Propensity Score [weighting] within complex survey


ehsan.karim@ubc.ca

Oct 7, 2020

SPPH 504/007

Propensity score Weighting (ATE + ATT)

IPW (inverse probability weighting)

How to conduct propensity score weighting?

Step 1: Specify PS & fit model → Exposure model (RA)

Step 2: ~~Match subjects by PS~~ Convert PS to IPW

Step 3: Covariate balance in ~~matched~~ weighted sample

Step 4: Estimate treatment effect → Outcome model (MI)

For the purposes of illustration, we will first assume that our data was collected via SRS.

IPW

Step 1: Fit PS model

A~L

Step 2: Convert PS = IPW(ATE)

IPW = 1/ps, if A = 1

IPW = 1/(1-ps), if A = 0

Step 3: Check balance

SMD in IPW-weighted data

Step 4: Outcome model with

Weight = IPW

IPW in complex survey

Step 1: Fit PS model

A~L (survey-weights as design variable / covariate)

Step 2: Convert PS = IPW(ATE)

IPW = 1/ps, if A = 1

IPW = 1/(1-ps), if A = 0

Step 3: Check balance

SMD (data weighted by w = IPW * survey-weights)

Step 4: Outcome model with

Weight = IPW * survey-weights

Generalizing observational study results: applying propensity score methods to complex surveys

EH DuGoff, M Schuler, EA Stuart - Health services research, 2014 - Wiley Online Library

Objective To provide a tutorial for using propensity score methods with complex survey data. Data Sources Simulated data and the 2008 Medical Expenditure Panel Survey. Study Design Using simulation, we compared the following methods for estimating the treatment effect: a naïve estimate (ignoring both survey weights and propensity scores), survey weighting, propensity score methods (nearest neighbor matching, weighting, and subclassification), and propensity score methods in combination with survey weighting ...

☆ 99 Cited by 219 Related articles All 11 versions

Propensity score analysis with survey weighted data

G Ridgeway, SA Kovalchik, BA Griffin... - Journal of Causal ..., 2015 - degruyter.com

"sampling weights in the propensity score estimation stage (as weights, not as a covariate)"

IPW

Step 1: Fit PS model

A~L

Step 2: Convert PS = IPW(ATT)

IPW = 1, if A = 1

IPW = ps/(1-ps), if A = 0

Step 3: Check balance

SMD in IPW-weighted data

Step 4: Outcome model with

Weight = IPW

IPW in complex survey (ATT)

Step 1: Fit PS model

A~L (survey-weights as design variable / covariate)

Step 2: Convert PS = IPW(ATT)

IPW = 1, if A = 1

IPW = ps/(1-ps), if A = 0

Step 3: Check balance

SMD (data weighted by w = IPW * survey-weights)

Step 4: Outcome model with

Weight = IPW * survey-weights

Reasonable approach (my summary)📄

- PS model: (population-level)
  - use design variables (cluster + strata + weight) to estimate ps (not as covariate)
  - Combined weight = ipw * survey weight
- Outcome model: (population-level)
  - use design features (strata+psu as well as combined weight) to get population level estimates

Estimates and conclusion

Adult <u>patients with RA</u> are at <u>increased risk for MI</u> in <u>US</u> (based on 2007-08 data)?

Risk of cardiovascular mortality in patients with rheumatoid arthritis: a meta-analysis of observational studies

JA Aviña-Zubieta, HK Choi... - Arthritis Care & ..., 2008 - Wiley Online Library

Objective To determine the magnitude of risk of cardiovascular mortality in patients with rheumatoid arthritis (RA) compared with the general population through a meta-analysis of observational studies. Methods We searched Medline, EMBase, and Lilacs databases from

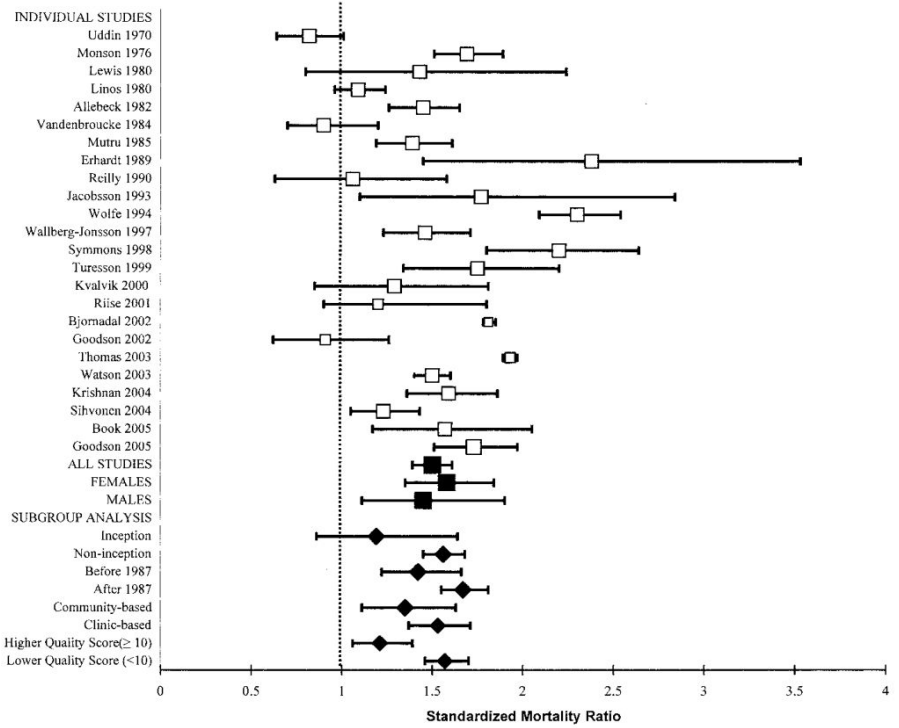

Figure 1. Meta-analysis of 24 studies on cardiovascular disease mortality in patients with rheumatoid arthritis.

50% increased risk of CVD death in patients with RA

Estimates from NHANES (2007-08) and conclusion

OR: population-based estimates, sample-based not shown

	Adjusted Regression	Matching (Zanutto)	Matching (DuGoff)	Matching (design in both stages)	Weighting (Ridgeway)	Weighting (DuGoff)
PATT		1.87 (0.86 4.07)	1.26 (0.55, 2.88)	1.66 (0.65, 4.28)	1.38 (0.71, 2.71)	1.37 (0.71, 2.67)
PATE	1.66* (0.71, 3.89)				1.51 (0.68, 3.35)	1.43 (0.62, 3.28)

* Also conditional estimates if further adjustment made; SE / CI width is a function of n.

NHANES vs. CCHS

- In the public release data, NHANES provides
  - masked variance pseudo-PSUs, and
  - masked variance pseudo-stratum

  to account for the complex survey design.

- CCHS public use microdata file (PUMF) does not contain PSU / Stratum information. Any SE calculation assumes SRS even if weights are used. RDC provides access to master data with these necessary information.

Short Reference and Textbook List📖

- Austin P. C., Jembere N. and Chiu M. (2018). Propensity score matching and complex surveys. Statistical Methods in Medical Research, 27(4), 1240-1257.
- DuGoff, E. H., Schuler, M., & Stuart, E. A. (2014). Generalizing observational study results: applying propensity score methods to complex surveys. Health services research, 49(1), 284-303.
- Heeringa, S. G., West, B. T., & Berglund, P. A. (2017). Applied survey data analysis. Chapman and Hall/CRC.
- Kim, Y., Park, S., Kim, N. S., & Lee, B. K. (2013). Inappropriate survey design analysis of the Korean National Health and Nutrition Examination Survey may produce biased results. Journal of preventive medicine and public health, 46(2), 96.
- Lenis, D., Nguyen, T. Q., Dong, N., & Stuart, E. A. (2017). It's all about balance: propensity score matching in the context of complex survey data. Biostatistics, 20(1), 147-163.
- Lewis, T. H. (2016). Complex survey data analysis with SAS. Chapman and Hall/CRC.
- Lumley, T. (2011). Complex surveys: a guide to analysis using R (Vol. 565). John Wiley & Sons.
- Lumley T. (2016). Survey: Analysis of Complex Survey Samples . R package version 3.31. https://cran.r-project.org/web/packages/survey/index.html.
- Ridgeway G., Kovalchik S. A., Griffin B. A. and Kabeto M. U. (2015). Propensity score analysis with survey weighted data. Journal of Causal Inference 3, 237-249.
- Zanutto E. L. (2006). A comparison of propensity score and linear regression analysis of complex survey data. Journal of Data Science 4, 67-91.

Thanks!


ehsan.karim@ubc.ca


www.ehsankarim.com