

Past-paper walk-through

Introduction to Planning a Study ■ 3.2

eXercise

Questions in this set

- 2024 Q3
- 2016 Q3

Reading the mark scheme

- **B1** a mark that stands on its own – the point is either made or not.
- **M1** a *method* mark: the right approach, even if the number is wrong.
- **A1** an *answer* mark, and it usually needs its M mark first.
- **C1** a working mark on the way to the answer.
- **//** separates answers the examiner accepts equally.
- **ecf** = error carried forward: a later part is marked on your own earlier answer.
- **owtte** = “or words to that effect” – the wording need not match.

Question 3

2024 ■ Q3

A car maker produces four different models of cars: A, B, C, and D. A group of researchers is investigating which model of car has the longest distance traveled per gallon of gas (mileage). Higher mileage is considered better than lower mileage. The researchers will conduct a study in which they contact several owners of each model of car and ask them to estimate their mileage.

(a) Is this an observational study or an experiment? Justify your answer in context.

Question 3

2024 ■ Q3

A car maker produces four different models of cars: A, B, C, and D. A group of researchers is investigating which model of car has the longest distance traveled per gallon of gas (mileage). Higher mileage is considered better than lower mileage. The researchers will conduct a study in which they contact several owners of each model of car and ask them to estimate their mileage.

(b) Model D has an autopilot feature, in which the car controls its own motion with human supervision. James owns a Model D car and will investigate whether using the autopilot feature results in higher mileage than not using the autopilot. James will drive his car on 70 different days to and from work, using the same route at the same time each day. James will record the mileage each day.

James will use a completely randomized design to conduct his investigation. Describe an appropriate method James could use to randomly assign the two treatments, driving

Question 3

using the autopilot feature and driving without using the autopilot feature, to 35 days each.

Question 3

2024 ■ Q3

A car maker produces four different models of cars: A, B, C, and D. A group of researchers is investigating which model of car has the longest distance traveled per gallon of gas (mileage). Higher mileage is considered better than lower mileage. The researchers will conduct a study in which they contact several owners of each model of car and ask them to estimate their mileage.

(c) After the investigation was completed, James verified that the conditions for inference were met and conducted a hypothesis test. He discovered the mean mileage when using the autopilot feature was significantly higher than the mean mileage when not using the autopilot feature.

James is a member of a Model D club with thousands of members who all drive Model D cars. He will give a presentation at a Model D club members' meeting later this year and would like to state that the results of his hypothesis test apply to all Model D cars in his club. Another member of the club who is a statistician tells James his findings

Question 3

do not apply to all Model D cars in the club. What change would James need to make to his original study to be able to generalize to all Model D cars in the club?

Solution 3

(a)

Mark scheme

- This is an observational study. The researchers had the car owners estimate their own mileage; the car owners were not randomly assigned a car model, so no treatment was imposed by the researchers. Full credit needs: (1) identifies "observational study" (stating it is an experiment scores Incorrect), (2) justification based on no treatment being imposed (e.g. "no random assignment of treatment"), (3) context – car models, drivers, cars, owners, or members (mileage alone is not sufficient context).

Solution 3

(b) Mark scheme

- Number the 70 days in the experiment from 1 to 70. Using a random number generator, generate 35 unique integers from 1 to 70 inclusive (no repeats). Assign the days with those 35 unique integers for James to drive the car WITH autopilot, and assign the remaining 35 days for James to drive WITHOUT autopilot. (Equivalent alternative: label 70 equally-sized slips of paper, 35 "with autopilot" and 35 "without autopilot," mix them in a bag, and each day select one slip without replacement to determine that day's driving method.) Full credit needs: (1) appropriate labels created for the units/treatments (the 70 days, and the two treatments), (2) a correctly described random process so every possible random assignment is equally likely (a coin-flip-with-stopping-rule method does NOT

Solution 3

satisfy this, since it biases the last few days), (3) the process results in exactly 35 days assigned to each treatment, using selection without replacement/without repeats.

Solution 3

(c) Mark scheme

- In order to generalize his findings to all Model D cars in his club, James would need to randomly select Model D cars from the club, then carry out a similar study using the Model D cars that were randomly sampled from the club. Full credit needs: (1) states that more cars must be sampled, (2) states that random sampling is required for generalization, (3) is in context, including sampling from the population of interest (Model D cars/owners/members in his club). Any discussion of experimental design beyond sample selection (e.g. conditions, two-sample design) is ignored in scoring.

Question 3

2016 ■ Q3

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:

Response variable:

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

Explain.

Question 3

(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.

Solution 3

(a)

Mark scheme

- Explanatory variable: the person's degree of cigarette smoking. Response variable: whether the person develops Alzheimer's disease during the course of the study. (Full credit requires both variables described with some indication of status/levels – e.g. smoking vs. not smoking, developing vs. not developing Alzheimer's – not just the bare topic names.)

Solution 3

(b)

Mark scheme

- Observational study – because the people in the study were not assigned a degree of cigarette smoking by the researchers. Instead, each person's smoking status was passively observed and recorded, not manipulated or assigned as a treatment.

Solution 3

(c) Mark scheme

- A confounding variable is related to the explanatory variable and possibly influences the response variable. Here, it is plausible that people who exercise more regularly are more health-conscious and therefore less likely to smoke cigarettes than people who do not exercise regularly (exercise status related to smoking status). It is also plausible that people who exercise more regularly are less likely to develop Alzheimer's disease than those who do not (exercise status may influence the response). If both relationships hold, smoking would appear associated with developing Alzheimer's simply because both are associated with exercise – even if there is no actual cause-and-effect relationship between smoking and Alzheimer's. Full credit requires both components: (1) a reasonable

Solution 3

explanation that exercise status is related to smoking status, and (2) a statement that exercise status might influence whether the person develops Alzheimer's disease.