Test Bank for Introduction to Statistics and Data Analysis 5th Peck

Access Complete Test Bank for Introduction to Statistics and Data Analysis 5th Peck

Chapter 2: Collecting Data
Concept Quiz

Name _____

The following questions are in a True / False format. The answers to these questions will frequently depend on remembering facts, understanding of the concepts, and knowing the statistical vocabulary. Before answering these questions, be sure to read them carefully!

T F 1. By definition, a simple random sample of size *n* is any sample that is selected in a manner to guarantee every individual in the population has an equal chance of selection.

T F 2. Clusters are non-overlapping subgroups of a population that have been identified as homogeneous.

T F 3. Blocking is a technique that can be used to filter out the effects of extraneous factors.

T F 4. Response bias can occur when responses are not actually obtained from all individuals selected for inclusion in the sample.

T F 5. Selection bias can occur if volunteers only are used in a study.

T F 6. Stratified sampling is a sampling method that in no way involves simple random sampling.

T F 7. Increasing sample size will generally eliminate bias in a sample.

T F 8. A placebo is identical in appearance to the treatment of interest, but contains no active ingredients.

T F 9. A study is an observational study if the investigator observes the behavior of a response variable when one or more factors are manipulated.

T F 10. In a well-designed experiment, the factors are confounded whenever possible.

Access Complete Test Bank for Introduction to Statistics and Data Analysis 5th Peck

Chapter 2: Collecting Data
Concept Quiz (Cont'd)

T F 11. As long as the sample size is small relative to the population, there is little practical difference between sampling with replacement and sampling without replacement.

T F 12. A treatment is any particular combination of values for the explanatory variables.

T F 13. Two factors are extraneous if their effects on the response variable cannot be distinguished from one another.

T F 14. Random assignment to treatments will guarantee groups that are exactly alike for experimental purposes.

T F 15. The method of control wherein an extraneous variable is held constant is called blocking.

T F 16. Random subpopulations of a population are called strata.

T F 17. A control group provides a baseline for comparison to a treatment group.

T F 18. Random assignment of volunteers should result in comparable experimental groups.

T F 19. If the subjects as well as the person measuring the response are aware of the treatment assigned to the subject, only single-blinding is being used.

T F 20. Replicating in an experiment means that the number of subjects is greater than 1.

Access Complete Test Bank for Introduction to Statistics and Data Analysis 5th Peck

Chapter 2, Concept Quiz Answer Key

1. F
2. F
3. T
4. F
5. T
6. F
7. F
8. T
9. F
10. F
11. T
12. T
13. F
14. F
15. F
16. F
17. T
18. T
19. F
20. T

Access Complete Test Bank for Introduction to Statistics and Data Analysis 5th Peck

Chapter 2: Collecting Data

Section 2.1-2.2

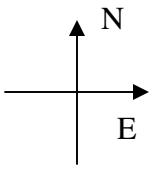
Name _____

1. A friend of yours, who is not taking statistics, wonders why it is that anyone would choose to take a sample. "Obviously," she says, "you would get better information from a census." In a short paragraph, explain why it is that statisticians take samples rather than taking a census.

2. The most basic sampling method studied in statistics is the simple random sample (SRS). In your own words, what is the correct definition of a simple random sample of size n?

Access Complete Test Bank for Introduction to Statistics and Data Analysis 5th Peck

3. The ZZZ chain of motels has a standard method of constructing their rooms to maximize the ease of parking for its customers. The rooms are arranged in adjacent buildings so that each customer can park outside the rented room. The layout for one of the with 48 rooms located along a famous highway is diagrammed below:



Route 66

Building A

Building B

The manager would like to survey customers in 12 of his rooms (one randomly selected customer for each room selected in the sample) to assess their satisfaction with the motel services. The surveys will be placed on the customers' beds before they check in to the motel. In order to make the directions easy to follow, he elects to use systematic sampling.

a) Explain how you would use random numbers to set up the systematic sampling process.

Access Complete Test Bank for Introduction to Statistics and Data Analysis 5th Peck

b) Write a short paragraph for the maids that helps them carry out your method in part (a).

Access Complete Test Bank for Introduction to Statistics and Data Analysis 5th Peck

4. Bias is a serious problem that sometimes arises when one takes a sample.

   a) In a few sentences, explain generally what bias is.

   b) What is the distinction between response bias and non-response bias?

Access Complete Test Bank for Introduction to Statistics and Data Analysis 5th Peck

Chapter 2, Quiz 1, Form A Answer Key

1. Although we may get better information from a census, it is usually far too costly and time consuming to contact every member of the population. A large random sample will be nearly as good for far less cost.

2. A simple random sample of size n is a sample that is selected from a population in a way that ensures that every different possible sample of the desired size has the same chance of being selected.
   Note: It is important that students not only state that each person has the same chance of being chosen, but also each possible sample of size n has the same chance of being chosen.

3. a) Since there are 48 units in the population and we want a sample of size 12, we want to choose every fourth room after randomly choosing one of the first four rooms to start with. If we are using a random digit table, we would go through the table until we get a number from 1 to 4. Then, we would keep adding 4 to that number until we get to the end of the hotel rooms. For example, if we come upon the number 3 first, we would survey the 3rd room, the 7th room, the 11th room, etc.

   b) Dear Maids, when you are placing the surveys in the rooms, please follow the following procedure. Starting at the northwest corner of building A and moving east, place a survey in the third room, the seventh room, and every fourth room thereafter, moving back and forth along the four rows of rooms.

4. a) Bias is the tendency for a sample to differ from the corresponding population in some systematic way.

   b) Non-response bias occurs when responses are not actually obtained from all individuals selected for the sample. With response bias, however, responses are obtained from the subjects, but the method of observation tends to produce values that systematically differ from the true population value in some way.

Access Complete Test Bank for Introduction to Statistics and Data Analysis 5th Peck

Chapter 2: Collecting Data

Section 2.1-2.2

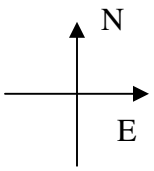
Name _____

1. A friend of yours, who is not taking statistics, wonders why it is that anyone would choose to take a sample. "Obviously," he says, "you would get better information from a census." In a short paragraph, explain why it is that statisticians take a sample rather than a census.

2. The most basic sampling method studied in statistics is the simple random sample (SRS). In your own words, what is the correct definition of a simple random sample of size n?

Access Complete Test Bank for Introduction to Statistics and Data Analysis 5th Peck

3. The ZZZ chain of motels has a standard method of constructing their rooms to maximize the ease of parking for its customers. The rooms are arranged in two adjacent buildings so that each customer can park outside the rented room. The layout for one of the motels in the chain along a famous highway is diagrammed below:



Route 66

Building A

Building B

The manager would like to survey customers in 12 of his rooms (one randomly selected customer for each selected room) to assess their satisfaction with the motel services. As he tries to decide what sampling strategy is appropriate, the manager realizes that the rooms in Building A might be noisier than Building B, and this should be taken into account.

a) What sampling strategy would you suggest to give the manager the best chance of estimating his customers' satisfaction?

Access Complete Test Bank for Introduction to Statistics and Data Analysis 5th Peck

b) Write a short paragraph for the maids that helps them carry out your method in part (a).

Access Complete Test Bank for Introduction to Statistics and Data Analysis 5th Peck

4. Bias is a serious problem that sometimes arises when one takes a sample.

   a) In a few sentences, explain generally what bias is.

   b) What is the distinction between selection bias and non-response bias?

Access Complete Test Bank for Introduction to Statistics and Data Analysis 5th Peck

Chapter 2, Quiz 1, Form B Answer Key

1. Although we may get better information from a census, it is usually far too costly and time consuming to contact every member of the population. A large random sample will be nearly as good for far less cost.

2. A simple random sample of size n is a sample that is selected from a population in a way that ensures that every different possible sample of the desired size has the same chance of being selected.
   Note: It is important that students not only state that each person has the same chance of being chosen, but also each possible sample of size n has the same chance of being chosen.

3. a) Stratified random sampling.

   b) Since the manager wants to take noise into account, it would be best to stratify by "rows." That is, one stratum would be the 12 rooms in building A facing the road, since they are the most similar with regard to noise. The next strata would be the 12 rooms on the other side of building A, etc. Then, in each of the 4 strata, we need to take a random sample of size 3. To do this we can write the 12 room numbers on 12 sheets of paper, mix them up in a hat, and choose three for each strata.

4. a) Bias is the tendency for a sample to differ from the corresponding population in some systematic way.

   b) Selection bias occurs when some part of the population is systematically excluded from the sample. However, non-response bias occurs when responses are not actually obtained from the individuals who were selected for the sample.

Access Complete Test Bank for Introduction to Statistics and Data Analysis 5th Peck

Chapter 2: Collecting Data

Section 2.3-2.4

Name _____

1. We have distinguished two types of studies: observational and experimental. Briefly explain the essential difference(s) between these two types of study.

2. In evaluating an experiment, how would you determine if a variable is an explanatory variable or an extraneous variable?

Access Complete Test Bank for Introduction to Statistics and Data Analysis 5th Peck

3. Suppose that two experiments were conducted to assess the effect of a new insect repellant. In Experiment A, a simple random sample was taken from the population of River City. In Experiment B, a simple random sample from a group of volunteers from the population of River City was used. The results of the experiments were the same: fewer insects landed on the arms that had been treated with the insect repellant. The volunteers were randomly assigned to the two treatments in both experiments.

   a) For each experiment, A and B, discuss whether one can legitimately infer a cause-and-effect relation between the use of the repellant and fewer insects landing from each of these experiments? Why or why not?

   b) For each experiment, A and B, discuss whether one can legitimately generalize to the population of River City from each of these experiments? Why or why not?

Access Complete Test Bank for Introduction to Statistics and Data Analysis 5th Peck

4. One of the tasks of wildlife biologists is to estimate bird populations. One method for doing this is to walk a "transect," typically a long randomly selected rectangle, and count the number of particular species that are seen. During hunting season the biologists wear brightly colored clothes as a protective measure. A theory, known as the "species-confidence hypothesis," predicts that birds may react to these brightly colored clothes. A robin, for example, when confronted with a biologist wearing an orange color may be less afraid than if the biologist were wearing a brown color. The idea is that the orange color gives the robin some "confidence" that the biologist is another (really big!) bird. This is a difficulty because a robin that is less afraid of the biologist may be easier to see and therefore "count" because the robin lets the biologist get closer. This would result in too high an estimate of the robin population. To test this theory a biologist performed a study in her area, wearing a bright orange vest or a brown vest while working. After training herself to accurately estimate distances she indicated on her notes how close the robins would let her get before they flew away (the "approach distance"), using this as a measure of their "confidence." She randomly selected her starting points in each trip through the transects and also carefully noted the weather conditions, specifically the wind speed.

(a) What is the explanatory variable (factor) for this experiment?

(b) What is the response variable for this experiment?

(c) After completing the study the biologist was examining her results and noticed that on those days she wore the brightly-colored vest the wind tended to be of greater velocity. How does this information affect the interpretation of the results of the experiment?

Access Complete Test Bank for Introduction to Statistics and Data Analysis 5th Peck

Chapter 2, Quiz 2, Form A Answer Key

1. In an experiment, researchers observe how a response variable behaves when they manipulate one or more factors. However, in an observational study, the researchers do not manipulate any factors. Instead, they observe characteristics of a subset of the members of one or more existing populations.

2. In an experiment, the explanatory variable is the one that researchers manipulate in order to observe changes in the response variable. An extraneous variable is any other variable which is thought to affect the response variable, but is not of interest in the study.

3. a) In each of these experiments we can infer a cause-and-effect relationship because the volunteers were randomly assigned to the two treatments.

   b) In experiment A, we can generalize the results of the experiment to the population of River City since the subjects were randomly selected from the population of River City. Thus, the subjects should be representative of the population in general. However, in experiment B, the random sample came from a group of volunteers from River City, who may not be representative of the general population of River City.

4. (a) The explanatory variable is the coloration of the vest (bright vs. drab).

   (b) The response variable is the "approach distance."

   (c) The wind velocity is a potential confounding variable. On a windy day the birds may not detect the researcher as easily, or they may be more hesitant to fly. Thus it might be that the results are at least partly the result of wind velocity, not the choice of garment color.

Access Complete Test Bank for Introduction to Statistics and Data Analysis 5th Peck

Chapter 2: Collecting Data

Section 2.3-2.4

Name _____

1. We have distinguished two types of studies: observational and experimental. Briefly explain the essential difference(s) between these two types of study.

2. In evaluating an experiment, how would you determine if a variable is an explanatory variable or a response variable?

Access Complete Test Bank for Introduction to Statistics and Data Analysis 5th Peck

3. Suppose that two experiments were conducted to compare the effects of two new insect repellants. In Experiment X, a simple random sample was taken from the population of River City, and the citizens were randomly assigned to repellant A and repellant B. In Experiment Y, a stratified random sample was taken from the population of River City, and the citizens were randomly assigned to repellant A and repellant B. (River City has 4 geographic areas of town, the NE, NW, SE, and SW quadrants; 100 people from each quadrant were sampled for Experiment B.)

   The results of the experiments were the same: fewer insects landed on the arms that had been treated with insect repellant A.

   a) For each experiment, X and Y, discuss whether one can legitimately infer from each of these experiments that repellant A is better than B? Why or why not?

   b) For each experiment, X and Y, discuss whether one can legitimately generalize to the population of River City from each of these experiments? Why or why not?

Access Complete Test Bank for Introduction to Statistics and Data Analysis 5th Peck

4. The likelihood that an animal will successfully avoid a predator's attack should increase if the presence of the predator is detected. Animals are, of course, known to warn others of a predator's presence – the well-known prairie-dog warning call is an example. Some aquatic animals are known to send chemical signals. For example the crayfish and Iowa darter may excrete ammonium from their gills or possibly in their urine during periods of heightened metabolic activity. In an experiment to see if red-legged frogs (*Rana aurora*) may send or respond to chemical signals, specimens in Oregon were collected as embryos and raised in an aquarium. As tadpoles they were separated into two pairs of aquaria. The "upper tank" in each pair was connected to the "lower tank" to create a (very!) small river. A wooden heron model was placed in one of the upper tanks and moved around for 30 seconds to simulate a predator attack. Both the sender (upper tank) and receiver (lower tank) animals were monitored for added activity (increased movement) indicating antipredator behavior.

(a) What is the explanatory variable (factor) for this experiment?

(b) What is the response variable for this experiment?

(c) During the course of the experiment the investigators were very careful with the wooden heron model not to come in contact with the glass of the aquaria or make noise in any other way. If they had been unsuccessful and their wooden heron made significant amounts of noise, how would that affect the interpretation of the results?

Access Complete Test Bank for Introduction to Statistics and Data Analysis 5th Peck

Chapter 2, Quiz 2, Form B Answer Key

1. In an experiment, researchers observe how a response variable behaves when they manipulate one or more factors. However, in an observational study, the researchers do not manipulate any factors. Instead, they observe characteristics of a subset of the members of one or more existing populations.

2. In an experiment, an explanatory variable is one whose value is manipulated or determined by the experimenter, while a response variable is one whose value is measured at the end of the experiment.

3. a) In each of these experiments we can infer a cause-and-effect relationship because the volunteers were randomly assigned to the two treatments.

   b) In each of these experiments, we can generalize the results of the experiment to the population of River City since the subjects were randomly selected from the population of River City. Thus, the subjects should be representative of the population in general.

4. (a) The explanatory variable is the presence/absence of the wooden heron model.

   (b) The response variable is the antipredator behavior.

   (c) The added noise would be a potential confounding variable. The tadpoles' response may be a startle response to a sudden change in their environment, and not specific to the detection of a predator.

Access Complete Test Bank for Introduction to Statistics and Data Analysis 5th Peck

Chapter 2: Collecting Data

Name _____

1. Bias, the tendency for samples to differ from the corresponding population in some systematic way, might be due to: (a) selection bias, (b) response bias, and/or (c) nonresponse bias. In a few sentences, discuss the differences among these different biases.

Access Complete Test Bank for Introduction to Statistics and Data Analysis 5th Peck

2. A pharmaceutical company wants to test its new drug that is designed to help balding men grow more hair. From their records of past customers, the company has data on about 5,000 men. The data contains information about the men's hair color, age, and percent of baldness. (A partial list is given below.) For their anticipated experiment, they want to take a sample that is representative of their customers.

.

Hair color	Age (yrs)	% Baldness
Light	67	83
Dark	62	73
Light	41	25
Dark	52	50
Dark	43	14
Light	69	96
Dark	56	57
…	…	…
Light	32	40

a) Briefly describe how you would select a simple random sample of size $n = 20$ from this list of customers.

Access Complete Test Bank for Introduction to Statistics and Data Analysis 5th Peck

b) Describe in a short paragraph why you might wish to use a stratified random sample.

3. The following paragraph describes an actual study. After reading the description, determine whether the study is an observational study or an experiment. Justify your answer with specific references to the information in the study.

> "We compared paired daytime and night counts of wild brook trout, brown trout, and rainbow trout made by the same snorkelers in five streams during August 1994. Overall, we counted 109 trout in the daytime and 333 trout at night. We speculate that trout counted at night were present during the daytime but were hidden from view. Biologists should consider that trout behavior and susceptibility to being seen might vary a great deal between daytime and night, even during summer. In some streams, the majority of trout may not be seen during the daytime."

Access Complete Test Bank for Introduction to Statistics and Data Analysis 5th Peck

4. In competitive sports, video recorders have been used more frequently in recent years. The idea behind the recorder is that coaches can replay training sessions for more effective feedback to the athlete. Some people believe video recording may make the athletes more nervous and actually decrease their performance. You have been asked to design an experiment to address this issue for competitive high school tennis players. You have decided to use the accuracy of tennis serves as your response variable, and the number of successful serves out of 100 as your performance measure. The subjects for your experiment are 60 high school male competitive tennis players of varying ability who have volunteered for the experiment.

a) Describe the treatments in your experiment

b) One possible confounding variable is the experience levels of the players. Explain how you would control this variable?

c) Can the results of this experiment could be generalized to all male tennis players? Why or why not?

Access Complete Test Bank for Introduction to Statistics and Data Analysis 5th Peck

Chapter 2, Test Form A Answer Key

1. Selection bias occurs when some part of the population is systematically excluded from the sample. Non-response bias occurs when responses are not actually obtained from all individuals who were selected for the sample. With response bias, however, responses are obtained from the subjects, but the method of observation tends to produce values that systematically differ from the true population value in some way.

2. a) To select a simple random sample of size 20, we could number the subjects from 0001-5000 and use a random digit table. On the table, we would look at sets of 4 digits until 20 numbers from 0001-5000 were selected (ignoring any repeats) and these would be the men selected.

   b) If the researchers anticipate an association between any of the variables listed (hair color, age, or % baldness) and the response variable, they should stratify by that variable so that the sample they get will not over- or under-represent a subgroup which may respond differently than the population in general.

3. This study is an observational study since the explanatory variable was not manipulated by the researchers and the subjects were not randomly assigned to different treatments. Instead, the researchers simply observed the number of trout visible during the day and during the evening.

4. a) The two treatments will be: 1. The subject is recorded and 2. The subject is not recorded.

   b) I would use blocking (pairing) to control the experience level of players. I would pair the two most experienced together as one block, the next two most experienced as the next block, and so on. Then the two members of each block would be randomly split into the two treatment groups. This way each treatment group should be roughly the equivalent with regard to experience level.

   c) No, the results of this study should not be generalized to all male tennis players for at least two reasons. One, competitive tennis players are presumably more used to playing in front of crowds and would be less bothered by video recording than the typical player. Two, volunteers are not generally representative of any larger population.
   Note: either reason should be sufficient to receive credit.

Access Complete Test Bank for Introduction to Statistics and Data Analysis 5th Peck

Chapter 2: Collecting Data

Name _____

1. Three methods for random sampling are: (a) simple random sampling, (b) stratified random sampling, and (c) cluster sampling. In a few sentences, discuss the similarities and differences among these sampling methods. Specifically, what sampling circumstances would lead you to choose each of these methods?

Access Complete Test Bank for Introduction to Statistics and Data Analysis 5th Peck

2. The two paragraphs below discuss aspects of two studies, each of which exhibit a bias. For each study, decide whether the problem is selection bias, response bias, or nonresponse bias, and in a few sentences explain why you chose your answer.

a) One part of the Nurses' Health Study is concerned with possible causes of skin cancer. Nurses were asked about different behaviors and aspects of their health when they entered the study. Then, the nurses were given the questionnaire again if they were diagnosed with cancer. When the questionnaires were analyzed, the investigators discovered that after the nurses were diagnosed with cancer they tended to report a reduced ability to tan. It is thought that the shift in reporting might be caused by an awareness of their diagnosis.

b) One part of the Demographic and Health Surveys Program is concerned with measures of malnutrition. Investigators measure physical aspects of growing children, and attempt to document the physical characteristics of a population at different ages. Sadly, in some countries many children die early, and thus a bias is introduced in the study when the investigators can not collect the data from the deceased children.

Access Complete Test Bank for Introduction to Statistics and Data Analysis 5th Peck

3. The following paragraph describes an actual study. After reading the description, determine whether the study is an observational or experimental study. Justify your answer with specific references to the information in the study.

"Before the opening of the new International Airport and the termination of the old airport, children near both sites were recruited into aircraft-noise groups (aircraft noise at present or pending) and control groups with no aircraft noise and closely matched for socio-economic status. A total of 326 children (mean age = 10.4 years) took part in three data-collection waves, one before and two after the switch-over of the airports. After the switch to the new airport, long-term memory and reading were found to be impaired in the noise group at the new airport, and improved in the formerly noise-exposed group at the old airport."

Access Complete Test Bank for Introduction to Statistics and Data Analysis 5th Peck

4. A common practice of teachers is to have students exchange their quizzes and grade each others. In addition to decreasing the teacher's work load, the reduced time between quiz and feedback is thought to be a plus for learning. Your U.S. History teacher, aware of your statistical prowess, has asked you to design an experiment to test this theory. You have decided to use the mid-term exam (not graded by students) as your response measure. Your history teacher has three classes, one early in the morning, one at noon, and one late in the afternoon. Each class contains 30 students.

   (a) Describe the treatments you will use in your experiment

   (b) One possible confounding variable is the time of day, since students may be more alert at certain times of the day than at other times. Describe a method would you use to control this variable? (Unfortunately you cannot change the student schedules!)

   (c) Do you feel the results of your experiment could be generalized to your statistics class? Why or why not?

Access Complete Test Bank for Introduction to Statistics and Data Analysis 5th Peck

Chapter 2, Test Form B Answer Key

1. In simple random sampling, every individual and every possible sample of size n has an equal chance of being selected for the study. In stratified random sampling, the population is divided into non-overlapping homogeneous groups (called strata) and a simple random sample is selected from each strata. In cluster sampling, the population is divided into non-overlapping (preferably heterogeneous) groups called clusters and then a random sample of clusters is selected and every member of the selected clusters is studied.

   Cluster sampling works best when the population is already divided into easily identifiable groups that are heterogeneous (i.e. each cluster can reasonably be assumed to be representative of the entire population). Stratified random sampling works best when there are easily identified groups in the population that are anticipated to have very different responses to the question of interest. Simple random sampling is best when neither of the circumstances listed above are present.

2. a) This is an example of response bias, since the awareness of their diagnosis may have caused them to change their response. It isn't non-response bias since they were able to obtain responses from the nurses and it isn't selection bias since they did not attempt to generalize to a larger population.

   b) This is an example of non-response bias, since some of the children selected for the study were not able to participate after they died. It is not selection bias since the children were not left out on purpose and it isn't response bias since the researchers were unable to obtain responses in the first place.

3. This is an observational study. The children were identified in place and the manipulation of sound level was not randomly assigned.

4. (a) Individual pairs of students would be randomly assigned to "trade papers" or "not trade papers" treatment groups. The non-trading students' work would be graded by the teacher each day and given back the next day.

   (b) Each class would be considered a "block." Within each block both treatments would be randomly assigned as indicated in part (a).

   (c) The results might be generalizable to other classes, but w/o doing the experiment in those classes there is no evidence suggest one could generalize. Statistics and history seem like they might be different enough that, although they are both classes with homework, the subject matter might be learned differently and the instant checking of the quizzes might be less or more of a help in one class or the other.

Access Complete Test Bank for Introduction to Statistics and Data Analysis 5th Peck