

If you had a CQUEST acct last term in sta220, you can use the same account right now, and Minitab should be functional; but if you did not take sta220 last term, you need to apply for an account at www.cquest.utoronto.ca, and wait one or two business days.

ASSIGNMENT 1: Due Jan 16, 2012, at tutorial (Quiz 1 on homework material, **don't forget to bring a calculator!**)

Read Lecture 1, and chapter 26 (pages 714-722) from the textbook (SDM).

Bring your complete solutions for the following problems to your tutorial for the class discussion.

Problem 1:

Waking versus bedtime symptoms. As part of the study on ongoing fright symptoms due to exposure to horror movies at a young age, the following table was presented to describe the lasting impact these movies have had during bedtime and waking life:

Bedtime symptoms	Waking symptoms	
	Yes	No
Yes	36	33
No	33	17

- What percent of the students have lasting waking-life symptoms?
- What percent of the students have both waking-life and bedtime symptoms?
- Test whether there is an association between waking-life and bedtime symptoms. State the null and alternative hypotheses, the X^2 statistic, and the P -value.

Problem 2: Construct a 3 x 2 table of counts where there is no apparent association between the row and column variables.

Problem 3:

Air pollution from a steel mill. One possible effect of air pollution is genetic damage. A study designed to examine this problem exposed one group of mice to air near a steel mill and another group to air in a rural area and compared the numbers of mutations in each group.¹⁸ Here are the data for a mutation at the *Hm-2* gene locus:

Mutation	Location	
	Steel mill air	Rural air
Yes	30	23
No		
Total	96	150

- Fill in the missing entries in the table.
- Summarize the data numerically and graphically.
- Is there evidence to conclude that the location is related to the occurrence of mutations? Perform the significance test and summarize the results.

Problem 4:

Identity theft. A study of identity theft looked at how well consumers protect themselves from this increasingly prevalent crime. The behaviors of 61 college students were compared with the behaviors of 59 nonstudents.²⁰ One of the questions was “When asked to create a password, I have used either my mother’s maiden name, or my pet’s name, or my birth date, or the last four digits of my social security number, or a series of consecutive numbers.” For the students, 22 agreed with this statement while 30 of the nonstudents agreed.

- Display the data in a two-way table and perform the chi-square test. Summarize the results.
- Reanalyze the data using the methods for comparing two proportions that we studied in the previous chapter. Compare the results and verify that the chi-square statistic is the square of the z statistic.
- The students in this study were junior and senior college students from two sections of a course in Internet marketing at a large northeastern university. The nonstudents were a group of individuals who were recruited to attend commercial focus groups on the West Coast conducted by a lifestyle marketing organization. Discuss how the method of selecting the subjects in this study relates to the conclusions that can be drawn from it.

Problem 5:

Secondhand stores. Shopping at secondhand stores is becoming more popular and has even attracted the attention of business schools. A study of customers' attitudes toward secondhand stores interviewed samples of shoppers at two secondhand stores of the same chain in two cities. The breakdown of the respondents by gender is as follows:²⁶

Gender	City 1	City 2
Men	38	68
Women	203	150
Total	241	218

Is there a significant difference between the proportions of women customers in the two cities?

- State the null hypothesis, find the sample proportions of women in both cities, do a two-sided z test, and give a P -value using Table A.
- Calculate the X^2 statistic and show that it is the square of the z statistic. Show that the P -value from Table F agrees (up to the accuracy of the table) with your result from (a).
- Give a 95% confidence interval for the difference between the proportions of women customers in the two cities.

More on secondhand stores. The study of shoppers in secondhand stores cited in the previous exercise also compared the income distributions of shoppers in the two stores. Here is the two-way table of counts:

Income	City 1	City 2
Under \$10,000	70	62
\$10,000 to \$19,999	52	63
\$20,000 to \$24,999	69	50
\$25,000 to \$34,999	22	19
\$35,000 or more	28	24

Verify that the X^2 statistic for this table is $X^2 = 3.955$. Give the degrees of freedom and the P -value. Is there good evidence that customers at the two stores have different income distributions?