DIGITALCOMMONS @WAYNESTATE

Journal of Modern Applied Statistical Methods

Volume 1 | Issue 2 Article 45

11-1-2002

Robust Estimation Of Multivariate Failure Data With Time-Modulated Frailty

Pingfu Fu
Case Western Reserve University, pxf16@case.edu

J. Sunil Rao
Case Western Reserve University, sunil@hal.cwru.edu

Jiming Jiang
University of California

Journal of Modern Applied Statistical Methods
Fall 2002, Vol. 1, No 2, 367-378

1538 – 9472/02/$30.00

Robust Estimation Of Multivariate Failure Data With Time-Modulated Frailty

Pingfu Fu
Epidemiology & Biostatistics
Case Western Reserve University

J. Sunil Rao
Epidemiology & Biostatistics
Case Western Reserve University

Jiming Jiang
Department of Statistics
University of California

A time-modulated frailty model is proposed for analyzing multivariate failure data. The effect of frailties, which may not be constant over time, is discussed. We assume a parametric model for the baseline hazard, but avoid the parametric assumption for the frailty distribution. The well-known connection between survival times and Poisson regression model is used. The parameters of interest are estimated by generalized estimating equations (GEE) or by penalized GEE. Simulation studies show that the procedure is successful to detect the effect of time-modulated frailty. The method is also applied to a placebo controlled randomized clinical trial of gamma interferon, a study of chronic granulomatous disease (CGD).

Key words: Frailty models; multivariate failure data; generalized linear models.

Introduction

In the analysis of failure time data, one of the common assumptions made is that the life histories for subjects under study are statistically independent (at least conditionally on the observed fixed-time covariates). This assumption may be violated when individuals within some subgroup (e.g. siblings or parents in the same family, litter mates in animal study) share common unmeasured factors. Frailty models have been widely used for correlated survival data after Vaupel et. al. (1979) introduced the concept of frailty for making adjustments for the over-dispersion (heterogeneity) in their mortality study.

A frailty is an unobserved random effect shared by subjects within a subgroup. These include shared frailty (Hougaard, 1986a), bivariate frailty (Xue, 1998) as well as correlated frailty (Yashin, et. al. 1995), but few of them deal with time-dependent frailty (Self, 1995; Yau and McGilchrist, 1998). Most papers in the literature assume that individuals in the same cluster are born at a certain level of relative frailty and stay at this level through out life. As mentioned by Vaupel et. al. (1979), this may not be true in reality, for example, in human population mortality study, the frailty of an individual is large during an early period of life, after which it stabilizes, followed by an increasing frailty due to the natural aging process. For univariate frailty model, there are several limitations, for example, the model only allows positive correlations within the cluster, and the unobserved factor (frailty) is the same within the cluster (Xue, 1998).

Typically we assume that the frailty acts multiplicatively on each individual's hazard rate. We propose a time-modulated frailty model to analyze multivariate failure time data. The proposed model is more general than other frailty models, having as special members regular frailty models, such as shared frailty and bivariate frailty models if we ignore the time-modulated component in the model. Using the well-known connection to Poisson regression (Aitkin and Clayton, 1980), the derived model is a generalized linear mixed model (glmm). We adopt a robust approach for estimating some parameters using the generalized estimating equations (GEE) in this Poisson regression setting. For other parameters, the estimating procedures are equivalent to a generalized penalized estimating equations

Pingfu Fu is senior instructor, email address: pxf16@po.cwru.edu. This is a portion of Fu's PhD dissertation. J. Sunil Rao is Associate professor, email address: sunil@hal.cwru.edu. Jiming Jiang is Associate professor, e-mail address: jiang@wald.ucdavis.edu.

(GPEE). Under this approach, we do not specify the exact distribution of frailty and in this sense, our approach is robust.

Model construction

Self (1995) introduced a time-dependent frailty model

$$\lambda_i(t) = Y_i(t)\varsigma_i(t)\lambda_0(t)\exp(\beta' x_i(t)),$$

where $Y_i(t)$ *and* $x_i(t)$ are predictable scalar and p-vector value processes, respectively, $\varsigma_i(t)$ is a stationary stochastic process with positive, continuous sample paths, $\beta = (\beta_1, \beta_2, ..., \beta_p)$ and $\lambda_0(t)$ are unknown parameters. Instead of putting a stochastic process $\varsigma_i(t)$, a time-dependent frailty process, in the hazard function, we introduce an "interaction" term between the frailty and time as a time-modulated frailty. In the following sections, we will give the model formulation in two different settings.

Single-level of clustering

The most common situation in the multivariate survival data is the time to the recurrence of some chronic disease for a patient, for example, breast cancer, or survival of litters of rats, survival of twins, etc. All these can be thought to consist of single-level clustering of data. The survival times in each cluster (patient, litter, twins) are correlated and the survival times between the clusters are assumed independent. Let the triple (T_{ik}, δ_{ik}, x_{ik}) represent the data, where i is the cluster index (i = 1, …, n) consisting of correlated survival times T_{ik} (k = 1, … , n_i). Thus, the kth individual in the ith group is modeled as

$$\lambda_{ik}(t) = w_i(t)\lambda_0(t)\exp(\beta' x_{ik}),$$

where $w_i(t) = t^{\theta}\xi_i$ and θ is unknown parameter. Here ξ_i are realizations of a nonnegative random variable with density function g (ξ).

Assume E (ξ_i)= 1 (see Nielsen et. al., 1992) and $\mathrm{var}(\xi_i) = \sigma^2$ for the distribution of the frailty ξ_i. When $\theta = 0$, the model is a shared frailty model, $\xi_i\lambda_0(t)\exp(\beta' x_{ik})$. The above model can also be easily generalized to the correlated individual frailty model studied by Yashin et. al. (1995) by specifying $w_i(t) = \xi_i + t^{\theta}\eta_i$ and letting $n_i = 2$ and $\theta = 0$.

Multiple-levels of clustering

In some studies it may be reasonable to expect more than one level of within-cluster association. For example, the association between a parent and child versus that two siblings in studies of familial disease aggregation, or the durations inside and outside of hospitals for a patient who is admitted into a hospital several times for the same disease (Xue, 1998). The single-level clustering model can be extended to allow for grouping defined by multiple nested factors.

Again, suppose the data consists of the usual triple (T_{ijk}, δ_{ijk}, x_{ijk}), using i to index the clusters (litters, families) (i = 1, 2, …, n). Each cluster contains two distinguishable subgroups (j = 1, 2). Within each cluster, individuals have correlated survival times T_{ijk} for k = 1, …, n_{ij}. When n_{ij} = 1, then (T_{i11}, T_{i21}) is bivariate survival time, for example, as used in the adult Danish twins study (Hougaard et. al., 1992). We will assume the frailty acts multiplicatively on the individual's hazard with following form

$$\lambda_{ijk}(t) = w_{ij}(t)\lambda_0(t)\exp(\beta' x_{ijk}),$$

where $w_{ij}(t) = t^{\theta}\eta_{ij}$ and η_{i1}, η_{i2} are the realizations of two correlated random variables with nonnegative values (with joint density function h (u, v)). The η_{ij} is the frailty for the ith cluster and jth subgroup. The frailties can be characterized by a parametric bivariate distribution, for example,

$$(\log(\eta_1), \log(\eta_2)) \sim N(0,0;\sigma_1^2, \sigma_2^2, \sigma_{12}).$$

We also assume E (η_{ij}) = 1, i = 1, …, n, j = 1,2, $\mathrm{var}(\eta_{ij}) = \sigma_j^2$ and $\mathrm{cov}(\eta_1, \eta_2) = \sigma_1\sigma_2\rho$. If θ = 0, then it is a case studied by Xue (1998); if $\theta > 0$ or $\theta < 0$, then we can see that the effect of frailty increases or decreases as time increases.

As we can see from the model construction in both single-level and multiple-level of clustering cases, given the frailty, its effect on the hazard changes over time.

For the exponential model, the baseline cumulative hazard is $\Lambda_0(t) = t$, and the hazard function becomes $\lambda_{ijk}(t \mid w_{ij}) = t^{\theta}\eta_{ij}\exp(\beta' x_{ijk})$. For the Weibull model, $\Lambda_0(t) = t^{\nu}$, and the hazard function is

$$\lambda_{ijk}(t \mid w_{ij}) = t^{\theta}\eta_{ij}\nu t^{\nu-1}\exp(\beta' x_{ijk}).$$

We assume that observations between different clusters are independent and given the frailty w_{ij} (namely η_{i1} and η_{i2}), the observations in each cluster are conditionally independent. It can be shown that, approximately,

$$\delta_{ijk} \mid (\eta_{i1}, \eta_{i2}) \sim Poisson(\mu_{ijk}),$$

where

$$\mu_{ijk}(t) = e^{\beta' x_{ijk}} \int_0^t w_{ij}(u)\lambda_0(u)du.$$

The details are given in Appendix 1.

Robust estimation procedures

As described in Appendix 1, we can treat the censoring variable as a correlated Poison random variable with degree of over-dispersion depending on its mean. Since the full likelihood method is not feasible without numerical integration, and because of the intractability of the marginal likelihood function, we may apply the generalized estimating equations (GEE) approach (Liang and Zeger, 1986), which only requires the specification of the first two moments of the responses for each individual.

As mentioned by Hougaard (1984), the choice of the frailty distribution is crucial since the results for the survival population will be rather different with different frailties. In the following section, we will examine this robust approach, which only requires up to second-order of moments of the frailty distribution. It is robust in the sense that the full likelihood is not required and a fully parametric assumption for the frailty is avoided. The following procedures are for the single-level of clustering case, but they can be easily generalized to the multiple-level clustering case.

Exponential case

Estimation of coefficients

We assume that the baseline hazard is from exponential distribution. Given the frailty ξ_i as mentioned before, $\delta_{ik} \mid \xi_i \sim Poisson(\tilde{\mu}_{ik}\xi_i)$, where $\tilde{\mu}_{ik} = e^{\beta' x_{ik}} \frac{t_{ik}^{\theta+1}}{\theta+1}$. It is easy to get following quantities from the formulae for the multiple-level of clustering case (see Appendix 1).

$$E(\delta_{ik}) = \tilde{\mu}_{ik} = e^{\beta' x_{ik}} \frac{t_{ik}^{\theta+1}}{\theta+1},$$

$$\mathrm{var}(\delta_{ik}) = E(\tilde{\mu}_{ik}\xi_i) + \mathrm{var}(\tilde{\mu}_{ik}\xi_i) = \tilde{\mu}_{ik} + \tilde{\mu}_{ik}^2\sigma^2$$

and the unconditional covariance

$$\begin{aligned}\mathrm{COV}(\delta_{ik}, \delta_{il}) &= \mathrm{COV}(\tilde{\mu}_{ik}\xi_i, \tilde{\mu}_{il}\xi_i) \\ &= \tilde{\mu}_{ik}\tilde{\mu}_{il}\sigma^2, k \neq l, \\ \mathrm{cov}(\delta_{ik}, \delta_{i'l}) &= \mathrm{cov}(\tilde{\mu}_{ik}\xi_i, \tilde{\mu}_{il}\xi_{i'}) = 0, i \neq i'.\end{aligned}$$

In order to get the estimates of the regression parameters, we apply the quasi-likelihood score equations in spirit of GEE, i.e.

$$U_{\beta}(\beta, \theta) = \sum_{i=1}^{n} \left(\frac{\partial \tilde{\mu}_i}{\partial \beta}\right)' Var(Y_i)^{-1}(Y_i - \tilde{\mu}_i) = 0, \quad (1)$$

where $Y_i' = (\delta_{i1}, ..., \delta_{in_i})$ and $\tilde{\mu}_i' = (\tilde{\mu}_{i1}, ..., \tilde{\mu}_{in_i})$. Note that $Var(Y_i) = Var(Y_i; \beta, \theta)$, which depends on β and θ in the above equations. Thus, we need the estimating procedure for θ.

Estimation of time-modulated frailty parameters

The estimate of the variance component σ^2 is treated as nuisance parameter, which is estimated by a method of moments defined as

$$\hat{\sigma}^2 = \frac{\sum_{i,n_i>1}\sum_{k\neq k'}(\delta_{ik} - \hat{\mu}_{ik})(\delta_{ik''} - \hat{\mu}_{ik'}) + \sum_{i,n_i=1}[(\delta_{i1} - \hat{\mu}_{i1})^2 - \hat{\mu}_{i1}]}{\sum_{i,n_i>1}\hat{\mu}_{ik}\hat{\mu}_{ik'} + \sum_{i,n_i=1}\hat{\mu}_{i1}^2}.$$

The conditional likelihood function has form:

$$L_{ik}(\beta \mid \xi_i) = (\xi_i \tilde{\mu}_{ik})^{\delta_{ik}} e^{-\xi_i \tilde{\mu}_{ik}} [\frac{w_i(t_{ik})\lambda_0(t_{ik})}{\int_0^{t_{ik}} w_i(u)\lambda_0(u)du}]^{\delta_{ik}}$$

the second term in the above equation equals to $[\frac{(\theta+1)}{t_{ik}}]^{\delta_{ik}}$. Thus, the log of the likelihood function can be approximated as

$$l \approx l_Q(\beta,\theta) + \sum_{i,k} \delta_{ik}[\log(\theta+1) - \log(t_{ik})], \quad (2)$$

where $l_Q(\beta,\theta)$ is the log of the quasi-likelihood function for correlated Poisson variates.

We then introduce the penalized score equation for the θ,

$$\sum_{i=1}^{n} (\frac{\partial \tilde{\mu}_i}{\partial \theta})' Var(Y_i)^{-1}(Y_i - \tilde{\mu}_i) + \kappa(n)\sum_{i,k} \frac{\delta_{ik}}{\theta+1} = 0, \quad (3)$$

the equation (3) can be viewed as a regularized generalized estimating equation with a penalty term $\kappa(n)\sum_{i,k} \frac{\delta_{ik}}{\theta+1}$, where $\kappa(n) = n^{-\tau}$ with $\tau > 0$. When the tuning parameter $\kappa(n) = 1$, the left hand side of equation (3) is the partial derivative $\frac{\partial l}{\partial \theta}$. The estimators for β and θ can be obtained by iterating between (1) and (3).

Weibull case

When the baseline hazard is assumed to have a Weibull distribution, the model is more flexible by introducing an additional scale parameter v.

Estimation of coefficients

As before, given the frailty,

$$\delta_{ik} \mid \xi_i \sim Poisson(\tilde{\mu}_{ik}\xi_i),$$

where $\tilde{\mu}_{ik} = e^{\beta' x_{ik}} \frac{v}{\theta+v} t_{ik}^{\theta+v}$. Similarly, we have

$$E(\delta_{ik}) = \tilde{\mu}_{ik} = e^{\beta' x_{ik}} \frac{v}{\theta+v} t_{ik}^{\theta+v},$$

$$\text{var}(\delta_{ik}) = E(\tilde{\mu}_{ik}\xi_i) + \text{var}(\tilde{\mu}_{ik}\xi_i) = \tilde{\mu}_{ik} + \tilde{\mu}_{ik}^2\sigma^2$$

and the unconditional covariance

$$\begin{aligned} \text{cov}(\delta_{ik}, \delta_{il}) &= \text{cov}(\tilde{\mu}_{ik}\xi_i, \tilde{\mu}_{il}\xi_i) \\ &= \tilde{\mu}_{ik}\tilde{\mu}_{il}\sigma^2, k \neq l, \\ \text{cov}(\delta_{ik}, \delta_{i'l}) &= \text{cov}(\tilde{\mu}_{ik}\xi_i, \tilde{\mu}_{il}\xi_{i'}) = 0, i \neq i'. \end{aligned}$$

The estimate of the regression parameters can be obtained by the following generalized estimating equations, i.e.

$$U_\beta(\beta,\theta,v) = \sum_{i=1}^{n} (\frac{\partial \tilde{\mu}_i}{\partial \beta})' Var(Y_i)^{-1}(Y_i - \tilde{\mu}_i) = 0, \quad (4)$$

where $Y_i' = (\delta_{i1},...,\delta_{in_i})$ and $\tilde{\mu}_i' = (\tilde{\mu}_{i1},...,\tilde{\mu}_{in_i})$. Note that $Var(Y_i) = Var(Y_i;\beta,\theta,v)$ which depends on β, θ and v in the above equations, as mentioned in exponential case, we have to get the $n^{1/2}$-consistent estimates for v and θ.

Estimation of other parameters

The estimate of the variance component σ^2 is defined the same way as the exponential case:

$$\hat{\sigma}^2 = \frac{\sum_{i,n_i>1}\sum_{k\neq k'}(\delta_{ik} - \hat{\mu}_{ik})(\delta_{ik'} - \hat{\mu}_{ik'}) + \sum_{i,n_i=1}[(\delta_{i1} - \hat{\mu}_{i1})^2 - \hat{\mu}_{i1}]}{\sum_{i,n_i>1}\hat{\mu}_{ik}\hat{\mu}_{ik'} + \sum_{i,n_i=1}\hat{\mu}_{i1}^2}.$$

The conditional likelihood function in this case has a form

$$L_{ik}(\beta \mid \xi_i) = (\xi_i \tilde{\mu}_{ik})^{\delta_{ik}} e^{-\xi_i \tilde{\mu}_{ik}} [\frac{w_i(t_{ik})\lambda_0(t_{ik})}{\int_0^{t_{ik}} w_i(u)\lambda_0(u)du}]^{\delta_{ik}}$$

with the second term in the above equation equals to $[\frac{(\theta+\nu)}{t_{ik}}]^{\delta_{ik}}$. Thus, the log of the likelihood function can be approximated as

$$l \approx l_Q(\beta,\theta,\nu) + \sum_{i,k} \delta_{ik}[\log(\theta+\nu) - \log(t_{ik})], \tag{5}$$

where $l_Q(\beta,\theta,\nu)$ is the log of the quasi-likelihood function for correlated Poisson variates. If we re-parameterized $\theta+\nu$ as φ, then

$$\frac{\partial \tilde{\mu}_{ik}}{\partial \varphi} = e^{\beta' x_{ik}}[-\frac{\nu}{\varphi^2} t_{ik}^{\varphi} + \frac{\nu}{\varphi} t_{ik}^{\varphi} \log(t_{ik})]$$
$$= \tilde{\mu}_{ik}[\log(t_{ik}) - \frac{1}{\varphi}],$$

and

$$\frac{\partial \tilde{\mu}_{ik}}{\partial \nu} = e^{\beta' x_{ik}} \frac{t_{ik}^{\varphi}}{\varphi} = \frac{\tilde{\mu}_{ik}}{\nu}$$

Thus, we introduce the penalized score equations for φ as we did in the exponential case,

$$U_{\varphi} = \sum_{i=1}^{n} (\frac{\partial \tilde{\mu}_i}{\partial \varphi})' Var(Y_i)^{-1}(Y_i - \tilde{\mu}_i) + \kappa(n)\sum_{i,k} \frac{\delta_{ik}}{\varphi} = 0, \tag{6}$$

where the tuning parameter $\kappa(n) = n^{-\tau}$, $\tau > 0$ and when $\tau = 0$, the left hand side of equation (6) is $\frac{\partial l}{\partial \varphi}$. Because ν is unidentifiable from the score equations, we use plug-in estimate for it. Notice that, if we have estimates of φ and β, then, from equation $\tilde{\mu}_{ik} = e^{\beta' x_{ik}} \frac{\nu}{\varphi} t_{ik}^{\varphi}$. we can obtain the estimate of ν by following formula,

$$\hat{\nu} = \frac{1}{N}\sum_{i,k} \frac{\tilde{\mu}_{ik}\varphi}{e^{\beta' x_{ik}} t_{ik}^{\varphi}}, \tag{7}$$

which is moment estimate if we replace $\tilde{\mu}_{ik}$ by its sample mean.

In summary, we propose following algorithm for the estimates of β, φ, ν and θ,

1. Given initial values of φ, ν: $\varphi^{(0)}$, $\nu^{(0)}$, and fit Poisson regression by generalized estimating equations (4) using log link function with offset equals to $\log(\frac{\nu^{(0)}}{\varphi^{(0)}} t_{ik}^{\varphi^{(0)}})$, and obtain $\beta^{(0)}$ and get (update) $\hat{\tilde{\mu}}_{ik}^{(0)}$.
2. Update $\varphi^{(0)}$ from equation (6).
3. Update ν by following formula: $\hat{\nu}^{(1)} = \frac{1}{N}\sum_{i,k} \frac{\hat{\tilde{\mu}}_{ik}^{(0)}\varphi^{(1)}}{e^{\beta'^{(0)} x_{ik}} t_{ik}^{\hat{\varphi}^{(1)}}}$.
4. Go to step 1, 2, and 3 again until the convergence criteria is satisfied.

Because $\hat{\theta}$ is consistent and $\text{var}_J(\hat{\theta})$ is asymptotically unbiased (see the results in Appendix 2 and 3), we can use statistic $\frac{\hat{\theta}-\theta}{[\text{var}_J(\hat{\theta})]^{1/2}}$, which is asymptotically $N(0,1)$ for inference; thus, the null hypothesis $\theta = 0$ can be tested. If we reject the null hypothesis from the test, then we claim that the effect of time-modulated frailty exists. In the following sections, we examine our method by simulation followed by analyzing CGD dataset.

Simulations

There is a difficulty with conducting simulations in this setting, since it's difficult to generate correlated survival times with time-modulated frailties as we can see it in the specification of the hazard function which involves time-modulated frailties.

We generate datasets of correlated Weibull (without time-modulated frailty, i.e. $\theta = 0$) by using positive mixing distributions (Hougaard, 1986a) along with the random effects approach. Let T_{ik} be the survival times of observation k of individual (cluster) i conditional on an observed covariate Z_i. In this setup we

assume that the T_{ik}'s in different clusters are independent. Now assume Z to be positive stable with index α. The Laplace transform for Z is $E(\exp(-sZ)) = \exp(-s^{\alpha})$. If we now define another random variable Y_{ik} to be Weibully distributed with scale parameter $\exp(\beta' x_{ik})$ and shape parameter a, then $T_{ik} = Y_{ik} Z_i^{-/a}$. Thus the T_{ik}'s within a cluster are multivariate Weibull with Weibull margins having scale $\exp(\alpha\beta' x_{ik})$ and shape αa. The correlation between $\log(T_{ik})$ and $\log(T_{il})$ is then just $1-\alpha^2$ for $k \neq l$. The generation of positive stable variates Z_i can be done using Splus which employs Chambers et. al.'s (1976) algorithm.

Instead of choosing different values of index of positive stable random variable, different cluster size and different percentage of censoring, we just generate two datasets with clusters 50 and 150. In each cluster, there are 5 observations and the index of positive stable random variable $\tilde{\alpha} = 0.6$, the coefficient of the linear predictor $\tilde{\beta} = 3$ and the shape parameter of the Weibull $\tilde{\nu} = 2,$ thus, the marginal distribution of the correlated Weibull is still Weibull with shape parameter $\nu = 1.2$ and the scale parameter $\beta = 1.8$ (actually $\exp(x'\beta)$) where x is from the design matrix which is 1 or 0 depending whether a random number from standard normal is nonnegative or negative. The survival times are censored at fixed value to achieve 10% censoring. The estimates of parameters interested are the means of 100 replicates. The tuning parameter in the penalized score equation is $\kappa(n) = 1$ and $\kappa(n) = n^{-1/30}$ which is arbitrarily picked. We understand that the optimal choice of the tuning parameter may be selected by many methods, for example, the cross validation approach.

In this correlated Weibull case, as we know, there is no time-modulated frailty in it. We still assume the time-modulated frailty model, and the frailty term is in the form of $w_{ik}(t) = t_{ik}^{\theta}\eta_i$, and $\lambda_0(t) = \nu t^{\nu-1}$ is the baseline hazard from the Weibull distribution.

As we can see from Table 1, the parameter estimate of θ is not significant from 0 for two different values of tuning parameter which means we can not reject null hypothesis $\theta = 0$ based on asymptotic Wald type test. Thus, there does not appear to be a time-modulated frailty effect in this dataset. The estimates of β and ν are very close to the true values.

Table 1: Results of fitting the correlated Weibull by time -dependent frailty model with two values of $\kappa(n)$ in the penalized score equation, number of clusters = 50 and 100 simulations.

Parameter	BC (no BC) Estimate	GJ Standard error	t Value	Pr > |t|
$\kappa(n) = 1$:				
β	1.783 (1.827)	0.2718	6.560	< 0.0001
θ	0.001 (-0.097)	0.3463	0.001	0.9998
ν	1.208 (1.316)	0.4406	2.742	0.0061
φ	1.208 (1.219)	0.1094	11.042	< 0.0001
$\kappa(n)=n^{-1/30}$:				
β	1.645 (1.696)	0.2578	6.381	< 0.0001
θ	0.150 (0.093)	0.1704	0.879	0.3793
ν	0.936 (1.016)	0.2605	3.593	0.00033
φ	1.086 (1.109)	0.1012	10.73	< 0.0001
β (SN, 1993)	1.781	0.3852	4.624	< 0.0001

Note: The true value of β is 1.8 and 1.2 for ν. BC stands for bias corrected, GJ for grouped jackknife, and SN for Segal and Neuhaus.

The estimate of β by our procedure is consistent with other two approaches. From the variance estimates of β, there is small gain in term of efficiency although there is no time-modulated frailty effect in this case.

The biased estimates (values in the `no BC' column) overestimated the parameters when the tuning parameter $\kappa(n) = 1$, and underestimated when $\kappa(n) = n^{-1/30}$. The optimal tuning parameter τ may be a positive value that is very close to zero. We can do further simulation for large number of clusters and for different values of τ, as well as other parameters, such as different percentage of censoring, different value of index in the positive stable distribution. The results from Table 2 are more close to the true values, this is

because we have larger number of clusters (150 clusters) and the estimates of β, φ and θ are consistent.

Table 2: Results of fitting the correlated Weibull by time-dependent frailty model with two values of $\kappa(n)$ in the penalized score equation, number of clusters = 150 and 100 simulations.

Parameter	BC (no BC) Estimate	GJ Standard error	t Value	Pr > |t|
$\kappa(n)=1$:				
β	1.808 (1.822)	0.1526	11.85	< 0.0001
θ	0.000 (-0.013)	0.1166	0.004	0.9968
v	1.205 (1.215)	0.1783	6.758	0.0001
φ	1.205 (1.203)	0.0644	18.71	< 0.0001
$\kappa(n)=n^{-1/200}$:				
β	1.779 (1.792)	0.1508	11.8	< 0.0001
θ	0.034 (0.014)	0.1017	0.332	0.7396
v	1.147 (1.171)	0.1626	7.054	< 0.0001
φ	1.181 (1.185)	0.0635	18.6	< 0.0001
β (SN)	1.781	0.3852	4.624	< 0.0001

Note: The true value of β is 1.8 and 1.2 for ν. BC stands for bias corrected, GJ for grouped jackknife, and SN for Segal and Neuhaus.

A real data example

The well-known Chronic Granulomatous Disease (CGD) dataset, which is described in the Appendix D of the book by Fleming and Harrington (1991), has been analyzed by many authors. CGD is a group of inherited rare disorders of the immune function characterized by recurrent pyogenic infections, which usually present early life and may lead to death in childhood. Phagocytes from CGD patients ingest microorganisms normally but fail to kill them, primarily due to the inability to generate a respiratory burst dependent on the production of superoxide and other toxic oxygen metabolites. Thus, it is the failure to generate microbicidal oxygen metabolites within the phagocytes of CGD patients.

There is evidence that gamma interferon is an important macrophage activating factor which could restore superoxide anion production and bacterial killing by phagocytes in CGD patients. In order to study the ability of gamma interferon to reduce the rate of serious infections, a double-blinded clinical trial was conducted in which patients were randomized to placebo vs. gamma interferon. The data we use here, which is a little different from the one used by Fleming and Harrington (1991) in the example at page 162, has 65 patients in placebo group, 63 in gamma interferon group, of 30 placebo patients who experienced at least one infection, 4 experienced 2, 4 experienced 3, 1 experienced 4, 1 experienced 5 and 1 experienced 7; of 14 treatment patients who experienced at least one infection, 4 experienced 2 and 1 experienced 3.

It is reasonable to assume that the patients' frailties are time-modulated, since the risk of infection may increase once a first failure event occurs. In this data set, we treat each patient as a cluster, and the frailty term is in the form of $w_{ik}(t) = t_{ik}^{\theta}\xi_i$.

Table 3. Results of fitting the CGD dataset by proposed method with other two models.

Parameter	BC (no BC) Estimate	GJ Standard error	t Value	Pr > |t|
$\kappa(n)=1$:				
β	-0.835 (-0.856)	0.2588	-3.207	0.0013
θ	1.293 (1.321)	0.1995	6.481	< 0.0001
φ	1.328 (1.357)	0.1945	6.828	< 0.0001
v	0.035 (0.037)	0.0184	1.944	0.052
$\kappa(n)=n^{-1/30}$:				
β	-0.822 (-0.845)	0.2468	-3.332	0.0009
θ	1.116 (1.169)	0.1809	6.169	< 0.0001
φ	1.148 (1.204)	0.1736	6.613	< 0.0001
v	0.032 (0.034)	0.01461	2.204	0.0275
β (SN, 1993)	-0.856	0.2489	-3.4389	0.00058

Note: BC stands for bias corrected, GJ for grouped jackknife, and SN for Segal and Neuhaus.

Table 3 provides estimates of β with several methods, the estimates of other parameters followed by standard error for case of $\kappa(n) = 1$ by our time-modulated frailty model are $\hat{v}$ =0.035 (0.0184), $\hat{\theta} = 1.293$ (0.1995), $\hat{\varphi} = 1.328$ (0.1945).

The negative value of $\hat{\beta}$ = -0.8353 means that the treatment (gamma interferon) effectively reduces the recurrence of pyogenic infections as compare to the placebo. The estimate of β is consistent to that from other approaches.

From the estimates of θ and its variance, we can see that there is a time-modulated frailty effect in this dataset as noticed by Self (1995) though we have different model formulations. The parameter estimate of the time-modulated frailty $\hat{\theta} = 1.293$ is statistically significant from 0; the positive sign also means that given the frailty, its effect on the hazard is increasing as the life goes on.

The estimate of the treatment effect β is consistent with other two approaches; all of them indicate a statistically significant difference between the gamma interferon and placebo. The time-modulated frailty model does not seem to improve the efficiency, but the proposed model does help us to understand the nature of the frailty. In CGD case, the existence of effect of time-modulated frailty means that if a patient has a large frailty at the beginning, then (s)he will have an increasing chance of recurrence of pyogenic infections.

Conclusion

Few results about time-modulated frailty models are available in the literature (Yau and McGilchrist, 1998; Self, 1995). Our model provides one way to detect whether there is a trend in the hazard function with time given the frailty. Our model is different from Yau and McGilchrist's (1998), which assumes a different frailty for each time period of recurrence of disease; and different from Self's (1995) which introduces a stochastic process of frailty in the hazard function. The models proposed can also be extended in more general case, for example, in the multiple-level of clustering case, the time-modulated frailty can have the following form $w_{ij}(t) = \xi_i + t^{\theta}\eta_{ij}$, where $\xi_i, (\eta_{i1}, \eta_{i2})$ are independent realizations of two independent random variables with positive values. The resulting models are more complex than the one we proposed. To fit this model, we may use techniques of nonlinear mixed-effects models (Pinheiro and Bates, 2000).

Clinically speaking, the significance of the model is to realize whether there is an effect of time-modulated frailty in some diseases. If it does exist, for example, the pyogenic infection case (CGD data), it will tell us that more frail patients (say, have recurrence at the beginning) are more likely to have recurrence late in their life, which may suggest that those patients need more aggressive treatment (e.g. high dosage).

References

Aitkin, M., & Clayton, D. G. (1980). The fitting of exponential, Weibull and extreme value distributions to complex censored survival data using GLIM. *Applied Statistics*, *29*, 156-163.

Chambers, J. M., & Stuck, B. W. (1976). A method for simulating stable random variables. *Journal of the American Statistical Association*, *71*, 340-344.

Clayton, D. G. (1983). Fitting a general family of failure-time distributions using GLIM. *Applied Statistics*, *32*, 102-109.

Diggle, P. J., Liang, K.-Y., & Zeger, S. L. (1994). *Analysis of longitudinal data*, Clarendon: Oxford University Press.

Fleming, T. R., & Harrington, D. P. (1991). *Counting processes and survival analysis*. New York: John Wiley & Sons.

Hougaard P. (1984). Life table methods for heterogeneous populations: Distributions describing the heterogeneity. *Biometrika*, *71*, 75-83.

Hougaard P. (1986a). A class of multivariate failure time distributions. *Biometrika*, *73*, 671-678.

Hougaard P. (1986b). Survival models for heterogeneous population derived from stable distributions. *Biometrika*, *73*, 387-396.

Hougaard, P., Harvald, B., & Holm, N. V. (1992). Measuring the similarities between the lifetimes of adult Danish twins born between 1881-1930. *Journal of the American Statistical Association*, *87*, 17-24.

Jiang, J., & Zhang, W. (in press). Robust estimation in generalized linear mixed models. *Biometrika*.

Laird, N., & Oliver, D. (1981). Covariance analysis of censored survival data using log-linear analysis techniques. *Journal of the American Statistical Association*, *76*, 3231-240.

Liang, K.-Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, *73*, 13-22.

McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models.* (2nd ed.). London: Chapman and Hall.

McGilchrist, C. A., & Aisbett, C. W. (1991). Regression with frailty in survival analysis. *Biometrics*, *47*, 461-466.

Nielsen, G. G., Gill, R. D., Andersen, P. K., & Sorensen, T. I. A.. A counting process approach to maximum likelihood estimation in frailty models. *Scandinavian Journal of Statistics*, *19*, 25-43.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and Splus*. New York: Springer-Verlag.

Segal, M. R., & Neuhaus, J. M. (1993). Robust inference for multivariate survival data. *Statistics in Medicine*, *12*, 1019-1031.

Self. S. (1995). A regression model for counting processes with a time dependent frailty. Technical report. Seattle, WA: Fred Hutchinson Cancer Center.

Therneau, T. M., & Grambsch, P. M. (2000). *Modeling survival data: Extending the Cox model*. New York: Springer-Verlag.

Therneau, T. M., & Hamilton, S. A. (1997). rhDNase as an example of recurrent event analysis. *Statistics in Medicine*, *16*, 2029-2047.

Vaupel, J. W., Manton, K.G., & Stallard E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, *16*, 439-454.

Xue, X. (1998). Multivariate survival data under bivariate frailty: An estimating equation approach. *Biometrics*, *54*, 1631-1637.

Yashin, A. I., Vaupel, J. W., & Iachine, I. A. (1995). Correlated individual frailty: An advanced approach to survival analysis of bivariate data. *Mathematical Population Studies*, *5*, 145-159.

Yau, K. K. W., & McGilchrist, C. A.. (1998). ML and REML estimation in survival analysis with time dependent correlated frailty. *Statistics in Medicine*, *17*, 1201-1213.

Appendix 1: Likelihood and moments

Likelihood construction. For the model with multiple-levels of clustering, the hazard function is

$$\lambda_{ijk}(t) = w_{ij}(t)e^{\beta' x_{ijk}}\lambda_0(t).$$

Its corresponding density and survival functions:

$$f_{ijk}(t \mid w_{ij}) = w_{ij}(t)\lambda_0(t)\exp(\beta' x_{ijk})\exp(-e^{\beta' x_{ijk}}\int_0^t w_{ij}(u)\lambda_0(u)du),$$

and

$$S_{ijk}(t \mid w_{ij}) = \exp(-e^{\beta' x_{ijk}}\int_0^t w_{ij}(u)\lambda_0(u)du).$$

Thus, the contribution of the *i*th individual to the conditional likelihood given frailty w_{ij} is

$$\begin{aligned} L_{ijk}(\theta,\beta \mid w_{ij}) &= f_{ijk}^{\delta_{ijk}}(S_{ijk})^{(1-\delta_{ijk})} \\ &= [w_{ij}(t)\lambda_0(t)\exp(\beta' x_{ijk})]^{\delta_{ijk}}\exp(-e^{\beta' x_{ijk}} \\ &\quad \int_0^t w_{ij}(u)\lambda_0(u)du) \\ &= [e^{\beta' x_{ijk}}\int_0^t w_{ij}(u)\lambda_0(u)du]^{\delta_{ijk}}\exp(-e^{\beta' x_{ijk}} \\ &\quad \int_0^t w_{ij}(u)\lambda_0(u)du)[\frac{w_{ij}(t)\lambda_0(t)}{\int_0^t w_{ij}(u)\lambda_0(u)du}]^{\delta_{ijk}} \\ &= [\mu_{ijk}^{\delta_{ijk}}e^{-\mu_{ijk}}][\frac{w_{ij}(t)\lambda_0(t)}{\int_0^t w_{ij}(u)\lambda_0(u)du}]^{\delta_{ijk}}, \end{aligned}$$

where $\mu_{ijk} = e^{\beta' x_{ijk}}\int_0^t w_{ij}(u)\lambda_0(u)du$ and $w_{ij} = t^{\theta}\eta_{ij}$. Because T_{ijk} are conditionally independent given w_{ij}, therefore the conditional likelihood is

$$\begin{aligned} &L(\theta,\beta \mid \eta_1,\eta_2) \\ &= \prod\{\mu_{ijk}^{\delta_{ijk}}e^{-\mu_{ijk}}[\frac{w_{ij}(t_{ijk})\lambda_0(t_{ijk})}{\int_0^{t_{ijk}} w_{ij}(u)\lambda_0(u)du}]^{\delta_{ijk}}\}, \end{aligned}$$

and the conditional log likelihood:

$$\log(L) = \sum \{\delta_{ijk} \log(\mu_{ijk}) - \mu_{ijk} + \delta_{ijk}[\log(w_{ijk}(t_{ijk})) + \log(\lambda_0(t_{ijk})) - \log(\int_0^{t_{ijk}} w_{ij}(u)\lambda_0(u)du)]\}.$$

Therefore, from the above arguments, we have the following :

Result: Given the frailties η_{i1} and η_{i2}, δ_{ijk} can be thought as a Poisson random variable with mean μ_{ijk}. We will focus on the baseline hazard from Weibull distribution since it has a fairly flexible hazard function; baseline hazards from other distributions can be modeled by piece-wise exponential distribution which is a special case of Weibull. Assume the hazard function from Weibull distribution is $\lambda(t) = \phi\nu t^{\nu-1}$, here ϕ is a scale parameter, and ν is a shape parameter. The Weibull distribution is flexible enough to accommodate increasing (ν >1), decreasing (ν < 1) or constant hazard rate (ν = 1). When we have Weibull baseline distribution, the above log likelihood becomes

$$\log(L) = \sum \{\delta_{ijk} \log(\frac{\nu}{\theta+\nu}\eta_{ij} t_{ijk}^{\theta+\nu} e^{\beta' x_{ijk}}) - \frac{\nu}{\theta+\nu}\eta_{ij} t_{ijk}^{\theta+\nu} e^{\beta' x_{ijk}}$$
$$+ \delta_{ijk}[\log(t_{ijk}^{\theta}\eta_{ij}) + \log(\nu) - \log(\frac{\nu}{\theta+\nu}\eta_{ij} t_{ijk}^{\theta+1})]\}.$$

Since for Weibull distribution, the baseline hazard

is $\lambda_0(t) = \nu t^{\nu-1}$ and $\mu_{ijk} = e^{\beta' x_{ijk}} \frac{\nu}{\theta+\nu}\eta_{ij} t_{ijk}^{\theta+\nu}$,

because $\int_0^{t_{ijk}} w_{ij}(u)\lambda_0(u)du = \frac{\nu}{\theta+\nu}\eta_{ij} t_{ijk}^{\theta+\nu}$.

Moments of censoring indicator variable

Under the Weibull baseline survival function, the hazard function for observation *k* of individual *j* in cluster *i* is

$$\lambda_{ijk}(t_{ijk}) = t_{ijk}^{\theta}\eta_{ij} e^{\beta' x_{ijk}} \nu t_{ijk}^{\nu-1}.$$

By the assumption that, conditional on the frailties, censoring is not informative of the frailties (Nielsen et. al., 1992), we have $\delta_{ijk} \mid (\xi_i, \eta_{ij}) \sim Poisson(\mu_{ijk})$, where

$$\mu_{ijk} = e^{\beta' x_{ijk}} \int_0^{t_{ijk}} u^{\theta}\eta_{ij}\nu u^{\nu-1} du$$
$$= e^{\beta' x_{ijk}} \frac{\nu}{\theta+\nu}\eta_{ij} t_{ijk}^{\theta+\nu}.$$

Notice that, for fixed j, $w_{1j},...,w_{nj}$ are independent. Thus $\mathrm{cov}(w_{il}, w_{jl}) = 0$, where $i \neq j$. For fixed *i*,

$$\mathrm{cov}(w_{i1}, w_{i2}) = \mathrm{cov}(t_{i1k}^{\theta}\eta_{i1}, t_{i2k'}^{\theta}\eta_{i2})$$
$$= (t_{i1k} t_{i2k'})^{\theta}\sigma_1\sigma_2\rho.$$

1. Unconditional Mean:

$$\tilde{\mu}_{ijk} = E(\delta_{ijk})$$
$$= E[E(\delta_{ijk} \mid w_{ij})] = E(\mu_{ijk})$$
$$= e^{\beta' x_{ijk}} E(\frac{\nu}{\theta+\nu}\eta_{ij} t_{ijk}^{\theta+\nu})$$
$$= e^{\beta'' x_{ijk}} \frac{\nu}{\theta+\nu} t_{ijk}^{\theta+\nu}.$$

2. Unconditional Variance:

$$\mathrm{var}(\delta_{ijk}) =$$
$$E[\mathrm{var}(\delta_{ijk} \mid w_{ij})] + \mathrm{var}[E(\delta_{ijk} \mid w_{ij})]$$
$$= E(\mu_{ijk}) + \mathrm{var}(\mu_{ijk})$$
$$= e^{\beta' x_{ijk}} \frac{\nu}{\theta+\nu} t_{ijk}^{\theta+\nu}$$
$$+ e^{2\beta' x_{ijk}} \mathrm{var}(\frac{\nu}{\theta+\nu}\eta_{ij} t_{ijk}^{\theta+\nu})$$
$$= e^{\beta' x_{ijk}} \frac{\nu}{\theta+\nu} t_{ijk}^{\theta+\nu} + e^{2\beta' x_{ijk}} (\frac{\nu}{\theta+\nu} t_{ijk}^{\theta+\nu})^2 \sigma_j^2.$$

3. Unconditional Correlation (covariance):
If $k \neq k'$,, note that given $w_{ij}, T_{ijk}, T_{ijk'}$ are independent and conditional on the frailties, censoring is uninformative of the frailties.

$$\begin{aligned}
&\mathrm{cov}(\delta_{ijk}, \delta_{ijk'}) \\
&= E(\mathrm{cov}(\delta_{ijk}, \delta_{ijk'}) \mid w_{ij}] \\
&\quad + \mathrm{cov}[E(\delta_{ijk} \mid w_{ij}), E(\delta_{ijk'} \mid w_{ij})] \\
&= \mathrm{cov}(\mu_{ijk}, \mu_{ijk'}) \\
&= e^{\beta'(x_{ijk}+x_{ijk'})}\,\mathrm{cov}\left(\frac{\nu}{\theta+\nu}\eta_{ij}t_{ijk}^{\theta+\nu}, \frac{\nu}{\theta+\nu}\eta_{ij}t_{ijk'}^{\theta+\nu}\right) \\
&= e^{\beta'(x_{ijk}+x_{ijk'})}\left(\frac{\nu}{\theta+\nu}\right)^2 (t_{ijk}t_{ijk'})^{\theta+\nu}\sigma_j^2.
\end{aligned}$$

If $j \neq j'$,

$$\begin{aligned}
&\mathrm{cov}(\delta_{ijk}, \delta_{ij'k'}) = 0 + \mathrm{cov}(\mu_{ijk}, \mu_{ij'k'}) \\
&= e^{\beta'(x_{ijk}+x_{ij'k'})}\,\mathrm{cov}\left(\frac{\nu}{\theta+\nu}\eta_{ij}t_{ijk}^{\theta+\nu}, \frac{\nu}{\theta+\nu}\eta_{ij'}t_{ij'k'}^{\theta+\nu}\right) \\
&= e^{\beta'(x_{ijk}+x_{ij'k'})}\left(\frac{\nu}{\theta+\nu}\right)^2 (t_{ijk}t_{ij'k'})^{\theta+\nu}\sigma_1\sigma_2\rho.
\end{aligned}$$

If $i \neq i'$,

$$\begin{aligned}
&\mathrm{cov}(\delta_{ijk}, \delta_{i'j'k'}) \\
&= \mathrm{cov}(\mu_{ijk}, \mu_{i'j'k'}) \\
&= e^{\beta'(x_{ijk}+x_{i'j'k'})}\,\mathrm{cov}\left(\frac{\nu}{\theta+\nu}\eta_{ij}t_{ijk}^{\theta+\nu}, \frac{\nu}{\theta+\nu}\eta_{ij'}t_{i'j'k'}^{\theta+\nu}\right) \\
&= 0
\end{aligned}$$

Thus, δ_{ijk}'s can be treated as a sequence of correlated Poisson variables with over-dispersion since the variance of δ_{ijk} is not constant.

Appendix 2: Asymptotic properties

As we can see that the variance matrices in equation (1), (4) involve parameters besides β. Consistent estimate of β can be obtained by replacing θ and v with their $n^{-1/2}-$consistent estimates (Liang and Zeger, 1986) and the asymptotic properties are well established in this case. As stated in Liang and Zeger (1986), under mild regularity conditions, the estimate of $\hat{\beta}$ from the generalized estimating equation (1) and (4) is consistent and $n^{1/2}(\hat{\beta}-\beta)$ is asymptotically multivariate Gaussian as $n \to \infty$, where β is the true value. For the estimates of variance of θ, φ and v, we adopt the grouped jackknife approach because the exact formulae are not available. The estimates are bias corrected and the asymptotic properties for φ, and θ will be shown in the following section, thus, we can use Wald type statistic $\hat{\theta}^2 / \mathrm{var}_J(\hat{\theta})$, to test the existence of time-modulated frailty, where $\mathrm{var}_J(\hat{\theta})$ is grouped jackknife variance estimate for $\hat{\theta}$.

For the estimate of φ from the penalized score equation (3) or (6), under mild regularity conditions, we have following theorem and give a semi-rigorous proof.

Theorem 1. The estimate $\hat{\varphi}$ of φ is consistent and $n^{1/2}(\hat{\varphi}-\varphi)$ is asymptotically normal as $n \to \infty$ if $\max_i(n_i) < M$, where M is a known integer.

Proof: Under the true values of β, v and φ,

$$\begin{aligned}
&E\left(\frac{U_\varphi}{n}\right) \\
&= \frac{1}{n}E\left\{\sum_{i=1}^{n}\left(\frac{\partial\tilde{\mu}_i}{\partial\varphi}\right)\mathrm{Var}(Y_i)^{-1}(Y_i - \tilde{\mu}_i) + \kappa(n)\sum_{i,k}\frac{\delta_{ik}}{\varphi}\right\} \\
&= \frac{1}{n}\kappa(n)E\left(\sum_{i,k}\frac{\delta_{ik}}{\varphi}\right) = \frac{1}{n}\kappa(n)O(n) = o(1),
\end{aligned}$$

as $n \to \infty$ since and $\kappa(n) = n^{-\tau}, \tau > 0$. By the law of large numbers, we have $\frac{U_\varphi}{n} - E\left(\frac{U_\varphi}{n}\right) \to 0$, in probability as $n \to \infty$. Therefore, from the above two equations, $\frac{1}{n}U_\varphi = o_p(1)$. Thus, $\hat{\varphi}$ is consistent estimate of φ. The asymptotical normality of $\hat{\varphi}$ can be obtained following the

proof in the appendix of Liang and Zeger (1986). Q.E.D.

Because $\hat{\nu}$ is moment estimate which is consistent and $\varphi = \nu + \theta$, thus, $\hat{\theta}$ is also consistent.

Appendix 3: Jackknife variance estimation and bias correction

For the parameter β, we can use the robust estimate building in the existing procedure. The parameter θ is indicator of the effect of time-modulated frailty, and it is our interest to see whether this effect exist, thus we cannot treat it as a nuisance parameter. First, we notice that the estimate of θ is not unbiased because of the penalty term in equation (3) or (6) and

$$E\frac{\partial l}{\partial \theta} \neq 0$$

(Page 28, McCullagh and Nelder, 1983). We will obtain the variance estimate as well an estimation of bias by grouped jackknife method (Therneau and Hamilton, 1997).

The grouped jackknife procedure is the following: Each time we delete the observations from each cluster (or a patient), say cluster i, and obtain the estimate, say $\hat{\theta}_{(i)}$, by applying above estimating procedure to the rest of the data. Let $\hat{\theta}$ be the estimate based on the all the observations, then the grouped jackknife estimation of variance for θ is

$$\mathrm{var}_J(\hat{\theta}) = \frac{(n-1)}{n}\sum_{i=1}^{n}(\hat{\theta}_{(i)} - \hat{\theta}_{(.)})^2,$$

where $\hat{\theta}_{(.)} = \sum_i \hat{\theta}_{(i)} / n.$

The bias estimate for θ is $\hat{B}_\theta = (n-1)(\hat{\theta}_{(.)} - \hat{\theta}))$. Thus, the bias corrected estimate for θ is

$$\tilde{\theta} = \hat{\theta} - \hat{B}_\theta = n\hat{\theta} - (n-1)\hat{\theta}_{(.)}.$$

The reason that we apply the grouped jackknife procedure is that we have correlated observations in each cluster and the observations from different clusters are independent.

Theorem 2. Under suitable conditions, the grouped jackknife estimates $\mathrm{var}_J(\hat{\varphi})$ and $\mathrm{var}_J(\hat{\theta})$ are asymptotically unbiased estimates of the variance of $\hat{\varphi}$ and variance of $\hat{\theta}$.

Proof: The arguments are similar to Grambsch and Therneau (2000). Q.E.D.