

Note: Your questions might have appeared in a different order.

1. (16 pts) On September 28, 2005, the news website MSNBC.com published an article about how consumers can redeem frequent flyer miles from airlines. Within the article was a link to participate in an on-line poll that asked “Do you use a credit card to accumulate airline miles?” Of the 1935 responses that had been received when I checked, 1606 (which is 83% of 1935) had answered “yes.”

a) (4 pts) Is 83% a parameter or a statistic? Explain briefly.

It is a statistic, because it refers to the sample of 1935 people who completed the on-line poll.

b) (6 pts) Use the “quick method” to calculate the margin-of-error for this poll.

The margin-of-error is $\frac{1}{\sqrt{n}} = \frac{1}{\sqrt{1935}}$, which is about .023, or 2.3%.

c) (6 pts) Comment on whether you suspect this sampling method is biased and, if so, in which direction. Explain briefly.

This sampling method is biased. People who choose to read such an article are probably interested in using a credit card to accumulate airline miles, and they probably tend to use such a credit card more than the general public. Thus, this poll almost certainly overestimates the proportion of adult Americans who use a credit card to accumulate airline miles.

2. (36 pts)

a) (6 pts) Suppose that the individuals in a study are the patients who arrive at a hospital’s emergency room in a given day. Name a categorical variable and a quantitative variable that could be recorded on these individuals.

Categorical:

Examples include gender, race, whether the patient arrived in ambulance or not, whether the person required stitches or not, ...

Quantitative:

Examples include age, number of stitches that the patient received, how long the patient waited before seeing a doctor, ...

Note: Summaries are not variables. For example, “number of patients who arrived” and “how many patients needed stitches” are not variables.

b) (6 pts) Recall the class study we conducted in which every student recorded how long it took a chocolate chip and a peanut butter chip to melt in his/her mouth. Even though this was a randomized comparative experiment, we pointed out that it still contained a confounding variable. What was that variable? Explain what it means to say that its effect on the response (time until melting) was confounded with that of the explanatory variable (type of chip).

The confounding variable was the size of the chip. All of the chocolate chips were smaller than all of the peanut butter chips, so we can not distinguish the effect of chip type (chocolate or peanut butter) on the melting time from the effect of chip size on the melting time.

c) (4 pts) Recall the memory study that we conducted in class, in which each student was asked to memorize as many letters as possible in 20 seconds. Half of the students were randomly assigned to see the letters in convenient three-letter groupings, and the other half saw the letters in less convenient groupings. Was this a matched pairs design? (Answer yes or no; no explanation is necessary.)

No, this was not a matched-pairs design, because every student was assigned to only one of the two groupings of letters.

d) (4 pts) A company used to require an IQ test of all of its job applicants, but this practice was declared illegal because IQ was not related to job performance. Is this a criticism of the IQ test's validity, reliability, or unbiasedness? (Choose one of these three; no explanation is necessary.)

This is a criticism of validity, because it concerns whether the IQ test is relevant and appropriate for measuring job performance.

e) (6 pts) Suppose that you want to collect data to investigate whether Cal Poly faculty tend to drive newer cars than Cal Poly students drive. Which would be an appropriate use of randomness in this study: i) random sampling from a population; ii) random assignment of subjects to treatment groups; iii) both; iv) neither? Clearly indicate your answer, and explain briefly.

We would use i), random sampling from the population of Cal Poly students and from the population of Cal Poly faculty. Random assignment is not appropriate because we are not assigning these people to groups or treatments.

f) (4 pts) New York City elected its first black mayor, David Dinkins, in 1989. Samples of voters leaving polling places in that election indicated that Dinkins would receive a higher percentage of the votes than he actually received. Identify the nonsampling error that this illustrates.

This is a response error issue; some voters were probably lying to disguise the fact that they did not vote for the black candidate. (See page 57.)

g) (6 pts) Would it be ethical to conduct an experiment to investigate whether smoking during pregnancy causes babies to weigh less at birth? Explain briefly.

No, because an experiment would require assigning some pregnant mothers to smoke, an activity that is widely believed to have harmful effects for the mother and baby.

3. (12 pts) According to the California Association of Realtors, the median home price in San Luis Obispo county in February of 2006 was \$604,170. One year earlier, in February of 2005, the median price had been \$496,430. Therefore, the median home price increased by \$107,740 over that year.

a) (6 pts) Calculate the percentage increase in the median home price over that year.

The percentage increase is $107,740 / 496,430 = .217$, or 21.7%.

b) (6 pts) If the median home price in Bakersfield increased by the same percentage as in San Luis Obispo over that year, do you think that it would increase by more or less than \$107,740? Explain briefly, and be sure to state any assumption that you make about home prices in Bakersfield.

Houses cost less in Bakersfield than in SLO, so the same percentage increase would result in an increase of less than 107,740.

4. (36 pts) A recent study examined the effectiveness of a nicotine lozenge for helping smokers to quit smoking. Newspaper advertisements sought volunteers who were smokers interested in quitting. Those selected to participate in the study were randomly assigned to receive either the nicotine lozenge or a placebo lozenge (with no active ingredient). At the end of six weeks, researchers asked the subjects whether they had successfully refrained from smoking for that time.

a) (4 pts) Who/what are the individuals in this study?

The individuals are the smokers who want to quit, respond to the ad, and are chosen for the study.

b) (8 pts) Identify the explanatory and response variable in this study. Also classify each as numerical or categorical.

Explanatory: Type of lozenge (nicotine or placebo) Type: categorical

Response: Whether the person successfully abstained from smoking for six weeks Type: categorical

c) (6 pts) Is this an observational study or an experiment? Explain briefly.

This is an experiment, because subjects were assigned to receive a certain type of lozenge.

d) (6 pts) Explain the purpose of the random assignment in this study.

Randomization balances out all variables (other than lozenge type) between the two groups. For example, the smokers' ages, genders, frequency of smoking, and all other variables should be even out between the nicotine and placebo groups. This way, if one group has a significantly higher rate of abstaining than the other, we can attribute that to the lozenge type.

e) (6 pts) Researchers found that smokers in the nicotine lozenge group were highly statistically significantly more likely to successfully refrain from smoking than those in the placebo group. Is it legitimate to conclude that the nicotine lozenge caused the increase in successful abstention from smoking? Explain briefly.

Yes. Because this was a randomized experiment, and because the increase was highly statistically significant, the only explanation for the two groups to differ significantly is that the nicotine lozenge was responsible for increasing the abstention rate.

f) (6 pts) Explain what the phrase "highly statistically significantly" means in this context.

This means that the difference in success rates between the groups would have been very unlikely to happen solely by random, chance variation if there were really no effect of the nicotine lozenge.