Answers to Comprehension Questions, Chapter 1

1. Chapter 1 distinguished between two different kinds of samples:

A. *Random samples* (selected randomly from a clearly defined population)

B. *Accidental or convenience* samples

a. Which type of sample (A or B) is more commonly reported in journal articles?

Answer: This depends on the type of journal. In laboratory based studies in psychology, for example, accidental or convenience samples are much more common. Reports from nationally representative surveys (e.g. Harris, Gallup) generally are based on random samples (and other procedures such as stratified sampling may be used to improve representativeness).

b. Which type of sample (A or B) is more likely to be representative of a clearly defined population?

Answer: A, random samples selected from a clearly defined population

c. What does it mean to say that a sample is "representative" of a population?

Answer: A sample is "representative" if the composition of the sample is similar to the composition of the population. For example, researchers generally want the age and gender composition of the sample to be similar to the population about which generalizations are to be made. Representativeness should be assessed for variables that are relevant (for example, a survey about voting intentions should be based on a sample that has a representative composition in terms of political party affiliation).

2. Suppose that a researcher tests the safety and effectiveness of a new drug on a convenience sample of male medical students between the ages of 24 and 30. If the drug appears to be effective and safe for this convenience sample, can the researcher safely conclude that the drug would be safe for women, children, and persons older than 70 years of age? Give reasons for your answer.

Answer: No, it is risky to assume that results obtained from a sample between ages 24 and 30 would apply to populations that differ greatly from the sample in age composition.

3. Given below are two applications of statistics. Identify which one of these is *descriptive* and which is *inferential* and explain why.

> *Case I:* An administrator at Corinth College looks at the verbal Scholastic Aptitude Test (SAT) scores for the entire class of students admitted in the fall of 2005 (mean = 660) and the verbal SAT scores for the entire class admitted in the fall of 2004 (mean = 540) and concludes that the class of students admitted to Corinth in 2005 had higher verbal scores than the class of students admitted in 2004.

Answer: Case I is descriptive, because the administrator used the data only to talk about the test performance of students in the two classes; the administrator did not make inferences about test performance for people not actually included in the two classes.

> *Case II:* An administrator takes a random sample of 45 Corinth College students in the fall of 2005 and asks them to self-report how often they engage in binge drinking. Members of the sample report an average of 2.1 binge drinking episodes per week. The administrator writes a report that says, "The average level of binge drinking among all Corinth College students is about 2.1 episodes per week."

Answer: Case II is inferential, because data were collected for a subset of Corinth students, but conclusions were drawn about drinking among all students at Corinth.

4. We will distinguish between two types of variables: *categorical* versus *quantitative*. For your answers to this question *do not* use any of the variables used in the textbook or in class as examples; think up your own example.

a. Give an example of a specific variable that is clearly a categorical type of measurement (e.g., Gender coded 1 = female and 2 = male is a categorical variable). Include the groups or levels (in the example given in the text, when Gender is the categorical variable, the groups are female and male).

This question has many possible answers.

b. Give an example of a specific variable that is clearly a quantitative type of measurement (e.g., IQ scores).

This question has many possible answers.

c. If you have scores on a categorical variable (e.g., Religion coded 1 = Catholic, 2 = Buddhist, 3 = Protestant, 4 = Islamic, 5 = Jewish, 6 = other religion), would it make sense to use these numbers to compute a mean? Give reasons for your answer.

Answer: No, because numbers serve only as labels for religious group membership. A "mean" religion computed by summing these scores would not be an interpretable number.

5. Using the guidelines given in this chapter (and Table 1.4), name an appropriate statistical analysis for each of the following imaginary studies. Also, state which variable is the predictor or independent variable and which variable is the outcome or dependent variable in each situation. Unless otherwise stated, assume that any categorical independent variable is between-S (rather than within-S), and assume that the conditions required for the use of parametric statistics are satisfied.

Case I: A researcher measures core body temperature for participants who are either male or female (i.e., Gender is a variable in this study). What statistics could the researcher use to see if mean body temperature differs between women and men?

Answer: independent samples t test; gender is the independent variable; core body temperature is the dependent variable.

Case II: A researcher who is doing a study in the year 2000 obtains yearbook photographs of women who graduated from college in 1970. For each woman, there are two variables. A rating is made of the "happiness" of her facial expression in the yearbook photograph. Each woman is contacted and asked to fill out a questionnaire that yields a score (ranging from 5 to 35) that measures her "life satisfaction" in the year 2000. The researcher wants to know whether "life satisfaction" in 2000 can be predicted from the "happiness" in the college yearbook photo back in 1970. Both variables are quantitative, and the researcher expects them to be linearly related, such that higher levels of happiness in the 1970 photograph will be associated with higher scores on life satisfaction in 2000. (Research on these variables has actually been done; see Harker & Keltner, 2001.)

Answer: Pearson correlation; observer rated "happiness" in the 1970 yearbook photograph is the independent variable (because this is measured at an earlier point in time); life satisfaction reported in 2000 is the dependent variable (because this is measured at a later point in time.)

Case III: A researcher wants to know whether preferred type of tobacco use (coded 1 = no tobacco use, 2 = cigarettes, 3 = pipe, 4 = chewing tobacco) is related to Gender (coded 1 = male, 2 = female).

Answer: set up a contingency table to report how many persons fall into these 8 groups (e.g., male cigarette smokers, females who do not use tobacco); do a Chi squared test of association (χ^2) or some other test of association that is used with contingency tables.

	No tobacco	**Cigarettes**	**Pipe**	**Chew**
Male				
Female				

Case IV: A researcher wants to know which of five drug treatments (Group 1 = Placebo, Group 2 = Prozac, Group 3 = Zoloft, Group 4 = Effexor, Group 5 = Celexa) is associated with the lowest mean score on a quantitative measure of depression (the Hamilton Depression Rating Scale).

Answer: Do a one way independent samples (Between-S) ANOVA. Independent variable is type of drug; dependent variable is depression score.

6. Draw a sketch of the standard normal distribution. What characteristics does this function have?

Answer: See Figure 1.4 for the shape of the distribution. The normal distribution is "bell shaped" and symmetric, with tails that gradually taper off at both ends of the distribution. There is a fixed relation between distance from the mean (in standard deviation units) and area under the curve, for example, approximately 68% of the area lies within 1 standard deviation above and below the mean. Although the tails of the mathematical function that defines the ideal normal distribution are infinite, the area under the curve is finite, and it can be scaled to equal 1.00 (or 100%).

7. Look at the standard normal distribution in Figure 1.4 to answer the following questions:

a. Approximately what proportion (or percentage) of scores in a normal distribution lie within a range from –2σ below the mean to +1σ above the mean?
Answer: .1359 + .3413 + .3413 = .8183 or about 81.83%

b. Approximately what percentage of scores lie above +3σ?

Answer: .0013 or about .13%

c. What percentage of scores lie inside the range –2σ below the mean to +2σ

above the mean? What percentage of scores lie outside this range?

Answer: about 95.44% of scores lie within the range from -2s below the mean to +2s above the mean; about (100 – 95.44) = 4.56% of the scores lie outside this range.

8. For what types of data would you use nonparametric versus parametric

statistics?

Answer: researchers prefer to use "nonparametric" statistics such as the Wilcoxon rank sum test, χ^2, Spearman r, or Kruskal- Wallis test when:

a. *Scores on the dependent variable are nominal or ordinal.*
b. *Scores on quantitative variables are not normally distributed.*
c. *The N of participants within each group is small.*
d. *The variances of scores are not equal across groups.*

Use of the parametric statistics described in this book (such as the independent samples t test, Pearson r, and ANOVA) generally require that:

a. *Scores on the dependent variable are interval-level ratio of measurement (in practice, however, many researchers do not enforce this requirement strictly; it is common to apply parametric statistics to scores obtained using 5 o5 7 point rating scales, in spite of the fact that these scores probably do not satisfy the requirements for equal interval level of measurement.*
b. *Scores on quantitative variables should be approximately normally distributed.*
c. *The N of participants within each group should be reasonably large.*
d. *Variances of scores should be roughly equal across groups (however, violations of the equal variance assumption do not typically cause serious problems unless they occur together with very small Ns in the groups, and/or the Ns are unequal across groups).*

In practice, many researchers prefer to use parametric statistics unless the assumptions for these are seriously violated.

9. What features of an experiment help us to meet the conditions for causal inference?

Answer: Temporal Precedence of X relative to Y: The researcher first manipulates the X "causal" variable and then subsequently measures or observes the Y outcome variable.

Covariation of X and Y: Researchers typically use statistics to assess whether scores on X and Y tend to covary or "go together".

Rule out all confounds or rival explanations: In experiments, researchers arrange the research situation so that other variables are held constant, or at least, kept the same across groups.

Even when these conditions are met, researchers cannot interpret the outcome of a single study as "proof" of causation. The results of any one study may be influenced by many types of error.

10. Briefly, what is the difference between internal and external validity?

Answer: A study has strong internal validity when the researcher can manipulate the variable that is thought to be "causal", control for nuisance variables, make sure that no other variables are confounded with the X variable, and the researcher can rule out alternative explanations for any observed differences in scores on Y. In other words, experiments can potentially have strong internal validity, if they are well designed.

11. For each of the following sampling methods, indicate whether it is random (i.e., whether it gives each member of a specific well-defined population an equal chance of being included) or accidental/convenience sampling. Does it involve stratification; that is, does it involve sampling from members of subgroups, such as males and females or members of various political parties? Is a sample obtained in this manner likely to be representative of some well-defined larger population? Or is it likely to be a biased sample, a sample that does not correspond to any clearly defined population?

a. A teacher administers a survey on math anxiety to the 25 members of her introductory statistics class. The teacher would like to use the results to make inferences about the average levels of math anxiety for all first-year college students in the United States.

Answer to 11a: this is a convenience or accidental sample; it does not involve stratified sampling; the sample is not likely to be representative of all first year college students.

b. A student gives out surveys on eating disorder symptoms to members of her teammates in gymnastics. She wants to be able to use her results to describe the correlates of eating disorders in all college women.

Answer to 11b: this is a convenience or accidental sample; it does not involve stratified sampling; the sample is not likely to be representative of all college women. Because gymnasts are at higher risk for eating disorder than college women in general, this would be a biased sample.

c. A researcher sends out 100,000 surveys on problems in personal relationships to mass mailing lists of magazine subscribers. She gets back about 5,000 surveys and writes a book in which she argues that the information provided by respondents is informative about the relationship problems of all American women.

Answer to 11c: In spite of the large number of surveys mailed out and returned, this is a convenience or accidental sample and not a random sample. It is likely that women who had stronger concerns and perhaps more negative attitudes toward men might have been

more likely to respond to this survey, than women who did not have negative attitudes toward men.

d. The Nielson television ratings organization selects a set of telephone area codes to make sure that its sample will include people taken from every geographical region that has its own area code; it then uses a random number generator to get an additional seven-digit telephone number. It calls every telephone number generated using this combination of methods (all area codes included and random dialing within an area code). If the person who answers the telephone indicates that this is a residence (not a business), that household is recruited into the sample, and the family members are mailed a survey booklet about their television-viewing habits.

Answer for 11d. This should be a fairly representative national sample; it is stratified by area code and involves random sampling within area code; however, people who don't have land-line telephones cannot be included, so the survey may systematically exclude people who don't have a land line telephone or who only use a cell phone.

12. Give an example of a specific sampling strategy that would yield a random sample of 100 students taken from the incoming freshman class of a state university. You may assume that you have the names of all 1,000 incoming freshmen on a spreadsheet.

Answer: You could generate a column of 1000 random numbers in the Excel spreadsheet, next to the column of 1000 names, and then use the last digit of the random numbers to select participants. You could select one digit at random (for example: 5) and contact all persons whose name is next to a random number that ends in "5".

13. Give an example of a specific sampling strategy that would yield a random sample of 50 households from the population of 2,000 households in that particular town. You may assume that the telephone directory for the town provides a nearly complete list of the households in the town.

Answer: This could be done in a number of ways. The researcher could randomly select pages of the telephone directory, and then randomly select one or two names from each page, by using a random number table. E.g. if the telephone directory has 99 pages, you could look at two digit random numbers. If you find this series of random numbers in your random number table: 57 11 21 82 etc. you could go to page 57 and take the 11th name, go to page 21 and take the 82nd name (if there is one), and so forth, until you have selected 50 names. Or, if all people in the town have the same area code and the same three digits for the telephone number (for example, 281-999-xxxx) you could use a random number table to identify the last four digits, for example, if you find 57 11 in the random number table, dial 281-999-5711 and find out if the telephone is a household or business, and whether it is in the town from which you want to sample, then attempt to recruit the person who answers the telephone.

14. Is each of the following variables categorical or quantitative? (Note that some variables could be treated as either categorical or quantitative.)

Number of children in a family

Answer: quantitative, but could be treated as categorical

Type of pet owned: 1 = none, 2 = dog, 3 = cat, 4 = other animal

Answer: categorical

IQ score

Answer: quantitative

Personality type (Type A, coronary prone; Type B, not coronary prone)

Answer: categorical

15. Do most researchers still insist on at least interval level of measurement as a condition for the use of parametric statistics?

Answer: No, in practice, many researchers apply parametric statistics such as the t test, Pearson r and ANOVA to scores obtained using 5 and 7 point rating scales; these scores probably fall short of satisfying strict requirements for interval level of measurement (as defined by Stevens). Authors cited in Section 1.6 argue that this practice is acceptable, however, some statisticians continue to adhere more strictly to the levels of measurement requirements outlined by Stevens.

16. How do categorical and quantitative variables differ?

Answer: When a variable is categorical (or nominal), the score serves only as a label for group membership. It would be nonsense to sum these scores and compute a mean, for example, it would be nonsense to compute a mean for scores on religion with scores of 1 = Buddhist, 2 = Catholic, 3 = Islamic, 4 = Jewish, 5 = Protestant, and 6 = Other religion. For categorical data we can examine the number (and proportion or percentage) of cases in each group, but it would be nonsense to compute a mean for scores on religion category. When a variable is quantitative, the score provides information about the location of each participant on some underlying characteristic (such as height). It is appropriate to compute means on scores for quantitative variables, such as height.

17. How do between-S versus within-S designs differ? Make up a list of names for imaginary participants, and use these names to show an example of between-S groups and within-S groups.

Answer: In a Between-S design each participant is assigned to, or observed in, one and only one group. In a Within-S design, each participant is tested or observed at several points in time and/or given several different treatments. The tables below show concrete examples.

Between – S or Independent Samples Design

Drug A	*Drug B*	*Drug C*
Bob	*Jane*	*John*
Carol	*Ken*	*Ed*
Ted	*Brenda*	*Cathy*
Alice	*Jack*	*Sandy*
Kim	*Linda*	*Mary*
Fred	*Vivian*	*Aidan*

Within S or Repeated Measures Design:

Drug A	*Drug B*	*Drug C*
Bob	*Bob*	*Bob*
Carol	*Carol*	*Carol*
Ted	*Ted*	*Ted*
Alice	*Alice*	*Alice*
Kim	*Kim*	*Kim*

Fred	*Fred*	*Fred*

18. When a researcher has an accidental or convenience sample, what kind of population can he or she try to make inferences about?

Answer: The ability to make inferences about populations may be quite limited when a researcher uses a convenience sample, because the convenience sample was not selected from any "real" population. At best, results from a convenience sample may be generalized (with caution) to populations that are similar in composition to the convenience sample (as argued by Campbell in his definition of the principal of "proximal similarity"). For example, a researcher who uses a convenience sample of males ages 18-22 should limit generalization of results to males in this age group; it would be risky to assume that results from this sample would be applicable to women or to persons from younger or older age groups.

19. Describe the guidelines for selecting an appropriate bivariate statistic. That is, what do you need to ask about the nature of the data and the research design in order to choose an appropriate analysis?

Answer:

a. What type of independent and dependent variables?

If both variables are categorical, set up a contingency table and do a test of association such as a χ^2; if both variables are quantitative, interval-ratio and linearly related, do a Pearson r; if both variables are ordinal, use Spearman r; if one variable is categorical and the other variable is quantitative, go on to consider other features of the design.

b. *Decide whether you need to use parametric or nonparametric statistics (based on issues such as whether scores on quantitative variables are normally distributed).*
c. *Decide whether the groups that are being compared are independent/ (Between – S) or non-independent (Within- S or repeated measures).*
d. *Note how many groups are being compared (only two groups, or more than two groups).*
e. *The decision tree shown in Table 1.4 can then be used to identify an appropriate statistical analysis.*