Statistics

* I can estimate population percentages using the Empirical Rule.

A college-entrance exam is designed so that scores are normally distributed with a mean of 500 and a standard deviation of 100.

1. Approximately what percent of exam scores are between 400 and 600?

Approximately 68%

2. What is the probability that a randomly chosen exam score is above 600?

 $P\left(Scores>600\right)≈0.16$

3. In a random sample size of 200, approximately how many scores would be below 300?

 About 27 scores

* I can determine if it is appropriate to use the Empirical Rule to approximate a population percentage from a set of data.

In an experiment, a die was rolled until a 3 appeared. The number of times the die was rolled until a 3 appeared was recorded. The experiment was performed 20 times and the following data was collected: {5, 3, 7, 7, 5, 3, 12, 5, 1, 13, 20, 1, 1, 3, 1, 12, 4, 17, 4, 6}. Make plot of the data and determine if it is appropriate to use the Empirical Rule.

A plot of the data will show a right-skewed distribution; therefore the Empirical Rule is not appropriate to use to approximate a population percentage from this set of data.

* I can estimate a mean and standard deviation from a normal distribution and I can use a mean and standard deviation to sketch a normal distribution.

The heights of Munchkins, in inches, is normally distributed as shown by the plot below. Estimate the mean and standard deviation of the heights of Munchkins (don’t forget the units).



The mean is approximately 52 inches with a standard deviation of about 6 inches.

Given that college entrance exam scores are normally distributed with a mean of 500 and a standard deviation of 100, sketch and label a normal distribution curve that represents this data.

Students should sketch a normal curve that has a center at about 500. The points of inflection define where the first standard deviation from the mean is, at about 400 and 600. That interval length is then used to approximate two and three standard deviations from the mean, at about 200, 300, 700, and 800.

* I can distinguish between random and non-random samples, and between representative and non-representative samples.

Determine a population and then give an example of each of the following types of samples that can be obtained from that population: For example: Population of dentists in Utah County. The

 following samples are just examples. Answers may vary.

 A random sample that represents the population

 Simple random sample taken from a list of the dentists in Utah County

 A non-random sample that represents the population

 Asking 5 dentists in from each city in Utah County

 A random sample that does not represent the population

 A simple random sample that only includes female dentists

 A non-random sample that does not represent the population

 Asking 20 dentists in Orem

(For example, given that Colgate wants to determine what percentage of dentists recommend their toothpaste.)

* I can recognize the purposes and differences between a sample survey, an observational study, and an experiment, and determine if/how randomization was used.

Antibiotic use in infants is linked to asthma and allergies, according to a study involving 1401 children. Researchers asked mothers how many doses of antibiotics their children received before 6 months of age as well as whether their children had developed asthma or allergies by age 6. Children who received just one dose of antibiotics were 40% more likely to develop asthma or allergies. The risk jumped to 70% for children who received two doses.

Is this a survey, an observational study, or an experiment? How do you know? Was randomization used in the research? If so, how?

This is an observational study because two variables, antibiotic use and occurrence of asthma and allergies, are compared but no treatment is imposed. This report does not say whether subjects were chosen at random.

Which of the following research topics are best addressed through an observational study? Which are best addressed through a sample survey, and which are best addressed with an experiment? (This is a discussion, and answers may vary depending on the purpose of the particular study.)

1. Does listening to loud music with headphones affect a person’s hearing?

Observational study

1. Does second-hand smoke affect the health of pets?

Observational study

1. Does a particular medication make seasonal allergy symptoms less severe?

Experiment

1. Do students prefer to take math classes before lunch or after lunch?

Sample survey

1. Does increasing the number of stoplights per mile on a road decrease the number of car accidents on the road?

Observational study

1. What type of music do students prefer?

Sample survey

1. Does drinking sports drinks before and while playing a game of baseball increase the number of runs scored?

Experiment

1. Does a certain toothpaste prevent cavities in children better than another toothpaste?

Experiment

* I can find an interval that estimates reasonably likely values for a population proportion or a population mean given a sample proportion/mean and standard deviation.

If a sample mean is 63 and the standard deviation is 6, what is the interval likely to contain the population mean? Approximately 51 to 75

Given a sample proportion of 0.4 and a margin of error of ± 0.04, what is the interval likely to contain the population proportion? Approximately 0.36 to 0.44

For the population of Nowheresville, the interval likely to contain the actual perceD.ntage of people who are left handed is 8% - 14%. If the sample proportion was 11%, which of the following could be the sample standard deviation?

 A. 6% B. 3½% C. 3% D. 1½% E. 1%

Why? D. 1½%; interval is $0.11\pm 2∙0.015$

* I can use results from a randomized experiment to compare two treatments and determine possible differences.

A class of 40 students measured their heart rates while sitting and while standing. Each student tossed a penny to randomize whether they measured their heart rate standing or sitting first. Each student found the difference between their standing and sitting heart rates (standing minus sitting). The differences were collected on a class dotplot.



Are any of the differences significant? Why? Typical heart rates appear to be from about $-1 to 4$. A heart rate of 7 is significant since it is so far from the rest of the heart rates. A heart rate of 5 may or may not be significant – we would need more data or a more precise definition of “significant” to decide.