**Assignment 7**

8.58

8.71

9.37

9.41

9.50

**8.58 Teeth and military service.**

In 1898 the United States and Spain fought a war over the U.S. intervention in the Cuban War of

Independence. At that time the U.S. military was concerned about the nutrition of its recruits. Many did not have a sufficient number of teeth to chew the food provided to soldiers. As a result, it was likely that they would be undernourished and unable to fulfill their duties as soldiers. The requirements at that time specified that a recruit must have “at least four sound double teeth, one above and one below on each side of the mouth, and so opposed” so that they could chew food. Of the 58,952 recruits who were under the age of 20,68 were rejected for this reason. For the 43,786 recruits who were 40 or over, 3801 were rejected.

(a) Find the proportion of rejects for each age group.

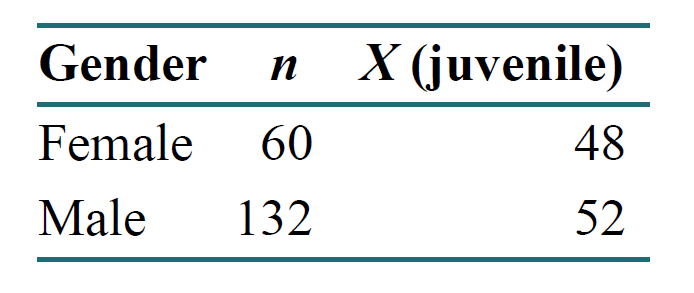
(b) Find a 99% confidence interval for the difference in the proportions.

(c) Use a significance test to compare the proportions. Write a short paragraph describing your results and conclusions.

(d) Are the guidelines for the use of the large-sample approach satisfied for your work in part (b) and (c)? Explain your answers.

**8.71 Gender bias in textbooks.**

To what extent do syntax textbooks, which analyze the structure of sentences, illustrate gender bias? A study of this question sampled sentences from 10 texts.23 One part of the study examined the use of the words “girl,” “boy,” “man,” and “woman.” We will call the first two words *juvenile* and the last two *adult*. Is the proportion of female references that are juvenile (girl) equal to the proportion of male references that are juvenile (boy)? Here are data from one of the texts:



(a) Find the proportion of juvenile references for females and its standard error. Do the same for the males.

(b) Give a 90% confidence interval for the difference and briefly summarize what the data show.

(c) Use a test of significance to examine whether the two proportions are equal.

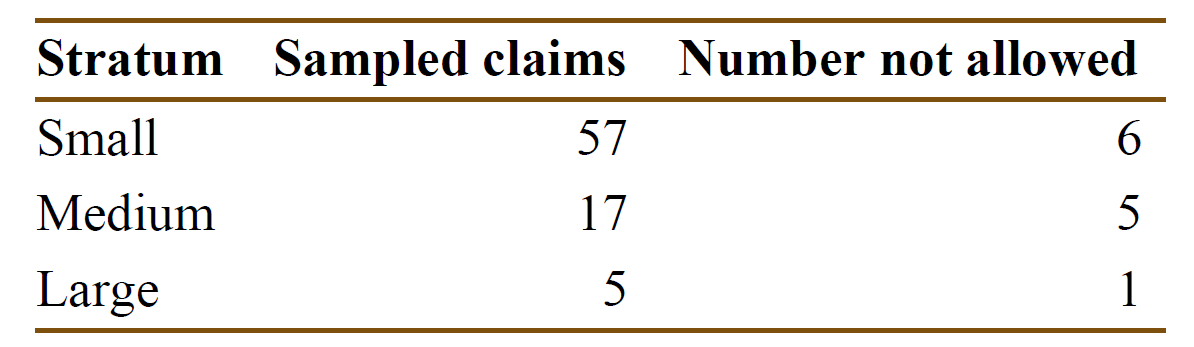
**9.37 Health care fraud**

Most errors in billing insurance providers for health care services involve honest mistakes by

patients, physicians, or others involved in the health care system. However, fraud is a serious

problem. The National Health Care Anti-fraud Association estimates that approximately $68 billion is lost to health care fraud each year.15 When fraud is suspected, an audit of randomly selected billings is often conducted. The selected claims are then reviewed by experts, and each claim is classified as allowed or not allowed. The distributions of the amounts of claims are frequently highly skewed, with a large number of small claims and a small number of large claims. Since simple random sampling would likely be overwhelmed by small claims and would tend to miss the large claims, stratification is often used. See the section on stratified sampling in Chapter 3.

Here are data from an audit that used three strata based on the sizes of the claims (small, medium, and large):



(a) Construct the 3 × 2 table of counts for these data that includes the marginal totals.

(b) Find the percent of claims that were not allowed in each of the three strata.

(c) To perform a significance test, combine the medium and large strata. Explain why we do this.

(d) State an appropriate null hypothesis to be tested for these data.

(e) Perform the significance test and report your test statistic with degrees of freedom and the *P*value. State your conclusion.

**9.41 When do Canadian students enter private career colleges?**

A survey of 13, 364 Canadian students who enrolled in private career colleges was conducted to

understand student participation in the private postsecondary educational system.19 In one part of the survey, students were asked about their field of study and about when they entered college.

Results Attached : **CANF**

In the table, the second column gives the number of students in each field of study. The next two columns give the marginal distribution of time of entry for each field of study.

(a) Use the data provided to make the 6 × 2 table of counts for this problem.

(b) Analyze the data.

(c) Write a summary of your conclusions. Be sure to include the results of your significance testing as well as a graphical summary.

**9.50 Goodness of fit to a standard Normal distribution**

Computer software generated 500 random numbers that should look as if they are from the standard Normal distribution. They are categorized into five groups: (1) less than or equal to −0.6; (2) greater than −0.6 and less than or equal to −0.1; (3) greater than −0.1 and less than or equal to 0.1; (4) greater than 0.1 and less than or equal to 0.6; and (5) greater than 0.6. The counts in the five groups are 139, 102, 41, 78, and 140, respectively. Find the probabilities for these five intervals using Table A. Then compute the expected number for each interval for a sample of 500. Finally, perform the goodness-of-fit test and summarize your results.