

Math 142: Wednesday, Aug 20th, Day 2 Notes

Agenda for Wednesday, Aug 20, 2024

Section 1.2 Data and Sampling	1
Group activity	2
Section 1.3 Experimental Design	8

Reminders

- ! Complete intro form (if haven't done so by now)



- Canvas Homeworks Section 1.1 and 1.2 will open today and be due next Wednesday

Section 1.1 Basic Definitions continued

Recall: Data, Descriptive statistics, Inferential Statistics, Population, Sample

A few more definitions:

- A **Parameter** is a value that represents a property of the population
 - use lower case Greek letters
 - A **Statistic** is a value that represents a property of the sample
 - $\bar{x}$ sample mean
 - s standard deviation
 - A **Variable** is a characteristic of interest for each item/member of the population
 - we use capital letters: X, Y, Z
- Handwritten notes:*
A bracket groups the Greek letters μ and σ . Next to μ is "(new) population mean (average)". Next to σ is "(sigma) pop standard deviation".

Example 1. Determine what the key terms refer to in the following study. We want to know the average amount of money first year college students spend at CLC on school supplies that do not include books. We randomly survey 100 first year students at the college. The average first year CLC student spends \$168 on school supplies that do not include books. The average of those surveyed spend \$174.

- Population: all first year CLC students
- Sample: 100 randomly selected students
- Parameter: $\mu = \$168$
- Statistic: $\bar{x} = 174$

$$n = 100$$

n = sample size

There are two types of variables:

- Numerical variable takes on numerical values
ex: age is a numerical variable
19, 20, 18, ...
- Categorical variable places each individual/object in a category
ex: phone type : iPhone, Android

Example 2. Is the variable numerical or categorical?

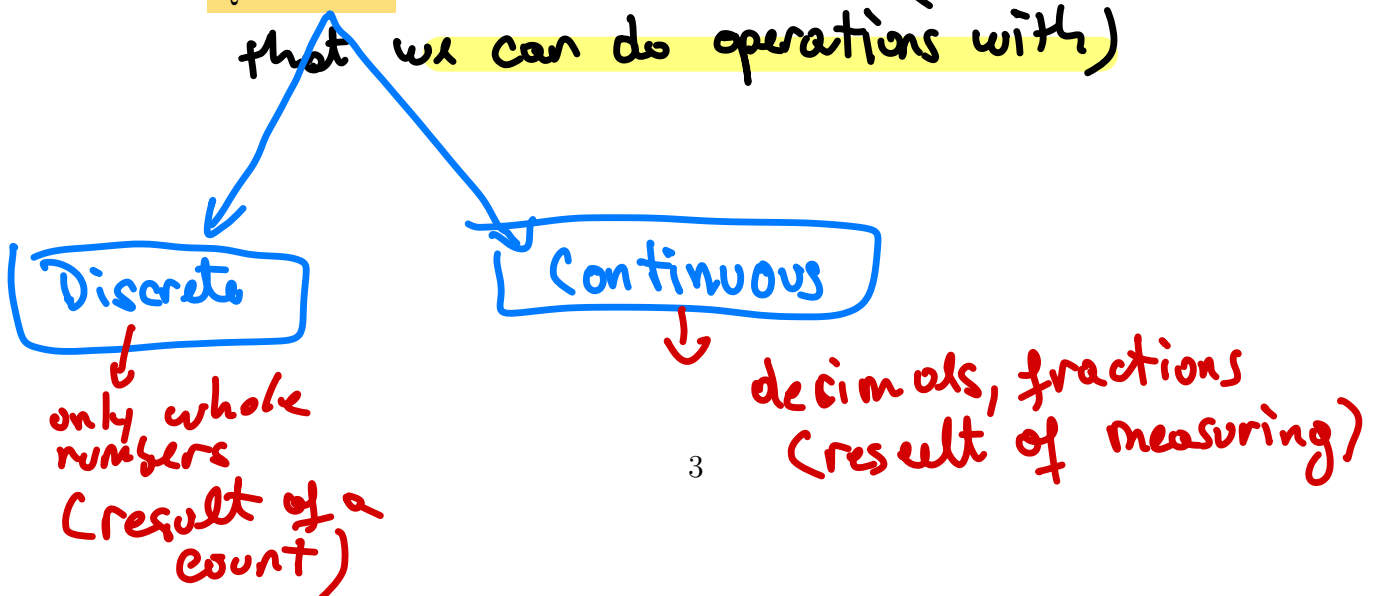
1. Let Y be the monthly cost to operate a business - Num.
2. Let P be a favorite pizza place - categ.
3. Let A be an Academic major - categ.

1.2 Data and Sampling

Actual value of a variable
result in data.

There are two types of data:

1. Qualitative non-numerical value
variable : favorite fast food place
qualitative data : Chipotle, Wendy's...
2. Quantitative - numeric values (real numbers
that we can do operations with)



Example 3. Determine if the data is qualitative or quantitative. If the data is quantitative, determine if it is continuous or discrete.

- a) The number of smartphones in a household *QD*
- b) The type of car a person drives *QL*
- c) Where one goes on vacation *QL*
- d) The distance it is from home to the nearest gas station *QTC*
- e) The tuition for a student's classes this semester *QTC*
- f) Movie ratings (PG, PG-13, G, ...) *QL*
- g) Political party preferences *QL*
- h) Weights of NFL wide receivers *QTC*
- i) Temperature of infants in the hospital nursery *QTC*
- j) Amount of money (in dollars) won playing video poker *QTC*
- k) Student ID number *QL*

Money is
always
quantitative
continuous
\$

Sampling Techniques

1. Simple Random Sample SRS - each member of the population has equal chance to be chosen as part of the sample.

2. Stratified Sample we first divide the population into strata (groups) based on certain characteristics and we randomly choose equal amounts or percentages from each group.

3. Cluster Sample Divide the population into clusters and then randomly select some of the clusters.

4. Systematic Sample - every 6th person entering a building.

5. Convenience Sample

- not good

Get in pairs and discuss Example 4.

Example 4. Determine the type of sampling used for the following:

Benefits package which includes health insurance, dental insurance, child care, vacation days, retirement plan, etc. You want to conduct a survey of benefits packages available in private business in the state of Illinois. You want a sample of size 200. Some sampling techniques are described below.

1. Assign each tax-id of the private business a number. Choose every 15th number to select to be included in the sample.

systematic

2. Pick a random sample of 10 zip code areas and then include all the businesses in each of the selected zip code area.

cluster

3. Assign each tax-id of the private business a number. Have a random generator select 200 the businesses to be included in the sample.

SRS

4. Group each private business into categories such as Medical, Retail, Financial, Restaurant, Hotel, Manufacturing, Education, other. Then select a random sample of 20 businesses from each type.

stratified

5. Send an online survey to each email associated with the tax ID registered to the state and choose the first 200 responses to the survey

convenience