
1.1 Descriptive and Inferential Statistics

- The current national average price per gallon of gas is \$2.39 (as of January 2017).
- In a survey of 1,000 Americans, 77% own smart phones (as of January 2017).
- Researchers find that calorie consumption is reduced by 6% when calories are posted for each item sold at chain restaurants.
- The 2017 Chevy Sonic, LTZ gets 34 (\pm 5) miles per gallon.

The preceding claims are examples of statistics that are commonly announced by various media. In this course, we will examine how studies are designed and performed that result in such claims. First, we will introduce some vocabulary.

Statistics

Statistics is the practice or science of collecting and analyzing numerical data in large quantities.

Probability

Probability is the extent to which an event is probable. The mathematics of probability is the tool by which an analysis of data from a sample can lead to information about an entire population.

| Definition | Examples |
|---|---|
| <ul style="list-style-type: none"> • Descriptive statistics is the discipline of quantitatively describing the main features of a collection of data. | <ul style="list-style-type: none"> • The statistics class earned a mean (average) midterm exam score of 89.5% with a standard deviation of 3.8%. • The SF Giants have earned an average of 3.8 runs batted in per game so far in the 2017 season. |
| <ul style="list-style-type: none"> • Inferential statistics is the process of drawing conclusions from data that is subject to random variation. | <ul style="list-style-type: none"> • Based on the midterm exam results of the statistics class, it is likely that the class will earn an average of 85% or better on the final exam. • The SF Giants are likely to maintain an average of 3.8 runs batted in per game in the 2017 season. |

Populations and Samples

| Definition | Examples |
|--|--|
| <p>A population is a set of people, places, things, or ideas, (sometimes referred to as subjects) from which we may collect data. Most often a population is very large, possibly too large or impossible to study in its entirety.</p> | <ul style="list-style-type: none"> • The number of homeruns hit by Buster Posey in every game of his entire career • All people residing in the U. S. |
| <p>A sample is a subset of a population. Often a sample is analyzed for the purpose of drawing conclusions about the entire population.</p> | <ul style="list-style-type: none"> • The number of homeruns hit by Buster Posey in the first 5 months of the 2017 season is {1, 6, 3, 1, 1} • A sample of 1,000 U. S. residents are surveyed |

Parameters and Statistics

| Definition | Examples |
|--|--|
| A parameter is a numerical measure of a characteristic of a given population | <ul style="list-style-type: none"> • The average number of homeruns hit by Buster Posey per game of his entire career • Of all people residing in the U. S., the number that own smartphones |
| A statistic is a numerical measure of a characteristic of a representative sample of a population | <ul style="list-style-type: none"> • The average number of homeruns hit by Buster Posey per month in the first 5 months of the 2017 season is 2.4 • In a sample of 1,000 U. S. residents surveyed, 770 own smartphones |

Data and Variables

| Definition | Examples |
|--|---|
| A variable is an attribute that describes a person, place, thing, or idea. The value of the variable can "vary" from one entity to another. | <ul style="list-style-type: none"> • Let X = the number of homeruns hit by Buster Posey in each month of the 2017 season. • Let Y = the number of students in each statistics class taught at this college. • Let C = the eye color of the students in this class |
| Data are facts or information that could be used in a study. | <ul style="list-style-type: none"> • Buster Posey hit 6 homeruns in the 2nd month of the 2017 season. • Scooby Doo ate 10 ounces of dog food this morning. • There are 34 students in this statistics class. |
| A data set is a set of related facts or information. | <ul style="list-style-type: none"> • Buster Posey hit 1 homerun in the 1st month of the 2017 season, 6 in the 2nd month, 3 in the 3rd month, 1 in the 4th month, 1 in the 5th month, This data set could be denoted by $\{1, 6, 3, 1, 1\}$. |
| A datum is a single value of a data set. | <ul style="list-style-type: none"> • Buster Posey hit 6 homeruns in the second month of the 2017 baseball season. |

| <i>Demonstration Problems</i> | <i>Practice Problems</i> |
|--|--|
| <p>Determine whether descriptive or inferential statistics were used to make the following claims (include why):</p> <p>1. (a) The average goals per match earned by the U. S. A. team during the 2014 World Cup tournament is 1.7.</p> <p>2. (a) A study done by Nieman DC, found that adults that exercise 1 to 4 days per week have a reduced risk of getting a cold.</p> | <p>Determine whether descriptive or inferential statistics were used to make the following claims (include why):</p> <p>1. (b) From an analysis of sales receipts for the month of July, a pet store determined that 10 fish per day were sold.</p> <p>2. (b) According to the U. S. Bureau of Labor Statistics, employment in all computer applications is expected to grow 22% by the year 2020.</p> |
| <p>Answers: 1. (b) Descriptive statistics, because this is an analysis of data collected at the pet shop.; 2. (b) Inferential, because it is not possible to determine the actual number of all computer application jobs that will occur in a future year.</p> | |