BIA 654 B: Experimental Design II

**Ephraim Schoenbrun**

CWID: 104-20-702

Date of Submission: **February 14, 2017**

Assignment:

**Assignment #3**

# **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 14, 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. In an NYT/CBS poll, 56% of 2,000 randomly selected voters in New York City said they would vote for the incumbent in a certain two-person race.
   1. Calculate a 95% confidence interval for the population proportion π. (Check whether the necessary `assumption' is met.)

**.56 ± 1.96\*sqrt(.56\*(1-.56)/2000) = .56 ± 0.022**

**The confidence interval is in between 0.538 and 0.582. Since np > 5 and n(1-p) > 5, this calculation is valid.**

* 1. Carefully interpret the meaning of the confidence interval obtained in (a).

**We are 95% confident that the true percentage of people who would vote for the incumbent is in between 53.8% and 58.2% of the population. Meaning, 95% of the intervals that we take will contain the true proportion.**

* 1. What is the margin of error?

**0.022 or 2.2%**

* 1. Assume we had no prior knowledge about the true proportion pi. We want to construct a 95% confidence interval for π with margin of error 2%. How large a sample is needed? How does the sample size change if we want to be 99% confident?

**Assuming the pilot sample yields .56, the aforementioned p for the population π:**

**95%: (1.96)2\*.56(1-.56)/.022 = 2367**

**99%: (2.575)2\*.56(1-.56)/.022 = 4088**

1. Many companies are experimenting with “flex-time," allowing employees to choose their schedules within broad limits set by management. Among other things, flex-time is supposed to reduce absenteeism. One firm knows that in the past few years, employees have averaged 6.3 days off from work (apart from vacations). This year, the firm introduces flex-time. Management chooses a simple random sample of 100 employees to follow in detail, and at the end of the year, these employees average 5.5 days off from work, and the sample standard deviation (SD) is 2.9 days.
   1. Did absenteeism really go down, or is this just chance variation? Formulate the null and alternative hypotheses and carry out the testing.

**H0: Absenteeism did not go down**

**H1: Absenteeism went down (one tailed test)**

**ZSTAT = (5.5 – 6.5)/(2.9/sqrt(100)) = -2.76 standard deviations from 6.5**

**Which means that there is a .03% chance that the sample was an error**

**So we reject the null hypothesis**

* 1. Repeat the above for a sample average of 5.9 days and an SD of 2.9 days.

**ZSTAT = (5.9 – 6.5)/(2.9/sqrt(100)) = -1.38 standard deviations from 6.5**

**Which means that there is an 8.4% chance that the sample was an error. Given a significance level of 5%, the probability is above α, so we fail to reject the null hypothesis.**

1. Open up the data file Therm.dat (or Therm.xls) uploaded in Canvas. Carry out the hypothesis testing to check whether the temperature measurement difference > 0 at significance level 0.01. (Don't forget to check whether an `underlying assumption' holds or not, e.g., normal probability plot and Goodness-of-Fit Test for normality checking.)

**Assuming normal distribution: chi-square X-squared = 6.6 | df = 19 | p-value = 0.9961**

**H0: Temperature difference = 0**

**H1: Temperature difference ≠ 0 (two tailed)**

**The mean of the difference is 1.12. The t Test shows that the confidence interval for a 99% test statistic is between 0.721 and 1.519. Since the mean falls within that range, we fail to reject the null hypothesis.**

1. (True, False) To make a t-test with 5 measurements, use Student's t-distribution with 5 degrees of freedom.

**FALSE – Degrees of Freedom are the number of observations that are free to vary AFTER sample mean has been calculated. Since d.f. = n-1, the proper t-distribution would use 4 degrees of freedom.**