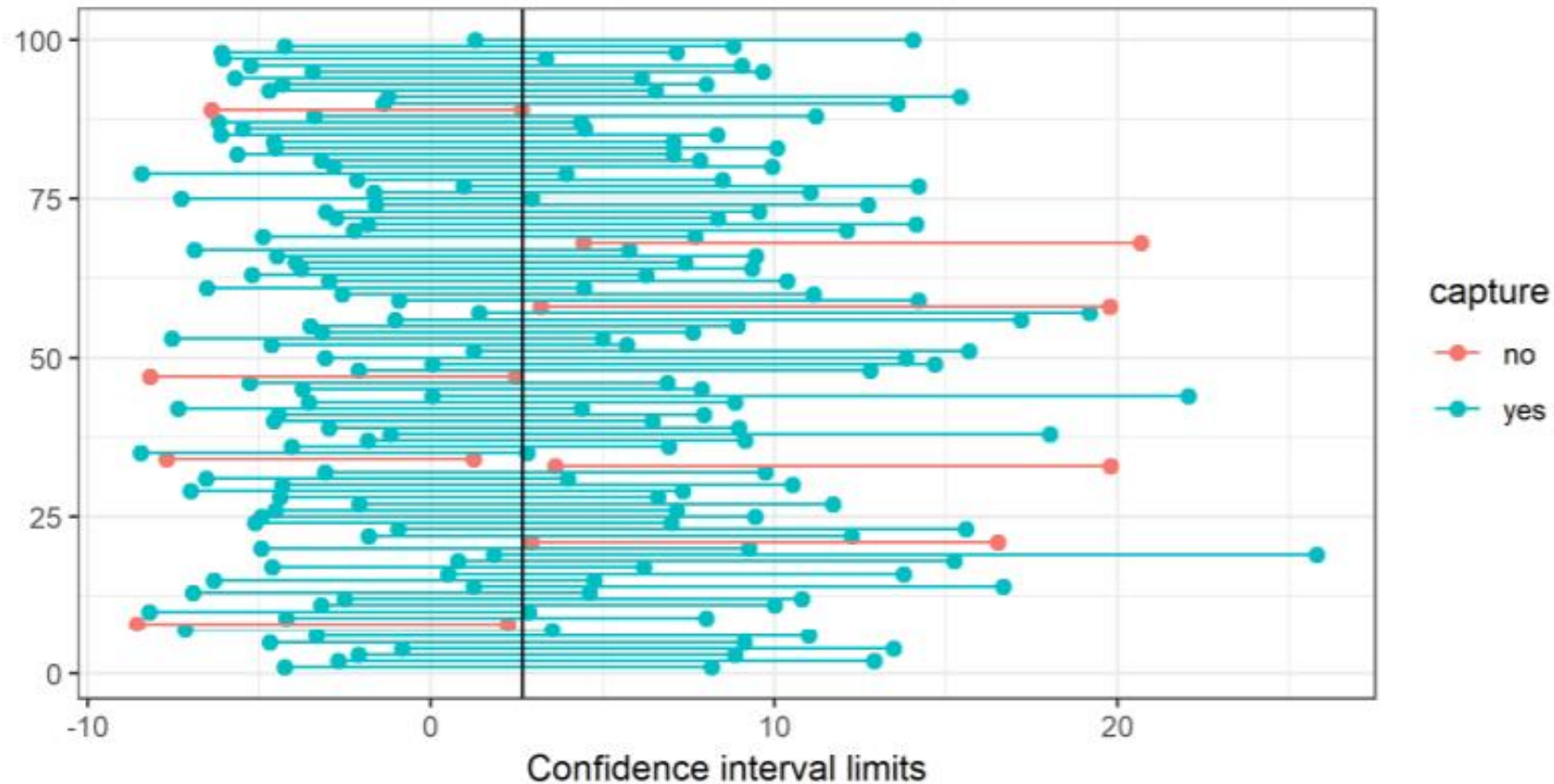
Steps of calculating confidence interval

- Randomly sample from your population.

- Calculate the estimated of the parameter you are interested in from the data in your sample from step 1 (this is called the **point estimate**, as a contrast to CI).

- By some methodology you will learn more in the 2nd year, you calculate the (1 - α)% Confidence interval. (Or by some R package)

- For now, you should get a big idea of what is a CI.

Some example

- Suppose you have a 95% confidence interval: [165cm, 182cm], how do you interpret it?

- Do the following statements correct?

- 1. There is 95% chance that the true mean of students' height in UofT is in [165cm, 182cm].

- 2. The probability that the true mean of students' height in UofT is in [165cm, 182cm] is 0.95

- Both are **wrong**, it is very temping to make such statements.
- We don't make probabilistic statement on CI!
- What is the correct interpretation?
- The true mean is either in [165cm, 182cm] or not in.

100 bootstrap confidence intervals for the mean, each calculated from a random sample from the population of size 200

How many of the confidence intervals capture the population mean?

- What does a 95% confidence interval of students' height in Uoft mean?

- Each time you sample from your population, you get a subset of student in UofT, you calculate the mean, using **the method,** you calculate a range of prediction of average height, which is your CI.

- Suppose you repeat the process 100 times, approximately 95% of time, the prediction interval you got will capture the true mean of students' height in UofT.

Week 8 presentation:

- In a **5-6 minute** presentation, describe the results of question1:
- Please use the following bullet points to prepare for your presentation:
- Describe the content of each histogram (mention the variable of interest)
  - What is the difference between histograms of a and b?
  - What is the difference between b (i) and b(ii)?
  - Describe the shape of the histograms (range, center, symmetry, skewness, number of points).
  - Are these bootstrap histograms or sampling histograms?
- How was one each histogram generated:
  - what statistic did you calculate from a particular sample?
- How can the generated distribution of median duration of flight be used for inference?
  - Which sample size would you choose (25 or 100)?
  - What do you conclude about sampling distribution of median?