

DS UA 112: Discussion 3

September 17, 2019

1 Study Design

Which of the following describe a proxy variable in study design? A proxy variable is ____ (check all that apply)

- ☐ A variable of interest
- ☐ Not a variable of interest
- ☐ Easy to measure
- ☐ Difficult to measure
- ☐ Related to a variable of interest
- ☐ Related to a population of interest

2 Census

Why are p values meaningless when taking a census?

- ☐ The chance that a census was drawn at random is large.
- ☐ A p value expresses the chance that some measured difference (or a larger one) would arise due to random sampling.
- ☐ There is a better alternative to a p value when taking a census.
- ☐ There is no need for statistical inference when directly measuring an entire population.

3 Sampling

3.1 Probability Sampling

If everyone in a probability sample has an equal chance of being included in the sample, this guarantees that the sample is a simple random sample (SRS).

☐ True.

☐ False.

3.2 Simple Random Sample

In a simple random sample (SRS), everyone has an equal chance of being included in the sample.

☐ True

☐ False

3.3 Unbiased Sample

Example

Kathy wants to know how many students in her city use the internet for learning purposes. She used an email poll. Based on the replies to her poll, she found that 83% of those surveyed used the internet. Kathy's sample is biased as she surveyed only the students those who use the internet. She should have randomly selected a few schools and colleges in the city to conduct the survey.

Question: Which one of the following sampling methods would give unbiased results, if you need to find out the number of people in your town liking vanilla or chocolate ice creams?

☐ Ask my neighbors

☐ Randomly select a few ice cream shops in town, and question people coming there

☐ Choose some ice cream shops in my street and survey people coming from there

☐ Ask few ice cream vendors