PRECISION MEDICINE METHODOLOGY DEVELOPMENT WITH APPLICATION TO SURVIVAL AND GENOMICS DATA

Hunyong Cho

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biostatistics in the Gillings School of Global Public Health.

Chapel Hill
2021

Approved by:

Michael R. Kosorok

Di Wu

Donglin Zeng

Stephen R. Cole

Michael I. Love

ABSTRACT

Hunyong Cho: Precision medicine methodology development
with application to survival and genomics data
(Under the direction of Michael R. Kosorok and Di Wu)

Precision medicine and genomics data provide chances for better decision making in the public health domain. In this dissertation, we develop some important elements of precision medicine and address some aspects of genomics data.

The first element is developing a nonparametric regression method for interval censored data. We develop a method called Interval Censored Recursive Forests (ICRF), an iterative random forest survival estimator for interval censored data. This method solves the splitting bias problem in tree-based methods for censored data. For this task, we develop consistent splitting rules and employ a recursion technique. This estimator is uniformly consistent and shows high prediction accuracy in simulations and data analyses.

Second, we develop an estimator of the optimal dynamic treatment regime (DTR) for survival outcomes with dependent censoring. When one wants to maximize the survival time or the survival probability of cancer patients who go through multiple rounds of chemotherapies, finding the dynamic optimal treatment regime is complicated by the incompleteness of the survival information. Some patients may drop out or face failure before going through all the preplanned treatment stages, which results in a different number of treatment stages for different patients. To address this issue, we generalize the Q-learning approach and the random survival forest framework. This new method also overcomes limitations of the existing methods—independent censoring or a strong modeling structure of the failure time. We show consistency of the value of the estimator and illustrate the performance of the method through simulations and analysis of the leukemia patient data and the national mortality data.

Third, we develop a method that measures gene-gene associations after adjusting for the dropout events in single cell RNA sequencing (scRNA-seq) data. Posing a bivariate zero-inflated negative binomial (BZINB) model, we estimate the dropout probability and measure the underlying correlation after controlling for the dropout effects. The gene-gene association measured in this way can serve as a building block of gene set testing methods. The BZINB model has a straightforward latent variable interpretation and is estimated using the EM algorithm.

To Jinyoung and my parents.

ACKNOWLEDGEMENTS

This work has been only possible with the guidance of my dissertation advisors and mentors of my life, Michael Kosorok and Di Wu. You have not only strengthened me as a researcher but have taught me how to embrace others as you have done to me. Meeting you is perhaps the largest gift that has made up for all the hardships during the pursuit of my Ph.D. degree.

The support of my committee has been precious. I thank Donglin Zeng, Michael Love, and Stephen Cole for their constructive suggestions that have enriched my work. Many thanks also go to Kimon Divaris, the PI of my GRA, who has supported me for years. Your warm smiles and brilliant ideas have always let me envision how an ideal leader would be.

I am grateful to have had opportunities to work with many excellent colleagues, Nicholas Jewell, John Preisser, Shannon Holloway, and many lab friends who made my Ph.D. life full of vigor and rigor.

My dear friends, Victor Ritter, Tamy Tsujimoto, Gilson Honvoh, and Paul Little, your company has been giving me a great vitality to my life as a rain shower does to a long drought.

My parents, you have always been dedicated to and proud of your son. I cannot thank you enough, ever.

Most importantly, I thank my love, Jinyoung, for your caring love and wisdom. Talking, laughing, eating, and sitting together, my life has been happier than ever. Love you.

Chapel Hill, North Carolina

June 1, 2021

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AFT	Accelerated Failure Time
AIPWE	Augmented Inverse Probability Weighting Estimator
BNB	Bivariate Negative Binomial
BNP	Bayesian Non-Parametric model
BZINB	Bivariate Zero-Inflated Negative Binomial
CART	Classification and Regression Trees
CI	Conditional Independence
CP	Coverage Probability
DTR	Dynamic Treatment Regime
EM	Expectation Maximiation
EM-ICM	Expectation Maximiation - Iterative Convex Minorant
ERT	Extremely Randomized Trees
GLR	Generalized Log-Rank test
GWRS	Generalized Wilcoxon's Rank-Sum test
ICRF	Interval Censored Recursive Forests
IMSE	Integrated Mean Squared Error
ITR	Individualized Treatment Rule
LR	Log-Rank test
MI	Mutual Information
MSE	Mean Squared Error
NB	Negative Binomial
NLMS	National Longitudinal Mortality Study
NP	Non-Parametric
NPMLE	Non-Parametric Maximum Likelihood Estimator
PC	Pearson Correlation
PCR	Polymerase chain reaction

PH	Proportional Hazards
RIST	Recursively Imputed Survival Trees by Zhu and Kosorok (2012)
RF	Random Forest
RSF	Random Survival Forest
scRNA-seq	single cell RNA-sequencing
SLR	Log-Rank-Score test
SMLE	Smoothed Maximum Likelihood Estimator
SWRS	Wilcoxon's Rank-Sum-Score test
UMI	Unique Molecular Identifier
WRS	Wilcoxon's Rank-Sum test
ZINB	Zero-Inflated Negative Binomial

INTRODUCTION

A notable paradigm shift from traditional medicine to precision medicine has brought a considerable impact on public health. This new paradigm has been made official through President Barack Obama's Precision Medicine Initiative in 2015 that aims to foster the related research. The increasing availability of patient information, including genetic features, provides essential ingredients for the growth of precision medicine. Although the machinery for storing, delivering, and processing the growing amount of data with complexity has been rapidly improved, there are a large number of method development problems that still need to be addressed. In this manuscript, we develop some important elements of precision medicine and genomics data analysis.

We outline three topics covered in this dissertation and introduce each of them in more detail with a literature review in the next section. The first two topics are precision medicine method developments in the survival data context. For these topics, we develop random forest methods suitable for 1) interval censored data problems and 2) multi-stage dynamic treatment regime estimation. The third topic is the genomics analysis method development for identifying functional pathways. For this, we develop a bivariate zero-inflated negative binomial model to account for the dropout events in single cell RNA sequencing (scRNA-seq) data.

In the first chapter, we develop a method called Interval Censored Recursive Forests (ICRF), an iterative random-forest-based regression estimator designed for interval censored survival data. This method handles the splitting bias problem prevalent in the existing tree-based methods. It recursively updates the survival estimates in a self-consistent way to minimize the bias. We generalize the log-rank and Wilcoxon rank-sum test for interval censored data and show that they are consistent. Their improved performances are illustrated through simulations. The

convergence of the self-consistency algorithm is managed via the out-of-bag (OOB) error monitoring, and kernel-smoothing is further applied. The ICRF is uniformly consistent, and both simulations and applications to the avalanche and national mortality data indicate its high prediction accuracy. For reproducible research, we developed an `R` package `icrf` available on CRAN, and the code for simulations and data analyses is available on `https://github.com/Hunyong/icrf`.

The second chapter focuses on the multi-stage dynamic treatment regime estimation problem for right censored data, where censoring is dependent on the failure time. We develop a nonparametric dynamic treatment regime estimator that maximizes either mean survival time or the survival probability at a certain time point. In these data, the number of treatments varies among patients due to censoring or failure before the terminal stage. This missing data issue complicates the application of the standard Q-learning approach. The existing methods that address the issue still have limitations, such as independent censoring and other strong modeling assumptions. To address the issue and the limitations of the existing methods, we propose a generalized random survival forest approach that extends the standard Q-learning framework. While standard Q-learning optimizes a scalar form of outcomes such as mean survival time, this new extension optimizes stochastic processes—the whole survival curve. Consistency of the estimator is shown using the empirical process theory. We show relatively high expected values of the estimated regime than the existing methods in many settings of the simulations and in the leukemia data analysis. An `R` package `dtrSurv` was developed and is available on CRAN, and the code for simulations and leukemia data analysis is available on `https://github.com/Hunyong/survQlearn`.

In the third chapter, we develop a bivariate count model with zero-inflation. Measuring gene-gene dependence in scRNA-seq count data is often of interest, as it can help identify gene sets and pathways. However, it remains challenging because an unidentified portion of the zero counts or "dropouts," caused by technical limitation of the sequencing procedures, may unduly moderate the degree of dependence. As a consequence, conventional statistical methods

that fail to account for the dropouts provide incorrect measures of dependence. To address this problem, we propose a bivariate zero-inflated negative binomial (BZINB) model, which is constructed using a bivariate Poisson-gamma mixture with dropout indicators for the excess zeros. Estimation is based on the expectation-maximization (EM) algorithm, and the underlying dependence is measured after decomposing the two sources of zeros—zeros before dropouts and the dropouts. This model has a simple latent variable interpretation, and its computation is feasible with large-scale data. Using a recent scRNA-seq dataset, we illustrate model fitting and compare the naive and proposed dependence measures. An `R` package `bzinb` is available on CRAN.

The rest of the dissertation is organized as the following. In the literature review part, we review the literature relevant to each of the three topics. In Chapters 1–3, we discuss the main content of topics 1 (ICRF), 2 (dynamic treatment regime estimator for right censored data), and 3 (BZINB) in order. Chapter 4 discusses future research directions of the three areas, followed by technical details of each section and references.

LITERATURE REVIEW

This part of the dissertation provides existing literature relevant to the methodologies developed in Chapters 1, 2, and 3.

Interval Censored Recursive Forests

In this section, we review interval censoring and tree-based methods for censored data and study the splitting bias problem in tree-based methods and how they were addressed in the literature.

Interval censoring is a widely observed censoring mechanism in survival analysis. In interval censored data, the failure time information is given as an interval that is known to contain the failure time. In that sense, right censored data, where the failure time is either exactly observed or is known to be later than a certain censoring time, is a special case of interval censored data. However, since interval censored data, in its narrow definition, are not given as exact failure times, analysis of such data is often challenging and unique. For instance, while the Kaplan-Meier estimator designed for right censored data has a closed-form solution, its counterpart for interval censored data, or the nonparametric maximum likelihood estimator (NPMLE), does not have a closed-form solution (Huang and Wellner, 1997). Also, estimation of the survival probability for interval censored data has been a challenge until recently. This can be seen from the fact that the algorithm for the NPMLE with a convergence guarantee was only developed in the early 1990's (Groeneboom and Wellner, 1992; Jongbloed, 1998). Particular challenges arise in current status data (also known as case-I censored data) where the survival status of a subject is inspected at a single random monitoring time, thus yielding an extreme form of interval censoring.

For censored data, tree-based methods have been widely used (Zhou and McArdle, 2015). Survival trees recursively partition data into two parts until they form small, homogeneous subgroups ('the terminal nodes') and estimate the marginal survival probabilities for each terminal node (Gordon and Olshen (1985), Segal (1988), Ciampi et al. (1991), LeBlanc and Crowley (1992), and LeBlanc and Crowley (1993)). The partitioning procedure is usually done by exhaustively examining the degree of heterogeneity at all possible cutoffs for each variable and selecting the cutoff that maximizes heterogeneity. Trees stop partitioning when the terminal nodes become smaller than a pre-defined size or when further splitting does not bring enough reduction in heterogeneity. As trees grow deeper, trees tend to have less bias and, as a trade-off, gain more variability.

Random survival forests are constructed by averaging a large number of diverse survival trees (Hothorn et al. (2004), Hothorn et al. (2005), Ishwaran et al. (2008), and Zhu and Kosorok (2012)). (Survival) trees in random (survival) forests are usually grown in full depth in an effort to minimize the bias. The individually highly variable trees, however, become much less variable when aggregated, thanks to the diversity across different trees. Diversity is induced by randomizations such as subsampling, random variable selection for splitting, and random cutoff selection (Geurts et al., 2006; Mentch and Zhou, 2020). That is, multiple trees are built 1) based on resampling of the data with or without replacement, 2) by only considering a random subset of variables for splitting at each node, and/or 3) by evaluating the heterogeneity of the daughter nodes at only a random subset of the cutoff values for each splitting variable. For example, in Geurts et al. (2006)'s extremely randomized trees (ERT), which is a generic algorithm and is applicable to the survival context, multiple trees are generated without resampling but by selecting a random subset of variables and one arbitrary cutoff point for each variable at each node. For a comprehensive review about survival trees and random survival forests, see Bou-Hamad et al. (2011), Ishwaran and Lu (2014), and Zhou and McArdle (2015).

One of the characteristic features of the tree-based methods in survival analysis is the way of incorporating censored information into measuring heterogeneity. While Classification

And Regression Trees (CARTs, Breiman et al. (1984)) and Random Forests (Breiman, 2001), designed for continuous outcomes, use mean squared error (MSE) for quantifying heterogeneity, in right-censored survival tree methods, alternative approaches such as the log-rank statistic (Ishwaran et al., 2008) and inverse probability weighting (Molinaro et al., 2004; Steingrimsson et al., 2019) are used.

Using the log-rank statistic, however, can cause bias for two reasons. First, the log-rank statistic assumes that censoring time is independent of failure time. In practice, censoring is often informative of failure time. Thus, when the independent censoring assumption is violated, survival trees built based on the log-rank statistic may not be able to identify the optimal partition. Second, even when censoring is independent of failure, the log-rank statistic does not account for heterogeneity within each daughter node. In other words, the log-rank statistic implicitly assumes that subjects within a daughter node share the same marginal hazards process over censored intervals, when in fact, they may have different hazard processes conditional on their covariate values. This discrepancy contradicts and, as a result, possibly undermines the purpose of the random survival forests—estimation of the covariate-conditional hazards. Thus, naive use of the log-rank statistics could incur significant bias by choosing sub-optimal partitions.

Most existing survival trees are subject to this bias. Often, at each node of survival trees, censored data are supplemented with information borrowed from the marginal survival probability of the node. By using such crude marginal information, however, the heterogeneity of individuals is not sufficiently accounted for, especially in the early phases of tree partitioning. Although as trees grow toward their terminal nodes, they utilize more covariate-conditional information and eventually form a finer partition, it is probable that the early stages of partitioning that utilize insufficient information might adversely affect subsequent splits resulting in potential bias. Thus, utilizing covariate-conditional survival probabilities for censored data from the beginning of the partitioning procedure is essential for reducing potential bias.

Zhu and Kosorok (2012) provided an intuitive solution to this problem by proposing recursively imputed survival trees (RIST) for right-censored data. The main idea is to guess

the censored failure time using conditional survival probabilities and to utilize it for splitting. Considering that the finest covariate-conditional survival probabilities are available only after the trees grow far enough towards their terminal nodes, they use a recursion technique so that the terminal node prediction is utilized to impute the censored subjects in the next iteration of the forest building process.

This issue, however, has yet to be fully addressed in the interval censored data literature. Moreover, tree-based regression methods for interval censored data are sparse; there are only a few tree-based methods available. Yin et al. (2002) and Fu and Simonoff (2017) developed tree models for interval censored data that use the likelihood ratio test and a modified log-rank test as a splitting criterion, respectively. Yao et al. (2019) recently extended the work of Fu and Simonoff (2017) to an ensemble method. Yang et al. (2021) proposed a survival tree method for current status data that applies the idea of censoring unbiased transformation (Steingrimsson et al., 2019). However, these methods have the aforementioned limitation of insufficient usage of covariate-conditional information.

To respond to this issue, we propose a tree-based nonparametric regression method for interval censored survival data in Chapter 1. This new estimator, which solves the splitting bias problem through recursion, is consistent for the true survival probability and shows high prediction accuracies in simulations and data analyses.

Dynamic Treatment Regime Estimation for Survival Outcomes

In this section, we review literature about dynamic treatment regime estimation, with a focus on survival outcomes. Multi-stage treatments are becoming more prevalent in medicine. Chronic conditions such as cancer, auto-immune disease, HIV, and heart disease, often require multiple treatments over a long period. For instance, cancer patients may receive multiple rounds of chemo- and/or biological therapies (Habermann et al., 2006; Huang et al., 2020). Auto-immune disease patients might receive additional treatments in response to relapses (Edwards and Cambridge, 2006; Hogan and Radhakrishnan, 2013). Traditionally such chronic diseases

have been treated following a one-size-fits-most approach by which the standard of care is selected as the treatment or treatment plan that will likely benefit most patients with a similar condition, i.e., the treatment that is optimal only in the sense of the marginal effects.

However, in recent years, a more personalized approach to the treatment of chronic diseases has become of interest. Under this approach, patients benefit from treatment rules that incorporate the patient heterogeneity of treatment effects, and thereby, treatment decisions are based on the individual characteristics of each patient. Dynamic treatment regimes formalize this data-driven approach to treatment by optimizing the overall outcome of interest and providing treatment suggestions that are based on the patient's information available at each point of the decision making process (Murphy, 2003; Kosorok and Moodie, 2015; Kosorok and Laber, 2019; Tsiatis et al., 2019). In part because they utilize all available information, such optimized dynamic treatment regimes are often found to be more beneficial than a traditional set of fixed treatment plans(Kidwell, 2015). Further, because optimal dynamic treatment regimes focus on the overall or long-term outcome, they are more advantageous than applying multiple single-stage treatment rules that only optimize the stage-level outcomes.

Finding an optimal dynamic treatment regime is a reinforcement learning problem (Sutton and Barto, 2018), wherein an action by an agent, or a physician, changes both the immediate reward and the environment of the next stage and the learner, or the statistician, searches for the rule of actions that brings the most overall reward. Like many statistical learning methods, reinforcement learning problems are often solved using the Q-learning algorithm (Watkins, 1989; Zhu et al., 2015; Zhao et al., 2011; Zhang et al., 2017; Qian and Murphy, 2011), which defines the value as the expected cumulative sum of discounted rewards and optimizes the set of rules using backward recursion. Next, we review a little more details of the Q-learning approach.

Review of Q-learning

In the precision medicine domain, multiple methods have also been developed using Q-learning (Murphy et al., 2007; Zhao et al., 2009, 2011; Schulte et al., 2014), where the best policy is

chosen so that under the best policy the accumulated reward given the initial state is maximized. The solution to reinforcement learning problem can be succinctly expressed as the solution to the Bellman equation (Bellman, 1966):

$$\pi_t^*(\boldsymbol{s}_t, \boldsymbol{a}_{t-1}) = \arg\max_{a_t} E[R_t + Q_{t+1}^*(\boldsymbol{S}_{t+1}, \boldsymbol{A}_t)|\boldsymbol{S}_t = \boldsymbol{s}_t, \boldsymbol{A}_t = \boldsymbol{a}_t],$$

where $Q_{t+1}^*(\boldsymbol{s}_{t+1}, \boldsymbol{a}_t) = E[\sum_{i=t+1}^{T} R_i|\boldsymbol{S}_{t+1} = \boldsymbol{s}_{t+1}, \boldsymbol{A}_t = \boldsymbol{a}_t]$ is the quality function, (S_t, A_t, R_t) denotes a tuple of state, action, and reward at stage t, bold symbols denote the vector of the corresponding history up to stage t, and the value function is defined as $V_t^\pi(s_t) = \int_{a_t} Q_t(s_t, a_t)\pi(a_t|s_t)da_t$. Q-learning solves the Bellman equations by estimating the quality functions and optimizing the policy using backward recursion. The policy estimates are given by $\hat{\pi}_t(\boldsymbol{s}_t, \boldsymbol{a}_{t-1}) = \arg\max_{a_t} \hat{Q}_t(\boldsymbol{s}_t, (\boldsymbol{a}_{t-1}, a_t))$ for $t = T, T-1, ..., 1$.

Q-learning in the survival analysis context

In precision medicine applications, the rewards are often defined in the form of survival outcomes, such as mean survival time or survival probabilities, which pose unique challenges. For survival outcomes, the failure time is often censored, and any approach should take into account the missing information. Several methods have been developed for this purpose, for example, redistribution of the failure probability as in the Kaplan-Meier Efron (1967) and Robertson and Uppuluri (1984) or inverse probability weighting (Robins et al., 1994; Robins and Rotnitzky, 1992; Wahed and Tsiatis, 2006; Orellana et al., 2010). Further, the treatment timing may not be fixed and may even depend on a patient's status or previous treatments having an association with the failure time. In such scenarios, the three events at each stage are failure, proceeding to the next treatment, and censoring, which can be restated in a competing risks framework, in which the three dependent events compete and only the earliest event is observed. However, the full joint distribution of the events is not identifiable without knowledge of their dependence structure (Tsiatis, 1975). Thus, the dependency between the events needs to be carefully considered. For example, a cancer patient whose physical condition has improved has a longer expected survival time and at the same time may want to take another round of

chemotherapy earlier than expected. If this is the case, the survival time and the treatment time have a negative association.

There are two additional challenges that arise in the analysis of survival outcomes that must be addressed. First, the total number of treatments received may vary among patients. For example, some patients may experience failure or be censored before taking all planned treatments. And finally, there could be different optimization criteria for survival times. For some patients and clinicians, the average survival time is of primary interest, while the six-month survival probability may be of greater concern for others. A method that can address different types of endpoints would be useful in practice.

Several methods regarding the estimation of dynamic treatment regimes for survival outcomes have been proposed in the literature. Goldberg and Kosorok (2012) first developed a dynamic treatment regime estimator for survival outcomes using the Q-learning framework. To address the problems introduced by a different number of treatments as well as by censoring, the authors proposed modifying the survival data so that each observation has the same number of treatment stages without missing values and showed that the value of a policy in the original problem can be represented as an expectation of the modified data. After the data modification, the problem was put into a standard Q-learning framework using inverse probability weighting to account for censoring.

To be more specific, in the original problem the outcome of interest is the sum $\sum_{j=1}^{\bar{T}} R_j$ of segment lengths up to the stage $\bar{T}$ that contains the failure time, $S_t = (Z_t, R_{t-1})$ is the state at stage t, and Z_t and A_t are the covariate and the treatment at the beginning of the stage for $t = 1, 2, ..., T$. The goal is to find a treatment rule that maximizes $E[\sum_{j=1}^{\bar{T}} R_j \wedge \tau]$ where τ is the study length. Here censoring time is denoted as C and is assumed independent of both covariates and event time. In the auxiliary problem, $S'_t = (Z'_t, R'_{t-1})$ is the state at stage t, $Z'_t = Z_t$, $R'_t = R_t$, and $A'_t = A_t$ for each t except that $Z'_t = \emptyset$, A'_t is a random draw from the treatment space $\mathcal{A}$ for stages t after the stage experiences the failure. Further modification is done to handle a situation where $\sum_{i=1}^{t} R'_i > \tau$. The main idea of all this translation into the

auxiliary problem is to put the irregular problem into a well-defined format without adding or losing additional information. Then instead of maximizing $E[\sum_{j=1}^{\tilde{T}} R_j \wedge \tau]$, one could set the value function as $V_\pi(s_1) = E[\sum_{j=1}^{T} R_j]$. This formulation enables solving the optimization by the Q-learning framework.

However, in this method, censoring was assumed to be completely independent of all covariates and event times. The estimated Q-function and the resulting decision rules would be thus subject to bias for data with dependent censoring.

Huang et al. (2014) addressed a similar problem using backward recursion. Their motivating problem was a recurrent disease clinical trial where patients receive an initial treatment followed by a salvage treatment if patients experience either treatment resistance or relapse. Simoneau et al. (2019) extended the dynamic treatment regime estimator for continuous outcomes via weighted least squares (Wallace and Moodie, 2015) into the censored time-to-event data setting. Both of these methods use the accelerated failure time model for the failure time distribution, which carries a risk of model misspecification. Thus, a more flexible model is needed to allow for more relaxed distributional assumptions. Wahed and Thall (2013) proposed a dynamic treatment regime estimator using a full specification of the likelihood and Xu et al. (2016) developed a Bayesian alternative, where the disease progression dynamic was modeled using a dependent Dirichlet process prior with Gaussian process measure. While the latter approach provides a more flexible modeling framework than the former, the class of regimes for both methods is refined to a fixed number of pre-defined treatment sequences.

Each of the existing methods described above has its limitations. First, not all methods allow for a flexible number of treatment stages and/or treatment levels. For example, the methods in Wahed and Thall (2013) and Huang et al. (2014) were designed for a two-stage decision problem, and the Simoneau et al. (2019) approach is restricted to binary treatments. Second, the outcomes of interest in most methods are limited to the mean survival time. Jiang et al. (2017) proposed an optimal dynamic treatment regime estimator that maximizes the t-time survival probability. However, none of the methods permit choosing the criterion except Wahed

and Thall (2013). Third, parametric models or strong structural assumptions are employed for all of these methods, with the exception of Goldberg and Kosorok (2012). Strong modeling assumptions, such as the accelerated failure time model or the proportional hazards models used by the existing methods, are subject to model-misspecification, may limit the policy class, and, consequently, may invoke loss in the value of the optimal policy. Finally, the censoring assumptions of Goldberg and Kosorok (2012) are restrictive. Censoring is associated with failure time in many medical applications, but it is often reasonable to assume independence of censoring after conditioning on patient historical information. See Table 0.1 for comparison of the methods.

Table 0.1: Assumptions of the existing and the proposed methods. Q, the number of stages; $|\mathcal{A}^{(q)}|$, the number of treatment arms at stage q; criterion, the target value to be optimized; T, failure time model; C, censoring assumption; $\{\boldsymbol{\pi}\}$, policy class; 2+, can be generalized to more than two; NP, non-parametric; AFT, accelerated failure time; PH, proportional hazards; BNP, Bayesian non-parametric; CI, conditional independence; Ind, independence; fixed, a fixed number of distinct treatment rules.

method	Q	$\lvert\mathcal{A}^{(q)}\rvert$	criterion	T	C	$\{\boldsymbol{\pi}\}$
the new method	finite	finite	$E[T \wedge \tau], S(t)$	NP	CI	flexible
Goldberg and Kosorok (2012)	finite	finite	$E[T \wedge \tau]$	NP	Ind	flexible
Huang et al. (2014)	2+	2+	$E[T]$	AFT	CI	linear
Jiang et al. (2017)	finite	2	$S(t)$	PH	CI	linear
Simoneau et al. (2019)	finite	2	$E[T]$	AFT	CI	linear
Wahed and Thall (2013)	finite	finite	$E[T]$, $S(t)$	AFT	CI	fixed
Xu et al. (2016)	finite	finite	$E[T]$	BNP	CI	fixed

In Chapter 2, we develop a general dynamic treatment regime estimator for censored time-to-failure outcomes that addresses all of the limitations discussed for the existing estimators. This method is nonparametric, allows covariate-independent censoring, and optimizes either the mean truncated survival time or the survival probability at a given time point.

Bivariate Zero-Inflated Negative Binomial Model

This section reviews existing methods that handle the dropout problem of single cell RNA sequencing data. We first describe why dropouts matter, outline two general approaches, and further study existing bivariate count models.

Single cell RNA sequencing (scRNA-seq) is a high throughput sequencing technology that profiles gene expression at a cell's resolution (Kolodziejczyk et al., 2015). This is in contrast to bulk RNA sequencing (RNA-seq), where a group of cells are sequenced altogether and consequently no cell-level information is available in data. As a price for the cell-level resolution, scRNA-seq loses some information by the so-called "dropout" phenomenon; during the sequencing steps (and the capturing steps) of scRNA-seq, a large amount of RNAs are undetected. Consequently, the observed count data include a greater number of zeros than would be expected given the number of molecules sequenced and our *a priori* knowledge of transcription rates at individual loci (Risso et al., 2018; Hicks et al., 2017; Huang et al., 2018). That is, an expressed gene in a cell might be recorded as zero due to low transcriptome capture and sequencing efficiency (Huang et al., 2018). In contrast, in a bulk RNA-seq, excess zeros are less frequently observed (Hicks et al., 2017). For these reasons, negative binomial models have been extensively used for bulk RNA-seq data (Love et al., 2014; Robinson et al., 2010), and ZINB models are typically used for scRNA-seq data (van den Berge et al., 2018; Risso et al., 2018).

There is a growing amount of literature that many scRNAseq data are not zero-inflated, and dropout events are primarily caused by PCR amplification that could be removed by the unique molecular identifiers (UMI) technique (Vieth et al., 2017; Townes et al., 2019; Svensson, 2020). While a good amount of comfort is available that there is no zero-inflation in the data for the droplet-based data such as 10X that uses UMI quantification, there is still a need to address dropouts in other platform-based scRNA-seq data as well as single cell proteomics and metatranscriptomics data. We show an example of zero-inflated scRNA-seq data in Section 3.4.

Statistical inferences at both individual gene level (Iacono et al., 2019; Yu, 2018) and gene set level, e.g., pathways, can be misleading without considering the excess zeros caused by dropouts. Inference of gene-gene dependence, e.g., the correlation-based method, has been widely used in pathway analysis of bulk RNA-seq data (Zhang and Horvath, 2005) and in recent scRNAseq data analyses (Iacono et al., 2019; Yu, 2018; Pont et al., 2019; Van Dijk et al., 2018; Eraslan et al., 2019). However, the conventional Pearson correlation of two genes with significant dropouts in the scRNAseq may not properly reflect the underlying gene-gene dependence.

For example, a pair of genes, of which expressions are highly correlated without dropouts, would have an attenuated correlation, based on the observed data, when only one of them have a large amount of dropouts. On the other hand, a pair of uncorrelated genes would have higher correlation, when both genes have dropouts in a substantial portion of the sample. The systematic bias will not vanish without adjusting for the effects of the dropout events, regardless of what dependence measure is used. This includes Pearson correlation, $PC(X,Y) = Cov(X,Y)/\sqrt{Var(X)Var(Y)}$, and mutual information, $MI(X,Y) := \int\int_{\mathcal{X}\times\mathcal{Y}} f(x,y) \log \frac{f(x,y)}{f(x)f(y)} dxdy$ (Mc Mahon et al., 2014; Chan et al., 2017).

Two strategies have been considered to address the bias in scRNA-seq data. Imputation methods (Li and Li (2018); Eraslan et al. (2019); Peng et al. (2019)) aim to provide expression levels free of the excess zeros by imputing them. While imputation methods are versatile in that they provide ready-to-use data, they are not deterministic, having different results for every implementation. The other strategy is estimation of the count distribution. Once having obtained information about the distribution of the expressions before dropouts, one can do downstream analyses such as measuring the dependence of the before-dropout expressions. Models such as SAVER (Huang et al., 2018) and DESCEND (Wang et al., 2018a) have been proposed to estimate the count distribution of scRNA-seq data in order to recover either dropouts or lower-than-expected expression levels. For example, correlations can be calculated from SAVER-recovered genes in unique molecule index (UMI)-based DropSeq scRNA-seq data

where its result is close to that measured from the "gold standard" RNA fluorescence in situ hybridization (FISH) (Huang et al., 2018). However, many of the methods taking this approach focus on modeling marginal distributions and they do not explicitly posit dependence structure between two genes.

Next we review a variety of existing bivariate models that fit bivariate zero-inflated count data with overdispersion: bivariate Poisson mixture models (Gurmu and Elder, 1999; Famoye, 2010; Jørgensen, 1987), bivariate generalized Poisson models (Famoye and Consul, 1995) and copula models (Cameron et al., 2004). These models can be further extended to flexibly accommodate excess zeros by introducing zero-inflation parameters or composing hurdle models. For a comprehensive survey of bivariate count models, refer to Cameron and Trivedi (2013) and Chou and Steenhard (2011).

Of a plethora of the proposed models in the literature, many of the bivariate Poisson mixture models and bivariate generalized Poisson models take overly complicated forms; they do not have simple marginal distributions (e.g., GBIVARNB model in Gurmu and Elder (1999)), their parameters are hard to interpret and/or computationally expensive to estimate. Copula-based bivariate models can be alternatives to the mixture models, but they depend on the underlying copula models, and the interpretation can be quite complicated.

Many of the existing bivariate negative binomial models are mostly designed specifically for modeling marginal means rather than pairwise dependence. For example, Gurmu and Elder (1999) discussed a bivariate negative binomial distribution (BIVARNB), but their model is specified by only four parameters, which may not provide sufficient flexibility to delineate diverse distributional structure. For a bivariate joint distribution, four parameters are needed to specify the first two marginal moments of each of the two independent variables, while another parameter is needed solely for modeling the dependence. Later on, Wang (2003) applied BIVARNB to a zero-inflated BIVARNB regression setting. In this model, zero-inflation is dictated by a single parameter, implying that when one variable either drops out or not, the other variable behaves exactly the same, which may not be the case for scRNA-seq data; one gene

can drop out, while the other does not. Instead, it is possible to have three free parameters for full joint zero-inflation probability (Li et al., 1999).

To overcome such limitations, we develop a new model called Bivariate Zero-Inflated Negative Binomial (BZINB) model that has a simple latent variable interpretation. With this model, which has the capability of modeling the dropout probability, the underlying gene-gene correlation can be measured. We further discuss this new development in Chapter 3.

CHAPTER 1: INTERVAL CENSORED RECURSIVE FORESTS

1.1 Introduction

In this chapter, we respond to the splitting bias problem raised in the literature review. We propose a tree-based nonparametric regression method for interval censored survival data. The method uses a recursion strategy (Zhu and Kosorok, 2012) which incorporates a self-consistency equation (Efron, 1967). In addition, we address additional challenges inherent to interval censored data: first, the self-consistency algorithm may not identify the global optimum for interval censored data, and second, the interval censored data are highly noisy. To overcome such additional concerns, the method is equipped with a convergence monitoring procedure over recursions, probabilistic provision of information rather than imputation, and smoothing along the time domain. The proposed method shows high prediction accuracy both on simulated data and on our illustrative examples of avalanche victims data (Haegeli et al., 2011; Jewell and Emerson, 2013) and national mortality data (Sorlie et al., 1995). An `R` package `icrf` is available on CRAN.

The rest of this chapter is organized as follows. In Section 1.2, we describe the data structure and modeling assumptions. In Section 1.3, the proposed methods are introduced and discussed in context. The uniform consistency of the method is derived in Section 1.4. The predictive accuracy of the proposed method is evaluated using simulations and analysis of two sets of data in Sections 1.5 and 1.6, respectively. In Section 1.7, we discuss choice of the modeling hyper-parameters.

1.2 Data setup and model

The proposed method is applicable to interval censored data that include right-censored and current status data as special cases. Current status data, also also known as case-I censoring, only include survival status of a subject inspected at a single random monitoring time. The event time, T, is only known to lie within an interval $I \equiv (L, R]$, where $L = T^-$ and $R = T$ for an exactly observed T. Let $F(t)$, $F(t|X)$, $F(t|I)$, and $F(t|X, I)$ denote the marginal, covariate-conditional, the interval-conditional, and the full-conditional distributions at time t, respectively, where $X \equiv (X_1, ..., X_p) \in \mathcal{X} \subset \mathbb{R}^p$ is a p-dimensional covariate with distribution function $F_X(\cdot)$. We use $S \equiv 1 - F$ to represent a corresponding (conditional) survival function. For the censoring mechanism, we consider covariate-conditional non-informative censoring which is defined as,

$$\Pr(T < t|L = l, R = r, L < T \leq R, X) = \Pr(T < t|l < T \leq r, X).$$

This implies that intervals do not provide any further information than the fact that the failure time lies in the interval given the covariate (Oller et al., 2004; Sun, 2007). The study length is denoted by $\tau < \infty$. A random vector $U = (U_1, U_2, ..., U_M)$ denotes the monitoring times at each element of which the survival status of the subject is identified. U follows a distribution F_U with maximum potential number of follow-up times $M > 0$. Among the M monitoring times, only one pair of two neighboring time points that includes T contributes to the likelihood. Thus we only consider $\{L, R\} = \{U_{(m)}, U_{(m+1)} : U_{(m)} < T \leq U_{(m+1)}, m = 0, 1, ..., M\}$ in the data analysis, where $U_{(m)}$ denotes the mth order statistic of the elements of U with $U_0 \equiv 0$ and $U_{(M+1)} \equiv \infty$. Current status data correspond to $M = 1$.

1.3 Interval censored recursive forests

1.3.1 Overview of the proposed method

We adopt the recursion strategy for interval censoring and address the challenges of interval censoring—higher noise and non-identifiability of self-consistency algorithm—by carrying the full conditional survival probabilities of censored subjects, employing kernel smoothing of the survival curves along time, and monitoring convergence over recursion.

We outline the high level idea of the proposed method before we give a detailed description in the following subsections. As an initial step, to provide rough information about the censored intervals, we estimate the marginal survival curve, $S^{(0)}(t|X) = \hat{S}(t)$, and obtain the estimate of the full conditional survival probability for each subject, $S(t|X_i, I_i)$ by projection. Instead of doing imputation as in RIST, we store the conditional probability information for each subject and use it in the splitting tests. In this way, we can avoid the Monte Carlo error resulting from the imputation procedures which can be significant for interval censored data. We develop the Generalized Wilcoxon's Rank Sum (GWRS) test and Generalized Log Rank (GLR) test that enable two-sample testing for interval censored data based on conditional probabilities. With one of those splitting rules selected, a predefined number of trees are built under a modified ERT algorithm. Unlike the original ERT algorithm, we subsample data to leave a small fraction ('the out-of-bag sample') of the data for later use. At each terminal node of the trees, a local survival probability estimate is obtained in two ways: 1) the NPMLE of the survival curve is obtained based on raw interval data without using the survival curve information, or 2) the full conditional survival curves are averaged. We call the former a "quasi-honest" approach, and the latter an "exploitative" approach. The tree survival probability estimates formed in this manner are averaged to obtain a forest survival probability estimate, $S^{(1)}(t|X)$, for the first iteration. Then $S^{(k-1)}(t|X)$ is used to update the full conditional survival curve of each subject $S^{(k)}(t|X_i, I_i)$ at the kth iteration, $k = 2, 3, ..., K$. For each k, $\tilde{S}^{(k-1)}(t|X)$ is obtained

by kernel-smoothing. The final prediction is then given by the smoothed survival curve at the iteration of the smallest out-of-bag error. A detailed pseudo-algorithm is given in Algorithm 1.

Result: $\tilde{S}(t|X) = \tilde{S}^{(k_{\text{opt}})}(t|X)$ where $k_{\text{opt}} = \arg\min_k \epsilon^{(k)}$;

initialize $S^{(0)}(t|X)$ and kernel smooth ($\tilde{S}^{(0)}(t|X)$), if INITIAL_SMOOTH is TRUE;

for k (*forest iteration*) $= 1, 2, ..., K$ **do**

Update $S^{(k-1)}(t|X_i, I_i)$ based on $S^{(k-1)}(t|X_i)$ and I_i for each i;

for b (*tree construction*) $= 1, 2, ..., K$ **do**

Sample $\mathbb{D}_b$ of size $s = \lceil 0.95n \rceil$ from the dataset $\mathbb{D}$ ($\mathbb{D}_b^{OOB} := \mathbb{D} \backslash \mathbb{D}_b$);

Recursively partitioning using GWRS based on $\{S^{(k-1)}(t|X_i)\}$:

At each node, randomly pick $\lceil \sqrt{p} \rceil$ variables, pick a random cut-off for each selected variable, and find the optimal cut-off suggested by GWRS;

if *QUASIHONEST* **then**

$S_{b,l}^{(k)}(t|A_{b,l}) = \text{NPMLE}(\{I_i : X_i \in A_{b,l}\})$;

else

$S_{b,l}^{(k)}(t|A_{b,l}) = \frac{1}{|A_{b,l}|} \sum_{X_i \in A_{b,l}} S^{(k-1)}(t|X_i, I_i)$;

Kernel smoothing: $\tilde{S}_{b,l}^{(k)} = \text{KERNELSMOOTH}(S_{b,l}^{k})$;

The conditional survival function for the tree:

$S_b^{(k)}(t|X) = \sum_{l=1}^{L_b} S_{b,l}^{(k)}(t|A_{b,l}) 1(X \in A_{b,l})$,

$\tilde{S}_b^{(k)}(t|X) = \sum_{l=1}^{L_b} \tilde{S}_{b,l}^{(k)}(t|A_{b,l}) 1(X \in A_{b,l})$;

The out-of-bag error for the tree: $\epsilon_b^{(k)} = IMSE(\tilde{S}_b^{(k)}, \mathbb{D}_b^{OOB})$;

Obtain the conditional survival function for the forest:

$S^{(k)}(t|X) = \frac{1}{n_{\text{tree}}} \sum_{b=1}^{n_{\text{tree}}} S_b^{(k)}(t|X)$, $\tilde{S}^{(k)}(t|X) = \frac{1}{n_{\text{tree}}} \sum_{b=1}^{n_{\text{tree}}} \tilde{S}_b^{(k)}(t|X)$;

Calculate the out-of-bag error for the forest: $\epsilon^{(k)} = \frac{1}{B} \sum_{b=1}^{B} \epsilon_b^{(k)}$;

Algorithm 1: Pseudo-algorithm for ICRF

1.3.2 Splitting rules

For right-censored data, Peto and Peto (1972) compared the two-sample test statistics including the Wilcoxon Rank Sum (WRS) test and the log-rank test. They showed that the

log-rank test is the most locally powerful test under Lehman-type alternative hypotheses while WRS also has strong power under Log-normal mean-shift alternative hypotheses. Thus, these tests can be considered as potential splitting rules with some modifications for interval censoring.

For interval censored data, we develop two splitting rules by extending the WRS and log-rank tests. We also consider two existing score tests proposed by Peto and Peto (1972) that are used by existing tree-based methods (Fu and Simonoff, 2017; Yao et al., 2019). Below we describe the four splitting rules and show the consistency property of the newly developed rules. Simulation results in Section 1.5.3.2 show that our developed rules have on average better performance than existing alternatives.

A. Generalized Wilcoxon's Rank Sum test (GWRS). The WRS test statistic,

$$\tilde{W}_n = \frac{1}{n_1 n_2} \sum_{i \in G_1} \sum_{j \in G_2} \xi(T_{1,i}, T_{2,j}),$$

estimates $\tilde{\theta} = \Pr(T_1 < T_2) + \frac{1}{2}\Pr(T_1 = T_2)$ where T_l is the survival time of a randomly chosen subject in group G_l, $\xi(T_{1,i}, T_{2,j}) = 1(T_{1,i} < T_{2,j}) + \frac{1}{2}1(T_{1,i} = T_{2,j})$, and n_1 and n_2 are the sample sizes of the two groups, respectively. The estimand can be alternatively expressed as $\tilde{\theta}(S) = 1 + \int_0^\infty \check{S}_{G_1}(t) dS_{G_2}(t)$, where $S_{G_l}(t) = \Pr(T_l > t | G_l), l = 1, 2$, is the marginal survival probability of the lth group and $\check{S}(t) = \frac{1}{2}S(t) + \frac{1}{2}S(t^-)$ where half of the probability mass in the left continuity point is shifted toward the right. In the presence of administrative censoring, $\mathring{W}_n = \frac{1}{n_1 n_2} \sum_{i \in G_1} \sum_{j \in G_2} \xi(\mathring{T}_{1,i}, \mathring{T}_{2,j})$ estimates $\theta(S) = \Pr(\mathring{T}_{1,i} < \mathring{T}_{2,j}) + \frac{1}{2}\Pr(\mathring{T}_{1,i} = \mathring{T}_{2,j}) = 1 + \int_0^\tau \check{S}_{G_1}(t) dS_{G_2}(t) - \frac{1}{2} S_{G_1}(\tau) dS_{G_2}(\tau)$, where $\mathring{T}_{l,i} = T_{l,i} \wedge \tau, l = 1, 2$.

We then generalize this statistic to allow non-informative interval censoring as follows:

$$W_n(S) = \frac{1}{n_1 n_2} \sum_{i \in G_1} \sum_{j \in G_2} \zeta(I_{1,i}, I_{2,j} | X_{1,i}, X_{2,j}; S),$$

where $\zeta(I_{1,i}, I_{2,j} | X_{1,i}, X_{2,j}; S) = \Pr(\mathring{T}_{1,i} < \mathring{T}_{2,j} | T_{1,i} \in I_{1,i}, T_{2,j} \in I_{2,j}, X_{1,i}, X_{2,j}; S) + \frac{1}{2}\Pr(\mathring{T}_{1,i} = \mathring{T}_{2,j} | T_{1,i} \in I_{1,i}, T_{2,j} \in I_{2,j}, X_{1,i}, X_{2,j}; S)$. Note $\zeta(I_{1,i}, I_{2,j} | X_{1,i}, X_{2,j}; S) = 1 + \int_0^\tau \check{S}(t | I_{1,i}, X_{1,i})$ $dS(t | I_{2,j}, X_{2,j}) - \frac{1}{2} S(\tau | I_{1,i}, X_{1,i}) S(\tau | I_{2,j}, X_{2,j})$.

By the following theorem, the GWRS statistic, $W_n(S_n)$, is shown to be consistent for $\theta(S_0)$, for a sequence S_n converging to the true survival function S_0. The proof of Theorem 1.1 is deferred to the Technical Details and the consistency conditions of S_n are provided in Theorem 1.3.

Theorem 1.1. *For a fixed pair of sets $G_l \subset \mathcal{X}$ and any sequence S_n such that $\sup_{t\in[0,\tau],x\in\mathcal{X}} |S_n(t|x) - S_0(t|x)| \to 0$ in probability as $n \to \infty$, $W_n(S_n) \to \theta(S_0)$ in probability as $n \to \infty$.*

B. Generalized Log-Rank test (GLR). The log-rank test statistic for uncensored data or right-censored data is given by

$$\widetilde{LR}_n = \frac{\sum_{j=1}^{J} \frac{Y_{2j}D_{1j}+Y_{1j}D_{2j}}{Y_{\cdot j}}}{\sqrt{\sum_{j=1}^{J} \frac{Y_{1j}Y_{2j}D_{\cdot j}(Y_{\cdot j}-D_{\cdot j})}{Y_{\cdot j}^2(Y_{\cdot j}-1)}}},$$

where J is the number of distinct observed time points, $Y_{l,j}$ and $D_{l,j}$ are the number of subjects at risk right before and the number of events at the jth time point in group l, respectively, for $l = 1, 2$; $Y_{\cdot j} = Y_{1,j} + Y_{2,j}$ and $D_{\cdot j} = D_{1,j} + D_{2,j}$.

Using the full-conditional survival probabilities $S_i(t) \equiv S(t|X_i, I_i)$, the log-rank test can be extended to a generalized log-rank test (GLR) for interval censored data:

$$LR_n(S) = \frac{\int_0^\tau \frac{Y_2(t;S)dN_1(t;S)+Y_1(t;S)dN_2(t;S)}{Y(t;S)}}{\sqrt{\int_0^\tau \frac{Y_1(t;S)Y_2(t;S)dN(t;S)(Y(t;S)-dN(t;S))}{Y(t;S)^3}}},$$

where $Y_l(t;S) = \frac{1}{n_l}\sum_{i\in G_l} S_i(t-)$, $N_l(t) = 1-\frac{1}{n_l}\sum_{i\in G_{n_l}} S_i(t), l = 1, 2, Y(t;S) = \lambda_{n,1}Y_1(t;S)+ \lambda_{n,2}Y_2(t;S)$, $N(t;S) = \lambda_{n,1}N_1(t;S) + \lambda_{n,2}N_2(t;S)$, $\lambda_{n,l} = \frac{n_l}{n}$, and $n = n_1 + n_2$. Note that the statistic $LR_n(S)$ is $\sqrt{n\lambda_{n,1}\lambda_{n,2}}$ times smaller in scale than $\widetilde{LR}_n$.

The following theorem establishes consistency of the GLR for

$$\rho(S_0) = -\frac{\int_0^\tau \frac{S_0(t-|G_2)dS_0(t|G_1)+S_0(t-|G_1)dS_0(t|G_2)}{S_0(t-|G_1\cup G_2)}}{\sqrt{-\int_0^\tau \frac{S_0(t-|G_1)S_0(t-|G_2)S(t|G_1\cup G_2)dS(t|G_1\cup G_2)}{S_0^3(t-|G_1\cup G_2)}}}$$

for some disjoint subsets $G_l \subset \mathcal{X}, l = 1, 2$, with the proof relegated to Technical Details.

Theorem 1.2. *For a fixed pair of sets $G_l \subset \mathcal{X}$ and any sequence S_n such that* $\sup_{t\in[0,\tau],x\in\mathcal{X}} |S_n(t|x) - S_0(t|x)| \to 0$ *in probability as* $n \to \infty$, $LR_n(S_n) \to \rho(S_0)$ *in probability as* $n \to \infty$.

C. WRS-score test (SWRS). Peto and Peto (1972) introduced asymptotic score statistics for interval censored data, one of which is the two sample WRS test. The test statistic is given by $\widetilde{SW}_n = \frac{1}{n_1}\sum_{i\in G_1} SW_{1,i} - \frac{1}{n_2}\sum_{i\in G_2} SW_{2,i}$, where $SW_{l,i} = \hat{S}_{G_l}(L_{l,i}) + \hat{S}_{G_l}(R_{l,i}) - 1$. To rely on the self-consistency scheme the test statistic is rewritten as $SW_n(S) = \frac{1}{n_1}\sum_{i\in G_1} SW_{1,i}(S) - \frac{1}{n_2}\sum_{i\in G_2} SW_{2,i}(S)$ with $SW_{l,i}(S) = S(L_{l,i}|X_{l,i}) + S(R_{l,i}|X_{l,i}) - 1,\ l = 1, 2.$

D. Log-Rank-score test (SLR). Another score statistic (SLR) based on the log-rank test was proposed by Peto and Peto (1972). This statistic, under the self-consistency algorithm, can be written as $SLR_n(S) = \frac{1}{n_1}\sum_{i\in G_1} SLR_{1,i}(S) - \frac{1}{n_2}\sum_{i\in G_2} SLR_{2,i}(S)$, where

$$SLR_{l,i}(S) = \begin{cases} \frac{S(L_{l,i}|X_{l,i})\log S(L_{l,i}|X_{l,i}) - S(R_{l,i}|X_{l,i})\log S(R_{l,i}|X_{l,i})}{S(L_{l,i}|X_{l,i}) - S(R_{l,i}|X_{l,i})} & S(L_{l,i}|X_{l,i}) > S(R_{l,i}|X_{l,i}), \\ \log S(L_{l,i}|X_{l,i}) + 1 & S(L_{l,i}|X_{l,i}) = S(R_{l,i}|X_{l,i}). \end{cases}$$

The best cut point is the one that maximizes $|W_n - \frac{1}{2}|$, LR_n, $|SW_n|$, or $|SLR_n|$. We use GWRS as our main splitting rule in the subsequent analyses. In Section 1.5.3.2, we illustrate how different splitting rules affect the prediction accuracy.

1.3.3 Self-consistent random forest and convergence monitoring

The proposed ICRF can be understood as a self-consistent estimator. The self-consistency algorithm (Efron, 1967) can be succinctly expressed as a solution to the equation $f(\cdot;\theta) = \mathbb{P}_n f(\cdot|\mathcal{Z};\theta)$, where $\mathbb{P}_n$ is the empirical average operator with respect to random quantities denoted as script letters, $\mathcal{Z}$ is the observed data, and $f(\cdot;\theta)$ is a functional parameter of interest. For instance, the non-parametric maximum likelihood estimator (NPMLE) for interval censored data is a self-consistent estimator for the marginal survival probability that solves for S in $S(t) = \mathbb{P}_n S(t|\mathcal{I})$, where $\mathcal{I}$ is the observed intervals.

This algorithm can also be extended to tree-based estimators for survival probabilities. Without the self-consistency scheme, survival forest estimators can generally be written as

$$\hat{S}(t|x) = \frac{1}{n_{\text{tree}}} \sum_{b=1}^{n_{\text{tree}}} \mathbb{P}_n \left[S_1(t|A_b(x; S_2)) \frac{1(\mathcal{X} \in A_b(x; S_2))}{|A_b(x; S_2)|/n} \right],$$

where $A_b(x)$ is the terminal node of the bth tree that contains x, $|A|$ is the sample size of node A, $S_1(\cdot|A_b)$ is the survival probability estimate of the terminal node A_b, S_2 is the survival probability that is used to support splitting decisions in trees, and the subscripts indicating the dependencies with the tree index b and the sample index n are suppressed in S_1 and S_2. Note that S_2 is needed for tree partitioning, only when the failure time is censored. If there is no censoring, survival forest estimators can be reduced to $\hat{S}(t|x) = \frac{1}{n_{\text{tree}}} \sum_{b=1}^{n_{\text{tree}}} \mathbb{P}_n \left[S_1(t|A_b(x)) \frac{1(\mathcal{X} \in A_b(x))}{|A_b(x)|/n} \right]$. Without censoring, the self-consistency of random survival forests can be achieved under certain smoothness assumptions by replacing $\hat{S}$ and S_1 with S and incorporating an appropriate splitting rule. Splitting rules bring consistency to tree or random forest estimators if every terminal node of the resulting tree partition has an arbitrarily small length in probability for every side that contains signal and at the same time has arbitrarily many sample points, as the sample size grows larger (Cui et al., 2017). Random splitting (Wager and Walther, 2015; Wager and Athey, 2018) is often used for theoretical purposes, as is assumed in Theorem 1.3, instead of the greedy splitting rules (Breiman, 2001). However, since this chapter is primarily directed at heuristic approaches based on the self-consistency concept, we will retain use of our modified ERT algorithm for splitting in simulations and data analyses.

Different survival tree methods assume disparate S_2 in the literature. For example, the marginal survival probability estimate $\hat{S}(t)$ is used in Fu and Simonoff (2017) and Yao et al. (2019) and a node marginal survival probability estimate $\hat{S}(t|A)$ is used in Ishwaran et al. (2008) and Yin et al. (2002). Note that most existing tree-based survival estimators have three survival quantities, $\hat{S}(t|x), S_1$ and S_2, that do not coincide with each other and thus, they are not self-consistent. This discrepancy between survival probabilities may cause a greater bias. Splitting based on crude information, e.g., using the marginal survival probability estimate $\hat{S}(t)$

or the intermediate node survival probability estimate $\hat{S}(t|A)$ as S_2 rather than using $S(t|x)$, results in greater finite sample bias. See the discussion of Cui et al. (2017), where the authors discuss the bias of random survival forests for which the splitting rule is based on the candidate node marginal survival probabilities.

Self-consistency can be derived by replacing $\hat{S}(t|x), S_1$ and S_2 with S. The ICRF estimator $\hat{S}$ solves for S in

$$S(t|x) = \frac{1}{n_{\text{tree}}} \sum_{b=1}^{n_{\text{tree}}} \mathbb{P}_n S(t|\mathcal{I}, \mathcal{X} \in A_b(x; S)) \frac{1(\mathcal{X} \in A_b(x; S))}{|A_b(x; S)|/n}. \tag{3.5.1}$$

The self-consistency equation can be solved by recursion. This self-consistent estimator makes sense when $S^{(k)}(t|x) \simeq S(t|x)$ for some large k. However, self-consistency equations in general may have multiple solutions (non-identifiability) and thus, recursion algorithms may not guarantee convergence to the truth; for example, This issue arises when estimating the NPMLE for interval censoring (Wellner and Zhan, 1997). For some initial guesses, the estimator may give an inconsistent estimate. Thus, it is crucial to make sure that an additional forest iteration brings reduction in error. To monitor this in the absence of knowing the true survival curve, the out-of-bag samples are used for estimating the accuracy. That is, for each tree in ERT, we randomly subsample a large fraction, e.g. 95%, for tree construction and evaluate the tree using the small (5%) hold-out sample. Using a metric that will be discussed in Section 1.5.2, we monitor the performance of the ERT's over a prespecified number of iterations, e.g. $K = 10$.

1.3.4 Quasi-honesty

Once partitioning procedures are done, the terminal node survival curves are estimated either i) by applying NPMLE to the raw interval data (*quasi-honest* prediction) or ii) by averaging the full conditional survival curves (*exploitative* prediction). The former approach is quasi-honest, as the survival probability of the previous iteration is only used in the partitioning procedure but not in the prediction procedure. It is not genuine honesty (Athey and Imbens, 2016), in the sense that ICRF still uses the same interval data in both partitioning and terminal node prediction.

The second approach is exploitative. This approach is computationally efficient, since the prediction does not require a complicated optimization procedures, it is computationally light. However, as is discussed in the following paragraphs, this approach tends to have higher bias, non-convergence, and dilution of signals. RIST, where imputed values containing the information about the covariate-conditional survival curve are used for both partitioning and terminal node prediciton, is hence exploitative.

The role of (quasi-) honesty in the prediction accuracy should be understood in terms of the bias-variance trade-off. While honesty induces higher variability by not utilizing the whole information at each procedure, it relaxes the overfitting problem and makes trees less biased by maintaining less dependence between the partitioning procedure and the terminal node prediction procedure. Hence, quasi-honesty may or may not be beneficial to interval censored survival analysis. A large amount of information about the true survival curve is lost due to interval censoring. This means that there might be a room for an exploitative approach to make up the information loss, since it more fully utilizes the information. However, it is also true that once the estimation moves in a wrong direction initially, then the exploitative approach may keep driving the estimation sequence in the wrong direction, while the quasi-honest approach may suffer less from such non-convergence.

Another property of the exploitative approach is dilution of signal. When the initial survival probability starts with the marginal survival distribution, even after partitioning, two different points in a feature space share a significant amount of information about the survival distribution. This results in lower variance and hence, sometimes, underfitting. This exploitative approach should therefore be used when the features do not contain a large amount of information about the failure time distribution. We compare the performances of these two approaches in Sections 1.5 and 1.6.

1.3.5 Smoothed forests

Random forests are relatively smoother than base learners with respect to features. However, they are still discrete in the time domain, especially for the NPMLE of interval censored data. Since in reality the survival function is unlikely to include step functions, it can be beneficial to assume some smoothness on the true survival function. Groeneboom et al. (2010) proposed two ways of estimating smooth survival curves for current status data. Although their first method, the maximum smoothed likelihood estimator (MSLE), may not apply to general interval censored data, one can easily use the second method, the smoothed maximum likelihood estimator (SMLE), for such data. The idea is to find a non-smooth nonparametric maximum likelihood estimator (NPMLE), $\hat{S}(t)$, and use kernel smoothing to obtain an SMLE: $\tilde{S}(t) = 1 + \int_0^t \int_{\mathbb{R}_+} \frac{1}{h} k_h(s-u) d\hat{S}(u) ds$, where k_h is a kernel function with bandwidth $h > 0$. For survival forests, the SMLE is computed for each terminal node of each tree: $\tilde{S}^{k,b}(t|x) = \sum_{l=1}^{L_{k,b}} \tilde{S}_l^{k,b}(t|A_l^{k,b}) 1(x \in A_l^{k,b})$, where $A_l^{k,b}$ is the lth terminal node in the bth tree of the kth forest iteration, $l = 1, 2, ..., L^{k,b}$, $b = 1, 2, ..., n_{\text{tree}}$, and $k = 1, 2, ..., K$. Then the smoothed random survival forest is $\tilde{S}^k(t|x) = \sum_{b=1}^{n_{\text{tree}}} \tilde{S}^{k,b}(t|x)$. In this chapter we use a Gaussian kernel with bandwidth $h = c{n_{\min}}^{-1/5}$ where we choose c to be the inter-quartile range of the marginal survival distribution estimate and $n_{\min}$ is the minimum size of the terminal nodes. For discussion on the choice of the bandwidth, see Groeneboom et al. (2010). For the boundary kernel near $t = 0$, we use a mirror kernel $\tilde{k}_h(t, u) = k_h(t, |u|)$ for $t \leq 4h$.

1.4 Uniform Consistency of ICRF

Although the recursion technique is intended for bias correction for finite samples, the large sample behaviour of ICRF is of interest. We present in Theorem 1.3 a uniform consistency result for the quasi-honest ICRF. The proof is provided in the Technical Details. We only consider case-II censoring, since this result can be generalized without much difficulty to case-K censoring for $K < \infty$ (Huang and Wellner, 1997).

Assumption 1.1 (Absolutely continuous measure)**.** *The probability measure of the failure time, T, is absolutely continuous with respect to that of the monitoring times (L, R). Specifically, the joint density of the monitoring times is positive ($g(l, r|x) > 0$), if $0 < S_0(r|x) < S_0(l|x) < 1$, for all $x \in \mathcal{X}$, where S_0 is the true survival probability.*

Assumption 1.2 (Lipschitz continuity of the failure and censoring survivor functions)**.** *There exist constants L_S and L_G such that $|S_0(t \mid x_1) - S_0(t \mid x_2)| \leq L_S \|x_1 - x_2\|_1$ and $|G(t_1, t_2 \mid x_1) - G(t_1, t_2 \mid x_2)| \leq L_G \|x_1 - x_2\|_1$ for all $x_1, x_2 \in \mathcal{X}$ and $t, t_1, t_2 \in [0, \tau]$, where G is the censoring survival distribution and g is its derivative with respect to time.*

Assumption 1.3 (Weakly dependent covariate values)**.** *The covariate space $\mathcal{X}$ is a p-dimensional unit hypercube, i.e., $X \in \mathcal{X} = [0, 1]^p$. X has a density f_X such that $\zeta^{-1} \leq f_X(x) \leq \zeta$ for all $x \in \mathcal{X}$ and some constant $\zeta \geq 1$.*

Assumption 1.4 (α-regular and random-split trees)**.** *Trees in the ICRF are random-split and α-regular according to Definitions 3 and 4 of Wager and Athey (2018).*

Assumption 1.5 (Terminal node size)**.** *The minimum size $n_{\min}$ of the terminal nodes in the ICRF trees grows at the following rate:*

$$n_{\min} \asymp n^{\beta}, \frac{1}{2} < \beta < 1,$$

where $a \asymp b$ implies both $a = \mathcal{O}(b)$ and $b = \mathcal{O}(a)$.

Theorem 1.3 (The uniform consistency of interval censored recursive forests)**.** *Suppose Assumptions 1.1–1.3 hold. Then the interval censored recursive forest $\hat{S}_n$ built based on Assumptions 1.4–1.5 and quasi-honesty is uniformly consistent. That is,*

$$\sup_{t \in [0,\tau], x \in \mathcal{X}} |\hat{S}_n(t \mid x) - S_0(t \mid x)| \to 0$$

in probability as $n \to \infty$.

1.5 Simulations

In this section, we run simulations in order to evaluate the prediction accuracy of ICRF in multiple aspects. We also discuss the computational cost of the method. The first set of simulations is to compare the prediction accuracy of ICRF to that of existing methods under multiple scenarios. The second set of simulations is to compare the performances of different splitting rules of ICRF and to compare the performances of quasi-honest and exploitative prediction rules. The final set shows the performance as sample size grows.

The competitors considered include the Cox proportional hazards model (Finkelstein, 1986) which is implemented using the R package `icenReg` (Anderson-Bergman, 2017b), the survival tree method for interval censored data (Fu and Simonoff, 2017), and the survival forest method for interval censored data (Yao et al., 2019).

All the models except the Cox model are implemented using an R package `icrf` (version 2.0.0). Note that since ICRF estimates are a weighted average of NPMLE's and the method of Yao et al. (2019) provides an NPMLE of weighted individuals, implementation of the latter by `icrf` might involve finite sample differences. Because for other methods than ICRF, the estimates are not identifiable at each time point but are uniquely obtained only as a set of probability masses in intervals, we interpolate the within-interval survival curve assuming a uniform density within those intervals. However, when the length of the intervals is not finite, an exponential density is assumed. That is, given the estimated probability $\hat{p}_{[a,\infty)}$ of the last unbounded interval, the interpolated survival estimate is given by $\hat{S}(t) = 1 - \hat{p}_{[a,\infty)}^{-t/a}$ for $a \leq t < \infty$. The NPMLE often assigns a probability mass to the last bounded time interval even when there are subjects known to have failure times in an unbounded interval. This can deflate survival curves in the tail drastically. Such non-regularity can be relaxed by posing structures to the estimator (Anderson-Bergman and Yu, 2016; Polyanskiy and Wu, 2020). In those cases, a correction is made so that the last probability mass is allocated exponentially over an unbounded interval. We further include smoothed versions of the existing methods for a fair comparison.

1.5.1 Generative models and tuning parameters

We first define the simulation settings by describing generative models and tuning parameters for the estimators. The basic framework for the generative models is largely taken from Zhu and Kosorok (2012).

Generative models. Six scenarios for two different monitoring times ($K = 1$ and $K = 3$) are studied. Scenario 1 (PH-L) assumes a proportional hazards model with linear hazards ratio, Scenario 2 (PH-NL) has a nonlinear hazards ratio (PH-NL) in place of Scenario 1, and the third (non-PH) is a non-proportional hazards model, where all three scenarios assume non-informative censoring. The fourth scenario (CNIC) has non-informative censoring conditional on X, and the fifth scenario (IC) has informative censoring. To further study how smoothed estimators behave under a non-smooth true survival curve, we further adopted Scenario 6 (non-SM), where the first scenario is modified so that the density of the event times is degenerate. The settings are defined more concretely in Table 1.2. The sample size n of the training sets is 300 and samples are independently drawn. The study period (τ) is set to 5 for all scenarios.

scenario	X	T	U_k	μ	ρ
1 PH-L	$N_{25}(0, \Sigma(\rho))$	$Exp(\mu)$	$Exp(\bar{\mu})$	$e^{0.1\sum_{j=11}^{20} X_j - 0.1}$	0.9
2 PH-NL	$U([0,1]^{10})$	$Exp(\mu)$	$U([0,\tau])$	$\sin(\pi X_1) + 2\lvert X_2 - \frac{1}{2}\rvert + X_3^3$	-
3 non-PH	$N_{25}(0, \Sigma(\rho))$	$G(\mu, 2)$	$U([0, \frac{3}{2}\tau])$	$0.5 + 0.3\lvert\sum_{j=11}^{15} X_j\rvert$	0.75
4 CNIC	$N_{25}(0, \Sigma(\rho))$	$LN(\mu)$	$LN(0.8\mu)$	$0.3\lvert\sum_{j=1}^{5} X_j\rvert + 0.3\lvert\sum_{j=21}^{25} X_j\rvert$	0.75
5 IC	$N_{10}(0, \Sigma(\rho))$	$Exp(\mu)$	$LN(T)$	$2expit(X_1 + X_2 + X_3)$	0.2
6 Non-SM	$N_{25}(0, \Sigma(\rho))$	$SDE(\mu)$	$Exp(\bar{\mu})$	$e^{0.1\sum_{j=11}^{20} X_j - 0.1}$	0.9

Table 1.2: Simulation settings; Independent samples of size $n = 300$, $X = (X_1, ..., X_P)$; $\Sigma(\rho) = \{\sigma_{ij}(\rho)\}, \sigma_{ij} = \rho^{|i-j|}$; $U = (U_{(1)}, ..., U_{(K)})$ with $K = 1, 3$ and elements U_k (conditionally) independent of each other; $N_P(\mu, \Sigma)$, the p-dimensional normal distribution with mean μ and variance Σ; $LN(\mu)$, the log-normal with mean μ and variance 1; $U(A)$, the uniform distribution over A; $Exp(\mu)$, the exponential distribution with mean μ; $G(\mu, \theta)$, the Gamma distribution with shape μ and scale θ; $SDE(\mu)$, the semi-discretized Exponential defined as $\frac{1}{2}(Exp(\mu) + \frac{1}{2}\lceil 2Exp(\mu)\rceil)$; $\bar{\mu}$, a constant near the sample average of the μ's.

Tuning parameters. The tuning parameters for the tree-based methods are summarized in Table 1.3. The minimum size of the terminal nodes is 6 for ensemble learners and 20 for the non-ensemble tree method. For ensemble learners, 300 trees are built by considering randomly

method	n_{fold}	n_{tree}	$mtry$	s	$replace$	$n_{\min}$
ICFR	10	300	$\lceil\sqrt{p}\rceil$	$\lceil 0.95n\rceil$	no	6
Fu	-	300	$\lceil\sqrt{p}\rceil$	$\lceil 0.632n\rceil$	yes	6
Yao	-	-	-	-	-	20

Table 1.3: Tuning parameters for the tree-based methods; Fu, the method of Fu and Simonoff (2017); Yao, the method of Yao et al. (2019); n_{fold}, the maximum number of iterations for ICFR; n_{tree}, the number of trees making up the random forests; $mtry$, the number of candidate features on which splitting tests are done at each node; s, the size of the random resample for a tree in random forests; $replace$, whether to resample with replacement or not; $n_{\min}$, the minimum number of observations in terminal nodes.

chosen $\lceil\sqrt{p}\rceil$ candidate variables at each node. The default splitting rule for ICRF is set as GWRS and both quasi-honest and exploitative prediction are used for terminal node predictions. However, other splitting rules are also compared. The marginal survival probability estimates are used as the initial guess. As for smoothing, the bandwidths are chosen to be $h = cn^{-1/5}$ with $c = \frac{1}{2}[\hat{S}^{-1}(0.25) - S^{-1}(0.75)]$.

1.5.2 Prediction Accuracy

To assess the prediction accuracy of the estimators, we use integrated absolute error and supremum absolute error over the study period. They are defined as $\epsilon^{INT}(\hat{S}_n) = \int_0^\tau |S_0(t) - \hat{S}_n(t)|dt$ and $\epsilon^{SUP}(\hat{S}_n) = \sup_{t\in[0,\tau]} |S_0(t) - \hat{S}_n(t)|$, respectively. These error measurements are obtainable only when the true survival curve S_0 is available. To measure the error in the absence of the true survival curve, we use the integrated mean squared errors type 1 (IMSE_1) and type 2 (IMSE_2) (Banerjee et al., 2019). IMSE_1 is defined as the squared discrepancy of the estimate from the actual survival status (observed $-$ model-predicted) averaged over the interval of the known survival status and then averaged over the sample. That is, $\text{IMSE}_1(\hat{S}_n|\mathbb{D}) = \frac{1}{n}\sum_{i=1}^n \frac{1}{\tau-(R_i\wedge\tau)+(L_i\wedge\tau)}\big\{\int_0^{L_i\wedge\tau}(1-\hat{S}_n(t|X_i))^2dt + \int_{R_i}^{R_i\wedge\tau}\hat{S}_n(t|X_i)^2dt\big\}$, where $\mathbb{D} = \{(L_1, R_1, X_1), ..., (L_n, R_n, X_n)\}$ is the test set. This can be regarded as a modified integrated Brier score (Graf et al., 1999).

IMSE_2 is defined over the whole time domain up to the study length, where the discrepancy over the censored interval is calculated by the difference between the covariate-conditional

survival curve and the full-conditional survival curve:

$$\text{IMSE}_2(\hat{S}_n|\mathbb{D}) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{\tau}\int_0^{\tau}(\hat{S}_n(t|X_i, I_i) - \hat{S}_n(t|X_i))^2 dt.$$

As mentioned in the previous section, IMSE_1 is used for convergence monitoring of ICFR, as it is a model-free measure. The out-of-bag samples are used as a test set for measuring IMSE_1. The error measurement for convergence monitoring is given by

$$\text{IMSE}_1^{ICFR}(\hat{S}_n|\mathbb{D}) = \frac{1}{n_{\text{tree}}}\sum_{b=1}^{n_{\text{tree}}}\text{IMSE}_1(\hat{S}_{n,b}|\mathbb{D}_b^{OOB}),$$

where $\mathbb{D}$ is the whole training data and $\mathbb{D}_b^{OOB}$ is the out-of-bag sample left for the bth tree.

1.5.3 Simulation results

1.5.3.1 Comparison with other methods

Simulations are done with $n_{sim} = 300$ replicates for each distinct setting. The simulation results based on quasi-honesty and GWRS rule are illustrated in Figure 1.1. The results in the left column are for Case-I censoring and those in the right column are for Case-II censoring. For convenience, we denote the ICRF estimator at the kth iteration by ICRF-k. The iteration with the best out-of-bag error among the ten iterations is denoted by A. In the results, ICRF-1, ICRF-2, ICRF-3, ICRF-5, ICRF-10, and ICRF-A are presented.

Comparison with other methods. For most of the scenarios, ICRF's have minimum or close-to-minimum integrated and supremum absolute errors. For Scenario 5 (both $M = 1, 3$), where Cox models have better integrated absolute errors than the ICRF's, ICRF's have as good supremum errors as the Cox models. Also noting that simpler models such as Fu and Simonoff (2017)'s method and the Cox models have better accuracy than the method of Yao et al. (2019) under Scenario 5, there is evidence that underfitting might be beneficial for settings where features contain weak signals, i.e., when $var(E(T|X))$ is low.

In Scenario 1, where data are generated under the proportional hazards model, ICRF's have better average accuracy than that of the Cox models. Although the Cox models eventually have higher accuracy for larger samples (see Figure 1.3), the results indicate that ICRF methods have a relatively high prediction accuracy.

Convergence monitoring. The ICRF's error rate often becomes smaller as the number of iterations increases on average. Although in general it decreases, it often fluctuates and sometimes increases. However, ICRF-A, the ICRF at the best iteration of IMSE_1 measured against the out-of-bag samples, have integrated and supremum absolute errors close to the minimums most of the time.

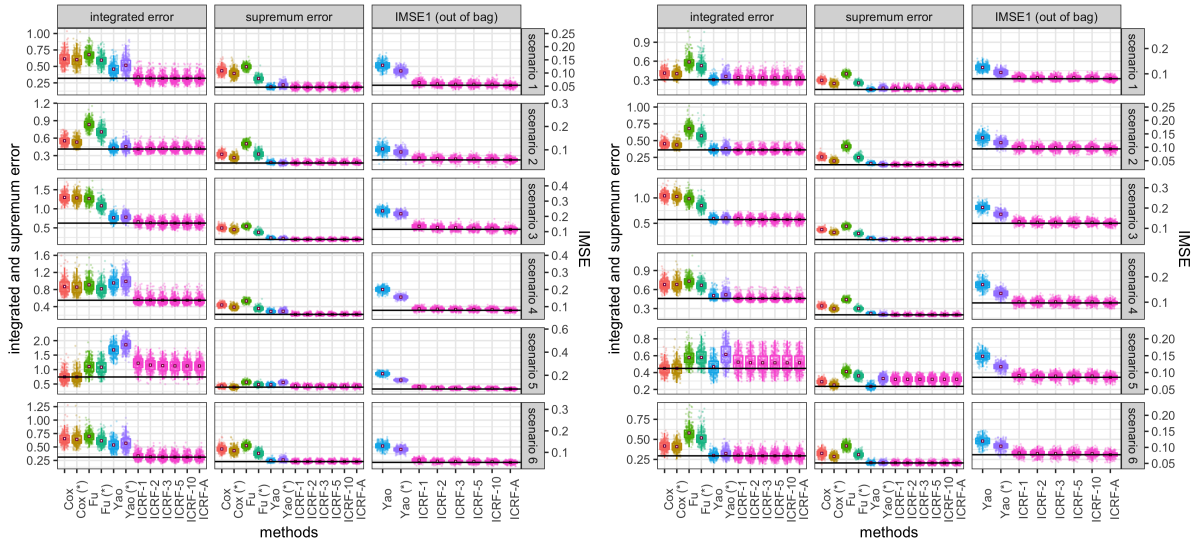

Figure 1.1: Prediction errors of methods under different simulation settings (the ICRF's are built in a quasi-honest manner); Fu, Fu and Simonoff (2017); Yao, Yao et al. (2019); (*), smoothed versions; The boxes on the left column are for case-I censoring ($M = 1$) and those on the right column are for case-II censoring ($M = 3$); For each setting, the horizontal line indicates the minimum of mean error levels of the methods.

1.5.3.2 Splitting rules and quasi-honesty

Four splitting rules (GWRS, GLR, SWRS, SLR) with quasi-honest versus exploitative predictions are compared in Figure 1.2 under six scenarios, with $M = 1$ monitoring time. Most of the time, the new splitting rules (GWRS and GLR) have on average less error than the

score-based rules (SWRS and SLR). Between GWRS and GLR, the two methods have about the same prediciton accuracy. The gap between the new splitting rules and the score-based rules might reflect the fact that score-based rules rely on approximation, while GWRS and GLR do not.

On the other hand, the comparison between quasi-honest and exploitative predictions is less consistent. One does not always beat the other. In Scenarios 2, the exploitative prediction has lower integrated absolute error, and in other scenarios, it has higher error rates. As mentioned in the last paragraph of Section 1.3.4, exploitative prediction tends to make weak contrasts between two feature values and is expected to perform well when the true distribution has faint signals. In contrast, quasi-honest prediction provides more precise estimates when the signal is strong. As can be seen Figure A1 in the Technical Details, exploitative prediction is computationally lighter than quasi-honest prediction and this difference overwhelms the difference made by different choices of the splitting rules. See the Technical Details for more discussion on computational cost.

1.5.3.3 Varying sample sizes

The prediction accuracy of each method is evaluated under different sample sizes for current status data ($M = 1$) under Scenario 1 (proportional hazards model). The integrated and supremum errors are measured. For ICRF, the last fold (10th) estimate is used for illustration. The mean, the 1st quartile, and the 3rd quartile of error measurements across 300 replicates are illustrated in Figure 1.3.

The Cox model, although it does not have the smallest errors for small sample sizes ($n = 100, 200$), has rapidly decreasing errors as the sample size grows larger for both integrated and supremum absolute errors. Among the nonparametric models, ICRF shows the highest prediction accuracy in terms of all error measures for most sample sizes. For $n = 1600$, the integrated error is lowest for the Cox model and has virtually a converged value for all ensemble-based methods.

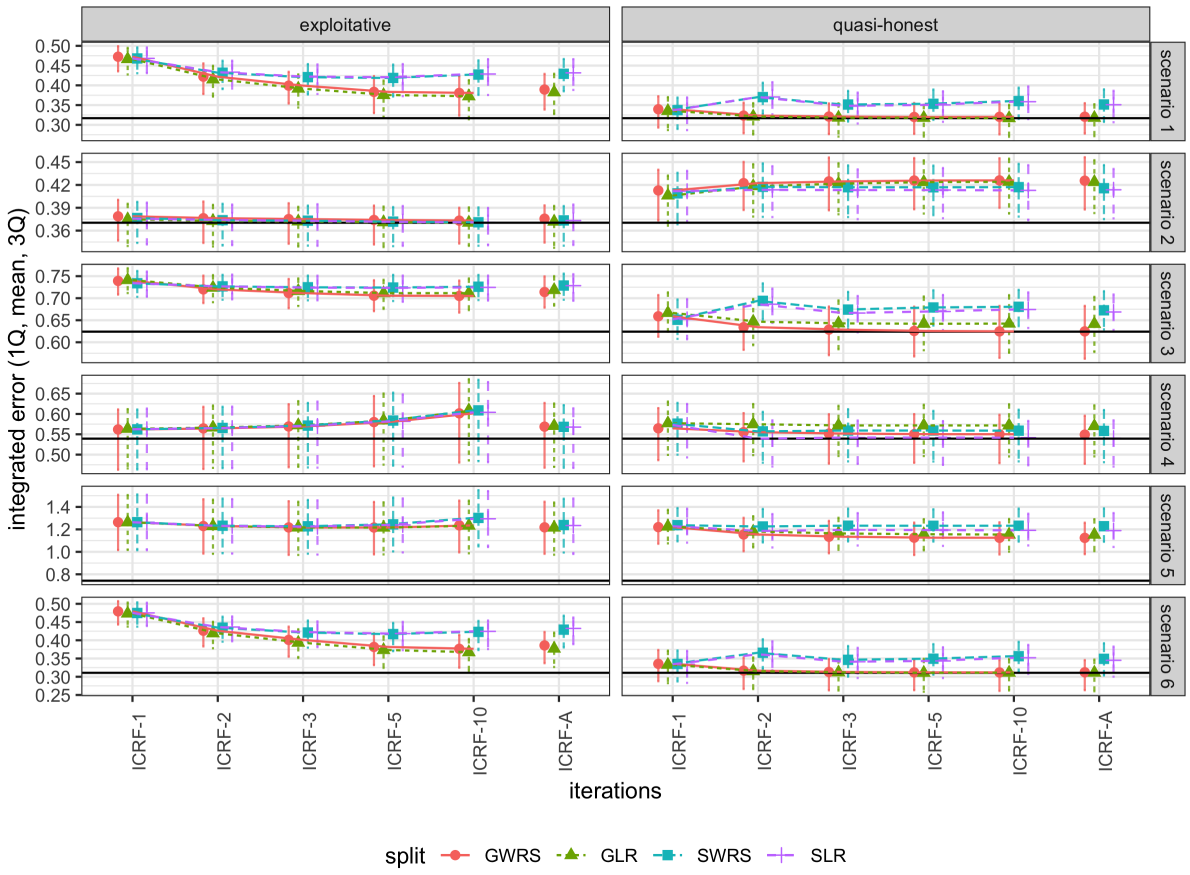

Figure 1.2: Mean and 1st and 3rd quartile ϵ^{INT} of splitting rules and prediction rules under Case-I censoring.

The computation time of ICRF increases in a mildly superlinear fashion with respect to the sample size. See Figure A2 and Web Section 4 of the Technical Details for this chapter.

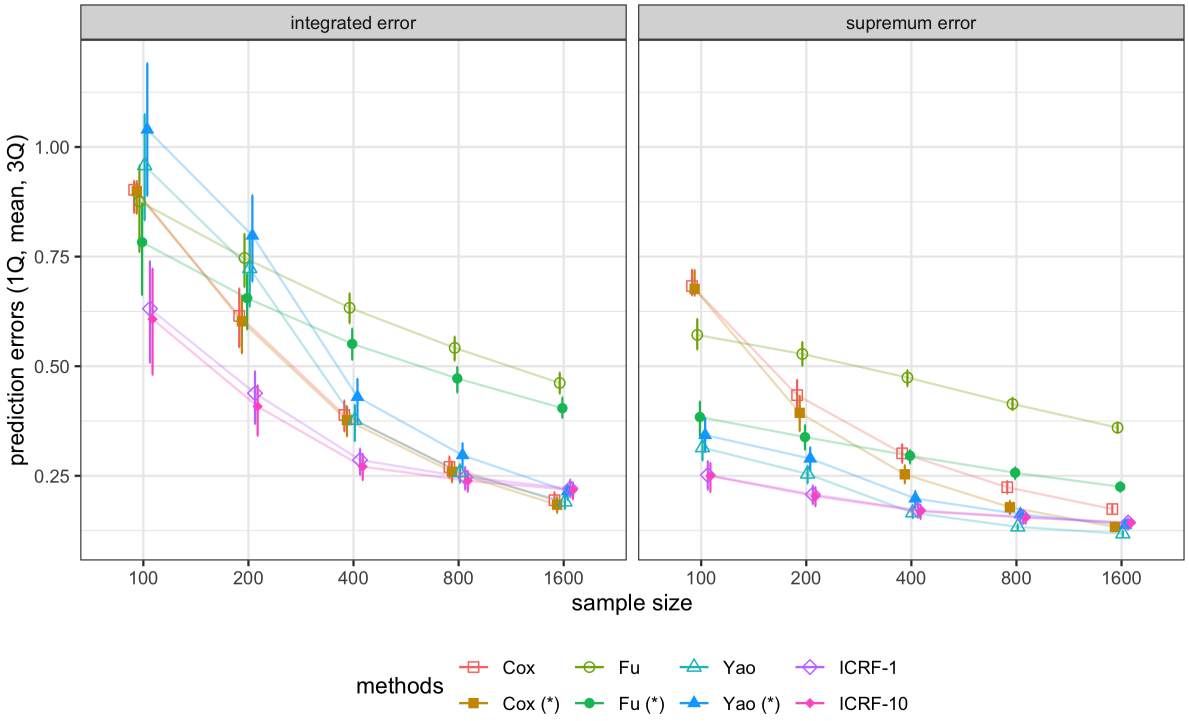

Figure 1.3: Prediction errors under different sample sizes for Scenario 1 and $K = 1$.

1.6 Data analyses

In this section, we apply ICRF (using 10 iterations), and three other methods—Fu and Simonoff (2017), Yao et al. (2019), and the Cox model—with the corresponding smoothed versions to two existing data sets: (i) avalanche victim data where the time of discovery and a victim's survival status were only observed (Jewell and Emerson (2013)), and (ii) data extracted from the National Longitudinal Mortality Study (Sorlie et al. (1995)).

1.6.1 The avalanche victims data

The data of 1,247 avalanche victims buried in Switzerland and Canada between October 1980 and September 2005 are analyzed ((Jewell and Emerson, 2013)). The dataset includes duration of burial and status of survival of the subjects, and thus can be regarded as current status data. The covariates include location, burial depth, and the type of outdoor activities involved. Approximately 10% of the observations have missing burial depth. The main quantity

of interest is the covariate-conditional survival probability where the event time is defined as time from burial to death. The event time here is counterfactual in a sense that the event time is the time until death had the person not been discovered (prior to death).

We use the following assumptions. First, burial duration is independent of the time to event. This assumption is feasible as avalanche recovery is usually performed in the absence of knowledge of the survival status of victims. Second, the missingness of burial depth is completely at random. Although this assumption may not be fully valid, we analyze the data using complete cases only for comparative purposes. Third, the survival of individual victims are independent. Since a single avalanche may involve multiple burials due to group activities, without a sufficient number of covariates, this assumption may not be valid. However, the point estimator of the survival function remains valid.

We randomly partition the complete data ($n = 1127$) into training ($n = 789$) and test ($n = 338$) datasets 300 times. The training sets are used for estimation of the survival curves, and the fitted models are evaluated using the corresponding test sets. The avalanche data is highly skewed (median = 30, mean = 2,932, 3rd quartile = 110, max = 342,720 in minutes). To make the estimation computationally feasible, a log-transformed time domain is used with a transformation $h : [0, \infty) \mapsto [0, \infty)$ where $h(t) = \log(t + 1)$, and the prediction accuracy is evaluated in the transformed time domain. The study length is set as $\tau_t = 14400$ minutes (10 days) or $\tau = \log(\tau_t + 1) = 9.58$. The analyses are implemented using the `R` package `icrf`. Preliminary parametric and semi-parametric regression analyses of the data are available in Jewell and Emerson (2013).

The prediction accuracy (IMSE) of the fitted models is summarized in Table 1.4 (LEFT). Among nonparametric methods, the ICRF with exploitative prediction has the best prediction accuracy. Although the smoothed Cox model shows the best prediction accuracy (IMSE_1 = 0.21, IMSE_2 = 0.19) among all available methods, the exploitative ICRF has a comparable performance (IMSE_1 = 0.22, IMSE_2 = 0.19).

Prediction error			Variable importance		
method	$IMSE_1$ (sd)	$IMSE_2$ (sd)	variable	quasi-honest	exploitative
ICRF (Q)	0.026 (0.0032)	0.026 (0.0038)	by $IMSE_1$ (multiplier = 0.0073)		
ICRF (E)	0.022 (0.0021)	0.019 (0.0013)	Burial depth	1.00	0.47
Fu	0.024 (0.0030)	0.027 (0.0042)	Group activity	0.17	0.41
Fu (*)	0.023 (0.0028)	0.020 (0.0032)	Location	0.16	0.24
Yao	0.025 (0.0031)	0.026 (0.0031)	by $IMSE_2$ (multiplier = 0.0041)		
Yao (*)	0.026 (0.0030)	0.026 (0.0030)	Burial depth	1.00	0.67
Cox	0.021 (0.0025)	0.019 (0.0021)	Group activity	0.27	0.46
Cox (*)	0.021 (0.0026)	0.019 (0.0022)	Location	0.55	0.27

Table 1.4: Average prediction error of the avalanche survival models for each method (LEFT) and variable importance of the ICRF model fitted on the first training set of the avalanche data (RIGHT). ICRF (Q), quasi-honest ICRF; ICRF (E), exploitative ICRF; The importance values are rescaled so that maximum values for each measure becomes 1. The multiplier is the original importance scale.

Figure 1.4 illustrates the expected truncated log survival time, $\int_0^\tau h(t)dS(h(t)) + \tau S(\tau)$, of avalanche victims estimated by each smoothed model. While the Cox model, by assumption, has a monotone expected survival time with respect to each of the covariates, nonparametric models show non-monotone curves. The expected truncated survival time curves of the two prediction rules have a significant difference in their model variability, or $var[E[T|X]]$. Quasi-honest ICRF, compared to exploitative ICRF, has a wigglier curve along burial depths and has wider gaps among different group activities.

For most models, burial depth seems to be the most important covariate. In general, the mean truncated survival time decreases as the burial depth increases. However, for the emsemble methods (ICRF, Yao et al. (2019)), the mean survival time increases for depths greater than 350 cm. This is considered to be an overfitting problem in a sparse data region. In many models, the location also plays as important a role as burial depth; In the Cox model, the mean survival time in Canada is on average smaller than in Switzerland. Unlike the Cox model, nonparametric models have different patterns of expected survival time curves for different countries.

Variable importance is formally quantified by measuring the increase in IMSE for a dataset where the values of each covariate in the original dataset are randomly permuted across the sample. The permutation is outside of the tree building procedure and does not affect the

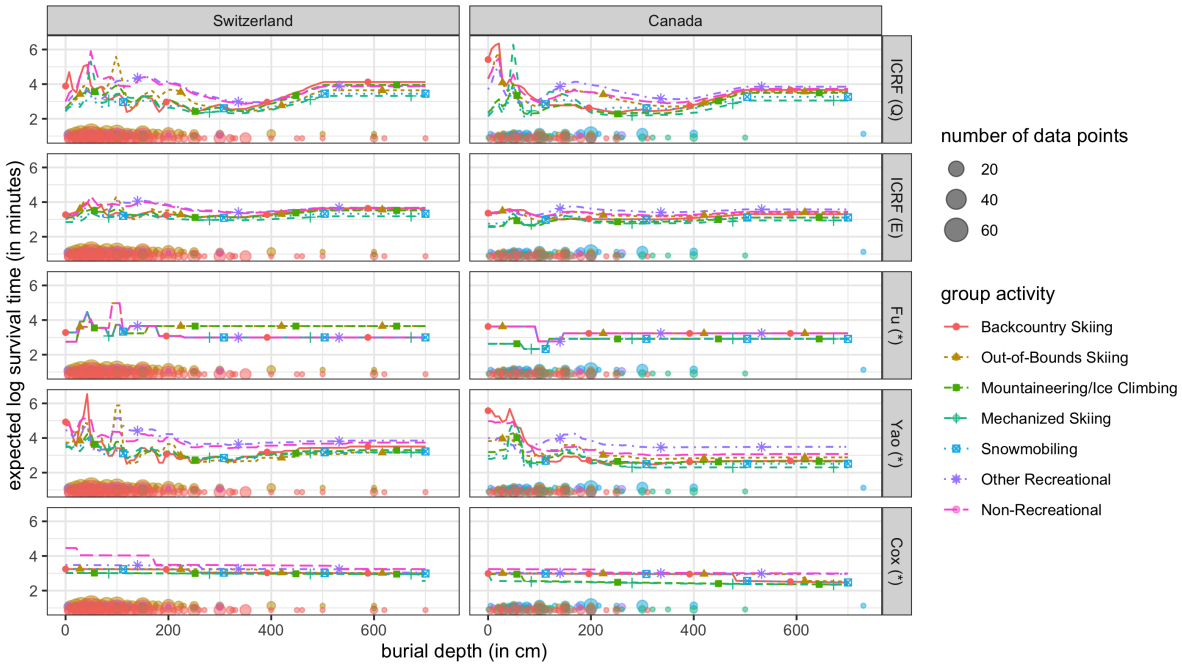

Figure 1.4: Estimated mean truncated log survival time in the avalanche data. The size of dots at the bottom of each box represents the number of sample data points.

final prediction. The increase in IMSE is averaged across ten sets of random permutations. A larger increase in error for a variable indicates higher importance of the variable. The variable importance calculated for the model fitted on the first training set of the avalanche data is presented in Table 1.4 (RIGHT). For either type of measurement (IMSE_1 or IMSE_2), burial depth is the most important variable explaining the survival probability. Group activity is chosen as more important than location except when importance is measured using IMSE_2 for the quasi-honest rule.

1.6.2 National Longitudinal Mortality Study

We use the National Longitudinal Mortality Study (NLMS) data to explore the ability of the proposed method to model rich covariate information for survival data. The NLMS is a collaborative effort between the US Census Bureau and the National Heart, Lung, and Blood Institute (NHLBI), National Cancer Institute (NCI), National Institute on Aging (NIA), and the National Center for Health Statistics (NCHS). The views expressed in this dissertation are

those of the authors and do not necessarily reflect the views of the Census Bureau, NHLBI, NCI, NIA, or NCHS. Among several data sets resulting from this extensive study, we use the dataset with six years of follow-up recorded around April 2002. The data are available at `https://biolincc.nhlbi.nih.gov/studies/nlms`.

The data include 0.7 million subjects with time to mortality, demographic information such as age, sex, and race, socioeconomic data such as income and housing tenure, and other covariates. Censoring is very high (97% survived six years), as this is a general population sample, but only administrative censoring was observed. We narrow our focus to the elderly (age ≥ 80 in years at entry) with complete covariate records ($n = 3,630$) and artificially induce current status censoring where the monitoring time depends only on age and household size. The proportion of missing covariate data is 20.7% for the whole data and 65.9% for the elderly subset. Thus, it should be noted that this data analysis is solely for performance comparison among the methods and that the results obtained from this regression analysis are limited to the selected population. The analysis framework is largely the same as for the avalanche data, except that with the increased sample size, the terminal node size was allowed to be larger ($n_{\min} = 20$ for random forests and $n_{\min} = 40$ for trees). We provide further detail about the data, the pre-processing pipeline, and the censoring mechanism in the Technical Details.

Table 1.5 (LEFT) provides the prediction accuracy (IMSE) of the models trained and evaluated based on 70:30 cross-validation. The methods of Yao et al. (2019) and the exploitative ICRF have similarly the lowest prediction errors among all methods including the Cox model. This indicates that strong assumptions such as proportional hazards and linearity may be violated in the data. Table 1.5 (RIGHT) lists the variable importance according to ICRF. Besides age, type of health insurance (HI-type) turns out to be the most important variable that explains the failure time distribution, followed by presence of a social security number (SSN), self-reported health status (health), and sex.

Prediction error		variable importance			
method	IMSE (sd)	quasi-honest		exploitative	
ICRF (Q)	0.113 (0.0038)	age	1.00	age	1.00
ICRF (E)	0.113 (0.0065)	HI-type	0.93	HI-type	0.71
Fu	0.135 (0.0057)	SSN	0.76	SSN	0.54
Fu (*)	0.134 (0.0057)	health	0.57	health	0.45
Yao	0.112 (0.0042)	sex	0.55	sex	0.31
Yao (*)	0.111 (0.0038)	race	0.20	weight	0.24
Cox	0.117 (0.0055)	tenure	0.15	relationship	0.18
Cox (*)	0.120 (0.0130)	(multiplier)	0.0169	(multiplier)	0.0151

Table 1.5: Average prediction error of the NLMS survival models for each method (LEFT) and variable importance of the ICRF model fitted on the first training set of the NLMS data based on IMSE_1 (RIGHT). For prediction error of the NLMS data, types 1 and 2 of the IMSE are equivalent. ICRF (Q), quasi-honest ICRF; ICRF (E), exploitative ICRF; The importance values are rescaled so that maximum values for each measure becomes 1. The multiplier is the original importance scale.

1.7 Discussion

In this chapter, we proposed a new tree-based iterative ensemble method for interval censored survival data. As interval censoring masks a huge amount of information, maximizing the use of available information can significantly improve the performance of estimators. Using an iterative fitting algorithm with convergence monitoring, ICRF solves the potential bias issue which most existing tree-based survival estimators have. Specifically, this bias issue arises from not fully utilizing the covariate-conditional survival probabilities in the early phases of the tree partitioning procedure for these methods, which causes the kernel estimate to incur significant bias. The WRS and log-rank tests were generalized for interval censored data and were used as splitting rules to fully utilize the hidden information. Quasi-honesty and exploitative rules were discussed for terminal node prediction. Smoothing adds another feature to ICRF.

We suggested many of the default modeling hyper-parameters, such as using GWRS or GLR as a splitting rule, the bandwidth of kernel smoothing, and the best iteration selection procedure by the out-of-bag IMSE_1 (or IMSE_2) measurement. However, the choice of the terminal node prediction rule remains unspecified. The quasi-honest and exploitative prediction

rules each have their own strengths. The quasi-honest rule induces higher model variability, while the exploitative rule tends to favor simpler models. Thus, they perform well under high and weak signal settings, respectively.

The challenge is that IMSE measurements are not always a good replacement for the true error measurement (ϵ^{INT} and ϵ^{SUP}). The out-of-bag IMSE_1 measurement recommends the exploitative prediction rule for most of the simulation settings, including scenario 3 where the quasi-honest rule has higher accuracy than the exploitative rule. Although the exploitative rule still beats the quasi-honest rule for five out of six scenarios and hence a decision rule based on out-of-bag IMSE_1 measurements may make sense, care must be taken.

This problem can be seen as a model selection problem balancing parsimony and flexibility. If the true model is thought to be smooth and simple, the exploitative rule should be employed. If the true model is believed to be complicated, the quasi-honest rule should be used. Unfortunately, the complexity or smoothness of true models is usually unknown. As model selection criteria such as AIC, BIC, and Mallow's Cp have been proposed in linear regression settings, new model selection criteria for interval censored survival models might greatly improve prediction accuracy.

The signal dilution property of the exploitative prediction rule might be caused by the fact that the marginal survival probability is shared by all censored subjects and the shared information is again carried forward to the next conditional survival probability estimate. This property might be mitigated by using non-marginal survival curves as the initial estimate. For example, the Cox model estimate or the first iteration of the quasi-honest ICRF estimate can be used as the initial estimate.

Technical Details for ICRF

The technical details for this chapter include 1) the proofs of Theorems 1.1–1.3, 2) computational cost, and 3) more details about the NLMS data analysis.

CHAPTER 2: DYNAMIC TREATMENT REGIME ESTIMATION FOR SURVIVAL OUTCOMES WITH DEPENDENT CENSORING

2.1 Introduction

In this chapter, we develop a general dynamic treatment regime estimator for censored time-to-failure outcomes that addresses all of the limitations discussed for the existing estimators in the literature review. Our estimator maximizes either the mean survival time or the survival probability at a certain time using backward recursion. As the objective may not be expressed as the sum of intermediate rewards, a standard Q-learning algorithm is not applicable. Instead, the conditional survival probability information, rather than the summarized Q-values, is appended to the previous stage information. A generalized random survival forest is developed for this task, where survival curves for each individual, instead of the observed survival or censoring time, is fed into the random survival forest. A general implementation of our method is available on CRAN, R package `dtrSurv`.

The key contributions of our work are the flexibility of the proposed method and the exposition of its theoretical properties. The proposed method affords significantly more flexibility than existing methods. The target value can be either the mean survival or the survival probability at a certain time. The method allows for a flexible number of treatment stages as well as a flexible number of treatment arms at each stage. Censoring can be conditionally independent. The conditional survival probability is modeled nonparametrically using a random survival forest-based algorithm. And finally, the proposed method does not require estimating the joint distribution of failure time and treatment time nor does it require any assumptions on their dependency, though these are allowed. We further show the consistency of the random survival forests and the estimated regime values.

The remainder of the manuscript is organized as follows. In Section 2.2, we give notation and describe the dynamic treatment rule estimator. Consistency of the estimator is derived in Section 2.3, and its performance is illustrated through simulations in Section 2.4. Section 2.5 is devoted to the application of the proposed method to a two-stage leukemia clinical trial. We conclude the chapter by discussing the implications of the modeling assumptions in Section 2.6.

2.2 The method

2.2.1 Data setup and notation

We assume that each of n patients can have a maximum of Q visits, or treatment stages, with a maximum study length $\tau > 0$. At the qth stage, $q = 1, 2, ..., Q$, the ith patient, among the $n^{(q)} \leq n$ available patients, receives treatment $A_i^{(q)} \in \mathcal{A}^{(q)}$, if he or she has survived and has not dropped out by the beginning of the stage, where $\mathcal{A}^{(q)}$ is the finite treatment space for the qth stage, $q = 1, 2, ..., Q$. Throughout this manuscript, we often drop the subject index i when it does not cause confusion. Our interest lies in estimating the survival distribution $S_{\boldsymbol{\pi}}$ of the overall failure time T of patients if they followed a dynamic treatment regime $\boldsymbol{\pi} = (\pi^{(1)}, \pi^{(2)}, ...\pi^{(Q)})^\top$, and then finding the 'best' rule $\boldsymbol{\pi}_*$, which is the rule that optimizes a certain criterion ϕ. For ϕ, we consider $\phi_\mu(S) = -\int_{t>0}(t \wedge \tau)dS(t)$ and $\phi_{\sigma,t}(S) = S(t)$ for some $0 < t \leq \tau$. Without loss of generality, we assume that a prolonged time to event is preferred; thus the objective is to maximize $\phi(S)$ over the regimes.

We allow the time to next treatment $U_i^{(q)}$ of the ith patient to depend on the historical information $\boldsymbol{H}_i^{(q)} \in \mathcal{H}^{(q)}$ available at the beginning of the qth stage and the treatment assignment $A_i^{(q)}$. At the qth stage, we observe either the ith patient's failure, advancement to next treatment, or censoring. The times from the beginning of the stage to each of these events are denoted as $T_i^{(q)}$, $U_i^{(q)}$, and $C_i^{(q)}$, respectively. The length of the stage $X_i^{(q)}$ for the ith patient is defined by the minimum of these times, and the hypothetically uncensored stage length is defined as $V_i^{(q)} = T_i^{(q)} \wedge U_i^{(q)}$. And, if $q = Q$, $V_i^{(Q)} = T_i^{(Q)}$. We denote the qth stage censoring and treatment indicators as $\delta_i^{(q)} = 1(V_i^{(q)} \leq C_i^{(q)})$ and $\gamma_i^{(q)} = 1(T_i^{(q)} \leq U_i^{(q)})$, respectively.

Note that when $C_i^{(q)} < V_i^{(q)}$, $\gamma_i^{(q)}$ is not observable. The baseline time at stage q is defined as $B_i^{(q)} = \sum_{k=1}^{q-1} X_i^{(k)}, q > 1$, and $B_i^{(1)} = 0$. Let T_i, C_i, $X_i = T_i \wedge C_i$, and $\delta_i = 1(T_i \leq C_i)$ denote the overall failure time, overall censoring time, overall observed time, and overall censoring indicator, respectively. We denote $\boldsymbol{Z}_i^{(q)} \in \mathcal{Z}^{(q)}$ as the covariate information of the ith patient that is newly available at the beginning of the qth stage. Thus, $\boldsymbol{H}_i^{(q)}$ may include the historical information such as $A_i^{(k)}$, $B_i^{(k)}$, $\boldsymbol{Z}_i^{(k)}$, $k = 1, 2, .., q-1$ and $\boldsymbol{Z}_i^{(q)}$; where $d^{(q)}$ denotes the dimension of $\boldsymbol{H}_i^{(q)}$ for all $i = 1, 2, ..., n^{(q)}; q = 1, 2, .., Q$. The number $n^{(q)}$ of patients eligible for treatment assignment at the beginning of stage q is $\sum_{i=1}^{n} 1(U_i^{(q)} \leq T_i^{(q)} \wedge C_i^{(q)})$. The notation is summarized in Table 2.6.

2.2.2 Overview of the method

As is the case in Q-learning, our estimator optimizes the intermediate stage decision assuming that all the later decisions are optimal, and thus backward recursion is used. Our method is based on the observation that the remaining life $L^{(q)}$ at stage q is defined recursively as a convolution of $L^{(q+1)}$ and the current stage length $V^{(q)}$ and that

$$S^{(q)}(t \mid \boldsymbol{H}^{(q)}, A^{(q)}, \delta^{(q)} = 1) = \int S^{(q+1)}(t - V^{(q)} \mid \boldsymbol{H}^{(q+1)}, A^{(q+1)}, \delta^{(q)} = 1) \\ dP(V^{(q)}, \boldsymbol{H}^{(q+1)}, A^{(q+1)} \mid \boldsymbol{H}^{(q)}, A^{(q)}, \delta^{(q)} = 1), \quad (2.1)$$

where $S^{(q)}(t \mid \cdot) = 1$ for all $t < 0$ and $S^{(q)}(t \mid \gamma^{(q-1)} = 1) = 1(t < 0)$, for $q = Q - 1, Q - 2, ..., 1$. In other words, assuming no censoring events, the survival probability $S^{(q)}$ of the remaining life $L^{(q)}$ at stage q can be obtained by augmenting the survival probability of the remaining life $L^{(q+1)}$ of the next stage by the current stage length $V^{(q)}$ and then marginalizing over all events that occur later than the beginning of stage q. Under the covariate-independent censoring assumption, the contribution of the censored cases is incorporated into the equation

Table 2.6: Summary of notation

	(Basic notation)
q	Treatment stage. $1, 2, ..., Q$
τ	Study length
$n^{(q)}$	Number of training examples available at stage q
	(Predictors)
$A^{(q)}$	Action at stage q.
$\boldsymbol{Z}^{(q)}$	Covariate information newly available at the beginning of stage q.
$\boldsymbol{H}^{(q)}$	History available at the beginning of stage q including $\boldsymbol{Z}^{(q)}$.
$d^{(q)}$	Dimension of history $\boldsymbol{H}^{(q)}$ at stage q.
$\pi^{(q)}(\boldsymbol{H}^{(q)})$	Decision rule at stage q given $\boldsymbol{H}^{(q)}$.
	(Stage-wise outcomes)
$T^{(q)}$	Failure time from the beginning of stage q.
$U^{(q)}$	Time to next treatment from the beginning of stage q.
$C^{(q)}$	Censoring time from the beginning of stage q.
$V^{(q)}$	Uncensored length of stage q. $V^{(q)} = T^{(q)} \wedge U^{(q)}$.
$X^{(q)}$	Observed length of stage q. $X^{(q)} = V^{(q)} \wedge C^{(q)}$.
$\delta^{(q)}$	Censoring indicator of stage q. $\delta^{(q)} = 1(V^{(q)} \leq C^{(q)})$.
$\gamma^{(q)}$	Treatment indicator of stage q. $\gamma^{(q)} = 1(T^{(q)} \leq U^{(q)})$.
$B^{(q)}$	Baseline time at stage q. $B^{(q)} = \sum_{k=1}^{q-1} X^{(k)}, q > 1$, $B^{(1)} = 0$.
$L^{(q)}$	Remaining life at stage q. $L^{(q)} = T - B^{(q)}$.
	(Overall outcomes)
T	Overall failure time. $T = \sum_{q=1}^{Q} V^{(q)}$
C	Overall censoring time
δ	Overall censoring indicator. $\delta = 1(T \leq C)$
	(Theoretical settings)
L_S, L_G	Lipschitz continuity constants of Assumption 2.6.
n_{min}	Minimum terminal node size.
β	Terminal node size rate in Assumption 2.8.
α	Regular split constant in Assumption 2.1
	(Survival functions and values)
S	Overall or generic failure survival function.
$S^{(q)}$	Survival function of remaining life at the beginning of stage q.
G	Overall or generic censoring survival function.
$\phi_\mu(S)$	Mean truncated survival time. $\phi_\mu(S) = E_S[T \wedge \tau]$.
$\phi_{\sigma,t}(S)$	Survival probability at time t. $\phi_{\sigma,t}(S) = S(t)$.
$\phi_\mu^{(q)}(S^{(q)})$	Mean truncated survival time, given $B^{(q)}$. $\phi_\mu^{(q)}(S^{(q)}) = B^{(q)} + E_{S^{(q)}}[L^{(q)} \wedge (\tau - B^{(q)})]$.
$\phi_{\sigma,t}^{(q)}(S^{(q)})$	Survival probability at time t, given $B^{(q)}$. $\phi_{\sigma,t}^{(q)}(S^{(q)}) = S^{(q)}(t - B^{(q)})$.

by redistributing the probability to the right (Efron, 1967):

$$
\begin{aligned}
& S^{(q)}(t \mid \boldsymbol{H}^{(q)}, A^{(q)}) \\
= & \int \Bigg[\delta^{(q)} S^{(q+1)}(t - X^{(q)} \mid \boldsymbol{H}^{(q+1)}, A^{(q+1)}) \\
& \quad + (1 - \delta^{(q)}) \Big\{ 1(X^{(q)} > t) + 1(X^{(q)} \leq t) \frac{S^{(q)}(t \mid \boldsymbol{H}^{(q+1)}, A^{(q+1)})}{S^{(q)}(X^{(q)} \mid \boldsymbol{H}^{(q+1)}, A^{(q+1)})} \Big\} \Bigg] \\
& \quad \times dP(X^{(q)}, \boldsymbol{H}^{(q+1)}, A^{(q+1)}, \delta^{(q)} \mid \boldsymbol{H}^{(q)}, A^{(q)}).
\end{aligned} \tag{2.2}
$$

Assuming the standard causal inference assumptions, which we will formally introduce in Subsection 2.3.2 (Assumption 2.13–2.15), we can translate $S^{(q)}(t \mid \boldsymbol{H}^{(q)}, A^{(q)} = a)$ into the counterfactual quantity $S_a^{(q)}(t \mid \boldsymbol{H}^{(q)})$. Thus, the regime is estimated using the following procedure: 1. estimating the conditional survival distribution $S^{(q)}(t \mid \boldsymbol{H}^{(q)}, A^{(q)})$ of the remaining life $L^{(q)}$ for $q = Q$; 2. optimizing the survival distribution over $A^{(q)}$ to have $\hat{\pi}^{(q)}$ and $\hat{S}_*^{(q)}(t \mid \boldsymbol{H}^{(q)})$ for $q = Q$; 3. augmenting the estimated optimal curves by the previous stage lengths, i.e., $\hat{S}_*^{(q)}(t - X^{(q-1)} \mid \boldsymbol{H}^{(q)}, A^{(q)})$ for $q = Q$; and, 4. repeating steps 1-3 for $q = Q-1, Q-2, ..., 1$. This procedure is summarized at a high level in Algorithm 2.

Result: A dynamic treatment regime estimate $\hat{\boldsymbol{\pi}} = (\hat{\pi}^{(1)}, \hat{\pi}^{(2)}, ..., \hat{\pi}^{(Q)})^\top$

for *stage* $q = Q, Q-1, ..., 1$ **do**

Obtain $\hat{S}^{(q)}(\cdot \mid \boldsymbol{H}^{(q)}, A^{(q)})$ via generalized random survival forest (Algorithm 3);

Obtain $\hat{\pi}^{(q)}(\boldsymbol{h}) = \arg\max_{a \in \mathcal{A}^{(q)}} \phi^{(q)} \Big\{ \hat{S}^{(q)}(\cdot \mid \boldsymbol{H}^{(q)} = \boldsymbol{h}, A^{(q)} = a) \Big\}$;

Obtain $\hat{S}_*^{(q)}(\cdot \mid \boldsymbol{H}^{(q)}) = \hat{S}^{(q)}(\cdot \mid \boldsymbol{H}^{(q)}, A^{(q)} = \hat{\pi}^{(q)}(\boldsymbol{H}^{(q)}))$;

Algorithm 2: the proposed dynamic treatment regime estimator

At the beginning of the last stage, there are $n^{(Q)}$ patients who have neither experienced failure nor been censored. The data for these patients are used to estimate the terminal stage covariate-conditional survival probability $S^{(Q)}(\cdot \mid \boldsymbol{H}^{(Q)}, A^{(Q)})$ for each treatment arm using random survival forests described in detail in Subsection 2.2.4. Then the optimal regime for the final stage is estimated by $\hat{\pi}^{(Q)} = \arg\max_A \phi^{(Q)} \Big\{ \hat{S}^{(Q)}(\cdot, A) \Big\}$, where $\hat{S}$ is the estimate of S and $\phi^{(q)}, q = 2, .., Q$ is defined adaptively for the intermediate and terminal survival

probabilities. Specifically, $\phi_{\mu}^{(q)}\Big\{S^{(q)}(\cdot \mid \boldsymbol{H}^{(q)})\Big\} = B^{(q)} - \int_{t>X^{(q)}}(t \wedge \tau)dS^{(q)}(t \mid \boldsymbol{H}^{(q)})$ and $\phi_{\sigma,t}^{(q)}\Big\{S^{(q)}(\cdot \mid \boldsymbol{H}^{(q)})\Big\} = S^{(q)}(t - B^{(q)} \mid \boldsymbol{H}^{(q)})$.

We denote the estimated optimized survival distribution of the qth stage as $\hat{S}_*^{(q)} = \hat{S}^{(q)}\{\cdot, \hat{\pi}^{(q)}(\cdot)\}$. The optimized survival information of the final stage is carried back to the previous stage. Specifically, the ith patient who is available at the $(Q-1)$st stage and has a stage length $X_i^{(Q-1)}$ is given a probabilistic augmentation so that the resulting survival probability of remaining life at the second to the last stage is $\hat{S}_*^{(Q-1)}(t - X_i^{(Q-1)} \mid \boldsymbol{H}_i^{(Q)})$. For those who have already experienced failure or censoring by the end of the stage, no augmentation is required, as either their remaining life has been determined already during the stage or the censoring indicator and the stage length contain all the necessary information.

Next, the conditional distribution of remaining life from the beginning of the $(Q-1)$st stage is estimated by the random survival forests based on the $n^{(Q-1)}$ patients. The distinction between the last stage and all the other stages is that given data, the patients' remaining life conditioning on the current historical information is observed in a scalar form at the last stage, while the remaining life for the other stages is not observed but is given in the form of a stochastic process. So a modification to the random survival forest estimator is required and is discussed in Subsection 2.2.4. Then estimation of $S^{(q)}, \pi^{(q)}$, and $S_*^{(q)}$ is recursively performed until we reach the first stage and obtain $\hat{\boldsymbol{\pi}} = (\hat{\pi}^{(1)}, \hat{\pi}^{(2)}, ..., \hat{\pi}^{(Q)})^{\top}$.

2.2.3 A few aspects of the optimization: Backward recursion and composite values

We discuss a few aspects of the optimization problem. First, we go over why backward recursion can be used in optimizing the truncated survival mean values (ϕ_{μ}) or survival probabilities ($\phi_{\sigma,t}$). For some criteria, such as the median $\phi_{med}(S) = S^{-1}(0.5)$, this backward recursion does not guarantee an overall optimum. However, for the mean survival time ϕ_{μ} and the survival probability $\phi_{\sigma,t}$ at time t, backward recursion is a legitimate method according to Proposition 2.1, which is proved in the Technical Details.

Proposition 2.1 (Legitimacy of backward recursion)**.** *Pick either ϕ_μ or $\phi_{\sigma,t}$ for ϕ. Let $\{\boldsymbol{\pi}\}$ be an arbitrary class of decision rules and $\boldsymbol{\pi}_* = (\pi_*^{(1)}, \pi_*^{(2)}, ..., \pi_*^{(Q)})^\top$ satisfy $\phi\Big\{S(\cdot \mid \boldsymbol{H}^{(1)}, \boldsymbol{\pi}_*)\Big\} \geq \phi\Big\{S(\cdot \mid \boldsymbol{H}^{(1)}, \boldsymbol{\pi})\Big\}$ for all $\boldsymbol{\pi} \in \{\boldsymbol{\pi}\}$. Then for any $q = 1, 2, ..., Q-1$,*

$$\pi_*^{(q)}(\boldsymbol{H}^{(q)}) = \arg\max_{a^{(q)} \in \mathcal{A}^{(q)}} \phi^{(q)}\Big\{S^{(q)}(\cdot \mid \boldsymbol{H}^{(q)}, \pi_*^{(q+1)}, \pi_*^{(q+2)}, ..., \pi_*^{(Q)})\Big\}.$$

Second, we study an alternative to the $\phi_{\sigma,t}$ criterion. When $\phi_{\sigma,t}(S) = S(t)$ is chosen as the optimization criterion, the treatment rule can not discriminate treatment arms at any point later than t. For example, when an investigator is interested in the six month survival probability and a patient takes her third round of treatment seven months after her onset of the disease, the optimal dynamic treatment rule could recommend any treatment option randomly, since it does not affect the patient's six month survival probability. Thus a composite criterion of first maximizing $S(t)$ and second maximizing $E[T \wedge \tau]$ could be more beneficial than the simple criterion $\phi_{\sigma,t}$. The composite criterion values $\phi_{\sigma,\mu,t}(S) = \{S(t), E[T \wedge \tau]\}$ are ordered lexicographically; i.e., $\phi_{\sigma,\mu,t}(S_{\boldsymbol{\pi}_1}) > \phi_{\sigma,\mu,t}(S_{\boldsymbol{\pi}_2})$ if and only if either of the following is true:

$$S_{\boldsymbol{\pi}_1}(t) > S_{\boldsymbol{\pi}_2}(t), \qquad \text{or}$$
$$S_{\boldsymbol{\pi}_1}(t) = S_{\boldsymbol{\pi}_2}(t) \text{ and } E[S_{\boldsymbol{\pi}_1} \wedge \tau] > E[S_{\boldsymbol{\pi}_2} \wedge \tau].$$

2.2.4 A generalized random survival forest

We develop a generalized random survival forest that takes survival probabilities as the outcome variable. The high level algorithm is provided in Algorithm 3. Suppose we are interested in estimating the marginal survival probability of a node $\mathcal{N}$ at stage q, where each individual's failure time information is given by the survival probability of the remaining life and the censoring status: $\{(S_i^{(q)}, \delta_i^{(q)})\}_{i=1}^{n_\mathcal{N}}$ where $n_\mathcal{N}$ is the size of node $\mathcal{N}$ and $S_i^{(q)}$ is shorthand

for $\hat{S}^{(q+1)}(t - X_i^{(q)} \mid \boldsymbol{H}_i^{(q+1)})$. Then we modify the Kaplan-Meier estimator as

$$\hat{S}^{(q)}(t \mid \mathcal{N}) = \prod_{s \geq 0}^{t} \left\{ 1 + \frac{\sum_{i \in \mathcal{N}} \delta_i^{(q)} dS_i^{(q)}(s)}{\sum_{i \in \mathcal{N}} S_i^{(q)}(s^-)} \right\}, \tag{2.3}$$

where $s^- = s - \epsilon$ with an arbitrarily small positive value ϵ. This modified Kaplan-Meier estimator will be used in both tree partitioning and terminal node survival probability estimation.

Result: generalized random survival forest for the qth stage $\hat{S}^{(q)}(\cdot \mid \boldsymbol{H}^{(q)}, A^{(q)})$

Parameters: the number of trees n_{tree}, the minimum terminal node size n_{min}, the minimum number of events in terminal nodes $n_{\text{min.E}}$;

Input data: $\{(X_i^{(q)}, \gamma_i^{(q)}, \delta_i^{(q)}, \boldsymbol{H}_i^{(q)}, A_i^{(q)})\}_{i=1}^{n^{(q)}}$, $\hat{S}_*^{(q+1)}(\cdot \mid \boldsymbol{H}^{(q+1)})$;

Stochastic augmentation: $S_i^{(q)} = \hat{S}_*^{(q+1)}(t - X_i^{(q)} \mid \boldsymbol{H}^{(q+1)} = \boldsymbol{h}_i^{(q+1)})$ with $S_i^{(q)}(t) = 1(t \leq X_i^{(q)})$ if $q = Q$ or if $\delta_i^{(q)} = 0$;

for *treatment arm* $a \in \mathcal{A}^{(q)}$ **do**

  for *tree* $b = 1, 2, ..., n_{tree}$ **do**

    Recursively partition the feature space with the augmented data $\{(S_i^{(q)}, \delta_i^{(q)}, \boldsymbol{H}_i^{(q)})\}_{\{i: a_i^{(q)} = a\}}$ via LR (2.4) or MD (2.5) until node size $< 2n_{\text{min}}$ or number of failures in each node $< 2n_{\text{min.E}}$;

    Obtain the terminal node survival probability $\hat{S}_{b,a}^{(q)}(t \mid \boldsymbol{H}^{(q)})$ via modified Kaplan-Meier (2.3);

  Obtain the random survival forest estimator $\hat{S}^{(q)}(\cdot \mid \boldsymbol{H}^{(q)}, A^{(q)} = a) = \frac{1}{n_{\text{tree}}} \sum_{b=1}^{n_{\text{tree}}} \hat{S}_{b,a}^{(q)}(t \mid \boldsymbol{H}^{(q)})$;

Algorithm 3: the generalized random survival forest estimator

Now we develop the splitting rules. Many random survival forest methods use the two-sample log-rank test statistic to partition the population into subgroups with a homogeneous survival distribution (Ishwaran et al., 2008; Hothorn et al., 2005). When failure time is not observed as a scalar but is given as a probabilistic process, the log-rank test can be generalized

to the generalized log-rank statistic (Cho et al., 2019);

$$LR(\mathcal{N}_1, \mathcal{N}_2) = \frac{\int_0^\tau \frac{Y_2(t)dN_1(t)+Y_1(t)dN_2(t)}{Y(t)}}{\sqrt{\int_0^\tau \frac{Y_1(t)Y_2(t)dN(t)(Y(t)-dN(t))}{Y(t)^3}}}, \tag{2.4}$$

where $\mathcal{N}_1$ and $\mathcal{N}_2$ are two candidate daughter nodes, $Y_l(t) = \frac{1}{|\mathcal{N}_l|}\sum_{i\in\mathcal{N}_l} S_i^{(q)}(t^-), N_l(t) = 1 - \frac{1}{|\mathcal{N}_l|}\sum_{i\in\mathcal{N}_l} S_i(t), l = 1, 2,\ Y(t) = \lambda_1 Y_1(t) + \lambda_2 Y_2(t),\ N(t) = \lambda_1 N_1(t) + \lambda_2 N_2(t)$, $\lambda_l = \frac{|\mathcal{N}_l|}{|\mathcal{N}_1|+|\mathcal{N}_2|}$, and $|\mathcal{N}| = \sum_{i=1}^{n^{(q)}} 1(i \in \mathcal{N})$. The generalized log-rank splitting rule chooses $(\mathcal{N}_1, \mathcal{N}_2) = \arg\max_{\mathcal{N}_1,\mathcal{N}_2} LR(\mathcal{N}_1, \mathcal{N}_2)$ as the best partition.

When the optimization criterion is the truncated mean survival time ϕ_μ, splitting can also be done by finding the split point that maximizes the difference of the truncated mean survival times. Define

$$MD(\mathcal{N}_1, \mathcal{N}_2) = \left| \int_{t>0} (t \wedge \tau) d\hat{S}(t \mid \mathcal{N}_1) - \int_{t>0} (t \wedge \tau) d\hat{S}(t \mid \mathcal{N}_2) \right|, \tag{2.5}$$

where $\hat{S}$ is the modified Kaplan-Meier estimator defined in Equation (2.3). The mean difference splitting rule finds $(\mathcal{N}_1, \mathcal{N}_2) = \arg\max_{\mathcal{N}_1,\mathcal{N}_2} MD(\mathcal{N}_1, \mathcal{N}_2)$ as the best split.

The recursive partitioning continues until each terminal node size is less than a predefined threshold k. At each terminal node $\mathcal{N}$ of the bth survival tree, the survival probability estimate $\hat{S}_b^{(q)}(t \mid \mathcal{N})$ is given by the modified Kaplan-Meier estimator in (2.3), $b = 1, 2, ..., n_{\text{tree}}$, where n_{tree} is the number of trees in the forest. The random survival forest estimate is then given by

$$\hat{S}^{(q)}(t \mid \boldsymbol{h}^{(q)}) = \frac{1}{n_{\text{tree}}} \sum_{b=1}^{n_{\text{tree}}} \sum_{l=1}^{\#\{\mathcal{N}_{b,l}\}_l} 1(\boldsymbol{h}^{(q)} \in \mathcal{N}_{b,l}) \hat{S}^{(q)}(t \mid \mathcal{N}_{b,l}),$$

where $\#\{\mathcal{N}_{b,l}\}_l$ is the size, or the number of terminal nodes, of the bth tree.

2.3 Theoretical properties

We show consistency of the proposed dynamic treatment regime estimator $\hat{\boldsymbol{\pi}}$ using empirical process theory. We first derive that the generalized random survival forest is a uniformly—over the feature space—consistent estimator. The uniformity is required since estimating the optimal treatment policy of the qth stage relies on integrating the estimated survival curve of $(q+1)$st stage across the newly accrued information $(\boldsymbol{H}^{(q+1)}, T^{(q)}, U^{(q)}) \mid \boldsymbol{H}^{(q)}$ between the two stages. Then we show that the value of the estimated regime is consistent for the value of the optimal regime, where the versions of the value are defined in Subsection 2.3.2.

2.3.1 Consistency of the generalized random survival forests

We derive consistency of the generalized random survival forest outside of the context of dynamic treatment regimes. We have the following assumptions on data distribution that are commonly used to achieve consistency of random forests or random survival forests (Meinshausen, 2006; Wager and Walther, 2015; Cui et al., 2017; Wager and Athey, 2018). These assumptions are tailored for use later in the dynamic treatment regime setting by adding stage indicators in Subsection 2.3.2.

Assumption 2.6 (Lipschitz continuous survival probability and censoring density)**.** *There exist constants L_S and L_G such that $|S(t \mid \boldsymbol{h}_1) - S(t \mid \boldsymbol{h}_2)| \leq L_S \|\boldsymbol{h}_1 - \boldsymbol{h}_2\|_1$ and $|g(t \mid \boldsymbol{h}_1) - g(t \mid \boldsymbol{h}_2)| \leq L_G \|\boldsymbol{h}_1 - \boldsymbol{h}_2\|_1$ for all $\boldsymbol{h}_1, \boldsymbol{h}_2 \in \mathcal{H}$ and $t \in [0, \tau]$, where G is the censoring survival distribution and g is its derivative with respect to time.*

Assumption 2.7 (Weakly dependent covariate values)**.** *The covariate information ($\boldsymbol{Z}$) is given as a d_Z-dimensional vector residing in a unit hypercube, i.e., $\boldsymbol{Z} \in [0,1]^{d_Z}$, with a density $f_{\boldsymbol{Z}}$ such that $\zeta^{-1} \leq f_{\boldsymbol{Z}}(\boldsymbol{z}) \leq \zeta$ for all but countably many points $\boldsymbol{z} \in [0,1]^{d_Z}$ and some constant $\zeta \geq 1$.*

Remark 2.1. *The distributional assumption was made for notational convenience and could be relaxed. For example, the results in this chapter still hold for covariates that are a multivariate*

normal random vector truncated by some constants since random forests are invariant to scale transformation. Also the results are still valid with categorical covariates with a finite number of levels after modification to Assumption 2.6: in this setting, the norm is calculated by excluding the categorical covariate coordinates. The case of categorical covariates will be discussed in the proof of Theorem 2.4.

We also have assumptions regarding the tree-based estimator terminal node size and splitting rules. The first assumption forces the terminal node size to be both sufficiently small and sufficiently large.

Assumption 2.8 (Terminal node size)**.** *The minimum size n_{min} of the terminal nodes grows at the following rate:*

$$n_{min} \asymp n^{\beta}, \frac{1}{2} < \beta < 1,$$

where $a \asymp b$ implies both $a = \mathcal{O}(b)$ and $b = \mathcal{O}(a)$.

Next, we give definitions of regular trees and random split trees that are often used for consistency proofs in the random forest literature (Meinshausen, 2006; Wager and Walther, 2015; Cui et al., 2017; Wager and Athey, 2018).

Definition 2.1 (random-split and α-regular trees and forests)**.** *A tree is called a random-split tree if each feature is given a minimum probability (φ/d) of being the split variable at each intermediate node, where $0 < \varphi < 1$. A tree is α-regular, if every daughter node has at least α fraction of the training sample in the corresponding parent node. A random forest is called a random-split (α-regular) forest, if each of its member trees is random-split (α-regular).*

Assumption 2.9 (α-regular and random split trees)**.** *Trees are α-regular and random-split with a constant $0 < \varphi < 1$.*

Now we give our uniform consistency results of the generalized random survival forests. This theorem is largely based on the Z-estimation lemma and Donskerness of the tree kernels. A detailed proof is relegated to the Technical Details.

Theorem 2.4 (Uniform consistency of a random survival forest)**.** *Suppose Assumptions 2.6-2.9 hold. Let $\tilde{S}'$ be an estimator of S' that is uniformly consistent in both $\boldsymbol{h} \in \mathcal{H}$ and $t \in [0, \tau]$, where S' satisfies $S(t \mid \boldsymbol{h}) = \int S'(t - (T \wedge U) \mid \boldsymbol{H}')dP(T, U, \boldsymbol{H}' \mid \boldsymbol{H} = \boldsymbol{h})$. Then, the generalized random survival forest $\hat{S}(t \mid \boldsymbol{h})$ built based on $\tilde{S}'$ is uniformly consistent. Specifically,*

$$\sup_{t \in [0,\tau], \boldsymbol{h} \in \mathcal{H}} |\hat{S}(t \mid \boldsymbol{h}) - S(t \mid \boldsymbol{h})| \to 0,$$

in probability as $n \to \infty$.

Consistency of a random survival forest sequence in a dynamic treatment regime setting requires adaptive versions of Assumptions 2.6-2.8 and an additional condition that guarantees an arbitrarily large sample size for the last stage with high probability. To achieve this condition, we assume that, at each stage except the last one, each patient receives the next treatment before experiencing failure or censoring with at least a certain minimum probability.

Assumption 2.10 (Lipschitz continuous survival probability and censoring density at each stage)**.** *There exist constants L_S and L_G such that $|S^{(q)}(t \mid \boldsymbol{h}_1) - S^{(q)}(t \mid \boldsymbol{h}_2)| \leq L_S \|\boldsymbol{h}_1 - \boldsymbol{h}_2\|$ and $|g^{(q)}(t \mid \boldsymbol{h}_1) - g^{(q)}(t \mid \boldsymbol{h}_2)| \leq L_G \|\boldsymbol{h}_1 - \boldsymbol{h}_2\|$ for all $\boldsymbol{h}_1, \boldsymbol{h}_2 \in \mathcal{H}^{(q)}$, $t \in [0, \tau - B^{(q)}]$, and $q = 1, 2, ..., Q$, where $G^{(q)}$ is the censoring survival distribution at the qth stage and $g^{(q)}$ is its derivative with respect to time.*

Assumption 2.11 (Weakly dependent historical information at each stage)**.** *The patient history information $\boldsymbol{H}^{(q)}$ at stage q is given as a $d^{(q)}$-dimensional vector lying in a subset $\mathcal{H}^{(q)}$ of $[0, 1]^{d^{(q)}}$, and its joint and marginal densities are bounded so that $\zeta^{-1} \leq f_{\boldsymbol{H}^{(q)}}(\boldsymbol{h}) \leq \zeta$ and $\zeta^{-1} \leq f_{\boldsymbol{H}^{(q)},j}(\boldsymbol{h}_j) \leq \zeta$ for all but a finite number of $\boldsymbol{h} \in \mathcal{H}^{(q)}$ and $j = 1, 2, ..., d^{(q)}$ for some constant $\zeta \geq 1$.*

Assumption 2.11 modifies Assumption 2.7 so that the support of the history space can be less than a hyper-cube, since, for example, the baseline survival times, $(B^{(1)} \leq B^{(2)} \leq B^{(3)} \leq ...)$, have a dependent support structure.

Assumption 2.12 (Probability of proceeding to the next stage)**.** *There exists a constant* $M \in (0,1)$ *such that for all* $h^{(q)} \in \mathcal{H}^{(q)}$ *and* $q = 1, 2, ..., Q-1$, $\Pr(U^{(q)} < T^{(q)} \wedge C^{(q)} \mid \boldsymbol{h}^{(q)}) > M$.

Corollary 2.1 (Uniform consistency of a random survival forest sequence)**.** *Suppose Assumptions 2.8–2.12 hold. Let* $\hat{\boldsymbol{S}} = (\hat{S}^{(1)}, ..., \hat{S}^{(q)}, ..., \hat{S}^{(Q)})$ *be the sequence of the generalized random survival forest estimators of* $\boldsymbol{S} = (S^{(1)}, ..., S^{(q)}, ..., S^{(Q)})$ *such that the* q*th stage random survival forest is built based on* $\hat{S}^{(q+1)}$ *for* $q = 1, 2, ..., Q-1$. *Then,*

$$\sup_{t \in [0,\tau], \boldsymbol{h} \in \mathcal{H}, q \in \{1,2,...,Q\}} |\hat{S}^{(q)}(t \mid \boldsymbol{h}) - S^{(q)}(t \mid \boldsymbol{h})| \to 0,$$

in probability as $n \to \infty$.

The proof of Corollary 2.1 is given in the Technical Details.

2.3.2 Consistency of the dynamic treatment regime estimator

We show consistency of the dynamic treatment regime estimator using the causal inference framework. With this framework, the theoretical properties are applicable to broader settings than sequentially randomized experiments. A counterfactual outcome is defined as an outcome that would have been obtained if a person had a treatment option contrary to his/her actual treatment. To denote the counterfactual outcomes pertaining to a counterfactual treatment or treatment policy, we add a subscript to the corresponding random variables or survival functions. For example, $S_{\pi}^{(q)}(t|\boldsymbol{H}^{(q)})$ is a counterfactual survival probability of the remaining life if treatment was given according to a treatment rule π to a person with history $\boldsymbol{H}^{(q)}$ at stage q. We assume the following three standard causal inference assumptions (Hernán and Robins, 2020; Rubin, 2005; Cole and Frangakis, 2009) known as the stable unit treatment value assumption, sequential ignorability, and positivity.

Assumption 2.13 (Stable unit treatment value assumption, or SUTVA)**.** *Each person's counterfactual dynamics, such as failure time, time to next treatment, and censoring, are not affected by*

the treatments or history of other patients. Moreover, each treatment has only one version, or if there are multiple versions, their differences are irrelevant.

Assumption 2.14 (Sequential ignorability)**.** *Treatment assignment is given independently of the counterfactual outcomes, conditionally on the individual's historical information, where the outcomes include all future random quantities such as failure time, time to next treatment, and censoring time.*

Assumption 2.15 (Positivity)**.** *For each stage $q = 1, 2, ..., Q$, given historical information, the probability of having each treatment $a \in \mathcal{A}^{(q)}$ is greater than a constant $L > 0$. Or,*

$$\Pr(A^{(q)} = a | \boldsymbol{H}^{(q)} = \boldsymbol{h}) > L$$

for all $a \in \mathcal{A}^{(q)}$, $\boldsymbol{h}$ such that $\Pr(\boldsymbol{H}^{(q)} = \boldsymbol{h}) > 0$.

The following theorem states that the dynamic treatment regime estimator has a value consistent for the value of the optimal regime. We define the value Φ of a treatment regime $\boldsymbol{\pi}$ as the criterion value of the survival probability if all individuals in a population follow the treatment regime. We use $\Phi(\boldsymbol{\pi}) = \phi(S_{\boldsymbol{\pi}})$ to denote the value of $\boldsymbol{\pi}$, and Φ inherits the subscript of ϕ.

Theorem 2.5 (Consistency of the dynamic treatment regime estimator)**.** *Suppose Assumptions 2.10–2.15 hold, and the optimal dynamic treatment regime estimator ($\hat{\boldsymbol{\pi}}$) is built based on the generalized random forest grown under Assumptions 2.8 and 2.9. Then, for each $\phi = \phi_\mu, \phi_{\sigma,t}$,*

$$|\Phi(\boldsymbol{\pi}) - \Phi(\hat{\boldsymbol{\pi}})| \to 0,$$

in probability as $n \to \infty$.

The proof is provided in the Technical Details.

2.4 Simulations

2.4.1 Simulation settings

We perform simulations to study the performance of the proposed method and compare those results to that of existing methods. The overall scheme of the simulations follows the pattern in Goldberg and Kosorok (2012), where they mimicked a cancer clinical trial using the tumor size and wellness dynamics. We adapt their data generating mechanism so that dependence among censoring, tumor size, and wellness is incorporated in the mechanism, and deterministic processes are relaxed into more realistic ones.

In the hypothetical trial, patients can go through up to $Q = 3$ treatment rounds during the trial with length of the trial, $\tau = 10$. Each patient has tumor size $\phi(t)$ and wellness $\omega(t)$ at time $t \in [0, 10]$ that affects the timing of failure or treatment. We use $[t^{(q)}, t^{(q+1)}]$ to denote the qth stage. For each stage, patients are given either a more aggressive ($A = 1$) or a less aggressive ($A = 0$) treatment.

At the beginning of the trial, patients have uniformly distributed tumor sizes and wellness: $\rho(0) \sim U(0.5, 1), \omega(0) \sim U(0.5, 1)$. At each treatment stage, the more aggressive treatment immediately reduces the tumor size more than the less aggressive one. However, the more aggressive treatment worsens patient wellness more than the less aggressive one:

$$\rho(t^{(q)}+) = \frac{\rho(t^{(q)})}{\omega(t^{(q)})\{10 - 6A\}} \vee 0$$
$$\omega(t^{(q)}+) = \omega(t^{(q)}) - 2^{A-2}.$$

After each treatment ($t^{(q)} < t \leq t^{(q+1)}$), the tumor size and wellness follow the Ornstein-Uhlenbeck processes with negative dependence between the two:

$$\frac{d\chi(t)}{dt} = -2\chi(t) + \eta_\chi(t),$$

where $\chi = \rho, \omega$ and the second term obeys a drifted Wiener process such that

$$\frac{d}{dt}\begin{pmatrix}\eta_\rho(t)\\ \eta_\omega(t)\end{pmatrix} \sim N\left(\begin{pmatrix}0.1\\ 0.05\end{pmatrix}, \begin{pmatrix}1 & -0.5\\ -0.5 & 1\end{pmatrix}\right).$$

The positive means in the normal distribution imply that on average both tumor size and wellness grow. When tumor size is greater than 1, the patient receives the next treatment. In other words, $t^{(q+1)} = \min_t \arg\max_{t>t^{(q)}} 1(\omega(t) > 1)$, for $q = 1, 2, .., Q-1$. The patient failure time is governed by the hazard process:

$$\lambda_F(c) = \frac{1}{5}\left\{\frac{\rho(u)}{\omega(u)} e^{\boldsymbol{g}(H^{(q)})^\top \boldsymbol{\beta}_F(A^{(q)})} + 1\Big(\omega(u) < 0.1\Big)\right\}, \tag{2.6}$$

where $g(H^{(q)}) = \{1, \log(B^{(q)}+1), \boldsymbol{Z}^{(q)^\top}\}^\top$, $\boldsymbol{\beta}_F(a)$ is the hazard ratio for the failure time given treatment a that is defined in Table 2.7, and the indicator accelerates the failure time when the wellness drops below the critical level, 0.1. The baseline covariates at the qth stage follow a $p-$variate normal distribution conditioning on the previous stage value: $\boldsymbol{Z}^{(q)} = \frac{1}{2}\boldsymbol{Z}^{(q-1)} + \frac{1}{2}\boldsymbol{Z}_0^{(q)}$, where $\boldsymbol{Z}_0^{(q)} \sim N_p(\mathbf{0}_p, 0.8I_p + 0.2 \cdot 1_p^p), q = 1, 2, ..., Q$, 1_p^p is the p-dimensional square matrix of ones, and $\boldsymbol{Z}^{(0)} = \mathbf{0}_p$.

For each stage, the censoring occurs according to the hazard function,

$$\lambda_C(u) = e^{\boldsymbol{g}(H^{(q)})^\top \boldsymbol{\beta}_C(A^{(q)})}, \tag{2.7}$$

where the censoring hazard ratio, $\boldsymbol{\beta}_C(a)$, is also defined in Table 2.7.

The treatment is allocated according to a Bernoulli trial with propensity

$$\pi(\boldsymbol{H}^{(q)}) = \Big\{1 + \exp(-g(\boldsymbol{H}^{(q)})^\top \boldsymbol{\beta}_\pi)\Big\}^{-1}, \tag{2.8}$$

where the coefficient is $\boldsymbol{\beta}_\pi = (-1, 1, -\frac{1}{2}\mathbf{1}_p^\top)^\top$ for the observational data setting and $\boldsymbol{\beta}_\pi = \mathbf{0}_{p+2}$ for the randomized trial setting.

Table 2.7: The failure time and censoring hazard coefficients and dimensions of covariates in Scenarios 1 to 4.

parameter	scenario 1 (base)	scenario 2 (small p)	scenario 3 (large p)	scenario 4 (moderate censoring)
p	5	2	10	5
$\boldsymbol{\beta}_F(1)$	$(0,-3,2,2,1,-1,-1)^\top$	$(0,-2,2,0)^\top$	$(0,-2,\mathbf{2}_4^\top,-\mathbf{1}_4^\top,-\mathbf{2}_2^\top)$	$(0,-3,2,2,1,-1,-1)^\top$
$\boldsymbol{\beta}_F(0)$	$(0,-1,1,1,1,1,1)^\top$	$(0,-1,1,1)^\top$	$(0,-1,\mathbf{1}_4^\top,\mathbf{1}_4^\top,\mathbf{0}_2^\top)$	$(0,-1,1,1,1,1,1)^\top$
$\boldsymbol{\beta}_C(1)$	$(-3,.2,\mathbf{.2}_3^\top,-\mathbf{.2}_2^\top)^\top$	$(-3,.2,.2,-.2)^\top$	$(-3,.2,\mathbf{.2}_4^\top,-\mathbf{.2}_3^\top,\mathbf{0}_3^\top)^\top$	$(-2,.2,\mathbf{.2}_3^\top,-\mathbf{.2}_2^\top)^\top$
$\boldsymbol{\beta}_C(0)$	$(-3,.1,\mathbf{.2}_3^\top,\mathbf{.2}_2^\top)^\top$	$(-3,.1,.2,.2)^\top$	$(-3,.1,\mathbf{.2}_4^\top,\mathbf{.2}_3^\top,\mathbf{0}_3^\top)^\top$	$(-2,.1,\mathbf{.2}_3^\top,\mathbf{.2}_2^\top)^\top$

$\boldsymbol{\beta}_F(\cdot)$ and $\boldsymbol{\beta}_C(\cdot)$, the log hazard ratios of the counterfactual failure and censoring times in (2.6) and (2.7), respectively.

Simulations are run with two sample sizes, $n = 300$ and 1000. Thus, we have in total 16 distinct factorial settings (four hazard/censoring scenarios, two propensity designs, and two sample sizes). For each setting, we have $n_{rep} = 200$ simulation replicates, and we estimate the optimal treatment rule based on both the ϕ_μ and $\phi_{\sigma,t=5,\mu}$ criteria. We compare the proposed method with the methods in Goldberg and Kosorok (2012) and Simoneau et al. (2019). We further compare the results with the observed policy characterized by (2.8) and the estimated optimal zero-order model, where all patients are given the same embedded treatment regime, which is the best on average over all embedded regimes. In particular, in implementing the Goldberg and Kosorok (2012) method, the linear Q-function space and the random forest-based approximation space were separately used. The method of Huang et al. (2014) is not included in these simulations because Simoneau et al. (2019) can be viewed as a doubly robust extension of their method. After estimation, the values—either the expected truncated survival time or the survival probability at $t = 5$—are evaluated by generating an uncensored sample of size $n_{eval} = 10000$ according to the estimated policies. The simulations were complete using `R` package `dtrSurv`. For reproducible research, the code is available at https://github.com/hunyong/survQlearn.

2.4.2 Simulation results

In Figure 2.5 LEFT, the simulation results show how the truncated mean survival time (ϕ_μ) of each policy behaves under different settings. For most of the settings, on average, all estimated regimes have greater values than the zero-order model, and the zero-order model has

a higher value than the observed policy. This implies that the standard of care improves the overall survival time of patients over the observed policies and that individualized treatment rules can further enhance the outcomes.

Figure 2.5 LEFT also shows that among the individualized treatment rules, the proposed method outperforms the other methods in many settings. Even when the method exhibits lower or about the same performance as the other methods for the small sample size, it often has higher performance than the others for the larger sample size. See, for example, the high censoring cases (the last row) of either the observational or the RCT setting. In the settings where the dimension of the covariate is high (the third row), the method of Simoneau et al. (2019) often has high values, although sometimes estimation is not possible. This might be because of the doubly robustness property of the method and the fact that the true optimal decision rule might be close to a linear function, as suggested by the linear interaction of the stage-wide hazard functions in (2.6). The method of Simoneau et al. (2019), however, often does not provide estimates when either the sample size of the terminal stage is small due to censoring or failure or the dimension of covariates is large.

Figure 2.5 RIGHT shows the simulation results in terms of the survival probability at $t = 5$, or $\phi_{\sigma,5}$. Note that the estimated regimes for the proposed method are distinct between LEFT and RIGHT but are the same for the other methods. The proposed estimator shows better performance in terms of the policy values under most settings, giving us similar interpretations as the previous results.

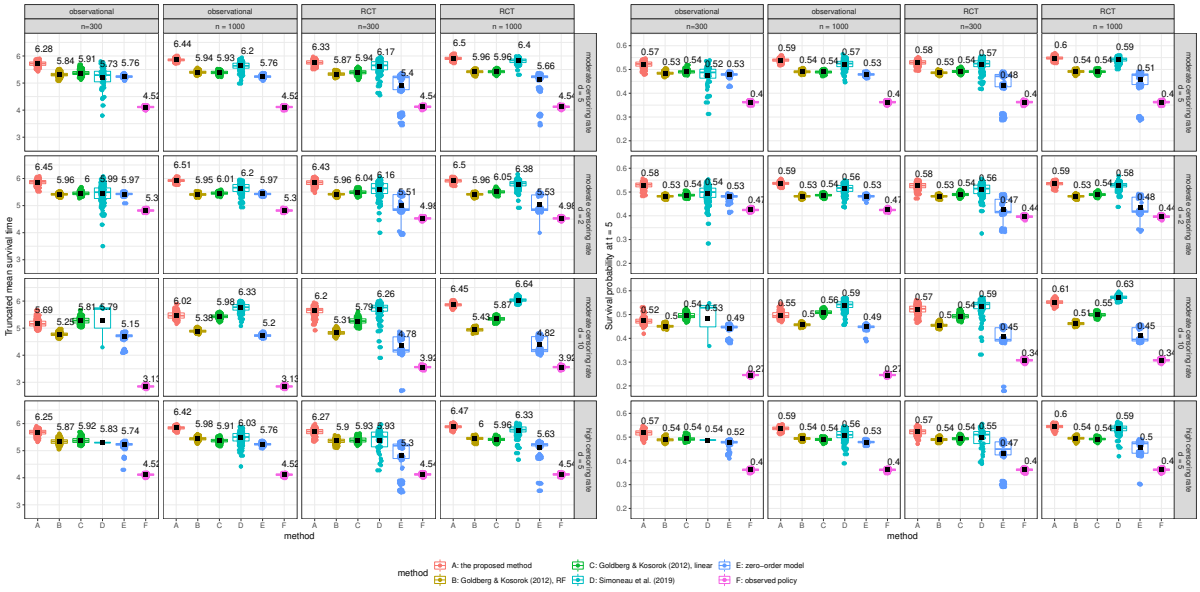

Figure 2.5: The estimated mean truncated survival time (LEFT) and the estimated survival probability at time $t = 5$ (RIGHT) under each estimated policy. Each dot represents one simulation replicate. The box plots are the quartile summary of the points and the black dots are the average values of each method. Each row of boxes represents a data generation scenarios, and each column of boxes represents a combination of a study design and a sample size.

2.5 Leukemia data example

We analyze acute myeloid leukemia patient survival time using the proposed and existing methods. We use a data set collected from a clinical trial in which 210 acute myeloid leukemia patients were randomized to a frontline treatment followed by a salvage treatment that was adaptively chosen based on the patient status (Wahed and Thall, 2013; Xu et al., 2016). One of four initial treatment options was given to each patient with an equal probability: combination of fludarabine, cytosine arabinoside, and idarubicin (FAI), combination of FAI and all-trans retinoic acid (ATRA) (FAIA), combination of FAI + granulocyte colony stimulating factor (G-CSF) (FAIG), and combination of FAI + ATRA + G-CSF (FAIAG). Response to the first stage treatment took one of the following forms: complete remission followed by disease progression, failure (either partial remission or remission with subsequent relapse), or resistance to treatment. Surviving patients that did not attain complete remission were given a salvage treatment based on physician judgement. The salvage treatments were labeled with one of two

categories depending on whether or not they contained high dose ara-C (HDAC) or not. We assume that the salvage treatment effects are homogeneous across different versions and that the current data set contains most of the necessary information to achieve Assumption 2.14.

The data set is composed of age at entry, type of leukemia characterized by cytogentic abnormality, treatment assignment, response to the initial treatment (failure, complete remission, or resistance), duration of each treatment stage, and survival and censoring status at each treatment stage. Because both age and type of leukemia are known to have high power in predicting overall survival time, they are the main factors taken into account by clinicians when prescribing treatment in practice thereby mitigating the potential violation of Assumption 2.14. Patient age at entry ranges from 21 to 87 years and is 61.1 years on average. The patients' cytogenetics abnormalities were grouped into three categories according to their prognosis (Wahed and Thall, 2013): 65 poor, 79 intermediate, and 66 good. Out of 210 patients, 198 experienced death during the study, and 12 were censored. The overall observed survival time ($T \wedge C$) was on average 414 days and had quartiles of 45, 184, and 443 days. We set $\tau = 450$ days. Figure 2.6 overviews the dynamics of the study.

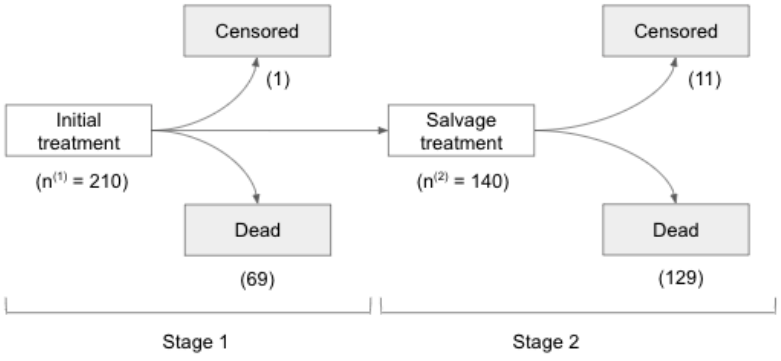

Figure 2.6: Treatment stages and patient status with the number of instances in the parentheses.

We derive the optimal treatment rules using the proposed method, the Goldberg and Kosorok (2012) method, the Simoneau et al. (2019) method, and the zero-order model. The value of the observed policy—the combination of randomization and physician's judgement—is estimated and is compared with the other rules. As in Section 2.4, the Goldberg and Kosorok (2012)

method was implemented using two different Q-function approximations: a linear model and a random forest. Because the Simoneau et al. (2019) method was implemented using `R` package `DTRreg` the current version (1.7) of which allows binary treatment arms only, the method was applied after dichotomizing the initial treatment according to inclusion of all-trans retinoic acid.

We compare the performance of the five models based on cross validation, where $K = 100$ training data sets of size 168 (80%) are sampled to estimate the treatment rules without replacement and stratified according to the initial treatment, and the remaining test sets are held out for evaluation. The value of each treatment rule is estimated by the inverse probability (censoring) weighting approach as follows.

Define

$$W_i(\boldsymbol{\pi}) = \frac{\prod_{q=1}^{2} 1(\pi^{(q)}(\boldsymbol{H}_i^{(q)}) = A_i^{(q)})\delta_i}{\widehat{\Pr}(A_i^{(1)})\widehat{\Pr}(A_i^{(2)} \mid \boldsymbol{H}_i^{(2)}) \prod_{q=1}^{2} \widehat{\Pr}(T_i^{(q)} > C_i^{(q)} \mid \text{age}_i)}, \tag{2.9}$$

where the propensity of the salvage treatment is modeled as $\Pr(A^{(2)} = a) = (1 + \exp(-\boldsymbol{H}^{(2)}\boldsymbol{\beta}))^{-1}$, $\boldsymbol{H}^{(2)}$ is a matrix with columns of one, $B^{(2)}$, $A^{(1)}$, age, cytogenetic leukemia type, and response to initial treatment. Because there was an insufficient number of censored patients in the study, the censoring propensity score is modeled using only the patient age at entry. Then the value of each policy $\boldsymbol{\pi}$ and each criterion $\phi = \phi_\mu, \phi_{\sigma,t}$ is estimated by

$$\hat{\phi}_\mu(S^{\boldsymbol{\pi}}) = \frac{\sum_{i:\text{test set}} (T_i \wedge \tau) W_i(\boldsymbol{\pi})}{\sum_{i:\text{test set}} W_i(\boldsymbol{\pi})}$$

and

$$\hat{\phi}_{\sigma,t}(S^{\boldsymbol{\pi}}) = \frac{\sum_{i:\text{test set}} 1(T_i > t) W_i(\boldsymbol{\pi})}{\sum_{i:\text{test set}} W_i(\boldsymbol{\pi})},$$

where for the second criterion, we use $t = 180$ to evaluate the six month survival probabilities. We estimate the value of the observed policy, i.e., a hybrid of randomized rule for the initial

treatment and the physician judgement for the salvage treatment, under the same evaluation procedure.

Figure 2.7 summarizes the value estimates of the estimated polices. Each box and scatter plot corresponds to a policy, and within each plot, each dot represents a cross-validation replicate of the estimated value (truncated mean survival time (LEFT) or six-month survival probability (RIGHT)) of the corresponding policy. The results show that the proposed method achieves higher values than the other estimated policies on average in terms of both the truncated mean survival time and the six month survival probability. This is true when compared with the zero-order model implying the advantage of the proposed individualized treatment rules over the standard of care. However, the estimated methods do not have higher values than the observed policy value on average. Recalling that the second treatment assignments were done according to physician judgement, this implies that their judgement was fairly successful and that a larger sample size and more patient information—available only to physicians but not included in this data set—could enhance the performance of the methods. Despite the gap, however, the proposed method yields the closest values to the observed policy values among the competing methods.

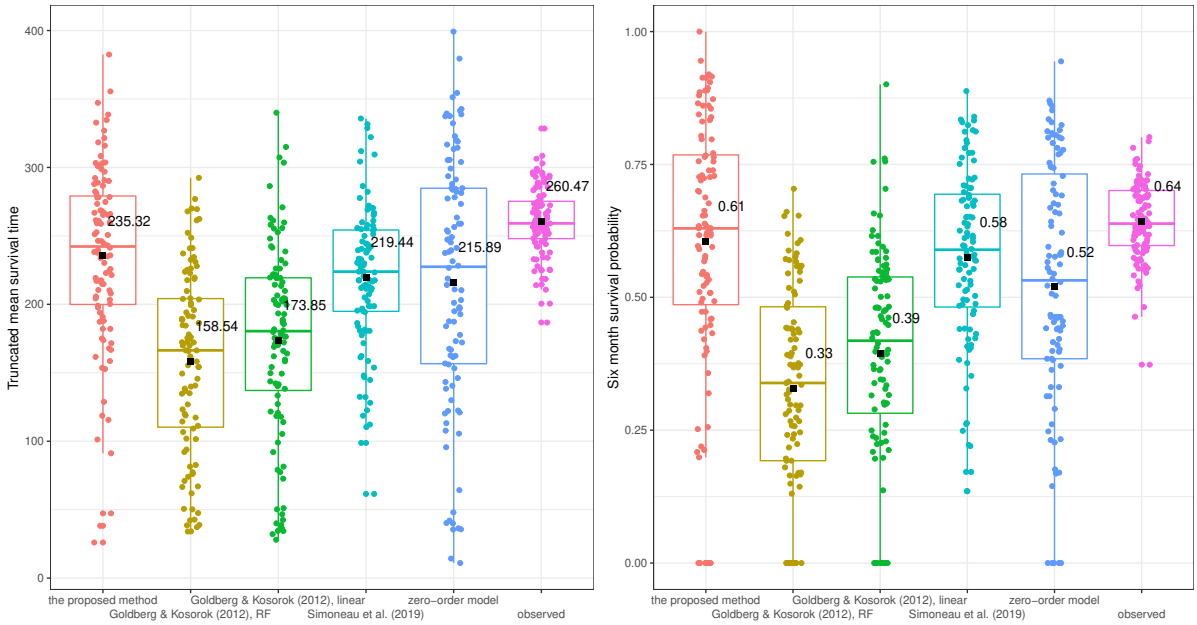

Figure 2.7: Estimated truncated mean survival time (LEFT) and estimated six-month survival probability (RIGHT) under the dynamic treatment regimes. Each dot represents one cross-validation replicate. The box plots are the quartile summary of the points and the black dots are the average values of each method.

2.6 Discussion

In this project, we introduced a new dynamic treatment regime estimator for survival outcomes. The estimator is versatile in the sense that it allows for a flexible number of treatment stages and arms, it gives a non-linear decision rule that maximizes either the mean survival time or the survival probability, and it permits dependent censoring. In this section, we discuss some considerations when using our method.

The proposed estimator assumes that the distribution is well characterized by the treatment stages rather than the natural passage of time. That is, the cohort at each treatment stage might include patients that received the stage treatment on day 10 as well as patients that received the stage treatment on day 1000. Thus, this method is effective when the disease dynamic is considered stationary between stage transitions or when the treatment effect contains relatively stronger signal than that of the baseline disease dynamics.

Another consideration is the choice of the study length, τ, which is chosen so that the survival probabilities are reliably estimated up to τ. For the stability of estimation, the chosen τ should not be too large; however, reducing τ results in information loss. If all patients are censored at a certain stage, the survival probability is not estimable. In this case, a smaller τ should be chosen so that every stage has a sufficient number of observed failure times.

Finally, in establishing the uniform consistency results, we use the fact that under the given assumptions, the minimum size of the terminal node becomes arbitrarily small in probability uniformly, as argued in Wager and Walther (2015). However, the uniform rate of convergence can be very low. As a random-split tree grows deeper and has an increasingly large number of terminal nodes, there exists, with a not-very-small probability, at least one terminal node created without enough splits on a certain covariate. To enhance the rate, splitting can potentially be done in a stratified fashion by assigning a certain minimum number of splits to each variable. The effect of the tree building rules on the rate of convergence would be an interesting future study especially in the presence of noise variables in high-dimensional data settings.

Technical Details for the survival DTR estimator

The technical details for this chapter include 1) the proofs of Proposition 2.1, Theorem 2.4, Corollary 2.1 and Theorem 2.5.

CHAPTER 3: BIVARIATE ZERO-INLFATED NEGATIVE BINOMIAL MODEL FOR MEASURING DEPENDENCE

3.1 Introduction

In this project, we study how the zero-inflation of single cell RNA-sequencing (scRNA-seq) data can be adjusted when measuring the gene-gene correlation is of interest. As discussed in the literature review, two strategies have been considered to address bias generally in scRNA-seq data. Imputation methods (Li and Li, 2018 and Peng et al., 2019) and estimation of the count distribution (Huang et al., 2018 and Wang et al., 2018b). Our proposed method takes the distribution estimation approach where a bivariate distribution explicitly addresses the dependence structure. Specifically, our method is built on a bivariate generalization of the zero-inflated negative binomial (ZINB) model. For univariate count data, zero-inflated negative binomial (ZINB) models have been well accepted and have greater capability than Poisson, zero-inflated Poisson, and negative binomial models in terms of handling augmented zeros and overdispersion. While negative binomial models have been extensively used for bulk RNA-seq data without much zero-inflation (Love et al., 2014, Robinson et al., 2010), ZINB models are typically used for scRNA-seq data (van den Berge et al., 2018, Risso et al., 2018). Therefore, we have a particular interest in a bivariate generalization of ZINB models to model dependence of two genes in scRNA-seq data.

We propose a bivariate zero-inflated negative binomial model with eight parameters: five parameters for the negative binomial part and another three free parameters for the zero-inflation part. This model allows analyzing the dependence of two zero-inflated count variables parametrically but with more flexibility than existing models.

The rest of the chapter is organized as follows. In Section 3.2, we describe how the model is constructed in the order of a Bivariate Negative Binomial model and a Bivariate Zero-inflated Negative Binomial model. We present the maximum likelihood estimator using expectation-maximization algorithm in Section 3.3. In Section 3.4, we illustrate how well the models fit data and how model-based dependence measures behave in contrast to naive measures using mouse paneth scRNA-seq data. Then in Section 3.5, we show how point and interval estimators perform based on simulations. In Section 3.6, we address limitation of the models and discuss potential extensions. Section 3.7 provides the software information.

3.2 The model

3.2.1 A Bivariate Negative Binomial Model

In constructing the BZINB model, to induce dependence and zero-inflation, layers of latent variables were used as in Kocherlakota and Kocherlakota (1992) and Li et al. (1999). We first introduce a simpler model, the Bivariate Negative Binomial (BNB) model in this subsection, and then generalize it to Bivariate Zero-Inflated Negative Binomial (BZINB) model in Subsection 3.2.2.

One of the key assumptions about the dependence structure of BNB (and BZINB) is that the mean parameters of two Poisson random variables are gamma random variables that share a common gamma random variable. Let $R_j \sim Gamma(\alpha_j, \beta)$ for $j = 0, 1, 2$, where α_j and β are the shape and scale parameters, respectively. Then $(R_0 + R_1, R_0 + R_2)$ is bivariate gamma distributed, denoted as $BGamma(\alpha_0, \alpha_1, \alpha_2, \beta)$. To account for heterogeneous scales of the two Poisson mean variables, we introduce an additional parameter $\delta \in \mathbb{R}^+$. Then, a pair (X_1, X_2) of Poisson variables with means $(R_0 + R_1, \delta(R_0 + R_2))$ follow a bivariate negative binomial distribution, denoted as

$$(X_1, X_2) \sim BNB(\alpha_0, \alpha_1, \alpha_2, \beta_1, \beta_2), \tag{3.10}$$

where we reparametrize (β, δ) as $(\beta_1, \beta_2) = (\beta, \delta\beta)$ and the observed density is given as,

$$
\begin{aligned}
&f_{BNB}(x_1, x_2) \\
&= \iiint_{\mathbb{R}^3_+} \frac{(R_0+R_1)^{x_1}(R_0+R_2)^{x_2} e^{-\frac{1+\beta_1+\beta_2}{\beta_1}R_0 - \frac{1+\beta_1}{\beta_1}R_1 - \frac{1+\beta_2}{\beta_1}R_2} R_0^{\alpha_0-1} R_1^{\alpha_1-1} R_2^{\alpha_2-1} \beta_2^{x_2}}{x_1! x_2! \Gamma(\alpha_0)\Gamma(\alpha_1)\Gamma(\alpha_2)\beta_1^{\alpha_0+\alpha_1+\alpha_2+x_2}} \prod_{j=0}^{2} dR_j \\
&\quad \times 1_{(x_1,x_2)\in\mathbb{N}_0^2} \\
&= \sum_{k=0}^{x_1}\sum_{m=0}^{x_2} \binom{\alpha_0+x_1+x_2-k-m-1}{\alpha_0+x_2-m-1} \binom{\alpha_0+x_2-m-1}{\alpha_0-1} \binom{\alpha_1+k-1}{\alpha_1-1} \binom{\alpha_2+m-1}{\alpha_2-1} \\
&\quad \times \frac{\beta_1^{x_1}\beta_2^{x_2}(\beta_1+\beta_2+1)^{k+m-x_1-x_2-\alpha_0}}{(\beta_1+1)^{k+\alpha_1}(\beta_2+1)^{m+\alpha_2}} 1_{(x_1,x_2)\in\mathbb{N}_0^2},
\end{aligned}
$$

where $\mathbb{N}_0$ denotes the nonnegative integer space, and superscripts represent the dimension of the product space. The support indicators will be omitted throughout this chapter when the context is clear.

This bivariate negative binomial model (BNB) is marginally negative binomial, as we know from the construction procedure that both X_1 and X_2 are Poisson random variables with means marginally Gamma distributed, respectively:

$$X_j \sim NB(\alpha_0 + \alpha_j, \frac{1}{\beta_j + 1}) \text{ for } j = 1, 2,$$

where the random variable $X \sim NB(\nu, \phi)$ can be interpreted as the minimum number of failures to have ν successes with probability of ϕ for each trial; i.e., its density is expressed as $f_{NB}(x; \nu, \phi) = \binom{x+\nu-1}{x}\phi^\nu(1-\phi)^x$.

Interpretation of the BNB parameters is straightforward: α_0, α_1, and α_2 are the shape parameters of latent variables, where the larger α_0 implies a larger amount of shared components in X_1 and X_2 and thus larger correlation; β_1 and β_2 controls the scale of X_1 and X_2, respectively. Note in scRNA-seq data context, X_1 and X_2 may represent the *before-dropout* expression level of each of two genes in a cell in the absence of dropout events, which we rarely observe in practice.

The first two moments and the correlation of a BNB random pair are given as,

$$
\begin{aligned}
E(X_j) &= (\alpha_0 + \alpha_j)\beta_j && j = 1, 2 \\
Var(X_j) &= (\alpha_0 + \alpha_j)\beta_j(\beta_j + 1) && j = 1, 2 \\
Cov(X_1, X_2) &= \alpha_0\beta_1\beta_2 \\
Cor(X_1, X_2) &= \frac{\alpha_0}{\sqrt{(\alpha_0 + \alpha_1)(\alpha_0 + \alpha_2)}} \sqrt{\frac{\beta_1\beta_2}{(\beta_1 + 1)(\beta_2 + 1)}} && (3.11)
\end{aligned}
$$

Note that this distribution only allows positive correlation. See Section 3.6 for more discussion.

We recognize that Maher (1990) developed another bivariate negative binomial distribution that is a constrained case of BNB in a sense that the marginal means and variances are the same for both variables.

One can further generalize this BNB model into a m-variate negative binomial model by adding common latent gamma variable(s) to the m gamma variables.

3.2.2 A Bivariate Zero-inflated Negative Binomial Model

In this subsection, we generalize BNB model to BZINB model by including zero-inflation components. Since BZINB is also a generalization of univariate zero-inflated negative binomial model (ZINB), we illustrate the construction of univariate ZINB model first and move to the bivariate version.

A univariate negative binomial model, $NB(\nu, \phi)$, can be generalized to allow zero-inflation by having an additional parameter, π: $ZINB(\nu, \phi, \pi)$. The zero-inflated negative binomial (ZINB) model has a latent variable interpretation. Let X follow $NB(\nu, \phi)$ and E denote the zero-inflation indicator having 1 with probability of π and 0 otherwise, independently of X. Then $Y \equiv (1 - E)X$ follows $ZINB(\nu, \phi, \pi)$ with the density of $f_{ZINB}(y; \nu, \phi, \pi) = (1 - \pi)f_{NB}(y; \nu, \phi) + \pi\zeta(y)$, where $\zeta(a) \equiv 1_{(a=0)}$.

Similarly, a multivariate zero-inflated random variable can be constructed using a latent variable that follows the multivariate Bernoulli distribution as in the Poisson case (Li et al., 1999). For a bivariate distribution, suppose we have a random vector $\boldsymbol{E} \equiv (E_1, E_2, E_3, E_4)^\top \sim MN(1, \boldsymbol{\pi})$, where $MN(1, \boldsymbol{\pi})$ denotes the multinomial distribution with a single trial and an associated probability of $\boldsymbol{\pi} \equiv (\pi_1, \pi_2, \pi_3, \pi_4)^\top$. Now the bivariate zero-inflated negative binomial distribution (BZINB) can be formulated as:

$$(Y_1, Y_2) := ((E_1 + E_2)X_1, (E_1 + E_3)X_2), \tag{3.12}$$

where $(X_1, X_2) \sim BNB(\alpha_0, \alpha_1, \alpha_2, \beta_1, \beta_2)$ and E_1, E_2, E_3 and E_4 are the indicators of observing both X_1 and X_2, only X_1, only X_2, and none of them, respectively. We say $(Y_1, Y_2) \sim BZINB(\alpha_0, \alpha_1, \alpha_2, \beta_1, \beta_2, \pi_1, \pi_2, \pi_3, \pi_4)$. A simpler model with a restriction of $\pi_2 = \pi_3 = 0$ can also be considered as in Wang (2003).

The density of a BZINB variable is

$$f_{BZINB}(y_1, y_2; \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\pi})$$

$$\begin{aligned} &= \pi_1 f_{BNB}(y_1, y_2; \alpha_0, \alpha_1, \alpha_2, \beta_1, \beta_2) + \pi_2 f_{NB}(y_1; \alpha_0 + \alpha_1, \frac{1}{\beta_1 + 1})\zeta(y_2) \\ &\quad + \pi_3 f_{NB}(y_2; \alpha_0 + \alpha_2, \frac{1}{\beta_2 + 1})\zeta(y_1) + \pi_4 \zeta(y_1 + y_2), \end{aligned}$$

where $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \alpha_2)^\top, \boldsymbol{\beta} = (\beta_1, \beta_2)^\top$, and $\boldsymbol{\pi} = (\pi_1, \pi_2, \pi_3, \pi_4)^\top$ with $\mathbf{1}^\top \boldsymbol{\pi} = 1$.

Here, the parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ has the same interpretation as in BNB but in the presence of dropouts, and $\boldsymbol{\pi}$ indicates the dropout probability, where π_1, π_2, π_3, and π_4 are the probability that none, Y_2 only, Y_1 only, and both were dropped out, respectively.

In scRNA-seq data,In scRNA-seq data, Y_1 and Y_2 are the *observed* number of expressions for each of two genes in a cell. The term *observed* was used in contrast to *before-dropout* in a sense that an unobserved subset of the zeros are excess zeros due to dropouts.

This BZINB distribution is marginally ZINB, since the latent random variables, X_1 and X_2, are marginally negative binomial random variables (from Subsection 3.2.1) with probabilities of being observed, $\pi_1 + \pi_2$ and $\pi_1 + \pi_3$, respectively:

$$Y_j \sim ZINB(\alpha_0 + \alpha_j, \frac{1}{\beta_j + 1}, \pi_{4-j} + \pi_4) \text{ for } j = 1, 2. \tag{3.13}$$

The first two moments of a BZINB pair are given as,

$$\begin{aligned}
E(Y_j) &= (\pi_1 + \pi_{j+1})(\alpha_0 + \alpha_j)\beta_j && j = 1, 2 \\
Var(Y_j) &= (\alpha_0 + \alpha_j)^2\beta_j^2(\pi_1 + \pi_{j+1})(1 - \pi_1 - \pi_{j+1}) \\
&\quad + (\alpha_0 + \alpha_j)\beta_j(\beta_j + 1)(\pi_1 + \pi_{j+1}) && j = 1, 2 \\
Cov(Y_1, Y_2) &= \{\alpha_0 + (\alpha_0 + \alpha_1)(\alpha_0 + \alpha_2)\}\beta_1\beta_2\pi_1 \\
&\quad - (\alpha_0 + \alpha_1)(\alpha_0 + \alpha_2)\beta_1\beta_2(\pi_1 + \pi_2)(\pi_1 + \pi_3),
\end{aligned}$$

and the correlation $\rho(Y_1, Y_2)$ is not further simplified than $Cov(Y_1, Y_2)/\sqrt{Var(Y_1)Var(Y_2)}$.

When dropouts are unwanted and need to be adjusted for, then the underlying correlation ρ^* of Y_1 and Y_2 under BZINB model is simply the correlation of X_1 and X_2 (Equation (3.11)), which is

$$\rho^*(Y_1, Y_2) = \frac{\alpha_0}{\sqrt{(\alpha_0 + \alpha_1)(\alpha_0 + \alpha_2)}}\sqrt{\frac{\beta_1\beta_2}{(\beta_1 + 1)(\beta_2 + 1)}}. \tag{3.14}$$

3.3 Estimation

With the natural interpretation of BZINB model as layers of latent variables, one can estimate the parameters by the expectation-maximization (EM) algorithm.

The complete density is given as,

$$
\begin{aligned}
&f(Y_1, Y_2, X_1, X_2, R_0, R_1, R_2, E_1, E_2, E_3, E_4) \\
&= f(X_1, X_2, R_0, R_1, R_2, E_1, E_2, E_3, E_4) \times 1_{(Y_1 = X_1(E_1+E_2), Y_2 = X_2(E_1+E_3))}
\end{aligned}
$$

with

$$
\begin{aligned}
&f(X_1, X_2, R_0, R_1, R_2, E_1, E_2, E_3, E_4) \\
&= \frac{(R_0+R_1)^{X_1}(R_0+R_2)^{X_2} R_0^{\alpha_0-1} R_1^{\alpha_1-1} R_2^{\alpha_2-1} \beta_2^{X_2} \prod_{k=1}^4 \pi_k^{E_k}}{X_1! X_2! \Gamma(\alpha_0)\Gamma(\alpha_1)\Gamma(\alpha_2) \exp\{R_0 \frac{1+\beta_1+\beta_2}{\beta_1} + R_1 \frac{1+\beta_1}{\beta_1} + R_2 \frac{1+\beta_2}{\beta_1}\} \beta_1^{X_2+\alpha_0+\alpha_1+\alpha_2}} \\
&\quad \times 1_{\sum_{k=1}^4 E_k = 1}.
\end{aligned}
$$

Thus, the full individual log-likelihood for the ith entry, or the ith cell, is

$$
\begin{aligned}
&l_i^{\text{Full}} \\
&= X_{1,i} \log(R_{0,i} + R_{1,i}) + X_{2,i} \log(R_{0,i} + R_{2,i}) \\
&\quad + (\alpha_0 - 1) \log R_{0,i} + (\alpha_1 - 1) \log R_{1,i} + (\alpha_2 - 1) \log R_{2,i}
\end{aligned}
$$

$$
\begin{aligned}
&+ X_{2,i} \log \beta_2 - (X_{2,i} + \alpha_0 + \alpha_1 + \alpha_2) \log \beta_1 + \sum_{k=1}^4 E_{k,i} \log \pi_k - \log X_{1,i}! - \log X_{2,i}! \\
&- \log \Gamma(\alpha_0) - \log \Gamma(\alpha_1) - \log \Gamma(\alpha_2) - R_{0,i} \frac{1+\beta_1+\beta_2}{\beta_1} - R_{1,i} \frac{1+\beta_1}{\beta_1} - R_{2,i} \frac{1+\beta_2}{\beta_1} \\
&+ \log 1_{(Y_{1,i} = X_{1,i}(E_{1,i}+E_{2,i}))} + \log 1_{(Y_{2,i} = X_{2,i}(E_{1,i}+E_{3,i}))} + \log 1_{\sum_{k=1}^4 E_k = 1}.
\end{aligned}
$$

The expected full log-likelihood conditional on the observed data is linear in $E[R_{j,i}|Y_{1,i}, Y_{2,i}; \boldsymbol{\theta}]$, $E[\log(R_{j,i}|Y_{1,i}, Y_{2,i}; \boldsymbol{\theta})]$, $E[E_{k,i}|Y_{1,i}, Y_{2,i}; \boldsymbol{\theta}]$, and $E[X_{2,i}|Y_{1,i}, Y_{2,i}; \boldsymbol{\theta}]$, where $\boldsymbol{\theta} \equiv (\boldsymbol{\alpha}^\top, \boldsymbol{\beta}^\top, \boldsymbol{\pi}^\top)^\top$, $j = 0, 1, 2$ and $k = 1, 2, 3, 4$. The formulae of the components are given in Technical details.

As the likelihood is the product of functions convex with respect to each of the parameters, the maximization can be achieved by solving a system of score equations. The individual scores are given as:

$$
\begin{aligned}
\partial_{\alpha_j} E[l_i^{Full}|\cdot] &= E[\log R_{j,i}|\cdot] - \log\beta_1 - \psi(\alpha_j) && j = 0,1,2 \\
\partial_{\beta_1} E[l_i^{Full}|\cdot] &= E[R_{0,i} + R_{2,i}|\cdot]\frac{1+\beta_2}{\beta_1^2} + \frac{E[R_{1,i}|\cdot]}{\beta_1^2} - \frac{\alpha_0+\alpha_1+\alpha_2+E[X_{2,i}|\cdot]}{\beta_1} \\
\partial_{\beta_2} E[l_i^{Full}|\cdot] &= -\frac{E[R_{0,i}+R_{2,i}|\cdot]}{\beta_1} + \frac{E[X_{2,i}|\cdot]}{\beta_2} \\
\partial_{\pi_j} E[l_i^{Full}|\cdot] &= \frac{E[E_{j,i}|\cdot]}{\pi_j} - \frac{1-E[E_{j,i}|\cdot]}{1-\pi_j} && j = 1,2,3,
\end{aligned}
$$

where the conditioning arguments $(\boldsymbol{Y}_1, \boldsymbol{Y}_2; \boldsymbol{\theta})$ are suppressed as $(\cdot)$ and can be replaced with $(Y_{1,i}, Y_{2,i}; \boldsymbol{\theta})$ where we assume a sample of independent entries, $\boldsymbol{Y}_l$ denotes $(Y_{l,1}, ..., Y_{l,n})^\top$ for $l = 1, 2$, n is the sample size, and $\partial_a b$ denotes the partial derivative of b with respect to a.

At the $k+1$st iteration of the EM algorithm, we get $\boldsymbol{\theta}^{(k+1)}$ by solving the score equations $\partial_{\boldsymbol{\theta}} \sum_i^n E[l_i^{Full}|\boldsymbol{Y}_1, \boldsymbol{Y}_2, \boldsymbol{\theta}^{(k)}] = \mathbf{0}$:

$$
\begin{aligned}
\frac{\beta_2^{(k+1)}}{\beta_1^{(k+1)}} &= \frac{\bar{E}[X_{2,i}|\cdot]}{\bar{E}[R_{0,i}+R_{2,i}|\cdot]} \\
\beta_1^{(k+1)} &= \frac{\bar{E}[R_{0,i}+R_{1,i}+R_{2,i}|\cdot]}{\alpha_0^{(k+1)}+\alpha_1^{(k+1)}+\alpha_2^{(k+1)}} \\
\pi_j^{(k+1)} &= \bar{E}[E_{j,i}|\cdot] && j = 1,2,3,4 \\
\alpha_j^{(k+1)} &= \psi^{-1}\{-\log\beta_1^{(k+1)} + \bar{E}[\log R_{j,i}|\cdot]\} && j = 0,1,2,
\end{aligned}
$$

where $\bar{E}[A|\cdot]$ denotes the empirical average of the conditional expectations, i.e., $\frac{1}{n}\sum_i^n E[A_i|\cdot]$, $\psi(\cdot)$ is the digamma function, and the conditioning arguments $(\boldsymbol{Y}_1, \boldsymbol{Y}_2, \boldsymbol{\theta}^{(k)})$ are again suppressed. The equations can be solved by solving the following through Newton-Raphson

algorithm:

$$\text{Solve for } \beta_1 = \frac{\bar{E}[R_0 + R_1 + R_2|\cdot]}{\sum_{k=0}^{2} \psi^{-1}(-\log \beta_1 + \bar{E}[\log R_k|\cdot])}.$$
$$\text{Then get } \alpha_j = \psi^{-1}(-\log \beta_1 + \bar{E}[\log R_j|\cdot]).$$

After iterations enough to observe convergence, the final updated parameter values serve as the maximum likelihood estimate.

The standard error of the maximum likelihood parameter estimates can be calculated using observed information. In Technical details, detailed formulae are given, and simulations illustrating the accuracy of standard error estimation are included in Section 3.5.

3.4 Model and measure comparisons based on mouse paneth data

3.4.1 Model comparison using mouse paneth data

In this section, we show how the BZINB model fits a scRNA-seq data set compared to its nested models (in Subsection 3.4.1) and present how model-based dependence measures can be different from naive measures (in Subsection 3.4.2). The data were collected from paneth cells of C57Bl6 mouse with a Sox9 gene knockout. The Fluidigm C1 system was used to capture single cells and generate Illumina libraries using manufacutrers' protocols. Illumina NextSeq sequencing platform was used for paired end sequencing. Reads per cell were demultiplexed using mRNASeqHT_demultiplex.pl, a script provided by Fluidigm. Low quality base calls and primers were removed using Trimmomatic (Bolger et al., 2014) and poly-A tails were removed using a custom perl script. Reads were aligned to the mouse genome (mm9) using STAR (`https://academic.oup.com/bioinformatics/article/29/1/15/272537`) and read per gene were counted using htseq-count (`https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4287950`). The data are composed of 23,425 genes for 800 cells, where all the cells came from a single mouse and have the same cell type. Over 90% of genes have more

than 90% zero counts and the average proportion of zero counts for a gene is 97.3% in these zero-inflated data.

We compare four nested models: BZINB, BNB, bivariate zero-inflated Poisson (BZIP), and bivariate Poisson (BP). BZIP has fixed mean values instead of latent gamma variables of BZINB, and BP further lacks zero-inflation components. The estimated densities of these models are compared with the empirical density for 50 gene pairs.

To systematically study the model performances, we performed stratified sampling of genes according to their proportion of zeros; strata H, M, L, and V include genes with $\geq 90\%$, 80% to 90%, 60% to 80%, and $< 60\%$ zeros, respectively. Genes with $\geq 98\%$ of zeros and genes with extremely large expression ($> 10,000$ counts for at least one cell) were screened out. After screening out those irregular genes, each group has 81.4%, 13.5%, 4.2%, and 0.9% of genes in the order.

We randomly selected 5 pairs from each possible combination of two strata (HH, MM, LL, VV, HM, HL, HV, ML, MV, and LV) without replacement. For each of the 50 pairs (5 pairs $\times$ 10 combinations), we estimated the parameters of the four nested models. Based on the parameter estimates, the distributions of the four models were compared. As it is not straightforward to compare the estimated model-based densities with the empirical density, we drew a random sample of size $n = 800$ from each estimated model and the resulting empirical densities were then compared (Figure 3.8 for several pairs and Web Figure 1 for all the 50 pairs). As we cannot preclude the chance of getting unlikely instances by doing Monte Carlo sampling, we added results of two more replicates in Web Figures 2 and 3. We furthermore illustrate the exact values of the estimated density in Figure 3.9 for a couple of pairs and in Web Figure 4 for all the pairs.

Figure 3.8 illustrates the true and the model-based empirical distributions for the first pairs of 10 combinations. The results including all 50 pairs and their replicates can be found in Web Figures 1 to 3. For any pair, the BP model obviously fails to address the overdispersion and zero-inflation, while the BZIP model could not properly mimic the overdispersion. BNB and BZINB seem to fairly mimic the true distribution in most of the pairs. The poor performances

of Poisson-based models and decently good performances of BNB and BZINB models can also be seen on Figure 3.9.

However, when genes have some large-valued counts and many zeros at the same time either marginally or jointly, BZINB has an apparent advantage over BNB model. Often, in BNB model, nonzero count pairs are highly concentrated on the diagonal line, while nonzero counts in BZINB model are more dispersed away from the diagonal line (LL1 in Figure 3.8 and more examples in Web Figures 1 to 3). This can be explained by the lack of flexibility of BNB model. When data are highly zero-inflated but overdispersed at the same time, BNB is forced to have small shape parameters ($\alpha_j,\ j = 0, 1, 2$) and large scale parameters ($\beta_j,\ j = 1, 2$) while keeping the mean of the latent Gamma variables, $E[R_j] = \alpha_j\beta_1$, close to zero. These latent Gamma variables, serving as mean parameters of Poisson variables, take on very small values most of the times and very large values with small chance. It is unlikely that both R_1 and R_2 have large numbers at the same time (CASE 1), but it is more frequent that R_0 alone has a large number (CASE 2). Thus, the latent Poisson variables, X_1 and X_2, are more likely to have similarly large numbers (resulting from CASE 2) than to have significantly different nonzero numbers (resulting from CASE 1).

3.4.2 Mouse paneth data example for dependence measures

When the excess zeros are believed to come from dropouts, BZINB model may uncover the underlying dependence using measures such as ρ^* and MI^*. Note that MI^* is defined similarly to that of ρ^* and can be estimated by first estimating the BZINB model parameters and by measuring the mutual information of the estimated distribution after replacing $\hat{\pi}$ with $(1, 0, 0, 0)^\top$.

For the same 50 pairs in the previous subsection, we estimated the dependence using naive measures – Pearson correlation (PC) and empirical mutual information (EMI) – and zero-inflation adjusted measures – underlying correlation (ρ^*) and underlying MI (MI^*) based on BZINB model. Figure 3.10 summarizes the estimates for all the pairs.

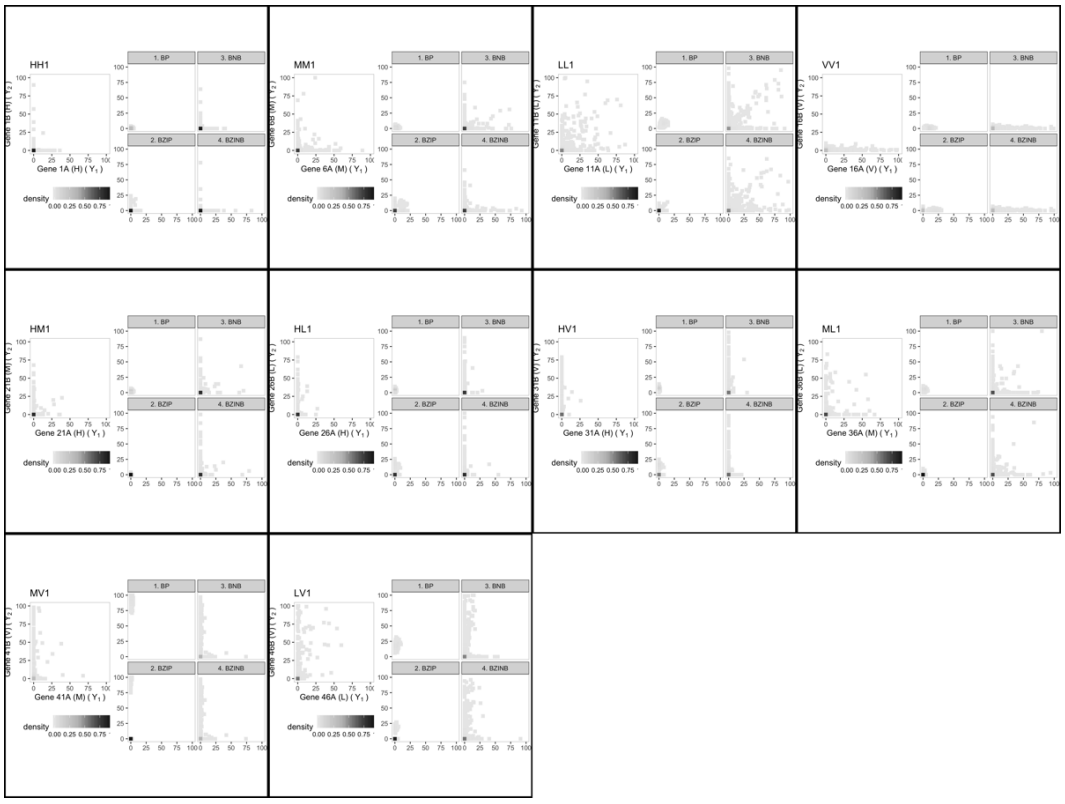

Figure 3.8: The bivariate distribution of true and simulated mouse paneth RNA count data. Each box corresponds to the first pair of each of the combination, HH1, MM1, LL1, VV1, HM1, HL1, HV1, ML1, MV1, and LV1, where letters represent stratum with varying proportions of zeros and the numbers represent the number of the pair in each combination. Each box has the empirical distribution (LEFT) which serves as truth for the simulations, and the four model-based simulated empirical distributions (RIGHT).

In Figrue 3.10 LEFT, we see that PC and ρ^* mostly behave in the same direction, but also that they can have values in the opposite directions (e.g., HL5 and HL4). If we judge whether two genes are correlated based on (naive) Pearson correlation (PC) with a certain threshold, say PC > 0.2, many genes might be missed (e.g., HL5) or falsely included (e.g., HL4).

Similar analyses can be done for MI-based measures. Both EMI and MI^* estimates are correlated, however, there are pairs that are located away from the tendency. For example the pair MV1 has highest MI^*, while its EMI is not one of the highest. Also note that the values of MI^* are in general less than those of EMI for scRNA-seq data. Heavy proportion of zero-zero pairs boosts naive EMI, while MI^* removes the effects of the co-zero-inflation. These results

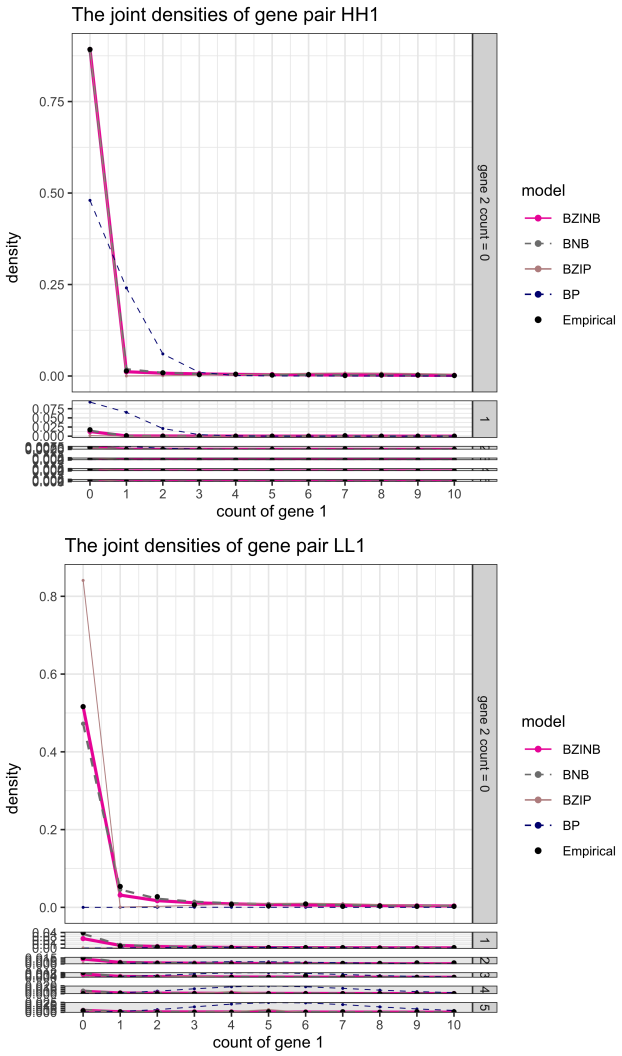

Figure 3.9: The model estimates of bivariate densities (lines) and the empirical densities (dots) of two gene pairs.

suggest that measures that fail to identify the excess zeros caused by the dropout events may be highly misleading.

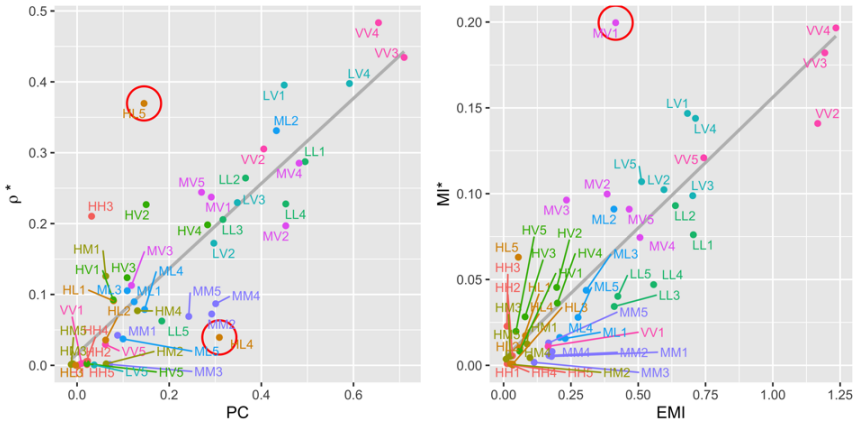

Figure 3.10: Estimated dependence measures of 50 pairs. Pearson correlation and underlying correlation estimates (LEFT). Empirical and underlying mutual information estimates (RIGHT).

3.5 Evaluation of estimators based on simulation

We ran simulations to study the performance of estimation of underlying correlation and the associated standard error under finite sample size. We considered 40 distinct sets of BZINB parameter values (Table 3.8). Note that for each of ρ^*'s there are two distinct sets of parameters $(\boldsymbol{\alpha}, \boldsymbol{\beta})$, the first (a) of which have lower $\boldsymbol{\alpha}$ values and the second (b) of which have higher $\boldsymbol{\alpha}$ values. For each parameter set $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\pi})$ and for $n = 250, 500, 800, 1500, 2500$, we generated random BZINB samples of size n, $n_{\text{sim}} = 1,000$ times.

For each k of n_{sim} simulation replicates, we got an estimate $\hat{\rho}_k^*$ of the parameter ρ^*, the standard error estimate $se(\hat{\rho}_k^*)$, and the logit-transformed 95% confidence interval (i.e., $\text{logit}^{-1}(\text{logit}(\hat{\rho}_k) \pm 1.96\frac{se(\hat{\rho}_k)}{\hat{\rho}_i(1-\hat{\rho}_k)})$). Then for each set of parameters, the following three quantities were calculated:

- the average estimated standard error (SE, $\bar{se}(\hat{\rho}^*)$)
- the standard deviation of the parameter estimates (SD, $sd(\hat{\rho}^*)$)
- the empirical coverage probability (CP, $\frac{1}{n_{\text{sim}}}\sum_{k=1}^{n_{\text{sim}}} 1_{\rho^* \in \text{CI}_k}$, where CI_k is the logit-transformed 95% confidence interval for the kth replicate).

Table 3.8: The set of parameters for simulation. Combination of $(\alpha_0, \alpha_1, \alpha_2, \beta_1, \beta_2)$ and $(\pi_1, \pi_2, \pi_3, \pi_4)$ below makes $40 (= 8 \times 5)$ sets in total.

underlying correlation		#	$(\alpha_0, \alpha_1, \alpha_2, \beta_1, \beta_2)$
1. high	($\rho^* = 0.6$)	1-a	(0.2, 0.05, 0.05, 3.0, 3.0)
		1-b	(2.0, 0.7, 0.1, 2.5, 2.5)
2. moderate	($\rho^* = 0.3$)	2-a	(1.0, 1.0, 1.0, 1.5, 1.5)
		2-b	(3.0, 2.0, 1.0, 1.5, 0.5)
3. low	($\rho^* = 0.1$)	3-a	(0.2, 0.3, 3.0, 2.0, 1.5)
		3-b	(0.5, 2.0, 2.0, 0.5, 3.0)
4. very low	($\rho^* = 0.01$)	4-a	(0.01, 0.1, 1.0, 0.5, 0.5)
		4-b	(0.05, 2.0, 3.0, 3.0, 0.5)

zero-inflation	$(\pi_1, \pi_2, \pi_3, \pi_4)$
i. low	(0.7, 0.1, 0.1, 0.1)
ii. moderate-balanced	(0.5, 0.15, 0.15, 0.2)
III. moderate-unbalanced	(0.5, 0.1, 0.3, 0.1)
iv. high-balanced	(0.2, 0.2, 0.2, 0.4)
v. high-unbalanced	(0.2, 0.1, 0.4, 0.3)

The simulation results are provided in Figures 3.11 and 3.12. First, the mean parameter estimates are close, or getting closer as sample size grows, to their true parameter values for each of the 40 scenarios. For most of the 40 parameter sets, CP was close to 0.95, and for those not close, CP gets closer to 0.95 with increasing sample size. In the same context, the average estimated standard error (SE) was close to the standard deviation of the parameter estimates (SD) especially when the sample size was large. However, when the underlying correlation was close to zero (i.e., 0.01 in our example), standard error estimation did not perform as well in terms of both CP and closeness of SE to SD. The parameter being to the boundary may be responsible for the poorer performance. Also, Scenarios iv and v have higher SE and SD than the others. This is perhaps because the effective sample size for those high zero-inflation scenarios is smaller than the other scenarios.

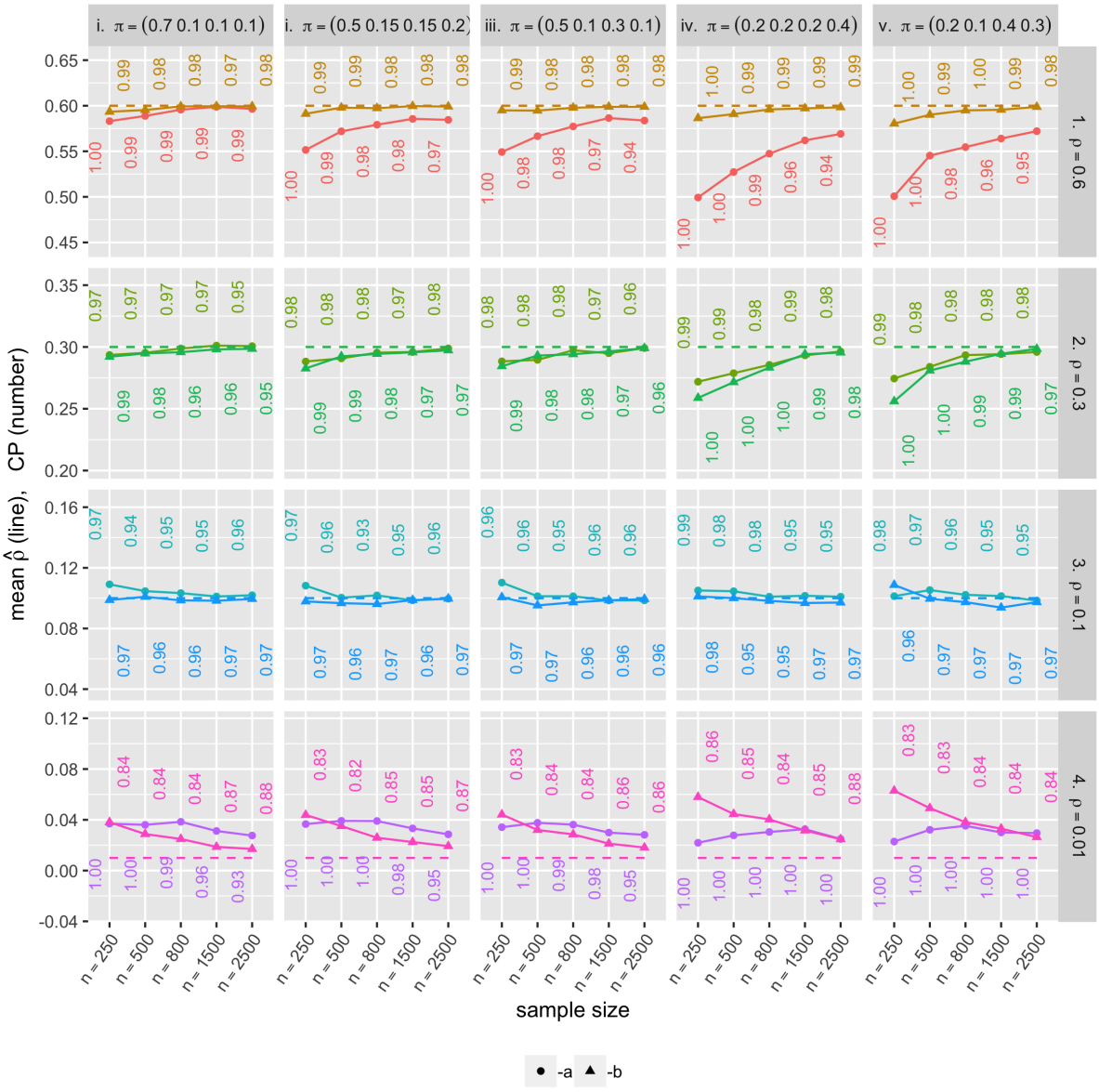

Figure 3.11: Mean parameter estimates ($\hat{\rho}^*$) and CP (each color represents distinct simulation scenarios.)

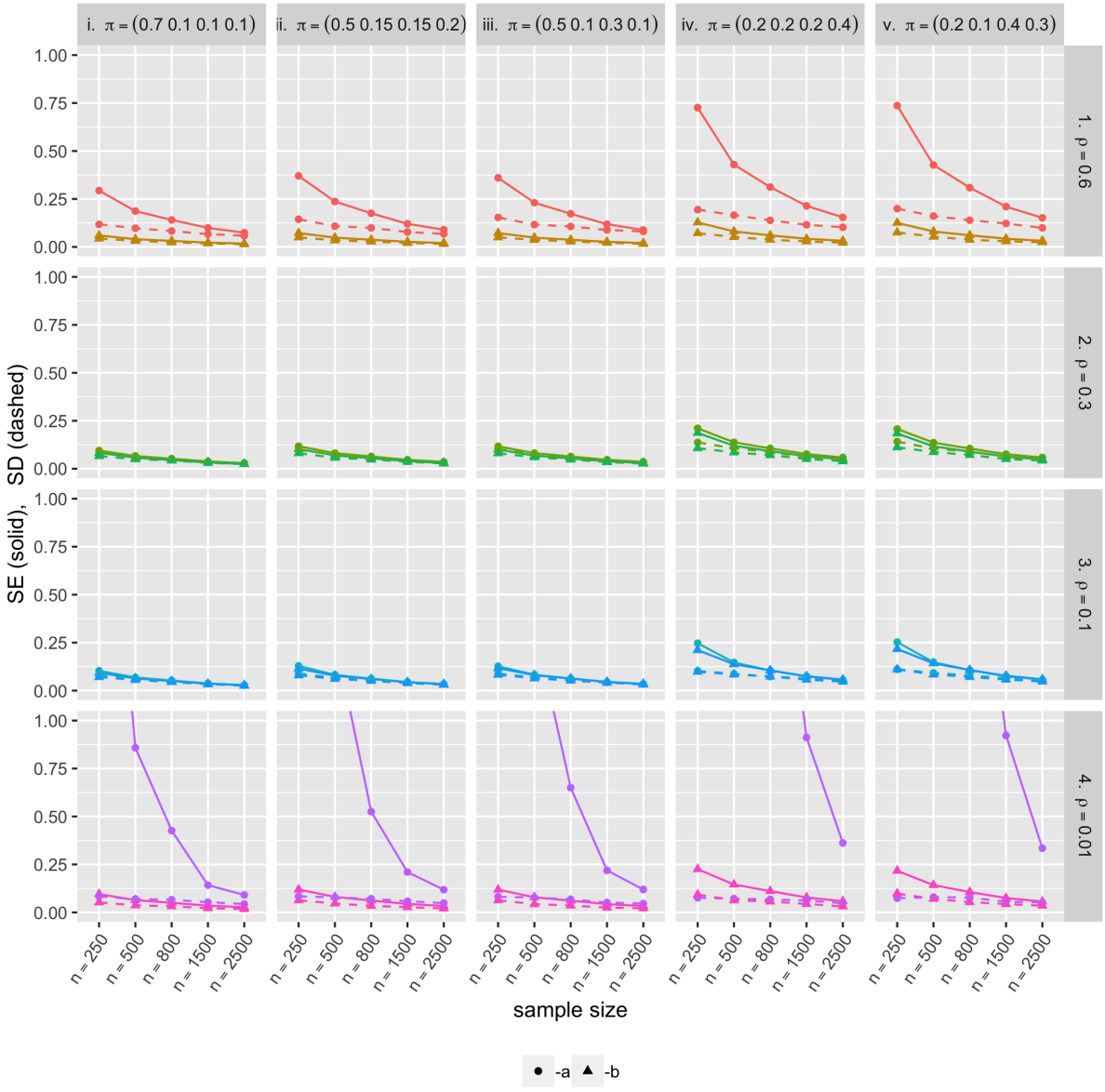

Figure 3.12: Standard error (SE, solid lines) and standard deviation (SD, dashed lines) of the BZINB-based underlying correlation estimates. Each color represents distinct simulation scenarios.

3.6 Discussion

In this chapter, we considered a BZINB model with application to scRNA-seq data where we assume an independent and identically distributed random bivarate sample of zero-inflated counts. One can generalize this homogeneous mean model to allow for subgroup analysis or joint conditional mean analysis by introducing the generalized linear model framework. As in univariate ZINB regression, the latent count variables (i.e., X_1 and X_2) can be modeled using linear predictors with some link function. Alternatively, when the zero-inflation is not cause by dropout but by frailty (e.g., there can be a group of people that do not have cavities at all) and thus has a meaningful interpretation, then the observed count variable (i.e., Y_1 and Y_2) can be directly modeled by the marginalized model framework (Preisser et al., 2016) or the zeros can be separately considered using hurdle models (Mullahy, 1986).

Allowing only positive ρ^* can be regarded as a limitation of the BZINB model. One justification is that the negative correlation of count data are not so prevalent in reality. For example, in genomics data, there are some genes that suppress other genes from being expressed, however, such genes either are relatively rare or have weak negative correlation with other genes. On the other hand, when we believe that the zeros are mostly not induced by dropout events, we can consider using the original correlation ($\rho(Y_1, Y_2)$) which allows for negative correlation, instead of $\rho^*(Y_1, Y_2)$.

An alternative to this fully parametric approach is to use the weighted Pearson correlation based on the parameter estimates. i.e., each pair $(Y_{1,i}, Y_{2,i})$ is given a weight of $w_i = \frac{f_{BNB}(Y_{1,i},Y_{2,i};\hat{\boldsymbol{\theta}})}{f_{BZINB}(Y_{1,i},Y_{2,i};\hat{\boldsymbol{\theta}})}$. Then the weighted Pearson correlation can be calculated as $\tilde{\rho} = \frac{Cov(\boldsymbol{Y}_1,\boldsymbol{Y}_2;\boldsymbol{w})}{Cov(\boldsymbol{Y}_1,\boldsymbol{Y}_1;\boldsymbol{w})Cov(\boldsymbol{Y}_2,\boldsymbol{Y}_2;\boldsymbol{w})}$, where $\boldsymbol{Y}_j = \{Y_{j,1}, ..., Y_{j,n}\}, \boldsymbol{w} = \{w_1, ..., w_n\}, j = 1, 2, Cov(\boldsymbol{Y}_1, \boldsymbol{Y}_2; \boldsymbol{w}) := \frac{\sum_{i=1}^n Y_{1,i}Y_{2,i}w_i}{\sum_{i=1}^n w_i} - \frac{\sum_{i=1}^n Y_{1,i}w_i}{\sum_{i=1}^n w_i}\frac{\sum_{i=1}^n Y_{2,i}w_i}{\sum_{i=1}^n w_i}$. Note that this correlation allows negative values.

As discussed before, the BZINB model can also be generalized to a multivariate zero-inflated negative binomial model. This model may have an exponentially increasing number of latent variables or parameters as the dimension gets large. Though the lack of parsimony may make the multivariate model look less attractive, the idea can be very practically used in

simulating multivariate zero-inflated count data. For instance, a genomic count data with large amount of zeros can be mimicked by a set of latent random layers along with the generalized linear model framework.

3.7 Software

An `R` package `bzinb` estimating BZINB parameters using EM algorithm was written in `R` version 3.5.1 (R Core Team, 2019), and is available on CRAN.

Technical Details for BZINB

The technical details for this chapter include 1) the components of the expected log-likelihood of the EM algorithm and 2) the Standard error formula.

CHAPTER 4: FUTURE RESEARCH

In this chapter, we discuss future research directions for the methods proposed in the three preceding chapters.

In Chapter 1, a tree-based ensemble method for interval censored data was proposed (ICRF). This research was initially motivated by another method development question—finding the optimal individualized treatment rule (ITR) for interval censored or current status data. Now that we have a suitable nonparametric regression estimator for the interval censored data, there are multiple ways of developing the ITR estimator. For example, the Q-learning approach is a natural extension, and adaptation of the outcome weighted learning (OWL) by Zhao et al. (2012) can be another possibility.

In Chapter 2, we introduced a new dynamic treatment regime estimator for right censored data. One interesting extension of the proposed method is optimizing a specific quantile of the survival distribution, such as the median survival time. However, this task introduces a unique challenge concerning backward recursion. Specifically, the optimal decisions made for later stages may not be optimal once earlier decisions moderate the distribution. Thus, an exhaustive search may be needed to find an optimal policy. An extension of the quantile-optimal dynamic treatment regime estimators developed by Wang et al. (2018b) and Linn et al. (2017) to right-censored data would be an interesting future work.

The proposed method assumes an unrestricted policy class. In practice, however, clinicians and patients may prefer understandable, simple rules. A linear rule, for example, is often of interest. By posing a Cox-type proportional hazards assumption on top of our method, the resulting policy class becomes a set of linear functions. The task then reduces to replacing the generalized random survival forest with the generalized Cox model, where the estimated

survival probabilities, instead of the observed time with censoring indicators, are entered as the outcome variable.

In Chapter 3, we developed a parametric bivariate count model based on the latent variable framework and the negative binomial distribution with the aim of measuring dependence between two genes. Being fully parametric, the model may be subject to misspecification. As an alternative, a weighted Pearson correlation can be proposed. Articulating and studying the theoretical nature of this alternative approach could be an intriguing research topic.

In this dissertation, the BZINB model was only estimated marginally, or without covariates. However, a regression estimator can be obtained under the generalized linear model framework. This may allow controlling the sequencing depths varying across sample and testing group differences.

The model estimation is done by the EM algorithm, of which computation time is relatively not ideal. Fast computation is essential for the practical application of this method. For example, when a set of thousand genes are to be analyzed, the number of pairs is about 500,000, which aggravates the computation problem. Thus, developing a computationally feasible solution could be the next step.

We discussed some future research directions, such as generalizations—quantile optimization, Cox decision rule, and the regression BZIN—and extensions—ITR based on ICRF and weighted correlations. We hope that these new research ideas could expand the realm and depth of precision medicine and genomics and, as a result, the general public could be benefited eventually.

APPENDIX 1: TECHNICAL DETAILS FOR CHAPTER 1

This chapter contains technical details including as assumptions, proofs, definitions, and other materials supplemental to the main text of Chapter 1.

A1.1 Proof of GWRS consistency

We prove Theorem 1 (consistency of GWRS).

$$
\begin{aligned}
&|W_n(S_n) - \theta(S_0)| \\
=&|W_n(S_n) - W_n(S_0)| + |W_n(S_0) - \theta(S_0)| \\
=&|(1A)| + |(1B)|,
\end{aligned}
$$

where in what follows we show each term is $o_P(1)$.

$$
\begin{aligned}
|(1A)| =& \left| \frac{1}{n_1 n_2} \sum_{i \in G_1} \sum_{j \in G_2} \zeta(I_{1,i}, I_{2,j} | X_{1,i}, X_{2,j}; S_n) - \zeta(I_{1,i}, I_{2,j} | X_{1,i}, X_{2,j}; S_n) \right| \\
=& \Bigg| \frac{1}{n_1 n_2} \sum_{i \in G_1} \sum_{j \in G_2} \Pr(\mathring{T}_{1,i} < \mathring{T}_{2,i} | \mathring{T}_{1,i} \in I_{1,i}, \mathring{T}_{2,i} \in I_{2,i}, X_{1,i}, X_{2,i}; S_n) \\
&\qquad + \frac{1}{2} \Pr(\mathring{T}_{1,i} = \mathring{T}_{2,i} | \mathring{T}_{1,i} \in I_{1,i}, \mathring{T}_{2,i} \in I_{2,i}, X_{1,i}, X_{2,i}; S_n) \\
&\qquad - \Pr(\mathring{T}_{1,i} < \mathring{T}_{2,i} | \mathring{T}_{1,i} \in I_{1,i}, \mathring{T}_{2,i} \in I_{2,i}, X_{1,i}, X_{2,i}; S_0) \\
&\qquad - \frac{1}{2} \Pr(\mathring{T}_{1,i} = \mathring{T}_{2,i} | \mathring{T}_{1,i} \in I_{1,i}, \mathring{T}_{2,i} \in I_{2,i}, X_{1,i}, X_{2,i}; S_0) \Bigg|
\end{aligned}
$$

$$
\begin{aligned}
=&\frac{1}{n_1 n_2}\Bigg| \sum_{i\in G_1}\sum_{j\in G_2}\int_0^\tau \{\check{S}_n(t|I_{1,i},X_{1,i})-\check{S}_0(t|I_{1,i},X_{1,i})\}dS_n(t|I_{2,i},X_{2,i}) \\
&\qquad -\frac{1}{2}\{S_n(\tau|I_{1,i},X_{1,i})-\check{S}_0(\tau|I_{1,i},X_{1,i})\}S_n(\tau|I_{2,i},X_{2,i}) \\
&\qquad +\int_0^\tau \check{S}_0(t|I_{1,i},X_{1,i})d\{S_n(t|I_{2,i},X_{2,i})-S_0(t|I_{2,i},X_{2,i})\} \\
&\qquad -\frac{1}{2}S_0(\tau|I_{1,i},X_{1,i})\{S_n(\tau|I_{2,i},X_{2,i})-S_0(\tau|I_{2,i},X_{2,i})\}\Bigg| \\
\le&\frac{1}{n_1 n_2}\sum_{i\in G_1}\sum_{j\in G_2}\sup_t\left|\check{S}_n(t|I_{1,i},X_{1,i})-\check{S}_0(t|I_{1,i},X_{1,i})\right| \\
&\qquad +\sup_t\left|S_n(t|I_{2,i},X_{2,i})-S_0(t|I_{2,i},X_{2,i})\right| \\
=&\frac{1}{n_1 n_2}\Bigg|\int_0^\tau \sum_{i\in G_1}\{\check{S}_n(t|I_{1,i},X_{1,i})-\check{S}_0(t|I_{1,i},X_{1,i})\}d\sum_{j\in G_2}S_n(t|I_{2,i},X_{2,i}) \\
&\qquad -\frac{1}{2}\sum_{j\in G_1}\{S_n(\tau|I_{1,i},X_{1,i})-\check{S}_0(\tau|I_{1,i},X_{1,i})\}\sum_{j\in G_2}S_n(\tau|I_{2,i},X_{2,i}) \\
&\qquad +\int_0^\tau \sum_{j\in G_1}\check{S}_0(t|I_{1,i},X_{1,i})d\sum_{j\in G_2}\{S_n(t|I_{2,i},X_{2,i})-S_0(t|I_{2,i},X_{2,i})\} \\
&\qquad -\frac{1}{2}\sum_{j\in G_1}S_0(\tau|I_{1,i},X_{1,i})\sum_{j\in G_2}\{S_n(\tau|I_{2,i},X_{2,i})-S_0(\tau|I_{2,i},X_{2,i})\}\Bigg| \\
\le&\sup_t\left|\frac{1}{n_1}\sum_{i\in G_1}\{\check{S}_n(t|I_{1,i},X_{1,i})-\check{S}_0(t|I_{1,i},X_{1,i})\}\right| \\
&\qquad +\sup_t\left|\frac{1}{n_2}\sum_{i\in G_2}\{S_n(t|I_{2,i},X_{2,i})-S_0(t|I_{2,i},X_{2,i})\}\right|.
\end{aligned}
$$

We further show that $\sup_{t\in[0,\tau]}\left|\frac{\mathbb{P}_n\check{S}_n(t|I,X)1(X\in G_l)}{\mathbb{P}_n 1(X\in G_l)}-\frac{\mathbb{P}_n\check{S}_0(t|I,X)1(X\in G_l)}{\mathbb{P}_n 1(X\in G_l)}\right|=o_P(1), l=1,2,$ where $\frac{\mathbb{P}_n\check{S}_n(t|I,X)1(X\in G_l)}{\mathbb{P}_n 1(X\in G_l)}=\frac{1}{n_1}\sum_{i\in G_1}\check{S}_n(t|I_{1,i},X_{1,i})$. We then have

$$\sup_{t\in[0,\tau]}\left|\frac{\mathbb{P}_n\check{S}_n(t|I,X)1(X\in G_l)}{\mathbb{P}_n1(X\in G_l)}-\frac{\mathbb{P}_n\check{S}_0(t|I,X)1(X\in G_l)}{\mathbb{P}_n1(X\in G_l)}\right| \tag{4.15}$$

$$=\sup_t\frac{\mathbb{P}_n\{(\check{S}_n(t|X,I)-\check{S}_0(t|X,I)\}1(X\in G_l)}{P1(X\in G_l)(1+o_P(1))} \tag{4.16}$$

$$\leq\sup_{t,x}|\mathbb{P}_n\{\check{S}_n(t|X,I)-\check{S}_0(t|X,I)\}|\lambda_l^{-1}(1+o_P(1))$$

$$\leq\sup_{t,x}\left|\mathbb{P}_n1(t\in[L,R))\left\{\frac{\check{S}_0(t|X)-\check{S}_0(R|X)}{\check{S}_0(L|X)-\check{S}_0(R|X)}-\frac{\check{S}_n(t|X)-\check{S}_n(R|X)}{\check{S}_n(L|X)-\check{S}_n(R|X)}\right\}\right|$$

$$\leq\sup_{t,x}\left|\mathbb{P}_n\frac{1(t\in[L,R))}{\check{S}_0(L|X)-\check{S}_0(R|X)}\left\{\check{S}_0(t|X)-\check{S}_0(R|X)-\check{S}_n(t|X)+\check{S}_n(R|X)\right\}\right|$$

$$+\sup_{t,x}\left|\mathbb{P}_n\frac{1(t\in[L,R))}{\check{S}_0(L|X)-\check{S}_0(R|X)}\left\{\check{S}_0(L|X)-\check{S}_0(R|X)-\check{S}_n(L|X)+\check{S}_n(R|X)\right\}\right|$$

$$\leq4\left|\sup_{t,x}\left\{\check{S}_n(t|x)-\check{S}_0(t|x)\right\}\right|\sup_t\mathbb{P}_n\frac{1(t\in[L,R))}{\check{S}_0(L|X)-\check{S}_0(R|X)}$$

$$\leq4\left|\sup_{t,x}\left\{\check{S}_n(t|x)-\check{S}_0(t|x)\right\}\right|(1+o_P(1))$$

$$=o_P(1),$$

where the last inequality is due to

$$\mathbb{P}_n\frac{1(t\in[L,R))}{\check{S}_0(L|X)-\check{S}_0(R|X)}$$

$$=\mathbb{P}_n\frac{1(t\in[L,R))}{S_0(L|X)-S_0(R|X)} \qquad \text{for continuous } S_0,$$

$$=\int\frac{1(l\leq t<r))}{\Pr(l\leq T<r|X=x)}d\mathbb{P}_n(l,r,x)$$

$$=\int\frac{1(l\leq T<r)}{\Pr(l\leq T<r|X=x)}\frac{1(l\leq t<r)}{1(l\leq T<r)}d\mathbb{P}_n(T,l,r,x)$$

$$\leq\sqrt{\int\frac{1(l\leq T<r)}{\Pr(l\leq T<r|X=x)}d\mathbb{P}_n(T,l,r,x)}\sqrt{\int\frac{1(l\leq t<r)}{1(l\leq T<r)}d\mathbb{P}_n(T,l,r,x)}$$

$$\leq \sqrt{\int \underbrace{\int \frac{1(l \leq T < r)}{\Pr(l \leq T < r | X = x)} d\mathbb{P}_n(T, l, r | X = x)}_{=1+o_P(1)} d\mathbb{P}_{nX}(x)}$$

$$\times \sqrt{\int \frac{1(l \leq t < r)}{1(l \leq T < r)} d\mathbb{P}_n(T, l, r, x)} \qquad \text{with the denominator} \leq 1 \text{ with probability 1.}$$

$$\leq 1 + o_P(1).$$

Now we show $(1B) = o_P(1)$ to conclude proof of Theorem 1.

$$\begin{aligned}
|(1B)| =& |W_n(S_0) - \theta(S_0)| \\
=& \Bigg| \int_0^\tau \left\{ \underbrace{\frac{1}{n_1} \sum_{i \in G_1}}_{=:\mathbb{P}_{n,1}} S_0(t | I_{1,i}, X_{1,i}) \right\} d \left\{ \underbrace{\frac{1}{n_2} \sum_{j \in G_2}}_{=:\mathbb{P}_{n,2}} S_0(t | I_{2,i}, X_{2,i}) \right\} \\
& \qquad - \frac{1}{2} \left\{ \frac{1}{n_1} \sum_{j \in G_1} S_0(\tau | I_{1,i}, X_{1,i}) \right\} \left\{ \frac{1}{n_2} \sum_{j \in G_2} S_0(\tau | I_{2,i}, X_{2,i}) \right\} \\
& - \int_0^\tau S_0(t | X \in G_1) dS_0(t | X \in G_2) + \frac{1}{2} S_0(\tau | X \in G_1) S_0(\tau | X \in G_2) \Bigg| \\
=& \Bigg| \int_0^\tau \left\{ P_{n,1} S_0(t | I, X) - S_0(t | X \in G_1) \right\} dP_{n,2} S_0(t | I, X) \\
& \qquad - \frac{1}{2} \left\{ P_{n,1} S_0(\tau | I, X) - S_0(\tau | X \in G_1) \right\} P_{n,2} S_0(\tau | I, X) \\
& \qquad + \int_0^\tau S_0(t | X \in G_1) d \{ P_{n,2} S_0(t | I, X) - S_0(t | X \in G_2) \} \\
& \qquad - \frac{1}{2} S_0(\tau | X \in G_1) \{ P_{n,2} S_0(t | I, X) - S_0(\tau | X \in G_2) \} \Bigg| \\
\leq & \sum_{l=1}^{2} \sup_t \sup_t |P_{n,l} S_0(t | I, X) - S_0(t | X \in G_l)| \\
=& o_P(1).
\end{aligned}$$

To see the last equality, we have

$$
\begin{aligned}
&\sup_t |P_{n,l} S_0(t|I,X) - S_0(t|X \in G_l)| \\
=&\sup_t \left| \frac{\mathbb{P}_n S_0(t|I,X) 1(X \in G_l)}{\mathbb{P}_n 1(X \in G_l)} - \frac{P S_0(t|I,X) 1(X \in G_l)}{P 1(X \in G_l)} \right| \\
\leq&\sup_t \left| \frac{(\mathbb{P}_n - P) S_0(t|I,X) 1(X \in G_l)}{\mathbb{P}_n 1(X \in G_l)} \right| + \sup_t \frac{P S_0(t|I,X) 1(X \in G_l)}{P 1(X \in G_l)} \left| \frac{(\mathbb{P}_n - P) 1(X \in G_l)}{\mathbb{P}_n 1(X \in G_l)} \right| \\
=&o_P(1),
\end{aligned}
$$

by the law of large numbers and Slutsky's lemma.

A1.2 Proof of GLR consistency

We prove Theorem 2 that states consistency of the GLR statistic.

Note that $LR_n(S_n) = g(\begin{pmatrix} Y_1(\cdot; S_n) \\ Y_2(\cdot; S_n) \end{pmatrix})$ and $\rho(S_0) = g(\begin{pmatrix} S_0(\cdot|G_1) \\ S_0(\cdot|G_2) \end{pmatrix})$, where g is a continuous map from $D_{[0,1]}[0,\tau]$ to $\mathbb{R}_+$ and $D_{[0,1]}[0,\tau]$ is the space of cadlag (right-continuous with left-hand limits) functions bounded by 0 and 1 with support $[0,\tau]$. The continuity of g can be shown without difficulty using convergence theorems for integration maps (see, e.g., Proposition 7.27 of Kosorok (2007)). If we show $\sup_{t,l} |Y_l(t; S_n) - S_0(t|G_l)| \to_p 0$, by the functional continuous mapping theorem, $LR_n(S_n) \to_p \rho(S_0)$. Thus, it remains to show $\sup_{t,l} |Y_l(t|S_n) - S_0(t|G_l)| \to_p 0$.

Let $\lambda_l = \lim_{n\to\infty} \lambda_{n,l}, l = 1, 2$. For each $l = 1, 2$, we have the following decomposition.

$$
\begin{aligned}
&Y_l(t; S_n) - S_0(t|G_l) \\
=&\frac{\mathbb{P}_n S_n(t|X,I) 1(X \in G_l)}{\mathbb{P}_n 1(X \in G_l)} - \frac{P S_0(t|X) 1(X \in G_l)}{P 1(X \in G_l)} \\
=&\frac{(\mathbb{P}_n - P) S_n(t|X,I) 1(X \in G_l)}{\mathbb{P}_n 1(X \in G_l)} + \qquad (2A)
\end{aligned}
$$

$$\frac{P\{(S_n(t|X,I) - S_0(t|X,I)\}1(X \in G_l)}{\mathbb{P}_n 1(X \in G_l)} + \tag{2B}$$

$$\frac{P\{S_0(t|X,I) - S_0(t|X)\}1(X \in G_l)}{\mathbb{P}_n 1(X \in G_l)} + \tag{2C}$$

$$PS_0(t|X)1(X \in G_l)\left\{\frac{1}{\mathbb{P}_n 1(X \in G_l)} - \frac{1}{P1(X \in G_l)}\right\}. \tag{2D}$$

In (2A), since $S_n(t|X,I)$ is a stochastic process that is monotone in t, by Lemma 9.10 of Kosorok (2007), this process $\{S_n(t|X,I) : t\}$ has VC-dimension 2 and, thus, is a Glivenko-Cantelli class. Also since any finite number of fixed sets form a Glivenko-Cantelli class and any collection of elementwise products of Glivenko-Cantelli classes that are bounded are again a Glivenko-Cantelli class, $\{S_n(t|X,I)1(X \in G_l) : t, l = 1, 2\}$ is Glivenko-Cantelli. Thus by the Glivenko-Cantelli Theorem,

$$\sup_t |(1A)| \leq \{(\mathbb{P}_n - P)S_n(t|X,I)1(X \in G_l)\}\,\lambda_l^{-1}(1 + o_P(1)) \to 0,$$

where we used the fact that $\mathbb{P}_n 1(X \in G_l) = \lambda_l(1 + o_P(1))$ and the numerator being asymptotically bounded by twice the denominator in absolute values.

$$\begin{aligned}
\sup_t |(2B)| =& \sup_t \frac{P\{(S_n(t|X,I) - S_0(t|X,I)\}1(X \in G_l)}{P1(X \in G_l)}(1 + o_P(1)) \qquad (4.17)\\
\leq& \sup_t |P\{S_n(t|X,I) - S_0(t|X,I)\}|\lambda_l^{-1}(1 + o_P(1))\\
\leq& \sup_t \left| P1(t \in [L,R))\left\{\frac{S_0(t|X) - S_0(R|X)}{S_0(L|X) - S_0(R|X)} - \frac{S_n(t|X) - S_n(R|X)}{S_n(L|X) - S_n(R|X)}\right\}\right|\\
\leq& \sup_t \left| P\frac{1(t \in [L,R))}{S_0(L|X) - S_0(R|X)}\{S_0(t|X) - S_0(R|X) - S_n(t|X) + S_n(R|X)\}\right|\\
&+ \sup_t \left| P\frac{1(t \in [L,R))}{S_0(L|X) - S_0(R|X)}\{S_0(L|X) - S_0(R|X) - S_n(L|X) + S_n(R|X)\}\right|
\end{aligned}$$

$$
\begin{aligned}
&\leq 4 \left| \sup_{t,x} \{S_n(t|x) - S_0(t|x)\} \right| \sup_t P \frac{1(t \in [L, R))}{S_0(L|X) - S_0(R|X)} \\
&\leq 4 \left| \sup_{t,x} \{S_n(t|x) - S_0(t|x)\} \right| \\
&= o_P(1),
\end{aligned}
$$

where the last inequality is due to

$$
\begin{aligned}
& P \frac{1(t \in [L, R))}{S_0(L|X) - S_0(R|X)} \\
&= \int \frac{1(l \leq t < r))}{\Pr(l \leq T < r | X = x)} dP(l, r, x) \\
&= \int \frac{1(l \leq T < r)}{\Pr(l \leq T < r | X = x)} \frac{1(l \leq t < r)}{1(l \leq T < r)} dP(T, l, r, x) \\
&\leq \sqrt{\int \frac{1(l \leq T < r)}{\Pr(l \leq T < r | X = x)} dP(T, l, r, x)} \sqrt{\int \frac{1(l \leq t < r)}{1(l \leq T < r)} dP(T, l, r, x)} \\
&\leq \sqrt{\int \underbrace{\int \frac{1(l \leq T < r)}{\Pr(l \leq T < r | X = x)} dP(T, l, r | X = x)}_{=1} dP_X(x)} \\
&\quad \times \sqrt{\int \frac{1(l \leq t < r)}{1(l \leq T < r)} dP(T, l, r, x)} \qquad \text{with the denominator} \leq 1 \text{ with probability 1.} \\
&\leq 1.
\end{aligned}
$$

$$
\begin{aligned}
\sup_t |(2C)| &= \sup_t \left\{ P\{S_0(t|X, I) - S_0(t|X)\} 1(X \in G_l) \right\} \lambda_l^{-1} (1 + o_P(1)) \\
&\leq \sup_t \left\{ P\{S_0(t|X, I) - S_0(t|X)\} \right\} \lambda_l^{-1} (1 + o_P(1)) \\
&= 0,
\end{aligned}
$$

where the last equality is from the fact that both $S_0(t|X, I)$ and $S_0(t|X)$ can be written as expectation with only difference in the conditioning argument that are marginalized out by the population average operator P (the double expectation). Finally,

$$\begin{aligned}
\sup_t |(2D)| =& \sup_t PS_0(t|X)1(X \in G_l) \left\{ \frac{1}{\mathbb{P}_n 1(X \in G_l)} - \frac{1}{P1(X \in G_l)} \right\} \\
\leq& P1(X \in G_l) \left\{ \frac{1}{\mathbb{P}_n 1(X \in G_l)} - \frac{1}{P1(X \in G_l)} \right\} \\
\leq& \frac{P1(X \in G_l) - \mathbb{P}_n 1(X \in G_l)}{\mathbb{P}_n 1(X \in G_l)} \\
=& o_P(1).
\end{aligned}$$

Therefore, the desired result holds.

A1.3 Proof of uniform consistency of interval censored recursive forests

A1.3.1 Overview of the proof of Theorem 3

It suffices to prove the theorem for a single iteration, because, for a large sample, terminal node size becomes arbitrarily small with the potential splitting bias being eliminated and as a result recursion does not add to bias reduction.

We borrow the strategy used in Cho et al. (2020) in establishing the uniform consistency of random survival forests, which uses empirical process theory for right censored data. There is a unique challenge in applying the approach to interval censored survival regression problems—namely the identifiability issue. In Cho et al. (2020), the Z-estimator theorem (Theorem 2.10) in Kosorok (2007) could be used without such an issue, since the self-consistency algorithm gives a unique solution for right-censored data. However, for interval censored data, the self-consistency algorithm may not identify the global maximum of the likelihood. Thus, a careful handling of the identifiability condition is required.

The main technique used to guarantee identifiability is to restrict the class of candidate survival functions to those which satisfy the identifiability condition given data. If one can show that the unrestricted class of the estimating equations is Glivenko-Cantelli (Kosorok, 2007; van der Vaart and Wellner, 2013), the resulting theoretical property—uniform convergence of empirical processes—is inherited to smaller, restricted classes. This is true even if the restriction is done in a data-dependent fashion, since any subset of a Glivenko-Cantelli class is also a Glivenko-Cantelli class. In this way, the desired result, or uniform consistency, can be established.

Noting that NPMLEs have uniform-over-time consistency (Groeneboom and Wellner, 1992) in the non-regression context and that the unique NPMLEs can be estimated through the iterative convex minorant (ICM) algorithms (Groeneboom, 1991; Jongbloed, 1998; Wellner and Zhan, 1997), the problem now reduces to incorporating the identifiability restriction into the estimating equation and extending the uniform consistency results to the regression context.

A1.3.2 The Z-estimator framework

Now we give a detailed proof of Theorem 3 with an introduction to some basic notation for self-consistency equations for non-regression settings and extend the notation to regression settings. The self-consistency equation, without covariates, for case-II censoring with two monitoring times can be expressed as

$$\mathbb{P}_n \psi_{S,t}^{(m)} = 0 \ \ \forall t \in [0, \tau], \tag{4.18}$$

where we put superscript (m) to denote that this is for marginal, or non-regression, settings, $\psi_{S,t}^{(m)} \equiv \eta_1 \frac{S(t)-S(U)}{1-S(U)} \vee 0 \ + \ \left(\eta_2 \frac{S(t)-S(V)}{S(U)-S(V)} \vee 0\right) \wedge 1 \ + \eta_3 \frac{S(t)}{S(V)} \wedge 1 - S(t)$, U and V are the ordered monitoring times, $\eta_1 = 1(T \leq U), \eta_2 = 1(U < T \leq V)$, $\eta_3 = 1 - \eta_1 - \eta_2$, $\wedge$ and $\vee$ are the minimum and the maximum operators, and $\mathbb{P}_n$ is the empirical measures of given data of size n. $\mathbb{P}_n$ is, at the same time, used to denote the sample average operator such that, given

a function $f : \mathcal{X} \mapsto \mathbb{R}$ that maps the sample space to the real space, $\mathbb{P}_n f = \int f(x) d\mathbb{P}_n(x) = \frac{1}{n}\sum_{i=1}^{n} f(X_i)$, where X_i is the ith random entry of the data, or $X_i = (U_i, V_i, \eta_{1,i}, \eta_{2,i})$ in this specific problem.

The following lemma is a restatement of Theorem 2.10 of Kosorok (2007). In the survival regression setting, we let $\Psi : \Theta \mapsto \mathbb{L}$ be a map between two normed spaces, where Θ is the space of all marginal survival functions with time ranging over $[0, \tau]$, $\mathbb{L}$ is a normed space of right-coninuous-over-time functions with support $[0, \tau]$ and range $[-1, 1]$, $\|\cdot\|_{\mathbb{L}}$ denotes the uniform norm over $[0, \tau]$, Ψ is a fixed map, and Ψ_n is a data-dependent map.

Lemma 4.1 (Consistency of Z-estimators)**.** *Let $\Psi(S_0) = 0$ for some $S_0 \in \Theta$, and assume $\|\Psi(S_n)\|_{\mathbb{L}} \to 0$ implies $\|S_n - S_0\|_{\mathbb{L}} \to 0$ for any sequence $\{S_n\} \in \Theta$. Then, if $\|\Psi_n(\hat{S}_n)\|_{\mathbb{L}} \to 0$ in probability for some sequence of estimators $\hat{S}_n \in \Theta$ and $\sup_{S\in\Theta} \|\Psi_n(S) - \Psi(S)\|_{\mathbb{L}} \to 0$ in probability, $\|\hat{S}_n - S_0\|_{\mathbb{L}} \to 0$ in probability, as $n \to \infty$.*

Now we adapt the lemma to address the identifiability issue by introducing a necessary and sufficient condition (Gentleman and Geyer, 1994) for an NPMLE S to be unique:

$$\mathbb{P}_n \phi_{S,t}^{(m)} \leq 1 \ \ \forall t \in [0, \tau], \tag{4.19}$$

where $\phi_{S,t}^{(m)} = \eta_1 \frac{1(t \leq U)}{1 - S(U)} + \eta_2 \frac{1(U < t \leq V)}{S(U) - S(V)} + \eta_3 \frac{1(t > V)}{S(V)}$. This condition guarantees that the NPMLE $\hat{S}$ that satisfies $\mathbb{P}_n \phi_{S,t}^{(m)} \leq 1$ is the global maximum and thus identifies the true S_0 at its limit given self-consistency. Thus, if we restrict the space to $\Theta_n = \{S : \sup_{t\in[0,\tau]} \mathbb{P}_n \phi_{S,t}^{(m)} \leq 1\}$, within the resticted space, only the unique NPMLE $S = \hat{S}_n$ satisfies the estimating equation (4.18). This space Θ_n is adaptively defined as it depends on a specific data set. This sequence of spaces always exists, because, given data, the NPMLE can be uniquely estimated via the ICM algorithm. Now the lemma is adapted in the following corollary to reflect the restriction and to be further used for the regression setting.

Corollary 4.1 (Consistency of Z-estimators)**.** *Let Θ be a class of all covariate-conditional survival functions $S : [0, \tau] \times \mathcal{X} \mapsto [0, 1]$ and let $\Psi : \Theta \mapsto \mathbb{L}$ where $\mathbb{L}$ is some normed space*

of functions $S : [0, \tau] \times \mathcal{X} \mapsto [-1, 1]$. *(i) Let* $\Psi(S_0) = 0$ *for some* $S_0 \in \Theta$. *(ii) Assume that there exists a sequence of subclasses* Θ_n *such that for any sequence* $\{S_n \in \Theta_n\}$, $\|\Psi(S_n)\|_{\mathbb{L}} \to 0$ *implies* $\|S_n - S_0\|_{\mathbb{L}} \to 0$. *(iii) Further assume that* $\|\Psi_n(\hat{S}_n)\|_{\mathbb{L}} \to 0$ *in probability for some sequence of estimators* $\{\hat{S}_n \in \Theta_n\}$. *Then, if (iv)* $\sup_{S \in \Theta_n} \|\Psi_n(S) - \Psi(S)\|_{\mathbb{L}} \to 0$ *in probability,* $\|\hat{S}_n - S_0\|_{\mathbb{L}} \to 0$ *in probability, as* $n \to \infty$.

We first define the regression-version estimating equations, Ψ and Ψ_n, as below.

$$\begin{aligned}\Psi(S) =& \frac{P\psi_{S,t}\delta_x}{P\delta_x} \equiv P_{\cdot|x}\psi_{S,t},\\ \Psi_n(S) =& \frac{\mathbb{P}_n\psi_{S,t}k_x}{\mathbb{P}_n k_x} \equiv \mathbb{P}_{n,\cdot|k_x}\psi_{S,t},\end{aligned}$$

where

$$\psi_{S,t} = \eta_1 \frac{S(t|X) - S(U|X)}{1 - S(U|X)} \vee 0 \ + \ \eta_2 \frac{S(t|X) - S(V|X)}{S(U|X) - S(V|X)} \vee 0 \wedge 1 \ + \eta_3 \frac{S(t|X)}{S(V|X)} \wedge 1 - S(t|X), \tag{4.20}$$

P is the population version of $\mathbb{P}_n$ so that, $Pf = \int f(x)dP(x)$, $\delta_x = I(\cdot = x)$ is the unnormalized Dirac measure, and k_x the unnormalized forest kernel. To be more specific, $k_x = \frac{1}{n_{\text{tree}}} \sum_{b=1}^{n_{\text{tree}}} 1(x \in L_b(x))$, where n_{tree} is the number of trees in the forest, $L_b(x)$ is the terminal node of the bth tree of the forest that contains the point x. We use the term 'unnormalized' to mean that they are not multiplied by the sample (or the population) size. By using the subscripts $\cdot \mid x$ and $\cdot \mid k_x$ we denote the conditional probability measures, where the latter is a probability measure weighted by the kernel k_x.

Note that for ease of theoretical exposition, we assume that the terminal node prediction of random survival forests is given by the NPMLEs of observations weighted by the forest kernels, instead of averaging the NPMLEs of each tree. This kernel weighting approach is often taken in the literature (Athey et al., 2019; Yao et al., 2019) and is equivalent to the average of tree predictions for non-censored mean outcomes. Although for censored data, these two approaches

are not equal in general, they can be shown equivalent for $B >> 1$, as the former can be seen as the average of estimates of random subsamples from the kernel-weighted population from which the latter is estimated.

The identifiability condition (4.19) in the marginal context is now replaced by

$$\frac{\mathbb{P}_n \phi_{S,t} k_x}{\mathbb{P}_n k_x} \leq 1 \ \ \forall t \in [0, \tau], x \in \mathbb{R}^d. \tag{4.21}$$

Note that the kernel k not only depends on given data but also has extra randomness due to subsampling, random variable subsetting, and/or random cut-off selection. In other words, given data, k is formed as a result of a realized partition of the trees or the forests. Similarly to the restriction done in the marginal setting, the class Θ of covariate-conditional survival curves can also be restricted to $\Theta_{n,k}$ given data and a specific partition (or the kernel k). In other words, $\Theta_{n,k} = \{S : S \in \Theta, \sup_{t \in [0,\tau], x \in \mathcal{X}} \frac{\mathbb{P}_n \phi_{S,t} k_x}{\mathbb{P}_n k_x} \leq 1\}$.

Hence Theorem 3 will follows if we can show that the conditions of Corollary A4.1 hold. First, that (i) $\Psi(S_0) = 0$ for some $S_0 \in \Theta$ is trivial. The second condition, (ii) existence of a restricted set Θ_n with which $\|\Psi(S_n)\|_{\mathbb{L}} \to 0$ implies $\|S_n - S_0\|_{\mathbb{L}} \to 0$ for any sequence $\{S_n \in \Theta_{n,k}\}$, can be shown to be satisfied by verifying the assumptions of Lemma 4.2 in Section A1.3.3. The third condition, (iii), is met, since the kernel-weighed NPMLE is the solution to $\|\Psi_n(\hat{S})n)\|_{\mathbb{L}}$. The last condition, (iv), is checked in Section A1.3.4 below.

A1.3.3 Uniform identifiability

We introduce additional notation for Lemma 4.2. Let $\mathcal{Q}$ denote the space of all survival functions on $S : [0, \tau] \mapsto [0, 1]$, $S_0 : \mathcal{X} \mapsto \mathcal{Q}$ denote the true survival functions, and $\mathcal{Q}_0 = \{S_0(x) : x \in \mathcal{X}\}$ be the collection of S_0's. Let $\Phi : \mathcal{Q} \times \mathcal{Q} \mapsto \mathbb{R}$ be the function that takes $S_1, S_2 \in \mathcal{Q}$ and computes the supremum over $[0, \tau]$ of the absolute value of a certain estimating equation, where S_2 is the true survival function and S_1 is the candidate survival functions.

Assumption 4.16 (Closed covariate space, compact and continuous true survival space)**.** *(i) $\mathcal{X}$ is closed, (ii) $\mathcal{Q}_0$ is compact with respect to the uniform norm on $\mathcal{Q}$, and (iii) for all sequence $\{x_n\} \in \mathcal{X}$ such that $x_n \to x_1$, $\|S_0(x_n) - S_0(x_1)\|_\infty \to 0$.*

Assumption 4.17 (local identifiability)**.** *For every sequence $S_n^* \in D$ (and also in Θ_n), and every seqeunce $\{x_n\} \in H : x_n \to x_1$, we have that $\Phi(S_n^*, S_0(x_n)) \to 0 \Rightarrow \|S_n^* - S_0(x_1)\|_\infty \to 0$.*

Lemma 4.2 (uniform identifiability)**.** *Assume Assumptions 4.16–4.17. Suppose $\forall x \in \mathcal{X}$, $S_n(x)$ is a sequence $\in \mathcal{Q}$ and suppose $\sup_{x\in\mathcal{X}} \Phi(S_n(x), S_0(x)) \to 0$. Then $\sup_{x\in\mathcal{X}} \|S_n(x) - S_0(x)\|_\infty \to 0$.*

Proof. Assume that $\sup_{x\in\mathcal{X}} \Phi(S_n(x), S_0(x)) \to 0$ but $\sup_{x\in\mathcal{X}} \|S_n(x) - S_0(x)\|_\infty \not\to 0$. Then there exists a subsequence n' and an associated sequence $x_{n'}$ such that $\|S_{n'}(x_{n'}) - S_0(x_{n'})\|_\infty \to c > 0$. Also, there exists, for this subsequence, n'' such that $x_{n''} \to x_1$ for some $x_1 \in \mathcal{X}$ (by compactness of $\mathcal{Q}_0$).

By Assumption 2, with $S_{n''}^* = S_{n''}(x_{n''})$, we obtain that $\Phi(S_{n''}^*, S_0(x_{n''})) \to 0 \Rightarrow \|S_{n''}(x_{n''}) - S_0(x_{n''})\|_\infty \to 0$. This is a contradiction. Thus, the conclusion of the lemma holds. ■

Assumptions 2 and 3 (Lipschitz continuity and bounded and closed covariate space) are sufficient for Assumption A4.16. Specifically, (ii) is obtained from the Ascoli-Arzelá Theorem. If we show that the interval censored recursive forest satisfies Assumption 4.17, the result of Lemma 4.2 holds. While this identifiability result is valid with Φ function, the second condition of Corollary A4.1 which relies on Ψ function, or the first term of the Φ function, is always satisfied within the restricted class Θ_n. Such a sequence Θ_n of spaces exists, because, given data, the NPMLE can be uniquely estimated via the ICM algorithm.

A1.3.4 Consistency of the estimating function

Finally, we show how the last condition is fulfilled. We decompose the quantity into three components and bound the error.

$$
\begin{aligned}
&\sup_{S\in\Theta_n} \|\Psi_n(S)) - \Psi(S)\|_{\mathbb{L}} \\
\leq &\sup_{S\in\Theta_n, t\in[0,\tau], x\in\mathcal{X}} \left| \frac{\mathbb{P}_n \psi_{S,t} k_x}{\mathbb{P}_n k_x} - \frac{P\psi_{S,t}k_x}{Pk_x} \right| + \sup_{S\in\Theta_n, t\in[0,\tau], x\in\mathcal{X}} \left| \frac{P\psi_{S,t}k_x}{Pk_x} - \frac{P\psi_{S,t}\delta_x}{P\delta_x} \right| \\
\leq &\underbrace{\sup_{S\in\Theta_n, t\in[0,\tau], x\in\mathcal{X}} \left| \frac{(\mathbb{P}_n - P)\psi_{S,t}k_x}{\mathbb{P}_n k_x} \right|}_{=(3A)} + \underbrace{\sup_{S\in\Theta_n, t\in[0,\tau], x\in\mathcal{X}} \left| P\psi_{S,t}k_x \{ \frac{1}{\mathbb{P}_n k_x} - \frac{1}{Pk_x} \} \right|}_{=(3B)} \\
&+ \underbrace{\sup_{S\in\Theta_n, t\in[0,\tau], x\in\mathcal{X}} \left| \frac{P\psi_{S,t}k_x}{Pk_x} - \frac{P\psi_{S,t}\delta_x}{P\delta_x} \right|}_{=(3C)}.
\end{aligned}
$$

We use empirical process theory to bound the error of $(3A)$. The class of functions $\{\psi_{S,t} : S \in \Theta_n, t \in [0,\tau]\}$ can be shown to be a Donsker class. To see this, notice that each of the four terms in (4.20) is a monotone stochastic process and, thus, is a VC class according to Lemma 9.10 of Kosorok (2007). As a finite sum of VC classes is a VC class and a VC class endowed with a bounded envelope—in this case $F = 1$—is Donsker, the class of $\psi_{S,t}$ functions is a Donsker class. Next, since k_x can be shown to be a Donsker class by Proposition 6 (Bounded entropy integral of the tree and forest kernels) of Cho et al. (2020) and is bounded above by 1, the class of their products $\{\psi_{S,t}\delta_x : S \in \Theta_n, t \in [0,\tau]\}$ is again Donsker. Consequently, $\sup_{S\in\Theta_n, t\in[0,\tau]} \left| (\mathbb{P}_n - P)\psi_{S,t}k_x \right| = O_P(n^{-1/2})$. Meanwhile since the denominator $\mathbb{P}_n k_x \asymp n^{\beta-1}$ by the assumption of terminal node size, we have $(3A) = o_P(1)$

$(3B) = o_P(1)$ can also be shown similarly.

$$
\begin{aligned}
(3B) &= \sup_{S\in\Theta_n, t\in[0,\tau], x\in\mathcal{X}} \left| P\psi_{S,t}k_x \{\frac{1}{\mathbb{P}_n k_x} - \frac{1}{Pk_x}\} \right| \\
&= \sup_{S\in\Theta_n, t\in[0,\tau], x\in\mathcal{X}} \left| \underbrace{\frac{P\psi_{S,t}k_x}{\mathbb{P}_n k_x}}_{=O_P(1)} \frac{(\mathbb{P}_n - P)k_x}{\mathbb{P}_n k_x} \right| \\
&= O_P(1) O_P(n^{-1/2-(\beta-1)}) = o_P(1).
\end{aligned}
$$

Finally, we show $(3C) = o_P(1)$. We first note that

$$
\begin{aligned}
&P_{\cdot\,|k_x}\psi_{S,x} \\
&= 1 - G_U(t|k_x) - \int_{u=t}^{\infty} \{1 - S_0(u|k_x)\} \frac{1 - S(t|x)}{1 - S(u|x)} dG_U(u|k_x) \\
&+ \int_{u=0}^{t}\int_{v=t}^{\infty} S_0(u|k_x) dG(u,v|k_x) + \int_0^t \frac{S(t|x)}{S(v|x)} S_0(v|k_x) dG_V(v|k_x) \\
&- \int_{u=0}^{t}\int_{v=t}^{\infty} \frac{S(u|x) - S(t|x)}{S(u|x) - S(v|x)} (S_0(u|k_x) - S_0(v|k_x)) dG(u,v|k_x) - S(t|x) \\
&= 1 - G_U(t|k_x) - S(t|x) \\
&- \int_{u=t}^{\infty} R_1(t,u,x)\{1 - S_0(u|k_x)\} dG_U(u|k_x) \\
&+ \int_{u=0}^{t}\int_{v=t}^{\infty} S_0(u|k_x) dG(u,v|k_x) + \int_0^t R_2(t,v,x) S_0(v|k_x) dG_V(v|k_x) \\
&- \int_{u=0}^{t}\int_{v=t}^{\infty} R_3(t,u,v,x)(S_0(u|k_x) - S_0(v|k_x)) dG(u,v|k_x),
\end{aligned}
$$

where $0 \le R_1(t,u,x), R_2(t,v,x), R_3(t,u,v,x) \le 1$ are decreasing in u and increasing in v. Thus,

$$
\begin{aligned}
(3C) &= \sup_{S\in\Theta,x\in\mathcal{X}} |P_{\cdot\,|k_x}\psi_{S,x} - P_{\cdot\,|x}\psi_{S,x}| \\
&= \sup_{S\in\Theta,x\in\mathcal{X}} \underbrace{\left| \int_{u=t}^{\infty} R_1(t,u,x)\{1-S_0(u|k_x)\}dG_U(u|k_x) - \int_{u=t}^{\infty} R_1(t,u,x)\{1-S_0(u|x)\}dG_U(u|x) \right|}_{=(3C1)} \\
&\quad + \sup_{S\in\Theta,x\in\mathcal{X}} \underbrace{\left| \int_{u=0}^{t}\int_{v=t}^{\infty} S_0(u|k_x)dG(u,v|k_x) - \int_{u=0}^{t}\int_{v=t}^{\infty} S_0(u|x)dG(u,v|x) \right|}_{=(3C2)} \\
&\quad + \sup_{S\in\Theta,x\in\mathcal{X}} \underbrace{\left| \int_{0}^{t} R_2(t,v,x)S_0(v|k_x)dG_V(v|k_x) - \int_{0}^{t} R_2(t,v,x)S_0(v|x)dG_V(v|x) \right|}_{=(3C3)} \\
&\quad + \sup_{S\in\Theta,x\in\mathcal{X}} \underbrace{\left| \int_{u=0}^{t}\int_{v=t}^{\infty} R_3(t,u,v,x)(S_0(u|k_x)-S_0(v|k_x))dG(u,v|k_x) - \int_{u=0}^{t}\int_{v=t}^{\infty} R_3(t,u,v,x)(S_0(u|x)-S_0(v|x))dG(u,v|x) \right|}_{=(3C4)}.
\end{aligned}
$$

We show $\sup_{S\in\Theta,x\in\mathcal{X}}(3C1) = o_P(1)$. Then $(3C2)$–$(3C4)$ can be shown to be $o_P(1)$ using similar arguments.

$$
\begin{aligned}
(3C1) \leq & \Bigg| \int_{u=t}^{\infty} R_1(t,u,x)\{1 - S_0(u|k_x)\} dG_U(u|k_x) \\
& \qquad\qquad - \int_{u=t}^{\infty} R_1(t,u,x)\{1 - S_0(u|x)\} dG_U(u|k_x) \Bigg| \\
& + \Bigg| \int_{u=t}^{\infty} R_1(t,u,x)\{1 - S_0(u|x)\} dG_U(u|k_x) \\
& \qquad\qquad - \int_{u=t}^{\infty} R_1(t,u,x)\{1 - S_0(u|x)\} dG_U(u|x) \Bigg| \\
= & \Bigg| \int_{u=t}^{\infty} \underbrace{R_1(t,u,x)}_{\in[0,1]} \{S_0(u|x) - S_0(u|k_x)\} dG_U(u|k_x) \Bigg| \\
& + \Bigg| \int_{u=t}^{\infty} \underbrace{R_1(t,u,x)\{1 - S_0(u|x)\}}_{\in[0,1]} d\{G_U(u|k_x) - G_U(u|x)\} \Bigg| \\
\leq & \int_{u=t}^{\infty} \sup_{u' \in [0,\infty)} |S_0(u'|x) - S_0(u'|k_x)| dG_U(u|k_x) \\
& + \Bigg| \Big[R_1(t,u,x)\{1 - S_0(u|x)\}\{G_U(u|k_x) - G_U(u|x)\} \Big]_{u=t}^{\infty} \Bigg| \\
& + \Bigg| \int_{u=t}^{\infty} \{G_U(u|k_x) - G_U(u|x)\} d\Big[R_1(t,u,x)\{1 - S_0(u|x)\} \Big] \Bigg| \\
\leq & \sup_{u \in [0,\infty)} \Big| S_0(u|x) - S_0(u|k_x) \Big| + \Big| G_U(t|k_x) - G_U(t|x) \Big| \\
& + \int_{u=t}^{\infty} \sup_{u' \in [0,\infty)} \Big| G_U(u'|k_x) - G_U(u'|x) \Big| d\Big[\underbrace{R_1(t,u,x)\{1 - S_0(u|x)\}}_{\text{increasing in } u\text{, bounded by 0 and 1.}} \Big] \\
\leq & \sup_{u \in [0,\infty)} \Big| S_0(u|x) - S_0(u|k_x) \Big| + \Big| G_U(t|k_x) - G_U(t|x) \Big| + \sup_{u \in [0,\infty)} \Big| G_U(u|k_x) - G_U(u|x) \Big| \\
& \leq \sup_{x \in \mathcal{X}} \sup_{x' \in k_x} L_S \|x - x'\|_1 + 2 \sup_{x \in \mathcal{X}} \sup_{x' \in k_x} L_G \|x - x'\|_1 \\
& = o_P(1).
\end{aligned}
$$

The last inequality comes from Assumption 2 (Lipschitz continuity) and the subsequent equations result from Assumptions 5 (shrinking terminal node) and 4 (random and regular splits). The derivation for a unit hypercube with a bounded density is given in the proof of Theorem 3 of Wager and Walther (2015).

A1.4 Computational cost

Computation of ICRF is affected by the choice of the splitting and splitting rules, the sample size, and the bandwidth in kernel-smoothing. We discuss the effects of choice of the rules and the sample size from the simulations done in Section 5. Larger bandwidths require more computation having a linear relationship with the number of operations.

First, we discuss the computational cost of ICRF with respect to the splitting and prediction rules. We use the Scenario 1 current status data ($M = 1$) with sample size of $n = 300$, 10 forest iterations, and 300 simulation replicates. First, different splitting rules do not make noticeable differences than having different prediction rules as can be seen in Figure 4.13. This is because the computationally expensive NPMLEs should be obtained for all n/n_{min} terminal nodes. The computation time for NPMLE is almost a quadratic function of sample size due to its $O(nk)$ EM-iterations and $O(n \log_2 n)$ steps for preprocessing (sorting and indexing), where n is the sample size which is $n_{\min}$ in our application and k is the number of Turnbull intervals (Anderson-Bergman, 2017a). Assuming $k \simeq cn_{\min}$ for some constant $c > 0$ and $n_{\min} = n^{\beta}$, the total computational burden of quasi-honest ICRF is $O(n_{\text{fold}} n^{1+\beta})$. In contrast, the exploitative ICRF implements the NPMLE computation only once at the initial step, saving the constant n_{fold} and the set-up cost of $n^{1-\beta}$ many NPMLE calculations.

Compared to existing splitting rules, SWRS and SLR, the new splitting rules cost slightly more computationally. However, as mentioned above, the prediction rule is the predominant determinant of computation over the splitting rule. Given a pair of two samples of sizes $n_l, l = 1, 2$ and k time points of evaluation for numerical integration, the computational burden of GWRS and GLR is $O(kn_1n_2)$ and $O(k(n_1 + n_2))$, respectively. Although GWRS and GLR do not make a large difference in computation time in Figure 4.13, for large samples, GWRS may be computationally more burdensome.

Figure 4.14 illustrates the trend of computation time in terms of sample size. The trend is mildly superlinear supporting the total burden of $O(n^{\gamma})$ for some $1 < \gamma \leq 2$.

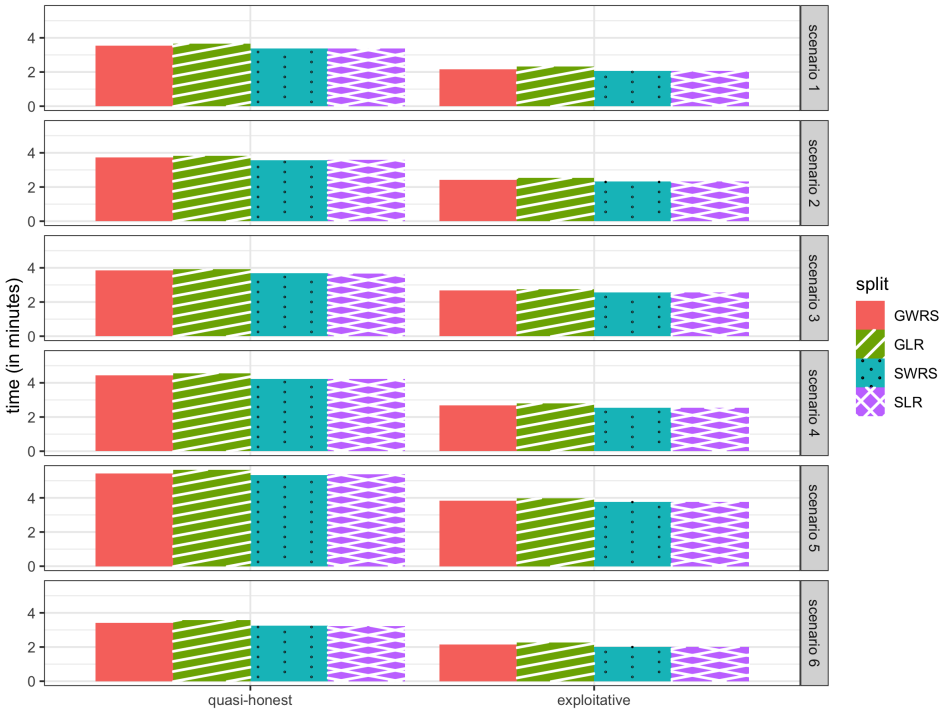

Figure 4.13: Running time for fitting ICRFs with different splitting and prediction rules.

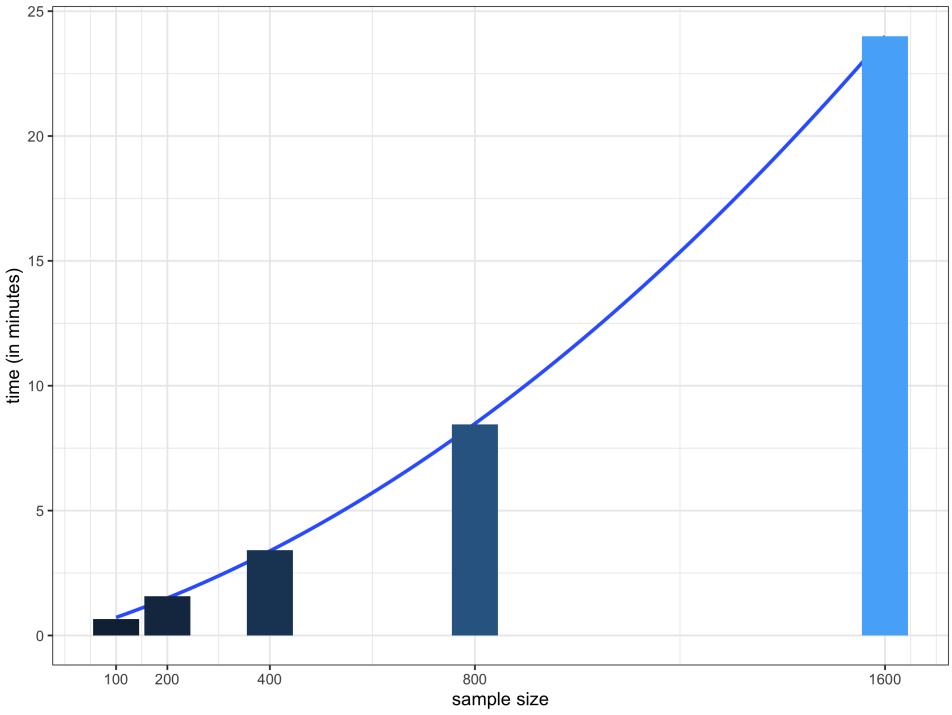

Figure 4.14: Running time for fitting ICRFs with different sample sizes. The blue curve is a quadratic line that minimizes the mean squared error.

A1.5 The National Longitudinal Mortality Study data analysis

We use the NLMS dataset with six years of follow-up recorded around April 2002. The dataset includes 745,162 subjects with their time to mortality, demographic information such as age, sex, and race, socioeconomic information such as income and housing tenure, and other covariates. The censoring rate of this dataset is very high (97% survived six years), as this is a general population, and only administrative censoring is observed. We narrow our focus to the aged population (age ≥ 80 in years) with complete covariate records ($n = 3,630$). The proportion of missing data is 20.7% for the whole data and 65.9% for the aged group data. Thus, it should be noted that this data analysis is for performance comparison among the methods and that the results obtained from this regression analysis are limited to the tracked population. The administrative censoring rate is 69.6%, and the distribution of the data is summarized in Table 4.9.

Since the observed failure time is sparse after the follow-up time of 1500, we set $\tau = 1500$. We induce current status censoring where the monitoring time is dependent on age and number of households. The monitoring time is randomly drawn from the model

$$C \sim N(1000 + 100(10 - X_{\text{age}}/10 + X_{\#\text{ households}}), 300^2) \vee 0 \wedge \tau.$$

The analysis framework is largely the same for the avalanche data analysis, except that with the large sample size, the terminal node size is allowed to be larger ($n_{min} = 20$ for random forests and $n_{min} = 40$ for trees).

variable	min	1Q	median	mean	3Q	max
failure time	2	315	633	694.9	1003	1739
censoring	censored 69.6%					
age (years)	81	83	85	84.8	86	90
number of households	1	1	2	1.9	2	13
adjusted weight	14	251	432	471	633	1982
sex	male 34.8%, female 65.2%					
race	white 85.6%, black 10.1%, others and unknown 4.4%					
Hispanic	Mexican 3.4%, other Hispanics 3.6%, non-Hispanics 93.0%					
relationship	A. 25.6%; B. 46.7%; C. 13.6%; D. 13.2%; E. 0.9					
adjusted income	1. 6.6%; 2. 11.4%; 3. 14.1%; 4. 11.6%; 5. 8.5%; 6. 12.6%; 7. 8.6%; 8. 5.7%; 9. 4.1%; 10. 3.5%; 11. 3.8%; 12. 2%; 13. 3.0%; 14. 4.6					
social security number	present 56.9%					
housing tenure	owner 76.4; rent 21.1; non-cash rent 2.5					
health in general	A. 5.1; B. 17.1; C. 32.5; D. 27.8; E. 17.4"					
health insurance type	A. 0.6; B. 75.2; C. 7.1; D. 7.5; E. 9.5					
urban	urban 73.1; rural 26.9					
citizenship	native citizen born in mainland US 88.3%; others 11.7%					

Table 4.9: The NLMS data. Failure time, non-censored time in days; Relationship, relationship to the reference person (A: reference person with other relatives in household, B. reference person with no other relatives in household, C. spouse of reference person, D. other relative of reference person, E. non-relative of reference person); Adjusted income, 1. $<$ \$5,000, 2. $<$ \$7,499, 3. $<$ \$10,000, 4. $<$, 4. $<$ \$12,500, 5. $<$ \$15,000, 6. $<$ \$20,000, 7. $<$ \$25,000, 8. $<$ \$30,000, 9. $<$ \$35,000, 10. $<$ \$40,000, 11. $<$ \$50,000, 12. $<$ \$60,000, 13. $<$ \$75,000, 14. $\geq$ \$75,000; health in general, A. Excellent, B. Very good, C. Good, D. Fair, E. Poor; health insurance type, A. Medicare, B. Medicaid, C. governmental healthcare (ChampUS, ChampVA, etc), D. employer-based, E. private non-employer-based.

APPENDIX 2: TECHNICAL DETAILS FOR CHAPTER 2

This chapter contains technical details including as assumptions, proofs, definitions, and other materials supplemental to the main text of Chapter 2.

A2.1 Proof of Proposition 2.1

Proof. We first prove the case when $\phi = \phi_\mu$. We use a recursion strategy starting from $q = Q-1$. Assume

$$\pi_*^{(q+1)}(\boldsymbol{H}^{(q+1)}) = \arg \max_{a^{(q+1)} \in \mathcal{A}^{(q+1)}} \phi^{(q+1)}\{S^{(q+1)}(\cdot \mid \boldsymbol{H}^{(q+1)}, \pi_*^{(q+2)}, ..., \pi_*^{(Q)})\}$$

is true for some $q = Q-1, Q-2, ..., 1$.

Then for any policy $\pi_*^{(q+1)}(\boldsymbol{H}^{(q+1)})$ for the qth stage,

$$\begin{aligned}
&\int_0^\tau S^{(q+1)}\{t - X^{(q+2)} \mid \boldsymbol{H}^{(q+1)}, A^{(q+1)} = \pi_*^{(q+1)}(\boldsymbol{H}^{(q+1)}), \pi_*^{(q+2)}, ..., \pi_*^{(Q)}\}dt \\
\geq &\int_0^\tau S^{(q+1)}\{t - X^{(q+2)} \mid \boldsymbol{H}^{(q+1)}, A^{(q+1)} = \pi^{(q+1)}(\boldsymbol{H}^{(q+1)}), \pi_*^{(q+2)}, ..., \pi_*^{(Q)}\}dt.
\end{aligned}$$

This implies that for any $a \in \mathcal{A}^{(q)}$,

$$\begin{aligned}
&\phi_\mu^{(q)}\{S^{(q)}(\cdot; \boldsymbol{H}^{(q)}, a^{(q)}, \pi_*^{(q+1)}, \pi_*^{(q+2)}, ..., \pi_*^{(Q)})\} \\
&= \int_0^\tau \int S^{(q+1)}\{t - X^{(q)} \mid \boldsymbol{H}^{(q+1)}, A^{(q+1)} = \pi_*^{(q+1)}(\boldsymbol{H}^{(q+1)})\} \\
&\qquad dP(X^{(q)}, \delta^{(q)}, \boldsymbol{H}^{(q+1)} \mid \boldsymbol{H}^{(q)}, A^{(q)} = a^{(q)})dt
\end{aligned}$$

$$
\begin{aligned}
&= \int \left[\int_0^\tau S^{(q+1)}\{t - X^{(q)} \mid \boldsymbol{H}^{(q+1)}, A^{(q+1)} = \pi_*^{(q+1)}(H^{(q+1)})\} dt \right] \\
&\qquad\qquad dP(X^{(q)}, \delta^{(q)}, H^{(q+1)} \mid \boldsymbol{H}^{(q)}, A^{(q)} = a^{(q)}) \\
&\geq \int \left[\int_0^\tau S^{(q+1)}\{t - X^{(q)} \mid \boldsymbol{H}^{(q+1)}, A^{(q+1)} = \pi^{(q+1)}(\boldsymbol{H}^{(q+1)})\} dt \right] \\
&\qquad\qquad dP(X^{(q)}, \delta^{(q)}, H^{(q+1)} \mid \boldsymbol{H}^{(q)}, A^{(q)} = a^{(q)}) \\
&= \phi_\mu^{(q)}\{S^{(q)}(\cdot \mid \boldsymbol{H}^{(q)}, a^{(q)}, \pi^{(q+1)}, \pi_*^{(q+2)}, ..., \pi_*^{(Q)})\}.
\end{aligned}
$$

This can be further extended to

$$
\begin{aligned}
&\phi_\mu\{S^{(q)}(\cdot; H^{(1)}, a^{(1)}, a^{(2)}, ..., a^{(q)}, \pi_*^{(q+1)}, \pi_*^{(q+2)}, ..., \pi_*^{(Q)})\} \\
&\geq \phi_\mu\{S^{(q)}(\cdot \mid H^{(1)}, a^{(1)}, a^{(2)}, ...a^{(q)}, \pi^{(q+1)}, \pi_*^{(q+2)}, ..., \pi_*^{(Q)})\},
\end{aligned}
$$

for any set $(a^{(1)}, a^{(2)}, ..., a^{(q)})^\top \in \otimes_{j=1}^q \mathcal{A}^{(j)}$ of treatments up to stage q. Since the optimal treatment at stage q is chosen independently of the regime of the previous stages, the desired result holds.

The proof for $\phi = \phi_{\sigma,t}$ can be done similarly and is not presented here. ■

A2.2 Proof of Theorem 2.4

The following Z-estimator lemma (Kosorok, 2007) plays a key role in this consistency proof. Let $\Psi : \Theta \mapsto \mathbb{L}$ and $\Psi_n : \Theta \mapsto \mathbb{L}$ be maps between two normed spaces, where Θ is a parameter space, $\mathbb{L}$ is some normed space, $\|\cdot\|_\mathbb{L}$ denotes the uniform norm, Ψ is a fixed map, and Ψ_n is a data-dependent map.

Lemma 4.3 (Consistency of Z-estimators)**.** *Let* $\Psi(S_0) = 0$ *for some* $S_0 \in \Theta$*, and assume* $\|\Psi(S_n)\|_\mathbb{L} \to 0$ *implies* $\|S_n - S_0\|_\mathbb{L} \to 0$ *for any sequence* $\{S_n\} \in \Theta$*. Then, if* $\|\Psi_n(\hat{S}_n)\|_\mathbb{L} \to 0$

in probability for some sequence of estimators $\hat{S}_n \in \Theta$ *and* $\sup_{S\in\Theta} \|\Psi_n(S) - \Psi(S)\|_{\mathbb{L}} \to 0$ *in probability,* $\|\hat{S}_n - S_0\|_{\mathbb{L}} \to 0$ *in probability, as* $n \to \infty$.

For our specific problem, let Θ be the space of all covariate-conditional survival functions. Define a normed space $\mathbb{L} = D_{[-1,1]}\{[0,\tau]\times\mathbb{R}^d\}$, where $D_A B$ is the space of all right-continuous left-limits functions with range A and support B.

Define

$$\psi_{S,t} = \delta S'(t - X \mid H') + (1-\delta)\{\frac{S(t \mid H)}{S(X \mid H)} \wedge 1\} - S(t \mid H), \tag{4.22}$$

$$\tilde{\psi}_{S,t,\tilde{S}'} = \delta \tilde{S}'(t - X \mid H') + (1-\delta)\{\frac{S(t \mid H)}{S(X \mid H)} \wedge 1\} - S(t \mid H) \tag{4.23}$$

$$\Psi(S) = \frac{P\psi_{S,t}\delta_h}{P\delta_h} \equiv P_{\cdot|h}\psi_{S,t} \qquad \text{and} \tag{4.24}$$

$$\Psi_n(S) = \frac{\mathbb{P}_n\tilde{\psi}_{S,t}k_h}{\mathbb{P}_n k_h} \equiv \mathbb{P}_{n,\cdot|k_h}\tilde{\psi}_{S,t,\tilde{S}'}, \tag{4.25}$$

where P is the population average of function values, i.e., $Pf = \int f(h)dP(h)$, $\mathbb{P}_n$ is the sample average of function values, i.e., $\mathbb{P}_n f = \int f(h)d\mathbb{P}_n(h) = \frac{1}{n}\sum_{i=1}^n f(H_i)$, S' is a fixed survival probability that is distinct from S, $\tilde{S}'$ is a data-dependent version of S' such that $\lim_{n\to\infty}\tilde{S}' \to S'$ in probability. In the context of dynamic treatment regimes, S' is the survival probability of the remaining life of the next stage and H' is the historical information available at the beginning of the next stage. Also, $\delta_h = I(\cdot = h)$ is the unnormalized Dirac measure, and k_h the unnormalized forest kernel. To be more specific, $k_h = \frac{1}{n_{\text{tree}}}\sum_{b=1}^{n_{\text{tree}}} 1(h \in L_b(h))$, where n_{tree} is the number of trees in the forest, $L_b(h)$ is the terminal node of the bth tree of the forest that contains the point h. We used the term 'unnormalized' to mean that they are not multiplied by the sample (or the population) size.

Then by Lemma 4.3 and the following four propositions (Propositions 4.2 to 4.5), $\|\hat{S} - S_0\|_{\mathbb{L}} \to 0$ in probability as $n \to \infty$.

Proposition 4.2. $\Psi(S_0) = 0$ *for* $S_0 \in \Theta$.

Proposition 4.3. *Assume Assumptions 2.6 to 2.9. As* $n \to \infty$, $\|\Psi(\hat{S})\|_{\mathbb{L}} \to 0$ *implies* $\|\hat{S} - S_0\|_{\mathbb{L}} \to 0$.

Proposition 4.4. *Assume Assumptions 2.6 to 2.9. Let* $\hat{S}$ *be the kernel-conditional modified Kaplan-Meier estimator, which is the modified Kaplan-Meier in (2.3) applied to the data with weights indicated by the forest kernel,* k_h. *As* $n \to \infty$, $\|\Psi_n(\hat{S})\|_{\mathbb{L}} \to 0$ *in probability.*

Proposition 4.5. *Assume Assumptions 2.6 to 2.9. As* $n \to \infty$, $\sup_{S \in \Theta} \|\Psi_n(S) - \Psi(S)\|_{\mathbb{L}} \to 0$ *in probability.*

We give the proofs of Propositions 4.2–4.5. The Proposition 4.2 follows immediately from (2.2).

Proof of Proposition 4.3. By hypothesis, $\|\Psi(S_n)\|_{\mathbb{L}} = \sup_{t \in [0,\tau], h \in \mathcal{H}} |P_{\cdot\,|h} \psi_{S_n,t}| \to 0$. By the following relation, it can be shown that $P_{\cdot\,|h} \psi_{S_n,t} \to 0$ uniformly over $h \in \mathcal{H}$ and $t \in [0, \tau]$ implies $\sup_{t,h} u_n(t|h) \to 0$ for u_n defined in (4.26) below.

$$
\begin{aligned}
P_{\cdot\,|h} \psi_{S_n,t} &= P_{\cdot\,|h} \Big\{ \delta S'(t - X \mid H') + (1-\delta) 1(X > t) + (1-\delta) 1(X \le t) \frac{S_n(t)}{S_n(X)} - S_n(t) \Big\} \\
&= \int_0^t \Big\{ \delta S'(t - X | H') + (1-\delta) 1(X > t) + (1-\delta) 1(X \le t) \frac{S_n(t|H=h)}{S_n(X|H=h)} \\
&\quad - S_n(t|H=h) \Big\} dP_{\cdot\,|H=h} \\
&= S(t|H=h) - S_n(t|H=h) \\
&\qquad + \int_0^t 1(X \le t) \Big\{ \frac{S_n(t|H=h)}{S_n(X|H=h)} - \frac{S(t|H=h)}{S(X|H=h)} \Big\} dP_{\cdot\,|H=h} \\
&= S(t|H=h) - S_n(t|H=h) - \int_0^t \frac{S(u|H=h)}{S_n(u|H=h)} dG(u|H=h) S_n(t|H=h) \\
&\quad + \int_0^t dG(u|H=h) S(t|H=h) \\
&= S_n(t|H=h) \Big\{ \underbrace{- \int \epsilon_n(u|H=h) dG(u|H=h) + \epsilon_n(t|H=h) G(t|H=h)}_{= u_n(t|h)} \Big\},
\end{aligned}
\tag{4.26}
$$

where the third equality comes from (2.2) and $\epsilon_n(t|h) = S(t|h)/S_n(t|h) - 1$. Since $\epsilon_n(t|h) = u_n(t|h)/G(t-) - \int_0^{t^-} \frac{u_n(s|h)dG(s|h)}{G(s|h)G(s^-|h)}$ and $G(t \mid h) \geq G(\tau \mid h) > 0$, $u_n(t|h) \to 0$ implies $\epsilon_n(t|h) \to 0$. The desired result now follows. ■

Proof of Proposition 4.4. We prove this by induction on $t \in [0, \tau]$. If $\sum_i^n \psi_{\hat{S},s,h}(X_i) = 0$ for $s < t$ implies $\sum_i^n \psi_{\hat{S},t}(X_i) = 0$, the desired result holds, since $\sum_i^n \psi_{\hat{S},t=0}(X_i) = 0$ is trivial. Rewrite $\psi_{S,t,h}(D_i) = S'(t - X_i) + (1 - \delta_i)\{1 - S'(t - X_i)\}\frac{S(t)}{S(U_i)} - S(t)$, where S', the survival probability of the remaining life after surviving X_i, is given. By hypothesis, if we assume $\sum_i^n \psi_{\hat{S},t^-,h}(X_i) = 0$ is true, then we have

$$
\begin{aligned}
&\sum_i^n \psi_{\hat{S},t} \\
=&\sum_i^n S'(t - X_i) + \sum_i^n \left\{ \frac{(1-\delta_i)(1 - S'(t - X_i))}{\hat{S}(X_i)} - 1 \right\} \hat{S}(t) \\
=&\underbrace{\left[\sum_i^n S'(t^- - X_i) + \sum_i^n \left\{ \frac{(1-\delta_i)(1 - S'(t^- - X_i))}{\hat{S}(X_i)} - 1 \right\} \hat{S}(t^-) \right]}_{=\sum_i^n \psi_{\hat{S},t^-} = 0 \text{ by hypothesis.}} \\
&+ \sum_i^n dS'(t - X_i) + \sum_i^n \underbrace{\left\{ \frac{(1-\delta_i)(1 - S'(t^- - X_i))}{\hat{S}(X_i)} - 1 \right\} \hat{S}(t^-)}_{=S'(t^- - X_i)} \frac{\delta_i dS'(t - X_i)}{S'(t^- - X_i)} \\
&- \sum_i^n \underbrace{\frac{(1-\delta_i)dS'(t - X_i)}{\hat{S}(X_i)} \hat{S}(t)}_{=-(1-\delta_i)1(X_i = t) \text{ by below.}} \\
=&\sum_i^n dS'(t - X_i) - \sum_i^n S'(t^- - X_i)\frac{\delta_i dS'(t - X_i)}{S'(t^- - X_i)} - \sum_i^n (1 - \delta_i)1(X_i = t) \\
=&\sum_i^n \delta_i dS'(t - X_i) + \sum_i^n \underbrace{(1-\delta_i)dS'(t - X_i)}_{=(1-\delta_i)1(X_i = t)} - \sum_i^n \delta_i dS'(t - X_i) - \sum_i^n (1 - \delta_i)1(X_i = t) \\
=&0,
\end{aligned}
$$

where the second equation is given by $\hat{S}(t) = \hat{S}(t-)\left\{1 + \frac{\sum_i^n \delta_i dS'(s-X_i)}{\sum_i^n S'(s^- - X_i)}\right\}$ from the definition of the modified Kaplan-Meier, and in the third equation, we used the following fact; when $\delta_i = 0$, $\begin{cases} dS' = -1, \hat{S}(X_i) = \hat{S}(t) & \text{if } X_i = t \\ dS' = 0 & \text{otherwise.} \end{cases}$ ■

The proof of Proposition 4.5 relies on the entropy calculation and the Donsker theorem. The strategy is to show that the tree kernels and consequently the forest kernels are Donsker, and to show that the ψ and $\tilde{\psi}$ functions are also Donsker, which we will show in Proposition 4.6 below. Then by empirical process theory and with Assumptions 2.6, 2.7, 2.8, and 2.9, the uniform consistency is derived. Hence, we provide the related lemma (4.6) regrading the entropy of the tree kernels before we move on to the proof of Proposition 4.5. The proof of Propositions 4.6 and the rest of the proof of Proposition 4.5 are given in Supplementary Material.

Proposition 4.6 (Bounded entropy integral of the tree and forest kernels)**.** *The collections of the unnormalized tree and forest kernel functions are Donsker, where the tree kernels $k^{tree}(\cdot)$ are axis-aligned random hyper-rectangles, and the forest kernels $k^{forest}(\cdot)$ are the mean of arbitrarily many (n_{tree}) tree kernels.*

Proof of Proposition 4.6. The Vapnik-Chervonenkis (VC)-dimension of any axis-aligned rectangles in $\mathbb{R}^d$ can be shown $2d$. By Theorem 9.2 of Kosorok (2007), for any VC-class of sets $\mathcal{C}$, the collection $1(\mathcal{C})$ of all the corresponding indicator functions has a bounded covering number $N(\epsilon, 1(\mathcal{C}), L_r(P)) \leq K_1 V(\mathcal{C})(4e)^{V(\mathcal{C})}\epsilon^{-r(V(\mathcal{C})-1)}$, where $V(\cdot)$ denotes the VC-dimension, and $K_1 < \infty$ is some universal constant. Thus, the collection $1(\mathcal{C}^{TREE})$ of the unnormalized tree kernels has a bounded covering number: $N(\epsilon, 1(\mathcal{C}^{TREE}), L_2(P)) \leq K_2\epsilon^{-(4d-2)}$, for some constant $K_2 < \infty$. This immediately implies that the class of tree kernels have bounded entropy integrals ($\int_0^1 N(\epsilon, 1(\mathcal{C}^{TREE}), L_2(P))d\epsilon \leq 1$) and thus is Donsker.

Next, since a forest kernel is simply a convex hull of the tree kernels, i.e., $\overline{conv}1(\mathcal{C}^{TREE}) = \{\sum_{i=1}^m \alpha_i 1(c_i) : \sum_{i=1}^m \alpha_i \leq 1, \alpha_i \geq 0, c_i \in \mathcal{C}^{TREE}, i = 1, 2, ..., m\}$, by Theorem 9.4 of

Kosorok (2007), $N\{\epsilon, \overline{conv}1(\mathcal{C}^{TREE}), L_2(P)\} \leq K_3\epsilon^{-(2-\frac{1}{d})}$ for some constant $K_3 < \infty$. This implies that the unnormalized forest kernels are also Donsker. ■

Proof of Proposition 4.5. We prove Proposition 4.5 by showing that the class of functions $\{\psi_{S,t} : S \in \Theta, t \in [0,\tau]\}$ and $\{\tilde{\psi}_{S,t,\tilde{S}'} : S, \tilde{S}' \in \Theta, t \in [0,\tau]\}$ are Donsker. First, $\{S'(t - \mathcal{X} \mid \mathcal{H}), t \in [0,\tau]\}$ is a stochastic process that is monotone in t for a fixed function S' where the randomness comes from $\mathcal{X}$ and $\mathcal{H}$. By Lemma 9.10 of Kosorok (2007), this process has VC-dimension 2 and is thus Donsker. Similarly, another monotone stochastic processes, $\{\frac{S(t|\mathcal{H})}{S(\mathcal{X}|\mathcal{H})} \wedge 1, t \in [0,\tau]\}$ and $\{S(t \mid \mathcal{H})\}$, are also both VC and Donsker, where $S \in \Theta$ itself is stochastic. Thus, $\{\psi_{S,t} : S \in \Theta, t \in [0,\tau]\}$ is Donsker, as all of its terms belong to uniformly bounded Donsker classes. The Donskerness of $\{\tilde{\psi}_{S,t,\tilde{S}'} : S, \tilde{S}' \in \Theta, t \in [0,\tau]\}$ can be shown by the same approach.

With Proposition 4.6 and the Donskerness of products of bounded Donsker classes, all of $\{\psi_{S,t}\delta_h : S \in \Theta, t \in [0,\tau], h\}$, $\{\psi_{S,t}k_h : S \in \Theta, t \in [0,\tau], h \in \mathcal{H}\}$, and $\{\tilde{\psi}_{S,t,\tilde{S}'}k_h : S, \tilde{S}' \in \Theta, t \in [0,\tau], h \in \mathcal{H}\}$ are Donsker. Also by Assumption 2.8, $\sup_h \mathbb{P}_n k_h \asymp n^{\beta-1}$. Thus,

$$
\begin{aligned}
&\sup_{S,\tilde{S}' \in \Theta, h \in \mathcal{H}} |\Psi_n(S) - \Psi(S)| \\
&\leq \underbrace{\sup_{S,\tilde{S}' \in \Theta, h \in \mathcal{H}} \left| \frac{\mathbb{P}_n \tilde{\psi}_{S,t,\tilde{S}'} k_h}{\mathbb{P}_n k_h} - \frac{\mathbb{P}_n \psi_{S,t} k_h}{\mathbb{P}_n k_h} \right|}_{(A)} + \underbrace{\sup_{S,\tilde{S}' \in \Theta, h \in \mathcal{H}} \left| \frac{\mathbb{P}_n \psi_{S,t} k_h}{\mathbb{P}_n k_h} - \frac{P\psi_{S,t} k_h}{P k_h} \right|}_{(B)} \\
&\quad + \underbrace{\sup_{S,\tilde{S}' \in \Theta, h \in \mathcal{H}} \left| \frac{P\psi_{S,t} k_h}{P k_h} - \frac{P\psi_{S,t} \delta_h}{P \delta_h} \right|}_{(C)} \\
&= o_P(1),
\end{aligned}
$$

where

$$
\begin{aligned}
(A) = & \sup_{S,\tilde{S}' \in \Theta, h \in \mathcal{H}} \left| \frac{\mathbb{P}_n \delta 1(X \leq t) \{ \tilde{S}'(t - X \mid H) - S'(t - X \mid H) \} k_h}{\mathbb{P}_n k_h} \right| \\
\leq & \underbrace{\sup_{h \in \mathcal{H}} \frac{\mathbb{P}_n \delta 1(X \leq t) k_h}{\mathbb{P}_n k_h}}_{\leq 1} \underbrace{\sup_{\tilde{S}' \in \Theta, h \in \mathcal{H}} |\tilde{S}'(t - X \mid H) - S'(t - X \mid H))|}_{= o_P(1)} \\
= & o_P(1),
\end{aligned}
$$

$$
\begin{aligned}
(B) = & \sup_{S,\tilde{S}' \in \Theta, h \in \mathcal{H}} \underbrace{\left| \frac{\mathbb{P}_n \psi_{S,t,\tilde{S}'} k_h - P \psi_{S,t} k_h}{\mathbb{P}_n k_h} \right|}_{\leq \frac{\mathcal{O}_P(n^{-1/2})}{n^{\beta - 1}} = o_P(1)} + \underbrace{\left| P \psi_{S,t} k_h \left\{ \frac{1}{\mathbb{P}_n k_h} - \frac{1}{P k_h} \right\} \right|}_{= \frac{P \psi_{S,t} k_h}{P k_h} \frac{(\mathbb{P}_n - P) k_h}{\mathbb{P}_n k_h} \leq \frac{\mathcal{O}_P(n^{-1/2})}{k/n}} \\
= & o_P(1),
\end{aligned}
$$

and

$$
\begin{aligned}
(C) = & \sup_{S \in \Theta, h \in \mathcal{H}} |P_{\cdot \mid k_h} \psi_{S,h} - P_{\cdot \mid h} \psi_{S,h}| \\
= & \sup_{S \in \Theta, h \in \mathcal{H}} \left| S(t \mid k_h) G(t \mid K_h) - \int_0^t \frac{S(u \mid k_h)}{S(u \mid h)} dG(u \mid k_h) S(t \mid h) - S(t \mid h) G(t \mid h) \right. \\
& \left. - \int_0^t \frac{S(u \mid h)}{S(u \mid h)} dG(u \mid h) S(t \mid h) \right| \\
\leq & \sup_{S \in \Theta, h \in \mathcal{H}} |S(t \mid k_h) G(t \mid K_h) - S(t \mid h) G(t \mid h)| + \\
& \sup_{S \in \Theta, h \in \mathcal{H}} \left| \int_0^t \{ S(u \mid k_h) dG(u \mid k_h) - S(u \mid h) dG(u \mid h) \} \frac{S(t \mid h)}{S(u \mid h)} \right|
\end{aligned}
$$

$$\leq \sup_{h \in \mathcal{H}} \underbrace{|S(t \mid k_h)G(t \mid K_h) - S(t \mid h)G(t \mid h)|}_{(C1)} +$$
$$\sup_{S \in \Theta, h \in \mathcal{H}} \underbrace{\left| \int_0^t \{S(u \mid k_h) - S(u \mid h)\} dG(u \mid k_h) \frac{S(t \mid h)}{S(u \mid h)} \right|}_{(C2)} +$$
$$\sup_{S \in \Theta, h \in \mathcal{H}} \underbrace{\left| \int_0^t S(u \mid h)\{dG(u \mid k_h) - dG(u \mid h)\} \frac{S(t \mid h)}{S(u \mid h)} \right|}_{(C3)},$$

$$\begin{aligned}
(C1) &\leq S(t \mid k_h) \sup_{h \in \mathcal{H}} \left| G(t \mid K_h) - G(t \mid h) \right| + G(t \mid h) \sup_{h \in \mathcal{H}} \left| S(t \mid k_h) - S(t \mid h) \right| \\
&\leq \sup_{h \in \mathcal{H}} \int_0^t |dG(u \mid K_h) - dG(u \mid h)| + \sup_{h \in \mathcal{H}} |S(t \mid K_h) - S(t \mid h)| \\
&\leq t L_G \sup_{h' \in k_h} \|h - h'\|_1 + L_S \sup_{h \in \mathcal{H}} \sup_{h' \in k_h} \|h - h'\|_1 \\
&= o_P(1), \\
(C2) &\leq \sup_{h \in \mathcal{H}, u \in [0,\tau]} |S(u \mid k_h) - S(u \mid h)| \\
&\leq \sup_{h \in \mathcal{H}} \sup_{h' \in k_h} L_S \|h - h'\|_1 \\
&= o_P(1), \\
(C3) &\leq \sup_{h \in \mathcal{H}} \int^t |dG(u \mid k_h) - dG(u \mid h)| \\
&\leq \sup_{h \in \mathcal{H}} \sup_{h' \in k_h} L_G \|h - h'\|_1 \\
&= o_P(1).
\end{aligned}$$

The last inequalities in (C1), (C2), and (C3) come from Assumption 2.6 and the subsequent equations result from Assumptions 2.8 and 2.9. The derivation for a unit hypercube with a bounded density is given in the proof of Theorem 3 of Wager and Walther (2015).

We further establish why the exception of countably many unbounded density points is allowed. First, countably many zero density points do not affect the proof of Lemma 2 of Meinshausen (2006). Second, we discuss the case of probability mass points. When and only when there is at least one point with a probability mass jointly for all the covariates, the number of observations in the corresponding terminal node cannot be bounded by an $o(n)$ term. However, since the point has zero Lebesgue measure, the number of observations at the point do not contribute to the maximum size of the node. Let $\mathcal{D} \in [0,1]^{dz}$ denote the set of the points that have a fully joint probability mass. Then the proof of Lemma 2 of Meinshausen (2006) is still valid after replacing "the number of observations" with "the number of observations outside of $\mathcal{D}$," since the Glivenko-Cantelli argument used in translating the empirical measure into the Lebesgue measure works in the presence of countably many discontinuity points.

The categorical covariates with finite levels can be understood in a similar manner. It is still different from the preceding setting in a sense that there is no continuous component. Thus, the number of splits made along the jth covariate in the terminal node that contains $\boldsymbol{h}$ cannot become arbitrarily large for a large n, or $P(S(\boldsymbol{h}, j, \theta) > g_n) \not\to 1$ according to the notation of Meinshausen (2006) where θ is a random parameter that governs partitioning and g_n is some sequence growing to infinity for large n. However, $S(\boldsymbol{h}, j, \theta)$ grows up to the number of categories of the j covariate almost surely, meaning that the categories are fully separated. Since the Lebesgue measure of each category is zero, the node size $I(\boldsymbol{h}, j, \theta)$ along the jth coordinate is zero almost surely, not affecting the result, $\max_j |I(\boldsymbol{h}, j, \theta)| = o_P(1)$ for $n \to \infty$.

■

Proof of Corollary 2.1. Corollary 2.1 follows from recursion once we show that the Qth stage random survival forest is consistent and $n^{(Q)} > cn$ for some constant c almost surely. For each $\boldsymbol{h} \in \mathcal{H}^{(Q)}$ and treatment $a \in \mathcal{A}^{(Q)}$, Assumption 2.12 guarantees that the probability of a patient making all planned visits is at least M^{Q-1}. That is, $n^{(Q)} \geq M^{(Q)} n$ almost surely. Uniform consistency of the stage-wise random survival forest again relies on arbitrarily small node size, $\sup_{h \in \mathcal{H}} \sup_{h' \in k_h} \|h - h'\|_1$. Although the support in Assumption 2.11 is different

from that in Assumption 2.7, the bounded marginal density in Assumption 2.7 is sufficient to follow the same lines of proof given in Meinshausen (2006) and Wager and Walther (2015). Now, by having $S'(t \mid \boldsymbol{h}) = \tilde{S}'(t \mid \boldsymbol{h}) = 1(t \leq 0)$ for all $\boldsymbol{h}$, Theorem 2.4 yields uniform consistency of $\hat{S}^{(Q)}(\cdot \mid \cdot)$. Consistency of the rest of the sequence then follows as $n^{(q)} \geq n^{(Q)}$ for $q = Q-1, Q-2, ..., 1$. ■

A2.3 Proof of Theorem 2.5

Proof of Theorem 2.5. The proof is largely based on the lines of proof used by Murphy (2005) and Goldberg and Kosorok (2012). We tailor their proofs based on the Q-learning framework into the context of this paper. We introduce notation for some intermediate stage quantities that are the analogs to the intermediate stage Q- and V- functions. Let $\Xi^{(q)}_{S,\boldsymbol{\pi}}(\boldsymbol{H}^{(q)}, a^{(q)}) = \phi^{(q)}\{S^{(q)}_{\boldsymbol{\pi}}(\cdot \mid \boldsymbol{H}^{(q)}, A^{(q)} = a^{(q)})\}$ and $\Phi^{(q)}_{S,\boldsymbol{\pi}}(\boldsymbol{H}^{(q)}, a^{(q)}) = \phi^{(q)}\{S^{(q)}_{\boldsymbol{\pi}}(\cdot \mid \boldsymbol{H}^{(q)}, A^{(q)} = \pi^{(q)}(\boldsymbol{H}^{(q)})\}$. The dependency of Ξ and Φ on the choice of ϕ is marked with the corresponding subscripts, e.g., $\Phi^{(q)}_{\cdot,\mu}(\cdot) = \phi^{(q)}_{\mu}(\cdot)$, or is suppressed, if it is clear by context.

We modify Lemmas 1 and 2 of Murphy (2005) to obtain Lemmas 4.4 and 4.5 below. Then we conclude the proof of Theorem 2.5 by bounding $E\left[\left\{\Xi^{(q)}_{S_*,\boldsymbol{\pi}_*}(\boldsymbol{H}^{(q)}, \boldsymbol{A}^{(q-1)}, A^{(q)}) - \hat{\Xi}^{(q)}(\boldsymbol{H}^{(q)}, \boldsymbol{A}^{(q-1)}, A^{(q)})\right\}^2\right]$ of Lemma 4.5 using a quantity

$$Err_{\Xi^{(q+1)}}(\Xi^{(q)}) = E\left\{\max_{a^{(q+1)}} \Xi^{(q+1)}(\boldsymbol{H}^{(q+1)}, A^{(q+1)}) - \Xi^{(q)}(\boldsymbol{H}^{(q)}, A^{(q)})\right\}^2,$$

where a similar strategy was used in Murphy (2005) and Goldberg and Kosorok (2012) and we drop the arguments of $\Xi^{(q)}$ and $\Xi^{(q+1)}$ as long as no confusion arises.

$$
\begin{aligned}
&Err_{\hat{\Xi}^{(q+1)}}(\hat{\Xi}^{(q)}) - Err_{\hat{\Xi}^{(q+1)}}(\Xi_*^{(q)}) \\
&= E[\hat{\Xi}^{(q)}]^2 - E[\Xi_*^{(q)}]^2 + 2E\left[(\Xi_*^{(q)} - \hat{\Xi}^{(q)}) \max_{a^{(q+1)}} \Xi_*^{(q+1)}\right] \\
&\qquad\qquad - 2E\left[(\max_{a^{(q+1)}} \Xi_*^{(q+1)} - \max_{a^{(q+1)}} \hat{\Xi}^{(q+1)})(\Xi_*^{(q)} - \hat{\Xi}^{(q)})\right] \\
&= E[\hat{\Xi}^{(q)}]^2 - E[\Xi_*^{(q)}]^2 + 2E\left[(\Xi_*^{(q)} - \hat{\Xi}^{(q)})\Xi_*^{(q)}\right] \\
&\qquad\qquad - 2E\left[(\max_{a^{(q+1)}} \Xi_*^{(q+1)} - \max_{a^{(q+1)}} \hat{\Xi}^{(q+1)})(\Xi_*^{(q)} - \hat{\Xi}^{(q)})\right] \\
&= E[\hat{\Xi}^{(q)} - \Xi_*^{(q)}]^2 - 2E\left[(\max_{a^{(q+1)}} \Xi_*^{(q+1)} - \max_{a^{(q+1)}} \hat{\Xi}^{(q+1)})(\Xi_*^{(q)} - \hat{\Xi}^{(q)})\right] \\
&\geq E[\hat{\Xi}^{(q)} - \Xi_*^{(q)}]^2 - 2E^{1/2}\left[\max_{a^{(q+1)}} \Xi_*^{(q+1)} - \max_{a^{(q+1)}} \hat{\Xi}^{(q+1)}\right]^2 E^{1/2}\left[\Xi_*^{(q)} - \hat{\Xi}^{(q)}\right]^2 \\
&\geq E[\hat{\Xi}^{(q)} - \Xi_*^{(q)}]^2 - 2\sqrt{LE\left[\Xi_*^{(q+1)} - \hat{\Xi}^{(q+1)}\right]^2} E^{1/2}\left[\Xi_*^{(q)} - \hat{\Xi}^{(q)}\right]^2 \\
&\geq E[\hat{\Xi}^{(q)} - \Xi_*^{(q)}]^2 - \frac{1}{2}\{4LE\left[\Xi_*^{(q+1)} - \hat{\Xi}^{(q+1)}\right]^2 + E\left[\Xi_*^{(q)} - \hat{\Xi}^{(q)}\right]^2\} \\
&= E[\hat{\Xi}^{(q)} - \Xi_*^{(q)}]^2 - 2LE[\hat{\Xi}^{(q+1)} - \Xi_*^{(q+1)}]^2,
\end{aligned}
$$

where in the second equality,

$$
E\left[\max_{a^{(q+1)}} \Xi_*^{(q+1)} = \phi^{(q+1)}\{S_*^{(q+1)}(\cdot \mid \boldsymbol{H}^{(q+1)}, A^{(q+1)})\} \mid \boldsymbol{H}^{(q)}, A^{(q)}\right] = \Xi_*^{(q)}(\boldsymbol{H}^{(q)}, A^{(q)})
$$

was used; in the first and the last inequalities, the Cauchy-Schwarz inequality and that fact that $ab \leq (a^2 + b^2)/2$ were used; and the second inequality holds true, since

$$
\begin{aligned}
&\{\max_{a^{(q+1)}} \Xi_*^{(q+1)}(a^{(q+1)}) - \max_{a^{(q+1)}} \hat{\Xi}^{(q+1)}(a^{(q+1)})\}^2 \\
&\leq \max_{a^{(q+1)}} \{\Xi_*^{(q+1)}(a^{(q+1)}) - \hat{\Xi}^{(q+1)}(a^{(q+1)})\}^2 \\
&\leq LE\{\Xi_*^{(q+1)}(a^{(q+1)}) - \hat{\Xi}^{(q+1)}(a^{(q+1)})\}^2.
\end{aligned}
$$

Then we have

$$E[\hat{\Xi}^{(q)} - \Xi_*^{(q)}]^2 \leq 2\sum_{q'=q}^{Q}(4L)^{q'-q}\big\{Err_{\hat{\Xi}^{(q+1)}}(\hat{\Xi}^{(q)}) - Err_{\hat{\Xi}^{(q+1)}}(\Xi_*^{(q)})\big\},$$

$q = 1, 2, ..., Q$. Thus, by Lemma 4.5, the error bound of the value function is bounded by

$$\begin{aligned}
&\Phi(\boldsymbol{\pi}_*) - \Phi(\hat{\boldsymbol{\pi}}) \\
\leq &\sum_{q=1}^{Q} 2L^{q/2}\sqrt{2\sum_{q'=q}^{Q}(4L)^{q'-q}\big\{Err_{\hat{\Xi}^{(q+1)}}(\hat{\Xi}^{(q)}) - Err_{\hat{\Xi}^{(q+1)}}(\Xi_*^{(q)})\big\}} \\
\leq &\sum_{q=1}^{Q}\sqrt{2^{3-2q}\sum_{q'=q}^{Q}L^{q'}\big\{Err_{\hat{\Xi}^{(q+1)}}(\hat{\Xi}^{(q)}) - Err_{\hat{\Xi}^{(q+1)}}(\Xi_*^{(q)})\big\}}
\end{aligned} \tag{4.27}$$

Now we obtain the bound of difference in Err values:

$$\begin{aligned}
&Err_{\hat{\Xi}^{(q+1)}}(\hat{\Xi}^{(q)}) - Err_{\hat{\Xi}^{(q+1)}}(\Xi_*^{(q)}) \\
&= E\left[\big\{(\max_{a^{(q+1)}}\hat{\Xi}^{(q+1)} - \hat{\Xi}^{(q)}) + (\max_{a^{(q+1)}}\hat{\Xi}^{(q+1)} - \Xi_*^{(q)})\big\}\big\{\Xi_*^{(q)} - \hat{\Xi}^{(q)}\big\}\right] \\
&\leq 2c(\phi)E\big[\Xi_*^{(q)} - \hat{\Xi}^{(q)}\big] \\
&\leq 2c(\phi)\sup_{\boldsymbol{h}^{(q)},a^{(q)}}\big\{\Xi_*^{(q)}(\boldsymbol{h}^{(q)},a^{(q)}) - \hat{\Xi}^{(q)}(\boldsymbol{h}^{(q)},a^{(q)})\big\} \\
&\leq 2c(\phi)\sup_{\boldsymbol{h}^{(q)},a^{(q)}}\left[\phi^{(q)}\left\{\left|S_*^{(q)}(\cdot \mid \boldsymbol{h}^{(q)},a^{(q)}) - \hat{S}^{(q)}(\cdot \mid \boldsymbol{h}^{(q)},a^{(q)})\right|\right\}\right] \\
&\leq 2c(\phi)\sup_{\boldsymbol{h}^{(q)},a^{(q)}}\left[\phi^{(q)}\left\{\sup_{t\in[0,\tau]}\left|S_*^{(q)}(\cdot \mid \boldsymbol{h}^{(q)},a^{(q)}) - \hat{S}^{(q)}(\cdot \mid \boldsymbol{h}^{(q)},a^{(q)})\right|\right\}\right] \\
&\leq 2c(\phi)\phi^{(q)}\left\{\sup_{\boldsymbol{h}^{(q)},a^{(q)},t\in[0,\tau]}\left|S_*^{(q)}(\cdot \mid \boldsymbol{h}^{(q)},a^{(q)}) - \hat{S}^{(q)}(\cdot \mid \boldsymbol{h}^{(q)},a^{(q)})\right|\right\},
\end{aligned}$$

where $c(\phi_\mu) = \tau$ and $c(\phi_{\sigma,t}) = 1$, and the first inequality comes from the triangular inequality and the fact that any Ξ function is bounded by $[0, \tau]$ for $\phi = \phi_\mu$ and $[0, 1]$ for $\phi = \phi_{\sigma,t}$, and, in the fourth inequality, Fatou's lemma was used.

Given that all assumptions of Corollary 2.1 hold, and that the space of history $\mathcal{H}$ incorporates the historical actions, the theorem gives us

$$\sup_{t\in[0,\tau],h^{(q)}\in\mathcal{H}^{(q)},a^{(q)}\in\mathcal{A}^{(q)},q\in\{1,2,\ldots,Q\}} |\hat{S}^{(q)}(t \mid h^{(q)}, a^{(q)}) - S_*^{(q)}(t \mid h^{(q)}, a^{(q)})| \to 0,$$

in probability as $n \to \infty$.

Since ϕ is either integration over the interval $[0, \tau]$ for $\phi = \phi_\mu$ or evaluation at t for $\phi = \phi_{\sigma,t}$, the error is bounded by $c(\phi)$ times the supremum error of the survival curves. Thus,

$$Err_{\hat{\Xi}^{(q+1)}}(\hat{\Xi}^{(q)}) - Err_{\hat{\Xi}^{(q+1)}}(\Xi_*^{(q)}) \to 0$$

in probability as $n \to \infty$ and since in (4.27) $\Phi(\boldsymbol{\pi}_*) - \Phi(\hat{\boldsymbol{\pi}})$ is bounded by finite summations and bounded transformations, the desired result holds. ■

Lemma 4.4. *For treatment regimes $\tilde{\boldsymbol{\pi}}$ and $\boldsymbol{\pi}$,*

$$\Phi(\tilde{\boldsymbol{\pi}}) - \Phi(\boldsymbol{\pi}) = -E_{\boldsymbol{\pi}}\left\{ \sum_{q=1}^{Q} \Delta^{(q)}_{S_*,\tilde{\boldsymbol{\pi}}}(\boldsymbol{H}^{(q)}, A^{(q)}) \right\},$$

where $\Delta^{(q)}_{S,\boldsymbol{\pi}}(\boldsymbol{h}^{(q)}, a^{(q)}) = \Xi_{S,\boldsymbol{\pi}}(\boldsymbol{h}^{(q)}, a^{(q)}) - \Phi_{S,\boldsymbol{\pi}}(\boldsymbol{h}^{(q)})$ is the advantage of treatment $a^{(q)}$ at stage q, and S_* is the true survival probability. This statement is true for both criteria $\phi = \phi_\mu, \phi_{\sigma,t}$.

Lemma 4.5. *For all functions,* $\hat{\Xi}^{(q)}$ *satisfying* $\hat{\pi}^{(q)}(\boldsymbol{h}^{(q)}) \in \arg\max_{a^{(q)}} \hat{\Xi}^{(q)}(\boldsymbol{h}^{(q)}, a^{(q)}), q = 1, 2, ..., Q$*, we have,*

$$\begin{aligned}
&\Phi(\boldsymbol{\pi}_*) - \Phi(\hat{\boldsymbol{\pi}}) \\
&\leq \sum_{q=1}^{Q} 2L^{q/2} \sqrt{E\left[\left\{\Xi^{(q)}_{S_*,\boldsymbol{\pi}_*}(\boldsymbol{H}^{(q)}, \boldsymbol{A}^{(q-1)}, A^{(q)}) - \hat{\Xi}^{(q)}(\boldsymbol{H}^{(q)}, \boldsymbol{A}^{(q-1)}, A^{(q)})\right\}^2\right]},
\end{aligned}$$

where $\boldsymbol{\pi}_* = (\pi^{(1)}, ..., \pi^{(Q)})$ *is the optimal decision rule for the survival probability* S_**. This statement is true for both criteria* $\phi = \phi_\mu, \phi_{\sigma,t}$ *.*

APPENDIX 3: TECHNICAL DETAILS FOR CHAPTER 3

This chapter contains technical details including as assumptions, proofs, definitions, and other materials supplemental to the main text of Chapter 3.

A3.1 EM algorithm

The components of expected log-likelihood

The expectation of the log-likelihood comprises the folowing terms with subscripts i being suppressed:

$$
\begin{aligned}
E(R_0|Y_1,Y_2;\boldsymbol{\theta}) &= \alpha_0\beta_1\frac{f_{BZINB}(Y_1,Y_2;\alpha_0+1,\alpha_1,\alpha_2,\boldsymbol{\beta},\boldsymbol{\pi})}{f_{BZINB}(Y_1,Y_2;\alpha_0,\alpha_1,\alpha_2,\boldsymbol{\beta},\boldsymbol{\pi})},\\
E(R_1|Y_1,Y_2;\boldsymbol{\theta}) &= \alpha_1\beta_1\frac{f_{BZINB}(Y_1,Y_2;\alpha_0,\alpha_1+1,\alpha_2,\boldsymbol{\beta},\boldsymbol{\pi})}{f_{BZINB}(Y_1,Y_2;\alpha_0,\alpha_1,\alpha_2,\boldsymbol{\beta},\boldsymbol{\pi})},\\
E(R_2|Y_1,Y_2;\boldsymbol{\theta}) &= \alpha_2\beta_1\frac{f_{BZINB}(Y_1,Y_2;\alpha_0,\alpha_1,\alpha_2+1,\boldsymbol{\beta},\boldsymbol{\pi})}{f_{BZINB}(Y_1,Y_2;\alpha_0,\alpha_1,\alpha_2,\boldsymbol{\beta},\boldsymbol{\pi})},\\
E(\log(R_0)|Y_1,Y_2;\boldsymbol{\theta}) &= \frac{1}{f_{BZINB}(Y_1,Y_2)}\times\\
&\Bigg[\sum_{k,m}H_0(k,m;Y_1,Y_2,\boldsymbol{\alpha},\boldsymbol{\beta})\{\psi(Y_1+Y_2-k-m+\alpha_0)+\\
&\qquad\log(\frac{\beta_1}{1+\beta_1+\beta_2})\}\pi_1+\\
&\sum_{k=0}^{Y_1}H_1(k;Y_1,\boldsymbol{\alpha},\boldsymbol{\beta})\{\psi(Y_1-k+\alpha_0)+\log(\frac{\beta_1}{1+\beta_1})\}\pi_2\zeta(Y_2)+\\
&\sum_{m=0}^{Y_2}H_2(m;Y_2,\boldsymbol{\alpha},\boldsymbol{\beta})\{\psi(Y_2-m+\alpha_0)+\log(\frac{\beta_1}{1+\beta_2})\}\pi_3\zeta(Y_1)+\\
&\{\psi(\alpha_0)+\log(\beta_1)\}\pi_4\zeta(Y_1+Y_2)\Bigg],
\end{aligned}
$$

$$
\begin{aligned}
E(\log(R_1)|Y_1,Y_2;\boldsymbol{\theta}) &= \frac{1}{f_{BZINB}(Y_1,Y_2)}\times \\
&\Bigg[\sum_{k,m} H_0(k,m;Y_1,Y_2,\boldsymbol{\alpha},\boldsymbol{\beta})\{\psi(k+\alpha_1)+\log(\frac{\beta_1}{1+\beta_1})\}\pi_1+ \\
&\sum_{k=0}^{Y_1} H_1(k;Y_1,\boldsymbol{\alpha},\boldsymbol{\beta})\{\psi(k+\alpha_1)+\log(\frac{\beta_1}{1+\beta_1})\}\pi_2\zeta(Y_2)+ \\
&\sum_{m=0}^{Y_2} H_2(m;Y_2,\boldsymbol{\alpha},\boldsymbol{\beta})\{\psi(\alpha_1)+\log(\beta_1)\}\pi_3\zeta(Y_1)+ \\
&\{\psi(\alpha_1)+\log(\beta_1)\}\pi_4\zeta(Y_1+Y_2)\Bigg], \\
E(\log(R_2)|Y_1,Y_2;\boldsymbol{\theta}) &= \frac{1}{f_{BZINB}(Y_1,Y_2)}\times \\
&\Bigg[\sum_{k,m} H_0(k,m;Y_1,Y_2,\boldsymbol{\alpha},\boldsymbol{\beta})\{\psi(m+\alpha_2)+\log(\frac{\beta_1}{1+\beta_2})\}\pi_1+ \\
&\sum_{k=0}^{Y_1} H_1(k;Y_1,\boldsymbol{\alpha},\boldsymbol{\beta})\{\psi(\alpha_2)+\log(\beta_1)\}\pi_2\zeta(Y_2)+ \\
&\sum_{m=0}^{Y_2} H_2(m;Y_2,\boldsymbol{\alpha},\boldsymbol{\beta})\{\psi(m+\alpha_2)+\log(\frac{\beta_1}{1+\beta_2})\}\pi_3\zeta(Y_1)+ \\
&\{\psi(\alpha_2)+\log(\beta_1)\}\pi_4\zeta(Y_1+Y_2)\Bigg], \\
E(E_1|Y_1,Y_2;\boldsymbol{\theta}) &= \frac{f_{BNB}(Y_1,Y_2;\boldsymbol{\alpha},\boldsymbol{\beta})\pi_1}{f_{BZINB}(Y_1,Y_2;\boldsymbol{\theta})}, \\
E(E_2|Y_1,Y_2;\boldsymbol{\theta}) &= \frac{f_{NB}(Y_1;\alpha_0+\alpha_1,\frac{\beta_1}{\beta_1+1})\pi_2\zeta(Y_2)}{f_{BZINB}(Y_1,Y_2;\boldsymbol{\theta})}, \\
E(E_3|Y_1,Y_2;\boldsymbol{\theta}) &= \frac{f_{NB}(Y_2;\alpha_0+\alpha_2,\frac{\beta_2}{\beta_2+1})\pi_3\zeta(Y_1)}{f_{BZINB}(Y_1,Y_2;\boldsymbol{\theta})}, \\
E(E_4|Y_1,Y_2;\boldsymbol{\theta}) &= \frac{\pi_4\zeta(Y_1+Y_2)}{f_{BZINB}(Y_1,Y_2;\boldsymbol{\theta})}, \qquad \text{and} \\
E(X_2|Y_1,Y_2;\boldsymbol{\theta}) &= Y_2+\frac{(\alpha_0+\alpha_2)\beta_2}{f_{BZINB}(Y_1,Y_2;\boldsymbol{\theta})}\times \\
&\Bigg[f_{NB}(Y_1;\alpha_0+\alpha_1,\frac{\beta_1}{\beta_1+1})\pi_2\zeta(Y_2)+\pi_4\zeta(Y_1+Y_2)\Bigg],
\end{aligned}
$$

where

$$H_0(k,m;Y_1,Y_2,\boldsymbol{\alpha},\boldsymbol{\beta}) := \binom{\alpha_0+Y_1+Y_2-k-m-1}{\alpha_0+Y_2-m-1}\binom{\alpha_0+Y_2-m-1}{\alpha_0-1}\times \binom{\alpha_1+k-1}{\alpha_1-1}\binom{\alpha_2+m-1}{\alpha_2-1}\frac{\beta_1^{Y_1}\beta_2^{Y_2}(\beta_1+\beta_2+1)^{k+m-Y_1-Y_2-\alpha_0}}{(\beta_1+1)^{k+\alpha_1}(\beta_2+1)^{m+\alpha_2}},$$

$$H_1(k;Y_1,\boldsymbol{\alpha},\boldsymbol{\beta}) := \binom{\alpha_0+Y_1-k-1}{\alpha_0-1}\binom{\alpha_1+k-1}{\alpha_1-1}\frac{\beta_1^{Y_1}}{(\beta_1+1)^{Y_1+\alpha_0+\alpha_1}},$$
$$H_2(m;Y_2,\boldsymbol{\alpha},\boldsymbol{\beta}) := \binom{\alpha_0+Y_2-m-1}{\alpha_0-1}\binom{\alpha_2+m-1}{\alpha_2-1}\frac{\beta_2^{Y_2}}{(\beta_2+1)^{Y_2+\alpha_0+\alpha_2}},$$

$\boldsymbol{\theta} \equiv (\boldsymbol{\alpha}^\top, \boldsymbol{\beta}^\top, \boldsymbol{\pi}^\top)^\top$, $\psi(\cdot)$ is the digamma function, and the parameters in density functions are written either in scalar, vector, or combination of both as needed without confusion.

A3.2 Standard error formula

Recall the density of BZINB is given as,

$$\begin{aligned} &f_{BZINB}(y_1,y_2;\boldsymbol{\theta}) \\ &= \pi_1 f_{BNB}(y_1,y_2;\boldsymbol{\theta}) + \pi_2 f_{NB}(y_1;\alpha_0+\alpha_1,\frac{1}{\beta_1+1})\zeta(y_2) \\ &\quad + \pi_3 f_{NB}(y_2;\alpha_0+\alpha_2,\frac{1}{\beta_2+1})\zeta(y_1) + (1-\pi_1-\pi_2-\pi_3)\zeta(y_1+y_2), \end{aligned}$$

where $\boldsymbol{\theta}$ here is redefined using free parameters only as $(\boldsymbol{\alpha}^\top, \boldsymbol{\beta}^\top, \pi_1, \pi_2, \pi_3)^\top$.

Then the observed information is given as,

$$\begin{aligned} I_{obs}(\hat{\boldsymbol{\theta}}) &= -\partial^2_{\boldsymbol{\theta}} l|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \\ &= -\sum_i^n \partial^2_{\boldsymbol{\theta}} \log f_{BZINB}(y_{1,i}, y_{2,i}; \boldsymbol{\theta})|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \\ &= -\sum_i^n \frac{\{\partial^2_{\boldsymbol{\theta}} f_{BZINB}(y_{1,i}, y_{2,i}; \boldsymbol{\theta})\}\{f_{BZINB}(y_{1,i}, y_{2,i}; \boldsymbol{\theta})\}}{\{f_{BZINB}(y_{1,i}, y_{2,i}; \boldsymbol{\theta})\}^2} \\ &\quad + \sum_i^n \frac{\{\partial_{\boldsymbol{\theta}} f_{BZINB}(y_{1,i}, y_{2,i}; \boldsymbol{\theta})\}^{\otimes 2}}{\{f_{BZINB}(y_{1,i}, y_{2,i}; \boldsymbol{\theta})\}^2}|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}, \end{aligned}$$

where $\boldsymbol{a}^{\otimes 2} = \boldsymbol{a}\boldsymbol{a}^\top$ and $\partial^2_{\boldsymbol{\theta}} l = \partial_{\boldsymbol{\theta}} \partial_{\boldsymbol{\theta}^\top} l$.

Then the large sample standard error estimate of the MLE $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ is $diag(I_{obs}(\hat{\boldsymbol{\theta}})^{-1})^{\frac{1}{2}}$, and that of the MLE $\hat{\rho}^*$ of ρ^* is $\left[(\nabla g(\hat{\boldsymbol{\theta}}))^\top I_{obs}(\hat{\boldsymbol{\theta}})^{-1} \nabla g(\hat{\boldsymbol{\theta}})\right]^{1/2}$, where $\nabla g(\hat{\boldsymbol{\theta}}) := \partial_{\boldsymbol{\theta}} \rho^*|_{\hat{\boldsymbol{\theta}}}$ is as follows:

$$\begin{aligned} \nabla g(\hat{\boldsymbol{\theta}}) = \hat{\rho}^{TRUE} \Bigg(&\left\{ \frac{1}{\hat{\alpha}_0} - \frac{1}{2(\hat{\alpha}_0 + \hat{\alpha}_1)} - \frac{1}{2(\hat{\alpha}_0 + \hat{\alpha}_2)} \right\}, \frac{-1}{2(\hat{\alpha}_0 + \hat{\alpha}_1)}, \frac{-1}{2(\hat{\alpha}_0 + \hat{\alpha}_2)}, \\ &\frac{1}{2\hat{\beta}_1(\hat{\beta}_1 + 1)}, \frac{1}{2\hat{\beta}_2(\hat{\beta}_2 + 1)}, 0, 0, 0 \Bigg)^\top. \end{aligned}$$

The observed information, however, can be approximated by the empirical observed information (e.g. Meilijson (1989)) which incurs only the first derivatives of the individual log-likelihood:

$$I_e(\hat{\boldsymbol{\theta}}) = \sum_{i=1}^n s(y_{1,i}, y_{2,i}; \hat{\boldsymbol{\theta}}) s(y_{1,i}, y_{2,i}; \hat{\boldsymbol{\theta}})^\top,$$

where

$$\begin{aligned} s(y_{1,i}, y_{2,i}; \boldsymbol{\theta}) &= \partial_{\boldsymbol{\theta}} \log f_{BZINB}(y_{1,i}, y_{2,i}; \boldsymbol{\theta}) \\ &= \frac{\partial_{\boldsymbol{\theta}} f_{BZINB}(y_{1,i}, y_{2,i}; \boldsymbol{\theta})}{f_{BZINB}(y_{1,i}, y_{2,i}; \boldsymbol{\theta})}, \end{aligned}$$

$$\partial_{\boldsymbol{\theta}} f_{BZINB}(y_1, y_2; \boldsymbol{\theta}) = \begin{pmatrix} \pi_1 D_1 + \pi_2 \zeta(y_2) D_6 + \pi_3 \zeta(y_1) D_8 \\ \pi_1 D_2 + \pi_2 \zeta(y_2) D_6 \\ \pi_1 D_3 + \pi_3 \zeta(y_1) D_8 \\ \pi_1 D_4 + \pi_2 \zeta(y_2) D_7 \\ \pi_1 D_5 + \pi_3 \zeta(y_1) D_9 \\ f_{BNB}(y_1, y_2; \boldsymbol{\alpha}, \boldsymbol{\beta}) - \zeta(y_1 + y_2) \\ \zeta(y_2) f_{NB}(y_1; \alpha_0 + \alpha_1, \frac{1}{\beta_1 + 1}) - \zeta(y_1 + y_2) \\ \zeta(y_1) f_{NB}(y_2; \alpha_0 + \alpha_2, \frac{1}{\beta_2 + 1}) - \zeta(y_1 + y_2) \end{pmatrix},$$

$$\begin{aligned} D_1 &= \partial_{\alpha_0} f_{BNB}(y_1, y_2; \boldsymbol{\alpha}, \boldsymbol{\beta}) \\ &= \sum_{k=0}^{y_1} \sum_{m=0}^{y_2} \left[H_0(k, m; y_1, y_2, \boldsymbol{\alpha}, \boldsymbol{\beta}) \Big\{ \psi(\alpha_0 + y_1 + y_2 - k - m) - \psi(\alpha_0) \right. \\ &\qquad \left. - \log(\beta_1 + \beta_2 + 1) \Big\} \right], \end{aligned}$$

$$\begin{aligned} D_2 &= \partial_{\alpha_1} f_{BNB}(y_1, y_2; \boldsymbol{\alpha}, \boldsymbol{\beta}) \\ &= \sum_{k=0}^{y_1} \sum_{m=0}^{y_2} \left[H_0(k, m; y_1, y_2, \boldsymbol{\alpha}, \boldsymbol{\beta}) \{ \psi(\alpha_1 + k) - \psi(\alpha_1) - \log(\beta_1 + 1) \} \right], \end{aligned}$$

$$\begin{aligned} D_3 &= \partial_{\alpha_2} f_{BNB}(y_1, y_2; \boldsymbol{\alpha}, \boldsymbol{\beta}) \\ &= \sum_{k=0}^{y_1} \sum_{m=0}^{y_2} \left[H_0(k, m; y_1, y_2, \boldsymbol{\alpha}, \boldsymbol{\beta}) \{ \psi(\alpha_2 + m) - \psi(\alpha_2) - \log(\beta_2 + 1) \} \right], \end{aligned}$$

$$\begin{aligned} D_4 &= \partial_{\beta_1} f_{BNB}(y_1, y_2; \boldsymbol{\alpha}, \boldsymbol{\beta}) \\ &= \sum_{k=0}^{y_1} \sum_{m=0}^{y_2} \left[H_0(k, m; y_1, y_2, \boldsymbol{\alpha}, \boldsymbol{\beta}) \left\{ \frac{y_1}{\beta_1} - \frac{\alpha_0 + y_1 + y_2 - k - m}{\beta_1 + \beta_2 + 1} - \frac{k + \alpha_1}{\beta_1 + 1} \right\} \right], \end{aligned}$$

$$\begin{aligned} D_5 &= \partial_{\beta_2} f_{BNB}(y_1, y_2; \boldsymbol{\alpha}, \boldsymbol{\beta}) \\ &= \sum_{k=0}^{y_1} \sum_{m=0}^{y_2} \left[H_0(k, m; y_1, y_2, \boldsymbol{\alpha}, \boldsymbol{\beta}) \left\{ \frac{y_2}{\beta_2} - \frac{\alpha_0 + y_1 + y_2 - k - m}{\beta_1 + \beta_2 + 1} - \frac{m + \alpha_2}{\beta_2 + 1} \right\} \right], \end{aligned}$$

$$
\begin{aligned}
D_6 &= \partial_{\alpha_j} f_{NB}(y_1; \alpha_0 + \alpha_1, \frac{1}{\beta_1 + 1}) \qquad\qquad j = 0, 1 \\
&= f_{NB}(y_1; \alpha_0 + \alpha_1, \frac{1}{\beta_1 + 1}) \Bigg[\psi(\alpha_0 + \alpha_1 + y_1) - \psi(\alpha_0 + \alpha_1) - \log(\beta_1 + 1) \Bigg], \\
D_7 &= \partial_{\beta_1} f_{NB}(y_1; \alpha_0 + \alpha_1, \frac{1}{\beta_1 + 1}) \\
&= f_{NB}(y_1; \alpha_0 + \alpha_1, \frac{1}{\beta_1 + 1}) \Bigg[\frac{y_1}{\beta_1} - \frac{\alpha_0 + \alpha_1 + y_1}{\beta_1 + 1} \Bigg], \\
D_8 &= \partial_{\alpha_j} f_{NB}(y_2; \alpha_0 + \alpha_2, \frac{1}{\beta_2 + 1}) \qquad\qquad j = 0, 2 \\
&= f_{NB}(y_2; \alpha_0 + \alpha_2, \frac{1}{\beta_2 + 1}) \Bigg[\psi(\alpha_0 + \alpha_2 + y_2) - \psi(\alpha_0 + \alpha_2) - \log(\beta_2 + 1) \Bigg], \quad \text{and} \\
D_9 &= \partial_{\beta_2} f_{NB}(y_2; \alpha_0 + \alpha_2, \frac{1}{\beta_2 + 1}) \\
&= f_{NB}(y_2; \alpha_0 + \alpha_2, \frac{1}{\beta_2 + 1}) \Bigg[\frac{y_2}{\beta_2} - \frac{\alpha_0 + \alpha_2 + y_2}{\beta_2 + 1} \Bigg].
\end{aligned}
$$

BIBLIOGRAPHY

Anderson-Bergman, C. (2017a). An efficient implementation of the emicm algorithm for the interval censored npmle. *Journal of Computational and Graphical Statistics*, 26(2):463–467.

Anderson-Bergman, C. (2017b). icenreg: Regression models for interval censored data in r. *Journal of Statistical Software*, 81(12):1–23.

Anderson-Bergman, C. and Yu, Y. (2016). Computing the log concave npmle for interval censored data. *Statistics and Computing*, 26(4):813–826.

Athey, S. and Imbens, G. (2016). Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences*, 113(27):7353–7360.

Athey, S., Tibshirani, J., Wager, S., et al. (2019). Generalized random forests. *The Annals of Statistics*, 47(2):1148–1178.

Banerjee, A., Yoganandan, N., Hsu, F.-C., Gayzik, S., and Pintar, F. A. (2019+). Evaluating performance of survival regression models with interval censored data in motorvehicle crash experiments. *A draft paper*.

Bellman, R. (1966). Dynamic programming. *Science*, 153(3731):34–37.

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinformatics*, 30(15):2114–2120.

Bou-Hamad, I., Larocque, D., Ben-Ameur, H., et al. (2011). A review of survival trees. *Statistics Surveys*, 5:44–71.

Breiman, L. (2001). Random forests. *Machine learning*, 45(1):5–32.

Breiman, L., Friedman, J., Olshen, R., and Stone, C. (1984). Classification and regression trees. wadsworth int. *Group*, 37(15):237–251.

Cameron, A. C., Li, T., Trivedi, P. K., and Zimmer, D. M. (2004). Modelling the differences in counted outcomes using bivariate copula models with application to mismeasured counts. *The Econometrics Journal*, 7(2):566–584.

Cameron, A. C. and Trivedi, P. K. (2013). *Regression analysis of count data*. Cambridge university press.

Chan, T. E., Stumpf, M. P., and Babtie, A. C. (2017). Gene regulatory network inference from single-cell data using multivariate information measures. *Cell systems*, 5(3):251–267.

Cho, H., Holloway, S. T., and Kosorok, M. R. (2020). Multi-stage optimal dynamic treatment regimes for survival outcomes with dependent censoring. *arXiv preprint arXiv:2012.03294*.

Cho, H., Jewell, N. P., and Kosorok, M. R. (2019). Interval censored recursive forests. *arXiv preprint arXiv:1912.09983*.

Chou, N. T. and Steenhard, D. (2011). Bivariate count data regression models - a SAS® macro program. Sas global forum - statistics and data analysis, SAS Institute.

Ciampi, A., du Berger, R., Taylor, H. G., and Thiffault, J. (1991). Recpam: A computer program for recursive partition and amalgamation for survival data and other situations frequently occurring in biostatistics. iii. classification according to a multivariate construct. application to data on haemophilus influenzae type b meningitis. *Computer methods and programs in biomedicine*, 36(1):51–61.

Cole, S. R. and Frangakis, C. E. (2009). The consistency statement in causal inference: a definition or an assumption? *Epidemiology*, 20(1):3–5.

Cui, Y., Zhu, R., Zhou, M., and Kosorok, M. (2017). Consistency of survival tree and forest models: splitting bias and correction. *arXiv preprint arXiv:1707.09631*.

Edwards, J. C. and Cambridge, G. (2006). B-cell targeting in rheumatoid arthritis and other autoimmune diseases. *Nature Reviews Immunology*, 6(5):394–403.

Efron, B. (1967). The two sample problem with censored data. in proceedings of the fifth berkeley symposium on mathematical statistics and probability. University of California Press, Berkeley, CA.

Eraslan, G., Simon, L. M., Mircea, M., Mueller, N. S., and Theis, F. J. (2019). Single-cell rna-seq denoising using a deep count autoencoder. *Nature communications*, 10(1):390.

Famoye, F. (2010). On the bivariate negative binomial regression model. *Journal of Applied Statistics*, 37(6):969–981.

Famoye, F. and Consul, P. C. (1995). Bivariate generalized poisson distribution with some applications. *Metrika*, 42(1):127–138.

Finkelstein, D. M. (1986). A proportional hazards model for interval-censored failure time data. *Biometrics*, 42(4):845–854.

Fu, W. and Simonoff, J. S. (2017). Survival trees for interval-censored survival data. *Statistics in medicine*, 36(30):4831–4842.

Gentleman, R. and Geyer, C. J. (1994). Maximum likelihood for interval censored data: Consistency and computation. *Biometrika*, 81(3):618–623.

Geurts, P., Ernst, D., and Wehenkel, L. (2006). Extremely randomized trees. *Machine learning*, 63(1):3–42.

Goldberg, Y. and Kosorok, M. R. (2012). Q-learning with censored data. *Annals of statistics*, 40(1):529.

Gordon, L. and Olshen, R. A. (1985). Tree-structured survival analysis. *Cancer treatment reports*, 69(10):1065–1069.

Graf, E., Schmoor, C., Sauerbrei, W., and Schumacher, M. (1999). Assessment and comparison of prognostic classification schemes for survival data. *Statistics in medicine*, 18(17-18):2529–2545.

Groeneboom, P. (1991). Nonparametric maximum likelihood estimators for interval censoring and deconvolution.

Groeneboom, P., Jongbloed, G., Witte, B. I., et al. (2010). Maximum smoothed likelihood estimation and smoothed maximum likelihood estimation in the current status model. *The Annals of Statistics*, 38(1):352–387.

Groeneboom, P. and Wellner, J. A. (1992). *Information bounds and nonparametric maximum likelihood estimation*, volume 19. Springer Science & Business Media.

Gurmu, S. and Elder, J. (1999). Generalized bivariate count data regression models. *Economics Letters*, 68(1):31–36.

Habermann, T. M., Weller, E. A., Morrison, V. A., Gascoyne, R. D., Cassileth, P. A., Cohn, J. B., Dakhil, S. R., Woda, B., Fisher, R. I., Peterson, B. A., et al. (2006). Rituximab-chop versus chop alone or with maintenance rituximab in older patients with diffuse large b-cell lymphoma. *Journal of clinical oncology*, 24(19):3121–3127.

Haegeli, P., Falk, M., Brugger, H., Etter, H.-J., and Boyd, J. (2011). Comparison of avalanche survival patterns in canada and switzerland. *Cmaj*, 183(7):789–795.

Hernán, M. A. and Robins, J. M. (2020). Causal inference: What if.

Hicks, S. C., Townes, F. W., Teng, M., and Irizarry, R. A. (2017). Missing data and technical variability in single-cell rna-sequencing experiments. *Biostatistics*, 19(4):562–578.

Hogan, J. and Radhakrishnan, J. (2013). The treatment of minimal change disease in adults. *Journal of the American Society of Nephrology*, 24(5):702–711.

Hothorn, T., Bühlmann, P., Dudoit, S., Molinaro, A., and Van Der Laan, M. J. (2005). Survival ensembles. *Biostatistics*, 7(3):355–373.

Hothorn, T., Lausen, B., Benner, A., and Radespiel-Tröger, M. (2004). Bagging survival trees. *Statistics in medicine*, 23(1):77–91.

Huang, H.-J., Yang, L.-Y., Tung, H.-J., Ku, F.-C., Wu, R.-C., Tang, Y.-H., Chang, W.-Y., Jung, S.-M., Wang, C.-C., Lin, C.-T., et al. (2020). Management and clinical outcomes of patients with recurrent/progressive ovarian clear cell carcinoma. *Journal of the Formosan Medical Association*, 119(4):793–804.

Huang, J. and Wellner, J. A. (1997). Interval censored survival data: a review of recent progress. In *Proceedings of the First Seattle Symposium in Biostatistics*, pages 123–169. Springer.

Huang, M., Wang, J., nd H. Dueck, E. T., Shaffer, S., Bonasio, R., ..., and Zhang, N. R. (2018). Saver: gene expression recovery for single-cell rna sequencing. *Nature methods*, 15(7):539–542.

Huang, X., Ning, J., and Wahed, A. S. (2014). Optimization of individualized dynamic treatment regimes for recurrent diseases. *Statistics in medicine*, 33(14):2363–2378.

Iacono, G., Massoni-Badosa, R., and Heyn, H. (2019). Single-cell transcriptomics unveils gene regulatory network plasticity. *Genome biology*, 20(1):110.

Ishwaran, H., Kogalur, U. B., Blackstone, E. H., Lauer, M. S., et al. (2008). Random survival forests. *The annals of applied statistics*, 2(3):841–860.

Ishwaran, H. and Lu, M. (2014). Random survival forests. *Wiley StatsRef: Statistics Reference Online*, pages 1–13.

Jewell, N. P. and Emerson, R. (2013). Current status data: An illustration with data on avalanche victims. *Handbook of Survival Analysis*, pages 391–412.

Jiang, R., Lu, W., Song, R., and Davidian, M. (2017). On estimation of optimal treatment regimes for maximizing t-year survival probability. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 79(4):1165–1185.

Jongbloed, G. (1998). The iterative convex minorant algorithm for nonparametric estimation. *Journal of Computational and Graphical Statistics*, 7(3):310–321.

Jørgensen, B. (1987). Exponential dispersion models. *Journal of the Royal Statistical Society: Series B (Methodological)*, 49(2):127–145.

Kidwell, K. M. (2015). Chapter 2: Dtrs and smarts: Definitions, designs, and applications. In *Adaptive Treatment Strategies in Practice: Planning Trials and Analyzing Data for Personalized Medicine*, pages 7–23. SIAM.

Kocherlakota, S. and Kocherlakota, K. (1992). *Bivariate Discrete Distributions*. Marcel Dekker: New York.

Kolodziejczyk, A. A., Kim, J. K., Svensson, V., Marioni, J. C., and Teichmann, S. A. (2015). The technology and biology of single-cell rna sequencing. *Molecular cell*, 58(4):610–620.

Kosorok, M. R. (2007). *Introduction to empirical processes and semiparametric inference*. Springer Science & Business Media.

Kosorok, M. R. and Laber, E. B. (2019). Precision medicine. *Annual review of statistics and its application*, 6:263–286.

Kosorok, M. R. and Moodie, E. E. (2015). *Adaptive treatment strategies in practice: planning trials and analyzing data for personalized medicine*. SIAM.

LeBlanc, M. and Crowley, J. (1992). Relative risk trees for censored survival data. *Biometrics*, pages 411–425.

LeBlanc, M. and Crowley, J. (1993). Survival trees by goodness of split. *Journal of the American Statistical Association*, 88(422):457–467.

Li, C., Lu, J., Park, J., Kim, K., Brinkley, P. A., and Peterson, J. P. (1999). Multivariate zero-inflated poisson models and their applications. *Technometrics*, 41(1):29–38.

Li, W. V. and Li, J. J. (2018). An accurate and robust imputation method scimpute for single-cell rna-seq data. *Nature communications*, 9(1):997.

Linn, K. A., Laber, E. B., and Stefanski, L. A. (2017). Interactive q-learning for quantiles. *Journal of the American Statistical Association*, 112(518):638–649.

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for rna-seq data with deseq2. *Genome biology*, 15(12):550.

Maher, M. J. (1990). A bivariate negative binomial model to explain traffic accident migration. *Accident Analysis & Prevention*, 22(5):487–498.

Mc Mahon, S. S., Sim, A., Filippi, S., Johnson, R., Liepe, J., Smith, D., and Stumpf, M. P. (2014). Information theory and signal transduction systems: from molecular information processing to network inference. volume 35 of *Seminars in cell & developmental biology*, pages 98–108. Academic Press.

Meinshausen, N. (2006). Quantile regression forests. *Journal of Machine Learning Research*, 7(Jun):983–999.

Mentch, L. and Zhou, S. (2020). Randomization as regularization: A degrees of freedom explanation for random forest success. *Journal of Machine Learning Research*, 21(171):1–36.

Molinaro, A. M., Dudoit, S., and Van der Laan, M. J. (2004). Tree-based multivariate regression and density estimation with right-censored data. *Journal of Multivariate Analysis*, 90(1):154–177.

Mullahy, J. (1986). Specification and testing of some modified count data models. *Journal of Econometrics*, 33(3):341–365.

Murphy, S. A. (2003). Optimal dynamic treatment regimes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(2):331–355.

Murphy, S. A. (2005). A generalization error for q-learning. *Journal of Machine Learning Research*, 6(Jul):1073–1097.

Murphy, S. A., Oslin, D. W., Rush, A. J., and Zhu, J. (2007). Methodological challenges in constructing effective treatment sequences for chronic psychiatric disorders. *Neuropsychopharmacology*, 32(2):257–262.

Oller, R., Gómez, G., and Calle, M. L. (2004). Interval censoring: model characterizations for the validity of the simplified likelihood. *Canadian Journal of Statistics*, 32(3):315–326.

Orellana, L., Rotnitzky, A., and Robins, J. M. (2010). Dynamic regime marginal structural mean models for estimation of optimal dynamic treatment regimes, part i: main content. *The international journal of biostatistics*, 6(2).

Peng, T., Zhu, Q., Yin, P., and Tan, K. (2019). Scrabble: single-cell rna-seq imputation constrained by bulk rna-seq data. *Genome biology*, 20(1):88.

Peto, R. and Peto, J. (1972). Asymptotically efficient rank invariant test procedures. *Journal of the Royal Statistical Society: Series A (General)*, 135(2):185–198.

Polyanskiy, Y. and Wu, Y. (2020). Self-regularizing property of nonparametric maximum likelihood estimator in mixture models. *arXiv preprint arXiv:2008.08244*.

Pont, F., Tosolini, M., and J, F. J. (2019). Single-cell signature explorer for comprehensive visualization of single cell signatures across scrna-seq data sets. *Nucleic Acids Research*, 47:e133.

Preisser, J. S., Das, K., Long, D. L., and Divaris, K. (2016). Marginalized zero-inflated negative binomial regression with application to dental caries. *Statistics in M edicine*, 35(10):1722–1735.

Qian, M. and Murphy, S. A. (2011). Performance guarantees for individualized treatment rules. *Annals of statistics*, 39(2):1180.

R Core Team (2019). *R: A language and environment for statistical computing*.

Risso, D., Perraudeau, F., Gribkova, S., Dudoit, S., and Vert, J. P. (2018). A general and flexible method for signal extraction from single-cell rna-seq data. *Nature communications*, 9(1):284.

Robertson, J. B. and Uppuluri, V. (1984). A generalized kaplan-meier estimator. *The Annals of Statistics*, pages 366–371.

Robins, J. M. and Rotnitzky, A. (1992). Recovery of information and adjustment for dependent censoring using surrogate markers. In *AIDS epidemiology*, pages 297–331. Springer.

Robins, J. M., Rotnitzky, A., and Zhao, L. P. (1994). Estimation of regression coefficients when some regressors are not always observed. *Journal of the American statistical Association*, 89(427):846–866.

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edger: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26(1):139–140.

Rubin, D. B. (2005). Bayesian inference for causal effects. *Handbook of statistics*, 25:1–16.

Schulte, P. J., Tsiatis, A. A., Laber, E. B., and Davidian, M. (2014). Q-and a-learning methods for estimating optimal dynamic treatment regimes. *Statistical science: a review journal of the Institute of Mathematical Statistics*, 29(4):640.

Segal, M. R. (1988). Regression trees for censored data. *Biometrics*, pages 35–47.

Simoneau, G., Moodie, E. E., Nijjar, J. S., Platt, R. W., and Investigators, S. E. R. A. I. C. (2019). Estimating optimal dynamic treatment regimes with survival outcomes. *Journal of the American Statistical Association*, (just-accepted):1–24.

Sorlie, P. D., Backlund, E., and Keller, J. B. (1995). Us mortality by economic, demographic, and social characteristics: the national longitudinal mortality study. *American Journal of Public Health*, 85(7):949–956.

Steingrimsson, J. A., Diao, L., and Strawderman, R. L. (2019). Censoring unbiased regression trees and ensembles. *Journal of the American Statistical Association*, 114(525):370–383.

Sun, J. (2007). *The statistical analysis of interval-censored failure time data*. Springer.

Sutton, R. S. and Barto, A. G. (2018). *Reinforcement learning: An introduction*. The MIT Press.

Svensson, V. (2020). Droplet scrna-seq is not zero-inflated. *Nature Biotechnology*, 38(2):147–150.

Townes, F. W., Hicks, S. C., Aryee, M. J., and Irizarry, R. A. (2019). Feature selection and dimension reduction for single-cell rna-seq based on a multinomial model. *Genome biology*, 20(1):1–16.

Tsiatis, A. (1975). A nonidentifiability aspect of the problem of competing risks. *Proceedings of the National Academy of Sciences*, 72(1):20–22.

Tsiatis, A. A., Davidian, M., T, H. S., and Laber, E. B. (2019). *Dynamic Treatment Regimes: Statistical Methods for Precision Medicine*. CRC press.

van den Berge, K., Perraudeau, F., Soneson, C., Love, M. I., Risso, D., Vert, J. P., ..., and Clement, L. (2018). Observation weights unlock bulk rna-seq tools for zero inflation and single-cell applications. *Genome biology*, 19(1):24.

van der Vaart, A. W. and Wellner, J. (2013). *Weak convergence and empirical processes: with applications to statistics*. Springer Science & Business Media.

Van Dijk, D., Sharma, R., Nainys, J., Yim, K., Kathail, P., Carr, A. J., Burdziak, C., Moon, K. R., Chaffer, C. L., Pattabiraman, D., et al. (2018). Recovering gene interactions from single-cell data using data diffusion. *Cell*, 174(3):716–729.

Vieth, B., Ziegenhain, C., Parekh, S., Enard, W., and Hellmann, I. (2017). powsimr: power analysis for bulk and single cell rna-seq experiments. *Bioinformatics*, 33(21):3486–3488.

Wager, S. and Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 113(523):1228–1242.

Wager, S. and Walther, G. (2015). Adaptive concentration of regression trees, with application to random forests. *arXiv preprint arXiv:1503.06388*.

Wahed, A. S. and Thall, P. F. (2013). Evaluating joint effects of induction–salvage treatment regimes on overall survival in acute leukaemia. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 62(1):67–83.

Wahed, A. S. and Tsiatis, A. A. (2006). Semiparametric efficient estimation of survival distributions in two-stage randomisation designs in clinical trials with censored data. *Biometrika*, 93(1):163–177.

Wallace, M. P. and Moodie, E. E. (2015). Doubly-robust dynamic treatment regimen estimation via weighted least squares. *Biometrics*, 71(3):636–644.

Wang, J., Huang, M., Torre, E., Dueck, H., Shaffer, S., Murray, J., Raj, A., Li, M., and Zhang, N. R. (2018a). Gene expression distribution deconvolution in single-cell rna sequencing. *Proceedings of the National Academy of Sciences*, 115(28):E6437–E6446.

Wang, L., Zhou, Y., Song, R., and Sherwood, B. (2018b). Quantile-optimal treatment regimes. *Journal of the American Statistical Association*, 113(523):1243–1254.

Wang, P. (2003). A bivariate zero-inflated negative binomial regression model for count data with excess zeros. *Economics Letters*, 78(3):373–378.

Watkins, C. J. C. H. (1989). *Learning from delayed rewards*. PhD thesis.

Wellner, J. A. and Zhan, Y. (1997). A hybrid algorithm for computation of the nonparametric maximum likelihood estimator from censored data. *Journal of the American Statistical Association*, 92(439):945–959.

Xu, Y., Müller, P., Wahed, A. S., and Thall, P. F. (2016). Bayesian nonparametric estimation for dynamic treatment regimes with sequential transition times. *Journal of the American Statistical Association*, 111(515):921–950.

Yang, C., Diao, L., and Cook, R. (2021). Survival trees for current status data. In *Survival Prediction-Algorithms, Challenges and Applications*, pages 83–94. PMLR.

Yao, W., Frydman, H., and Simonoff, J. S. (2019). An ensemble method for interval-censored time-to-event data. *Biostatistics*.

Yin, Y., Anderson, S. J., Parran Hall, G., Street, D., et al. (2002). Nonparametric tree-structured modeling for interval-censored survival data. In *Joint Statistical Meeting*.

Yu, T. (2018). A new dynamic correlation algorithm reveals novel functional aspects in single cell and bulk rna-seq data. *PLoS computational biology*, 14(8):e1006391.

Zhang, B. and Horvath, S. (2005). A general framework for weighted gene co-expression network analysis. *Statistical applications in genetics and molecular biology*, 4(1).

Zhang, Y., Laber, E. B., Davidian, M., and Tsiatis, A. A. (2017). Estimation of optimal treatment regimes using lists. *Journal of the American Statistical Association,(just-accepted)*.

Zhao, Y., Kosorok, M. R., and Zeng, D. (2009). Reinforcement learning design for cancer clinical trials. *Statistics in medicine*, 28(26):3294–3315.

Zhao, Y., Zeng, D., Rush, A. J., and Kosorok, M. R. (2012). Estimating individualized treatment rules using outcome weighted learning. *Journal of the American Statistical Association*, 107(499):1106–1118.

Zhao, Y., Zeng, D., Socinski, M. A., and Kosorok, M. R. (2011). Reinforcement learning strategies for clinical trials in nonsmall cell lung cancer. *Biometrics*, 67(4):1422–1433.

Zhou, Y. and McArdle, J. J. (2015). Rationale and applications of survival tree and survival ensemble methods. *Psychometrika*, 80(3):811–833.

Zhu, R. and Kosorok, M. R. (2012). Recursively imputed survival trees. *Journal of the American Statistical Association*, 107(497):331–340.

Zhu, R., Zeng, D., and Kosorok, M. R. (2015). Reinforcement learning trees. *Journal of the American Statistical Association*, 110(512):1770–1784.