The What and the Why of Statistics

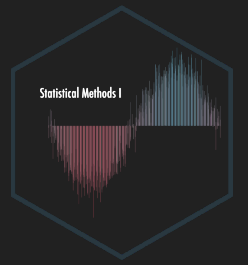
EDP 613

Week 1

Basic Ideas

> - **Parameter** - A number describing an entire *population*
> - **Statistic** - A number describing a slice, or a *sample* of a *population*

Polar Views of the World



Frequentist Statisticians believe that there is one and only one correct parameter that can be found by using multiple samples.

- parameters are fixed and data vary
- there is a single truth that can be found with enough indicators
- only objectibity can be used

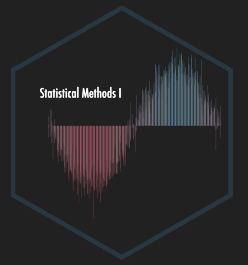
Bayesian Statisticians believe that multiple parameters exist which are all based on varying probabilities.

- parameters vary and the data is fixed
- there are multiple truths and getting to any one is based on chance
- subjectivity is a built feature

Learning Statistics

- You likely do not know enough about probability so for now assume that the frequentist point-of-view is correct.
- It is easier to begin to learn statistics if you don't have to consider multiple outcomes in superposition.
- We will come back to the Bayesian vs. Frequentist arguement

Start



Descriptive Statistics - Mathematical techniques for organizing and summarizing a set of numerical data

Finish

Inferential Statistics - Generalizing from a sample to a population

Definitions

- Information is collected on *elements* or *individuals*
- The characteristics of the individuals about which we collect information are called *variables*
- The values of the variables that we obtain are called *data*

Overarching Types of Data

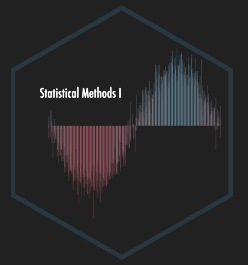
- *Qualitative variables* (aka *categorical variables*) classify elements into categories.
- *Quantitative variables* tell how much or how many of something there is.

Example

Which of the following variables are qualitative and which are quantitative?

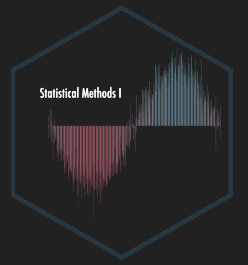
	Situation	Type
1	The name of the schools in your district.	
2	The number of schools in your district.	
3	The amount of each ingredient in a cake.	
4	The ingredients in a cake.	

Solution



	Situation	Type
1	The name of the schools in your district.	Qualitative
2	The number of schools in your district.	Quantitative
3	The amount of each ingredient in a cake.	Quantitative
4	The ingredients in a cake.	Qualitative

Levels of Measurement

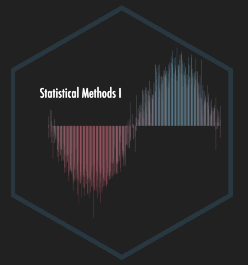


	Nominal	Ordinal	Interval	Ratio
Naming, labeling, or classifying observations	✓	✓	✓	✓
Ranks categories in order		✓	✓	✓
Known equal intervals			✓	✓
Includes a natural zero point				✓

Note

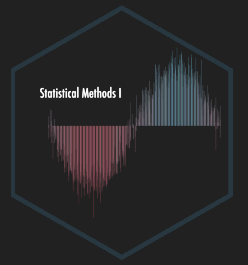
Your textbook pools interval and ratio together as *interval-ratio*.

Example



	Situation	Type
1	The (typical) letter grade distribution in a school	
2	Toppings on a cheeseburger	
3	Social economic status	
4	A telephone number	
5	Time	

Solution



	Situation	Type
1	The (typical) letter grade distribution in a school	Ordinal
2	Toppings on a cheeseburger	Nominal
3	Social economic status	Ordinal
4	A telephone number	Ordinal
5	Time	Interval Ratio

Discrete and Continuous

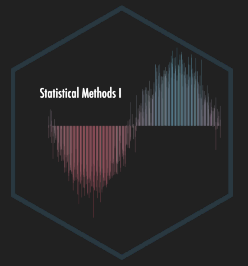
- *Discrete variables* are quantitative variables whose possible values can be listed
    - possibly infinite
    - obtained by counting
- *Continuous variables* are quantitative variables that can take on any value in some interval.
    - possibly infinite
    - obtained by measuring

Example

Which of the following variables are discrete or continuous?

	Situation	Type
1	Time it takes to get to school	
2	Water temperature	
3	Ratings on a 5-point rating scale	
4	Number of cars currently in a parking lot	

Solution



	Situation	Type
1	Time it takes to get to school	Continuous
2	Water temperature	Continuous
3	Ratings on a 5-point rating scale	Discrete
4	Number of cars currently in a parking lot	Discrete

That's it. Take a break before our R session!