

STATISTICS COLLOQUIUM

MONDAY, FEBRUARY 4, 2013

TALK: 4:15 PM — SCIENCE CENTER RM. 705

RECEPTION: 3:50 PM — SCIENCE CENTER, 7TH FLOOR

“Using Mixed Integer Programming for Matching in Observational Studies: Effect of the 2010 Chilean Earthquake on Posttraumatic Stress”

**Jose Zubizarreta
Department of Statistics
University of Pennsylvania**

ABSTRACT

Matching is a widely used method of adjustment for observed covariates in observational studies. With matching, one attempts to replicate a randomized experiment by finding matched groups that look alike in terms of their observed covariates, as if they were randomly assigned to treatment. However, most matching methods involve a considerable amount of guesswork because they do not target covariate balance directly. In a recent paper (Zubizarreta 2012), I describe a new matching method based on mixed integer programming that overcomes this issue and targets covariate balance directly. By either optimizing or constraining several measures of covariate imbalance simultaneously, this new method can directly balance univariate moments (such as means, variances, and skewness), multivariate moments (such as correlations), and statistics (such as the Kolmogorov-Smirnov statistic). Furthermore, while balancing several of these measures, it makes possible to match with fine balance for more than one nominal covariate, whereas most matching algorithms can finely balance only a single nominal covariate. By virtue of optimality, matching based on mixed integer programming also tells whether certain forms of covariate balance are feasible with the data at hand, and therefore whether the data supports certain causal comparisons. In this talk I will go over the basics of matching, and illustrate this new method in an observational study about the effect of the 2010 Chilean earthquake on posttraumatic stress (Zubizarreta et al 2013). A new R package called *mipmatch* implements this method with a wide range of applications in the medical and social sciences.

- Zubizarreta, J. R., (2012), “Using Mixed Integer Programming for Matching in an Observational Study of Acute Kidney Injury after Surgery,” *Journal of the American Statistical Association*, **107**, 1360–1371.
- Zubizarreta, J. R., Cerda, M., and Rosenbaum, P. R. (2013), “Effect of the 2010 Chilean Earthquake on Posttraumatic Stress: Designing an Observational Study to be Less Sensitive to Unmeasured Biases,” *Epidemiology*, **24**, 79–87.