

Basic Statistics and Probability (2018-19)
Science & Engineering Program Boston University-Faculty of Science UAM

Chapter 1: Statistics, Data and Statistical Thinking

The following problems are from McClave, J. and Sincich, T. (2017), *Statistics*, 13th. edition, Pearson. The number preceding the exercise is the corresponding one from this textbook. The problems also appear in previous editions of the book, possibly with different numbering.

1.15 Ground motion of earthquakes. In the *Journal of Earthquake Engineering* (Nov. 2004), a team of civil and environmental engineers studied the ground motion characteristics of 15 earthquakes that occurred around the world since 1940. Three (of many) variables measured on each earthquake were the type of ground motion (short, long, or forward directive), the magnitude of the earthquake (on the Richter scale), and peak ground acceleration (feet per second). One of the goals of the study was to estimate the inelastic spectra of any ground motion cycle.

- a. Identify the experimental units for this study.
- b. Do the data for the 15 earthquakes represent a population or a sample? Explain.
- c. Define the variables measured and classify them as quantitative or qualitative.

1.20 Extinct birds. Biologists at the University of California (Riverside) are studying the patterns of extinction in the New Zealand bird population. (*Evolutionary Ecology Research*, July 2003.) At the time of the Maori colonization of New Zealand (prior to European contact), the following variables were measured for each bird species:

- a. Flight capability (volant or flightless)
- b. Type of habitat (aquatic, ground terrestrial, or aerial terrestrial)
- c. Nesting site (ground, cavity within ground, tree, cavity above ground)
- d. Nest density (high or low)
- e. Diet (fish, vertebrates, vegetables, or invertebrates)
- f. Body mass (grams)
- g. Egg length (millimeters)
- h. Extinct status (extinct, absent from island, present)

Identify each variable as quantitative or qualitative.

1.27 CT scanning for lung cancer. A new type of screening for lung cancer, computed tomography (CT), has been developed. Medical researchers believe that CT scans are more sensitive than regular X-rays in pinpointing small tumors. The H. Lee Moffitt Cancer Center at the University of South Florida conducted a clinical trial of 50,000 smokers nationwide to compare the effectiveness of CT scans with X-rays for detecting lung cancer. (*Today's Tomorrows*, Fall 2002.) Each participating smoker was randomly assigned to one of two screening methods, CT or chest X-ray, and his or her progress tracked over time. The age at which the scanning method first detects a tumor is the variable of interest.

- a. Identify the data collection method used by the cancer researchers.
- b. Identify the experimental units of the study.
- c. Identify the type (quantitative or qualitative) of the variable measured.
- d. Identify the population and sample.
- e. What is the inference that will ultimately be drawn from the clinical trial?

1.30 Insomnia and education. Is insomnia related to education status? Researchers at the Universities of Memphis, Alabama at Birmingham, and Tennessee investigated this question in the *Journal of Abnormal Psychology* (Feb. 2005). Adults living in Tennessee were selected to participate in the study, which used a random-digit telephone dialing procedure. Two of the many variables measured for each of the 575 study participants were number of years of education and insomnia status (normal sleeper or chronic insomniac). The researchers discovered that the fewer the years of education, the more likely the person was to have chronic insomnia.

- a. Identify the population and sample of interest to the researchers.
- b. Identify the data collection method. Are there any potential biases in the method used?
- c. Describe the variables measured in the study as quantitative or qualitative.
- d. What inference did the researchers make?

1.35 Your choice for a mom. After running stories on current First Lady Michelle Obama and former vice presidential candidate Sarah Palin on consecutive weeks, *USA Weekend* magazine asked its readers on Mother's Day, "Who would you rather have as your mom, Sarah Palin or Michelle Obama?" Readers were asked to vote online at www.usaweekend.com. Based on over 34,000 votes cast, the results were: Obama–71%, Palin–29%. (*USA Weekend Magazine* press release, May 11, 2010.)

- a. What type of data collection method is used in this study?
- b. Is the data collected quantitative or qualitative? Explain.
- c. Discuss the validity of the study results. What are the potential problems with running a poll where voting is done online?

Activity. Data in the News. Scan your daily newspaper or weekly news magazine, or search the Internet for articles that contain data. The data might be a summary of the results of a public opinion poll, the results of a vote by the U.S. Senate, or a list of crime rates, birth or death rates, etc. For each article you find, answer the following questions:

- a. Do the data constitute a sample or an entire population? If a sample has been taken, clearly identify both the sample and the population; otherwise, identify the population.
- b. What type of data (quantitative or qualitative) has been collected?
- c. What is the source of the data?
- d. If a sample has been observed, is it likely to be representative of the population?
- e. If a sample has been observed, does the article present an explicit (or implied) inference about the population of interest? If so, state the inference made in the article.