

Review Exercises

Key

MULTIPLE-CHOICE QUESTIONS

1. A marketing company offers to pay \$25 to the first 100 persons who respond to their advertisement and complete a questionnaire regarding displays of their client's product. This situation is an example of which of the following?

- (A) simple random sample
- (B) convenience sample
- (C) voluntary response sample
- (D) multistage cluster sample
- (E) systematic sample

2. The AP Test Committee would like to survey principals in large urban school districts throughout New England regarding the AP curriculum. A simple random sample of large urban school districts was selected. Within each selected district, a simple random sample of high schools was chosen and each principal was interviewed. Which of the following statements regarding this design is true?

- (A) This is an example of a cluster sample.
- (B) Results from the interviews cannot be used to infer responses of the population of interest.
- (C) The population of interest is the set of all high school principals from large urban school districts in New England.
- (D) The population of interest is the set of all large urban school districts in New England.
- (E) Every subset of principals has the same chance of selection.

3. Which does *not* represent a stratified design of a study involving students at your high school? For each answer choice, assume the results of the separate samples will be combined into a single sample.

- (A) Two simple random samples are conducted: one of the boys in the student body and the other of the girls in the student body.
- (B) Four simple random samples are conducted: one in each of the four classes.
- (C) A simple random sample of homerooms is conducted. Within each selected homeroom, two random samples are conducted: one of the boys and the other of the girls.
- (D) Two random samples are conducted: one of students whose GPA's are 2.5 or higher and the other of students whose GPA's are less than 2.5.
- (E) All of these statements are stratified designs.

Use the following information to answer questions 4 and 5.

In order to assess the membership's attitudes about a new Supreme Court decision, a local bar association selects a simple random sample of 100 lawyers from its membership list. Surveys are delivered to the selected lawyers. Only 63 of the lawyers return their surveys.

4. Which of the following is of great concern in this situation?

- (A) Nothing is known about the parameters of population of interest.
- (B) Nothing is stated regarding the methodology of the simple random sample.
- (C) There may be a problem with the sampling frame.
- (D) There may be a problem with non-response bias.
- (E) None of these statements is of concern.

5. If 43 of the respondents disagree with the new ruling, which of the following statements is true?

- (A) The sample is the membership of the bar association.
- (B) The sample is the 100 selected lawyers.
- (C) The sample is the 63 respondents.
- (D) The sample is the 43 who disagree with the new ruling.
- (E) None of these is true.

Use the following excerpt from a random digit table to answer questions 6–8.

21052 65031 45074 92846 67815 78231
01548 20235 56410 82713

6. Which of the following statements regarding this excerpt of the given random digit table is true?

- (A) This table can only be used for data coded with five-digit numbers.
- (B) In order to use this table for a population of 100 names, the names could be coded 00 to 99; then distinct two-digit numbers would be selected sequentially to identify chosen names.
- (C) It is critical to begin the selection of random digits at the beginning of the list.
- (D) This table can only be used if the data labels include all of the digits in the table.
- (E) None of these statements is true.

7. If data are labeled: 1. Chevy; 2. Plymouth; 3. Lincoln; 4. Volkswagen; 5. Porsche; 6. Ford; and *single-digit* random digit selection begins at the left side of the first row, which cars would be included in a simple random sample of three cars?

- B
- (A) Plymouth, Lincoln, Chevy
 - (B) Plymouth, Chevy, Porsche
 - (C) Plymouth, Ford, Porsche
 - (D) Lincoln, Plymouth, Porsche
 - (E) Lincoln, Ford, Plymouth

8. Which of the following statements is true?

- C
- (A) A second sample of three beginning at a different position in the row would always produce the same set of selections.
 - (B) A second sample of three beginning at a different position in the row would never include any of the cars selected in question 7.
 - (C) Every subset of three cars would have the same chance of selection as the one identified in question 7.
 - (D) The sample size is too large relative to the population size to conduct the sample.
 - (E) None of these statements is true.

FREE-RESPONSE QUESTIONS

Open-Ended Questions

1. Consider the variable of the heights of male adults in inches. Below are listed 100 heights of male adults. Suppose that these 100 heights are the population of interest. The list is organized in five rows of twenty each.

62 65 65 67 72 74 62 68 65 59 71 70 65 69 68 63 62 65 64 67
 68 70 72 65 61 64 64 68 62 72 73 68 64 58 58 74 60 62 64 68
 69 72 75 52 54 68 67 64 64 62 60 61 71 75 72 63 63 62 68 64
 67 62 58 59 74 76 60 69 70 61 68 68 72 70 59 58 65 54 72 63
 66 67 66 62 61 68 69 71 70 75 63 61 69 69 67 72 74 60 61 69

- a. Code each of the 100 heights by its position number using one of the following options:
 - (1) If you use a random number generator on a calculator or a computer, code the heights using numbers from 1 to 100, using 1–20 in row 1, 21–40 in row 2, etc.
 - (2) If you use a random digits table, code the heights using numbers from 00 to 99, using 00–19 in row 1, 20–39 in row 2, etc.
- b. Select a simple random sample of 10 values from the set of 100. Indicate both the code number and the corresponding height in your sample.

Code	Value
1	2
3	4
5	6
7	8
9	10

- c. Calculate the mean height of your sample.
- d. If the mean height of the population is 65.89 inches, comment on how accurately the mean height of your sample estimates the population mean height.

For each of the sampling procedures listed, comment on the compliance with the two conditions for valid probability-based sampling:

2. Voluntary Response

- F a. Interviewers do not choose participants in a subjective manner, and subjects do not self-select themselves.
- F b. Selection based on probability.

3. Convenience Sampling

- T a. Interviewers do not choose participants in a subjective manner, and subjects do not self-select themselves.
- F b. Selection based on probability.

X Quota Sampling

- a. Interviewers do not choose participants in a subjective manner, and subjects do not self-select themselves.
- b. Selection based on probability.

5. Simple Random Sampling

- T a. Interviewers do not choose participants in a subjective manner, and subjects do not self-select themselves.
- T b. Selection based on probability.

6. Stratified Random Sampling

- T a. Interviewers do not choose participants in a subjective manner, and subjects do not self-select themselves.
- T b. Selection based on probability.

X Multi-stage Random Sampling

- a. Interviewers do not choose participants in a subjective manner, and subjects do not self-select themselves.
- b. Selection based on probability.

8. Systematic Sampling

- T a. Interviewers do not choose participants in a subjective manner, and subjects do not self-select themselves.
- T b. Selection based on probability.

FRAPPY! Free Response AP® Problem, Yay!

The following problem is modeled after actual Advanced Placement Statistics free response questions. Your task is to generate a complete, concise response in 15 minutes. After you generate your response, view two example solutions and determine whether you feel they are "complete", "substantial", "developing" or "minimal". If they are not "complete", what would you suggest to the student who wrote them to increase their score? Finally, you will be provided with a rubric. Score your response and note what, if anything, you would do differently to increase your own score.

A large school district is interested in determining student attitudes about their co-curricular offerings such as athletics and fine arts. The district consists of students attending 4 elementary schools (2000 students total), 1 middle school (1000 students total), and 2 high schools (2000 students total).

The administration is considering two sampling plans. The first consists of taking a simple random sample of students in the district and surveying them. The second consists of taking a stratified random sample of students and surveying them.

- (a) Describe how you would select a simple random sample of 200 students in the district.

Write each person's name on a slip of paper. Place them in a hat, mix well and select 200 slips of paper.

OR ~~Use~~ Label each student alphabetically from 0001 to 6000. Use a random # generator or table to produce 200 unique numbers. The sample will be the individuals that correspond to #s.

- (b) Describe how you would select a stratified random sample consisting of 200 students.

We should stratify by level (Elem, Middle, High) b/c attitudes may differ by level. Label students at each level and randomly select 80 elem, 40 middle, and 80 high students using a random # generator. This ensures each level is represented in the same proportion as the population.

- (c) Describe the statistical advantage of using a stratified random sample over the simple random sample in this study.

Stratifying ensures every level is adequately represented in the sample. If selecting a SRS, it is possible (although unlikely) to choose all students from the same level or no students from a level.

If this were to occur, the sample would not accurately represent the opinions of the population.