

AP Statistics
Ch. 4 Review #2

Name _____

1. Which of the following is a key distinction between well designed experiments and observational studies?

- (A) More subjects are available for experiments than for observational studies.
- (B) Ethical constraints prevent large-scale observational studies.
- (C) Experiments are less costly to conduct than observational studies.
- (D) An experiment can show a direct cause-and-effect relationship, whereas an observational study cannot.
- (E) Tests of significance cannot be used on data collected from an observational study.

2. A high school statistics class wants to conduct a survey to determine what percentage of students in the school would be willing to pay a fee for participating in after-school activities. Twenty students are randomly selected from each of the freshman, sophomore, junior and senior classes to complete the survey. This plan is an example of which type of sampling?

- (A) Cluster
- (B) Convenience
- (C) Simple random
- (D) Stratified random
- (E) Systematic

3. Jason wants to determine how age and gender are related to political party preference in his town. Voter registration lists are stratified by gender and age-group. Jason selects a simple random sample of 50 men from the 20 to 29 age-group and records their age, gender, and party registration (Democratic, Republican, neither). He also selects an independent simple random sample of 60 women from the 40 to 49 age-group and records the same information. Of the following, which is the most important observation about Jason's plan?

- (A) The plan is well conceived and should serve the intended purpose.
- (B) His samples are too small.
- (C) He should have used equal sample sizes.
- (D) He should have randomly selected the two age groups instead of choosing them non-randomly.
- (E) He will be unable to tell whether a difference in party affiliation is related to differences in age of to the difference in gender.

4. A study of existing records of 27,000 automobile accidents involving children in Michigan found that about 10 percent of children who were wearing a seatbelt (group SB) were injured and that about 15 percent of children who were not wearing a seatbelt (group NSB) were injured. Which of the following statements should NOT be included in a summary report about this study?

- (A) Driver behavior may be a potential confounding factor.
- (B) The child's location in the car may be a potential confounding factor.
- (C) This study was not an experiment, and cause-and-effect inferences are not warranted.
- (D) This study demonstrates clearly that seat belts save children from injury.
- (E) Concluding that seatbelts save children from injury is risky, at least until the study is independently replicated.

5. A new medication has been developed to treat sleep-onset insomnia (difficulty in falling asleep). Researchers want to compare this drug to a drug that has been used in the past by comparing the length of time it takes subjects to fall asleep. Of the following, which is the best method for obtaining this information?

- (A) Have subjects choose which drug they are willing to use, then compare the results.
- (B) Assign the two drugs to the subjects on the basis of their past sleep history without randomization, then compare the results.
- (C) Give the new drug to all subjects on the first night. Give the old drug to all subjects on the second night. Compare the results.
- (D) Randomly assign the subjects to two groups, giving the new drug to one group and no drug to the other group, and then compare results.
- (E) Randomly assign the subjects to two groups, giving the new drug to one group and the old drug to the other group, and then compare results.

6. In which of the following situations would it be most difficult to use a census?

- (A) To determine what proportion of licensed bicycles on a university campus have lights.
- (B) To determine what proportion of students in a high school support wearing uniforms.
- (C) To determine what proportion of registered students enrolled in a college are employed more than 20 hours each week.
- (D) To determine what proportion of single-family dwellings in a small town have two-car garages.
- (E) To determine what proportion of fish in Lake Michigan are bass.

7. A television news editor would like to know how local registered voters would respond to the question, "Are you in favor of the school bond measure that will be voted on in an upcoming special election?" A television survey is conducted during a break in the evening news by listing two telephone numbers side by side on the screen, one for viewers to call if they approve of the bond measure, and the other to call if they disapprove. This survey method could produce biased results for a number of reasons. Which one of the following is the most obvious reason?

- (A) It uses a stratified sample rather than a simple random sample.
- (B) People who feel strongly about the issue are more likely to respond.
- (C) Viewers should be told about the issues before the survey is conducted.
- (D) Some registered voters who call might not vote in the election.
- (E) The wording of the question is biased.

8. A researcher wishes to test a new drug developed to treat hypertension (high blood pressure). A group of 40 hypertensive men and 60 hypertensive women is to be used. The experimenter randomly assigns 20 of the men and 30 of the women to the placebo and assigns the rest to the treatment. The major reason for separate assignment for men and women is that

- (A) it is a large study with 100 subjects
- (B) the new drug may affect men and women differently
- (C) the new drug may affect hypertensive and non-hypertensive people differently
- (D) this design uses matched pairs to detect the new drug effect
- (E) there must be an equal number of subjects in both the placebo group and the treatment group

9. Which of the following is NOT a characteristic of stratified random sampling?

- (A) Random sampling is part of the sampling procedure.
- (B) The population is divided into groups of units that are similar on some characteristic.
- (C) The strata are based on facts known before the sample is selected.
- (D) Each individual unit in the population belongs to one and only one of the strata.
- (E) Every possible subset of the population, of the desired sample size, has an equal chance of being selected.

10. Automobile brake pads are either metallic or nonmetallic. An experiment is to be conducted to determine whether the stopping distance is the same for both types of brake pads. In previous studies, it was determined that car size (small, medium, large) is associated with stopping distance, but car type (sedan, wagon, coupe) is not associated with stopping distance. The experiment would be best done

- (A) by blocking on car size
- (B) by blocking on car type
- (C) by blocking on stopping distance
- (D) by blocking on brake pad type
- (E) without blocking

11. A group of students has 60 houseflies in a large container and needs to assign 20 to each of the three groups labeled A, B, and C for an experiment. They can capture the flies one at a time when the flies enter a side chamber in the container that is baited with food. Which of the following methods will be most likely to result in three comparable groups of 20 houseflies each?

- (A) Label the first 20 flies caught as Group A, the second 20 caught as group B, and the third 20 caught as group C.
- (B) Write the letters A, B, and C on separate slips of paper. Randomly pick one of the slips of paper and assign the first 20 flies caught to that group. Pick another slip and assign the next 20 flies caught to that group. Assign the remaining flies to the remaining group.
- (C) When each fly is caught, roll a die. If the die shows an even number, the fly is labeled A. If the die shows an odd number, the fly is labeled B. When 20 flies have been labeled A and 20 have been labeled B, the remaining flies are then labeled C.
- (D) Place each fly in its own numbered container (numbered from 1 to 60) in the order that it was caught. Write the numbers from 1 to 60 on slips of paper, put the slips in a jar, and mix them well. Pick 20 numbers out of the jar. Assign the flies in the containers with those numbers to group A. Pick 20 more numbers and assign the flies in the containers with those numbers to group B. Assign the remaining 20 flies to group C.
- (E) When each fly is caught, roll a die. If the die shows a 1 or 2, the fly is labeled A. If the die shows a 3 or 4, the fly is labeled B. If the die shows a 5 or 6, the fly is labeled C. Repeat this process for all 60 flies.

12. A large simple random sample of people aged nineteen to thirty living in the state of Colorado was surveyed to determine which of two MP3 players just developed by a new company was preferred. To which of the following populations can the results of this survey be safely generalized?

- (A) Only people aged nineteen to thirty living in the state of Colorado who were in this survey.
- (B) Only people aged nineteen to thirty living in the state of Colorado
- (C) All people living in the state of Colorado
- (D) Only people aged nineteen to thirty living in the United States
- (E) All people living in the United States

13. Publishers of a magazine wish to determine what proportion of the magazine's 50,000 subscribers are pleased with their subscription. The publishers intend to mail a survey to 1,000 subscribers randomly selected from those who have received the magazine for 5 years or more. This introduces selection bias, since long-subscribing customers are more likely to be pleased with their subscription. Which of the following would best eliminate selection bias?

- (A) Mail surveys to 2,000 subscribers randomly selected from those who have received the magazine for 5 years or more.
- (B) Mail surveys to 1,000 subscribers randomly selected from those who have received the magazine for 1 year or less.
- (C) Mail surveys to 1,000 subscribers randomly selected from all subscribers.
- (D) Mail surveys to 1,000 subscribers randomly selected from those who have received the magazine for 5 years or more and 1,000 subscribers randomly selected from those who have received the magazine for 1 year or less.
- (E) Mail surveys to 500 subscribers randomly selected from a group who have received a free six-month subscription within this past year.

14. A new restaurant is interested in determining the best time-temperature combination for roasting a five-pound cut of lamb. The times to be tested are 45 minutes, 60 minutes, and 90 minutes at temperature of 350 degrees Fahrenheit and 425 degrees Fahrenheit for each time, with the exception of the 90-minute-425 degree combination. That combination is being eliminated because it will overcook the lamb, which leaves five combinations remaining. From 10 identical cuts of lamb, 2 are randomly selected to roast using each of the time-temperature combinations in the same oven. The quality of the finished product is evaluated for each roast. Which of the following is true?

- (A) The explanatory variable is the quality of the finished lamb.
- (B) The response variable is the roasting temperature for the lamb.
- (C) If the experiment is repeated, identical results will be expected.
- (D) There should be a control group (i.e. a group in which no treatment is given).
- (E) The two cuts that are being roasted for each time-temperature combination are an example of replication.

15. A dog food company wishes to test a new high-protein formula for puppy food to determine whether it promotes faster weight gain than the existing formula for that puppy food. Puppies participating in an experiment will be weighed at weaning (when they begin to eat puppy food) and will be weighed at one-month intervals for one year. In designing this experiment, the investigators wish to reduce the variability due to natural differences in puppy growth rates. Which of the following strategies is most appropriate for accomplishing this?

- (A) Block on dog breed and randomly assign puppies to existing and new formula group within each breed.
- (B) Block on geographic location and randomly assign puppies within each breed. Then assign puppies by breed to either the existing or the new formula.
- (C) Stratify on dog breed and randomly sample puppies within each breed. Then assign puppies by breed to either the existing or the new formula.
- (D) Stratify on geographic location of the puppies and randomly sample puppies within each geographic area. Then assign puppies by geographic area to either the existing or the new formula.
- (E) Stratify on gender and randomly sample puppies within gender groups. Then assign puppies by gender to either the existing or the new formula.

16. The manager of a public swimming pool wants to compare the effectiveness of two laundry detergents, Detergent A and Detergent B, in cleaning the towels that are used daily. As each dirty towel is turned in, it is placed into the only washing machine on the premises. When the washing machine contains 20 towels, the manager flips a coin to determine whether Detergent A or Detergent B will be used for that load. The cleanliness of the load of towels is rated on a scale of 1 to 10 by a person who does not know which detergent was used. The manager continues this experiment for many days. Which of the following best describes the manager's study?

- (A) A completely randomized design
- (B) A randomized block design with Detergent A and Detergent B as blocks
- (C) A randomized block design with the washing machine as the block
- (D) A matched-pairs design with Detergent A and Detergent B as the pair
- (E) An observational study

17. A compact disc (CD) manufacturer wanted to determine which of two different cover designs for a newly released CD will generate more sales. The manufacturer chose 70 stores to sell the CD. Thirty-five of these stores were randomly assigned to sell CDs with one of the cover designs and the other 35 were assigned to sell the CDs with the other cover design. The manufacturer recorded the number of CDs sold at each of the stores and found a significant difference between the mean number of CDs sold for the two cover designs. Which of the following gives the conclusion that should be made based on the results and provides the best explanation for the conclusion?

- (A) It is not reasonable to conclude that the difference in sales was caused by the different cover designs because this was not an experiment.
- (B) It is not reasonable to conclude that the difference in sales was caused by the different cover designs because there was no control group for comparison.
- (C) It is not reasonable to conclude that the difference in sales was caused by the different cover designs because the 70 stores were not randomly chosen.
- (D) It is reasonable to conclude that the difference in sales was caused by the different cover designs because the cover designs were randomly assigned to stores.
- (E) It is reasonable to conclude that the difference in sales was caused by the different cover designs because the sample size was large.

18. In the design of a survey, which of the following best explains how to minimize response bias?

- (A) Increase the sample size.
- (B) Decrease the sample size.
- (C) Randomly select the sample.
- (D) Increase the number of questions in the survey.
- (E) Carefully word and field-test survey questions.

19. A polling firm is interested in surveying a representative sample of registered voters in the United States. The firm has automated its sampling so that random phone numbers within the United States are called. Each time a number is called, the procedure below is followed.

- If there is no response or if an answering machine is reached, another number is automatically called.
- If a person answers, a survey worker verifies that the person is at least 18 years of age.
- If the person is not at least 18 years of age, no response is recorded, and another number is called.
- If the person is at least 18 years of age, that person is surveyed.

Some people claim the procedure being used does not permit the results to be extended to all registered voters. Which of the following is NOT a legitimate concern about the procedure being used?

- (A) Registered voters with children under the age of 18 years may be underrepresented in the sample.
- (B) Registered voters with unlisted telephone numbers may be underrepresented in the sample.
- (C) Registered voters who have more than one telephone number may be overrepresented in the sample.
- (D) Registered voters who live in households consisting of more than one voter may be underrepresented.
- (E) People who are not registered to vote may bias the sample results.

20. A biologist is interested in studying the effect of growth-enhancing nutrients and different salinity (salt) levels in water on the growth of shrimps. The biologist has ordered a large shipment of young tiger shrimps from a supply house for use in the study. The experiment is to be conducted in a laboratory where 10 tiger shrimps are placed randomly into each of 12 similar tanks in a controlled environment. The biologist is planning to use 3 different growth-enhancing nutrients (A, B, and C) and two different salinity levels (low and high).

(a.) List the treatments the biologist plans to use in this experiment.

(b) Using the treatments listed in part (a), describe a completely randomized design that will allow the biologist to compare the shrimps' growth after 3 weeks.

(c) Give one statistical advantage to having only tiger shrimp in the experiment. Explain why this is an advantage.

(d) Give one statistical disadvantage to having only tiger shrimp in the experiment. Explain why this is a disadvantage.

Answers:

1. D 2. D 3. E 4. D 5. E 6. E 7. B
 8. B 9. E 10. A 11. D 12. B 13. C 14. E
 15. A 16. A 17. D 18. E 19. B

20. Part(a): The three different growth-enhancing nutrients (A, B, and C) and two different salinity levels (low and high) yield a total of $3 \times 2 = 6$ different treatment combinations.

Part (b): Since 10 tiger shrimps have already been randomly placed into each of 12 similar tanks in a controlled environment, we must randomly assign the treatment combinations to the tanks. Each treatment combination will be randomly assigned to 2 of the 12 tanks. One way to do this is to generate a random number for each tank. The treatment combinations are then assigned by sorting the random numbers from smallest to largest.

Treatment Combination	Nutrient	Salinity Level
1	A	Low
2	A	High
3	B	Low
4	B	High
5	C	Low
6	C	High

After three weeks the weight gain (after – before) is computed for each tank, and the treatments are compared using appropriate averages.

There are other methods of random assignment of treatments. Make sure randomization was used and how the treatments are assigned is explained fully.

Treatment Combination	Nutrient	Salinity Level	Tanks with
1	A	Low	Smallest and second smallest random numbers
2	A	High	Third and fourth smallest random numbers
3	B	Low	Fifth and sixth smallest random numbers
4	B	High	Seventh and eighth smallest random numbers
5	C	Low	Ninth and tenth smallest random numbers
6	C	High	Next to largest and largest random numbers

Part (c): Using only tiger shrimp will reduce a source of variation in the experimental units, the tanks of shrimp in this experiment. By eliminating this possible source of variation, type of shrimp, we are better able to isolate the variability due to the factors of interest to us (nutrient and salinity level). This will make it easier to identify any treatment effects that may be present.

Part (d): Using only tiger shrimp will limit the scope of inference for the biologist. Ideally, the biologist would like to identify the treatment combination that leads to the most growth for all shrimp. However, the biologist will only be able to identify the best treatment combination for tiger shrimp because other types of shrimp may respond differently to the treatments.