

**Advising the Queen:
Simulating Real World Significance Testing in about an Hour**

Ann Cannon and Jim Freeman
Cornell College

Abstract: Students doing significance testing homework are often confused by the sample data included in the statement of the problem (data yielding large p-value, for example). This laboratory exercise takes students through the process of determining the question to examine, experimental design and collection of data, sampling, analysis, and reporting results.

Students have trouble grasping significance tests. The complete process, from design to data collection to data analysis, is time consuming and requires different intellectual processes. In addition, a test of significance does not ask for a definitive answer which is very unsettling to students, and most homework problems in texts on significance testing tend to be too prescriptive in their statement. This is not surprising as the sample data is included with the statement of the problem. The confusion this causes is most apparent when the sample data yields a large p-value, which most students interpret as meaning they had their hypotheses set up wrong. This is less of a problem if the hypotheses are decided upon before any sample information is presented. This is very hard to do in a problem presented in a book. Simulating the actual real-life experience of doing significance testing is more effective.

A group oriented exercise called "Rectalinea" attempts to address these issues. The students act as the statistical advisor to the ruler of Rectalinea, Queen Quadra Lateral. She has three counties for which she must decide on what actions to take. Students provide a written report to the Queen advising the actions to take and why. Students must decide on the size and sampling method to use, then sample the residents of Rectalinea through a web interface, analyze the data and make written recommendations to the Queen. Students can complete the sampling through data analysis in around an hour and write up their report outside of the laboratory environment if they wish. Once all groups report their advice to the Queen a lively classroom discussion can occur as students see that different groups may come up with different conclusions. The instructor is also able to have students discuss how they designed the sampling protocol, bringing together many ideas from the course in one exercise.

This exercise has been used either to introduce students to the concept of significance testing or as a capstone type of experience. By manipulating the simulated data, the instructor has control over the results likely to be found by students. The current data being used at Cornell yields p-values which tend to be small for one question for the Queen, large for the second and moderate for the third. The web interface to the data is easy to implement on any web server with minimal programming expertise. The web interface slows down sampling, so there is a cost to students for taking a large sample.

Welcome to Rectalinea, a province of the world of Flatland, a two dimensional world. The inhabitants of flatland are planar figures like triangles and rectangles. Rectalinea is a region entirely inhabited by rectangles. Queen Quadra Lateral rules the three counties of Rectalinea, namely Alpha, Beta and Gamma. You are the chief advisor to Queen Quadra Lateral and have been ordered to give expert advice on three issues that she must decide.

Before describing the issues, you will need to know a unique aspect of Rectalinea. It is a communal land, where the well being of a county is determined by the mean area of all the inhabitants of that county. So, the greater the mean area of the inhabitants, the more prosperous the county.

Queen Quadra Lateral is facing three issues, one for each county, and needs a written report by the end of the class period.

Issue One: Any county whose mean area of its inhabitants is less than 68 is considered impoverished and will receive special government subsidies for as long as its mean area remains below 68. At the last census, 5 years ago, Alpha had a mean area of 63 and has been receiving subsidies ever since. Queen Quadra Lateral believes that these subsidies have succeeded in improving the conditions in Alpha to the point that the subsidies are no longer necessary. The next census is not for another 5 years. She wants to know today whether or not it is reasonable and fair to stop the subsidies now.

Issue Two: The greater the mean area, the greater the prestige of the county and the more political clout the county has. In fact, if a county has a mean area of 140 or greater, it is law that they receive special representation in the governing council for Rectalinea. Beta was the most prosperous county in the last census with a mean area of 140 thus achieving the level necessary for the special representative. Quadra Lateral is tired of dealing with the representative from Beta and hopes that Beta is not as prosperous as it once was. She wants to know if she can safely send the representative home.

Issue Three: Gamma is a bit of a hypochondriac county. At the last census, their mean area was 95. They do not like to change anything in their county and are always pestering Quadra Lateral for help in maintaining the status quo. She believes there is no evidence that the status quo has changed in Gamma. Does her belief appear to be correct?

So those are the issues on which Quadra Lateral is expecting a written report by the end of the period. She expects you to clearly argue how she should address each of these issues and to explain exactly how you obtained your conclusions.

Before you go, one other item. The next census is not for five more years and could not be accomplished in the time you have today, so you will have to sample. Be sure that you have discussed how you will sample, the size of your sample, and your sampling techniques before starting this project. You have been given a list of the social security numbers of all of the residents of each county. You can use this list to choose your sample, but then you must actually contact each resident to find out his or her size. Residents from county Alpha can be contacted using the address www.cornellcollege.edu/perl/stat/alpha.pl. The residents of the other two counties can be contacted through their pages by substituting beta or gamma for the alpha in the address given above.

Queen Quadra Lateral is waiting impatiently for your report. You know how rulers can get and the consequences to those who disappoint them.