Regression Analysis of Mean Residual Life Function

Shufang Liu
PRA International, Charlottesville, VA
Sujit K. Ghosh*
Department of Statistics, North Carolina State University, Raleigh, NC 27695
**email*: ghosh@stat.ncsu.edu

Institute of Statistical Mimeo Series # 2613

Abstract

The mean residual life function (mrlf) of a subject is defined as the expected remaining lifetime of the subject given that the subject has survived up to a given time. The commonly used regression models as proportional mean residual life (PMRL) and linear mean residual life (LMRL) have limited applications due to adhoc restriction on the parameter space. The regression model we propose does not have any constraints. It turns out that the proposed proportional scaled mean residual life (PSMRL) model is equivalent to the accelerated failure time (AFT) model, which provides an alternative way to estimate the regression parameters of the AFT model and to interpret the regression parameters estimated from the AFT model in terms of the mrlf instead of the survival function. We use full likelihood by nonparametrically estimating the baseline mrlf using the smooth scale mixture estimator of the mrlf based on a single sample of iid observations to develop the statistical inference for the regression parameters, which are estimated using an iterative procedure. A simulation study is carried out to assess the properties of the estimators of the regression coefficients. We illustrate our regression model by applying it to the well-known Veteran's Administration lung cancer data.

Keywords: AFT model; Mean residual life function; PSMRL model; Regression analysis.

1 Introduction

The survival function and the hazard function are important in understanding the survival or aging process. However, to assess the remaining life expectancy, it is desirable to study the mean residual life function, which can sometimes serve as a more useful tool than the survival function and the hazard function. For

instance, people might be more interested to know how many more years they are expected to survive as compared to their instantaneous survival chance given that they began a treatment at a certain time ago (Elandt-Johnson and Johnson, 1980). The lifetime random variable T is assumed to have finite expectation, i.e., $E(T) = \int_0^\infty t dF(t) = \int_0^\infty S(t)dt < \infty$. Then, the mean residual life function (mrlf), which gives a measure of the remaining lifetime of a subject given that the subject has survived up to time t, is defined as

$$m(t) \equiv E[T - t|T > t] = \begin{cases} \int_t^\infty \frac{S(u)}{S(t)} du, & S(t) > 0, \\ 0 \qquad\qquad , & \text{otherwise}, \end{cases} \tag{1}$$

To measure the effect of covariates on the survival time, Cox (1972) proposed the the proportional hazards (PH) model and Kalbfleisch and Prentice (1980) proposed the AFT model. To evaluate the effect of covariates on the remaining lifetime, Oaks and Dasu (1990) proposed a semiparametric model called the proportional mean residual life (PMRL) model, which is similar to the PH model and is defined as

$$m(t|\boldsymbol{z}) = m_0(t) \exp\{\boldsymbol{z}^\tau \boldsymbol{\beta}\}, \tag{2}$$

where $m(t|\boldsymbol{z}) = E[T-t|T > t, \boldsymbol{Z} = \boldsymbol{z}]$ is the conditional mean residual life function (cmrlf) and $m_0(t) = m(t|\boldsymbol{Z} = 0)$ is the baseline mrlf, which is generally unspecified and can be interpreted as the mrlf for the population of subjects with $\boldsymbol{Z} = 0$. The PMRL model has been studied by Maguluri and Zhang (1994), Chen and Jewell (2002), Chen et al.(2005), Chen and Cheng (2005), and so on. The PMRL model directly describes the distribution of remaining lifetime and holds appealing interpretation in life expectancy. However, this model has some limitations, such as the life expectancies of T at t given by $e_0(t) = m_0(t) + t$ and $e(t|\boldsymbol{z}) = m(t|\boldsymbol{z}) + t$ must be non-decreasing (see the Characterization Theorem of Hall and Wellner, 1981). From these limitations, we can be see that

(1) for an increasing $m_0(\cdot)$, $m(t|\boldsymbol{z})$ is a proper mrlf for all $\boldsymbol{\beta}$; and

(2) for a decreasing and non-monotonic $m_0(\cdot)$, $m(t|\boldsymbol{z})$ is a proper mrlf for all $\boldsymbol{\beta}$ only when $\exp(\boldsymbol{z}^\tau \boldsymbol{\beta}) > 1$ is satisfied.

Thus, in order to obtain a proper estimator of $m(t|\boldsymbol{z})$, the above two conditions have to be checked after obtaining $\hat{\boldsymbol{\beta}}$ and $\hat{m}_0(\cdot)$. Even Oakes and Dasu (1990) pointed out that $m(t|\boldsymbol{z})$ is an mrlf for all $\boldsymbol{\beta}$ if $m_0(\cdot)$ itself is monotonically non-decreasing when they proposed the PMRL model.

To ensure that the life expectancy $e(t|\boldsymbol{z})$ is montonically non-decreasing, Chen and Cheng (2006) proposed a linear mean residual life (LMRL) model, which is defined as

$$m(t|\boldsymbol{z}) = m_0(t) + \boldsymbol{z}^\tau \boldsymbol{\beta}. \tag{3}$$

However, this model also has constraints such as $\boldsymbol{z}^\tau \boldsymbol{\beta}$ is non-negative and $e_0(t)$ is non-decreasing.

It can be seen that both the PMRL and the LMRL have some limitations. In Section 2, we propose a proportional scaled mean residual life model, which does not have any restrictions. In Section 3, we present a simulation study to evaluate the performance of the proposed regression model. In Section 4, we apply the proposed regression model to a real data set. In Section 5, we do some discussion.

2 Proportional Scaled Mean Residual Life Model

Let T_i and C_i denote the survival time and the censoring time of the ith subject for $i = 1, \ldots, n$, respectively, i.e., we only observe $X_i = \min\{T_i, C_i\}$ and the indicator of censoring of $\Delta_i = I(C_i - T_i)$. Let $Z_i = (Z_{i1}, \ldots, Z_{ip})^\tau$ denote a vector of p explanatory variables. For instance, Z_i are the prognostic factors for the real data set in Section 4 and Z_{i1} = treatment, Z_{i2} = karno, and so on. The observed data set is given by $\mathcal{D}_R = \{(X_i, \Delta_i, Z_i), i = 1, \ldots, n\}$. We assume that for each subject given Z_i, the T_i's and C_i's are conditionally independent and the triplets (T_i, C_i, Z_i) are also independent across subjects, $i = 1, \ldots, n$.

The motivation behind the proportional scaled mean residual life (PSMRL) model is based on the following theorem.

THEOREM **1.** *Let $m(t) = E[T - t|T > t]$ be an mrlf. Then $\frac{m(t\theta)}{\theta}$ is a proper mrlf for any $\theta > 0$.*

Proof: Notice that, for $\theta > 0$,

$$\begin{aligned} \frac{m(t\theta)}{\theta} &= E[T - t\theta|T > t\theta]\frac{1}{\theta} \\ &= E\left[\frac{T}{\theta} - t\middle|\frac{T}{\theta} > t\right]. \end{aligned}$$

Hence $\frac{m(t\theta)}{\theta}$ is the mrlf of $\frac{T}{\theta}$.

We assume that $m_0(t)$ is continuously differentiable on $[0, T_{(n)}]$, where $T_{(n)} = \max(T_i, i = 1, \ldots, n)$. Based on Theorem 1, the PSMRL model which is proper for all $\boldsymbol{\beta}$, is defined as

$$m(t|\boldsymbol{z}) = \exp\{g(\boldsymbol{z}, \boldsymbol{\beta})\} m_0(t \exp\{-g(\boldsymbol{z}, \boldsymbol{\beta})\}), \tag{4}$$

where the link function $g(\boldsymbol{z}, \boldsymbol{\beta}) = \boldsymbol{z}^\tau \boldsymbol{\beta}$ is chosen in this study, i.e., the PSMRL model is then given by

$$m(t|\boldsymbol{z}) = \exp\{\boldsymbol{z}^\tau \boldsymbol{\beta}\} m_0(t \exp\{-\boldsymbol{z}^\tau \boldsymbol{\beta}\}). \tag{5}$$

The AFT model can be obtained directly from our PSMRL model based on the inversion formula given by

$$S(t|\boldsymbol{z}) = \Pr(T \geq t, \boldsymbol{Z} = \boldsymbol{z}) = \frac{m(0|\boldsymbol{z})}{m(t|\boldsymbol{z})} \exp\left\{ - \int_0^t \frac{1}{m(u|\boldsymbol{z})} du \right\}. \tag{6}$$

This indicates that the PSMRL model is equivalent to the AFT model in every sense. Naturally, the regression parameter $\boldsymbol{\beta}$ in the PSMRL model has the same direct interpretation as the one in the AFT model as the change in the logarithm of the survival times per unit change in $\boldsymbol{Z}$. The PSMRL model also has its own physical interpretation. This model reflects both the time scale changing effect and the proportional mean residual life effect, because

$$\frac{m(te^{\boldsymbol{\delta}^\tau \boldsymbol{\beta}} | \boldsymbol{z} + \boldsymbol{\delta})}{m(t|\boldsymbol{z})} = e^{\boldsymbol{\delta}^\tau \boldsymbol{\beta}},$$

where $\boldsymbol{\delta}$ is a p-dimensional vector.

With $\delta_k = 1$ unit increase in the k-th covariate Z_k while other covariates being held fixed, we can obtain

$$\frac{m(te^{\beta_k} | Z_k + 1)}{m(t|Z_k)} = e^{\beta_k}.$$

If $\beta_k > 0$, the expected remaining lifetime at a larger time point for a subject with $Z_k + 1$ is e^{β_k} (where $e^{\beta_k} > 1$) times of the expected remaining lifetime at a smaller time point for a subject with Z_k. Let Z_k and $Z_k + 1$ denote a control treatment and a new treatment for a certain disease. If $\beta_k > 0$, subjects taking the new treatment even for a longer time have a longer average remaining lifetime than subjects taking the control treatment for a shorter time. This indicates that the new treatment has improved the remaining lifetime of subjects. If $\beta_k < 0$,

the expected remaining lifetime at a smaller time point for a subject with $Z_k + 1$ is e^{β_k} (where $e^{\beta_k} < 1$) times of the expected remaining lifetime at a larger time point for a subject with Z_k. Let Z_k and $Z_k + 1$ denote low blood pressure and high blood pressure, respectively. If $\beta_k < 0$, subjects with higher blood pressure have a shorter average remaining lifetime at a younger age than subjects with lower blood pressure at older age.

The AFT model has been well studied since its proposition. The least square based, the rank-based, and the M-estimation methods have been advanced in theory, but the complicated computation has kept them from being applied in software. The semiparametric method of Jin et al. (2003) was the first to be applied in the `rankreg` package of `R`, where the Gehan estimates and the logrank estimates for the regression parameters and their covariance matrices from resampling can be obtained. However, $S_0(\cdot)$ and $S(\cdot|\boldsymbol{z})$ can not be estimated from their approach. In this study, we propose a nonparametric maximum likelihood method based on the PSMRL model to estimate $\boldsymbol{\beta}$ and $m_0(\cdot)$, and hence $S_0(\cdot)$, $S(\cdot|\boldsymbol{z})$, and $m(\cdot|\boldsymbol{z})$ can be obtained.

The conditional density function $f(t|\boldsymbol{z}) = h(t|\boldsymbol{z})^{\Delta} S(t|\boldsymbol{z})$ is used to build a full likelihood function. Appendix A gives the procedures to derive the loglikelihood function, which is expressed as

$$\begin{aligned} ll(\boldsymbol{\beta}, m_0(\cdot)) &= n\log(m_0(0)) - \sum_{i=1}^{n} \Delta_i \boldsymbol{Z}_i^{\tau}\boldsymbol{\beta} - \sum_{i=1}^{n} (\Delta_i + 1)\log(m_0(X_i \exp\{-\boldsymbol{Z}_i^{\tau}\boldsymbol{\beta}\})) \\ &\quad + \sum_{i=1}^{n} \Delta_i \log(m_0'(X_i \exp\{-\boldsymbol{Z}_i^{\tau}\boldsymbol{\beta}\}) + 1) - \sum_{i=1}^{n} \int_0^{\exp\{-\boldsymbol{Z}_i^{\tau}\boldsymbol{\beta}\}X_i} \frac{1}{m_0(v)} dv. \end{aligned}$$

To obtain the estimate of $\boldsymbol{\beta}$ from the above loglikelihood function, we have to know $m_0(t)$. Because of a special property of the PSMRL model, $m_0(t)$ emerges as the mrlf of the transformed survival time $X^* = X \exp\{-\boldsymbol{z}^{\tau}\boldsymbol{\beta}\}$ and the details of derivation are given in Appendix B. For the transformed data $\{(X_i^*, \Delta_i), i = 1, \ldots, n\}$, $\hat{m}_0(t)$ can be calculated using the closed analytical form for the smooth estimator of the mrlf based on a single sample (Liu and Ghosh, 2008), which is also given in Appendix B. Since $\hat{m}_0(t)$ can be calculated based on the transformed survival time, which involves $\boldsymbol{\beta}$, we propose a 2-step estimation scheme to obtain $\hat{\boldsymbol{\beta}}$ and $\hat{m}_0(t)$ iteratively:

- **STEP 0: Set initial values**

Set $\hat{\boldsymbol{\beta}}^{(0)} = 0$ and the transformed survival time is still $\{X_i, i = 1, \ldots, n\}$.
Calculate $\hat{m}_0^{(0)}(t)$ based on $\{(X_i, \Delta_i), i = 1, \ldots, n\}$.
Obtain $\hat{\boldsymbol{\beta}}^{(1)}$ based on $\{(X_i, \Delta_i, Z_i), i = 1, \ldots, n\}$ by maximizing $ll\left(\boldsymbol{\beta}, m_0^{(0)}(\cdot)\right)$.
Set k=1.

- **STEP 1: Calculate** $\hat{m}_0(\cdot)$
  Transform the survival time, $X^{*(k)} = X \exp\left\{-\boldsymbol{z}^\tau \hat{\boldsymbol{\beta}}^{(k-1)}\right\}$.
  Calculate $\hat{m}_0^{(k)}(t)$ based on $\left\{\left(X_i^{*(k)}, \Delta_i\right), i = 1, \ldots, n\right\}$.

- **STEP 2: Estimate** $\boldsymbol{\beta}$
  Obtain $\hat{\boldsymbol{\beta}}^{(k+1)}$ based on $\{(X_i, \Delta_i, Z_i), i = 1, \ldots, n\}$ by maximizing $ll\left(\boldsymbol{\beta}, m_0^{(k)}(\cdot)\right)$.
  Set k=k+1 and return to **STEP 1**.

STEP 1 and **STEP 2** are iterated until the convergence criterion, which is set to be that the difference between two successive estimates is smaller than 0.01, is satisfied. The module `OPTIM` in `R` can be used to obtain $\hat{\boldsymbol{\beta}}$ by maximizing $ll(\boldsymbol{\beta}, \hat{m}_0(\cdot))$. After convergence, we can obtain $\hat{m}_0(t)$ from the last iteration and $\hat{m}(t|\boldsymbol{z})$ as

$$\hat{m}(t|\boldsymbol{z}) = \exp\{\boldsymbol{z}^\tau \hat{\boldsymbol{\beta}}\} \hat{m}_0(t \exp\{-\boldsymbol{z}^\tau \hat{\boldsymbol{\beta}}\}).$$

In Section 4, we apply the above equation to obtain $\hat{m}(t|\boldsymbol{z})$ for two different sets of covariates (see Figure 1). Accordingly,$\hat{S}_0(t), \hat{S}(t|\boldsymbol{z}), \hat{h}_0(t)$, and $\hat{h}(t|\boldsymbol{z})$ can be calculated.

3 A Simulation Study

A simulation study was conducted to assess the properties of the proposed estimation procedures. Some preliminary simulation studies showed that the parameter k_n used to calculate $\hat{m}_0(t)$ (8) is needed to be tuned for highly censored data. For simplicity, in this simulation study, we just consider complete data with covariates, i.e., $\mathcal{D}_R = \{(T_i, Z_i), i = 1, \ldots, n\}$. The simulation study was carried out under the following conditions:

- $m_0(t) = 1$ is Exponential distribution with the density function $f_0(t) = e^{-t} I(t)$.

- The covariates are $\boldsymbol{Z} = (Z_1, Z_2)^\tau$, where $Z_1 \sim$ Ber(0.5) and Z_1 is then centered to have values of ± 0.5 and $Z_2 \sim$ N(0, 1). Therefore, a discrete variable and a continuous variable are used.

- The lifetime T is generated according to our PSMRL model (5), where the true regression parameters $\boldsymbol{\beta}$ are $(1, 1)^\tau$ or $(0, 1)^\tau$.

- A sample size of $n = 100$ and a Monte Carlo sample size of $N = 1000$ are used.

- The tuning parameter k_n used to calculate $\hat{m}_0(t)$ is set to 2.

Table 1: Results based on the simulation study

	Proposed model		Jin et al. method Gehan estimate		Jin et al. method Logrank estimate	
	Z1	Z2	Z1	Z2	Z1	Z2
$\boldsymbol{\beta} = (1, 1)^\tau$						
N	995	995	1000	1000	995	995
Bias	-0.012	-0.006	-0.251	-0.253	0.312	0.324
SE	0.019	0.009	0.009	0.007	0.014	0.012
P-value	0.264	0.252	< 0.001	< 0.001	< 0.001	< 0.001
$\boldsymbol{\beta} = (0, 1)^\tau$						
N	995	995	1000	1000	1000	1000
Bias	0.008	-0.014	0.001	-0.269	0.014	0.273
SE	0.018	0.009	0.007	0.006	0.011	0.010
P-value	0.328	0.06	0.443	< 0.001	0.101	< 0.001
SE is the standard deviation of parameter estimates.						

The generated data were analyzed using the 2-step estimation procedures in `R` and using the module `aft.fun` in the `rankreg` package of `R` based on the approaches of Jin et al. (2003) for the AFT model. The simulation results are summarized in Table 1, including N, Bias, SE, and P-value. N is the number of the estimates used to calculate Bias and SE. If the absolute value of an estimate exceeded 3, the estimate was excluded to calculate Bias and SE. P-value is used to test if Bias is significantly different from 0.

From Table 1, it can be seen that N is at least 995 out of 1000, which indicates that we do not have many extreme estimates. From Table 1, it can be observed that the biases of the estimates for β_1 and β_2 under our approach are close to 0, while the Gehan estimates tend to underestimate and the logrank estimates tend to overestimate the parameters with true values of 1. Further the P-values for the biases of the Gehan estimates and the logrank estimates for the parameters with true values of 1 are less than 0.001, which indicates these biases are significantly different from 0.

It can be concluded that the 2-step estimation procedures based on the PSMRL model is better than the semiparametric approach of the AFT model of Jin et al. (2003) in terms of reducing bias based on the simulation study.

4 Application to a Lung Cancer Trial

The data set of the Veterans' Administration lung cancer trial (Prentice, 1973) is often analyzed in survival analysis. There are 137 subjects in the data set with a low censoring rate of 6.6%. Therefore, the tuning parameter k_n used to calculate $\hat{m}_0(t)$ is still set to 2 as in the simulation study. The response variable is the survival time ranging from 1 to 999 days. The first covariate we consider is treatment with Levels 1 and 2. The second covariate is karno (Karnofsky Performance Score), which ranges from 10 to 99. Treatment is a discrete variable and karno is a continuous variable. To maintain numerical stability, we rescaled the response variable by dividing its standard deviation, and centered treatment to take values ± 0.5, and standardized karno to take values roughly in $(-3, 3)$.

The data set was analyzed by the 2-step estimation procedures based on the PSMRL model and the semiparametric approach of Jin et al. (2003) based on the AFT model. The estimation results are summarized in Table 2, including Estimate, SE from resampling of the data for 100 times, and P-value. It can be seen the two approaches have similar results, i.e., the two levels of treatment do not have significant different effects while karno has a significant effect on the survival time based on P-values. It can also be seen from Table 2 that the three estimates for karno are quite different, even though they are supposed to be the same because the PSMRL model is equivalent to the AFT model. Since treatment is a discrete variable, it is possible for our approach or the AFT approach of Jin et al. (2003) not to converge. In this case, our approach converged and the semiparametric

approach of Jin et al. did not converge, which implies that the solution from Jin et al. (2003) is not unique. This might explain the discrepancy among the estimates. Since our approach converged, it is more reasonable to interpret the estimates from our approach.

As we mentioned in Section 2, $\hat{\boldsymbol{\beta}}$ and $\hat{m}_0(t)$ can be simultaneously obtained through the 2-step procedures. Hence we can calculate $\hat{m}(t|\boldsymbol{z})$, which can not obtained through the approach of Jin et al. (2003). If we consider that two subjects A and B are both taking treatment with Level 2. Subject A has a karno of 55, while Subject B has a karno of 50. Then, $\hat{m}(t|\boldsymbol{z})$ for Subject A (treatment=Leval 2, karno=55) and Subject B (treatment=Leval 2, karno=50) can be calculated. The changes of expected remaining lifetime of these two subjects over time are shown in Figure 1.

Table 2: Parameter estimates and estimated standard errors for the lung cancer data

	Proposed model		Jin et al. method Gehan estimate		Jin et al. method Logrank estimate	
	treatment	karno	treatment	karno	treatment	karno
estimate	-0.084	0.706	-0.019	0.267	0.029[1]	0.351[1]
SE[2]	0.089	0.049	0.067	0.027	0.119	0.059
P-value	0.372	< 0.001	0.388	< 0.001	0.404	< 0.001

[1]The estimate is not unique

[2]SE is the standard derivation of estimates from resampling the data 100 times

Figure 1 shows the time scale changing effect and the proportional mean residual lifetime effect as we expected for the PSMRL model. It can be seen from Figure 1 that Subject A is always having a longer expected remaining lifetime than Subject B. For Subject A, the expected remaining lifetime increases with the days of survival before 230 days and decreases after 230 day. For Subject B, the expected remaining lifetime increases with the days of survival before 190 days and decreases after 190 day. The reason why the expected remaining lifetime increase for all subjects at an early time point might be the patients who entered this study were very sick and the survived ones were strong ones. Once they survived, the treatments would improve their expected remaining lifetime. The reason why the

expected remaining lifetime decreases for all subjects is that subjects would be expected to live shorter time when they are having a lung cancer for a longer time.

Now, let us interpret the estimate value 0.706 for the standardized karno under our approach. We transform it back to the original scale and obtain an estimate value 0.035 for the original karno. From Figure 1, we can see that Subject B is expected to live 55 more days if he or she has survived up to 500 days. We also can see that Subject A is expected to live 66 more days if he or she has survived up to 595 days. This also can be obtained from the following equation:

$$\frac{m(500e^{5*0.035}|Leval=2,karno=55)}{m(500|Leval=2,karno=50)}=\frac{m(595|Leval=2,karno=55)}{m(500|Leval=2,karno=50)}=e^{5*0.035}=1.19.$$

It can be concluded, a higher value of Karnofsky Performance Score a subject has, the longer the subject can live.

Let us interpret the estimate value 0.035 for the original karno under the AFT model. The lifetime (825 days)of Subject A with a karno of 55 is $e^{5*0.035}=1.19$ times of the lifetime (690 days) of Subject B. This interpretation is in accordance with the interpretation from our PSMRL model approach.

5 Discussion and Future Work

Unlike the PMRL model and the LMRL model, our PSMRL model is not restricted by any parameter constraints like $e^{\boldsymbol{z}^\tau\boldsymbol{\beta}}>1$ or $\boldsymbol{z}^\tau\boldsymbol{\beta}>0$. One of the advantages of our regression method is that we estimate the regression parameters and $\hat{m}_0(\cdot)$ simultaneously, hence we can obtain $\hat{m}(t|\boldsymbol{z}), \hat{S}_0(t), \hat{S}(t|\boldsymbol{z}), \hat{h}_0(t),$ and $\hat{h}(t|\boldsymbol{z})$, which can not be obtained through the approaches (e.g., Jin et al., 2003) for the AFT model. Another obvious advantage is that we can apply some existing estimation methods of the AFT model to our model, since our PSMRL model is equivalent to the AFT model. Another advantage is that our PSMRL model has a special physical interpretation for the regression parameters in addition to the same interpretation as the AFT model.

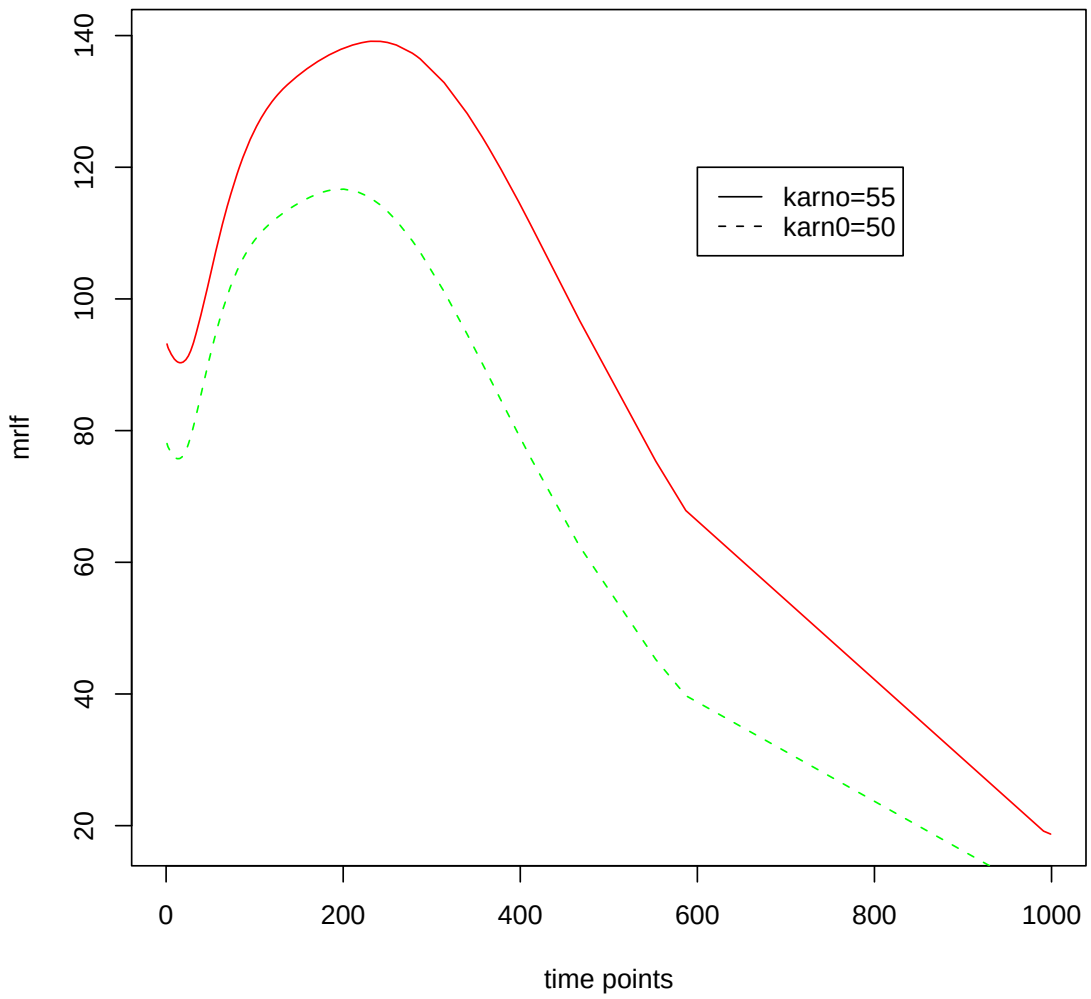

Figure 1: Comparison of the conditional mean residual life functions for the lung cancer data. The solid line is the cmrlf for people with treatment Level 2 and karno of 55. The dash line is the cmrlf for people with treatment Level 2 and karno of 50.

References

Chen, Y.Q. and Cheng, S. (2005). Semiparametric Regression Analysis of Mean Residual Life with Censored Survival Data. Biometrika 92, 19-29.

Chen, Y.Q. and Cheng, S. (2006). Linear Life Expectancy Regression with Censored Data. Biometrika 93, 303-313.

Chen, Y.Q. and Jewell, N.P. (2002). Inference for Proportional Mean Residual Life Model in Presence of Censoring.U.C. Berkeley Division of Biostatistics Working Paper Series, paper 108.

Chen, Y.Q., Jewell, N.P., Lei, X., and Cheng, S.C. (2005). Semiparametric Estimation of Proportional Mean Residual Life Model in Presence of Censoring. Biometrics 61, 170-178.

Cox, D.R. (1972). Regression Models and Life-tables (with discussion). Journal of Royal Statistical Society, Series B 34,187-220.

Elandt-Johnson, R.C. and Johnson, N.L. (1980). Survival Models and Data Analysis. Wiley, New York.

Ghorai, J., Susarla, A., Susarla, V., and Van-Ryzin, J. (1982). Nonparametric Estimation of Mean Residual Life Time with Censored Data. Nonparametric Statistical Inference. Vol.I, Colloquia Mathematica Societatis, 32. North-Holland, Amsterdam-New York, 269-291.

Hall, W.J. and Wellner, J.A. (1981). Mean Residual Life. Statistics and Related Topics. Edited by M. Csorgo, D.A. Dawson, J.N.K. Rao , and A.K.Md.E. Saleh. North-Holland, Amsterdam. 169-184.

Liu, S. and Ghosh, K. S. (2008). Nonparametric Estimation of Mean Residual Life Function Using Scale Mixtures. In Process.

Jin, Z., Lin, D. Y., and Ying, Z. (2003). Rank-based Inference for the Accelerated Failure Time Model. Biometrika, 90, 341-353.

Kalbfleisch, J.D. and Prentice, R.L. (1980). The Statistical Analysis of Failure Time Data. New York: Wiley.

Maguluri, G. and Zhang, C. H. (1994). Estimation in the Mean Residual Life Regression Model. J.R.Statist. Statist. Sinica 2, 17-46.

Oakes, D. and Dasu, T. (1990). A Note on Residual Life. Biometrika 77, 409-10.

Prentice, R.L. (1973). Exponential Survival with Censoring and Explanatory Variables. Biometrika 60, 279-288.

Yang, G.L. (1978). Estimation of a Biometric Function. Annals of Statistics 6, 112-116.

Appendix A

Derivation of the Loglikelihood Function:

The conditional distribution function $f(t|\boldsymbol{z}) = h(t|\boldsymbol{z})^{\Delta} S(t|\boldsymbol{z})$ is used to build the likelihood function, which indicates that $h(t|\boldsymbol{z})$ and $S(t|\boldsymbol{z})$ must be expressed in terms of $m(t|\boldsymbol{z})$. It is well known that there is a one-to-one relationship between the hazard function and the mrlf, i.e.,

$$h(t|\boldsymbol{z}) = \frac{m'(t|\boldsymbol{z}) + 1}{m(t|\boldsymbol{z})}. \tag{7}$$

Notice that under our PSMRL model, the conditional hazard function in (7) can be expressed as

$$h(t|\boldsymbol{z}) = \frac{m_0'(t\exp\{-\boldsymbol{z}^\tau\boldsymbol{\beta}\}) + 1}{\exp\{\boldsymbol{z}^\tau\boldsymbol{\beta}\} m_0(t\exp\{-\boldsymbol{z}^\tau\boldsymbol{\beta}\})}.$$

Also it follows from (6) that

$$S(t|\boldsymbol{z}) \;=\; \frac{m_0(0)}{m_0(t\exp\{-\boldsymbol{z}^\tau\boldsymbol{\beta}\})} \exp\left\{ -\int_0^{\exp\{-\boldsymbol{z}^\tau\boldsymbol{\beta}\}t} \frac{1}{m_0(v)} dv \right\}.$$

The loglikelihood function of $\boldsymbol{\beta}$ is then expressed as,

$$\begin{aligned}
ll(\boldsymbol{\beta}, m_0(\cdot)) &= \log\left(\prod_{i=1}^n f(X_i|\boldsymbol{Z}_i)\right) = \log\left(\prod_{i=1}^n h(X_i|\boldsymbol{Z}_i)^{\Delta_i} S(X_i|\boldsymbol{Z}_i)\right) \\
&= \sum_{i=1}^n \{\Delta_i \log h(X_i|\boldsymbol{Z}_i) + \log S(X_i|\boldsymbol{Z}_i)\} \\
&= n\log(m_0(0)) - \sum_{i=1}^n \Delta_i \boldsymbol{Z}_i^\tau\boldsymbol{\beta} - \sum_{i=1}^n (\Delta_i + 1)\log(m_0(X_i\exp\{-\boldsymbol{Z}_i^\tau\boldsymbol{\beta}\})) \\
&\quad + \sum_{i=1}^n \Delta_i \log(m_0'(X_i\exp\{-\boldsymbol{Z}_i^\tau\boldsymbol{\beta}\}) + 1) - \sum_{i=1}^n \int_0^{\exp\{-\boldsymbol{Z}_i^\tau\boldsymbol{\beta}\}X_i} \frac{1}{m_0(v)} dv.
\end{aligned}$$

Appendix B

Derivation of $m_0(\cdot)$ and its Closed Analytical Form:

The PSMRL model has the following property:

$$E[T-t|T>t,\boldsymbol{z}] = \exp\{\boldsymbol{z}^\tau\boldsymbol{\beta}\}m_0(t\exp\{-\boldsymbol{z}^\tau\boldsymbol{\beta}\})$$

$$\Rightarrow E\left[\frac{T}{\exp\{\boldsymbol{z}^\tau\boldsymbol{\beta}\}} - \frac{t}{\exp\{\boldsymbol{z}^\tau\boldsymbol{\beta}\}}\middle|\frac{T}{\exp\{\boldsymbol{z}^\tau\boldsymbol{\beta}\}} > \frac{t}{\exp\{\boldsymbol{z}^\tau\boldsymbol{\beta}\}}\right] = m_0\left(\frac{t}{\exp\{\boldsymbol{z}^\tau\boldsymbol{\beta}\}}\right)$$

$$\Rightarrow E\left[\frac{T}{\exp\{\boldsymbol{z}^\tau\boldsymbol{\beta}\}} - s\middle|\frac{T}{\exp\{\boldsymbol{z}^\tau\boldsymbol{\beta}\}} > s\right] = m_0\left(s\right).$$

Therefore, $m_0(\cdot)$ emerges as the mrlf of the transformed survival time $T^* = T\exp\{-\boldsymbol{z}^\tau\boldsymbol{\beta}\}$ for complete data. Since the empirical mrlf for complete data (Yang, 1978) and the empirical mrlf for the censored data (Ghorai et al., 1980) both converge to $m_0(t)$, $m_0(t)$ also emerges as the mrlf of the transformed survival time $\{(X_i^*, \Delta_i), i = 1, \ldots, n\}$, $\hat{m}_0(\cdot)$ can be calculated using the closed analytical form for the smooth estimator of the mrlf based on a single sample (Liu and Ghosh), which is given by

$$\begin{aligned}\hat{m}_0(t|\boldsymbol{\beta}) &= \sum_{l=0}^{n-1}\left[\left(\frac{\sum_{j=l+1}^{n} X_j^* w_j}{\sum_{j=l+1}^{n} w_j}\right)\left(F\left(X_{l+1}^*\middle|k_n, \frac{k_n}{t}\right) - F\left(X_l^*\middle|k_n, \frac{k_n}{t}\right)\right)\right] \\ &\quad -tF\left(X_n^*\middle|k_n+1, \frac{k_n}{t}\right), \qquad (8)\end{aligned}$$

where k_n is the function of n and $k_n \to \infty$ as $n \to \infty$, and the weight funciton is $w_j = F_n(X_j^*) - F_n(X_j^*-)$ and $F_n(X_j^*)$ is the KM estimate, and $F\left(\cdot\middle|k_n, \frac{t}{k_n}\right)$ is the cdf of $\mathrm{Ga}\left(k_n, \frac{t}{k_n}\right)$.