

Sampling

Math 222 - Workshop

A student project group (Norquest, Bayer, & McConville, 2004) wanted to determine whether there was a preference for the primary presidential candidates (George Bush and John Kerry) among first-year students at their university. They believed that these younger students would be more liberal and more inclined to vote for the democratic candidate (Kerry). They took a sample of 30 students from the 705 first-years at their school by assigning the first-year residence halls a number between 1 and 5 (inclusive) and rolling a die to select a residence hall. Then from the selected residence hall they took every seventh room to be part of the sample. The survey was distributed by going from room to room with the surveys, giving each resident of the room a copy of the survey to fill out, and immediately collecting the surveys after they were completed. If one of the residents was not in the room at the time of the first visit, the group repeatedly returned to the room until the person was contacted. The surveys were anonymous and the group members did not look at the completed surveys until they were randomly scrambled. The survey contained three questions, the first question was whether the student planned to vote for Kerry or Bush, and the other two questions concerned other social issues. For half of the surveys Kerry was listed first and for half of the surveys Bush was listed first.

1. Identify the individuals and the variable variable of interest in this study. Is the variable quantitative or categorical?
2. Identify the sample, the population, and the sampling frame (a *sampling frame* is a list of individuals from which the sample is taken).

3. How did the researchers try to avoid sample bias? Did they use a SRS or some other method?
4. What additional steps did the researchers take to avoid non-sample bias (*non-sample bias* is bias that occurs after the sample is selected, examples include leading questions, measurement error, embarrassing questions, the order in which questions occur, etc.)?

In this study the sample results were that 22 of the 30 first-years planned to vote for John Kerry, whereas 8 planned to vote for George Bush.

5. What is the parameter of interest in this study? What symbol would represent it?
6. Can the researchers be confident that bias will not be a problem in this study? If not, describe a potential source of bias. Would it be sample bias or non-sample bias?