Project Assignment for Statistics 371, Professor Wardrop, Spring Semester, 2012

You may perform either or both of the projects described below and submit a written report for each project. **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 should be submitted by the 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 2.0 points extra credit.**

I have 240 students; thus, I could receive as many as 480 reports. Sometimes reports disappear. Some of those times, it is due to my negligence. In any event, you must keep a copy of your report in case the original is lost.

**Project 1. Due before 5:00 pm on Tuesday, April 10.** (I will accept Project 1 as late as 5:00 pm on Thursday, April 12, with a penalty of one-half point. I will not accept Project 1 after April 12.)

Perform 100 dichotomous trials. Your trials **might** be Bernoulli Trials and they **might not** be Bernoulli Trials. Part of your fun will be to investigate this issue (see below). We will discuss possible choices for trials in lecture. Your report should describe the activity that generated the trials and you should explain (briefly) why this activity interests you. I am not looking for originality, but I want to know why the activity is interesting to you. Here are two examples:

1. “I have always loved to play basketball, so I want to study my shooting.” (Free throws; 3-point shots; etc.)

2. “My friends and I play quarters to see who must drink Diet Snapple. My trials will be attempts to bounce a quarter into a juice glass.”

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 \( p = 0.53 \).”)

(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 Course Notes Chapter 6.** (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.

Let’s make this clear. To check assumption two you should create the first half/second half table, discuss what it reveals and do Fisher’s Test. To check assumption three you should create the memory table, discuss what it reveals and do Fisher’s Test.
(c) On the assumption you have Bernoulli trials—regardless of what you concluded above—use the standard normal curve to obtain the approximate two-sided 95 percent confidence interval for \( p \). Is your point guess from part (a) in your CI? What does the CI tell you about 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). Note: I do not want you to check the assumptions of BT’s with this second set of data.

Project 2. Due before 5:00 pm on Thursday, April 26. (Note: I will accept Project 2 as late as 5:00 pm on Thursday, May 3, with a penalty of one-half point. I will not accept Project 2 after May 3.)

Perform/observe 100 dichotomous trials under each of two conditions. Thus, you will perform/observe a total of 200 trials. The idea is to decide which condition is better at yielding the desired response. You must assign trials to conditions (treatments) by randomization. This, of course, limits your choice of conditions. Carefully, but briefly, explain how you randomized.

Before you collect data, conjecture on which condition will perform better. Make a guess as to how much better. For example, if your trials are shooting three point baskets and your conditions are location (front and left corner) you might conjecture that you will shoot better from the front and you might guess that the number of successes you achieve from the front will be eight larger than the number from the left corner.

After collecting your data, discuss your conjecture and your guess.

After collecting your data, perform Fisher’s test for the one-sided alternative that corresponds to your conjecture. (If your conjecture was that the conditions were equally good, use the two-sided alternative.) Comment on what you have learned from your data and your P-value.

Assuming that your data are the realizations of two independent sequences of Bernoulli Trials, calculate the 90% confidence interval for \( p_1 - p_2 \). Comment on what you learn from this confidence interval.

Grading of Reports. Here is a guide to how the projects will be graded. Basically, I look for errors. If an error is particularly egregious, I will count it as two or perhaps three errors. Egregious errors are rare and will occur only if you neglect to do some big part of the project or you totally mess up some big part of the project. Occasionally, I will erase some errors if your project is wonderful in some way. But don’t aim for wonderful; it is more productive to avoid errors!

The table below shows the relationship between the extra credit points you will receive and your total number of errors.

<table>
<thead>
<tr>
<th>Number of Errors:</th>
<th>0 or 1</th>
<th>2 or 3</th>
<th>4 or 5</th>
<th>6 or 7</th>
<th>8 or more</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extra Credit Points:</td>
<td>2.0</td>
<td>1.5</td>
<td>1.0</td>
<td>0.5</td>
<td>0.0</td>
</tr>
</tbody>
</table>

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.
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.** Model projects 4 and 5 are examples of your project 1 assignment. Model projects 1 and 2 are examples of your project 2 assignment in which each trial is asking a different person a question and the two conditions are two wordings of the question.

Be careful in using these models. Do not simply rewrite one of them, ‘plugging in’ your numbers and treatments/trials. This is because the specific questions I ask you to answer in the description 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 ....

See examples 1– 11 on pages 3–5 and examples 25–30 on pages 8 and 9 for ideas for project topics. Or talk to me about it.

Finally, it is fine if you study the same ‘topic’ in Projects 1 and 2. For example, Project 1 could study your free throw shooting with a men’s basketball. Project 2 could study shooting free throws with the conditions ‘men’s basketball’ and ‘women’s basketball’ to investigate whether you shoot better with the smaller ball. Of course, there is no need for the topics of the two projects to be related in any way.