Project Assignment for Statistics 301, Professor Wardrop, Spring Semester, 2009

You must perform one or two of the three projects described below and submit a written report for each project you select. **You may not submit both projects 1A and 1B.** Reports must be submitted on paper, not electronically.

You may work alone or as a member of a team of **three or fewer persons**. If you work as a team, only one report per project should be submitted by a team (remember to include everyone's name on the report!) and all members of the team will receive the same grade. Neither my expectations for nor grading of reports is influenced by the size of the team.

Please submit legible reports. For many of you, this will mean words typed, and "nonwords" (e.g. tables, figures, formulas, and computations) drawn by hand. Effective use of color and illustrations can greatly enhance a report.

The maximum score for each project is 18 points.

I have 288 students; thus, I could receive as many as 576 projects. Sometimes projects disappear. Some of those times, it is due to my negligence. In any event, you must keep a copy of your project in case the original is lost.

Restriction. This is extremely important. Read it. I want to encourage you to submit team projects. If all (1 or 2) of your projects are team projects, then you will receive an additional one-half point bonus on the project with the lower score.

You must notify me of the members of your project team at least one week before the project report is due. Failure to notify me of the team members will result in a loss of one-half point on the project and you will not be eligible for the bonus mentioned above.

Project 1A. Due before 5:01 pm on Wednesday, March 11. Perform a balanced completely randomized design on 100 trials with a dichotomous response to investigate some issue. Explain why the issue interests you and describe how you randomized. Present and summarize the data using techniques discussed in the course. Choose an alternative for Fisher's test and explain your choice. Calculate two P-values for the data and alternative: (1) Use the standard normal curve with the continuity correction; and (2) Use the website to obtain the exact P-value. Compare your two P-values and comment. Your report should include a summary of what you have learned from your study.

Project 1B. Due before 5:01 pm on Wednesday, March 11. Perform a balanced completely randomized design on 50 persons with a dichotomous response to investigate some issue. Indicate how you obtained your subjects and why these particular persons are of interest to you. Explain why the issue interests you and describe how you randomized. Present and summarize the data using techniques discussed in the course. Choose an alternative for Fisher's test and explain your choice. Calculate two P-values for the data and alternative: (1) Use the standard normal curve with the continuity correction; and (2) Use the website to obtain the exact P-value. Compare your two P-values and comment. Your report should include a summary of what you have learned from your study.

Project 2. Due before 5:01 pm on Tuesday, April 14. Perform 100 dichotomous trials under "fixed conditions" related to an activity of interest to you. Exercises 6–14 on pages 177–178 of the text present summaries of studies of this type that were performed by other students. Your report should describe the activity, why it interests you, and why you are interested in the fixed conditions you study.

Note that if a success is extremely difficult or extremely easy to obtain, 100 trials are not likely to reveal much of interest. You will obtain more interesting data if you select conditions that make the probability of success on a trial be a number between 0.25 and 0.75.

(a) Before you collect any data make a *point guess* of the value of p. (Point means single number. For example, if I were shooting free throws, before I collect data I would say, "My best guess is that $p = 0.58$.")

(b) **Use the data from your 100 trials to investigate the validity of Assumptions 2 and 3 of Bernoulli trials, using the techniques presented in lecture for Chapter 5.** (Note: I put this in bold-face type because, historically, a large proportion of students have neglected to do this when told with regular-face type!) Your report must contain the time-order results of your 100 trials so that I can check your analysis.

(c) On the assumption you have Bernoulli trials, find the 95 percent confidence interval for p. Is your point guess from part (a) in your CI? What does the CI tell you about the accuracy of your guess?

(d) On the assumption you have Bernoulli trials, find the point prediction and 95 percent prediction interval for the number of successes in 100 future trials.

(e) Perform the 100 additional trials and compare your results to your predictions from part (d).

Grading of Reports. A report will not receive full credit simply because it contains no errors. Reports will be graded on effort, creativity, report presentation (primarily clear writing), accuracy and appropriateness of analysis, and interpretation of results.

Very important, read this. Based on a problem that arose recently, I want to make the following very clear.

1. If you share data with another individual/team, you will receive a score of 0 on your project.
2. If you use data from a report submitted previously by a student in any of my classes, you will receive a score of 0 on your project.
3. If you fabricate your data, you will receive a score of 0 on your project.

It is ok to replicate an earlier project. For example, many students have studied shooting free throws in basketball; it is acceptable for you to do it too. But it is wrong to use data collected by another or to fabricate your free throw data.

Apart from issues of integrity and honesty, here is my point: I assign these projects so that you can get the experience of doing a study and so that you have the possibility of seeing that Statistics might be useful in your life.

By the way, it is my judgment that will determine whether any of the forbidden actions listed has occurred. In particular, I don't need to prove it in any legal or statistical sense, but I will, of course, be willing to discuss the matter with you if you think I am mistaken.

Model projects are available on my webpage. Note that model projects 1 and 2 are examples of project 1B; model projects 4 and 5 are examples of project 2; model project 3 can be safely ignored. Be careful in using these models. Do not simply rewrite one of them, 'plugging in' your numbers and treatments/trials. This is b/c the specific questions I ask you to answer in the descriptions above do not exactly match what I asked previous students.

Need some ideas? On my webpage, click onto

Research Papers: $>$ Papers on Statistical Education $>$ Small Student . . . , pp. 3–9

for ideas for project topics.