

**ST 810A, SPRING 2005**  
**PREPARATION FOR STATISTICAL RESEARCH**  
**ASSIGNMENT 2, DUE TUESDAY, 3/15/05**

*Background:* Simulation studies are invaluable in Statistics. Most often, simulations are carried out when analytical derivation of the properties of statistical methods is intractable; instead, the properties are deduced empirically. A common situation is that in which exact, finite-sample results are unavailable; large-sample properties have been derived; and insight into the relevance of these large-sample properties in sample sizes likely to be encountered in practice is desired. Another is that where properties may be known under certain assumptions (e.g., normality), but behavior under departures from these assumptions is unknown and difficult to derive.

In this assignment, you will make the first steps involved in a project that will evolve over the rest of the semester. The overall project will consist of the following:

1. Identify a problem where properties of a statistical method (e.g., an estimator or competing estimators, or a testing procedure or procedures) need empirical study. For example, only large-sample results may be available, and you may wish to see how well these approximate finite-sample behavior. Alternatively, you may wish to evaluate properties of a method or methods under conditions where usual assumptions under which they are derived do not hold. It is up to you to identify a problem and method(s) to study. *This need not be “original research” studying cutting-edge techniques! You may consider widely-used, familiar techniques or a problem in an area in which you are interested.*
2. Write a proposal for the simulation study you intend to carry out. This document should introduce the problem you will consider at a level that someone with general training in statistics could understand; provide background and rationale for why the problem and associated methods need study; and state clearly the objectives of the study, including what is likely to be learned and how the results will advance understanding. This should be a self-contained document, so you should introduce a consistent notation in which you describe the problem and the specific estimators or tests you will study.

For the proposal, you do not have to spell out a *specific* design for the simulations that you will carry out, but you should include some information on the general type of studies you envision you will carry out and why these general choices are relevant. As with any scientific communication, citation to relevant references should be given, and a “references” section should be included. Also, the document should follow an organized format (to be determined by you).

3. Design a specific set of simulation studies that address the objectives you have identified in your proposal, write relevant computer programs and carry out these studies, and write a full paper, in the form of a journal article, reporting on the studies and the results. The paper should integrate the background and objectives from your proposal into the presentation in whatever way you see fit. It should also describe in detail the design and why it was chosen. The results should be clearly presented and discussed, and the paper should offer interpretation of the results for general practice and a summary of what was learned. This document should follow the format and conventions of a regular journal article, including an abstract and references.

Thus, the paper should be a self-contained document in the spirit of a journal article that takes the reader through introduction of the problem, rationale for studying it, what was done and why, and what conclusions can be drawn.

4. Develop a 15-minute oral presentation of the material in your paper in item 3. Your presentation should be accessible to a broad audience with general training in statistics (but perhaps not in the particular problem area) and should communicate the main ideas and messages. Oral presentations will take place during the final exam period for the class.

*For this assignment:* You should carry out tasks 1 and 2 above as follows:

- Identify a problem and the particular issue you would like to study. *Please send me an email giving a brief description of the problem you have chosen and the objective of your proposed study. If you are having trouble coming up with a problem, please make an appointment to see me (by email) as soon as possible.*
- *Once you have received approval from me regarding your problem, using L<sup>A</sup>T<sub>E</sub>X, prepare a 5–10 page proposal document (using 12 point type with no more than 25 lines/page and 1-inch margins). When writing your proposal, it will help you later if you keep in mind that eventually you will be integrating the content into the full paper in 3 above. Keep in mind also that your final paper will be required to be no longer than 20 pages, including title, abstract, body, and references.*

*You will find it helpful to consult the document “Simulation Study in Statistics” by Li, Boos, and Gumpertz, available on the class web page. You may also find it useful to browse recent issues of statistical journals to see how authors motivate and describe simulation studies and display and interpret results. Additionally, you will find on the web page an example of a problem that is in the spirit of this project and general remarks on questions that could be investigated through simulation studies.*

**IMPORTANT:** The problem you choose should *not* be something that you have already done! So, for example, if you have already taken the written prelim exam, you should not use the same topic here; you should identify a completely new problem. Similarly, if you are carrying out simulations as part of your dissertation research already, you should not use this as the basis for your project.