

Math 141

Identification of Problem Types

Recall the following notation:

- X : categorical variable with 2 groups
- Y : categorical variable with 3+ groups
- W : continuous variable

For each of the following problems,

- a. identify the model type (e.g., $X_1 \sim X_2$),
- b. sketch out what the data frame might look like (you'll have to come up with sensible names for the data set and variables),
- c. draw a sketch of an effective visualization, give the name of that type of plot, and provide pseudo R-code to create that plot using ggplot2,
- d. determine the most common type of hypothesis test for this sort of data,
- e. write the null and alternative hypotheses, and
- f. assuming conditions are met, provide the named distribution (e.g., $t(df = 22)$) that could be used to approximate the null distribution.
- g. write out pseudo R code to approximate the null distribution using a computational technique: permutation or simulation (use the same names as in part b).

-
1. During labor, mothers whose fetuses show abnormal heart rate patterns are often administered oxygen in the hope of increasing the percentage of oxygen in the blood of the fetus. A recent study recorded this fetal oxygen for 24 mothers whose fetuses showed abnormal heart patterns. These measurements were taken with room air as the baseline (statistic of 43.54) and after administering 40% oxygen to the mother (statistic of 48.42). Test whether fetus oxygen levels were significantly different, on average, after the mother was administered 40% oxygen than at baseline.
 2. Cocaine addicts need the drug to feel pleasure. Perhaps giving them a medication that fights depression will help them stay off cocaine. A three-year study compared an antidepressant called desipramine with lithium (a standard treatment for cocaine addiction) and a placebo. The subjects were 72 randomly selected chronic users of cocaine who wanted to break their drug habit. Twenty-four of the subjects were randomly assigned to each treatment. Of those given the desipramine treatment, 14 did not relapse; of the lithium, 6 did not relapse; and, of the placebo, 4 did not relapse. Are these data good evidence that the proportions of successes for the three treatments differ in the population of all cocaine addicts who no longer wish to be addicts?
 3. The Survey of Study Habits and Attitudes (SSHA) is a psychological test that measures students' study habits and attitude toward school. Scores range from 0 to 200. The mean score for U.S. College students is about 115. A teacher suspects that older students have different attitudes toward school than the national average. She gives the SSHA to 25 randomly selected students who are at least 30 years old. The mean of these 25 scores is 118.6 and the standard deviation is about 30.
 4. Different kinds of companies compensate their key employees in different ways. Established companies may pay higher salaries, while new companies may offer stock options that will be valuable if the company succeeds. One study looked at a random sample of 200 companies. Of these, 91 were listed in the *Directory of Public High Technology Corporations* and 109 were not listed. Treat these two groups

as SRSs of high-tech and non-high-tech companies. Seventy-three of the high-tech companies and 75 of the non-high-tech companies offered incentive stock options to key employees. Is there evidence that a higher proportion of high-tech companies offer stock options?

5. Gift cards have become a popular present. Retailers like these cards because they are easier to process than paper gift certificates and more difficult to forge. Customers appreciate the convenience; the cards make great stocking stuffers and are easy to mail. A random sample of 14 credit card-type gift certificates from Nordstrom was obtained, and the purchased value of each was recorded (average of \$24.07, standard deviation of \$6.12). A different random sample of size 16 was collected at Macy's, and the dollars were noted as well (average of \$26.61, standard deviation of \$7.08). Is there any evidence to suggest that the true mean values differ between Macy's and Nordstrom?
6. If you are a dog lover, perhaps having your dog along reduces the effect of stress. To examine the effect of pets in stressful situations, researchers recruited 45 women who said they were dog lovers. Fifteen subjects were randomly assigned to one of three groups to do a stressful task alone, with a good friend present, or with their dog present (each group had 15 subjects). The stressful task was to count backward by 13s or 17s. The subject's heart rate during the task is one measure of the effect of stress. The data below provides summary measures from the experiment. Can we believe that the heart rate of dog lovers is related to stress when working on a task alone, with a friend, or with a dog?

| Group | Mean | Standard Deviation |
|---------|--------|--------------------|
| Control | 82.524 | 9.242 |
| Friend | 81.325 | 8.341 |
| Pet | 73.483 | 9.970 |

7. Jill is testing an octahedral die to see if it is biased. She rolled the die 80 times with the following results:

| Score | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|------------------|----------|-----------|-----------|----------|-----------|-----------|-----------|----------|
| Frequency | 7 | 10 | 11 | 9 | 12 | 10 | 14 | 7 |

8. There are many urban legends involving a full moon and human behavior. Research at the University of Basel in Switzerland suggests that sleep is associated with the lunar cycle. In a new sleep study, 120 random adults were selected and studied during a full moon phase. Melatonin levels were used to determine whether each person experienced a deep sleep or had trouble sleeping during the full moon. 76 people were found to have trouble sleeping. Is there any evidence to suggest that a majority of all people have trouble sleeping during a full moon?