

Part 3 Free Response
AP[®] STATISTICS Practice
2011 SCORING GUIDELINES

Question 3

Answers

1 Intent of Question

The primary goals of this question were to assess students' ability to (1) describe a process for implementing cluster sampling; (2) describe a statistical advantage of stratified sampling over cluster sampling in a particular situation.

Solution

Part (a):

The following two-step process can be used to select the eight apartments.

Step 1: Generate a random integer between 1 and 9, inclusive, using a calculator, a computer program, or a table of random digits. Select all four apartments on the floor corresponding to the selected integer.

Step 2: Generate another random integer between 1 and 9, inclusive. If the generated integer is the same as the integer generated in step 1, continue generating random integers between 1 and 9 until a different integer appears. Again select all four apartments on the floor corresponding to the second selected integer.

The cluster sample consists of the eight apartments on the two randomly selected floors.

Part (b):

Because the amount of wear on the carpets in apartments with children could be different from the wear on the carpets in apartments without children, it would be advantageous to have apartments with children represented in the sample. The cluster sampling procedure in part (a) could produce a sample with no children in the selected apartments; for example, a cluster sample of the apartments on the third and sixth floors would consist entirely of apartments with no children. Stratified random sampling, where the two strata are apartments with children and apartments without children, guarantees a sample that includes apartments with and without children, which, in turn, would yield sample data that are representative of both types of apartments.

Scoring

Parts (a) and (b) are scored as essentially correct (E), partially correct (P) or incorrect (I).

Part (a) is scored as follows:

Essentially correct (E) if the response correctly addresses the following two components:

1. Indication that two floors are randomly selected, with all four apartments on each of the selected floors forming the sample (or that the entire floors should be carpeted).
2. Description of a valid random sampling procedure for selecting two floors that could be implemented after reading the response (so that two knowledgeable statistics users would use the same method to select the floors).

Partially correct (P) if the response includes exactly one of the two components listed above.

AP[®] STATISTICS

2011 SCORING GUIDELINES

Question 3 (continued)

Incorrect (I) if the response includes neither of the two components listed above *OR* the response does not involve taking a random sample of two floors out of the nine.

Note: Some possible errors in component 2 include the following:

- Using 10 random digits rather than nine
- Failing to explicitly deal with the issue of potentially repeated random numbers

Part (b) is scored as follows:

Essentially correct (E) if the response indicates the following two components:

1. The amount of carpet wear could be different for apartments with and without children.
2. The stratified random sample ensures that some apartments with children will be selected.

Partially correct (P) if the response includes exactly one of the two components listed above.

Incorrect (I) if the response fails to meet the criteria for E or P.

Notes

- If the response in part (b) says that this stratified sampling method guarantees proportional representation of apartments with and without children, then the second component is satisfied.
- If the sampling procedure in part (a) divides the floors into two groups — those that have apartments with children and those that do not (“prestratification”) — and then selects one floor from each group, score part (b) based on the degree to which a statistical advantage of the stratified sampling in part (b) is addressed.

4 Complete Response

Both parts essentially correct

3 Substantial Response

One part essentially correct and one part partially correct

2 Developing Response

One part essentially correct and one part incorrect

OR

Two parts partially correct

1 Minimal Response

One part partially correct and one part incorrect

AP[®] STATISTICS
2009 SCORING GUIDELINES

Question 3

Intent of Question

2 The primary goals of this question were to assess a student's ability to (1) describe a randomization process required for comparing two groups in a randomized experiment; and (2) describe a potential consequence of using self-selection instead of randomization.

Solution

Part (a) (completely randomized design):

Each student will be assigned a unique random number using a random number generator on a calculator, statistical software, or a random number table. The assigned numbers will be listed in ascending order. The students with the lowest 12 numbers in the ordered list will receive the instructional program that requires physically dissecting frogs. The students with the highest 12 numbers will receive the instructional program that uses computer software to simulate the dissection of a frog.

Part (a) *alternative* (randomized block design):

Students will be paired or placed into blocks of size two, based on having similar pretest scores. So, the first block will contain the two students with the two lowest pretest scores, the second block will contain the two students with the third- and fourth-lowest pretest scores, and so on, with the last block containing the two students with the two highest pretest scores. In each block, the students will be assigned a unique random number using a random number generator on a calculator, statistical software, or a random number table. The student in each block with the lower random number will receive the instructional program that requires physically dissecting frogs, and the student with the higher random number will receive the instructional program that uses computer software to simulate the dissection of a frog.

Part (b):

By not randomizing and allowing the students to self-select, there is a potential for changes to occur in the differences between pretest and posttest scores for a particular group because of the characteristics of students who choose a particular instructional method, not because of the instructional method itself. For example, suppose frog-loving students already know a lot about frog anatomy; one would therefore expect these students to be less likely to show a large change between the pretest and posttest scores. Suppose the frog-loving students tend to select the computer simulation method (perhaps because they do not like the notion of dissecting the frogs they love). The possible low change between pretest and posttest scores for the computer simulation group might then be attributed to the students' already knowing a lot about frog anatomy beforehand, not to the instructional method itself. The frog dissection group might see a larger change in scores because the students entering this group are those with the lower pretest scores (less prior knowledge) and who are thus more likely to show greater improvement between pretest and posttest scores.

Scoring

Parts (a) and (b) are scored as essentially correct (E), partially correct (P), or incorrect (I).

AP[®] STATISTICS

2009 SCORING GUIDELINES

Question 3 (continued)

Part (a) is scored as follows:

Essentially correct (E) if a proper method of randomization is described that (1) creates two groups of equal size; *AND* (2) assigns the named treatments to the groups in a manner that knowledgeable statistics users would employ to assign the students to the two instructional groups.

Partially correct (P) if only one of the two criteria above is met.

Incorrect (I) if neither criterion is met.

Notes:

- Coin tossing (or equivalent method) using a stopping rule to obtain equal sample sizes requires placing the students in the class in a random order. If this method does not include a random order, at best, part (a) is scored as partially correct.
- In using a random number table, if numbers are specified, the student must work with two-digit numbers. For example, if using the first 24 integers, the student must use 01–24, not 1–24. If the student uses numbers such as 1–24, a solution that would otherwise be essentially correct becomes partially correct, and a partially correct response becomes incorrect.

Part (a) alternative is scored as follows:

Essentially correct (E) if (1) blocks are formed based on students' having similar pretest scores; *AND* (2) the two students in each block are assigned to different treatments; *AND* (3) the method of randomization used to assign the students in each block to the treatments is correct and can be implemented after reading the student's response (in a manner that knowledgeable statistics users would employ to assign the students to the two instructional groups).

Partially correct (P) if two of the three components above are presented correctly.

Incorrect (I) if no more than one of the three components is presented correctly.

Part (b) is scored as follows:

Essentially correct (E) if (1) the example gives a reasonable characteristic of the self-selected students in the study; *AND* (2) explains how this characteristic could be associated with changes in the differences between the pretest and posttest scores.

Partially correct (P) if (1) the example gives a reasonable characteristic of the self-selected students in the study; *AND* (2) a weak explanation is provided of how this characteristic could be associated with changes in the differences between pretest and posttest scores.

Note: A *weak* explanation of how a characteristic could be associated with changes in the differences between pretest and posttest scores must at least mention test scores or state that one group will perform better than the other. (Simply mentioning a behavioral difference is not sufficient.)

AP[®] STATISTICS
2009 SCORING GUIDELINES

Question 3 (continued)

Incorrect (I) if an incorrect or no explanation is provided of how a characteristic could be associated with changes in the differences between pretest and posttest scores

OR

the example does not give a reasonable characteristic of the self-selected students in the study

OR

a student says that there must be an equal number of students in the class assigned to each treatment.

4 Complete Response

Both parts essentially correct

3 Substantial Response

One part essentially correct and the other part partially correct

2 Developing Response

One part essentially correct and the other part incorrect

OR

Both parts partially correct

1 Minimal Response

No part essentially correct and only one part partially correct

AP[®] STATISTICS
2007 SCORING GUIDELINES

Question 2

Intent of Question

The three primary goals of this question are to assess a student's ability to: (1) clearly explain the importance of a control group in the context of an experiment; (2) describe the randomization process required for three groups; and (3) reduce variability by grouping experimental units as homogeneously as possible.

Solution

Part (a):

A control group gives the researchers a comparison group to be used to evaluate the effectiveness of the treatments. The control group allows the impact of the normal aging process on joint and hip health to be measured with appropriate response variables. The effects of glucosamine and chondroitin can be assessed by comparing the responses for these two treatment groups with those for the control group.

Part (b):

Each dog will be assigned a unique random number, 001–300, using a random number generator on a calculator, statistical software, or a random number table. The numbers will be sorted from smallest to largest. The dogs assigned the first 100 numbers in the ordered list will receive glucosamine. The dogs with the next 100 numbers in the ordered list will be assigned to the control group. Finally, the dogs with the numbers 201–300 will receive chondroitin.

Part (c):

The key question is which variable has the strongest association with joint and hip health. The goal of blocking is to create groups of homogeneous experimental units. It is reasonable to assume that most clinics will see all kinds and breeds of dogs so there is no reason to suspect that joint and hip health will be strongly associated with a clinic. On the other hand, different breeds of dogs tend to come in different sizes. The size of a dog is associated with joint and hip health, so it would be better to form homogeneous groups of dogs by blocking on breed.

Scoring

Parts (a), (b), and (c) are scored as essentially correct (E), partially correct (P), or incorrect (I).

Part (a) is scored as essentially correct (E) if an advantage of using a comparison group is described in the context of this study.

Part (a) is scored as partially correct (P) if an advantage of using a control group is described but not in the context of this study.

Part (a) is scored as incorrect (I) if the student says that control groups should always be used but gives no further explanation *OR* an incorrect explanation.

AP[®] STATISTICS 2007 SCORING GUIDELINES

Question 2 (continued)

Note: Since “treatment” and “control” are standard terms in design, a comparison of specific aspects of the study is needed to establish context.

Part (b) is scored as essentially correct (E) if randomization is used correctly, and the method of randomization can be implemented after reading the student response (so that two knowledgeable statistics users would use the same method to assign dogs to treatment groups).

Part (b) is scored as partially correct (P) if randomization or chance is used, but the method could not be implemented after reading the student response.

Part (b) is scored as incorrect (I) if randomization or chance is not used in a planned way *OR* the solution does not yield a completely randomized design.

Part (c) is scored as essentially correct (E) if:

the student argues that the variable with the stronger relationship to joint and hip health should be used as the blocking variable;

OR

the student states that the variable with the larger anticipated variability in the response measure should be used as the blocking variable so that units within blocks are as homogeneous as possible. A rationale is required, but a variable does not have to be selected.

Part (c) is scored as partially correct (P) if:

the student indicates that the purpose of blocking is to create groups of homogeneous experimental units but makes an error in the application to this experiment;

OR

the student does not acknowledge that there may be more variability associated in the response variable with one of the variables (breed or clinic) than the other;

OR

the student does not recognize that both variables are associated with variation in the response variable.

Part (c) is scored as incorrect (I) if the student does not exhibit an understanding of the purpose of blocking.

4 Complete Response

All three parts essentially correct

3 Substantial Response

Two parts essentially correct and one part partially correct

2 Developing Response

Two parts essentially correct and no parts partially correct

OR

One part essentially correct and two parts partially correct

OR

Three parts partially correct

AP[®] STATISTICS
2007 SCORING GUIDELINES

Question 2 (continued)

1 Minimal Response

One part essentially correct and either zero or one part partially correct
OR
No parts essentially correct and two parts partially correct