On Estimation of Time-varying Population Attributable Fractions for Population-based Case-control Studies

Wei Zhao

A dissertation[†]
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2014

Reading Committee:

Ying Qing Chen, Chair

Li Hsu, Chair

Susanne May

Ching-Yun Wang

Program Authorized to Offer Degree:
Biostatistics

[†]an egocentric imitation, actually

University of Washington

Abstract

On Estimation of Time-varying Population Attributable Fractions for Population-based Case-control Studies

Wei Zhao

Co-Chairs of the Supervisory Committee:
Affiliate Professor Ying Qing Chen
Department of Biostatistics

Affiliate Professor Li Hsu
Department of Biostatistics

The population attributable fraction (PAF) is an important measure in evaluating the contribution of risk factors to disease burden or mortality. It is a useful tool for planning of prevention actions. Recent development of PAF has extended the classic static measure to a more dynamic time-varying functional form, which may provide more details of the underlying quantity.

This dissertation focuses on estimation of the time-varying PAF for population-based case-control studies. We consider both situations with or without adjustment for confounders. The underlying model is assumed to be the Cox model with possibly time-varying covariates. We propose three kernel-type estimators of the time-varying PAF for case-control studies. We study large sample properties of these estimators, and prove strong uniform consistency and asymptotic normality. We also study the finite sample properties of these estimators, and derive variance estimators with corrections for the finite sampling. We also propose a resampling based approach for calculating the simultaneous confidence bands. Simulation studies show that all the point and variance estimators perform well under all the designed scenarios. We apply our proposed methods the Genetics and Epidemiology of Colorectal Cancer Consortium (GECCO) smoking data to understand the risk of colorectal cancer attributable to smoking.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my dissertation advisors, Ying Qing Chen, and Li Hsu, for their inspiration and guidance. I could not have asked for more brilliant and supportive mentors. I would also like to thank the members of my supervisory committee, Susanne May, Ching-Yun Wang, and Chongzhi Di for their invaluable inputs into this dissertation. I want to thank Polly Newcomb and Thomas Vaughan for serving as my Graduate School Representative. I also want to extend my thanks to the GECCO study group for providing the data.

Also, I would like to acknowledge my wife Shanshan. Without her constant love and support, I could not have completed this journey.

DEDICATION

to my parents

Chapter 1

INTRODUCTION

1.1 Population Attributable Fraction

Disease prevention programs often have to prioritize risk factors for public health preventive intervention. Assessing the impact associated with exposure to the risk factors thus is critical. For this purpose, the population attributable fraction (PAF) has been widely used to quantify the impact of exposure on disease burden at the population level, and helps to guide public health policies. The concept of PAF was first introduced by Levin [20] to quantify the impact of smoking on lung cancer occurrence, and thereafter became a useful tool for measuring impact of risk factors on disease burden. There are as many as 16 different terms in the literature to denote the PAF, among which the most commonly used terms other than PAF are "attributable risk", "population attributable risk", and "etiologic fraction" [29, 3]. The PAF of an exposure to a disease is defined as the proportional reduction in the probability of disease by comparing the current population with a hypothetical population which is the same as the current population but the exposure had been eliminated. In the simple case of binary disease (D) and exposure status (Z), PAF can be written as [20]

$$PAF = \frac{P(D=1) - P(D=1|Z=0)}{P(D=1)}, \tag{1.1}$$

where $P(D=1)$ is the probability of being diseased in the population, and $P(D=1|Z=0)$ is the hypothetical probability of being diseased in the same population but with the exposure eliminated. Therefore, the PAF measures the proportion of the excessive disease probability in the population that is associated with the presence of exposure. Consequently, PAF is used to assess the potential impact of public health actions aimed at eliminating exposure from the population. The probabilities in equation (1.1) usually refer to disease risk or incidence, with emphasis on assessing impact of exposure to disease occurrence. For cross-sectional studies, these probabilities could also be referred to as disease prevalence, with which PAF quantifies the impact of exposure on the overall burden of disease rather than disease occurrence.

PAF combines both the strength of association between exposure and disease and the prevalence of exposure in the population. It is therefore a more appropriate measure of the population level impact of risk factors than association measures such as relative risk or rate ratio. To see this property of PAF, it is helpful to rewrite equation (1.1) into other two forms. By noting that $P(D=1) = P(D=1|Z=1)P(Z=1) + P(D=1|Z=0)P(Z=0)$ and the relative risk or rate ratio $RR = P(D=1|Z=1)/P(D=1|Z=0)$, we divide both the numerator and denominator by $P(D=1|Z=0)$ and can rewrite the formula for PAF as [10, 21],

$$PAF = \frac{P(Z=1)(RR-1)}{1+P(Z=1)(RR-1)}, \tag{1.2}$$

which is a function of the prevalence of exposure in the population, $P(Z=1)$, and the relative risk or rate ratio, RR. Another approach is to express $P(D=1)$ by $P(D=1) = P(D=1|Z=1)P(Z=1)/P(Z=1|D=1)$ using Bayes' theorem, and rewrite the formula for PAF as [21]

$$PAF = \frac{P(Z=1|D=1)(RR-1)}{RR} = P(Z=1|D=1)\left(1-\frac{1}{RR}\right), \tag{1.3}$$

which a function of the prevalence of exposure in diseased individuals and the relative risk. Formulas (1.2) and (1.3) explicitly describe how the measure of PAF integrates the association measure and the prevalence of exposure. A risk factor with high relative risk but low prevalence of exposure could have low PAF, whereas a risk factor with low relative risk but high prevalence could have high PAF. The range of PAF lies between 0 and 1 when the exposure is a risk factor ($RR > 1$). PAF increases with increase of the relative risk (or rate ratio) as well as with increase of the prevalence of exposure. When the prevalence of exposure reaches the maximum value 1 or 100%, the value of PAF equals to $(RR-1)/RR$. When RR goes to infinity, the value of PAF goes to 1 as long as the exposure is present in the population. PAF equals to 0 when either there is no association between exposure and disease ($RR = 1$) or the exposure is not present in the population, $P(Z=1) = 0$.

The formulas for PAF presented so far consider only one dichotomous risk factor. In a more realistic situation, risk factors can have multiple levels of exposure or there are multiple risk factors. When multilevel or multivariate exposures are considered, the definition of PAF by formula (1.1) still holds as long as the reference group ($Z = 0$) is well defined. To see how formulas (1.2) and (1.3) could be generalized to a multifactoral setting, let Z be the vector of all categorical exposures of interest, and suppose Z takes J levels according to cross-classification of these categorical variables, $z_0, z_1, ..., z_J$, where z_0 is the reference level. Then equation (1.2) and (1.3) can be generalized as [30]

$$PAF = \frac{\sum_{i=1}^{J} P(Z = i)(RR_i - 1)}{1 + \sum_{i=1}^{J} P(Z = i)(RR_i - 1)} \tag{1.4}$$

and [21]

$$PAF = \sum_{i=1}^{J} P(Z = i|D = 1)\left(1 - \frac{1}{RR_i}\right) \tag{1.5}$$

respectively, where $RR_i = P(D = 1|Z = i)/P(D = 1|Z = 0)$ is the relative risk or rate ratio of developing a disease in group z_i comparing with the reference group. It is also of interest to consider the situation when Z contains exposure variables that are continuous or semi-continuous. By the total probability theorem for continuous variable, equation (1.2) can be generalized as [5]

$$PAF = \frac{\int_{\mathcal{Z}} \{RR_z - 1\} dF_Z(z)}{1 + \int_{\mathcal{Z}} \{RR_z - 1)\} dF_Z(z)} = 1 - \left\{ \int_{\mathcal{Z}} RR_z dF_Z(z) \right\}^{-1} \tag{1.6}$$

and (1.3) can be generalized as

$$PAF = \int_{\mathcal{Z}} \left\{1 - RR_z^{-1}\right\} dF_{Z|D=1}(z) = 1 - \int_{\mathcal{Z}} RR_z^{-1} dF_{Z|D=1}(z) \tag{1.7}$$

where $RR_z = P(D = 1|Z = z)/P(D = 1|Z = 0)$, $F_Z(z)$ is the distribution function of the exposure in the population, and $F_{Z|D=1}(z|D = 1)$ is the conditional distribution function of the exposure in diseased individuals. These formulas facilitate estimation of PAF with various types of exposures.

PAF can be estimated from major study designs, such as cross-sectional, cohort, and case-control studies. In cross-sectional studies, all probabilities in formula (1.1) can be directly estimated, so the point estimate of PAF can be obtained from plugging in empirical estimates of these probabilities to formula (1.1) [30]. The same methods can be used for estimation of PAF in cohort studies with a fixed follow-up time. In cohort studies with unfixed follow-up time which yield time-to-event outcome, survival models such as the Cox proportional hazards model are needed to obtain unbiased estimates. PAF can be obtained through formula (1.2) or (1.3) by incorporating estimates of rate ratio from survival models and estimates of the prevalence $P(Z = 1)$ or $P(Z = 1|D = 1)$ from the sample. In case-control studies, formula (1.2) or (1.3) can be used to estimate PAF by estimating odds ratio to approximate the relative risk RR and estimating the prevalence $P(Z = 1)$ from the controls (by assuming a rare disease) or $P(Z = 1|D = 1)$ from the cases. It is worth noting that the interpretation of the PAF measure in the cross-sectional studies is different from those in cohort or case-control studies. Since no follow-up is involved in cross-sectional studies, the probabilities in the PAF formula represent prevalence rather than incidence, and the PAF in this situation measures the impact of exposure on the overall burden of disease rather than disease occurrence. Variance estimates of the aforementioned point estimates of PAF can be obtained from either applying the delta method [23] and considering appropriate distributions, or from resampling based methods such as the bootstrap [18].

1.2 Adjusted PAF to account for confounding

The PAF estimates introduced in section 1.1 fail to take into account other factors that may confound the association between the risk factor of interest and disease. Since confounders cannot be avoided in most observational studies, the unadjusted PAF estimates are in general biased [21, 30, 31]. The unadjusted estimates of PAF can be misleading in some situation. For example, Walter [31] compared unadjusted and adjusted PAF estimates for a case-control study assessing the association between alcohol, smoking and oral cancer. He found that the differences between unadjusted and adjusted estimates were both large for smoking (0.51 vs 0.31) and alcohol (0.52 vs 0.37).

An appropriate measure of adjusted PAF would be the proportion of disease risk attributable to an exposure given the same level of other (confounding) factors. There are three main approaches for adjusting for confounders in the estimation of PAF: the weighted-sum approach, the Mantel-Haenszel approach, and the regression modeling approach. The weighted-sum approach is a full stratification approach that corresponds to a saturated model. With this approach, PAF is written as a weighted sum of the stratum-specific PAFs over strata formed by the adjustment factors [30, 33],

$$PAF = \sum_{j=1}^{J} w_j PAF_j = \sum_{j=1}^{J} w_j \frac{P(D=1|U=u_j) - P(D=1|Z=0,U=u_j)}{P(D=1|U=u_j)} \tag{1.8}$$

where U is the vector of adjustment factors forming J levels, and w_j and PAF_j denote the weight and the stratum-specific PAF of level u_j. Different measures of adjusted PAF could be produced by choosing different weights. A common choice of the weights is the proportion of diseased individuals in stratum j (i.e., $w_j = P(U = u_j|D = 1)$) which yields an unbiased estimator [33]. An alternative choice of weights is given by setting w_j as the inverse variance of the PAF estimate in level j over the sum of inverse variances over all levels. However, this choice of weights has been shown to result in inconsistent estimate except in some special cases [2]. Estimation of weighted-sum adjusted PAF has been focused on estimating the stratum-specific PAF in each stratum as estimation of the unadjusted PAF, and estimating the weights from the sample observation. The weighted-sum approach is appealing because of its straightforward structure and application. However, as the number of adjustment and exposure levels increases, it suffers from sample size bias and cannot guarantee a reasonable estimator. Furthermore, stratification requires that both the risk factors of interest and the confounding factors be categorical, which may result in loss of information.

In the the Mantel-Haenszel approach, estimate of PAF is obtained by replacing RR in formula (1.3) with an adjusted RR (odds ratio in case-control study) estimated by a Mantel-Haenszel estimator [15]. The assumption of a common relative risk or odds ratio cross strata formed by adjustment factors is required for this method to be valid. Therefore, the Mantel-Haenszel approach is a stratification approach with restriction on the relative risks across adjustment strata, and corresponds to a model with only main effects for exposure and adjustment factors and not any interactions. For this reason, this approach cannot be used to adjust for effect modifiers, but in cases without interactions it is more efficient than the weighted-sum approach.

The two stratification approaches for adjusting for PAF estimates are based on opposite options. The weighted-sum approach does not assume any structure on the relative risk across adjustment factors, while the Mantel-Haenszel approach assumes a common relative risk across all the adjustment strata. In a modeling point of view, both of these options are too extreme, and there could be some intermediate situation, for example, when confounders and effect modifiers both present. To overcome the difficulties of the stratification based approaches, the regression modeling approach have been developed [30, 33, 15]. This approach takes advantage of flexibility and efficiency of regression models, and includes the stratification based approaches as special cases. With this approach, PAF is expressed as [33, 7]

$$\begin{aligned} PAF &= \frac{P(D=1) - \sum_{j=1}^{J} P(D=1|Z=0, U=u_j)P(U=u_j)}{P(D=1)} \qquad (1.9) \\ &= 1 - \sum_{j=1}^{J}\sum_{i=0}^{I} \frac{P(Z=z_i, U=u_j|D=1)}{RR_{i|j}} \end{aligned}$$

where $RR_{i|j} = P(D=1|Z=z_i, U=u_j)/P(D=1|Z=0, U=u_j)$ is the relative risk or rate ratio for level i of exposure given level j of adjustment factors. It is worth noting that this formula coincides with the weighted-sum adjustment formula (1.8) by setting the weights $w_j = P(U=u_j|D=1)$. However, the model based approach makes use regression models rather than using a full stratification in the estimation procedure, therefore yields more flexible and more efficient estimates. The relative risk or rate ratio can be estimated through a regression model in cohort or cross-sectional studies, and odds ratio can be used to approximate relative risk in case-control studies (by assuming rare disease). The quantity $P(D=1|Z=z_i, U=u_j)$ can be estimated either through observed proportions or through regression models [16]. Variance estimators have been developed based on asymptotic methods [4, 5, 1] or resampling methods such as bootstrapping [18, 17].

The classic PAF introduced above is a quantity defined over a pre-specified time interval. To make a valid interpretation of PAF, the time interval needs to be defined such that the disease risk or incidence rate are meaningful. Therefore, the PAF is a static measure that quantifies the overall impact of exposure over a specific period. From a dynamic point of view, the impact of an exposure to disease could change over time, because the two components of PAF, the prevalence of exposure and the relative risk, can both change over time in a realistic situation. A measurement that captures this dynamic pattern would provide additional information on when and how the exposure impact varies over time and such information is helpful in deciding the timing of actions/interventions. Therefore, it is necessary to extend the concept of PAF to a time-varying measure and develop the estimators for commonly used study designs. We will introduce the extension of PAF to time varying measures in Section 4. Since the survival analysis of time-to-event data is the focus of those measures, we will first briefly review some major concepts and techniques in survival analysis below.

1.3 Time-to-event Data and the Cox Model

The time-to-event data analysis has important applications in clinical and epidemiological studies to evaluate association between disease and exposure to risk factors. In the time-to-event data analysis, the outcome of interest is time to the occurrence of an event (e.g. disease occurrence). The occurrence of the event is often referred to as "failure", and the time to the occurrence of the event is referred to as "failure time". Sometimes the exact failure time for some subjects may not be observed, due to either expected or unexpected reasons such as loss of followup or end of study before the event occurs, yet the failure time is known to be greater or less than certain values. This situation is referred to as "censoring". The most common type of censoring is the "right censoring", which means the failure time is unknown but is known to be greater than a certain value. Time-to-event data often come from prospective studies such as cohort or case-cohort studies. They could also be derived from retrospective studies such as case-control studies. For example, in a case-control study in which ages of subjects are recorded, the outcome may be age-at-onset, which can be regarded as the failure time - time from birth (start point) to disease occurrence, and the age at study of controls (non-diseased subjects) can be regarded as right-censored.

Hazard rate is usually used to describe the distribution of the failure time. The hazard rate at time t is defined as the instantaneous rate of failure in an arbitrarily short interval Δt around time t, provided that the event has not occurred at time t. It can be written as:

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} \tag{1.10}$$

where T is the failure time. The hazard ratio between hazard rate with certain exposures and hazard rate with reference exposures is usually used to measure the association between exposures and disease risk. The Cox proportional hazards model [12] (often referred as the Cox model) is commonly used to estimate the hazard ratio for right-censored time-to-event data. The Cox model links the hazard rate at time t with covariate effect through a baseline hazard function, by assuming the hazard ratio is independent of time. The equation can be expressed as,

$$\lambda(t|Z) = \lambda_0(t) \exp(\beta^T Z) \tag{1.11}$$

where Z is a vector of covariates, $\lambda_0(t)$ is an unspecified baseline hazard rate at time t, $\lambda(t|Z)$ is the hazard rate at time t with covariates indicated by Z, and β is a vector of log hazard ratios. As its name indicates, one essential property of the Cox model is the proportional hazards assumption, which implies that $\exp(\beta^T Z)$ in model (1.11) is free of time t.

Because of the nonparametric component $\lambda_0(t)$, a partial likelihood approach is used to estimate β. Suppose that the sample of n subjects consists k failure time (uncensored) $t_1 < t_2 < \cdots < t_k$, and there is no tie in failure time (i.e. only one subject fails in each unique failure time). The remaining $n - k$ subjects are right-censored. Let j denote the subject failing at time t_j and $R(t)$ denote the set of subjects at risk for failure at time t, then the partial likelihood is defined as

$$L(\beta) = \prod_{j=1}^{k} \frac{\exp(\beta^T Z_j)}{\sum_{l \in R(t_j)} \exp(\beta^T Z_l)} \tag{1.12}$$

The point estimator of regression parameter β is obtained by solving the score equation of the partial likelihood function, $\partial log L(\beta)/\beta = 0$, and the variance estimator is obtained through the second derivative of the log partial likelihood.

The proportional hazards assumption required by the Cox model may not be realistic in practice, and the interpretation can be quite tricky. To relax the proportional hazards assumption, the Cox model can be extended to include time-varying covariate $Z(t)$ and/or time-varying coeffiecient $\beta(t)$. When the covariates are time varying while the coefficient is independent of time, one can still use the partial-likelihood approach to obtain valid estimation of coefficient β. When the coeffiecient is time-varying, the estimation procedure needs special considerations, such as histogram sieve estimation with assumption that $\beta(t)$ is piecewise constant, or local fitting techniques coupled with the partial likelihood estimation.

1.4 Time-varying PAF

Chen et al. (2006) firstly extended PAF to allow it to vary with time for cohort studies [9]. Under the survival analysis framework for time-to-event data, a natural extension of PAF is to substitute the probabilities of disease in formula (1.1) with the respective cumulative distribution functions of failure time T [9]:

$$\Phi(t) = \frac{P(T \leq t) - P(T \leq t|Z=0)}{P(T \leq t)} = \frac{F(t) - F(t|Z=0)}{F(t)}. \tag{1.13}$$

This measure focuses on the attributable risk in a cumulative manner - the PAF of disease risk due to an exposure over a period from baseline to time t. When the instantaneous incidence rate is considered, an alternative measure called "attributable hazard function" can be used. The attributable hazard function is defined as the proportional reduction of hazard rate comparing the current population with a population with no exposure [9]:

$$\phi(t) = \frac{\lambda(t) - \lambda_0(t)}{\lambda(t)}, \tag{1.14}$$

where $\lambda(t)$ is the composite hazard function defined as the instantaneous rate of failing at time t, and $\lambda_0(t) = \lambda(t|Z=0)$ is the baseline hazard function. An extended measure of $\phi(t)$ is the average attributable hazard function, i.e. $\overline{\phi}(t) = t^{-1}\int_0^t \phi(u)du$. Under a rare disease scenario, $\Phi(t)$ and $\phi(t)$ numerically approximate each other [9], whereas simulations show that the two measures could be substantially different when the disease is not rare [26, 11]. The explicit relationship between the two measures is not clear yet and calls for further study.

Estimators of $\Phi(t)$ for cohort studies have been considered in several papers. Samuelsen (2008) [26] introduced an estimator based on the Cox proportional hazards model. Laaksonen (2010) [19] considered censoring due to competing risks and proposed an estimator based on a piecewise constant hazard model. Chen (2010) [8] proposed both a non-parametric estimator of $\Phi(t)$ based on the Kaplan-Meier estimator and a semi-parametric estimator based on a transformation model of the Cox model. An estimator of $\phi(t)$ for cohort studies has been proposed by Chen et al. (2006) [9] under the Cox proportional hazards model, by plugging in the Breslow estimator for the baseline hazard function $\lambda_0(t)$ and the Nelson-Aalon estimator for the marginal hazard function $\lambda(t)$.

Besides the two measures aforementioned, several other time-varying measures have been proposed to extend the concept of PAF in specific study designs. Samuelsen (2008) [26] focused on intervention program with varying lengths of follow-up, and proposed two variants of $\phi(.)$ to measure the impact of lack of intervention. The two measures both assess the consequences of an intervention on the instantaneous risk of disease at time t, but one considers the case that the intervention takes place at time t, and the other considers the case that the intervention takes place at time zero (program start point). Cox (2009) [11] focused on the survival time attributable to the intervention, and introduced two new measures called "attributable survival" and "attributable survival time" respectively, and compares these measures with $\phi(.)$ and $\Phi(.)$.

So far the estimation of either $\Phi(t)$ or $\phi(t)$ has been only focused on cohort studies, and development of time-varying PAF for case-control studies is completely missing. Case-control studies have been the backbone in clinical and epidemiological studies because of its time and cost efficiency, and are very popular particularly for rare and chronic diseases such as cancer. Therefore, it is of great interest to develop methods for time-varying PAF for case-control studies. We find that $\phi(t)$ is a natural extension of the classic PAF to a instantaneous measure. Building upon the hazard function, $\phi(t)$ measures the instantaneous attributable risk at time t, which provides fundamental description of a time dependent process. Because of this appealing feature, it is of particular interest to develop its estimation for case-control studies.

Since the kernel estimator plays an important role in the estimation procedure of $\phi(t)$, we will briefly review this technique in the next section. Then we will review the data set to which we apply the proposed methods.

1.5 Kernel Estimation

Kernel estimation or kernel smoother refers to a class of techniques of non-parametric estimation of functions. It provides a way of describing the structure in the data set without imposing any parametric model. One of the most fundamental applications of kernel estimation to estimate the probability density function. Suppose that X_1, X_2, ..., X_n is a set of continuous random variables with a common density function f. A parametric estimation of f assumes an explicit parametric form of f, and estimates the unknown parameters in it. In contrast, a non-parametric estimation does not assume any functional form of f. The kernel density estimate of f at the point x is given by

$$\hat{f}(x;h) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right). \tag{1.15}$$

Here K is the kernel function satisfying $\int K(x)dx = 1$, and h is a positive number and usually known as the bandwidth. A more compact notation of the kernel density estimation formula is

$$\hat{f}(x;h) = \frac{1}{n}\sum_{i=1}^{n} K_h\left(x - X_i\right), \tag{1.16}$$

by introducing $K_h(u) = K(u/h)/h$. In practice K is usually chosen to be a unimodal probability density function symmetric about zero, which ensures that $\hat{f}(x;h)$ is itself also a density function.

A widely used criteria for the performance of the kernel estimator is the mean integrated squared error (MISE), which is given by

$$\text{MISE}\{\hat{f}(x;h)\} = E\left[\int \{\hat{f}(x;h) - f(x)\}^2 dx\right]. \tag{1.17}$$

Assuming that the underlying density is sufficiently smooth, the kernel has finite fourth moment, and the bandwidth satisfies $\lim_{n\to\infty} h = 0$ and $\lim_{n\to\infty} nh = \infty$, it can be shown by using Taylor series that [32]

$$\text{MISE}\{\hat{f}(x;h)\} = AMISE\{\hat{f}(x;h)\} + o\{(nh)^{-1} + h^4\}, \tag{1.18}$$

where AMISE$\{\hat{f}(x;h)\}$ is called the asymptotic mean integrated squared error, and has the expression

$$\text{AMISE}\{\hat{f}(x;h)\} = \frac{1}{nh}R(K) + \frac{h^4}{4}\mu_2(K)^2 R(f''), \tag{1.19}$$

where

$$R(K) = \int K^2(u)du, \qquad R(f'') = \int \{f''(u)\}^2 du, \qquad \mu_2(K) = \int u^2 K(u)du.$$

Therefore, the asymptotically optimal bandwidth implying minimal AMISE satisfies

$$h_{\text{AMISE}} = \left\{\frac{R(K)}{\mu_2(K)^2 R(f'')}\right\}^{1/5} n^{-1/5}, \tag{1.20}$$

and the minimal AMISE for estimation of f using kernel K is

$$\text{AMISE}_{opt}\{\hat{f}(x;h)\} = \frac{5}{4}\{R(K)^4 \mu_2(K)^2 R(f'')\}^{1/5} n^{-4/5}. \tag{1.21}$$

In (1.21), the term $R(K)^4\mu_2(K)^2$ is about the kernel function. The kernel that minimize this term is $K(x) = 0.75(1-x^2)I_{|x|<1}$, which is usually called the Epanechnikov kernel [28]. Other commonly used kernels include Gaussian, uniform, triangular, biweight, and triweight kernels. These kernels all have a very close efficiency comparing with the Epanechnikov kernel [32].

It is known that the kernel choices usually have little effect on the accuracy of the kernel estimator, whereas that the bandwidth choices have a strong impact [28]. The bandwidth can be manually selected by the data analyst by examining the fitted curves. A more sophisticated way is to use an automatic bandwidth selector. There are roughly three classes of bandwidth selectors: quick and simple methods, cross-validation methods, and plug-in methods. The principle of the quick and simple methods is based on replacing the unknown term $R(f'')$ in (1.20) with its value from a parametric family (e.g., Gaussian) and estimate the parameters in the parametric model using the data. The cross-validation methods make use of the cross-validation techniques to minimize a predicted MISE, such as in the least squares cross-validation (LSCV) [25, 6], or a predicted AMISE, such as in the biased cross-validation (BCV) [27]. The plug-in methods are based on replacing $R(f'')$ with a kernel estimator $R(\hat{f}''_g)$ [34], where g is called a pilot bandwidth . Then the remaining problem is how to choose the pilot bandwidth g. When no structure of g is imposed, one needs another 'pilot bandwidth' g_2 for choosing g, then g_3 for choosing g_2, and so on. One solution to this problem is to choose a pilot bandwidth using the quick and simple or the cross-validation methods. The direct plug-in method (DPI) chooses a pilot bandwidth some stage (e.g. g_3) using the quick and simple methods without assuming any explicit form of the bandwidth. The solve-the-equation method assumes a functional form of g depending on h, and solves the equation in which a quick and simple pilot bandwidth selector is also involved. The smoothed cross-validation method choose g by using LSCV or BCV.

1.6 Introduction of Application Data Set

We will apply our methods in analyzing the Genetics and Epidemiology of Colorectal Cancer Consortium (GECCO) data, and here we provide a brief introduction of the data set. The GECCO is a well established consortium, which is comprised of a coordinating center at the Fred Hutchinson Cancer Research Center (FHCRC) and investigators from twelve well characterized cohort and casecontrol studies conducted in North America and Europe. This consortium aims to both accelerate the discovery of colorectal cancer related variants and perform thorough epidemiologic evaluations of new susceptibility loci via gene-environmental interaction analyses. Key clinical and environmental data have been harmonized across all studies. The GECCO smoking data set used in our application consists of five population-based case-control studies in U.S. with frequency match on ages between cases and controls. The primary outcome is the case control status of colorectal cancer.

The other variables include age (age at onset for cases and age at enrollment for controls), gender, weight, height, BMI, medical history, dietary, alcohol use, smoking, exercises, cancer status, family history of colorectal cancer, and history of diabetes. The data analysis will focus on understanding the association of risk factors to the colorectal cancer risk, and quantify the PAF and time-varying PAF to the colorectal cancer risk of these risk factors.

1.7 Overview of Dissertation

The main objective of this dissertation is to derive estimators of $\phi(t)$ for population-based case-control studies. We derive the estimators based on a time axis of observed age, considering its wide use and interest in public health studies, although these estimators are also applicable to other time scales. We consider various case-control designs including simple (unmatched), frequency matched, and exact matched designs.

In Chapter 2, we derive unadjusted estimators of $\phi(t)$ based on the Cox model without considering confounders. We propose two separate point estimators by combining the β estimate from logistic model and a kernel smoother. We also propose a combined estimator of these two estimators. We develop consistency and asymptotic normality of these estimators, and study their finite sampling properties. We propose variance estimators and study their properties. We also propose a resampling-based approach for calculating the simultaneous confidence bands.

In Chapter 3, we derive adjusted estimators of $\phi(t)$ with present of confounders. We study the asymptotic and study finite sampling properties of these estimators.

In Chapter 4, we evaluate the performance of the proposed point and variance estimators as well as the proposed approach for confidence bands calculation via Monte Carlo simulation studies. We consider the situations with binary, continuous, or time-varying covariates, and various kernel and bandwidth choices.

In Chapter 5, we apply our methods to the GECCO smoking data. An analysis of the classic PAF and time-varying PAF ($\phi(t)$) for various risk factors is provided.

In Chapter 6, we summarize the dissertation with a discussion of the direction of future research.

Chapter 2

ESTIMATION OF TIME-VARYING PAF FOR CASE-CONTROL STUDIES

2.1 Overview

The attributable hazard function $\phi(t)$ proposed by Chen et al. (2006)[9] is a natural extension of PAF to an instantaneous time-varying measure of the attributable risk. In this chapter, we describe the statistical methods for estimation of $\phi(t)$ for case-control studies. To keep the problem simple, we start with the situation that no confounder is present. The generalization to include confounders is provided in Chapter 3. In the rest of this chapter, we first formulate the estimation problem for two common case-control sampling schemes, and discuss the estimation of $\phi(t)$ for case-control data. Then we propose a set of estimators for $\phi(t)$, and show the uniform consistency and asymptotic normality of these estimators, as well as the finite sampling properties. Followed by the theoretical derivation, we propose the variance estimators and the method for the pointwise confidence intervals and the simultaneous confidence bands. Finally we discuss the kernel and bandwidth choices.

2.2 Formulation of Estimation Problem

2.2.1 Estimation Problem for Case-control Data

We first formulate the statistical problem for the case-control data. Without loss of generality, we consider case-control data containing the observed age, which is age-at-onset if the subject is diseased (case) and is the current age if the subject is not diseased (control). Consider a case-control study which consists of n_1 cases and n_2 controls, in total n ($n = n_1 + n_2$) subjects. Let $\mathbf{X} = \{X, Z(.), \Delta\} \sim P_0$ represent the experiment unit and the corresponding distribution in the population of interest, where X is observed age, $Z(.)$ is a vector of possibly time-varying covariates, and Δ is the binary disease status. We will distinguish between two types of case-control sampling: unmatched case-control sampling and age-matched case-control sampling.

Unmatched Case-control Sampling: In an unmatched case-control sampling with designed case ratio of π_0, to sample a total of n subjects, one randomly samples $n_1 = \pi_0 n$ subjects from the cases population $\{X, Z(.)|\Delta = 1\}$, and randomly samples $n_2 = n - n_1$ controls from the control population $\{X, Z(.)|\Delta = 0\}$. This procedure forms a sample set with $n_1 + n_2 = n$ experimental units subject to a sampling distribution P^* with π_0 proportion of cases:

$$\mathbf{X}^* = \{X_i, Z_i(.), i = 1, ..., n\} \sim P^*,$$

In this case the n subjects consisting a case-control data set can be regarded as n independent and identically distributed observations sampled from P^*, where the marginal distributions of cases and controls are specified by P_0.

Matched Case-control Sampling: Without loss of generality, we consider the one-to-one age-matched case-control sampling ($n_1 = n_2 = n/2$). Under this sampling, one first samples a case by random sampling (X_1, Z_1) from the conditional distribution of $\{X, Z(.)|\Delta = 1\}$. Then one samples a control (X_0, Z_0) from the conditional distribution of $\{X, Z(.)|\Delta = 0, X = X_1\}$. This procedure forms an experimental unit $\mathbf{X}^*$ subject to a sampling distribution P^*:

$$\mathbf{X}^* = \{X_1, Z_1(.), \Delta_1 = 1, X_0 = X_1, Z_0(.), \Delta_0 = 1\} \sim P^*.$$

Thus the matched case-control data set with n subjects will consist of $n/2$ independent and identically distributed experimental units $\mathbf{X}_1^*, ..., \mathbf{X}_{n/2}^*$ with sampling distribution P^*. The marginal distribution of cases is specified by P_0.

In the following sections, we will first study the problem based on unmatched case-control sampling, then describe how those results can be applied to matched case-control sampling. We will use E^* and P^* to indicate expectation and probability with respect to the sampling population, and use E_0 and P_0 for expectation and probability with respect to the target population.

2.2.2 Estimability of Time-varying PAF for Case-control Studies

Recall that the time-varying PAF measure $\phi(t)$ is defined as

$$\phi(t) = 1 - \frac{\lambda_0(t)}{\lambda(t)}.$$

To facilitate the estimation of $\phi(t)$ for case-control data, we need to rewrite $\phi(t)$ into other forms. By using the relationship $\lambda(t) = f_T(t)/S_T(t)$ and applying total probability theorem on $S_T(t)$, we have

$$\begin{aligned}
\phi(t) &= 1 - \frac{\lambda_0(t)S_T(t)}{f_T(t)} \\
&= 1 - \frac{\int \lambda_0(t)S_{T|Z}(t|z)f_Z(z)dz}{f_T(t)} \\
&= 1 - \int_{\mathcal{Z}} \frac{\lambda_0(t)}{\lambda(t|z)} dF_{Z|T}(z|t), \qquad (2.1)
\end{aligned}$$

or alternatively,

$$\begin{aligned}
\phi(t) &= 1 - \frac{\lambda_0(t)S_T(t)}{f_T(t)} \\
&= 1 - \frac{\lambda_0(t)S_T(t)}{\int f_{T|Z}(t|z)f_Z(z)dz} \\
&= 1 - \frac{\lambda_0(t)S_T(t)}{\int \lambda(t|z)S_{T|Z}(t|z)f_Z(z)dz} \\
&= 1 - \left\{ \int_{\mathcal{Z}} \frac{\lambda(t|z)}{\lambda_0(t)} dF_{Z|T\geq t}(z) \right\}^{-1}. \qquad (2.2)
\end{aligned}$$

Note that these two formulas explicitly relate the time-varying PAF $\phi(t)$ to the classic Levin's PAF. Recall that the Levin's PAF can be written as

$$PAF = 1 - \int_{\mathcal{Z}} RR_z^{-1} dF_{Z|D=1}(z), \qquad (2.3)$$

or

$$PAF = 1 - \left\{ \int_{\mathcal{Z}} RR_z dF_Z(z) \right\}^{-1}. \qquad (2.4)$$

By comparing equation (2.1) with (2.3), and (2.2) with (2.4), we can see that $\phi(t)$ equals to Levin's PAF with relative risk replaced with hazard ratio, and the prevalence of exposure in cases replaced with the density of exposure for subjects who fail at time t, or the prevalence of exposure in controls replaced with the density of exposure for subjects who are at-risk at time t. Thus, $\phi(t)$ can be regarded as the instantaneous evaluation of Levin's PAF at each time point.

Now consider the Cox model with possible external time-varying covariates,

$$\lambda\{t|Z(t)\} = \lambda_0(t)\exp\{\beta^T Z(t)\}, \tag{2.5}$$

where $Z(t) = \{Z_1(t), ..., Z_p(t)\}^T$ is a $p \times 1$ vector of external time-dependent covariates, and $\beta = (\beta_1, ..., \beta_p)^T$ is a vector of regression parameters assumed to be independent of t. Under model (2.5), we can substitute $\lambda(t|z)/\lambda_0(t)$ with $\exp(\beta^T z)$ in equation (2.1) and (2.2), and have

$$\phi(t) = 1 - \int_{\mathcal{Z}} exp(-\beta^T z) dF_{Z|T}(z|t), \tag{2.6}$$

and

$$\phi(t) = 1 - \Big\{ \int_{\mathcal{Z}} exp(\beta^T z) dF_{Z|T\geq t}(z) \Big\}^{-1}. \tag{2.7}$$

Now we describe how $\phi(t)$ could be estimated on basis of these formulas. One important property of case-control data is that although they are collected retrospectively, they can be analyzed as if they were prospectively collected using a logistic model[22]. For age- matched case-control studies, consistent estimate of β can be obtained from a conditional logistic model; for frequency matched case-control studies, β can be estimated from a regular logistic model with adjustment for age frequency group, and for unmatched case-control studies, β can be estimated from a regular logistic model with adjustment for age. The estimation of β in the latter two situations are generally not consistent, but the approximation yields little bias. The remaining is to estimate the conditional distribution function $F_{Z|T}(z|t)$ or $F_{Z|T\geq t}(z)$. Under the random censoring assumption that the censoring time is independent of both the failure time and covariates[35], we can show that $F_{Z|T}(z|t)$ equals to $F(z|T = t, T \leq C) = P\{Z(t) \leq z|T = t, T \leq C\}$, which can be estimated from the cases, and $F_{Z|T\geq t}(z)$ equals to $F(z|C = t, T \geq C) = P\{Z(t) \leq z|C = t, T \geq C\}$, a quantity that can be estimated from the controls. Therefore, either $F_{Z|T}(z|t)$ or $F_{Z|T\geq t}(z)$ can be estimated by its empirical estimator. However, these empirical estimators would have poor performance because of few subjects at time t. To improve the performance, we use a kernel smoother to estimate these distribution functions. We describe the proposed estimators in details below.

2.2.3 Point Estimators of $\phi(t)$

We propose two estimators of $\phi(t)$, corresponding to the two representations of $\phi(t)$ ((2.6) and (2.7)). The first estimator, corresponding to formula (2.6), is

$$\widehat{\phi}_+(t;\hat{\beta}) = 1 - \frac{\sum_{i=1}^n \exp\{-\hat{\beta}^T Z_i(t)\}\Delta_i K_h(t-X_i)}{\sum_{i=1}^n \Delta_i K_h(t-X_i)}, \tag{2.8}$$

where $K_h(x) = K(x/h)/h$, $K(.)$ is a kernel function, and h is the bandwidth that controls the spread of weighting window. $\hat{\beta}$ is the estimate of β obtained from the logistic regression with adjustment for age (for unmatched case-control studies). The following results are developed under unmatched case-control studies. We show results for matched case-control studies at the end of this chapter.

The second estimator, corresponding to formula (2.7), is

$$\widehat{\phi}_-(t;\hat{\beta}) = 1 - \frac{\sum_{i=1}^n (1-\Delta_i) K_h(t-X_i)}{\sum_{i=1}^n \exp\{\hat{\beta}^T Z_i(t)\}(1-\Delta_i) K_h(t-X_i)}. \tag{2.9}$$

Since the kernel smoother of the two estimators are based on cases and controls separately, combining the two estimators could potentially improve the efficiency. We therefore propose a weighted estimator as follows:

$$\widehat{\phi}_w(t;\hat{\beta}) = w(t)\widehat{\phi}_+(t;\hat{\beta}) + \{1-w(t)\}\widehat{\phi}_-(t;\hat{\beta}),$$

where $w(t)$ is a weighting function with value between 0 and 1. A natural choice of the weight is to let $w(t)$ equal to the proportion of cases in the sample. For example, in a $1:1$ case-control study, π_0 would be 0.5. Let $\pi_0 = P^*(T \le C)$ be the the proportion of cases in the sample, then the weighted estimator is

$$\widehat{\phi}_w(t;\hat{\beta}) = \pi_0\widehat{\phi}_+(t;\hat{\beta}) + (1-\pi_0)\widehat{\phi}_-(t;\hat{\beta}). \tag{2.10}$$

2.3 Large Sample Properties

In this section we will show the strong uniform consistency and pointwise asymptotic normality of $\widehat{\phi}_+(t;\hat{\beta})$, $\widehat{\phi}_-(t;\hat{\beta})$, and $\widehat{\phi}_w(t;\hat{\beta})$. We first list the assumptions and conditions required for consistency and asymptotic normality.

Assumptions.

A1. Random censoring: the censoring time C is independent of both the failure time T and covariates $Z(.)$.

A2. The time t is in a range of $(0, \tau)$ for a constant $\tau > 0$ such that the density of failure time $f_T(t)$, the density of censoring time $f_C(t)$ and their survival functions, $S_T(t)$ and $S_C(t)$ all take positive real values on $[0, \tau]$.

A3. The density of failure time $f_T(t)$ and the density of censoring time $f_C(t)$ are both continuous, uniformly bounded, and have second derivatives on $[0, \tau]$.

A4. The bandwidth satisfies $h = n^d h_0$ for constants $-1/2 < d < -1/5$ and $h_0 > 0$.

A5. The kernel function K(.) has bounded variation and satisfies the following conditions,

$$\int_{-\infty}^{\infty} K(u)du = 1, \qquad \int_{-\infty}^{\infty} K^2(u)du < \infty,$$
$$\int_{-\infty}^{\infty} uK(u)du = 0, \qquad \int_{-\infty}^{\infty} u^2K(u)du < \infty.$$

A6. $Z(.)$ is bounded almost surely and has uniformly bounded total variation on $[0, \tau]$.

A7. The number of cases n_1 and total number of subjects n satisfies $n_1/n \to \pi_0$ as $n \to \infty$, where $0 < \pi_0 < 1$.

We establish the asymptotic properties of $\widehat{\phi}_+(t;\hat{\beta})$ in the next two subsections, followed by the asymptotic properties of $\widehat{\phi}_-(t;\hat{\beta})$.

2.3.1 Uniform Consistency of $\widehat{\phi}_+(t;\hat\beta)$

We first introduce the following notation that will be used in the proof.

$$\begin{aligned}
p_0 &= P_0(T \le C), \pi_0 = P^*(T \le C),\\
A_n(t;\beta) &= \frac{1}{n}\sum_{i=1}^n \Delta_i K_h(t-X_i)\{1-e^{-\beta^T Z_i(t)}\},\\
B_n(t) &= \frac{1}{n}\sum_{i=1}^n \Delta_i K_h(t-X_i),\\
C_n(t) &= \frac{1}{n}\sum_{i=1}^n \Delta_i K_h(t-X_i)\{e^{-\beta_0^T Z_i(t)} - e^{-\hat\beta^T Z_i(t)}\},\\
D_n(t;\beta) &= \frac{1}{n}\sum_{i=1}^n \Delta_i K_h(t-X_i) Z_i(t) e^{-\beta^T Z_i(t)}.
\end{aligned}$$

The limits of $A_n(t;\beta)$, $B_n(t)$, and $D_n(t;\beta)$ are denoted by

$$\begin{aligned}
A(t) &= \phi(t) f_T(t) S_C(t)\pi_0/p_0,\\
B(t) &= f_T(t) S_C(t)\pi_0/p_0,\\
D(t;\beta) &= \Big\{\int_{\mathcal{Z}} z e^{-\beta^T z} dF_{Z|T}(z|t)\Big\} f_T(t) S_C(t)\pi_0/p_0,
\end{aligned}$$

respectively.

We establish the strong uniform consistency of $\widehat{\phi}_+(t;\hat\beta)$ in the following theorem.

Theorem 1. Suppose assumptions A1-A7 are satisfied. Then,

$$\sup_{t\in(0,\tau)} |\widehat{\phi}_+(t;\hat\beta) - \phi(t)| \to 0 \text{ a.s.}$$

as $n \to \infty$.

Theorem 1 is proved with the help of the following lemmas.

Lemma 1.1. Suppose assumptions A1-A7 are satisfied. Then,

$$\sup_{t\in(0,\tau)} |B_n(t) - B(t)| \to 0 \text{ a.s.}$$

as $n \to \infty$.

Proof. This lemma is proved with help of Lemma 1 and Lemma 2 in Nadaraya (1970), which are stated below.

Settings in Nadaraya (1970). Let (X_i, Y_i), $i = 1, ..., n$ be iid two-dimensional random variables with density function $f(x, y)$. Let the marginal density of X be $f(x)$, $y(x) = \int_{-\infty}^{\infty} y f(y|x) dy$ be the regression curve of Y with respect to X, and $\varphi(x) = f(x)y(x)$. Assume $f(x)$ and $y(x)$ are continuous over the real line. Let

$$y_n(x) = \frac{\sum_{i=1}^n Y_i K_h(x - X_i)}{\sum_{i=1}^n K_h(x - X_i)},$$

and

$$\varphi_n(x) = \frac{1}{n}\sum_{i=1}^n Y_i K_h(x - X_i), \qquad f_n(x) = \frac{1}{n}\sum_{i=1}^n K_h(x - X_i),$$

where $K_h(x) = K(x/h)/h$, and $K(.)$ is a kernel function satisfying assumption A3, and h is the bandwidth satisfying $h \to 0$ with increasing n.

Lemma 1 (Nadaraya 1970). $\sup_{a \le x \le b} |E\varphi_n(x) - \varphi(x)| \to 0$ with increasing n. a and b are real numbers satisfying $a \le b$, and $E\varphi_n(x)$ can be written as

$$E\varphi_n(x) = \frac{1}{h}\int_{-\infty}^{\infty} K\left(\frac{x - t}{h}\right)\varphi(t)dt.$$

Lemma 2 (Nadaraya 1970). If $K(x)$ is a function of bounded variation and Y is bounded almost surely, then for any $\lambda > 0$

$$P\{\sup_{a \le x \le b} |\varphi_n(x) - E\varphi_n(x)| > \lambda\} \le c_0 \exp(-\alpha_1 n h^2) + c_1 \exp(-\alpha_2 n h^2)$$

for some constants α_1, α_2, c_0, and c_1.

Now we begin the proof of Lemma 1. To simplify the notation, unless otherwise specified, the integration with respect to x is from 0 to ∞, and the integration with respect to z is on its support $\mathcal{Z}$ through out the proofs in this chapter. First note that

$$E^*\{B_n(t)\} = E^*\{\Delta K_h(t - X)\} = E^*\{K_h(t - X)|\Delta = 1\}P^*(\Delta = 1).$$

Since the cases are randomly sampled from the cases in the population of interest, the sampling distribution for $\{X, Z(.)|\Delta = 1\}$ is the same as the conditional distribution in the population of interest. Thus we have $E^*\{K_h(t-X)|\Delta=1\} = E_0\{K_h(t-X)|\Delta=1\}$, and so that

$$\begin{aligned} E^*\{B_n(t)\} &= E_0\{K_h(t-X)|\Delta=1\}P^*(\Delta=1) \\ &= E_0\{K_h(t-T)|T\le C\}\pi_0 \\ &= \frac{\pi_0}{h}\int K\left(\frac{t-x}{h}\right) f_{T|T\le C}(x)dx \\ &= \frac{\pi_0}{h}\int K\left(\frac{t-x}{h}\right) \frac{f_T(x)S_C(x)}{P_0(T\le C)}dx \\ &= \frac{1}{h}\int K\left(\frac{t-x}{h}\right) B(x)dx. \end{aligned} \tag{2.11}$$

Since $B(.)$ is continuous given that $f_T(.)$ and $S_C(.)$ are both continuous, by Lemma 1 in Nadaraya (1970), we have

$$\sup_{t\in(0,\tau)} |E^*B_n(t) - B(t)| \to 0 \text{ with increasing n.} \tag{2.12}$$

It remains to check the convergence of $\sup_{t\in(0,\tau)} |B_n(t) - E^*B_n(t)|$. Since $K(.)$ is of bounded variation, and Δ is bounded almost surely, by Lemma 2 in Nadaraya (1970) we have, for any $\lambda > 0$,

$$P^*\{\sup_{t\in(0,\tau)} |B_n(t) - E^*B_n(t)| > \lambda\} \le c_1\exp(-\alpha_1 nh^2) + c_2\exp(-\alpha_2 nh^2) \tag{2.13}$$

for positive finite constants c_1, c_2, α_1, and α_2. Moreover, the series $\sum_{n=1}^{\infty}\exp(-\gamma nh^2)$ converges for any $\gamma > 0$ under condition A4. To see this, let m be any integer larger than $1/(1+2d)$, then it is not hard to verify the following relationship,

$$\begin{aligned} \int_1^\infty \exp(-\gamma h_0^2 x^{1+2d})dx &< \int_1^\infty \exp(-\gamma h_0^2 x^{1/m})dx \\ &= \frac{m}{(\gamma h_0^2)^m}\int_{\gamma h_0^2}^\infty y^{m-1}\exp(-y)dy \\ &= \frac{m}{(\gamma h_0^2)^m}\Big\{ -\exp(-y)\sum_{i=0}^{m-1}\frac{(m-1)!}{(m-1-i)!}y^{m-1-i}\Big|_{y=\gamma h_0^2}^{\infty}\Big\} < \infty, \end{aligned}$$

then $\sum_{n=1}^{\infty}\exp(-\gamma nh^2)$ converges by the integral test. Hence it follows by the fact

$$\sum_{n=1}^{\infty} P^*\{\sup_{t\in(0,\tau)} |B_n(t) - E^*B_n(t)| > \lambda\} \le \sum_{n=1}^{\infty} c_1\exp(-\alpha_1 nh^2) + \sum_{n=1}^{\infty} c_2\exp(-\alpha_2 nh^2) < \infty,$$

and the Borel-Cantelli lemma that there exists $N < \infty$ such that

$$P^*\{\sup_{t\in(0,\tau)} |B_n(t) - E^*B_n(t)| > \lambda\} = 0$$

for all $n > N$, which is equivalent to say that

$$\sup_{t\in(0,\tau)} |B_n(t) - E^*B_n(t)| \to 0 \text{ a.s.} \tag{2.14}$$

as $n \to \infty$. On the basis of (2.12) and (2.14) and the continuous mapping theorem, we can conclude that

$$\begin{aligned}\sup_{t\in(0,\tau)} |B_n(t) - B(t)| &= \sup_{t\in(0,\tau)} |\{B_n(t) - E^*B_n(t)\} + \{E^*B_n(t) - B(t)\}| \\ &\le \sup_{t\in(0,\tau)} |B_n(t) - E^*B_n(t)| + \sup_{t\in(0,\tau)} |E^*B_n(t) - B(t)| \\ &\to 0 \text{ a.s.}\end{aligned}$$

as $n \to 0$. Lemma 1.1 is proved.

Lemma 1.2. Suppose assumptions A1-A7 are satisfied. Then,

$$\sup_{t\in(0,\tau)} \left|\widehat{\phi}_+(t;\beta_0) - \phi(t)\right| \to 0 \text{ a.s.}$$

as $n \to \infty$.

Proof. First we derive the expectation of $A_n(t;\beta_0)$:

$$\begin{aligned}E^*\{A_n(t;\beta_0)\} &= E^*\left[\Delta K_h(t-X)\{1-e^{-\beta_0^T Z(t)}\}\right] \\ &= E^*\left[K_h(t-T)\{1-e^{-\beta_0^T Z(t)}\}|T \le C\right] P^*(T \le C) \\ &= \frac{\pi_0}{h}\iint \{1-\exp(-\beta_0^T z)\}K\left(\frac{t-x}{h}\right) f_{T,Z|T\le C}(x,z)dxdz \\ &= \frac{\pi_0}{h}\iint \{1-\exp(-\beta_0^T z)\}K\left(\frac{t-x}{h}\right) \frac{f_{T|Z}(x|z)S_C(x)f_Z(z)}{P(T\le C)}dxdz \\ &= \frac{1}{h}\frac{\pi_0}{p_0}\int K\left(\frac{t-x}{h}\right) S_C(x)\int \{1-\exp(-\beta_0^T z)\}f_{T|Z}(x|z)f_Z(z)dzdx \end{aligned} \tag{2.15}$$

From (2.6) it is straightforward to get the relationship

$$\int \{1-\exp(-\beta_0^T z)\}f_{T|Z}(x|z)f_Z(z)dz = \phi(x)f_T(x).$$

Substitute it into (2.15) we have

$$E^*\{A_n(t;\beta_0)\} = \frac{1}{h}\frac{\pi_0}{p_0}\int K\left(\frac{t-x}{h}\right) S_C(x)\phi(x)f_T(x)dx$$
$$= \frac{1}{h}\int K\left(\frac{t-x}{h}\right) A(x)dx. \tag{2.16}$$

Then using the same arguments in Lemma 1.1, we can conclude that

$$\sup_{t\in(0,\tau)} |A_n(t;\beta_0) - A(t)| \to 0 \text{ a.s.} \tag{2.17}$$

as $n \to 0$, given $K(.)$ has bounded variation, $Z(.)$ is bounded almost surely, and $\sum_{n=1}^{\infty}\exp(-\gamma n h^2)$ converges for any $\gamma > 0$.

Let $\inf_{t\in(0,\tau)}|B(t)| = \mu > 0$. Then for any $0 < \lambda < \mu$, we have the following relationship:

$$P^*\{\sup_{t\in(0,\tau)}\left|\widehat{\phi}_+(t;\beta_0) - \phi(t)\right| > \lambda\} = P^*\{\sup_{t\in(0,\tau)}\left|\frac{A_n(t;\beta_0)}{B_n(t)} - \phi(t)\right| > \lambda\}$$
$$\leq P^*\{\frac{\sup_{t\in(0,\tau)}|A_n(t;\beta_0) - \phi(t)B_n(t)|}{\inf_{t\in(0,\tau)}|B_n(t)|} > \lambda\}$$
$$= P^*\{\sup_{t\in(0,\tau)}|A_n(t;\beta_0) - \phi(t)B_n(t)| \geq \lambda \inf_{t\in(0,\tau)}|B_n(t)|, \inf_{t\in(0,\tau)}|B_n(t)| \geq \mu - \lambda\}$$
$$+ P^*\{\sup_{t\in(0,\tau)}|A_n(t;\beta_0) - \phi(t)B_n(t)| \geq \lambda \inf_{t\in(0,\tau)}|B_n(t)|, \inf_{t\in(0,\tau)}|B_n(t)| < \mu - \lambda\}$$
$$\leq P^*\{\sup_{t\in(0,\tau)}|A_n(t;\beta_0) - \phi(t)B_n(t)| \geq \lambda(\mu - \lambda)\} + P^*\{\inf_{t\in(0,\tau)}|B_n(t)| < \mu - \lambda\}.$$

Since

$$P^*\{\inf_{t\in(0,\tau)}|B_n(t)| < \mu - \lambda\} = P^*\{\inf_{t\in(0,\tau)}|B(t)| - \inf_{t\in(0,\tau)}|B_n(t)| > \lambda\} \tag{2.18}$$
$$\leq P^*\{\sup_{t\in(0,\tau)}|B_n(t) - B(t)| > \lambda\},$$

we have

$$
\begin{aligned}
&P^*\{\sup_{t\in(0,\tau)}\left|\widehat{\phi}_+(t;\beta_0)-\phi(t)\right|>\lambda\} && (2.19)\\
\leq &P^*\{\sup_{t\in(0,\tau)}|A_n(t;\beta_0)-\phi(t)B_n(t)|\geq\lambda(\mu-\lambda)\}+P^*\{\sup_{t\in(0,\tau)}|B_n(t)-B(t)|>\lambda\}\\
\leq &P^*\{\sup_{t\in(0,\tau)}|A_n(t;\beta_0)-\phi(t)B(t)|+\sup_{t\in(0,\tau)}|\phi(t)B_n(t)-\phi(t)B(t)|\geq\lambda(\mu-\lambda)\}\\
&+P^*\{\sup_{t\in(0,\tau)}|B_n(t)-B(t)|>\lambda\}\\
\leq &P^*\{\sup_{t\in(0,\tau)}|A_n(t;\beta_0)-A(t)|\geq\lambda(\mu-\lambda)\}\\
&+P^*\{\sup_{t\in(0,\tau)}|B_n(t)-B(t)|\geq\lambda(\mu-\lambda)/d\}\\
&+P^*\{\sup_{t\in(0,\tau)}|B_n(t)-B(t)|>\lambda\}. && (2.20)
\end{aligned}
$$

Using (2.14) and (2.17) combined with 2.19, we can conclude that

$$\sup_{t\in(0,\tau)}\left|\widehat{\phi}_+(t;\beta_0)-\phi(t)\right|\to 0 \text{ a.s.} \tag{2.21}$$

as $n\to\infty$. Lemma 1.2 is proved.

Lemma 1.3. Suppose assumptions A1-A7 are satisfied. Then,

$$\sup_{t\in(0,\tau)}\left|\widehat{\phi}_+(t;\hat{\beta})-\widehat{\phi}_+(t;\beta_0)\right|\to 0 \text{ a.s.}$$

as $n\to\infty$.

Proof. Note that

$$\widehat{\phi}_+(t;\hat{\beta})-\widehat{\phi}_+(t;\beta_0)=\frac{C_n(t)}{B_n(t)}.$$

We first study the convergence of $\sup_{t\in(0,\tau)}|C_n(t)|$. By Taylor's theorem we can write $C_n(t)$ as

$$\begin{aligned}
C_n(t) &= \frac{1}{n}\sum_{i=1}^{n}\left[\exp\{-\beta_0^T Z_i(t)\} - \exp\{-\hat{\beta}^T Z_i(t)\}\right]\Delta_i K_h(t-X_i)\\
&= \frac{1}{n}\sum_{i=1}^{n}(\hat{\beta}-\beta_0)^T Z_i(t)\exp\{-\beta_0^T Z_i(t)\}\Delta_i K_h(t-X_i) + o_p(1)\\
&= (\hat{\beta}-\beta_0)^T\left[\frac{1}{n}\sum_{i=1}^{n} Z_i(t)\exp\{-\beta_0^T Z_i(t)\}\Delta_i K_h(t-X_i)\right] + o_p(1)\\
&= (\hat{\beta}-\beta_0)^T D_n(t;\beta_0) + o_p(1),
\end{aligned}$$

It is easy to see that for any constant β,

$$\begin{aligned}
E^*\{D_n(t;\beta)\} &= E^*\left[Z_i(t)\exp\{-\beta^T Z_i(t)\}\Delta_i K_h(t_0-X_i)\}\right]\\
&= E^*\left[Z_i(t)\exp\{-\beta^T Z_i(t)\}K_h(t_0-X_i)|T\le C\right]P^*(T\le C)\\
&= \frac{\pi_0}{h}\iint\{z\exp(-\beta^T z)\}K\left(\frac{t-x}{h}\right)f_{T,Z|T\le C}(x,z)dxdz\\
&= \frac{\pi_0}{h}\iint\{z\exp(-\beta^T z)\}K\left(\frac{t-x}{h}\right)\frac{f_{T|Z}(x|z)S_C(x)f_Z(z)}{P(T\le C)}dxdz\\
&= \frac{1}{h}\frac{\pi_0}{p_0}\int K\left(\frac{t-x}{h}\right)\int z\exp(-\beta^T z)f_{Z|T}(z|x)dzf_T(x)S_C(x)dx\\
&= \frac{1}{h}\int K\left(\frac{t-x}{h}\right)D(x;\beta)dx
\end{aligned}\tag{2.22}$$

using the same arguments in Lemma 1.1, we can conclude that

$$\sup_{t\in(0,\tau)}|D_n(t;\beta)-D(t;\beta)|\to 0 \text{ a.s.}\tag{2.23}$$

as $n\to 0$, given that $D(.)$ is continuous, $K(.)$ has bounded variation, $Z(t)$ is bounded almost surely, and $\sum_{n=1}^{\infty}\exp(-\gamma nh^2)$ converges for any $\gamma>0$. So that $D_n(t;\beta_0)$ is bounded almost surely. By the strong consistency of $\hat{\beta}$ and continuous mapping theorem, we can conclude that

$$\sup_{t\in(0,\tau)}|C_n(t)| = |\hat{\beta}-\beta_0|^T\{\sup_{t\in(0,\tau)}|D_n(t;\beta_0)|\} + o_p(1)\to 0 \text{ a.s.}\tag{2.24}$$

as $n \to 0$. Use similar arguments in (2.19) we can get the following relationship

$$\begin{aligned} P\{\sup_{t\in(0,\tau)} \left|\frac{C_n(t)}{B_n(t)}\right| > \lambda\} \leq & P^*\{\sup_{t\in(0,\tau)} |C_n(t;\beta_0)| \geq \lambda(\mu-\lambda)\} \\ & + P^*\{\sup_{t\in(0,\tau)} |B_n(t)-B(t)| \geq \lambda(\mu-\lambda)/d\} \\ & + P^*\{\sup_{t\in(0,\tau)} |B_n(t)-B(t)| > \lambda\}, \end{aligned}$$

which in conjunction with the Borel-Cantelli lemma will imply that

$$\sup_{t\in(0,\tau)} \left|\frac{C_n(t)}{B_n(t)}\right| \to 0 \text{ a.s.} \tag{2.25}$$

Lemma 1.3 is proved.

Now we can prove Theorem 1 with help of the lemma.

Proof of Theorem 1. Using the triangle inequality, we can easily obtain

$$\begin{aligned} \sup_{t\in(0,\tau)} \left|\widehat{\phi}_+(t;\hat{\beta}) - \phi(t)\right| &= \sup_{t\in(0,\tau)} \left|\left\{\widehat{\phi}_+(t;\hat{\beta}) - \widehat{\phi}_+(t;\beta_0)\right\} + \left\{\widehat{\phi}_+(t;\beta_0) - \phi(t)\right\}\right| \\ &\leq \sup_{t\in(0,\tau)} \left|\widehat{\phi}_+(t;\hat{\beta}) - \widehat{\phi}_+(t;\beta_0)\right| + \sup_{t\in(0,\tau)} \left|\widehat{\phi}_+(t;\beta_0) - \phi(t)\right|. \end{aligned}$$

Then use Lemma 1.2 and Lemma 1.3 we can conclude that

$$\sup_{t\in(0,\tau)} |\widehat{\phi}_+(t;\hat{\beta}) - \phi(t)| \to 0 \text{ a.s.}$$

as $n \to \infty$. Theorem 1 is proved.

2.3.2 Asymptotic Normality of $\widehat{\phi}_+(t;\hat{\beta})$

The pointwise asymptotic normality of $\widehat{\phi}_+(t;\hat{\beta})$ is stated in the following theorem.

Theorem 2. Suppose assumptions A1-A7 are satisfied. Then,

$$S_n(t_0) = \sqrt{nh}\{\widehat{\phi}_+(t_0;\hat{\beta}) - \phi(t_0)\} \to_d N(0,\sigma^2(t_0))$$

as $n \to \infty$ for any $t_0 \in (0,\tau)$.

Proof. For simplicity, we will use Z to denote $Z(t_0)$ in the following proof.

Write $S_n(t_0)$ into two parts,

$$S_n(t_0) = \sqrt{nh}\{\widehat{\phi}_+(t_0;\hat{\beta}) - \phi(t_0;\beta_0)\} = \sqrt{nh}\{\widehat{\phi}_+(t_0;\hat{\beta}) - \widehat{\phi}_+(t_0;\beta_0)\} + \sqrt{nh}\{\widehat{\phi}_+(t_0;\beta_0) - \phi(t_0;\beta_0)\}$$

We first study the second part.

$$\begin{aligned}\sqrt{nh}\{\widehat{\phi}_+(t_0;\beta_0) - \phi(t_0;\beta_0)\} &= \frac{\sqrt{nh}\sum_{i=1}^n \{1-\phi(t_0;\beta_0) - \exp(-\beta_0^T Z_i)\}\Delta_i K_h(t - X_i)}{\sum_{i=1}^n \Delta_i K_h(t_0 - X_i)} \\ &= \frac{\sqrt{nh}\{A_n(t_0;\beta_0) - \phi(t_0)B_n(t_0)\}}{B_n(t_0)}.\end{aligned}$$

Given the uniform consistency of $B_n(t)$, it suffices to study the asymptotic normality of $\sqrt{nh}\{A_n(t_0;\beta_0) - \phi(t_0)B_n(t_0)\}$ and apply Slutsky's theorem.

Denote

$$a_{n,i}(t_0) = \sqrt{\frac{h}{n}}\{1-\phi(t_0;\beta_0) - \exp(-\beta_0^T Z_i)\}\Delta_i K_h(t_0 - X_i),$$

so that $\sqrt{nh}\{A_n(t_0;\beta_0) - \phi(t_0)B_n(t_0)\} = \sum_{i=1}^n a_{n,i}(t_0)$. Then it suffices to show that $a_{n,i}(t_0)$ satisfies the Lyapunov's condition. We will set forth the condition in the following steps. First note that

$$\begin{aligned}
&E^*\{a_{n,i}(t_0)\} \\
=&\sqrt{\frac{h}{n}}E^*\left[\{1-\phi(t_0;\beta_0) - \exp(-\beta_0^T Z)\}\Delta K_h(t_0 - X)\right] \\
=&\pi_0\sqrt{\frac{h}{n}}E_0\left[\{1-\phi(t_0;\beta_0) - \exp(-\beta_0^T Z)\}K_h(t_0 - X)\middle| T \le C\right] \\
=&\frac{\pi_0}{\sqrt{nh}}\iint \{1-\phi(t_0;\beta_0) - \exp(-\beta_0^T z)\}K\left(\frac{t_0 - x}{h}\right) f_{T,Z|T\le C}(x,z)dxdz \\
=&\frac{\pi_0}{\sqrt{nh}}\iint \{1-\phi(t_0;\beta_0) - \exp(-\beta_0^T z)\}K\left(\frac{t_0 - x}{h}\right) \frac{f_{T|Z}(x|z)S_C(x)f_Z(z)}{P(T\le C)}dxdz \\
&\text{by the random censoring assumption,} \\
=&\frac{\pi_0}{p_0}\sqrt{\frac{h}{n}}\iint \{1-\phi(t_0;\beta_0) - \exp(-\beta_0^T z)\}K(u)f_{T|Z}(t_0 - uh|z)S_C(t_0 - uh)f_Z(z)dudz \\
&\text{ by letting } u = \frac{t_0 - x}{h}, \\
=&\frac{\pi_0}{p_0}\sqrt{\frac{h}{n}}\left[\int \{1-\phi(t_0;\beta_0) - \exp(-\beta_0^T z)\}f_{T|Z}(t_0|z)S_C(t_0)f_Z(z)dz + h^2 r(t_0) + O(h^3)\right] \\
&\text{by the Taylor's expansion,}
\end{aligned}$$

where

$$r(t_0) = \int u^2 K(u)du \int \{1-\phi(t_0;\beta_0) - \exp(-\beta_0^T z)\}\{f'_{T|Z}(t_0|z)S'_C(t_0) + \frac{1}{2}f_{T|Z}(t_0|z)S''_C(t_0) + \frac{1}{2}f''_{T|Z}(t_0|z)S_C(t_0)\}f_Z(z)dz.$$

The first part in the brackets equals zero by noting that

$$\begin{aligned}
&\int \{1-\phi(t_0;\beta_0) - \exp(-\beta_0^T z)\}f_{T|Z}(t_0|z)S_C(t_0)f_Z(z)dz \\
=&S_C(t_0)\{1-\phi(t_0;\beta_0)\}\int f_{T|Z}(t_0|z)f_Z(z)dz - S_C(t_0)\int \exp(-\beta_0^T z)f_{T|Z}(t_0|z)f_Z(z)dz \\
=&S_C(t_0)\{1-\phi(t_0;\beta_0)\}f_T(t_0) - S_C(t_0)\int \exp(-\beta_0^T z)f_{T|Z}(t_0|z)f_Z(z)dz \\
=&S_C(t_0)\frac{\int \exp(-\beta_0^T z)f_{T|Z}(t_0|z)f_Z(z)dz}{f_T(t_0)}f_T(t_0) - S_C(t_0)\int \exp(-\beta_0^T z)f_{T|Z}(t_0|z)f_Z(z)dz \\
=&0
\end{aligned}$$

Hence $E^*\{a_{n,i}(t_0)\} = \sqrt{h^5/n}r(t_0)\pi_0/p_0 + O(\sqrt{h^7}/n) = O(\sqrt{h^5/n})$.

Similarly, we can show that

$$\begin{aligned}
&E^*\{a_{n,i}^2(t_0)\} \\
=&\frac{h}{n}E^*\left[\{1-\phi(t_0;\beta_0) - \exp(-\beta_0^T Z)\}^2 \Delta K_h^2(t_0 - X)\right] \\
=&\frac{\pi_0 h}{n}E_0\left[\{1-\phi(t_0;\beta_0) - \exp(-\beta_0^T Z)\}^2 K_h^2(t_0 - X)\middle| T \le C\right] \\
=&\frac{\pi_0}{nh}\iint \{1-\phi(t_0;\beta_0) - \exp(-\beta_0^T z)\}^2 K^2\left(\frac{t_0 - x}{h}\right) f_{T,Z|T\le C}(x,z)dxdz \\
=&\frac{\pi_0}{nh}\iint \{1-\phi(t_0;\beta_0) - \exp(-\beta_0^T z)\}^2 K^2\left(\frac{t_0 - x}{h}\right) \frac{f_{T|Z}(x|z)S_C(x)f_Z(z)}{P(T\le C)}dxdz \\
&\text{by the random censoring assumption,} \\
=&\frac{\pi_0}{np_0}\iint \{1-\phi(t_0;\beta_0) - \exp(-\beta_0^T z)\}^2 K^2(u) f_{T|Z}(t_0 - uh|z)S_C(t_0 - uh)f_Z(z)dudz \\
&\text{ by letting } u = \frac{t_0 - x}{h}, \\
=&\frac{\pi_0}{np_0}\left[\int \{1-\phi(t_0;\beta_0) - \exp(-\beta_0^T z)\}^2 f_{T|Z}(t_0|z)S_C(t_0)f_Z(z)dz \int K^2(u)du + O(h)\right] \\
&\text{by the Taylor's expansion,}
\end{aligned}$$

and

$$
\begin{aligned}
&E^*\{a_{n,i}^3(t_0)\}\\
=&(\frac{h}{n})^{3/2}E^*\left[\{1-\phi(t_0;\beta_0)-\exp(-\beta_0^T Z)\}^3 \Delta K_h^3(t_0-X)\right]\\
=&\pi_0(\frac{h}{n})^{3/2}E_0\left[\{1-\phi(t_0;\beta_0)-\exp(-\beta_0^T Z)\}^3 K_h^3(t_0-X)\middle| T\le C\right]\\
=&\pi_0(\frac{1}{nh})^{3/2}\iint\{1-\phi(t_0;\beta_0)-\exp(-\beta_0^T z)\}^3 K^3\left(\frac{t_0-x}{h}\right) f_{T,Z|T\le C}(x,z)dxdz\\
=&\pi_0(\frac{1}{nh})^{3/2}\iint\{1-\phi(t_0;\beta_0)-\exp(-\beta_0^T z)\}^3 K^3\left(\frac{t_0-x}{h}\right) \frac{f_{T|Z}(x|z)S_C(x)f_Z(z)}{P(T\le C)}dxdz\\
&\text{by the random censoring assumption,}\\
=&\frac{\pi_0}{p_0 n\sqrt{nh}}\iint\{1-\phi(t_0;\beta_0)-\exp(-\beta_0^T z)\}^3 K^3(u) f_{T|Z}(t_0-uh|z)S_C(t_0-uh)f_Z(z)dudz\\
&\text{by letting } u=\frac{t_0-x}{h},\\
=&\frac{\pi_0}{p_0 n\sqrt{nh}}\left[\int\{1-\phi(t_0;\beta_0)-\exp(-\beta_0^T z)\}^3 f_{T|Z}(t_0|z)S_C(t_0)f_Z(z)dz\int K^3(u)du+O(h)\right]\\
&\text{by the Taylor's expansion,}
\end{aligned}
$$

Hence we have,

$$
\begin{aligned}
&\sum_{i=1}^n E^*\{a_{n,i}(t_0)-E^*a_{n,i}(t_0)\}^3\\
=&nE^*\{a_{n,i}(t_0)^3\}-3nE^*\{a_{n,i}(t_0)^2\}E^*\{a_{n,i}(t_0)\}+2n\left[E^*\{a_{n,i}(t_0)\}\right]^3\\
=&O(1/\sqrt{nh})+nO(1/n)\cdot O(\sqrt{h/n})+nO((h/n)^{3/2})\\
=&O(1/\sqrt{nh}),
\end{aligned}
$$

and

$$
\begin{aligned}
&\sum_{i=1}^n Var^*\{a_{n,i}(t_0)\}=nE^*\{a_{n,i}^2(t_0)\}-n[E^*\{a_{n,i}(t_0)\}]^2\\
=&\frac{\pi_0}{p_0}\int\{1-\phi(t_0;\beta_0)-\exp(-\beta_0^T z)\}^2 f_{T|Z}(t_0|z)S_C(t_0)f_Z(z)dz\int K^2(u)du+O(h^2).
\end{aligned}
$$

Denote the limiting variance as

$$
\sigma_A^2(t)=\frac{\pi_0}{p_0}\int\{1-\phi(t;\beta_0)-\exp(-\beta_0^T z)\}^2 f_{T|Z}(t|z)S_C(t)f_Z(z)dz\int K^2(u)du,
$$

then

$$\sum_{i=1}^{n} Var^*\{a_{n,i}(t_0)\} = \sigma_A^2(t_0) + O(h^2),$$

The Lyapunov's condition satisfies

$$\frac{\sum_{i=1}^{n} E\{a_{n,i}(t) - Ea_{n,i}(t)\}^3}{[\sum_{i=1}^{n} Var\{a_{n,i}(t)\}]^{3/2}} \to 0$$

as $\sqrt{nh} \to \infty$. By Lyapunov's Central Limit Theorem, we have

$$\frac{\sqrt{nh}\{A_n(t_0;\beta_0) - \phi(t_0)B_n(t_0)\} - \sum_{i=1}^{n} E^*\{a_{n,i}(t_0)\}}{[\sum_{i=1}^{n} Var^*\{a_{n,i}(t_0)\}]^{1/2}} \to_d N(0,1),$$

and so that

$$\sqrt{nh}\{A_n(t_0;\beta_0) - \phi(t_0)B_n(t_0)\} \to_d N(0, \sigma_A^2(t_0))$$

by Slutsky's theorem and the facts that

$$\sum_{i=1}^{n} E^*\{a_{n,i}(t_0)\} = nE^*\{a_{n,i}(t_0)\} = \sqrt{nh^5}\pi_0 r(t_0)/p_0 \to 0,$$
$$\sum_{i=1}^{n} Var^*\{a_{n,i}(t_0)\} \to \sigma_A^2(t_0),$$

Remark. A usual choice of bandwidth is to let $h = h_0 n^{-1/5}$ which minimizes the asymptotic integrated mean square error. This bandwidth choice would lead to asymptotic bias for $\sqrt{nh}\{A_n(t_0;\beta_0) - \phi(t_0)B_n(t_0)\}$ (thus for $S_n(t_0)$), which in our case is $\sqrt{h_0}r(t)\pi_0/p_0$. We choose to use a faster converging bandwidth i.e. $h = h_0 n^d$ with $d < -1/5$ so that the bias vanishes as n gets large, although this would compromise some efficiency.

Then we can conclude from the above results that,

$$\sqrt{nh}\{\widehat{\phi}_+(t_0;\beta_0) - \phi(t_0;\beta_0)\} \to_d N(0, \sigma^2(t_0)).$$

where $\sigma^2(t) \equiv \sigma_A^2(t)/B^2(t)$.

This completes the asymptotic normality of the second part. In the following we will show that the first part vanish as $h \to 0$. Note that

$$\sqrt{nh}\{\widehat{\phi}_+(t_0;\hat{\beta}) - \widehat{\phi}_+(t_0;\beta_0)\} = \frac{\sqrt{nh}C_n(t_0)}{B_n(t_0)},$$

and by Taylor's theorem, $\sqrt{nh}C_n(t_0)$ can be written as

$$\sqrt{nh}C_n(t_0) \tag{2.26}$$

$$=\sqrt{\frac{h}{n}}\sum_{i=1}^{n}\{\exp(-\beta_0^T Z_i) - \exp(-\hat{\beta}^T Z_i)\}\Delta_i K_h(t_0 - X_i) \tag{2.27}$$

$$=\sqrt{\frac{h}{n}}\sum_{i=1}^{n}\{\exp(-\beta_0 Z_i)(\hat{\beta} - \beta_0)^T Z_i + O_p(1/n)\}\Delta_i K_h(t_0 - X_i) \tag{2.28}$$

$$=\sqrt{\frac{h}{n}}\sum_{i=1}^{n}\exp(-\beta_0 Z_i)(\hat{\beta} - \beta_0)^T Z_i \Delta_i K_h(t_0 - X_i) + O_p(1/n)\cdot\sqrt{\frac{h}{n}}\sum_{i=1}^{n}\Delta_i K_h(t_0 - X_i) \tag{2.29}$$

$$=\sqrt{h}\{\sqrt{n}(\hat{\beta} - \beta_0)\}^T\Big\{\frac{1}{n}\sum_{i=1}^{n}\exp(-\beta_0 Z_i)Z_i\Delta_i K_h(t_0 - X_i)\Big\} + O_p(\sqrt{h/n})B_n(t_0) \tag{2.30}$$

$$=\sqrt{h}\{\sqrt{n}(\hat{\beta} - \beta_0)\}^T D_n(t_0;\beta_0) + O_p(\sqrt{h/n})B_n(t_0). \tag{2.31}$$

By the properties of MLE, we have

$$\sqrt{n}(\hat{\beta} - \beta_0) \to_d N(0, I^{-1}(\beta_0))$$

where $I(\beta_0)$ is the Fisher information with respect to β. Combined with the convergence of $D_n(t_0, \beta_0)$ and $B_n(t_0)$, we see that $\sqrt{nh}C_n(t_0) \to 0$ as $h \to 0$, so that

$$\sqrt{nh}\{\widehat{\phi}_+(t_0;\hat{\beta}) - \widehat{\phi}_+(t_0;\beta_0)\} \to 0$$

as $h \to 0$. This completes the proof of theorem 2.

2.3.3 Uniform Consistency of $\widehat{\phi}_-(t;\hat{\beta})$

The derivation of asymptotic results for $\widehat{\phi}_-(t;\hat{\beta})$ is similar to that of $\widehat{\phi}_+(t;\hat{\beta})$. We first introduce the following notation

$$
\begin{aligned}
\psi(t) &= \int \exp(\beta^T z) f_{Z|T\geq t}(z)dz,\\
\widehat{\psi}(t;\hat{\beta}) &= \frac{\sum_{i=1}^n \exp\{\hat{\beta}^T Z_i(t)\}(1-\Delta_i)K_h(t-X_i)}{\sum_{i=1}^n (1-\Delta_i)K_h(t-X_i)},\\
\overline{A}_n(t;\beta) &= \frac{1}{n}\sum_{i=1}^n (1-\Delta_i)K_h(t-X_i)\exp\{\beta^T Z_i(t)\},\\
\overline{B}_n(t) &= \frac{1}{n}\sum_{i=1}^n (1-\Delta_i)K_h(t-X_i),\\
\overline{C}_n(t) &= \frac{1}{n}\sum_{i=1}^n (1-\Delta_i)K_h(t-X_i)\{e^{\hat{\beta}^T Z_i(t)} - e^{\beta_0^T Z_i(t)}\},\\
\overline{D}_n(t;\beta) &= \frac{1}{n}\sum_{i=1}^n (1-\Delta_i)K_h(t-X_i)Z_i(t)e^{\beta^T Z_i(t)},\\
\overline{A}(t) &= \psi(t) f_C(t) S_T(t)(1-\pi_0)/(1-p_0),\\
\overline{B}(t) &= f_C(t) S_T(t)(1-\pi_0)/(1-p_0),\\
\overline{D}(t;\beta) &= \Big\{\int z e^{\beta^T z} dF_{Z|T\geq t}(z)\Big\} f_C(t) S_T(t)(1-\pi_0)/(1-p_0).
\end{aligned}
$$

We state the uniform consistency of $\widehat{\phi}_-(t;\hat{\beta})$ in the following theorem:

Theorem 3. Suppose assumptions A1-A7 are satisfied. Then,

$$\sup_{t\in(0,\tau)} |\widehat{\phi}_-(t;\hat{\beta}) - \phi(t)| \to 0 \text{ a.s.}$$

as $n \to \infty$.

Proof. Since $\widehat{\phi}_-(t;\hat{\beta}) = 1 - \widehat{\psi}(t;\hat{\beta})^{-1}$, it suffices to show that $\widehat{\psi}(t;\hat{\beta})$ is uniformly consistent for $\psi(t)$ for $t \in (0,\tau)$ and then apply the continuous mapping theorem. Note that

$$\widehat{\psi}(t;\hat{\beta}) = \frac{\overline{A}_n(t;\beta_0)}{\overline{B}_n(t)} + \frac{\overline{C}_n(t)}{\overline{B}_n(t)}, \tag{2.32}$$

so it suffices to prove the consistency of $\overline{B}_n(t)$, $\overline{A}_n(t;\beta_0)$, $\overline{C}_n(t)$, then use the similar argument in Theorem 1 to show the consistency of $\widehat{\psi}(t;\hat{\beta})$. Indeed,

$$\begin{aligned}
E^*\{\overline{B}_n(t)\} &= E^*\{(1-\Delta)K_h(t-X)\} \\
&= E^*\{K_h(t-X)|\Delta=0\}P^*(\Delta=0) \\
&= E_0\{K_h(t-C)|\Delta=0\}(1-\pi_0) \\
&= E_0\{K_h(t-C)|C\leq T\}(1-\pi_0) \\
&= \frac{1-\pi_0}{h}\int K\left(\frac{t-x}{h}\right) f_{C|C\leq T}(x)dx \\
&= \frac{1-\pi_0}{1-p_0}\frac{1}{h}\int K\left(\frac{t-x}{h}\right) f_C(x)S_T(x)dx \\
&= \frac{1}{h}\int K\left(\frac{t-x}{h}\right)\overline{B}(x)dx, \qquad (2.33)
\end{aligned}$$

and

$$\begin{aligned}
E^*\{\overline{A}_n(t;\beta_0)\} &= E^*\left[(1-\Delta)K_h(t-X)\exp\{\beta_0^T Z_i(t)\}\right] \\
&= E^*\left[K_h(t-X)\exp\{\beta_0^T Z(t)\}|\Delta=0\right]P^*(\Delta=0) \\
&= E_0\left[K_h(t-C)\exp\{\beta_0^T Z(t)|C\leq T\right](1-\pi_0) \\
&= \frac{1-\pi_0}{h}\iint \exp(\beta_0^T z)K\left(\frac{t-x}{h}\right) f_{C,Z|C\leq T}(x,z)dxdz \\
&= \frac{1-\pi_0}{1-p_0}\frac{1}{h}\iint \exp(\beta_0^T z)K\left(\frac{t-x}{h}\right) f_C(x)S_T(x)f_{Z|T\geq x}(z)dxdz \\
&= \frac{1-\pi_0}{1-p_0}\frac{1}{h}\int K\left(\frac{t-x}{h}\right)\psi(x)f_C(x)S_T(x)dx \\
&= \frac{1}{h}\int K\left(\frac{t-x}{h}\right)\overline{A}(x)dx. \qquad (2.34)
\end{aligned}$$

Therefore, with similar arguments in the proofs of Lemma 1.1 and 1.2, we can conclude that

$$\sup_{t\in(0,\tau)}\left|\frac{\overline{A}_n(t;\beta_0)}{\overline{B}_n(t)}-\psi(t)\right|\to 0 \text{ a.s.} \qquad (2.35)$$

as $n \to \infty$.

Additionally, we have

$$\begin{aligned}
\overline{C}_n(t) &= \frac{1}{n}\sum_{i=1}^{n}\left[\exp\{\hat{\beta}^T Z_i(t)\} - \exp\{\beta_0^T Z_i(t)\}\right](1-\Delta_i)K_h(t-X_i) \\
&= \frac{1}{n}\sum_{i=1}^{n}(\hat{\beta}-\beta_0)^T Z_i(t)\exp\{\beta_0^T Z_i(t)\}(1-\Delta_i)K_h(t-X_i) + o_p(1) \\
&= (\hat{\beta}-\beta_0)^T \overline{D}_n(t;\beta_0) + o_p(1), \qquad (2.36)
\end{aligned}$$

and

$$\begin{aligned}
E^*\{\overline{D}_n(t;\beta)\} &= E^*\left[Z(t)\exp\{\beta^T Z(t)\}(1-\Delta)K_h(t_0-X)\}\right] \\
&= E^*\left[Z(t)\exp\{\beta^T Z(t)\}K_h(t_0-C)|C \le T\right](1-p_0) \\
&= \frac{1-\pi_0}{h}\iint\{z\exp(\beta^T z)\}K\left(\frac{t-x}{h}\right) f_{C,Z|C\le T}(x,z)dxdz \\
&= \frac{1-\pi_0}{1-p_0}\frac{1}{h}\iint\{z\exp(\beta^T z)\}K\left(\frac{t-x}{h}\right) f_C(x)S_T(x)f_{Z|T\ge x}(z)dxdz \\
&= \frac{1}{h}\int K\left(\frac{t-x}{h}\right)\overline{D}(x;\beta)dx \qquad (2.37)
\end{aligned}$$

for any β. Then we conclude using similar arguments in proof of Lemma 1.3 that

$$\sup_{t\in(0,\tau)}\left|\frac{\overline{C}_n(t)}{\overline{B}_n(t)}\right| \to 0 \text{ a.s.} \qquad (2.38)$$

as $n \to \infty$. With (2.35) and (2.38) and by using triangle inequality we can conclude that

$$\sup_{t\in(0,\tau)}|\widehat{\psi}_a(t;\hat{\beta}) - \psi(t)| \to 0 \text{ a.s.}$$

as $n \to \infty$. Theorem 3 is proved.

2.3.4 Asymptotic Normality of $\widehat{\phi}_-(t;\hat{\beta})$

The asymptotic property of $\widehat{\phi}_-(t;\hat{\beta})$ is stated in the following theorem.

Theorem 4. Suppose that assumptions A1-A7 are satisfied. Then $\sqrt{nh}\{\widehat{\phi}_-(t;\hat{\beta}) - \phi(t)\}$ converges weakly to a zero-mean Gaussian process for $t \in [0,\tau]$.

Proof. It suffices to prove the convergence of $\sqrt{nh}\{\widehat{\psi}(t;\hat{\beta}) - \psi(t)\}$ then apply the delta method. We will first prove the pointwise (finite dimensional) convergence, then verify the Stone's condition for tightness.

Let t_0 be any constant between 0 and τ. Write

$$\begin{aligned}\sqrt{nh}\{\widehat{\psi}(t_0;\hat{\beta})-\psi(t_0)\} &= \sqrt{nh}\{\widehat{\psi}(t_0;\hat{\beta})-\widehat{\psi}(t_0;\beta_0)\}+\sqrt{nh}\{\widehat{\psi}(t_0;\beta_0)-\psi(t_0)\}\\ &= \sqrt{nh}\left(\frac{\overline{C}_n(t_0)}{\overline{B}_n(t_0)}\right)+\sqrt{nh}\left(\frac{\overline{A}_n(t_0;\beta_0)}{\overline{B}_n(t_0)}-\psi(t_0)\right).\end{aligned}$$

We first show the convergence of $\sqrt{nh}\{\overline{A}_n(t_0;\beta_0)\overline{B}_n(t_0)^{-1}-\psi(t_0)\}$. Given the uniform consistency for $\overline{B}_n(t)$, it sufficed to show the asymptotic normality of $\sqrt{nh}(\overline{A}_n(t_0;\beta_0)-\psi(t_0)\overline{B}_n(t_0))$. Denote

$$\overline{a}_{n,i}(t_0)=\sqrt{\frac{h}{n}}\{\exp(\beta_0^T Z_i)-\psi(t_0)\}(1-\Delta_i)K_h(t_0-X_i),$$

so that $\sum_{i=1}^n \overline{a}_{n,i}(t_0)=\sqrt{nh}(\overline{A}_n(t_0;\beta_0)-\psi(t_0)\overline{B}_n(t_0))$. We verify the Lyapunov condition in the following steps. First note that

$$\begin{aligned}
&E^*\{\overline{a}_{n,i}(t_0)\}\\
=&\sqrt{\frac{h}{n}}E^*\left[\{\exp(-\beta_0^T Z)-\psi(t_0;\beta_0)\}(1-\Delta)K_h(t_0-X)\right]\\
=&(1-\pi_0)\sqrt{\frac{h}{n}}E_0\left[\{\exp(\beta_0^T Z)-\psi(t_0;\beta_0)\}(1-\Delta)K_h(t_0-C)\middle|C\le T\right]\\
=&\frac{1-\pi_0}{\sqrt{nh}}\iint\{\exp(\beta_0^T z)-\psi(t_0;\beta_0)\}K\left(\frac{t_0-x}{h}\right)f_{C,Z|C\le T}(x,z)dxdz\\
=&\frac{1-\pi_0}{\sqrt{nh}(1-p_0)}\iint\{\exp(\beta_0^T z)-\psi(t_0;\beta_0)\}K\left(\frac{t_0-x}{h}\right)f_C(x)S_{T|Z}(x|z)f_Z(z)dxdz\\
=&\frac{1-\pi_0}{1-p_0}\sqrt{\frac{h}{n}}\iint\{\exp(\beta_0^T z)-\psi(t_0;\beta_0)\}K(u)f_C(t_0-uh)S_{T|Z}(t_0-uh|z)f_Z(z)dudz\\
&\text{by letting } u=\frac{t_0-x}{h},\\
=&\frac{1-\pi_0}{1-p_0}\sqrt{\frac{h}{n}}\left[\int\{\exp(\beta_0^T z)-\psi(t_0;\beta_0)\}f_C(t_0)S_{T|Z}(t_0|z)f_Z(z)dz+h^2\overline{r}(t_0)+O(h^3)\right]\\
&\text{by the Taylor's expansion,}\\
=&\frac{1-\pi_0}{1-p_0}\sqrt{\frac{h}{n}}\{h^2\overline{r}(t_0)+O(h^3)\},
\end{aligned}$$

where

$$\bar{r}(t_0) = \int u^2 K(u)du \int \{\exp(\beta_0^T z) - \psi(t_0;\beta_0)\}\{S'_{T|Z}(t_0|z)f'_C(t_0) + \frac{1}{2}S_{T|Z}(t_0|z)f''_C(t_0) + \frac{1}{2}S''_{T|Z}(t_0|z)f_C(t_0)\}f_Z(z)dz.$$

Similarly, we have

$$\begin{aligned}
&E^*\{\bar{a}_{n,i}^2(t_0)\}\\
=&\frac{h}{n}E^*\left[\{\exp(-\beta_0^T Z) - \psi(t_0;\beta_0)\}^2(1-\Delta)K_h^2(t_0 - X)\right]\\
=&\frac{1-\pi_0}{nh}\iint \{\exp(\beta_0^T z) - \psi(t_0;\beta_0)\}^2 K^2\left(\frac{t_0 - x}{h}\right) f_{C,Z|C\le T}(x,z)dxdz\\
=&\frac{1-\pi_0}{nh(1-p_0)}\iint \{\exp(\beta_0^T z) - \psi(t_0;\beta_0)\}^2 K^2\left(\frac{t_0 - x}{h}\right) f_C(x)S_{T|Z}(x|z)f_Z(z)dxdz\\
=&\frac{1-\pi_0}{n(1-p_0)}\iint \{\exp(\beta_0^T z) - \psi(t_0;\beta_0)\}^2 K^2(u)f_C(t_0 - uh)S_{T|Z}(t_0 - uh|z)f_Z(z)dudz\\
&\text{ by letting } u = \frac{t_0 - x}{h},\\
=&\frac{1-\pi_0}{n(1-p_0)}\left[\int \{\exp(\beta_0^T z) - \psi(t_0;\beta_0)\}^2 f_C(t_0)S_{T|Z}(t_0|z)f_Z(z))dz \int K^2(u)du + O(h)\right]\\
&\text{by the Taylor's expansion,}
\end{aligned}$$

and

$$\begin{aligned}
&E^*\{\bar{a}_{n,i}^3(t_0)\}\\
=&\left(\frac{h}{n}\right)^{3/2} E^*\left[\{\exp(\beta_0^T Z) - \psi(t_0;\beta_0)\}^3(1-\Delta)K_h^3(t_0 - X)\right]\\
=&\frac{1-\pi_0}{1-p_0}(\frac{1}{nh})^{3/2}\iint \{\exp(\beta_0^T z) - \psi(t_0;\beta_0)\}^3 K^3\left(\frac{t_0 - x}{h}\right) f_{C,Z|C\le T}(x,z)dxdz\\
=&\frac{1-\pi_0}{1-p_0}(\frac{1}{nh})^{3/2}\iint \{\exp(\beta_0^T z) - \psi(t_0;\beta_0)\}^3 K^3\left(\frac{t_0 - x}{h}\right) f_C(x)S_{T|Z}(x|z)f_Z(z)dxdz\\
=&\frac{1-\pi_0}{n\sqrt{nh}(1-p_0)}\iint \{\exp(\beta_0^T z) - \psi(t_0;\beta_0)\}^3 K^3(u)f_C(t_0 - uh)S_{T|Z}(t_0 - uh|z)f_Z(z)dudz\\
&\text{ by letting } u = \frac{t_0 - x}{h},\\
=&\frac{1-\pi_0}{n\sqrt{nh}(1-p_0)}\left[\int \{\exp(\beta_0^T z) - \psi(t_0;\beta_0)\}^3 f_C(t_0)S_{T|Z}(t_0|z)f_Z(z)dz \int K^3(u)du + O(h)\right]\\
&\text{by the Taylor's expansion,}
\end{aligned}$$

Given that

$$\begin{aligned}
&\sum_{i=1}^{n} E^*\{\overline{a}_{n,i}(t_0) - E^*\overline{a}_{n,i}(t_0)\}^3 \\
=& nE^*\{\overline{a}_{n,i}(t_0)^3\} - 3nE^*\{\overline{a}_{n,i}(t_0)^2\}E^*\{\overline{a}_{n,i}(t_0)\} + 2n\left[E^*\{\overline{a}_{n,i}(t_0)\}\right]^3 \\
=& O(1/\sqrt{nh}) + 3nO(1/n)\cdot O(\sqrt{h/n}) + 2nO((h/n)^{3/2}) \\
=& O(1/\sqrt{nh}),
\end{aligned}$$

and

$$\begin{aligned}
&\sum_{i=1}^{n} Var^*\{\overline{a}_{n,i}(t_0)\} = nE^*\{\overline{a}^2_{n,i}(t_0)\} - n[E^*\{\overline{a}_{n,i}(t_0)\}]^2 \\
=&(1-\pi_0)\int \{\exp(\beta_0^T z) - \psi(t_0;\beta_0)\}^2 f_C(t_0) S_{T|Z}(t_0|z) f_Z(z)dz \int K^2(u)du + O(h) + O(h^5) \\
=&\sigma^2_{\overline{A}}(t_0) + O(h),
\end{aligned}$$

where

$$\sigma^2_{\overline{A}}(t) = \overline{B}(t)\int K^2(u)du \int \{\exp(\beta_0^T z) - \psi(t;\beta_0)\}^2 dF_{Z|T\geq t}(z), \tag{2.39}$$

the Lyapunov's condition satisfies by verifying

$$\frac{\sum_{i=1}^{n} E^*\{\overline{a}_{n,i}(t) - E^*\overline{a}_{n,i}(t)\}^3}{[\sum_{i=1}^{n} Var^*\{\overline{a}_{n,i}(t)\}]^{3/2}} = O(1/\sqrt{nh}) \to 0$$

as $\sqrt{nh} \to \infty$. Then by Lyapunov's Central Limit Theorem,

$$\frac{\sqrt{nh}(\overline{A}_n(t_0;\beta_0) - \psi(t_0)\overline{B}_n(t_0)) - \sum_{i=1}^{n} E^*\{\overline{a}_{n,i}(t_0)\}}{[\sum_{i=1}^{n} Var^*\{\overline{a}_{n,i}(t_0)\}]^{1/2}} \to_d N(0,1).$$

Since

$$\sum_{i=1}^{n} E^*\{\overline{a}_{n,i}(t_0)\} = nE^*\{\overline{a}_{n,i}(t_0)\} = O(\sqrt{nh^5}) \to 0,$$

and

$$\sum_{i=1}^{n} Var^*\{\overline{a}_{n,i}(t_0)\} = \sigma^2_{\overline{A}}(t_0) + O(h^2),$$

we conclude that

$$\sqrt{nh}\{\overline{A}_n(t_0;\beta_0) - \psi(t_0)\overline{B}_n(t_0)\} \to_d N(0, \sigma^2_{\overline{A}}(t_0)),$$

and thereafter by Slutsky's theorem

$$\sqrt{nh}\left(\frac{\overline{A}_n(t_0;\beta_0)}{\overline{B}_n(t_0)} - \psi(t_0)\right) \to_d N(0, \sigma_\psi^2(t_0))$$

where

$$\sigma_\psi^2(t_0) = \frac{\sigma_{\overline{A}}^2(t_0)}{\overline{B}^2(t_0)}.$$

The next step is to show the convergence of $\sqrt{nh}\frac{\overline{C}_n(t_0)}{\overline{B}_n(t_0)}$. Note that from (2.36), we have

$$\sqrt{nh}\overline{C}_n(t_0) = \sqrt{h}\{\sqrt{n}(\hat{\beta} - \beta_0)^T\}\overline{D}_n(t;\beta_0) + o_p(1).$$

Because $\sqrt{n}(\hat{\beta} - \beta_0)$ converges and $\overline{D}_n(t;\beta_0)$ is bounded almost surely, $\sqrt{nh}\overline{C}_n(t_0)$ vanishes as h goes to 0. Therefore,

$$\sqrt{nh}\{\widehat{\psi}(t_0;\hat{\beta}) - \psi(t_0)\} \to_d N(0, \sigma_\psi^2(t_0)), \tag{2.40}$$

and by the delta method, we have

$$\sqrt{nh}\{\widehat{\phi}_-(t_0;\hat{\beta}) - \phi(t_0)\} \to_d N(0, \sigma_-^2(t_0)), \tag{2.41}$$

where $\sigma_-^2(t_0) = \sigma_\psi^2(t_0)\psi^{-4}(t_0)$. Theorem 4 is proved.

2.3.5 Uniform Consistency and Asymptotic Normality of $\widehat{\phi}_w(t;\hat{\beta})$

We now study the asymptotic properties of $\widehat{\phi}_w(t;\hat{\beta})$. Recall that

$$\widehat{\phi}_w(t;\hat{\beta}) = \pi_0\widehat{\phi}_+(t;\hat{\beta}) + (1-\pi_0)\widehat{\phi}_-(t;\hat{\beta}). \tag{2.42}$$

Since $\widehat{\phi}_w(t;\hat{\beta})$ is the weighted average of $\widehat{\phi}_+(t;\hat{\beta})$ and $\widehat{\phi}_-(t;\hat{\beta})$ with known constant weights, the uniform consistency of $\widehat{\phi}_w(t;\hat{\beta})$ immediately follows the uniform consistency of both $\widehat{\phi}_+(t;\hat{\beta})$ and $\widehat{\phi}_-(t;\hat{\beta})$. Now we prove the asymptotic normality of $\widehat{\phi}_w(t;\hat{\beta})$ with the following theorem.

Theorem 5. Suppose that assumptions A1-A7 are satisfied. Then $\sqrt{nh}\{\widehat{\phi}_w(t;\hat{\beta}) - \phi(t)\}$ converges weakly to a zero-mean Gaussian process for $t \in [0,\tau]$.

Proof. Let $t_0 \in (0,\tau)$. Note that

$$\sqrt{nh}\{\widehat{\phi}_w(t_0;\hat{\beta}) - \phi(t_0)\} = \sqrt{nh}\{\widehat{\phi}_+(t_0;\hat{\beta}) - \phi(t_0)\}\pi_0 + \sqrt{nh}\{\widehat{\phi}_-(t_0;\hat{\beta}) - \phi(t_0)\}(1-\pi_0).$$

We have proved that

$$\begin{aligned}\sqrt{nh}\{\widehat{\phi}_+(t_0;\hat{\beta})-\phi(t_0)\} &= \sqrt{nh}\{A_n(t_0;\beta_0)-\phi(t_0)B_n(t_0)\}B_n^{-1}(t_0)+\sqrt{nh}\{C_n(t_0)\}B_n^{-1}(t_0)\\ &= \sqrt{nh}\{A_n(t_0;\beta_0)-\phi(t_0)B_n(t_0)\}B^{-1}(t_0)+o_p(1),\end{aligned}$$

and

$$\begin{aligned}&\sqrt{nh}\{\widehat{\phi}_-(t_0;\hat{\beta})-\phi(t_0)\}\\ =&\sqrt{nh}\{\widehat{\psi}(t_0;\hat{\beta})-\psi(t_0)\}\psi^{-2}(t_0)+o_p(1)\\ =&\sqrt{nh}\{\overline{A}_n(t_0;\beta_0)-\psi(t_0)\overline{B}_n(t_0)\}\overline{B}_n^{-1}(t_0)+\sqrt{nh}\{\overline{C}_n(t_0)\}\overline{B}_n^{-1}(t_0)+o_p(1)\\ =&\sqrt{nh}\{\overline{A}_n(t_0;\beta_0)-\psi(t_0)\overline{B}_n(t_0)\}\overline{B}^{-1}(t_0)+o_p(1).\end{aligned}$$

Since $A_n(t_0;\beta_0)-\phi(t_0)B_n(t_0)$ involves only the cases, and $\overline{A}_n(t_0;\beta_0)-\psi(t_0)\overline{B}_n(t_0)$ involves only the controls, the two quantities are independent. Thus $\sqrt{nh}\{\widehat{\phi}_+(t_0;\hat{\beta})-\phi(t_0)\}$ and $\sqrt{nh}\{\widehat{\phi}_-(t_0;\hat{\beta})-\phi(t_0)\}$ are asymptotically independent. Given the asymptotic normality of $\sqrt{nh}\{\widehat{\phi}_+(t_0;\hat{\beta})-\phi(t_0)\}$ and $\sqrt{nh}\{\widehat{\phi}_-(t_0;\hat{\beta})-\phi(t_0)\}$ and the property of normal distribution, we can conclude that

$$\sqrt{nh}\{\widehat{\phi}_w(t_0;\hat{\beta})-\phi(t_0)\}\to_d N(0,\sigma_w^2(t_0)),$$

where

$$\sigma_w^2(t)=\pi_0^2\sigma_+^2(t)+(1-\pi_0)^2\sigma_-^2(t). \tag{2.43}$$

Theorem 5 is proved.

2.4 Finite Sample Properties

2.4.1 Variance estimates of $\widehat{\phi}_+(t;\hat{\beta})$ with correction for finite sample

The asymptotic properties derived in the last section are very useful in estimating the variances and confidence intervals. However, variance estimates relying solely on the asymptotic results may perform poorly due to the slow vanishing rate of $\sqrt{h}$. In fact, in real practice with finite samples, $\sqrt{h}$ might not be close to zero even with large sample size. In this case the first part of $S_n(t)$, $\sqrt{nh}\{\widehat{\phi}_+(t;\hat{\beta})-\widehat{\phi}_+(t;\beta_0)\}$ cannot be regarded as zero. Instead we need to consider its contribution to the asymptotic results by taking the bandwidth h as a known non-zero constant.

It can be easily verified that the Lyapunov's condition for asymptotic normality of $\sqrt{nh}A_n(t;\beta_0)$ does not depend on the bandwidth assumption A4, so the asymptotic normality still holds, yet the variance estimate needs to be corrected. The correction term comes from the variance of $\sqrt{nh}\{\widehat{\phi}_+(t_0;\hat{\beta}) - \widehat{\phi}_+(t;\beta_0)\}$ and the covariance between $\sqrt{nh}\{\widehat{\phi}_+(t;\hat{\beta}) - \widehat{\phi}_+(t;\beta_0)\}$ and $\sqrt{nh}\{\widehat{\phi}_+(t;\beta_0) - \phi(t;\beta_0)\}$. These are stated in Theorem 6, after introduction of the follow notation:

$$\sigma_C^2(t;h) = hD(t;\beta_0)^T I^{-1}(\beta_0)D(t;\beta_0),$$
$$\sigma_{AC}(t;h) = hD(t;\beta_0)^T E\left[\Delta_i K_h(t-X_i)\{1-\phi(t;\beta) - e^{-\beta_0 Z_i}\}\tilde{l}_\beta(\mathbf{X_i})\right],$$

where $\tilde{l}_\beta(\mathbf{X_i})$ is the efficient influence function for β in the logistic model.

Theorem 6 Suppose assumptions A1-A3, A5-A7 are satisfied. Then the limiting variance of $\sqrt{nh}\{\widehat{\phi}_+(t_0;\hat{\beta}) - \widehat{\phi}_+(t_0;\beta_0)\}$ equals to $\sigma_C^2(t_0;h)B^{-2}(t_0)$ for any $t_0 \in (0,\tau)$. Moreover, the limiting covariance between $\sqrt{nh}\{\widehat{\phi}_+(t;\hat{\beta}) - \widehat{\phi}_+(t;\beta_0)\}$ and $\sqrt{nh}\{\widehat{\phi}_+(t;\beta_0) - \phi(t;\beta_0)\}$ is $\sigma_{AC}(t_0;h)B^{-2}(t_0)$.

Proof. Since condition A4 is not assumed here, we regard the bandwidth as a known constant. From (2.26) we have

$$\sqrt{nh}C_n(t_0) = \sqrt{h}\{\sqrt{n}(\hat{\beta}-\beta_0)\}^T D_n(t_0;\beta_0) + o_p(1),$$

for any $t_0 \in (0,\tau)$. When h is a pre-specified constant, it is straightforward to obtain

$$\sqrt{nh}C_n(t_0) \to_d N(0, \sigma_C^2(t_0;h)).$$

Therefore, the limiting variance of $\sqrt{nh}\{\widehat{\phi}_+(t_0;\hat{\beta}) - \widehat{\phi}_+(t_0;\beta_0)\}$ equals to $\sigma_C^2(t_0)/B^2(t_0)$. The next step is to study the covariance between $\sqrt{nh}\{A_n(t_0;\beta_0) - \phi(t_0)B_n(t_0)\}$ and $\sqrt{nh}C_n(t_0)$. Write $\sqrt{n}(\hat{\beta}-\beta_0)$ into asymptotic linear form,

$$\sqrt{n}(\hat{\beta}-\beta_0) = \frac{1}{\sqrt{n}}\sum_{i=1}^n \tilde{l}_\beta(\mathbf{X_i}) + o_p(1)$$

and substitute it into (2.26), we can write both $\sqrt{nh}\{A_n(t_0;\beta_0)-\phi(t_0)B_n(t_0)\}$ and $\sqrt{nh}C_n(t_0)$ into asymptotic iid form:

$$\begin{aligned}\sqrt{nh}\{A_n(t_0;\beta_0)-\phi(t_0)B_n(t_0)\} &= \sum_{i=1}^n a_{n,i}(t_0)\\ \sqrt{nh}C_n(t_0;h) &= \sum_{i=1}^n c_{n,i}(t_0)+o_p(1),\end{aligned}$$

where

$$\begin{aligned}a_{n,i}(t_0) &= \sqrt{\frac{h}{n}}\{1-\phi(t_0)-\exp(-\beta_0^T Z_i)\}\Delta_i K_h(t_0-X_i),\\ c_{n,i}(t_0) &= \sqrt{\frac{h}{n}}\tilde{l}_\beta(\mathbf{X_i})^T D(t_0).\end{aligned}$$

Thus the covariance between $\sqrt{nh}\{A_n(t_0;\beta_0)-\phi(t_0)B_n(t_0)\}$ and $\sqrt{nh}C_n(t_0)$ is

$$\begin{aligned}&nCov\{a_{n,i}(t_0),c_{n,i}(t_0)\}+o_p(1)\\ =&nE^*\{a_{n,i}(t_0)c_{n,i}(t_0)\}+o_p(1)\\ =&\sigma_{AC}(t_0;h)+o_p(1).\end{aligned}$$

Therefore, the limiting covariance between between $\sqrt{nh}\{\widehat{\phi}_+(t;\hat{\beta})-\widehat{\phi}_+(t;\beta_0)\}$ and $\sqrt{nh}\{\widehat{\phi}_+(t;\beta_0)-\phi(t;\beta_0)\}$ is $\sigma_{AC}(t_0;h)/B^2(t_0)$.
Theorem 6 is proved.

The variance estimator of $\widehat{\phi}_+(t;\hat{\beta})$ could be established based on the asymptotic variance with finite sampling correction. We propose the following estimators for $\sigma_A^2(t)$, $\sigma_C^2(t;h)$, and $\sigma_{AC}(t;h)$ respectively:

$$\begin{aligned}\widehat{\sigma}_A^2(t) &= \frac{h}{n}\sum_{i=1}^n \Delta_i K_h(t-X_i)^2\{1-\widehat{\phi}_+(t;\hat{\beta})-e^{-\hat{\beta}^T Z_i(t)}\}^2,\\ \widehat{\sigma}_C^2(t;h) &= hD_n(t;\hat{\beta})^T\widehat{I}^{-1}(\hat{\beta})D_n(t;\hat{\beta}),\\ \widehat{\sigma}_{AC}(t;h) &= \frac{h}{n}\sum_{i=1}^n \Delta_i K_h(t-X_i)\{1-\widehat{\phi}_+(t;\hat{\beta})-e^{-\hat{\beta}^T Z_i(t)}\}\widehat{\tilde{l}}_\beta(\mathbf{X_i})^T D_n(t;\hat{\beta}),\end{aligned}$$

where $\widehat{I}^{-1}(\hat{\beta})$ is the estimator of the Fisher's information and $\widehat{\tilde{l}}_\beta(\mathbf{X_i})$ is the estimator of the efficient influence function from the logistic model. Combining these together we have the variance estimator for $\sqrt{nh}\{\widehat{\phi}_+(t;\hat{\beta})-\phi(t)\}$ that

$$\widehat{\sigma}_+^2(t;h)=\frac{\widehat{\sigma}_A^2(t)+2\widehat{\sigma}_{AC}(t;h)+\widehat{\sigma}_C^2(t;h)}{B_n^2(t)}. \tag{2.44}$$

2.4.2 Variance estimates of $\widehat{\phi}_-(t;\hat{\beta})$ with correction for finite sample

Similar to $\widehat{\phi}_+(t_0;\hat{\beta})$, the finite sample performance of $\widehat{\phi}_-(t_0;\hat{\beta})$ depends on the bandwidth, and the variance estimation through asymptotic derivation needs to be corrected in the finite sampling. The idea is to regard the bandwidth as a non-zero constant and calculate the correction term from the variance of $\sqrt{nh}\{\widehat{\psi}(t_0;\hat{\beta})-\widehat{\psi}(t;\beta_0)\}$ and the covariance between $\sqrt{nh}\{\widehat{\psi}(t;\hat{\beta})-\widehat{\psi}(t;\beta_0)\}$ and $\sqrt{nh}\{\widehat{\psi}(t;\beta_0)-\psi(t;\beta_0)\}$, and apply this to the variance estimation. We state the main result in the following theorem.

Theorem 7 Denote

$$\begin{aligned}
\sigma_{\overline{C}}^2(t;h) &= h\overline{D}^T(t;\beta_0)I^{-1}(\beta_0)\overline{D}(t;\beta_0),\\
\sigma_{\overline{AC}}(t;h) &= h\overline{D}(t;\beta_0)^T E\left[(1-\Delta_i)K_h(t-X_i)\{\exp(\beta_0^T Z_i)-\psi(t)\}\tilde{l}_\beta(\mathbf{X_i})\right].
\end{aligned}$$

Suppose assumptions A1-A3, A5-A7 are satisfied. Then the limiting variance of $\sqrt{nh}\{\widehat{\psi}(t_0;\hat{\beta})-\widehat{\psi}(t_0;\beta_0)\}$ equals to $\sigma_{\overline{C}}^2(t_0;h)\overline{B}^{-2}(t_0)$ for any $t_0\in(0,\tau)$. Moreover, the limiting covariance between $\sqrt{nh}\{\widehat{\psi}(t;\hat{\beta})-\widehat{\psi}(t;\beta_0)\}$ and $\sqrt{nh}\{\widehat{\psi}(t;\beta_0)-\psi(t;\beta_0)\}$ is $\sigma_{\overline{AC}}(t_0;h)\overline{B}^{-2}(t_0)$.

Proof. Since Condition A4 is not assumed here, we regard the bandwidth as a known constant. We have established the relationship

$$\sqrt{nh}\overline{C}_n(t_0)=\sqrt{h}\{\sqrt{n}(\hat{\beta}-\beta_0)\}^T\overline{D}_n(t_0;\beta_0)+o(1),$$

Since

$$\begin{aligned}
\sqrt{n}(\hat{\beta}-\beta_0) &\to_d N(0,I^{-1}(\beta_0)),\\
\overline{D}_n(t_0;\beta_0) &\to_p \overline{D}(t_0),
\end{aligned}$$

we have

$$\sqrt{nh}\overline{C}_n(t_0)\to_d N(0,\sigma_{\overline{C}}^2(t_0;h)),$$

and so that the limiting variance of $\sqrt{nh}\{\widehat{\psi}(t_0;\hat{\beta})-\widehat{\psi}(t_0;\beta_0)\}=\sqrt{nh}\{\overline{C}_n(t_0)/\overline{B}_n(t_0)\}$ is $\sigma_{\overline{C}}^2(t_0;h)/\overline{B}^2(t_0)$. Next we will study the covariance between $\sqrt{nh}\{\widehat{\psi}(t;\hat{\beta})-\widehat{\psi}(t;\beta_0)\}$ and

$\sqrt{nh}\{\widehat{\psi}(t;\beta_0)-\psi(t;\beta_0)\}$. Since

$$\sqrt{n}(\hat{\beta}-\beta_0)=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\tilde{l}_{\beta}(\mathbf{X_i})+o_p(1),$$

we can write both $\sqrt{nh}\{\overline{A}_n(t_0;\beta_0)-\psi(t_0)\overline{B}_n(t_0)\}$ and $\sqrt{nh}\overline{C}_n(t_0)$ in asymptotic iid form:

$$\begin{aligned}\sqrt{nh}\{\overline{A}_n(t_0;\beta_0)-\psi(t_0)\overline{B}_n(t_0)\}&=\sum_{i=1}^{n}\overline{a}_{n,i}(t_0)\\ \sqrt{nh}\overline{C}_n(t_0;h)&=\sum_{i=1}^{n}\overline{c}_{n,i}(t_0)+o_p(1).\end{aligned}$$

where

$$\begin{aligned}\overline{a}_{n,i}(t_0)&=\sqrt{\frac{h}{n}}\{\exp(\beta_0^TZ_i)-\psi(t_0)\}(1-\Delta_i)K_h(t_0-X_i),\\ \overline{c}_{n,i}(t_0)&=\sqrt{\frac{h}{n}}\tilde{l}_{\beta}(\mathbf{X_i})^T\overline{D}(t_0).\end{aligned}$$

Thus the limiting covariance between $\sqrt{nh}\{\overline{A}_n(t_0;\beta_0)-\psi(t_0)\overline{B}_n(t_0)\}$ and $\sqrt{nh}\overline{C}_n(t_0;h)$ is

$$\begin{aligned}&nCov\{\overline{a}_{n,i}(t_0),\overline{c}_{n,i}(t_0)\}+o_p(1)\\ =&nE^*\{\overline{a}_{n,i}(t_0)\overline{c}_{n,i}(t_0)\}+o_p(1)\\ =&\sigma_{\overline{AC}}(t_0)+o_p(1),\end{aligned}$$

and the limiting covariance between $\sqrt{nh}\{\widehat{\phi}_-(t;\hat{\beta})-\widehat{\phi}_-(t;\beta_0)\}$ and $\sqrt{nh}\{\widehat{\phi}_+(t;\beta_0)-\phi(t;\beta_0)\}$ is $\sigma_{\overline{AC}}(t_0;h)/\overline{B}^2(t_0)$.

Theorem 7 is proved.

We propose the following estimators for $\sigma^2_{\overline{A}}(t)$, $\sigma^2_{\overline{C}}(t;h)$, and $\sigma_{\overline{AC}}(t;h)$ respectively:

$$\begin{aligned}\widehat{\sigma}^2_{\overline{A}}(t)&=\frac{h}{n}\sum_{i=1}^{n}(1-\Delta_i)K_h^2(t-X_i)\{e^{\hat{\beta}^TZ_i(t)}-\widehat{\psi}(t;\hat{\beta})\}^2,\\ \widehat{\sigma}^2_{\overline{C}}(t;h)&=h\overline{D}_n(t;\hat{\beta})^T\widehat{I}^{-1}(\hat{\beta})\overline{D}_n(t;\hat{\beta}),\\ \widehat{\sigma}_{\overline{AC}}(t;h)&=\frac{h}{n}\sum_{i=1}^{n}(1-\Delta_i)K_h(t-X_i)\{e^{\hat{\beta}^TZ_i(t)}-\widehat{\psi}(t;\hat{\beta})\}\widehat{\tilde{l}}_{\beta}(\mathbf{X_i})^T\overline{D}_n(t;\hat{\beta}),\end{aligned}$$

and thereafter a variance estimator for $\sqrt{nh}\{\widehat{\phi}_-(t;\hat{\beta})-\phi(t)\}$ is

$$\widehat{\sigma}^2_-(t;h)=\frac{\widehat{\sigma}^2_{\overline{A}}(t)+2\widehat{\sigma}_{\overline{AC}}(t;h)+\widehat{\sigma}^2_{\overline{C}}(t;h)}{\overline{B}_n^2(t)\widehat{\psi}^4(t;\hat{\beta})}. \tag{2.45}$$

2.4.3 Variance estimates of $\widehat{\phi}_w(t;\hat{\beta})$ with correction for finite sample

The asymptotic variance of $\widehat{\phi}_w(t;\hat{\beta})$ is a combination of the asymptotic variances of $\widehat{\phi}_+(t;\hat{\beta})$ and $\widehat{\phi}_-(t;\hat{\beta})$. The same finite sample correction on the variance estimates of $\widehat{\phi}_+(t;\hat{\beta})$ and $\widehat{\phi}_-(t;\hat{\beta})$ would be applied in the same way on $\widehat{\phi}_w(t;\hat{\beta})$. Furthermore, when the bandwidth h is not close to zero, $\sqrt{nh}\{\widehat{\phi}_+(t_0;\hat{\beta})-\phi(t_0)\}$ and $\sqrt{nh}\{\widehat{\phi}_-(t_0;\hat{\beta})-\phi(t_0)\}$ cannot be regarded as approximately independent, and their covariance needs to be considered. Here we derive the limiting covariance of $\sqrt{nh}\{\widehat{\phi}_+(t_0;\hat{\beta})-\phi(t_0)\}$ and $\sqrt{nh}\{\widehat{\phi}_-(t_0;\hat{\beta})-\phi(t_0)\}$ under finite sampling. Denote

$$
\begin{aligned}
\sigma_{A\overline{C}}(t;h) &= h\overline{D}(t;\beta_0)^T E\left[\Delta_i K_h(t-X_i)\{1-\phi(t;\beta)-e^{-\beta_0 Z_i}\}\tilde{l}_\beta(\mathbf{X_i})\right],\\
\sigma_{\overline{A}C}(t;h) &= hD(t;\beta_0)^T E\left[(1-\Delta_i)K_h(t-X_i)\{\exp(\beta_0^T Z_i)-\psi(t)\}\tilde{l}_\beta(\mathbf{X_i})\right],\\
\sigma_{C\overline{C}}(t;h) &= hD^T(t;\beta_0)I^{-1}(\beta_0)\overline{D}(t;\beta_0),
\end{aligned}
$$

and let

$$
\sigma_{+-}(t;h)=\frac{\sigma_{A\overline{C}}(t;h)+\sigma_{\overline{A}C}(t;h)+\sigma_{C\overline{C}}(t;h)}{B(t)\overline{B}(t)\psi^2(t)}.
$$

Theorem 8 Suppose assumptions A1-A3, A5-A7 are satisfied. Then the limiting covariance between $\sqrt{nh}\{\widehat{\phi}_+(t_0;\hat{\beta})-\phi(t_0)\}$ and $\sqrt{nh}\{\widehat{\phi}_-(t_0;\hat{\beta})-\phi(t_0)\}$ is $\sigma_{+-}(t;h)$.

Proof. Recall that

$$
\begin{aligned}
&\sqrt{nh}\{\widehat{\phi}_+(t_0;\hat{\beta})-\phi(t_0)\}\\
=&\sqrt{nh}\{A_n(t_0;\beta_0)-\phi(t_0)B_n(t_0)\}B^{-1}(t_0)+\sqrt{nh}\{C_n(t_0)\}B^{-1}(t_0)+o_p(1),
\end{aligned}
$$

and

$$
\begin{aligned}
&\sqrt{nh}\{\widehat{\phi}_-(t_0;\hat{\beta})-\phi(t_0)\}\\
=&\sqrt{nh}\{\overline{A}_n(t_0;\beta_0)-\psi(t_0)\overline{B}_n(t_0)\}\overline{B}^{-1}(t_0)\psi^{-2}(t_0)+\sqrt{nh}\{\overline{C}_n(t_0)\}\overline{B}^{-1}(t_0)\psi^{-2}(t_0)+o_p(1).
\end{aligned}
$$

Since $A_n(t_0;\beta_0)-\phi(t_0)B_n(t_0)$ and $\overline{A}_n(t_0;\beta_0)-\psi(t_0)\overline{B}_n(t_0)$ are independent, the limiting covariance between $\sqrt{nh}\{\widehat{\phi}_+(t_0;\hat{\beta})-\phi(t_0)\}$ and $\sqrt{nh}\{\widehat{\phi}_-(t_0;\hat{\beta})-\phi(t_0)\}$ comes from the limiting covariances between $\{A_n(t_0;\beta_0)-\phi(t_0)B_n(t_0)\}$ and $\overline{C}_n(t_0)$, $\{\overline{A}_n(t_0;\beta_0)-\psi(t_0)\overline{B}_n(t_0)\}$

and $C_n(t_0)$, and $C_n(t_0)$ and $\overline{C}_n(t_0)$. Using the results in the proofs of Theorem 6 and Theorem 7, it is straightforward to show that these limiting covariances are $\sigma_{A\overline{C}}(t_0;h)$, $\sigma_{\overline{A}C}(t_0;h)$, and $\sigma_{C\overline{C}}(t_0;h)$, respectively; hence, the limiting covariance between $\sqrt{nh}\{\widehat{\phi}_+(t_0;\hat{\beta})-\phi(t_0)\}$ and $\sqrt{nh}\{\widehat{\phi}_-(t_0;\hat{\beta})-\phi(t_0)\}$ is $\sigma_{+-}(t;h)$. Theorem 8 is proved.

We propose the following estimators for $\sigma_{A\overline{C}}(t_0;h)$, $\sigma_{\overline{A}C}(t_0;h)$, and $\sigma_{C\overline{C}}(t_0;h)$, respectively:

$$\widehat{\sigma}_{A\overline{C}}(t;h) = \frac{h}{n}\sum_{i=1}^{n}(1-\Delta_i)K_h(t-X_i)\{1-\widehat{\phi}(t;\hat{\beta})-e^{-\hat{\beta}Z_i(t)}\}\widehat{\tilde{l}}_\beta(\mathbf{X_i})^T\overline{D}_n(t;\hat{\beta}),$$

$$\widehat{\sigma}_{\overline{A}C}(t;h) = \frac{h}{n}\sum_{i=1}^{n}(1-\Delta_i)K_h(t-X_i)\{e^{\hat{\beta}^TZ_i(t)}-\widehat{\psi}(t;\hat{\beta})\}\widehat{\tilde{l}}_\beta(\mathbf{X_i})^T D_n(t;\hat{\beta}),$$

$$\widehat{\sigma}_{C\overline{C}}(t;h) = hD_n(t;\hat{\beta})^T\widehat{I}^{-1}(\hat{\beta})\overline{D}_n(t;\hat{\beta}).$$

Then the estimator for $\sigma_{+-}(t;h)$ is

$$\widehat{\sigma}_{+-}(t;h) = \frac{\widehat{\sigma}_{A\overline{C}}(t;h)+\widehat{\sigma}_{\overline{A}C}(t;h)+\widehat{\sigma}_{C\overline{C}}(t;h)}{B_n(t)\overline{B}_n(t)\widehat{\psi}^2(t;\hat{\beta})}.$$

Combining these results, we have the following estimator for $\sigma^2_w(t)$:

$$\widehat{\sigma}^2_w(t) = \pi_0^2\widehat{\sigma}^2_+(t;h)+2\pi_0(1-\pi_0)\widehat{\sigma}_{+-}(t;h)+(1-\pi_0)^2\widehat{\sigma}^2_-(t;h). \tag{2.46}$$

2.5 Calculation of Simultaneous Confidence Bands

2.5.1 Confidence Bands for $\widehat{\phi}_+(t;\hat{\beta})$

In addition to the pointwise confidence intervals, it is also of practical interest to consider simultaneous $100(1-\alpha)$th percentile confidence bands. It is not straightforward to calculate the theoretical confidence bands given the structure of the estimator. Here we provide an approach using resampling techniques adapted from Lin et al. (1994). Let

$$A_i(t;\hat{\beta}) = \{1-\widehat{\phi}_+(t;\hat{\beta})-\exp(-\hat{\beta}^TZ_i)\}\Delta_iK_h(t-X_i).$$

Now consider a process

$$\widehat{S}_+(t) = \frac{\sqrt{h/n}\sum_{i=1}^{n}\left\{A_i(t;\hat{\beta})+\hat{\tilde{l}}_\beta(\mathbf{X_i})D_n(t;\hat{\beta})\right\}M_i}{B_n(t)}$$

with $\{M_i\}, n = 1, ..., n$ be iid random variables with Normal(1, 1). Conditional on the data, $\widehat{S}_+(t)$ converges to a zero-mean Gaussian process by Lindeberg-Feller CLT, and the variance

$$\begin{aligned}
&E[\widehat{S}_+^2(t)|\{\Delta_i, X_i, Z_i(.)\}]\\
=&B_n^{-2}(t)\frac{h}{n}\sum_{i=1}^{n}\Big\{A_i(t;\hat{\beta}) + \hat{\tilde{l}}_\beta(\mathbf{X_i})D_n(t;\hat{\beta})\Big\}^2\\
=&B_n^{-2}(t)\Bigg\{\frac{h}{n}\sum_{i=1}^{n}A_i^2(t;\hat{\beta}) + \frac{h}{n}\sum_{i=1}^{n}\hat{\tilde{l}}_\beta^2(\mathbf{X_i})D_n^2(t;\hat{\beta}) + 2\frac{h}{n}\sum_{i=1}^{n}A_i(t;\hat{\beta})\hat{\tilde{l}}_\beta(\mathbf{X_i})D_n(t;\hat{\beta})\Bigg\}\\
=&B_n^{-2}(t)\Big\{\widehat{\sigma}_A^2(t) + \widehat{\sigma}_C^2(t;h) + 2\widehat{\sigma}_{AC}(t;h)\Big\}\\
=&\widehat{\sigma}_+^2(t).
\end{aligned}$$

Thus the distribution of $S_n(t)$ can be approximated by simulating a large number of realizations from $\widehat{S}_+(t)$ through repeatedly generating $\{M_i\}$. Then the critical value for $100(1-\alpha)$th percentile simultaneous confidence bands can be calculated based on these simulations:

$$P\Bigg\{\sup_{t\in[0,\tau]}\frac{|\widehat{S}_+(t)|}{\widehat{\sigma}_+(t;h)} \le \widetilde{z}_{1-\alpha/2}\Bigg\} \doteq 1-\alpha,$$

and the corresponding confidence bands are

$$\widehat{\phi}_+(t;\widehat{\beta}) \mp \widetilde{z}_{1-\alpha/2}(nh)^{-1/2}\widehat{\sigma}_+(t;h).$$

2.5.2 Confidence Bands for $\widehat{\phi}_-(t;\hat{\beta})$

The confidence bands for $\widehat{\phi}_-(t;\widehat{\beta})$ can be constructed using the same resampling technique. Let

$$\overline{A}_i(t;\hat{\beta}) = \{\exp(\hat{\beta}^T Z_i) - \widehat{\psi}(t;\hat{\beta})\}(1-\Delta_i)K_h(t-X_i),$$

and $\{M_i\}, n = 1, ..., n$ be iid standard normal random variables. Then we consider the following process

$$\widehat{S}_-(t) = \frac{\sqrt{h/n}\sum_{i=1}^{n}\Big\{\overline{A}_i(t;\hat{\beta}) + \hat{\tilde{l}}_\beta(\mathbf{X_i})\overline{D}_n(t;\hat{\beta})\Big\}M_i}{\overline{B}_n(t)\widehat{\psi}^2(t;\hat{\beta})}.$$

Conditional on the data, $\widehat{S}_{-}(t)$ converges to a zero-mean Gaussian process by Lindeberg-Feller CLT, and the variance

$$
\begin{aligned}
& E[\widehat{S_{-}}(t)|\{\Delta_i, X_i, Z_i(.)\}] \\
=& \overline{B}_n^{-2}(t)\widehat{\psi}^{-4}(t;\hat{\beta})\frac{h}{n}\sum_{i=1}^{n}\left\{\overline{A}_i(t;\hat{\beta})+\hat{\hat{l}}_{\beta}(\mathbf{X_i})\overline{D}_n(t;\hat{\beta})\right\}^2 \\
=& \overline{B}_n^{-2}(t)\widehat{\psi}^{-4}(t;\hat{\beta})\left\{\frac{h}{n}\sum_{i=1}^{n}\overline{A}_i^2(t;\hat{\beta})+\frac{h}{n}\sum_{i=1}^{n}\hat{\hat{l}}_{\beta}^2(\mathbf{X_i})\overline{D}_n^2(t;\hat{\beta})+2\frac{h}{n}\sum_{i=1}^{n}\overline{A}_i(t;\hat{\beta})\hat{\hat{l}}_{\beta}(\mathbf{X_i})\overline{D}_n(t;\hat{\beta})\right\} \\
=& \overline{B}_n^{-2}(t)\widehat{\psi}^{-4}(t;\hat{\beta})\left\{\widehat{\sigma}_{\overline{A}}^2(t)+\widehat{\sigma}_{\overline{C}}^2(t;h)+2\widehat{\sigma}_{\overline{AC}}(t;h)\right\} \\
=& \widehat{\sigma}_{-}^2(t).
\end{aligned}
$$

The critical value for $100(1-\alpha)$th percentile simultaneous confidence bands can be calculated by

$$
P\left\{\sup_{t\in[0,\tau]}\frac{|\widehat{S}_{-}(t)|}{\widehat{\sigma}_{-}(t;h)}\leq\widetilde{z}_{1-\alpha/2}\right\}\doteq 1-\alpha,
$$

and the corresponding confidence bands are

$$
\widehat{\phi}_{-}(t;\widehat{\beta})\mp\widetilde{z}_{1-\alpha/2}(nh)^{-1/2}\widehat{\sigma}_{-}(t;h).
$$

2.5.3 Confidence Bands for $\widehat{\phi}_w(t;\hat{\beta})$

To calculate the simultaneous confidence bands for $\widehat{\phi}_w(t;\hat{\beta})$, we can construct the following process:

$$
\widehat{S}_w(t)=\sqrt{\frac{h}{n}}\sum_{i=1}^{n}\left\{\pi_0\frac{A_i(t;\hat{\beta})+\hat{\hat{l}}_{\beta}(\mathbf{X_i})D_n(t;\hat{\beta})}{B_n(t)}+(1-\pi_0)\frac{\overline{A}_i(t;\hat{\beta})+\hat{\hat{l}}_{\beta}(\mathbf{X_i})\overline{D}_n(t;\hat{\beta})}{\overline{B}_n(t)\widehat{\psi}^2(t;\hat{\beta})}\right\}M_i,
$$

where $\{M_i\}, n=1,...,n$ are iid standard normal random variables. Conditional on the data, $\widehat{S}_w(t)$ converges to a zero-mean Gaussian process and the variance equals to $\widehat{\sigma}_w^2(t)$. So that the critical value for $100(1-\alpha)$th percentile simultaneous confidence bands can be calculated by

$$
P\left\{\sup_{t\in[0,\tau]}\frac{|\widehat{S}_w(t)|}{\widehat{\sigma}_w(t;h)}\leq\widetilde{z}_{1-\alpha/2}\right\}\doteq 1-\alpha,
$$

and the corresponding confidence bands are

$$
\widehat{\phi}_w(t;\widehat{\beta})\mp\widetilde{z}_{1-\alpha/2}(nh)^{-1/2}\widehat{\sigma}_w(t;h).
$$

2.6 Bandwidth Selection

2.6.1 Bandwidth selection for $\widehat{\phi}_{+}(t;\hat{\beta})$

It is known that in kernel estimation bandwidth choice is critical to the adequacy of the estimator. It is of practical interest to develop an automatic bandwidth selector for the estimator of time-varying PAF. A well-working approach for optimizing the bandwidth is the "leave-one-out" cross-validation approach(Bierens 1983). The idea is to split the data into two parts, the larger part that contains all but one subject, and the smaller part that contains one subject, then use the larger part of the data for estimation and the smaller part for evaluation of accuracy. The optimal bandwidth would minimize the average prediction squared errors.

In particular, let

$$\widehat{\phi}_{+}^{(-j)}(X_j;h) = 1 - \frac{\sum_{i=1}^{n} I(i \neq j)\exp(-\hat{\beta}^T Z_i)\Delta_i K_h(X_j - X_i)}{\sum_{i=1}^{n} I(i \neq j)\Delta_i K_h(X_j - X_i)},$$

which is the estimator of $\phi(X_j)$ computed with all measurements of the jth subject deleted. Then the cross-validation criterion is given by

$$CV(h) = \frac{1}{n}\sum_{j=1}^{n} \Delta_j \left[1 - \exp\{-\hat{\beta}^T Z_j(X_j)\} - \widehat{\phi}_{+}^{(-j)}(X_j;h)\right]^2$$

A cross-validation bandwidth is obtained by minimizing $CV(h)$ with respect to h,

$$\hat{h}_{CV} = \arg\min_{h>0} CV(h).$$

This approach generally works well in practice, however, the theoretical properties of $\hat{h}_{CV}$ as well as asymptotic properties of $\widehat{\phi}_{+}^{(-j)}(X_j;\hat{h}_{CV})$ are hard to track. In addition, the process may be very computationally demanding. Instead, we could use a variant of this approach adopting finite grid search, which has very close adequacy yet is much less computationally intensive. That is, we optimize h on a set of prespecified grid points,

$$CV(\hat{h}_{CV}) = \inf\{CV(h_i), i = 1, 2, ..., K\}$$

where $h_1 < h_2 < ... < h_K$ are the grid points.

2.6.2 Bandwidth selection for $\widehat{\phi}_{-}(t;\hat{\beta})$

We can apply the same cross-validation based bandwidth selection approach on $\widehat{\phi}_{-}(t;\hat{\beta})$. Let

$$\widehat{\phi}_{-}^{(-j)}(X_j;h) = 1 - \frac{\sum_{i=1}^{n} I(i \neq j)(1-\Delta_i)K_h(t-X_i)}{\sum_{i=1}^{n} I(i \neq j)\exp\{\hat{\beta}^T Z_i(t)\}(1-\Delta_i)K_h(t-X_i)},$$

which is the estimator of $\phi(X_j)$ computed with all measurements of the jth subject deleted. Since $\{1-\widehat{\phi}_{-}(t;\hat{\beta})\}^{-1}$ estimates $\int_{\mathcal{Z}} exp(\beta z) dF_{Z|T\geq t}(z)$, the cross-validation criterion for $\widehat{\phi}_{-}(t;\hat{\beta})$ is

$$CV(h) = \frac{1}{n}\sum_{j=1}^{n}(1-\Delta_j)\left[\{1-\widehat{\phi}_{-}^{(-j)}(X_j;h)\}^{-1} - \exp\{\hat{\beta}^T Z_j(X_j)\}\right]^2,$$

and the finite grid search approach yields the optimal bandwidth

$$CV(\hat{h}_{CV}) = \inf\{CV(h_i), i = 1,2,...,K\}$$

for pre-specified grid points $h_1 < h_2 < ... < h_K$.

2.6.3 Bandwidth selection for $\widehat{\phi}_{w}(t;\hat{\beta})$

The cross-validation based bandwidth selector for $\widehat{\phi}_{w}(t;\hat{\beta})$ is a combination of the selectors for $\widehat{\phi}_{+}(t;\hat{\beta})$ and $\widehat{\phi}_{-}(t;\hat{\beta})$:

$$\begin{aligned} CV(h) = \frac{1}{n}\sum_{j=1}^{n}\Bigg(&\pi_0\Delta_j\left[1-\exp\{-\hat{\beta}^T Z_j(X_j)\} - \widehat{\phi}_{+}^{(-j)}(X_j;h)\right] \\ &+ (1-\pi_0)(1-\Delta_j)\left[\{1-\widehat{\phi}_{-}^{(-j)}(X_j;h)\}^{-1} - \exp\{\hat{\beta}^T Z_j(X_j)\}\right]\Bigg)^2. \end{aligned}$$

Then the optimal bandwidth can be obtained by taking the minimal criterion over a series of pre-specified grid points,

$$CV(\hat{h}_{CV}) = \inf\{CV(h_i), i = 1,2,...,K\}$$

for pre-specified grid points $h_1 < h_2 < ... < h_K$.

Chapter 3

ESTIMATION OF TIME-VARYING PAF FOR CASE-CONTROL STUDIES WITH ADJUSTMENT FOR CONFOUNDERS

3.1 Overview

In Chapter 2 we studied the estimation of $\phi(t)$ without considering confounders. Like the classic PAF, the crude estimates of $\phi(t)$ are in general biased in the presence of confounders. In this chapter we study the adjusted estimators of $\phi(t)$ that take into account of risk factors that may confound the association between the exposure of interest and disease. Recall that for classic PAF, the most commonly used approach for adjusted estimate is the regression modeling approach. Under this approach, the PAF formula is written as [33, 7]

$$\begin{aligned} PAF &= \frac{P(D=1) - \sum_{j=1}^{J} P(D=1|Z=0, U=u_j)P(U=u_j)}{P(D=1)} \\ &= 1 - \sum_{j=1}^{J}\sum_{i=0}^{I} \frac{P(Z=z_i, U=u_j|D=1)}{RR_{i|j}} \end{aligned} \tag{3.1}$$

where U is the vector of adjustment factors forming J levels, and $RR_{i|j} = P(D=1|Z=z_i, U=u_j)/P(D=1|Z=0, U=u_j)$.

In this study we will show how the adjusted time-varying PAF $\phi(t)$ can be assessed through regression modeling approaches. We propose a set of adjusted estimators for $\phi(t)$, and study their asymptotic and finite sampling properties.

3.2 Formulation of Estimation Problem

3.2.1 Definition of Time-varying PAF with Adjustment for Confounders

We followed the same formulation for case-control data as in Chapter 2. The only difference is that here we let the data consist of confounding factors U. Therefore, the case-control data consisting of n subjects will be represented by the notation $\mathbf{X} = (\{X_i, \Delta_i, Z_i(.), U_i(.)\}, i = 1, ..., n)$.

We define the adjusted time-varying PAF of Z in the presence of U as

$$\phi_{adj}(t) = \frac{\lambda(t) - E_{U|T\geq t}\{\lambda(t|Z=0,U)\}}{\lambda(t)}, \tag{3.2}$$

which is the proportional reduction of hazard at time t comparing the current population and the hypothetical population with no exposure and stratified by the potential confounder (U) among subject at risk.

Consider the following Cox model,

$$\lambda\{t|Z(t), U(t)\} = \lambda_0(t)\exp\{\beta^T Z(t) + \gamma^T U(t)\}, \tag{3.3}$$

where $Z(t) = \{Z_1(t), ..., Z_p(t)\}^T$ is a $p \times 1$ vector of external time-dependent covariates, $\beta = (\beta_1, ..., \beta_p)^T$ is a vector of regression parameters for $Z(t)$, $U(t) = \{U_1(t), ..., U_q(t)\}^T$ is a $q \times 1$ vector of adjustment factors, and $\gamma = (\gamma_1, ..., \gamma_q)^T$ is a vector of regression parameters for U. Under model (3.3), (3.2) can be expressed as

$$\begin{aligned}\phi_{adj}(t) &= 1 - \frac{\int_{\mathcal{U}} \lambda(t|Z=0,u)dF_{U|T\geq t}(u)}{\lambda(t)} \\ &= 1 - \frac{\lambda_0(t)}{\lambda(t)}\int_{\mathcal{U}} \exp(\gamma^T u) f_{U|T\geq t}(u)du. \end{aligned}\tag{3.4}$$

By using the relationship $\lambda(t) = f_T(t)/S_T(t)$ and applying total probability theorem on $S_T(t)$, we can further express (3.4) as

$$\phi_{adj}(t) = 1 - \int_{\mathcal{U}}\int_{\mathcal{Z}} \exp(-\beta^T z - \gamma^T u) f_{Z,U|T}(z,u|t)dzdu \int_{\mathcal{U}} \exp(\gamma^T u) f_{U|T\geq t}(u)du, \tag{3.5}$$

or

$$\phi_{adj}(t) = 1 - \Big\{\int_{\mathcal{U}}\int_{\mathcal{Z}} \exp(\beta^T z + \gamma^T u) f_{Z,U|T\geq t}(z,u)dzdu\Big\}^{-1} \int_{\mathcal{U}} \exp(\gamma^T u) f_{U|T\geq t}(u)du. \tag{3.6}$$

Equation 3.6 explicitly show the impact of U. For example, when $\beta = 0$, $\phi_{adj}(t) = 0$ for all t. When $\gamma = 0$ or Z and U are independent, we have $\phi_{adj}(t) = \phi(t)$, the unadjusted time-varying PAF.

3.2.2 Point Estimators of $\phi(t)$

Let

$$\varphi(t) = \int_{\mathcal{U}}\int_{\mathcal{Z}} exp(-\beta^T z - \gamma^T u) f_{Z,U|T}(z, u|t) dz du,$$
$$\psi(t) = \int_{\mathcal{U}}\int_{\mathcal{Z}} exp(\beta^T z + \gamma^T u) f_{Z,U|T\geq t}(z, u) dz du,$$
$$\upsilon(t) = \int_{\mathcal{U}} \exp(\gamma^T u) f_{U|T\geq t}(u) du.$$

Then $\phi_{adj}(t)$ can be expressed as

$$\phi_{adj}(t) = 1 - \varphi(t)\upsilon(t),$$

or

$$\phi_{adj}(t) = 1 - \psi^{-1}(t)\upsilon(t).$$

We propose the following estimators for $\varphi(t)$, $\psi(t)$, and $\upsilon(t)$ respectively:

$$\widehat{\varphi}(t; \hat{\beta}, \hat{\gamma}) = \frac{\sum_{i=1}^{n} \exp\{-\hat{\beta}^T Z_i(t) - \hat{\gamma}^T U_i(t)\}\Delta_i K_h(t - X_i)}{\sum_{i=1}^{n} \Delta_i K_h(t - X_i)},$$
$$\widehat{\psi}(t; \hat{\beta}, \hat{\gamma}) = \frac{\sum_{i=1}^{n} \exp\{\hat{\beta}^T Z_i(t) + \hat{\gamma}^T U_i(t)\}(1 - \Delta_i) K_h(t - X_i)}{\sum_{i=1}^{n} (1 - \Delta_i) K_h(t - X_i)},$$
$$\widehat{\upsilon}(t; \hat{\gamma}) = \frac{\sum_{i=1}^{n} \exp\{\hat{\gamma}^T U_i(t)\}(1 - \Delta_i) K_h(t - X_i)}{\sum_{i=1}^{n} (1 - \Delta_i) K_h(t - X_i)},$$

where $K_h(x) = K(x/h)/h$, $K(.)$ is a kernel function that is a weighting function that satisfies $\int K(x)dx = 1$, and h is the bandwidth that controls the spread of weighting window. $\hat{\beta}$ and $\hat{\gamma}$ are the estimates of β and γ obtained from the logistic regression or conditional logistic regression.

With these estimators, we have the following two estimators for $\phi_{adj}(t)$, corresponding to formula (3.5) and (3.6) respectively:

$$\widehat{\phi}_{adj+}(t; \hat{\beta}, \hat{\gamma}) = 1 - \widehat{\varphi}(t; \hat{\beta}, \hat{\gamma})\widehat{\upsilon}(t; \hat{\gamma}), \tag{3.7}$$

and

$$\widehat{\phi}_{adj-}(t; \hat{\beta}, \hat{\gamma}) = 1 - \widehat{\psi}^{-1}(t; \hat{\beta}, \hat{\gamma})\widehat{\upsilon}(t; \hat{\gamma}), \tag{3.8}$$

and a weighted estimator of the two,

$$\widehat{\phi}_{adjw}(t;\hat{\beta},\hat{\gamma}) = \pi_0\widehat{\phi}_{adj+}(t;\hat{\beta},\hat{\gamma}) + (1-\pi_0)\widehat{\phi}_{adj-}(t;\hat{\beta},\hat{\gamma}). \tag{3.9}$$

We study the properties of these estimators in the next sections.

3.3 Large Sample Properties

In this section we will show that under certain assumptions and regularity conditions, both $\widehat{\phi}_{adj+}(t;\hat{\beta},\hat{\gamma})$ and $\widehat{\phi}_{adj-}(t;\hat{\beta},\hat{\gamma})$ are uniformly consistent for $\phi_{adj}(t)$ for $t \in (0,\tau)$. In addition, both $\sqrt{nh}\{\widehat{\phi}_{adj+}(t;\hat{\beta},\hat{\gamma}) - \phi_{adj}(t)\}$ and $\sqrt{nh}\{\widehat{\phi}_{adj-}(t;\hat{\beta},\hat{\gamma}) - \phi_{adj}(t)\}$ converge weakly to zero-mean Gaussian processes. Assumptions A1-A7 are required for the consistency and asymptotic normality. Additionally, the following assumption is needed:

Assumptions.

A8. U is bounded almost surely, and is independent of censoring time C.

3.3.1 Uniform Consistency of $\widehat{\phi}_{adj+}(t;\hat\beta,\hat\gamma)$, $\widehat{\phi}_{adj-}(t;\hat\beta,\hat\gamma)$, and $\widehat{\phi}_{adjw}(t;\hat\beta,\hat\gamma)$

We first introduce the following notation that will be used in the proof. To save notation, we recycle some notation used in Chapter 2.

$$
\begin{aligned}
p_0 &= P_0(T \le C), \pi_0 = P^*(T \le C),\\
V_i(t) &= (Z_i^T(t), U_i^T(t))^T, \theta = (\beta^T, \gamma^T)^T,\\
A_n(t;\beta,\gamma) &= \frac{1}{n}\sum_{i=1}^n \Delta_i K_h(t-X_i)\exp\{-\beta^T Z_i(t) - \gamma^T U_i(t)\},\\
B_n(t) &= \frac{1}{n}\sum_{i=1}^n \Delta_i K_h(t-X_i),\\
C_n(t) &= \frac{1}{n}\sum_{i=1}^n \Delta_i K_h(t-X_i)\left[\exp\{-\beta_0^T Z_i(t) - \gamma_0^T U_i(t)\} - \exp\{-\hat\beta^T Z_i(t) - \hat\gamma^T U_i(t)\}\right],\\
D_n(t;\theta) &= \frac{1}{n}\sum_{i=1}^n \Delta_i K_h(t-X_i) V_i(t)\exp\{-\theta^T V_i(t)\},\\
\overline{A}_n(t;\beta,\gamma) &= \frac{1}{n}\sum_{i=1}^n (1-\Delta_i) K_h(t-X_i)\exp\{\beta^T Z_i(t) + \gamma^T U_i(t)\},\\
\overline{B}_n(t) &= \frac{1}{n}\sum_{i=1}^n (1-\Delta_i) K_h(t-X_i),\\
\overline{C}_n(t) &= \frac{1}{n}\sum_{i=1}^n (1-\Delta_i) K_h(t-X_i)\left[\exp\{\hat\beta^T Z_i(t) + \hat\gamma^T U_i(t)\} - \exp\{\beta_0^T Z_i(t) + \gamma_0^T U_i(t)\}\right],\\
\overline{D}_n(t;\theta) &= \frac{1}{n}\sum_{i=1}^n (1-\Delta_i) K_h(t-X_i) V_i(t)\exp\{\theta^T V_i(t)\}.
\end{aligned}
$$

The limits of $A_n(t;\beta,\gamma)$, $B_n(t)$, and $D_n(t;\theta)$ are denoted by

$$
\begin{aligned}
A(t) &= \varphi(t) f_T(t) S_C(t)\pi_0/p_0,\\
B(t) &= f_T(t) S_C(t)\pi_0/p_0,\\
D(t;\theta) &= \left\{\int_{\mathcal{V}} v\exp(-\theta^T v)dF_{V|T}(v|t)\right\} f_T(t) S_C(t)\pi_0/p_0,
\end{aligned}
$$

respectively, the limits of $\overline{A}_n(t;\beta,\gamma)$, $\overline{B}_n(t)$, and $\overline{D}_n(t;\beta,\gamma)$ are denoted by

$$
\begin{aligned}
\overline{A}(t) &= \psi(t) f_C(t) S_T(t)(1-\pi_0)/(1-p_0),\\
\overline{B}(t) &= f_C(t) S_T(t)(1-\pi_0)/(1-p_0),\\
\overline{D}(t;\theta) &= \left\{\int_{\mathcal{V}} v\exp(-\theta^T v)dF_{V|T}(v|t)\right\} f_C(t) S_T(t)(1-\pi_0)/(1-p_0).
\end{aligned}
$$

We state the main results in the following two theorems.

Theorem 7. Suppose assumptions A1-A8 are satisfied. Then

$$\sup_{t\in(0,\tau)} |\widehat{\phi}_{adj+}(t;\hat{\beta},\hat{\gamma}) - \phi_{adj}(t)| \to 0 \text{ a.s.}$$

as $n \to \infty$.

Theorem 8. Suppose assumptions A1-A8 are satisfied. Then

$$\sup_{t\in(0,\tau)} |\widehat{\phi}_{adj-}(t;\hat{\beta},\hat{\gamma}) - \phi_{adj}(t)| \to 0 \text{ a.s.}$$

as $n \to \infty$.

Theorem 9. Suppose assumptions A1-A8 are satisfied. Then

$$\sup_{t\in(0,\tau)} |\widehat{\phi}_{adjw}(t;\hat{\beta},\hat{\gamma}) - \phi_{adj}(t)| \to 0 \text{ a.s.}$$

as $n \to \infty$.

Theorem 7, 8, 9 are proved with the help of the following lemmas.

Lemma 8.1. Suppose assumptions A1-A8 are satisfied. Then,

$$\sup_{t\in(0,\tau)} \left|\widehat{\varphi}(t;\hat{\beta},\hat{\gamma}) - \varphi(t)\right| \to 0 \text{ a.s.}$$

as $n \to \infty$.

Proof. The consistency of $B_n(t)$ has been shown in Lemma 1.1 in Chapter 2. Hence as argued in Lemma 1.2 in Chapter 2, it suffices to show the consistency of the numerator $A_n(t;\hat{\beta},\hat{\gamma})$. To simplify the notation, unless otherwise specified, the integration with respect to x is from 0 to ∞, integration with respect to z is on its support $\mathcal{Z}$ and integration with respect to u is on its support $\mathcal{U}$ through out the proofs in this chapter.

Note that $A_n(t;\hat{\beta},\hat{\gamma}) = A_n(t;\beta_0,\gamma_0) - C_n(t)$. The first term has expectation

$$\begin{aligned}
&E^*\{A_n(t;\beta_0,\gamma_0)\} &(3.10)\\
=&E^*\left[\Delta K_h(t-X)\exp\{-\beta_0^T Z(t) - \gamma_0^T U(t)\}\right]\\
=&E^*\left[K_h(t-T)\exp\{-\beta_0^T Z(t) - \gamma_0^T U(t)\}|T\le C\right]P^*(T\le C)\\
=&\frac{\pi_0}{h}\iiint \exp(-\beta_0^T z - \gamma_0^T u)K\left(\frac{t-x}{h}\right) f_{T,Z,U|T\le C}(x,z,u)dxdzdu\\
=&\frac{\pi_0}{h}\iiint \exp(-\beta_0^T z - \gamma_0^T u)K\left(\frac{t-x}{h}\right) \frac{f_{T|Z,U}(x|z,u)S_C(x)f_{Z,U}(z,u)}{P(T\le C)}dxdzdu\\
=&\frac{1}{h}\int K\left(\frac{t-x}{h}\right) f_T(x)S_C(x)\pi_0/p_0\iint \exp(-\beta_0^T z - \gamma_0^T u)f_{Z,U|T}(z,u|x)dzdudx,\\
=&\frac{1}{h}\int K\left(\frac{t-x}{h}\right)A(x)dx. &(3.11)
\end{aligned}$$

Using the same arguments as in Lemma 1.1 and 1.2 in Chapter 2, we can conclude that

$$\sup_{t\in(0,\tau)}|A_n(t;\beta_0,\gamma_0) - A(t)| \to 0 \text{ a.s.} \tag{3.12}$$

as $n\to 0$, given $K(.)$ has bounded variation, and $Z(.)$ and $U(.)$ are both bounded almost surely.

By using Taylor's expansion, the second term $C_n(t)$ can written as

$$\begin{aligned}
C_n(t) &= \frac{1}{n}\sum_{i=1}^n \left[\exp\{-\theta_0^T V_i(t)\} - \exp\{-\hat{\theta}^T V_i(t)\}\right]\Delta_i K_h(t-X_i)\\
&= (\hat{\theta}-\theta_0)^T D_n(t;\theta_0) + o_p(1).
\end{aligned}$$

Consistency of $D_n(t;\theta)$ can be established similarly as in the proof of Lemma 1.3 in Chapter 2, by noticing that

$$\begin{aligned}
E^*\{D_n(t;\theta)\} &= E^*\left[V_i(t)\exp\{-\theta^T V_i(t)\}\Delta_i K_h(t_0-X_i)\}\right]\\
&= E^*\left[V_i(t)\exp\{-\theta^T V_i(t)\}K_h(t_0-X_i)|T\le C\right]P^*(T\le C)\\
&= \frac{\pi_0}{h}\iint\{v\exp(-\theta^T v)\}K\left(\frac{t-x}{h}\right)f_{T,V|T\le C}(x,v)dxdv\\
&= \frac{1}{h}\int K\left(\frac{t-x}{h}\right)D(x;\theta)dx, &(3.13)
\end{aligned}$$

$D(t;\theta)$ is continuous for t, and $V(.)$ is bounded almost surely. By the strong consistency of $\hat{\theta}$ and the continuous mapping theorem, we have

$$\sup_{t\in(0,\tau)} |C_n(t)| \to 0 \text{ a.s.} \tag{3.14}$$

as $n \to 0$. Putting (3.10) and (3.14) together, we conclude that

$$\sup_{t\in(0,\tau)} \left| A_n(t;\hat{\beta},\hat{\gamma}) - A(t) \right| \to 0 \text{ a.s.} \tag{3.15}$$

as $n \to 0$. Lemma 8.1 is proved.

Lemma 8.2. Suppose assumptions A1-A8 are satisfied. Then,

$$\sup_{t\in(0,\tau)} \left| \widehat{\psi}(t;\hat{\beta},\hat{\gamma}) - \psi(t) \right| \to 0 \text{ a.s.}$$

as $n \to \infty$.

Proof. Consistency of $\overline{B}_n(t)$ has been shown in the proof of Theorem 3, so it suffices to show the consistency of the numerator $\overline{A}_n(t;\hat{\beta},\hat{\gamma})$. Note that $\overline{A}_n(t;\hat{\beta},\hat{\gamma}) = \overline{A}_n(t;\beta_0,\gamma_0) + \overline{C}_n(t)$. The expectation of $\overline{A}_n(t;\beta_0,\gamma_0)$ equals

$$\begin{aligned} E^*\{\overline{A}_n(t;\beta_0,\gamma_0)\} &= E^*\left[(1-\Delta)K_h(t-X)\exp\{\beta_0^T Z(t) + \gamma_0^T U(t)\}\right] \\ &= E^*\left[K_h(t-T)\exp\{\beta_0^T Z(t) + \gamma_0^T U(t)\} | C \le T\right] P^*(C \le T) \\ &= \frac{1-\pi_0}{h} \iiint \exp(\beta_0^T z + \gamma_0^T u) K\left(\frac{t-x}{h}\right) f_{C,Z,U|C\le T}(x,z,u)dxdzdu \\ &= \frac{1-\pi_0}{1-p_0}\frac{1}{h} \iiint \exp(\beta_0^T z + \gamma_0^T u) K\left(\frac{t-x}{h}\right) \frac{f_C(x)S_T(x)f_{Z,U}(z,u)}{P(T\le C)} dxdzdu \\ &= \frac{1}{h}\int K\left(\frac{t-x}{h}\right) A(x)dx. \end{aligned} \tag{3.16}$$

Using the same arguments as in Lemma 1.1 and 1.2 in Chapter 2, we conclude that

$$\sup_{t\in(0,\tau)} \left|\overline{A}_n(t;\beta_0,\gamma_0) - \overline{A}(t)\right| \to 0 \text{ a.s.} \tag{3.17}$$

as $n \to 0$, given $K(.)$ has bounded variation, and $Z(.)$ and $U(.)$ are both bounded almost surely.

By using Taylor's expansion, the second term $\overline{C}_n(t)$ can written as

$$\begin{aligned}\overline{C}_n(t) &= \frac{1}{n}\sum_{i=1}^{n}\left[\exp\{\hat{\theta}^T V_i(t)\} - \exp\{\theta_0^T V_i(t)\}\right](1-\Delta_i)K_h(t-X_i)\\ &= (\hat{\theta}-\theta_0)^T\overline{D}_n(t;\theta_0) + o_p(1).\end{aligned}$$

The consistency of $\overline{D}_n(t;\theta)$ can be established similarly as in the proof of Lemma 8.1, and by the strong consistency of $\hat{\theta}$ and the continuous mapping theorem, we conclude that

$$\sup_{t\in(0,\tau)}\left|\overline{C}_n(t)\right| \to 0 \text{ a.s.} \tag{3.18}$$

as $n \to 0$. Putting (3.16) and (3.18) together, we conclude that

$$\sup_{t\in(0,\tau)}\left|\overline{A}_n(t;\hat{\beta},\hat{\gamma}) - \overline{A}(t)\right| \to 0 \text{ a.s.} \tag{3.19}$$

as $n \to 0$. Lemma 8.2 is proved.

Lemma 8.3. Suppose assumptions A1-A8 are satisfied. Then,

$$\sup_{t\in(0,\tau)}|\widehat{\upsilon}(t;\hat{\gamma}) - \upsilon(t)| \to 0 \text{ a.s.}$$

as $n \to \infty$.

Proof. The proof is done by replacing θ with γ, and $V_i(t)$ with $U_i(t)$, and following the same argument in the proof of Lemma 8.2.

Proof of Theorem 7. By applying Lemma 8.1 and 8.3, and using the continuous mapping theorem, we can say that $\widehat{\phi}_{adj+}(t;\hat{\beta},\hat{\gamma})$ is consistent for $\phi_{adj}(t)$ uniformly on $t \in (0,\tau)$. Theorem 7 is proved.

Proof of Theorem 8. By applying Lemma 8.2 and 8.3, and using the continuous mapping theorem, we can say that $\widehat{\phi}_{adj-}(t;\hat{\beta},\hat{\gamma})$ is consistent for $\phi_{adj}(t)$ uniformly on $t \in (0,\tau)$. Theorem 8 is proved.

Proof of Theorem 9. By Theorem 7 and 8 and the continuous mapping theorem, we can say that $\widehat{\phi}_{adjw}(t;\hat{\beta},\hat{\gamma})$ is consistent for $\phi_{adj}(t)$ uniformly on $t \in (0,\tau)$. Theorem 9 is proved.

3.3.2 Asymptotic Normality of $\widehat{\phi}_{adj+}(t;\hat{\beta},\hat{\gamma})$, $\widehat{\phi}_{adj-}(t;\hat{\beta},\hat{\gamma})$, and $\widehat{\phi}_{adjw}(t;\hat{\beta},\hat{\gamma})$

In this subsection we study the asymptotic normality of the proposed estimators. We state the results in the following theorems.

Theorem 10. Suppose that assumptions A1-A8 are satisfied. Then

$$\sqrt{nh}\{\widehat{\phi}_{adj+}(t_0;\hat{\beta},\hat{\gamma}) - \phi(t_0)\} \to_d N(0,\sigma_+^2(t_0))$$

for $t_0 \in (0,\tau)$.

Proof. Let $t_0 \in (0,\tau)$. First note that

$$\begin{aligned}
&\sqrt{nh}\{\widehat{\phi}_{adj+}(t_0;\hat{\beta},\hat{\gamma}) - \phi(t_0)\}\\
=&\sqrt{nh}\{\widehat{\varphi}(t_0;\hat{\beta},\hat{\gamma})\widehat{\upsilon}(t_0;\hat{\gamma}) - \varphi(t_0)\upsilon(t_0)\}\\
=&\sqrt{nh}\{\widehat{\varphi}(t_0;\hat{\beta},\hat{\gamma})\widehat{\upsilon}(t_0;\hat{\gamma}) - \widehat{\varphi}(t_0;\hat{\beta},\hat{\gamma})\upsilon(t_0) + \widehat{\varphi}(t_0;\hat{\beta},\hat{\gamma})\upsilon(t_0) - \varphi(t_0)\upsilon(t_0)\}\\
=&\sqrt{nh}\{\widehat{\upsilon}(t_0;\hat{\gamma}) - \upsilon(t_0)\}\widehat{\varphi}(t_0;\hat{\beta},\hat{\gamma}) + \sqrt{nh}\{\widehat{\varphi}(t_0;\hat{\beta},\hat{\gamma}) - \varphi(t_0)\}\upsilon(t_0).
\end{aligned}$$

We first study the asymptotic normality of $\sqrt{nh}\{\widehat{\upsilon}(t_0;\hat{\gamma}) - \upsilon(t_0)\}$. Let

$$\begin{aligned}
\overline{E}_n(t;\gamma) &= \frac{1}{n}\sum_{i=1}^{n}(1-\Delta_i)K_h(t-X_i)\exp\{\gamma^T U_i(t)\},\\
\overline{F}_n(t) &= \frac{1}{n}\sum_{i=1}^{n}(1-\Delta_i)K_h(t-X_i)\left[\exp\{\hat{\gamma}^T U_i(t)\} - \exp\{\gamma_0^T U_i(t)\}\right],\\
\overline{G}_n(t;\gamma) &= \frac{1}{n}\sum_{i=1}^{n}(1-\Delta_i)K_h(t-X_i)U_i(t)\exp\{\gamma^T U_i(t)\},\\
\overline{E}(t) &= \upsilon(t)f_C(t)S_T(t)(1-\pi_0)/(1-p_0),\\
\overline{G}(t;\gamma) &= \Big\{\int_{\mathcal{U}} u\exp(-\gamma^T u)dF_{U|T}(u|t)\Big\}f_C(t)S_T(t)(1-\pi_0)/(1-p_0).
\end{aligned}$$

Since the format of the relevant quantities are the same as those in Chapter 2, to save space and avoid redundancy, we skip the details of the derivation which could be easily transferred from that in Chapter 2.

We write $\sqrt{nh}\{\widehat{\upsilon}(t_0;\hat{\gamma})-\upsilon(t_0)\}$ into a summation of asymptotic iid terms,

$$\begin{aligned}\sqrt{nh}\{\widehat{\upsilon}(t_0;\hat{\gamma})-\upsilon(t_0)\} &= \sqrt{nh}\{\overline{E}_n(t_0;\gamma_0)-\overline{B}_n(t_0)\upsilon(t_0)\}\overline{B}_n^{-1}(t_0)+\sqrt{nh}\overline{F}_n(t_0)\overline{B}_n^{-1}(t_0)\\ &= \sum_{i=1}^n \overline{e}_{n,i}(t_0)\{\overline{B}^{-1}(t_0)+o_p(1)\}+\{\sum_{i=1}^n \overline{f}_{n,i}(t_0)+o_p(1)\}\{\overline{B}^{-1}(t_0)+o_p(1)\}\\ &= \sum_{i=1}^n \overline{e}_{n,i}(t_0)\overline{B}^{-1}(t_0)+\sum_{i=1}^n \overline{f}_{n,i}(t_0)\overline{B}^{-1}(t_0)+o_p(1),\end{aligned}$$

where

$$\begin{aligned}\overline{e}_{n,i}(t_0) &= \sqrt{\frac{h}{n}}\{\exp(\gamma_0^T U_i)-\upsilon(t_0)\}(1-\Delta_i)K_h(t_0-X_i),\\ \overline{f}_{n,i}(t_0) &= \sqrt{\frac{h}{n}}\tilde{l}_\gamma(\mathbf{X_i})^T\overline{G}(t_0;\gamma_0),\end{aligned}$$

where $\tilde{l}_\gamma(\mathbf{X_i})$ is the efficient influence function for γ in the logistic model. Using the same argument as in the proof of Theorem 4, replacing β and $Z_i(.)$ with γ and $U_i(.)$ respectively, we have

$$\sqrt{nh}\{\widehat{\upsilon}(t_0;\gamma_0)-\upsilon(t_0)\}=\sqrt{nh}\left(\frac{\overline{E}_n(t_0;\gamma_0)}{\overline{B}_n(t_0)}-\upsilon(t_0)\right)\to_d N\left(0,\sigma_{\overline{E}}^2(t_0)\overline{B}_n^{-2}(t_0)\right),$$

where

$$\sigma_{\overline{E}}^2(t)=\overline{B}(t)\int K^2(u)du\int\{\exp(\gamma_0^T u)-\upsilon(t;\gamma_0)\}^2 dF_{U|T\geq t}(u).$$

We also have

$$\sqrt{nh}\{\widehat{\upsilon}(t_0;\hat{\gamma})-\widehat{\upsilon}(t_0;\gamma_0)\}=\sqrt{nh}\left(\frac{\overline{F}_n(t_0;\gamma_0)}{\overline{B}_n(t_0)}\right)=\sqrt{h}\{\sqrt{n}(\hat{\gamma}-\gamma_0)\}\overline{G}_n(t_0;\gamma_0)+o_p(1),$$

which converges to 0 as $h\to 0$ given the convergence of $\sqrt{n}(\hat{\gamma}-\gamma_0)$ and $\overline{G}_n(t_0;\gamma_0)$. Therefore,

$$\sqrt{nh}\{\widehat{\upsilon}(t_0;\hat{\gamma})-\upsilon(t_0)\}\to_d N\left(0,\sigma_{\overline{E}}^2(t_0)\overline{B}_n^{-2}(t_0)\right).$$

The asymptotic normality of $\sqrt{nh}\{\widehat{\varphi}(t_0;\hat{\beta},\hat{\gamma})-\varphi(t_0)\}$ can be obtained using the same argument as in the proof of Theorem 2, replacing β and $Z_i(.)$ with θ and $V_i(.)$. Specifically, we have

$$\sqrt{nh}\{\widehat{\varphi}(t_0;\hat{\beta},\hat{\gamma})-\varphi(t_0)\}=\sqrt{nh}\left(\frac{A_n(t_0;\beta_0,\gamma_0)}{B_n(t_0)}-\varphi(t_0)\right)\to_d N\left(0,\frac{\sigma_A^2(t_0)}{B^2(t_0)}\right),$$

where

$$\sigma_A^2(t) = B(t)\int K^2(u)du \iint \{\exp(-\beta_0^T z - \gamma_0^T u) - \varphi(t;\beta_0,\gamma_0)\}^2 dF_{Z,U|T}(z,u),$$

and

$$\sqrt{nh}\{\widehat{\varphi}(t_0;\hat\beta,\hat\gamma) - \widehat{\varphi}(t_0;\beta_0,\gamma_0)\} = \sqrt{nh}\left(\frac{A_n(t_0;\beta_0\gamma_0)}{B_n(t_0)}\right) = \sqrt{h}\{\sqrt{n}(\hat\theta - \theta_0)\}D_n(t_0;\theta_0) + o_p(1),$$

which converges to 0 as $h \to 0$ given the convergence of $\sqrt{n}(\hat\theta - \theta_0)$ and $D_n(t_0;\theta_0)$. Therefore we conclude that

$$\sqrt{nh}\{\widehat{\varphi}(t_0;\hat\beta,\hat\gamma) - \varphi(t_0)\} \to_d N\left(0, \sigma_A^2(t_0)B^{-2}(t_0)\right).$$

Now we look at the joint distribution of $\sqrt{nh}\{\widehat{\upsilon}(t_0;\hat\gamma) - \upsilon(t_0)\}\widehat{\varphi}(t_0;\hat\beta,\hat\gamma)$ and $\sqrt{nh}\{\widehat{\varphi}(t_0;\hat\beta,\hat\gamma) - \varphi(t_0)\}\upsilon(t_0)$. Since

$$\begin{aligned}
\widehat{\varphi}(t_0;\hat\beta,\hat\gamma) &= \varphi(t_0) + O_p(1/\sqrt{nh}), \\
\sqrt{nh}\{\widehat{\upsilon}(t_0;\hat\gamma) - \upsilon(t_0)\} &= \sqrt{nh}\{\widehat{\upsilon}(t_0;\gamma_0) - \upsilon(t_0)\} + o_p(1), \\
\sqrt{nh}\{\widehat{\varphi}(t_0;\hat\beta,\hat\gamma) - \varphi(t_0)\} &= \sqrt{nh}\{\widehat{\varphi}(t_0;\beta_0,\gamma_0) - \varphi(t_0)\} + o_p(1),
\end{aligned}$$

it suffices to study the joint distribution of $\sqrt{nh}\{\widehat{\upsilon}(t_0;\gamma_0) - \upsilon(t_0)\}$ and $\sqrt{nh}\{\widehat{\varphi}(t_0;\beta_0,\gamma_0) - \varphi(t_0)\}$. Note that $\widehat{\upsilon}(t_0;\gamma_0)$ involves only controls and $\widehat{\varphi}(t_0;\beta_0,\gamma_0)$ involves only cases, so that they are independent. Therefore, with the property of normal distribution, we can conclude that

$$\sqrt{nh}\{\widehat{\upsilon}(t_0;\hat\gamma) - \upsilon(t_0)\}\widehat{\varphi}(t_0;\hat\beta,\hat\gamma) + \sqrt{nh}\{\widehat{\varphi}(t_0;\hat\beta,\hat\gamma) - \varphi(t_0)\}\upsilon(t_0) \to_d N(0, \sigma_+^2(t_0)), \quad (3.20)$$

where

$$\sigma_+^2(t_0) = \frac{\overline{\sigma}_E^2(t_0)\varphi^2(t_0)}{\overline{B}^2(t_0)} + \frac{\sigma_A^2(t_0)\upsilon^2(t_0)}{B^2(t_0)}. \quad (3.21)$$

Theorem 10 is proved.

Theorem 11. Suppose that assumptions A1-A8 are satisfied. Then

$$\sqrt{nh}\{\widehat{\phi}_{adj-}(t_0;\hat\beta,\hat\gamma) - \phi(t_0)\} \to_d N(0, \sigma_-^2(t_0))$$

for $t_0 \in (0, \tau)$.

Proof. Let $t_0 \in (0, \tau)$. Write

$$\begin{aligned}\sqrt{nh}\{\widehat{\phi}_{adj-}(t_0;\hat{\beta},\hat{\gamma}) - \phi(t_0)\} &= \sqrt{nh}\{\widehat{\psi}^{-1}(t_0;\hat{\beta},\hat{\gamma})\widehat{\upsilon}(t_0;\hat{\gamma}) - \psi^{-1}(t_0)\upsilon(t_0)\} \\ &= \sqrt{nh}\left\{\frac{\overline{E}_n(t_0;\hat{\gamma})}{\overline{A}_n(t_0;\hat{\beta},\hat{\gamma})} - \psi^{-1}(t_0)\upsilon(t_0)\right\} \\ &= \sqrt{nh}\left\{\frac{\overline{E}_n(t_0;\hat{\gamma}) - \psi^{-1}(t_0)\upsilon(t_0)\overline{A}_n(t_0;\hat{\beta},\hat{\gamma})}{\overline{A}_n(t_0;\hat{\beta},\hat{\gamma})}\right\}.\end{aligned}$$

The consistency of $\overline{A}_n(t_0;\hat{\beta},\hat{\gamma})$ has been shown in the proof of Lemma 8.2, so it suffices to study the asymptotic normality of $\sqrt{nh}\{\overline{E}_n(t_0;\hat{\gamma}) - \psi^{-1}(t_0)\upsilon(t_0)\overline{A}_n(t_0;\hat{\beta},\hat{\gamma})\}$. Using the results in the previous proofs, we note that

$$\begin{aligned}&\sqrt{nh}\{\overline{E}_n(t_0;\hat{\gamma}) - \psi^{-1}(t_0)\upsilon(t_0)\overline{A}_n(t_0;\hat{\beta},\hat{\gamma})\} \\ =&\sqrt{nh}\{\overline{E}_n(t_0;\gamma_0) - \psi^{-1}(t_0)\upsilon(t_0)\overline{A}_n(t_0;\beta_0,\gamma_0)\} + \sqrt{nh}\{\overline{F}_n(t_0) - \psi^{-1}(t_0)\upsilon(t_0)\overline{C}_n(t_0)\} \\ =&\sqrt{nh}\{\overline{E}_n(t_0;\gamma_0) - \psi^{-1}(t_0)\upsilon(t_0)\overline{A}_n(t_0;\beta_0,\gamma_0)\} + o_p(1),\end{aligned}$$

so the asymptotic normality of $\sqrt{nh}\{\overline{E}_n(t_0;\hat{\gamma}) - \psi^{-1}(t_0)\upsilon(t_0)\overline{A}_n(t_0;\hat{\beta},\hat{\gamma})\}$ only depends on $\sqrt{nh}\{\overline{E}_n(t_0;\gamma_0) - \psi^{-1}(t_0)\upsilon(t_0)\overline{A}_n(t_0;\beta_0,\gamma_0)\}$. Let

$$e_{n,i}(t_0) = \sqrt{\frac{h}{n}}\{\exp(\gamma_0^T U_i) - \psi^{-1}(t_0)\upsilon(t_0)\exp(\beta_0^T Z_i + \gamma_0 U_i)\}(1-\Delta_i)K_h(t_0 - X_i),$$

so that $\sqrt{nh}\{\overline{E}_n(t_0;\gamma_0) - \psi^{-1}(t_0)\upsilon(t_0)\overline{A}_n(t_0;\beta_0,\gamma_0)\} = \sum_{i=1}^n e_{n,i}(t_0)$. We will set forth the Lyapunov' conditions. Let $g(u,z;t_0) = \exp(\gamma_0^T u) - \psi^{-1}(t_0)\upsilon(t_0)\exp(\beta_0^T z + \gamma_0^T u)$. Then we have

$$
\begin{aligned}
& E^*\{e_{n,i}(t_0)\} \\
=& \sqrt{\frac{h}{n}} E^* \left[\{\exp(\gamma_0^T U) - \psi^{-1}(t_0)\upsilon(t_0) \exp(\beta_0^T Z + \gamma_0^T U)\}(1-\Delta) K_h(t_0 - X) \right] \\
=& \frac{1-\pi_0}{\sqrt{nh}} \iiint g(u,z;t_0) K\left(\frac{t_0 - x}{h}\right) f_{C,Z,U|C\le T}(x,z,u) dx dz du \\
=& \frac{1-\pi_0}{\sqrt{nh}(1-p_0)} \iiint g(u,z;t_0) K\left(\frac{t_0 - x}{h}\right) f_C(x) S_{T|Z,U}(x|z,u) f_{Z,U}(z,u) dx dz du \\
=& \frac{1-\pi_0}{1-p_0} \sqrt{\frac{h}{n}} \iiint g(u,z;t_0) K(y) f_C(t_0 - yh) S_{T|Z,U}(t_0 - yh|z,u) f_{Z,U}(z,u) dy dz du \\
& \text{ by letting } y = \frac{t_0 - x}{h}, \\
=& \frac{1-\pi_0}{1-p_0} \sqrt{\frac{h}{n}} \left\{ \iint g(u,z;t_0) f_C(t_0) S_{T|Z,U}(t_0|z,u) f_{Z,U}(z,u) dz du + O(h^2) \right\} \\
& \text{by the Taylor's expansion,} \\
=& \frac{1-\pi_0}{1-p_0} \sqrt{\frac{h}{n}} O(h^2) = O(\sqrt{h^5/n}).
\end{aligned}
$$

The last step comes from the fact that

$$
\begin{aligned}
& \iint g(u,z;t_0) f_C(t_0) S_{T|Z,U}(t_0|z,u) f_{Z,U}(z,u) dz du \\
=& \iint \{\exp(\gamma_0^T u) - \psi^{-1}(t_0)\upsilon(t_0) \exp(\beta_0^T z + \gamma_0^T u)\} f_C(t_0) S_{T|Z,U}(t_0|z,u) f_{Z,U}(z,u) dz du \\
=& f_C(t_0) S_T(t_0) \iint \{\exp(\gamma_0^T u) - \psi^{-1}(t_0)\upsilon(t_0) \exp(\beta_0^T z + \gamma_0^T u)\} f_{Z,U|T\ge t}(z,u) dz du \\
=& f_C(t_0) S_T(t_0) \{\upsilon(t_0) - \psi^{-1}(t_0)\upsilon(t_0) \cdot \psi(t_0)\} \\
=& 0.
\end{aligned}
$$

Similarly, we have

$$
\begin{aligned}
&E^*\{e^2_{n,i}(t_0)\}\\
=&\frac{h}{n}E^*\left[\{\exp(\gamma_0^T U)-\psi^{-1}(t_0)\upsilon(t_0)\exp(\beta_0^T Z+\gamma_0^T U)\}^2(1-\Delta)^2K_h^2(t_0-X)\right]\\
=&\frac{1-\pi_0}{nh}\iiint g^2(u,z;t_0)K^2\left(\frac{t_0-x}{h}\right)f_{C,Z,U|C\le T}(x,z,u)dxdzdu\\
=&\frac{1-\pi_0}{nh(1-p_0)}\iiint g^2(u,z;t_0)K^2\left(\frac{t_0-x}{h}\right)f_C(x)S_{T|Z,U}(x|z,u)f_{Z,U}(z,u)dxdzdu\\
=&\frac{1-\pi_0}{n(1-p_0)}\iiint g^2(u,z;t_0)K^2(y)f_C(t_0-yh)S_{T|Z,U}(t_0-yh|z,u)f_{Z,U}(z,u)dydzdu\\
&\text{ by letting } y=\frac{t_0-x}{h},\\
=&\frac{1-\pi_0}{n(1-p_0)}\left\{\iint g^2(u,z;t_0)f_C(t_0)S_{T|Z,U}(t_0|z,u)f_{Z,U}(z,u)dzdu\int K^2(y)dy+O(h)\right\}\\
&\text{by the Taylor's expansion,}
\end{aligned}
$$

and

$$
\begin{aligned}
&E^*\{e^3_{n,i}(t_0)\}\\
=&\left(\frac{h}{n}\right)^{3/2}E^*\left[\{\exp(\gamma_0^T U)-\psi^{-1}(t_0)\upsilon(t_0)\exp(\beta_0^T Z+\gamma_0^T U)\}^3(1-\Delta)^3K_h^3(t_0-X)\right]\\
=&\frac{1-\pi_0}{1-p_0}(\frac{1}{nh})^{3/2}\iiint g^3(u,z;t_0)K^3\left(\frac{t_0-x}{h}\right)f_{C,Z,U|C\le T}(x,z,u)dxdzdu\\
=&\frac{1-\pi_0}{1-p_0}(\frac{1}{nh})^{3/2}\iiint g^3(u,z;t_0)K^3\left(\frac{t_0-x}{h}\right)f_C(x)S_{T|Z,U}(x|z,u)f_{Z,U}(z,u)dxdzdu\\
=&\frac{1-\pi_0}{n\sqrt{nh}(1-p_0)}\iiint g^3(u,z;t_0)K^3(y)f_C(t_0-yh)S_{T|Z,U}(t_0-yh|z,u)f_{Z,U}(z,u)dydzdu\\
&\text{ by letting } y=\frac{t_0-x}{h},\\
=&\frac{1-\pi_0}{n\sqrt{nh}(1-p_0)}\left\{\iint g^3(u,z;t_0)f_C(t_0)S_{T|Z,U}(t_0|z,u)f_{Z,U}(z,u)dzdu\int K^3(y)dy+O(h)\right\}\\
&\text{by the Taylor's expansion.}
\end{aligned}
$$

Given that

$$\begin{aligned}
&\sum_{i=1}^{n} E^*\{e_{n,i}(t_0) - E^* e_{n,i}(t_0)\}^3 \\
=&nE^*\{e_{n,i}(t_0)^3\} - 3nE^*\{e_{n,i}(t_0)^2\}E^*\{e_{n,i}(t_0)\} + 2n\left[E^*\{e_{n,i}(t_0)\}\right]^3 \\
=&O(1/\sqrt{nh}) + 3nO(1/n)\cdot O(\sqrt{h/n}) + 2nO((h/n)^{3/2}) \\
=&O(1/\sqrt{nh}),
\end{aligned}$$

and

$$\begin{aligned}
&\sum_{i=1}^{n} Var^*\{e_{n,i}(t_0)\} = nE^*\{e_{n,i}^2(t_0)\} - n[E^*\{e_{n,i}(t_0)\}]^2 \\
=&(1-\pi_0)\iint g^2(u,z;t_0) f_C(t_0) S_{T|Z,U}(t_0|z,u) f_{Z,U}(z,u)dzdu \int K^2(y)dy + O(h) + O(h^5) \\
=&\sigma_E^2(t_0) + O(h),
\end{aligned}$$

where

$$\sigma_E^2(t) = \overline{B}(t)\int K^2(u)du \iint \{\exp(\gamma_0^T u) - \psi^{-1}(t_0)\upsilon(t_0)\exp(\beta_0^T z + \gamma_0^T u)\}^2 f_{Z,U|T\geq t}(z,u)dzdu, \tag{3.22}$$

the Lyapunov's condition satisfies by verifying

$$\frac{\sum_{i=1}^{n} E^*\{e_{n,i}(t) - E^* e_{n,i}(t)\}^3}{[\sum_{i=1}^{n} Var^*\{e_{n,i}(t)\}]^{3/2}} = O(1/\sqrt{nh}) \to 0$$

as $\sqrt{nh} \to \infty$. Then by Lyapunov's Central Limit theorem and Slutsky's theorem, we conclude that

$$\sqrt{nh}\{\overline{E}_n(t_0;\gamma_0) - \psi^{-1}(t_0)\upsilon(t)\overline{A}_n(t_0;\beta_0,\gamma_0)\} \to_d N(0,\sigma_E^2(t_0)), \tag{3.23}$$

and thereafter

$$\sqrt{nh}\{\widehat{\phi}_{adj-}(t_0;\hat{\beta},\hat{\gamma}) - \phi(t_0)\} \to_d N(0,\sigma_-^2(t_0)), \tag{3.24}$$

where $\sigma_-^2(t_0) = \sigma_E^2(t_0)\overline{A}_n^{-2}(t_0;\beta_0,\gamma_0)$. Theorem 11 is proved.

Theorem 12. Suppose that assumptions A1-A8 are satisfied. Then

$$\sqrt{nh}\{\widehat{\phi}_{adjw}(t_0;\hat{\beta},\hat{\gamma}) - \phi(t_0)\} \to_d N(0,\sigma_w^2(t_0))$$

for $t_0 \in (0, \tau)$.

Proof. Let $t_0 \in (0, \tau)$. Write

$$
\begin{aligned}
&\sqrt{nh}\{\widehat{\phi}_{adjw}(t_0; \hat{\beta}, \hat{\gamma}) - \phi(t_0)\} \\
=&\pi_0\sqrt{nh}\{\widehat{\upsilon}(t_0; \hat{\gamma}) - \upsilon(t_0)\}\widehat{\varphi}(t_0; \hat{\beta}, \hat{\gamma}) + \pi_0\sqrt{nh}\{\widehat{\varphi}(t_0; \hat{\beta}, \hat{\gamma}) - \varphi(t_0)\}\upsilon(t_0) \\
+&(1 - \pi_0)\sqrt{nh}\{\widehat{\phi}_{adj-}(t_0; \hat{\beta}, \hat{\gamma}) - \phi(t_0)\}.
\end{aligned}
$$

The asymptotic normality of these quantities has been studied in Theorem 10 and 11. Now we study the covariance between them. We have shown in the proof of Theorem 10 and 11 that the asymptotic normality of $\sqrt{nh}\{\widehat{\varphi}(t_0; \hat{\beta}, \hat{\gamma}) - \varphi(t_0)\}$ is determined by the cases, and that of $\sqrt{nh}\{\widehat{\phi}_{adj-}(t_0; \hat{\beta}, \hat{\gamma}) - \phi(t_0)\}$ is determined by the controls. Therefore they are asymptotically independent. The asymptotic covariance comes only from $\sqrt{nh}\{\widehat{\upsilon}(t_0; \hat{\gamma}) - \upsilon(t_0)\}$ and $\sqrt{nh}\{\widehat{\phi}_{adj-}(t_0; \hat{\beta}, \hat{\gamma}) - \phi(t_0)\}$. Write them in iid form, we have

$$
\begin{aligned}
\sqrt{nh}\{\widehat{\upsilon}(t_0; \hat{\gamma}) - \upsilon(t_0)\} &= \sum_{i=1}^{n} \overline{e}_{n,i}(t_0)\overline{B}^{-1}(t_0) + o_p(1), \\
\sqrt{nh}\{\widehat{\phi}_{adj-}(t_0; \hat{\beta}, \hat{\gamma}) - \phi(t_0)\} &= \sum_{i=1}^{n} e_{n,i}(t_0)\overline{A}^{-1}(t_0; \beta_0, \gamma_0) + o_p(1).
\end{aligned}
$$

Therefore, the limiting covariance is

$$
\sigma_{+-}(t_0) = \sigma_{E\overline{E}}(t_0)\overline{B}^{-1}(t_0)\overline{A}^{-1}(t_0; \beta_0, \gamma_0),
$$

where

$$
\sigma_{E\overline{E}}(t_0) = nE\{\overline{e}_{n,i}(t_0)e_{n,i}(t_0)\},
$$

and hence by multivariate central limit theorem,

$$
\sqrt{nh}\{\widehat{\phi}_{adjw}(t_0; \hat{\beta}, \hat{\gamma}) - \phi(t_0)\} \to_d N(0, \sigma_w^2(t_0)),
$$

where

$$
\sigma_w^2(t) = \pi_0^2\sigma_+^2(t) + 2\pi_0(1 - \pi_0)\sigma_{+-}(t) + (1 - \pi_0)^2\sigma_-^2(t).
$$

Theorem 12 is proved.

3.4 *Finite Sample Properties*

3.4.1 *Variance estimates of $\widehat{\phi}_{adj+}(t;\hat{\beta};\hat{\gamma})$ with correction for finite sample*

As we have discussed in Chapter 2, the bandwidth has a substantial effect on the variance estimate under finite sampling. Here we show to estimate the variance of $\widehat{\phi}_{adj+}(t;\hat{\beta};\hat{\gamma})$ with finite sampling correction. Recall that

$$\sqrt{nh}\{\widehat{\phi}_{adj+}(t_0;\hat{\beta},\hat{\gamma})-\phi(t_0)\}=\sqrt{nh}\{\widehat{\upsilon}(t_0;\hat{\gamma})-\upsilon(t_0)\}\widehat{\varphi}(t_0;\hat{\beta},\hat{\gamma})+\sqrt{nh}\{\widehat{\varphi}(t_0;\hat{\beta},\hat{\gamma})-\varphi(t_0)\}\upsilon(t_0).$$

We write $\sqrt{nh}\{\widehat{\upsilon}(t_0;\hat{\gamma})-\upsilon(t_0)\}$ and $\sqrt{nh}\{\widehat{\varphi}(t_0;\hat{\beta},\hat{\gamma})-\varphi(t_0)\}$ into a summation of asymptotic iid terms,

$$\sqrt{nh}\{\widehat{\upsilon}(t_0;\hat{\gamma})-\upsilon(t_0)\}=\sum_{i=1}^{n}\overline{e}_{n,i}(t_0)\overline{B}^{-1}(t_0)+\sum_{i=1}^{n}\overline{f}_{n,i}(t_0)\overline{B}^{-1}(t_0)+o_p(1),$$

and

$$\begin{aligned}\sqrt{nh}\{\widehat{\varphi}(t_0;\hat{\beta},\hat{\gamma})-\varphi(t_0)\}&=\sqrt{nh}\{A_n(t_0;\beta_0,\gamma_0)-B_n(t_0)\varphi(t_0)\}B_n^{-1}(t_0)-\sqrt{nh}C_n(t_0)B_n^{-1}(t_0)\\&=\sum_{i=1}^{n}a_{n,i}(t_0)B^{-1}(t_0)-\sum_{i=1}^{n}c_{n,i}(t_0)B^{-1}(t_0)+o_p(1),\end{aligned}$$

where

$$\begin{aligned}a_{n,i}(t_0)&=\sqrt{\frac{h}{n}}\{\exp(-\beta_0^TZ_i-\gamma_0^TU_i)-\varphi(t_0)\}\Delta_iK_h(t_0-X_i),\\c_{n,i}(t_0)&=\sqrt{\frac{h}{n}}\tilde{l}_\theta(\mathbf{X_i})^TD(t_0;\theta_0).\end{aligned}$$

Under finite sampling, we regard the bandwidth h as a known constant, rather than letting it go to zero. Let

$$\begin{aligned}\sigma_C^2(t_0;h)&=hD(t_0;\theta_0)^TI^{-1}(\theta_0)D(t_0;\theta_0),\\\sigma_{AC}(t_0;h)&=hD(t_0;\theta_0)^TE\left[\Delta_iK_h(t_0-X_i)\{\exp(-\beta_0^TZ_i-\gamma_0^TU_i)-\varphi(t_0)\}\tilde{l}_\theta(\mathbf{X_i})\right],\\\sigma_{\overline{F}}^2(t_0;h)&=h\overline{G}(t_0;\gamma_0)^TI^{-1}(\gamma_0)\overline{G}(t_0;\gamma_0),\\\sigma_{\overline{EF}}(t_0;h)&=h\overline{G}(t_0;\gamma_0)^TE\left[(1-\Delta_i)K_h(t_0-X_i)\{\exp(\gamma_0^TU_i)-\upsilon(t_0)\}\tilde{l}_\gamma(\mathbf{X_i})\right],\\\sigma_{A\overline{F}}(t_0;h)&=h\overline{G}(t_0;\gamma_0)^TE\left[\Delta_iK_h(t_0-X_i)\{\exp(-\beta_0^TZ_i-\gamma_0^TU_i)-\varphi(t_0)\}\tilde{l}_\gamma(\mathbf{X_i})\right],\\\sigma_{C\overline{E}}(t_0;h)&=hD(t_0;\theta_0)^TE\left[(1-\Delta_i)K_h(t_0-X_i)\{\exp(\gamma_0^TU_i)-\upsilon(t_0)\}\tilde{l}_\theta(\mathbf{X_i})\right],\\\sigma_{C\overline{F}}(t_0;h)&=hE\left[\{\tilde{l}_\theta(\mathbf{X_i})^TD(t_0;\theta_0)\}\{\tilde{l}_\gamma(\mathbf{X_i})^T\overline{G}(t_0;\gamma_0)\}\right].\end{aligned}$$

Thus the limiting variance of $\sqrt{nh}\{\widehat{\phi}_{adj+}(t_0;\hat{\beta},\hat{\gamma})-\phi(t_0)\}$ under finite sampling is

$$\begin{aligned}\sigma_+^2(t_0;h) &= \frac{\sigma_A^2(t_0)\upsilon^2(t_0)}{B^2(t_0)} + \frac{\sigma_C^2(t_0)\upsilon^2(t_0)}{B^2(t_0)} - \frac{2\sigma_{AC}(t_0)\upsilon^2(t_0)}{B^2(t_0)}\\ &+ \frac{\sigma_{\overline{E}}^2(t_0)\varphi^2(t_0)}{\overline{B}^2(t_0)} + \frac{2\sigma_{\overline{EF}}(t_0)\varphi^2(t_0)}{\overline{B}^2(t_0)} + \frac{\sigma_{\overline{F}}^2(t_0)\varphi^2(t_0)}{\overline{B}^2(t_0)}\\ &+ \frac{2\sigma_{A\overline{F}}(t_0)\upsilon(t_0)\varphi(t_0)}{B(t_0)\overline{B}(t_0)} - \frac{2\sigma_{C\overline{E}}(t_0)\upsilon(t_0)\varphi(t_0)}{B(t_0)\overline{B}(t_0)} - \frac{2\sigma_{C\overline{F}}(t_0)\upsilon(t_0)\varphi(t_0)}{B(t_0)\overline{B}(t_0)}.\end{aligned}$$

An estimator of $\sigma_+^2(t_0;h)$ could be obtained by plugging in the following estimators:

$$\begin{aligned}\widehat{\sigma}_A^2(t) &= \frac{h}{n}\sum_{i=1}^n \Delta_i K_h(t-X_i)^2\left[\exp\{-\hat{\beta}^T Z_i(t)-\hat{\gamma}^T U_i(t)\}-\widehat{\varphi}(t;\hat{\beta},\hat{\gamma})\right]^2,\\ \widehat{\sigma}_C^2(t;h) &= hD_n(t;\hat{\theta})^T\widehat{I}^{-1}(\hat{\theta})D_n(t;\hat{\theta}),\\ \widehat{\sigma}_{AC}(t;h) &= \frac{h}{n}\sum_{i=1}^n \Delta_i K_h(t-X_i)\left[\exp\{-\hat{\beta}^T Z_i(t)-\hat{\gamma}^T U_i(t)\}-\widehat{\varphi}(t;\hat{\beta},\hat{\gamma})\right]\widehat{\tilde{l}}_\theta(\mathbf{X_i})^T D_n(t;\hat{\beta}),\\ \widehat{\sigma}_{\overline{E}}^2(t) &= \frac{h}{n}\sum_{i=1}^n (1-\Delta_i)K_h(t-X_i)^2\left[\exp\{\hat{\gamma}^T U_i(t)\}-\widehat{\upsilon}(t;\hat{\gamma})\right]^2,\\ \widehat{\sigma}_{\overline{F}}^2(t;h) &= h\overline{G}_n(t;\hat{\gamma})^T\widehat{I}^{-1}(\hat{\gamma})\overline{G}_n(t;\hat{\gamma}),\\ \widehat{\sigma}_{\overline{EF}}(t;h) &= \frac{h}{n}\sum_{i=1}^n (1-\Delta_i)K_h(t-X_i)\left[\exp\{\hat{\gamma}^T U_i(t)\}-\widehat{\upsilon}(t;\hat{\gamma})\right]\widehat{\tilde{l}}_\gamma(\mathbf{X_i})^T\overline{G}_n(t;\hat{\gamma}),\\ \widehat{\sigma}_{A\overline{F}}(t;h) &= \frac{h}{n}\sum_{i=1}^n \Delta_i K_h(t-X_i)\left[\exp\{-\hat{\beta}^T Z_i(t)-\hat{\gamma}^T U_i(t)\}-\widehat{\varphi}(t;\hat{\beta},\hat{\gamma})\right]\widehat{\tilde{l}}_\gamma(\mathbf{X_i})^T\overline{G}_n(t;\hat{\gamma}),\\ \widehat{\sigma}_{C\overline{E}}(t;h) &= \frac{h}{n}\sum_{i=1}^n (1-\Delta_i)K_h(t-X_i)\left[\exp\{\hat{\gamma}^T U_i(t)\}-\widehat{\upsilon}(t;\hat{\gamma})\right]\widehat{\tilde{l}}_\theta(\mathbf{X_i})^T D_n(t;\hat{\beta}),\\ \widehat{\sigma}_{C\overline{F}}(t;h) &= \frac{h}{n}\sum_{i=1}^n \left\{\widehat{\tilde{l}}_\theta(\mathbf{X_i})^T D_n(t;\hat{\theta})\right\}\left\{\widehat{\tilde{l}}_\gamma(\mathbf{X_i})^T\overline{G}_n(t;\widehat{\gamma})\right\}.\end{aligned}$$

So that

$$\begin{aligned}\widehat{\sigma}_+^2(t_0;h) &= \frac{\widehat{\sigma}_A^2(t_0)\widehat{\upsilon}^2(t_0;\hat{\gamma})}{B_n^2(t_0)} + \frac{\widehat{\sigma}_C^2(t_0;h)\widehat{\upsilon}^2(t_0;\hat{\gamma})}{B_n^2(t_0)} - \frac{2\widehat{\sigma}_{AC}(t_0;h)\widehat{\upsilon}^2(t_0;\hat{\gamma})}{B_n^2(t_0)}\\ &+ \frac{\widehat{\sigma}_{\overline{E}}^2(t_0)\widehat{\varphi}^2(t_0;\hat{\beta},\hat{\gamma})}{\overline{B}_n^2(t_0)} + \frac{\widehat{\sigma}_{\overline{F}}^2(t_0;h)\widehat{\varphi}^2(t_0;\hat{\beta},\hat{\gamma})}{\overline{B}_n^2(t_0)} + \frac{2\widehat{\sigma}_{\overline{EF}}(t_0;h)\widehat{\varphi}^2(t_0;\hat{\beta},\hat{\gamma})}{\overline{B}_n^2(t_0)}\\ &+ \frac{2\widehat{\sigma}_{A\overline{F}}(t_0;h)\widehat{\upsilon}(t_0;\hat{\gamma})\widehat{\varphi}(t_0;\hat{\beta},\hat{\gamma})}{B_n(t_0)\overline{B}_n(t_0)} - \frac{2\widehat{\sigma}_{C\overline{E}}(t_0;h)\widehat{\upsilon}(t_0;\hat{\gamma})\widehat{\varphi}(t_0;\hat{\beta},\hat{\gamma})}{B_n(t_0)\overline{B}_n(t_0)}\\ &- \frac{2\widehat{\sigma}_{C\overline{F}}(t_0;h)\widehat{\upsilon}(t_0;\hat{\gamma})\widehat{\varphi}(t_0;\hat{\beta},\hat{\gamma})}{B_n(t_0)\overline{B}_n(t_0)}.\end{aligned}$$

3.4.2 Variance estimates of $\widehat{\phi}_{adj-}(t;\hat{\beta};\hat{\gamma})$ with correction for finite sample

Similarly we can estimate the variance of $\widehat{\phi}_{adj-}(t;\hat{\beta};\hat{\gamma})$ with finite sampling correction. Recall that

$$\sqrt{nh}\{\widehat{\phi}_{adj-}(t_0;\hat{\beta},\hat{\gamma})-\phi(t_0)\}=\sqrt{nh}\{\overline{E}_n(t_0;\hat{\gamma})-\psi^{-1}(t_0)\upsilon(t_0)\overline{A}_n(t_0;\hat{\beta},\hat{\gamma})\}\overline{A}_n^{-1}(t_0;\hat{\beta},\hat{\gamma}).$$

Let

$$\begin{aligned}\widetilde{E}_n(t;\beta,\gamma)&=\overline{E}_n(t;\gamma)-\psi^{-1}(t)\upsilon(t)\overline{A}_n(t;\beta,\gamma)\\&=\frac{1}{n}\sum_{i=1}^{n}(1-\Delta_i)K_h(t-X_i)\left[\exp\{\gamma^TU_i(t)\}-\psi^{-1}(t)\upsilon(t)\exp\{\beta^TZ_i(t)+\gamma^TU_i(t)\}\right],\\ \widetilde{F}_n(t)&=\widetilde{E}_n(t;\hat{\beta},\hat{\gamma})-\widetilde{E}_n(t;\beta_0,\gamma_0)=\overline{F}_n(t)-\psi^{-1}(t)\upsilon(t)\overline{C}_n(t).\end{aligned}$$

Then

$$\begin{aligned}\sqrt{nh}\{\overline{E}_n(t_0;\hat{\gamma})-\psi^{-1}(t_0)\upsilon(t_0)\overline{A}_n(t_0;\hat{\beta},\hat{\gamma})\}&=\sqrt{nh}\widetilde{E}_n(t_0;\beta_0,\gamma_0)+\sqrt{nh}\widetilde{F}_n(t_0)\\&=\sum_{i=1}^{n}e_{n,i}(t_0)+\sum_{i=1}^{n}f_{n,i}(t_0)+o_p(1),\end{aligned}$$

where

$$\begin{aligned}e_{n,i}(t_0)&=\sqrt{\frac{h}{n}}(1-\Delta_i)K_h(t-X_i)\left[\exp\{\gamma_0^TU_i(t_0)\}-\psi^{-1}(t)\upsilon(t)\exp\{\beta_0^TZ_i(t_0)+\gamma_0^TU_i(t_0)\}\right],\\ f_{n,i}(t_0)&=\sqrt{\frac{h}{n}}\{\tilde{l}_\gamma(\mathbf{X_i})^T\overline{G}(t_0;\gamma_0)-\psi^{-1}(t_0)\upsilon(t_0)\tilde{l}_\theta(\mathbf{X_i})^T\overline{D}(t_0;\theta_0)\}.\end{aligned}$$

Therefore, the limiting variance of $\sqrt{nh}\{\overline{E}_n(t_0;\hat{\gamma})-\psi^{-1}(t_0)\upsilon(t_0)\overline{A}_n(t_0;\hat{\beta},\hat{\gamma})\}$ under finite sampling is $\sigma_E^2(t_0)+\sigma_F^2(t_0;h)+2\sigma_{EF}(t_0;h)$, where

$$\begin{aligned}\sigma_E^2(t)&=\overline{B}(t)\int K^2(u)du\iint\{\exp(\gamma_0^Tu)-\psi^{-1}(t)\upsilon(t)\exp(\beta_0^Tz+\gamma_0^Tu)\}^2f_{Z,U|T\geq t}(z,u)dzdu,\\ \sigma_F^2(t;h)&=hE\{\tilde{l}_\gamma(\mathbf{X_i})^T\overline{G}(t;\gamma_0)-\psi^{-1}(t)\upsilon(t)\tilde{l}_\theta(\mathbf{X_i})^T\overline{D}(t;\theta_0)\}^2,\\ \sigma_{EF}(t;h)&=nE\{e_{n,i}(t)f_{n,i}(t)\}.\end{aligned}$$

We could estimate these quantities with the following estimators

$$\widehat{\sigma}_E^2(t) = \frac{h}{n}\sum_{i=1}^{n}(1-\Delta_i)K_h(t-X_i)^2\left[\exp\{\hat{\gamma}^T U_i(t)\} - \widehat{\psi}^{-1}(t;\hat{\beta},\hat{\gamma})\widehat{\upsilon}(t;\hat{\gamma})\exp\{\hat{\beta}^T Z_i(t)+\hat{\gamma}^T U_i(t)\}\right]^2,$$

$$\widehat{\sigma}_F^2(t;h) = \frac{h}{n}\sum_{i=1}^{n}\{\widehat{\tilde{l}}_\gamma(\mathbf{X_i})^T\overline{G}_n(t;\hat{\gamma}) - \widehat{\psi}^{-1}(t;\hat{\beta},\hat{\gamma})\widehat{\upsilon}(t;\hat{\gamma})\widehat{\tilde{l}}_\theta(\mathbf{X_i})^T\overline{D}_n(t;\hat{\theta})\}^2,$$

$$\widehat{\sigma}_{EF}(t;h) = \frac{h}{n}\sum_{i=1}^{n}(1-\Delta_i)K_h(t-X_i)\left[\exp\{\hat{\gamma}^T U_i(t)\} - \widehat{\psi}^{-1}(t;\hat{\beta},\hat{\gamma})\widehat{\upsilon}(t;\hat{\gamma})\exp\{\hat{\beta}^T Z_i(t)+\hat{\gamma}^T U_i(t)\}\right]\cdot$$
$$\{\widehat{\tilde{l}}_\gamma(\mathbf{X_i})^T\overline{G}_n(t;\hat{\gamma}) - \widehat{\psi}^{-1}(t;\hat{\beta},\hat{\gamma})\widehat{\upsilon}(t;\hat{\gamma})\widehat{\tilde{l}}_\theta(\mathbf{X_i})^T\overline{D}_n(t;\hat{\theta})\}.$$

The limiting variance of $\sqrt{nh}\{\widehat{\phi}_{adj-}(t_0;\hat{\beta},\hat{\gamma})-\phi(t_0)\}$ is then

$$\sigma_-^2(t_0;h) = \frac{\sigma_E^2(t_0)+\sigma_F^2(t_0;h)+2\sigma_{EF}(t_0;h)}{\overline{A}^2(t_0)}, \tag{3.25}$$

and the corresponding estimator is

$$\widehat{\sigma}_-^2(t_0;h) = \frac{\widehat{\sigma}_E^2(t_0)+\widehat{\sigma}_F^2(t_0;h)+2\widehat{\sigma}_{EF}(t_0;h)}{\overline{A}_n^2(t_0;\hat{\beta},\hat{\gamma})} \tag{3.26}$$

3.4.3 Variance estimates of $\widehat{\phi}_{adjw}(t;\hat{\beta};\hat{\gamma})$ with correction for finite sample

To obtain the limiting variance of $\widehat{\phi}_{adjw}(t;\hat{\beta};\hat{\gamma})$ under finite sampling, take the bandwidth h as a known constant, and write

$$\begin{aligned}&\sqrt{nh}\{\widehat{\phi}_{adjw}(t_0;\hat{\beta},\hat{\gamma})-\phi(t_0)\}\\ =&\pi_0\sqrt{nh}\{\widehat{\upsilon}(t_0;\hat{\gamma})-\upsilon(t_0)\}\widehat{\varphi}(t_0;\hat{\beta},\hat{\gamma}) + \pi_0\sqrt{nh}\{\widehat{\varphi}(t_0;\hat{\beta},\hat{\gamma})-\varphi(t_0)\}\upsilon(t_0)\\ +&(1-\pi_0)\sqrt{nh}\{\widehat{\phi}_{adj-}(t_0;\hat{\beta},\hat{\gamma})-\phi(t_0)\}.\end{aligned}$$

Write the components into asymptotic iid forms

$$\sqrt{nh}\{\widehat{\upsilon}(t_0;\hat{\gamma})-\upsilon(t_0)\} = \sum_{i=1}^{n}\overline{e}_{n,i}(t_0)\overline{B}^{-1}(t_0) + \sum_{i=1}^{n}\overline{f}_{n,i}(t_0)\overline{B}^{-1}(t_0) + o_p(1),$$

$$\sqrt{nh}\{\widehat{\varphi}(t_0;\hat{\beta},\hat{\gamma})-\varphi(t_0)\} = \sum_{i=1}^{n}a_{n,i}(t_0)B^{-1}(t_0) - \sum_{i=1}^{n}c_{n,i}(t_0)B^{-1}(t_0) + o_p(1),$$

$$\sqrt{nh}\{\widehat{\phi}_{adj-}(t_0;\hat{\beta},\hat{\gamma})-\phi(t_0)\} = \sum_{i=1}^{n}e_{n,i}(t_0)\overline{A}^{-1}(t_0;\beta_0,\gamma_0) + \sum_{i=1}^{n}f_{n,i}(t_0)\overline{A}^{-1}(t_0;\beta_0,\gamma_0) + o_p(1).$$

The variances of these terms, and covariance between the first two terms have been studied in the variance under finite sampling of $\widehat{\phi}_{adj+}(t;\hat{\beta};\hat{\gamma})$ and $\widehat{\phi}_{adj-}(t;\hat{\beta};\hat{\gamma})$. The covariance between $\sqrt{nh}\{\widehat{\upsilon}(t_0;\hat{\gamma})-\upsilon(t_0)\}$ and $\sqrt{nh}\{\widehat{\phi}_{adj-}(t_0;\hat{\beta},\hat{\gamma})-\phi(t_0)\}$ is

$$\begin{aligned}&\sigma_{E\overline{E}}(t_0)\overline{A}^{-1}(t_0;\beta_0,\gamma_0)\overline{B}^{-1}(t_0)+\sigma_{E\overline{F}}(t_0)\overline{A}^{-1}(t_0;\beta_0,\gamma_0)\overline{B}^{-1}(t_0)\\&+\sigma_{\overline{E}F}(t_0)\overline{A}^{-1}(t_0;\beta_0,\gamma_0)\overline{B}^{-1}(t_0)+\sigma_{F\overline{F}}(t_0)\overline{A}^{-1}(t_0;\beta_0,\gamma_0)\overline{B}^{-1}(t_0),\end{aligned}$$

and the covariance between $\sqrt{nh}\{\widehat{\varphi}(t_0;\hat{\beta},\hat{\gamma})-\varphi(t_0)\}$ and $\sqrt{nh}\{\widehat{\phi}_{adj-}(t_0;\hat{\beta},\hat{\gamma})-\phi(t_0)\}$ is

$$\begin{aligned}&\sigma_{AF}(t_0)\overline{A}^{-1}(t_0;\beta_0,\gamma_0)B^{-1}(t_0)-\sigma_{CE}(t_0)\overline{A}^{-1}(t_0;\beta_0,\gamma_0)B^{-1}(t_0)\\&-\sigma_{CF}(t_0)\overline{A}^{-1}(t_0;\beta_0,\gamma_0)B^{-1}(t_0),\end{aligned}$$

where

$$\begin{aligned}\sigma_{E\overline{E}}(t)&=nE\{e_{n,i}(t)\overline{e}_{n,i}(t)\},\\\sigma_{E\overline{F}}(t)&=nE\{e_{n,i}(t)\overline{f}_{n,i}(t)\},\\\sigma_{\overline{E}F}(t)&=nE\{\overline{e}_{n,i}(t)f_{n,i}(t)\},\\\sigma_{F\overline{F}}(t)&=nE\{f_{n,i}(t)\overline{f}_{n,i}(t)\},\\\sigma_{AF}(t)&=nE\{a_{n,i}(t)f_{n,i}(t)\},\\\sigma_{CE}(t)&=nE\{c_{n,i}(t)e_{n,i}(t)\},\\\sigma_{CF}(t)&=nE\{c_{n,i}(t)f_{n,i}(t)\}.\end{aligned}$$

So that the limiting variance of $\sqrt{nh}\{\widehat{\phi}_{adjw}(t_0;\hat{\beta},\hat{\gamma})-\phi(t_0)\}$ under finite sampling is

$$\sigma^2_{adjw}(t_0)=\pi_0^2\sigma^2_{adj+}(t_0)+2\pi_0(1-\pi_0)\sigma_{adj+-}(t_0)+(1-\pi_0)^2\sigma^2_{adj-}(t_0),$$

where

$$\begin{aligned}\sigma_{adj+-}(t_0)=\;&\sigma_{E\overline{E}}(t_0)\overline{A}^{-1}(t_0;\beta_0,\gamma_0)\overline{B}^{-1}(t_0)+\sigma_{E\overline{F}}(t_0)\overline{A}^{-1}(t_0;\beta_0,\gamma_0)\overline{B}^{-1}(t_0)\\&+\sigma_{\overline{E}F}(t_0)\overline{A}^{-1}(t_0;\beta_0,\gamma_0)\overline{B}^{-1}(t_0)+\sigma_{F\overline{F}}(t_0)\overline{A}^{-1}(t_0;\beta_0,\gamma_0)\overline{B}^{-1}(t_0)\\&+\sigma_{AF}(t_0)\overline{A}^{-1}(t_0;\beta_0,\gamma_0)B^{-1}(t_0)-\sigma_{CE}(t_0)\overline{A}^{-1}(t_0;\beta_0,\gamma_0)B^{-1}(t_0)\\&-\sigma_{CF}(t_0)\overline{A}^{-1}(t_0;\beta_0,\gamma_0)B^{-1}(t_0).\end{aligned}$$

We propose the following estimators for these terms:

$$\begin{aligned}
\widehat{\sigma}_{E\overline{E}}(t) =& \frac{h}{n}\sum_{i=1}^{n}(1-\Delta_i)K_h(t-X_i)^2\left[\exp\{\hat{\gamma}^T U_i(t)\} - \widehat{\psi}^{-1}(t;\hat{\beta},\hat{\gamma})\widehat{v}(t;\hat{\gamma})\exp\{\hat{\beta}^T Z_i(t)+\hat{\gamma}^T U_i(t)\}\right]\cdot \\
&\left[\exp\{\hat{\gamma}^T U_i(t)\} - \widehat{v}(t;\hat{\gamma})\right], \\
\widehat{\sigma}_{E\overline{F}}(t;h) =& \frac{h}{n}\sum_{i=1}^{n}(1-\Delta_i)K_h(t-X_i)\left[\exp\{\hat{\gamma}^T U_i(t)\} - \widehat{\psi}^{-1}(t;\hat{\beta},\hat{\gamma})\widehat{v}(t;\hat{\gamma})\exp\{\hat{\beta}^T Z_i(t)+\hat{\gamma}^T U_i(t)\}\right]\cdot \\
&\{\widehat{\tilde{l}}_\gamma(\mathbf{X_i})^T\overline{G}_n(t;\hat{\gamma})\}, \\
\widehat{\sigma}_{\overline{E}F}(t;h) =& \frac{h}{n}\sum_{i=1}^{n}(1-\Delta_i)K_h(t-X_i)\left[\exp\{\hat{\gamma}^T U_i(t)\} - \widehat{v}(t;\hat{\gamma})\right]\cdot \\
&\{\widehat{\tilde{l}}_\gamma(\mathbf{X_i})^T\overline{G}_n(t;\hat{\gamma}) - \widehat{\psi}^{-1}(t;\hat{\beta},\hat{\gamma})\widehat{v}(t;\hat{\gamma})\widehat{\tilde{l}}_\theta(\mathbf{X_i})^T\overline{D}_n(t;\hat{\theta})\}, \\
\widehat{\sigma}_{F\overline{F}}(t;h) =& \{\widehat{\tilde{l}}_\gamma(\mathbf{X_i})^T\overline{G}_n(t;\hat{\gamma}) - \widehat{\psi}^{-1}(t;\hat{\beta},\hat{\gamma})\widehat{v}(t;\hat{\gamma})\widehat{\tilde{l}}_\theta(\mathbf{X_i})^T\overline{D}_n(t;\hat{\theta})\}\{\widehat{\tilde{l}}_\gamma(\mathbf{X_i})^T\overline{G}_n(t;\hat{\gamma})\}, \\
\widehat{\sigma}_{AF}(t;h) =& \frac{h}{n}\sum_{i=1}^{n}\Delta_i K_h(t-X_i)\left[\exp\{-\hat{\beta}^T Z_i(t) - \hat{\gamma}^T U_i(t)\} - \widehat{\varphi}(t;\hat{\beta},\hat{\gamma})\right]\cdot \\
&\{\widehat{\tilde{l}}_\gamma(\mathbf{X_i})^T\overline{G}_n(t;\hat{\gamma}) - \widehat{\psi}^{-1}(t;\hat{\beta},\hat{\gamma})\widehat{v}(t;\hat{\gamma})\widehat{\tilde{l}}_\theta(\mathbf{X_i})^T\overline{D}_n(t;\hat{\theta})\}, \\
\widehat{\sigma}_{CE}(t;h) =& \frac{h}{n}\sum_{i=1}^{n}(1-\Delta_i)K_h(t-X_i)\left[\exp\{\hat{\gamma}^T U_i(t)\} - \widehat{\psi}^{-1}(t;\hat{\beta},\hat{\gamma})\widehat{v}(t;\hat{\gamma})\exp\{\hat{\beta}^T Z_i(t)+\hat{\gamma}^T U_i(t)\}\right]\cdot \\
&\big\{\widehat{\tilde{l}}_\theta(\mathbf{X_i})^T D_n(t;\hat{\theta})\big\}, \\
\widehat{\sigma}_{CF}(t;h) =& \frac{h}{n}\sum_{i=1}^{n}\big\{\widehat{\tilde{l}}_\theta(\mathbf{X_i})^T D_n(t;\hat{\theta})\big\}\{\widehat{\tilde{l}}_\gamma(\mathbf{X_i})^T\overline{G}_n(t;\hat{\gamma}) - \widehat{\psi}^{-1}(t;\hat{\beta},\hat{\gamma})\widehat{v}(t;\hat{\gamma})\widehat{\tilde{l}}_\theta(\mathbf{X_i})^T\overline{D}_n(t;\hat{\theta})\},
\end{aligned}$$

and hence

$$\widehat{\sigma}^2_{adjw}(t_0;h) = \pi_0^2\widehat{\sigma}^2_{adj+}(t_0;h) + 2\pi_0(1-\pi_0)\widehat{\sigma}_{adj+-}(t_0;h) + (1-\pi_0)^2\widehat{\sigma}^2_{adj-}(t_0;h),$$

where

$$\begin{aligned}
\widehat{\sigma}_{adj+-}(t_0;h) =& \widehat{\sigma}_{E\overline{E}}(t_0;h)\overline{A}_n^{-1}(t_0;\hat{\beta},\hat{\gamma})\overline{B}_n^{-1}(t_0) + \widehat{\sigma}_{E\overline{F}}(t_0;h)\overline{A}_n^{-1}(t_0;\hat{\beta},\hat{\gamma})\overline{B}_n^{-1}(t_0) \\
&+ \widehat{\sigma}_{\overline{E}F}(t_0;h)\overline{A}_n^{-1}(t_0;\hat{\beta},\hat{\gamma})\overline{B}_n^{-1}(t_0) + \widehat{\sigma}_{F\overline{F}}(t_0;h)\overline{A}_n^{-1}(t_0;\hat{\beta},\hat{\gamma})\overline{B}_n^{-1}(t_0) \\
&+ \widehat{\sigma}_{AF}(t_0;h)\overline{A}_n^{-1}(t_0;\hat{\beta},\hat{\gamma})B_n^{-1}(t_0) - \widehat{\sigma}_{CE}(t_0;h)\overline{A}_n^{-1}(t_0;\hat{\beta},\hat{\gamma})B_n^{-1}(t_0) \\
&- \widehat{\sigma}_{CF}(t_0;h)\overline{A}_n^{-1}(t_0;\hat{\beta},\hat{\gamma})B_n^{-1}(t_0).
\end{aligned}$$

Chapter 4

SIMULATION STUDIES

In this chapter we conduct simulation studies to evaluate the finite sample performance of the proposed statistical procedures in Chapter 2 and Chapter 3. We consider the frequency matched case-control studies with univariate risk factors. Each simulation consists of generating a target population, sampling case-control data from the target population, and calculating the point and variance estimators as well as the simultaneous confidence bands using the case-control data. In our simulations, the kernel function is taken to be Epanechnikov kernel, i.e., $K(x) = 0.75(1-x^2)I_{|x|<1}$. The bandwidth will be automatically selected from the proposed cross-validation approach, and fixed bandwidths from a broad range will be used to evaluate the performance of the automatic bandwidth selector. The details of the simulation studies are described below.

4.1 Unadjusted Time-varying PAF

The simulation studies in this section are to evaluate the estimators proposed in Chapter 2. We consider the following Cox model:

$$\lambda\{t|Z(t)\} = \lambda_0(t)\exp\{\beta Z(t)\}, \tag{4.1}$$

where $Z(t)$ is a univariate possibly time-dependent covariate, β is the regression parameter, and the baseline hazard $\lambda_0(t)$ is assumed to be a Weibull hazard:

$$\lambda_0(t) = \frac{\nu}{\eta}\left(\frac{t}{\eta}\right)^{\nu-1}.$$

We consider both time-independent covariate, $Z(t) = Z$, and time-dependent covariate, $Z(t) = cZt$, where c is a known parameter larger than zero. For time-independent covariate, the true $\phi(t)$ is calculated using the following formulas:

$$\begin{aligned}
\phi(t) &= \frac{\lambda(t)-\lambda_0(t)}{\lambda(t)} = 1 - \frac{\lambda_0(t)S(t)}{f_T(t)}, \\
S(t) &= \int_Z S(t|z)dF_Z(z) = \int_Z \exp\{-\Lambda(t|z)\}dF_Z(z) = \int_Z \exp\{-\Lambda_0(t)\exp(\beta z)\}dF_Z(z), \\
f_T(t) &= -dS(t)/dt = \lambda_0(t)\int_Z \exp(\beta^T z)\exp\{-\Lambda_0(t)\exp(\beta z)\}dF_Z(z),
\end{aligned}$$

with pre-specified ν, η, β, and $f_Z(z)$. For time-dependent covariate, the true $\phi(t)$ is calculated using the following formulas:

$$\begin{aligned}
\phi(t) &= \frac{\lambda(t)-\lambda_0(t)}{\lambda(t)} = 1 - \frac{\lambda_0(t)S(t)}{f_T(t)}, \\
S(t) &= \int_Z S(t|z)dF_Z(z) = \int_Z \exp\Big\{-\int_0^t \lambda_0(s)\exp(\beta czs)ds\Big\}dF_Z(z), \\
f_T(t) &= -dS(t)/dt = \lambda_0(t)\int_Z \exp(\beta czt)\exp\Big\{-\int_0^t \lambda_0(s)\exp(\beta czs)ds\Big\}dF_Z(z).
\end{aligned}$$

Each simulation consisted of generating a population of interest, sampling case-control data from the population, and calculating the estimators using the case-control data. A population with 20,000 subjects was first generated. For each subject, a univariate covariate Z was generated from a pre-specified distribution, the failure time T was generated based on model (4.1), and the censoring time C was independently generated from a truncated normal distribution in $[1, 100]$. The observed age X was set to be the minimum of T and C, and the binary disease status $\Delta = 1$ if $T \leq C$ and $\Delta = 0$ if $T > C$. The case-control sample was obtained by randomly sampling 1000 cases ($\Delta = 1$) from the population, and 1000 controls with observed age matched with that of the cases within five-year intervals (frequency match). The data for analysis included X, Z and Δ, while the failure time and the censoring time were masked. Then $\widehat{\phi}_+(t;\hat{\beta})$, $\widehat{\phi}_-(t;\hat{\beta})$ and their corresponding variance estimators were obtained using the procedures in Chapter 2.

The above simulation procedure was repeated for 2,000 times, and produced 2,000 pairs of point and variance estimates for each of $\widehat{\phi}_+(t;\hat{\beta})$ and $\widehat{\phi}_-(t;\hat{\beta})$. The mean estimates were calculated by taking the average of the 2,000 point estimates, the mean variance estimates were calculated as the average of the 2,000 variance estimates, and empirical variance estimates were taken as the sample variance of the 2,000 point estimates. The bias was calculated by taking the absolute difference between the mean estimates and the theoretical $\phi(t)$.

We consider four scenarios: 1) early onset common disease with binary exposure; 2) late onset rare disease with binary exposure; 3) late onset rare disease with continuous exposure; 4) late onset rare disease with continuous time-dependent exposure.

4.1.1 Early onset common disease with binary time-independent exposure

Simulation Setting. This scenario was to mimic the early onset common diseases such as infectious diseases in high-risk populations. We took 1,000 cases and 1,000 controls with age frequency matched with five-year intervals. The baseline hazard parameters were set as $\eta = 40$, $\nu = 2$. The covariate Z was Bernoulli with success probability of $P_z = 0.3$ and $P_z = 0.6$. The odds ratio parameter was set as $\beta_0 = \log 2$ and $\log 3$. Censoring time was truncated normal between 1 and 100 with standard deviation of 20, and the mean was set to yield 30% censoring probability. These settings yielded an average age at onset of about 26 years.

Simulation Results. Figure 4.1 shows the true $\phi(t)$ curves in solid lines and the mean estimated curves of $\widehat{\phi}_+(t;\hat{\beta})$ and the averages of their corresponding pointwise 95% confidence intervals in dashed and dotted lines, respectively. For each simulation, the cross-validation bandwidth is used for computing the estimated curves. The true $\phi(t)$ is a monotone decreasing function on $[0, 1]$. For fixed odds ratio, when the prevalence of exposure (P_z) is higher, the $\phi(t)$ curve shifts higher with similar decreasing rate. For fixed prevalence of exposure, when the odds ratio is higher, the $\phi(t)$ curve takes a higher initial value and has a steeper decrease. The curves of $\widehat{\phi}_w(t;\hat{\beta})$ generally overlap with the true $\phi(t)$ curve, indicating that these estimators and cross-validation bandwidths give generally satisfactory results.

The simulation results for $\widehat{\phi}_w(t;\hat{\beta})$ with various bandwidth choices are summarized in Table 4.1 and Table 4.2. h_{CV} denotes cross-validation bandwidth. The results include ages ranged from 10 to 50. The estimates out of this age range are not stable because of small sample, hence are not shown. Comparing the results across different bandwidths, when a smaller bandwidth is used, the biases tend to be smaller while the variances tend to be larger, and when a larger bandwidth is used, the biases tend to be larger, especially at late ages, and the variances tend to be smaller. The cross-validation bandwidths balance the bias and variance, and yield satisfactory estimates. We also want to point out that while we observe trends of bias and variance in association of bandwidth selection, the differences are generally minor. This suggests that the proposed estimators are fairly robust over a wide range of bandwidths in our settings. The standard error estimators are close to the sampling standard deviations, and the coverage of 95% pointwise confidence intervals are close to 95%, suggesting that the variance estimators perform well.

Results for $\widehat{\phi}_+(t;\hat{\beta})$, $\widehat{\phi}_-(t;\hat{\beta})$, and $\widehat{\phi}_w(t;\hat{\beta})$ with cross-validation bandwidth are listed in Table 4.3. The three estimators all have satisfactory control on bias. Also as we expected, the combined estimator ($\widehat{\phi}_w(t;\hat{\beta})$) has smallest variances on almost all time points, followed by $\widehat{\phi}_+(t;\hat{\beta})$, and $\widehat{\phi}_-(t;\hat{\beta})$ has the worst performance.

4.1.2 Late onset rare disease with binary time-independent exposure

Simulation Setting. This scenario was to mimic the late onset rare diseases such as chronic diseases or cancers. We took 1,000 cases and 1,000 controls with age frequency matched with five-year intervals. The baseline hazard parameters were set as $\eta = 70$, $\nu = 9$. The covariate Z was Bernoulli with success probability of $P_z = 0.3$ and $P_z = 0.6$. The odds ratio parameter was set as $\beta_0 = \log 2$ and $\log 3$. Censoring time was truncated normal between 1 and 100 with standard deviation of 20, and the mean was set to yield 70% censoring probability. These settings yielded an average age at onset of about 60 years.

Simulation Results. Figure 4.2 shows the true $\phi(t)$ curves in solid lines and the mean es-

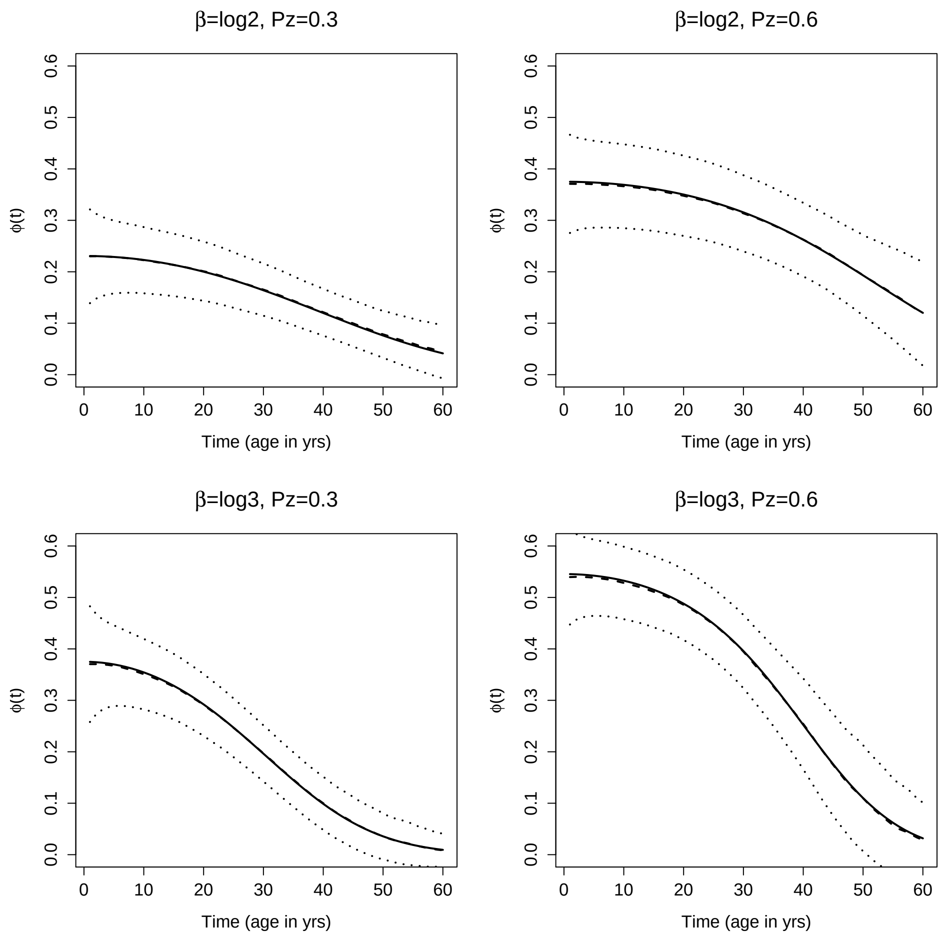

Figure 4.1: Simulation results with cross-validation based bandwidth for early onset disease scenario with binary exposure. The theoretical function $\phi(t)$ (solid lines), the mean estimated curve of $\phi(t)$ (dash lines), and the averages of their 95% pointwise confidence intervals (dotted lines).

timated curves of $\widehat{\phi}_+(t;\hat{\beta})$ and the averages of their corresponding pointwise 95% confidence intervals in dashed and dotted lines. For each simulation, the cross-validation bandwidth is used for computing the estimated curves. The true $\phi(t)$ is a monotone non-increasing function on $[0, 1]$. It remains flat at the early stage, and starts to drop at age of 50-60 years and then converges to zero around age of 80 years. Same as in the early onset disease example, $\phi(t)$ takes higher values with higher prevalence of exposure or larger odds ratio. The curves of $\widehat{\phi}_+(t;\hat{\beta})$ generally overlap with the true $\phi(t)$ curve, indicating that these estimators and cross-validation bandwidths give satisfactory results.

The simulation results for $\widehat{\phi}_w(t;\hat{\beta})$ with various bandwidth choices are summarized in Table 4.4 and Table 4.5. These results shows the same pattern as in the early onset example. The point and variance estimators both give satisfactory results.

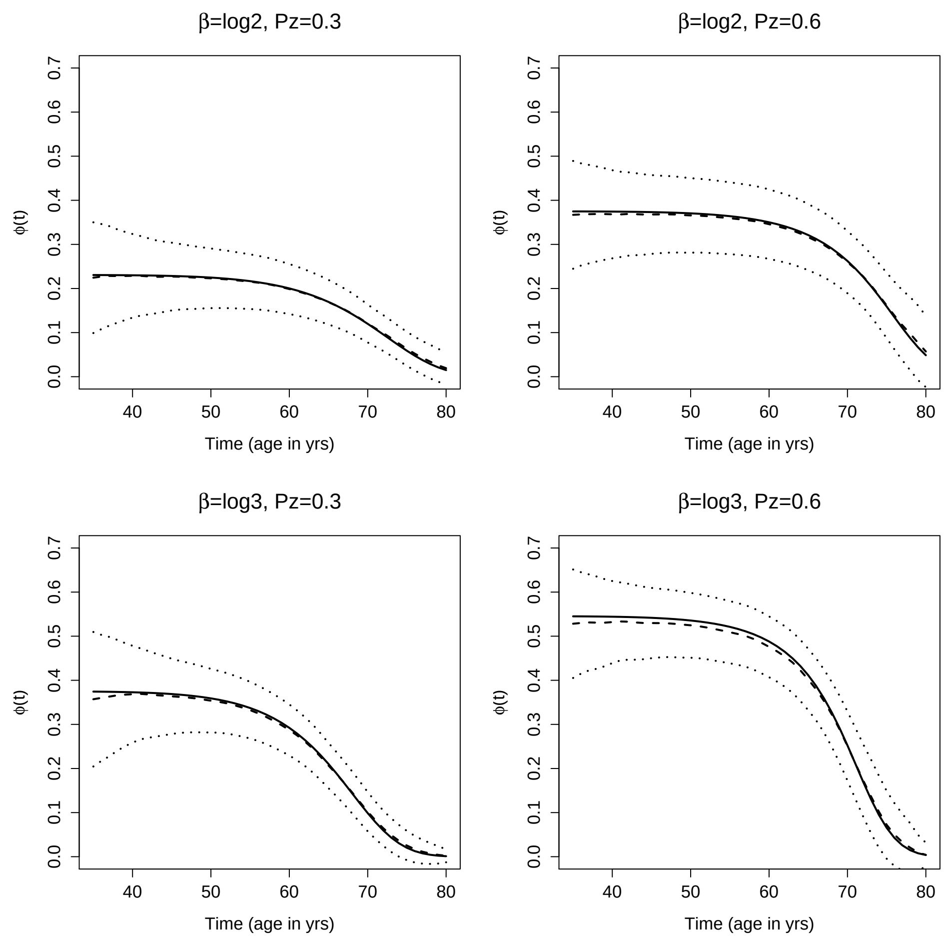

Figure 4.2: Simulation results with cross-validation based bandwidth for late onset disease scenario with binary exposure. The theoretical function $\phi(t)$ (solid lines), the mean estimated curve of $\phi(t)$ (dash lines), and the averages of their 95% pointwise confidence intervals (dotted lines).

4.1.3 Late onset rare disease with continuous time-independent exposure

Simulation Setting. This scenario was to show how the proposed estimators work with continuous covariate. We took 1000 cases and 1000 controls with age frequency matched with five-year intervals. The baseline hazard parameters were set as $\eta = 70$, $\nu = 9$. The covariate Z was uniform on $[0, 4]$. The odds ratio parameter was set as $\beta_0 = \log(1.3)$ and $\log(1.6)$. Censoring time was truncated normal between 1 and 100 with standard deviation of 25, and the mean was set to yield 70% censoring probability. These settings yielded an average age at onset of about 58 years.

Simulation Results. The results are shown in Figure 4.3 and Table 4.6 and Table 4.7. From these results we see that the point and variance estimators both give satisfactory results.

4.1.4 Late onset rare disease with continuous time-dependent exposure

Simulation Setting. This scenario was to show how the proposed estimators work with time-dependent covariates. We took 1000 cases and 1000 controls with age frequency matched with five-year intervals. The baseline hazard parameters were set as $\eta = 70$, $\nu = 9$. The time-dependent covariate was set as $Z(t) = 0.02Zt$ with Z be uniform $[0, 4]$. The odds ratio parameter was set as $\beta_0 = \log(1.3)$ and $\log(1.6)$. Censoring time was truncated normal between 1 and 100 with standard deviation of 25, and the mean was set to yield 70% censoring probability. These settings yielded an average age at onset of about 58 years.

Simulation Results. The results are shown in Figure 4.4 and Table 4.8 and Table 4.9. From these results we see that the point and variance estimators both give satisfactory results. It is worth noting that $\phi(t)$ is a concave function rather than a monotone function.

4.1.5 Simulation results for simultaneous confidence bands

We conducted extensive simulations to evaluate the performance of the proposed methods for calculating the simultaneous confidence bands. The results for the late onset disease

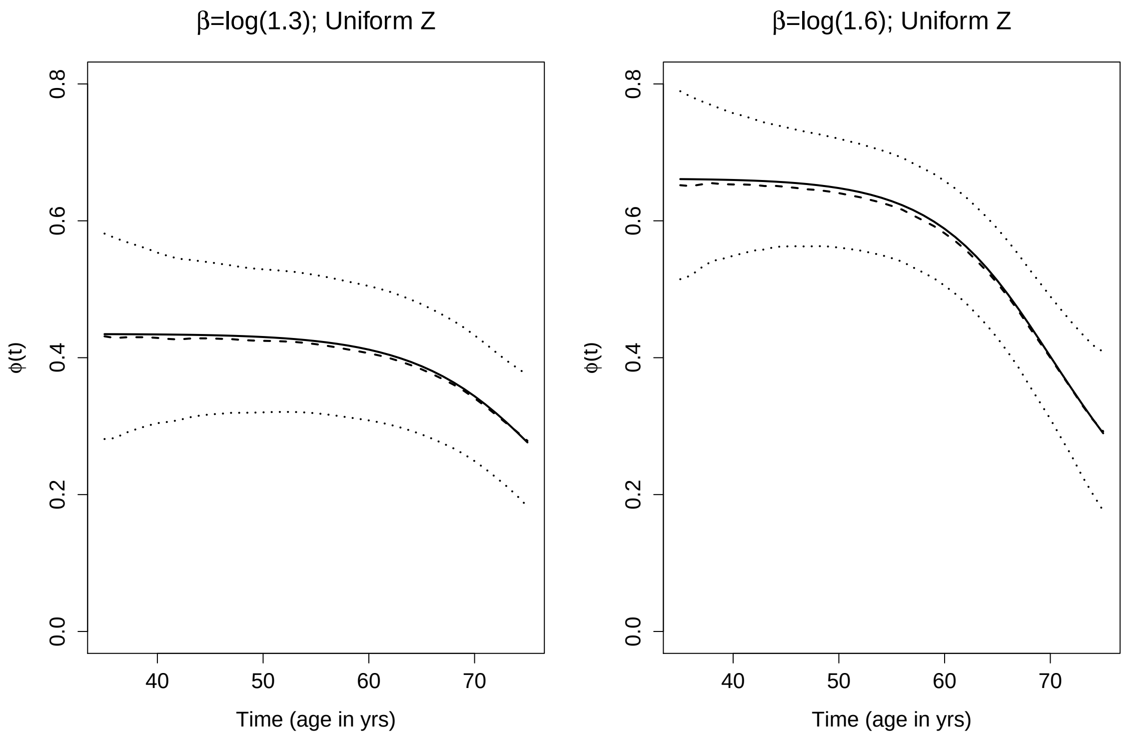

Figure 4.3: Simulation results with cross-validation based bandwidth for late onset disease scenario with continuous exposure. The theoretical function $\phi(t)$ (solid lines), the mean estimated curve of $\phi(t)$ (dash lines), and the averages of their 95% pointwise confidence intervals (dotted lines).

with continuous exposure are shown in Table 4.10. The results for other scenarios are similar. These results suggests the accuracy of the proposed confidence bands estimators.

4.2 Adjusted Time-varying PAF

In this section we evaluate the estimators proposed in Chapter 3 via simulations. We consider the following Cox model:

$$\lambda\{t|Z(t), U(t)\} = \lambda_0(t) \exp\{\beta Z(t) + \gamma U(t)\}, \tag{4.2}$$

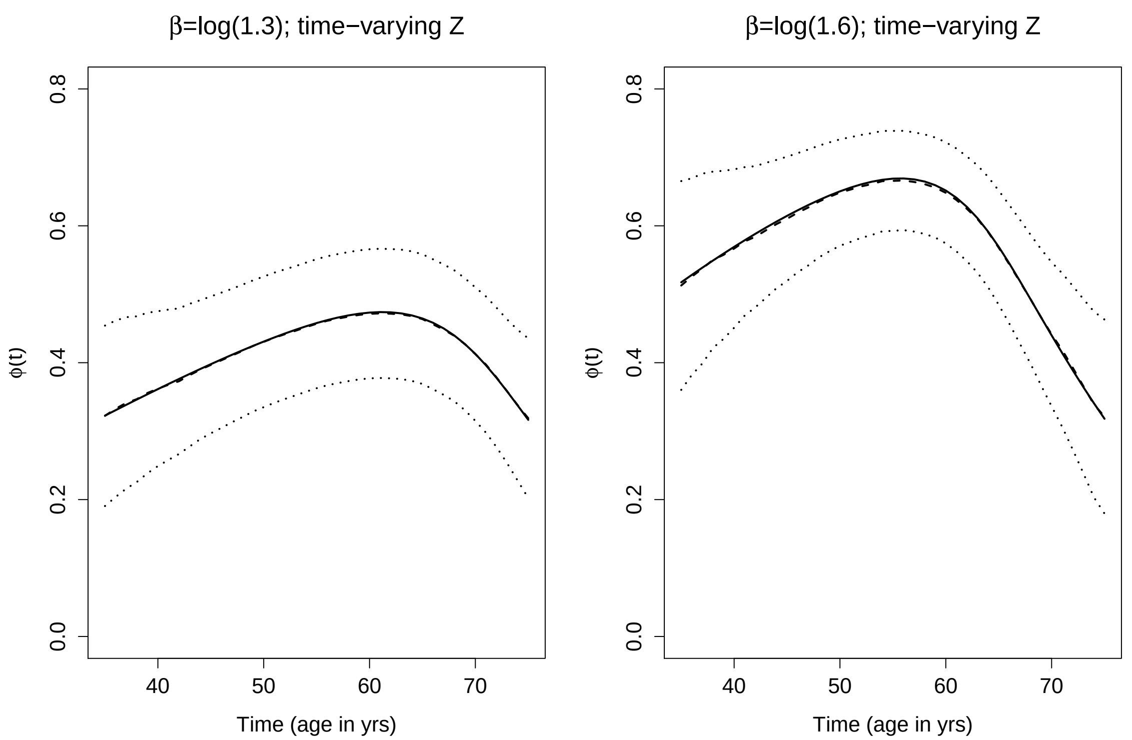

Figure 4.4: Simulation results with cross-validation based bandwidth for late onset disease scenario with continuous time-varying exposure. The theoretical function $\phi(t)$ (solid lines), the mean estimated curve of $\phi(t)$ (dash lines), and the averages of their 95% pointwise confidence intervals (dotted lines).

where $Z(t)$ is a univariate possibly time-dependent covariate, $U(t)$ is a univariate possibly time-dependent adjustment factor, and β and γ are the regression parameters. The baseline hazard $\lambda_0(t)$ is assumed to be a Weibull hazard:

$$\lambda_0(t) = \frac{\nu}{\eta}\left(\frac{t}{\eta}\right)^{\nu-1}.$$

To calculate the true $\phi_{adj}(t)$, we first write

$$\phi_{adj}(t) = 1 - \frac{\lambda_0(t)S_T(t)}{f_T(t)}\int_{\mathcal{U}} \exp(\gamma u) f_{U|T\geq t}(u)du,$$

and $S_T(t)$, $f_T(t)$, $f_{U|T\geq t}(u)$ can be calculated by using the following formulas

$$\begin{aligned} S_T(t) &= \int_U \int_Z \exp\{-\Lambda_0(t)\exp(\beta z + \gamma u)\} f_{Z,U}(z,u)dzdu, \\ f_T(t) &= \lambda_0(t) \int_U \int_Z \exp(\beta z + \gamma u)\exp\{-\Lambda_0(t)\exp(\beta z + \gamma u)\} f_{Z,U}(z,u)dzdu, \\ f_{U|T\geq t}(u) &= \frac{S_{T|U}(t|u) f_U(u)}{S_T(t)}, \\ S_{T|U}(t|u) &= \int_Z S_{T|Z,U}(t|z,u) f_{Z,U}(z,u)dz / f_U(u). \end{aligned}$$

Simulation Setting. For the simulation studies for $\phi_{adj}(t)$, we considered the late onset disease with continuous Z and U. We took 1,000 cases and 1,000 controls with age frequency matched with five-year intervals. The baseline hazard parameters were set as $\eta = 70$, $\nu = 9$. The covariate Z and confounder U were bivariate normal with means $\mu_Z = 2$ and $\mu_U = 1$, standard deviations $\sigma_Z = 2$ and $\sigma_U = 1$, and correlation $\rho = 0.6$. β_0 was log(1.3) and log(1.6). Censoring time was truncated normal between 1 and 100 with standard deviation of 25, and the mean was set to yield 70% censoring probability. These settings yielded an average age at onset of about 62 years.

Simulation Results. Figure 4.5 shows the true $\phi(t)$ curves in solid lines and the mean estimated curves of $\widehat{\phi}_{adjw}(t;\hat{\beta},\hat{\gamma})$ and the averages of their corresponding pointwise 95% confidence intervals in dashed and dotted lines. For each simulation, the cross-validation bandwidth is used for computing the estimated curves. The true $\phi_{adj}(t)$ is a monotone decreasing function with values less than 1. In this case the true $\phi_{adj}(t)$ can take negative values, since the covariate Z can take negative values. These estimators and cross-validation bandwidths give generally satisfactory results.

The simulation results for $\widehat{\phi}_{adjw}(t;\hat{\beta},\hat{\gamma})$ with various bandwidth choices are summarized in Table 4.11 and Table 4.12. The results include ages ranged from 40 to 70. From these results we see that the point and variance estimators both give satisfactory results for cross-validation bandwidth. Inappropriate fixed bandwidth may result in large bias and poor coverage (e.g. $h = 8$, $h = 10$).

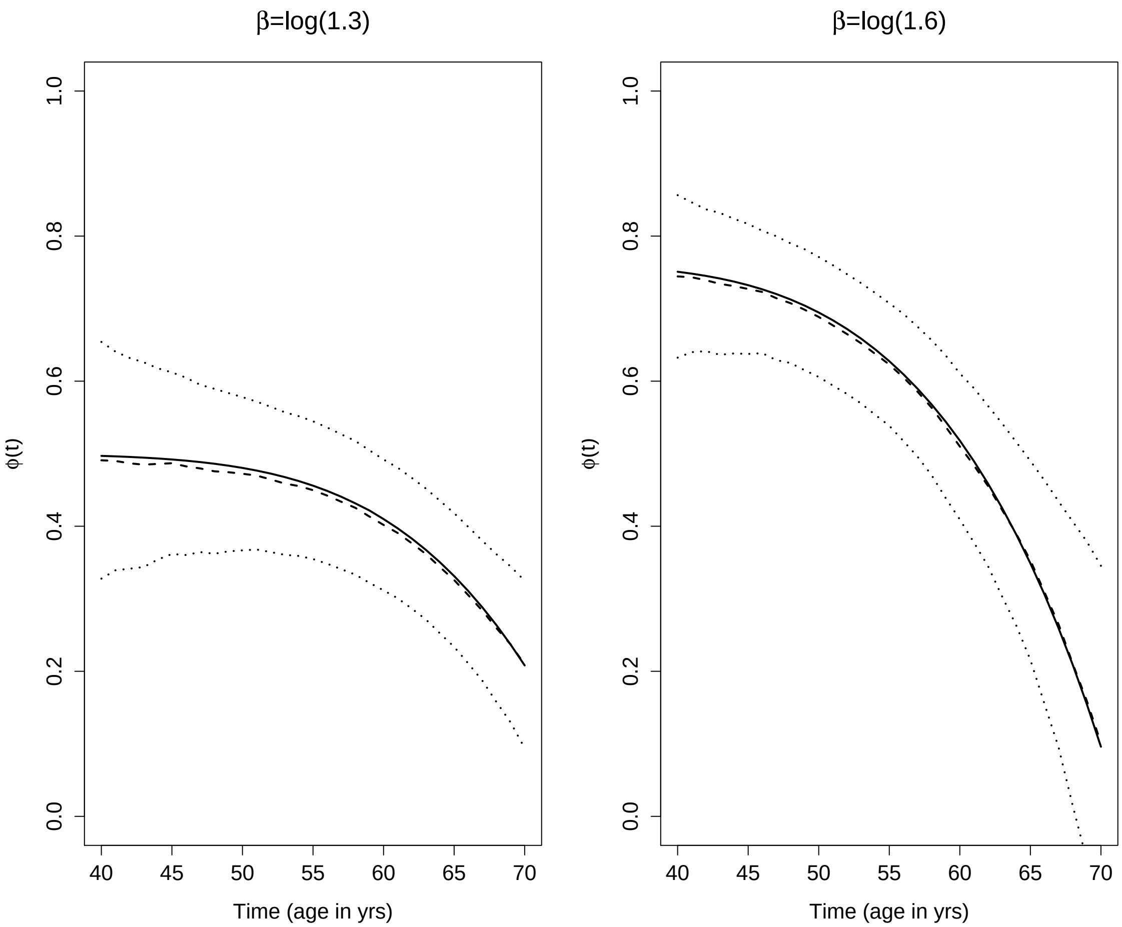

Figure 4.5: Simulation results with cross-validation based bandwidth for continuous exposure, adjusted for confounders. The theoretical function $\phi_{adj}(t)$ (solid lines), the mean estimated curve of $\phi_{adj}(t)$ (dash lines), and the averages of their 95% pointwise confidence intervals (dotted lines).

Table 4.1: Summary statistics for $\widehat{\phi}_w(t;\hat{\beta})$ with binary Z for the early onset scenario.

		h_{CV}				h=2				h=4			
Parameter	Age (yrs)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)
$\beta_0 = \log 2$, $p_z = 0.3$	10	0.0010	0.0327	0.0329	94.9	0.0015	0.0384	0.0386	94.5	0.0010	0.0338	0.0342	95.0
	20	0.0002	0.0289	0.0294	95.2	0.0008	0.0339	0.0340	94.7	0.0006	0.0300	0.0304	95.0
	30	0.0001	0.0261	0.0259	94.0	0.0002	0.0317	0.0315	94.5	0.0000	0.0273	0.0272	95.0
	40	0.0012	0.0231	0.0233	94.5	0.0014	0.0315	0.0315	93.1	0.0002	0.0248	0.0253	94.6
$\beta_0 = \log 2$, $p_z = 0.6$	10	0.0027	0.0407	0.0416	95.5	0.0026	0.0430	0.0439	95.4	0.0025	0.0410	0.0420	95.7
	20	0.0023	0.0391	0.0398	95.4	0.0024	0.0412	0.0419	95.3	0.0024	0.0392	0.0401	95.5
	30	0.0015	0.0372	0.0378	95.5	0.0022	0.0405	0.0411	95.0	0.0018	0.0378	0.0383	95.2
	40	0.0004	0.0360	0.0365	95.2	0.0017	0.0433	0.0439	94.7	0.0005	0.0372	0.0378	95.1
$\beta_0 = \log 3$, $p_z = 0.3$	10	0.0030	0.0360	0.0350	94.0	0.0028	0.0432	0.0420	94.0	0.0027	0.0363	0.0357	94.1
	20	0.0015	0.0310	0.0307	94.8	0.0020	0.0375	0.0375	94.2	0.0015	0.0316	0.0313	94.2
	30	0.0007	0.0290	0.0279	93.6	0.0014	0.0382	0.0371	93.5	0.0001	0.0295	0.0288	93.5
	40	0.0011	0.0265	0.0262	92.5	0.0025	0.0395	0.0366	86.4	0.0004	0.0279	0.0274	90.8
$\beta_0 = \log 3$, $p_z = 0.6$	10	0.0033	0.0358	0.0359	95.3	0.0030	0.0374	0.0381	95.3	0.0031	0.0354	0.0357	95.5
	20	0.0021	0.0356	0.0349	94.1	0.0025	0.0386	0.0376	93.8	0.0018	0.0351	0.0346	94.6
	30	0.0012	0.0367	0.0364	94.1	0.0038	0.0423	0.0425	94.2	0.0012	0.0354	0.0356	94.5
	40	0.0005	0.0474	0.0451	93.0	0.0038	0.0603	0.0580	92.8	0.0007	0.0450	0.0436	93.0

Bias, the sampling bias; SD, the sampling standard deviation; ASE, the sampling mean of the standard error estimator; CP, the coverage probability of the 95% pointwise confidence interval.

Table 4.2: Summary statistics for $\widehat{\phi}_w(t;\hat{\beta})$ with binary Z for the early onset scenario, continued.

		h=6				h=8				h=10			
Parameter	Age (yrs)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)
$\beta_0 = \log 2$, $p_z = 0.3$	10	0.0011	0.0321	0.0326	94.7	0.0016	0.0313	0.0317	94.7	0.0023	0.0308	0.0311	95.0
	20	0.0004	0.0285	0.0292	95.2	0.0000	0.0278	0.0285	95.4	0.0004	0.0274	0.0280	95.2
	30	0.0003	0.0256	0.0256	94.5	0.0006	0.0247	0.0247	94.3	0.0011	0.0242	0.0243	94.6
	40	0.0011	0.0224	0.0228	94.8	0.0022	0.0212	0.0215	95.5	0.0035	0.0204	0.0207	95.7
$\beta_0 = \log 2$, $p_z = 0.6$	10	0.0026	0.0404	0.0413	95.7	0.0030	0.0401	0.0410	95.5	0.0036	0.0399	0.0407	95.7
	20	0.0022	0.0387	0.0396	95.6	0.0022	0.0384	0.0393	95.5	0.0024	0.0383	0.0391	95.4
	30	0.0013	0.0368	0.0374	95.5	0.0007	0.0363	0.0369	95.5	0.0001	0.0360	0.0367	95.3
	40	0.0007	0.0350	0.0356	95.0	0.0023	0.0339	0.0345	95.5	0.0043	0.0332	0.0339	95.5
$\beta_0 = \log 3$, $p_z = 0.3$	10	0.0032	0.0338	0.0333	94.3	0.0047	0.0324	0.0320	94.4	0.0068	0.0316	0.0312	94.2
	20	0.0014	0.0293	0.0290	94.3	0.0015	0.0281	0.0277	94.2	0.0017	0.0274	0.0270	94.5
	30	0.0016	0.0261	0.0254	94.1	0.0035	0.0243	0.0236	94.0	0.0058	0.0231	0.0225	94.0
	40	0.0023	0.0231	0.0233	94.1	0.0054	0.0205	0.0208	94.9	0.0092	0.0190	0.0193	94.3
$\beta_0 = \log 3$, $p_z = 0.6$	10	0.0038	0.0349	0.0348	95.2	0.0047	0.0347	0.0344	94.8	0.0061	0.0345	0.0341	94.6
	20	0.0017	0.0340	0.0336	94.7	0.0018	0.0335	0.0331	94.8	0.0020	0.0332	0.0328	94.7
	30	0.0008	0.0331	0.0331	94.0	0.0033	0.0318	0.0318	94.5	0.0064	0.0311	0.0311	94.3
	40	0.0054	0.0389	0.0373	93.5	0.0110	0.0352	0.0336	92.8	0.0182	0.0328	0.0313	90.8

Bias, the sampling bias; SD, the sampling standard deviation; ASE, the sampling mean of the standard error estimator; CP, the coverage probability of the 95% pointwise confidence interval.

Table 4.3: Summary statistics for $\widehat{\phi}_+(t;\hat{\beta})$, $\widehat{\phi}_-(t;\hat{\beta})$, and $\widehat{\phi}_w(t;\hat{\beta})$, cross-validation bandwidth, with binary Z for the early onset scenario.

		$\widehat{\phi}_+(t;\hat{\beta})$				$\widehat{\phi}_-(t;\hat{\beta})$				$\widehat{\phi}_w(t;\hat{\beta})$			
Parameter	Age (yrs)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)
$\beta_0 = \log 2,\ p_z = 0.3$	10	0.0008	0.0366	0.0359	94.0	0.0015	0.0367	0.0365	94.2	0.0010	0.0327	0.0329	94.9
	20	0.0012	0.0310	0.0317	95.2	0.0006	0.0322	0.0321	94.8	0.0002	0.0289	0.0294	95.2
	30	0.0000	0.0288	0.0285	94.2	0.0001	0.0303	0.0299	94.8	0.0001	0.0261	0.0259	94.0
	40	0.0016	0.0270	0.0272	93.8	0.0005	0.0308	0.0301	93.8	0.0012	0.0231	0.0233	94.5
$\beta_0 = \log 2,\ p_z = 0.6$	10	0.0025	0.0426	0.0435	95.2	0.0028	0.0420	0.0428	95.8	0.0027	0.0407	0.0416	95.5
	20	0.0019	0.0406	0.0414	95.3	0.0027	0.0402	0.0408	95.5	0.0023	0.0391	0.0398	95.4
	30	0.0009	0.0397	0.0402	95.2	0.0017	0.0391	0.0398	95.6	0.0015	0.0372	0.0378	95.5
	40	0.0004	0.0421	0.0415	94.6	0.0000	0.0415	0.0421	94.9	0.0004	0.0360	0.0365	95.2
$\beta_0 = \log 3,\ p_z = 0.3$	10	0.0019	0.0413	0.0403	94.5	0.0038	0.0426	0.0411	94.3	0.0030	0.0360	0.0350	94.0
	20	0.0007	0.0355	0.0349	94.8	0.0022	0.0385	0.0372	93.8	0.0015	0.0310	0.0307	94.8
	30	0.0014	0.0333	0.0324	93.9	0.0005	0.0408	0.0389	93.6	0.0007	0.0290	0.0279	93.6
	40	0.0018	0.0310	0.0306	93.5	0.0007	0.0434	0.0415	89.7	0.0011	0.0265	0.0262	92.5
$\beta_0 = \log 3,\ p_z = 0.6$	10	0.0024	0.0393	0.0392	94.6	0.0040	0.0373	0.0374	94.8	0.0033	0.0358	0.0359	95.3
	20	0.0014	0.0393	0.0385	94.2	0.0027	0.0381	0.0371	94.2	0.0021	0.0356	0.0349	94.1
	30	0.0004	0.0433	0.0428	95.0	0.0023	0.0442	0.0436	94.7	0.0012	0.0367	0.0364	94.1
	40	0.0029	0.0573	0.0545	93.1	0.0021	0.0682	0.0663	92.8	0.0005	0.0474	0.0451	93.0

Bias, the sampling bias; SD, the sampling standard deviation; ASE, the sampling mean of the standard error estimator; CP, the coverage probability of the 95% pointwise confidence interval.

Table 4.4: Summary statistics for $\widehat{\phi}_w(t;\hat{\beta})$ with binary Z for the late onset scenario.

		h_{CV}				h=2				h=4			
Parameter	Age (yrs)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)
$\beta_0 = \log 2$, $p_z = 0.3$	40	0.0038	0.0500	0.0482	92.3	0.0056	0.0610	0.0577	90.5	0.0033	0.0477	0.0463	92.5
	50	0.0027	0.0349	0.0345	94.7	0.0028	0.0382	0.0381	94.6	0.0029	0.0340	0.0338	94.7
	60	0.0026	0.0293	0.0289	94.7	0.0025	0.0312	0.0311	94.8	0.0027	0.0289	0.0285	94.7
	70	0.0006	0.0219	0.0220	94.7	0.0011	0.0252	0.0252	93.4	0.0009	0.0211	0.0213	95.2
$\beta_0 = \log 2$, $p_z = 0.6$	40	0.0071	0.0511	0.0509	94.5	0.0071	0.0546	0.0545	94.4	0.0065	0.0480	0.0480	95.0
	50	0.0062	0.0432	0.0431	95.0	0.0066	0.0441	0.0442	95.0	0.0058	0.0425	0.0423	94.7
	60	0.0059	0.0401	0.0402	95.0	0.0056	0.0408	0.0409	94.8	0.0061	0.0396	0.0397	94.8
	70	0.0034	0.0365	0.0360	94.2	0.0049	0.0383	0.0379	93.4	0.0020	0.0351	0.0346	94.8
$\beta_0 = \log 3$, $p_z = 0.3$	40	0.0052	0.0592	0.0560	93.0	0.0081	0.0721	0.0673	92.3	0.0044	0.0532	0.0520	93.7
	50	0.0051	0.0357	0.0368	95.0	0.0052	0.0405	0.0416	94.8	0.0048	0.0340	0.0352	95.6
	60	0.0044	0.0296	0.0295	94.5	0.0034	0.0334	0.0331	93.8	0.0046	0.0284	0.0283	94.5
	70	0.0031	0.0232	0.0227	92.5	0.0014	0.0280	0.0276	90.8	0.0042	0.0210	0.0210	93.7
$\beta_0 = \log 3$, $p_z = 0.6$	40	0.0114	0.0473	0.0474	94.5	0.0119	0.0505	0.0501	94.2	0.0108	0.0423	0.0426	94.3
	50	0.0097	0.0374	0.0375	94.7	0.0095	0.0382	0.0384	94.5	0.0101	0.0354	0.0359	95.2
	60	0.0108	0.0348	0.0351	94.2	0.0108	0.0359	0.0360	93.9	0.0110	0.0333	0.0337	94.4
	70	0.0000	0.0422	0.0400	93.2	0.0046	0.0446	0.0435	92.8	0.0072	0.0342	0.0339	94.5

Bias, the sampling bias; SD, the sampling standard deviation; ASE, the sampling mean of the standard error estimator; CP, the coverage probability of the 95% pointwise confidence interval.

Table 4.5: Summary statistics for $\widehat{\phi}_w(t;\hat{\beta})$ with binary Z for the late onset scenario, continued.

		h=6				h=8				h=10			
Parameter	Age (yrs)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)
$\beta_0 = \log 2,\, p_z = 0.3$	40	0.0030	0.0424	0.0414	92.8	0.0028	0.0396	0.0384	93.0	0.0029	0.0376	0.0364	93.3
	50	0.0031	0.0326	0.0322	94.5	0.0036	0.0317	0.0313	94.5	0.0046	0.0310	0.0306	94.7
	60	0.0035	0.0278	0.0276	94.6	0.0047	0.0272	0.0270	94.7	0.0062	0.0268	0.0266	94.6
	70	0.0041	0.0199	0.0201	95.5	0.0078	0.0196	0.0198	94.0	0.0115	0.0196	0.0197	91.4
$\beta_0 = \log 2,\, p_z = 0.6$	40	0.0064	0.0454	0.0457	95.0	0.0063	0.0440	0.0443	94.8	0.0064	0.0433	0.0434	95.0
	50	0.0060	0.0418	0.0416	95.1	0.0066	0.0414	0.0412	95.2	0.0074	0.0411	0.0409	95.1
	60	0.0067	0.0393	0.0392	94.9	0.0077	0.0390	0.0389	94.8	0.0091	0.0388	0.0387	94.9
	70	0.0024	0.0342	0.0337	94.4	0.0075	0.0339	0.0335	94.5	0.0127	0.0339	0.0336	93.5
$\beta_0 = \log 3,\, p_z = 0.3$	40	0.0038	0.0460	0.0455	94.0	0.0041	0.0413	0.0415	94.7	0.0047	0.0383	0.0387	94.8
	50	0.0057	0.0316	0.0328	95.3	0.0073	0.0302	0.0314	95.2	0.0098	0.0293	0.0305	94.7
	60	0.0058	0.0266	0.0265	94.2	0.0079	0.0255	0.0255	93.6	0.0108	0.0246	0.0248	93.2
	70	0.0134	0.0183	0.0184	91.6	0.0241	0.0170	0.0174	76.2	0.0355	0.0166	0.0170	44.2
$\beta_0 = \log 3,\, p_z = 0.6$	40	0.0108	0.0394	0.0398	94.8	0.0111	0.0380	0.0383	94.2	0.0115	0.0368	0.0372	94.5
	50	0.0110	0.0345	0.0350	95.0	0.0123	0.0340	0.0346	95.0	0.0140	0.0336	0.0342	94.7
	60	0.0119	0.0324	0.0329	94.8	0.0133	0.0320	0.0325	94.5	0.0154	0.0318	0.0322	93.9
	70	0.0237	0.0306	0.0303	89.2	0.0421	0.0292	0.0288	68.5	0.0608	0.0288	0.0283	43.2

Bias, the sampling bias; SD, the sampling standard deviation; ASE, the sampling mean of the standard error estimator; CP, the coverage probability of the 95% pointwise confidence interval.

Table 4.6: Summary statistics for $\widehat{\phi}_w(t;\hat{\beta})$ with continuous Z for the late onset scenario.

		h_{CV}				h=2				h=4			
Parameter	Age (yrs)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)
$\beta_0 = \log 1.3$	40	0.0083	0.0558	0.0558	95.1	0.0091	0.0593	0.0585	94.2	0.0080	0.0543	0.0546	95.3
	50	0.0080	0.0509	0.0512	95.5	0.0080	0.0517	0.0520	95.2	0.0080	0.0506	0.0509	95.3
	60	0.0073	0.0489	0.0491	95.5	0.0071	0.0495	0.0495	95.9	0.0072	0.0487	0.0489	95.5
	70	0.0042	0.0443	0.0447	95.0	0.0060	0.0450	0.0455	95.5	0.0036	0.0439	0.0443	95.2
$\beta_0 = \log 1.6$	40	0.0102	0.0409	0.0419	95.0	0.0105	0.0438	0.0444	95.3	0.0099	0.0390	0.0400	95.0
	50	0.0104	0.0359	0.0368	95.3	0.0104	0.0366	0.0376	95.6	0.0104	0.0352	0.0361	95.2
	60	0.0097	0.0347	0.0359	94.8	0.0097	0.0355	0.0367	94.8	0.0097	0.0341	0.0353	94.8
	70	0.0008	0.0364	0.0368	95.2	0.0055	0.0384	0.0394	94.8	0.0027	0.0334	0.0348	95.5

Bias, the sampling bias; SD, the sampling standard deviation; ASE, the sampling mean of the standard error estimator; CP, the coverage probability of the 95% pointwise confidence interval.

Table 4.7: Summary statistics for $\widehat{\phi}_w(t;\hat{\beta})$ with continuous Z for the late onset scenario, continued.

		h=6				h=8				h=10			
Parameter	Age (yrs)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)
$\beta_0 = \log 1.3$	40	0.0077	0.0525	0.0532	95.8	0.0078	0.0516	0.0523	95.7	0.0078	0.0511	0.0518	95.7
	50	0.0081	0.0501	0.0505	95.5	0.0084	0.0499	0.0503	95.6	0.0090	0.0497	0.0501	95.7
	60	0.0076	0.0484	0.0486	95.4	0.0084	0.0483	0.0485	95.4	0.0094	0.0482	0.0483	95.6
	70	0.0001	0.0440	0.0441	94.7	0.0042	0.0442	0.0442	94.8	0.0083	0.0445	0.0444	93.8
$\beta_0 = \log 1.6$	40	0.0100	0.0372	0.0384	95.5	0.0102	0.0361	0.0375	95.8	0.0106	0.0355	0.0369	95.8
	50	0.0111	0.0346	0.0356	95.2	0.0124	0.0342	0.0354	95.5	0.0143	0.0339	0.0352	94.8
	60	0.0095	0.0337	0.0348	94.5	0.0096	0.0335	0.0346	95.0	0.0102	0.0334	0.0345	95.0
	70	0.0149	0.0320	0.0333	93.0	0.0290	0.0315	0.0328	84.6	0.0442	0.0315	0.0327	73.0

Bias, the sampling bias; SD, the sampling standard deviation; ASE, the sampling mean of the standard error estimator; CP, the coverage probability of the 95% pointwise confidence interval.

Table 4.8: Summary statistics for $\widehat{\phi}_w(t;\hat{\beta})$ with continuous time-varying exposure for the late onset scenario.

		h_{CV}				h=2				h=4			
Parameter	Age (yrs)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)
$\beta_0 = \log 1.3$	40	0.0035	0.0482	0.0470	93.8	0.0022	0.0494	0.0482	94.4	0.0038	0.0451	0.0437	93.2
	50	0.0012	0.0461	0.0452	94.8	0.0010	0.0463	0.0456	94.8	0.0016	0.0451	0.0443	94.1
	60	0.0007	0.0467	0.0464	94.8	0.0003	0.0471	0.0466	95.0	0.0009	0.0461	0.0459	94.6
	70	0.0007	0.0452	0.0456	94.7	0.0003	0.0452	0.0462	94.9	0.0037	0.0440	0.0440	95.4
$\beta_0 = \log 1.6$	40	0.0025	0.0456	0.0433	93.2	0.0027	0.0459	0.0435	93.1	0.0004	0.0391	0.0380	93.8
	50	0.0024	0.0349	0.0350	94.9	0.0024	0.0350	0.0351	94.9	0.0019	0.0332	0.0336	94.7
	60	0.0044	0.0353	0.0345	94.4	0.0043	0.0354	0.0345	94.3	0.0054	0.0338	0.0331	94.9
	70	0.0001	0.0433	0.0418	94.3	0.0005	0.0432	0.0420	94.7	0.0106	0.0359	0.0360	94.8

Bias, the sampling bias; SD, the sampling standard deviation; ASE, the sampling mean of the standard error estimator; CP, the coverage probability of the 95% pointwise confidence interval.

Table 4.9: Summary statistics for $\widehat{\phi}_w(t;\hat{\beta})$ with continuous time-varying exposure for the late onset scenario, continued.

		h=6				h=8				h=10			
Parameter	Age (yrs)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)
$\beta_0 = \log 1.3$	40	0.0066	0.0436	0.0421	92.8	0.0109	0.0429	0.0415	92.4	0.0164	0.0425	0.0412	90.7
	50	0.0028	0.0445	0.0439	93.7	0.0047	0.0443	0.0438	93.7	0.0073	0.0442	0.0438	93.3
	60	0.0019	0.0458	0.0456	94.7	0.0034	0.0456	0.0454	95.1	0.0052	0.0454	0.0452	94.8
	70	0.0087	0.0437	0.0435	94.7	0.0141	0.0438	0.0434	92.8	0.0191	0.0439	0.0435	91.3
$\beta_0 = \log 1.6$	40	0.0037	0.0368	0.0359	94.0	0.0093	0.0353	0.0347	93.0	0.0165	0.0344	0.0340	91.1
	50	0.0019	0.0329	0.0330	94.9	0.0014	0.0327	0.0328	95.0	0.0008	0.0326	0.0326	95.1
	60	0.0068	0.0333	0.0327	95.1	0.0087	0.0330	0.0325	94.7	0.0110	0.0328	0.0324	93.7
	70	0.0264	0.0336	0.0338	86.9	0.0454	0.0326	0.0328	70.7	0.0657	0.0323	0.0325	48.2

Bias, the sampling bias; SD, the sampling standard deviation; ASE, the sampling mean of the standard error estimator; CP, the coverage probability of the 95% pointwise confidence interval.

Table 4.10: Simultaneous confidence bands coverage (continuous time-independent Z, late onset disease).

β_0	$h = h_{CV}$	h=2	h=4	h=6	h=8	h=10
$\log 1.3$	93.2%	92.2%	94.3%	95.1%	94.8%	95.0%
$\log 1.6$	92.1%	92.7%	93.6%	93.3%	89.6%	80.9%

Table 4.11: Summary statistics for $\widehat{\phi}_{adjw}(t;\hat{\beta},\hat{\gamma})$ with continuous Z and U.

		h_{CV}				h=2				h=4			
Parameter	Age (yrs)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)
$\beta_0 = \log 1.3$	40	0.0136	0.0585	0.0579	94.4	0.0152	0.0612	0.0612	94.5	0.0121	0.0507	0.0531	95.9
	50	0.0111	0.0440	0.0465	96.0	0.0108	0.0455	0.0478	96.3	0.0112	0.0415	0.0448	96.4
	60	0.0104	0.0387	0.0406	95.1	0.0106	0.0397	0.0416	95.0	0.0105	0.0375	0.0393	95.3
	70	0.0010	0.0378	0.0377	95.1	0.0040	0.0412	0.0413	94.8	0.0029	0.0317	0.0328	95.1
$\beta_0 = \log 1.6$	40	0.0133	0.0413	0.0406	95.1	0.0138	0.0447	0.0442	94.2	0.0124	0.0364	0.0379	95.5
	50	0.0121	0.0334	0.0349	95.6	0.0115	0.0357	0.0368	95.4	0.0120	0.0311	0.0334	95.5
	60	0.0091	0.0368	0.0366	94.3	0.0110	0.0393	0.0395	94.3	0.0076	0.0333	0.0337	94.5
	70	0.0055	0.0783	0.0708	90.5	0.0092	0.0881	0.0839	92.1	0.0115	0.0610	0.0588	91.0

Bias, the sampling bias; SD, the sampling standard deviation; ASE, the sampling mean of the standard error estimator; CP, the coverage probability of the 95% pointwise confidence interval.

Table 4.12: Summary statistics for $\widehat{\phi}_{adjw}(t;\hat{\beta},\hat{\gamma})$ with continuous Z and U, continued.

		h=6				h=8				h=10			
Parameter	Age (yrs)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)
$\beta_0 = \log 1.3$	40	0.0117	0.0472	0.0500	95.5	0.0116	0.0450	0.0481	96.0	0.0122	0.0432	0.0468	95.7
	50	0.0123	0.0399	0.0437	96.5	0.0141	0.0391	0.0430	96.7	0.0167	0.0386	0.0425	96.5
	60	0.0109	0.0366	0.0384	95.8	0.0121	0.0360	0.0380	95.6	0.0138	0.0354	0.0376	95.4
	70	0.0124	0.0284	0.0299	94.1	0.0240	0.0273	0.0289	90.0	0.0366	0.0273	0.0288	76.9
$\beta_0 = \log 1.6$	40	0.0130	0.0328	0.0355	96.5	0.0146	0.0308	0.0341	96.3	0.0170	0.0297	0.0332	95.6
	50	0.0135	0.0295	0.0322	95.3	0.0161	0.0286	0.0315	95.1	0.0201	0.0283	0.0312	94.1
	60	0.0035	0.0312	0.0316	94.5	0.0011	0.0302	0.0306	94.3	0.0052	0.0295	0.0300	94.3
	70	0.0367	0.0486	0.0472	83.6	0.0663	0.0410	0.0404	57.8	0.0991	0.0362	0.0358	21.2

Bias, the sampling bias; SD, the sampling standard deviation; ASE, the sampling mean of the standard error estimator; CP, the coverage probability of the 95% pointwise confidence interval.

Chapter 5
DATA ANALYSIS

In this chapter, we apply our proposed methods to three case-control studies from the Genetics and Epidemiology Colorectal Cancer Consortium (GECCO). The aim of this data analysis is to estimate the time-varying population attributable fraction of various risk factors to colorectal cancer (CRC).

5.1 Data Description

The GECCO study is supported by the U.S. National Cancer Institute and it is composed of twelve well-characterized prospective cohorts and case-control studies of colorectal cancer. This consortium aims to both accelerate the discovery of colorectal cancer related variants and perform thorough epidemiologic evaluations of new susceptibility loci via gene-environmental interaction analyses. Key clinical and environmental data have been harmonized across all studies. The GECCO smoking data set we use for our analysis include a total of 5498 subjects (2742 cases and 2756 controls) from three population-based case-control GECCO studies. The primary outcome is the case control status of colorectal cancer. The other variables include age (age at onset for cases and age at enrollment for controls), gender, weight, height, BMI, medical history, dietary, alcohol use, smoking, exercises, cancer status, family history of colorectal cancer, and history of diabetes. The details of these studies are provided below and the number of subjects is summarized in table 5.1.

1. **Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial (PLCO).** PLCO is a large population-based randomized trial to determine the effectiveness of screening to reduce cancer mortality. The trial enrolled 154,934 participants (men and women) aged between 55 and 74 years at ten centers from 1993 to 2001. Details of this study are available online (http://prevention.cancer.gov/plco). The GECCO PLCO data included a subset of 991 colorectal cancer cases and 937 controls. The cases were

self-reported as being non-Hispanic White with available DNA samples, questionnaire data, and appropriate consent for ancillary epidemiologic studies. Cases were excluded if they had a history of inflammatory bowel disease, polyps, polyposis syndrome or cancer (excluding basal or squamous cell skin cancer). Controls came from the Cancer Genetic Markers of Susceptibility (CGEMS) prostate cancer scan (all male) and the Genome-wide Association Study (GWAS) of Lung Cancer and Smoking along with an additional 92 non-Hispanic White female controls. Controls were frequency matched 1:1 to cases without replacement. Matching criteria were age at enrollment (two year blocks), enrollment date (two year blocks), sex, race / ethnicity, trial arm, and study year of diagnosis (i.e. controls must be cancer free into the case's year of diagnosis).

2. **VITamins And Lifestyle (VITAL).**
   The VITAL study enrolled 77,738 Washington State men and women aged 50 to 76 years from 2000 to 2002 to investigate the association of supplement use with cancer risk. All subjects completed a questionnaire about detailed information on use of 38 supplements over the past 10 years, diet, non-steroidal anti-inflammatory drug use, anthropometrics, exercise, health history and cancer risk factors. 54,045 participants also provided DNA sample for genetic studies. Subjects are followed for cancer by linkage to the western Washington Surveillance, Epidemiology, and End Results (SEER) cancer registry and are censored when they move out of the area covered by the registry or at time of death. In GECCO, a nested case-control set was genotyped. Samples included colorectal cancer cases with DNA, excluding subject with colorectal cancer before baseline, in situ cases, (large cell) neuroendocrine carcinoma, squamous cell carcinoma, carcinoid tumor, Goblet cell carcinoid, any type of lymphoma, including non-Hodgkin, Mantle cell, large B-cell, or follicular lymphoma. Controls were matched on age at enrollment (within one year), enrollment date (within one year), sex, and race / ethnicity. One control was randomly selected per case among all controls that matched on the four factors above and where the control follow-up time was greater than follow-up time of the case until diagnosis. A total of 285 cases and 288 controls were included.

3. **Women's Health Initiative (WHI).**
WHI is a long-term health study of 161,808 post-menopausal women aged 50 to 79 years recruited from 1993 to 1998 at 40 clinical centers throughout the U.S. WHI comprises a clinical trial (CT) arm, an Observational Study (OS) arm, and several extension studies. The details of WHI are available online (https://cleo.whi.org/SitePages/Home.aspx). The GECCO WHI data included 1466 cases selected from centrally adjudicated colon cancer cases from both arms, and 1531 controls from controls previously genotyped as part of a Hip Fracture GWAS conducted within the WHI OS. The controls were matched to cases on age (within three years) enrollment date (within 365 days), hysterectomy status, and prevalent conditions at baseline. For 37 cases, there was not a control match in the Hip Fracture GWAS. For these participants, matched controls were identified in the WHI OS based on same criteria. In addition, case and control participants were subject to the following exclusion criteria: a prior history of colorectal cancer at baseline, IRB approval not available for data submission into dbGaP, and not sufficient DNA available. Matching criteria included age (within years), race/ethnicity, WHI date (within three years), WHI Calcium and Vitamin D study date (within three years), and randomization arms (OSag, hormone therapy assignments, dietary modification assignments, calcium/vitamin D assignments). In addition, they were matched on the four regions of randomization centers. Each case was matched with one control (1:1) that exactly met the matching criteria. Control selection was done in a time-forward manner, selecting one control for each case from the risk set at the time of the case event. The matching algorithm was allowed to select the closest match based on a criterion to minimize an overall distance measure. Each matching factor was given the same weight. Additional available controls that were genotyped as part of the Hip Fracture GWAS were included to improve power.

5.2 Statistical Analysis

In this analysis, we focused on three smoking variables: smoking status (ever-smoker vs non-smokers), pack-years ($>= 20$ vs < 20 and non-smoker), and years since quit smoking

Table 5.1: Study information

Study	N Cases	N Controls	Total
PLCO	991	937	1928
VITAL	285	288	573
WHI	1466	1531	2997
Total	2742	2756	5498

($\leq$ 10 years vs $>$ 10 years and non-smokers). We also investigated the association between risk of CRC and other variables including obesity (BMI larger than 30 kg/m^2), alcohol intake (more than 1g/day vs 0-1g/day), family history of CRC, and history of diabetes. Time-varying PAF was estimated for those with significant associations. For each of these variables, logistic regression models were used to estimate its association with CRC, and the proposed kernel estimators were used to obtain the estimated functional time-varying PAF $\phi(t)$. Epanechnikov kernel and cross-validation bandwidth were used for the kernel estimation. Both the unadjusted estimates and the estimates adjusting for potential confounders were provided. For the unadjusted model, β was estimated from a logistic model with adjustment for observed age, gender, and study. For the adjusted model, β was estimated from a logistic model was adjusted for observed age, gender, study, and other potential confounders.

5.3 Results

The descriptive statistics of the study population are summarized in Table 5.2. We see that the age, gender, and mean BMI are balanced between cases and controls, while obesity, alcohol use, family history of CRC, history of diabetes, and ever smoking have higher prevalence in cases. The mean of pack-year is higher in cases, and the mean of year since quit smoking is higher in controls.

The sample prevalence, estimated odds ratios, and the estimated classic PAFs for key risk factors are listed in Table 5.3. These variables are all significantly associated with CRC risk, with odds ratios (OR) between 1.19 and 1.65. History of diabetes has the highest OR of 1.65, however, the prevalence of exposure is low with 7.5% in cases and 4.7% in controls, resulting in an estimated PAF of only 2.9. The composite variable ever-smoking+obesity has the highest PAF of 10.4, followed by the single variable of ever-smoking, with a PAF of 9.0.

Figure 5.1 shows the estimated curves of $\phi(t)$ for the three smoking variables, and Figure 5.2 shows the estimated curves of $\phi(t)$ for family history of CRC, diabetes history, and obesity. Since $\widehat{\phi}_w(t;\hat{\beta})$ generally provides better results than either $\widehat{\phi}_+(t;\hat{\beta})$ or $\widehat{\phi}_-(t;\hat{\beta})$, we show only $\widehat{\phi}_w(t;\hat{\beta})$, and the results for $\widehat{\phi}_+(t;\hat{\beta})$ and $\widehat{\phi}_-(t;\hat{\beta})$ are provided in the Appendix. From these plots, we see that the attributable risk of smoking to CRC decreases over time. The $\widehat{\phi}_w(t;\hat{\beta})$ for ever-smoking is about 10% at age 50 and drops slowly to 8% at age 80, and further to 5% at 90 years of age. The estimated curve for year since quit smoking has a steeper decreasing pattern, from 9% at age 50 to 1% at age 90. The estimated curve for pack-years is approximately flat from age 50 to 70, and then slowly decreases from 8% to 5% at age 90. The estimated curves for family history of CRC and history of diabetes both have a increasing pattern. The estimated PAF increases from 2% to 4.8% for family history of CRC, and from 0.5% to 3% for history of diabetes, over 50 to 90 years of age. The estimated PAF for obesity has a decreasing pattern, from 8% to 4%. The estimated time-varying PAF for the composite variable of ever smoking and obesity takes greater values than both estimates of the single variable, and has a steeper decreasing pattern, from 14% to 5%. The estimates for the composite variable of pack-years and obesity and the composite variable of year since quit smoking and obesity both take slightly smaller values than their corresponding single smoking variables.

Figure 5.3 shows the estimated curves of $\phi(t)$ for the three smoking variables with adjustment for age, gender, study, and obesity. These curves has similar pattern as the unadjusted ones.

Table 5.2: Summary statistics of study populations

Characteristic	Cases	Controls	Total
Age (Mean, range)	70.8 (50 - 91)	70.9 (50 - 91)	70.9 (50 - 91)
Female (%)	73.9	74.2	74.1
Ever smoking (%)	55.2	50.8	53.0
Pack-year (Mean, SD)	31.1 (26.9)	28.6 (28.3)	29.9 (27.6)
Year since quit smoking (Mean, SD)	17.9 (13.9)	20.3 (14.4)	19.1 (14.2)
BMI (Mean, SD)	28.1 (5.4)	27.5 (5.1)	27.8 (5.2)
BMI Categories (%)			
Underweight	2.2	2.2	2.2
Normal	30.1	32.8	31.4
Overweight	38.8	39.7	39.2
Obese	29.0	25.2	27.2
Alcohol use (%)	47.9	50.7	49.4
Family history of CRC (%)	17.6	15.1	16.3
History of diabetes (%)	7.5	4.7	6.1

Age: age at onset for cases and age at selection for controls; Pack-year: only for ever-smokers; Year since quit smoking: only for ever-smokers.

5.4 Summary

In this chapter, we present an analysis of the GECCO smoking data. We provide estimation of both classic PAF and the time-varying PAF with or without adjustment of confounders. The unadjusted and adjusted models give similar results. The the time-varing PAF of the three smoking variables and obesity all show downward patterns. The PAF for year since

Table 5.3: Sample prevalences, estimated odds ratios, and the estimated PAFs

Risk factor	Prevalence in cases (%)	Prevalence in controls (%)	ORs (95% CI)	$\widehat{PAF}$
Family history of CRC	17.6	15.1	1.21 (1.04, 1.40)	3.0
History of diabetes	7.5	4.7	1.65 (1.31, 2.07)	2.9
Obesity	29.5	25.7	1.21 (1.07, 1.36)	5.1
Ever smoking	55.2	50.8	1.19 (1.07, 1.33)	9.0
Pack-year (>= 20)	30.0	24.1	1.35 (1.20, 1.52)	7.8
Year since quit smoking ($\leq$ 10 yrs)	18.9	14.9	1.33 (1.15, 1.54)	4.7

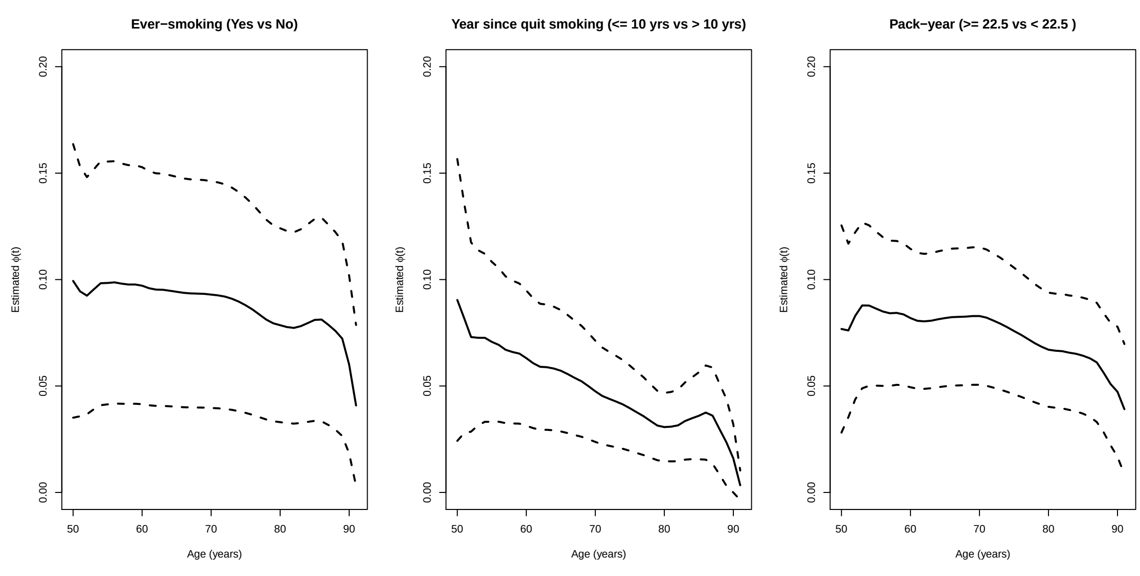

Figure 5.1: Estimated curves of $\widehat{\phi}_w(t;\hat{\beta})$ for smoking variables. Estimated curve (solid lines) and estimated 95% pointwise confidence intervals (dash lines).

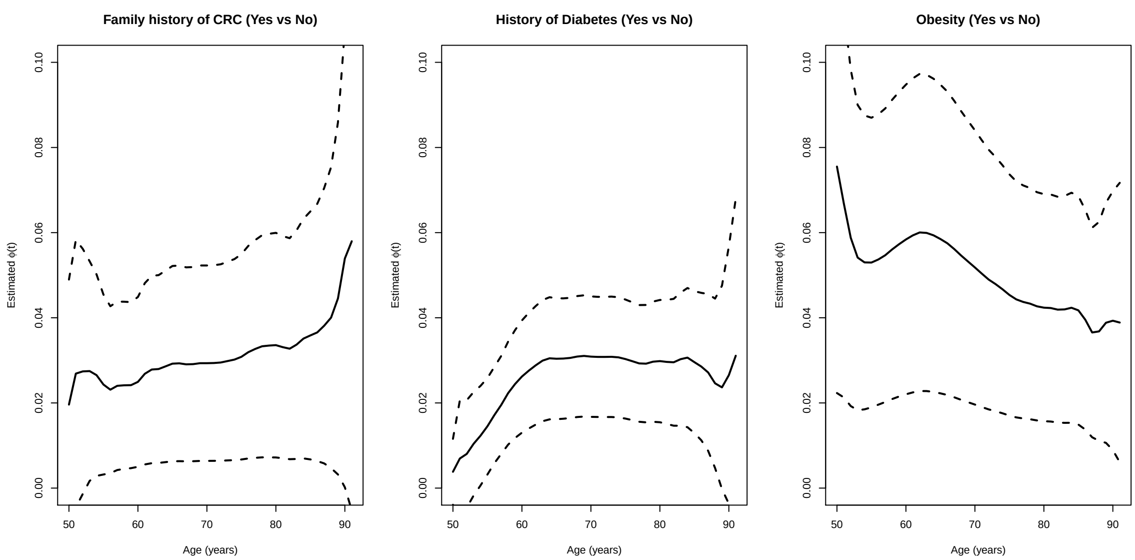

Figure 5.2: Estimated curves of $\widehat{\phi}_w(t;\hat{\beta})$ for family history of CRC, history of diabetes, and obesity. Estimated curve (solid lines) and estimated 95% pointwise confidence intervals (dash lines).

quit smoking deceases more than 50% over 40 years age. These estimates are based on the Cox model where we assume a constant regression coefficient. Therefore, the overall time-varying pattern in the time-varying PAF should be determined by $F_{Z|T}(z|t)$, i.e., the time-varying proportion of the risk factor Z in the cohort. People with adverse exposure (e.g. smoking) has a higher risk for CRC, therefore as time progress, the people with higher risk tends to fail prior to t and shall not be considered in $F_{Z|T}(z|t)$, and the cohort thus tends to be 'healthier'. This may explain the decreasing pattern of the PAF curves. The PAF curves for family history of CRC and history of diabetes have upward patterns, which can be explained by the fact that senior people are naturally more likely to have these exposures.

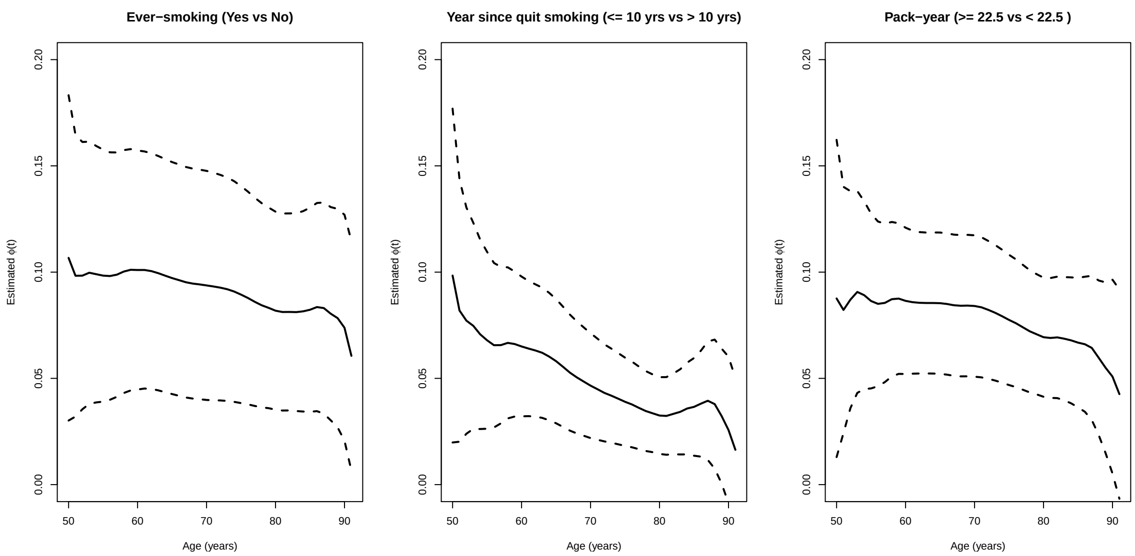

Figure 5.3: Estimated curves of $\widehat{\phi}_{adjw}(t;\hat{\beta})$ for smoking variables. Model adjusted for age, study, gender, and obesity. Estimated curve (solid lines) and estimated 95% pointwise confidence intervals (dash lines).

Chapter 6

SUMMARY AND FUTURE WORK

6.1 Summary

The population attributable fraction (PAF) is an important measure in evaluating the contribution of risk factors to disease burden or mortality. It is a useful tool for planning of prevention strategies. Recent development of PAF has extended the classic static measure to a time-varying function, which provides a dynamic picture of how PAF for an exposure varies over time. In this dissertation we propose three kernel-type estimators of the time-varying PAF under the Cox model for case-control studies. These estimators are constructed by combining a logistic regression model based estimator and a kernel smoother. The first estimator incorporates a kernel smoother on the cases data, the second estimator incorporates a kernel smoother on the control data, and the third is a combined estimator that sums the first two estimators with weight. We study the large sample properties of these estimators, and establish the strong uniform consistency and asymptotic normality by using modern asymptotic theories. We also study the finite sample properties of these estimators, and note that the asymptotic variances are not adequate to explain the actual variances under finite sampling. The extra variability comes from the variability of the regression coefficient estimators that vanishes as the bandwidth goes to zero but contributes considerable variability when the bandwidth is not small enough. We derive variance corrections under finite sampling by considering non-zero bandwidths. We propose empirical variance estimators based on the asymptotic normality with correction for finite sampling. We also propose a resampling based approach for calculating the simultaneous confidence bands.

We also consider a more realistic situation when there exists other risk factors that may confound the association between the exposure of interest and the disease. We propose the adjusted time-varying PAF, and three kernel-type estimators for it. We study both the large and finite sample properties of these estimators. We note, that any associations identified by estimating time-varying PAF might be the result of underlying mechanisms that relate to timing of exposure to different risk factors (see e.g. Rothman and Poole, 1988 [24]). Unfortunately, identification of the timing of exposures and further details regarding the underlying mechanisms is difficult if not impossible in the case-control design.

We conduct extensive Monte Carlo simulation studies to evaluate the proposed statistical procedures. We design simulations for four scenarios, each with multiple parameter settings, to assess the performance of the estimators under these scenarios. We consider early onset common disease with binary exposure, late onset rare disease with binary exposure, late onset rare disease with continuous exposure, and late onset rare disease with time-varying continuous exposure. We find that the point and variance estimators perform well under all the situations tested. As expected, the combined estimator generally is more efficient than the other two estimators.

Finally, we illustrate our proposed methods by applying them to the GECCO data. We observe PAF increases with age for the positive family history of colorectal cancer and positive history of diabetes, and it decreases with age for smoking and obesity.

6.2 Future Work

In this work we focus on the concept of time-varying PAF and estimators based on the kernel smoother. There are at least two aspects of our estimators that can be improved by applying modification to the kernel smoother. The first is to deal with the boundary effect. It is known that the kernel smoother can be dramatically biased near the boundary of the underlying functions. A common remedy to this issue is to use boundary kernels within the boundary region. The second aspect is to use a locally-varying bandwidth. The kernel estimator with fixed bandwidth is susceptible to local bumpiness where density of data points is low. The bandwidths which vary according to the density of local data points may have better performances, especially when the local density of the data points changes substantially.

Besides the kernel smoother refinement, one possibly interesting direction is to study more on the combined estimator. The proposed combined estimator is a special case within a more general class of weighted estimators. Specially, one may consider the following weighted estimator:

$$\widehat{\phi}_w(t;\hat{\beta}) = w(t)\widehat{\phi}_+(t;\hat{\beta}) + \{1 - w(t)\}\widehat{\phi}_-(t;\hat{\beta}). \tag{6.1}$$

where $w(t)$ could be any weighting function. Our estimator is a special case with $w(t) = \pi_0$. Since the combined estimator generally has better performances than its two components, some optimization on the weighting could potentially make it even better. For example, one could also use a inverse variance weight, where

$$w(t) = \frac{\mathrm{Var}\{\widehat{\phi}_+(t;\hat{\beta})\}^{-1}}{\mathrm{Var}\{\widehat{\phi}_+(t;\hat{\beta})\}^{-1} + \mathrm{Var}\{\widehat{\phi}_-(t;\hat{\beta})\}^{-1}}.$$

Since this weighting includes variance estimates from the data, the large sample properties and the bandwidth selection are not as straightforward as those for the constant weighting, and requires substantial further development.

PAF is also an important quantity in estimating the age-specific absolute risk from population-based case-control studies [13]. Estimation of the baseline hazard is a crucial step in age-specific absolute risk estimation. One common approach for estimating the baseline hazard for case-control study is through the relationship $\lambda_0(t) = \lambda^*(t)\{1 - \phi(t)\}$, where $\lambda^*(t)$ is the composite incidence rate which can be obtained from an external registry such as the National Cancer Institute's Surveillance Epidemiology and End Results (SEER). The existing approach for estimating $\phi(t)$ is to approximate it by using the classic PAF [14]. Since the classic PAF is a static measure, this approximation is only valid when $\phi(t)$ remains roughly constant over time. The development in this dissertation will be helpful to obtain a more accurate estimation of $\phi(t)$, and hence $\lambda_0(t)$, which results in a more accurate absolute risk estimation.

BIBLIOGRAPHY

[1] S. Basu and J. R. Landis. Model-based estimation of population attributable risk under cross-sectional sampling. *American Journal of Epidemiology*, 142(12):1338–1343, 1995.

[2] J. Benichou. A review of adjusted estimators of attributable risk. *Statistical Methods in Medical Research*, 10(3):195–216, 2001.

[3] J. Benichou. Attributable risk. *Encyclopedia of Environmetrics*, 2005.

[4] J. Benichou and M. H. Gail. A delta method for implicitly defined random variables. *The American Statistician*, 43(1):41–44, 1989.

[5] J. Benichou and M. H. Gail. Variance calculations and confidence intervals for estimates of the attributable risk based on logistic models. *Biometrics*, pages 991–1003, 1990.

[6] Adrian W Bowman. An alternative method of cross-validation for the smoothing of density estimates. *Biometrika*, 71(2):353–360, 1984.

[7] P. Bruzzi, S. B. Green, D. P. Byar, L. A. Brinton, and C. Schairer. Estimating the population attributable risk for multiple risk factors using case-control data. *American journal of epidemiology*, 122(5):904–914, 1985.

[8] L. Chen, D. Y. Lin, and D. L. Zeng. Attributable fraction functions for censored event times. *Biometrika*, 97(3):713–726, 2010.

[9] Y. Q. Chen, C. Hu, and Y. Wang. Attributable risk function in the proportional hazards model for censored time-to-event. *Biostatistics*, 7(4):515–529, 2006.

[10] P. Cole and B MacMahon. Attributable risk percent in case-control studies. *British journal of preventive & social medicine*, 25(4):242–244, 1971.

[11] C. Cox, H. Chu, and A. Muñoz. Survival attributable to an exposure. *Statistics in medicine*, 28(26):3276–3293, 2009.

[12] D. R. Cox. Regression models and life-tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 187–220, 1972.

[13] L. A. Gail, M. H.and Brinton, D. P. Byar, D. K. Corle, S. B. Green, C. Schairer, and J. J. Mulvihill. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *Journal of the National Cancer Institute*, 81(24):1879–1886, 1989.

[14] Mitchell H Gail. Personalized estimates of breast cancer risk in clinical practice and public health. *Statistics in medicine*, 30(10):1090–1104, 2011.

[15] S. Greenland. Variance estimators for attributable fraction estimates consistent in both large strata and sparse data. *Statistics in medicine*, 6(6):701–708, 1987.

[16] S. Greenland and K. Drescher. Maximum likelihood estimation of the attributable fraction from logistic models. *Biometrics*, pages 865–872, 1993.

[17] S. Greenland, O. Gefeller, C. Kooperberg, and D. B. Petitti. The bootstrap method for standard errors and confidence intervals of the adjusted attributable risk. *Epidemiology*, 3(3):271–273, 1992.

[18] D. B. Kooperberg, C.and Petitti. Using logistic regression to estimate the adjusted attributable risk of low birthweight in an unmatched case-control study. *Epidemiology*, pages 363–366, 1991.

[19] M. A. Laaksonen, T. Härkänen, P. Knekt, E. Virtala, and H. Oja. Estimation of population attributable fraction (paf) for disease occurrence in a cohort study design. *Statistics in medicine*, 29(7-8):860–874, 2010.

[20] M. L. Levin. The occurrence of lung cancer in man. *Acta-Unio Internationalis Contra Cancrum*, 9(3):531, 1953.

[21] O. S. Miettinen. Proportion of disease caused or prevented by a given exposure, trait or intervention. *American journal of epidemiology*, 99(5):325–332, 1974.

[22] R. L. Prentice and R. Pyke. Logistic disease incidence models and case-control studies. *Biometrika*, 66(3):403–411, 1979.

[23] C. R. Rao. *Linear statistical inference and its applications*, volume 22. Wiley-Interscience, 2009.

[24] KENNETH J Rothman and CH Poole. A strengthening programme for weak associations. *International journal of epidemiology*, 17(4):955–959, 1988.

[25] Mats Rudemo. Empirical choice of histograms and kernel density estimators. *Scandinavian Journal of Statistics*, pages 65–78, 1982.

[26] S. O. Samuelsen and G. E. Eide. Attributable fractions with survival data. *Statistics in medicine*, 27(9):1447–1467, 2008.

[27] David W Scott and George R Terrell. Biased and unbiased cross-validation in density estimation. *Journal of the american Statistical association*, 82(400):1131–1146, 1987.

[28] Bernard W Silverman. *Density estimation for statistics and data analysis*, volume 26. CRC press, 1986.

[29] W. Uter and A. Pfahlberg. The concept of attributable risk in epidemiological practice. *Biometrical journal*, 41(8):985–993, 1999.

[30] S. D. Walter. The estimation and interpretation of attributable risk in health research. *Biometrics*, pages 829–849, 1976.

[31] S. D. Walter. Prevention for multifactorial diseases. *American Journal of Epidemiology*, 112(3):409–416, 1980.

[32] MP Wand and MC Jones. Kernel smoothing. 1995. *Chapman&Hall, London*, 1995.

[33] A. S. Whittemore. Statistical methods for estimating attributable risk from retrospective data. *Statistics in Medicine*, 1(3):229–243, 1982.

[34] Michael Woodroofe et al. On choosing a delta-sequence. *Ann. Math. Statist*, 41(5):1665–1671, 1970.

[35] Ronghui Xu and John O'Quigley. Proportional hazards estimate of the conditional survival function. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 62(4):667–680, 2000.

Appendix A

SIMULATION RESULTS FOR UNADJUSTED TIME-VARYING PAF

We reported the simulation results for the combined estimators, $\widehat{\phi}_w(t;\hat{\beta})$ and $\widehat{\phi}_{adjw}(t;\hat{\beta},\hat{\gamma})$, in Chapter 4. Now we report the results for $\widehat{\phi}_+(t;\hat{\beta})$, $\widehat{\phi}_-(t;\hat{\beta})$, $\widehat{\phi}_{adj+}(t;\hat{\beta},\hat{\gamma})$, and $\widehat{\phi}_{adj-}(t;\hat{\beta},\hat{\gamma})$.

A.0.1 Early onset common disease with binary time-independent exposure

The simulation results for $\widehat{\phi}_+(t;\hat{\beta})$ with various bandwidth choices are summarized in Table A.1 and Table A.2. The simulation results for $\widehat{\phi}_-(t;\hat{\beta})$ with various bandwidth choices are summarized in Table A.3 and Table A.4. These results show similar patterns as those for $\widehat{\phi}_w(t;\hat{\beta})$.

A.0.2 Late onset rare disease with binary time-independent exposure

The simulation results for $\widehat{\phi}_+(t;\hat{\beta})$ with various bandwidth choices are summarized in Table A.5 and Table A.6. The simulation results for $\widehat{\phi}_-(t;\hat{\beta})$ with various bandwidth choices are summarized in Table A.7 and Table A.8. These results show similar patterns as those for $\widehat{\phi}_w(t;\hat{\beta})$.

A.0.3 Late onset rare disease with continuous time-independent exposure

The simulation results for $\widehat{\phi}_+(t;\hat{\beta})$ with various bandwidth choices are summarized in Table A.9 and Table A.10. The simulation results for $\widehat{\phi}_-(t;\hat{\beta})$ with various bandwidth choices are summarized in Table A.11 and Table A.12. These results show similar patterns as those for $\widehat{\phi}_w(t;\hat{\beta})$.

A.0.4 Late onset rare disease with continuous time-dependent exposure

The simulation results for $\widehat{\phi}_+(t;\hat{\beta})$ with various bandwidth choices are summarized in Table A.13 and Table A.14. The simulation results for $\widehat{\phi}_-(t;\hat{\beta})$ with various bandwidth choices are summarized in Table A.15 and Table A.16. These results show similar patterns as those for $\widehat{\phi}_w(t;\hat{\beta})$.

Table A.1: Summary statistics for $\widehat{\phi}_{+}(t;\hat{\beta})$ with binary Z for the early onset scenario.

		h_{CV}				h=2				h=4			
Parameter	Age (yrs)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)
$\beta_0 = \log 2,\ p_z = 0.3$	10	0.0008	0.0366	0.0359	94.0	0.0001	0.0440	0.0445	94.0	0.0003	0.0371	0.0374	93.9
	20	0.0012	0.0310	0.0317	95.2	0.0015	0.0370	0.0384	94.9	0.0012	0.0318	0.0328	95.3
	30	0.0000	0.0288	0.0285	94.2	0.0000	0.0373	0.0364	93.3	0.0001	0.0302	0.0298	94.3
	40	0.0016	0.0270	0.0272	93.8	0.0002	0.0382	0.0379	92.5	0.0010	0.0289	0.0291	93.1
$\beta_0 = \log 2,\ p_z = 0.6$	10	0.0025	0.0426	0.0435	95.2	0.0020	0.0472	0.0481	95.1	0.0022	0.0432	0.0441	95.0
	20	0.0019	0.0406	0.0414	95.3	0.0015	0.0450	0.0455	94.8	0.0021	0.0411	0.0419	95.5
	30	0.0009	0.0397	0.0402	95.2	0.0013	0.0456	0.0463	94.7	0.0013	0.0406	0.0410	94.9
	40	0.0004	0.0421	0.0415	94.6	0.0012	0.0537	0.0532	94.1	0.0005	0.0432	0.0432	94.5
$\beta_0 = \log 3,\ p_z = 0.3$	10	0.0019	0.0413	0.0403	94.5	0.0017	0.0504	0.0495	94.2	0.0019	0.0405	0.0399	94.0
	20	0.0007	0.0355	0.0349	94.8	0.0004	0.0434	0.0433	94.3	0.0007	0.0346	0.0345	94.8
	30	0.0014	0.0333	0.0324	93.9	0.0009	0.0421	0.0424	94.0	0.0014	0.0323	0.0320	94.0
	40	0.0018	0.0310	0.0306	93.5	0.0006	0.0417	0.0418	91.0	0.0015	0.0298	0.0302	93.3
$\beta_0 = \log 3,\ p_z = 0.6$	10	0.0024	0.0393	0.0392	94.6	0.0020	0.0433	0.0437	94.2	0.0023	0.0386	0.0385	94.6
	20	0.0014	0.0393	0.0385	94.2	0.0016	0.0444	0.0436	94.1	0.0014	0.0379	0.0377	94.8
	30	0.0004	0.0433	0.0428	95.0	0.0019	0.0511	0.0521	95.3	0.0003	0.0413	0.0412	94.7
	40	0.0029	0.0573	0.0545	93.1	0.0025	0.0742	0.0710	91.5	0.0034	0.0542	0.0519	93.1

Bias, the sampling bias; SD, the sampling standard deviation; ASE, the sampling mean of the standard error estimator; CP, the coverage probability of the 95% pointwise confidence interval.

Table A.2: Summary statistics for $\widehat{\phi}_{+}(t;\hat{\beta})$ with binary Z for the early onset scenario, continued.

		h=6				h=8				h=10			
Parameter	Age (yrs)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)
$\beta_0 = \log 2$, $p_z = 0.3$	10	0.0008	0.0346	0.0347	94.5	0.0015	0.0331	0.0332	94.7	0.0023	0.0321	0.0323	94.8
	20	0.0009	0.0298	0.0307	95.4	0.0003	0.0287	0.0295	95.5	0.0003	0.0280	0.0288	95.5
	30	0.0001	0.0275	0.0273	94.6	0.0004	0.0261	0.0260	94.8	0.0009	0.0252	0.0252	94.5
	40	0.0019	0.0251	0.0256	94.3	0.0028	0.0231	0.0237	95.1	0.0039	0.0218	0.0225	95.5
$\beta_0 = \log 2$, $p_z = 0.6$	10	0.0025	0.0418	0.0427	95.5	0.0030	0.0412	0.0419	95.3	0.0036	0.0407	0.0415	95.7
	20	0.0019	0.0398	0.0407	95.5	0.0020	0.0393	0.0400	95.2	0.0023	0.0389	0.0396	95.1
	30	0.0010	0.0385	0.0391	95.5	0.0005	0.0374	0.0382	95.3	0.0002	0.0368	0.0376	95.5
	40	0.0008	0.0392	0.0393	95.0	0.0024	0.0371	0.0373	95.0	0.0041	0.0356	0.0360	95.2
$\beta_0 = \log 3$, $p_z = 0.3$	10	0.0027	0.0367	0.0361	94.3	0.0043	0.0347	0.0340	94.7	0.0065	0.0333	0.0327	94.5
	20	0.0011	0.0313	0.0311	94.8	0.0014	0.0296	0.0292	94.8	0.0018	0.0284	0.0280	94.8
	30	0.0023	0.0280	0.0277	94.8	0.0037	0.0255	0.0253	94.5	0.0056	0.0239	0.0239	94.5
	40	0.0037	0.0249	0.0252	94.2	0.0063	0.0223	0.0224	95.0	0.0095	0.0207	0.0206	94.5
$\beta_0 = \log 3$, $p_z = 0.6$	10	0.0032	0.0371	0.0367	94.2	0.0043	0.0362	0.0357	94.0	0.0057	0.0357	0.0351	94.0
	20	0.0015	0.0357	0.0355	94.5	0.0018	0.0346	0.0344	94.8	0.0021	0.0341	0.0337	94.8
	30	0.0011	0.0371	0.0369	94.7	0.0030	0.0349	0.0346	94.5	0.0055	0.0333	0.0331	94.2
	40	0.0067	0.0450	0.0436	93.3	0.0113	0.0395	0.0388	93.6	0.0173	0.0361	0.0356	93.2

Bias, the sampling bias; SD, the sampling standard deviation; ASE, the sampling mean of the standard error estimator; CP, the coverage probability of the 95% pointwise confidence interval.

Table A.3: Summary statistics for $\widehat{\phi}_{-}(t;\hat{\beta})$ with binary Z for the early onset scenario.

		h_{CV}				h=2				h=4			
Parameter	Age (yrs)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)
$\beta_0 = \log 2,\ p_z = 0.3$	10	0.0015	0.0367	0.0365	94.2	0.0030	0.0474	0.0472	92.8	0.0017	0.0383	0.0389	94.6
	20	0.0006	0.0322	0.0321	94.8	0.0001	0.0417	0.0410	93.6	0.0001	0.0339	0.0341	94.8
	30	0.0001	0.0303	0.0299	94.8	0.0005	0.0407	0.0406	93.8	0.0002	0.0321	0.0324	94.7
	40	0.0005	0.0308	0.0301	93.8	0.0029	0.0448	0.0444	91.0	0.0006	0.0337	0.0337	92.8
$\beta_0 = \log 2,\ p_z = 0.6$	10	0.0028	0.0420	0.0428	95.8	0.0033	0.0459	0.0469	95.0	0.0028	0.0423	0.0434	95.5
	20	0.0027	0.0402	0.0408	95.5	0.0032	0.0440	0.0445	94.6	0.0028	0.0404	0.0414	95.5
	30	0.0017	0.0391	0.0398	95.6	0.0031	0.0451	0.0457	95.2	0.0023	0.0399	0.0407	95.5
	40	0.0000	0.0415	0.0421	94.9	0.0023	0.0537	0.0549	95.4	0.0005	0.0433	0.0441	95.0
$\beta_0 = \log 3,\ p_z = 0.3$	10	0.0038	0.0426	0.0411	94.3	0.0039	0.0546	0.0541	94.0	0.0035	0.0430	0.0424	94.2
	20	0.0022	0.0385	0.0372	93.8	0.0035	0.0505	0.0502	93.2	0.0022	0.0389	0.0384	93.7
	30	0.0005	0.0408	0.0389	93.6	0.0036	0.0567	0.0554	92.9	0.0013	0.0415	0.0407	93.3
	40	0.0007	0.0434	0.0415	89.7	0.0056	0.0636	0.0586	79.1	0.0023	0.0451	0.0439	88.5
$\beta_0 = \log 3,\ p_z = 0.6$	10	0.0040	0.0373	0.0374	94.8	0.0040	0.0405	0.0409	94.5	0.0039	0.0367	0.0370	95.3
	20	0.0027	0.0381	0.0371	94.2	0.0034	0.0426	0.0416	94.0	0.0023	0.0371	0.0365	94.5
	30	0.0023	0.0442	0.0436	94.7	0.0057	0.0546	0.0542	94.5	0.0020	0.0418	0.0423	95.0
	40	0.0021	0.0682	0.0663	92.8	0.0100	0.0899	0.0874	90.5	0.0021	0.0640	0.0640	93.3

Bias, the sampling bias; SD, the sampling standard deviation; ASE, the sampling mean of the standard error estimator; CP, the coverage probability of the 95% pointwise confidence interval.

Table A.4: Summary statistics for $\widehat{\phi}_{-}(t;\hat{\beta})$ with binary Z for the early onset scenario, continued.

		h=6				h=8				h=10			
Parameter	Age (yrs)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)
$\beta_0 = \log 2,\ p_z = 0.3$	10	0.0014	0.0350	0.0357	95.2	0.0017	0.0334	0.0340	94.8	0.0023	0.0323	0.0329	94.5
	20	0.0000	0.0310	0.0315	95.0	0.0002	0.0295	0.0301	95.5	0.0005	0.0287	0.0293	95.3
	30	0.0005	0.0288	0.0291	95.0	0.0007	0.0271	0.0273	95.2	0.0013	0.0260	0.0263	95.3
	40	0.0004	0.0290	0.0291	93.7	0.0016	0.0264	0.0264	94.8	0.0032	0.0247	0.0247	95.2
$\beta_0 = \log 2,\ p_z = 0.6$	10	0.0026	0.0413	0.0422	96.0	0.0029	0.0408	0.0416	95.8	0.0035	0.0404	0.0412	96.2
	20	0.0026	0.0394	0.0403	95.6	0.0025	0.0389	0.0398	95.5	0.0026	0.0386	0.0395	95.3
	30	0.0016	0.0382	0.0389	95.5	0.0010	0.0374	0.0380	95.3	0.0003	0.0368	0.0375	95.0
	40	0.0007	0.0393	0.0399	94.7	0.0022	0.0370	0.0377	95.1	0.0044	0.0356	0.0364	95.1
$\beta_0 = \log 3,\ p_z = 0.3$	10	0.0038	0.0381	0.0377	94.5	0.0051	0.0353	0.0352	94.5	0.0070	0.0338	0.0337	94.2
	20	0.0017	0.0340	0.0337	94.2	0.0015	0.0313	0.0311	94.2	0.0015	0.0297	0.0295	94.8
	30	0.0009	0.0355	0.0342	93.3	0.0032	0.0318	0.0305	93.7	0.0061	0.0292	0.0281	94.0
	40	0.0009	0.0369	0.0366	91.2	0.0044	0.0320	0.0322	93.3	0.0088	0.0290	0.0291	93.7
$\beta_0 = \log 3,\ p_z = 0.6$	10	0.0043	0.0355	0.0357	95.2	0.0051	0.0350	0.0350	94.7	0.0065	0.0347	0.0346	94.8
	20	0.0019	0.0353	0.0348	94.2	0.0018	0.0344	0.0339	94.2	0.0018	0.0338	0.0334	94.5
	30	0.0005	0.0374	0.0376	94.6	0.0035	0.0349	0.0351	94.7	0.0074	0.0334	0.0335	94.8
	40	0.0041	0.0536	0.0530	94.0	0.0108	0.0474	0.0463	93.8	0.0190	0.0431	0.0417	91.9

Bias, the sampling bias; SD, the sampling standard deviation; ASE, the sampling mean of the standard error estimator; CP, the coverage probability of the 95% pointwise confidence interval.

Table A.5: Summary statistics for $\widehat{\phi}_+(t;\hat{\beta})$ with binary Z for the late onset scenario.

		h_{CV}				h=2				h=4			
Parameter	Age (yrs)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)
$\beta_0 = \log 2$, $p_z = 0.3$	40	0.0028	0.0640	0.0628	92.2	0.0026	0.0775	0.0751	90.8	0.0021	0.0579	0.0570	92.7
	50	0.0016	0.0399	0.0394	94.5	0.0009	0.0444	0.0444	93.9	0.0017	0.0372	0.0371	94.7
	60	0.0017	0.0319	0.0314	95.3	0.0016	0.0344	0.0343	95.0	0.0019	0.0305	0.0301	95.3
	70	0.0013	0.0256	0.0252	94.8	0.0006	0.0294	0.0290	93.9	0.0016	0.0237	0.0235	94.7
$\beta_0 = \log 2$, $p_z = 0.6$	40	0.0045	0.0638	0.0628	93.1	0.0035	0.0687	0.0676	92.5	0.0046	0.0558	0.0557	93.8
	50	0.0058	0.0471	0.0468	94.5	0.0059	0.0488	0.0485	94.0	0.0054	0.0448	0.0444	94.6
	60	0.0051	0.0426	0.0424	95.2	0.0046	0.0438	0.0434	95.0	0.0052	0.0411	0.0409	94.9
	70	0.0027	0.0414	0.0408	93.8	0.0037	0.0437	0.0431	93.6	0.0012	0.0379	0.0373	94.8
$\beta_0 = \log 3$, $p_z = 0.3$	40	0.0004	0.0817	0.0777	93.2	0.0003	0.0930	0.0880	91.8	0.0010	0.0662	0.0652	93.3
	50	0.0040	0.0444	0.0450	94.6	0.0037	0.0486	0.0496	94.5	0.0038	0.0385	0.0397	94.8
	60	0.0019	0.0349	0.0345	94.5	0.0013	0.0378	0.0376	94.5	0.0032	0.0308	0.0307	94.5
	70	0.0031	0.0270	0.0268	93.3	0.0008	0.0302	0.0304	92.9	0.0056	0.0225	0.0226	94.0
$\beta_0 = \log 3$, $p_z = 0.6$	40	0.0088	0.0664	0.0654	91.6	0.0090	0.0671	0.0660	91.5	0.0090	0.0531	0.0522	93.5
	50	0.0074	0.0439	0.0440	94.6	0.0074	0.0441	0.0442	94.7	0.0081	0.0385	0.0389	94.9
	60	0.0079	0.0403	0.0402	94.3	0.0079	0.0405	0.0404	94.2	0.0085	0.0358	0.0359	94.2
	70	0.0011	0.0513	0.0511	94.0	0.0008	0.0518	0.0515	93.9	0.0097	0.0389	0.0389	94.5

Bias, the sampling bias; SD, the sampling standard deviation; ASE, the sampling mean of the standard error estimator; CP, the coverage probability of the 95% pointwise confidence interval.

Table A.6: Summary statistics for $\widehat{\phi}_{+}(t;\hat{\beta})$ with binary Z for the late onset scenario, continued.

		h=6				h=8				h=10			
Parameter	Age (yrs)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)
$\beta_0 = \log 2$, $p_z = 0.3$	40	0.0023	0.0499	0.0489	93.1	0.0028	0.0449	0.0441	93.4	0.0030	0.0415	0.0409	93.7
	50	0.0023	0.0345	0.0344	95.0	0.0031	0.0329	0.0328	94.9	0.0042	0.0318	0.0318	95.0
	60	0.0030	0.0289	0.0285	94.8	0.0044	0.0279	0.0276	94.7	0.0060	0.0273	0.0270	94.5
	70	0.0041	0.0218	0.0216	94.4	0.0073	0.0209	0.0208	93.5	0.0111	0.0206	0.0205	92.3
$\beta_0 = \log 2$, $p_z = 0.6$	40	0.0049	0.0505	0.0509	94.4	0.0051	0.0475	0.0482	94.8	0.0054	0.0459	0.0464	95.0
	50	0.0056	0.0431	0.0430	95.0	0.0063	0.0422	0.0422	95.4	0.0070	0.0417	0.0416	95.3
	60	0.0061	0.0402	0.0399	94.7	0.0072	0.0397	0.0394	95.2	0.0086	0.0392	0.0390	95.2
	70	0.0025	0.0360	0.0354	95.0	0.0071	0.0351	0.0347	94.5	0.0122	0.0347	0.0344	94.0
$\beta_0 = \log 3$, $p_z = 0.3$	40	0.0014	0.0543	0.0550	94.7	0.0022	0.0478	0.0488	94.5	0.0030	0.0436	0.0446	95.0
	50	0.0046	0.0345	0.0357	95.1	0.0064	0.0325	0.0334	95.2	0.0090	0.0311	0.0319	94.8
	60	0.0053	0.0280	0.0280	94.8	0.0081	0.0264	0.0265	93.8	0.0112	0.0253	0.0254	93.0
	70	0.0131	0.0193	0.0196	92.7	0.0229	0.0178	0.0183	79.5	0.0340	0.0172	0.0177	52.7
$\beta_0 = \log 3$, $p_z = 0.6$	40	0.0092	0.0468	0.0464	93.8	0.0095	0.0435	0.0432	93.8	0.0098	0.0411	0.0410	94.5
	50	0.0091	0.0366	0.0370	94.8	0.0105	0.0355	0.0359	95.0	0.0122	0.0348	0.0352	94.7
	60	0.0100	0.0341	0.0343	94.3	0.0121	0.0332	0.0334	94.3	0.0145	0.0326	0.0328	94.0
	70	0.0231	0.0339	0.0339	91.1	0.0394	0.0314	0.0315	77.7	0.0570	0.0301	0.0302	53.5

Bias, the sampling bias; SD, the sampling standard deviation; ASE, the sampling mean of the standard error estimator; CP, the coverage probability of the 95% pointwise confidence interval.

Table A.7: Summary statistics for $\widehat{\phi}_{-}(t;\hat{\beta})$ with binary Z for the late onset scenario.

		h_{CV}				h=2				h=4			
Parameter	Age (yrs)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)
$\beta_0 = \log 2,\, p_z = 0.3$	40	0.0060	0.0645	0.0617	91.6	0.0086	0.0818	0.0771	90.4	0.0046	0.0605	0.0591	91.7
	50	0.0040	0.0402	0.0393	94.5	0.0047	0.0467	0.0460	93.8	0.0041	0.0385	0.0381	94.4
	60	0.0035	0.0322	0.0316	94.0	0.0034	0.0360	0.0359	93.0	0.0035	0.0313	0.0309	94.2
	70	0.0001	0.0277	0.0275	93.5	0.0028	0.0336	0.0338	93.0	0.0003	0.0259	0.0265	94.5
$\beta_0 = \log 2,\, p_z = 0.6$	40	0.0091	0.0588	0.0572	94.2	0.0108	0.0647	0.0631	93.5	0.0083	0.0529	0.0527	95.0
	50	0.0069	0.0447	0.0449	95.5	0.0072	0.0462	0.0468	95.5	0.0062	0.0436	0.0436	94.9
	60	0.0068	0.0409	0.0413	94.8	0.0066	0.0422	0.0427	94.7	0.0070	0.0401	0.0404	94.9
	70	0.0043	0.0416	0.0407	93.0	0.0061	0.0450	0.0445	92.8	0.0028	0.0385	0.0380	93.6
$\beta_0 = \log 3,\, p_z = 0.3$	40	0.0099	0.0806	0.0742	92.3	0.0160	0.1006	0.0928	90.5	0.0078	0.0708	0.0686	93.2
	50	0.0067	0.0438	0.0442	94.1	0.0068	0.0519	0.0529	94.5	0.0059	0.0406	0.0416	94.3
	60	0.0063	0.0356	0.0355	93.7	0.0055	0.0421	0.0427	94.3	0.0061	0.0332	0.0334	94.0
	70	0.0025	0.0367	0.0354	91.5	0.0036	0.0461	0.0443	87.1	0.0028	0.0330	0.0326	92.5
$\beta_0 = \log 3,\, p_z = 0.6$	40	0.0138	0.0534	0.0531	95.2	0.0149	0.0583	0.0579	94.8	0.0126	0.0460	0.0467	95.6
	50	0.0119	0.0392	0.0394	94.8	0.0116	0.0408	0.0410	94.7	0.0122	0.0366	0.0372	95.0
	60	0.0136	0.0370	0.0373	93.5	0.0136	0.0389	0.0390	93.8	0.0135	0.0345	0.0351	93.5
	70	0.0040	0.0611	0.0566	92.1	0.0101	0.0667	0.0637	91.6	0.0047	0.0477	0.0466	93.9

Bias, the sampling bias; SD, the sampling standard deviation; ASE, the sampling mean of the standard error estimator; CP, the coverage probability of the 95% pointwise confidence interval.

Table A.8: Summary statistics for $\widehat{\phi}_{-}(t;\hat{\beta})$ with binary Z for the late onset scenario, continued.

		h=6				h=8				h=10			
Parameter	Age (yrs)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)
$\beta_0 = \log 2,\ p_z = 0.3$	40	0.0038	0.0523	0.0512	92.0	0.0029	0.0473	0.0459	93.2	0.0027	0.0438	0.0422	93.5
	50	0.0038	0.0358	0.0351	94.2	0.0040	0.0341	0.0334	94.2	0.0049	0.0329	0.0322	94.0
	60	0.0040	0.0293	0.0290	94.0	0.0050	0.0281	0.0279	94.5	0.0065	0.0275	0.0272	94.0
	70	0.0042	0.0231	0.0237	95.3	0.0082	0.0221	0.0224	94.4	0.0118	0.0216	0.0218	92.7
$\beta_0 = \log 2,\ p_z = 0.6$	40	0.0079	0.0488	0.0489	94.6	0.0075	0.0466	0.0466	94.4	0.0074	0.0453	0.0451	94.5
	50	0.0064	0.0425	0.0424	95.1	0.0070	0.0419	0.0418	94.8	0.0078	0.0416	0.0413	95.2
	60	0.0073	0.0396	0.0396	94.8	0.0082	0.0392	0.0392	95.0	0.0096	0.0389	0.0389	95.2
	70	0.0023	0.0362	0.0358	94.4	0.0079	0.0352	0.0350	94.2	0.0132	0.0349	0.0347	93.0
$\beta_0 = \log 3,\ p_z = 0.3$	40	0.0063	0.0598	0.0583	94.2	0.0059	0.0518	0.0514	94.4	0.0064	0.0464	0.0464	94.2
	50	0.0067	0.0359	0.0371	94.3	0.0082	0.0331	0.0345	95.0	0.0107	0.0313	0.0327	94.9
	60	0.0063	0.0297	0.0297	93.5	0.0078	0.0277	0.0277	93.8	0.0103	0.0262	0.0265	93.5
	70	0.0136	0.0275	0.0274	94.6	0.0252	0.0244	0.0248	87.4	0.0370	0.0227	0.0232	67.8
$\beta_0 = \log 3,\ p_z = 0.6$	40	0.0123	0.0421	0.0427	95.5	0.0127	0.0399	0.0404	95.2	0.0132	0.0383	0.0388	95.3
	50	0.0129	0.0351	0.0359	95.0	0.0142	0.0344	0.0352	94.8	0.0159	0.0339	0.0347	94.1
	60	0.0137	0.0329	0.0337	93.8	0.0146	0.0323	0.0331	94.1	0.0163	0.0319	0.0327	93.7
	70	0.0243	0.0393	0.0389	91.0	0.0447	0.0355	0.0350	76.3	0.0646	0.0334	0.0328	50.8

Bias, the sampling bias; SD, the sampling standard deviation; ASE, the sampling mean of the standard error estimator; CP, the coverage probability of the 95% pointwise confidence interval.

Table A.9: Summary statistics for $\widehat{\phi}_+(t;\hat{\beta})$ with continuous Z for the late onset scenario.

		h_{CV}				h=2				h=4			
Parameter	Age (yrs)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)
$\beta_0 = \log 1.3$	40	0.0068	0.0640	0.0627	93.8	0.0073	0.0687	0.0667	93.0	0.0067	0.0589	0.0591	94.5
	50	0.0076	0.0528	0.0531	95.5	0.0074	0.0541	0.0543	95.6	0.0075	0.0517	0.0520	95.3
	60	0.0064	0.0499	0.0501	95.4	0.0061	0.0506	0.0508	95.3	0.0066	0.0493	0.0494	95.3
	70	0.0038	0.0466	0.0466	94.7	0.0047	0.0477	0.0480	94.9	0.0027	0.0452	0.0455	95.1
$\beta_0 = \log 1.6$	40	0.0089	0.0535	0.0527	93.2	0.0085	0.0541	0.0534	93.2	0.0085	0.0444	0.0449	94.2
	50	0.0087	0.0402	0.0402	94.4	0.0087	0.0403	0.0405	94.2	0.0089	0.0370	0.0375	94.8
	60	0.0074	0.0379	0.0387	95.0	0.0074	0.0382	0.0390	95.2	0.0079	0.0354	0.0364	94.7
	70	0.0025	0.0433	0.0451	94.9	0.0031	0.0441	0.0457	94.8	0.0032	0.0366	0.0382	95.5

Bias, the sampling bias; SD, the sampling standard deviation; ASE, the sampling mean of the standard error estimator; CP, the coverage probability of the 95% pointwise confidence interval.

Table A.10: Summary statistics for $\widehat{\phi}_{+}(t;\hat{\beta})$ with continuous Z for the late onset scenario, continued.

		h=6				h=8				h=10			
Parameter	Age (yrs)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)
$\beta_0 = \log 1.3$	40	0.0066	0.0555	0.0561	94.8	0.0066	0.0537	0.0544	95.3	0.0067	0.0527	0.0534	95.5
	50	0.0078	0.0506	0.0512	95.9	0.0082	0.0501	0.0508	95.9	0.0087	0.0499	0.0505	95.7
	60	0.0073	0.0488	0.0490	95.2	0.0082	0.0486	0.0487	95.2	0.0093	0.0483	0.0485	95.5
	70	0.0003	0.0448	0.0448	95.0	0.0039	0.0447	0.0446	94.5	0.0079	0.0448	0.0447	93.3
$\beta_0 = \log 1.6$	40	0.0084	0.0406	0.0415	95.5	0.0088	0.0384	0.0397	95.5	0.0093	0.0372	0.0386	95.7
	50	0.0099	0.0357	0.0365	95.1	0.0113	0.0349	0.0359	95.5	0.0132	0.0345	0.0356	95.0
	60	0.0086	0.0346	0.0354	94.7	0.0096	0.0341	0.0350	94.7	0.0105	0.0337	0.0347	94.8
	70	0.0137	0.0342	0.0355	93.7	0.0269	0.0331	0.0343	88.1	0.0415	0.0326	0.0338	76.8

Bias, the sampling bias; SD, the sampling standard deviation; ASE, the sampling mean of the standard error estimator; CP, the coverage probability of the 95% pointwise confidence interval.

Table A.11: Summary statistics for $\widehat{\phi}_{-}(t;\hat{\beta})$ with continuous Z for the late onset scenario.

		h_{CV}				h=2				h=4			
Parameter	Age (yrs)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)
$\beta_0 = \log 1.3$	40	0.0099	0.0608	0.0602	94.5	0.0109	0.0662	0.0650	93.7	0.0093	0.0581	0.0582	94.7
	50	0.0086	0.0524	0.0525	95.0	0.0087	0.0538	0.0539	94.9	0.0084	0.0517	0.0518	94.7
	60	0.0080	0.0495	0.0497	95.4	0.0081	0.0507	0.0506	94.9	0.0078	0.0491	0.0494	95.3
	70	0.0053	0.0461	0.0466	95.0	0.0073	0.0476	0.0486	95.2	0.0045	0.0452	0.0457	94.8
$\beta_0 = \log 1.6$	40	0.0116	0.0459	0.0464	95.2	0.0125	0.0500	0.0504	94.7	0.0114	0.0424	0.0434	95.0
	50	0.0118	0.0369	0.0384	95.3	0.0121	0.0385	0.0399	95.2	0.0118	0.0360	0.0373	95.2
	60	0.0116	0.0366	0.0377	94.7	0.0119	0.0380	0.0393	94.1	0.0115	0.0353	0.0365	95.0
	70	0.0016	0.0468	0.0448	93.0	0.0079	0.0522	0.0498	91.6	0.0022	0.0410	0.0410	94.5

Bias, the sampling bias; SD, the sampling standard deviation; ASE, the sampling mean of the standard error estimator; CP, the coverage probability of the 95% pointwise confidence interval.

Table A.12: Summary statistics for $\widehat{\phi}_{-}(t;\hat{\beta})$ with continuous Z for the late onset scenario, continued.

		h=6				h=8				h=10			
Parameter	Age (yrs)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)
$\beta_0 = \log 1.3$	40	0.0089	0.0548	0.0557	95.5	0.0090	0.0532	0.0541	95.5	0.0089	0.0523	0.0530	95.7
	50	0.0083	0.0509	0.0511	95.1	0.0087	0.0505	0.0507	95.3	0.0094	0.0502	0.0504	95.6
	60	0.0080	0.0486	0.0489	95.5	0.0085	0.0484	0.0487	95.6	0.0096	0.0483	0.0485	95.7
	70	0.0002	0.0448	0.0450	94.7	0.0044	0.0449	0.0449	94.2	0.0087	0.0450	0.0449	93.8
$\beta_0 = \log 1.6$	40	0.0115	0.0397	0.0408	95.2	0.0116	0.0380	0.0392	95.2	0.0119	0.0369	0.0382	95.1
	50	0.0123	0.0351	0.0363	94.8	0.0134	0.0346	0.0359	95.0	0.0153	0.0342	0.0356	94.7
	60	0.0103	0.0343	0.0356	95.2	0.0097	0.0339	0.0351	95.2	0.0099	0.0337	0.0349	95.0
	70	0.0161	0.0368	0.0376	92.6	0.0311	0.0349	0.0360	86.1	0.0469	0.0340	0.0352	73.7

Bias, the sampling bias; SD, the sampling standard deviation; ASE, the sampling mean of the standard error estimator; CP, the coverage probability of the 95% pointwise confidence interval.

Table A.13: Summary statistics for $\widehat{\phi}_+(t;\hat{\beta})$ with continuous time-varying exposure for the late onset scenario.

		h_{CV}				h=2				h=4			
Parameter	Age (yrs)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)
$\beta_0 = \log 1.3$	40	0.0028	0.0557	0.0555	93.8	0.0028	0.0574	0.0574	93.4	0.0033	0.0496	0.0488	93.7
	50	0.0016	0.0490	0.0477	93.8	0.0015	0.0495	0.0482	93.7	0.0016	0.0467	0.0456	93.6
	60	0.0002	0.0478	0.0476	94.1	0.0002	0.0480	0.0480	94.4	0.0008	0.0464	0.0465	94.2
	70	0.0018	0.0485	0.0485	95.0	0.0011	0.0490	0.0494	95.0	0.0041	0.0459	0.0455	94.5
$\beta_0 = \log 1.6$	40	0.0005	0.0578	0.0543	91.3	0.0006	0.0580	0.0544	91.3	0.0010	0.0457	0.0442	93.0
	50	0.0022	0.0385	0.0381	94.1	0.0022	0.0386	0.0382	94.1	0.0018	0.0352	0.0350	93.8
	60	0.0039	0.0372	0.0366	94.4	0.0039	0.0372	0.0367	94.4	0.0054	0.0345	0.0341	94.7
	70	0.0024	0.0507	0.0496	93.3	0.0024	0.0508	0.0497	93.3	0.0115	0.0407	0.0401	93.7

Bias, the sampling bias; SD, the sampling standard deviation; ASE, the sampling mean of the standard error estimator; CP, the coverage probability of the 95% pointwise confidence interval.

Table A.14: Summary statistics for $\widehat{\phi}_{+}(t;\hat{\beta})$ with continuous time-varying exposure for the late onset scenario, continued.

		h=6				h=8				h=10			
Parameter	Age (yrs)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)
$\beta_0 = \log 1.3$	40	0.0066	0.0465	0.0456	93.8	0.0110	0.0448	0.0440	93.0	0.0160	0.0438	0.0431	92.1
	50	0.0029	0.0456	0.0448	94.1	0.0048	0.0450	0.0444	94.3	0.0070	0.0447	0.0442	93.7
	60	0.0019	0.0459	0.0459	94.4	0.0032	0.0457	0.0456	94.8	0.0051	0.0455	0.0453	95.1
	70	0.0079	0.0449	0.0443	94.4	0.0129	0.0446	0.0439	93.3	0.0179	0.0446	0.0437	91.9
$\beta_0 = \log 1.6$	40	0.0046	0.0411	0.0401	93.1	0.0095	0.0383	0.0378	92.8	0.0157	0.0366	0.0364	91.5
	50	0.0016	0.0341	0.0339	94.6	0.0011	0.0335	0.0334	94.6	0.0007	0.0331	0.0330	94.8
	60	0.0073	0.0337	0.0332	95.2	0.0095	0.0332	0.0328	95.0	0.0121	0.0328	0.0325	94.4
	70	0.0250	0.0365	0.0365	88.4	0.0425	0.0344	0.0346	76.2	0.0617	0.0332	0.0336	54.4

Bias, the sampling bias; SD, the sampling standard deviation; ASE, the sampling mean of the standard error estimator; CP, the coverage probability of the 95% pointwise confidence interval.

Table A.15: Summary statistics for $\widehat{\phi}_{-}(t;\hat{\beta})$ with continuous time-varying exposure for the late onset scenario.

		h_{CV}				h=2				h=4			
Parameter	Age (yrs)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)
$\beta_0 = \log 1.3$	40	0.0035	0.0548	0.0530	92.6	0.0016	0.0571	0.0553	93.1	0.0043	0.0493	0.0479	93.1
	50	0.0012	0.0478	0.0471	94.7	0.0006	0.0484	0.0479	94.8	0.0017	0.0458	0.0455	94.4
	60	0.0011	0.0482	0.0474	94.8	0.0008	0.0488	0.0479	94.9	0.0010	0.0471	0.0465	94.5
	70	0.0003	0.0487	0.0498	95.2	0.0018	0.0494	0.0513	95.5	0.0033	0.0460	0.0466	95.1
$\beta_0 = \log 1.6$	40	0.0043	0.0522	0.0511	93.9	0.0048	0.0527	0.0516	93.8	0.0003	0.0427	0.0427	94.2
	50	0.0025	0.0368	0.0372	95.4	0.0026	0.0370	0.0373	95.2	0.0021	0.0341	0.0347	95.2
	60	0.0048	0.0379	0.0370	93.4	0.0047	0.0382	0.0372	93.4	0.0055	0.0352	0.0344	94.0
	70	0.0024	0.0565	0.0547	91.9	0.0034	0.0564	0.0552	92.1	0.0097	0.0430	0.0444	95.1

Bias, the sampling bias; SD, the sampling standard deviation; ASE, the sampling mean of the standard error estimator; CP, the coverage probability of the 95% pointwise confidence interval.

Table A.16: Summary statistics for $\widehat{\phi}_{-}(t;\hat{\beta})$ with continuous time-varying exposure for the late onset scenario, continued.

		h=6				h=8				h=10			
Parameter	Age (yrs)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)
$\beta_0 = \log 1.3$	40	0.0066	0.0466	0.0452	93.8	0.0108	0.0451	0.0436	92.7	0.0168	0.0444	0.0429	91.3
	50	0.0027	0.0451	0.0447	94.9	0.0046	0.0447	0.0444	94.2	0.0075	0.0446	0.0443	93.7
	60	0.0020	0.0464	0.0459	94.5	0.0035	0.0460	0.0456	94.5	0.0052	0.0457	0.0454	94.8
	70	0.0095	0.0450	0.0451	93.8	0.0153	0.0447	0.0446	91.9	0.0203	0.0445	0.0444	90.1
$\beta_0 = \log 1.6$	40	0.0028	0.0396	0.0393	94.7	0.0090	0.0375	0.0371	93.5	0.0173	0.0361	0.0357	91.1
	50	0.0022	0.0334	0.0338	95.6	0.0017	0.0331	0.0333	95.9	0.0008	0.0329	0.0330	96.0
	60	0.0064	0.0341	0.0334	94.7	0.0078	0.0336	0.0330	94.0	0.0098	0.0334	0.0328	93.9
	70	0.0278	0.0388	0.0396	88.9	0.0482	0.0365	0.0373	74.4	0.0697	0.0356	0.0360	51.2

Bias, the sampling bias; SD, the sampling standard deviation; ASE, the sampling mean of the standard error estimator; CP, the coverage probability of the 95% pointwise confidence interval.

Appendix B

SIMULATION RESULTS FOR ADJUSTED TIME-VARYING PAF

The simulation results for $\widehat{\phi}_{adj+}(t; \hat{\beta}, \hat{\gamma})$ with various bandwidth choices are summarized in Table B.1 and Table B.2. The simulation results for $\widehat{\phi}_{adj-}(t; \hat{\beta}, \hat{\gamma})$ with various bandwidth choices are summarized in Table B.3 and Table B.4. These results show similar patterns as those for $\widehat{\phi}_{adjw}(t; \hat{\beta}, \hat{\gamma})$.

Table B.1: Summary statistics for $\widehat{\phi}_{adj+}(t;\hat{\beta},\hat{\gamma})$ with continuous Z and U.

		h_{CV}				h=2				h=4			
Parameter	Age (yrs)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)
$\beta_0 = \log 1.3$	40	0.0060	0.0858	0.0833	92.7	0.0056	0.0854	0.0839	92.5	0.0069	0.0645	0.0652	94.4
	50	0.0080	0.0536	0.0539	95.3	0.0078	0.0537	0.0541	94.3	0.0090	0.0459	0.0464	95.3
	60	0.0081	0.0459	0.0460	94.5	0.0078	0.0461	0.0464	94.9	0.0079	0.0404	0.0405	94.4
	70	0.0013	0.0630	0.0589	93.9	0.0005	0.0616	0.0599	93.9	0.0044	0.0437	0.0441	94.7
$\beta_0 = \log 1.6$	40	0.0054	0.0580	0.0571	91.4	0.0052	0.0592	0.0595	91.8	0.0068	0.0457	0.0451	93.4
	50	0.0071	0.0425	0.0424	94.4	0.0069	0.0441	0.0435	93.9	0.0084	0.0353	0.0353	95.1
	60	0.0063	0.0523	0.0509	94.2	0.0069	0.0534	0.0522	93.6	0.0046	0.0418	0.0404	94.5
	70	0.0071	0.1367	0.1234	89.2	0.0016	0.1385	0.1325	89.7	0.0146	0.0953	0.0922	89.8

Bias, the sampling bias; SD, the sampling standard deviation; ASE, the sampling mean of the standard error estimator; CP, the coverage probability of the 95% pointwise confidence interval.

Table B.2: Summary statistics for $\widehat{\phi}_{adj+}(t;\hat{\beta},\hat{\gamma})$ with continuous Z and U, continued.

		h=6				h=8				h=10			
Parameter	Age (yrs)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)
$\beta_0 = \log 1.3$	40	0.0079	0.0560	0.0573	94.8	0.0084	0.0512	0.0527	95.0	0.0094	0.0479	0.0495	95.3
	50	0.0106	0.0429	0.0435	95.3	0.0124	0.0412	0.0419	96.1	0.0148	0.0402	0.0408	95.7
	60	0.0091	0.0383	0.0383	95.1	0.0110	0.0372	0.0371	94.5	0.0133	0.0363	0.0363	94.3
	70	0.0116	0.0369	0.0378	94.2	0.0214	0.0336	0.0345	90.1	0.0329	0.0318	0.0326	82.8
$\beta_0 = \log 1.6$	40	0.0086	0.0397	0.0394	94.7	0.0108	0.0361	0.0359	94.8	0.0134	0.0335	0.0336	95.1
	50	0.0109	0.0321	0.0321	93.9	0.0143	0.0302	0.0305	92.8	0.0190	0.0293	0.0295	91.1
	60	0.0035	0.0371	0.0356	93.4	0.0018	0.0345	0.0330	92.9	0.0000	0.0327	0.0314	92.7
	70	0.0333	0.0733	0.0735	87.2	0.0568	0.0602	0.0619	77.4	0.0842	0.0518	0.0538	58.7

Bias, the sampling bias; SD, the sampling standard deviation; ASE, the sampling mean of the standard error estimator; CP, the coverage probability of the 95% pointwise confidence interval.

Table B.3: Summary statistics for $\widehat{\phi}_{adj-}(t;\hat{\beta},\hat{\gamma})$ with continuous Z and U.

		h_{CV}				h=2				h=4			
Parameter	Age (yrs)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)
$\beta_0 = \log 1.3$	40	0.0218	0.0755	0.0672	91.3	0.0247	0.0781	0.0731	89.7	0.0174	0.0603	0.0596	92.9
	50	0.0143	0.0481	0.0476	94.3	0.0138	0.0512	0.0503	94.2	0.0134	0.0442	0.0446	94.4
	60	0.0132	0.0424	0.0416	93.3	0.0135	0.0442	0.0437	93.0	0.0132	0.0398	0.0390	94.0
	70	0.0036	0.0467	0.0469	95.1	0.0076	0.0522	0.0526	93.9	0.0015	0.0379	0.0395	95.2
$\beta_0 = \log 1.6$	40	0.0201	0.0524	0.0461	90.9	0.0224	0.0566	0.0523	90.3	0.0179	0.0431	0.0421	92.1
	50	0.0165	0.0380	0.0363	92.0	0.0160	0.0417	0.0397	92.1	0.0156	0.0344	0.0334	92.2
	60	0.0113	0.0422	0.0408	93.8	0.0151	0.0468	0.0460	92.8	0.0105	0.0367	0.0367	93.8
	70	0.0002	0.1013	0.0891	91.3	0.0200	0.1116	0.1069	92.6	0.0083	0.0772	0.0744	92.8

Bias, the sampling bias; SD, the sampling standard deviation; ASE, the sampling mean of the standard error estimator; CP, the coverage probability of the 95% pointwise confidence interval.

Table B.4: Summary statistics for $\widehat{\phi}_{adj-}(t;\hat{\beta},\hat{\gamma})$ with continuous Z and U, continued.

		h=6				h=8				h=10			
Parameter	Age (yrs)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)	Bias	SD	ASE	CP(%)
$\beta_0 = \log 1.3$	40	0.0155	0.0541	0.0539	93.5	0.0149	0.0499	0.0502	93.5	0.0149	0.0469	0.0475	94.4
	50	0.0139	0.0416	0.0424	94.2	0.0158	0.0402	0.0412	94.3	0.0186	0.0393	0.0404	93.6
	60	0.0127	0.0381	0.0374	93.6	0.0131	0.0369	0.0366	94.2	0.0144	0.0361	0.0361	93.7
	70	0.0131	0.0328	0.0344	94.0	0.0266	0.0302	0.0320	90.1	0.0404	0.0296	0.0309	76.9
$\beta_0 = \log 1.6$	40	0.0174	0.0379	0.0377	92.7	0.0184	0.0347	0.0349	92.3	0.0207	0.0328	0.0331	91.0
	50	0.0160	0.0317	0.0310	91.4	0.0179	0.0303	0.0299	91.0	0.0212	0.0296	0.0293	89.2
	60	0.0034	0.0331	0.0330	94.6	0.0040	0.0315	0.0312	93.9	0.0104	0.0306	0.0302	92.8
	70	0.0400	0.0617	0.0587	85.7	0.0758	0.0515	0.0495	63.7	0.1139	0.0445	0.0431	24.3

Bias, the sampling bias; SD, the sampling standard deviation; ASE, the sampling mean of the standard error estimator; CP, the coverage probability of the 95% pointwise confidence interval.

VITA

Wei Zhao was born and grew up in Yantai, Shandong, China. He received a Bachelor of Science in Biosciences from the University of Science and Technology of China in 2006, and a Master of Science in Biotechnology from the University of California, Irvine in 2008. He received a Doctor of Philosophy in Biostatistics from the University of Washington 2014.