Weighted Accelerated Failure Time Model

by

© **Ayesha Madhushani Rathnayake Nayaka Bandaralage**

A thesis submitted to the School of Graduate Studies
in partial fulfillment of the requirements for the degree
of Master of Science in Statistics.

Department of Mathematics and Statistics
Memorial University

May 2025

St. John's, Newfoundland and Labrador, Canada

Abstract

The accelerated failure time (AFT) model is widely used in survival analysis and auxiliary information can be used to improve the efficiency of the model. We developed a weighted AFT model by using empirical likelihood probabilities as weights based on information from previous studies. The proposed model effectively overcomes the challenges associated with managing censored observations, resulting in more reliable and accurate estimates. Theoretical justifications of the proposed model are developed.

A comprehensive simulation study was conducted to assess the effectiveness of the proposed weighted models, incorporating both partial and complete auxiliary information. Both the Standard Accelerated Failure Time (AFT) and AFT with Generalized Estimating Equations (AFTGEE) models were employed for this comparative analysis. The simulation results suggest that when estimating coefficients, weighted models incorporating complete or partial auxiliary information on the linked covariate provide more accurate estimates compared to the model without any weights. Finally, the proposed method was implemented on a real dataset, illustrating its ability to accurately determine coefficients, minimize standard errors, and enhance significance levels by incorporating auxiliary information.

This thesis is dedicated to my family, teachers, and friends, whose unwavering support brought my dream academic journey to life.

Acknowledgements

I express my sincere appreciation to my supervisor, Dr. Asokan Mulayath Variyath, and Dr. Zhaozhi Fan from the Department of Mathematics and Statistics. Their remarkable proficiency and expertise significantly contributed to the success of this thesis, providing me with an exceptionally enriching research experience.

I would like to express genuine gratitude to the School of Graduate Studies, the Department of Mathematics and Statistics, Dr. Variyath, and Dr. Fan for their substantial financial support through graduate assistantships and teaching assistantships. Additionally, I appreciate the assistance provided by the Department of Mathematics and Statistics staff.

My sincere gratitude goes to my family, whose unwavering encouragement has been a constant source of stability throughout my academic pursuit.

I want to express appreciation to my friends and colleagues for their valuable contributions that enriched my academic journey.

Last but not least, I want to express my deep gratitude to all individuals who have directly or indirectly contributed to the completion of this work.

Table of contents

List of tables

List of figures

Chapter 1

Introduction

1.1 Survival Analysis

Survival analysis is a statistical methodology used to analyze and interpret time-to-event data, involves various models and techniques, including the key tools in analyzing time-to-event data such as Kaplan-Meier estimator, Cox Proportional Hazards model, Parametric Survival Models, and Accelerated Failure Time (AFT) models. These techniques aim not only to estimate the probability of an event occurring over time but also to identify factors associated with the occurrence of the event. Survival analysis was initially developed for assessing patient survival rates in medical research, this analytical method has expanded well beyond its original scope. It now finds widespread use in numerous fields, including finance, engineering, sociology, and others.

Ajay et al. [2021] discussed the applications of survival analysis, highlighting its broad utility across health and economic studies. Emmerson and Brown [2021] provide an overview of survival analysis in clinical trials, focusing on its application to time-to-event data that often involves censoring. The researchers explain commonly used techniques, such as Kaplan-Meier plots for visualizing survival curves, as well as statistical tests like the log-rank and Wilcoxon tests to evaluate the significance of differences between groups. Additionally, they highlight the usage of hazard ratios as a measure for comparing the impact of treatments on survival outcomes.

Censoring is a crucial concept in survival analysis that addresses situations where the exact time of the event of interest is not observed for all subjects in the study. The situation may occur due to various factors like participant dropout before experiencing the event, the study ending before the event occurs for all individuals, or missing follow-up data. Censoring is classified into three main types right censoring, left censoring, and interval censoring. Right censoring is the most common, where participants haven't experienced the event by the study's end, often due to premature departure, early study conclusion, or incomplete follow-up information. Left censoring involves events occurring before the study starts, with unknown timing, while interval censoring is when the event happened within a known time range but with uncertain timing.

In this study, we employed right censoring, where subjects' survival times (Y) were observed as the minimum of either the censoring time (C) or the failure time (T), whichever occurred first. The recorded data consists of three elements $(Y, \delta, \boldsymbol{X})$, where δ is the censoring indicator defined as $I(T \leq C)$, and $\boldsymbol{X}$ represents the covariate vector. Let T be a continuous random variable, with $f(t)$ representing the probability density function and $F(t) = P(T \leq t)$ as the cumulative distribution function. The probability of an individual's survival beyond a given time t can be mathematically expressed using the survival function.

$$S(t) = P(T > t) = 1 - F(t) = \int_t^{\infty} f(u)du$$

The hazard function denoted as $h(t)$ represents the instantaneous failure rate within a short interval $[t, t + \Delta t)$, given that the subject has survived up to time t.

Mathematically, the hazard function is expressed as,

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} = \frac{f(t)}{S(t)} = -\frac{d}{dt} \log S(t)$$

Several statistical methods have been developed to address the challenges caused by censored data in survival analysis. The methods include non-parametric techniques such as the Kaplan-Meier method and log-rank test, semi-parametric techniques like

the Cox Proportional Hazard (PH) model, and parametric models such as the Parametric PH model and Accelerated Failure Time (AFT) model. While the Cox PH model is widely employed in modeling survival data, the Parametric AFT model introduces a distinctive perspective. The AFT model expresses survival time logarithmically as a linear function of covariates, providing an alternative approach that diverges from the proportional hazard assumptions inherent in the Cox PH model.

Turkson et al. [2021] conducted a study that investigated various methods for addressing censoring in survival analysis. Their research highlighted the potential for bias and reduced statistical power when censoring is not effectively managed. They explored diverse approaches, emphasizing the importance of integrating incomplete information to enhance the comprehension of the study and effectively reduce biases.

Another method in survival analysis is the Buckley–James method, introduced by Buckley and James [1979], a few years after Sir David Cox proposed the Cox proportional hazards model (David et al. [1972]). Both methods can be utilized for analyzing survival data. However, the former method mainly focuses on the computation of the expected value of the survival time, while the later method focuses on determining the relative risk of explanatory variables on the failure event. Currently, the Cox model is the prevailing method for analyzing survival data (Cui [2005]).

The study conducted by Ali et al. [2015] compared AFT models with Cox Proportional Hazard (PH) models for analyzing the survival of gastric cancer patients. They discovered that when the proportional hazards assumption is violated, the results from the Cox PH model may become unreliable and biased. To address this limitation, the study recommends using AFT models as an alternative approach. AFT models with error distributions generalized gamma, log-logistic, log-normal, Gompertz, Weibull, and exponential do not rely on the proportional hazards assumption.

Furthermore, the AFT model in survival analysis has gained significant attention and is currently recognized as a valuable alternative to Cox models. It offers a more natural and straightforward approach for describing the impact of covariates on event times compared to Cox models (Kalbfleisch and Prentice [2011]).

1.2 AFT Model

The AFT model is used to investigate the association between covariates and the duration until an event of interest occurs (Kalbfleisch and Prentice [1980]). The AFT model expresses the logarithm of survival time as a linear function involving covariates. The model's adaptability and interpretability make it a useful tool for studying time-to-event data. The general expression for the AFT model is as follows.

$$\log(T_i) = X_i'\boldsymbol{\beta} + \epsilon_i, \quad i = 1, \ldots, n \tag{1.1}$$

Here, $\log(T_i)$ denotes logarithm of failure time of the i^{th} subject, X_i represents the covariate vector for the i^{th} subject with dimensions $p \times 1$, where p is the number of covariates with dimensions $p \times 1$, $\boldsymbol{\beta}$ denotes the vector of regression coefficients, and ϵ_i represents the error term specific to each observation. The error terms, denoted as ϵ_i, are typically assumed to be independent and identically distributed random variables. These errors generally follow specific distributions, such as the normal, extreme value, or log-logistic distributions. The AFT model is classified as parametric when the error terms adhere to a recognized statistical distribution. Alternatively, if the error distribution is not specified the AFT model is classified as semi-parametric.

Swindell [2009] used the AFT model to analyze data from 16 survivorship experiments investigating the effects of genetic manipulations on mouse lifespan. This study revealed that the majority of genetic modifications had a multiplicative effect on survivorship, which was consistent across different ages and age groups, accurately reflected by the "deceleration factor" of the AFT model.

Recently, a study in the medical field was conducted to utilize the Weibull AFT model for predicting the time until a health-related event occurs (Liu et al. [2023]). The study demonstrates the application of this model in providing a more comprehensible estimate of survival time compared to conventional probability based methods. The larynx cancer dataset was used to illustrate the implementation of the proposed method.

The AFT model, implemented with Generalized Estimating Equations (GEE), is a statistical approach that combines components of survival analysis and generalized estimating equations for datasets that indicate clustering or correlation. This model

aims to understand the relationship between covariates and the time until an event of interest while accounting for potential correlations within clusters of observations. Chiou et al. [2014b] introduced the AFTGEE method, which provides an alternative solution to the challenges associated with the less utilized AFT models, providing improved reliability and computational efficiency.

1.3 Weighted Accelerated Failure Time Model (Weighted AFT Model)

The Weighted AFT model is an improved version of the standard AFT model. It enhances accuracy of coefficient estimation by including weights that are assigned to individual observations based on their qualities or significance within the dataset. This weighting approach is very efficient in reducing the impact of censoring in survival data analysis. Researchers have explored various methods to derive these weights, aiming to improve the precision and reliability of their analyses.

For instance, Mustefa and Chen [2021] compared the performance of weighted least-squares estimation against classical methods using a real dataset of patients undergoing Antiretroviral Therapy. The results offer more precise estimations of how covariates impact outcomes and identify important associations with patient survival factors.

In another approach, Dong et al. [2023] proposed a novel weighted least squares model averaging method for AFT models with right censored data. This method, using Mallows criterion-based weights, demonstrated superior performance in model selection and averaging, particularly in cases of misspecified candidate models.

Furthermore, LASSO and threshold-gradient-directed regularization (TGDR) were investigated by Huang et al. [2006] as two regularization techniques for the Stute estimator in the AFT model with multiple covariates. They employed a weighted least squares approach in the estimator, incorporating Kaplan-Meier weights to handle censoring efficiently. Both methods aim to achieve simultaneous variable selection and estimation, with LASSO penalizing the L_1 norm ($\|\boldsymbol{\beta}\|_1 = \sum_{j=1}^{p} |\beta_j|$), where the vector $\boldsymbol{\beta}$ represents the regression coefficients and p is the number of covariates, of regression coefficients, and TGDR using cross-validation to determine the number

of gradient search steps and threshold value τ. Simulation studies and a real data example demonstrate the effectiveness of these methods in handling high-dimensional covariates and censored failure time data. Additionally, they explore the asymptotic distributions of the estimators and evaluate the performance of the bootstrap method for variance estimation.

In our study, we utilize auxiliary information from previous studies to compute weights aiming to enhance the effectiveness of the AFT model.

1.3.1 Auxiliary Information

Many studies are continuous or refined versions of previous studies. Properly integrating information from previous studies as auxiliary could enhance the efficiency of parameter estimation and statistical inference when fitting a model to relevant data. For example, in a study examining two covariates, if we have information available for one covariate from previous studies, it will be helpful to utilize this partial auxiliary information in the data analysis.

Let $\boldsymbol{X}$ be the covariates from the previous study, and let $\boldsymbol{X}_d$ be a subset of $\boldsymbol{X}$. Consider the association between logarithm of survival time (Y) and $\boldsymbol{X}_d$, expressed as,

$$Y = f(\boldsymbol{X}_d; \boldsymbol{\phi})$$

The understanding of this relationship might be considered as auxiliary information. Generally, the auxiliary information can be formulated as $E\{g(Z; \boldsymbol{\phi})\} = 0$, where Z denotes the observed data derived from present study. Here, $\boldsymbol{\phi} \in \mathbb{R}^d$, and the function $g(Z; \boldsymbol{\phi}) \in \mathbb{R}^q$, where $q \geq d$. The parameter $\boldsymbol{\phi}$ might not be known initially and can be estimated by utilizing available information from prior studies.

Granville and Fan [2014] proposed a nonparametric approach using the Buckley-James estimator to estimate regression parameters in accelerated failure time models incorporating auxiliary covariates. By employing kernel smoothing techniques and utilizing auxiliary covariates, this approach effectively handles missing or mismeasured data. Application of this estimator to the entire study cohort allows for robust

inference on covariate effects, supported by bootstrapping to estimate standard deviations of regression coefficients. The application of this approach to the PBC data illustrated its practical utility.

Auxiliary information, which can come in various forms such as additional covariates, known relationships between certain covariates, or established relationships between covariates and the response based on past experience or records, serves as a valuable asset (Vasudevan et al. [2019]).

We utilized the empirical likelihood approach in our study to compute weights, which were subsequently incorporated into the AFT model to derive the weighted AFT model.

1.3.2 Empirical Likelihood

Empirical likelihood is a nonparametric inference method with sampling properties similar to the bootstrap. However, instead of depending on resampling techniques like the bootstrap, it generates a multinomial probability by utilizing the observed sample data (Owen [1991]). The characteristics of empirical likelihood in independent and identically distributed scenarios are explained in Owen [1990], Hall [1990], and DiCiccio et al. [1991]. In a later stage, Owen [1991] extended the empirical likelihood method to regression problems, addressing both fixed and random regressors, as well as robust and heteroscedastic regressions.

Empirical likelihood approaches have become more widely used in various fields due to their nonparametric framework and strong statistical properties. These methods have also been extensively applied to survival data for their notable benefits in dealing with censored and complex data structures.

Li and Wang [2003] developed empirical likelihood methods for linear regression analysis of right censored data. They formulated an empirical likelihood for the regression coefficients vector using synthetic data. The adjusted empirical likelihood exhibits a central chi-squared limiting distribution, allowing for inference using standard chi-square tables. A simulation study was conducted to compare the performance of the adjusted empirical likelihood (ADEL) and estimated empirical likelihood (EEL) methods with the normal approximation method (Koul et al. [1981] ; Lai et al. [1995]). The findings indicate that empirical likelihood confidence intervals tend to offer more

precise coverage compared to intervals based on normal theory.

Later, Fang et al. [2013] proposed an innovative empirical likelihood method for semiparametric linear regression. This method focused on dealing with scenarios where the error distributions were completely unknown and involved right-censored survival data. Their method involved constructing an estimated empirical likelihood based on the Buckley-James estimating equation (Buckley and James [1979]) and integrated auxiliary information. They conducted simulations to compare their method with the synthetic data empirical likelihood approach proposed by Li and Wang [2003]. Additionally, illustrate the proposed method using the Stanford heart transplantation data.

Further extending the application of empirical likelihood methods, Wu [2005] presented computational algorithms for the pseudo empirical likelihood method in analyzing complex survey data. These algorithms are designed to determine maximum pseudo empirical likelihood estimators and construct pseudo empirical likelihood ratio confidence intervals. The algorithms are executed using R functions for practical use.

Vasudevan et al. [2019] introduced a weighted quantile regression method based on EL. The proposed method aims to improve the efficiency of censored quantile regression estimates by utilizing auxiliary data. This approach greatly enhances estimation accuracy by converting previous population information into probabilities based on empirical likelihood. The EL-based data driven probability computation was designed for scenarios with both known and unknown prior information about population parameters. Integrated into the regression model, these probabilities improve consistency and yield lower standard errors than the standard method, particularly when using all available covariates. The method also maintains reliable coverage probability and demonstrates efficacy in both heteroscedastic and homoscedastic models.

1.4 Motivation

Many survival studies are repetitive, and data from previous research are often readily accessible. Utilizing this historical information provides an opportunity to refine current models by incorporating insights from similar past studies. By calculating and incorporating weights derived based on information from previous studies, we can effectively enhance the overall accuracy of the models.

Exploring nonparametric methods in survival analysis offers a chance to improve model flexibility and robustness. However, there is a notable lack of research on how previous study information can be utilized in a nonparametric setting within AFT models. This research aims to fill that gap by incorporating nonparametric techniques with empirical likelihood-based weights, providing a new and innovative approach to survival analysis.

The proposed method extends the capabilities of traditional AFT models and AFTGEE models by integrating empirical likelihood, using weights based on previous study information to improve parameter estimation. Theoretical justifications for this approach have been developed. Simulation studies have demonstrated that the proposed method significantly improves the efficiency of parameter estimation.

The structure of the remaining part of the thesis is outlined as follows. Chapter 2 provides an overview of the Standard AFT Model, the incorporation of Generalized Estimating Equations (GEE) into the AFT model, the method for computing weights using the Empirical Likelihood function, and the Weighted AFT model. The last section of the chapter presents asymptotic of the proposed method. In Chapter 3, synthetic data was generated for a simulation study, offering a detailed explanation of the process. In the end, the chapter covers the summary and interpretation of simulation results, along with the illustration of the proposed method using a real dataset. Finally, Chapter 4 presents the conclusion and discussion of the entire study.

Chapter 2

Methodology

2.1 The Accelerated Failure Time (AFT) Model

The AFT model, employed in survival analysis to examine time-to-event data, offers numerous advantages compared to other survival models such as the Cox proportional hazards model. The key benefit of this method is that it provides coefficients that are more easily interpreted and directly related to the logarithm of the time-to-event variable, improving the understanding of predictor effects on survival time. This feature enhances predictability when analyzing events. AFT models possess remarkable flexibility in accommodating a wide variety of survival time distributions, including Exponential, Weibull, Log-logistic, and Log-normal distributions. This adaptability empowers researchers to select the distribution that best fits their data.

When survival data is right-censored, the relative risk model (Cox [1972]) and the AFT model (Kalbfleisch and Prentice [2002]) are commonly employed as regression models (Chiou et al. [2014b]). The AFT models effectively manage data that is right-censored and exhibit robustness to outliers and extreme values, addressing concerns that may impact other models used in survival analysis. These models also demonstrate resilience in handling the common issue of missing values frequently seen in real-world datasets. These qualities collectively make AFT models a valuable tool for survival analysis in various research contexts.

The AFT models find widespread application in diverse fields for analyzing time-to-event data. In the field of medical research, they are crucial for examining patient

survival and treatment effects. In addition, various other fields of study, such as economics, utilize AFT models to examine the duration of unemployment, patterns of retirement, and rates of loan repayment. Engineers utilize these models for predictive analysis, forecasting maintenance requirements, and enhancing system efficiency. AFT models are utilized in pharmacology to optimize drug dosages, and in environmental science to assess the duration of resources for sustainable management. These examples highlight the versatile and essential role of AFT models in addressing time-to-event data analysis challenges across various fields.

The AFT model can be defined mathematically as follows.

$$Y_i = \log(T_i) = X_i'\boldsymbol{\beta} + \epsilon_i, \quad i = 1, \ldots, n \tag{2.1}$$

where:

$\log(T_i)$ represents the logarithm of the survival time of the i^{th} subject,

X_i' denotes the transpose of the covariate vector with dimension $1 \times p$,

$\boldsymbol{\beta}$ is a vector of coefficients associated with the covariates with dimension $p \times 1$,

ϵ_i is a random error term.

2.1.1 Parameter Estimation

Parameter estimation in the Accelerated Failure Time (AFT) model with a Weibull distribution typically involves the following steps (Klein et al. [2003]).

The survival function for the Weibull distribution, given the scale parameter λ and the shape parameter α, is expressed as:

$$S_T(t) = \exp(-\lambda t^{\alpha})$$

The hazard rate is expressed as:

$$h_T(t) = \lambda \alpha t^{\alpha - 1}$$

The survival function for $Y = \log(T)$ is:

$$S_Y(y) = \exp(-\lambda e^{\alpha y})$$

If parameters are redefined as $\lambda = \exp(-\mu/\sigma)$ and $\sigma = 1/\alpha$, then:

$$Y = \log(T) = \mu + \sigma W$$

Here, W follows the extreme value distribution with the probability density function:

$$f_W(w) = \exp(w - e^w)$$

and survival function:

$$S_W(w) = \exp(-e^w)$$

The probability density function of Y is given by:

$$f_Y(y) = \frac{1}{\sigma} \exp\left[\frac{y-\mu}{\sigma} - e^{\left(\frac{y-\mu}{\sigma}\right)}\right]$$

and the survival function:

$$S_Y(y) = \exp\left[-e^{\left(\frac{y-\mu}{\sigma}\right)}\right]$$

The likelihood function for right-censored data is:

$$\begin{aligned} L &= \prod_{j=1}^{n} \left[f_Y(y_j)\right]^{\delta_j} \left[S_Y(y_j)\right]^{(1-\delta_j)} \\ &= \prod_{j=1}^{n} \left[\frac{1}{\sigma} f_W\left(\frac{y_j-\mu}{\sigma}\right)\right]^{\delta_j} \left[S_W\left(\frac{y_j-\mu}{\sigma}\right)\right]^{1-\delta_j} \end{aligned}$$

The maximum likelihood estimators of λ and α are:

$$\hat{\lambda} = \exp\left(-\frac{\hat{\mu}}{\hat{\sigma}}\right) \quad \text{and} \quad \hat{\alpha} = \frac{1}{\hat{\sigma}}$$

2.2 Accelerated Failure Time Model with GEE

When the error distribution is not specified, the AFT model is referred to as the semiparametric AFT model. This model has undergone thorough investigation and serves as an alternative to the relative risk model, particularly when error distribution is unspecified. Two widely used approaches for fitting such models have gained popularity. One method is the rank-based approach, which is inspired by the inversion of the weighted log-rank test (Prentice [1978]). The other follows the least squares principle, such as the Buckley-James (BJ) estimator (Buckley and James [1979]). Both approaches were not widely used in practice until recently due to the lack of efficient and reliable computing algorithms (Jin et al. [2003]; Jin et al. [2006a]).

Chiou et al. [2014a] explored the practical application of the AFTGEE model in routine survival analysis. The authors provide a thorough analysis of the method, emphasizing its flexibility in constructing AFT models. Their work discusses various modeling strategies supported by the AFTGEE approach, highlighting its effectiveness in handling diverse survival data sets, distributions, and covariate effects. The AFTGEE method offers convenient access to AFT models, utilizing both rank-based and least squares techniques.

In survival analysis, multiple computational methods are available for estimating parameters in AFT models. These methods are implemented in various R packages, including **survival** (Therneau and Lumley [2015], **rms** Harrell Jr [2014], and **eha** Broström [2014]). Parametric AFT models face the critical issue of potential misspecification of error distributions, which can result in biased estimations and draw misleading conclusions.

Semiparametric AFT models, which do not specify an error distribution, offer an alternative approach (Harrell Jr [2014]; Huang and Jin [2007]). However, existing methods for these models also have their limitations. For instance, the Buckley-James (BJ) estimator produces variance estimators using only non-censored observations. While this method shows favorable results in simulation studies, it lacks solid theoretical justification. Furthermore, the BJ estimator exhibits slow and non-guaranteed convergence, and it is specifically designed for univariate failure time data (Chiou et al. [2014b]).

The **lss** package in R is designed for fitting AFT model with right-censored data

using rank-based estimators with Gehan's weight derived from linear programming methods, along with iterative techniques for least squares estimation starting with the rank-based estimator (Huang and Jin [2007]). The variance estimators for both methods are based on bootstrap resampling, and their validity is theoretically justified. The process of estimating variance using bootstrap methodology is quite time-consuming. The rank-based estimator is constrained to Gehan's weight, which may be the optimal. The linear programming technique for rank-based estimator is quite computationally demanding, which has an impact on the least squares estimator as well as the initial estimator. In addition, the package does not provide functionality for specifying user-defined initial values for the least squares estimator. When dealing with clustered failure times, this method assumes that each cluster functions are independent and ignores any dependence within the cluster. However, this approach may result in a loss of efficiency, particularly when there is a high dependence within the cluster.

To address these limitations, the AFTGEE model provides more comprehensive tools for practical survival analysis. This approach significantly enhances computational speed compared to linear programming based methods, without compromising accuracy. It also offers efficient sandwich variance estimators as faster alternatives to full bootstrap variance estimation. By utilizing rapid rank-based estimators as initial estimates, this method employs an iterative least squares procedure that extends GEE to handle clustered censored data. Additionally, these methodologies can be extended to incorporate additional sampling weights to manage missing data and diverse sampling schemes. These features make the AFTGEE model an attractive choice for analysts seeking to fit AFT models seamlessly in their routine survival data analysis.

2.2.1 Buckley-James estimators

The Buckley-James estimator is the most suitable extension of least squares estimation for right-censored regression models (Kong and Yu [2007]). It is computed using a combination of iterative methods and numerical integration. Considering survival data that incorporates right censoring, Buckley and James [1979] replaced each response T_i with the conditional expectation $\hat{Y}_i(\beta) = E_\beta(T_i \mid Y_i, \delta_i, \mathbf{X}_i)$, where the expectation is determined based on regression coefficients β. Here, T_i is redefined as the

logarithm of failure time and Y_i is redefined as the logarithm of survival time to align with the methodology.

Specifically, the estimator is given by

$$\hat{Y}_i(\boldsymbol{b}) = \delta_i Y_i + (1-\delta_i)\left[\frac{\int_{e_i(\boldsymbol{b})}^{\infty} t\, d\hat{F}_{e_i(\boldsymbol{b})}(t)}{1-\hat{F}_{e_i(\boldsymbol{b})}[e_i(\boldsymbol{b})]} + \boldsymbol{X}_i'\boldsymbol{b}\right] \tag{2.2}$$

where $\boldsymbol{b}$ represents the estimated vector of regression coefficients, δ_i is an indicator variable equal to 1 for observed events and 0 for right-censored observations, Y_i is the survival time, $e_i(\boldsymbol{b}) = Y_i - X_i\boldsymbol{b}$ and $\hat{F}_{e_i(\boldsymbol{b})}$ is the Kaplan-Meier estimator based on the censored residual $e_i(\boldsymbol{b})$. That is,

$$\hat{F}_{e_i(\boldsymbol{b})}(t) = 1 - \prod_{i:e_i(\boldsymbol{b})<t}\left[1 - \frac{\delta_i}{\sum_{j=1}^{n} I_{(e_j(\boldsymbol{b})\geq e_i(\boldsymbol{b}))}}\right] \tag{2.3}$$

Although many researchers have extensively studied the theoretical properties of the BJ estimator, its practical use remains rare due to numerous challenges. An alternative approach has been suggested, emphasizing a more realistic solution by deriving a least squares estimator from a particular estimating equation, using an initial estimator $\boldsymbol{b}_n$ of $\boldsymbol{\beta}$. The least squares estimator is derived by solving the following estimating equation.

$$U_{n,\mathrm{ls}}(\boldsymbol{\beta}, \boldsymbol{b}) = \sum_{i=1}^{n}\left(\boldsymbol{X}_i - \bar{X}\right)'\left(\hat{Y}_i(\boldsymbol{b}) - \boldsymbol{X}_i\boldsymbol{\beta}\right) = 0 \tag{2.4}$$

where $\bar{X} = n^{-1}\sum_{i=1}^{n}\boldsymbol{X}_i$ is the mean vector.

The BJ estimator is given as the solution to the equation $U_{n,\mathrm{ls}}(\boldsymbol{\beta}, \boldsymbol{\beta}) = 0$. The benefit of setting the beginning value b_n is to prevent numerical complexities arising from solving Equation 2.4, which is neither continuous nor monotonic in $\boldsymbol{\beta}$ (Jin et al. [2006b]) developed an iterative algorithm, $\hat{\boldsymbol{\beta}}_{\mathrm{n,ls}}^{(m)} = L_n\left(\hat{\boldsymbol{\beta}}_{\mathrm{n,ls}}^{(m-1)}\right)$ for $m > 1$ with $\hat{\boldsymbol{\beta}}_{\mathrm{n,ls}}^{(0)} = \boldsymbol{b}_n$

where,

$$L_n(\boldsymbol{b}) = \left[\sum_{i=1}^{n}(\boldsymbol{X}_i - \bar{X})'(\boldsymbol{X}_i - \bar{X})\right]^{-1}\left[\sum_{i=1}^{n}(\boldsymbol{X}_i - \bar{X})'\left(\hat{Y}_i(\boldsymbol{b}) - \bar{Y}(\boldsymbol{b})\right)\right] \tag{2.5}$$

with $\hat{\beta}^{(m)} = L(\hat{\beta}^{(m-1)})$ and $\bar{Y}(\boldsymbol{b}) = \frac{1}{n}\sum_{i=1}^{n}\hat{Y}_i(\boldsymbol{b})$.

If the initial estimator $\boldsymbol{b}_n$ is both consistent and asymptotically normal, the $\hat{\boldsymbol{\theta}}_{\text{ls}}^{(m)}$ is also both consistent and asymptotically normal for every m (Jin et al. [2006a]). The induced smoothing Gehan estimator is a suitable candidate for the initial estimator. An estimation of the variance of the resulting estimate can be obtained by employing a resampling approach (Jin et al. [2006a]).

2.2.2 GEE approach incorporating with weights

The weighted least squares estimator is also determined through a combination of iterative techniques and numerical integration. Let $\boldsymbol{b}$ be an initial estimator of $\boldsymbol{\beta}$. The expression for the weighted least squares estimator can be formulated as follows (Zhang [2019]).

$$U_{n,\text{GEE}}(\boldsymbol{\beta}, \boldsymbol{b}) = \sum_{i=1}^{n} w_i \left(\boldsymbol{X}_i - \bar{X}\right)' \left(\hat{Y}_i(\boldsymbol{b}) - \boldsymbol{X}_i\boldsymbol{\beta}\right) = 0, \quad i = 1, \ldots, n \tag{2.6}$$

where,

$$\bar{X} = \frac{\sum_{i=1}^{n} w_i \boldsymbol{X}_i}{\sum_{i=1}^{n} w_i}$$

Here, w_i represents the weight for each observation. The weighted form of the function $L_n(\boldsymbol{b})$ can be expressed as $L_n^*(\boldsymbol{b})$.

$$L_n^*(\boldsymbol{b}) = \left[\sum_{i=1}^{n} w_i(\boldsymbol{X}_i - \bar{X})'(\boldsymbol{X}_i - \bar{X})\right]^{-1} \left[\sum_{i=1}^{n} w_i(\boldsymbol{X}_i - \bar{X})' \left(\hat{Y}_i^*(\boldsymbol{b}) - \bar{Y}^*(\boldsymbol{b})\right)\right] \tag{2.7}$$

Here, $\hat{Y}_i(\boldsymbol{b})$ and $\bar{Y}(\boldsymbol{b})$ are defined as follows.

$$\hat{Y}_i^*(\boldsymbol{b}) = \delta_i Y_i + (1 - \delta_i)\left[\frac{\int_{e_i(\boldsymbol{b})}^{\infty} t\, d\hat{F}^*_{e_i(\boldsymbol{b})}(t)}{1 - \hat{F}^*_{e_i(\boldsymbol{b})}[e_i(\boldsymbol{b})]} + \boldsymbol{X}_i'\boldsymbol{b}\right] \tag{2.8}$$

$$\bar{Y}_i^*(\boldsymbol{b}) = n^{-1} \sum_{i=1}^{n} \hat{Y}_i^*(\boldsymbol{b}) \tag{2.9}$$

The Kaplan-Meier estimator $\hat{F}^*_{e_i(\boldsymbol{b})}$ is computed as follows.

$$\hat{F}^*_{e_i(\boldsymbol{b})}(t) = 1 - \prod_{i:e_i(\boldsymbol{b})<t} \left[1 - \frac{\delta_i}{\sum_{j=1}^{n} \boldsymbol{I}_{(e_j(\boldsymbol{b}) \geq e_i(\boldsymbol{b}))}} \right] \tag{2.10}$$

The iterative procedure for updating the estimator is given by,

$$\hat{\boldsymbol{\beta}}^{*(m)} = L_n \left(\hat{\boldsymbol{\beta}}^{*(m-1)} \right), \quad \text{for} \quad m > 1$$

In both the standard AFT modeling and the AFTGEE method, the EL method can be utilized to determine weights by incorporating auxiliary information. In the next section, we will explore the detailed process of calculating these weights using the EL method.

2.3 Empirical Likelihood

Maximum Likelihood Estimation (MLE) encounters challenges in statistical modeling, particularly when dealing with inaccuracies in the actual distributions. MLE depends on specific parametric assumptions about the distribution function, and any inaccuracies in these assumptions can result in less efficient or precise estimations. While incorrect model selection can be effective in certain situations, such as estimating normal means using the Central Limit Theorem, it may fail when attempting to estimate normal variances. This failure can lead to inefficiencies in determining the outcomes of statistical tests. In addressing these challenges, non-parametric methods such as Empirical Likelihood prove to be a valuable alternative that effectively overcomes the limitations associated with MLE.

The empirical likelihood method, introduced by Owen [1988], has gained significant popularity and is commonly employed as a flexible nonparametric statistical tool. It efficiently overcomes practical challenges in various fields of study. One of its significant benefits is its capacity to effectively handle the complexity of various datasets,

manage uncertainties caused by unknown distributions, address non-parametric scenarios, and consistently handle outliers. Furthermore, empirical likelihood is a powerful method for estimating parameters, constructing data-driven confidence intervals, and conducting hypothesis tests with minimal reliance on distributional assumptions. Hence, its significant benefit becomes evident in areas where data naturally shows variability or deviates from standard distributions.

2.3.1 Weight estimation using Empirical Likelihood

In our study, We developed a method to transform auxiliary information into data-driven probabilities based on EL. These probabilities were then utilized as weights for our AFT models and AFTGEE models. This decision was motivated by the inherent advantages of empirical likelihood. Specifically, EL is robust in managing complex data and flexible in accommodating non-standard distributions. By incorporating these weights, our goal is to enhance the reliability and adaptability of our survival analysis.

Let $X_1, X_2 \ldots, X_n$ be iid observations from an unspecified distribution function F. The empirical distribution function $F_n(x)$ is a reliable estimator of the distribution F, and it can be considered as a non-parametric maximum likelihood estimate of F.

The cumulative distribution function is $F(x) = P(X_i \leq x_i)$ and $F(x_i) - F(x_i^-) = P(X_i = x_i)$. So $P(X_i = x_i) = F(x_i) - F(x_i^-)$.

The empirical cumulative distribution function (ECDF) of $X_1, X_2, \ldots, X_n \in \mathbb{R}$ is

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} 1_{X_i \leq x}, \quad for -\infty < x < \infty$$

Given $X_1, X_2, \ldots, X_n \in \mathbb{R}$, the non-parametric likelihood function of F is given by,

$$\begin{aligned} L_n(F) &= \prod_{i=1}^{n} [F(x_i) - F(x_i-)] \\ &= \prod_{i=1}^{n} P(X_i = x_i) \\ &= \prod_{i=1}^{n} p_i \end{aligned} \tag{2.11}$$

with $p_i = P(X = x_i)$. $p_i \geq 0, \quad i = 1, \ldots n$ and $\sum_{i=1}^{n} p_i = 1$

The non-parametric empirical log-likelihood function $l_n(F)$ is defined as follows.

$$l_n(F) = \sum_{i=1}^{n} \log p_i \tag{2.12}$$

subject to the constraints $\sum_{i=1}^{n} p_i = 1$ and $p_i > 0, i = 1, 2, \ldots, n$.

Let $L_n(F)$ be the likelihood of the observed data under the distribution function F and $L_n(F_n)$ be the likelihood of the observed data under the empirical distribution function F_n. Then the $L_n(F)$ is maximized when F is equal to the empirical distribution function F_n. The Empirical Likelihood (EL) ratio can be defined as,

$$\begin{aligned} R_n(F) &= \frac{L_n(F)}{L_n(F_n)} \\ &= \frac{\prod_{i=1}^{n} p_i}{(1/n)^n} \\ &= \prod_{i=1}^{n} (np_i) \end{aligned} \tag{2.13}$$

The Empirical log-likelihood ratio can be derived as $r_n(F)$,

$$r_n(F) = \sum_{i=1}^{n} \log(np_i) \tag{2.14}$$

2.3.2 Profile Empirical Likelihood

Assume we aim to investigate inference on the parameters under the assumptions that F belongs to a nonparametric distribution family $\mathcal{F}$, denoted as $\boldsymbol{\phi} = T(F)$, where T is some functional of the distribution.

Given a likelihood value at $\boldsymbol{\phi}$, we can make inferences about $\boldsymbol{\phi}$ using the likelihood approach. For each given value of $\boldsymbol{\phi}$, there are many members of $\mathcal{F}$ such that $T(F) = \boldsymbol{\phi}$. We must determine which F best represents $\boldsymbol{\phi}$. The idea behind profile empirical likelihood is to identify the F for which the empirical likelihood reaches its maximum among the set satisfying $T(F) = \boldsymbol{\phi}$.

The profile empirical likelihood is defined as follows.

$$L_n(\boldsymbol{\phi}) = \sup \{L_n(F) \mid T(F) = \boldsymbol{\phi}; F \in \mathcal{F}\} \tag{2.15}$$

The likelihood inference on $\boldsymbol{\phi}$ can be constructed using $L_n(\boldsymbol{\phi})$. This likelihood shares similar properties with its parametric counterpart.

Since $L_n(\boldsymbol{\phi}) \leq n^{-n} = L_n(F_n)$, it is convenient to standardize $L_n(\boldsymbol{\phi})$ by defining the ratio function.

$$\begin{aligned} R_n(\boldsymbol{\phi}) &= \frac{L_n(\boldsymbol{\phi})}{L_n(F_n)} \\ &= n^n L_n(\boldsymbol{\phi}) \end{aligned}$$

The empirical log-likelohood ratio function is,

$$r_n(\boldsymbol{\phi}) = n \log n + \log L_n(\boldsymbol{\phi}) = n \log n + l_n(\boldsymbol{\phi})$$

Let $\phi_0 = E(X_1)$ and $\mathrm{Var}(X_1) < \infty$. Then,

$$-2 \log [R_n(\phi_0)] \xrightarrow{D} \chi^2_d \quad \text{as} \quad n \to \infty$$

Where $d = \dim(X) = \dim(\boldsymbol{\phi})$ (Owen [2001]).

Based on the above results, we derive the $100(1-\alpha)\%$ confidence interval for $\boldsymbol{\phi}$:

$$\{\phi : -2\log\left[R_n(\phi_0)\right] \leq \chi^2_{d,1-\alpha}\}$$

2.3.3 Optimization

Suppose the parameter $\boldsymbol{\phi} = (\phi_1, \phi_2, \ldots, \phi_d)$ is defined by an estimating equation $E\{g(Z, \boldsymbol{\phi})\} = 0$, where $g(.)$ is a real-valued function.

EL based on general estimating equations was developed by Qin and Lawless [1994]. It applies to a random sample $\{T_i, Y_i, \delta_i, X_{di}\}_{i=1}^n$, denoted as $\{Z_i{}_{i=1}^n\}$, and an estimating function $g(Z_i; \boldsymbol{\phi})$ with parameter $\boldsymbol{\phi}$. The maximum empirical likelihood is given by,

$$L_{\mathrm{EL}}(\boldsymbol{\phi}) = \sup\left\{\prod_{i=1}^n p_i : p_i \geq 0, \sum_{i=1}^n p_i g(Z_i; \boldsymbol{\phi}) = 0\right\}$$

The Lagrange multiplier method is highly effective for solving this constrained maximization problem.

Define,

$$G(p_1, p_2, \ldots, p_n, s, \lambda, \boldsymbol{\phi}) = \sum_{i=1}^n \log(p_i) + s\left(\sum_{i=1}^n p_i - 1\right) - n\lambda\left(\sum_{i=1}^n p_i g(Z_i, \boldsymbol{\phi})\right)$$

where λ (vector valued) and s are Lagrange multipliers.

Now, setting to zero the partial derivative of G with respect to p_i gives the following results.

$$\frac{\partial G}{\partial p_i} = s + n \tag{2.16}$$

So,

$$0 = \sum_{i=1}^n p_i \frac{\partial G}{\partial p_i} = \frac{1}{p_i} + s - n\lambda g(Z_i, \boldsymbol{\phi}) \tag{2.17}$$

giving $s = -n$.

Thus,

$$\hat{p}_i = \frac{1}{n\{1 + \hat{\lambda}' g(Z_i, \boldsymbol{\phi})\}}, \quad i = 1, 2, \ldots, n. \tag{2.18}$$

The value of λ can be determined using numerical search for given $\boldsymbol{\phi}$.

We know that $\lambda = \lambda(\boldsymbol{\phi})$ is the solution of,

$$\frac{1}{n} \sum_{i=1}^{n} \frac{g(Z_i, \boldsymbol{\phi})}{1 + \hat{\lambda}' g(Z_i, \boldsymbol{\phi})} = 0$$

The profile EL corresponding to a given parameter $\boldsymbol{\phi}$ is expressed as:

$$l_n(\boldsymbol{\phi}) = -n \log(n) - \sum_{i=1}^{n} \log\{1 + \hat{\lambda}' g(Z_i, \boldsymbol{\phi})\}$$

and the EL ratio function is,

$$r_n(\boldsymbol{\phi}) = -\sum_{i=1}^{n} \log\{1 + \hat{\lambda}' g(Z_i, \boldsymbol{\phi})\}$$

Furthermore, the EL ratio statistic for the given $\boldsymbol{\phi}$ is defined as:

$$\mathcal{W}(\boldsymbol{\phi}) = -2r_n(\boldsymbol{\phi}) = -2 \sum_{i=1}^{n} \log\{1 + \hat{\lambda}' g(Z_i, \boldsymbol{\phi})\}, \quad \text{which converges in distribution to } \chi^2_d \quad \text{as} \quad n \to \infty.$$

where d is the dimension of $\boldsymbol{\phi}$.

Computation of Lagrange Multipliers

The determination of Lagrange multipliers will be discussed in detail in the this section. Chen et al. [2002] introduced an adapted Newton-Raphson algorithm to compute the Lagrange multiplier corresponding to a given value of the parameter. Recently, several methods have been developed for computing Lagrange multipliers

(Zhou [2023]; Kim et al. [2024]).

The Lagrange multiplier λ is determined by solving the equation ,

$$\sum_{i=1}^{n} \frac{g_i(\boldsymbol{\phi})}{1 + \lambda' g_i(\boldsymbol{\phi})} = 0 \tag{2.19}$$

for the given set of vectors $g_i(\boldsymbol{\phi})$, $i = 1, 2, \ldots, n$.

The equation 2.19 is the derivative of R with respect to λ for a given $\boldsymbol{\phi}$, where

$$R = \sum_{i=1}^{n} \log \{1 + \lambda' g_i(\boldsymbol{\phi})\} \tag{2.20}$$

In the empirical likelihood estimation, it is necessary that the solution satisfies the condition

$$1 + \lambda' g_i(\boldsymbol{\phi}) > 0, i = 1, 2, \ldots, n. \tag{2.21}$$

The modified Newton-Raphson algorithm for estimating λ for a given value of $\boldsymbol{\phi}$ can be outlined as follows:

Step 1 : Set $\lambda^c = 0$, $c = 0$, γ^k, $\epsilon = 10^{-8}$ and $\boldsymbol{\phi} = \boldsymbol{\phi}^0$

Step 2 : Let R^{λ} and $R^{\lambda\lambda}$ denote the first and second partial derivatives of R given in 2.20 with respect to λ, which are given by

$$R^{\lambda} = \sum_{i=1}^{n} \left[\frac{g_i(\boldsymbol{\phi})}{1 + \lambda' g_i(\boldsymbol{\phi})} \right],$$

$$R^{\lambda\lambda} = -\sum_{i=1}^{n} \left[\frac{g_i(\boldsymbol{\phi}) g_i'(\boldsymbol{\phi})}{1 + \lambda' g_i(\boldsymbol{\phi})^2} \right],$$

Compute R^{λ} and $R^{\lambda\lambda}$ for $\lambda = \lambda^c$ and let $\Delta(\lambda^c) = -[R^{\lambda\lambda}]^{-1} R^{\lambda}$

If $\|\Delta(\lambda^c)\| < \epsilon$ terminate the algorithm and report λ^c ; otherwise, continue.

Step 3 : Calculate $\delta^c = \gamma^c \Delta(\lambda^c)$. If $1 + \lambda^c - \delta^c g_i(\boldsymbol{\phi}) \leq 0$ for some i, then set $\gamma^c = \gamma^c/2$ and go to Step 2.

Step 4 : Set $\lambda^{c+1} = \lambda^c - \delta^c$, $c = c + 1$, and $\gamma^{c+1} = (c+1)^{-1/2}$ and go to Step 2. Step 2 will guarantee that $p_i > 0$ and the optimization is performed in the right direction.

2.3.4 Determining Weights Utilizing Auxiliary Information

This section provides a brief overview of how weights are determined using partial and complete auxiliary information.

Auxiliary information based on both X_1 and X_2

Suppose both X_1 and X_2 are used as auxiliary information. Derive the loss function after fitting a model with Y_0 and covariates with estimated coefficients θ_0, θ_1, and θ_2 from previous data. Here, Y_0 denotes the Y values from the current data.

$$L = (Y_0 - \boldsymbol{X}\hat{\boldsymbol{\phi}})'(Y_0 - \boldsymbol{X}\hat{\boldsymbol{\phi}})) \tag{2.22}$$

Where $\boldsymbol{X} = (X_0, X_1, X_2)$ and $\hat{\boldsymbol{\phi}} = (\theta_0, \theta_1, \theta_2)$. Here, X_0 typically represents a column vector of ones that is used to account for the intercept term in the model.

$$L_i = (Y_{0i} - (\theta_0 + \theta_1 X_{1i} + \theta_2 X_{2i}))'(Y_{0i} - (\theta_0 + \theta_1 X_{1i} + \theta_2 X_{2i})), \quad i = 1, 2, \ldots, n \tag{2.23}$$

To minimize L_i, we compute the partial derivatives of L_i to each parameter θ and set them to zero.

Partial derivatives with respect to θ_0,

$$\frac{\mathrm{d}L_i}{\mathrm{d}\theta_0} = -2X_{0i} * (Y_{0i} - (\theta_0 + \theta_1 X_{1i} + \theta_2 X_{2i})) = 0 \tag{2.24}$$

Partial derivatives with respect to θ_1,

$$\frac{\mathrm{d}L_i}{\mathrm{d}\theta_1} = -2X_{1i} * (Y_{0i} - (\theta_0 + \theta_1 X_{1i} + \theta_2 X_{2i})) = 0 \tag{2.25}$$

Partial derivatives with respect to θ_2,

$$\frac{\mathrm{d}L_i}{\mathrm{d}\theta_2} = -2X_{2i} * (Y_{0i} - (\theta_0 + \theta_1 X_{1i} + \theta_2 X_{2i})) = 0 \tag{2.26}$$

The functions in the equations 2.24, 2.25, and 2.26 were utilized to calculate empirical likelihood and derive weights based on auxiliary information based on X_1

and X_2.

Auxiliary information based on X_1

Consider a real-world scenario where we only have auxiliary information based on a single covariate X_1 and the estimated coefficients $\boldsymbol{\phi} = (\phi_0, \phi_1)$ are denoted as γ_0 and γ_1. Now, the loss function can be expressed as follows.

$$L = (Y_0 - \boldsymbol{X}\hat{\boldsymbol{\phi}})'(Y_0 - \boldsymbol{X}\hat{\boldsymbol{\phi}})) \tag{2.27}$$

Where $\boldsymbol{X} = (X_0, X_1)$ and $\hat{\boldsymbol{\phi}} = (\gamma_0, \gamma_1)$.

$$L_i = (Y_{0i} - (\gamma_0 + \gamma_1 X_{1i}))'(Y_{0i} - (\gamma_0 + \gamma_1 X_{1i}), \quad i = 1, 2, \ldots, n \tag{2.28}$$

To minimize L_i, we calculate the partial derivatives of L_i and equate them to zero.

Partial derivatives with respect to γ_0,

$$\frac{\mathrm{d}L_i}{\mathrm{d}\gamma_0} = -2X_{0i} * (Y_{0i} - (\gamma_0 + \gamma_1 X_{1i})) = 0 \tag{2.29}$$

Partial derivatives with respect to γ_1,

$$\frac{\mathrm{d}L_i}{\mathrm{d}\gamma_1} = -2X_{1i} * (Y_{0i} - (\gamma_0 + \gamma_1 X_{1i})) = 0 \tag{2.30}$$

The functions in the equations 2.29 and 2.30 were employed to calculate empirical likelihood and derive weights based on auxiliary information based on X_1.

Auxiliary information based on X_2

Now, consider a scenario where we have auxiliary information based on a single covariate X_2, and the estimated coefficients for $\boldsymbol{\phi} = (\phi_0, \phi_2)$ are denoted as δ_0 and δ_2. The loss function, which evaluates the fit of the model using these estimated coefficients,

can be expressed as follows.

$$L = (Y_0 - \boldsymbol{X}\hat{\boldsymbol{\phi}})'(Y_0 - \boldsymbol{X}\hat{\boldsymbol{\phi}})) \tag{2.31}$$

Where $\boldsymbol{X} = (X_0, X_2)$ and $\hat{\boldsymbol{\phi}} = (\delta_0, \delta_2)$.

$$L_i = (Y_{0i} - (\delta_0 + \delta_2 X_{2i}))'(Y_{0i} - (\delta_0 + \delta_2 X_{2i}), \quad i = 1, 2, \ldots, n \tag{2.32}$$

Once again, to minimize L_i, we calculate the partial derivatives of L_i and set them equal to zero.

Partial derivatives with respect to δ_0,

$$\frac{\mathrm{d}L_i}{\mathrm{d}\delta_0} = -2X_{0i} * (Y_{0i} - (\delta_0 + \delta_2 X_{2i})) = 0 \tag{2.33}$$

Partial derivatives with respect to δ_1,

$$\frac{\mathrm{d}L_i}{\mathrm{d}\delta_1} = -2X_{2i} * (Y_{0i} - (\delta_0 + \delta_2 X_{2i})) = 0 \tag{2.34}$$

The functions in equations 2.33 and 2.34 were utilized to calculate empirical likelihood and derive weights based on auxiliary information from X_2.

2.4 The Weighted Accelerated Failure Time Model (Weighted AFT Model)

It is well known that introducing weights in the models of survival analysis could improve the efficiency of the statistical inference, for example the inverse probability weights. Zhang [2019] proposed a weighted least-squares method for estimating parameters in semiparametric AFT models. This methodology can estimate parameters for mixture cure and case-cohort data, and can be extended to handle clustered data using generalized estimating equations (GEE). They used inverse probability weights (IPW) to address sampling bias and validated the method through large-scale simulations. In this thesis we propose to utilize weights obtained through empirical likelihood

method towards the AFT model.

Our study aims to enhance the precision of parameter estimates by utilizing weights and auxiliary information. This methodology employs empirical likelihood probabilities to calculate weights, integrating information from a prior dataset to improve the weighting procedure. By considering previous information, we address the challenges of censoring, enabling a more detailed examination of survival times.

The derived weights can be utilized in both parametric AFT models and AFTGEE models.

2.5 Parameter Estimation and Asymptotic Properties

Recalling equation 2.4, the least squares estimating equation can be expressed as follows :

$$U_{n,\mathrm{ls}}(\boldsymbol{\beta}, \boldsymbol{b}) = \sum_{i=1}^{n} \left(\boldsymbol{X}_i - \bar{\boldsymbol{X}}\right)' \left(\hat{Y}_i(\boldsymbol{b}) - \boldsymbol{X}_i\boldsymbol{\beta}\right) = 0 \tag{2.35}$$

where :

$\boldsymbol{\beta}$ is a vector of coefficients associated with the covariates,

$\boldsymbol{b}$ is a initial estimator of $\boldsymbol{\beta}$,

$\bar{\boldsymbol{X}}$ is the mean vector, which is defined as:

$$\bar{\boldsymbol{X}} = n^{-1} \sum_{i=1}^{n} \boldsymbol{X}_i$$

,

$$\hat{Y}_i(\boldsymbol{b}) = \delta_i Y_i + (1 - \delta_i) \left[\frac{\int_{e_i(b)}^{\infty} t \, d\hat{F}_{e_i(\boldsymbol{\beta})}(t)}{1 - \hat{F}_{e_i(\boldsymbol{\beta})}[e_i(\boldsymbol{b})]} + \boldsymbol{X}_i'\boldsymbol{b} \right],$$

$$e_i(\boldsymbol{b}) = Y_i - \boldsymbol{X}_i\boldsymbol{b},$$

$\hat{F}_{e_i(\boldsymbol{b})}$ the Kaplan-Meier estimator based on censored residual $e_i(\boldsymbol{b})$ where,

$$\hat{F}_{e_i(\boldsymbol{b})}(t) = 1 - \prod_{i:e_i(\boldsymbol{b})<t} \left[1 - \frac{\delta_i}{\sum_{j=1}^{n} I_{(e_j(\boldsymbol{b})\geq e_i(\boldsymbol{b}))}}\right].$$

Now, starting with an initial estimator $\boldsymbol{b}$ of $\boldsymbol{\beta}$, the proposed weighted least squares estimator can be obtained by solving the estimating equation.

$$U_{n,\text{wls}}(\boldsymbol{\beta}, \boldsymbol{b}) = \sum_{i=1}^{n} w_i \left(\boldsymbol{X}_i - \bar{\boldsymbol{X}}\right)' \left(\hat{Y}_i(\boldsymbol{b}) - \boldsymbol{X}_i\boldsymbol{\beta}\right) = 0 \tag{2.36}$$

where,

$$w_i = n\hat{p}_i$$

,

$$\bar{\boldsymbol{X}} = \frac{\sum_{i=1}^{n} w_i \boldsymbol{X}_i}{\sum_{i=1}^{n} w_i}$$

and

$$\hat{Y}_i^*(\boldsymbol{b}) = \delta_i Y_i + (1-\delta_i)\left[\frac{\int_{e_i(\boldsymbol{b})}^{\infty} t\, d\hat{F}^*_{e_i(\boldsymbol{b})}(t)}{1 - \hat{F}^*_{e_i(\boldsymbol{b})}[e_i(\boldsymbol{b})]} + \boldsymbol{X}_i'\boldsymbol{b}\right]$$

The Kaplan-Meier estimator $\hat{F}^*_{e_i(\boldsymbol{b})}$ is determined using the following equation.

$$\hat{F}^*_{e_i(\boldsymbol{b})}(t) = 1 - \prod_{i:e_i(\boldsymbol{b})<t} \left[1 - \frac{\delta_i}{\sum_{j=1}^{n} I_{(e_j(\boldsymbol{b})\geq e_i(\boldsymbol{b}))}}\right]$$

The least squares estimating equation and the weighted estimation equation can be solved through numerical methods. Next, it is necessary to demonstrate that the weighted estimating equation is asymptotically normal and consistent.

Asymptotic Properties

Now we wish to show that the resulting estimates of the regression parameters are consistent and asymptotically normal.

Define

$$F(t \mid \boldsymbol{X}) = \Pr(Y \le t \mid \boldsymbol{X}), \quad \bar{F}(t \mid \boldsymbol{X}) = \Pr(Y > t \mid \boldsymbol{X}), \quad \tilde{F}(t \mid \boldsymbol{X}) = \Pr(Y \le t, \delta = 1 \mid \boldsymbol{X}),$$

$$\bar{f}(y \mid \boldsymbol{X}) = -f(y \mid \boldsymbol{X}) = -\frac{dF(y \mid \boldsymbol{X})}{dy}, \quad \tilde{f}(y \mid \boldsymbol{X}) = \frac{d\tilde{F}(y \mid \boldsymbol{X})}{dy}.$$

Define $\boldsymbol{V_i} = \lambda'_{\phi_0} g(Z_i, \boldsymbol{\phi}_0)\boldsymbol{X}_i$, $i = 1, 2, \ldots, n$ as a p - vector.

Regularity Conditions :

***R*.1** ϵ and (Z, C) are independent.

***R*.2** (ϵ, C, Z) takes on finitely many values.

***R*.3** The observations, denoted as Z_i for $i = 1, 2, \ldots, n$, are independent and identically distributed (iid) samples from a certain distribution F. We make the assumption, without loss of generality, that $(Y_i, \delta_i, \boldsymbol{X}'_{di})' \subset Z_i$ for all $i = 1, 2, \ldots, n$.

***R*.4** There exists $\boldsymbol{\phi}_0$ such that $\mathbb{E}\{g(Z_i; \boldsymbol{\phi}_0)\} = 0$, the matrix $\boldsymbol{\Sigma}(\boldsymbol{\phi}_0) = \mathbb{E}\{g(Z_i; \boldsymbol{\phi}_0)g(Z_i; \boldsymbol{\phi}_0)'\}$ is positive definite, $\frac{\partial g(z;\boldsymbol{\phi})}{\partial \boldsymbol{\phi}}$ is continuous in the neighborhood of ϕ_0. The matrix $\mathbb{E}\left[\frac{\partial g(z;\boldsymbol{\phi})}{\partial \boldsymbol{\phi}}\right]$ is of full rank. Furthermore, there exist functions $H_{lj}(z)$ such that for ϕ in the neighborhood of $\boldsymbol{\phi}_0$:

(a) $\left|\frac{\partial g_l(z;\boldsymbol{\phi})}{\partial \boldsymbol{\phi}_j}\right| \le H_{lj}(z)$,

(b) For a constant C, $\mathbb{E}\{H^2(Z)\} \le C < \infty$ for $l = 1, \ldots, q$ and $j = 1, \ldots, d$.

***R*.5** $\max_i \|\boldsymbol{X}_i\|^2 = o(\sqrt{n})$ and $\max_i \|\boldsymbol{X}_i Y_{iG}\| = o(\sqrt{n})$, a.s.

***R*.6** $\sup_i \|\boldsymbol{X}_i\| < \infty$

Now, we state the following theorems.

Theorem 2.5.1: Assuming that the regularity conditions ***R*.1** - ***R*.6** hold, the estimator $\|\hat{\beta} - \beta_0\| \xrightarrow{Pr} 0$ as the sample size $n \to \infty$.

Theorem 2.5.2: Assuming that the regularity conditions ***R*.1** - ***R*.6** hold, the estimator $n^{1/2}\{\hat{\beta}-\beta_0\}$ converges in distribution to $N(0,\sigma^2_{\beta_0})$ as the sample size $n \to \infty$.

To prove Theorems 2.5.1 and 2.5.2, we need to show that $\max_{1\le i\le n} |\lambda'_{\boldsymbol{\phi}_0} g(Z_i;\boldsymbol{\phi}_0)| = o_p(1)$. Here, $g(Z_i;\boldsymbol{\phi})$ contains the censored observations.

Under the regularity condition ***R*.5**, Qin and Jing [2001] proved the following for the function $g(.)$ with censored observations.

$$\max_{1\le i\le n} |\lambda_{\boldsymbol{\phi}_0} g(Z_i;\boldsymbol{\phi}_0)| = o_p(1) \tag{2.37}$$

Following Owen [2001] and using Taylor's series expansion of weights, p_i's can be rewritten as follows.

$$\begin{aligned} p_i(\boldsymbol{\phi}_0) &= \frac{1}{n[1+\lambda'_{\boldsymbol{\phi}_0} g(Z_i;\boldsymbol{\phi}_0)]} \\ &= \frac{1}{n}[1-\lambda'_{\boldsymbol{\phi}_0} g(Z_i;\boldsymbol{\phi}_0)]\{1+o_p(1)\}; \quad i=1,2,\ldots,n \end{aligned} \tag{2.38}$$

Now $w_i(\boldsymbol{\phi}_0)$ can be calculated as follows.

$$\begin{aligned} w_i(\boldsymbol{\phi}_0) &= np_i(\boldsymbol{\phi}_0) \\ &= [1-\lambda'_{\boldsymbol{\phi}_0} g(Z_i;\boldsymbol{\phi}_0)]\{1+o_p(1)\}; \quad i=1,2,\ldots,n \end{aligned}$$

By equation 2.36, the weighted least square estimating function can be rewritten as follows.

$$\begin{aligned} U_{n,\mathrm{wls}}(\boldsymbol{\beta},\boldsymbol{b}) &= \sum_{i=1}^{n} w_i(\mathbf{X}_i-\bar{\mathbf{X}})'(\hat{Y}_i(\boldsymbol{b})-\mathbf{X}_i\boldsymbol{\beta}) \\ &= \sum_{i=1}^{n}[1-\lambda'_{\theta_0} g(Z_i;\theta_0)](\mathbf{X}_i-\bar{\mathbf{X}})'(\hat{Y}_i(\boldsymbol{b})-\mathbf{X}_i\boldsymbol{\beta})+o_p(\sqrt{n}) \\ &= \sum_{i=1}^{n}(\mathbf{X}_i-\bar{\mathbf{X}})'(\hat{Y}_i(\boldsymbol{b})-\mathbf{X}_i\boldsymbol{\beta}) - \sum_{i=1}^{n}[\lambda'_{\theta_0} g(Z_i;\theta_0)](\mathbf{X}_i-\bar{\mathbf{X}})'(\hat{Y}_i(\boldsymbol{b})-\mathbf{X}_i\boldsymbol{\beta})+o_p(\sqrt{n}) \\ &= \sum_{i=1}^{n}(\mathbf{X}_i-\bar{\mathbf{X}})'(\hat{Y}_i(\boldsymbol{b})-\mathbf{X}_i\boldsymbol{\beta}) - \sum_{i=1}^{n} V_i(\mathbf{X}_i-\bar{\mathbf{X}})'(\hat{Y}_i(\boldsymbol{b})-\mathbf{X}_i\boldsymbol{\beta})+o_p(\sqrt{n}) \end{aligned}$$

Asymptotically, by 2.37, we have $\|\boldsymbol{V}_i\| = o_p(1); i = 1, 2, \ldots, n$. So,

$$U_{n,\text{wls}}(\boldsymbol{\beta}, \boldsymbol{b}) = \sum_{i=1}^{n} (\mathbf{X}_i - \bar{\mathbf{X}})'(\hat{Y}_i(\boldsymbol{b}) - \mathbf{X}_i'\boldsymbol{\beta}) + o_p(\sqrt{n})$$

Compared to the first part of the equation, $o_p(\sqrt{n})$ is ignorable. Therefore, our weighted least square estimating function is asymptotically equivalent to the least square estimating function. Jin et al. [2006a] have shown that the least square estimator is both consistent and asymptotically normal. This implies that the weighted least square estimator also possesses these desirable properties of consistency and asymptotic normality.

Chapter 3

Numerical Studies

In this chapter, simulation studies are conducted to compare the performance of standard AFT models with AFTGEE models. Further investigation is also performed to compare the weighted versions of both methods using empirical likelihood. A real dataset is used to illustrate the methods and verify the findings.

3.1 Simulation Study

3.1.1 Data generation and steps for the simulation studies

In the simulation study, we fitted the regression model by including fixed covariates X_1 and X_2, and the logarithm of failure time T. We assumed a linear relationship between T and both X_1 and X_2.

$$T_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \epsilon_i, \quad \text{for} \quad i = 1, 2, \ldots, n \tag{3.1}$$

where:

$X_{11}, X_{12} \ldots X_{1n}$ generated from $Ber(p = 0.5)$ distribution,

$X_{21}, X_{22} \ldots X_{2n}$ generated from $N(\mu = 40, \sigma = 5)$ distribution,

$\epsilon_i \overset{\text{iid}}{\sim} N(0, \sigma_\epsilon)$.

The simulation study was divided into two main phases. In the initial phase, we generated a finite population and used it as auxiliary information for subsequent steps in the study. From this auxiliary data, we computed probability weights P_i based on Empirical Likelihood (EL). Subsequently, the study proceeded to estimate parameters using various random samples drawn from the previously specified population and different percentages of censoring. Both the standard AFT method and the AFTGEE method were employed in these estimations, incorporating EL-based weights into fitting the standard AFT and AFTGEE models. The evaluation of model performance involved the comparison of average bias, average standard deviations, and coverage probabilities.

To provide a more comprehensive explanation of the simulation study, the description of Phase I and Phase II is as follows.

3.1.2 Steps for the Simulation Study - Phase I

Step 1 : For the simulation study, a regression model was fitted incorporating X_1, X_2, and the logarithm of the failure time T. The formulated equation is expressed as follows. Here, X_{1i} was generated from a Bernoulli distribution with a probability of success of 0.5. X_{2i} was generated from a Normal distribution with a mean of 40 and a standard deviation of 5, and the resulting values were rounded up to the nearest integer.

$$T_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \epsilon_i, \quad \text{for} \quad i = 1, 2, \ldots, n \tag{3.2}$$

where, $\sigma_\epsilon = (0.4, 1.0)$, $\beta_0 = 6.9$, $\beta_1 = 1.2$ and $\beta_2 = -0.12$.

Step 2 : Censoring time is generated from a uniform distribution $U(0, \tau)$ to achieve the desired censoring percentage. The status indicator function $\delta_i = \mathbb{I}(T_i \leq C_i)$ is determined using the logarithm of the failure time (T_i), and the logarithm of censoring time (C_i). To reach a specific censoring rate, τ is gradually increased until the proportion of censored data matches the target percentage. This process obtains the optimal τ values for each combination of sample size and σ_ϵ values.

Step 3 : For a sample size of $N = 50,000$, $\boldsymbol{X_1}$ and $\boldsymbol{X_2}$ are generated as described in Step

1, and the regression model is fitted according to equation 3.2. The failure time can be determined by exponentiating T_i. Censoring times are generated using a uniform distribution ranging from 0 to the τ value determined based on the sample size $N = 50,000$ and the corresponding σ_ϵ.

Step 4 : Observed time is determined by selecting the minimum value between the failure time and censoring time, $Y = \min(T_{0i}, C_{0i})$, where T_{0i} is logarithm of failure time and C_{0i} is logarithm of censoring time. The censoring rate is calculated using the status indicator $\delta_{i0} = \mathbb{I}(T_{i0} \leq C_{i0})$.

Step 5 : The linear regression models were fitted using Y with each covariate individually, as well as with Y and both covariates together. Coefficients were computed using the least squares approach. The individual models were fitted as follows:

- A linear regression was fitted between Y and X_1 as $Y \sim X_1$ and estimating the coefficients γ_0 and γ_1.
- A linear regression was fitted between Y and X_2 as $Y \sim X_2$ and estimated the coefficients δ_0 and δ_2.
- A linear regression was fitted between Y with X_1 and X_2 as $Y \sim X_1 + X_2$ and estimated the coefficients θ_0, θ_1 and θ_2 .

Step 6 : Steps 3, 4, and 5 were repeated for varying censoring rates of 10%, 20%, and 30% and different values of σ_ϵ.

3.1.3 Steps for the Simulation Study - Phase II

Step 1 : For the sample size ($n = 100, n = 200$ or $n = 500$) , fixed X_1 and X_2 were generated, and regression models were fitted according to Step 1 in Phase I.

Step 2 : Failure time, censoring time, observed time, and censoring rate were calculated as described in Step 3 and Step 4. To determine the censoring time, the corresponding τ value was used based on the sample size and σ_ϵ.

Step 3 : Weights were obtained using the empirical likelihood method as described in Section 2.3.2 of Chapter 2. The **el.test.wt2** function from the **emplik** package was used to calculate the weights, incorporating the partial derivatives described

in that section. Weight $\boldsymbol{W_1}$ was derived using auxiliary data from X_1, weight $\boldsymbol{W_2}$ was obtained using auxiliary information from X_2, and weight $\boldsymbol{W_{12}}$ was calculated using auxiliary information from both X_1 and X_2.

Step 4 : Coefficients were estimated using the weighted model incorporating $\boldsymbol{W_1}$, $\boldsymbol{W_2}$, and $\boldsymbol{W_{12}}$, as well as without any weight, in both AFT models and AFTGEE models.

Step 4 : For the fixed covariates, the simulation study was repeated 1,000 times to calculate the mean of the estimated coefficients. Standard errors were computed using the bootstrap method, and coverage probabilities were determined based on these standard errors. Bias, standard deviations, and coverage probabilities were subsequently used to compare the performance of the models.

As discussed in Step 3 of Phase II, to compute the weights using empirical likelihood, it is essential to have information on population parameters ϕ or their estimated values. When both X_1 and X_2 are used as auxiliary information, the estimated coefficients for $\phi = (\phi_0, \phi_1, \phi_2)$ can be obtained as θ_0, θ_1, and θ_2 following the process detailed in Step 5 of Phase I. These estimates were derived using the least squares approach in ordinary linear regression, conducted with a finite population size of $N = 50,000$.

To proceed with the remaining part of the simulation study, after calculating the weights, the regression coefficients, and its standard errors can be calculated as described in Phase II.

3.1.4 Implementation of R functions

AFT Model

The **survreg** function in the **survival** package is used to fit AFT models (Therneau and Lumley [2015]). The response is constructed using the **Surv** object, which comprises two columns. The first column represents the survival time or censored time, while the second column is the censoring indicator, which indicates the right censored data. The convergence of the procedure in **survreg** is controlled by the relative tolerance and the maximum number of iterations. These parameters can be defined

using the **survreg.control()** function. The iteration process terminates and produces the outcome when either the relative tolerance is met, or the iteration reaches the predetermined maximum number of iterations.

survreg(formula, data, weights, subset, na.action, dist = "weibull", init = NULL, scale = 0, control, parms = NULL, model = FALSE, x = FALSE, y = TRUE, robust = FALSE, cluster, score = FALSE, ...)

AFTGEE Model

The **aftgee** function in the **aftgee** package is used to fit AFTGEE models (Chiou et al. [2014a]). Similar to the **survreg** function in the survival package, the **aftgee** function requires both the formula and data arguments.

The iteration process stops and generates results when either the given tolerance is achieved or the maximum number of iterations is reached. The settings can be modified by use the **aftgee.control()** method.

aftgee(formula, data, subset, id = NULL, contrasts = NULL, weights = NULL, margin = NULL, corstr = "independence", binit = "srrgehan", B = 100, control = aftgee.control())

In situations where there is no censoring, and the independent working correlation structure is specified in the **aftgee** function, the model will return an ordinary least squares estimate.

Empirical Likelihood Probabilities

Empirical likelihood probabilities can be calculated using the **el.test.wt2** function from the **emplik** package (Zhou [2023]). The function returns several key outputs, including the Lagrange multiplier at the solution, the vector of weights used in the empirical likelihood calculation and the probabilities that maximize the weighted empirical likelihood under the mean constraint.

el.test.wt2(x, wt, mu, maxit = 25, gradtol = 1e-07, Hessian = FALSE, svdtol = 1e-09, itertrace = FALSE)

In the following section, we focus on the findings of the simulation study using various combinations.

Simulation Study Combinations

In the simulation study, the performance of the proposed method is compared under different sample sizes, censoring percentages, and standard deviation values of the error term. The study comprises 18 cases, each with these varying parametric values. These cases are detailed below.

Summary of simulation study cases	
Case number	Description
Case I	$n = 100$, $\sigma_\epsilon = 0.4$ and 10% Censored data
Case II	$n = 100$, $\sigma_\epsilon = 0.4$ and 20% Censored data
Case III	$n = 100$, $\sigma_\epsilon = 0.4$ and 30% Censored data
Case IV	$n = 200$, $\sigma_\epsilon = 0.4$ and 10% Censored data
Case V	$n = 200$, $\sigma_\epsilon = 0.4$ and 20% Censored data
Case VI	$n = 200$, $\sigma_\epsilon = 0.4$ and 30% Censored data
Case VII	$n = 500$, $\sigma_\epsilon = 0.4$ and 10% Censored data
Case VIII	$n = 500$, $\sigma_\epsilon = 0.4$ and 20% Censored data
Case IX	$n = 500$, $\sigma_\epsilon = 0.4$ and 30% Censored data
Case X	$n = 100$, $\sigma_\epsilon = 1.0$ and 10% Censored data
Case XI	$n = 100$, $\sigma_\epsilon = 1.0$ and 20% Censored data
Case XII	$n = 100$, $\sigma_\epsilon = 1.0$ and 30% Censored data
Case XIII	$n = 200$, $\sigma_\epsilon = 1.0$ and 10% Censored data
Case XIV	$n = 200$, $\sigma_\epsilon = 1.0$ and 20% Censored data
Case XV	$n = 200$, $\sigma_\epsilon = 1.0$ and 30% Censored data
Case XVI	$n = 500$, $\sigma_\epsilon = 1.0$ and 10% Censored data
Case XVII	$n = 500$, $\sigma_\epsilon = 1.0$ and 20% Censored data
Case XVIII	$n = 500$, $\sigma_\epsilon = 1.0$ and 30% Censored data

The performance of the standard AFT and AFTGEE models was assessed in terms of bias, standard deviation (sd), and coverage probabilities (CP) at 90%, 95%, and 99% confidence levels of the regression parameters. These results were utilized to compare the performance of weighted models to models without weights. Each case is divided into two parts, denoted as (a) and (b), to record $\hat{\beta}_1$ and $\hat{\beta}_2$ data, respectively.

3.1.5 Scenario 01 : Case I (Sample size 100 , $\sigma_\epsilon = 0.4$ and 10% Censored data)

We first set the sample size to 100, with a censoring percentage of 10% and $\sigma_\epsilon = 0.4$. We generated 1000 datasets and estimated the parameters β_1 and β_2 for both AFT and AFTGEE models using different weighting schemes, including no weights, W_1, W_2, and W_{12}. We then computed the bias, standard deviation (sd), and coverage probabilities at 90%, 95%, and 99% confidence levels, as detailed in Table 3.1 and Table 3.2 for β_1 and β_2, respectively.

Standard AFT -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0019	0.0826	89.4	93.7	98.3
W_1	-0.0015	0.0752	89.1	93.6	98.9
W_2	0.0040	0.0835	89.6	93.9	98.3
W_{12}	-0.0032	0.0650	87.3	91.1	96.0
AFTGEE -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0021	0.0826	89.4	93.8	98.4
W_1	-0.0016	0.0753	89.0	93.8	98.9
W_2	0.0039	0.0834	89.6	94.1	98.3
W_{12}	-0.0033	0.0650	87.3	90.9	95.7

Table 3.1: Comparative Analysis of $\hat{\beta}_1$: Standard AFT vs AFTGEE with 10% censored data for sample size $n = 100$ and $\sigma_\epsilon = 0.4$ (Case I)

Standard AFT -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0004	0.0086	89.0	93.3	98.0
W_1	-0.0009	0.0088	88.1	93.5	97.9
W_2	-0.0002	0.0080	87.2	93.3	98.5
W_{12}	0.0006	0.0071	86.4	91.6	95.8
AFTGEE -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0004	0.0086	89.2	93.3	97.9
W_1	-0.0009	0.0088	88.3	93.5	97.7
W_2	-0.0002	0.0079	87.2	93.3	98.5
W_{12}	0.0006	0.0071	86.3	91.6	95.7

Table 3.2: Comparative Analysis of $\hat{\beta}_2$: Standard AFT vs AFTGEE with 10% censored data for sample size $n = 100$ and $\sigma_\epsilon = 0.4$ (Case I)

From Table 3.1, we see that model with W_{12} weight have the lowest standard deviation for β_1 among the models in both the AFTGEE and standard AFT models. W_1 follows with the next lowest standard deviation. The model incorporating W_2 does not show any improvement over the model without weights. Models using W_1 weights consistently maintaining coverage probabilities across various confidence levels (90%, 95%, and 99%). Although models with W_{12} also perform well, they show slightly lower coverage probabilities compared to those with W_1, but still remain at an acceptable level. This suggests that, when estimating β_1, the weighted models incorporating auxiliary information based on X_1, as well as those incorporating information from both X_1 and X_2, provide better estimates for β_1.

Table 3.2 presents the data based on estimates for β_2. From this table we see that models utilizing the weight W_{12} have the lowest standard deviation in both the standard AFT and AFTGEE models. Although models associated with W_{12} exhibit slightly lower coverage probabilities compared to other models , these probabilities remain within an acceptable range of the nominal levels. This suggests that the weighted models that include auxiliary information from X_2, as well as auxiliary information from both covariates X_1 and X_2, provide good estimations for β_2.

3.1.6 Scenario 02 : Impact of the censoring

In this scenario, we examine how different levels of censoring percentages affect the performance of weighted AFT and AFTGEE models. Using a fixed sample size of 100 and $\sigma_\epsilon = 0.4$, we evaluated model performance by increasing the censoring percentage to 20% and 30% to explore the impact of censoring on the estimation of parameters β_1 and β_2. Table 3.3 and Table 3.4 display the results for data with 20% censoring, while Table 3.5 and Table 3.6 present the results for data with 30% censoring.

Standard AFT -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0043	0.0908	88.2	94.0	99.1
W_1	-0.0039	0.0831	88.7	94.3	98.8
W_2	0.0005	0.0906	88.8	94.3	98.9
W_{12}	0.0014	0.0822	88.3	93.5	98.3
AFTGEE -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0045	0.0907	87.9	94.1	99.0
W_1	-0.0041	0.0830	88.9	94.6	98.7
W_2	0.0004	0.0905	88.6	94.6	98.9
W_{12}	0.0011	0.0821	88.4	93.7	98.4

Table 3.3: Comparative Analysis of $\hat{\beta}_1$: Standard AFT vs AFTGEE with 20% censored data for sample size $n = 100$ and $\sigma_\epsilon = 0.4$ (Case II)

Standard AFT -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0004	0.0091	88.4	94.3	98.8
W_1	-0.0008	0.0092	88.5	94.8	99.2
W_2	-0.0001	0.0085	89.5	94.6	99.2
W_{12}	0.0000	0.0085	90.1	94.7	98.1
AFTGEE -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0004	0.0092	88.7	94.4	98.8
W_1	-0.0008	0.0092	88.2	94.6	99.1
W_2	-0.0001	0.0086	89.3	94.4	99.2
W_{12}	0.0000	0.0086	90.0	94.6	98.0

Table 3.4: Comparative Analysis of $\hat{\beta}_2$: Standard AFT vs AFTGEE with 20% censored data for sample size $n = 100$ and $\sigma_\epsilon = 0.4$ (Case II)

Standard AFT -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	0.0013	0.0924	90.7	95.6	98.9
W_1	0.0033	0.0872	90.3	95.7	99.0
W_2	0.0063	0.0923	91.1	95.9	99.0
W_{12}	0.0086	0.0861	91.0	95.8	99.4
AFTGEE -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	0.0010	0.0923	90.3	95.4	99.0
W_1	0.0030	0.0871	90.8	95.7	98.9
W_2	0.0062	0.0923	91.1	95.6	98.9
W_{12}	0.0084	0.0861	91.1	95.7	99.5

Table 3.5: Comparative Analysis of $\hat{\beta}_1$: Standard AFT vs AFTGEE with 30% censored data for sample size $n = 100$ and $\sigma_\epsilon = 0.4$ (Case III)

Standard AFT -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0007	0.0107	85.2	91.5	97.7
W_1	-0.0012	0.0108	86.4	92.2	97.6
W_2	-0.0005	0.0101	85.5	91.3	98.4
W_{12}	-0.0008	0.0101	87.0	93.7	98.0
AFTGEE -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0007	0.0107	85.7	91.4	97.7
W_1	-0.0011	0.0108	87.1	91.9	97.5
W_2	-0.0005	0.0101	86.2	91.8	98.3
W_{12}	-0.0008	0.0101	87.4	93.4	97.9

Table 3.6: Comparative Analysis of $\hat{\beta}_2$: Standard AFT vs AFTGEE with 30% censored data for sample size $n = 100$ and $\sigma_\epsilon = 0.4$ (Case III)

When comparing β_1 and β_2 estimations, as the censoring percentage increases, models with W_{12} show results that approach the nominal levels more closely. This improvement is observed across all confidence levels (90%, 95%, and 99%) compared to the situation with the 10% censoring rate.

Even with a high censoring percentage, the results indicate that models incorporating auxiliary information based on X_1 and both X_1 and X_2 provide better estimates for β_1. Additionally, when estimating β_2, models using auxiliary information based on X_2 and both X_1 and X_2 offer more precise estimates.

When comparing the Standard AFT and AFTGEE models, no significant differences are observed in the standard deviation or coverage probabilities between the two methods.

As the censoring percentage increases, the proposed weighted model that utilizes either partial or complete auxiliary information specific to the relevant covariate achieves more precise estimations.

3.1.7 Scenario 03 : Impact of the sample size

In this scenario, we increased the sample size to 200 and 500 to examine the effect of sample size on the performance of weighted AFT and AFTGEE models. We compared the impact of sample size across all censoring levels of 10%, 20%, and 30% with $\sigma_\epsilon = 0.4$. This analysis aims to determine how varying sample sizes influence the estimation of parameters β_1 and β_2. Table 3.7 to Table 3.12 display results for a sample size of 200, while Table 3.13 to Table 3.18 show results for a sample size of 500, at different censoring percentages.

Standard AFT -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0024	0.0597	88.5	94.5	98.9
W_1	-0.0017	0.0522	89.4	94.6	98.7
W_2	0.0007	0.0598	88.2	94.7	98.8
W_{12}	-0.0030	0.0456	88.5	94.1	97.8
AFTGEE -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0024	0.0597	88.7	94.3	98.9
W_1	-0.0018	0.0522	89.1	94.4	98.7
W_2	0.0007	0.0598	88.2	94.7	98.6
W_{12}	-0.0030	0.0456	88.1	94.0	97.9

Table 3.7: Comparative Analysis of $\hat{\beta}_1$: Standard AFT vs AFTGEE with 10% censored data for sample size $n = 200$ and $\sigma_\epsilon = 0.4$ (Case IV)

Standard AFT -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0004	0.0061	88.4	93.7	98.7
W_1	-0.0007	0.0061	88.1	93.5	98.9
W_2	-0.0005	0.0054	88.7	94.1	98.5
W_{12}	-0.0003	0.0051	87.2	93.1	97.4
AFTGEE -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0004	0.0061	88.5	93.9	98.7
W_1	-0.0007	0.0061	88.1	93.5	98.8
W_2	-0.0005	0.0054	88.7	94.1	98.3
W_{12}	-0.0003	0.0051	87.0	93.1	97.3

Table 3.8: Comparative Analysis of $\hat{\beta}_2$: Standard AFT vs AFTGEE with 10% censored data for sample size $n = 200$ and $\sigma_\epsilon = 0.4$ (Case IV)

Standard AFT -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0020	0.0633	88.5	93.9	98.6
W_1	-0.0014	0.0571	89.1	93.6	98.9
W_2	0.0009	0.0636	88.2	94.0	98.6
W_{12}	-0.0002	0.0553	89.0	94.6	98.6
AFTGEE -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0022	0.0632	88.3	93.6	98.6
W_1	-0.0016	0.0571	89.3	93.5	98.8
W_2	0.0007	0.0635	88.5	93.7	98.7
W_{12}	-0.0004	0.0552	88.9	94.3	98.6

Table 3.9: Comparative Analysis of $\hat{\beta}_1$: Standard AFT vs AFTGEE with 20% censored data for sample size $n = 200$ and $\sigma_\epsilon = 0.4$ (Case V)

Standard AFT -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0003	0.0064	89.3	94.5	98.7
W_1	-0.0006	0.0064	89.3	95.1	99.0
W_2	-0.0004	0.0060	88.9	94.3	98.9
W_{12}	-0.0005	0.0060	87.9	94.2	98.8
AFTGEE -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0003	0.0064	89.4	94.6	98.7
W_1	-0.0006	0.0064	89.7	95.0	98.8
W_2	-0.0004	0.0060	88.9	94.6	98.9
W_{12}	-0.0005	0.0060	88.1	94.1	99.0

Table 3.10: Comparative Analysis of $\hat{\beta}_2$: Standard AFT vs AFTGEE with 20% censored data for sample size $n = 200$ and $\sigma_\epsilon = 0.4$ (Case V)

Standard AFT -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0003	0.0674	88.4	93.8	98.6
W_1	0.0003	0.0624	89.5	94.2	98.9
W_2	0.0023	0.0679	88.3	93.6	98.5
W_{12}	0.0017	0.0624	89.1	94.4	98.4
AFTGEE -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0005	0.0673	88.9	93.8	98.8
W_1	0.0001	0.0624	89.7	94.1	98.8
W_2	0.0021	0.0678	88.4	93.9	98.6
W_{12}	0.0015	0.0623	89.4	94.6	98.7

Table 3.11: Comparative Analysis of $\hat{\beta}_1$: Standard AFT vs AFTGEE with 30% censored data for sample size $n = 200$ and $\sigma_\epsilon = 0.4$ (Case VI)

Standard AFT -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0002	0.0069	89.3	94.0	98.7
W_1	-0.0005	0.0069	89.1	94.2	98.5
W_2	-0.0003	0.0066	89.9	94.5	98.2
W_{12}	-0.0005	0.0066	90.6	94.3	98.2
AFTGEE -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0002	0.0069	89.3	94.1	98.5
W_1	-0.0005	0.0069	89.1	94.3	98.5
W_2	-0.0003	0.0065	89.4	94.6	98.3
W_{12}	-0.0005	0.0065	90.6	94.4	98.3

Table 3.12: Comparative Analysis of $\hat{\beta}_2$: Standard AFT vs AFTGEE with 30% censored data for sample size $n = 200$ and $\sigma_\epsilon = 0.4$ (Case VI)

Standard AFT -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0017	0.0396	86.4	93.8	98.5
W_1	-0.0011	0.0347	87.4	93.9	98.8
W_2	-0.0005	0.0396	86.5	94.1	98.6
W_{12}	-0.0017	0.0295	89.7	94.4	98.3
AFTGEE -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0017	0.0396	86.3	93.9	98.5
W_1	-0.0011	0.0347	87.5	93.7	98.8
W_2	-0.0005	0.0396	86.4	94.0	98.6
W_{12}	-0.0017	0.0295	89.6	94.4	98.3

Table 3.13: Comparative Analysis of $\hat{\beta}_1$: Standard AFT vs AFTGEE with 10% censored data for sample size $n = 500$ and $\sigma_\epsilon = 0.4$ (Case VII)

Standard AFT -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0003	0.0038	89.7	94.3	98.4
W_1	-0.0004	0.0038	88.8	94.2	98.6
W_2	-0.0002	0.0034	89.4	94.1	98.5
W_{12}	-0.0001	0.0030	91.4	95.4	98.6
AFTGEE -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0003	0.0038	89.8	94.3	98.5
W_1	-0.0004	0.0038	88.7	94.1	98.6
W_2	-0.0002	0.0034	89.3	94.0	98.5
W_{12}	-0.0001	0.0030	91.4	95.5	98.6

Table 3.14: Comparative Analysis of $\hat{\beta}_2$: Standard AFT vs AFTGEE with 10% censored data for sample size $n = 500$ and $\sigma_\epsilon = 0.4$ (Case VII)

Standard AFT -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0013	0.0410	87.4	93.5	98.4
W_1	-0.0006	0.0378	87.6	93.7	98.6
W_2	-0.0002	0.0410	87.3	94.0	98.6
W_{12}	-0.0005	0.0354	88.7	94.0	98.5
AFTGEE -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0013	0.0410	87.5	93.5	98.5
W_1	-0.0007	0.0378	87.6	93.7	98.6
W_2	-0.0002	0.0410	87.4	94.2	98.6
W_{12}	-0.0006	0.0353	88.6	94.3	98.6

Table 3.15: Comparative Analysis of $\hat{\beta}_1$: Standard AFT vs AFTGEE with 20% censored data for sample size $n = 500$ and $\sigma_\epsilon = 0.4$ (Case VIII)

Standard AFT -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0001	0.0040	89.9	94.4	98.5
W_1	-0.0002	0.0041	89.7	94.4	98.6
W_2	0.0000	0.0037	89.2	93.4	98.7
W_{12}	0.0000	0.0035	89.9	94.9	99.4
AFTGEE -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0001	0.0040	90.0	94.1	98.5
W_1	-0.0002	0.0040	89.7	94.3	98.6
W_2	0.0000	0.0037	89.2	93.5	98.7
W_{12}	0.0000	0.0035	90.0	94.9	99.5

Table 3.16: Comparative Analysis of $\hat{\beta}_2$: Standard AFT vs AFTGEE with 20% censored data for sample size $n = 500$ and $\sigma_\epsilon = 0.4$ (Case VIII)

Standard AFT -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0015	0.0441	88.0	94.3	98.8
W_1	-0.0008	0.0417	87.0	93.3	99.1
W_2	-0.0006	0.0441	88.2	94.3	98.8
W_{12}	-0.0006	0.0409	87.8	93.8	98.6
AFTGEE -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0016	0.0441	88.1	94.4	98.8
W_1	-0.0009	0.0418	86.9	93.5	99.1
W_2	-0.0006	0.0441	88.3	94.3	98.8
W_{12}	-0.0007	0.0409	87.8	93.8	98.7

Table 3.17: Comparative Analysis of $\hat{\beta}_1$: Standard AFT vs AFTGEE with 30% censored data for sample size $n = 500$ and $\sigma_\epsilon = 0.4$ (Case IX)

Standard AFT -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	0.0001	0.0043	90.8	94.6	98.4
W_1	0.0000	0.0043	90.5	94.8	98.6
W_2	0.0001	0.0041	88.7	94.7	99.1
W_{12}	0.0001	0.0039	91.3	96.1	99.1

AFTGEE -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	0.0001	0.0043	90.6	94.6	98.6
W_1	0.0000	0.0043	90.6	94.6	98.6
W_2	0.0001	0.0041	88.5	94.7	99.0
W_{12}	0.0001	0.0039	91.3	96.2	99.1

Table 3.18: Comparative Analysis of $\hat{\beta}_2$: Standard AFT vs AFTGEE with 30% censored data for sample size $n = 500$ and $\sigma_\epsilon = 0.4$ (Case IX)

After increasing the sample size to 200 and 500, the study provides evidence that biases remain insignificant across all models. A comprehensive analysis reveals that, with larger sample sizes, the AFTGEE models produce outcomes that are more similar to those of the Standard AFT models.

Across all sample sizes and censoring percentages, models associated with W_{12} exhibit the lowest standard deviations in both the Standard AFT and AFTGEE models, consistent with the findings from scenarios 1 and 2. When estimating β_1, the analysis consistently reveals that models with weights W_1 exhibit the second lowest standard deviations in both the Standard AFT and AFTGEE models. In contrast, models with weights W_2 show the second lowest standard deviations when estimating β_2.

When comparing coverage probabilities across all models, they consistently reach nominal levels of 90%, 95%, and 99%. However, the weighted model with W_{12} achieves the best coverage probabilities, particularly when the sample size is large.

Similar to the results presented in Scenarios 1 and 2, it can be concluded that models incorporating auxiliary information based on X_1 and both X_1 and X_2 provide good estimates for β_1, whereas models incorporating auxiliary information based on X_2 offer better estimates for β_2.

Furthermore, considering all the tables, the proposed weighted models demonstrate significant improvements in results with high censoring percentages and large sample sizes.

This indicates that with increased sample sizes, the proposed weighted methods perform well in both Standard AFT and AFTGEE models.

3.1.8 Scenario 04 : Impact of the σ_ϵ

To investigate the impact of error variability on model performance, we increased σ_ϵ to 1.0. We repeated the entire simulation study that was originally conducted with $\sigma_\epsilon = 0.4$. This allows us to examine how increased error variability influences the performance of weighted AFT and AFTGEE models, particularly in terms of parameter estimation and the effects of censoring percentages and sample sizes. The simulation results with $\sigma_\epsilon = 1.0$ are presented in Table A.1 to Table A.18 in Appendix A for varying sample sizes and censoring percentages.

The results from both the Standard AFT and AFTGEE models show that the biases in estimating the parameters $\hat{\beta}_1$ and $\hat{\beta}_2$ are are negligible and tend towards zero, even for significant σ_ϵ.

Increasing the error standard deviation from 0.4 to 1.0 leads to higher standard deviations across all models. This outcome was expected as a result of the increased variation in error.

The models with weight W_{12}, which incorporate auxiliary information from both X_1 and X_2, provide the lowest standard deviation and consistently approach nominal levels of 90%, 95%, and 99%. Moreover, as the error variance increases, the gap in standard deviations between these W_{12} weighted models and others becomes higher.

Models incorporating auxiliary information from both X_1 and X_2 consistently perform better than other models, especially with larger sample sizes and higher censoring rates.

Similar to the scenario with lower error standard deviation, the weighted models incorporating auxiliary information from X_1 alone and from both X_1 and X_2 provide better estimates for β_1. Conversely, the weighted models incorporating information from X_2 alone and from both X_1 and X_2 offer better estimates for β_2. Considering the

overall results, we can conclude that the proposed weighted method remains stable under varying levels of error variability.

3.2 Application of the proposed Weighted AFT Model to Real-Time Data

The Leukemia Survival dataset explores the survival outcomes of 1,043 patients diagnosed with acute myeloid leukemia, initially examined by Henderson et al. [2002]. The focus of the investigation is to identify potential spatial variations in survival, considering established individual prognostic factors such as **age**, **sex**, white blood cell count (**wbc**) at diagnosis, and the Townsend score (**tpi**), where higher values signify less affluent areas.

In this study, we utilize data encompassing survival duration (measured in years), final status at the end of observation (0 - right-censored, 1- dead), **age** (in years), **sex** (0 for female, 1 for male), white blood cell count recorded at diagnosis (**wbc** limited to 500), and the Townsend score (**tpi**).

Let's define the variables as follows. X_1 represents the variable for **sex**, X_2 relates to the variable for **age**, X_3 designates the variable for white blood cell count (**wbc**), and X_4 represents the variable for the Townsend score (**tpi**).

Among the 1,043 observations in the dataset, 164 were censored, making up 15.72% of the total. Next, the dataset was split into two distinct parts. The first part, including 200 observations, served as auxiliary data with a censoring rate of 12.50%. The remaining 843 observations were included in the analysis, and this subset had a censoring percentage of 16.49%.

We derived weights from the covariates **age** (W_2), white blood cell count (W_3), and Townsend score (W_4), along with their combinations (W_{23}, W_{24}, W_{34}, W_{234}). Due to its insignificance, the **sex** covariate was only included in the weights with all covariates model (W_{1234}).

Variables in the models	**EL Weights**
$\boldsymbol{X_2}$	W_2
$\boldsymbol{X_3}$	W_3
$\boldsymbol{X_4}$	W_4
$\boldsymbol{X_2}$ and $\boldsymbol{X_3}$	W_{23}
$\boldsymbol{X_2}$ and $\boldsymbol{X_4}$	W_{24}
$\boldsymbol{X_3}$ and $\boldsymbol{X_4}$	W_{34}
$\boldsymbol{X_2}$, $\boldsymbol{X_3}$ and $\boldsymbol{X_4}$	W_{234}
$\boldsymbol{X_1}$, $\boldsymbol{X_2}$, $\boldsymbol{X_3}$ and $\boldsymbol{X_4}$	W_{1234}

Table 3.19: Weights associated with variable combinations

When evaluating the effectiveness of different weights for demonstrating a proposed weighted approach, our attention turns to the comparison of five particular models: No Weights, W_2, W_{23}, W_{234}, and W_{1234}.

Table 3.20 shows a comparison between the Standard AFT model and various weighted models. The results are presented for the model without weights, as well as for weighted models with W_2, W_{23}, W_{234}, and W_{1234}. For each weighting scheme, the table reports the estimated coefficients ($\hat{\beta}$), SE, and p-values for the covariates sex, age, wbc, and tpi.

Table 3.21 compares the Standard AFTGEE model with several weighted models, with weights W_2, W_{23}, W_{234}, and W_{1234}. The table presents the outcomes for both the unweighted model and weighted models utilizing W_2, W_{23}, W_{234}, and W_{1234}. Each model offers the estimated coefficients, SE, and p-values for the covariates.

Tables 3.20 and 3.21 show that the use of the W_{1234} weight consistently results in the lowest standard error among the different models. This suggests that the precision of estimation is enhanced when utilizing auxiliary information based on all covariates for both the standard AFT and AFTGEE models.

The p-values of the covariate **tpi** (X_4) become statistically more significant when the W_{234} and W_{1234} weights are included in the models. Furthermore, the p-values of the covariate **sex** (X_4) become more insignificant with W_{1234} weights. This indicates that this study accurately captures past data properties, as the **tpi** covariate is significant and the **sex** covariate is insignificant in past data.

The findings strongly show that using W_{1234} weight not only improves the accuracy of the estimates but also enhances the significance of essential variables in the models.

Table 3.20: Comparison of Standard AFT with different Weights

	Standard AFT Model (No Weights)		
Covariates	$\hat{\boldsymbol{\beta}}$	***SE***	***P-value***
sex (X_1)	-0.0692	0.1414	0.6246
age (X_2)	-0.0588	0.0038	0.0000
wbc (X_3)	-0.0063	0.0011	0.0000
tpi (X_4)	-0.0568	0.0199	0.0044

	AFT model with weight W_2		
Covariates	$\hat{\boldsymbol{\beta}}$	***SE***	***P-value***
sex (X_1)	-0.0537	0.1565	0.7314
age (X_2)	-0.0442	0.0017	0.0000
wbc (X_3)	-0.0068	0.0012	0.0000
tpi (X_4)	-0.0611	0.0232	0.0086

	AFT model with weight W_{23}		
Covariates	$\hat{\boldsymbol{\beta}}$	***SE***	***P-value***
sex (X_1)	-0.0520	0.1528	0.7337
age (X_2)	-0.0456	0.0016	0.0000
wbc (X_3)	-0.0082	0.0002	0.0000
tpi (X_4)	-0.0601	0.0225	0.0076
	AFT model with weight W_{234}		
Covariates	$\hat{\boldsymbol{\beta}}$	***SE***	***P-value***
sex (X_1)	-0.0470	0.1549	0.7613
age (X_2)	-0.0459	0.0015	0.0000
wbc (X_3)	-0.0079	0.0002	0.0000
tpi (X_4)	-0.0733	0.0048	0.0000
	AFT model with weight W_{1234}		
Covariates	$\hat{\boldsymbol{\beta}}$	***SE***	***P-value***
sex (X_1)	-0.0025	0.0348	0.9425
age (X_2)	-0.0458	0.0017	0.0000
wbc (X_3)	-0.0079	0.0001	0.0000
tpi (X_4)	-0.0729	0.0046	0.0000

Table 3.21: Comparison of AFTGEE with different Weights

	Standard AFTGEE model (No Weights)		
Covariates	$\hat{\boldsymbol{\beta}}$	***SE***	***P-value***
sex(X_1)	-0.0723	0.1418	0.6103
age (X_2)	-0.0590	0.0037	0.0000
wbc (X_3)	-0.0063	0.0011	0.0000
tpi (X_4)	-0.0576	0.0200	0.0039

	AFTGEE model with weight W_2		
Covariates	$\hat{\boldsymbol{\beta}}$	***SE***	***P-value***
sex (X_1)	-0.0585	0.1561	0.7081
age (X_2)	-0.0445	0.0016	0.0000
wbc (X_3)	-0.0069	0.0012	0.0000
tpi (X_4)	-0.0617	0.0232	0.0077

	AFTGEE model with weight W_{23}		
Covariates	$\hat{\boldsymbol{\beta}}$	***SE***	***P-value***
sex (X_1)	-0.0566	0.1527	0.7107
age (X_2)	-0.0460	0.0016	0.0000
wbc (X_3)	-0.0082	0.0002	0.0000
tpi (X_4)	-0.0607	0.0225	0.0070
	AFTGEE model with weight W_{234}		
Covariates	$\hat{\boldsymbol{\beta}}$	***SE***	***P-value***
sex (X_1)	-0.0516	0.1550	0.7391
age (X_2)	-0.0463	0.0015	0.0000
wbc (X_3)	-0.0079	0.0002	0.0000
tpi (X_4)	-0.0737	0.0051	0.0000
	AFTGEE model with weight W_{1234}		
Covariates	$\hat{\boldsymbol{\beta}}$	***SE***	***P-value***
sex (X_1)	-0.0067	0.0358	0.8509
age (X_2)	-0.0462	0.0017	0.0000
wbc (X_3)	-0.0079	0.0002	0.0000
tpi (X_4)	-0.0734	0.0049	0.0000

The following figures illustrate the residual diagnostics for standard AFT models and AFTGEE models, as well as their weighted models with W_{1234}. This analysis helps assess whether residuals exhibit randomness and independence, which are essential for validating model assumptions.

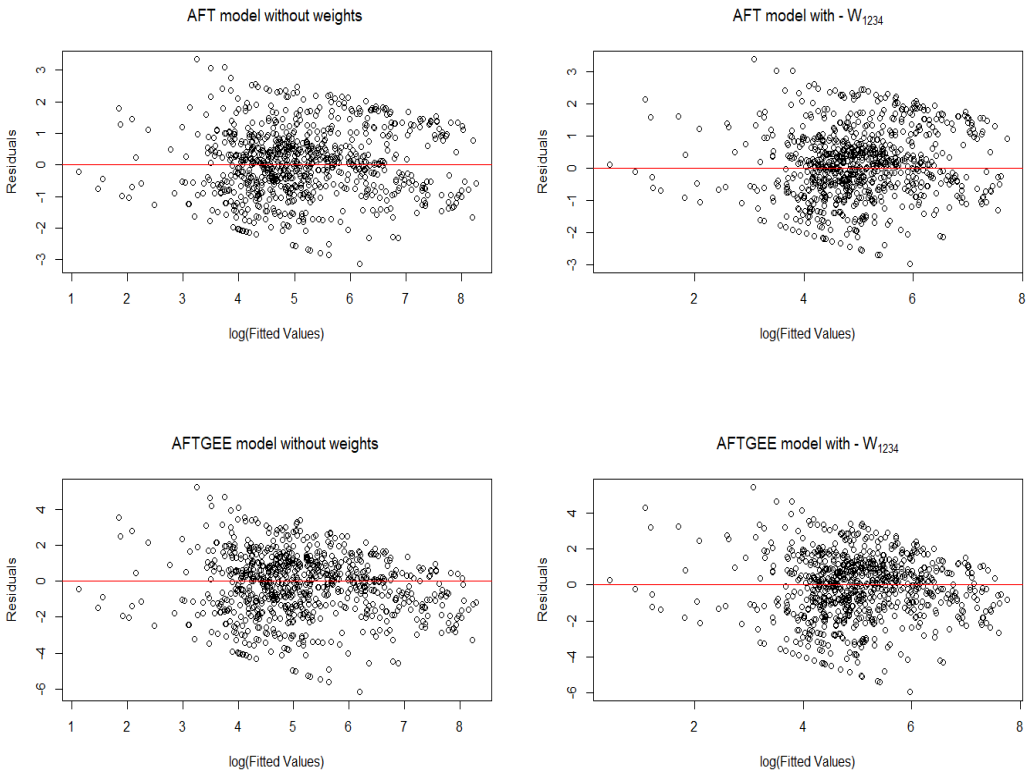

Figure 3.1: Residuals vs. log(Fitted values) plots

In Figure 3.1, we plotted the logarithm of the fitted values against the residuals for all models. The graphs show that the residuals are randomly distributed in a horizontal band centered around the zero line, with no clear pattern, indicating that the models fit the data appropriately and the error terms are independent. Even with the weighted models, there is no systematic pattern of positive and negative values, suggesting that the assumptions of linearity, constant variance, and independence of error terms are met for both standard and weighted models.

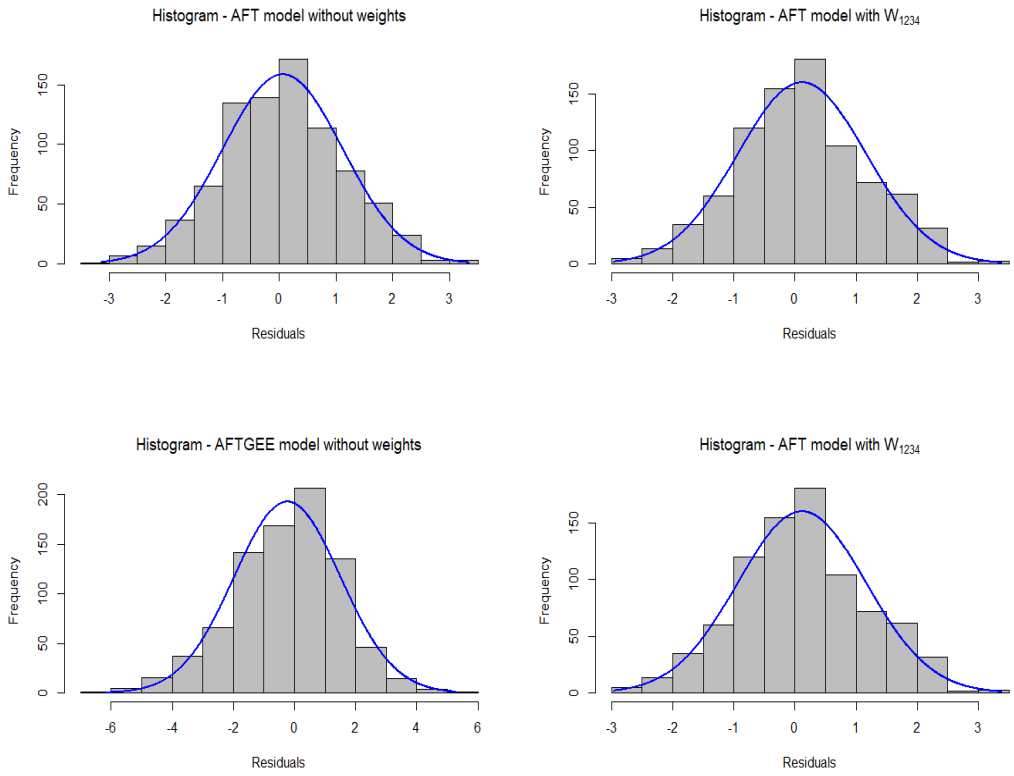

Figure 3.2: Histograms

The Figure 3.2 shows that both the models without weights and the models with weights W_{1234} exhibit roughly symmetric, bell-shaped distributions centered around zero. There are no obvious clusters or outliers in the residual distributions, suggesting uniform variance across the range of residuals in all models. Based on these observations, we can conclude that both the standard models and the models with W_{1234} weights are approximately normally distributed.

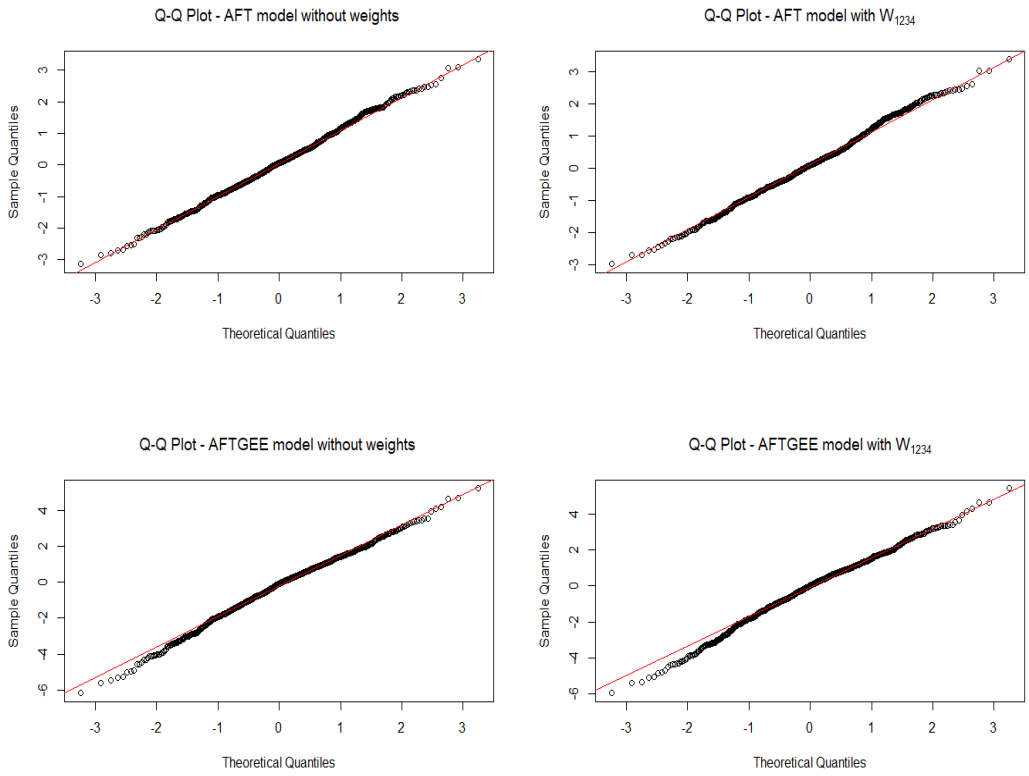

Figure 3.3: Q-Q Plots

Considering the Q-Q plots in Figure 3.3, it is evident that the residuals of both standard and weighted models lie almost entirely along the straight line in the Q-Q plots, indicating that the residuals are approximately normally distributed.

Chapter 4

Summary and Future Work

4.1 Summary

This research propose a Weighted AFT model, enhancing the Standard AFT model and AFTGEE model by incorporating auxiliary information. The models utilize empirical likelihood probabilities as weights derived from previous studies, improving accuracy by considering individual observation qualities. This methodology effectively overcomes the challenges associated with managing censored observations, resulting in more reliable and accurate estimates.

An extensive simulation study was conducted to evaluate the effectiveness of weighted models with both partial and complete auxiliary information. Standard AFT and AFTGEE models were employed to test these weighted models, along with models without weights.

This study examined both discrete and continuous covariates, using three different sample sizes: $n = 100, 200$ & 500, with different error standard deviations, such as $\sigma_\epsilon = 0.4$ & 1.0. Each sample size had varying levels of censoring at 10%, 20%, and 30%.

According to the simulation study, all biases were found to be negligible, suggesting that the estimated parameters closely approximate the true values for all models. Therefore, comparisons between models were made using standard deviation and coverage probabilities. When evaluating β_1 estimates associated with discrete covariates,

both weighted models with auxiliary information based on X_1 and auxiliary information based on X_1 and X_2 can be considered good estimates compared to other models.

As the sample size increases, the W_{12} models consistently demonstrate the lowest standard deviation and the highest coverage probability, suggesting that they offer more precise estimates. However, it's worth noting that in cases of small sample sizes and low censoring percentages, models with W_1 often display accurate coverage probabilities compared to W_{12} models. When the censoring percentage increases, weighted models with W_{12} provide the most accurate estimates in both standard AFT models and AFTGEE models.

When evaluating estimates for β_2 associated with a continuous variable, both weighted models incorporating auxiliary information based on X_2 and auxiliary information based on X_1 and X_2 are considered more accurate than alternative models. However, models with auxiliary information based on X_1 and X_2 can be identified as the best estimate when considering standard deviation and coverage probabilities across all confidence levels. As the sample size increases and the censoring percentage rises, models with the W_{12} weight consistently provide the best results.

When σ_ϵ is low, both standard AFT and AFTGEE models yield very similar estimates. However, as the standard deviation of the error increases, they produce two distinct results.

When estimating coefficients using weighted models, it is prefered to incorporate weights based on all covariate as auxiliary information. Excluding the covariate from the weighting process negatively impacts the outcome, resulting in an increase in the standard deviation rather than improving the estimation

The application of the proposed method to real data, specifically using "The Leukemia Survival Data," serves as a valuable illustration of the weighted AFT approach. The results demonstrate that utilizing the correct weight allows for the estimation of coefficients with smaller standard errors compared to the model with no weights. Additionally, the analysis reveals that the proposed method enhances the significance level by incorporating auxiliary information.

Overall results reveal that the Weighted AFT model, which incorporates auxiliary information through empirical likelihood, is highly beneficial for enhancing the

precision and efficiency of estimates.

4.2 Future Work

Based on the findings of this study, there are several opportunities to refine and expand the use of the Weighted AFT model. One such direction is to enhance the model to handle left-censored and interval-censored data, which are frequently encountered in survival analysis. Furthermore, the model could be applied to multivariate survival data. These extensions would make the Weighted AFT model more flexible, precise, and suitable for a broader range of survival analysis applications.

Bibliography

Ajay, A., Singh, S., et al. (2021). Analytical models of survival analysis: Concepts and their applications. *Human Journals*.

Ali, Z., Hosseini, M., Mahmoodi, M., Mohammad, K., Zeraati, H., and Naieni, K. H. (2015). A comparison between accelerated failure-time and cox proportional hazard models in analyzing the survival of gastric cancer patients. *Iranian journal of public health*, 44(8):1095.

Broström, G. (2014). *eha: Event History Analysis*. R package version 2.4-1. URL http://CRAN.R-project.org/package=eha.

Buckley, J. and James, I. (1979). Linear regression with censored data. *Biometrika*, 66(3):429–436.

Chen, J., Sitter, R., and Wu, C. (2002). Using empirical likelihood methods to obtain range restricted weights in regression estimators for surveys. *Biometrika*, 89(1):230–237.

Chiou, S. H., Kang, S., and Yan, J. (2014a). *aftgee: Accelerated Failure Time Model with Generalized Estimating Equations*. R package version 1.0-0. URL http://CRAN.R-project.org/package=aftgee.

Chiou, S. H., Kang, S., and Yan, J. (2014b). Fitting accelerated failure time models in routine survival analysis with r package aftgee. *Journal of Statistical Software*, 61:1–23.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, 34(2):187–202.

Cui, J. (2005). Buckley–james method for analyzing censored data, with an application to a cardiovascular disease and an hiv/aids study. *The Stata Journal*, 5(4):517–526.

David, C. R. et al. (1972). Regression models and life tables (with discussion). *Journal of the Royal Statistical Society*, 34(2):187–220.

DiCiccio, T., Hall, P., and Romano, J. (1991). Empirical likelihood is bartlett-correctable. *the Annals of Statistics*, pages 1053–1061.

Dong, Q., Liu, B., and Zhao, H. (2023). Weighted least squares model averaging for accelerated failure time models. *Computational Statistics & Data Analysis*, 184:107743.

Emmerson, J. and Brown, J. (2021). Understanding survival analysis in clinical trials. *Clinical Oncology*, 33(1):12–14.

Fang, K.-T., Li, G., Lu, X., and Qin, H. (2013). An empirical likelihood method for semiparametric linear regression with right censored data. *Computational and Mathematical Methods in Medicine*, 2013(1):469373.

Granville, K. and Fan, Z. (2014). Buckley-james estimator of aft models with auxiliary covariates. *PloS one*, 9(8):e104817.

Hall, P. (1990). Pseudo-likelihood theory for empirical likelihood. *The Annals of Statistics*, pages 121–140.

Harrell Jr, F. E. (2014). *rms: Regression Modeling Strategies*. R package version 4.2-1. URL http://CRAN.R-project.org/package=rms.

Henderson, R., Shimakura, S., and Gorst, D. (2002). Modeling spatial variation in leukemia survival data. *Journal of the American Statistical Association*, 97(460):965–972.

Huang, J., Ma, S., and Xie, H. (2006). Regularized estimation in the accelerated failure time model with high-dimensional covariates. *Biometrics*, 62(3):813–820.

Huang, L. and Jin, Z. (2007). Lss: An s-plus/r program for the accelerated failure time model to right censored data based on least-squares principle. *Computer methods and programs in biomedicine*, 86(1):45–50.

Jin, Z., Lin, D., Wei, L., and Ying, Z. (2003). Rank-based inference for the accelerated failure time model. *Biometrika*, 90(2):341–353.

Jin, Z., Lin, D., and Ying, Z. (2006a). On least-squares regression with censored data. *Biometrika*, 93(1):147–161.

Jin, Z., Lin, D., and Ying, Z. (2006b). Rank regression analysis of multivariate failure time data based on marginal linear models. *Scandinavian Journal of Statistics*, 33(1):1–23.

Kalbfleisch, J. D. and Prentice, R. L. (1980). *The Statistical Analysis of Failure Time Data*. John Wiley & Sons.

Kalbfleisch, J. D. and Prentice, R. L. (2002). *The Statistical Analysis of Failure Time Data*, volume 360. John Wiley & Sons.

Kalbfleisch, J. D. and Prentice, R. L. (2011). *The statistical analysis of failure time data*. John Wiley & Sons.

Kim, E., MacEachern, S. N., and Peruggia, M. (2024). melt: Multiple empirical likelihood tests in r. *Journal of Statistical Software*, 108:1–33.

Klein, J. P., Moeschberger, M. L., et al. (2003). *Survival analysis: techniques for censored and truncated data*, volume 1230. Springer.

Kong, F. and Yu, Q. (2007). Asymptotic distributions of the buckley-james estimator under nonstandard conditions. *Statistica Sinica*, 17(1):341–360.

Koul, H., Susarla, V., and Van Ryzin, J. (1981). Regression analysis with randomly right-censored data. *The Annals of statistics*, pages 1276–1288.

Lai, T. L., Ying, Z., and Zheng, Z. (1995). Asymptotic normality of a class of adaptive statistics with applications to synthetic data methods for censored regression. *Journal of Multivariate Analysis*, 52(2):259–279.

Li, G. and Wang, Q.-H. (2003). Empirical likelihood regression analysis for right censored data. *Statistica Sinica*, pages 51–68.

Liu, E., Liu, R. Y., and Lim, K. (2023). Using the weibull accelerated failure time regression model to predict time to health events. *Applied Sciences*, 13(24):13041.

Mustefa, Y. A. and Chen, D.-G. (2021). Accelerated failure-time model with weighted least-squares estimation: application on survival of hiv positives. *Archives of Public Health*, pages 1–7.

Owen, A. (1988). Empirical likelihood ratio confidence intervals for a single functional. *Biometrika*, 75(2):237–249.

Owen, A. (1990). Empirical likelihood ratio confidence regions. *The annals of statistics*, 18(1):90–120.

Owen, A. (1991). Empirical likelihood for linear models. *The Annals of Statistics*, pages 1725–1747.

Owen, A. (2001). *Empirical likelihood*. Chapman and Hall/CRC.

Prentice, R. L. (1978). Linear rank tests with right censored data. *Biometrika*, 65(1):167–179.

Qin, G. and Jing, B.-Y. (2001). Empirical likelihood for censored linear regression. *Scandinavian Journal of Statistics*, 28(4):661–673.

Qin, J. and Lawless, J. (1994). Empirical likelihood and general estimating equations. *the Annals of Statistics*, 22(1):300–325.

Swindell, W. R. (2009). Accelerated failure time models provide a useful statistical framework for aging research. *Experimental gerontology*, 44(3):190–200.

Therneau, T. M. and Lumley, T. (2015). Package 'survival'. *R Top Doc*, 128(10):28–33. `https://cran.r-project.org/package=survival`.

Turkson, A. J., Ayiah-Mensah, F., and Nimoh, V. (2021). Handling censoring and censored data in survival analysis: a standalone systematic literature review. *International journal of mathematics and mathematical sciences*, pages 1–16.

Vasudevan, C., Variyath, A. M., and Fan, Z. (2019). Weighted censored quantile regression. *Survey Methodology 45-1*, 45(1):127–144.

Wu, C. (2005). Algorithms and r codes for the pseudo empirical likelihood method in survey sampling. *Survey Methodology*, 31(2):239.

Zhang, D. (2019). *Weighted Accelerated Failure Time Models and Their Applications in Clustered Data*. The University of Texas at Dallas.

Zhou, M. (2023). Empirical likelihood ratio for censored/truncated data. *Comprehensive R Archive Network (CRAN)*. `https://www.ms.uky.edu/~mai/EmpLik.html`.

Appendix A

Simulation Study Results: $\sigma_\epsilon = 1.0$

Standard AFT -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0075	0.2021	90.3	95.1	99.3
W_1	-0.0082	0.1215	90.8	95.3	99.1
W_2	0.0059	0.2048	90.5	94.9	98.9
W_{12}	-0.0074	0.0804	88.6	93.2	97.8

AFTGEE -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0090	0.2017	90.5	94.8	99.3
W_1	-0.0097	0.1214	90.7	95.1	99.0
W_2	0.0043	0.2043	90.6	94.6	99.2
W_{12}	-0.0090	0.0807	88.2	92.4	97.8

Table A.1: Comparative Analysis of $\hat{\beta}_1$: Standard AFT vs AFTGEE with 10% censored data for sample size $n = 100$ and $\sigma_\epsilon = 1.0$ (Case X)

Standard AFT -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	0.0007	0.0217	89.6	94.0	97.9
W_1	-0.0007	0.0220	89.8	93.3	97.9
W_2	-0.0004	0.0135	89.7	94.8	99.2
W_{12}	-0.0005	0.0095	88.3	93.5	96.8

AFTGEE -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	0.0008	0.0217	89.3	93.6	98.0
W_1	-0.0006	0.0220	89.7	93.1	98.0
W_2	-0.0003	0.0136	89.7	94.8	99.3
W_{12}	-0.0004	0.0095	87.8	93.4	96.8

Table A.2: Comparative Analysis of $\hat{\beta}_2$: Standard AFT vs AFTGEE with 10% censored data for sample size $n = 100$ and $\sigma_\epsilon = 1.0$ (Case X)

Standard AFT -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0152	0.2191	88.1	93.1	98.2
W_1	-0.0079	0.1458	89.7	94.5	98.3
W_2	-0.0015	0.2228	88.7	94.3	98.4
W_{12}	-0.0103	0.1213	88.4	94.0	98.1

AFTGEE -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0152	0.2198	88.3	93.1	98.1
W_1	-0.0079	0.1468	89.3	94.7	98.2
W_2	-0.0017	0.2236	89.0	93.9	98.5
W_{12}	-0.0104	0.1225	88.7	93.6	97.7

Table A.3: Comparative Analysis of $\hat{\beta}_1$: Standard AFT vs AFTGEE with 20% censored data for sample size $n = 100$ and $\sigma_\epsilon = 1.0$ (Case XI)

Standard AFT -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0003	0.0219	88.5	93.8	98.8
W_1	-0.0016	0.0222	89.7	93.8	98.8
W_2	-0.0002	0.0155	91.0	95.5	98.7
W_{12}	0.0002	0.0137	88.5	94.1	98.0

AFTGEE -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0002	0.0220	88.7	93.6	98.8
W_1	-0.0015	0.0223	89.5	93.8	98.7
W_2	-0.0001	0.0156	91.0	95.4	98.8
W_{12}	0.0003	0.0137	88.2	93.7	98.0

Table A.4: Comparative Analysis of $\hat{\beta}_2$: Standard AFT vs AFTGEE with 20% censored data for sample size $n = 100$ and $\sigma_\epsilon = 1.0$ (Case XI)

Standard AFT -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0088	0.2289	87.1	94.4	98.8
W_1	-0.0005	0.1655	90.7	95.8	98.9
W_2	0.0028	0.2317	88.0	94.5	98.9
W_{12}	-0.0017	0.1573	88.8	93.5	98.4

AFTGEE -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0091	0.2293	87.7	94.2	98.8
W_1	-0.0006	0.1663	91.0	95.6	98.8
W_2	0.0022	0.2322	88.6	94.4	99.0
W_{12}	-0.0023	0.1582	88.3	93.3	98.6

Table A.5: Comparative Analysis of $\hat{\beta}_1$: Standard AFT vs AFTGEE with 30% censored data for sample size $n = 100$ and $\sigma_\epsilon = 1.0$ (Case XII)

Standard AFT -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0010	0.0226	89.5	94.9	99.0
W_1	-0.0022	0.0229	90.0	94.7	98.9
W_2	-0.0005	0.0177	90.8	95.4	99.2
W_{12}	-0.0003	0.0168	88.3	93.5	98.4

AFTGEE -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0009	0.0226	88.9	94.7	98.8
W_1	-0.0021	0.0229	90.6	94.7	98.8
W_2	-0.0003	0.0178	91.0	94.9	99.1
W_{12}	-0.0002	0.0169	87.8	92.9	98.3

Table A.6: Comparative Analysis of $\hat{\beta}_2$: Standard AFT vs AFTGEE with 30% censored data for sample size $n = 100$ and $\sigma_\epsilon = 1.0$(Case XII)

Standard AFT -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0069	0.1456	88.9	94.4	99.1
W_1	-0.0043	0.0852	90.8	95.2	99.3
W_2	0.0009	0.1468	88.8	94.4	99.1
W_{12}	-0.0082	0.0579	88.0	93.0	96.9

AFTGEE -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0076	0.1455	88.4	94.4	99.0
W_1	-0.0050	0.0853	90.8	95.5	99.2
W_2	0.0003	0.1466	89.0	94.3	99.0
W_{12}	-0.0087	0.0582	87.9	93.0	96.8

Table A.7: Comparative Analysis of $\hat{\beta}_1$: Standard AFT vs AFTGEE with 10% censored data for sample size $n = 200$ and $\sigma_\epsilon = 1.0$ (Case XIII)

Standard AFT -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0008	0.0148	87.8	94.8	98.7
W_1	-0.0017	0.0148	88.6	94.3	98.8
W_2	-0.0010	0.0089	91.6	95.8	99.1
W_{12}	-0.0007	0.0068	87.6	93.0	97.6

AFTGEE -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0008	0.0148	88.0	94.8	98.8
W_1	-0.0017	0.0149	88.4	94.2	98.8
W_2	-0.0010	0.0090	91.8	95.9	99.1
W_{12}	-0.0007	0.0069	87.5	93.0	97.7

Table A.8: Comparative Analysis of $\hat{\beta}_2$: Standard AFT vs AFTGEE with 10% censored data for sample size $n = 200$ and $\sigma_\epsilon = 1.0$ (Case XIII)

Standard AFT -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0052	0.1514	88.9	94.3	99.0
W_1	-0.0034	0.0992	90.9	94.5	98.8
W_2	0.0026	0.1530	89.0	94.3	99.1
W_{12}	-0.0071	0.0822	89.6	93.9	98.3

AFTGEE -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0057	0.1510	88.6	94.4	98.9
W_1	-0.0038	0.0991	90.8	94.6	98.9
W_2	0.0022	0.1527	88.7	94.5	99.0
W_{12}	-0.0075	0.0826	89.7	93.5	98.1

Table A.9: Comparative Analysis of $\hat{\beta}_1$: Standard AFT vs AFTGEE with 20% censored data for sample size $n = 200$ and $\sigma_\epsilon = 1.0$ (Case XIV)

Standard AFT -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0008	0.0155	89.4	93.8	98.7
W_1	-0.0017	0.0155	88.9	93.7	98.7
W_2	-0.0010	0.0109	88.9	94.4	98.2
W_{12}	-0.0007	0.0096	87.8	92.3	98.2

AFTGEE -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0007	0.0155	89.6	93.5	98.7
W_1	-0.0016	0.0156	89.4	93.6	98.6
W_2	-0.0009	0.0109	88.7	94.2	98.3
W_{12}	-0.0005	0.0097	87.3	92.5	98.0

Table A.10: Comparative Analysis of $\hat{\beta}_2$: Standard AFT vs AFTGEE with 20% censored data for sample size $n = 200$ and $\sigma_\epsilon = 1.0$ (Case XIV)

Standard AFT -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0007	0.1595	88.9	94.2	98.5
W_1	0.0004	0.1151	90.7	94.9	98.6
W_2	0.0071	0.1611	88.8	94.8	98.8
W_{12}	-0.0025	0.1060	89.4	94.9	98.9

AFTGEE -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0013	0.1593	88.9	93.9	98.9
W_1	-0.0002	0.1151	90.4	94.9	98.7
W_2	0.0064	0.1610	88.8	94.3	98.8
W_{12}	-0.0031	0.1063	89.1	94.7	98.7

Table A.11: Comparative Analysis of $\hat{\beta}_1$: Standard AFT vs AFTGEE with 30% censored data for sample size $n = 200$ and $\sigma_\epsilon = 1.0$ (Case XV)

Standard AFT -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0009	0.0161	89.5	94.5	98.8
W_1	-0.0018	0.0162	89.5	95.3	98.5
W_2	-0.0011	0.0125	88.4	94.2	98.5
W_{12}	-0.0009	0.0119	88.7	94.0	98.8

AFTGEE -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0008	0.0161	89.2	94.7	98.9
W_1	-0.0017	0.0162	89.4	95.2	98.6
W_2	-0.0010	0.0125	89.1	93.8	98.4
W_{12}	-0.0008	0.0119	87.7	93.5	98.7

Table A.12: Comparative Analysis of $\hat{\beta}_2$: Standard AFT vs AFTGEE with 30% censored data for sample size $n = 200$ and $\sigma_\epsilon = 1.0$ (Case XV)

Standard AFT -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0049	0.097	87.5	94.0	98.7
W_1	-0.0032	0.0546	91.1	95.7	99.1
W_2	-0.0017	0.0976	87.3	94.2	98.7
W_{12}	-0.0056	0.0383	88.2	93.2	97.9

AFTGEE -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0049	0.0970	87.4	94.2	98.7
W_1	-0.0032	0.0547	90.8	95.5	99.3
W_2	-0.0017	0.0975	87.6	94.4	98.7
W_{12}	-0.0057	0.0383	88.1	93.5	97.9

Table A.13: Comparative Analysis of $\hat{\beta}_1$: Standard AFT vs AFTGEE with 10% censored data for sample size $n = 500$ and $\sigma_\epsilon = 1.0$ (Case XVI)

Standard AFT -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0005	0.0092	88.2	94.2	98.6
W_1	-0.0009	0.0093	88.3	94.0	98.8
W_2	-0.0006	0.0056	90.5	95.1	98.4
W_{12}	-0.0004	0.0040	90.1	94.6	97.9

AFTGEE -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0006	0.0092	88.2	94.2	98.7
W_1	-0.0009	0.0093	88.2	93.8	98.8
W_2	-0.0006	0.0056	90.7	95.1	98.4
W_{12}	-0.0004	0.0040	90.0	94.9	98.0

Table A.14: Comparative Analysis of $\hat{\beta}_2$: Standard AFT vs AFTGEE with 10% censored data for sample size $n = 500$ and $\sigma_\epsilon = 1.0$ (Case XVI)

Standard AFT -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0054	0.1001	87.5	93.8	98.5
W_1	-0.0033	0.0644	89.0	94.3	98.9
W_2	-0.0023	0.1004	87.2	94.1	98.7
W_{12}	-0.0052	0.0549	88.2	94.3	98.1

AFTGEE -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0054	0.0999	87.8	94.3	98.6
W_1	-0.0032	0.0643	89.0	94.0	99.0
W_2	-0.0022	0.1003	87.4	94.1	98.8
W_{12}	-0.0052	0.0547	88.7	94.4	98.2

Table A.15: Comparative Analysis of $\hat{\beta}_1$: Standard AFT vs AFTGEE with 20% censored data for sample size $n = 500$ and $\sigma_\epsilon = 1.0$ (Case XVII)

Standard AFT -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0004	0.0096	89.4	93.6	98.8
W_1	-0.0008	0.0096	89.7	94.3	99.1
W_2	-0.0005	0.0066	90.2	94.2	98.4
W_{12}	-0.0003	0.0057	90.4	95.5	98.3

AFTGEE -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0004	0.0096	89.6	93.7	99.0
W_1	-0.0007	0.0096	89.4	94.2	99.1
W_2	-0.0005	0.0066	89.6	94.4	97.9
W_{12}	-0.0003	0.0057	90.2	95.4	98.2

Table A.16: Comparative Analysis of $\hat{\beta}_2$: Standard AFT vs AFTGEE with 20% censored data for sample size $n = 500$ and $\sigma_\epsilon = 1.0$ (Case XVII)

Standard AFT -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0022	0.1046	87.9	93.2	98.1
W_1	-0.0004	0.0755	87.4	94.0	98.8
W_2	0.0007	0.1048	87.5	93.3	98.3
W_{12}	-0.0019	0.0703	88.9	93.6	98.1

AFTGEE -$\hat{\beta}_1$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	-0.0023	0.1045	88.3	93.4	98.0
W_1	-0.0006	0.0753	87.7	93.9	99.0
W_2	0.0006	0.1046	88.4	93.6	98.2
W_{12}	-0.0021	0.0700	89.6	93.7	98.2

Table A.17: Comparative Analysis of $\hat{\beta}_1$: Standard AFT vs AFTGEE with 30% censored data for sample size $n = 500$ and $\sigma_\epsilon = 1.0$ (Case XVIII)

Standard AFT -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	0.0000	0.0100	89.9	93.7	99.0
W_1	-0.0003	0.0100	90.4	94.4	98.9
W_2	0.0000	0.0075	89.1	94.2	98.8
W_{12}	0.0001	0.0070	90.1	95.1	98.5

AFTGEE -$\hat{\beta}_2$					
Weights	**Bias**	**sd**	**90% CP**	**95% CP**	**99% CP**
No W	0.0001	0.0100	90.4	93.9	99.1
W_1	-0.0002	0.0100	89.8	94.3	99.0
W_2	0.0000	0.0075	89.1	94.0	99.0
W_{12}	0.0002	0.0070	90.7	95.3	98.5

Table A.18: Comparative Analysis of $\hat{\beta}_2$: Standard AFT vs AFTGEE with 30% censored data for sample size $n = 500$ and $\sigma_\epsilon = 1.0$ (Case XVIII)