STATS 300A Theory of Statistics
Stanford University, Fall 2015

Problem Set 8

Due: Thursday, December 3, 2015 (**Except Problem 1**)

Instructions:

- You may appeal to any result proved in class or proved in the course textbooks.
- Any request to "find" requires proof that all requested properties are satisfied.

Problem 1 (Modern-Day Decision Problems)**.** This problem, due on **Monday, November 30, 2015**, will expose you to modern statistical literature and allow you to apply the tools learned in class to a decision problem of current interest. You must work on this problem **alone** and select a unique paper to analyze. Each student will present his or her findings in class during the final week of classes. Please read all prompts before beginning this problem.

1. Find a research paper published in 2013, 2014, or 2015 that describes a point estimation task with loss function L and that proposes or makes use of an estimator δ for that task. Single-parameter estimation tasks, simultaneous estimation tasks, and prediction problems, like regression or classification, are all valid choices of estimation problems.

   You should select your paper based on interest and on your ability to answer the questions to follow. Google Scholar and the following journals and conference proceedings may be good starting points for your search (but your paper need not come from these sources):

   - Annals of Statistics, Annals of Applied Statistics,
     Journal of the American Statistical Association,
     Journals of the Royal Statistical Society
   - Advances in Neural Information Processing Systems,
     International Conference on Machine Learning,
     Conference on Artificial Intelligence and Statistics.

   When you have chosen your paper, add your name and paper selection to the following spreadsheet: `http://bit.ly/1fR70q3`

   You **may not** select a paper that has already been chosen by another.

2. Formulate the estimation problem tackled in the paper as a decision theoretic problem. Your formulation should include the data-generating model $\mathcal{P}$, the estimand $g(\theta)$, the decision space $\mathcal{D}$, and the loss function L.

3. Answer at least three of the following questions concerning the estimator δ:

(a) Is the estimator admissible?

(b) Is the estimator UMRU?

(c) Is the estimator minimum risk location equivariant? (Do not attempt to answer this question if your decision problem is not location invariant.)

(d) Is the estimator minimum risk location-scale equivariant? (Do not attempt to answer this question if your decision problem is not location-scale invariant.)

(e) Is the estimator MRE with respect to another group of transformations (see TPE 3.2)? (Do not attempt to answer this question if your decision problem is not invariant under this group.)

(f) Is the estimator Bayes under some prior distribution Λ? If so, under which Λ?

(g) Is the estimator minimax?

4. Perform one of the following tasks:

   (a) Derive an estimator δ' that satisfies an optimality criterion not satisfied by δ.

   (b) Derive an estimator δ' that has strictly better performance than δ under an applicable optimality criterion from part 3 (for example, find a better unbiased estimator or a dominating estimator).

   (c) For one criterion from part 3, show that no estimator exists satisfying the optimality criterion. (Do not choose an MRE criterion if your decision problem is not invariant.)

5. Which estimator, δ or δ', would you prefer to use for this task? Why?

6. In addition to your normal problem write-up, prepare no more than 2 presentation slides (e.g., in PowerPoint, Keynote, or LaTeX beamer) containing

   - Your name
   - The title and authors of your selected paper
   - Brief motivation: why are the authors considering this estimation problem?
   - The decision problem formulation of the estimation task, including $\mathcal{P}$, $g(\theta)$, $\mathcal{D}$, and L
   - The estimator δ discussed in your paper along with the optimality criteria from part 3 that it satisfies
   - Your alternative estimator δ' or the statement that no optimal δ' exists under your criterion from part 4
   - Briefly, why you prefer δ or δ'

   Save your slide(s) as a **PDF file**, include **your Stanford email handle** in the file name (e.g., lmackey.pdf), and email it with the **subject line "300A Slides"** to `stats300a-aut1516-staff@lists` by **11:59PM on Monday, November 30**. Any proofs or derivations of results should be submitted with the remainder of the problem set in class on Thursday, December 3. You'll find example slides on the course website.

Each student will have 3 minutes to present his or her slide(s) in class. If the first letter in your last name is the range A - L, you will present on Tuesday, December 1. If the first letter in your last name is the range M - Z, you will present on Thursday, December 3.

Problem 2 (TSH 4.20)**.**

Problem 3 (TSH 4.25)**.**

Problem 4 (TSH 6.10, part (i))**.** You may find the discussion in TSH Section 6.4 relevant. You may assume the conclusion of Thm. 6.5.3 part (iii).

Problem 5 (TSH 6.18)**.** You may assume the conclusion of Thm. 6.5.3 part (iii).