Statistica Sinica **14**(2004), 945-947

SYSTEMATIZING THE EVALUATION OF PARTIALLY CONTROLLED STUDIES USING PRINCIPAL STRATIFICTAION: FROM THEORY TO PRACTICE

Constantine E. Frangakis and Ravi Varadhan

Johns Hopkins University

Abstract: We often wish to evaluate treatments and other factors in studies where only some of those factors are directly controlled by the investigators. The framework of "principal stratification" has been proposed recently for evaluating partially controlled studies under such broader settings (Frangakis and Rubin (2002)). We have developed a software package, "PSpack", with appropriate documentation containing examples for implementing principal stratification. We hope that this helps bridge theory with practice, and that it systematizes the use of principal stratification in evaluating partially controlled studies. In this article we provide an introduction to using the software we have developed, which is available with documentation at the web site: http://biosun01.biostat.jhsph.edu/∼cfrangak/papers/ps.html.

Key words and phrases: Causal inference, needle exchange, principal stratification, software.

Background to Problem and Software

In problems that involve the evaluation of treatments and other factors, usually investigators can have direct control of only some of those factors. For some of these partially controlled studies, the more standard approach of instrumental variables (IV) may be used in some limited settings characterized by one particular set of assumptions. See, for example, Card (1993) and McClelan, McNeil and Newhouse (1994) who use proximity of individuals to colleges and hospitals, respectively, as an IV, to evaluate, respectively, effects of education on income, and effects of health care on health outcomes. Most partially controlled studies, however, have broader settings, including more than one uncontrolled factors, or other more complex data structures. In such cases, the standard IV is not adequate to evaluate treatment effects. The framework of "principal stratification" has been proposed recently for evaluating partially controlled studies under such broader settings (Frangakis and Rubin (2002)).

An example of such studies arises in the evaluation of the impact that needle exchange programs (NEP) have in reducing HIV transmission among injection

drug users. Such studies can control location of the NEP sites, and hence proximity of the sites to drug users. Moreover, proximity can affect both who actually exchanges at the NEP and who gives blood measurements to monitor HIV status, and this can be used to evaluate the effect of exchange on HIV transmission. However, such studies do not directly control either who exchanges or who gives blood measurements for HIV monitoring, which, therefore, are factors that are both only partially controlled through proximity. Frangakis, Brookmeyer, Varadhan, Safaeian, Vlahov and Strathdee (2004) show that standard IV is not generally appropriate to evaluate the effect of exchange on HIV and they describe how principal stratification can be used to better evaluate this effect.

The general idea in principal stratification with respect to a partially controlled (post-treatment) variable is the explicit definition of the strata (called principal strata) that are a cross-classification of subjects defined by the joint potential values of that post-treatment variable. The first key property of principal strata is that they are not affected by the controlled treatment, and hence can be used, in principle, just as any pre-treatment covariate, such as gender. For example, in the NEP studies, a principal stratum of exposure for a participant is defined as the entire vector of the values that indicate if the participant would exchange (or not) at the NEP for every potential distance of the NEP from his residence. The second key property of principal stratification is that comparisons of potential outcomes at different values of the controlled variable *conditionally* on principal strata are well defined causal effects (Frangakis and Rubin (2002)). These effects, called principal effects, can be used to evaluate the effect that the controlled treatment has on the outcome and that is attributable to the uncontrolled factor. For the NEP example, a principal effect of particular interest is the effect that proximity to the NEP has on HIV transmission, but restricted to the (partly unobserved) principal stratum of subjects for whom proximity would affect their needle exchange behavior at the NEP. This effect, under conditions, describes, better than standard estimands, the effect that proximity has on HIV transmission and that is attributable to actual exchange at the NEP (Frangakis et al. (2004)). In other recent work using principal stratification, Barnard, Frangakis, Hill and Rubin (2003) estimate effects of school vouchers in student performance; Zhang and Rubin (2003) show how to address censoring of outcomes by death; and Little, Lin and Long (2003) show how to evaluate treatments in choice based experimental designs.

An important, more general advantage of using principal stratification is that any assumptions it makes are explicit and so, in principle, it is more accessible to scientists. On the other hand, to address the special latent structure of the principal strata, actual estimation of models with principal stratification requires considerable work from any individual user, but of which a large part is common across problems and can be therefore be avoided by appropriate programming.

To systematize the use of principal stratification, we have developed a software package, "PSpack", with appropriate documentation including examples. "PSpack" runs in the "R" environment (R Development Core Team (2004)), which allows the user to interact and modify functions of the software; for information on how to obtain and install "R", check the website: www.r-project.org. PSpack is currently capable of handling binary outcomes and multiple levels of the controlled factor (e.g., distance). It is to be extended to handle ordinal outcomes with more than two levels. In parallel to PSpack's documentation, the user of the software is encouraged to read the methodology that justifies these procedures (Frangakis et al. (2004)). The documentation and software are available from the web site: http://biosun01.biostat.jhsph.edu/∼cfrangak/papers/ps.html.

References

Angrist, J. D., Imbens, G. W. and Rubin, D. B. (1996). Identification of causal effects using instrumental variables (with discussion). *J. Amer. Statist. Assoc.* **91**, 444-472.

Barnard, J., Frangakis, C. E., Hill, L. and Rubin, D. B. (2003). Principal stratification approach in broken randomized experiments: A case study of School Choice vouchers in NY City (with discussion). *J. Amer. Statist. Assoc.* **98**, 299-323.

Card, D. (1993). Using geographic variation in college proximity to estimate the return to schooling. National Bureau of Economic Research, Paper No. 4483.

Frangakis, C. E. and Rubin, D. B. (2002). Principal stratification in causal inference. *Biometrics* **58**, 20-29.

Frangakis, C. E., Brookmeyer, R. S., Varadhan, R., Safaeian, M., Vlahov, D. and Strathdee, S. A. (2004). Methodology for evaluating a partially controlled longitudinal treatment using principal stratification, with application to a Needle Exchange Program. *J. Amer. Statist. Assoc.* **98**, 239-249.

Little, R. J., Lin, X. and Long, Q. (2003). Causal inference in choice based intervention studies. University of Michigan Department of Biostatistics Research Report.

McClellan, M., McNeil, B. J. and Newhouse, J. P. (1994). Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? Analysis using instrumental variables. *J. Amer. Medical Assoc.* **272**, 859-866.

R Development Core Team (2004). *R : A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org.

Zhang, J. and Rubin, D. B. (2003). Censoring by death via principal stratification. Harvard University Department of Statistics Research Report.

Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimare, Maryland 21205-2179, U.S.A.

E-mail: cfrangak@jhsph.edu

Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimare, Maryland 21205-2179, U.S.A.

E-mail: rvaradha@jhsph.edu

(Received January 2002; accepted January 2003)