STA130 Winter 2022, Tutorial 5

Ente Kang

University of Toronto

2022-02-14

Key vocabulary words for statistical inference

- Population
  - The 'objects' that you are interested in and want to make a conclusion about
  - Canadian census $\implies$ interested in all Canadians
- Parameter
  - A trait or characteristic about the population that you wish to study
  - μ, σ^2, p
- Estimation
  - Methods we will use to answer of question of interest

- Sample
  - A subset or collection of the population
  - May or may not be biased or representative or random
  - A **random** sample is as the word suggests - random
- Sampling distribution
  - A way to explore randomness
  - Recall repeated sampling & multiple experiments $\implies$ a distribution $\implies$ reject or fail to reject H_0
- Quantile
  - The xth quantile means x percent of values are below it

Confidence intervals

Let's suppose we have a γ% confidence level for μ, [0, 100]

What this does not mean

- There is γ% that $\mu \in$ [0, 100]

What this does mean

- If we kept taking *repeated* samples, in γ% of those, $\mu \in$ [0, 100]

What do we choose γ to be?

- By convention: 90%, 95% 97.5%, 99%

Why not a 100% confidence interval for μ?

- This is like saying, oh yeah $\mu \in \mathbb{R}$!
- Clearly, one can see issues with feasibility

An example

> Your supervisor asked you to determine the average price of a cup of Americano $\mu_{americano}$ sold in coffee shops in Canada.
>
> You **do not** want to say:
>
> It is some real number! :(

Hence you do not see uses of 100% confidence intervals

Furthermore, in statistics, we are never 100% sure of anything

- Not because we are indecisive!

Thank you, and see you at the next session