PARAMETRIC INFERENCE ON QUANTILE RESIDUAL LIFE

by

Kidane B. Ghebrehawariat

MS, Washington State University, 2009

MS, University of Idaho, 2005

BSc, Asmara University, Eritrea, 1998

Submitted to the Graduate Faculty of

the Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2015

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This dissertation was presented

by

Kidane B. Ghebrehawariat

It was defended on

December 1, 2015

and approved by

Advisor:

Jong-Hyeon Jeong, PhD

Professor

Department of Biostatistics

Graduate School of Public Health

University of Pittsburgh

Co-Advisor:

Ying Ding, PhD

Assistant Professor of Biostatistics

Department of Biostatistics

Graduate School of Public Health

University of Pittsburgh

Committee Member:

Daniel Normolle, PhD

Associate Professor

Department of Biostatistics

Graduate School of Public Health

University of Pittsburgh

Committee Member:

Yu Cheng, PhD

Associate Professor

Department of Statistics

University of Pittsburgh

PARAMETRIC INFERENCE ON QUANTILE RESIDUAL LIFE

Kidane B. Ghebrehawariat, PhD

University of Pittsburgh, 2015

ABSTRACT

The need for residual life analysis arises in many fields including medicine and life testing. For instance, in medicine, a clinician and a patient would be interested in knowing by how long a new drug can extend the life span of that patient. Problems in remaining life time after surviving up to a certain time are often framed and addressed statistically in terms of mean, hazard rate or quantile. The quantile approach enjoys some practical advantages over the other approaches such as robustness, ease of interpretation, and existence. Most of the methodological work on quantile residual life in the literature has been semi-parametric or non-parametric. However, parametric approaches are expected to be optimal or asymptotically efficient under a correct specification of the model. Furthermore, the parametric approach does not require nonparametric estimation of the probability density function of the underlying distribution under informative or noninformative censoring to evaluate the variance of the quantile estimator. In this dissertation, parametric inference procedures for the quantile residual life under competing and non-competing risks settings are developed for the one-sample, two-sample and regression cases. We adopt the accelerated failure time (AFT) framework to incorporate covariates for the regression case. The finite sample properties of the proposed methods are studied through extensive simulations. The simulation results indicate that the proposed methods perform well. The proposed methods are applied to a breast cancer data set.

PUBLIC HEALTH SIGNIFICANCE: The results established in this dissertation will provide new parametric methods to researchers and investigators in public health who conduct quantile

residual life analysis, which will facilitate efficient communication between researchers and stakeholders regarding the efficacy of new interventions.

Keywords: Quantile residual life, Parametric, Breast Cancer, Competing Risk, Accelerated Failure Time.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

ACKNOWLEDGMENTS

I would like to thank my advisors, Dr. Jeong and Dr. Ding, for their guidance throughout the dissertation study. I am also very grateful to members of my PhD dissertation committee, Dr. Normolle and Dr. Cheng, for their valuable feedback on the dissertation and for serving on my dissertation committee. I would like to thank the Department of Biostatistics for providing me financial support during my graduate study. I am also thankful to all the staff of the Department of Biostatistics for all their help.

Finally, I would like to thank my family for their love and support.

1.0 INTRODUCTION

Residual life analyses are used in many fields, including medicine and life-testing. These analyses have been undertaken via mean, quantile and hazard-based approaches. The quantile approach has some theoretical and practical advantages over the other approaches. Indeed, quantile inference methods are more interpretable than hazard based approaches. Quantiles are more stable to tail behavior than means are. Quantiles always exist whereas means many not exist.

Most of the proposed inference procedures for quantile residual life in the literature are non-parametric or semi-parametric. These include one-sample and two-sample procedures for quantile residual life under competing and non-competing risks by Jeong et al. (2008), and Jeong and Fine (2013); and a semi-parametric regression method for inference on quantile residual life by Jung et al. (2009) .

Comparatively, there has been less work on the parametric inference method for quantile residual life. This is despite the fact that parametric procedures are expected to result in a gain in efficiency, simplicity in inference as there is no need for non-parametric density estimation, and simplicity in analysis and interpretation (Koenker and Bassett, 1978; Lin and Spiekerman, 1996).

Parametric procedures for quantile residual life in the literture include a one-sample method and regresson method for inference on the cumulative incidence function by Jeong and Fine (2006, 2007). More recently, Lee and Fine (2011) proposed one-sample and two-sample parametric inferential procedures for cause-specific quantile life time.

In this dissertation, we will develop parametric inferential procedures for the quantile residual life time under non-competing and competing risks scenarios. Three settings will be covered: a single sample, a two sample and a regression setting. The procedures will be illustrated numerically and via simulations using the Weibull distribution. The Weibull distribution is a flexible and widely used parametric survival model.

The dissertation is organized as follows. In Chapter 2, a brief review of the literature is given. In Chapter 3, likelihood based quantile residual life inference procedures for one-sample, two-sample and regression cases under non-competing risks setting are developed. The results are illustrated with simulations and applied to a real data set. In Chapter 4, the results in Chapter 3 are generalized to the competing risks setting. The results are also described by simulations and applied to a real data set.

2.0 LITERATURE REVIEW

We briefly review recent works on inference procedures for the quantile (residual) life time.

Non-parametric procedures include that of Jeong et al. (2008), Jeong and Fine (2013) and Jung et al. (2009). Jeong et al. (2008) proposed one-sample and two-sample inferential procedure for median residual life time for non-competing risks data. Jeong and Fine (2013) extended this procedure to a method for quantile residual life time and competing risks data. Jung et al. (2009) proposed a semi-parametric regression method for inference on quantile residual life time.

Parametric procedures include a one-sample method and regression method for inference on cumulative incidence function by Jeong and Fine (2006, 2007). Since the cumulative incidence function is the inverse of the cause-specific quantile life life time, these methods are related to inference on quantile life time. More recently, Lee and Fine (2011) proposed a parametric inferential procedure for cause-specific quantile life time. All of the above parametric methods are based on the asymptotic properties of the maximum likelihood estimator (MLE) ((in particular, consistency and asymptotic normality)) and the delta method.

In this dissertation we use a parametric approach to develop one-sample, two-sample and regression parametric procedures for the quantile residual life time in both non-competing risks and competing risks setting. When the quantile residual life does not have a closed-form representation, which can arise even in simple settings, its estimation is not obvious. We propose a simple numerical scheme to estimate it.

3.0 NON-COMPETING RISKS

Much of survival analysis deals with the analysis of failure data in which there is a well-defined and single failure type that may be subject to a censoring process that is unrelated to the failure process. We will term this type of failure process the *non-competing risks setting*. This is in contrast to the *competing risks setting*, where there are two or more distinct failure types in addition to a random censoring process (Prentice et al., 1978).

In this chapter, likelihood-based inference procedures for the quantile residual life are derived for the non-competing risks setting. Point estimators and asymptotic variance formulas for the quantile residual life time are given. The procedures are illustrated using simulations and real data analysis. Results for the one-sample, two-sample, and regression cases are presented in Sections 3.2, 3.3, and 3.4, respectively.

We review some basic concepts from survival analysis. These concepts are used in subsequent sections.

3.1 INTRODUCTION

Suppose $T \geq 0$ is failure time random variable with an absolutely continuous distribution function $F(t)$. For example, T can be time to death from diagnosis of a breast cancer patient. Let $f(t) = \frac{dF(t)}{dt}$ and $S(t) = 1 - F(t)$ denote the density and survival functions, respectively,

of T. The hazard function $h(t)$ of T is defined as

$$\begin{aligned} h(t) &= \lim_{\Delta t \to 0} \frac{P[t \leq T < t + \Delta t | T \geq t]}{\Delta t} \\ &= \frac{f(t)}{S(t)}, \end{aligned} \tag{3.1.01}$$

whereas its cumulative hazard function $H(t)$ is given by

$$H(t) = \int_0^t h(u) du.$$

The hazard function quantifies the rate of failure at time t among subjects that are still at risk of failure at time t.

When F is a proper distribution function, i.e., $F(\infty) = 1$, there is a 1-1 correspondence among the different survival quantities. We will point out two such correspondences that we will use later. The survival function $S(t)$ can be written as $S(t) = \exp(-H(t))$; thus, F has the following equivalent representations,

$$\begin{aligned} F(t) = \int_0^t f(u)\, du &= \int_0^t S(u) h(u)\, du \\ &= \int_0^t \exp(-H(u)) h(u)\, du, \end{aligned} \tag{3.1.02}$$

where we have used Equation 3.1.01 in the second equality. Equation 3.1.02 suggests that F can be specified in variously equivalent ways. The last representation in the equation (3.1.02) indicates that the hazard function completely determines the distribution function.

Given F, the τ-quantile of the distribution of T , $Q(\tau)$, is defined as

$$Q(\tau) = F^{-1}(\tau), \quad \tau \in (0, 1);$$

and the τ-quantile residual or remaining life time at time t_0, $Q_{t_0}(\tau)$, is defined as

$$Q_{t_0}(\tau) = F_{t_0}^{-1}(\tau), \quad \tau \in (0, 1);$$

where $F_{t_0}(t)$ denotes the distribution function of $T - t_0 | T > t_0$. The τ-quantile residual life at time t_0, $Q_{t_0}(\tau)$, can be written in terms of $Q(\tau)$. Indeed, by the definition of τ-quantile residual life time ,

$$\begin{aligned} F_{t_0}(t) &= P(T - t_0 \leq t \mid T \geq t_0) \\ &= \frac{P(t_0 < T \leq t + t_0)}{P(T > t_0)}. \end{aligned}$$

This implies

$$F_{t_0}(t) = \frac{F(t + t_0) - F(t_0)}{1 - F(t_0)}, \tag{3.1.03}$$

for $t \geq 0$. Thus,

$$Q_{t_0}(\tau) = F^{-1}(\tau(1 - F(t_0)) + F(t_0)) - t_0 \tag{3.1.04}$$

where $0 < \tau < 1$. Equation 3.1.04 is equivalent to:

$$Q_{t_0}(\tau) = Q(\tau(1 - F(t_0)) + F(t_0)) - t_0 \tag{3.1.05}$$

Parametric inference on Q_{t_0} (Equations 3.1.04) can be undertaken by making a parametric assumption about $F_{t_0}(t)$ or on $F(t)$ (or equivalently on $h(t)$). The latter approach is used as it provides a more flexible generalization to the competing risks setting. The dependency of F on the unknown parameter vector θ can be made more explicit by writing it as $F(t; \theta)$. In the regression setting, we can write it as $F(t; \theta, Z)$ to indicate its dependency on θ and the baseline covariate vector Z as well. However, for notational convenience, we will sometimes suppress the dependency of the survival quantities on their arguments, if no confusion may arise.

Finally, we will assume that regularity conditions needed to make the asymptotics work hold. In particular, we will assume that $Q_{t_0}(\tau)$ is differentiable with respect to θ.

3.2 ONE SAMPLE

In this section, a parametric inference on the true residual quantile life time based on a random sample of right-censored data from a homogeneous population is considered. Section 3.2.1 presents the basic results, Section 3.2.2 considers application of the results to the Weibull model, Section 3.2.3 presents simulation results using the Weibull model, and Section 3.2.4 applies the results to real data set.

3.2.1 Theory

Let $\{T_i\}_{i=1}^n$ be i.i.d failure times with distribution function $F(.;\theta)$ assumed to be known up to an unknown parameter $\theta = (\theta_1, \theta_2, \ldots, \theta_p)$, and $\{C_i\}_{i=1}^n$ be i.i.d censoring times with distribution function $G(.)$. G is assumed not have any parameter in common with F. Thus, we are making a parametric assumption about F. To ensure random censorship, we assume the failure times and censoring times are mutually independent. We only observe $\{(X_i, \delta_i)\}_{i=1}^n$, where

$$X_i = \min\{T_i, C_i\} \quad \text{and} \quad \delta_i = I(T_i \leq C_i).$$

These are right-censored data, the most common form of censored data in clinical trial and biomedical studies (Lagakos, 1979). The goal is to make inference on Q_{t_0}, the true quantile residual life time at time t_0,

$$Q_{t_0}(\tau;\theta) = F^{-1}(\tau(1 - F(t_0;\theta)) + F(t_0;\theta)) - t_0 \tag{3.2.11}$$

using the observed data. In particular, we are interested in point and large sample interval estimates of Q_{t_0}.

While inference on Q_{t_0} can be made in other ways, likelihood based inference provides asymptotically optimal estimates when the assumed model is correct (Boos and Stefanski, 2013). The delta method, along with the asymptotic property of the MLE, are used to derive likelihood-based point and confidence interval estimators.

Under random censorship, the contribution of $(X_i, \delta_i = 1)$ to the likelihood is $f(X_i;\theta)$ and that of $(X_i, \delta_i = 0)$ is $S(X_i;\theta)$. Thus, the likelihood function for the observed data is:

$$\begin{aligned} L(\theta \mid \{(X_i,\delta_i)\}_{i=1}^n) &= \prod_{i=1}^n f(X_i;\theta)^{\delta_i} S(X_i;\theta)^{1-\delta_i} \\ &= \prod_{i=1}^n h(X_i;\theta)^{\delta_i} S(X_i;\theta) \quad \text{(by Equation 3.1.01)}. \end{aligned}$$

The log-likelihood in terms of the latter form of the likelihood has the form:

$$\ell_n(\theta) = \sum_{i=1}^n \delta_i \ln h(X_i;\theta) + \sum_{i=1}^n \ln S(X_i;\theta).$$

The score function,

$$U(\theta) = \frac{\partial}{\partial\theta}\ell_n(\theta) \tag{3.2.12}$$

contains the elements:

$$\frac{\partial}{\partial\theta_j}\ell_n(\theta) = \sum_{i=1}^n \left\{ \delta_i \frac{1}{h(X_i;\theta)} \frac{\partial}{\partial\theta_j} h(X_i;\theta) + \frac{1}{S(X_i;\theta)} \frac{\partial}{\partial\theta_j} S(X_i;\theta) \right\} \quad j = 1, \cdots, p.$$

A maximum likelihood estimator (MLE) of θ, $\hat{\theta}$, is any solution of $U(\theta) = 0$. Under regularity conditions, the MLE is unique and consistent (Borgan, 1984).

Since $Q_{t_0}(\tau;\theta)$ is a differentiable function of θ, a consistent estimator of it is,

$$\hat{Q}_{t_0}(\tau;\theta) = Q_{t_0}(\tau,\hat{\theta}) = F^{-1}(\tau(1 - F(t_0;\hat{\theta})) + F(t_0;\hat{\theta})) - t_0. \tag{3.2.13}$$

The asymptotic distribution of $\hat{Q}_{t_0}(\tau;\theta)$ follows by the multivariate-version of the delta method. The application of the delta method requires that $\hat{\theta}$ be consistent for θ and asymptotically normally distributed. By the asymptotic properties of the MLEs (Borgan, 1984), as $n \longrightarrow \infty$,

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N(0, \mathcal{I}^{-1}(\theta)) \tag{3.2.14}$$

where $\mathcal{I}^{-1}(\theta)$ is the inverse of the expected information matrix,

$$\mathcal{I}(\theta) = -E\left\{ \frac{1}{n} \frac{\partial^2}{\partial\theta\partial\theta^T} \ell_n(\theta) \right\}.$$

The expected information matrix, even under random censorship, depends on the unknown censoring distribution and could be difficult or impossible to calculate even when the censoring

distribution is known (Kalbfleisch and Prentice, 2002). As a result, in applications the expected information matrix is replaced by its consistent estimator, the sample information matrix,

$$I(\theta) = -\frac{\partial^2}{\partial\theta\partial\theta^T}\ell_n(\theta).$$

A consistent estimator of $\mathcal{I}^{-1}(\theta)$ is $nI^{-1}(\hat{\theta})$, where $I^{-1}(\hat{\theta})$ is the inverse of the sample information matrix evaluated at the MLE of θ. Now, applying the delta method yields the asymptotic distribution of $\hat{Q}_{t_0}$,

$$\sqrt{n}(\hat{Q}_{t_0} - Q_{t_0}) \xrightarrow{d} N(0, nAvar(\hat{Q}_{t_0})) \tag{3.2.15}$$

where

$$Avar(\hat{Q}_{t_0}) = \frac{1}{n}\left\{\nabla_{\hat{\theta}} Q_{t_0}(\tau;\hat{\theta})_{|\hat{\theta}=\theta}\right\}^T \mathcal{I}^{-1}(\theta)\left\{\nabla_{\hat{\theta}} Q_{t_0}(\tau;\hat{\theta})_{|\hat{\theta}=\theta}\right\}. \tag{3.2.16}$$

For inference (testing and confidence intervals) on the quantile residual life time, the unknown quantities in Equation 3.2.16 are estimated by their consistent estimators. Hence, a consistent estimator for the asymptotic variance of the quantile residual life estimator is:

$$\widehat{Avar(\hat{Q}_{t_0})} = \left\{\nabla_{\hat{\theta}} Q_{t_0}(\tau;\hat{\theta})_{|\hat{\theta}=\hat{\theta}}\right\}^T I^{-1}(\hat{\theta})\left\{\nabla_{\hat{\theta}} Q_{t_0}(\tau;\hat{\theta})_{|\hat{\theta}=\hat{\theta}}\right\}. \tag{3.2.17}$$

Using the point estimate $\hat{Q}_{t_0}$ and the estimator of its asymptotic variance, Wald-type approximate tests and confidence intervals (CIs) can be constructed in the usual way. An approximate two-sided $100(1-\alpha)\%$ Wald confidence interval(CI) for $Q_{t_0}(\tau;\theta)$ is:

$$\left(\hat{Q}_{t_0}(\tau;\theta) \pm z_{1-\alpha/2}\sqrt{\widehat{Avar(\hat{Q}_{t_0})}}\right). \tag{3.2.18}$$

In the subsequent sections, we will consider the application of the above result to the the Weibull distribution.

3.2.2 Application to the Weibull Model

We will use the two-parameter Weibull model for simulations and data analyses throughout this dissertation. Some of its basic properties are briefly reviewed here. There are different equivalent parametrization of the two-parameter Weibull. We will use the following parametrization of it:

$$F(t) = 1 - \exp(-\lambda t^{\alpha}), \tag{3.2.21}$$

with shape parameter $\alpha > 0$ and scale parameter $\lambda > 0$. Its survival function is given by

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

and its hazard function is given by,

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

with $\alpha > 1, \alpha < 1$, and, $\alpha = 1$ corresponding to increasing, decreasing and constant hazard, respectively. This flexibility, along with the simplicity of the Weibull model, makes it a popular model for parametric modeling. The Weibull distribution is the only parametric distribution that has both the accelerated failure time and proportional hazard representations (Klein and Moeschberger, 2003).

The likelihood of right-censored data $\{(X_i, \delta_i)\}_{i=1}^n$ based on Weibull failure times under random censorship is:

$$L(\alpha, \lambda | \{(X_i, \delta_i)\}_{i=1}^n) = \exp\left(-\lambda \sum_{i=1}^n X_i\right) \lambda^{\sum_{i=1}^n \delta_i}.$$

Thus, its log-likelihood is of the form:

$$\ell_n(\alpha, \lambda) = (\ln \alpha + \ln \lambda) \sum_i^n \delta_i + (\alpha - 1) \sum_{i=1}^n \delta_i \ln(X_i) - \lambda \sum_{i=1}^n X_i^{\alpha}.$$

The score vector, $U(\alpha, \lambda)$, has the components:

$$U_\alpha(\alpha,\lambda) = \frac{\partial}{\partial\alpha}\ell_n(\alpha,\lambda) = \frac{\sum_{i=1}^n \delta_i}{\alpha} + \sum_i^n \delta_i \ln(X_i) - \lambda\sum_{i=1}^n X_i^\alpha \ln(X_i)$$

$$U_\lambda(\alpha,\lambda) = \frac{\partial}{\partial\lambda}\ell_n(\alpha,\lambda) = \frac{\sum_{i=1}^n \delta_i}{\lambda} - \sum_{i=1}^n X_i^\alpha \ln(X_i).$$

The maximum likelihood estimator (MLE) of α and λ is the solution of the score vector equation: $U(\alpha,\lambda) = (U_\alpha(\alpha,\lambda), U_\lambda(\alpha,\lambda)) = 0$. Thus, the MLE of α, $\hat{\alpha}$, is a solution to the following non-linear equation:

$$\sum_{i=1}^n \delta_i + \alpha\sum_{i=1}^n \delta_i \ln(X_i) - \alpha\left(\sum_{i=1}^n \delta_i\right)\left(\sum_i^n X_i^\alpha\right)^{-1}\sum_{i=1}^n X_i^\alpha \ln(X_i) = 0.$$

The MLE of λ can be expressed in terms of the MLE of α as:

$$\hat{\lambda} = \frac{\sum_{i=1}^n \delta_i}{\sum_{i=1}^n X_i^{\hat{\alpha}}}.$$

The sample information matrix contains the following entries:

$$I(\alpha,\lambda) = \begin{bmatrix} \alpha^2\sum_{i=1}^n \delta_i + \lambda\sum_{i=1}^n X_i^\alpha(\ln(X_i))^2 & \sum_{i=1}^n X_i^\alpha \ln(X_i) \\ \sum_{i=1}^n X_i^\alpha \ln(X_i) & \frac{\sum_{i=1}^n \delta_i}{\lambda^2} \end{bmatrix}.$$

The MLE of the quantile residual life time function for the Weibull distribution at t_0, $\hat{Q}_{t_0}$, is:

$$\left[t_0^{\hat{\alpha}} - \hat{\lambda}^{-1}\ln(1-\tau)\right]^{1/\hat{\alpha}} - t_0. \tag{3.2.22}$$

The gradient vector of the MLE of the quantile residual life function,$\nabla\hat{Q}_{t_0}$, contains the following components,

$$\frac{\partial}{\partial\hat{\alpha}}Q_{t_0}(\tau;\hat{\alpha},\hat{\lambda}) = \hat{\alpha}^{-2}(t_0^{\hat{\alpha}} - \hat{\lambda}^{-1}\ln(1-\tau))^{1/\hat{\alpha}}\left\{\frac{\hat{\alpha}t_0^{\hat{\alpha}}\ln(t_0)}{[t_0^{\hat{\alpha}} - \hat{\lambda}\ln(1-\tau)]} - \ln[t_0^{\hat{\alpha}} - \hat{\lambda}\ln(1-\tau)]\right\}$$
$$\frac{\partial}{\partial\hat{\lambda}}Q_{t_0}(\tau;\hat{\alpha},\hat{\lambda}) = \hat{\alpha}^{-1}\hat{\lambda}^{-2}[t_0^{\hat{\alpha}} - \hat{\lambda}^{-1}\ln(1-\tau)]^{-1+1/\hat{\alpha}}\ln(1-\tau),$$

where we have used the convention $t_0^{\hat{\alpha}}\ln(t_0) = 0$ for $t_0 = 0$ in the first equation.

The estimated asymptotic variance of the MLE of the quantile residual life time function has the form,

$$\widehat{Avar(\hat{Q}_{t_0})} = n\left\{\nabla\hat{Q}_{t_0}(\tau;\hat{\alpha},\hat{\lambda})\right\}^T I(\hat{\alpha},\hat{\lambda})^{-1}\left\{\nabla\hat{Q}_{t_0}(\tau;\hat{\alpha},\hat{\lambda})\right\}. \tag{3.2.23}$$

By specializing the above discussion to the $Weibull(1,\lambda)$, which is the exponential distribution with scale parameter, λ, we get

$$\hat{Q}_{t_0}(\tau) = \left[\frac{-\ln(1-(\tau(1-\exp(-\hat{\lambda}t_0))+1-\exp(-\hat{\lambda}t_0)))}{\hat{\lambda}}\right] - t_0$$

which simplifies to:

$$\frac{-\ln(1-\tau)}{\hat{\lambda}}.$$

The above expression is free of t_0. Thus, the quantile and quantile residual life functions of the exponential distribution are identical. This is due to the memory-less property of the exponential distribution. Plugging 1 for $\hat{\alpha}$ and $\frac{\sum_{i=1}^n \delta_i}{\sum_{i=1}^n X_i}$ for $\hat{\lambda}$ in Equation 3.2.23, we obtain the asymptotic variance of the quantile residual life estimator for the exponential distribution,

$$\begin{aligned}\widehat{Avar(\hat{Q}_{t_0})} &= \left[\frac{\ln(1-\tau)}{\hat{\lambda}^2}\right]\frac{\hat{\lambda}^2}{\sum_{i=1}^n \delta_i}\left[\frac{\ln(1-\tau)}{\hat{\lambda}^2}\right] \\ &= \left(\ln(1-\tau)\sum_{i=1}^n X_i\right)^2\left(\sum_{i=1}^n \delta_i\right)^{-3},\end{aligned}$$

which is again free of t_0.

3.2.3 Simulation Study

We conducted a series of simulation studies to examine the performance of the proposed method. The censoring proportion was set at 0.1. To simulate different failure time scenarios, we generated failure times from two Weibull distributions: $Weibull(2, 0.5)$ and $Weibull(2, 0.25)$, where $Weibull(\alpha, \lambda)$ is the Weibull distribution with shape parameter α and scale parameter λ. Censoring times were generated, respectively, from $Unif(0, 5.5048)$,and $Unif(0, 12.5331)$ to get the desired level of censoring. For each failure time scenario, we simulated $N = 5000$ data sets of sample sizes $n = 400$ and $n = 100$. Two quantile levels $\tau = 0.25$ and $\tau = 0.5$, and two time points $t_0 = 0$ and $t_0 = 1$ were considered. All simulations and data analyses in this section and throughout this dissertation were performed using R statistical software.

Simulation results for the one-sample case are presented in Tables 3.2.1-3.2.2. In each of the tables, the estimated quantile residual life time ($\hat{Q}_{t_0}$), estimated asymptotic variance ($\widehat{Avar(\hat{Q}_{t_0})}$), empirical variance ($\widehat{var(\hat{Q}_{t_0})}$), empirical bias ($\widehat{Bias}$), empirical MSE ($\widehat{MSE}$), and 95% empirical coverage probability (CP) are shown.

The Monte Carlo variance estimates and asymptotic variance estimates are similar across all the simulation settings, suggesting that the asymptotic variance formula works well. Coverage probability appears to be at the nominal-level. The quantile residual life time estimates are unbiased and the empirical MSEs are small.

Table 3.2.1: One-sample simulation result, Weibull($\alpha = 2$, $\lambda = 0.5$)

n	t_0	τ	$\widehat{Avar(\hat{Q}_{t_0})}$	$\widehat{var(\hat{Q}_{t_0})}$	$\widehat{Bias}$	$\widehat{MSE}$	CP
400	0	0.25	0.001	0.001	0.001	0.001	0.952
		0.50	0.001	0.001	0.001	0.001	0.952
	1	0.25	0.000	0.000	0.000	0.000	0.944
		0.50	0.001	0.001	-0.000	0.001	0.945
100	0	0.25	0.004	0.004	0.004	0.004	0.945
		0.50	0.005	0.005	0.001	0.005	0.943
	1	0.25	0.001	0.001	-0.000	0.001	0.944
		0.50	0.003	0.002	-0.002	0.002	0.941

Table 3.2.2: One-sample simulation result, Weibull($\alpha = 2$, $\lambda = 0.25$)

n	t_0	τ	$\widehat{Avar(\hat{Q}_{t_0})}$	$\widehat{var(\hat{Q}_{t_0})}$	$\widehat{Bias}$	$\widehat{MSE}$	CP
400	0	0.25	0.002	0.002	0.001	0.002	0.948
		0.50	0.003	0.003	0.000	0.003	0.949
	1	0.25	0.000	0.000	0.000	0.000	0.950
		0.50	0.001	0.001	-0.000	0.001	0.948
100	0	0.25	0.002	0.002	0.002	0.002	0.949
		0.50	0.003	0.003	0.001	0.003	0.949
	1	0.25	0.000	0.000	0.001	0.000	0.947
		0.50	0.001	0.001	0.001	0.001	0.946

3.2.4 Data Analysis

Breast cancer survival data from the B-04 randomized trial run by the National Surgical Adjuvant Breast and Bowel Project (NSABP) will be used to apply the proposed methods to real data. This study was initiated in 1971 to investigate the effectiveness on survival of three breast cancer surgery procedures: radical mastectomy, total mastectomy without radiation therapy, and total mastectomy with radiation therapy. Participants in the study were 1765 women with primary operable breast cancer. At the beginning of the study, the patients were assessed for the presence or absence of clinical nodal involvement in their cancer and were, accordingly, stratified as clinically node positive or negative patients. 1079 of the patients were found to be clinically node positive and 586 were found to be clinically node negative. The radical and less extensive surgeries were compared within each nodal group. A major result of the study reported in Fisher et al. (1977) and confirmed again in a later analysis reported in Fisher et al. (2002) was that there was no difference among the three treatments in their effects on various survival endpoints including overall survival.

For the data analyses in this chapter, we will use death from all causes (breast cancer or other causes) as the survival outcome for the non-competing risks analysis. For the one-sample analysis, we will perform separate quantile residual life analysis for the two nodal groups. For the two-sample analysis, we will compare the quantile residual life of the two groups. For the regression analysis, we will perform quantile residual life analysis using age (year), nodal status and tumor size (mm) as covariates.

Summary statistics on the analysis variables considered in this chapter are presented in Table 3.2.3. Due to missingness in the variable node type indicator variable, 66 subjects were excluded from the analysis. Node positive patients have higher mortality than node negative patients. Node positive patients tend to die earlier than node negative patients. The observed censoring proportion is about 0.23. The median age at baseline is 55 years for node positive patients and 57 years for node negative patients. Node positive patients have higher median tumor size than node-negative patients.

For the one-sample residual life analysis, we fitted separate Weibull models to the survival data for the node negative and node-positive patient groups. The quantile residual life time

Table 3.2.3: Summary statistics on mortality, observed time, age, tumor size by node status

Node	n	Death	Median Time	Median Age	Median Tumor Size
Positive	561	0.84	9.58	55.00	32.00
Negative	1038	0.74	12.72	57.00	30.00
All	1599	0.77	11.62	56.00	30.00

for each group was estimated at four time points $t_0 = 0, 2, 4, 6$ and three quantile levels $\tau = 0.1, 0.2, 0.5$. The results for these analysis are shown in Table 3.2.4 for the node-positive patients and in Table 3.2.5 for the node-negative patients. Each table presents point estimates of the quantile residual life, estimate of the asymptotic variance of the quantile residual life estimate, and 95% confidence interval estimate (CI_{LL} for lower limit and CI_{UL} for upper limit) of the quantile residual life at four time points $t_0 = 0, 2, 4, 6$ and three quantile levels $\tau = 0.1, 0.2, 0.5$.

The results in both tables indicate that quantile residual life times do not seem to vary much with time. The MLE for the fitted Weibull model for the node-positive group was $(\hat{\alpha} = 0.93, \hat{\lambda} = 0.11)$; and for that of the node-negative group was $(\hat{\alpha} = 1.07, \hat{\lambda} = 0.047)$. The $\hat{\alpha}$ s for both estimated models are close to 1, suggesting that the fitted model might not be "far" from the exponential model and this might explain the lack of appreciable dependency of quantile residual life estimates on time points from both models. The median residual life estimates (for $\tau = 0.5$) in Table 3.2.4 and Table 3.2.5 are within reasonable range of the non-parameteric median residual life estimates reported in Table 3 of Jeong et al. (2008).

The plausibility of the fitted Weibull models was checked by comparing them with Kaplan-Meier(KM) curves. Quantile-comparisons for the fitted and KM models are shown in Figure 3.2.1 for the positive patient group and Figure 3.2.2 for the negative patient group. It appears that the assumed Weibull models do not show significant lack of fit.

Table 3.2.4: Quantile residual life analysis for node positive patients

t_0	τ	$\hat{Q}_{t_0}$	$\widehat{Avar}(\hat{Q}_{t_0})$	CI_{LL}	CI_{UL}
0	0.1	0.992	0.012	0.777	1.208
	0.2	2.234	0.035	1.865	2.602
	0.5	7.604	0.170	6.796	8.413
2	0.1	1.151	0.003	1.036	1.267
	0.2	2.477	0.015	2.239	2.716
	0.5	8.019	0.153	7.253	8.786
4	0.1	1.203	0.003	1.093	1.313
	0.2	2.573	0.015	2.335	2.811
	0.5	8.229	0.180	7.397	9.060
6	0.1	1.236	0.004	1.119	1.354
	0.2	2.637	0.017	2.380	2.895
	0.5	8.382	0.214	7.474	9.289

Table 3.2.5: Quantile residual life analysis for node negative patients

t_0	τ	$\hat{Q}_{t_0}$	$\widehat{Avar(\hat{Q}_{t_0})}$	CI_{LL}	CI_{UL}
0	0.1	2.099	0.023	1.801	2.397
	0.2	4.231	0.053	3.781	4.681
	0.5	12.194	0.183	11.355	13.034
2	0.1	1.916	0.007	1.750	2.081
	0.2	3.980	0.024	3.675	4.284
	0.5	11.831	0.167	11.029	12.632
4	0.1	1.847	0.005	1.713	1.980
	0.2	3.863	0.019	3.594	4.132
	0.5	11.616	0.180	10.785	12.446
6	0.1	1.803	0.004	1.676	1.929
	0.2	3.783	0.019	3.517	4.050
	0.5	11.453	0.200	10.575	12.331

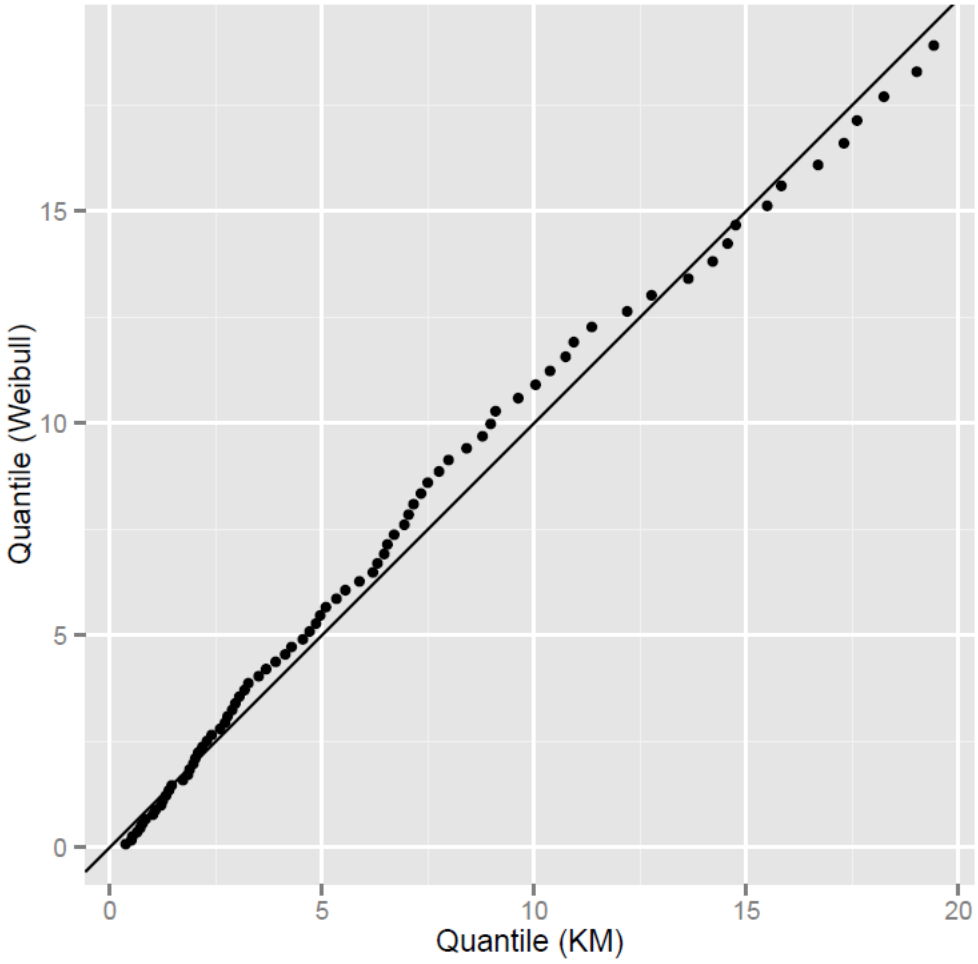

Figure 3.2.1: Kaplan Meier Quantile vs. Weibull fitted Quantile plot (node positive patients)

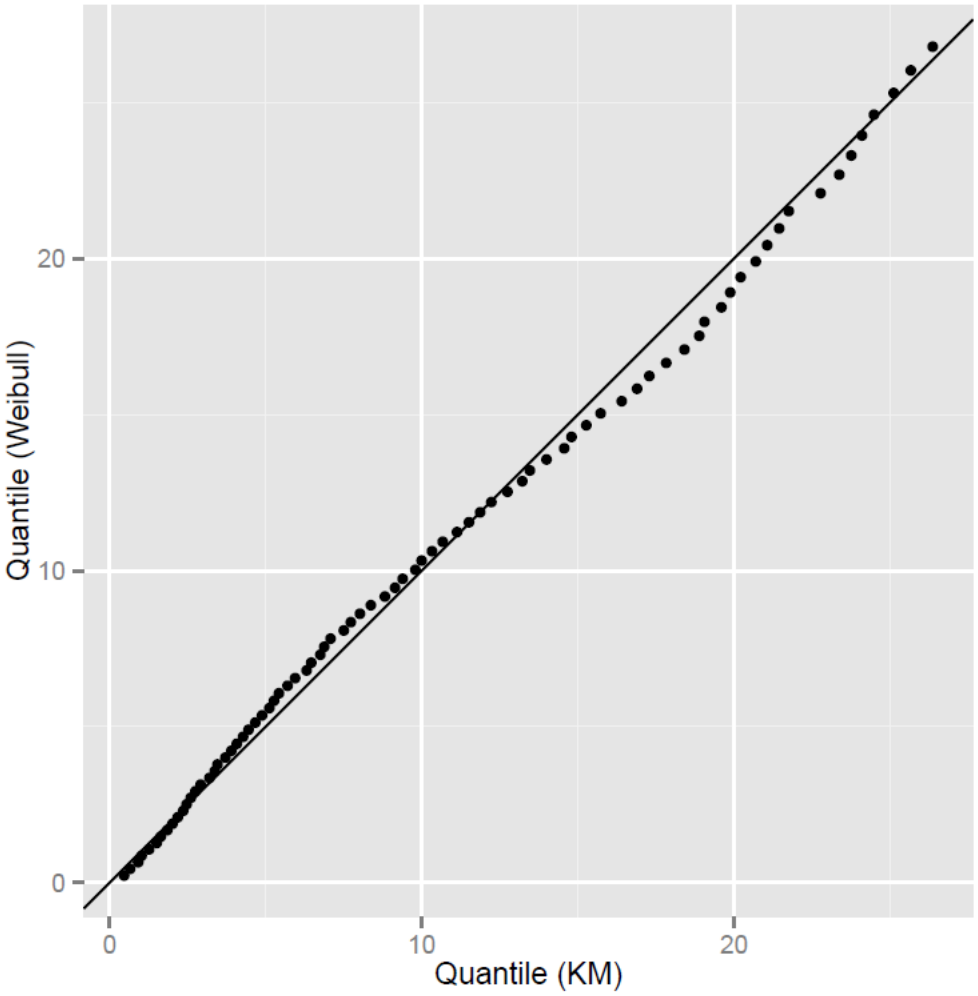

Figure 3.2.2: Kaplan Meier Quantile vs. Weibull fitted Quantile plot (node negative patients)

3.3 TWO SAMPLE

In this section, we extend the one-sample results in Section 3.2 to two independent groups. The results obtained in this section can be easily generalized to more than two groups. Other than the number of groups, the setting is similar to one-sample case. In Section 3.3.1, basic results are presented. Section 3.3.2 presents simulation results using the Weibull model, and Section 3.3.3 applies the results to real data set.

3.3.1 Theory

We have a random sample of right censored data from each of the two populations. We assume that the failure times in the two groups come from the same parametric family. Hence, the random sample of failure times $T_{1i}, i = 1, \ldots, n_1$ for group 1 comes from $F(;\theta^{(1)})$. We also have corresponding random sample of censoring times $C_{1i} \sim G_1(.), i = 1, \ldots, n_1$. Similarly, the failure times $T_{2i}, i = 1, \ldots, n_2$ for group 2 comes from $F(;\theta^{(2)})$ with a corresponding random sample of censoring times from $C_{2i} \sim G_2(.), i = 1, \ldots, n_2$. No assumption on the functional dependency between $\theta^{(1)}$ and $\theta^{(2)}$ is made. We observe $\{(X_{ji}, \delta_{ji})\}_{i=1}^{n_j}$, $j = 1, 2$, where

$$X_{ji} = \min\{T_{ji}, C_{ji}\} \quad \text{and} \quad \delta_{ji} = I(T_{ji} \leq C_{ji}).$$

Using the observed data, we want to make inferences about the difference between the true quantile residual residual life times at time t_0,

$$Q_{t_0}^{(1)}(\tau;\theta^{(1)}) - Q_{t_0}^{(2)}(\tau;\theta^{(2)}) \tag{3.3.11}$$

Likelihood based point and confidence interval estimators are derived below. The likelihood for the observed data is:

$$\begin{aligned} L(\theta \mid \{(X_{ji}, \delta_{ji})\}_{i=1}^{n}) &= \prod_{j=1}^{2}\prod_{i=1}^{n_j} f(X_{ji};\theta)^{\delta_{ji}} S(X_{ji};\theta)^{1-\delta_{ji}} \\ &= \prod_{j=1}^{2}\prod_{i=1}^{n_j} h(X_{ji};\theta)^{\delta_{ji}} S(X_{ji};\theta)\,, \end{aligned}$$

where $\theta = \left(\theta^{(1)}, \theta^{(2)}\right)$ is a vector of parameter vectors $\theta^{(1)}$ and $\theta^{(2)}$ for the distributions of T_{1i} and T_{2i}, respectively. Thus, the log-likelihood is given by

$$\ell_n(\theta) = \sum_{i=1}^{n_1} \delta_{1i} \ln h(X_{1i}; \theta^{(1)}) + \ln S(X_{1i}; \theta^{(1)}) + \sum_{i=1}^{n_2} \delta_{2i} \ln h(X_{2i}; \theta^{(2)}) + \ln S(X_{2i}; \theta^{(2)}).$$

As a result of the assumed independence between the two groups, the joint log-likelihood separates into the log-likelihood for each group. Thus, likelihood quantities and inference follow directly from the one-sample case.

By the result in Section 3.2, and the invariance property of the MLEs, the MLE of $Q_{t_0}^{(1)}\left(\tau; \theta^{(1)}\right) - Q_{t_0}^{(2)}\left(\tau; \theta^{(2)}\right)$ is:

$$Q_{t_0}^{(1)}\left(\tau; \hat{\theta}^{(1)}\right) - Q_{t_0}^{(2)}\left(\tau; \hat{\theta}^{(2)}\right) \tag{3.3.12}$$

where $\hat{\theta}^{(1)}$ and $\hat{\theta}^{(2)}$ are MLEs of $\theta^{(1)}$ and $\theta^{(2)}$, respectively.

The asymptotic distribution of $Q_{t_0}^{(1)}\left(\tau; \hat{\theta}^{(1)}\right) - Q_{t_0}^{(2)}\left(\tau; \hat{\theta}^{(2)}\right)$ also follows by the delta method. By the result in Section 3.2, $\hat{Q}_{t_0}^{(j)}$ is consistent for $Q_{t_0}^{(j)}$, $j = 1, 2$, and asymptotically normally distributed. Now, applying the delta method on the function $g(\hat{Q}_{t_0}^{(1)}, \hat{Q}_{t_0}^{(2)}) = \hat{Q}_{t_0}^{(1)} - \hat{Q}_{t_0}^{(2)}$, we get

$$\hat{Q}_{t_0}^{(1)} - \hat{Q}_{t_0}^{(2)} \sim AN\left(Q_{t_0}^{(1)} - Q_{t_0}^{(2)}, Avar\left(\hat{Q}_{t_0}^{(1)}\right) + Avar\left(\hat{Q}_{t_0}^{(2)}\right)\right) \tag{3.3.13}$$

where $Avar(\hat{Q}_{t_0}^{(j)})$, $j = 1, 2$, is as in Equation 3.2.16; and its consistent estimator, $\widehat{Avar(\hat{Q}_{t_0}^{(j)})}$, is as in Equation 3.2.17, $j = 1, 2$.

Using the point estimate $\hat{Q}_{t_0}^{(1)} - \hat{Q}_{t_0}^{(2)}$ and the estimator of its asymptotic variance, Wald-type approximate tests and CIs can be constructed. An approximate two-sided $100(1-\alpha)\%$ Wald confidence interval(CI) for $Q_{t_0}^{(1)} - Q_{t_0}^{(2)}$ is:

$$\left(\hat{Q}_{t_0}^{(1)} - \hat{Q}_{t_0}^{(2)} \pm z_{1-\alpha/2}\sqrt{\widehat{Avar(\hat{Q}_{t_0}^{(1)})} + \widehat{Avar(\hat{Q}_{t_0}^{(2)})}}\right). \tag{3.3.14}$$

3.3.2 Simulation Study

We ran simulation studies to investigate the performance of the proposed method. The simulation were conducted as in the one-sample case except for the following differences: two samples were independently generated from the same population at each simulation run; sample sizes of $n = 100$ and $n = 200$ per sample were used. Simulation results for the two-sample setting are presented in Tables 3.3.1-3.3.2.

In each of the tables, asymptotic variance, Monte Carlo variance, empirical bias, and empirical MSE for the estimated difference in quantile residual lifetime $\left(\Delta\hat{Q}_{t_0} = \hat{Q}_{1,t_0} - \hat{Q}_{2,t_0}\right)$, and 95% empirical coverage (CP) for the true difference in quantile residual life time are shown.

The simulation results indicate that the method works well under all the simulation scenarios considered. The asymptotic variance and the Monte Carlo variance are similar to each other indicating that the model-based asymptotic variance formula performs well. The quantile residual lifetime estimator is unbiased and the empirical coverage probability is close to the nominal value.

Table 3.3.1: Two-sample simulation result, Weibull($\alpha_1 = \alpha_2 = 2$, $\lambda_1 = \lambda_2 = 0.5$)

n	t_0	τ	$\widehat{Avar(\Delta\hat{Q}_{t_0})}$	$\widehat{var(\Delta\hat{Q}_{t_0})}$	$\widehat{Bias}$	$\widehat{MSE}$	CP
200	0	0.25	0.004	0.004	-0.001	0.004	0.948
		0.50	0.005	0.005	-0.002	0.005	0.953
	1	0.25	0.001	0.001	-0.001	0.001	0.955
		0.50	0.002	0.002	-0.002	0.002	0.950
100	0	0.25	0.008	0.008	-0.000	0.008	0.944
		0.50	0.010	0.010	0.000	0.010	0.947
	1	0.25	0.001	0.001	0.000	0.001	0.955
		0.50	0.005	0.005	0.001	0.005	0.954

Table 3.3.2: Two-sample simulation result, Weibull($\alpha_1 = \alpha_2 = 2$, $\lambda_1 = \lambda_2 = 0.25$)

n	t_0	τ	$\widehat{Avar(\Delta\hat{Q}_{t_0})}$	$\widehat{var(\Delta\hat{Q}_{t_0})}$	$\widehat{Bias}$	$\widehat{MSE}$	CP
200	0	0.25	0.008	0.008	-0.003	0.008	0.951
		0.50	0.010	0.010	-0.003	0.010	0.955
	1	0.25	0.002	0.002	-0.001	0.002	0.956
		0.50	0.006	0.005	-0.002	0.005	0.954
100	0	0.25	0.016	0.016	0.001	0.016	0.948
		0.50	0.020	0.020	0.002	0.020	0.950
	1	0.25	0.004	0.004	0.001	0.004	0.953
		0.50	0.011	0.011	0.001	0.011	0.946

3.3.3 Data Analysis

The analyses started in Section 4.2.5 is carried on here, where we apply the proposed method for the two-sample setting on the B04-study data using nodal status as a grouping variable. The difference in the quantile residual life time between the two nodal groups are estimated at four time points $t_0 = 0, 2, 4, 6$ and three quantile levels $\tau = 0.1, 0.2, 0.5$. The results are shown in Table 3.3.3. It is clearly seen that the nodal positive group has a shorter remaining life time than the nodal negative group. This is consistent with the one-sample result presented in Table 3.2.3.

Table 3.3.3: Quantile residual life time difference between node-negative and node-positive patients

t_0	τ	$\Delta\hat{Q}_{t_0}$	$\widehat{Avar(\Delta\hat{Q}_{t_0})}$	CI_{LL}	CI_{UL}
0	0.1	1.107	0.035	0.739	1.474
	0.2	1.997	0.088	1.416	2.579
	0.5	4.590	0.354	3.425	5.755
2	0.1	0.764	0.011	0.563	0.966
	0.2	1.502	0.039	1.116	1.889
	0.5	3.811	0.320	2.702	4.921
4	0.1	0.644	0.008	0.470	0.817
	0.2	1.290	0.034	0.930	1.649
	0.5	3.387	0.360	2.212	4.563
6	0.1	0.566	0.008	0.394	0.739
	0.2	1.146	0.036	0.775	1.517
	0.5	3.071	0.415	1.809	4.334

3.4 REGRESSION - ACCELERATED FAILURE TIME MODEL

So far, one sample and two sample settings in which the data in each sample are assumed to be a random sample from an underlying distribution (homogeneous population) were considered. The one-sample and two-sample procedures are valid for homogeneous populations.

In many applications, such as clinical trials and observational studies, there is heterogeneity among the study subjects that affects or is associated with failure time (Klein and Moeschberger, 2003; Prentice et al., 1978). Thus, it is important to develop a procedure that accommodates variability in the characteristics of individuals in a population. In this section, the one-sample and the independent two-sample quantile residual life procedures developed in previous sections are generalized to the regression case. There are two standard regression approaches for modeling the effect of covariates on survival times: the proportional hazard (PH) and the accelerated failure time (AFT) models.

Under the PH assumption, the effect of covariates is to scale up (down) the baseline hazard function. Under the AFT approach, the effect of covariates is to scale up (down) the baseline time-scale. This makes the AFT model a natural model for quantile inference. The accelerated failure time approach provides an alternative approach to the PH approach (Wei, 1992); includes some PH models; may be more appropriate in some circumstances (Patel et al., 2006; Kay and Kinnersley, 2002). We will use the AFT model to incorporate covariate information.

3.4.1 Theory

Suppose on the log scale the failure time T can be written as a linear model

$$\ln T = \mathbf{Z}^T\beta + \sigma\varepsilon \tag{3.4.11}$$

where $\varepsilon \sim f_\varepsilon(e)$. This log-linear model contains two types of parameters: β, the regression coefficient and b, baseline parameter. The baseline parameter contains all parameters in the baseline distribution, i.e., the distribution of T when $\mathbf{Z} = 0$. In the above model, σ is part of

the baseline parameter b. From equation 3.4.11, the survival function of T, $S(x;\mathbf{Z})$ is given by

$$\int_x^\infty f_\varepsilon\left(\frac{\ln T-\mathbf{Z}^T\beta}{\sigma}\right)\frac{1}{T\sigma}\,dT.$$

Making the change of variable $\varepsilon=\frac{\ln T-\mathbf{Z}^T\beta}{\sigma}$, we obtain

$$\begin{aligned}S(x;\mathbf{Z}) &= \int_{\frac{\ln x-\mathbf{Z}^T\beta}{\sigma}}^\infty f_\varepsilon(\varepsilon)\,d\varepsilon\\ &= S_\varepsilon\left(\sigma^{-1}\ln x\exp(-\mathbf{Z}^T\beta);\mathbf{Z}\right).\end{aligned}$$

Denoting the survival function of T for the population with baseline covariate $\mathbf{Z}=0$ by $S_0(x)$, the above equation implies the following relationship between the baseline survival function $S_0(x)$ and the survival function given a covariate $\mathbf{Z}$, $S(x;\mathbf{Z})$,

$$S(x;\mathbf{Z})=S_0(x\exp[-\mathbf{Z}^T\beta]). \tag{3.4.12}$$

This is an accelerated failure time (AFT) representation of the log-linear model in equation 3.4.11. The AFT model in 3.4.12 is equivalent to the following characterizations

$$h(x;\mathbf{Z})=h_0(x\exp[-\beta^T\mathbf{Z}])\exp(-\beta^T\mathbf{Z}) \tag{3.4.13}$$

$$H(x;\mathbf{Z})=H_0(x\exp[-\beta^T\mathbf{Z}]). \tag{3.4.14}$$

Using Equation 3.4.12 or Equation 3.4.14, the quantile residual life time for a population with covariate vector $\mathbf{Z}$ is related to that of the baseline population as follows:

$$Q_{t_0}(\tau;\mathbf{Z},\theta)=Q_{t_0}(\tau;\mathbf{Z}=\mathbf{0},\theta)\exp(\beta^T\mathbf{Z}).$$

Thus, the effect of covariates on the quantiles is proportional. Likelihood-based inference of $Q_{t_0}(\tau;\mathbf{Z},\theta)$ using observed failure times from a heterogeneous population is considered next.

The observed data from a heterogeneous population are similar to the homogeneous population. In this case, the observed time X_i and censoring indicator δ_i for the i^{th} subject are defined given the subject's covariate vector $\mathbf{Z}_i=(Z_{1i},Z_{2i},\cdots,Z_{pi})$. The failure time T_i

given Z_i is assumed to be independent of C_i. Under this assumption, the likelihood of the random sample $\{(X_i, \delta_i, \mathbf{Z}_i)\}_{i=1}^n$ is given by,

$$L(\theta \mid \{(X_i, \delta_i, \mathbf{Z}_i)\}_{i=1}^n) = \prod_{i=1}^n f(X_i; \theta, \mathbf{Z}_i)^{\delta_i} S(X_i; \theta, \mathbf{Z}_i)^{1-\delta_i},$$

where $\theta = (\mathbf{b}, \beta)$ is a vector of regression parameter vectors. $\mathbf{b}$ is a parametric vector for the baseline distribution and β is a vector regression parameter. Using the AFT assumption, the likelihood can be written as

$$\begin{aligned} L(\theta \mid \{(X_i, \delta_i, \mathbf{Z}_i)\}_{i=1}^n) &= \prod_{i=1}^n \left[h_0(X_i \exp(-\beta^T \mathbf{Z}_i)) \exp(-\beta^T \mathbf{Z}_i)\right]^{\delta_i} S_0(X_i \exp(-\beta^T \mathbf{Z}_i)) \\ &= \prod_{i=1}^n \left[h_0(X_i \exp(-\beta^T \mathbf{Z}_i)) \exp(-\beta^T \mathbf{Z}_i)\right]^{\delta_i} \exp\left[-H_0(X_i \exp(-\beta^T \mathbf{Z}_i))\right]. \end{aligned}$$

Thus, the log-likelihood takes the following form,

$$\ell_n(\theta) = \sum_{i=1}^n \delta_i \ln\left[h_0(X_i \exp(-\beta^T \mathbf{Z}_i))\right] - \sum_{i=1}^n \delta_i \beta^T \mathbf{Z}_i - \sum_{i=1}^n H_0(X_i \exp(-\beta^T \mathbf{Z}_i)).$$

The score vector contains

$$\frac{\partial \ell_n}{\partial \mathbf{b}} = \sum_{i=1}^n \delta_i \frac{\partial \ell_n}{\partial \mathbf{b}} \ln\left[h_0(X_i \exp(-\beta^T \mathbf{Z}_i))\right] - \sum_{i=1}^n \frac{\partial \ell_n}{\partial \mathbf{b}} H_0(X_i \exp(-\beta^T \mathbf{Z}_i)), \tag{3.4.15}$$

the score sub-vector corresponding to the baseline parameter vector $\mathbf{b}$ and

$$\frac{\partial \ell_n}{\partial \beta} = \sum_{i=1}^n \delta_i \mathbf{Z}_i - \sum_{i=1}^n \left[\delta_i \frac{h_0'(X_i \exp(-\beta^T \mathbf{Z}_i))}{h_0(X_i \exp(-\beta^T \mathbf{Z}_i))} - H_0'(X_i \exp(-\beta^T \mathbf{Z}_i))\right] X_i \exp(-\beta^T \mathbf{Z}_i) \mathbf{Z}_i, \tag{3.4.16}$$

the score sub-vector corresponding to the regression coefficient vector β. The MLE of $\theta = (\mathbf{b}, \beta)$, $\hat{\theta} = (\hat{\mathbf{b}}, \hat{\beta})$ is found by solving

$$\frac{\partial \ell_n}{\partial \mathbf{b}} = 0, \quad \text{and} \quad \frac{\partial \ell_n}{\partial \beta} = 0$$

for $(\mathbf{b}, \beta)$. The sample information matrix is given by

$$I(\theta) = -\frac{\partial^2 \ell_n}{\partial \theta \partial \theta^T}, \quad \text{where} \quad \theta = (\mathbf{b}, \beta). \tag{3.4.17}$$

A consistent estimator of $Q_{t_0}(\tau;\mathbf{Z},\theta)$ is $\hat{Q}_{t_0} = Q_{t_0}(\tau;\mathbf{Z},\hat{\theta})$ with its asymptotic variance formula given by

$$\left\{\nabla\hat{Q}_{t_0}(\tau;\mathbf{Z})\right\}^T \mathcal{I}^{-1}(\theta)\left\{\nabla\hat{Q}_{t_0}(\tau;\mathbf{Z})\right\}, \tag{3.4.18}$$

where $\nabla\hat{Q}_{t_0}(\tau;\mathbf{Z})$ is:

$$\nabla\hat{Q}_{t_0}(\tau;\mathbf{Z}) = \begin{pmatrix}\nabla_{\mathbf{b}}\hat{Q}_{t_0}(\tau;\mathbf{Z})\\ \nabla_{\beta}\hat{Q}_{t_0}(\tau;\mathbf{Z})\end{pmatrix} = \begin{pmatrix}\exp(\beta^T\mathbf{Z})\nabla_{\mathbf{b}}\hat{Q}_{t_0}(\tau;\mathbf{Z}=\mathbf{0})\\ Q_{t_0}(\tau;\mathbf{Z}=\mathbf{0})\exp(\beta^T\mathbf{Z})\mathbf{Z}\end{pmatrix}.$$

3.4.2 Application to the Weibull Model

The likelihood of right-censored data $\{(X_i,\delta_i,\mathbf{Z}_i)\}_{i=1}^n$ based on baseline Weibull failure times under random censorship is of the form:

$$L(\alpha,\lambda,\beta) = \prod_{i=1}^n [\alpha\lambda\{X_i\exp(\beta^T\mathbf{Z}_i)\}^{\alpha-1}\exp(\beta^T\mathbf{Z}_i)]^{\delta_i}\exp[-\lambda\{X_i\exp(\beta^T\mathbf{Z}_i)\}^{\alpha}],$$

where the log-likelihood is given by

$$\ell_n(\alpha,\lambda,\beta) = (\ln\alpha+\ln\lambda)\sum_{i=1}^n\delta_i + (\alpha-1)\sum_{i=1}^n\delta_i\ln X_i + \alpha\sum_{i=1}^n\delta_i\beta^T\mathbf{Z_i} - \lambda\sum_{i=1}^n\{X_i\exp(\beta^T\mathbf{Z}_i)\}^{\alpha}.$$

Its score function is given by

$$\begin{aligned}
\frac{\partial\ell_n}{\partial\alpha} &= \alpha^{-1}\sum_{i=1}^n\delta_i + \sum_{i=1}^n\delta_i(\ln X_i+\beta^T\mathbf{Z}_i) - \lambda\sum_{i=1}^n\{X_i\exp(\beta^T\mathbf{Z}_i)\}^{\alpha}(\ln X_i+\beta^T\mathbf{Z}_i)\\
\frac{\partial\ell_n}{\partial\lambda} &= \lambda^{-1}\sum_{i=1}^n\delta_i - \sum_{i=1}^n\{X_i\exp(\beta^T\mathbf{Z}_i)\}^{\alpha}\ln(X_i\exp(\beta^T\mathbf{Z}_i))\\
\frac{\partial\ell_n}{\partial\beta_k} &= \left[\alpha\sum_{i=1}^n\delta_i - \lambda\sum_{i=1}^n\alpha\{X_i\exp(\beta^T\mathbf{Z}_i)\}^{\alpha}\right]Z_{ik} \quad \text{where} \quad k=1,2,\ldots,p.
\end{aligned}$$

The information matrix is

$$\begin{aligned}
\frac{\partial^2 \ell_n}{\partial^2 \alpha} &= -\alpha^{-2}\sum_{i=1}^{n}\delta_i - \lambda\sum_{i=1}^{n}\{X_i\exp(\beta^T\mathbf{Z}_i)\}^{\alpha}(\ln X_i + \beta^T\mathbf{Z}_i)^2 \\
\frac{\partial^2 \ell_n}{\partial\beta_k\partial\alpha} &= \left[\sum_{i=1}^{n}\delta_i - \lambda\sum_{i=1}^{n}\{X_i\exp(\beta^T\mathbf{Z}_i)\}^{\alpha} - \lambda\alpha\sum_{i=1}^{n}\{X_i\exp(\beta^T\mathbf{Z}_i)\}^{\alpha}(\ln X_i + \beta^T\mathbf{Z}_i)\right]Z_{ik} \\
\frac{\partial^2 \ell_n}{\partial\lambda\partial\beta_k} &= \sum_{i=1}^{n}\{X_i\exp(\beta^T\mathbf{Z}_i)\}^{\alpha}Z_{ik} \\
\frac{\partial^2 \ell_n}{\partial^2 \lambda} &= -\lambda^{-2}\sum_{i=1}^{n}\delta_i \\
\frac{\partial^2 \ell_n}{\partial\lambda\partial\beta_k} &= \sum_{i=1}^{n}\{X_i\exp(\beta^T\mathbf{Z}_i)\}^{\alpha}Z_{ik} \\
\frac{\partial^2 \ell_n}{\partial\beta_k\partial\beta_l} &= \left[-\lambda\sum_{i=1}^{n}\alpha^2\{X_i\exp(\beta^T\mathbf{Z}_i)\}^{\alpha}\right]Z_{ik}Z_{il} \quad \text{where} \quad k,l = 1,2,\ldots,p.
\end{aligned}$$

The asymptotic variance formula is given by

$$\left\{\nabla\hat{Q}_{t_0}(\tau;\mathbf{Z})\right\}^T\mathcal{I}^{-1}(\theta)\left\{\nabla\hat{Q}_{t_0}(\tau;\mathbf{Z})\right\} \tag{3.4.21}$$

where $\nabla\hat{Q}_{t_0}(\tau;\mathbf{Z})$ is:

$$\nabla\hat{Q}_{t_0}(\tau;\mathbf{Z}) = \begin{pmatrix} \exp(-\hat{\beta}^T\mathbf{Z})\nabla_{\hat{\mathbf{b}}}\hat{Q}_{t_0}(\tau;\mathbf{Z}=\mathbf{0}) \\ -\hat{Q}_{t_0}(\tau;\mathbf{Z}=\mathbf{0})\exp(-\hat{\beta}^T\mathbf{Z})\mathbf{Z} \end{pmatrix}$$

where

$$\nabla_{\hat{\mathbf{b}}}\hat{Q}_{t_0}(\tau;\mathbf{Z}=\mathbf{0}) = \begin{pmatrix} \hat{\alpha}^{-2}(t_0^{\hat{\alpha}} - \hat{\lambda}^{-1}\ln(1-\tau))^{1/\hat{\alpha}}\left(\frac{\hat{\alpha}t_0^{\hat{\alpha}}\ln(t_0)}{(t_0^{\hat{\alpha}}-\hat{\lambda}\ln(1-\tau))} - \ln(t_0^{\hat{\alpha}} - \hat{\lambda}\ln(1-\tau))\right) \\ \hat{\alpha}^{-1}\hat{\lambda}^{-2}(t_0^{\hat{\alpha}} - \hat{\lambda}^{-1}\ln(1-\tau))^{-1+1/\alpha}\ln(1-\tau) \end{pmatrix}$$

and

$$\hat{Q}_{t_0}(\tau;\mathbf{Z}=0) = \left(\left[t_0^{\hat{\alpha}} - \hat{\lambda}^{-1}\ln(1-\tau)\right]^{1/\hat{\alpha}} - t_0\right).$$

3.4.3 Simulation Study

We conducted simulation studies to investigate the performance of the proposed method. Weibull($\alpha = 2$, $\lambda = 0.5$) was used as the baseline distribution. A bivariate vector (Z_1, Z_2) was generated by simulating Z_1 from $Bernoulli(0.5)$ and Z_2 from $N(0, 1)$. The regression coefficient for Z_1 was set at $\beta_1 = 0.5$ and for Z_2 was set at $\beta_2 = 1$. A censoring proportion of 0.3 was used. A sample size of 200 and Monte Carlo sample size of 5000 were used.

Inference for the true residual quantile life at the two values of the dichotomous variable $(Z_1 = 0, 1)$ and at two values of the continuous covariates $(Z_2 = 0, 1)$ were investigated. The simulation results for these settings are given in Tables 3.4.1-3.4.4. The results suggest that the proposed procedure works well. The coverage probability is close to the nominal level, the estimates are unbiased, and the asymptotic and Monte Carlo variances are close to each other.

Table 3.4.1: Regression simulation result $(Z_1 = 0, Z_2 = 0)$

t_0	τ	$\hat{Q}_{t_0}$	$\widehat{Avar(\hat{Q}_{t_0})}$	Q_{t_0}	$\widehat{var(\hat{Q}_{t_0})}$	$\widehat{Bias}$	$\widehat{MSE}$	CP
0	0.25	0.764	0.003	0.759	0.003	0.006	0.003	0.945
	0.50	1.179	0.005	1.177	0.005	0.002	0.005	0.944
1	0.25	0.254	0.001	0.255	0.001	-0.001	0.001	0.943
	0.50	0.541	0.003	0.545	0.003	-0.003	0.003	0.942

Table 3.4.2: Regression simulation result ($Z_1 = 1, Z_2 = 0$)

t_0	τ	$\hat{Q}_{t_0}$	$\widehat{Avar(Q_{t_0})}$	Q_{t_0}	$\widehat{var(\hat{Q}_{t_0})}$	$\widehat{Bias}$	$\widehat{MSE}$	CP
0	0.25	1.259	0.009	1.251	0.009	0.008	0.009	0.951
	0.50	1.943	0.017	1.941	0.016	0.002	0.016	0.948
1	0.25	0.418	0.001	0.421	0.001	-0.002	0.001	0.938
	0.50	0.891	0.006	0.898	0.006	-0.007	0.006	0.938

Table 3.4.3: Regression simulation result ($Z_1 = 0, Z_2 = 1$)

t_0	τ	$\hat{Q}_{t_0}$	$\widehat{Avar(Q_{t_0})}$	Q_{t_0}	$\widehat{var(\hat{Q}_{t_0})}$	$\widehat{Bias}$	$\widehat{MSE}$	CP
0	0.25	2.081	0.037	2.062	0.038	0.020	0.038	0.953
	0.50	3.212	0.076	3.201	0.077	0.011	0.077	0.948
1	0.25	0.692	0.008	0.694	0.009	-0.001	0.009	0.939
	0.50	1.474	0.035	1.481	0.036	-0.007	0.036	0.937

Table 3.4.4: Regression simulation result ($Z_1 = 1, Z_2 = 1$)

t_0	τ	$\hat{Q}_{t_0}$	$\widehat{Avar(Q_{t_0})}$	Q_{t_0}	$\widehat{var(\hat{Q}_{t_0})}$	$\widehat{Bias}$	$\widehat{MSE}$	CP
0	0.25	3.443	0.115	3.399	0.115	0.043	0.117	0.951
	0.50	5.310	0.245	5.277	0.246	0.033	0.247	0.949
1	0.25	1.142	0.018	1.143	0.018	-0.001	0.018	0.939
	0.50	2.432	0.081	2.441	0.081	-0.010	0.081	0.938

3.4.4 Data Analysis

For the regression analysis, we fitted a Weibull regression model using the combined (node negative and node positive patients) dataset. We used the covariates node status, age (in years) and tumor size (in mm) as covariates. Based on the fitted model, estimates of the quantile residual life for a node-negative patient at the median age (57 yr) and median tumor size (30 mm) are presented in Table 3.4.5. The corresponding estimates for a node-positive patient at the median age (55 yr) and median tumor size (32 mm) are given in Table 3.4.6. The results are similar to the results presented in the one-sample setting. We also observe the lack of appreciable dependency of the estimated quantile residual life times on the time points, suggesting that the baseline shape parameter is close to 1 and exponential regression might work for the data.

We checked the plausibility of the assumed Weibull model using the Quantile-Quantile plot of the Kaplan Meier estimator and the fitted Weibull model (see Figure 3.4.1). The assumed model appears to fit the data well.

Table 3.4.5: Quantile residual life analysis for a node-negative patient at the median age (57 yrs) and tumor size (30 mm)

t_0	τ	$\hat{Q}_{t_0}$	$\widehat{Avar(\hat{Q}_{t_0})}$	CI_{LL}	CI_{UL}
0	0.1	1.926	0.014	1.692	2.160
	0.2	3.985	0.038	3.605	4.365
	0.5	11.949	0.185	11.106	12.792
2	0.1	1.826	0.005	1.693	1.959
	0.2	3.840	0.019	3.570	4.111
	0.5	11.724	0.178	10.898	12.550
4	0.1	1.793	0.004	1.668	1.917
	0.2	3.781	0.018	3.516	4.046
	0.5	11.603	0.196	10.734	12.471
6	0.1	1.772	0.004	1.644	1.901
	0.2	3.742	0.020	3.465	4.019
	0.5	11.515	0.220	10.595	12.434

Table 3.4.6: Quantile residual life analysis for a node-positive patient at median age (55 yrs) and tumor size (32 mm)

t_0	τ	$\hat{Q}_{t_0}$	$\widehat{Avar}(\hat{Q}_{t_0})$	CI_{LL}	CI_{UL}
0	0.1	1.361	0.010	1.168	1.553
	0.2	2.815	0.028	2.486	3.144
	0.5	8.441	0.158	7.662	9.220
2	0.1	1.290	0.004	1.165	1.415
	0.2	2.713	0.017	2.459	2.967
	0.5	8.282	0.145	7.534	9.029
4	0.1	1.266	0.004	1.150	1.383
	0.2	2.671	0.015	2.427	2.915
	0.5	8.196	0.150	7.438	8.954
6	0.1	1.252	0.003	1.136	1.367
	0.2	2.643	0.016	2.398	2.889
	0.5	8.134	0.158	7.356	8.912

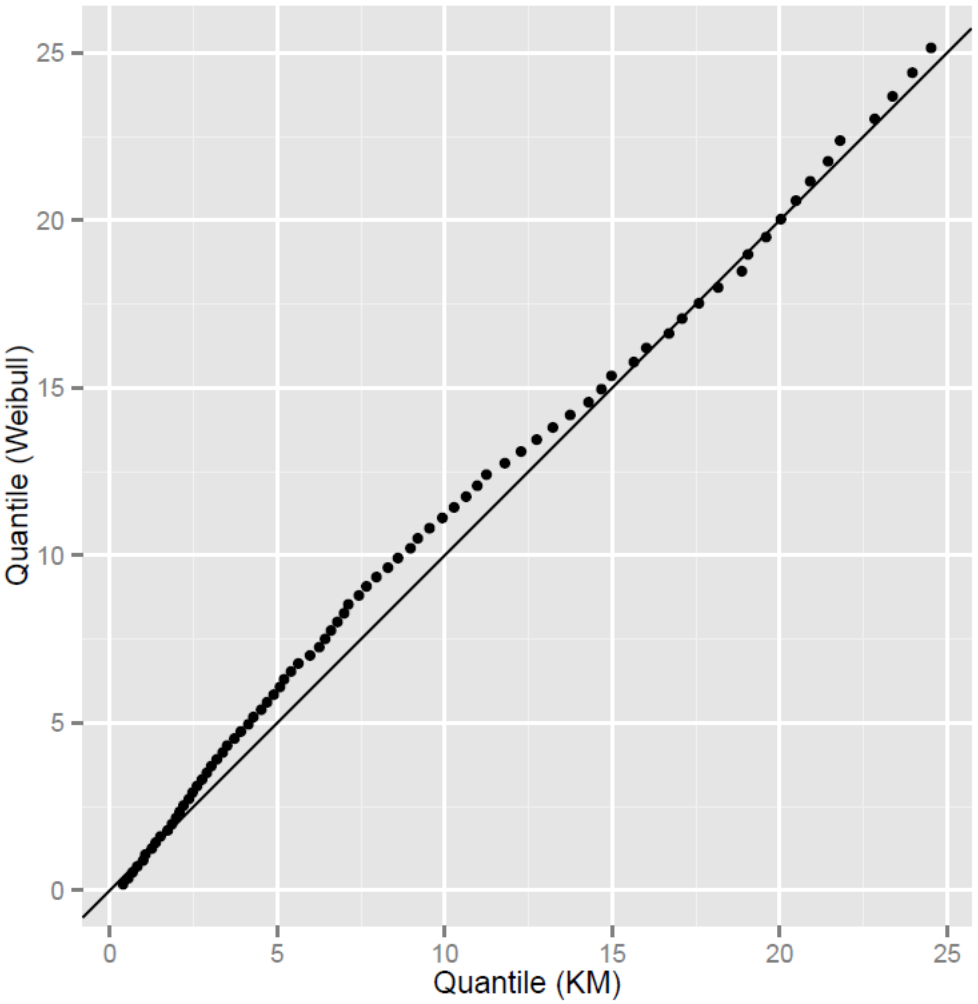

Figure 3.4.1: Kaplan Meier Quantile vs. Weibull fitted Quantile plot

4.0 COMPETING RISKS

In Chapter 3, parametric inference for the residual quantile life function of a failure time was considered. In some applications, it is important to make inference on the quantile residual life function of a failure time from a particular cause of failure in the presence of at least one other cause of failure. This is a competing risks problem.

In this chapter, parametric procedures developed previously are extended to the competing risks scenario. Parametric inferences for one sample, two-sample, and regression cases are developed. The performance of the methods is investigated using simulations. The methods are applied to a real data set.

4.1 INTRODUCTION

In Section 3.1, survival quantities that are important to describe a failure time distribution were presented. In this section, we will present corresponding survival quantities for the competing risks situation. In some survival studies, we have failures due to several causes. For instance, in survival study of cancer patients, where the primary event of interest is time to death due to cancer, mortality from non-cancer causes occurs precluding (censoring) the primary event from occurring. In such a scenario, we observe a failure time and failure type (death due to cancer and death due non-cancer causes). This is an example of a competing risks problem. Generally, competing risk data are of the type (T, δ), where T is a failure time and δ is a failure type variable. In medical studies, T is subject to a random right censoring. Under random censorship, the only survival quantities that are identifiable from

the competing risks data are cause-specific hazards or functions of them Prentice et al. (1978). This fact is the basis for the cause-specific hazard modeling framework, where the hazards for each cause are modeled separately.

The survival quantities presented in Section 3.1 are generalized to competing risks below. They are defined by cause type. Without loss of generality, we will assume that there are two distinct types or causes of failure.

The k^{th} cause-specific hazard function is defined as

$$\begin{aligned} h_k(t) &= \lim_{\Delta t \to 0} \frac{P[t \leq T < t + \Delta t, \delta = k | T \geq t]}{\Delta t} \\ &= \frac{f_k(t)}{S(t)} \quad k = 1, 2, \end{aligned} \tag{4.1.01}$$

where $f_k(t)$ is the density of $(T, \delta = k)$ and $S(t)$ is the all cause survival function. The k^{th} cause hazard function quantifies the rate at which subjects who are still at risk of failing at time t from any cause are failing from the k^{th} cause at time t. Because the two event types are mutually exclusive, the cause-specific hazard functions add up to the overall hazard function,

$$h(t) = h_1(t) + h_2(t). \tag{4.1.02}$$

Upon integrating both sides of Equation 4.1.02, we get

$$\begin{aligned} \int_0^t h(u)\, du &= \int_0^t h_1(u)\, du + \int_0^t h_2(u)\, du \\ H(t) &= H_1(t) + H_2(t). \end{aligned} \tag{4.1.03}$$

That is, the cause-specific cumulative hazard functions add up to the overall cumulative hazard function. Using the fact that $S(t) = \exp(-H(t))$ and the decomposition in Equation 4.1.03, the all cause survival function can be written as,

$$S(t) = \exp(-H_1(t) - H_2(t)) = S_1(t)S_2(t), \tag{4.1.04}$$

where we have denoted $\exp(-H_k(t))$ by $S_k(t)$ for $k = 1, 2$.

To define the k^{th} cause τ-quantile residual life, we need the subdistribution function or cumulative incidence function for the k^{th} cause,

$$F_k(t) = P(T \leq t, \delta = k) = \int_0^t f_k(u)du = \int_0^t S(u)h_k(u)du$$
$$= \int_0^t \exp(-H(u))h_k(u)du, \quad k = 1, 2 \qquad (4.1.05)$$

where we have used Equation 4.1.01 in the third equality. Equation 4.1.05 suggests that $F_k(t)$ is completely determined by the cause-specific hazard functions. Using Equation 4.1.05, we can easily see that distribution function of T is related to the k^{th} cause cumulative incidence functions as $F(t) = F_1(t) + F_2(t)$. Similarly, the density function of T is related to the densities of the k^{th} cause cumulative incidence function as $f(t) = f_1(t) + f_2(t)$.

Given F and F_k, the τ-quantile residual life time for the k^{th} cause, $Q_{k,t_0}(\tau)$, is defined as

$$Q_{k,t_0}(\tau) = F^{-1}_{k,t_0}(\tau), \quad \tau \in (0, F_k(\infty));$$

where $F_{k,t_0}(t)$ denotes the distribution function of $(T - t_0, \delta = k)|T > t_0$. The k^{th} cause τ-quantile residual life at time t_0, $Q_{k,t_0}(\tau)$, can be written in terms of the k^{th} cause τ-quantile residual life at time $t_0 = 0$, $Q_k(\tau)$. Indeed,

$$F_{k,t_0}(t) = P(T - t_0 \leq t, \delta = k \mid T \geq t_0)$$
$$= \frac{P(t_0 < T \leq t + t_0, \delta = k)}{P(T > t_0)}.$$

This implies

$$F_{k,t_0}(t) = \frac{F_k(t + t_0) - F_k(t_0)}{1 - F(t_0)} \qquad (4.1.06)$$

when $T > t_0$. Thus,

$$Q_{1,t_0}(\tau) = F_1^{-1}(\tau(1 - F(t_0)) + F_1(t_0)) - t_0 \qquad (4.1.07)$$

where $0 < \tau < 1$. Equation 4.1.07 is equivalent to:

$$Q_{1,t_0}(\tau) = Q_1(\tau(1 - F(t_0)) + F_1(t_0)) - t_0. \qquad (4.1.08)$$

For parametric inference on $Q_{1,t_0}(\tau)$, we will place a parametric assumption on $F_k(t)$ via parameterizing the cause-specific hazards $h_k(t)(k = 1, 2)$, which are estimable from competing risks data and completely determine the competing risks process (Beyersmann et al., 2009). Since there is little restriction on the hazard function, this approach allows us to flexibly choose the forms of the cause-specific hazards. This approach also automatically gurantees the improperness of $F_k(t)$ $(k = 1, 2)$ and their additivity to $F(t)$ (see Shi et al. (2013)). Moreover, when the cause-specific hazards do not share any parameter, the methods of Chapter 3 can be directly used for constructing and estimating the likelihood quantities.

As in Chapter 3, we will assume that regularity conditions needed to make the asymptotics work hold. In particular, we will assume that $Q_{1,t_0}(\tau)$ is differentiable with respect to θ.

4.2 ONE SAMPLE

In this section, parametric inference on the true residual quantile life time based on a random sample of right-censored data from a homogeneous population under a competing-risks setting is considered.

4.2.1 Theory

Without loss of generality, we will assume that there are two types of failure, $k = 1, 2$. Suppose we have a random sample of potential failure times $\{T_i\}_{i=1}^n$ from a distribution function $F(\cdot;\theta)$ and a corresponding sample of potential censoring times $\{C_i\}_{i=1}^n$ from a distribution function $G(\cdot)$. If the observed time is a failure time, the failure type is one of the two types. The observed data are a random sample $\{(X_i, \delta_{ki})\}_{i=1}^n$ where $X_i = \min(T_i, C_i)$ and δ_{ki} are the observed time and failure type indicator respectively for the i^{th} observation. $\delta_{ki} = 1$, if the failure type is k, and $X_i = T_i$. Otherwise, $\delta_{ik} = 0$. If $\delta_{i1} = \delta_{i2} = 0$, then the i^{th} observation is censored. Based on the observed data, inference on Q_{1,t_0}, the true quantile

residual life time for failure times of failure type or cause 1 at time t_0,

$$Q_{1,t_0}(\tau;\theta) = F_1^{-1}(\tau(1 - F(t_0;\theta)) + F_1(t_0;\theta)) - t_0 \tag{4.2.11}$$

is considered. Likelihood based point and confidence interval estimators are developed below.

The likelihood function for the competing risks data is

$$L(\theta \mid \{(X_i, \delta_{ki})\}_{i=1}^n) = \prod_{i=1}^n f_1(X_i;\theta)^{\delta_{1i}} f_2(X_i;\theta)^{\delta_{2i}} S(X_i;\theta)^{1-(\delta_{1i}+\delta_{2i})}$$

where $\theta = (\theta_1, \theta_2)$ with $\theta_k = (\theta_{k1}, \theta_{k2}, \ldots, \theta_{kp_k})$ being the vector of parameters corresponding to the k^{th} cause.

The likelihood can be written in terms of the cause-specific hazards as

$$L(\theta \mid \{(X_i, \delta_{ki})\}_{i=1}^n) = \prod_{i=1}^n h_1(X_i;\theta_1)^{\delta_{1i}} h_2(X_i;\theta_2)^{\delta_{2i}} S(X_i;\theta) \tag{4.2.12}$$

$$= \prod_{i=1}^n h_1(X_i;\theta_1)^{\delta_{1i}} S_1(X_i;\theta_1) \prod_{i=1}^n h_2(X_i;\theta_2)^{\delta_{2i}} S_2(X_i;\theta_2), \tag{4.2.13}$$

where we have used the identity $S(X_i;\theta) = S_1(X_i;\theta_1)S_2(X_i;\theta_2)$ to go from 4.2.12 to 4.2.13. Due to the factorization of the likelihood into the likelihood contribution for each cause, the likelihood quantities for each cause can be estimated separately. Equation 4.2.13 implies that likelihood quantities for the k^{th} cause can be estimated in the usual way by censoring failure times from the other causes at their observed times. Thus, likelihood quantities (log-likelihood, score vectors, MLEs, and information matrix) for each cause are exactly as in Section 3.2.

By the consistency of $\hat{\theta}$, the MLE of $\theta = (\theta_1, \theta_2)$, and the differentiability of $Q_{1,t_0}(\tau;\theta)$, a consistent estimator of the quantile residual life time for the failure time of failure type 1 is,

$$\hat{Q}_{1,t_0}(\tau;\theta) = Q_{1,t_0}(\tau;\hat{\theta}) = F_1^{-1}(\tau(1 - F(t_0;\hat{\theta})) + F_1(t_0;\hat{\theta})) - t_0. \tag{4.2.14}$$

Using the same argument as in Section 3.2, the asymptotic distribution of $\hat{Q}_{1,t_0}(\tau;\theta)$ follows

$$\hat{Q}_{1,t_0} \sim AN\Big(Q_{1,t_0}, Avar(\hat{Q}_{1,t_0})\Big), \tag{4.2.15}$$

where

$$Avar(\hat{Q}_{1,t_0}) = \frac{1}{n}\left\{\nabla_{\hat{\theta}} Q_{1,t_0}(\tau;\hat{\theta})_{|\hat{\theta}=\theta}\right\}^T \mathcal{I}^{-1}(\theta)\left\{\nabla_{\hat{\theta}} Q_{t_0}(\tau;\hat{\theta})_{|\hat{\theta}=\theta}\right\}. \tag{4.2.16}$$

$\mathcal{I}^{-1}(\theta)$ in Equation 4.2.16 is the inverse of the expected information matrix and is consistently estimable by $nI^{-1}(\hat{\theta})$. $I^{-1}(\hat{\theta})$ is the inverse of the sample Fisher information matrix evaluated at $\hat{\theta}$, the MLE of θ. The sample Fisher information matrix is given by

$$I(\theta) = -\frac{\partial^2}{\partial\theta\partial\theta^T}\ell_n(\theta).$$

For inference (testing and confidence intervals) on the quantile residual life time of the failure type 1, the unknown quantities in Equation 4.2.16 are estimated by their consistent estimators. A consistent estimator of the asymptotic variance (Equation‘4.2.16) is given by

$$\widehat{Avar(\hat{Q}_{1,t_0})} = \left\{\nabla_{\hat{\theta}} Q_{1,t_0}(\tau;\hat{\theta})_{|\hat{\theta}=\hat{\theta}}\right\}^T I^{-1}(\hat{\theta})\left\{\nabla_{\hat{\theta}} Q_{1,t_0}(\tau;\hat{\theta})_{|\hat{\theta}=\hat{\theta}}\right\}. \tag{4.2.17}$$

Using the point estimate $\hat{Q}_{1,t_0}$ and the estimator of its asymptotic variance, Wald-type approximate tests and CIs for Q_{1,t_0} can be constructed. An approximate two-sided $100(1-\alpha)\%$ Wald confidence interval(CI) for $Q_{1,t_0}(\tau;\theta)$ is:

$$\left(\hat{Q}_{t_{1,0}}(\tau;\theta) \pm z_{1-\alpha/2}\sqrt{\widehat{Avar(\hat{Q}_{1,t_0})}}\right). \tag{4.2.18}$$

4.2.2 Computational Issues

In Section 4.2.1, the applicability of the formulas for the cause-specific quantile residual life (Equation 4.2.14) and asymptotic variance of the cause-specific quantile residual life (Equation 4.2.16) relies on their computability. Generally, the cause-specific quantile residual life does not have a closed form. This means the gradient of the cause-specific quantile residual life in the asymptotic variance formula cannot be calculated analytically. This lack of closed-form representation of the quantile residual life also complicates the simulation of competing risks data. These computational issues do not normally arise in the non-competing risks settings. We have not seen any reference or solutions to these problems in the competing-risks literature. We will propose a solution to this problem in this section.

4.2.2.1 Estimation of the Quantile Residual Life

The k^{th}-cause quantile residual life time $Q_{k,t_0}(\tau)$ satisfies the following equation,

$$F_{k,t_0}(Q_{k,t_0}(\tau)) = \frac{F_k(Q_{k,t_0}(\tau)+t_0) - F_k(t_0)}{S(t_0)} = \tau.$$

This implies

$$F_k(Q_{k,t_0}(\tau)+t_0) - F_k(t_0) - S(t_0)\tau = 0,$$

which can be written as

$$\int_{t_0}^{Q_{k,t_0}(\tau)+t_0} f_k(u)\,du - S(t_0)\tau = 0. \tag{4.2.21}$$

The quantile residual life time can be determined by solving Equation 4.2.21 for $Q_{k,t_0}(\tau)$. In cases where the integral has has a closed form representation, the quantile residual life time can be solved analytically. Generally, the quantile residual life time has to be determined numerically. We will suggest an algorithm for obtaining the quantile residual life time numerically. To do so, we will write the left-hand-side of Equation 4.2.21 as a function in $q^\star$,

$$R(q^\star) = \int_{t_0}^{q^\star} f_k(u)\,du - S(t_0)\tau, \tag{4.2.22}$$

where $q^\star = Q_{k,t_0}(\tau)+t_0$. The quantile residual life time is the root of $R(q^\star)$ minus t_0. The root of $R(q^\star)$ can be found using the Newton-Raphson algorithm. Indeed, given an initial value $q_0^\star$, the first iterate $q_1^\star$ is given by

$$q_1^\star = q_0^\star - \frac{R(q_0^\star)}{f_k(q_0^\star)}.$$

Generally, given the n^{th} iterate $q_n^\star$, the $(n+1)^{th}$ iterate $q_{n+1}^\star$ is given by

$$q_{n+1}^\star = q_n^\star - \frac{R(q_n^\star)}{f_k(q_n^\star)}.$$

This iteration continues until some convergence criterion, say, $|q_{n+1}^\star - q_n^\star| \le c$, for c a tolerance limit, is met. Since $R(q^\star)$ is not available analytically, it should also be estimated numerically at each iteration. Suppose the final iterate is denoted by $q_{n^\star}^\star$, the quantile residual life is given by $Q_{k,t_0}(\tau) = q_{n^\star}^\star - t_0$.

All the quantile functions and density functions above are functions of a distributional parameter vector, θ. When θ is a known (for example, in a simulation study), the quantile residual life determined by the above scheme is the true quantile residual life for the population with that parameter value. When the θ is replaced by its estimator $\hat{\theta}$, the quantile residual life determined by the above scheme is an estimate of the true quantile residual life.

4.2.2.2 Estimation of the gradient of the Quantile Residual Life

The asymptotics of the quantile residual life requires the determination of the gradient of the estimated quantile residual life. The determination of the gradient formula is given below.

Consider finding the gradient of both sides of Equation 4.2.21, i.e.,

$$\nabla_{\hat{\theta}} \int_{t_0}^{Q_{k,t_0}(\tau)+t_0} f_k(u)\, du - \tau \nabla_{\hat{\theta}} S(t_0) = 0.$$

Applying Leibniz's rule for differentiation under the integral sign, we obtain

$$\int_{t_0}^{Q_{k,t_0}(\tau)+t_0} \nabla_{\hat{\theta}} f_k(u;.)\, du + f_k(Q_{j,t_0}(\tau)+t_0) \nabla_{\hat{\theta}} Q_{k,t_0}(\tau) - \tau \nabla_{\hat{\theta}} S(t_0) = 0.$$

Interchange integration and differentiation to get

$$\nabla_{\hat{\theta}} \int_{t_0}^{Q_{k,t_0}(\tau)+t_0} f_k(u;.)\, du + f_k(Q_{k,t_0}(\tau)+t_0) \nabla_{\hat{\theta}} Q_{k,t_0}(\tau) - \tau \nabla_{\hat{\theta}} S(t_0) = 0,$$

which can be written as

$$\nabla_{\hat{\theta}} \Big\{ F_k(Q_{k,t_0}(\tau)+t_0;.) - F_k(t_0;.) \Big\} + f_k(Q_{k,t_0}(\tau)+t_0) \nabla_{\hat{\theta}} Q_{k,t_0}(\tau;.) - \tau \nabla_{\hat{\theta}} S(t_0;.) = 0.$$

Solving for the gradient of the quantile residual life time, we get

$$\nabla_{\hat{\theta}} Q_{k,t_0}(\tau) = f_k(Q_{k,t_0}(\tau)+t_0)^{-1} \left\{ \nabla_{\hat{\theta}} F_k(t_0;.) - \nabla_{\hat{\theta}} F_k(Q_{k,t_0}(\tau)+t_0;.) - \tau \nabla_{\hat{\theta}} F(t_0;.) \right\}. \tag{4.2.23}$$

In Equation 4.2.23, the gradient with respect to $\hat{\theta}$ of the k^{th} cause cumulative incidence function is obtained by first differentiating the k^{th} cause sub-density and numerically integrating it. This interchange of integration and differentiation allows us to numerically obtain the gradient of $F_k(t;.)$ (and, hence, that of $Q_{k,t_0}(\tau)$) when neither is available in closed form. Thus, the asymptotic variance of the cause-specific quantile residual life time can be expressed in terms

of the gradient of the cumulative incidence functions. This formula is particularly useful when the cause-specific quantile residual life time does not have an explicit representation (i.e., when it is a result of a numerical iteration).

In the case of non-competing risks setting, Equation 4.2.23 is reduced to

$$\nabla_{\hat{\theta}} Q_{t_0}(\tau) = f(Q_{t_0}(\tau) + t_0)^{-1} \left\{(1-\tau)\nabla_{\hat{\theta}} F(t_0; .) - \nabla_{\hat{\theta}} F(Q_{t_0}(\tau) + t_0; .)\right\}.$$

4.2.2.3 Simulating Competing Risks Data The lack of a closed-form expression of the cumulative incidence function (hence, the cause-specific quantiles) complicates the simulation of competing risks data. We will show this difficulty by reviewing competing risks simulation procedures in use in the literature. Competing risks data are mostly simulated using the latent failure time approach (Beyersmann et al., 2009). This approach has limitations because, among other things, it requires the specification of dependency structure among the cause-specific latent failure times and has physical and statistical identifiability problems (Prentice et al., 1978). We won't consider this approach in the review below. Competing risk data that don't use the latent failure time approach are simulated in either of the following two methods.

Method 1:

This method uses a mixture representation of the distribution of the all cause failure time T in terms of the sub-distributions of the competing causes. In the case where there are only two competing causes ($k = 1, 2$) of failure, we have

$$\begin{aligned} P(T \leq t) &= P(T \leq t, \delta = 1) + P(T \leq t, \delta = 2) \\ &= P(\delta = 1)P(T \leq t|\delta = 1) + P(\delta = 2)P(T \leq t|\delta = 2) \end{aligned} \tag{4.2.24}$$

where

$$P(T \leq t, \delta = k) = \int_0^t S(u)h_k(u)du,$$

and

$$P(\delta = k) = \int_0^\infty S(u)h_k(u)du, \tag{4.2.25}$$

with $P(\delta = 1) + P(\delta = 2) = 1$.

This implies

$$P(T \leq t|\delta = k) = \int_0^t P(\delta = k)^{-1} S(u)h_k(u)du.$$

Thus, the density of $T|\delta = k$ is:

$$P(\delta = k)^{-1} S(t)h_k(t). \tag{4.2.26}$$

Method 1 uses the mixture representation in Equation (4.2.24) to first generate the failure type from a $Bernoulli(1, P(\delta = k))$ and then generate a failure time from $T|\delta = k$, where its density is given in Equation (4.2.26). This method has been used, for example, by Jeong and Fine (2013), Jeong and Fine (2009) and Ng and McLachlan (2003) where they assume that $T|\delta = k$ follow Weibull, Gompertz, and exponential distributions, respectively. This simplifying assumption makes simulating from $T|\delta = k$ straightforward.

Under the cause-specific hazard modelling approach used in this dissertation, Equation (4.2.26) does not have simple distributional form even in simple settings where we assume that each cause-specific hazard can be independently modeled by Weibull hazard functions. The application of Method 1 to the cause-specific hazard modelling set-up requires the determination of $P(\delta = k)$. This marginal probability of failure type 1 is not generally available analytically. Thus, it has to be usually estimated numerically. Moreover, once $P(\delta = k)$ is found, we should be able to simulate from $P(\delta = k)^{-1} S(u)h_k$, which maybe not be amenable to standard simulation procedures such as the inverse transform method. We propose below the use of the Accept-Reject algorithm to simulate from $T|\delta = k$.

To utilize the Accept-Reject algorithm, we need a density that is easy to simulate from and that dominates the target density. To come up with a suitable dominating density, we write Equation (4.2.26) as

$$P(\delta = k)^{-1} \exp(-H_1(t) - H_2(t))h_k(t) \leq P(\delta = k)^{-1} \exp(-H_k(t))h_k(t), \tag{4.2.27}$$

where the upper bound follows from the fact that $S_k(t) \leq 1 \quad (k = 1, 2)$. The expression $\exp(-H_k(t))h_k(t)$ in the upper bound is the so-called instrumental (or dominating) density and $P(\delta = k)^{-1}$ is the expected number of attempts for each acceptance (see Casella and Robert (2004)).

To simulate T from the target density $P(\delta = k)^{-1} \exp(-H_1(t) - H_2(t))h_k(t)$ using the Accept-Reject algorithm, we generate X from the dominating density $\exp(-H_k(x))h_k(x)$, and conditional on $X = x$ we simulate a uniform random variate $U|x$ from $Unif(0, P(\delta = k)^{-1} \exp(-H_k(x))h_k(x))$. If $0 < u < P(\delta = k)^{-1} \exp(-H_1(x) - H_2(x))h_k(x)$, we take $T = x$ (i.e, x is accepted). Otherwise, this procedure is started over (i.e., x is rejected).

We illustrate this method using the Weibull model. When the cause-specific shape parameters in the Weibull-model, α_1 and α_2, are allowed to vary independently, there is no closed-form expression for the k^{th} cause cumulative incidence function nor for the density of $T|\delta = k$. In this case, to simulate competing risks data from $T|\delta = k$, we first estimate $P(\delta = k)^{-1}$ numerically. Then, we apply the Accept-Reject algorithm to simulate from the density of $T|\delta = k$. For the Weibull model, Equation 4.2.27 is given by,

$$P(\delta = k)^{-1} \exp(-\lambda_1 t^{\alpha_1} + -\lambda_2 t^{\alpha_2})\lambda_k \alpha_k t^{\alpha_k - 1} \leq P(\delta = k)^{-1} \exp(-\lambda_k t^{\alpha_k})(\lambda_k \alpha_k t^{\alpha_k - 1}),$$

where $\exp(-\lambda_k t^{\alpha_k})(\lambda_k \alpha_k t^{\alpha_k - 1})$ is the instrumental (or dominating) density and $P(\delta = k)^{-1}$ is the expected number of attempts for each acceptance. The dominating density is that of a Weibull random variable.

Method 2:

This method, discussed in Beyersmann et al. (2009), suggests simulating first from the distribution function for the all cause failure time T, and conditional on the observed failure time $T = t$ assigning a failure type according to a random draw of causes from a $Bernoulli\left(1, \frac{h_k(t)}{h_1(t)+h_2(t)}\right)$ with $k = 1, 2$. When it is easy to simulate from the distribution function of the all cause failure time T, this method can be conveniently used to generate competing risks data. Simulating competing risks data using this method when the cause-specific hazards are parameterized independently is not obvious.

4.2.3 Application to the Weibull Model

4.2.3.1 Independent shape and scale parameters The hazard function for the Weibull distribution is: $\alpha\lambda t^{\alpha-1}$. Thus, the k^{th} cause-specific hazard function is

$$h_k(t;\alpha_k,\lambda_k) = \alpha_k\lambda_k t^{\alpha_k-1}.$$

Thus, the all cause survival function is $S(t;\alpha_1,\lambda_1,\alpha_2,\lambda_2) = \exp(-\lambda_1 t^{\alpha_1} - \lambda_2 t^{\alpha_2})$. Hence, the k^{th} cause cumulative incidence function is given by,

$$F_k(t;\alpha_1,\lambda_1,\alpha_2,\lambda_2) = \int_0^t \exp(-\lambda_1 u^{\alpha_1} - \lambda_2 u^{\alpha_2})\alpha_k\lambda_k u^{\alpha_k-1}du \quad , k = 1,2. \tag{4.2.31}$$

The k^{th} cumulative incidence function (Equation 4.2.31) does not have a closed form except in the case of common shape parameter, i.e., $\alpha_1 = \alpha_2 = \alpha$. Thus, in the case of independent shape parameters, the quantile residual life time and its gradient should be evaluated numerically using the procedure discussed in Section 4.2.2.1 and Section 4.2.2.2, respectively.

Using the result in Section 4.1, the likelihood based on a right-censored competing risks data is:

$$\begin{aligned} &L(\alpha_1,\lambda_1,\alpha_2,\lambda_2 \mid \{(X_i,\delta_{ki})\}_{i=1}^n) \\ &\propto \prod_{i=1}^n (\alpha_1\lambda_1 X_i^{\alpha_1-1})^{\delta_{1i}} \exp(-\lambda_1 X_i^{\alpha_1} - \lambda_2 X_i^{\alpha_2}) \prod_{i=1}^n (\alpha_2\lambda_2 X_i^{\alpha_2-1})^{\delta_{2i}} \exp(-\lambda_1 X_i^{\alpha_1} - \lambda_2 X_i^{\alpha_2}). \end{aligned}$$

As was pointed out in Section 4.1, the joint likelihood factorizes into likelihoods for one-sample Weibull likelihood of Section 3.2. Thus, all likelihood quantities (log-likelihood, score vectors, MLEs, and information matrix) for each cause are the same as in Section 3.2.2. The k^{th} cause-specific quantile residual life, Q_{k,t_0} , has to be determined numerically using Equation 4.2.22 where f_k and $S(t_0)$ in the equation are

$$f_k(t;\alpha_1,\lambda_1,\alpha_2,\lambda_2) = \exp(-\lambda_1 u^{\alpha_1} - \lambda_2 u^{\alpha_2})\alpha_k\lambda_k u^{\alpha_k-1}du \quad , k = 1,2, \tag{4.2.32}$$

and

$$S(t_0;\alpha_1,\lambda_1,\alpha_2,\lambda_2) = \exp(-\lambda_1 t_0^{\alpha_1} - \lambda_2 t_0^{\alpha_2}). \tag{4.2.33}$$

The gradient of the cause-specific quantile residual life is found using Equation 4.2.23 where the elements of the gradient of $F_1(t;.)$ are

$$\begin{aligned}
\frac{\partial F_1(t;.)}{\partial \lambda_1} &= \int_0^t \alpha_1 u^{\alpha_1-1}\left[1-\lambda_1 u^{\alpha_1}\right]\exp(\alpha_1\exp(-\lambda_1 u^{\alpha_1}-\lambda_2 u^{\alpha_2})\,du \\
\frac{\partial F_1(t;.)}{\partial \lambda_2} &= \int_0^t -\lambda_1\alpha_1 u^{\alpha_1+\alpha_2-1}\exp(-\lambda_1 u^{\alpha_1}-\lambda_2 u^{\alpha_2})\,du \\
\frac{\partial F_1(t;.)}{\partial \alpha_1} &= \int_0^t \lambda_1 u^{\alpha_1-1}\left(1+\alpha_1\left[1-\lambda_1 u^{\alpha_1}\right]\left[\ln(u)\right]\right)\exp(-\lambda_1 u^{\alpha_1}-\lambda_2 u^{\alpha_2})\,du \\
\frac{\partial F_1(t;.)}{\partial \alpha_2} &= \int_0^t -\lambda_1\lambda_2\left[\ln(u)\right]\alpha_1 u^{\alpha_1+\alpha_2}\exp\left(-\lambda_1 u^{\alpha_1}-\lambda_2 u^{\alpha_2}\right)\,du,
\end{aligned}$$

and the elements of the gradient of $F(t,.)$ are,

$$\begin{aligned}
\frac{\partial F(t;.)}{\partial \lambda_k} &= t^{\alpha_k}\exp(-\lambda_1 t^{\alpha_1}-\lambda_2 t^{\alpha_2}) \\
\frac{\partial F(t;.)}{\partial \alpha_k} &= \lambda_k\left[\ln(t)\right]t^{\alpha_1}\exp(-\lambda_1 t^{\alpha_1}-\lambda_2 t^{\alpha_2}), \quad k=1,2.
\end{aligned}$$

4.2.3.2 Common shape parameter ($\alpha_1=\alpha_2$) Under the assumption that $\alpha_1=\alpha_2=\alpha$, the sub-distribution function (cumulative incidence function) for the failure time of type k is given by

$$F_k(t;\alpha,\lambda_1,\lambda_2)=\frac{\lambda_k}{\lambda_1+\lambda_2}(1-\exp(-[\lambda_1+\lambda_2]t^{\alpha})).$$

Thus, in this case, the quantile residual life and its gradient can be found analytically. Competing risks data can be simulated easily from this model using either Method 1 or Method 2.

The likelihood is

$$L(\alpha,\lambda_1,\lambda_2|\{(X_i,\delta_{ki})\}_{i=1}^n,k=1,2)=\prod_{i=1}^n(\alpha\lambda_1 X_i^{\alpha-1})^{\delta_{1i}}(\alpha\lambda_2 X_i^{\alpha-1})^{\delta_{2i}}\exp[-(\lambda_1+\lambda_2)X_i^{\alpha}],$$

whereas the log-likelihood is

$$\ell_n(\alpha,\lambda_1,\lambda_2)=\ln\lambda_1\sum_{i=1}^n\delta_{1i}+\ln\lambda_2\sum_{i=1}^n\delta_{2i}+\ln\alpha\sum_{i=1}^n(\delta_{1i}+\delta_{2i})+(\alpha-1)\sum_{i=1}^n(\delta_{1i}+\delta_{2i})\ln X_i-(\lambda_1+\lambda_2)\sum_{i=1}^n X_i^{\alpha}.$$

The score vector contains the following elements:

$$\frac{\partial}{\partial\alpha}\ell_n(\alpha,\lambda_1,\lambda_2)=\frac{\sum_{i=1}^n\delta_{1i}+\delta_{2i}}{\alpha}+\sum_i^n(\delta_{1i}+\delta_{2i})\ln(X_i)-(\lambda_1+\lambda_2)\sum_{i=1}^n X_i^\alpha\ln(X_i)$$

$$\frac{\partial}{\partial\lambda_k}\ell_n(\alpha,\lambda_1,\lambda_2)=\frac{\sum_{i=1}^n\delta_{ji}}{\lambda_k}-\sum_{i=1}^n X_i^\alpha,\qquad k=1,2.$$

The maximum likelihood estimator (MLE) of $\lambda_k\,(k=1,2)$ is

$$\hat{\lambda}_k=\frac{\sum_{i=1}^n\delta_{ki}}{\sum_{i=1}^n X_i^{\hat{\alpha}}},$$

whereas the MLE of α is a solution of the following non-linear equation:

$$\sum_{i=1}^n(\delta_{1i}+\delta_{2i})+\alpha\sum_{i=1}^n(\delta_{1i}+\delta_{2i})\ln(X_i)-\alpha\left(\sum_{i=1}^n\delta_{1i}+\delta_{2i}\right)\left(\sum_i^n X_i^\alpha\right)^{-1}\sum_{i=1}^n X_i^\alpha\ln(X_i)=0.$$

The elements of the the sample information matrix are:

$$I_{11}(\alpha,\lambda_1,\lambda_2)=\left[\alpha^{-2}\sum_{i=1}^n(\delta_{1i}+\delta_{2i})+(\lambda_1+\lambda_2)\sum_{i=1}^n X_i^\alpha(\ln(X_i))^2\right]$$

$$I_{22}(\alpha,\lambda_1,\lambda_2)=\left[\lambda_1^{-2}\sum_{i=1}^n\delta_{1i}\right]$$

$$I_{21}(\alpha,\lambda_1,\lambda_2)=I_{12}(\alpha,\lambda_1,\lambda_2)=\sum_{i=1}^n X_i^\alpha\ln(X_i)$$

$$I_{33}(\alpha,\lambda_1,\lambda_2)=\left[\lambda_2^{-2}\sum_{i=1}^n\delta_{2i}\right]$$

$$I_{31}(\alpha,\lambda_1,\lambda_2)=I_{13}(\alpha,\lambda_1,\lambda_2)=\sum_{i=1}^n X_i^\alpha\ln(X_i)$$

$$I_{32}(\alpha,\lambda_1,\lambda_2)=I_{23}(\alpha,\lambda_1,\lambda_2)=0.$$

The quantile residual life time function for the sub-distribution of failure time of type 1 is:

$$Q_{1,t_0}(\tau;\alpha,\lambda_1,\lambda_2)=\left[t_0^\alpha-[\lambda_1+\lambda_2]^{-1}\ln\left(1-\frac{\tau(\lambda_1+\lambda_2)}{\lambda_1}\right)\right]^{1/\alpha}-t_0,$$

which, upon letting $\lambda_1=\lambda$ and $\lambda_2=0$, is reduced to $Q_{t_0}(\tau;\alpha,\lambda_1)$, the quantile residual life of the Weibull model. The maximum likelihood estimator of $Q_{1,t_0}(\tau;\alpha,\lambda_1,\lambda_2)$ is $Q_{1,t_0}(\tau;\hat{\alpha},\hat{\lambda}_1,\hat{\lambda}_2)$.

The gradient vector of the MLE of the quantile residual life function contains the following entries,

$$\frac{\partial}{\partial\hat{\alpha}}Q_{1,t_0}(\tau;\alpha,\hat{\lambda}_1,\hat{\lambda}_2) = \hat{\alpha}^{-2}\left(t_0^{\hat{\alpha}} - \tilde{\lambda}^{-1}\ln[1-\tau\exp(-\tilde{\lambda})]\right)^{1/\hat{\alpha}} \left(\frac{\hat{\alpha}t_0^{\hat{\alpha}}\ln(t_0)}{\left\{t_0^{\hat{\alpha}} - \tilde{\lambda}\ln[1-\tau\exp(-\tilde{\lambda})]\right\}} - \ln\left\{t_0^{\hat{\alpha}} - \tilde{\lambda}\ln[1-\tau\exp(-\tilde{\lambda})]\right\}\right)$$

$$\frac{\partial}{\partial\hat{\lambda}_1}Q_{1,t_0}(\tau;\hat{\lambda}_1,\hat{\lambda}_2) = \hat{\alpha}^{-1}\hat{\lambda}_2\tilde{\lambda}^{-2}\left\{t_0^{\hat{\alpha}} - \tilde{\lambda}^{-1}\ln[1-\tau\exp(-\tilde{\lambda})]\right\}^{-1+1/\hat{\alpha}}\ln[1-\tau\exp(-\tilde{\lambda})]$$

$$\frac{\partial}{\partial\hat{\lambda}_2}Q_{1,t_0}(\tau;\hat{\lambda}_1,\hat{\lambda}_2) = \hat{\alpha}^{-1}\hat{\lambda}_1\tilde{\lambda}^{-2}\left\{t_0^{\hat{\alpha}} - \tilde{\lambda}^{-1}\ln[1-\tau\exp(-\tilde{\lambda})]\right\}^{-1+1/\hat{\alpha}}\ln[1-\tau\exp(-\tilde{\lambda})],$$

where $\tilde{\lambda} = \hat{\lambda}_1 + \hat{\lambda}_2$.

The estimator of asymptotic variance of the MLE of the quantile (residual) life function is given by:

$$\left\{\nabla Q_{1,t_0}(\tau;\hat{\alpha},\hat{\lambda}_1,\hat{\lambda}_2)\right\}^T I^{-1}(\hat{\alpha},\hat{\lambda}_1,\hat{\lambda}_2)\left\{\nabla Q_{1,t_0}(\tau;\hat{\alpha},\hat{\lambda}_1,\hat{\lambda}_2)\right\}.$$

4.2.4 Simulation Study

Simulation Method 2 was used to generate the competing risks data. The all cause failure time T was generated from the Weibull$(2, \lambda = \lambda_1 + \lambda_2)$, where λ_k is the k^{th} cause scale parameter. Given an observed time $T = t$ from the all cause failure time distribution, the failure time is a failure of type 1 with probability $\frac{\lambda_1}{\lambda_1+\lambda_2}$ and of a failure type 2 with probability $\frac{\lambda_2}{\lambda_1+\lambda_2}$. Two λ levels, $\lambda = 0.5$ with $\lambda_1 = 0.375$ and $\lambda_2 = 0.125$, and $\lambda = 0.25$ with $\lambda_1 = 0.1875$ and $\lambda_2 = 0.0625$; two quantile levels $\tau = 0.25$, 0.5 ; two time levels $t_0 = 0$, 1; and two sample sizes $n = 100$, 400 were used. The censoring proportion was set at 0.1, where the censoring times were generated from uniform distribution. 5000 data sets were simulated at each simulation setting.

Simulation results are in shown in Tables 4.2.1-4.2.2. In each of the tables, the estimated quantile residual life time for cause 1, its estimated asymptotic variance, its empirical variance, empirical bias, empirical MSE and 95% coverage probability of the true quantile residual life time for cause 1 are shown.

The results indicate that the procedure performs well under all simulation settings. The asymptotic and Monte Carlo variances are similar to each other, and the cause 1 quantile residual life time estimator is unbiased and has small MSE. The empirical coverage probability is close to the nominal coverage probability (95%) .

Table 4.2.1: One-sample simulation result($\alpha = 2$, $\lambda = 0.375$ cause 1 and $\lambda = 0.125$ cause 2)

n	t_0	τ	$\hat{Q}_{1,t_0}$	$\widehat{Avar(\hat{Q}_{1,t_0})}$	$\widehat{var(\hat{Q}_{1,t_0})}$	$\widehat{Bias}$	$\widehat{MSE}$	CP
400	0	0.25	0.902	0.001	0.001	0.001	0.001	0.949
		0.50	1.484	0.003	0.003	0.002	0.003	0.951
	1	0.25	0.346	0.000	0.000	0.000	0.000	0.949
		0.50	0.790	0.003	0.003	0.002	0.003	0.946
100	0	0.25	0.906	0.006	0.006	0.005	0.006	0.941
		0.50	1.489	0.014	0.014	0.007	0.014	0.953
	1	0.25	0.347	0.002	0.002	0.001	0.002	0.944
		0.50	0.793	0.011	0.011	0.005	0.011	0.943

Table 4.2.2: One-sample simulation result($\alpha = 2$, $\lambda = 0.1875$ cause 1 and $\lambda = 0.0625$ cause 2)

n	t_0	τ	$\hat{Q}_{1,t_0}$	$\widehat{Avar(\hat{Q}_{1,t_0})}$	$\widehat{var(\hat{Q}_{1,t_0})}$	$\widehat{Bias}$	$\widehat{MSE}$	CP
400	0	0.25	1.277	0.003	0.003	0.004	0.003	0.950
		0.50	2.101	0.007	0.007	0.005	0.007	0.949
	1	0.25	0.621	0.001	0.001	0.002	0.001	0.945
		0.50	1.326	0.006	0.005	0.004	0.005	0.946
100	0	0.25	1.277	0.011	0.011	0.004	0.011	0.945
		0.50	2.099	0.028	0.028	0.002	0.028	0.947
	1	0.25	0.620	0.004	0.004	0.001	0.004	0.943
		0.50	1.324	0.023	0.023	0.001	0.023	0.940

4.2.5 Data Analysis

In Chapter 3, we used death from all causes (breast cancer or non-cancer causes) as the survival endpoint. The B-04 data set contains information on cause of death (i.e, death due to breast cancer or non-cancer causes) for each subject. We will use this additional information to perform breast-cancer related death quantile residual life analysis in the presence of other causes of death using the proposed method in this chapter. Summary statistics on mortality and censoring by node status are shown in Table 4.2.3. The mortality rate due to cancer in the node-positive patients is higher than in the node-negative patients. The mortality rate due to non-cancer death is lower in the node-positive patients than in the node-negative patients. The observed censoring proportion in the node positive group is about 0.16, whereas in the node negative group is about 0.26.

Table 4.2.3: Summary statistics on mortality and censoring by node status

Node	Event Type	n	Prop.
Positive	Censored	92	0.16
	Death (cancer)	332	0.59
	Death (non-cancer)	137	0.24
Negative	Censored	271	0.26
	Death (cancer)	431	0.42
	Death (non-cancer)	336	0.32

For the one-sample competing risks quantile residual life analysis, we treated breast cancer related death as the primary event of interest and non-cancer death as the competing event. We performed separate analyses for each nodal group. We fitted the Weibull model, where we assumed that the hazard of death due to each cause could by independently modeled by the Weibull hazard function. The results for these analyses are shown in Table 4.2.4 for the node-negative patients and in Table 4.2.5 for the node-positive patients. Each table shows the point estimate of the quantile residual life for breast cancer related death, the estimate of the asymptotic variance of the breast cancer-related death quantile residual life, and the 95% confidence interval estimate of the breast cancer-related death quantile residual life at three time points $t_0 = 0$, 2, 4 and three quantile levels $\tau = 0.1$, 0.2, 0.3. The cause-specific quantile residual life estimates in Table 4.2.4 and Table 4.2.5 are within reasonable range of the non-parameteric cause-specific quantile residual life estimates reported in Table 4 of Jeong and Fine (2013).

Table 4.2.4: Quantile residual life for cancer-related death, node negative patient

t_0	τ	$\hat{Q}_{1,t_0}$	$\widehat{Avar}(\hat{Q}_{1,t_0})$	CI_{LL}	CI_{UL}
0	0.1	2.827	0.068	2.316	3.338
	0.2	6.679	0.221	5.757	7.601
	0.3	11.973	0.642	10.403	13.543
2	0.1	3.190	0.031	2.848	3.532
	0.2	7.351	0.172	6.538	8.164
	0.3	13.188	0.776	11.461	14.915
4	0.1	3.350	0.030	3.012	3.689
	0.2	7.721	0.204	6.836	8.606
	0.3	13.975	1.060	11.957	15.993

Table 4.2.5: Quantile residual life for non-cancer death, node positive patient

t_0	τ	$\hat{Q}_{1,t_0}$	$\widehat{Avar}(\hat{Q}_{1,t_0})$	CI_{LL}	CI_{UL}
0	0.1	1.070	0.020	0.791	1.349
	0.2	2.732	0.075	2.195	3.270
	0.3	4.991	0.184	4.152	5.831
2	0.1	1.589	0.009	1.402	1.775
	0.2	3.626	0.051	3.181	4.070
	0.3	6.304	0.186	5.459	7.149
4	0.1	1.777	0.011	1.572	1.983
	0.2	4.018	0.068	3.507	4.529
	0.3	6.962	0.263	5.956	7.968

4.3 TWO SAMPLE

In this section, we will generalize the result for the one-sample inference in Section 4.2 to the case of two independent samples. The result presented in this section can be easily extended to more than two groups. Except for the number of groups, the setting is similar to that of the one-sample case. In Section 4.3.1 basic results are given. In Section 4.2.4, simulation results using the Weibulll model are presented, and in Section 4.3.3 proposed method is used for data analysis.

4.3.1 Theory

The observed data is as follows:

A random sample of competing risks data from population 1, $\{(X_{1i}, \delta_{1ki})\}_{i=1}^{n_1}$, k=1,2 type of failure. $\delta_{1ki} = 1$ if the failure type is k, and $T_{1i} = X_{1i}$. Otherwise, $\delta_{1ki} = 0$. If $\delta_{11i} = \delta_{12i} = 0$, then the i^{th} observation is censored. $T_1 \sim F(;\theta^{(1)})$.

We also have a random sample of competing risks data from population 2, $\{(X_{2i}, \delta_{2ki})\}_{i=1}^{n_2}$, k=1,2 type of failure. $\delta_{2ki} = 1$ if the failure type is k, and $T_{2i} = X_{2i}$. Otherwise, $\delta_{2ki} = 0$. If $\delta_{21i} = \delta_{22i} = 0$, then the i^{th} observation is censored. $T_2 \sim F(;\theta^{(2)})$. No assumption about the functional dependency between $\theta^{(1)}$ and $\theta^{(2)}$ is made. Using the observed data, inference on the difference between the true cause 1 quantile residual life times of the two populations is considered,

$$Q_{1,t_0}^{(1)}(\tau;\theta^{(1)}) - Q_{1,t_0}^{(2)}(\tau;\theta^{(2)}). \tag{4.3.11}$$

Likelihood based point and confidence interval estimators are given below. The likelihood for the observed competing risks data is:

$$L(\theta \mid \{(X_{ji}, \delta_{jki})\}_{i=1}^{n}, j = 1,2, k = 1,2) = \prod_{j=1}^{2}\prod_{i=1}^{n_j} f_1(X_{ji};\theta^{(j)})^{\delta_{j1i}} f_2(X_{ji};\theta^{(j)})^{\delta_{j2i}} S(X_{ji};\theta^{(j)}),$$

where $\theta = (\theta^{(1)}, \theta^{(2)})$ is a vector of parameter vectors $\theta^{(1)}$ and $\theta^{(2)}$ for the distributions of T_{1i} and T_{2i}, respectively. The joint likelihood (Equation 4.3.11) factorizes into the likelihood of each group. As a result, the likelihood quantities and inference follow directly from the

one-sample case. Since the MLEs of $\theta^{(1)}$ and $\theta^{(2)}$, $\hat{\theta}^{(1)}$ and $\hat{\theta}^{(1)}$, respectively, are consistent and the cause-1 quantile residual life of each group are differentiable,

$$\hat{Q}^{(1)}_{1,t_0}(\tau;\hat{\theta}^{(1)}) - Q^{(2)}_{1,t_0}(\tau;\hat{\theta}^{(2)}) \tag{4.3.12}$$

is a consistent estimator of $Q^{(1)}_{1,t_0}(\tau;\theta^{(1)}) - Q^{(2)}_{1,t_0}(\tau;\theta^{(2)})$. In Equation 4.3.12, the MLE $\hat{\theta}^{(1)}$ and $\hat{\theta}^{(1)}$ are estimated separately from each sample.
The asymptotic variance formula for $Q^{(1)}_{1,t_0}(\tau;\theta^{(1)}) - Q^{(2)}_{1,t_0}(\tau;\theta^{(2)})$ is:

$$Avar(\hat{Q}^{(1)}_{1,t_0}) + Avar(\hat{Q}^{(2)}_{1,t_0}) \tag{4.3.13}$$

where $Avar(\hat{Q}^{(j)}_{t_0})$

$$= \frac{1}{n_j}\left\{\nabla_{\hat{\theta}^{(j)}} Q_{1,t_0}(\tau;\hat{\theta}^{(j)})_{|\hat{\theta}^{(j)}=\theta^{(j)}}\right\}^T \mathcal{I}^{-1}(\theta^{(j)}) \left\{\nabla_{\hat{\theta}^{(j)}} Q_{t_0}(\tau;\hat{\theta}^{(j)})_{|\hat{\theta}^{(j)}=\theta^{(j)}}\right\} \quad j = 1, 2$$

which are consistently estimated using MLEs of $\theta^{(1)}$ and $\theta^{(2)}$ and the observed information matrices from the two samples using the same procedure presented in Section 4.2. The results in Section 4.2 are directly applicable to the two sample situation. MLEs and asymptotic variance formulas can be derived using the results in Sections 4.2.

4.3.2 Simulation Study

A simulation study was performed to examine the performance of the proposed two-sample procedure. The simulation was run as in the one-sample case except for the following differences: two samples were independently generated from the same population at each simulation run; sample sizes of $n = 100$ and $n = 200$ per sample were used.

Simulation results for the two-sample case are presented in Table 4.3.1 and Table 4.3.2. In each of the tables, the estimated difference in cause 1 quantile residual lifetime($\Delta\hat{Q}_{1,t_0}$) of the two groups, estimated asymptotic variance, empirical variance, empirical bias, empirical MSE, and 95% empirical coverage (CP) are shown.

The simulation results suggest that the proposed method works well for all the simulation scenarios considered. The asymptotic and Monte Carlo variances are similar to each other indicating that the model-based asymptotic variance formula performs well. The quantile

residual life time estimator is unbiased and has small MSE. The empirical coverage probabilities are close to the 95% nominal coverage probability.

Table 4.3.1: Two-sample simulation result, Weibull($\alpha = 2$, $\lambda = 0.375$ cause 1 and $\lambda = 0.125$ cause 2)

n	t_0	τ	$\Delta\hat{Q}_{1,t_0}$	$\widehat{Avar(\Delta\hat{Q}_{1,t_0})}$	$\widehat{var(\Delta\hat{Q}_{1,t_0})}$	$\widehat{Bias}$	$\widehat{MSE}$	CP
200	0	0.25	0.000	0.006	0.006	0.000	0.006	0.950
		0.50	0.001	0.014	0.014	0.001	0.014	0.959
	1	0.25	0.000	0.002	0.002	0.000	0.002	0.955
		0.50	0.001	0.010	0.010	0.001	0.010	0.960
100	0	0.25	-0.002	0.011	0.011	-0.002	0.011	0.951
		0.50	-0.004	0.029	0.028	-0.004	0.028	0.958
	1	0.25	-0.001	0.003	0.003	-0.001	0.003	0.951
		0.50	-0.003	0.022	0.021	-0.003	0.021	0.961

Table 4.3.2: Two-sample simulation result, Weibull($\alpha = 2$, $\lambda = 0.1875$ cause 1 and $\lambda = 0.0625$ cause 2)

n	t_0	τ	$\Delta\hat{Q}_{1,t_0}$	$\widehat{Avar(\Delta\hat{Q}_{1,t_0})}$	$\widehat{var(\Delta\hat{Q}_{1,t_0})}$	$\widehat{Bias}$	$\widehat{MSE}$	CP
200	0	0.25	-0.001	0.011	0.011	-0.001	0.011	0.945
		0.50	-0.000	0.027	0.027	-0.000	0.027	0.954
	1	0.25	0.000	0.004	0.004	0.000	0.004	0.952
		0.50	0.001	0.023	0.022	0.001	0.022	0.955
100	0	0.25	0.000	0.022	0.023	0.000	0.023	0.942
		0.50	-0.002	0.057	0.058	-0.002	0.058	0.956
	1	0.25	-0.001	0.009	0.009	-0.001	0.009	0.948
		0.50	-0.003	0.048	0.048	-0.003	0.048	0.962

4.3.3 Data Analysis

To apply the two-sample result in this section to the B-04 study data, we used nodal status as a grouping variable and breast cancer-related death as the primary event of interest. Death from other causes was regarded as the competing event. The difference in the quantile residual life due to death from breast cancer between the two nodal groups was estimated at three time points $t_0 = 0$, 2, 4 and three quantile levels $\tau = 0.1$, 0.2, 0.3. The results are shown in Table 4.3.3. The results indicate that the nodal positive group has a shorter remaining life time than the nodal negative group for death due to breast-cancer.

Table 4.3.3: Difference in quantile residual life for breast-cancer death between node negative and node positive patient

t_0	τ	$\Delta\hat{Q}_{1,t_0}$	$\widehat{Avar}(\Delta\hat{Q}_{1,t_0})$	CI_{LL}	CI_{UL}
0	0.10	1.757	0.088	1.175	2.339
	0.20	3.947	0.297	2.879	5.014
	0.30	6.982	0.825	5.202	8.762
2	0.10	1.602	0.040	1.212	1.991
	0.20	3.725	0.224	2.798	4.652
	0.30	6.884	0.962	4.962	8.807
4	0.10	1.573	0.041	1.176	1.969
	0.20	3.703	0.272	2.681	4.725
	0.30	7.013	1.323	4.758	9.268

4.4 REGRESSION - ACCELERATED FAILURE TIME MODEL

The goal of this section is to extend the quantile residual inference for the one-sample case to the regression case while staying within the accelerated failure time (AFT) regression and the cause-specific hazard modeling framework.

4.4.1 Theory

In competing risks study, in addition to data on failure time and failure type, data on baseline covariates $\mathbf{Z} = (Z_1, Z_2, \cdots, Z_p)$ are collected with the object of associating these covariates to the competing risks process. The effect of covariate $\mathbf{Z}$ on the k^{th} cause hazard function $h_k(t; \mathbf{Z})$ can be modelled using AFT model as

$$h_k(t; \mathbf{Z}) = h_{0k}\{t \exp(\mathbf{Z}^T \beta_k)\} \exp(\mathbf{Z}^T \beta_k), \tag{4.4.11}$$

for $k = 1, 2$ causes of failure (Prentice et al. (1978)). The k^{th} cause baseline hazard $h_{0k}(t)$ can be specified as in the case of the one-sample competing risks setting. This model allows the effect of the covariate to vary by cause.

Consider a random sample of competing risks data $\{(X_i, \delta_{ki}, \mathbf{Z}_i)\}_{i=1}^n$, where $\mathbf{Z}_i = (Z_{i1}, Z_{i2}, \cdots, Z_{ip})$ is the vector of covariates measured on the i^{th} subject. The likelihood based on this competing risks data is given by

$$L(\theta \mid \{(X_i, \delta_{ki}, \mathbf{Z}_i)\}_{i=1}^n) = \prod_{i=1}^n f_1(X_i; \theta, \mathbf{Z}_i)^{\delta_{1i}} f_2(X_i; \theta, \mathbf{Z}_i)^{\delta_{2i}} S(X_i; \theta, \mathbf{Z}_i)^{1-(\delta_{1i}+\delta_{2i})},$$

where $\theta = (\theta_1, \theta_2)$ is the total parameter vector for causes 1 and 2; and $\theta_k = (\mathbf{b}_k, \beta_k)$ where $\mathbf{b}_k$ is a parametric vector for the k^{th} cause baseline hazard and β_k is the regression parameter vector for the k^{th}-cause hazard function.

Because the parameters for each cause are independent of the parameters of the other cause, the likelihood factorizes into the likelihood contribution for each cause. Thus, likelihood quantities (log-likelihood, score vectors, MLEs, and information) for each cause are exactly as those obtained in the AFT regression model for the non-competing risks setting.

4.4.2 Computational Issues

Numerical evaluation of the k^{th} cause-specific quantile residual life $Q_{k,t_0}(\tau;\theta,\mathbf{Z})$ and its gradient can be obtained by using the procedure in Section 4.2.2 as the covariate $\mathbf{Z}$ is known. Simulation of competing risks data under the AFT assumption can also be performed using the simulation procedures outlined there.

4.4.3 Application to the Weibull Model

4.4.3.1 Independent shape and scale parameters Under the AFT assumption, the k^{th} cause-specific cumulative-hazard for the population with covariate $\mathbf{Z}$, $H(.;\mathbf{Z})$, is related to the k^{th} cause-specific cumulative hazard of baseline population (i.e. the population with $\mathbf{Z}=0$) as follows,

$$\begin{aligned} H_k(t;\mathbf{Z}) &= \int_0^t h_{0j}\left(u\exp\left(\mathbf{Z}^T\beta_k\right)\right)\exp(\mathbf{Z}^T\beta_k)\,du \\ &= H_{0k}\left(t\exp\left(\mathbf{Z}^T\beta_k\right)\right), \quad k=1,2. \end{aligned}$$

For the Weibull distribution, we have

$$\begin{aligned} H_{0k}\left(t\exp\left(\mathbf{Z}^T\beta_k\right)\right) &= \lambda_k\left(t\exp\left(\mathbf{Z}^T\beta_k\right)\right)^{\alpha_k} \\ &= \lambda_k\exp\left(\alpha_k\mathbf{Z}^T\beta_k\right)t^{\alpha_k}, \quad k=1,2. \end{aligned}$$

Thus, the k^{th} $(k=1,2)$ cause-specific cumulative incidence function for the population with the covariate $\mathbf{Z}$, $F_k(t;\alpha_1,\alpha_2,\lambda_1,\lambda_2,\mathbf{Z})$, is

$$\int_0^t \exp\left(-\lambda_1\exp\left(\alpha_1\mathbf{Z}^T\beta_1\right)u^{\alpha_1}-\lambda_2\exp\left(\alpha_2\mathbf{Z}^T\beta_2\right)u^{\alpha_2}\right)\lambda_k\exp\left(\alpha_k\mathbf{Z}^T\beta_k\right)\alpha_k u^{\alpha_k-1}\,du. \tag{4.4.31}$$

The k^{th} cumulative incidence function (Equation 4.4.31) does not have a closed form except in the case of common shape parameter, i.e., $\alpha_1=\alpha_2=\alpha$. Thus, in the case of independent shape parameters, the k^{th} quantile residual life time and its gradient should be evaluated numerically using the procedure discussed in Section 4.2.2.1 and Section 4.2.2.2, respectively.

The cause 1 quantile residual life for the subject with covariate Z, $Q_{k,t_0}(\tau;\mathbf{Z})$ has to be determined numerically using Equation 4.2.22 where $f_k(u)$ and $S(t_o)$ in the equation are

$$f_k(t;\alpha_1,\alpha_2,\lambda_1,\lambda_2,\mathbf{Z}) = \exp\left(-\lambda_1 \exp\left(\alpha_1\mathbf{Z}^T\beta_1\right) u^{\alpha_1} - \lambda_2 \exp\left(\alpha_2\mathbf{Z}^T\beta_2\right) u^{\alpha_2}\right) \lambda_j \exp\left(\alpha_k\mathbf{Z}^T\beta_k\right) \alpha_k u^{\alpha_k-1},$$

and

$$S(t_0;\alpha_1,\alpha_2,\lambda_1,\lambda_2,\mathbf{Z}) = \exp\left(-\lambda_1 \exp\left(\alpha_1\mathbf{Z}^T\beta_1\right) t_0^{\alpha_1} - \lambda_2 \exp\left(\alpha_2\mathbf{Z}^T\beta_2\right) t_0^{\alpha_2}\right).$$

The gradient of the cause 1 quantile residual life given covariate Z is found using Equation 4.2.23, where the elements of the gradient of $F_1(t;.)$ are

$$\begin{aligned}
\frac{\partial F_1(t;\mathbf{Z})}{\partial\lambda_1} &= \int_0^t \alpha_1 u^{\alpha_1-1}\left[1-\lambda_1^\star u^{\alpha_1}\right]\exp(\alpha_1\mathbf{Z}^T\beta_1)\exp(-\lambda_1^\star u^{\alpha_1}-\lambda_2^\star u^{\alpha_2})\,du\\
\frac{\partial F_1(t;\mathbf{Z})}{\partial\lambda_2} &= \int_0^t -\lambda_1^\star\alpha_1 u^{\alpha_1+\alpha_2-1}\exp(\alpha_1\mathbf{Z}^T\beta_1+\alpha_2\mathbf{Z}^T\beta_2)\exp(-\lambda_1^\star u^{\alpha_1}-\lambda_2^\star u^{\alpha_2})\,du\\
\frac{\partial F_1(t;\mathbf{Z})}{\partial\alpha_1} &= \int_0^t \lambda_1^\star u^{\alpha_1-1}\left(1+\alpha_1\left[1-\lambda_1^\star u^{\alpha_1}\right]\left[\mathbf{Z}^T\beta_1+\ln(u)\right]\right)\exp(-\lambda_1^\star u^{\alpha_1}-\lambda_2^\star u^{\alpha_2})\,du\\
\frac{\partial F_1(t;\mathbf{Z})}{\partial\alpha_2} &= \int_0^t -\lambda_1^\star\lambda_2^\star\left[\mathbf{Z}^T\beta_2+\ln(u)\right]\alpha_1 u^{\alpha_1+\alpha_2}\exp\left(-\lambda_1^\star u^{\alpha_1}-\lambda_2^\star u^{\alpha_2}\right)\,du\\
\frac{\partial F_1(t;\mathbf{Z})}{\partial\beta_k} &= \lambda_k\alpha_k\frac{\partial F_1(t;\mathbf{Z})}{\partial\lambda_k}\mathbf{Z}, \quad k=1,\,2,
\end{aligned}$$

and the elements of the gradient of all cause distribution function $F(t,.)$ are,

$$\begin{aligned}
\frac{\partial F(u;\mathbf{Z})}{\partial\lambda_k} &= u^{\alpha_k}\exp(\alpha_k\mathbf{Z}^T\beta_k)\exp(-\lambda_1^\star u^{\alpha_1}-\lambda_2^\star u^{\alpha_2})\\
\frac{\partial F(u;\mathbf{Z})}{\partial\alpha_k} &= \lambda_k^\star\left[\mathbf{Z}^T\beta_k+\ln(u)\right]u^{\alpha_k}\exp(-\lambda_1^\star u^{\alpha_1}-\lambda_2^\star u^{\alpha_2})\\
\frac{\partial F(u;\mathbf{Z})}{\partial\beta_k} &= \lambda_k\alpha_k\frac{\partial F(u;\mathbf{Z})}{\partial\lambda_k}\mathbf{Z}, \quad k=1,\,2,
\end{aligned}$$

where $\lambda_k^\star = \lambda_k\exp(\alpha_k\mathbf{Z}^T\beta_k), \quad k=1,2.$

4.4.3.2 Common shape parameter ($\alpha_1 = \alpha_2$)

The MLEs for the Weibull-model with common alpha are obtained by finding the root of the score vector whose elements are,

$$\frac{\partial}{\partial\alpha}\ell_n(\theta) = \sum_{i=1}^{n}\delta_{1i}\mathbf{Z}_i^T\beta_1 + (\alpha^{-1})\sum_{i=1}^{n}\delta_{1i} + \sum_{i=1}^{n}\delta_{1i}\ln(X_i) + \sum_{i=1}^{n}\delta_{2i}\mathbf{Z}_i^T\beta_2 + (\alpha^{-1})\sum_{i=1}^{n}\delta_{2i} + \sum_{i=1}^{n}\delta_{2i}\ln(X_i)$$
$$-\lambda_1\left[\sum_{i=1}^{n}X_i^\alpha\exp(\alpha\mathbf{Z}_i^T\beta_1)(\mathbf{Z}_i^T\beta_1 + \ln(X_i))\right] - \lambda_2\left[\sum_{i=1}^{n}X_i^\alpha\exp(\alpha\mathbf{Z}_i^T\beta_2)(\mathbf{Z}_i^T\beta_2 + \ln(X_i))\right]$$

$$\frac{\partial}{\partial\lambda_1}\ell_n(\theta) = (\lambda_1^{-1})\sum_{i=1}^{n}\delta_{1i} - \sum_{i=1}^{n}X_i^\alpha\exp(\alpha\mathbf{Z}_i^T\beta_1)$$

$$\frac{\partial}{\partial\lambda_2}\ell_n(\theta) = (\lambda_2^{-1})\sum_{i=1}^{n}\delta_{2i} - \sum_{i=1}^{n}X_i^\alpha\exp(\alpha\mathbf{Z}_i^T\beta_2)$$

$$\frac{\partial}{\partial\beta_1}\ell_n(\theta) = \alpha\sum_{i=1}^{n}\left[\delta_{1i} - \lambda_1\alpha X_i^\alpha\exp(\alpha\mathbf{Z}_i^T\beta_1)\right]\mathbf{Z}_i$$

$$\frac{\partial}{\partial\beta_2}\ell_n(\theta) = \alpha\sum_{i=1}^{n}\left[\delta_{2i} - \lambda_2\alpha X_i^\alpha\exp(\alpha\mathbf{Z}_i^T\beta_2)\right]\mathbf{Z}_i.$$

The information matrix contains the following entries,

$$\frac{\partial^2 \ell_n}{\partial^2 \lambda_k} = -\lambda_k^2 \sum_{i=1}^{n} \delta_{1i}, \quad k = 1, 2$$

$$\frac{\partial^2 \ell_n}{\partial \lambda_1 \partial \lambda_2} = 0$$

$$\frac{\partial^2 \ell_n}{\partial \alpha \partial \lambda_k} = -\sum_{i=1}^{n} (\mathbf{Z}_i^T \beta_k + \ln(X_i)) X_i^{\alpha} \exp(\alpha \mathbf{Z}_i^T \beta_k), \quad k = 1, 2$$

$$\frac{\partial^2 \ell_n}{\partial \lambda_1 \partial \beta_1} = -\sum_{i=1}^{n} \alpha X_i^{\alpha} \exp(\alpha \mathbf{Z}_i^T \beta_1) \mathbf{Z}_i$$

$$\frac{\partial^2 \ell_n}{\partial \lambda_2 \partial \beta_2} = -\sum_{i=1}^{n} \alpha X_i^{\alpha} \exp(\alpha \mathbf{Z}_i^T \beta_2) \mathbf{Z}_i$$

$$\frac{\partial^2 \ell_n}{\partial \beta_2 \partial \lambda_1} = \frac{\partial^2 \ell_n}{\partial \beta_1 \partial \lambda_2} = 0$$

$$\frac{\partial^2 \ell_n}{\partial \beta_1 \partial \beta_2} = 0$$

$$\frac{\partial^2 \ell_n}{\partial \beta_k^T \partial \beta_k} = -\sum_{i=1}^{n} \lambda_k \exp(\alpha \mathbf{Z}_i^T \beta_1) \alpha^2 X_i^{\alpha} \mathbf{Z}_i \mathbf{Z}_i^T, \quad k = 1, 2$$

$$\frac{\partial^2 \ell_n}{\partial \alpha \partial \beta_k} = -\sum_{i=1}^{n} \left[\lambda_k \exp(\alpha \mathbf{Z}_i^T \beta_1) [\mathbf{Z}_i^T \beta_k + \ln(X_i)] \alpha X_i^{\alpha} + [\delta_{2i} - \lambda_k \exp(\alpha \mathbf{Z}_i^T \beta_1) \alpha X_i^{\alpha}] \right] \mathbf{Z}_i, \quad k = 1, 2$$

$$\begin{aligned} \frac{\partial^2 \ell_n}{\partial^2 \alpha} = & -\alpha^2 \sum_{i=1}^{n} (\delta_{1i} + \delta_{2i}) - \sum_{i=1}^{n} [\lambda_{1i}^{\star} (\mathbf{Z}_i^T \beta_1)^2 + \lambda_{2i}^{\star} \mathbf{Z}_i^T \beta_2)^2 + (\lambda_{1i}^{\star} + \lambda_{2i}^{\star}) (\ln(X_i))^2] X_i^{\alpha} \\ & -2 \sum_{i=1}^{n} [\lambda_{1i}^{\star} \mathbf{Z}_i^T \beta_1 + \lambda_{2i}^{\star} \mathbf{Z}_i^T \beta_2] X_i^{\alpha} \ln(X_i) \quad \text{where} \quad \lambda_{ki}^{\star} = \lambda_k \exp(\alpha \mathbf{Z}_i^T \beta_k), \quad k = 1, 2. \end{aligned}$$

The asymptotic variance formula for the estimator of the k^{th} quantile residual life given $\mathbf{Z}$ is given by,

$$\left\{ \nabla \hat{Q}_{k,t_0}(\tau; \mathbf{Z}) \right\}^T \mathcal{I}^{-1}(\theta) \left\{ \nabla \hat{Q}_{k,t_0}(\tau; \mathbf{Z}) \right\}. \tag{4.4.32}$$

The gradient of the k^{th} cause quantile residual life given covariate $\mathbf{Z}$, $\nabla \hat{Q}_{k,t_0}(\tau;\mathbf{Z})$, is found using Equation 4.2.23, where the elements of the gradient of $F_1(t;.)$ in closed form are

$$\begin{aligned}
\frac{\partial F_1(t;\mathbf{Z})}{\partial \lambda_1} &= \frac{\lambda_1^\star t^\alpha \exp(\alpha \mathbf{Z}^T\beta_1)}{(\lambda_1^\star+\lambda_2^\star)} \exp(-(\lambda_1^\star+\lambda_2^\star)t^\alpha) \\
\frac{\partial F_1(t;\mathbf{Z})}{\partial \lambda_2} &= \frac{-\lambda_1^\star \exp(\alpha \mathbf{Z}^T\beta_2)}{(\lambda_1^\star+\lambda_2^\star)^2} \{1-[1+(\lambda_1^\star+\lambda_2^\star)t^\alpha]\exp(-(\lambda_1^\star+\lambda_2^\star)t^\alpha)\} \\
\frac{\partial F_1(t;\mathbf{Z})}{\partial \alpha} &= \frac{-\lambda_1^\star \mathbf{Z}^T(\beta_1-\beta_2)}{(\lambda_1^\star+\lambda_2^\star)^2} [1-\exp(-(\lambda_1^\star+\lambda_2^\star)t^\alpha)] \quad + \\
&\qquad \frac{\lambda_1^\star}{\lambda_1^\star+\lambda_2^\star} \exp(-(\lambda_1^\star+\lambda_2^\star)t^\alpha) \left\{ \left[\lambda_1^\star \mathbf{Z}^T\beta_1 + \lambda_2^\star \mathbf{Z}^T\beta_2\right] t^\alpha + (\lambda_1^\star+\lambda_2^\star)t^\alpha \ln(t) \right\} \\
\frac{\partial F_1(t;\mathbf{Z})}{\partial \beta_1} &= \alpha\lambda_1 \frac{\partial F_1(t;\mathbf{Z})}{\partial \lambda_1}\mathbf{Z} \\
\frac{\partial F_1(t;\mathbf{Z})}{\partial \beta_2} &= \alpha\lambda_2 \frac{\partial F_1(t;\mathbf{Z})}{\partial \lambda_2}\mathbf{Z},
\end{aligned}$$

and the elements of the gradient of all cause distribution function $F(t,.)$ in closed form are,

$$\begin{aligned}
\frac{\partial F(u;\mathbf{Z})}{\partial \lambda_1} &= u^\alpha \exp(\alpha \mathbf{Z}^T\beta_1) \exp(-(\lambda_1^\star+\lambda_2^\star)u^\alpha) \\
\frac{\partial F(u;\mathbf{Z})}{\partial \lambda_2} &= u^\alpha \exp(\alpha \mathbf{Z}^T\beta_2) \exp(-(\lambda_1^\star+\lambda_2^\star)u^\alpha) \\
\frac{\partial F(u;\mathbf{Z})}{\partial \alpha} &= \lambda_1^\star \left[\mathbf{Z}^T\beta_1 + \ln(u)\right] u^\alpha \exp(-(\lambda_1^\star+\lambda_2^\star)u^\alpha) \quad + \\
&\qquad \lambda_2^\star \left[\mathbf{Z}^T\beta_2 + \ln(u)\right] u^\alpha \exp(-(\lambda_1^\star+\lambda_2^\star)u^\alpha) \\
\frac{\partial F(u;\mathbf{Z})}{\partial \beta_1} &= \alpha\lambda_1 \frac{\partial F(u;\mathbf{Z})}{\partial \lambda_1}\mathbf{Z} \\
\frac{\partial F(u;\mathbf{Z})}{\partial \beta_2} &= \alpha\lambda_2 \frac{\partial F(u;\mathbf{Z})}{\partial \lambda_2}\mathbf{Z},
\end{aligned}$$

where $\lambda_k^\star = \lambda_k \exp(\alpha \mathbf{Z}^T\beta_k), \quad k = 1,\, 2.$

4.4.4 Simulation Study

Simulation Method 2 was used to generate competing risks data from the Weibull AFT regression model with common shape parameter α where we used $(\alpha_1, \lambda_1, \beta_{11}, \beta_{12}) = (2, 1, 0.2, 0.15)$ for cause 1 and $(\alpha_2, \lambda_2, \beta_{21}, \beta_{22}) = (2, 0.5, -0.15, 0.2)$ for cause 2. We simulated a binary covariate Z_1 from $Bernoulli(0.5)$ distribution and a continuous covariate Z_2 from the the standard normal distribution. Two time points $t_0 = 0$, 1, two quantile levels $\tau = 0.1$, 0.3, and a sample size of $n = 200$ were used. The censoring proportion was set at 0.3, where censoring times were generated from uniform distributions. 5000 data sets were generated at each simulation setting. The true cause 1 quantile residual life at $\mathbf{Z} = (0, 0)$, $\mathbf{Z} = (1, 0)$, $\mathbf{Z} = (0, 1)$, and $\mathbf{Z} = (0, 0)$ were used to investigate the performance of the proposed procedure. The coverage probability (CP) is for 95% confidence intervals.

The results in Table 4.4.1 suggest that coverage probability is close to the nominal 95% coverage level for all simulation settings. The Monte Carlo variance is also similar to the asymptotic variance indicating that the model-based asymptotic variance formula works well under the simulation settings considered. The cause 1 quantile residual life estimator is unbiased.

Table 4.4.1: AFT competing risks regression, common shape parameter

(Z_1, Z_2)	t_0	τ	$\widehat{Avar(Q_{1,t_0})}$	$\widehat{var(\hat{Q}_{1,t_0})}$	$\widehat{Bias}$	CP
$(0,0)$	0	0.10	0.001	0.001	-0.006	0.949
	1		0.000	0.000	0.000	0.929
	0	0.30	0.003	0.003	-0.006	0.954
	1		0.001	0.001	0.000	0.931
$(1,0)$	0	0.10	0.001	0.001	-0.003	0.947
	1		0.000	0.000	0.000	0.931
	0	0.30	0.002	0.002	-0.002	0.948
	1		0.000	0.000	0.001	0.930
$(0,1)$	0	0.10	0.001	0.001	-0.006	0.945
	1		0.000	0.000	-0.000	0.929
	0	0.30	0.004	0.004	-0.008	0.952
	1		0.001	0.001	-0.002	0.928
$(1,1)$	0	0.10	0.001	0.001	-0.004	0.948
	1		0.000	0.000	0.000	0.930
	0	0.30	0.002	0.002	-0.004	0.952
	1		0.000	0.000	0.001	0.932

4.4.5 Data Analysis

For data analyses, we modeled the cause-specific hazards due to each cause (cancer and non-caner death) independently using the Weibull model. We used the variables node (a binary variable), age and tumor size as covariates for AFT Weibull regression modeling. Age and tumor size were centered around their overall means before the analyses. The baseline distribution is that of the node-positive patient at the mean age and mean tumor size.

The results in Table 4.4.2 indicate that bigger tumor size and positive node status significantly increase the hazard of death from breast cancer, whereas higher baseline age significantly increases the hazard of death from non-cancer causes. Higher baseline age increases non-significantly the hazard of death from cancer, and positive node status and bigger tumor size increase non-significantly the hazard of death from non-cancer causes.

Table 4.4.2: Parameter estimates for the cause-specific Weibull hazard functions

Cause of death	α	λ	Node	Age	Tumor Size
Cancer	0.869(0.027)	0.085(0.007)	-0.673(0.085)	0.002(0.004)*	0.009(0.002)
Non-cancer	1.468(0.055)	0.006(0.001)	-0.04(0.069)*	0.042(0.003)	0.002(0.002)*

*Non-significant at 0.05 significance level. Standard errors are given in parentheses.

For cause-specific quantile residual life analyses, we used breast-cancer death as the main event and non-cancer death as the competing event. Quantile residual life times for death from cancer were estimated for node positive patient at the median age (55 yrs) and tumor size (32 mm), and node negative patient at the median age (57 yrs) and tumor size (30 mm). The results in Table 4.4.3 and Table 4.4.4 are consistent with the results obtained for the one-sample analyses.

Table 4.4.3: Quantile residual life for death due to cancer for node negative with median age and tumor size

t_0	τ	$\hat{Q}_{1,t_0}$	$\widehat{Avar}(\hat{Q}_{1,t_0})$	CI_{LL}	CI_{UL}
0	0.10	2.638	0.192	1.779	3.497
	0.20	6.461	1.083	4.422	8.501
	0.30	11.829	4.174	7.825	15.833
2	0.10	3.191	0.190	2.336	4.046
	0.20	7.460	1.203	5.311	9.610
	0.30	13.597	5.448	9.022	18.172
4	0.10	3.433	0.206	2.543	4.324
	0.20	8.013	1.369	5.721	10.306
	0.30	14.784	6.994	9.600	19.967

Table 4.4.4: Quantile residual life for death due to cancer for node positive patient with median age and tumor size

t_0	τ	$\hat{Q}_{1,t_0}$	$\widehat{Avar}(\hat{Q}_{1,t_0})$	CI_{LL}	CI_{UL}
0	0.10	1.315	0.016	1.067	1.563
	0.20	3.151	0.062	2.661	3.640
	0.30	5.512	0.162	4.723	6.300
2	0.10	1.683	0.010	1.488	1.878
	0.20	3.765	0.052	3.316	4.213
	0.30	6.400	0.170	5.591	7.208
4	0.10	1.819	0.011	1.613	2.025
	0.20	4.045	0.061	3.560	4.530
	0.30	6.870	0.210	5.972	7.767

5.0 SUMMARY

In this dissertation, we proposed a parametric inference for quantile residual life for one-sample, two-sample and regression cases under both competing and non-competing risks settings. For the competing risks set-up, we adopted a cause-specific hazard modeling approach, where the hazard function for each cause is independently parameterized.

While this modeling approach is flexible, it raises computational challenges that have not been identified and addressed in the existing parametric competing risks literature. The main computational problems were lack of closed form expression for the cause-specific quantile residual life and its gradient, and the difficulty of simulating competing risks data. We proposed numerical solutions for determining the cause-specific quantile residual life and its gradient, and derived an asymptotic variance formula for the estimator of the cause-specific quantile residual life time. We also suggested a simulation procedure for generating competing risks data under the cause-specific hazard modeling approach.

We investigated the performance of the proposed procedures using simulation studies. The simulation results indicated that the proposed procedures work well. We applied the proposed methods to analyze a breast cancer survival data.

We used the Weibull model for both simulation and data analyses. We chose this model for illustration because it is a simple and flexible model. Under ideal circumstance, alternative models should be compared for their appropriateness for the application on hand.

BIBLIOGRAPHY

Beyersmann, Jan, Aurelien Latouche, Anika Buchholz, and Martin Schumacher (2009), "Simulating competing risks data in survival analysis." *Statistics in Medicine*, 28, 956–971.

Boos, Dennis D and LA Stefanski (2013), "Large sample theory: The basics." In *Essential Statistical Inference*, 207–274, Springer.

Borgan, Ørnulf (1984), "Maximum likelihood estimation in parametric counting process models, with applications to censored failure time data." *Scandinavian Journal of Statistics*, 1–16.

Casella, George and CP Robert (2004), "Monte carlo statistical methods." *Springer&Verlag, New York*.

Fisher, Bernard, Jong-Hyeon Jeong, Stewart Anderson, John Bryant, Edwin R Fisher, and Norman Wolmark (2002), "Twenty-five-year follow-up of a randomized trial comparing radical mastectomy, total mastectomy, and total mastectomy followed by irradiation." *New England Journal of Medicine*, 347, 567–575.

Fisher, Bernard, Eleanor Montague, Carol Redmond, Bruce Barton, Donna Borland, Edwin R Fisher, Melvin Deutsch, George Schwarz, Richard Margolese, William Donegan, et al. (1977), "Comparison of radical mastectomy with alternative treatments for primary breast cancer: a first report of results from a prospective randomized clinical trial." *Cancer*, 39, 2827–2839.

Jeong, J. H. and J. P. Fine (2013), "Nonparametric inference on cause-specific quantile residual life." *Biom J*, 55, 68–81.

Jeong, J-H and JP Fine (2009), "A note on cause-specific residual life." *Biometrika*, 96, 237–242.

Jeong, Jong-Hyeon and Jason Fine (2006), "Direct parametric inference for the cumulative incidence function." *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 55, 187–200.

Jeong, Jong-Hyeon and Jason P Fine (2007), "Parametric regression on cumulative incidence function." *Biostatistics*, 8, 184–196.

Jeong, Jong-Hyeon, Sin-Ho Jung, and Joseph P Costantino (2008), "Nonparametric inference on median residual life function." *Biometrics*, 64, 157–163.

Jung, Sin-Ho, Jong-Hyeon Jeong, and Hanna Bandos (2009), "Regression on quantile residual life." *Biometrics*, 65, 1203–1212.

Kalbfleisch, John D and Ross L Prentice (2002), *The statistical analysis of failure time data*, volume 360. John Wiley & Sons.

Kay, Richard and Nelson Kinnersley (2002), "On the use of the accelerated failure time model as an alternative to the proportional hazards model in the treatment of time to event data: a case study in influenza." *Drug information journal*, 36, 571–579.

Klein, John P and Melvin L Moeschberger (2003), *Survival analysis: techniques for censored and truncated data*. Springer Science & Business Media.

Koenker, Roger and Gilbert Bassett (1978), "Regression quantiles." *Econometrica: journal of the Econometric Society*, 33–50.

Lagakos, SW (1979), "General right censoring and its impact on the analysis of survival data." *Biometrics*, 139–156.

Lee, Minjung and Jason P Fine (2011), "Inference for cumulative incidence quantiles via parametric and nonparametric approaches." *Statistics in medicine*, 30, 3221–3235.

Lin, DY and CF Spiekerman (1996), "Model checking techniques for parametric regression with censored data." *Scandinavian journal of statistics*, 157–177.

Ng, SK and GJ McLachlan (2003), "An em-based semi-parametric mixture model approach to the regression analysis of competing-risks data." *Statistics in Medicine*, 22, 1097–1111.

Patel, Katie, Richard Kay, and Lucy Rowell (2006), "Comparing proportional hazards and accelerated failure time models: an application in influenza." *Pharmaceutical statistics*, 5, 213–224.

Prentice, Ross L, John D Kalbfleisch, Arthur V Peterson Jr, Nancy Flournoy, VT Farewell, and NE Breslow (1978), "The analysis of failure times in the presence of competing risks." *Biometrics*, 541–554.

Shi, Haiwen, Yu Cheng, and Jong-Hyeon Jeong (2013), "Constrained parametric model for simultaneous inference of two cumulative incidence functions." *Biometrical Journal*, 55, 82–96.

Wei, LJ (1992), "The accelerated failure time model: a useful alternative to the cox regression model in survival analysis." *Statistics in medicine*, 11, 1871–1879.