

Statistics I, Fall 2012

Homework 11

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

- (10 points; modified from Problem 8.20 in the textbook) Some fast-food chains offer a lower-priced combination meal in an effort to attract budget-conscious customers. One chain tested a burger, fries, and a drink combination for \$1.71. The weekly sales volume for these meals was impressive. Suppose the chain wants to estimate the average amount its customers spent on a meal at their restaurant while this combination offer was purchased. An analyst gathers data from 20 randomly selected customers who have purchased the combination. The data (in \$) are recorded in the sheet “FastFood” of the MS Excel file “StatFa12_hw11.xlsx”. Use these data to construct a 90% confidence interval to estimate the average amount of dollars spent by customers who buy this combination. Assume the amounts spent are normally distributed.
- (10 points; modified from Problem 8.25 in the textbook) How much experience do supply chain transportation managers have in their field? Suppose in an effort to estimate this, 40 supply chain managers are surveyed and asked how many years of managerial experience they have in this field. Survey results (in years) are recorded in the sheet “SupplyChain” of the MS Excel file “StatFa12_hw11.xlsx”. Use these data to construct a 99% confidence interval to estimate the mean number of years of experience in supply chain.
- (10 points) For a particular candidate in an election, a survey has been conducted for the proportion of registered voters supporting him. In this survey, the responses of 1000 randomly selected voters have been recorded in the sheet “Vote” of the MS Excel file “StatFa12_hw11.xlsx”, with 1 for “support” and 0 for “against”. With these data, construct a 95% confidence interval for the proportion of voters supporting this candidate.
- (10 points, 5 points each; modified from Problem 8.28 in the textbook) The Universal Music Group is the music industry leader worldwide in sales according to Nielsen Sound Scan. Suppose a researcher wants to determine what market share the company holds in the city of St. Louis by randomly selecting 1003 people who purchased a CD last month. Among all the money they spent, 27% were for products manufactured by the Universal Music Group.
 - Based on these data, construct a 99% confidence interval to estimate the Universal Music Group’s market share of the CD sales market in St. Louis.
 - Suppose that survey had been taken with 5000 people. Recompute the confidence interval and compare your result with the first one. How does the sample size affect the length of the confidence interval?
- (10 points, 5 points each; modified from Problem 8.35 in the textbook) For the sample data contained in the sheet “Variance” of the MS Excel file “StatFa12_hw11.xlsx”, construct the requested confidence interval. Assume the data come from normally distributed populations.
 - For the data in column A, construct the 99% confidence interval for the population variance.
 - For the data in column B, construct the 95% confidence interval for the population standard deviation.
- (10 points; modified from Problem 8.39 in the textbook) Suppose a random sample of 14 people that are between 30 to 39 years of age produced the household income shown in the sheet “Income” of the MS Excel file “StatFa12_hw11.xlsx”. Use these data to determine a 95% confidence interval for the population variance. Assume household income are normally distributed.
- (10 points; modified from Problem 8.44) A group of investors wants to open a restaurant. One thing they want to do is to estimate the average monthly electric bill. Before they sample some existing

similar restaurants, they have decided that they want to be 90% confident of their results and want the error of the interval estimate to be no more than \$100 (i.e., the interval size should be no larger than \$200). Suppose the population standard deviation is estimated to be \$500. Determine the sample size.

Note. Though not required for solving this problem, reading Section 8.5 in the textbook helps.

8. (10 points; modified from Problem 8.45) Suppose a production facility purchases a particular component part in large lots from a supplier. The production manager wants to estimate the proportion of defective parts received from this supplier. She wants to be within 0.02 of the true proportion of defected parts with a 90% confidence level. What is the *largest* sample size that she needs to take?

Hint 1. Note that the population proportion is unknown!

Hint 2. For a variable $x \in [0, 1]$, the quantity $x(1 - x)$ is maximized at $x = 0.5$.

Note. Though not required for solving this problem, reading Section 8.5 in the textbook helps.

9. (20 points) Consider the sheet “Sampling” of the MS Excel file “StatFa12.hw11.xlsx”. There you may find a set of population data whose size is 2000. In this problem, we will numerical verify that a confidence interval will really cover the parameter with a probability around the confidence level.

- (a) (2 points) Calculate the population mean and variance.
- (b) (3 points) Graphically show that the population is normal.
- (c) (5 points) Repeatedly sample 20 values from the population for 200 times¹. For each sample, compute a 95% confidence interval for the population mean by assuming that the population variance is *unknown*.
 - i. (2 points) Which sampling distribution should we use? Why?
 - ii. (3 points) For each sample, check whether the interval really covers the population mean. Among the 200 intervals, how many cover the population mean?
- (d) (5 points) Repeatedly sample 40 values from the population for 200 times. For each sample, compute a 90% confidence interval for the population mean by assuming that the population variance is *known*.
 - i. (2 points) Which sampling distribution should we use? Why?
 - ii. (3 points) For each sample, check whether the interval really covers the population mean. Among the 200 intervals, how many cover the population mean?
- (e) (5 points) Repeatedly sample 30 values from the population for 200 times. For each sample, compute a 95% confidence interval for the population variance.
 - i. (2 points) Which sampling distribution should we use? Why?
 - ii. (3 points) For each sample, check whether the interval really covers the population variance. Among the 200 intervals, how many cover the population variance?

Note. In answering Parts (c.ii), (d.ii), and (e.ii), you do not need to list all the samples and intervals you get. In other words, if you just want to get credits, you can even write down your answers without doing sampling. Nevertheless, going through the process definitely helps you understand the topic better.

¹It is certainly possible that you do not need to use the “Sampling” function in “Data Analysis” for 200 times.