Statistical Science
2020, Vol. 35, No. 3, 361–366
https://doi.org/10.1214/19-STS740
Main article: https://doi.org/10.1214/19-STS699
© Institute of Mathematical Statistics, 2020

Comment: Matching Methods for Observational Studies Derived from Large Administrative Databases

Mark M. Fredrickson, Josh Errickson and Ben B. Hansen

1. INTRODUCTION

In the era of big data, finding a comparable control group for a set of treated units provides new opportunities and challenges. When controls vastly outnumber treated subjects, there will likely be many good potential matches for each treated subject. On the other hand, with larger data sets, increased computation time prevents applying existing methods to find the best possible match. Yu et al. propose a fast caliper solution that restricts the possible controls for each treated subject, making matching with large databases tractable. Their results on determining the narrowest caliper that is compatible with pair matching (without replacement) will be of great practical use.

We take issue with the labeling of this caliper as "optimal." The label is accurate in a certain sense—it does minimize an objective of caliper width, subject to the constraint that pair matching remain feasible while no treatment group member is discarded—but these are quite different objectives and constraints from those otherwise targeted in the course of optimal matching. The meaning of "optimal" in "optimal matching" is already obscure to many, as Yu and Rosenbaum have themselves acknowledged (Yu and Rosenbaum, 2019). Adding a new and distinct connotation seems a step in the wrong direction.

It happens that Yu et al.'s optimal caliper can have the surprising result of forcing matches to be *suboptimal*, at least for the matching problem's original objective. We demonstrate this phenomenon in a small stylized example (Section 2). Full matching is less affected; also, the narrowest caliper that is compatible with full matching is simple to describe and quick to calculate (Section 3). In another large surgical outcomes study, a form of full matching with restrictions is shown to generate matches faster and with better optimality properties, while still maintaining a structure similar to pairs (Section 4). These critiques of pair matching notwithstanding, we expect caliper width to continue to be a leading determinant of matching speed, even as optimal matching in statistics assimilates algorithmic developments from related fields (Section 5).

Mark M. Fredrickson is Postdoctoral Research Fellow, Department of Statistics, University of Michigan, 311 West Hall, 1085 South University Ave, Ann Arbor, MI 48109 (e-mail: mfredric@umich.edu). Josh Errickson is Statistician Senior, Consulting for Statistics, Computing and Analytics Research (CSCAR), University of Michigan 3550 Rackham Bldg, Ann Arbor MI 48109 (e-mail: jerrick@umich.edu). Ben B. Hansen is Associate Professor, Department of Statistics, University of Michigan, 323 West Hall, 1085 South University Ave, Ann Arbor, MI 48109 (e-mail: bbh@umich.edu).

2. OPTIMALITY AND THE EYE OF THE BEHOLDER

For a set of treated units $\mathcal{T}$ and a set of control units $\mathcal{C}$, the match $\boldsymbol{m} = (m_{ij} : (i, j) \in \mathcal{T} \times \mathcal{C})^T$ has $m_{ij} = 1$ if treated unit i is matched to control unit j, 0 otherwise. The total distance of $\boldsymbol{m}$ is given by

$$f(\boldsymbol{m}) = \sum_{i \in \mathcal{T}} \sum_{j \in \mathcal{C}} m_{ij} d_{ij}, \tag{1}$$

where d_{ij} is the distance between i and j. Yu et al. contribute to a body of work that uses careful application of network flow algorithms or advances in integer programming to minimize f (Rosenbaum, 1989, Hansen and Klopfer, 2006, Yang et al., 2012, Zubizarreta, 2012, Pimentel et al., 2015, Pimentel, Yoon and Keele, 2015, Rosenbaum, 2017, Pimentel et al., 2018). Yu et al. focus on pair matching, minimization of $f(\cdot)$ over

$$\left\{ \boldsymbol{m} \in \mathcal{T} \times \mathcal{C} : \sum_{j \in \mathcal{C}} m_{ij} = 1, \text{ all } i \in \mathcal{T}; \right. \\ \left. \sum_{i \in \mathcal{T}} m_{ij} \leq 1, \text{ all } j \in \mathcal{C} \right\}. \tag{2}$$

Matched sets corresponding to such $\boldsymbol{m}$ maintain a strict $1:1$ ratio of treated to control subjects. In contrast, *full matching* minimizes over strictly broader classes of $\boldsymbol{m}$, permitting both $i \in \mathcal{T}$ with $\sum_{j \in \mathcal{C}} m_{ij} > 1$ and $j \in \mathcal{C}$ with $\sum_{i \in \mathcal{T}} m_{ij} > 1$, while requiring that: for each $i \in \mathcal{T}$, $\sum_{j \in \mathcal{C}} m_{ij} \geq 1$; if $\sum_{j \in \mathcal{C}} m_{i'j} > 1$ then $\sum_{i \in \mathcal{T}} m_{ij'} = 1$ for each $j' \in \mathcal{C}$ s.t. $m_{i'j'} = 1$; and similarly if $\sum_{i \in \mathcal{T}} m_{ij'} > 1$ then $\sum_{j \in \mathcal{C}} m_{i'j} = 1$ for each $i' \in \mathcal{T}$ such that $m_{i'j'} = 1$ (Rosenbaum, 1991). That is, both many-one and one-many matched sets are permitted, as are $1:1$ matched pairs; however, many-many configurations are excluded. Since optimal pair matching minimizes (1) over (2) and optimal full matching minimizes the same objective over a strictly broader domain, it is clear that for objective (1),