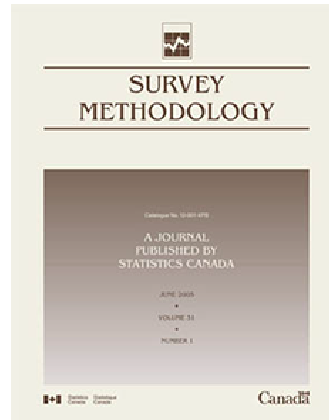
Catalogue no. 12-001-X
ISSN 1492-0921

Survey Methodology

Comments on "Exchangeability assumption in propensity-score based adjustment methods for population mean estimation using non-probability samples":

Causal inference, non-probability sample, and finite population



by Takumi Saegusa

Release date: June 25, 2024

How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website, www.statcan.gc.ca.

You can also contact us by

Email at infostats@statcan.gc.ca

Telephone, from Monday to Friday, 8:30 a.m. to 4:30 p.m., at the following numbers:

• Statistical Information Service	1-800-263-1136
• National telecommunications device for the hearing impaired	1-800-363-7629
• Fax line	1-514-283-9350

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, Statistics Canada has developed standards of service that its employees observe. To obtain a copy of these service standards, please contact Statistics Canada toll-free at 1-800-263-1136. The service standards are also published on www.statcan.gc.ca under "Contact us" > "Standards of service to the public."

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued co-operation and goodwill.

Published by authority of the Minister responsible for Statistics Canada

An HTML version is also available.

Cette publication est aussi disponible en français.

Survey Methodology, June 2024
Vol. 50, No. 1, pp. 65-69
Statistics Canada, Catalogue No. 12-001-X

Comments on “Exchangeability assumption in propensity-score based adjustment methods for population mean estimation using non-probability samples”:

Causal inference, non-probability sample, and finite population

Takumi Saegusa[1]

Abstract

In some of non-probability sample literature, the conditional exchangeability assumption is considered to be necessary for valid statistical inference. This assumption is rooted in causal inference though its potential outcome framework differs greatly from that of non-probability samples. We describe similarities and differences of two frameworks and discuss issues to consider when adopting the conditional exchangeability assumption in non-probability sample setups. We also discuss the role of finite population inference in different approaches of propensity scores and outcome regression modeling to non-probability samples.

Key Words: Causal inference; Finite population; Non-probability sample.

1. Introduction

I congratulate Professor Yan Li on another important addition to her active research on non-probability samples. In her paper, Professor Li classified existing research on non-probability samples into (1) the propensity score weighting methods and (2) the propensity score matching methods, and discussed that the conditional exchangeability (CE) assumption is required for the former. After reviewing existing methods in view of the CE assumption, Professor Li proposed the novel adaptive balancing score to ensure that the CE assumption holds. As the crystallization of accumulating literature on non-probability samples and causal inference, her paper demands a fair amount of background knowledge in order to understand complex concepts. The focus of our discussion here is to examine basic concepts and foundational issues which Professor Li’s sophisticated presentation touched only lightly.

This discussion is organized as follows. In Section 2, we review the conditional exchangeability assumption in causal inference. We describe differences of probabilistic frameworks in causal inference and non-probability samples, and discuss issues to consider when adopting the conditional exchangeability assumption in non-probability samples. In Section 3, we describe two major approaches in missing data problems including causal inference. Then we discuss issues of the role of finite population inference arising from the conditional exchangeability assumption in different approaches.

1. Takumi Saegusa, Department of Mathematics, University of Maryland, College Park, Maryland 20742, United States of America. E-mail: tsaegusa@umd.edu.

2. Causal inference

First, we discuss the relationship between the CE assumption and causal inference. In the paper, the CE assumption is formulated as the equality

$$E[y \mid b(x), C] = E[y \mid b(x), U] \tag{2.1}$$

where $b(x)$ is a function of covariates x referred as a balancing score, U is a finite population, and $C \subset U$ is a non-probability sample. Though simply defined, the criterion of its choice in the paper indicates that the balancing score seems to be implicitly defined to satisfy the CE assumption. Moreover, it is stated as a fact without much discussion that any quantity (including the propensity score) finer than the propensity score satisfies the CE assumption as a balancing score. An important literature that helps understand these concepts is the one coauthored by Professor Li (Wang, Graubard, Katki and Li, 2022), which is, to the best of our knowledge, the first paper that explicitly introduced balancing scores and conditional exchangeability in causal inference to the non-probability sample literature. In Wang, Graubard, Katki and Li (2022), however, these concepts were directly borrowed from the work of Rosenbaum and Rubin (1983) on causal inference, and results on propensity scores were claimed to hold in the non-probability setting without formal discussion. Because the definitions of the CE assumption and balancing score in the paper are different from those in Rosenbaum and Rubin (1983), and because the counterfactual framework of Rosenbaum and Rubin (1983) is fairly different from the setting of non-probability samples, it is worthwhile to pay a close attention to similarities and differences between causal inference and non-probability samples.

To this end, we first briefly summarize Rosenbaum and Rubin (1983) where variables of interest are potential outcomes $(Y(0), Y(1))$, covariates X, and treatment assignment $Z \in \{0,1\}$. The balancing score $b(x)$ in Rosenbaum and Rubin (1983) was defined as the function of covariates $X = x$ that satisfies the conditional independence between X and treatment assignment Z given $b(X)$ (i.e., $X \perp Z \mid b(X)$). It was shown that the propensity score into treatment is a balancing score, and that any function of x that can get mapped into the propensity score is also a balancing score. As the definition suggests, there is no requirement on the relationship between potential outcomes and covariates. The assumption that connects these variables is the conditional exchangeability with respect to covariates (or strong ignorability of Rosenbaum and Rubin (1983)), defined differently as the conditional independence between the potential outcomes and treatment assignment given covariates (i.e., $(Y(0), Y(1)) \perp Z \mid X$). The main result is that conditional exchangeability with respect to covariate X implies conditional exchangeability with respect to a balancing score $b(X)$. In other words, starting from the key conditional exchangeability assumption given covariates x one can reduce the information of x to a balancing score. Balancing scores $b(x)$ are only meaningful in the presence of conditional exchangeability with respect to covariates x. An implication of this result is that the difference between two potential outcomes are explained only by treatment assignment.

A natural way to apply these results to the non-probability sample setting is to consider selection to the non-probability sample as treatment assignment, and outcomes in the non-probability sample C and the rest in the finite population (i.e., $U \setminus C$) as two potential outcomes. In this setting, the conditional exchangeability of Rosenbaum and Rubin (1983) implies the conditional exchangeability with respect to the propensity score so that C and $U \setminus C$ are comparable given the propensity score. In contrast, Professor Li immediately assumes comparability of C and U given the propensity score. From the causal inference perspective, comparability of Rosenbaum and Rubin (1983) is a consequence of a conceptually checkable assumption while Professor Li begins with the desired comparability by assuming it. If, instead, one starts from conditional exchangeability as in Rosenbaum and Rubin (1983), a result still may not be satisfactory because two samples (i.e., C and $U \setminus C$) remain different by "treatment" of participation in a non-probability sample. For example, if non-probability samples are hospital records or participants of a certain educational program, both samples differ due to receipt of care by the hospital or the educational effect. Even if we do not find such "treatment" that differentiates the non-probability sample and the rest, the conditional comparability between C and $U \setminus C$ does not necessarily correspond to the finite population U. To achieve the correct target population, one needs to obtain a distribution of the propensity score in the finite population U. This task is not simple to carry out as described below in relation to the odds representation of the propensity score.

Another approach is to deviate from causal inference by starting from the conditional independence between Y and selection Z into C given X instead of the conditional exchangeability with potential outcomes. In this case, all derivations in fact remain valid to conclude the result that $Y \perp Z \mid X$ implies $Y \perp Z \mid b(X)$ as desired. However, a new conditional independence assumption is simply the standard missing at random (MAR) assumption in the missing data problem, which is also adopted by Chen, Li and Wu (2020) on their non-probability sample research. The MAR assumption is familiar to many statisticians and easier to examine than the conditional exchangeability assumption of Professor Li. If this approach is the one implicitly adopted in Wang, Graubard, Katki and Li (2022), as well as the current paper, it is worthwhile to discuss additional benefits of this approach over the MAR assumption in addition to the discrepancy between $U \setminus C$ and U for comparability. If a different approach is adopted, an unverified relationship between balancing scores and the CE assumption (2.1) should be explicitly derived. As an aside, we would like to point out that Chen, Li and Wu (2020), is not the only literature that does not use the CE assumption of Professor Li for the propensity score weighting methods (see e.g. Kim and Morikawa (2023) for the non-ignorable missing case).

As mentioned above, the comparability of C and $U \setminus C$ allows reliable estimation of the regression model based on C for items in $U \setminus C$ but the estimation of $\bar{Y}_N$ requires consistent estimation of propensity scores for U to bridge regression given X to the entire population U. However, simple estimation of the propensity score is not possible because X is not available for all items in $U \setminus C$. The variable X is available in a reference sample S from U with a known design but S is not a simple alternative to $U \setminus C$ because items in S can be also in a non-probability sample C. To address this challenging issue, Wang, Valliant and Li (2021) found the relationship between the propensity score into C relative to U and the propensity score into C relative to the stacked sample of C and U where the same items in C and S are treated differently (for a rigorous derivation, see Savitsky, Williams, Gershunskaya, Beresovsky and Johnson (2023)). Using this relationship, Professor Li modeled the latter propensity score by binary regression to estimate the former. The event for the latter propensity score for a stacked sample is artificially constructed and conceptually difficult to model. This issue enhances the higher possibility of model misspecification, which would invalidate design-consistent estimation of $\bar{Y}_N$. The event for the former propensity score is the original event, and is natural to model. This approach was adopted by Savitsky, Williams, Gershunskaya, Beresovsky and Johnson (2023).

3. Finite population inference

Another concept we want to discuss is the role of the finite population in non-probability samples. The goal of the paper is to develop a design-consistent estimator of the finite population average $\bar{Y}_N$. For design consistency, one assumes series of conditions on the sequence of finite populations with all variables except selection into samples treated non-random. In contrast, the model-based approach treats the finite population as a random realization from the super population, and models the stochastic relationship among variables. In the missing data research, on the other hand, two major approaches (and their combinations) for estimation are the propensity score modeling and the outcome regression modeling. A more suitable approach to the design-based approach is the propensity score modeling that models selection into samples given covariates. This is because one can consider random selections while all other variables can be treated fixed. On the other hand, the outcome regression modeling assumes a distribution for Y given X, and is suitable for the model-based approach.

Professor Li made a difficult attempt to bridge the outcome regression approach to the design-based approach. Note that the conditional expectation can be considered as regression with conditioning variables as covariates. From this view, the approach in the paper seems to be purely the model-based approach based on the outcome regression. However, Professor Li attempted to carefully develop the conditional expectation step by step beginning a finite population and a non-probability sample. If the condition was purely model-based, the variable y in the condition (2.1) is simply a random variable from the super population. In the conditional approach of the paper, this variable y should be clearly defined in relation to the finite population U and the non-probability sample C through indices. If y is a random choice of a variable from a sample S from U, $E_S[y \mid U] = \sum_{i \in U} \pi_i Y_i$ where π_i is the inclusion probability for the unit i. In this case, the self-weighting sample S satisfies $E_S[y \mid U] = \bar{Y}_N$ but a stratified sample S, for example, does not satisfy this equality in general. In other words, the claimed issue of bias may not be unique to a non-probability sample. To fully appreciate the conditional exchangeability condition, a clear definition of y in C and/or U is much desired. Moreover, it is desirable to elucidate how the model-based condition of the CE assumption leads to the design-based result despite conceptual discrepancy.

References

Chen, Y., Li, P. and Wu, C. (2020). Doubly robust inference with nonprobability survey samples. *Journal of the American Statistical Association*, 115(532), 2011-2021. Retrieved from https://doi-org.proxy-um.researchport.umd.edu/10.1080/01621459.2019.1677241. Doi: 10.1080/01621459.2019.1677241.

Kim, J., and Morikawa, K. (2023). An empirical likelihood approach to reduce selectionbias in voluntary samples. To appear in *Calcutta Statistical Association Bulletin*, 35.

Rosenbaum, P.R., and Rubin, D.B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41-55. Retrieved from https://doi-org.proxy-um.researchport.umd.edu/10.1093/biomet/70.1.41. Doi: 10.1093/biomet/70.1.41.

Savitsky, T.D., Williams, M.R., Gershunskaya, J., Beresovsky, V. and Johnson, N.G. (2023). *Methods for Combining Probability and Nonprobability Samples Under Unknown Overlaps*.

Wang, L., Graubard, B.I., Katki, H.A. and Li, Y. (2022). Efficient and robust propensity-score-based methods for population inference using epidemiologic cohorts. *International Statistical Review*, 90, 146-164.

Wang, L., Valliant, R. and Li, Y. (2021). Adjusted logistic propensity weighting methods for population inference using nonprobability volunteer-based epidemiologic cohorts. *Statistics in Medicine*, 40(24), 5237-5250. Retrieved from https://doi-org.proxy-um.researchport.umd.edu/10.1002/sim.9122. Doi: 10.1002/sim.9122.