Available online at www.sciencedirect.com
ScienceDirect
Statistics & Probability Letters 78 (2008) 144–149
STATISTICS & PROBABILITY LETTERS
www.elsevier.com/locate/stapro

Sharp bounds on the causal effects in randomized experiments with "truncation-by-death"☆

Kosuke Imai*

Department of Politics, Princeton University, Princeton, NJ 08544, USA

Received 7 December 2006; received in revised form 17 March 2007; accepted 23 May 2007
Available online 13 June 2007

Abstract

Many randomized experiments suffer from the "truncation-by-death" problem where potential outcomes are not defined for some subpopulations. For example, in medical trials, quality-of-life measures are only defined for surviving patients. In this article, I derive the sharp bounds on causal effects under various assumptions. My identification analysis is based on the idea that the "truncation-by-death" problem can be formulated as the contaminated data problem. The proposed analytical techniques can be applied to other settings in causal inference including the estimation of direct and indirect effects and the analysis of three-arm randomized experiments with noncompliance.

Keywords: Average treatment effect; Causal inference; Direct and indirect effect; Identification; Principal stratification; Quantile treatment effect

1. Introduction

In a randomized medical trial, researchers are often interested in estimating the causal effects of a treatment on various quality-of-life measures. However, these outcomes are only measured for surviving patients and are not even defined for those individuals who die during the trial. Thus, the causal quantities of interest are defined only for those patients who would survive under both the treatment and control conditions. A similar "truncation-by-death" problem also arises in other experiments. For example, suppose that one conducts a randomized experiment to estimate the causal effects of a certain teaching program on students' test scores. Since the test scores are not defined for those who drop out of school during the experiment, one can estimate the causal effects only for those who would stay in school under both the treatment and control conditions.

Zhang and Rubin (2003) analyzes this "truncation-by-death" problem using the principal stratification approach of Frangakis and Rubin (2002), and derives the bounds on the average treatment effect. In this article, I provide an alternative proof that these bounds are sharp (i.e., the shortest bounds possible without

☆This research is in part supported by the National Science Foundation (SES-0550873) and Princeton University Committee on Research in the Humanities and Social Sciences.
*Tel.: +1 609 258 6610; fax: +1 973 556 1929.
E-mail address: kimai@princeton.edu
URL: http://imai.princeton.edu

0167-7152/$ - see front matter
doi:10.1016/j.spl.2007.05.015

additional assumptions). I also show how the proposed proof can be generalized to derive the sharp bounds on the quantile treatment effect, which is another commonly used causal quantity of interest (e.g., Abadie et al., 2002). In particular, I demonstrate that the "truncation-by-death" problem can be formulated as the contaminated data problem and apply the identification results of Horowitz and Manski (1995) to derive the sharp bounds. Under additional assumptions, the bounds of Zhang and Rubin (2003) can be simplified and are shown to have a closed-form expression unlike the expression given in the original paper which involves numerical optimization. Finally, I briefly discuss the other potential applications of the analytical techniques employed in this article.

2. Framework

Consider a random sample of size n from a population. Let T_i be a binary treatment variable which is equal to 1 if unit i is treated and is equal to 0 if unit i belongs to the control group. Following the potential outcome framework (e.g., Holland, 1986), let $W_i(t)$ denote the potential "truncation-by-death" indicator variable under the treatment status $T_i = t$ for $t = 0, 1$. For example, $W_i(1) = 1$ means that unit i dies under the treatment and hence the outcome variable is not defined for this unit under the treatment arm. The observed "truncation-by-death" indicator, therefore, is given by $W_i = T_i W_i(1) + (1 - T_i) W_i(0)$. Next, let $Y_i(w, t)$ denote the potential outcome under the treatment status $T_i = t$ and the truncation status $W_i(t) = w$ for $t, w = 0, 1$. While $Y_i(0, t)$ exists, $Y_i(1, t)$ is not defined. Similarly, the observed outcome, Y_i, exists only for units with $W_i = 0$. Throughout this article, I assume that the treatment is randomized,

Assumption 1 (*Randomization of the treatment*).

$$(Y_i(0,0),\ Y_i(0,1),\ W_i(1),\ W_i(0)) \perp\!\!\!\perp T_i \quad \text{for all } i.$$

3. Sharp bounds without assumptions

Applying the principal stratification approach of Frangakis and Rubin (2002), Zhang and Rubin (2003) show that the unit causal effect is defined only for the subpopulation of units whose outcome would not be truncated under both the treatment and control conditions, i.e., $Y_i(0, 1) - Y_i(0, 0)$. They then derive the bounds on the average treatment effect (ATE) for this subpopulation,

$$\tau_{\text{ATE}} = E[Y_i(0,1) - Y_i(0,0) \mid W_i(0) = 0, W_i(1) = 0]. \tag{1}$$

Here, I prove that the bounds on τ_{ATE} given in Zhang and Rubin (2003) are sharp and derive the sharp bounds on the quantile treatment effect (QTE) for the same subpopulation,

$$\tau_{\text{QTE}}(\alpha) = q_{00|1}(\alpha) - q_{00|0}(\alpha), \tag{2}$$

where $q_{00|t}(\alpha) \equiv \inf\{y: P_{00|t}(y) \geqslant \alpha\}$ is the α-quantile of the distribution $P_{00|t}$ for $\alpha \in (0, 1)$. For example, if $\alpha = 0.5$, then τ_{QTE} represents the median treatment effect.

Let $P_t \equiv P(y \mid W_i = 0, T_i = t)$ be the conditional distribution of the observed outcome variable given the treatment status. I use $P_{w_0 0|1} \equiv P(y(0,1) \mid W_i(0) = w_0, W_i(1) = 0, T_i = 1)$ and $P_{0 w_1|0} \equiv P(y(0,0) \mid W_i(0) = 0, W_i(1) = w_1, T_i = 0)$ to represent the conditional distributions of the two potential outcomes given the treatment status for a corresponding subpopulation. Finally, define $\pi_{w_0 w_1} \equiv \Pr(W_i(0) = w_0, W_i(1) = w_1)$ for $w_0,\ w_1 = 0, 1$. I assume $\pi_{00} > 0$ so that the causal quantities of interest in Eqs. (1) and (2) are well defined.

As shown in Zhang and Rubin (2003), the observed outcome distribution can be written as a mixture of the two unobserved distributions,

$$P_0 = \frac{\pi_{00}}{1 - p_0} P_{00|0} + \left(1 - \frac{\pi_{00}}{1 - p_0}\right) P_{01|0}, \tag{3}$$

$$P_1 = \frac{\pi_{00}}{1 - p_1} P_{00|1} + \left(1 - \frac{\pi_{00}}{1 - p_1}\right) P_{10|1}, \tag{4}$$

where $p_t = \Pr(W_i = 1|T_i = t)$. Note that the sampling process only identifies P_t and p_t for $t = 0, 1$, while only the following sharp bounds are identified for π_{00},

$$\pi_{00} \in \Pi \equiv (0, 1] \cap [1 - p_0 - p_1, \min(1 - p_0, 1 - p_1)]. \tag{5}$$

I now establish the sharp bounds on the ATE and the QTE by applying the identification results of Horowitz and Manski (1995).

Proposition 1 (*Sharp bounds without assumptions*). *Let Ω be the sample space of Y. Define $r_t(\alpha) = \inf\{y: P_t[-\infty, y] \geqslant \alpha\}$ if $0 < \alpha < 1$, $r_t(\alpha) = \inf\{y: y \in \Omega\}$ if $\alpha \leqslant 0$, and $r_t(\alpha) = \sup\{y: y \in \Omega\}$ if $\alpha \geqslant 1$. Then, given Assumption 1, the following sharp bounds can be derived*:

1. *Average treatment effect*:

$$\tau_{\text{ATE}} \in \left[\min_{\pi_{00}\in\Pi}\left(\int y\,\mathrm{d}L_{1-\pi_{00}/(1-p_1)|1} - \int y\,\mathrm{d}U_{1-\pi_{00}/(1-p_0)|0}\right), \max_{\pi_{00}\in\Pi}\left(\int y\,\mathrm{d}U_{1-\pi_{00}/(1-p_1)|1} - \int y\,\mathrm{d}L_{1-\pi_{00}/(1-p_0)|0}\right)\right],$$

where the distributions, $L_{\gamma|t}$ and $U_{\gamma|t}$ for $t = 0, 1$, are defined as follows:

$$L_{\gamma|t}[-\infty, y] \equiv \begin{cases} P_t[-\infty, y]/(1-\gamma) & \text{if } y < r_t(1-\gamma), \\ 1 & \text{if } y \geqslant r_t(1-\gamma), \end{cases}$$

$$U_{\gamma|t}[-\infty, y] \equiv \begin{cases} 0 & \text{if } y < r_t(\gamma), \\ (P_t[-\infty, y] - \gamma)/(1-\gamma) & \text{if } y \geqslant r_t(\gamma). \end{cases}$$

2. *Quantile treatment effect*:

$$\tau_{\text{QTE}} \in \left[\min_{\pi_{00}\in\Pi}\left\{r_1\left(\frac{\alpha\pi_{00}}{1-p_1}\right) - r_0\left(1 - \frac{(1-\alpha)\pi_{00}}{1-p_0}\right)\right\}, \max_{\pi_{00}\in\Pi}\left\{r_1\left(1 - \frac{(1-\alpha)\pi_{00}}{1-p_1}\right) - r_0\left(\frac{\alpha\pi_{00}}{1-p_0}\right)\right\}\right].$$

The proof is given in Appendix.

4. Sharp bounds with additional assumptions

Zhang and Rubin (2003) consider the two assumptions that can be used to tighten the sharp bounds in Proposition 1. First, it may be assumed that those individuals who would survive under both the treatment and control conditions are healthier than those who die under one of the two conditions.

Assumption 2 (*Stochastic dominance*).

$$P_{00|0}[-\infty, y] \leqslant P_{01|0}[-\infty, y] \quad \text{and} \quad P_{00|1}[-\infty, y] \leqslant P_{10|1}[-\infty, y], \text{ for all } y \in \Omega.$$

Under this assumption, as the following proposition shows, it is possible to obtain a closed-form expression of the sharp bounds,

Proposition 2 (*Sharp bounds under the stochastic dominance assumption*). *Suppose that $p_0 + p_1 < 1$. Then, under Assumptions 1 and 2, the following sharp bounds can be derived*:

1. *Average treatment effect*:

$$\tau_{\text{ATE}} \in \left[\overline{Y}_1 - \int y\,\mathrm{d}U_{p_1/(1-p_0)|0}, \int y\,\mathrm{d}U_{p_0/(1-p_1)|1} - \overline{Y}_0\right],$$

where $\overline{Y}_t = E(Y_i|W_i = 0, T_i = t)$ for $t = 0, 1$.

2. *Quantile treatment effect*:

$$\tau_{\text{QTE}} \in \left[r_1(\alpha) - r_0\left\{1 - \frac{(1-\alpha)p_1}{(1-p_0)^2}\right\}, r_1\left\{1 - \frac{(1-\alpha)p_0}{(1-p_1)^2}\right\} - r_0(\alpha)\right].$$

The proof is given in Appendix. Note that $\overline{Y}_t$ is identifiable from the data. The closed-form expression of the sharp bounds on the ATE, derived in Proposition 2, simplifies the expression given in Zhang and Rubin (2003), which involves numerical optimization.

Secondly, Zhang and Rubin (2003) consider another assumption, which states that there exists no unit whose outcome would be truncated only under the treatment.

Assumption 3 (*Monotonicity*).

$$W_i(1) \leqslant W_i(0) \quad \text{for all } i,$$

which implies $\pi_{00} = 1 - p_0$.

Because of the bounds on π_{00} given in Eq. (5), Assumption 3 has a testable implication, i.e., $p_1 \leqslant p_0$, which can be checked using the observed data. Under this assumption, the closed-form sharp bounds can be obtained by substituting $\pi_{00} = 1 - p_0$ into the sharp bounds in Proposition 1,

$$\tau_{\text{ATE}} \in \left[\int y \, \mathrm{d}L_{(p_0-p_1)/(1-p_1)|1} - \overline{Y}_0, \int y \, \mathrm{d}U_{(p_0-p_1)/(1-p_1)|1} - \overline{Y}_0\right], \tag{6}$$

$$\tau_{\text{QTE}} \in \left[r_1\left\{\frac{\alpha(1-p_0)}{1-p_1}\right\} - r_0(\alpha), r_1\left\{1 - \frac{(1-\alpha)(1-p_0)}{1-p_1}\right\} - r_0(\alpha)\right]. \tag{7}$$

Finally, if both Assumptions 2 and 3 hold (in addition to Assumption 1), then the sharp bounds can be further simplified as follows:

$$\tau_{\text{ATE}} \in \left[\overline{Y}_1 - \overline{Y}_0, \int y \, \mathrm{d}U_{(p_0-p_1)/(1-p_1)|1} - \overline{Y}_0\right], \tag{8}$$

$$\tau_{\text{QTE}} \in \left[r_1(\alpha) - r_0(\alpha), r_1\left\{1 - \frac{(1-\alpha)(1-p_0)}{1-p_1}\right\} - r_0(\alpha)\right]. \tag{9}$$

The closed-form expressions of the sharp bounds on the ATE shown in Eqs. (6) and (8) again simplify the expressions given in Zhang and Rubin (2003). Furthermore, the bounds in Eqs. (8) and (9) imply that under Assumptions 2 and 3, a naive analysis based on the units with $W_i = 0$ underestimates the ATE and QTE.

5. Concluding remarks and possible applications

In this article, building on the works of Zhang and Rubin (2003) and Horowitz and Manski (1995), I show how to derive sharp bounds on the average treatment effect and the quantile treatment effect in randomized experiments with the "truncation-by-death" problem under several assumptions. A similar analytical strategy can be applied to three-arm (or more) randomized trials with noncompliance (Imai, 2005; Cheng and Small, 2006) as well as randomized experiments with both noncompliance and "truncation-by-death" in order to derive the sharp bounds on various causal quantities of interest. Yet, another application is the identification of direct and indirect effects in randomized experiments (Rubin, 2004). This line of research is currently undertaken by the author (Imai, 2007).

Acknowledgment

I thank a referee for helpful comments.

Appendix A. Proofs of the Propositions

The proof of Proposition 1 uses the following lemma, which shows that the identification results with the contaminated data, specifically those in Horowitz and Manski (1995, Proposition 4) can be directly applied to the situation where the lower bound as well as the upper bound on the error probability are known.

Lemma 1 (*Identification with the contaminated data when both upper and lower bounds are known*). *Let (Y_0, Y_1, Z) be random variables where $Y_0, Y_1 \in \Omega$ and $Z \in \{0,1\}$. Assume that only $Y \equiv Y_0(1-Z) + Y_1 Z$ is observed. Moreover, let $Q \equiv Q(y)$ be the distribution of the observable Y, and $P_{ij} \equiv P_i(y_i|z=j)$ denote the distribution of Y_i given $Z=j$ for $i,j = 0,1$. The sampling process identifies Q but neither P_{11} nor P_{00} is identified, where $Q = (1-p)P_{11} + pP_{00}$ and $p \in (0,1)$ is the error probability.*

Suppose that $0 \leqslant \lambda_0 \leqslant p \leqslant \lambda_1 < 1$, and that both λ_0 and λ_1 are assumed to be known. If $g(\cdot)$ is a function that respects stochastic dominance, then the following bounds are sharp:

$$g(P_{11}) \in [g(L_{\lambda_1}), g(U_{\lambda_1})],$$

where

$$L_\gamma[-\infty, y] \equiv \begin{cases} Q[-\infty, y]/(1-\gamma) & \text{if } y < r(1-\gamma), \\ 1 & \text{if } y \geqslant r(1-\gamma), \end{cases}$$

$$U_\gamma[-\infty, y] \equiv \begin{cases} 0 & \text{if } y < r(\gamma), \\ (Q[-\infty, y] - \gamma)/(1-\gamma) & \text{if } y \geqslant r(\gamma), \end{cases}$$

and, $r(\alpha) = \inf\{y\colon Q[-\infty, y] \geqslant \alpha\}$ *if* $0 < \alpha < 1$, $r(\alpha) = \inf\{y\colon y \in \Omega\}$ *if* $\alpha \leqslant 0$, *and* $r(\alpha) = \sup\{y\colon y \in \Omega\}$ *if* $\alpha \geqslant 1$.

Proof of Lemma 1. Suppose that p is known. Then, the same argument as the one given in the proof of Proposition 4 of Horowitz and Manski (1995) can be applied to show that $g(P_{11}) \in [g(L_p), g(U_p)]$ and that these restrictions are sharp. Next, I show that for any $\varepsilon > 0$ and p such that $p + \varepsilon < 1$, $L_p[-\infty, y] \leqslant L_{p+\varepsilon}[-\infty, y]$ and $U_p[-\infty, y] \geqslant U_{p+\varepsilon}[-\infty, y]$ hold for all $y \in \Omega$. If $y < r(1-p-\varepsilon)$, then $L_{p+\varepsilon}[-\infty, y] - L_p[-\infty, y] = \varepsilon Q[-\infty, y]/[(1-p)(1-p-\varepsilon)] \geqslant 0$. If $r(1-p-\varepsilon) \leqslant y < r(1-p)$, then $L_{p+\varepsilon}[-\infty, y] - L_p[-\infty, y] = 1 - Q[-\infty, y]/(1-p) \geqslant 0$. If $y \geqslant r(1-p)$, then $L_{p+\varepsilon}[-\infty, y] - L_p[-\infty, y] = 0$. A similar calculation can be used to establish $U_p[-\infty, y] \geqslant U_{p+\varepsilon}[-\infty, y]$. Thus, since $g(\cdot)$ respects stochastic dominance, $[g(L_p), g(U_p)] \subset [g(L_{p+\varepsilon}), g(U_{p+\varepsilon})]$ for any p and $\varepsilon > 0$ with $p + \varepsilon < 1$. Finally, if $0 \leqslant \lambda_0 \leqslant p \leqslant \lambda_1 < 1$ and λ_0 and λ_1 are known, then $g(P_{11}) \in \bigcup_{\lambda_0 \leqslant p \leqslant \lambda_1} [g(L_p), g(U_p)] = [g(L_{\lambda_1}), g(U_{\lambda_1})]$. □

Next, I provide the proof of Proposition 1. The proof is given only for the bounds on the ATE. A similar proof can be constructed for the QTE.

Proof of Proposition 1. Assume that the value of π_{00} is known. Since the mean respects stochastic dominance, applying Lemma 1 to Eqs. (3) and (4) yields the following sharp bounds, $E[Y_i(0,t)|W_i(0) = 0, W_i(1) = 0] \in [\int y\,\mathrm{d}L_{\pi_{00}|t}, \int y\,\mathrm{d}L_{\pi_{00}|t}]$, for $t = 0,1$. Then, the bounds on the ATE are given by $\tau_{\mathrm{ATE}} \in [\int y\,\mathrm{d}L_{\pi_{00}|1} - \int y\,\mathrm{d}U_{\pi_{00}|0}, \int y\,\mathrm{d}U_{\pi_{00}|1} - \int y\,\mathrm{d}L_{\pi_{00}|0}]$. To prove that these bounds are sharp, note that the sampling process does not yield any restriction on the joint distribution of the potential outcomes $(Y_i(0,0), Y_i(0,1))$ because only one of them is observed for each unit. Thus, the sampling process does not provide any restriction on $(P_{00|0}, P_{00|1})$ other than Eqs. (3) and (4). Then, given a fixed value of π_{00}, the bounds on the ATE are sharp. Finally, since the value of π_{00} is unknown, minimizing (maximizing) the lower (upper) bound over the possible range of π_{00} gives the desired sharp bounds on the ATE. □

To prove Proposition 2, I first prove the following lemma:

Lemma 2 (*Stochastic dominance*). *In Lemma 1, assume further that P_{00} is stochastically dominated by P_{11}, i.e., $P_{11}[-\infty, y] \leqslant P_{00}[-\infty, y]$ for all $y \in \Omega$. Then, Lemma 1 holds by modifying the definition of the distribution, L_γ, as $L_\gamma[-\infty, y] \equiv Q[-\infty, y]$, which no longer depends on the value of γ.*

Proof of Lemma 2. Let Ψ be the set of all possible probability distributions for Y_i with $i = 0, 1$. The identification region for (P_{11}, P_{00}) is given by, $(P_{11}, P_{00}) \in \{(\psi_{11}, \psi_{00}) \in \Psi \times \Psi : Q = (1 - \lambda_1)\psi_{11} + \lambda_1\psi_{00}, \psi_{11}[-\infty, y] \leqslant \psi_{00}[-\infty, y]$ for all $y \in \Omega\}$. Thus, the identification region for $P_{11}[-\infty, y]$ is given by, $P_{11}[-\infty, y] \in \Psi_{11}(\lambda_1) \equiv \{(Q[-\infty, y] - \lambda_1\psi_{00}[-\infty, y])/(1 - \lambda_1) : \psi_{00} \in \{\psi_{00} \in \Psi : \psi_{00} \geqslant Q[-\infty, y]\}\}$. This implies that every member of $\Psi_{11}(\lambda_1)$ stochastically dominates L_{λ_1}, i.e., $L_{\lambda_1}[-\infty, y] = Q[-\infty, y] \geqslant \psi_{11}[-\infty, y]$ with $\psi_{11}[-\infty, y] \in \Psi_{11}(\lambda_1)$ for all $y \in \Omega$. Using the same argument given in the proof of Proposition 4 of Horowitz and Manski (1995), one can also show that U_{λ_1} stochastically dominates every member of $\Psi_{11}(\lambda_1)$. Since $g(\cdot)$ respects stochastic dominance, the results follow immediately. □

Now, the proof of Proposition 2 is immediate. The proof is given only for the bounds on the ATE. A similar proof can be constructed for the QTE.

Proof of Proposition 2. Lemma 2 implies that the sharp lower bound of $E[Y_i(0, t) | W_i(0) = 0, W_i(1) = 0]$ for $t = 0, 1$ equals $\overline{Y}_t$ and does not depend on the value of π_{00}, while its upper bound remains the same as before. This yields the desired results. □

References

Abadie, A., Angrist, J., Imbens, G., 2002. Instrumental variables estimates of the effect of subsidized training on the quantiles of trainess earnings. Econometrica 70 (1), 91–117.

Cheng, J., Small, D., 2006. Bounds on causal effects in three-arm trials with noncompliance. J. Roy. Statist. Soc. Ser. B Methodol. 68 (5), 815–836.

Frangakis, C.E., Rubin, D.B., 2002. Principal stratification in causal inference. Biometrics 58 (1), 21–29.

Holland, P.W., 1986. Statistics and causal inference (with discussion). J. Amer. Statist. Assoc. 81, 945–960.

Horowitz, J.L., Manski, C.F., 1995. Identification and robustness with contaminated and corrupted data. Econometrica 63 (2), 281–302.

Imai, K., 2005. Do get-out-the-vote calls reduce turnout?: the importance of statistical methods for field experiments. Amer. Polit. Sci. Rev. 99 (2), 283–300.

Imai, K., 2007. Identification analysis for randomized experiments with noncompliance and truncation by death. Technical Report, Department of Politics, Princeton University.

Rubin, D.B., 2004. Direct and indirect causal effects via potential outcomes (with discussions). Scandinavian J. Statist. 31 (2), 161–170.

Zhang, J.L., Rubin, D.B., 2003. Estimation of causal effects via principal stratification when some outcomes are truncated by “death”. J. Educational Behavioral Statist. 28 (4), 353–368.