Leveraging External Individualized Prediction Models in Bayesian Survival Analysis

Mi-Ok Kim[1], Yena Jeon[1], Yunxiang Huang[1], Hangjoon Kim[2],

[1]*Department of Epidemiology and Biostatistics, University of California San Francisco, USA*

[2]*Division of Statistics and Data Science, University of Cincinnati, USA*

ABSTRACT

Individualized risk prediction algorithms, such as the Prostate Cancer Risk Assessment tool, are increasingly used to predict cancer relapse or progression. Since these algorithms are typically trained on large datasets, effectively integrating their outputs can enhance the efficiency of analyzing individual studies. In this research, we consider Cox regression analysis for right-censored time-to-event outcomes, incorporating external information provided by large-scale prediction models. We adopt a Bayesian inference in estimating the baseline hazard at each distinct time point. External information is integrated through the Kullback–Leibler (KL) divergence, leading to informative priors for Bayesian analysis. The performance of the proposed model is demonstrated through simulation studies and an application to data from a prostate cancer clinical trial.

Keywords: Informative Priors; External Prediction Information

Back to Sessions List

Propensity Weighting Plus Adjustment in Proportional Hazards Model Is Not Doubly Robust

Erin Evelyn Gabriel*[1],   Michael C Sachs[1] , Ingeborg Waernbaum[2] , Els Goetghebeur[3] , Paul F Blanche[1] , Stijn Vansteelandt[3] , Arvid Sjölander[4] , Thomas Scheike[1]

[1]*Section of Biostatistics, Department of Public Health, University of Copenhagen, København 1353, Denmark*

[2]*Department of Statistics, Uppsala University, Uppsala 75120, Sweden*

[3]*Department of Applied Mathematics, Computer Science and Statistics, Ghent University, Gent 9000, Belgium*

[4]*Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm 17177, Sweden*

ABSTRACT

Recently, it has become common for applied works to combine commonly used survival analysis modeling methods, such as the multivariable Cox model and propensity score weighting, with the intention of forming a doubly robust estimator of an exposure effect hazard ratio that is unbiased in large samples when either the Cox model or the propensity score model is correctly specified. This combination does not, in general, produce a doubly robust estimator, even after regression standardization, when there is truly a causal effect. We demonstrate via simulation this lack of double robustness for the semiparametric Cox model, the Weibull proportional hazards model, and a simple proportional hazards flexible parametric model, with both the latter models fit via maximum likelihood. We provide a novel proof that the combination of propensity score weighting and a proportional hazards survival model, fit either via full or partial likelihood, is consistent under the null of no causal effect of the exposure on the outcome under particular censoring mechanisms if either the propensity score or the outcome model is correctly specified and contains all confounders. Given our results suggesting that double robustness only exists under the null, we outline 2 simple alternative estimators that are doubly robust for the survival difference at a given time point (in the above sense), provided the censoring mechanism can be correctly modeled, and one doubly robust method of estimation for the full survival curve. We provide R code to use these estimators for estimation and inference in the supporting information.

Keywords: causal inference; Cox model; double robustness; inverse probability of treatment weighting; parametric proportional hazards

Event History Regression with Pseudo-Observations: Computational Approaches and Causal Inference

Michael C Sachs[1], Erin E Gabriel[1]

Section of Biostatistics and the Pioneer Centre for SMARTbiomed, University of Copenhagen

ABSTRACT

Due to tradition and ease of estimation, the vast majority of clinical and epidemiological papers with time-to-event data report hazard ratios from Cox proportional hazards regression models. Although hazard ratios are well known, they can be difficult to interpret, particularly as causal contrasts, in many settings. Nonparametric or fully parametric estimators allow for the direct estimation of more easily causally interpretable estimands such as the cumulative incidence and restricted mean survival. However, modeling these quantities as functions of covariates is limited to a few categorical covariates with nonparametric estimators, and often requires simulation or numeric integration with parametric estimators. Combining pseudo-observations based on non-parametric estimands with parametric regression on the pseudo-observations allows for the best of these two approaches and has many nice properties. In this talk, we review a user friendly, easy to understand way of doing event history regression for the cumulative incidence and the restricted mean survival, using the pseudo-observation framework for estimation. The method uses the well known formulation of a generalized linear model and allows for extensions to double-robust estimation.

Keywords: survival analysis; competing risks; pseudo observations; regression; causal inference

Back to Sessions List

Transporting Evidence from and to External Studies by Leveraging Aggregate Data

Ying Sheng[1], Yifei Sun[2], **Chiung-Yu Huang[3]**

[1]*State Key Laboratory of Mathematical Sciences, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China*

[2]*Department of Biostatistics, Mailman School of Public Health, Columbia University, New York, USA*

[3]*Department of Epidemiology & Biostatistics, University of California at San Francisco, California, USA*

ABSTRACT

With the increasing availability of data in the public domain, there has been a growing interest in exploiting information from various sources to facilitate the decision-making processes. However, in real-world applications, particularly those dealing with sensitive areas such as healthcare and finance, individual-level data are often unavailable, leaving only aggregate data from external sources. In this talk, I will demonstrate how one can leverage the external aggregate data to improve the estimation efficiency in smaller-scale studies via the empirical likelihood framework. This approach can accommodate the heterogeneity and uncertainty in external information simultaneously. Under a similar framework, I will also introduce a novel approach for transporting evidence from clinical studies to target populations using only covariate summary statistics to account for distributional shifts and uncertainty in the external information. Conditions to ensure the validity of the proposed estimators will be examined.

Keywords: Covariate shift; Entropy balancing; Exponential tilting; Meta-Analysis; Transportability

Back to Sessions List