

University of Sydney

ECMT1010 Written Assignment (2016, Semester 2)

Lecturer: Simon Kwok

DUE DATE: 5:00pm, 28 October 2016 (Friday)

General instructions:

- **NO plagiarism!** This assignment must be all your own work. You will sign a formal declaration to that effect on your assignment cover sheet. Any form of cheating (including copying the work of others) is strictly prohibited and will lead to severe penalty. Do your own work and be fair to your fellow classmates.
- This assignment is worth a total of 10 marks and counts towards 10% of your final grade. Maximum marks are indicated in the questions below.
- The assignment is due by **5:00pm on 28 October 2016 (Friday)**. Late submissions will be penalized.
- The assignment must be **submitted in electronic copy by the due date and time**. Submit your assignment to the blackboard through Turnitin by following the instructions under “Written assignment” on the blackboard.
- Review the Academic Honesty Compliance Declaration under “Written assignment” on the blackboard, and click on the “Mark Reviewed” button.
- Different students will use different data sets. The data sets are numbered from 0 to 9. They can be downloaded under “Written assignment” on the blackboard. **Use the data set that has a number equal to the last digit of your SID.**
- Show all numerical answers to 2 decimal places. Where written explanations are required, you should provide **no more than 40 words**.
- Keep your answers concise and confine your assignment submission **to no more than 5 pages**.
- Attempt all questions.

Aim: By the time you finish this written assignment, you should be able to appreciate the usefulness of various statistical techniques in economic applications. You will have the chance to make use of computing software (e.g., Excel, StatKey) in analyzing a real economic data set.

Data description: You have been assigned a data set consisting of 100 randomly-selected employed workers from the May 1978 current population survey conducted by the U.S. Department of Commerce. This is a survey of over 50,000 households conducted monthly, and it serves as the basis for U.S. employment and unemployment statistics. Data are collected on a number of individual characteristics as well as employment status. This data extract contains information on wage, gender, and years of education for the 100 employed workers in the sample.

The data set is available in the Excel spreadsheet **wages.xlsx**. It contains 4 columns and 101 rows. The first row contains the variable names; the remaining 100 rows contain the information for each of the 100 workers in your sample. The column Obs identifies each worker (and can be ignored), WAGE is each worker's hourly wage rate (in 1978 U.S. dollars), FEM is gender (0 = male, 1 = female), and EDUC is years of education (i.e., total years of schooling).

Hints: One handy way to import the data into StatKey is by direct copy-and-paste. To do this, copy any required columns in the data set to a new spreadsheet. Now with all the required columns side-by-side, copy and paste them all at once to the text-input region in the "Edit Data" pop-up window.

Assignment Questions

1. Using Excel or other appropriate software, produce two separate histograms for the wage rates of males and females respectively. Use an appropriate number of bins for your histogram and remember to label the axes. Describe and compare the two histograms, including the central location, dispersion and skewness. [1 mark]
2. Compute the sample means and sample standard deviations of wage rates earned by males and females, respectively. No need to show your computational steps. [1 mark]
3. Using Excel or other appropriate software, produce a scatterplot of wage rates and years of education using your sample. Compute the sample correlation. No need to show your computational steps. [1 mark]

Amy claims that the wage rates and years of education are linearly related. To investigate this claim formally, you carry out a hypothesis test with the help of a regression model.

4. Write down the regression equation, and set up the null and alternative hypotheses. Define the notation(s) clearly. [1 mark]
5. What is the predicted wage rate for an individual with 12 years of education. [1 mark]
6. Using the data in the sample, test Amy's claim at a 5% significance level. [1 mark]

Ben claims that men and women are paid different amounts on average. To investigate the claim formally, you carry out a hypothesis test – this time by constructing a bootstrap confidence interval.

7. Set up the null and alternative hypotheses. Define the notation(s) clearly. [1 mark]
8. Using StatKey or other appropriate software, produce a dotplot of the bootstrap distribution (with at least 5,000 bootstrap samples) of the appropriate sample statistic. Comment on its shape (e.g., central location, degree of symmetry). [1 mark]
9. With reference to the bootstrap distribution in question 8, construct two 95% bootstrap confidence intervals, one using the percentile method and another one based on normal distributions. What can you conclude about Ben's claim at a 5% significance level? [1 mark]
10. Do you think men and women are paid differently because of the gender difference? Explain. [1 mark]

Remember to submit the assignment along with a completed cover sheet in front. Please submit your assignment electronically through Turnitin on the blackboard.

DUE DATE: 5:00pm, 28 October 2016 (Friday)