ADM 2304 -- ASSIGNMENT 1

Due Date: *Saturday, February 2, 2013 at midnight.*

Instructions:

1. *The first question must be completed on MyStatLab. The remainder of the assignment should be uploaded in a separate document using the "Document Sharing" option in MyStatLab (please use the "Assignment 1" folder if possible). Please submit* ***.pdf*** *files only, and name your file using the 'Assignment1' label plus your name. Please do not include characters like "#" in your filename as MyStatLab will not accept such files. The TA will mark your assignments with comments in the answer sections of the custom questions on MyStatLab, provided you do Question 1.*
2. *You may use Minitab or other software for any calculations. However, you must show your manual calculations when asked. You may paste your output onto your assignment to show your use of Minitab; however, this output does not replace any of the steps outlined in 3. This means that answers that are exclusively Minitab output will only get very partial marks.*
3. *If you are performing a hypothesis test, make sure you state the hypotheses, the level of significance, the decision rule in terms of the critical value, the test statistic or p-value, your decision (whether to reject or not to reject the null hypothesis), and your conclusion in managerial terms. These steps must be completed in addition to any Minitab output.*
4. *The data in the Minitab data files can be found in the file* ***Assign1Data.xls****.*
5. *Remember to include your integrity statement.*

1. This question must be completed on MyStatLab. Do your calculations as accurately as possible and round to the nearest dollar. You should check your answer using Minitab before submitting it as you may not be able to revisit this question.

2. In the 2011 federal election, 39.62% of the electorate voted for the Conservative party, 30.63% for the NDP, and 18.91% for the Liberal Party.

(a) In the fall, a Leger Marketing poll of 1500 respondents found 31% would support the federal Liberals under Justin Trudeau and 31% would support the Conservative Party under Stephen Harper. Test whether this is sufficient evidence to show that Justin Trudeau would help decrease the Conservative vote by more than 5% from their share of the popular vote in 2011. Use the .01 level of significance and show your manual calculations.

(b) Suppose you wanted to estimate the national level of support for the Conservatives (assuming that Justin Trudeau were the Liberal leader) using a 99% 2-sided confidence interval with a margin of error of ±1%. What sample size would be required?

(c) Suppose that, in a random sample of 16 University of Ottawa students, only 3 indicated a preference for the Conservatives. Test whether this is sufficient evidence to indicate that the level of support for the Conservatives among U of O students is lower than the 39.62% share of the popular vote in 2011. Use the .05 level of significance and show how you would calculate the p-value for this test. Why does this not allow us to infer anything about the national level of support for the Conservatives?

3. A file on DocDepot in the assignments folder on doc-depot called **Toronto.mtp** contains data on the median (**medinc**) and average incomes (**avginc**) of census dissemination areas in the Toronto Census Division.

(a) Treating this set of data as the population, use Minitab to calculate the population mean for the **avginc** variable. Set aside all population information until part e.
(b) Examine a boxplot and histogram of the population data. Explain if the mean of a random sample of size 31 would have a normal distribution.
(c) Now use Minitab (Calc Menu – Random Data – Sample from Columns) to draw twenty samples of size n = 31 from the Toronto **avginc** population. This procedure must be replicated twenty times (note that if you open up the same sampling dialog box each time from the menu, then you only have to replace the last destination column with the next one). Using each sample, use Minitab to calculate a 95% confidence interval estimate for the population mean, assuming you do not know the population standard deviation (this interval estimation can be done in one operation on all twenty columns).
(d) For the first sample, confirm the Minitab generated interval by calculating the interval manually. Display the sample data graphically and comment on whether the relevant assumption regarding the population distribution is warranted (state clearly the assumption needed to justify the interval estimation).
(e) Count the number of intervals out of your twenty, that contain the true value of the population mean from part a.
(f) Based on your observations regarding the probability distribution in (b) and the sample distribution in (e), comment on whether you should expect that 95% of the confidence intervals calculated by all students collectively will cover the true value of the population mean.

4. A second file on DocDepot named **MarksF12.mtp** contains the final exam marks (**FX**) of students in ADM2304 from the Fall 2012 semester. Assume that the class can be considered a random sample from the population of commerce students. The **Attend** variable counts the number of classes each student attended out of five classes when attendance was recorded. The FX data have been "unstacked" into six separate columns, with the FX_3 and FX_5 columns comprising final exam marks for students who came to class 3 times out of 5 and 5 times out of 5, respectively.

(a) Assuming unequal population variances, test whether there is a difference in average final exam marks between the students who come to class 60% of the time and those who come to class 100% of the time in the population. Use the critical value approach and the .01 level of significance. Perform the test manually after using Minitab to summarize the data and/or to calculate the degrees of freedom.
(b) Now calculate manually the 99% confidence interval for the true difference between the two population means.
(c) Explain how the result in (b) is consistent with your conclusion in (a).
(d) Examine the distributions of the sample data to determine if the 2 sample t-test is valid. Comment.
(e) If the 2-sample t-test were not valid, then identify another test you could use. Explain.