SUPPLEMENTARY EXERCISE SOLUTIONS, CHAPTER 3

DESIGNING STUDIES AND COLLECTING DATA

S3.7. A study directed by Dr. Michael Gaziano, a heart specialist at the Harvard Medical School-affiliated Brigham and Women's Hospital in Boston, suggests that tea may reduce the chance of a heart attack by 44 percent. Gaziano's study examined 340 men and women who had suffered heart attacks and matched them by age, sex and neighborhood with people who had never had a heart attack. It then investigated their coffee- and tea-drinking habits over the course of a year. Gaziano thinks that the reason for the reduced incidence of heart attacks among tea drinkers is that tea contains powerful amounts of flavonoids, natural substances that make blood cells less prone to clots, which can cause heart attacks. The study was adjusted for factors that could have skewed the results, such as smoking, exercise, alcohol intake and family history of heart trouble. Total calories consumed, intake of fatty foods and body mass index which compares the girth of people of different heights to determine obesity was about the same across the board.

    a. What kind of a study was this? Be as precise as you can, and justify your answer.

        ANS: *A retrospective observational study, since subjects are classified into groups after the fact, and causal factors sought.* **(5 points)**

    b. Based on this study, is it valid to conclude that drinking tea causes a reduction in the chance of a heart attack? Why or why not?

        ANS: *No, because, even though the researchers tried to adjust for extraneous factors, this is not a controlled experiment.* **(5 points)**

S3.8. A researcher conducted a study of the effectiveness of an SAT preparation course. To do so, she randomly selected a group of 100 high school juniors from a large high school. She randomly divided them into two groups of 50 each. Both groups took the SAT on the same date in the fall. Then, one group was given the SAT preparation course and the other was not. Finally, both groups took the SAT on the same date in the spring. To evaluate the results, the differences in the two SAT scores for each student were computed, and these differences were compared for differences between the two groups.

    a. Explain why this a controlled experiment.
    ANS: *It is a controlled experiment since treatments (prep course/no prep course) were imposed on experimental units (students) to observe a response (difference in SAT scores).* **(5 points)**
    b. How is blocking used in this experiment? What do you think its effect will be?
    ANS: *Students are blocks in which both the before and after SAT score are observed. It should eliminate between-student variation, allowing the researcher to focus on the effect of the prep course.* **(5 points)**
    c. Randomization is used in two ways in this experiment. Explain how it is used and why.
    ANS: *Randomization is used to select the sampling units from the larger population. This will enable the results to be extrapolated to that larger population. It is also used to assign treatments to experimental units. This helps guard against unsuspected sources of bias in the two groups.* **(5 points)**
    d. Describe how replication is used in this experiment.
    ANS: *There are 100 experimental units: 50 in each group. These are replicates.* **(5 points)**

S3.10. Three studies are described below. Tell as completely as possible what type of study each is. Justify your answers.

    a. Scientists randomly divide volunteers into two groups. In one group, each individual is given a geometric puzzle to solve, while each individual in the other group is given no stimulation. After five minutes, magnetic resonance imaging (MRI) measurements of the individual's brain activity are obtained.
    ANS: *This is a* **controlled experiment** *since treatments (puzzle or no activity) are assigned to experimental units (subjects) in order to observe a response (MRI measurements).* **(5 points)**

b. Crop yields are estimated by satellite. First the US is divided into sectors that can be viewed by the satellite in a single image. Images are obtained of a set of sectors selected by a probability sampling scheme. These images are analyzed to estimate the total US crop yield.

ANS: *This is a* **sample survey** *since a set of sampling units (sectors) is selected from the population of sectors to obtain information about that population.* **(5 points)**

c. In order to identify possible causes of tread separation on a particular model of SUV tire, quality engineers compare a sample of failed tires with a sample of unfailed tires with respect to a number of variables.

ANS: *This is a retrospective observational study, since tires are grouped by end result (failure or not) and differences in hypothesized causes sought.* **(5 points)**

S3.11. Describe how the first study in exercise S3.10 could be improved by the use of blocking. Be sure to tell how the blocking would help.

ANS: *Use each subject as a block and measure each under treatment (puzzle) and control (no stimulation). This would help reduce subject to subject variation.* **(5 points)**

S3.12. Are the results of the last study in exercise S3.10 necessarily applicable to the population of all tires of the model studied? If you conclude 'yes,' tell why. If you conclude 'no,' tell why not and indicate what must be done to make the results applicable to that population.

ANS: *They are not necessarily applicable, since the samples may not have been chosen in a way that makes it reasonable to assume they are representative of the population. The tires should be chosen according to an appropriate probability sampling scheme in order to justify extending the results to the population.* **(5 points)**