

Name \_\_\_\_\_

Key

Date \_\_\_\_\_

Pd \_\_\_\_\_

Statistics – Chapter 1 Test Review

Vocabulary

Sample	Systematic Sampling	Interval
Random Sampling	Qualitative Variables	Block
Simulation	Sampling Frame	Lurking Variables
Stratified Sampling	Ordinal	Individuals
Multistage Sampling	Randomized Experiment	Variables
Statistics	Replication	Census
Experiment	Non-sampling Error	Ratio
Control Group	Placebo Effect	Nominal
Double-blind Experiment	Sampling Error	Cluster Sampling
Population	Quantitative Variables	Convenience Sampling
Observational Study	Undercoverage	Confounding Variables

Use a word from the word bank to complete the sentence.

1. Systematic sampling is a sampling technique in which all members of the population are numbered sequentially. Then starting from a random point, every kth member is included in the sample.
2. Interval is a level of measurement that can be arranged in order and the differences between data values are meaningful, but there is no true zero.
3. experiment is when a treatment is deliberately imposed on individuals in order to observe a possible change in the response or the variable being measured.
4. control group is the group of individuals that does not receive the treatment, but may instead receive the "dummy treatment".
5. nominal is a level of measurement in which the data has no numerical value.
6. Sampling error is the difference between measurement from the sample and the corresponding measurement of the respective population.
7. census is a measurement or observation of the entire population.

8. Simulation is a numerical facsimile or representation of a real-world phenomenon.
9. population is ALL individuals of interest.
10. sample is only SOME of the individuals of interest.
11. double-blind is a type of experiment in which neither the doctor nor the patient is aware if they are providing/receiving the treatment or the "dummy treatment".
12. cluster sampling is a sampling technique in which the entire population is divided into pre-existing segments or clusters, often geographically, then every member from a specific cluster is used in the sample.
13. Qualitative sampling are variables that have no numerical value, only categorical.
14. Ordinal is a level of measurement in which the data can be arranged in order; however, the differences between data are meaningless.
15. multistage sampling is a sampling technique in which we use a variety of sampling methods to create smaller groups at each stage.
16. individuals the people or objects included in the study sample.
17. observational study is a type of study in which observations and measurement of individuals are conducted in a way that does not change the response or the variable being measured.
18. lurking variables are variables in which no data has been collected but nevertheless has an influence on other variables in the study.
19. convenience sampling is a sampling technique in which we create our sample using the data from the population members that are readily available.
20. quantitative variables are variables that have numerical value.
21. ratio is a level of measurement in which data can be arranged in order, differences and ratios of the data are meaningful because there is a true zero.
22. placebo effect occurs when a subject receives no treatment, but (incorrectly) believes that he or she is in fact receiving the treatment, and responds favorably.
23. random sampling is a basic sampling technique in which a group of subjects are chosen for the sample, and each individual has an equal chance of being selected.
24. replication is when we redo an experiment to be sure the results of the experiment did or did not happen by chance.
25. statistics is the study of how to collect, organize, analyze, and interpret numerical information from data.

26. stratified sampling is a sampling technique in which the population is divided into distinct subgroups based on specific characteristic, then random samples are drawn from each subgroup.
27. confounding variables are two variables in which the effects from each variable cannot be distinguished between each other.
28. non-sampling error occurs as a result of poor sample design, sloppy data collection, faulty measuring instruments, bias questions, and so on.
29. block is a group of individuals that share common features that might affect the treatment.
30. variables are the characteristics of the individuals to be measured or observed.
31. Sampling frame is the list of individuals from which the sample is selected.
32. randomized experiment is an experiment in which a random process is used to assign individuals to one of the treatments.
33. undercoverage is the result of omitting population members from the sample frame.

**Practice Problem s: Free Response-** justify your answers in complete sentences when applicable.

1. Does it make sense to fill in blanks square in a Sudoku puzzle using a random-number table?

the rules for sudoku require one # per row and column so random #'s would not suffice

2. Alesha wants to do a statistical study to determine how long it takes people to complete a Sudoku puzzle. Her plan is:

- ① Download 10 different puzzles, varying in difficulty, from the Internet.
- ② Find 10 friends willing to participate.
- ③ Ask each friend to complete one of the puzzles and time him/her.
- ④ Gather the completion times from each friend.

Describe some of the problems with Alesha's plan for the study. What method of sampling could she follow instead?

- ① Different levels means different difficulty = longer time
- ② "willing" friends introduces bias
- ③ Volunteer and who checks for accuracy
- ④ Self-timing is not uniform

3. You are conducting a study of students doing work-study jobs on your campus. Among the questions on the survey instrument are:

- A. How many hours are you scheduled to work each week? Answer to the nearest hour.
- B. How applicable is this work experience to your future employment goals? Respond using this scale: 1 = not at all, 2 = somewhat, 3 = very
- a) Suppose you take random samples from each grade level. What kind of sampling technique are you using? (SRS, stratified, systematic, cluster, multistage, or convenience)

stratified sampling

b) Describe the individuals of this study.

Students on campus w/ work-study jobs.

c) What is the variable for questions A? Classify the variable as qualitative or quantitative and state the level of measurement.

Hours scheduled; quantitative; ratio.

d) What is the variable for questions B? Classify the variable as qualitative or quantitative and state the level of measurement.

Rating of applicability of work experience to future employment; qualitative; ordinal.

4. A radio talk show host asked listeners to respond either yes or no to the question, "IS the candidate who spends the most on a campaign the most likely to win?" Fifteen people called in and nine said yes. What is the implied population? What is the variable? Can you detect any bias in the selection of the sample?

All listeners

opinion of a caller

yes, bias occurs from voluntary response

5. One cable station knows that approximately 30% of its viewers have DVR and can easily skip over advertising breaks. You are to design a simulation of how a random sample of seven station viewers would respond to the question, "Do you have DVR?" How would you assign the random digits 0-9 to the responses "YES" or "NO" to the DVR questions? Use your random-digit assignment and the random-number table to generate the responses from a random sample of seven station viewers.

0-2 = yes  
3-9 = No

Line 1 block 1  
N Y N N N Y N N

6. Categorize the type of sampling (list in 3a) used in each of the following situations.
- To conduct a pre-election opinion poll on a proposed amendment to the state constitution, a random sample of 10 telephone prefixes (first three digits of the phone number) was selected, and all households from the phone prefixes selected were called.

Cluster

- To conduct a study on depression among the elderly, a sample of 30 patients in one nursing home was used.

Convenience

- To maintain quality control in a brewery, every 20<sup>th</sup> bottle of beer coming off the production line was opened and tested.

Systematic

- Subscribers to the magazine *Sound Alive* were assigned numbers. Then a sample of 30 subscribers was selected by using a random-number table. The subscribers in the sample were invited to rate new Spotify for a "What the Subscribers Think" column.

Random

- To judge the appeal of a proposed TV sitcom, a random sample of 10 people from each of the three different age categories was selected and those chosen were asked to rate the pilot show.

Stratified

7. Which technique for gathering data (observational study or experiment) do you think was used in the following studies? Explain.

- The U.S. Census Bureau tracks population age. In 1900, the percentage of the population that was 19 years old or younger was 44.4%. In 1930, the percentage was 38.8%; in 1970, the percentage was 37.9%; and in 2000, the percentage in that age group was down to 28.5%.

Observational Study

- After receiving the same lessons, a class of 100 students was randomly divided into two groups of 50 each. One group was given a multiple-choice exam covering the material in the lessons. The average test scores for the two groups were then compared.

Experiment

8. How would you use a completely randomized experiment in each of the following settings? Is a placebo being used or not? Be specific and give details.

- a. A charitable nonprofit organization wants to test two methods of fundraising. From a list of 1000 past donors, half will be sent literature about the successful activities of the charity and asked to make another donation. The percentage of people from each group who make a new donation will be compared.

randomly select have to solicit by mail.  
no placebo

- b. A tooth-whitening gel is to be tested for effectiveness. A group of 85 adults have volunteered to participate in the study. Of these, 43 are to be given a gel that contains the tooth-whitening chemicals. The remaining 42 are to be given a similar-looking package of gel that does not contain the tooth-whitening chemicals. A standard method will be used to evaluate the whiteness of teeth for all participants. Then the results for the two groups will be compared. How could this experiment be designed to be double-blind?

randomly select 43 volunteers check results  
placebo used. DB: no one knows who received  
the placebo gel

- c. Consider the experiment described in part (a). Describe how you would use a randomized block experiment with blocks based on age. Use three blocks: donors under 30 years old, donors 30 to 59 years old, donors 60 and over.

Block based on age, randomly split block btwn  
mail/phone. Compare results No placebo.

9. Suppose you are conducting a study to compare firefly populations exposed to normal daylight/darkness conditions with firefly populations exposed to continuous light (24 hours a day). You set up two firefly colonies in a laboratory environment. The two colonies are identical except that one colony is exposed to normal daylight/darkness conditions and the other is exposed to continuous light. Each colony is populated with the same number of mature fireflies. After 72 hours, you count the number of living fireflies in each colony.

- a. Is this an experiment or an observation study? Explain.

experiment, since treatment is imposed on  
one colony.

- b. Is there a control group? Is there a treatment group?

control group receives normal daylight/  
darkness conditions. The treatment has  
24 hours per day of light.

- c. What is the variable in this study?

The # of fireflies living at the end  
of 72 hours

- d. What is the level of measurement (nominal, interval, ordinal, or ratio) of the variable?

Ratio