Multiple choice. Write the letter of the best available choice in the blank provided.

1. _____ A well-designed experiment should have which of the following characteristics?

     I. Subjects assigned randomly to treatments
     II. A control group or at least two treatment groups
     III. Replication

     A) I only   B) I and II only   C) I and III only   D) II and III only   E) I, II, and III

2. _____ Suppose your local school district decides to randomly test high school students for attention deficit disorder (ADD). There are three high schools in the district, each with grades 9 – 12. The school board pools all of the students together and randomly samples 250 students. Is this a simple random sample?

     A) Yes, because the students were chosen at random.
     B) Yes, because each student is equally likely to be chosen.
     C) Yes, because every possible combination of 250 students is equally likely to be chosen.
     D) No, because we can't guarantee that there are students from each school in the sample.
     E) No, because we can't guarantee that there are students from each grade in the sample.

3. _____ A food and dessert blogger is testing the efficiency of baking cupcakes in different oven and pan types. Four types of pans will be tested, and each type of pan will be tested in 3 types of ovens. How many total treatments are there in this study?

     A) Two   B) Three   C) Four   D) Seven   E) Twelve

4. _____ A chemistry professor at Podunk University (known as PU) who teaches a large lecture class surveys his students who attend his class. He wants to get some ideas about how he can make the class more interesting since he notices about 25% of the class is "absent" on any given day. During one class period, he distributes a survey for his students to complete. This survey method suffers from

     A) voluntary response   B) nonresponse bias   C) response bias   D) undercoverage   E) None of these

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

6. _____ To study a cause-and-effect relationship between variables, which of the following methods of data collection should be used?

    A) observational study

    B) sample survey

    C) census

    D) randomized experiment

7. _____ A researcher conducting an email survey is concerned about possible sources of bias. Of the following, which is the best example of nonresponse bias?

    A) The wording of the questions in the survey leads people to respond in a certain way.

    B) The behavior of the interviewer leads people to respond in a certain way.

    C) People might be uncomfortable with the survey questions and, as a result, might not always respond to those questions truthfully.

    D) Many of the people selected to participate in the survey who do not respond might have opinions different from those who do respond.

    E) People without access to the internet are overlooked in the sampling procedure used to determine who is surveyed.

8. _____ 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 is 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) Randomly assign the subjects to two groups, giving the new drug to one group and no drug to the other group, 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.

    D) Assign the two drugs to the subjects on the basis of their past sleep history without randomization, then compare the results.

    E) Randomly assign the subjects to two groups, giving the new drug to one group and the old drug to the other group, then compare the results.

9. _____ A regional transportation authority is interested in estimating the mean number of minutes working adults in the region spend commuting to work on a typical day. A random sample of working adults will be selected from each of three strata: urban, suburban, and rural. Selected individuals will be asked the number of minutes they spend commuting to work on a typical day. Why is stratification used in this situation?

    A) To remove bias when estimating the proportion of working adults living in urban, suburban, and rural areas

    B) To remove bias when estimating the mean commuting time

    C) To reduce bias when estimating the mean commuting time

    D) To decrease the variability in estimates of the proportion of working adults living in urban, suburban, and rural areas

    E) To decrease the variability in estimates of the mean commuting time

10. _____ Jacob has friends that will be attending college at Texas, Texas A&M, Texas Tech, and Oklahoma, and is interested in finding out which of those schools is the most popular in the entire state of Texas. To investigate, Jacob takes a systematic random sample of 76 people as they are leaving a Texas A&M football game in College Station, and asks each of them to pick their favorite school out of the four. 73 of the 76 people agree to participate in the survey. Which of the following conclusions is correct?

   A) Proper use of chance as evidenced by the systematic random sample makes this a well-designed survey.
   B) The high response rate makes this a well-designed survey.
   C) Selection bias makes this a poorly designed survey.
   D) A voluntary response study like this gives too much emphasis to persons with strong opinions.
   E) Lack of anonymity makes this a poorly designed survey.

11. _____ What fault do all these sampling designs have in common?

   I. The Wall Street Journal plans to make a prediction for a presidential election based on a survey of its readers.
   II. A radio talk show asks people to phone in their views on whether the United States should pay off its huge debt to the United Nations.
   III. A police detective, interested in determining the extent of drug use by teenagers, randomly picks a sample of high school students and interviews each one about any illegal drug use by the student during the past year.

   A) All the designs make improper use of stratification.
   B) All the designs have errors that can lead to strong bias.
   C) All the designs confuse *association* with *cause and effect.*
   D) None of the designs satisfactorily controls for sampling error.
   E) None of the designs makes use of chance in selecting a sample.

12. _____ A randomized block design will be used in an experiment to compare two lotions that protect people from getting sunburned. Which of the following should guide the formation of the blocks?

   A) Participants in the same block should receive the same lotion.
   B) Participants should be randomly assigned to the blocks.
   C) Participants should be kept blind as to which block they are in.
   D) Participants within each block should be as similar as possible with respect to how easily they get sunburned.
   E) Participants within each block should be as different as possible with respect to how easily they get sunburned.

13. _____ A newspaper wants to know public opinion of a town regarding the construction of a new library in a downtown location. It is decided that 48 people will be surveyed using a simple random sample. Which of the following is the best method to produce a simple random sample?

   A) Randomly select 12 people from each of the northwest, northeast, southwest, and southeast sections of the city.
   B) Survey every fourth person who enters the current library until 48 people have responded.
   C) Randomly select 48 people from the city phone directory.
   D) Number the residents using the census data. Use a random number generator to pick 48 people.
   E) Record the opinions of the first 48 people who visit the newspaper's web site.

FREE RESPONSE & SHORT ANSWER QUESTIONS

14. The Westwood Student Council wishes to estimate the percent of freshman that support eliminating class rank from high school transcripts. A survey given to a random sample of 64 freshman students reveals that 47% of the sample supports eliminating class rank. Suppose that, in fact, 52 percent of all Westwood freshman students support eliminating class rank.

Is 52% a parameter or statistic ? (circle answer)   What symbol should be used to represent it? _____

Is 47% a parameter or statistic ? (circle answer)   What symbol should be used to represent it? _____

15. According to recent census data, the mean number of children per American family is 2.6. In a sample survey of 200 families that were randomly selected from major U.S. cities, the mean number of children per family was 2.83.

Is "2.83" a parameter or statistic ? (circle answer)   What symbol should be used to represent it? _____

Is "2.6" a parameter or statistic ? (circle answer)   What symbol should be used to represent it? _____

16. A researcher wonders if meat in the diet may be a factor in high blood pressure. She selected a random sample of 80 adults in San Francisco, and compares the blood pressures of the 40 adults who choose to be vegetarians, to those of the 40 adults who choose to eat meat. The mean blood pressure of the 40 vegetarians was found to be statistically significantly lower than the mean blood pressure of the 40 meat-eaters.

a) Based on this study, is it reasonable to conclude that being a vegetarian will cause an adult to have lower blood pressure? Explain why or why not.

b) Can the results of this study be generalized to all adults in San Francisco? Explain.

c) The researcher notes that sodium intake – among other factors – may have been a confounding factor in this study. Carefully explain how sodium intake could be a confounding variable in this study.

17. High cholesterol levels in people can be reduced by exercise, diet, and medication. Twenty middle-aged males with cholesterol readings between 220 and 240 milligrams per deciliter (mg/dL) of blood at a large local hospital volunteered for this study. Ten of the 20 males were randomly assigned to group A, advised on appropriate exercise and diet, and also received a placebo. The other 10 males were assigned to group B, received the same advice on appropriate exercise and diet, but received a drug intended to reduce cholesterol instead of a placebo. After three months, the mean cholesterol reduction for the males in group B was significantly higher than for the males in group A.

    a) Would it be reasonable to conclude that taking this cholesterol drug (in addition to exercise and diet) will lead to a greater reduction in cholesterol level than exercise and diet alone? Explain.

    b) Can the results of this study be generalized to all such male patients at this large local hospital? Explain.

18. Athletes who had suffered hamstring injuries are to be randomly assigned to one of two exercise programs – program A (static stretching) and program B (agility and trunk stabilization). Researchers wish to see if the time it takes these athletes to be able to return to sports is different depending on which of the two exercise programs they are engaged in.

    a) A wise statistician suggests adding a third group: A control group of athletes who do not participate in either special exercise program. Explain the advantage of including such a group for this experiment.

    b) Suppose the researchers conducting this experiment decide to utilize a control group, and that 72 athletes are available for this study. Explain an appropriate process to randomly assign the 72 athletes to the three groups for a completely randomized experiment. For this experiment, it is required that the three groups each have an equal number of experiment units.

19. In the village of Podunk, a local parents group is concerned that the children in the community are not reading as much as kids in other communities. Thus, they have requested funding from the city for additional after-school reading programs, in hopes that this will encourage school children to read more books in their free time.

    The city has agreed to provide funding, but only if a survey of school children shows that the average number of books read per year by a school child in Podunk is less than the average for surrounding communities. If the city's survey shows that Podunk children are NOT reading fewer books per year than in other communities, then they will NOT provide funding for additional reading programs.

    The city gets an intern to conduct a survey of 50 school children in Podunk, and based on the data from this survey, the city will make a decision about funding for the requested reading programs.

    The intern visits a local bookstore where a children's book club is meeting that afternoon. There are 320 children at this book club meeting, and all of them find tremendous enjoyment in reading books in their free time. The intern decides to take a simple random sample of 50 of these children and survey them about how many books they have read in the past year, and this data is taken back to the city.

    a) Identify the population of interest in this situation.

    b) Identify the parameter of interest in this situation.

    c) Are there any issues with bias in the way the intern collected his data? Explain.

    d) After complaints from the parents group, the city decides to fire the intern and hire a high school Statistics student – named Alex – to conduct the survey. In order to get a precise estimate, Alex wishes to use a stratified sampling procedure. He is not sure, however, whether he should stratify his sample by gender, or by ethnicity/race. It is widely thought that – for whatever reason – girls in Podunk tend to read more books than boys. It is also thought that a child's ethnicity/race has very little association with the number of books that a child reads. Based on this information, which variable should Alex choose to stratify by: gender or ethnicity/race? Explain.

20. A nutritionist wants to study the effect of storage time (6, 12, and 18 months) on the amount of vitamin C present in dried fruit when stored for these lengths of time. Vitamin C is measured in milligrams per 100 milligrams of fruit. The researcher has a random sample of 60 total packages of dried fruit: 30 of them are dried apple, and the other 30 are dried peach.

a) Identify the experimental units:

the explanatory variable(s) or treatment(s):

the response variable(s):

b) Explain how you would carry out a completely randomized experiment for this study.

c) Describe the changes that would be made to your experiment in part (B) by incorporating blocking.

d) Can the nutritionist generalize his/her findings to all types of dried fruit? Explain.

Review Unit I – ANSWERS

1. E
2. C
3. E
4. D
5. C

6. D
7. D
8. E
9. E

10. C
11. B
12. D
13. D

14. 52% is a parameter; symbol: p

   47% is a statistic; symbol: $\hat{p}$

15. 2.83 is a statistic; symbol: $\bar{x}$
    2.6 is a parameter; symbol: μ

16. a) No – subjects CHOSE their type of diet, hence this is an observational study (from which cause and effect relationships cannot be justified)
    b) Yes, since she took a random sample of adults in San Francisco (the study might reveal an <u>association</u> between type of diet and blood pressure, even if cause-and-effect is not justified). Thus the sample should be representative of all adults in San Francisco.
    c) Perhaps those who eat meat regularly also consume more sodium than those who choose to be vegetarians. The higher sodium intake may be what is causing higher blood pressure among the "meat-eaters" when compared to the blood pressure of vegetarians. [Grader's note: The student MUST make the connection between the confounding variable and BOTH the explanatory AND response variables!]

17. a) Yes – a cause-and-effect relationship between the drug treatment and cholesterol reduction is justified since the men were <u>randomly assigned</u> to the two groups.
    b) No, since the men were not randomly selected from all such men at the hospital, and thus might not be representative of all such men at the hospital in terms of cholesterol levels.

18. a) A 3rd group serving as the control group would allow researchers to see if either treatment (A or B) is more effective in reducing recovery time than just rest (a.k.a., no treatment). Note: The control group is <u>not</u> necessary to determine whether treatment A is more effective than treatment B (or vice-versa).
    b) Assign each subject a number alphabetically from 1 to 72.
    Then use a random number generator to select 24 numbers from within that range, ignoring repeats. Volunteers with their numbers selected are assigned to program A.
    Use the random number generator to select another 24 numbers in that range, IGNORING numbers that were already selected and ignoring repeats.
    The remaining 24 subjects are assigned to the control group.
    [Answers may vary, but a proper randomization process must be given with sufficient details for another knowledgeable statistician to be able to follow]

19. a) All children in Podunk

b) Mean (or average) number of books read per year

c) Yes, even though he selected a random sample, his sampling frame was biased (the 320 children at the bookstore all enjoy reading). Thus, this will likely lead to an overestimate of the mean number of books read by all children in Podunk. This makes it less likely that the city will provide funding for additional reading programs.

d) Alex should choose the variable that is associated with greater variability in number of books read. Since it is believed that the number of books read varies more by gender than by ethnicity/race, Alex should stratify by gender.
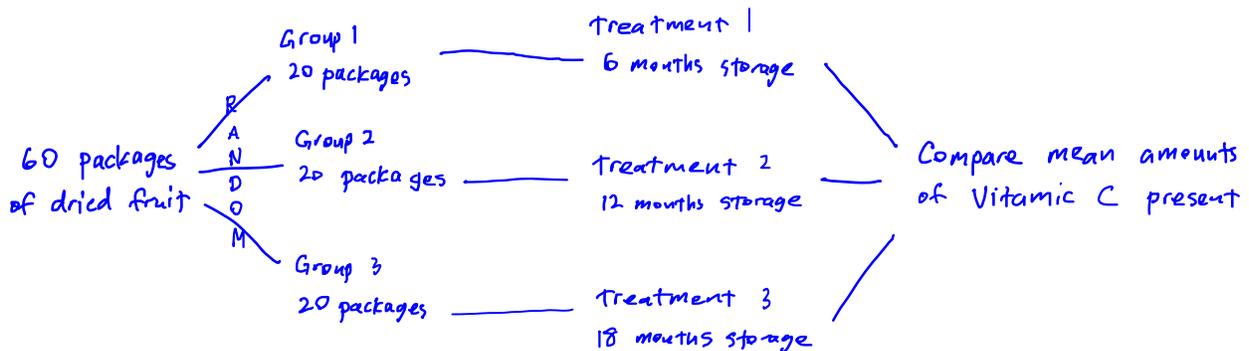
20.

a) Identify the experimental units: Packages of dried fruit

the explanatory variable(s) or treatment(s): Storage time of 6 months, 12 months, or 18 months
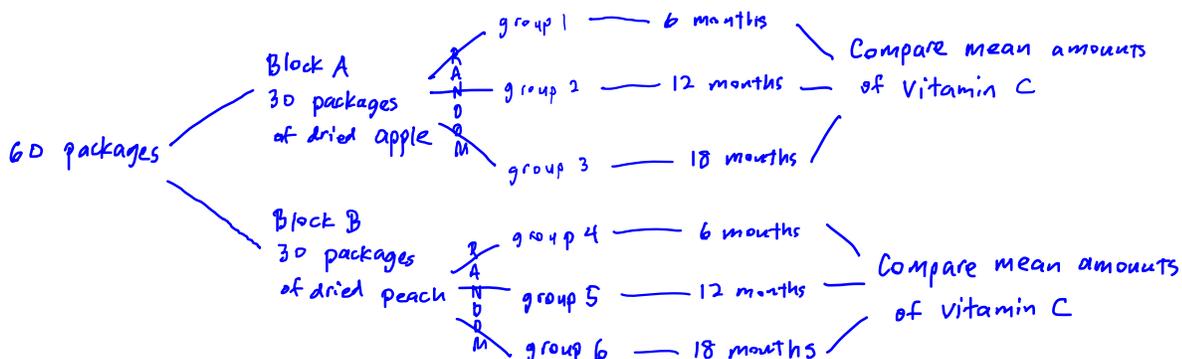
the response variable(s): amount of vitamin C

b) Explain how you would carry out a completely randomized experiment for this study.

60 packages of dried fruit — RANDOM —

Group 1
20 packages ——— Treatment 1
6 months storage

Group 2
20 packages ——— Treatment 2
12 months storage

Group 3
20 packages ——— Treatment 3
18 months storage

→ Compare mean amounts of Vitamin C present

Randomization Procedure: Assign each package a number from 1 – 60. Write each number on a slip of paper and put all 60 slips in a hat (make sure each slip of paper is the same size, and be sure to stir the slips in the hat). Without looking, draw 20 slips of paper. The packages of fruit with those numbers are assigned to group 1 (6 months of storage). Draw 20 more slips from the hat and place those packages of fruit into group 2 (12 months of storage). The remaining 20 packages of fruit are placed into group 3 (18 months of storage).

c) Describe the changes that would be made to your experiment in part (B) by incorporating blocking.

Since the different types of fruit may differ in amounts of Vitamin C, we will block by type of fruit.

60 packages

Block A
30 packages of dried apple — RANDOM —
group 1 — 6 months
group 2 — 12 months
group 3 — 18 months
→ Compare mean amounts of Vitamin C

Block B
30 packages of dried peach — RANDOM —
group 4 — 6 months
group 5 — 12 months
group 6 — 18 months
→ Compare mean amounts of vitamin C

d) Can the nutritionist generalize his/her findings to all types of dried fruit? Explain.

No. Because this experiment was conducted with a random sample of dried apple and dried peach, and these two types of fruit might not be representative of all types of dried fruits – in terms of amounts of vitamin C. Thus we can only generalize the findings to dried apple and dried peach (and not to other types of dried fruit).