Exercises

Prerequisites
• All material presented in Chapter: “Introduction”

1. A teacher wishes to know whether the males in his/her class have more conservative attitudes than the females. A questionnaire is distributed assessing attitudes and the males and the females are compared. Is this an example of descriptive or inferential statistics?

2. A cognitive psychologist is interested in comparing two ways of presenting stimuli on subsequent memory. Twelve subjects are presented with each method and a memory test is given. What would be the roles of descriptive and inferential statistics in the analysis of these data?

3. If you are told only that you scored in the 80th percentile, do you know from that description exactly how it was calculated? Explain.

4. A study is conducted to determine whether people learn better with spaced or massed practice. Subjects volunteer from an introductory psychology class. At the beginning of the semester 12 subjects volunteer and are assigned to the massed-practice condition. At the end of the semester 12 subjects volunteer and are assigned to the spaced-practice condition. This experiment involves two kinds of non-random sampling: (1) Subjects are not randomly sampled from some specified population and (2) subjects are not randomly assigned to conditions. Which of the problems relates to the generality of the results? Which of the problems relates to the validity of the results? Which problem is more serious?

5. Give an example of an independent and a dependent variable.

6. Categorize the following variables as being qualitative or quantitative:
   - Rating of the quality of a movie on a 7-point scale
   - Age
   - Country you were born in
   - Favorite Color
   - Time to respond to a question
**Summation Notation**

by David M. Lane

**Prerequisites**
- None

**Learning Objectives**
1. Use summation notation to express the sum of all numbers
2. Use summation notation to express the sum of a subset of numbers
3. Use summation notation to express the sum of squares

Many statistical formulas involve summing numbers. Fortunately there is a convenient notation for expressing summation. This section covers the basics of this summation notation.

Let's say we have a variable $X$ that represents the weights (in grams) of 4 grapes. The data are shown in Table 1.

**Table 1. Weights of 4 grapes.**

<table>
<thead>
<tr>
<th>Grape</th>
<th>$X$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.6</td>
</tr>
<tr>
<td>2</td>
<td>5.1</td>
</tr>
<tr>
<td>3</td>
<td>4.9</td>
</tr>
<tr>
<td>4</td>
<td>4.4</td>
</tr>
</tbody>
</table>

We label Grape 1's weight $X_1$, Grape 2's weight $X_2$, etc. The following formula means to sum up the weights of the four grapes:

$$
\sum_{i=1}^{4} X_i
$$

The Greek letter $\Sigma$ indicates summation. The "$i = 1$" at the bottom indicates that the summation is to start with $X_1$ and the 4 at the top indicates that the summation will end with $X_4$. The "$X_i$" indicates that $X$ is the variable to be summed as $i$ goes from 1 to 4. Therefore,
7. Specify the level of measurement used for the items in Question 6.

8. Which of the following are linear transformations?
   Converting from meters to kilometers
   Squaring each side to find the area
   Converting from ounces to pounds
   Taking the square root of each person's height.
   Multiplying all numbers by 2 and then adding 5
   Converting temperature from Fahrenheit to Centigrade

9. The formula for finding each student’s test grade (g) from his or her raw score (s) on a test is as follows: 
   \[ g = 16 + 3s \]
   Is this a linear transformation?

   If a student got a raw score of 20, what is his test grade?

10. For the numbers 1, 2, 4, 16, compute the following:
    \[ \Sigma X \]
    \[ \Sigma X^2 \]
    \[ (\Sigma X)^2 \]

11. Which of the frequency polygons has a large positive skew? Which has a large negative skew?

   A. 
   B. 
   C. 

12. What is more likely to have a skewed distribution: time to solve an anagram problem (where the letters of a word or phrase are rearranged into another
Treatment Effects of a Drug on Cognitive Functioning in Children with Mental Retardation and ADHD

RESEARCH CONDUCTED BY: Pearson et al. (2003, see reference below)

CASE STUDY PREPARED BY: David Lane and Emily Zitek

OVERVIEW

This study investigated the cognitive effects of stimulant medication in children with mental retardation and Attention-Deficit/Hyperactivity Disorder. This case study shows the data for the Delay of Gratification (DOG) task. Children were given various dosages of a drug, methylphenidate (MPH) and then completed this task as part of a larger battery of tests. The order of doses was counterbalanced so that each dose appeared equally often in each position. For example, six children received the lowest dose first, six received it second, etc. The children were on each dose one week before testing.

This task, adapted from the preschool delay task of the Gordon Diagnostic System (Gordon, 1983), measures the ability to suppress or delay impulsive behavioral responses. Children were told that a star would appear on the computer screen if they waited "long enough" to press a response key. If a child responded sooner than four seconds after their previous response, they did not earn a star, and the 4-second counter restarted. The DOG differentiates children with or without ADHD of normal intelligence (e.g., Mayes et al., 2001), and is sensitive to MPH treatment in these children (Hall Kataria, 1992).

QUESTIONS TO ANSWER

Does higher dosage lead to higher cognitive performance (measured by the number of correct responses to the DOG task)?

DESIGN ISSUES

This is a repeated-measures design because each participant performed the task after each dosage.

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>d0</td>
<td>Number of correct responses after taking a placebo</td>
</tr>
<tr>
<td>d15</td>
<td>Number of correct responses after taking .15 mg/kg of the drug</td>
</tr>
<tr>
<td>d30</td>
<td>Number of correct responses after taking .30 mg/kg of the drug</td>
</tr>
<tr>
<td>d60</td>
<td>Number of correct responses after taking .60 mg/kg of the drug</td>
</tr>
</tbody>
</table>

REFERENCES


http://onlinestatbook.com/2/case_studies/adhd.html

3/17/2016
EXERCISES
1. What is the independent variable of this experiment? How many levels does it have?
2. What is the dependent variable? On what scale (nominal, ordinal, interval, ratio) was it measured?
3. Create a line graph of the data. Do certain dosages appear to be more effective than others?
4. What is the mean number of correct responses of the participants after taking the placebo (0 mg/kg)?
5. Create a stem and leaf plot of the number of correct responses of the participants after taking the placebo (d0 variable). What can you say about the shape of the distribution?
6. Create a normal quantile plot of d0. What does it tell you and what aspect of the plot reveals this information?
7. What are the standard deviation and the interquartile range of the d0 condition?
8. What is the correlation between the participants' correct number of responses after taking the placebo and their correct number of responses after taking 0.60 mg/kg of MPH?
9. Perform a repeated measures ANOVA.
   a. How many degrees of freedom does the independent variable have?
   b. What are the F and p values?
   c. Does the performance of the children on the DOG task differ depending on the dose?
10. Test the difference between the d0 and d60 conditions using a sign test.
word or phrase like “dear” and “read” or “funeral” and “real fun”) or scores on a vocabulary test?

Questions from Case Studies

Angry Moods (AM) case study

13. (AM) Which variables are the participant variables? (They act as independent variables in this study.)

14. (AM) What are the dependent variables?

15. (AM) Is Anger-Out a quantitative or qualitative variable?

Teacher Ratings (TR) case study

16. (TR) What is the independent variable in this study?

ADHD Treatment (AT) case study

17. (AT) What is the independent variable of this experiment? How many levels does it have?

18. (AT) What is the dependent variable? On what scale (nominal, ordinal, interval, ratio) was it measured?
a. Inmates in a correctional facility are offered good behavior credit in return for participation in a study.
b. A research study is designed to investigate a new children's allergy medication.
c. Participants in a study are told that the new medication being tested is highly promising, but they are not told that only a small portion of participants will receive the new medication. Others will receive placebo treatments and traditional treatments.

**HOMEWORK**

1.1 Definitions of Statistics, Probability, and Key Terms

For each of the following eight exercises, identify: a. the population, b. the sample, c. the parameter, d. the statistic, e. the variable, and f. the data. Give examples where appropriate.

42. A fitness center is interested in the mean amount of time a client exercises in the center each week.

43. Ski resorts are interested in the mean age that children take their first ski and snowboard lessons. They need this information to plan their ski classes optimally.

44. A cardiologist is interested in the mean recovery period of her patients who have had heart attacks.

45. Insurance companies are interested in the mean health costs each year of their clients, so that they can determine the costs of health insurance.

46. A politician is interested in the proportion of voters in his district who think he is doing a good job.

47. A marriage counselor is interested in the proportion of clients she counsels who stay married.

48. Political pollsters may be interested in the proportion of people who will vote for a particular cause.

49. A marketing company is interested in the proportion of people who will buy a particular product.

Use the following information to answer the next three exercises: A Lake Tahoe Community College instructor is interested in the mean number of days Lake Tahoe Community College math students are absent from class during a quarter.

50. What is the population she is interested in?
   a. all Lake Tahoe Community College students
   b. all Lake Tahoe Community College English students
   c. all Lake Tahoe Community College students in her classes
   d. all Lake Tahoe Community College math students

51. Consider the following:
   \( X = \) number of days a Lake Tahoe Community College math student is absent

   In this case, \( X \) is an example of a:

   a. variable.
   b. population.
   c. statistic.
   d. data.

52. The instructor's sample produces a mean number of days absent of 3.5 days. This value is an example of a:

   a. parameter.
   b. data.
   c. statistic.
   d. variable.

1.2 Data, Sampling, and Variation in Data and Sampling

For the following exercises, identify the type of data that would be used to describe a response (quantitative discrete, quantitative continuous, or qualitative), and give an example of the data.

53. number of tickets sold to a concert

54. percent of body fat

55. favorite baseball team

56. time in line to buy groceries

57. number of students enrolled at Evergreen Valley College

58. most-watched television show
59. brand of toothpaste
60. distance to the closest movie theatre
61. age of executives in Fortune 500 companies
62. number of competing computer spreadsheet software packages

Use the following information to answer the next two exercises: A study was done to determine the age, number of times per week, and the duration (amount of time) of resident use of a local park in San Jose. The first house in the neighborhood around the park was selected randomly and then every 8th house in the neighborhood around the park was interviewed.

63. “Number of times per week” is what type of data?
   a. qualitative
   b. quantitative discrete
   c. quantitative continuous

64. “Duration (amount of time)” is what type of data?
   a. qualitative
   b. quantitative discrete
   c. quantitative continuous

65. Airline companies are interested in the consistency of the number of babies on each flight, so that they have adequate safety equipment. Suppose an airline conducts a survey. Over Thanksgiving weekend, it surveys six flights from Boston to Salt Lake City to determine the number of babies on the flights. It determines the amount of safety equipment needed by the result of that study.
   a. Using complete sentences, list three things wrong with the way the survey was conducted.
   b. Using complete sentences, list three ways that you would improve the survey if it were to be repeated.

66. Suppose you want to determine the mean number of students per statistics class in your state. Describe a possible sampling method in three to five complete sentences. Make the description detailed.

67. Suppose you want to determine the mean number of cans of soda drunk each month by students in their twenties at your school. Describe a possible sampling method in three to five complete sentences. Make the description detailed.

68. List some practical difficulties involved in getting accurate results from a telephone survey.

69. List some practical difficulties involved in getting accurate results from a mailed survey.

70. With your classmates, brainstorm some ways you could overcome these problems if you needed to conduct a phone or mail survey.

71. The instructor takes her sample by gathering data on five randomly selected students from each Lake Tahoe Community College math class. The type of sampling she used is
   a. cluster sampling
   b. stratified sampling
   c. simple random sampling
   d. convenience sampling

72. A study was done to determine the age, number of times per week, and the duration (amount of time) of residents using a local park in San Jose. The first house in the neighborhood around the park was selected randomly and then every eighth house in the neighborhood around the park was interviewed. The sampling method was:
   a. simple random
   b. systematic
   c. stratified
   d. cluster

73. Name the sampling method used in each of the following situations:
   a. A woman in the airport is handing out questionnaires to travelers asking them to evaluate the airport’s service. She does not ask travelers who are hurrying through the airport with their hands full of luggage, but instead asks all travelers who are sitting near gates and not taking naps while they wait.
   b. A teacher wants to know if her students are doing homework, so she randomly selects rows two and five and then calls on all students in row two and all students in row five to present the solutions to homework problems to the class.
   c. The marketing manager for an electronics chain store wants information about the ages of its customers. Over the next two weeks, at each store location, 100 randomly selected customers are given questionnaires to fill out asking for information about age, as well as about other variables of interest.
   d. The librarian at a public library wants to determine what proportion of the library users are children. The librarian has a tally sheet on which she marks whether books are checked out by an adult or a child. She records this data for every fourth patron who checks out books.
"The percentage of people we interview – out of all we try to interview – has been declining over the past decade or more."[7]

a. What are some reasons for the decline in response rate over the past decade?
b. Explain why researchers are concerned with the impact of the declining response rate on public opinion polls.

1.3 Frequency, Frequency Tables, and Levels of Measurement

Fifty part-time students were asked how many courses they were taking this term. The (incomplete) results are shown below:

<table>
<thead>
<tr>
<th># of Courses</th>
<th>Frequency</th>
<th>Relative Frequency</th>
<th>Cumulative Relative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>30</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1.33 Part-time Student Course Loads

a. Fill in the blanks in Table 1.33.

b. What percent of students take exactly two courses?
c. What percent of students take one or two courses?

81. Sixty adults with gum disease were asked the number of times per week they used to floss before their diagnosis. The (incomplete) results are shown in Table 1.34.

<table>
<thead>
<tr>
<th># Flossing per Week</th>
<th>Frequency</th>
<th>Relative Frequency</th>
<th>Cumulative Relative Freq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>27</td>
<td>0.4500</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td>0.9333</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>0.0500</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>0.0167</td>
<td></td>
</tr>
</tbody>
</table>

Table 1.34 Flossing Frequency for Adults with Gum Disease

a. Fill in the blanks in Table 1.34.
b. What percent of adults flossed six times per week?
c. What percent flossed at most three times per week?

82. Nineteen immigrants to the U.S were asked how many years, to the nearest year, they have lived in the U.S. The data are as follows: 2; 5; 7; 2; 2; 10; 20; 15; 0; 7; 0; 20; 5; 12; 15; 12; 4; 5; 10.

Table 1.35 was produced.

<table>
<thead>
<tr>
<th>Data</th>
<th>Frequency</th>
<th>Relative Frequency</th>
<th>Cumulative Relative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2</td>
<td>$\frac{2}{19}$</td>
<td>0.1053</td>
</tr>
</tbody>
</table>

Table 1.35 Frequency of Immigrant Survey Responses


This content is available for free at http://cnx.org/content/col11562/1.17
### Data Frequency Relative Frequency Cumulative Relative Frequency
2 | 3 | \(\frac{3}{19}\) | 0.2632 |
4 | 1 | \(\frac{1}{19}\) | 0.3158 |
5 | 3 | \(\frac{3}{19}\) | 0.4737 |
7 | 2 | \(\frac{2}{19}\) | 0.5789 |
10 | 2 | \(\frac{2}{19}\) | 0.6842 |
12 | 2 | \(\frac{2}{19}\) | 0.7895 |
15 | 1 | \(\frac{1}{19}\) | 0.8421 |
20 | 1 | \(\frac{1}{19}\) | 1.0000 |

Table 1.35 Frequency of Immigrant Survey Responses

**a.** Fix the errors in Table 1.35. Also, explain how someone might have arrived at the incorrect number(s).

**b.** Explain what is wrong with this statement: “47 percent of the people surveyed have lived in the U.S. for 5 years.”

**c.** Fix the statement in b to make it correct.

**d.** What fraction of the people surveyed have lived in the U.S. five or seven years?

**e.** What fraction of the people surveyed have lived in the U.S. at most 12 years?

**f.** What fraction of the people surveyed have lived in the U.S. fewer than 12 years?

**g.** What fraction of the people surveyed have lived in the U.S. from five to 20 years, inclusive?

### How much time does it take to travel to work? Table 1.36 shows the mean commute time by state for workers at least 16 years old who are not working at home. Find the mean travel time, and round off the answer properly.

| 24.0 | 24.3 | 25.9 | 18.9 | 27.5 | 17.9 | 21.8 | 20.9 | 16.7 | 27.3 |
| 24.7 | 24.6 | 28.1 | 24.9 | 22.6 | 23.9 | 18.0 | 31.4 | 22.3 | 24.0 | 25.5 |
| 24.7 | 24.6 | 28.1 | 24.9 | 22.6 | 23.9 | 26.0 | 16.3 | 23.1 | 21.4 | 21.5 |
| 27.0 | 27.0 | 18.6 | 31.7 | 23.3 | 30.1 | 22.9 | 23.3 | 21.7 | 18.6 |

Table 1.36

### Forbes magazine published data on the best small firms in 2012. These were firms which had been publicly traded for at least a year, have a stock price of at least $5 per share, and have reported annual revenue between $5 million and $1 billion. Table 1.37 shows the ages of the chief executive officers for the first 60 ranked firms.

<table>
<thead>
<tr>
<th>Age</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>40–44</td>
<td>3</td>
</tr>
<tr>
<td>45–49</td>
<td>11</td>
</tr>
<tr>
<td>50–54</td>
<td>13</td>
</tr>
<tr>
<td>55–59</td>
<td>16</td>
</tr>
</tbody>
</table>

Table 1.37
<table>
<thead>
<tr>
<th>Age</th>
<th>Frequency</th>
<th>Relative Frequency</th>
<th>Cumulative Relative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>60-64</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>65-69</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>70-74</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1.37

What is the frequency for CEO ages between 54 and 65?
What percentage of CEOs are 65 years or older?
What is the relative frequency of ages under 50?
What is the cumulative relative frequency for CEOs younger than 55?
Which graph shows the relative frequency and which shows the cumulative relative frequency?

![Graph A](image)

![Graph B](image)

Use the following information to answer the next two exercises: Table 1.38 contains data on hurricanes that have made direct hits on the U.S. Between 1851 and 2004. A hurricane is given a strength category rating based on the minimum wind speed generated by the storm.

<table>
<thead>
<tr>
<th>Category</th>
<th>Number of Direct Hits</th>
<th>Relative Frequency</th>
<th>Cumulative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>109</td>
<td>0.3993</td>
<td>0.3993</td>
</tr>
<tr>
<td>2</td>
<td>72</td>
<td>0.2637</td>
<td>0.6630</td>
</tr>
<tr>
<td>3</td>
<td>71</td>
<td>0.2601</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>18</td>
<td></td>
<td>0.9890</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>0.0110</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

Total = 273

Table 1.38 Frequency of Hurricane Direct Hits

85. What is the relative frequency of direct hits that were category 4 hurricanes?
   a. 0.0768
   b. 0.0659
   c. 0.2601
   d. Not enough information to calculate

86. What is the relative frequency of direct hits that were AT MOST a category 3 storm?
   a. 0.3480
   b. 0.9231

This content is available for free at http://cnx.org/content/col11562/1.17
c. 0.2601

d. 0.3370

1.4 Experimental Design and Ethics

87. How does sleep deprivation affect your ability to drive? A recent study measured the effects on 19 professional drivers. Each driver participated in two experimental sessions: one after normal sleep and one after 27 hours of total sleep deprivation. The treatments were assigned in random order. In each session, performance was measured on a variety of tasks including a driving simulation. Use key terms from this module to describe the design of this experiment.

88. An advertisement for Acme Investments displays the two graphs in Figure 1.14 to show the value of Acme's product in comparison with the Other Guy's product. Describe the potentially misleading visual effect of these comparison graphs. How can this be corrected?

![Figure 1.14](image)

**Figure 1.14** As the graphs show, Acme consistently outperforms the Other Guys!

89. The graph in Figure 1.15 shows the number of complaints for six different airlines as reported to the US Department of Transportation in February 2013. Alaska, Pinnacle, and Airtran Airlines have far fewer complaints reported than American, Delta, and United. Can we conclude that American, Delta, and United are the worst airline carriers since they have the most complaints?

![Figure 1.15](image)

**Figure 1.15**

**BRINGING IT TOGETHER: HOMEWORK**

90. Seven hundred and seventy-one distance learning students at Long Beach City College responded to surveys in the 2010-11 academic year. Highlights of the summary report are listed in Table 1.39.
What percent of the students surveyed do not have a computer at home?

About how many students in the survey live at least 16 miles from campus?

If the same survey were done at Great Basin College in Elko, Nevada, do you think the percentages would be the same? Why?

91. Several online textbook retailers advertise that they have lower prices than on-campus bookstores. However, an important factor is whether the Internet retailers actually have the textbooks that students need in stock. Students need to be able to get textbooks promptly at the beginning of the college term. If the book is not available, then a student would not be able to get the textbook at all, or might get a delayed delivery if the book is back ordered.

A college newspaper reporter is investigating textbook availability at online retailers. He decides to investigate one textbook for each of the following seven subjects: calculus, biology, chemistry, physics, statistics, geology, and general engineering. He consults textbook industry sales data and selects the most popular nationally used textbook in each of these subjects. He visits websites for a random sample of major online textbook sellers and looks up each of these seven textbooks to see if they are available in stock for quick delivery through these retailers. Based on his investigation, he writes an article in which he draws conclusions about the overall availability of all college textbooks through online textbook retailers.

Write an analysis of his study that addresses the following issues: Is his sample representative of the population of all college textbooks? Explain why or why not. Describe some possible sources of bias in this study, and how it might affect the results of the study. Give some suggestions about what could be done to improve the study.

REFERENCES

1.1 Definitions of Statistics, Probability, and Key Terms

1.2 Data, Sampling, and Variation in Data and Sampling
Data from http://www.bookofodds.com/Relationships-Society/Articles/A0374-How-George-Gallup-Picked-the-President