Exam #1 Review

Ryan Miller

Q1

True or false:

Suppose we take a **random sample** of 10,000 households in the United States. Due to random sampling, the distribution of household income in this sample will be symmetric and bell-shaped.

Be prepared to explain your answer.

Q1 Solution

FALSE random sampling means the sample will be **representative** of the population, if the population is skewed then the sample will be skewed

Q2

A 2005 study found the correlation between insulin and BMI to be 0.67. Furthermore, they found that insulin had a mean of 93.78 and a standard deviation of 9.36 while BMI had a mean of 25.01 and a standard deviation of 4.68.

Suppose an individual's insulin is 0.4 standard deviations below average.

1. What is the z-score for this individual?
2. How many standard deviations above or below average do you expect their BMI to be?

Be prepared to explain your answers.

Q2 Solution

1. We are given the z-score, it's -0.4
2. Using the correlation coefficient -0.4*0.67 = -0.268, so we predict their BMI is 0.268 standard deviations below average

Q3

In one study, participants were randomly assigned to receive a session of cognitive behavioral therapy (CBT) or receive no treatment. They then reported whether they experienced an improvement in sleep quality over the next three months.

1. Was this experiment **double-blind**? Was it **placebo controlled**?
2. A media outlet reported a "strong correlation between CBT and improved sleep quality", why is this an incorrect use of the statistical term *correlation*?

Be prepared to explain your answers.

Q3 Solution

1. It was neither blinded (since participants knew their treatment group) nor placebo controlled (since those who weren't being treated received nothing).
2. These variables are categorical. Correlation specifically describes a linear relationship between quantitative variables. The word that the outlet should have used is *association*.

Q4

A 2011 analyzed brain scans of city-dwellers and country-dwellers as they took a difficult math test. To increase the stress of the test, the researchers tried to humiliate the participants by telling them how poorly they were doing on the test. The brain scans showed different levels of activity in stress centers of the brain, with urban-dwellers having greater brain activity than rural-dwellers in regions that react to stress.

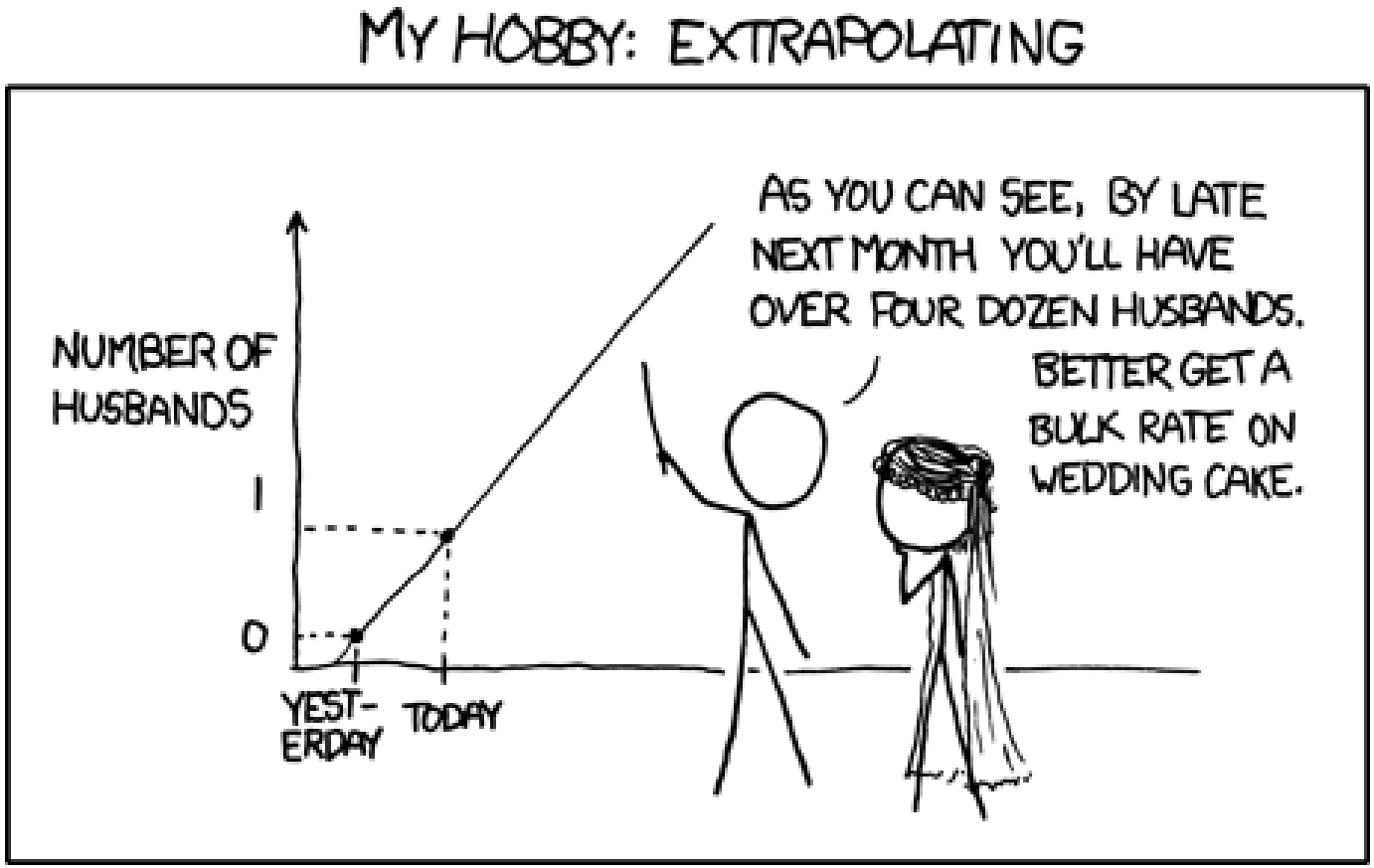
- Is this study and experiment or an observational study?

Q4 Solution

This is an **observational study**, since the researchers did not manipulate the explanatory variable (whether the participant was a city-dweller or country-dweller)

Q5

Explain the mistake the husband is making in this xkcd comic:



Q5 Solution

He shouldn't be inferring anything beyond the range of the observed data