

Guide for Mini-Projects for STAT course (Advanced Programs)

Nguyen Chanh Tu (FAST-DUT)

A. Introduction:

- Project will involve taking *random samples* from a well-defined population and measuring several variables concerning to parameters (population means, proportions, or a different of two means from normal distributions...).

Before collecting the data, you should state your research problems or make conjectures as to the value of the population parameters.

-Using data in the excel file about 4 groups of 2021 freshmen of FAST as your collected samples for your research projects. You can add other variables of “unreal” information into the file if you like.

-The first part of the project is to use all method of descriptive statistics as in Chapter 1 to summary and visualize the given data.

-The second part of the project is using all methods of inferential statistics as estimations, hypothesis testing, regression/correlation to find answers for your problems.

You are free to make assumptions or choose your own topics from given samples in excel file, or you can use following suggestions:

+means of height/weight of DUT male/female students is more than 165/150 cm.

+the means of entering scores of freshman’s students of DUT is more than 21.

+the mean of self-study time are 2.5 hours per week.

+the different of means of heights of male and female students of freshman’s student of DUT is less than 10cm; +DUT freshman’s students participate more than 3 social networks

.....

B.-Final Report:

-This should be a typed report, written collaboratively by all team members and presentation.

-Deadlines for submitting report is one day before the last class via LMS. The presentation would be on the last class. Each member will have the same chance to speak about his/her work in the project.

-Your report should be written as if will be read by other student researchers. Make sure it includes at least:

I. Introduction to problems

-You should describe the population parameter of interest, an initial conjecture for its value (that makes sense in the context) and whether you suspect the actual value is higher or lower (of just different) than this conjectured value.

-Describe member names and the task of each member.

II. Data Collection Explain:

Describe the way you have sample and what you add on the provided data.

III. Analysis of Results

1.Descriptive Statistics. Use R to present and analyze your data.

-You will need to make choices as to which numerical and graphical summaries are most relevant.

-Fully describe your sample, sample size, and report the sample statistic and whether it supports your conjecture.

2. Inferential Statistics

Carrying the estimator, the confident intervals, the test of hypothesis, correlations, ...

- define the population and parameter in words
- state your conjectured value about the parameter and what it signifies.
- state whether you suspected (before you saw the data) whether you thought the actual value of the parameter was higher or lower than this conjectured value. If you had no prior direction in mind, then you will calculate a two-sided p-value.
- state what a type I and a type II error would represent in this setting.
- calculate probability to represent the p-value corresponding to the direction of your conjecture. Include an interpretation of what this p-value represents.
- use R to calculate P-value, confidence intervals, test of hypothesis... to describe the plausible values of your population parameter.
- state your conclusions in context.

IV. Conclusion

-Pay particular attention to whether the conditions were satisfied for you to generalize your sample to the larger population.

-References

Danang Oct, 2023