Statistics and Data Analysis
Homework 5: Hypothesis Testing

Instructor: Ling-Chieh Kung
Department of Information Management
National Taiwan University

1. According to Whatever.com, the average price charged to a customer to have a 12' by 18' wall-to-wall carpet shampoo cleaned is about \$50. Suppose that a start-up carpet-cleaning company believes that in the region in which they operate, the average price for this service is higher.

   (a) Write down the statistical hypothesis for testing whether the average price is really higher than \$50.

   (b) The carpet-cleaning company randomly contacts 25 customers who have recently had a 12' by 18' wall-to-wall carpet shampoo cleaned and asked the customers how much they were charged for the job. The resulting data are given in the sheet "Carpet" of the MS Excel file "SDA-Fa14_hw05_data.xlsx". Suppose the population standard deviation price is known to be \$3.25 and these prices are normally distributed in the population. What is the p-value?

   (c) For a 95% confidence level, what is the decision? How about 99%?

2. A hole-punch machine is set to punch a hole 1.9 centimeters in diameter in a strip of sheet metal in a manufacturing process. The strip of metal is then creased and sent on to the next phase of production, where a metal rod is slipped through the hole. It is wondering whether the holes are really punched to the specified diameter of 1.9 cm in average.

   (a) Write down the statistical hypothesis for testing whether the average diameters of the holes is 1.9 cm.

   (b) To test punching accuracy, technicians have randomly sampled 10 punched holes and measured the diameters. The data (in centimeters) are given in the sheet "Hole" of the MS Excel file "SDA-Fa14_hw05_data.xlsx". Assume the punched holes are normally distributed in the population. What is the p-value?

   (c) Use a 90% confidence level to determine whether the holes are being punched an average of 1.9 centimeters. How about 99%?

3. Based on population figures and other general information on the U.S. population, suppose it has been estimated that, on average, a family of four in the United States spends about \$1,200 annually on dental expenditures. Suppose further that a regional dental association wants to test to determine if this figure is accurate for their area of the country.

   (a) Write down the statistical hypothesis for testing whether the average expenditure is really \$1,200.

   (b) To test this, 25 families of four are randomly selected form the population in that area of the country and a log is kept of the family's dental expenditures for one year. The resulting data are given in the sheet "Dental" of the MS Excel file "SDA-Fa14_hw05_data.xlsx". What is the p-value?

   (c) Assuming that dental expenditures are normally distributed in the population, what is the p-value?

   (d) Use a 95% confidence level to determine whether the holes are being punched an average of 1.9 centimeters. How about 99%?

4. A study by Hewitt Associates in 2010 showed that 76% of consumers in a specific region consume at least one bottle of milk per month. Suppose a market researcher believes that the percentage has become lower. If this is true, a new promotion plan will be launched.

   (a) Write down the appropriate statistical hypothesis.

(b) The researcher randomly selects 315 consumers and conducts interviews. If a consumer buys at least one bottle per month, a "1" is recorded in the MS Excel file "SDA-Fa14_hw05_data.xlsx". Otherwise, a "0" is recorded. May we apply the z test for this set of sample data? Why?

(c) Use a 99% confidence level to complete the test. Should a new promotion plan be launched?