Journal of the Korean Data & Information Science Society
2009, **20**(5), 895–903

한국데이터정보과학회지

A correction of SE from penalized partial likelihood in frailty models†

Il Do Ha[1]

[1]Department of Asset Management, Daegu Haany University

Received 17 May 2009, revised 30 July 2009, accepted 7 August 2009

Abstract

The penalized partial likelihood based on restricted maximum likelihood method has been widely used for the inference of frailty models. However, the standard-error estimate for frailty parameter estimator can be downwardly biased. In this paper we show that such underestimation can be corrected by using hierarchical likelihood. In particular, the hierarchical likelihood gives a statistically efficient procedure for various random-effect models including frailty models. The proposed method is illustrated via a numerical example and simulation study. The simulation results demonstrate that the corrected standard-error estimate largely improves such bias.

Keywords: Frailty Models, hierarchical likelihood, marginal likelihood, penalized partial likelihood, random effects, restricted maximum likelihood.

1. Introduction

Frailty models, Cox's (1972) proportional hazard models allowing random-effect terms, have been widely used for the analysis of various correlated survival data (Hougaard, 2000; Duchateau and Janssen, 2008). Here, the frailty means an unobserved random effect in the hazard models. For the inferences the marginal likelihood, which is obtained by integrating out the frailties, has been often used (Nielsen *et al.*, 1992; Vaida and Xu, 2000). However, the integration is generally intractable except for a shared gamma frailty model (Nielsen *et al.*, 1992; Ha *et al.*, 2001; Ha, 2006).

As an alternative, the penalized partial likelihood (PPL) has been suggested (McGilchrist 1993; Ripatti and Palmgren, 2000), which is constructed using partial likelihood (Cox, 1972; Breslow, 1974) and a penality function (e.g. gamma or normal density) for frailty. The implementation of PPL method is relatively easy because the difficult integration is not required. The PPL method based on restricted maximum likelihood (REML) in the normal frailty model (McGilchrist, 1993) works well. However, the variance estimate of PPL frailty-parameter estimator is substantially underestimated (Ripatti and Palmgren,

† This work was supported by the Korea Research Foundation Grant funded by the Korean Government (KRF-2008-521-C00057).

[1] Professor, Department of Asset Management, Daegu Haany University, Gyeongsan 712-715, Korea. E-mail: idha@dhu.ac.kr

2000) due to ignoring an important derivative term in Section 3.2, even though that of regression parameter estimator performs well. The hierarchical likelihood (h-likelihood; Lee and Nelder, 1996; Ha *et al.*, 2001) is defined from the joint density function of response variable and unobserved random effect, and also avoids the integration itself as in the PPL. In particular, the h-likelihood provides a statistically efficient estimation procedure for various random-effect models including frailty models (Lee *et al.*, 2006; Ha, 2008). Thus, in this paper we propose how to correct such underestimation using h-likelihood.

The paper is organized as follows. In Section 2 we review the PPL method in frailty models. In Section 3 we study the relationship between PPL and the h-likelihood and then show how to correct the variance of frailty-parameter estimator. The proposed method is demonstrated with a numerical example based on a well-known real data set in Section 4, followed by simulation study. Finally, some remarks are given in Section 5.

2. Penalized partial likelihood in frailty models

2.1. Frailty models

Assume that data consist of censored time-to-event observations from q subjects, each with n_i observations, $i = 1, \dots, q$. Let T_{ij} be the survival time for the j th observation of the i th subject and C_{ij} be the corresponding censoring time ($i = 1, \dots, q,\ j = 1, \dots, n_i,\ n = \sum_i n_i$). Let the observable random variables be

$$y_{ij} = \min(T_{ij}, C_{ij}) \text{ and } \delta_{ij} = I(T_{ij} \le C_{ij}),$$

where $I(\cdot)$ is the indicator function. Denote by v_i the unobserved log-frailty (or random effect) for the i th subject. Given $v_{i,}$ the conditional hazard function of T_{ij} takes the form

$$\lambda_{ij}(t|v_i) = \lambda_0(t)\exp(\eta_{ij}), \tag{2.1}$$

where $\lambda_0(t)$ is an unspecified baseline hazard function, and

$$\eta_{ij} = x_{ij}^T\beta + v_i$$

is the linear predictor for the hazards, β is a $p \times 1$ vector of unknown regression parameters corresponding to fixed covariates $x_{ij} = (x_{ij1}, \dots, x_{ijp})^{\mathrm{T}}$. Here, the term $x_{ij}^T\beta$ in (2.1) does not include an intercept term because of identifiable purposes. In this paper, for the distribution of independent random effects v_i 's we assume a normal distribution with mean $E(v_i) = 0$ and var $(v_i) = \alpha$ (McGilchrist, 1993; Ha *et al.*, 2001). In particular, the normal assumption is very useful for modelling multi-component (Ha *et al.*, 2007) or correlated frailties (Vaida and Xu, 2000). For the v_i other frailty distribution such as log-gamma can be assumed (Hougaard, 2000).

Let C_{ij} be censoring time corresponding to survival time T_{ij}. Then we have the observable random variables:

$$y_{ij} = \min(T_{ij}, C_{ij}) \ and\ \delta_{ij} = I(T_{ij} \le C_{ij}),$$

where $I(\cdot)$ is the indicator function. Since the functional form of $\lambda_0(t)$ is unknown, following Breslow (1972, 1974) and Ha *et al.* (2001) we consider the baseline cumulative hazard function $\Lambda_0(t)$ to be a step function with jumps at the r distinct observed death times,

$$\Lambda_0(t) = \sum_{k:y_{(k)}\leq t} \lambda_{0k},$$

where $y_{(k)}$ is the k th ($k = 1,\cdots,r$) smallest distinct death time among the y_{ij}'s, and $\lambda_{0k} = \lambda_0(y_{(k)})$.

2.2. Penalized partial likelihood

For the inference of model (2.1) Ripatti and Palmgren (2000) and Therneau and Grambsch (2000) proposed the use of PPL, defined by

$$\ell_{PPL} = \ell_{PL} - \ell_{Pen}, \tag{2.2}$$

where

$$\ell_{PL} = \sum_{ij} \delta_{ij}\eta_{ij} - \sum_{k} d_k \log\left\{\sum_{(i,j)\in R_{(k)}} \exp(\eta_{ij})\right\}$$

is the partial log-likelihood of Breslow (1974) and

$$\ell_{Pen} = -\sum_{i} \ell_{2i}$$

is a penality function having the logarithm of the normal density function for v_i with parameter α

$$\ell_{2i} = \ell_{2i}(\alpha; v_i) = -\frac{1}{2}\log(2\pi\alpha) - \frac{1}{2\alpha}v_i^2. \tag{2.3}$$

Here $d_{(k)}$ is the number of deaths at $y_{(k)}$ and $R_{(k)} = R(y_{(k)}) = \{(i,j) : y_{ij} \geq y_{(k)}\}$ is the risk set at $y_{(k)}$.

The estimation procedure of fixed parameters (β, α) and random effects $v = (v_{1,} \ldots, v_q)^T$ is as follows. Given α, the estimation of fixed and random effects, $\tau = (\beta, v)^T$, is obtained by solving

$$\partial\ell_{PPL}/\partial\tau = 0. \tag{2.4}$$

Note that the asymptotic covariance matrix of $\widehat{\beta}$ is obtained from the inverse of information matrix, $-\partial^2\ell_{PPL}/\partial\tau^2$.

For the estimation of α the PPL method uses the first-order Laplace approximation (Therneau and Grambsch, 2000) for the modified marginal likelihood

$$m_P = \log\left\{\int \exp(\ell_{PPL})dv\right\}.$$

That is, as an approximation of m_P Ripatti and Palmgren (2000) proposed to use

$$\ell_{PPL1} = [\ell_{PPL} - \frac{1}{2}\log|K^{''}(v)|]|_{v=\widehat{v}} \tag{2.5}$$

where $K^{''}(v) = -\partial^2\ell_{PPL}/\partial v^2$ and $\widehat{v}$ solves $\partial\ell_{PPL}/\partial v = 0$ in (2.4). McGilchrist (1993) also suggested a restricted likelihood to use $K^{''}(\tau)$, instead of $K^{''}(v)$ in (2.5), defined by

$$\ell_{PPL2} = [\ell_{PPL} - \frac{1}{2}\log|K^{''}(\tau)|]|_{\tau=\widehat{\tau}} \tag{2.6}$$

where $K^{''}(\tau) = -\partial^2\ell_{PPL}/\partial\tau^2$ and $\widehat{\tau}$ solves $\partial\ell_{PPL}/\partial\tau = 0$. Note that the maximizations of ℓ_{PPL1} and ℓ_{PPL2} give an approximated ML and REML estimators, respectively:

$$\widehat{\alpha}_{PPL1} = \sum_i \widehat{v}_i/(q - r_1) \ \textit{and} \ \ \widehat{\alpha}_{PPL2} = \sum_i \widehat{v}_i/(q - r_2), \tag{2.7}$$

where $r_1 = \text{tr}(K_{11})/\widehat{\alpha}$ with $K_{11} = -\partial^2\ell_{PPL}/\partial v^2$, and $r_2 = \text{tr}(K_{22})/\widehat{\alpha}$ and K_{22} is the matrix given by the bottom right-hand corner of the inverse of $K^{''}(\tau)$. Following McGilchrist (1993) and Ripatti and Palmgren (2000), the corresponding asymptotic variances are, respectively, given by

$$\text{var}(\widehat{\alpha}_{PPL1}) = 2\alpha^2[q - 2r_1 + \alpha^{-2}\text{tr}(K_{11}^2)]^{-1}, \tag{2.8}$$

$$\text{var}(\widehat{\alpha}_{PPL2}) = 2\alpha^2[q - 2r_2 + \alpha^{-2}\text{tr}(K_{22}^2)]^{-1}. \tag{2.9}$$

McGilchrist (1993) showed via simulation studies that $\widehat{\alpha}_{PPL2}$ (i.e. REML estimator) provides better results in terms of biases, especially in the frailty parameter: see also Ha and Lee (2003). Thus, in this paper we consider the REML estimator for frailty parameter. However, the estimate of $\text{var}(\widehat{\alpha}_{PPL2})$ is still downwardly biased as in $\text{var}(\widehat{\alpha}_{PPL1})$ (McGilchrist, 1993; Ripatti and Palmgren, 2000). We shall show how to correct the REML variance below.

3. Correction of SE in PPL

In this section we study the relationship between PPL and the h-likelihood. From this we propose how to correct the PPL variance in (2.9).

3.1. Relationship between PPL and h-likelihood

Following Ha *et al.* (2001), the h-likelihood for frailty model (2.1) is defined by

$$h = h(\beta, \lambda_0, \alpha) = \sum_{ij}\ell_{1ij} + \sum_i \ell_{2i}, \tag{3.1}$$

where

$$\begin{aligned}\sum_{ij}\ell_{1ij} &= \sum_{ij}\delta_{ij}\left\{\log\lambda_0(y_{ij}) + \eta_{ij}\right\} - \sum_{ij}\Lambda_0(y_{ij})\exp(\eta_{ij}) \\ &= \sum_k d_{(k)}\log\lambda_{0k} + \sum_{ij}\delta_{ij}\eta_{ij} - \sum_k \lambda_{0k}\left\{\sum_{(i,j)\in R_{(k)}}\exp(\eta_{ij})\right\},\end{aligned}$$

$\ell_{1ij} = \ell_{1ij}(\beta, \lambda_0; y_{ij}, \delta_{ij}|v_i)$ is the logarithm of the conditional density function for y_{ij} and δ_{ij} given v_i, $\ell_{2i} = \ell_{2i}(\alpha; v_i)$ is given in (2.3) and $\lambda_0 = (\lambda_{01}, \dots, \lambda_{0r})^T$.

We now show the relationship between PPL and h-likelihood. Since the dimension of λ_0 increases with sample size n, for the estimation of (β, v) Ha *et al.* (2001) proposed to use the profile h-likelihood h^* with λ_0 eliminated:

$$h^* = h|_{\lambda_0 = \widehat{\lambda_0}}, \tag{3.2}$$

where

$$\widehat{\lambda}_{0k} = \frac{d_{(k)}}{\sum_{(i,j)\in R(y_{(k)})} \exp(x_{ij}^T\beta + v_i)} \tag{3.3}$$

are solutions of the estimating equations, $\partial h/\partial\lambda_{0k} = 0$, for $k = 1, \dots, r$. Substituting (3.1) and (3.3) into (3.2) gives

$$h^* = \sum_{ij} \delta_{ij}\eta_{ij} - \sum_k d_k \log\left\{ \sum_{(i,j)\in R_{(k)}} \exp(\eta_{ij}) \right\} + \sum_k d_{(k)} \left\{\log d_{(k)} - 1\right\} + \sum_i \ell_{2i}, \tag{3.4}$$

From (3.2) and (3.4) we see that h^* is proportional to the PPL

$$h^* = \ell_{PPL} + \sum_k d_{(k)} \left\{\log d_{(k)} - 1\right\}.$$

Since $\partial h^*/\partial\tau = \partial\ell_{PPL}/\partial\tau = 0$, for the estimation of (β, v) given α we also see that the PPL and h-likelihood lead to the same estimation results. Furthermore, we have that $-\partial^2 h^*/\partial\tau^2 = -\partial^2\ell_{PPL}/\partial\tau^2$. In particular, Ha *et al.* (2001) and Ha and Lee (2003) showed that the inverse of Hessian matrix $H^* = -\partial^2 h^*/\partial\tau^2$ gives the asymptotic covariance matrix of $\widehat{\beta}$ and $\widehat{v} - v$, given by

$$H^*(\beta, v) = -\begin{pmatrix} \partial^2 h^*/\partial\beta^2 & \partial^2 h^*/\partial\beta\partial v \\ \partial^2 h^*/\partial v\partial\beta & \partial^2 h^*/\partial v^2 \end{pmatrix} = \begin{pmatrix} X^T W^* X & X^T W^* Z \\ Z^T W^* X & Z^T W^* Z + R \end{pmatrix} \tag{3.5}$$

where X is the $n \times p$ matrix whose i th row vector is x_{ij}^T, Z is the $n \times q$ group indicator matrix whose i th row vector is z_{ij}^T, $W^* = W^*(\beta, v)$ is the $n \times n$ weight matrix given in Appendix 2 of Ha and Lee (2003), and $R = \text{diag}\left\{-\partial^2\ell_{2i}/\partial v_i^2\right\} = \alpha^{-1} I_q$ is the $q \times q$ diagonal matrix with q dimensional identity matrix I_q. Note that the upper left-hand corner of H^{-1} in (3.5), provides a covariance matrix of $\widehat{\beta}$, given by

$$\text{var}(\widehat{\beta}) = (X^T V^{-1} X)^{-1} \textit{ with } V = W^{*-1} + Z R^{-1} Z^T. \tag{3.6}$$

Note that this covariance matrix is the same as that of PPL estimator vector of β, which performs well (Ha *et al.*, 2001; Ha and Lee, 2003, 2005).

Next, for the estimation of the frailty parameter α we use Lee and Nelder's (2001) adjusted profile h-likelihood (i.e. extended restricted likelihood), defined by

$$p_\tau(h^*) = [h^* - \frac{1}{2}\log\det\left\{H^*(\tau)/(2\pi)\right\}]|_{\tau=\widehat{\tau}}, \tag{3.7}$$

where $H^*(\tau) = -\partial^2 h^*/\partial\tau^2$ and $\widehat{\tau}$ solves $\partial h^*/\partial\tau = 0$. Note here that $\widehat{\tau} = \widehat{\tau}(\alpha) = (\widehat{\beta}(\alpha), \widehat{v}(\alpha))^T$. Thus, the $p_\tau(h^*)$ is the function of α only which eliminate β and v from h^*, a profile h-likelihood from which the nuisance parameters λ_0 have already been eliminated. The h-likelihood estimator for α, a REML estimator, is obtained by solving

$$\partial p_\tau(h^*)/\partial\alpha = 0. \tag{3.8}$$

This also yields McGilchrist's (1993) REML estimator of α (Ha *et al.*, 2001).

In summary, the h-likelihood estimator and its variance for β are the same as those of PPL. The h-likelihood estimator for α is also the same as PPL estimator based on the REML.

3.2. Corrected SE

Following the relationship between h-likelihood and PPL, the variance of $\widehat{\alpha}_{PPL2}$ in (2.9) can be derived by ignoring the $\partial\widehat{\tau}/\partial\alpha$ term in computing $-\partial^2 p_\tau(h^*)/\partial\alpha^2$, leading to an underestimation of the true variance of $\widehat{\alpha}_{PPL2}$.

We now show how to correct it using $p_\tau(h^*)$ in (3.7). The $p_\tau(h^*)$ can be written as

$$p_\tau(h^*) = \widetilde{h} - \frac{1}{2}\log\det(\widetilde{H}) \tag{3.9}$$

where $\widetilde{h} = h^*|_{\tau=\widehat{\tau}} = h^*(\widehat{\tau}(\alpha), \alpha)$ and $\widetilde{H} = H^*|_{\tau=\widehat{\tau}} = H^*(\widehat{\tau}(\alpha), \alpha)$ depend on α, and $c = \{(p+q)/2\}\log(2\pi)$ is a constant. Following Lee and Nelder (2001) and Ha and Lee (2003), we allow for $\partial\widehat{v}/\partial\alpha$ in computing the $\partial^2 p_\tau(h^*)/\partial\alpha^2$ terms in (3.9), but not for $\partial\widehat{\beta}/\partial\alpha$. Then the derivation and computation procedure of the Hessian matrix, $-\partial^2 p_\tau(h^*)/\partial\alpha^2$, is as follows.

$$-\frac{\partial^2 p_\tau(h^*)}{\partial\alpha^2} = -\frac{\partial^2\widetilde{h}}{\partial\alpha^2} + \frac{1}{2}\mathrm{tr}\left\{\widetilde{H}^{-1}\frac{\partial^2\widetilde{H}}{\partial\alpha^2} - (\widetilde{H}^{-1}\frac{\partial\widetilde{H}}{\partial\alpha})(\widetilde{H}^{-1}\frac{\partial\widetilde{H}}{\partial\alpha})\right\}. \tag{3.10}$$

Following Appendix 2 of Ha *et al.* (2001), we have that

$$\begin{aligned}\frac{\partial\widetilde{h}}{\partial\alpha} &= \left\{\frac{\partial h^*}{\partial\alpha} + \frac{\partial h^*}{\partial v}\cdot\frac{\partial\widehat{v}}{\partial\alpha}\right\}\Big|_{v=\widehat{v}} \\ &= \frac{\partial h^*}{\partial\alpha}\Big|_{v=\widehat{v}}\end{aligned}$$

since $\frac{\partial h^*}{\partial v}|_{v=\widehat{v}} = 0$. Thus, the first term on the right hand side of (3.10) is given by

$$\frac{\partial^2\widetilde{h}}{\partial\alpha^2} = \left\{\frac{\partial^2 h^*}{\partial\alpha^2} + \frac{\partial^2 h^*}{\partial\alpha\partial v}\cdot\frac{\partial\widehat{v}}{\partial\alpha}\right\}\Big|_{v=\widehat{v}}.$$

Here, since $\frac{\partial h^*}{\partial\alpha} = \frac{\partial}{\partial\alpha}(\sum_i \ell_{2i})$ we obtain $\frac{\partial^2 h^*}{\partial\alpha^2} = \sum_i(\frac{1}{2}\alpha^{-2} - \alpha^{-3}v_i^2)$ and $\frac{\partial^2 h^*}{\partial\alpha\partial v} = \alpha^{-2}v$.

Note that following Lee *et al.* (2006, pp. 316) and (3.5), we also have that

$$\frac{\partial \widehat{v}}{\partial \alpha} = -\left(\frac{-\partial^2 h^*}{\partial v^2}\right)^{-1}\left(\frac{-\partial^2 h^*}{\partial v \partial \alpha}\right)\Big|_{v=\widehat{v}} = -(Z^T\widehat{W}Z+\widehat{R})^{-1}(-\alpha^{-2}\widehat{v}), \tag{3.11}$$

where $\widehat{W} = W^*|_{v=\widehat{v}} = W^*(\widehat{v}(\alpha),\alpha)$ and $\widehat{R} = \alpha^{-1}I_q = R$. The second term of (3.10) is calculated with

$$\frac{\partial \widetilde{H}}{\partial \alpha} = \begin{pmatrix} X^TW'X & X^TW'Z \\ Z^TW'X & Z^TW'Z+R' \end{pmatrix} \; and \; \frac{\partial^2 \widetilde{H}}{\partial \alpha^2} = \begin{pmatrix} X^TW''X & X^TW''Z \\ Z^TW''X & Z^TW''Z+R'' \end{pmatrix},$$

where $W' = \partial\widehat{W}/\partial\alpha$, $R' = -\alpha^{-2}I_q$, $W'' = \partial^2\widehat{W}/\partial\alpha^2$ and $R'' = 2\alpha^{-3}I_q$. Notice that for the computation of W' and W'', the $\partial\widehat{v}/\partial\alpha$ and $\partial^2\widehat{v}/\partial\alpha^2$ terms should be considered. Thus, the corrected variance of $\widehat{\alpha}$ is obtained from the inverse of $-\partial^2 p_\tau(h^*)/\partial\alpha^2$ based on (3.10) and (3.11). That is, the corresponding SE is given by $\sqrt{\{-\partial^2 p_\tau(h^*)/\partial\alpha^2\}^{-1}}$. We investigate its performance by simulation below.

4. Illustration

For the illustration, we compare the performance of the corrected SE in (3.10) with that of the PPL's SE (i.e. standard SE) based on REML in (2.9). For this we present a numerical example and a simulation study. For the model fitting and computation we used SAS/IML.

4.1. Numerical example

The kidney infection data (McGilchrist and Aisbett, 1991) consist of times to the first and second recurrences of infection in 38 kidney patients using a portable dialysis machine. Here, each survival time is time to infection since insertion of the catheter. The survival times from the same patient are likely to be related because of frailty describing the patient's effect. We use a single covariate, the sex of the patients, coded as 1 for male and 0 for female. The estimation results are presented in Table 4.1.

Table 4.1 The PPL estimation results for frailty parameter α in the kidney infection data

$\widehat{\alpha}$	Standard SE	Corrected SE	$\widehat{\beta}$ (SE)
0.509	0.303	0.333	-1.368 (0.427)

As expected, the corrected SE is larger than the standard SE even though $\widehat{\alpha}$ is somewhat small as in $\widehat{\alpha} = 0.509$. This result indicates that the standard SE may be underestimated.

4.2. Simulation study

Simulated studies, based upon 200 replications of simulated data, are presented to evaluate the performances of the standard and corrected SEs for the PPL estimator for α. The simulation scheme is similar to that of McGilchrist (1993). That is, we generate data from the frailty model (1) assuming the exponential baseline hazard $\lambda_0(t) = 0.1$, regression parameter $\beta = 0.5$ and the variance $\alpha = 0.5, 1, 4$. Here, we set a single covariate x_{ij} to be 0 for the first $q/2$ individuals (control group), and x_{ij} to be 1 for the remaining $q/2$ individuals (treatment group). We also set the sample size $n = \sum_{i=1}^{q} n_i = 90$ or 270, which corresponds to $q = 30$ or 90 with $n_i = 3$, respectively. The corresponding censoring times C_{ij} are generated from an exponential distribution with parameter empirically determined to achieve approximately the right censoring rate, around 20%. For the 200 replications we computed the bias, standard deviation (SD), the mean of the estimated SEs for $\widehat{\alpha}$. The standard and corrected SEs are obtained from (2.9) and (3.10), respectively. For the regression parameter β the corresponding mean, SE and SD are also given.

Table 4.2 Simulations results for the PPL estimation of parameters

		$\widehat{\alpha}$				$\widehat{\beta}$		
n	α	Bias	Standard SE	Corrected SE	SD	Bias	SE	SD
90	0.5	0.019	0.271	0.351	0.354	-0.012	0.361	0.344
	1.0	0.021	0.411	0.540	0.511	-0.027	0.442	0.452
	4.0	-0.212	1.194	1.475	1.378	-0.078	0.750	0.776
270	0.5	0.013	0.149	0.202	0.198	-0.011	0.220	0.233
	1.0	-0.020	0.216	0.237	0.245	-0.017	0.263	0.252
	4.0	-0.145	0.643	0.840	0.832	-0.013	0.432	0.436

The results of PPL estimates are summarized in Table 4.2. The larger the frailty parameter α, the bigger the bias in its estimation as in simulation results by McGilchrist (1993) and Ripatti and Palmgren (2000). However, the estimates for fixed parameters (α, β) overall work well as sample size n increases. In Table 4.2, for α the SD is the estimate of the true $\{\text{var}(\widehat{\alpha})\}^{1/2}$ and the standard or corrected SE is the average of SE estimates corresponding to $\widehat{\alpha}$. The corrected SE performs well as judged by the very good agreement between corrected SE and SD, while the standard SE does not. That is, the standard SE is substantially underestimated: see also McGilchrist (1993) and Ripatti and Palmgren (2000). In addition, for β the estimated SE also works well, which confirms simulation results of Ha *et al.* (2001) and Ha and Lee (2003).

5. Remarks

We have showed that the correction for the SE of the PPL is possible by using h-likelihood. Care is necessary in implementing the SE of frailty parameter estimator when the number of nuisance parameters λ_{0k} 's in the baseline hazards increases with sample size n. Here, the consideration of the $\partial\widehat{v}/\partial\alpha$ term is important. The proposed method was developed in a simple frailty model (2.1) with one random component. Thus, extension of our method to correlated or multi-component frailty models (Ripatti and Palmgren, 2000; Yau, 2001; Ha *et al.*, 2007) would be an interesting further work.

References

Breslow, N. E. (1972). Discussion of professor Cox's paper.
Breslow, N. E. (1974). Covariance analysis of censored survival data. *Biometrics*, **30**, 89-99.

Cox, D. R. (1972). Regression models and life-tables (with discussion). *Journal of the Royal Statistical Society*, B, **34**, 187-220.

Duchateau, L. and Janssen, P. (2008). *The frailty model,* Springer-Verlag, New York.

Ha, I. D. (2006). On profile likelihood for gamma frailty models. *Journal of Korean Data & Information*, **17**, 999-1007.

Ha, I. D. (2008). On estimation of random effect in Poisson HGLMs. *Journal of Korean Data & Information*, **19**, 375-383.

Ha, I. D. and Lee, Y. (2003). Estimating frailty models via Poisson hierarchical generalized linear models. *Journal of Computational Graphical Statistics*, **12**, 663-681.

Ha, I. D. and Lee, Y. (2005). Comparison of hierarchical likelihood versus orthodox best linear unbiased predictor approaches for frailty models. *Biometrika*, **92**, 717-723.

Ha, I. D., Lee, Y. and MacKenzie, G. (2007). Model selection for multi-component frailty models. *Statistics in Medicine*, **26**, 4790-4807.

Ha, I. D., Lee, Y. and Song, J. K. (2001). Hierarchical likelihood approach for frailty models. *Biometrika*, **88**, 233-243.

Hougaard, P. (2000). *Analysis of multivariate survival Data,* Springer-Verlag, New York.

Lee, Y. and Nelder, J. A. (1996). Hierarchical generalized linear models (with discussion). *Journal of the Royal Statistical Society* B, **58**, 619-678.

Lee, Y. and Nelder, J. A. (2001). Hierarchical generalised linear models: A synthesis of generalised linear models, random-effect models and structured dispersions. *Biometrika*, **88**, 987-1006.

Lee, Y. and Nelder, J. A. and Pawitan, Y. (2006). *Generalised linear models with random effects: Unified analysis via h-likelihood,* Chapman and Hall, London.

McGilchrist, C. A. (1993). REML estimation for survival models with frailty. *Biometrics*, **49**, 221-225.

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

Nielsen, G. G., Gill, R. D., Andersen, P. K. and Sorensen, T. I. A. (1992). A counting process approach to maximum likelihood estimation in frailty models. *Scandinavian Journal of Statistics,* **19**, 25-44.

Ripatti, S. and Palmgren, J. (2000). Estimation of multivariate frailty models using penalized partial likelihood. *Biometrics*, **56**, 1016-1022.

Therneau, T. M. and Grambsch, P. M. (2000). *Modelling survial data: extending the Cox model,* Springer-Verlag, New York.

Vaida, F. and Xu, R. (2000). Proportional hazards model with random effects. *Statistics in Medicine*, **19**, 3309-3324.

Yau, K. K. W. (2001). Multilevel models for survival analysis with random effects. *Biometrics*, **57**, 96-102.