

MATH 598/782 – PROJECT 5

Please submit your solutions by 11.59 pm (EDT) on Monday 29th April by uploading a single pdf to myCourses.

The paper

Joseph D. Y. Kang. Joseph L. Schafer. **"Demystifying Double Robustness: A Comparison of Alternative Strategies for Estimating a Population Mean from Incomplete Data (with Discussion)."** *Statistical Science*. 22 (4) 523 - 539, November 2007. <https://doi.org/10.1214/07-STS227>

<https://www.math.mcgill.ca/dstephens/598-Causal-2024/Exercises/KangSchafer2007.pdf>

is a famous paper discussing issues of doubly robust estimation. The paper is couched in the language of *missing data* rather than causal inference, however, all of the methods described and implemented have exact parallels in the causal inference literature¹.

This final (term) project involves writing a summary of the issues addressed in the paper, with the following guide points:

- (a) Summarize the contribution of the main paper, describing the key issues raised including the model set up, estimation methods and simulation studies. 6 Marks
- (b) Summarize the main points raised in the Discussion contributions, and the Rejoinder made by the authors. 6 Marks
- (c) Reproduce (as far possible) one of the simulation studies mentioned in the paper, and comment on what it is attempting to illustrate. 6 Marks
- (d) Give your analysis and/or critique of the contribution of the paper. 2 Marks

Your submitted project should comprise at least 4 pdf pages, and no more than 12 pdf pages.

¹Many authors, including DB Rubin, have contended that causal inference is essentially a missing data problem