Cheng, Yu-Jen; **Wang, Mei-Cheng**

Causal estimation using semiparametric transformation models under prevalent sampling.
(English) Zbl 1390.62243
Biometrics 71, No. 2, 302-312 (2015).

Summary: This article presents methods and inference for causal estimation in semiparametric transformation models for the prevalent survival data. Through the estimation of the transformation models and covariate distribution, we propose a few analytical procedures to estimate the causal survival function. As the data are observational, the unobserved potential outcome (survival time) may be associated with the treatment assignment, and therefore there may exist a systematic imbalance between the data observed from each treatment arm. Further, due to prevalent sampling, subjects are observed only if they have not experienced the failure event when data collection began, causing the prevalent sampling bias. We propose a unified approach, which simultaneously corrects the bias from the prevalent sampling and balances the systematic differences from the observational data. We illustrate in the simulation study that standard analysis without proper adjustment would result in biased causal inference. Large sample properties of the proposed estimation procedures are established by techniques of empirical processes and examined by simulation studies. The proposed methods are applied to the Surveillance, Epidemiology, and End Results (SEER) and Medicare-linked data for women diagnosed with breast cancer.

MSC:

62P10 Applications of statistics to biology and medical sciences; meta analysis
62G07 Density estimation
62N05 Reliability and life testing

Keywords:

causal estimation; dependent truncation; prevalent sampling; survival analysis

Software:

timereg

Full Text: DOI

References:

[1] Brookmeyer, Biases in prevalent cohorts, Biometrics 43 pp 739– (1987) · Zbl 0715.62221 · doi:10.2307/2531529

[2] Buchholz, Radiation therapy for early-stage breast cancer after breast-conserving surgery, The New England Journal of Medicine 360 pp 63– (2009) · doi:10.1056/NEJMct0803525

[3] Chan, Estimating incident population distribution from prevalent data, Biometrics 68 pp 521– (2012) · Zbl 1251.62040 · doi:10.1111/j.1541-0420.2011.01708.x

[4] Chen, Semiparametric analysis of transformation models with censored data, Biometrika 89 pp 659– (2002) · Zbl 1039.62094 · doi:10.1093/biomet/89.3.659

[5] Chen, Causal inference on the difference of the restricted mean lifetime between two groups, Biometrics 57 pp 1030– (2001) · Zbl 1209.62267 · doi:10.1111/j.0006-341X.2001.01030.x

[6] Cheng, Analysis of transformation models with censored data, Biometrika 82 pp 867– (1995) · Zbl 0861.62071 · doi:10.1093/biomet/82.4.835

[7] Cheng, Predicting survival probabilities with semiparametric transformation models, Journal of the American Statistical Association 92 pp 227– (1997) · Zbl 0889.62080 · doi:10.1080/01621459.1997.10473620

[8] Cheng, Estimating propensity scores and causal survival functions using prevalent survival data, Biometrics 68 pp 707– (2012) · Zbl 1271.62065 · doi:10.1111/j.1541-0420.2012.01754.x

[9] Cole, The consistency assumption in causal inference: A definition or an assumption, Epidemiology 20 pp 3– (2009) · doi:10.1097/EDE.0b013e31818ef366

[10] Cox , D. R. 1958 Planning of Experiments · Zbl 0084.15802

[11] Greenland, Model-based estimation of relative risks and other epidemiologic measures in studies of common outcomes and in case-control studies, American Journal of Epidemiology 160 pp 301– (2004) · doi:10.1093/aje/kwh221

[12] Hahn, Conditions for sample-continuity and central limit theorem, Annals of Probability 5 pp 351– (1977) · Zbl 0388.60043 · doi:10.1214/aop/1176995796

[13] Holland, Statistics and causal inference, Journal of the American Statistical Association 81 pp 945– (1986) · Zbl 0607.62001 · doi:10.1080/01621459.1986.10478354

[14] Hudgens, Toward causal inference with interference, Journal of the American Statistical Association 103 pp 832– (2008) · Zbl 05564536 · doi:10.1198/016214508000000292

[15] Kaplan, Nonparametric estimation from incomplete observations, Journal of the American Statistical Association 53 pp 457– (1958) · Zbl 0089.14801 · doi:10.1080/01621459.1958.10501452

[16] Lane, Analysis of covariance and standardization as instances of prediction, Biometrics 38 pp 613– (1982) · doi:10.2307/2530043

[17] Lynden-Bell, A method of allowing for known obervational selection in small samples applied to 3cr quasars, Monographs of the Monthly Notices of the Royal Astronomical Society 155 pp 95– (1971) · doi:10.1093/mnras/155.1.95

[18] Martinussen , T. Scheike , T. H. 2006 Dynamic Regression Models for Survival Data · Zbl 1096.62119

[19] Neyman, Roczniki Nauk Roiniczych, Tom X pp 1– (1923)

[20] Rebolledo, Central limit theorems for local martingales, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete 51 pp 269– (1980) · Zbl 0432.60027 · doi:10.1007/BF00587353

[21] Rosenbaum, Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome, Journal of the Royal Statistical Society, Series B 45 pp 212– (1983)

[22] Rubin, Estimating causal effects of treatments in randomized and nonrandomized studies, Journal of Educational Psychology 66 pp 688– (1974) · doi:10.1037/h0037350

[23] Rubin, Bayesian inference for causal effects: The role of randomization, Annals of Statistics 6 pp 34– (1978) · Zbl 0383.62021 · doi:10.1214/aos/1176344064

[24] Rubin, Journal of the American Statistical Association 75 pp 591– (1980)

[25] Rubin, Formal models of statistical inference for causal effects, Journal of Statistical Planning and Inference 25 pp 279– (1990) · doi:10.1016/0378-3758(90)90077-8

[26] Rubin, Estimating causal effects from large data sets using propensity scores, Annals of Internal Medicine 127 pp 757– (1997) · doi:10.7326/0003-4819-127-8_Part_2-199710151-00064

[27] Rubin, Teaching statistical inference for causal effects in experiments and observational studies, Journal of Educational and Behavioral Statistics 29 pp 343– (2004) · doi:10.3102/10769986029003343

[28] Sobel, What do randomized studies of housing mobility demonstrate?, Journal of the American Statistical Association 101 pp 1398– (2006) · Zbl 1171.62365 · doi:10.1198/016214506000000636

[29] Tan, A distributional approach for causal inference using propensity scores, Journal of the American Statistical Association 101 pp 1619– (2006) · Zbl 1171.62320 · doi:10.1198/016214506000000023

[30] VanderWeele, Concerning the consistency assumption in causal inference, Epidemiology 20 pp 880– (2009) · doi:10.1097/EDE.0b013e3181bd5638

[31] VanderWeele, Unmeasured confounding for general outcomes, treatments, and confounders: Bias formulas for sensitivity analysis, Epidemiology 22 pp 42– (2011) · doi:10.1097/EDE.0b013e3181f74493

[32] Wang, Nonparametric estimation from cross-sectional survival data, Journal of the American Statistical Association 86 pp 130– (1991) · Zbl 0739.62026 · doi:10.1080/01621459.1991.10475011

[33] Wang, Statistical models for prevalent cohort data, Biometrics 49 pp 1– (1993) · Zbl 0771.62079 · doi:10.2307/2532597

[34] Wang, Asymptotic properties of the product limit estimate under random truncation, Annals of Statistics 14 pp 1597– (1986) · Zbl 0656.62048 · doi:10.1214/aos/1176350180

[35] Warren, Overview of the seer-medicare data: Content, research applications, and generalizability to the united states elderly population, Medical Care 40 pp 3– (2002) · doi:10.1097/00005650-200208001-00002

[36] Woodroofe, Estimating a distribution function with truncated data, Annals of Statistics 13 pp 163– (1985) · Zbl 0574.62040 · doi:10.1214/aos/1176346584

[37] Zeng, Efficient estimation of semiparametric transformation models for counting processes, Biometrika 93 pp 627– (2006) · Zbl 1108.62083 · doi:10.1093/biomet/93.3.627

[38] Zucker, A pseudo-partial likelihood method for semiparametric survival regression with covariate errors, Journal of the American Statistical Association 100 pp 1264– (2005) · Zbl 1117.62463 · doi:10.1198/016214505000000538