BIA 654 B: Experimental Design II

**Ephraim Schoenbrun**

CWID: 104-20-702

Date of Submission: **February 7, 2017**

Assignment:

**Assignment #2**

# **Ethical Conduct**

|  |
| --- |
| The following statement is printed in the Stevens Graduate Catalog and applies to all students taking Stevens courses, on and off campus.  “Cheating during in-class tests or take-home examinations or homework is, of course, illegal and immoral. A Graduate Academic Evaluation Board exists to investigate academic improprieties, conduct hearings, and determine any necessary actions. The term ‘academic impropriety’ is meant to include, but is not limited to, cheating on homework, during in-class or take home examinations and plagiarism.”  Consequences of academic impropriety are severe, ranging from receiving an “F” in a course, to a warning from the Dean of the Graduate School, which becomes a part of the permanent student record, to expulsion.  *Reference: The Graduate Student Handbook, Academic Year 2003-2004 Stevens*  *Institute of Technology, page 10.*  Consistent with the above statements, all homework exercises, tests and exams that are designated as individual assignments MUST contain the following signed statement before they can be accepted for grading. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  I pledge on my honor that I have not given or received any unauthorized assistance on this assignment/examination. I further pledge that I have not copied any material from a book, article, the Internet or any other source except where I have expressly cited the source.  Signature: Date: February 7, 2017  Please note that assignments in this class may be submitted to [www.turnitin.com](http://www.turnitin.com/), a web-based anti-plagiarism system, for an evaluation of their originality. |

1. **Utility Company**
   1. A utility company serves 50,000 households. As part of a survey of customer attitudes, they take a simple random sample (meaning drawing at random without replacement) of 750 of these households. The average number of television sets in the sample households turns out to be 1.86 and the standard deviation is 0.80. If possible, find a 95% confidence interval for the average number of television sets in all 50,000 households. If this isn't possible, explain why not.

**Assumption: Population µ and σ unknown**

**Method of hypothesis testing: t-test (although as n approaches ∞, t 🡪 z)**

**Confidence Interval Estimate =**

**with 750 – 1 degrees of freedom and = 0.05: ≈ 1.96**

**95% C.I. = = 1.803 to 1.917**

* 1. Now, due to a budget constraint, the survey was conducted for only 18 households. The average number of television sets in the sample households turns out to be 1.86 and the standard deviation is 0.80. If possible, find a 95% confidence interval for the average number of television sets in all 50,000 households. If this isn't possible, explain why not.

**CANNOT RUN TEST SINCE NO ASSUMPTION OF NORMAL DISTRIBUTION**

**IF assumed normally distributed**

**Assumption: Population µ and σ unknown**

**Method of hypothesis testing: t-test (n < 30)**

**Confidence Interval Estimate =**

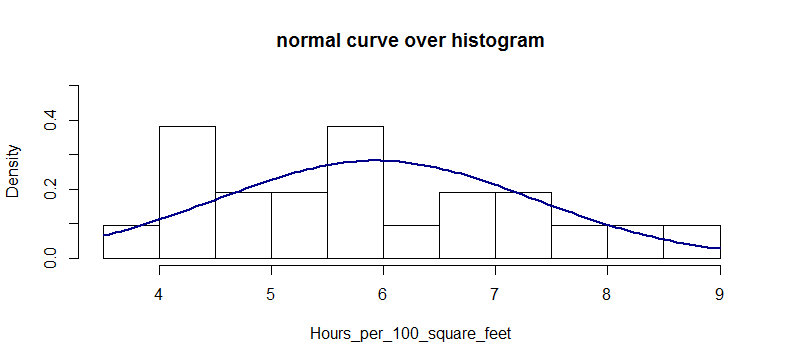
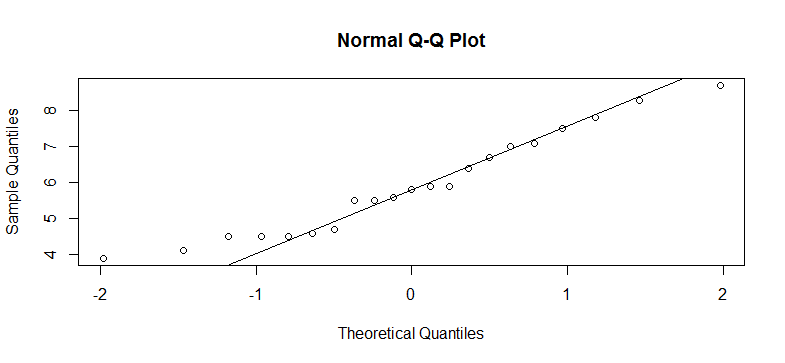
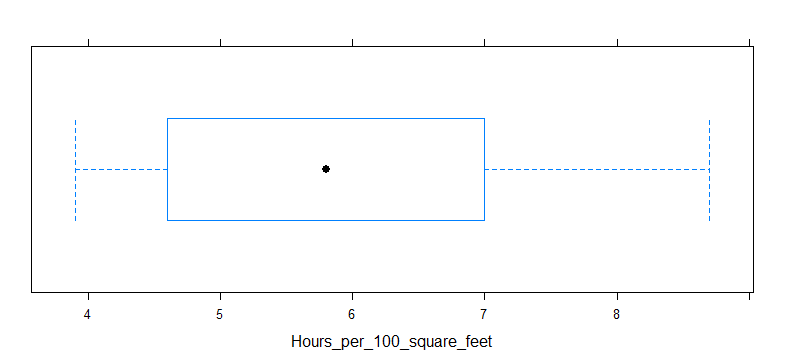
**with 18 – 1 degrees of freedom and = 0.05: 2.11**

**95% C.I. = = 1.449 to 2.271**

1. **Cleaning Service**
   1. A cleaning business operates in the city of New York and works for the companies that lease office space in the city. The business contracts to clean office space in increments of 100 square feet. The business determines its margins by determining how long it takes a crew to clean 100 square feet of office space, and bases its rates on this information. Because the company is relatively new, it has to estimate the time it takes to clean a 100 square feet of office space. The company estimates that it should take 5.7 hours to clean 100 square feet. The company starts its business with this expectation and works for a week straight, collecting data as it proceeds in order to be certain that it is neither over- nor under-charging its clients. The data collected by the company can be seen in the data file attached (see Canvas Assignments folder). After collecting this data, the company wants to determine if the time originally estimated to clean 100 square feet of office space was reasonable. Check this by computing a 95 percent confidence interval. (Hint: Notice the sample size is relatively small. So, one has to justify the assumption of normal distribution, by examining (a) normal quantile plot and (b) a Goodness-of-fit test, e.g., Shapiro-Wilk test. Perform these and present the result. Once this population normality assumption is met, then use a t-value with n - 1 degrees of freedom.)

**See visualizations and statistics below to justify normal distribution:**

|  |
| --- |
| Shapiro-Wilk Normality Test  data: Hours\_per\_100\_square\_feet  W = 0.94823, p-value = 0.3153 |
| Chi-Squared Test for Given Probabilities  data: Hours\_per\_100\_square\_feet  X-squared = 6.7407, df = 20, p-value = 0.9975 |



**Assumption: Population µ and σ unknown**

**Method of hypothesis testing: t-test (n < 30, justified as normally distributed)**

**Confidence Interval Estimate =**

**with 21 – 1 degrees of freedom and = 0.05: 2.086**

**95% C.I. = = 5.285 to 6.572**

**FAIL TO REJECT NULL HYPOTHESIS**

* 1. If the cleaning company from the previous question had a sample of 8 rather than a sample of 21 upon which to base its conclusions, what would be the boundaries of the 95 percent confidence interval for the estimate of the number of hours?

Assume that the sample mean and standard deviation are equal to those calculated above. Assume also the number of hours are distributed according to a normal distribution.

**Assumption: Population µ and σ unknown**

**Method of hypothesis testing: t-test (n < 30, assumed as normally distributed)**

**Confidence Interval Estimate =**

**with 8 – 1 degrees of freedom and = 0.05: 2.365**

**95% C.I. = = 4.747 to 7.110**

**FAIL TO REJECT NULL HYPOTHESIS**