Financial Lit Name: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_

WS Assessment

Target 23:

Intro to statistic

**I can:**

* Define and identify the population, parameter, sample and statistic
* Identify four sampling methods: simple random sample (SRS), stratified, systematic and convenience
* Identify and discuss types of bias association with sampling
* Distinguish between experimental and observational studies
* Explain margin of error and confidence intervals

**Unit 11 Math Topics:**

* Statistical Process
* Summary of Statistic: Measure of center and spread

We are bombarded by information and statistics every day. But if we cannot distinguish credible information from misleading information, then we are vulnerable to manipulation and making decisions that are not in our best interest. Statistics provides tools for us to evaluate information critically. In this sense, statistics is one of the most important things to know about.



**Population**

Before we begin gathering any data to analyze, we need to identify the population we are studying. The population of a study is the group we want to know something about. The population could be people, auto parts or tomato plants.

A newspaper website contains a poll asking people their opinion on a recent news article. What is the population? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Parameter**

A parameter is the value (percentage, average, etc.) that we are interested in for the whole population. Since it is often too time-consuming, expensive and/or impossible to get data for the entire population, the parameter is usually a theoretical quantity that we are trying to estimate.

**Sample**

To estimate the parameter, we select a sample, which is a smaller subset of the entire population. It is very important that we choose a representative sample, one that matches the characteristics of the population, to have a good estimate.

**Statistic**

To get our data, we would then ask each student in the sample how much they spent on textbooks and record the answers, or raw data. Then we could calculate the average, which is our statistic. A statistic is a value (percentage, average, etc.) calculated using data from a sample.

A researcher wants to know how the citizens of Portland feel about a voter initiative. To study this, they go downtown to the Pioneer Place Mall and survey 500 shoppers. Sixty percent indicate they are supportive of the initiative. Identify

|  |  |  |  |
| --- | --- | --- | --- |
| population | parameter | sample | statistic |
|  |  |  |  |

**Sampling Methods**

* Simple random sample (SRS) is one in which every member of the population has an equal probability of being chosen.
* Stratified sampling, a population is divided into a number of subgroups (or strata). Random samples are then taken from each subgroup. It is often desirable to make the sample sizes proportional to the size of each subgroup in the population.
* Systematic sampling, every nth member of the population is selected to be in the sample. The starting position is often chosen at random.
* Convenience sampling is when samples are chosen by selecting whomever is convenient. This is the worst type of sampling because it does not use randomness.

Matching

|  |  |  |
| --- | --- | --- |
| Every 4th person in the class was selected |  | Simple random sample |
| Viewers of a new show are asked to vote on the show’s website |  | Stratified sampling |
| A website randomly selects 50 of their customers to send a satisfaction survey to |  | Systematic sampling |
| To survey voters in a town, a polling company randomly selects 10 city blocks, and interviews everyone who lives on those blocks |  | Convenience sampling |
| A sample was selected to contain 25 men and 35 women |  |  |

**Types of bias**

* Sampling bias – when the sample is not representative of the population
* Voluntary response bias – the sampling bias that often occurs when the sample is made up of volunteers
* Self-interest study – bias that can occur when the researchers have an interest in the outcome
* Response bias – when the responder gives inaccurate responses for any reason
  + Perceived lack of anonymity – when the responder fears giving an honest answer might negatively affect them
  + Loaded questions – when the question wording influences the responses
* Non-response bias – when people refuse to participate in a study or drop out of an experiment, we can no longer be certain that our sample is representative of the population

Matching

|  |  |  |
| --- | --- | --- |
| A survey asks how many sexual partners a person has had in the last year |  | Sampling bias |
| A radio station asks readers to phone in their choice in a daily poll. |  | Voluntary response bias |
| High school students are asked if they have consumed alcohol in the last two weeks |  | Self-interest study |
| The Beef Council releases a study stating that consuming red meat poses little cardiovascular risk. |  | Response bias |
| A poll asks “Do you support a new transportation tax, or would you prefer to see our public transportation system fall apart?” |  | Non-response bias |
| A substitute teacher wants to know how students in the class did on their last test. The teacher asks the 10 students sitting in the front row to state their latest test score. |  |  |

Is each scenario describing an observational study or an experiment?

a. The weights of 30 randomly selected people are measured \_\_\_\_\_\_\_\_\_\_\_\_\_\_

b. Subjects are asked to do 20 jumping jacks, and then their heart rates are measured \_\_\_\_\_\_\_\_\_\_

c. Twenty coffee drinkers and twenty tea drinkers are given a concentration test \_\_\_\_\_\_\_\_\_\_\_\_

What is the treatment for the following scenario

a. A pharmaceutical company tests a new medicine for treating Alzheimer’s disease by administering the drug to 50 elderly patients with recent diagnoses. The treatment here is \_\_\_\_\_\_\_\_\_\_\_\_\_\_.

b. A gym tests out a new weight loss program by enlisting 30 volunteers to try out the program. The treatment here is \_\_\_\_\_\_\_\_\_\_.

c. You test a new kitchen cleaner by buying a bottle and cleaning your kitchen.

The treatment is \_\_\_\_\_\_\_\_\_

d. A psychology researcher explores the effect of music on temperament by measuring people’s temperament while listening to different types of music. The treatment is \_\_\_\_\_\_\_\_\_\_\_\_\_\_

When conducting experiments, it is essential to isolate the treatment being tested.

**Confounding** occurs when there are two potential variables that could have caused the outcome and it is not possible to determine which actually caused the result.

What is the cofounding factors for the following scenario

a. Suppose a middle school (junior high) finds that their students are not scoring well on the state’s standardized math test. They decide to run an experiment to see if an alternate curriculum would improve scores. To run the test, they hire a math specialist to come in and teach a class using the new curriculum. To their delight, they see an improvement in test scores.

Factors that affect test score?

A drug company study about a weight loss pill might report that people lost an average of 8 pounds while using their new drug. However, in the fine print you find a statement saying that participants were encouraged to also diet and exercise.

Factors that affect the weight loss?

There are a number of measures that can be introduced to help reduce the likelihood of confounding. The primary measure is to use a **control group**. For example

To determine if a two-day prep course would help high school students improve their scores on the SAT test, a group of students was randomly divided into two subgroups. The first group, the **treatment group**, was given a two-day prep course. The second group, the **control group**, was not given the prep course. Afterwards, both groups were given the SAT.



Example

a. In a study for a new medicine that is dispensed in a pill form, **a sugar pill** could be used as a placebo.

b. In a study on the effect of alcohol on memory, **a non-alcoholic** beer might be given to the control group as a placebo.

**Blind studies**

A blind study is one that uses a placebo and the participants do not know whether they are receiving the treatment or a placebo. A double-blind study is one in which the subjects and those interacting with them don’t know who is in the treatment group and who is in the control group.

**Margin of Error and Confidence Intervals**

You’ve probably heard something like, “The candidate has 54 percent of the likely voters, plus or minus three percent.” The 3% is called the margin of error, so the true percentage is somewhere between 51% and 57%, with a certain level of confidence. To write this as a confidence interval, we place the numbers in parentheses from smallest to largest, separated by a comma: (51%, 57%).

A political scientist surveys 28 of the current 106 representatives in a state's congress. Of them, 14 said they were supporting a new education bill, 12 said there were not supporting the bill, and 2 were undecided.

a. Who is the population of this survey?

b. What is the size of the population?

c. What is the size of the sample?

d. Give the statistic for the percentage of representatives surveyed who said they were supporting the education bill.

e. If the margin of error was 5%, give the confidence interval for the percentage of representatives we might we expect to support the education bill.

Which sampling method is being described?

a. In a study, the sample is chosen by separating all cars by size and selecting 10 of each size grouping.

b. In a study, the sample is chosen by writing everyone’s name on a playing card, shuffling the deck, then choosing the top 20 cards.

c. Every 4th person on the class roster was selected.

**Assessment Target 23**

**I can…** define population, sample and other statistical terminology

To determine the average length of trout in a lake, researchers catch 20 fish and measure them. Describe the population and sample of this study.

The city of Raleigh has 9,500 registered voters. There are two candidates for city council in an upcoming election: Brown and Feliz. The day before the election, a telephone poll of 350 randomly selected registered voters was conducted. 112 said they'd vote for Brown, 207 said they'd vote for Feliz, and 31 were undecided.

a. Who is the population of this survey?

b. What is the size of the population?

c. What is the size of the sample?

d. Give the statistic for the percentage of voters surveyed who said they'd vote for Brown.

e. If the margin of error was 3.5%, give the confidence interval for the percentage of voters surveyed that we might we expect to vote for Brown

For the clinical trials of a weight loss drug containing *Garcinia Cambogia* the subjects were randomly divided into two groups. The first received an inert pill along with an exercise and diet plan, while the second received the test medicine along with the same exercise and diet plan. The patients do not know which group they are in, nor do the fitness and nutrition advisors.

a. Which is the treatment group?

b. Which is the control group (if there is one)?

c. Is this study blind, double-blind, or neither?

d. Is this best described as an experiment, a controlled experiment, or a placebo-controlled experiment?

A poll found that 30%, plus or minus 5% of college freshmen prefer morning classes to afternoon classes.

a. What is the margin of error?

b. Write the survey results as a confidence interval.