PRACTICE XIII

1. Nine sales representatives, 6 men and 3 women, at a small company wanted to attend a national convention. There were only enough travel funds to send 3 people. The manager selected 3 people to attend and stated that the people were selected at random. The 3 people selected were women. There were concerns that no men were selected to attend the convention.

   (a) Calculate the probability that randomly selecting 3 people from a group of 6 men and 2 women will result in selecting 3 women.

   • there are 8 people total with 6 men and 2 woman

   $$P(\text{selecting 3 woman}) = \frac{3 \cdot 2 \cdot 1}{8 \cdot 7 \cdot 6}$$

   $= P(\text{first is woman}) \cdot P(\text{second is a woman} \mid \text{first is a woman}) \cdot P(\text{third is woman} \mid \text{first two women})$

   $$= \left(\frac{3}{8}\right)\left(\frac{2}{7}\right)\left(\frac{1}{6}\right) \approx 0.012$$

   (b) Based on your answer to part (a), is there reason to doubt the manager's claim that 3 people were selected at random? Explain.

   there is a 1.2% chance of randomly selecting three woman. This provides reason to doubt that the manager was choosing completely at random.

An alternative to calculating the exact probability is to conduct a simulation to estimate the probability. A proposed simulation process is described below:

Each trial in the simulation consists of rolling three fair, six-sided dice, one die for each of the convention attendees. For each die, rolling a 1, 2, 3, or 4 represents selecting a man; rolling a 5 or 6 represents selecting a woman. After 1,000 trials, the number of times the dice indicate selecting 3 women is recorded.

(c) Does the proposed process correctly simulate the random selection of 3 women from a group of 9 people consisting of 6 men and 3 women? Explain why or why not.

- No the process does not correctly simulate the random selection of 3 women from a group of 9 consisting of 6 men and 3 women.
- The random selection of three people among 9 is done without replacement. In the simulation with the dice, the three dice rolls in any given trial are independent of one another, indicating a selection process done with replacement.

2. As part of its twenty fifth reunion celebration,theclassof1988(students who graduated in 1988) at a state university held a reception on campus. In an informal survey, the director of alumni development asked 50 of the attendees about their incomes. The director computed the mean income of the 50 attendees to be $189,952. In a news release, the director announced, "The members of our class of 1988 enjoyed resounding success. Last year's mean income of its members was $189,952!"

   (a) What would be a statistical advantage of using the median of the reported incomes, rather than the mean, as the estimate of the typical income?

The median is less affected by skewness and outliers than the mean. With a variable such as income, a small number of very large incomes could dramatically increase the mean but not the median.

Therefore, the median would provide a better estimate of a typical income value.

(b) The director felt the members who attended the reception may be different from the class as a whole. A more detailed survey of the class was planned to find a better estimate of the income as well as other facts about the alumni. The staff developed two methods based on the available funds to carry out the survey.

- Method 1: Send out an e-mail to all 6,826 members of the class asking them to complete an online form. The staff estimates that at least 600 members will respond.
- Method 2: Select a simple random sample of members of the class and contact the selected members directly by phone. Follow up to ensure that all responses are obtained. Because method 2 will require more time than method 1, the staff estimates that only 100 members of the class could be contacted using method 2.

Which of the two methods would you select for estimating the average yearly income of all 6,826 members of the class of 1988 ? Explain your reasoning by comparing the two methods and the effect of each method on the estimate.

Method 2 is better than Method 1. A sample obtained from method 1 could be biased because of the voluntary nature of the response.

It is plausible that class members with larger incomes might be more likely to return the form than class members with smaller incomes.

The mean income for such a sample would overestimate the mean income of all class members.

With Method 2, despite the smaller sample size, the random selection is likely to result in a sample that is more representative of the entire class and produce an unbiased estimate of mean yearly income of all class members..