Statistics 214: Probability and Statistical Models
Final Project

Overview
In this final project you will explore in detail one of the modeling or computation topics discussed in the course, implementing and applying it in the context of a specific application or methodological study. This will help you gain a deeper understanding of the topic of your choice, as well as to gain experience in translating these ideas into practice.

You should begin by choosing an application domain, data set, or methodological question of interest to you. This may be a topic related to your own research, or simply a topic about which you are curious. You are encouraged to choose problems relevant to your own research interests and activities, although the project should be self-contained and the work done for Stat 214 clearly defined. Please see me with any questions.

The project consists of two parts: (1) a 1-page project proposal on which I will provide comments to help you identify and scope an appropriate project for the course; and (2) a final paper, written in the style of a journal or conference paper.

Project Proposal (due 11-9-09)
Prior to beginning work on your project you must provide me with a brief description of your proposed project for approval.

1. Find three (3) relevant papers which support the importance of your problem, discuss previous work on modeling/analysis in the area, and cover technical aspects of the methods from class you will apply. If there are more than three, select the three most relevant to include in the proposal.

2. Prepare a brief (≤1pg) description of your proposed project. Explicitly address each of the following:
   (a) Why the problem is of interest
   (b) What work has been done in this area previously
   (c) Goals of your project - what you hope to learn or demonstrate
   (d) The models or computational methods you will implement. Give models or likelihoods explicitly.
   (e) Datasets you will use
   (f) How will you evaluate success.
   (g) Citations for the three papers you have chosen

I will return your proposal with comments and suggestions by the following class. I will also provide you with one additional paper reference for you to read and incorporate into your final write-up.
Final Paper (due 12-10-09)

The final write-up should be in the style of a journal or conference paper. Please observe carefully the following specifications:

1. Formatting: Single space, ≥11pt font, ≥1-inch margins. LaTeX article defaults are fine.

2. Length: 7 page maximum, including references and figures.

3. Your paper should contain the following sections:
   
   (a) Abstract
   
   (b) Introduction - Outline the general problem and its importance, providing any required domain-specific background. Describe the origin of any datasets used.
   
   (c) Previous work - Summarize previous approaches to this problem, their results and limitations, motivating your approach.
   
   (d) Methods - Describe the models and computational methods used in your project, emphasizing the details necessary for adapting to your problem, and any novel aspects of your approach.
   
   (e) Results - Report the results of your analysis, providing appropriate tables/plots/figures. Indicate areas where your results agree or disagree with previous work.
   
   (f) Conclusions - Interpret your results. What have you learned and contributed. Suggest topics for further work.
   
   (g) References - Bibliography of papers cited. Should contain the 4 papers from your proposal; no more than 8 papers total.

4. Sections (b), (c), and (d) should cite the appropriate references from your proposal, and any additional relevant references you have found (up to a total of 8).

5. Do not submit code or software output in bulk - extract relevant materials and include them as tables or figures in the body of the paper. Keep within the page limit.

6. Make sure to spell-check and grammar-check your paper before submitting.

7. Please see the following university resources on proper methods for citation of sources:
   
   http://www.lib.duke.edu/libguide/citing.htm
   
   http://www.lib.duke.edu/libguide/plagiarism.htm

Please talk to me about any questions or clarifications.

Note: No late projects will be accepted, no exceptions. No Incomplete grades will be given without a Dean’s excuse.

Example Projects

Past projects have included the application of HMMs and stochastic volatility models to modeling of time series datasets in engineering and finance, the development and MCMC fitting of hierarchical Bayesian models for spatial applications in environmental science, variable selection methods for predictive modeling in bioinformatics, comparative studies of the efficiency of different MCMC algorithms, and many others. These example should by no means limit your creativity - a wide range of other topics and approaches are possible. There is no requirement that you utilize any particular model or technique, rather you should choose those from class most appropriate to your problem.