Available online at www.sciencedirect.com
ScienceDirect
Computational Statistics and Data Analysis 52 (2008) 4091–4103
COMPUTATIONAL STATISTICS & DATA ANALYSIS
www.elsevier.com/locate/csda

Notes on interval estimation of risk difference in stratified noncompliance randomized trials: A Monte Carlo evaluation

Kung-Jong Lui*

Department of Mathematics and Statistics, College of Sciences, San Diego State University, San Diego, CA 92182-7720, USA

Received 7 February 2007; received in revised form 12 November 2007; accepted 2 February 2008
Available online 7 February 2008

Abstract

When comparing an experimental treatment with a standard treatment in a randomized clinical trial (RCT), we often use the risk difference (RD) to measure the efficacy of an experimental treatment. In this paper, we have developed four asymptotic interval estimators for the RD in a stratified RCT with noncompliance. These include an asymptotic interval estimator based on the weighted-least-squares (WLS) estimator of the RD, an asymptotic interval estimator using $\tanh^{-1}(x)$ transformation with the WLS optimal weight, an asymptotic interval estimator derived from Fieller's Theorem, and an asymptotic interval estimator using a randomization-based approach. Based on Monte Carlo simulations, we have compared these four asymptotic interval estimators with the asymptotic interval estimator recently proposed elsewhere. We have found that when the probability of compliance is high, the interval estimator using a randomization-based approach is probably more accurate than the others, especially when the stratum size is not large. When the probability of compliance is moderate, the interval estimator using $\tanh^{-1}(x)$ transformation is likely to be the best among all interval estimators considered here. We note that the interval estimator proposed elsewhere can be of use when the underlying RD is small, but lose accuracy when the RD is large. We also note that when the number of patients per assigned treatment is large, the four asymptotic interval estimators developed here are essentially equivalent; they are all appropriate for use. Finally, to illustrate the use of these interval estimators, we consider the data taken from a large field trial studying the effect of a multifactor intervention program on reducing the mortality of coronary heart disease in middle-aged men.

1. Introduction

Noncompliance may often occur in a randomized clinical trial (RCT) due to ethicality, a change of patient's mind to accept a treatment after randomization, or other reasons (Sommer and Zeger, 1991; Sato, 2001; Lui, 2006; Walter et al., 2006). When a patient assigned to an experimental treatment has serious side effects, we should allow the patient to switch to receive the standard treatment. For studying the effects of certain treatments, the traditional RCT can be inapplicable due to ethical concerns. For example, it is unethical to randomly assign high-risk patients to receive either a flu vaccine or a placebo in a flu shot trial (Zhou and Li, 2006). Similarly, when studying the effect of smoking, we cannot simply assign patients randomly either to quit smoking or to continue smoking for ethical consideration (Multiple Risk factor Intervention Trial Research Group, 1982). To alleviate the above ethical concerns,

* Tel.: +1 619 594 6191; fax: +1 619 594 6746.
E-mail address: kjl@rohan.sdsu.edu.

0167-9473/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.csda.2008.02.001

Table 1
The frequency data and the corresponding proportion in percentage (in parenthesis) regarding the status of cigarette smoking and the death incidence of coronary heart disease (CHD) during 7 years of follow-up in a multiple risk factor intervention trial

Number of cigarettes per day at baseline	Group:		Intervention			Control		
	Smoking:		Quit	Not quit	Total	Quit	Not quit	Total
<30	CHD	Yes	6	19	25	3	23	26
			(0.47)	(1.51)	(1.98)	(0.25)	(1.95)	(2.2)
		No	448	787	1235	156	998	1154
			(35.56)	(62.46)	(98.02)	(13.22)	(84.58)	(97.80)
		Total	454	806	1260	159	1021	1180
			(36.03)	(63.97)	(100)	(13.47)	(86.53)	(100)
≥30	CHD	Yes	5	39	44	1	47	48
			(0.19)	(1.52)	(1.71)	(0.04)	(1.77)	(1.81)
		No	532	1997	2529	214	2388	2602
			(20.68)	(77.61)	(98.60)	(8.08)	(90.11)	(98.19)
		Total	537	2036	2573	215	2435	2650
			(20.87)	(79.13)	(100)	(8.12)	(91.88)	(100)

the randomized encouragement design (RED), in which patients are randomly assigned either to the intervention group in which each patient receives an encouragement of accepting an experimental treatment or to the control group in which each patient receive no encouragement, is often employed (Matsui, 2005; Zhou and Li, 2006). The basic idea behind use of the RED is to increase the number of patients receiving the experimental treatment through encouragement in the intervention group without affecting the choice of a treatment in the control group. Thus, when there is a difference in the response rates between the intervention and control groups, we may attribute this difference to the effect of the experimental treatment. To illustrate the use of RED in which noncompliance can often occur, consider the data (Table 1) taken from the multiple risk factors intervention trial (MRFIT) to study the effect of a multiple intervention program on mortality from coronary heart disease (CHD) in middle-aged men (Multiple Risk factor Intervention Trial Research Group, 1982). Participants were randomly assigned to either the intervention or the control group. Participants assigned to the intervention group were provided with dietary advice to reduce blood cholesterol, smoking cessation counseling, and hypertension medication, while participants assigned to the control group were referred to their usual physicians for treatment. Since smoking is the only risk factor that has been effectively reduced by the intervention, we restrict our attention to the effects of quitting smoking as done elsewhere (Mark and Robins, 1993; Sato, 2000; Matsui, 2005). Note that patients with noncompliance considered in this RED example correspond to those who received smoking cessation counseling but did not quit smoking, and those who did not receive smoking cessation counseling but quitted smoking. Although there is a subtle difference in the meaning of "noncompliance" between the traditional RCT and RED, the results presented in this paper are applicable to both of these two designs. Note also that as seen from Table 1, for patients with the baseline of cigarettes smoked per day <30, there were approximately 64% (noncompliance) of patients assigned to the intervention group who did not quit smoking, while there were approximately 13% (noncompliance) of patients assigned to the control group who quitted smoking. By contrast, for patients with the baseline of cigarettes smoked per day ≥30, the former increases to approximately 79%, while the latter decreases to approximately 8%. When we assess a treatment effect, it is important to account for the differential rates of noncompliance between strata to avoid a possibly misleading inference. In fact, despite of the random assignment of patients to treatments in a traditional RCT, we may still encounter data with potential confounders that are unbalanced between two treatment groups by chance. Stratified analysis to control the confounding effects due to these unbalanced confounders is often suggested in this case as well. These motivate us to study interval estimation of the risk difference (RD) in a stratified trial with noncompliance.

Since noncompliance frequently occurs at non-random (Angrist et al., 1996; Frangakis and Rubin, 1999), analyzing data as treated or as per protocol can often lead us to a biased estimate of the treatment effect. The intention-to-treat (ITT) analysis, in which patients are compared according to the treatment to which they are originally designated rather than the treatment which they actually receive, is probably the most commonly-used approach to study the treatment