Chapter Four

Proportional Mean Residual Life Model for Gap Time Distributions of Recurrent Events

4.1. Introduction

In survival studies, it is often of interest to analyze the mean residual life function to characterize the stochastic behavior of survival over time. The mean residual life function $m(t)$ defined in (1.8) is interpreted as the average remaining lifetime of an individual given that the individual has survived up to time $t$. Although the hazard function, mean residual life function and survival function are in one-to-one correspondence with each other, Muth (1977) considered the mean residual life to be a superior concept than the hazard function on the following grounds:

a) Regarding the ageing phenomena the two concepts are not equivalent. A decreasing mean residual life does not imply an increasing hazard function, though the converse is true. Thus the decreasing mean residual life is more general in character.

b) The hazard function accounts only for the immediate future in assessing failure phenomenon as described by the derivative of $S(t)$, where as the latter is descriptive of the entire future implied through the integral of $S(t)$ over $t$ to $\infty$. A consequence of this is that a component may experience deterioration though its hazard function may be zero at a certain point.

c) It is advantageous to use the mean residual life function as a decision making criterion for replacement or maintenance policies. The expected remaining life of the component gives an indication of whether to replace or to re-schedule and

The results in this Chapter have been published as entitled “Proportional Mean Residual Life Model for Gap Time Distributions of Recurrent Events”, Metron, Vol. LXV, n.3 (see Sreeja and Sankaran (2007)).
this could be more useful than the hazard function to formulate maintenance policies.

Oakes and Dasu (1990) considered a proportional mean residual life model as an alternative to Cox (1972) proportional hazards model to assess the effects of covariates on the survival time. The model is defined by

\[ m(t \mid z) = m_0(t)e^{\beta'z} \]  

(4.1)

where \( m(t \mid z) \) is the mean residual life corresponding to the \( r \)-vector of covariates \( z \), \( m_0(t) \) is the unknown baseline mean residual life function when \( z = 0 \) and \( \beta \) is the vector of regression parameters. Generally, there is no direct relationship between the proportional mean residual life model and the Cox proportional hazards model. However, Oakes and Dasu (1990) proved that, when a model satisfies both the proportional hazards and the proportional mean residual life assumptions, its underlying distribution then belongs to the Hall-Wellner class of distributions with linear mean residual life function (Hall and Wellner, 1981).

Previous work on the mean residual life has focused on single-sample and two-sample cases; see Oakes and Dasu (2003). In regression analysis, Robins and Rotnitzky (1992) and Maguluri and Zang (1994) developed estimation procedures for the model (4.1) under uncensored and censored set up respectively. Recently, Chen and Cheng (2005) developed semi-parametric inference procedures for the regression model (4.1) using martingale theory of counting processes, in the presence of censoring. Later, Chen and Cheng (2006) considered a linear mean residual life model and developed inference procedures under right censoring. The analysis of gap times for recurrent event data using mean residual life function is more appropriate in many practical situations, as seen in Section 4.5. Motivated by this, we propose a bivariate proportional mean residual life model to assess the relationship between mean residual life function and covariates for gap time of recurrent events. Note that the focus will be on the development of the regression model of the duration times, when the recurrent events are of same type.

In Section 4.2, we introduce a bivariate proportional mean residual life model to assess the relationship between mean residual lifetime and covariates for gap time of recurrent events. Estimators of the parameter vectors as well as baseline
mean residual life function are discussed in Section 4.3. Asymptotic properties of
the estimators are studied. A simulation study is carried out to assess the
performance of the estimators in Section 4.4. In Section 4.5, we illustrate the
procedure using kidney dialysis data given in Lawless (2003). Conclusions of the
study is given in Section 4.6.

4.2. Bivariate Proportional Mean Residual Life Model

Suppose that an individual may experience \( k \) consecutive events at times
\( X_1 < X_2 < \ldots < X_k \). Let \( T_1, T_2, \ldots, T_k \) represents the gap times where \( T_1 = X_1, T_2 = X_2 - X_1 \) and \( T_k = X_k - X_{k-1} \). As in Chapter 3, we assume that the follow up
time is subject to independent right censoring by \( C \) which implied that
\( (X_1, X_2, \ldots X_k) \) are independent of \( C \). We now consider the regression problem in
which the marginal and conditional mean residual lifetime functions of \( (T_1, T_2, \ldots T_k) \)
depend on certain covariates. We consider the case where \( k = 2 \). The extension to
higher dimensions is direct.

Let \( S(t_1,t_2) = P[T_1 \geq t_1, T_2 \geq t_2] \) be the joint survival function of \( T_1 \) and \( T_2 \).
Let \( m_i(t_i) \) be the mean residual life function of \( T_i \), which is defined as
\[
m_i(t_i) = E[T_i - t_i | T_i \geq t_i].
\] (4.2)
In the context of recurrent events, \( m_i(t_i) \) can be interpreted as the expected
remaining gap time of \( T_i \) given that \( T_i \) is larger than or equal to \( t_i \). For the recurrent
events, the occurrence of the second event depends on the occurrence of the first
one. Accordingly, we can consider mean residual life function of \( T_2 \) given \( T_1 \geq t_1 \).
The mean residual life function of \( T_2 \) given \( T_1 \geq t_1 \) is defined as
\[
m_2(t_1,t_2) = E[T_2 - t_2 | T_1 \geq t_1, T_2 \geq t_2].
\] (4.3)
The expression (4.3) can be interpreted as the average remaining gap time of \( T_2 \)
given that \( T_1 \) is larger than or equal to \( t_1 \) and \( T_2 \) is larger than or equal to \( t_2 \). We use
the term mean residual life time for average remaining gap time through out this
chapter. Note that (4.3) is the second component of the vector MRL in the bivariate
set up, defined in Arnold and Zahedi (1988).
Then the survival function $S(t_1, t_2)$ can be determined from (4.2) and (4.3) by the identity

$$S(t_1, t_2) = \frac{m_1(0) m_2(t_1, 0)}{m_1(t_1) m_2(t_1, t_2)} \exp \left[ - \int_0^{t_1} \frac{du}{m_1(u)} - \int_0^{t_2} \frac{du}{m_2(t_1, u)} \right].$$

(4.4)

Note that the hazard functions given in (1.34) and (1.35), and mean residual life functions are related by the identities

$$\lambda_1(t) = \frac{1 + m_1'(t_1)}{m_1(t_1)}$$

(4.5)

and

$$\lambda_2(t_1, t_2) = \frac{1 + m_2'(t_1, t_2)}{m_2(t_1, t_2)}.$$  

(4.6)

where $m_1'(t_1)$ is the derivative of $m_1(t_1)$ with respect to $t_1$ and $m_2'(t_1, t_2)$ is the derivative for $m_2(t_1, t_2)$ with respect to $t_2$.

For the analysis of gap times of recurrent event data using MRL, one possible method is to consider marginal mean residual life functions of $T_1$ and $T_2$ and then apply ideas from generalized estimating functions to calculate an appropriate combination of the two marginal estimates. This can be done in the case of homogeneity among two regression coefficients. In many practical situations as shown in Section 4.4, the conditional mean residual life function of $T_2$ given $T_1 \geq t_1$ is meaningful than the marginal mean residual life function of $T_2$ to explain the joint dependence structure of pair of lifetimes on the covariate vector. Accordingly, we define a bivariate proportional mean residual life model for $T_1$ and $T_2$ as

$$m_1(t_1 | z) = m_{10}(t_1) e^{\beta_1^z}$$

(4.7)

and

$$m_2(t_1, t_2 | z) = m_{20}(t_1, t_2) e^{\beta_2^z}.$$  

(4.8)

In model (4.7), $m_1(t_1 | z)$ is the mean residual life function at time $t_1$ when the $r \times 1$ covariate vector $z$ is given and $m_{10}(t_1)$ is the baseline mean residual life function, which is the mean residual life function when $z = 0$. Here $\beta_1$ and $\beta_2$ are $r \times 1$
vector of parameters. For the model (4.7), the ratio $\frac{m_1(t_1 \mid \mathbf{z}^{(1)})}{m_1(t_1 \mid \mathbf{z}^{(2)})}$ of the mean residual life functions of two individuals with covariate vectors $\mathbf{z}^{(1)}$ and $\mathbf{z}^{(2)}$ does not vary with time $t_1$. The model (4.8) implies that the ratio $\frac{m_2(t_1, t_2 \mid \mathbf{z}^{(1)})}{m_2(t_1, t_2 \mid \mathbf{z}^{(2)})}$ of the mean residual life functions of two pairs of individuals with covariate vectors $\mathbf{z}^{(1)}$ and $\mathbf{z}^{(2)}$ does not vary with $t_1$ and $t_2$.

Using (4.5), (4.6), (4.7) and (4.8), we can have

$$m_{t_0}(t_1) \wedge_1 (dt_1) = e^{-\beta \mathbf{z}^T \mathbf{1}} dt_1 + m_{t_0}(dt_1) \tag{4.9}$$

and

$$m_{20}(t_1, t_2) \wedge_2 (t_1, dt_2) = e^{-\beta \mathbf{z}^T \mathbf{1}} dt_2 + m_{20}(t_1, dt_2) \tag{4.10}$$

where $\wedge_1(t_1)$ and $\wedge_2(t_1, t_2)$ represent the cumulative hazard functions corresponding to $\lambda_1(t_1)$ and $\lambda_2(t_1, t_2)$ respectively.

From (4.4), (4.7) and (4.8), we obtain

$$S(t_1, t_2 \mid \mathbf{z}) = \frac{m_{t_0}(0)}{m_{t_0}(t_1)} \frac{m_{20}(t_1, 0)}{m_{20}(t_1, t_2)} \exp \left[ -\int_0^{t_1} \frac{du}{m_{t_0}(u)} e^{\beta \mathbf{z}^T \mathbf{1}} - \int_0^{t_2} \frac{du}{m_{20}(t_1, u)} e^{\beta \mathbf{z}^T \mathbf{1}} \right]. \tag{4.11}$$

### 4.3. Inference Procedures

The observed data set consists of $n$ i.i.d. sets of $(\mathbf{X}_i, \delta_i, \mathbf{z}_i)$, $j = 1, 2$, where $\mathbf{X}_i = \min(T_{i1}, C_i)$, $\mathbf{X}_{2i} = \min(T_{2i}, C_i - X_{i1})$, $\delta_i = I(T_{i1} < C_i)$, $\delta_{2i} = I(T_{2i} < C_i - \mathbf{X}_{i1})$ and $\mathbf{z}_i$ is the covariate vector for $i = 1, 2, \ldots, n$. Here $I(.)$ is the indicator function. Given $\mathbf{z}_i$, $T_{ij}$ and $C_i$'s are assumed to be independent. Let

$$N_{1i}(t_1) = I(\mathbf{X}_{i1} \leq t_1) \delta_{i1}, \ N_{2i}(t_1, t_2) = I(\mathbf{X}_{i1} \geq t_1, \mathbf{X}_{2i} \leq t_2) \delta_{2i}, \ Y_{1i}(t_1) = I(\mathbf{X}_{i1} \geq t_1) \text{ and } Y_{2i}(t_1, t_2) = I(\mathbf{X}_{i1} \geq t_1, \mathbf{X}_{2i} \geq t_2), \ i = 1, 2, \ldots, n. \text{ From Fleming and Harrington (1991), we can write }$$

$$E\left[ N_{1i}(dt_1) \mid F_{t_1}; \beta, m_{t_0}(.) \right] = Y_{1i}(t_1) \wedge_{i1} (dt_1, \beta_{i1}, m_{t_0}) \tag{4.12}$$

and for fixed $t_1$,
where $\mathbb{F}_t$ belongs to the right-continuous filtrations $\{\mathbb{F}_t : t \geq 0\}$ and for fixed $t_1$, $\mathbb{F}_{t_1,t_2}$ belongs to the right-continuous filtrations $\{\mathbb{F}_{t_1,t_2} : t_1 \geq 0, t_2 \geq 0\}$ with $\mathbb{F}_t$ and $\mathbb{F}_{t_1,t_2}$ are defined by

\[
\mathbb{F}_t = \sigma\{N_{i_1}(u), Y_{i_1}(u+), \mathbb{Z}_i : 0 \leq u \leq t, i = 1, 2, ..., n\}
\]

and

\[
\mathbb{F}_{t_1,t_2} = \sigma\{N_{i_2}(t_1, v), Y_{i_2}(t_1, v+), \mathbb{Z}_i : 0 \leq v \leq t_2, i = 1, 2, ..., n\}.
\]

Denoting

\[
M_{i_1}(t_1, \beta, m_{i_0}) = N_{i_1}(t_1) - \int_0^{t_1} Y_{i_1}(s) \wedge_{i_1} (ds, \beta, m_{i_0})
\]

and

\[
M_{i_2}(t_1, t_2, \beta, m_{i_0}) = N_{i_2}(t_1) - \int_0^{t_1} Y_{i_2}(t_1, s) \wedge_{i_2} (t_1, ds, \beta, m_{i_0}), i = 1, 2, ..., n,
\]

$\{M_{i_1}(t_1, \beta, m_{i_0})\}$ is zero mean $\mathbb{F}_t$ martingale and for fixed $t_1$, $\{M_{i_2}(t_1, t_2, \beta, m_{i_0})\}$ is zero mean $\mathbb{F}_{t_1,t_2}$ martingale. Therefore the estimates of $\beta$ and $m_{i_0}(t_1)$ are obtained from the following partial score equations;

\[
\sum_{i=1}^{n} \left[ m_{i_0}(t_1) N_{i_1}(dt_1) - Y_{i_1}(t_1) \left\{ e^{-\beta \cdot \mathbb{Z}} dt_1 + m_{i_0}(dt_1) \right\} \right] = 0 \tag{4.14}
\]

and

\[
\sum_{i=1}^{n} \int_0^{t_1} \left[ m_{i_0}(t_1) N_{i_1}(dt_1) - Y_{i_1}(t_1) \left\{ e^{-\beta \cdot \mathbb{Z}} dt_1 + m_{i_0}(dt_1) \right\} \right] = 0, \quad 0 \leq t_1 < \tau_i. \tag{4.15}
\]

where $0 < \tau_i = \inf \{t_i : P[\tilde{X}_i > t_i] < \infty\}$.

It is easy to note that, equation (4.14) is a first order linear ordinary differential equation in $m_{i_0}(t_1)$, which can be written as

\[
\begin{bmatrix}
\sum_{i=1}^{n} N_{i_1}(dt_1) \\
\sum_{i=1}^{n} Y_{i_1}(t_1)
\end{bmatrix}
\begin{bmatrix}
m_{i_0}(t_1) \\
m_{i_0}(dt_1)
\end{bmatrix}
= Q_i(t_1, \beta) dt_1, \quad 0 \leq t_1 \leq \tau_i \tag{4.16}
\]
where \( Q_i(t_1, \beta) = \frac{\sum_{i=1}^{n} Y_{ii}(t_1)e^{-\beta^T z_i}}{\sum_{i=1}^{n} Y_{ii}(t_1)} \).

Then the solution of (4.16) is obtained as

\[
\hat{m}_{10}(t_1, \beta) = \hat{S}_{NA}^{(1)}(t_1)^{-1} \int_{0}^{t_1} \hat{S}_{NA}^{(1)}(u)Q_i(u, \beta) \, du
\]

(4.17)

where

\[
\hat{S}_{NA}^{(1)}(t_1) = \exp \left[ -\int_{0}^{t_1} \frac{\sum_{i=1}^{n} N_{ii}(du)}{\sum_{i=1}^{n} Y_{ii}(u)} \right],
\]

is the Nelson-Aalen estimator of the survival function for the pooled observations.

To estimate \( \beta \), we replace \( m_{10}(t_1) \) with \( \hat{m}_{10}(t_1, \beta) \) in (4.15) and then divide the resulting equation by \( n \). This will leads to score function

\[
U_i(\beta) = \frac{1}{n} \sum_{j=1}^{n} \left\{ \left( \sum_{i=1}^{n} N_{ii}(dt_1) - Y_{ii}(t_1)e^{-\beta^T z_i} \right) \right\} = 0
\]

(4.18)

with \( \bar{z}_i(t_1) = \frac{\sum_{i=1}^{n} Y_{ii}(t_1)Z_i}{\sum_{i=1}^{n} Y_{ii}(t_1)} \).

Here, \( \hat{m}_{10}(t_1) \) serves as a role similar to a weight function on each individual term in the summation. Then \( \hat{\beta} \) is the solution of (4.18). From Chen and Cheng (2005), it follows that under some regularity conditions, the random vector \( n^{1/2}(\hat{\beta} - \beta) \) converges weakly to a \( r \)-vector normal variable with mean zero and covariance matrix \( A_i^{-1}V_iA_i^{-1} \), where matrices \( A_i \) and \( V_i \) are given as

\[
A_i = \int_{0}^{t_1} \left[ E \left\{ \left( Z_i - \mu^{(1)}(t_1) \right) \right\} \otimes \hat{S}_{21}^{(1)}(t_1 | z_i)e^{-\beta^T z_i} \right] \, dt_1
\]

(4.19)

and

\[63\]
In addition, $A_1$ and $V_1$ can be consistently estimated by the empirical counterparts, and respectively, where $v^\otimes = vv'$ for a vector $v$. Inferences for $\beta_1$ can then be made through this large sample distribution of $\beta_1$. As shown by Chen and Cheng (2005), one can increase the efficiency of the estimator of $\beta_1$ by using the weighted version of the estimating equation.

To estimate $\beta_2$ and $m_20(t_1,t_2)$, for fixed $t_1$, the equations parallel to the partial score equations (4.14) and (4.15) are

$$\sum_{i=1}^n \left[ m_{20}(t_1,t_2)N_{2i}(t_1,dt_2) - Y_{2i}(t_1,t_2) \left\{ e^{-\beta_2'zdt_2} + m_{20}(t_1,dt_2) \right\} \right] = 0$$

and

$$\sum_{i=1}^n \int_{0}^{\tau_2} \left[ m_{20}(t_1,t_2)dN_{2i}(t_1,dt_2) - Y_{2i}(t_1,t_2) \left\{ e^{-\beta_2'zdt_2} + m_{20}(t_1,dt_2) \right\} \right] = 0, 0 \leq t_2 \leq \tau_2,$$

where for fixed $t_1$, $0 < \tau_2 = \inf \{ t_2 : P[X_2 > t_2 | X_1 > t_1] < \infty \}$.

Similarly, from (4.23), we obtain the first order linear ordinary differential equation in $m_{20}(t_1,t_2)$ as

$$\frac{m_{20}(t_1,t_2) - m_{20}(t_1,dt_2)}{\sum_{i=1}^n Y_{2i}(t_1,t_2)} = \frac{Q_2(t_1,t_2,\beta_2)dt_2, 0 \leq t_2 < \tau_2}{\sum_{i=1}^n N_{2i}(t_1,dt_2)}$$

(4.25)
where $Q_2(t_1, t_2, \beta_2) = \frac{\sum_{i=1}^{n} Y_{2i}(t_1, t_2) e^{-\beta_2 z_i}}{\sum_{i=1}^{n} Y_{2i}(t_1, t_2)}$.

Thus, for fixed $t_1$, the solution of (4.25) is given by

$$\hat{m}_{20}(t_1, t_2, \beta_2) = \hat{S}_{NA}^{(2)}(t_1, t_2)^{-1} \int_{t_2}^{t_2} \hat{S}_{NA}^{(2)}(t_1, u) Q_2(t_1, u, \beta_2) \, du$$

(4.26)

where $\hat{S}_{NA}^{(2)}(t_1, t_2) = \exp \left[ - \int_{0}^{t_2} \frac{\sum_{i=1}^{n} N_{2i}(t_1, du)}{\sum_{i=1}^{n} Y_{2i}(t_1, u)} \, du \right]$.

From (4.24) and (4.26), for fixed $t_1$, we obtain the score function as

$$U_2(\beta_2) = \frac{1}{n} \sum_{i=1}^{n} \left\{ \left[ z_i - \bar{z}_2(t_1, t_2) \right] \left\{ \hat{m}_{20}(t_1, t_2, \beta_2) dN_{2i}(t_1, t_2) - Y_{2i}(t_1, t_2) e^{-\beta_2 z_i} \, dt_2 \right\} \right\} = 0,$$

(4.27)

with $\bar{z}_2(t_1, t_2) = \frac{\sum_{i=1}^{n} Y_{2i}(t_1, t_2) z_i}{\sum_{i=1}^{n} Y_{2i}(t_1, t_2)}$.

The solution of the equation (4.27) provides the estimate of $\beta_2$.

For fixed $t_1$, the asymptotic properties of $\hat{\beta}_2$ and $\hat{m}_2(t_1, t_2)$ can be established by extending the proofs for $\hat{\beta}_1$ and $\hat{m}_1(t_1)$, given in Chen and Cheng (2005). To see this, let $H(.)$ be the marginal distribution function of $z$ and let $S^*(t_2 | t_1)$ be the conditional survival function of $T_2$ given $T_1 \geq t_1$. Since $m_{20}(t_1, t_2)$ is the mean residual life function of $T_2$ given $T_1 \geq t_1$, it follows from Arnold and Zahedi (1988) and (4.10) that

$$m_{20}(t_1, t_2) S^*(t_2 | t_1, z) = e^{-\beta_2 z} m_{20}(t_1, t_2) S^*(t_2 | t_1, z)$$

$$= e^{-\beta_2 z} \int_{t_2}^{t_2} S^*(u | t_1, z) \, du$$
for any possible $Z = z \in \text{Supp}\{z \in \mathbb{R}^p; H(z)\}$.

For fixed $t_1$, using Baye's theorem,

$$m_{20}(t_1, t_2) = \frac{1}{S^*(t_2 | t_1)} \int \int m_2(t_1, t_2 | z) S^*(t_2 | t_1, z) dH(z)$$

$$= \frac{1}{S^*(t_2 | t_1)} \int \left\{ e^{-\beta^*_z z} \int t_z u | t_1, z \right\} dH(z)$$

$$= \frac{1}{S^*(t_2 | t_1)} \int t_z u | t_1, z \int e^{-\beta^*_z z} dH(z | t_2 \geq u) du.$$

(4.28)

If we replace $S^*(t_2 | t_1)$ and $\int e^{-\beta^*_z z} dH(z | t_2 \geq u)$ with $S_{NA}^{(2)}(t_1, t_2)$ and $Q_2(t_1, u, \beta_2)$ respectively in (4.28), we obtain the same estimator for $m_{20}(t_1, t_2, \beta_2)$ as (4.26). As in the univariate case, under appropriate regularity conditions, any consistent estimates of these quantities will lead to a consistent estimator for $m_{20}(t_1, t_2)$.

Under certain regularity conditions, the random vector $n^{1/2}(\hat{\beta}_2 - \beta_2)$ is asymptotically $r$-vector normal with mean zero vector and covariance matrix $A_2^{-1}V_2A_2^{-1}$, where the matrices $A_2$ and $V_2$ are given by

$$A_2 = \int_0^{T_2} E \left[ (z - \mu^{(2)}_z(t_1, t_2))^\otimes S^*(t_1, t_2, z) e^{-\beta^*_z z} \right] dt_2$$

(4.29)

and

$$V_2 = \int_0^{T_2} E \left[ (z - \mu^{(2)}_z(t_1, t_2))^\otimes S^{(2)}_z(t_1, t_2 | z) m_{20}(t_1, t_2) \right] e^{-\beta^*_z z} dt_2 + m_{20}(t_1, dt_2)$$

(4.30)

with $\mu^{(2)}_z(t_1, t_2)$ is the limit of $\hat{z}_z(t_1, t_2)$ as $n \to \infty$.

Then $A_2$ and $V_2$ can be consistently estimated by their empirical counterparts,

$$ \hat{A}_2 = \frac{1}{n} \sum_{i=1}^{n} \int_0^{T_2} (\hat{z}_i - \hat{z}_z(t_1, t_2))^\otimes Y_2(t_1, t_2) e^{-\hat{\beta}^*_z z} dt_2$$

(4.31)
respectively.

One may often be interested in estimating the survival function \( S(t_1, t_2 \mid z_0) \) of gap times with a fixed covariate \( z_0 \). From (4.11), the survival function \( S(t_1, t_2 \mid z_0) \) can be estimated as

\[
\hat{S}(t_1, t_2 \mid z_0) = \frac{\hat{m}_{10}(0) \hat{m}_{20}(t_1, 0)}{\hat{m}_{10}(t_1) \hat{m}_{20}(t_1, t_2)} \exp \left[ -\int_0^{t_1} \frac{du}{\hat{m}_{10}(u)e^{\hat{\beta}_1 u}} - \int_0^{t_2} \frac{du}{\hat{m}_{20}(t_1, u)e^{\hat{\beta}_2 u}} \right].
\]  (4.33)

Asymptotic distribution theory is difficult for non-parametric bivariate survival function estimates and the most attractive approach to variance or confidence interval estimation is through resampling methods. The naive bootstrap procedure of resampling the observed data units \((X_{1i}, X_{2i}, \delta_{1i}, \delta_{2i}, z_0)\) with replacement will be satisfactory under fairly mild conditions (see Efron and Tibshirani, 1993).

The most challenging part in this procedure is the tail due to potential censoring. If the underlying recurrence times are heavily right censored, it is not possible to estimate the mean residual life functions on the whole positive real line without additional assumptions. One possible approach is to modify the fully unspecified \( m_{10}(t_1) \) and \( m_{20}(t_1, t_2) \) by including a parametric component in the tail. For example, when \( \tau_0 \) is a pre-specified truncation time, we assume that

\[
m_{10}(t_1) = m_{10}(t_1)I(t_1 < \tau_0) + m_{1a}I(t_1 > \tau_0)
\]  (4.34)

and for fixed \( t_1 \),

\[
m_{20}(t_1, t_2) = m_{20}(t_1, t_2)I(t_2 < \tau_0) + m_{2a}I(t_2 > \tau_0),
\]  (4.35)

where \( m_{1a} \) and \( m_{2a} \) are some positive constants. Thus the \( m_{10}(t_1) \) and \( m_{20}(t_1, t_2) \) are unspecified up to the time \( \tau_0 \), while it becomes exponential after \( \tau_0 \), so that the techniques discussed earlier can be extended to the whole positive real line.
4.4. Simulation Study

In this section, we carried out a simulation study to evaluate the performance of the aforementioned inference procedures. We consider a Gumbel's (1960) bivariate exponential distribution with survival function

\[ S(t_1, t_2) = \exp(-t_1 - t_2 - \gamma_1 t_2) \], \( t_1, t_2 > 0 \), \( 0 \leq \gamma \leq 1 \) \hfill (4.36)

We considered a single covariate \( z \), which is generated from uniform \((0, 1)\) distribution. We generated observations from Gumbel's bivariate exponential distribution for different values of \( \gamma \) using algorithm given in Devroye (1986). Independent censoring times are generated from the uniform distribution \((0, b)\), where the constant \( b \) is taken in such a way that 30% of the observations are censored. We compute estimates of \( \hat{\beta}_1 \) and \( \hat{\beta}_2 \) for 1000 simulations and then calculate empirical bias and variance of \( \hat{\beta}_1 \) and \( \hat{\beta}_2 \), which are given in Table 4.1 and Table 4.2. The empirical bias and variance of the estimates of the baseline mean residual functions along with coverage probabilities are given in Table 4.3 and Table 4.4. As \( n \) increases, both bias and variance of the estimates decreases.

<table>
<thead>
<tr>
<th>( \beta_1 )</th>
<th>( \beta_2 )</th>
<th>( \gamma )</th>
<th>( n )</th>
<th>( \text{Bias} \hat{\beta}_1 )</th>
<th>( \text{Var} \hat{\beta}_1 )</th>
<th>( \text{Cov. Prob} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.6</td>
<td>0.6</td>
<td>50</td>
<td>-0.0448</td>
<td>0.0920</td>
<td>0.939</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>250</td>
<td>-0.0408</td>
<td>0.0545</td>
<td>0.949</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.8</td>
<td>50</td>
<td>-0.0376</td>
<td>0.0964</td>
<td>0.942</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>250</td>
<td>-0.0355</td>
<td>0.0626</td>
<td>0.948</td>
</tr>
<tr>
<td>-0.6</td>
<td>1.1</td>
<td>0.6</td>
<td>50</td>
<td>0.0577</td>
<td>0.0829</td>
<td>0.939</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>250</td>
<td>0.0546</td>
<td>0.0412</td>
<td>0.941</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.8</td>
<td>50</td>
<td>0.0554</td>
<td>0.0499</td>
<td>0.938</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>250</td>
<td>0.0525</td>
<td>0.0289</td>
<td>0.943</td>
</tr>
</tbody>
</table>

Table 4.1. Bias and variance of \( \hat{\beta}_1 \)
Table 4.2. Bias and variance of $\hat{p}_2$

<table>
<thead>
<tr>
<th>$\beta_1$</th>
<th>$\beta_2$</th>
<th>$\gamma$</th>
<th>n</th>
<th>Bias $\hat{p}_2$</th>
<th>Var $\hat{p}_2$</th>
<th>Cov. Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.6</td>
<td>0.6</td>
<td>50</td>
<td>-0.0487</td>
<td>0.0406</td>
<td>0.942</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>250</td>
<td>-0.0241</td>
<td>0.0194</td>
<td>0.945</td>
</tr>
<tr>
<td>0.5</td>
<td>0.6</td>
<td>0.8</td>
<td>50</td>
<td>-0.0510</td>
<td>0.0461</td>
<td>0.946</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>250</td>
<td>-0.0498</td>
<td>0.0162</td>
<td>0.952</td>
</tr>
<tr>
<td>-0.6</td>
<td>1.1</td>
<td>0.6</td>
<td>50</td>
<td>-0.0971</td>
<td>0.0631</td>
<td>0.939</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>250</td>
<td>-0.0850</td>
<td>0.0422</td>
<td>0.941</td>
</tr>
<tr>
<td>-0.6</td>
<td>1.1</td>
<td>0.8</td>
<td>50</td>
<td>-0.0958</td>
<td>0.0756</td>
<td>0.948</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>250</td>
<td>-0.0897</td>
<td>0.0561</td>
<td>0.949</td>
</tr>
</tbody>
</table>

Table 4.3. Bias and variance of $\hat{m}_{10}(t_1)$

<table>
<thead>
<tr>
<th>$\beta_1$</th>
<th>$\beta_2$</th>
<th>$t_1$</th>
<th>$\gamma$</th>
<th>n</th>
<th>Bias $\hat{m}_{10}(t_1)$</th>
<th>Var $\hat{m}_{10}(t_1)$</th>
<th>Cov. Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.6</td>
<td>0.05</td>
<td>0.6</td>
<td>50</td>
<td>0.0578</td>
<td>0.0206</td>
<td>0.943</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>250</td>
<td>0.0102</td>
<td>0.0112</td>
<td>0.945</td>
</tr>
<tr>
<td>0.5</td>
<td>0.6</td>
<td>0.09</td>
<td>0.6</td>
<td>50</td>
<td>0.0989</td>
<td>0.0288</td>
<td>0.947</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>250</td>
<td>0.0539</td>
<td>0.0150</td>
<td>0.949</td>
</tr>
<tr>
<td>0.5</td>
<td>0.6</td>
<td>0.09</td>
<td>0.8</td>
<td>50</td>
<td>0.0526</td>
<td>0.0248</td>
<td>0.951</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>250</td>
<td>0.0101</td>
<td>0.0116</td>
<td>0.954</td>
</tr>
<tr>
<td>0.5</td>
<td>0.6</td>
<td>0.09</td>
<td>0.8</td>
<td>50</td>
<td>0.0955</td>
<td>0.0503</td>
<td>0.936</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>250</td>
<td>0.0499</td>
<td>0.0218</td>
<td>0.940</td>
</tr>
<tr>
<td>-0.6</td>
<td>1.1</td>
<td>0.05</td>
<td>0.6</td>
<td>50</td>
<td>0.0587</td>
<td>0.0736</td>
<td>0.941</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>250</td>
<td>0.0105</td>
<td>0.0166</td>
<td>0.950</td>
</tr>
<tr>
<td>-0.6</td>
<td>1.1</td>
<td>0.09</td>
<td>0.6</td>
<td>50</td>
<td>0.0561</td>
<td>0.0197</td>
<td>0.942</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>250</td>
<td>0.0103</td>
<td>0.0135</td>
<td>0.953</td>
</tr>
<tr>
<td>-0.6</td>
<td>1.1</td>
<td>0.09</td>
<td>0.8</td>
<td>50</td>
<td>0.0581</td>
<td>0.0243</td>
<td>0.949</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>250</td>
<td>0.0105</td>
<td>0.0168</td>
<td>0.957</td>
</tr>
<tr>
<td>-0.6</td>
<td>1.1</td>
<td>0.09</td>
<td>0.8</td>
<td>50</td>
<td>0.0992</td>
<td>0.0315</td>
<td>0.951</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>250</td>
<td>0.0558</td>
<td>0.0199</td>
<td>0.955</td>
</tr>
</tbody>
</table>
Table 4.4. Bias and variance $\hat{m}_{20}(t_1, t_2)$

<table>
<thead>
<tr>
<th>$\beta_1$</th>
<th>$\beta_2$</th>
<th>$(t_1, t_2)$</th>
<th>$\gamma$</th>
<th>n</th>
<th>Bias $\hat{m}_{20}(t_1, t_2)$</th>
<th>Var $\hat{m}_{20}(t_1, t_2)$</th>
<th>Cov. Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.6</td>
<td>(0.05, 0.06)</td>
<td>0.6</td>
<td>50</td>
<td>0.0232</td>
<td>0.0324</td>
<td>0.950</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>250</td>
<td>0.0157</td>
<td>0.0129</td>
<td>0.952</td>
</tr>
<tr>
<td>0.5</td>
<td>0.6</td>
<td>(0.09, 0.07)</td>
<td>0.8</td>
<td>50</td>
<td>0.0163</td>
<td>0.0138</td>
<td>0.948</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>250</td>
<td>0.0123</td>
<td>0.0101</td>
<td>0.953</td>
</tr>
<tr>
<td>-0.6</td>
<td>1.1</td>
<td>(0.05, 0.06)</td>
<td>0.6</td>
<td>50</td>
<td>0.0223</td>
<td>0.0269</td>
<td>0.939</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>250</td>
<td>0.0144</td>
<td>0.0177</td>
<td>0.940</td>
</tr>
<tr>
<td>-0.6</td>
<td>1.1</td>
<td>(0.09, 0.07)</td>
<td>0.8</td>
<td>50</td>
<td>0.0978</td>
<td>0.0216</td>
<td>0.935</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>250</td>
<td>0.0156</td>
<td>0.0102</td>
<td>0.939</td>
</tr>
</tbody>
</table>

4.5. Data Analysis

For the illustration of the procedure, we consider a data given in Lawless (2003, p.531). The data shows the recurrent times to infection at the point of insertion of the catheter for 38 persons undergoing kidney dialysis. Data for the first two occurrences of infection are given; either one or both may be censored, because catheters were sometimes removed for causes other than infection. The covariate considered in our study is kidney disease type ($0 = $glomerulo nephritis, $1 = $acute nephritis, $2 = $polycystic kidney disease, $3 = $other). $T_1$ and $T_2$ represents the first two occurrences of infection.

We compute the estimates of $\hat{\beta}_1$ and $\hat{\beta}_2$ by the method given in Section 4.3. The estimates $\hat{\beta}_1$ and $\hat{\beta}_2$ are given in Table 4.5. We then estimate $m_{10}(t_1)$ and $m_{20}(t_1, t_2)$. Finally, we estimate $S(t_1, t_2 \mid z)$ by substituting the estimates of $\hat{\beta}_1$, $\hat{\beta}_2$, $m_{10}(t_1)$ and $m_{20}(t_1, t_2)$ in (4.33). The estimates are given in Table 4.5. From Table 4.5, it is easy to see that the values of both $\hat{\beta}_1$ and $\hat{\beta}_2$ are negative. The disease type
has negative effect on the first and second recurrence times of the individuals respectively. As expected, \( \hat{m}_{10}(t_1) \) is decreasing in \( t_1 \). However, \( \hat{m}_{20}(t_1, t_2) \) depends both on \( t_1 \) and \( t_2 \).

There are natural restrictions on both \( m_{10}(t_1) \) and \( m_{20}(t_1, t_2) \) such that \( m_{10}(t_1) \) must be monotonically non-decreasing in \( t_1 \) and \( m_{20}(t_1, t_2) + t_2 \) should be monotonically non-decreasing in \( t_2 \) for every fixed \( t_1 \). These constraints are satisfied in this data example. To check the model adequacy of (4.10) and (4.11), the estimated mean residual life functions of \( T_1 \) and that of \( T_2 \) given \( T_1 \geq t_1 \), without adjusting for any of the covariates and the estimated baseline mean residual life functions are plotted. On the log scale, their lowess curves are parallel to each other and their difference is roughly constant (see Figures 4.1 and 4.2). This suggests a reasonable goodness of fit of the proportionality assumption given in (4.7) and (4.8). Figure 4.3 shows the estimates of the survival function.

**Table 4.5.** Estimates of \( \beta_1, \beta_2, m_{10}(t_1) \) and \( m_{20}(t_1, t_2) \)

<table>
<thead>
<tr>
<th>((t_1, t_2))</th>
<th>(\hat{\beta}_1)</th>
<th>(\hat{\beta}_2)</th>
<th>(z)</th>
<th>(\hat{m}_{10}(t_1))</th>
<th>(\hat{m}_{20}(t_1, t_2))</th>
<th>(\hat{S}(t_1, t_2 \mid z))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(7,9)</td>
<td>-0.0774</td>
<td>-0.0440</td>
<td>0</td>
<td>20.3853</td>
<td>11.5136</td>
<td>0.2965</td>
</tr>
<tr>
<td>(8,16)</td>
<td>-0.0774</td>
<td>-0.0440</td>
<td>3</td>
<td>19.6691</td>
<td>9.2308</td>
<td>0.2242</td>
</tr>
<tr>
<td>(2,25)</td>
<td>-0.0774</td>
<td>-0.0440</td>
<td>0</td>
<td>21.5899</td>
<td>6.1896</td>
<td>0.0108</td>
</tr>
<tr>
<td>(30,12)</td>
<td>-0.0774</td>
<td>-0.0440</td>
<td>3</td>
<td>13.7672</td>
<td>11.0371</td>
<td>0.0061</td>
</tr>
<tr>
<td>(12,40)</td>
<td>-0.0774</td>
<td>-0.0440</td>
<td>1</td>
<td>18.0828</td>
<td>9.5078</td>
<td>0.0012</td>
</tr>
<tr>
<td>(34,30)</td>
<td>-0.0774</td>
<td>-0.0440</td>
<td>1</td>
<td>13.2308</td>
<td>12.0054</td>
<td>0.0010</td>
</tr>
<tr>
<td>(22,28)</td>
<td>-0.0774</td>
<td>-0.0440</td>
<td>3</td>
<td>17.0786</td>
<td>3.5833</td>
<td>0.0001</td>
</tr>
</tbody>
</table>
Figure 4.1. Log of mean residual lifetimes for $T_1$

Figure 4.2. Log of mean residual lifetimes for $T_2$
4.6. Conclusion

We introduced a bivariate proportional mean residual life model to assess the effect of covariates on the mean residual life function for the gap time distribution of recurrent events. The model is a transparent extension of the mean residual lifetime model developed by Chen and Cheng (2005) for univariate survival data. The estimation of parameter vectors and baseline mean residual life functions were done using counting process theory. The proposed method can directly be extended to the higher dimensions by considering the multivariate mean residual life function of Arnold and Zahedi (1988).

The efficiency of the proposed method depends on the conditionally i.i.d. assumption on the bivariate recurrence times as well as the independent censoring assumption. In the presence of trend on the bivariate recurrence times, the i.i.d. assumption will be violated and therefore, the proposed method would not be appropriate. The independent assumption could also fail if the observed data is terminated by information drop out or a failure event. Both assumptions should be examined carefully when applying the proposed method.