

AP[®] STATISTICS

2016 SCORING GUIDELINES

Question 3

Intent of Question

The primary goals of this question were to assess a student's ability to (1) identify explanatory and response variables from the description of a research study; (2) indicate and justify whether a study is observational or experimental; and (3) explain what confounding means in the context of a particular study with a specific confounding variable.

Solution

Part (a):

The explanatory variable is the person's degree of cigarette smoking. The response variable is whether the person develops Alzheimer's disease during the course of the study.

Part (b):

This is an observational study because the people in the study were not assigned to a certain degree of cigarette smoking. Rather, the degree of cigarette smoking for each person was passively observed and recorded, not manipulated by the researchers.

Part (c):

A confounding variable is one that is related to the explanatory variable and possibly influences the response variable. In this case it seems plausible that people who exercise more regularly might be more health conscious, therefore, less likely to smoke cigarettes than people who do not exercise regularly. Similarly, it's possible that people who exercise more regularly are less likely to develop Alzheimer's disease than people who do not exercise regularly. If both of these relationships turn out to be true, then smoking cigarettes would be associated with developing Alzheimer's disease due to the association of both variables with exercise, even if there were no cause-and-effect relationship between smoking cigarettes and developing Alzheimer's disease.

Scoring

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

Part (a) is scored as follows:

Essentially correct (E) if both variables are described correctly. A correct description includes some degree of status of the variables, such as smoking versus not smoking and developing Alzheimer's versus not developing Alzheimer's.

Partially correct (P) if one variable is described correctly and one is not described correctly;

OR

if neither variable is described correctly but smoking is mentioned as the explanatory variable, and Alzheimer's is mentioned as the response variable;

OR

if the explanatory and response variables are correctly described but are interchanged.

Incorrect (I) if the response does not meet the criteria for E or P.

AP[®] STATISTICS

2016 SCORING GUIDELINES

Question 3 (continued)

Note: Describing the variables only as smoking and Alzheimer's without a reference to levels/status is not sufficient for E. But making the connection between smoking and explanatory and between Alzheimer's and response is sufficient for P.

Part (b) is scored as follows:

Essentially correct (E) if the response selects the correct type of study (observational) and provides the correct explanation that smoking status was not assigned *OR* that smoking status was only observed.

Partially correct (P) if the response selects the correct type of study (observational) and provides a correct explanation but does not refer to smoking status as the variable that would have been assigned had it been an experiment; for instance, by simply stating that treatment was not assigned;

OR

if the response selects the correct type of study (observational) and provides an explanation that says that smoking is observed (or that smoking status and Alzheimer's are observed) without a modifier for observed (such as, only, just, merely, simply) *AND* without indicating that treatments were not assigned.

Incorrect (I) if the response does not meet the criteria for E or P.

Notes:

- A response that states the explanatory variable was not assigned without naming smoking status is sufficient for E if the explanatory variable is correctly defined in part (a).
- A response that states that smoking status is *ONLY* observed (or that smoking status and Alzheimer's are *ONLY* observed) is sufficient for E.
- A response that says that Alzheimer's is observed without mentioning smoking status is scored I.
- A response that provides an incorrect statistical explanation (such as, "the study is observational because an experiment must have a control group") lowers the score in part (b) by one level (from E to P or from P to I).
- If Alzheimer's is given as the explanatory variable and smoking is given as the response variable in part (a), then part (b) should be scored accordingly with the two variables interchanged.
- Any phrase that refers to the "effect of smoking on Alzheimer's" or "the association between smoking and Alzheimer's" (rather than smoking status) should be ignored.

Part (c) is scored as follows:

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

1. Provides a reasonable explanation that exercise status is related to smoking status.
2. States that exercise status might influence whether the person develops Alzheimer's disease.

Partially correct (P) if the response describes only one of the two components;

OR

if the response only describes how smoking and exercise jointly influence whether the person develops Alzheimer's;

OR

AP[®] STATISTICS

2016 SCORING GUIDELINES

Question 3 (continued)

if the response mentions both of the associations in each component, but also includes an incorrect statistical statement. For instance, part (a) is correct, but in part (c), the explanatory and response variables are interchanged.

Incorrect (I) if the response does not meet the criteria for E or P.

4 Complete Response

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 one or two parts partially correct

OR

Three parts partially correct

1 Minimal Response

One part essentially correct

OR

No parts essentially correct and two parts partially correct

3. Alzheimer's disease results in a loss of cognitive ability beyond what is expected with typical aging. A local newspaper published an article with the following headline.

Study Finds Strong Association Between Smoking and Alzheimer's

The article reported that a study tracked the medical histories of 21,123 men and women for 23 years. The article stated that, for those who smoked at least two packs of cigarettes a day, the risk of developing Alzheimer's disease was 2.57 times the risk for those who did not smoke.

- (a) Identify the explanatory and response variables in the study.

Explanatory variable:

The explanatory variable was whether or not the individual smoked at least two packs of cigarettes a day.

Response variable:

The response variable was whether or not the individual developed Alzheimer's disease.

- (b) Is the study described in the article an observational study or an experiment? Explain.

The study described in the article is an observational study, not an experiment. This is because the researchers merely tracked the individuals by analyzing their medical histories. An experiment would need to impose a treatment on the subjects (e.g. assigning them to smoke or not to smoke), which this study did not do, making it an observational study.

- (c) Exercise status (regular weekly exercise versus no regular weekly exercise) was mentioned in the article as a possible confounding variable. Explain how exercise status could be a confounding variable in the study.

Exercise could be a confounding variable as those who exercise regularly may not be as at risk for Alzheimer's disease. Additionally, smokers may be more likely not to have regular weekly exercise compared to nonsmokers. Therefore the ^{observed} association is unknown to have been produced as a result of smoking or a lack of weekly exercise, making exercise status a potential confounding variable.

3. Alzheimer's disease results in a loss of cognitive ability beyond what is expected with typical aging. A local newspaper published an article with the following headline.

Study Finds Strong Association Between Smoking and Alzheimer's

The article reported that a study tracked the medical histories of 21,123 men and women for 23 years. The article stated that, for those who smoked at least two packs of cigarettes a day, the risk of developing Alzheimer's disease was 2.57 times the risk for those who did not smoke.

- (a) Identify the explanatory and response variables in the study.

Explanatory variable: The explanatory variable ~~was~~ is whether or not the women & men involved smoked at least two packs of cigarettes a day or not.

Response variable: The response variable is developing Alzheimer's disease.

- (b) Is the study described in the article an observational study or an experiment? Explain.

The study described in the article is an observational study because treatments were not assigned to the men & women involved in the study. In other words, an experimenter did not randomly select individuals to smoke at least two packs of cigarettes a day, the participants either already smoked that much or didn't.

- (c) Exercise status (regular weekly exercise versus no regular weekly exercise) was mentioned in the article as a possible confounding variable. Explain how exercise status could be a confounding variable in the study.

Exercise status could be a confounding variable in the study because exercise could be linked to a lower risk of developing Alzheimer's disease. In addition, individuals who exercise regularly may be more concerned with their health and less likely to smoke. Therefore, the strong association between smoking and Alzheimer's could not be accurate if more people in the non-smoking category exercise regularly.

3. Alzheimer's disease results in a loss of cognitive ability beyond what is expected with typical aging. A local newspaper published an article with the following headline.

Study Finds Strong Association Between Smoking and Alzheimer's

The article reported that a study tracked the medical histories of 21,123 men and women for 23 years. The article stated that, for those who smoked at least two packs of cigarettes a day, the risk of developing Alzheimer's disease was 2.57 times the risk for those who did not smoke.

- (a) Identify the explanatory and response variables in the study.

Explanatory variable:

Smoking vs. not smoking

Response variable:

Development of Alzheimer's Disease

- (b) Is the study described in the article an observational study or an experiment? Explain.

Observational study because the researchers did not apply any treatment; instead they watched and took data from subjects from two different groups. Since neither group received any form of treatment, this was not an experiment.

- (c) Exercise status (regular weekly exercise versus no regular weekly exercise) was mentioned in the article as a possible confounding variable. Explain how exercise status could be a confounding variable in the study.

It is impossible to determine which members of each group performed weekly exercise. It is also impossible to determine whether weekly exercise either deterred the development of Alzheimer's or exacerbated it. This means that the development of the disease could be affected by either smoking OR exercise, and the researchers do not know which affected the results.

AP[®] STATISTICS

2016 SCORING COMMENTARY

Question 3

Overview

The primary goals of this question were to assess a student's ability to (1) identify explanatory and response variables from the description of a research study; (2) indicate and justify whether a study is observational or experimental; and (3) explain what confounding means in the context of a particular study with a specific confounding variable.

Sample: 3A

Score: 4

In part (a) the explanatory variable is correctly described as a degree of smoking. The description refers both to those individuals who smoke at least two packs of cigarettes per day and to those who do not smoke two packs per day. The response variable is also correctly described as individuals who develop Alzheimer's disease and individuals who do not. Part (a) was scored as essentially correct. The response in part (b) correctly identifies the study as observational and justifies the choice with the explanation that to be an experiment, treatments would have had to have been imposed, and the treatments would have to be levels of the explanatory variable, smoking status. Part (b) was scored as essentially correct. The response in part (c) correctly describes how exercise status could be a confounding variable in this study. The first sentence describes how exercise status could influence the development of Alzheimer's disease, satisfying component 2. The second sentence describes a possible relationship between exercise status and smoking status by explaining that smokers might be less likely to exercise than nonsmokers, and component 1 is satisfied. With both components satisfied, part (c) was scored as essentially correct. Because all three parts were scored as essentially correct, the response earned a score of 4.

Sample: 3B

Score: 3

In part (a) the explanatory variable is correctly described as a degree of smoking. The description indicates that there are two possible categories of the variable: people who smoke at least two packs of cigarettes a day and people who do not. The response variable refers only to those people who are developing Alzheimer's disease and neglects to describe that the variable also includes those who do not develop Alzheimer's disease. The response variable is not correctly described. With only one of the two variables described correctly, part (a) was scored as partially correct. In part (b) the response correctly identifies the study as observational. The response justifies the choice by explaining that the study is observational because treatments were not assigned. The explanation includes a description of the correct treatments based on smoking status that would have been assigned had it been an experiment. Part (b) was scored as essentially correct. The response in part (c) describes how exercise status could influence the development of Alzheimer's disease by explaining that exercise could decrease the risk of developing Alzheimer's, satisfying component 2. The next sentence describes the relationship between exercise status and smoking status by explaining that those who exercise might be less likely to smoke, satisfying component 1. The response satisfies both components of part (c) and was scored as essentially correct. Because two parts were scored as essentially correct, and one part was scored as partially correct, the response earned a score of 3.

AP[®] STATISTICS
2016 SCORING COMMENTARY

Question 3 (continued)

Sample: 3C
Score: 2

The explanatory variable is correctly described in part (a). The description indicates that there is a degree of smoking; that is, people who smoke and people who do not smoke. The response variable is incorrectly described; the description refers only to those people who develop Alzheimer's disease and does not refer to those who do not develop Alzheimer's disease. With only one of the two variables described correctly, part (a) was scored as partially correct. The response correctly identifies the study as observational in part (b). The justification correctly explains that the study cannot be an experiment because treatments were not assigned. However, the response does not explain that the treatments would have been levels of the explanatory variable, smoking status. Because levels of the explanatory variable are not identified in context as the potential treatments, the response in part (b) was scored as partially correct. The response in part (c) explains how exercise could influence the development of Alzheimer's disease by stating that "It is impossible to determine . . . whether weekly exercise either deterred the development of Alzheimer's or exacerbated it." The statement satisfies component 2. The response continues to explain that one cannot determine if it is exercise status or smoking status that influences the development of Alzheimer's disease, but the possible relationship between exercise status and the explanatory variable, smoking status, is never addressed. Therefore component 1 is not satisfied. With only one of the two components satisfied, part (c) was scored as partially correct. Because all three parts were scored partially correct, the response earned a score of 2.