

frailtyHL: A Package for Fitting Frailty Models with H-likelihood

by Il Do Ha, Maengseok Noh and Youngjo Lee

Abstract We present the **frailtyHL** package for fitting semi-parametric frailty models using h-likelihood. This package allows lognormal or gamma frailties for random-effect distribution, and it fits shared or multilevel frailty models for correlated survival data. Functions are provided to format and summarize the **frailtyHL** results. The estimates of fixed effects and frailty parameters and their standard errors are calculated. We illustrate the use of our package with three well-known data sets and compare our results with various alternative R-procedures.

Introduction

Frailty models with a non-parametric baseline hazard have been widely adopted for the analysis of survival data (Hougaard, 2000; Duchateau and Janssen, 2008). The **frailtyHL** package (Ha et al., 2012) implements h-likelihood procedures (Ha et al., 2001; Ha and Lee, 2003, 2005) for frailty models. The package fits Cox proportional hazards models (PHMs) with random effects, or frailties. The lognormal or gamma distribution can be adopted as the frailty distribution, corresponding to the normal or log-gamma distribution for the log frailties. H-likelihood obviates the need for marginalization over the frailty distribution, thereby providing a statistically efficient procedure for various random-effect models (Lee et al., 2006).

The main function is `frailtyHL()`. For instance,

```
> frailtyHL(Surv(time, status) ~ x + (1|id),
+           RandDist = "Normal",
+           mord = 0, dord = 1,
+           Maxiter = 200, convergence = 10^-6,
+           varfixed = FALSE, varinit = 0.1)
```

fits a lognormal frailty model. The first argument is a formula object, with the response on the left of a `~` operator, and the terms for the fixed and random effects on the right. The response is a survival object as returned by the `Surv` function (Therneau, 2011). Here, `time` and `status` denote survival time and censoring indicator taking value 1 or 0 for uncensored or censored observations, respectively; `x` denotes a fixed covariate and `id` denotes the subject identifier. The expression `(1|id)` specifies a random intercept model (`(x|id)` would specify a random intercept, random slope model). The parameters `mord` and `dord` are the orders of Laplace approximations to fit the mean parameters (`mord = 0` or `1`) and the dispersion parameters (`dord = 1` or `2`), respectively. The

`Maxiter` parameter specifies the maximum number of iterations and `convergence` specifies the tolerance of the convergence criterion. If `varfixed` is specified as `TRUE` (or `FALSE`), the value of one or more of the variance terms for the frailties is fixed (or estimated) with starting value (e.g. 0.1) given by `varinit`.

Previously, frailty models have been implemented in several R functions such as the `coxph()` function in the **survival** package (Therneau, 2010) and the `coxme()` function in the **coxme** package (Therneau, 2011), based on penalized partial likelihood (PPL), the `phmm()` function in the **phmm** package (Donohue and Xu, 2012), based on a Monte Carlo EM (MCEM) method, and the `frailtyPenal()` function in the **frailtypack** package (Gonzalez et al., 2012), based on penalized marginal likelihood. The **phmm** package fits one-component frailty models, although it does allow multivariate frailties. The `coxme()` function can fit the multi-component model as shown in Example 2. Results from **frailtyHL** are compared with those from **survival**, **coxme** and **phmm**.

Recently, the h-likelihood procedures of Lee and Nelder (1996) for fitting hierarchical generalized linear models (HGLMs) have been implemented using the `hglm()` function in the **hglm** package (Alam et al., 2010), the `HGLMfit()` function in the **HGLMMM** package (Molas, 2010) and the `dhglmfit()` function in the **dhglm** package (Noh and Lee, 2011). The **frailtyHL** package is for fitting frailty models with a non-parametric baseline hazard, providing estimates of fixed effects, random effects, and variance components as well as their standard errors. In addition, it provides a statistical test for the variance components of frailties and also three AIC criteria for the model selection.

Frailty models

The **frailtyHL** package makes it possible to

1. fit models with log-normal and gamma frailty distributions and
2. estimate variance components when the frailty structure is shared or nested.

For illustration, we present two models below and show how to fit these models using `frailtyHL()` in the **Examples** section.

One-component frailty models

Suppose that data consist of right censored time-to-event observations from q subjects (or clusters), with

n_i observations each; $i = 1, \dots, q$. Let T_{ij} be the survival time for the j th observation of the i th subject; $j = 1, \dots, n_i$. Here, $n = \sum_i n_i$ is the total sample size and n_i is the cluster size. Let C_{ij} be the corresponding censoring time and let $y_{ij} = \min\{T_{ij}, C_{ij}\}$ and $\delta_{ij} = I(T_{ij} \leq C_{ij})$ be the observable random variables, where $I(\cdot)$ is the indicator function.

Given the common unobserved frailty for the i th subject u_i , the conditional hazard function of T_{ij} is of the form

$$\lambda_{ij}(t|u_i) = \lambda_0(t) \exp(x_{ij}^T \beta) u_i, \quad (1)$$

where $\lambda_0(\cdot)$ is an unspecified baseline hazard function and $\beta = (\beta_1, \dots, \beta_p)^T$ is a vector of regression parameters for the fixed covariates x_{ij} . Here, the term $x_{ij}^T \beta$ does not include an intercept term because of identifiability. Assume that the frailties u_i are independent and identically distributed (i.i.d) random variables with a frailty parameter α ; the gamma and lognormal frailty models assume gamma and lognormal distributions for u_i , respectively. That is, our package allows

- gamma frailty with $E(u_i) = 1$ and $\text{var}(u_i) = \alpha$, and
- lognormal frailty having $v_i = \log u_i \sim N(0, \alpha)$.

The model represented by (1) is known as a shared or one-component frailty model (Hougaard, 2000).

Multi-component frailty models

We can fit a multi-component model as below (Ha et al., 2007):

$$X\beta + Z^{(1)}v^{(1)} + Z^{(2)}v^{(2)} + \dots + Z^{(k)}v^{(k)}, \quad (2)$$

X is the $n \times p$ model matrix, $Z^{(r)}$ ($r = 1, 2, \dots, k$) are the $n \times q_r$ model matrices corresponding to the $q_r \times 1$ frailties $v^{(r)}$, and $v^{(r)}$ and $v^{(l)}$ are independent for $r \neq l$.

For example, the CGD data (Fleming and Harrington, 1991) have a multilevel structure in which patients nested within centers have recurrent infection times.

Later, we analyze these data using model (2) with $k = 2$. Here, the frailty structures are:

$$\begin{aligned} v^{(1)}: & \text{center frailty} \sim N(0, \alpha_1 A_1), \\ v^{(2)}: & \text{patient frailty} \sim N(0, \alpha_2 A_2), \end{aligned}$$

where $A_r = I_r$ ($r = 1, 2$) are the $q_r \times q_r$ identity matrices, and q_1 and q_2 are the number of centers and patients, respectively. Notice that the corresponding $Z^{(1)}$ to $v^{(1)}$ (or $Z^{(2)}$ to $v^{(2)}$) is, respectively, a matrix of indicator variables such that $Z_{st}^{(1)} = 1$ (or $Z_{st}^{(2)} = 1$) if observation s is a member of center (or patient) t and 0 otherwise (Therneau and Grambsch, 2000). This is called the multilevel frailty model (Yau, 2001; Ha et al., 2007).

H-likelihood theory

The h-likelihood h (Ha et al., 2001) for frailty model (1) is defined by

$$h = h(\beta, \lambda_0, \alpha) = \ell_0 + \ell_1, \quad (3)$$

where

$$\begin{aligned} \ell_0 &= \sum_{ij} \log f(y_{ij}, \delta_{ij} | u_i; \beta, \lambda_0) \\ &= \sum_{ij} \delta_{ij} \{ \log \lambda_0(y_{ij}) + \eta_{ij} \} - \sum_{ij} \Lambda_0(y_{ij}) \exp(\eta_{ij}) \end{aligned}$$

is the sum of conditional log densities for y_{ij} and δ_{ij} given u_i , and

$$\ell_1 = \sum_i \log f(v_i; \alpha)$$

is the sum of log densities for $v_i = \log u_i$ with parameter α . Here, $\eta_{ij} = x_{ij}^T \beta + v_i$ is the linear predictor for the hazards, and

$$\Lambda_0(t) = \int_0^t \lambda_0(k) dk$$

is the baseline cumulative hazard function.

The functional form of $\lambda_0(t)$ in (1) is unknown; following Breslow (1972), we consider $\Lambda_0(t)$ to be a step function with jumps at the observed event times:

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

where $y_{(k)}$ is the k th ($k = 1, \dots, l$) smallest distinct event time among the y_{ij} 's, and $\lambda_{0k} = \lambda_0(y_{(k)})$. However, the dimension of $\lambda_0 = (\lambda_{01}, \dots, \lambda_{0l})^T$ increases with the sample size n . For inference, Ha et al. (2001) proposed the use of the profile h-likelihood with λ_0 eliminated, $h^* \equiv h|_{\lambda_0 = \hat{\lambda}_0}$, given by

$$h^* = h^*(\beta, \alpha) = \ell_0^* + \ell_1, \quad (4)$$

where

$$\begin{aligned} \ell_0^* &= \sum_{ij} \log f^*(y_{ij}, \delta_{ij} | u_i; \beta) \\ &= \sum_{ij} f(y_{ij}, \delta_{ij} | u_i; \beta, \hat{\lambda}_0) \end{aligned}$$

does not depend on λ_0 , and

$$\hat{\lambda}_{0k}(\beta, v) = \frac{d_{(k)}}{\sum_{(i,j) \in R_{(k)}} \exp(\eta_{ij})},$$

are solutions of the estimating equations, $\partial h / \partial \lambda_{0k} = 0$, for $k = 1, \dots, l$. Here, $d_{(k)}$ is the number of events at $y_{(k)}$ and

$$R_{(k)} = R(y_{(k)}) = \{(i, j) : y_{ij} \geq y_{(k)}\}$$

is the risk set at $y_{(k)}$. Therneau and Grambsch (2000) and Ripatti and Palmgren (2000) proposed an h-likelihood (3), while using the partial likelihood (Breslow, 1974) for ℓ_0 . They call it the penalized partial likelihood (PPL) h_p , defined by

$$h_p(\beta, v, \alpha) = \sum_{ij} \delta_{ij} \eta_{ij} - \sum_k d_{(k)} \log \left\{ \sum_{ij \in R_{(k)}} \exp(\eta_{ij}) \right\} + \ell_1.$$

Furthermore, Ha et al. (2001) and Ha et al. (2010) have shown that h^* is proportional to the PPL h_p because

$$\begin{aligned} h^* &= \sum_k d_{(k)} \log \hat{\lambda}_{0k} + \sum_{ij} \delta_{ij} \eta_{ij} - \sum_k d_{(k)} + \ell_1 \\ &= h_p + \sum_k d_{(k)} \{ \log d_{(k)} - 1 \}, \end{aligned}$$

where $\sum_k d_{(k)} \{ \log d_{(k)} - 1 \}$ is a constant that does not depend upon unknown parameters. Notice here that PPL h_p does not depend on nuisance parameters λ_0 . Thus, Lee and Nelder's (1996; 2001) h-likelihood procedure for HGLMs can be directly used to fit frailty models based on h_p (Ha et al., 2010).

Review of estimation procedures

Lee and Nelder (1996, 2001) have proposed the use of the Laplace approximation based on the h-likelihood when the marginal likelihood, $m = \log \{ \int \exp(h) dv \}$, is hard to obtain. To reduce the bias in small cluster sizes higher-order approximations for the mean (β) and the frailty variance (α) have been developed. The lower-order approximation is computationally efficient, but could have large biases when cluster sizes are small (Ha and Lee, 2003, 2005; Ha et al., 2010).

The h-likelihood procedures use the Breslow method for handling tied event times, while the PPL procedures allow the Efron method. For estimating β , the h-likelihood methods allow the Laplace approximation $p_v(h_p)$ to a marginal partial likelihood $m_p = \log \{ \int \exp(h_p) dv \}$, but the PPL procedures do not. For estimating α , the PPL methods use adjusted profile h-likelihoods $p_v(h_p)$ and $p_{\beta,v}(h_p)$ which give maximum partial likelihood (MPL) and restricted maximum partial likelihood (REMPPL) estimators, respectively. In contrast, the h-likelihood method uses the restricted partial likelihood (based upon the first-order Laplace approximation $p_{\beta,v}(h_p)$ or the second-order Laplace approximation $s_{\beta,v}(h_p)$) for REMPL estimators. Here $p_v(h_p)$ and $p_{\beta,v}(h_p)$ are defined as follows:

$$p_v(h_p) = \left[h_p - \frac{1}{2} \log \det \{ D(h_p, v) / (2\pi) \} \right] \Big|_{v=\hat{v}},$$

where $D(h_p, v) = -\partial^2 h_p / \partial v^2$ and \hat{v} solves $\partial h_p / \partial v = 0$, which is the first-order Laplace approximation of

m_p , and

$$p_{\beta,v}(h_p) = \left[h_p - \frac{1}{2} \log \det \left\{ \frac{D(h_p, (\beta, v))}{2\pi} \right\} \right] \Big|_{\beta=\hat{\beta}, v=\hat{v}},$$

where $D(h_p, (\beta, v)) = -\partial^2 h_p / \partial (\beta, v)^2$ and $(\hat{\beta}, \hat{v})$ solves $\partial h_p / \partial (\beta, v) = 0$, which becomes the Cox and Reid (1987) adjusted profile marginal likelihood, eliminating fixed effects β by conditioning their asymptotic sufficient statistics $\hat{\beta}$, in addition to eliminating random effects v by the first-order Laplace approximation (Ha and Lee, 2003; Lee et al., 2006). The corresponding second-order approximation is

$$s_{\beta,v}(h_p) = p_{\beta,v}(h_p) - \{ F(h_p) / 24 \},$$

with

$$F(h_p) = \text{tr} \left[- \left\{ 3 \frac{\partial^4 h_p}{\partial v^4} + 5 \frac{\partial^3 h_p}{\partial v^3} D(h_p, v)^{-1} \frac{\partial^3 h_p}{\partial v^3} \right\} \times D(h_p, v)^{-2} \right] \Big|_{v=\hat{v}}.$$

To reduce the computational burden Ha et al. (2010) used $F(h)$ instead of $F(h_p)$.

Table 1: Estimation criteria for h-likelihood (HL(mord, dord)), PPL (coxph(), coxme()) and marginal likelihood (phmm()) for lognormal (LN) and gamma frailty models (FM)

Method	Criterion		Literature
	β	α	
HL(0,1)	h_p	$p_{\beta,v}(h_p)$	Ha and Lee (2003)
HL(0,2)	h_p	$s_{\beta,v}(h_p)$	Ha and Lee (2003)
HL(1,1)	$p_v(h_p)$	$p_{\beta,v}(h_p)$	Ha et al. (2012)
HL(1,2)	$p_v(h_p)$	$s_{\beta,v}(h_p)$	Ha et al. (2012)
coxph()	h_p	$p_{\beta,v}(h_p)$	Therneau (2010) for LN FM
coxph()	h_p	m	Therneau (2010) for gamma FM
coxme()	h_p	$p_v(h_p)$	Therneau (2011) for LN FM
phmm()	m	m	Donohue and Xu (2012) for LN FM

Table 1 shows historical developments of estimation criteria for frailty models. The frailtyHL() function provides estimators based on the h-likelihood. As the orders in mord and dord increase, the bias of estimator is reduced, but the calculation becomes computationally intensive due to the extra terms.

We recommend the use of HL(1,1) for the lognormal frailty and of HL(1,2) for the gamma frailty. However, HL(0,1) for the lognormal frailty and HL(0,2) for the gamma frailty often perform well if α is not large. Note that the variance matrices of $\hat{\tau} = (\hat{\beta}, \hat{v})$ and $\hat{\alpha}$ are directly obtained from the Hessian matrices, $\{-\partial^2 h_p / \partial \tau^2\}^{-1}$ and $\{-\partial^2 p_{\beta,v}(h_p) / \partial \alpha^2\}^{-1}$, respectively; the **frailtyHL** package provides the standard errors (SEs) of $\hat{\alpha}$ as well as $\hat{\beta}$. For the use of standard errors of $\hat{v} - v$, see Lee and Ha (2010), Lee et al (2011) and Ha et al. (2011).

Based on the PPL methods, the `coxph()` and `coxme()` functions, respectively, implement REMPL and MPL estimators for α in the lognormal frailty model, and the `coxph()` function the maximum likelihood (ML) estimators, maximizing the marginal likelihood m , for α in the gamma frailty model. For comparison, we present the Breslow and Efron methods for handling ties in survival times in the `coxph()` and `coxme()` functions; Therneau (2010) recommended the Efron method. For the lognormal frailty the ML estimator maximizing m is available via the `phmm()` function, but care must be taken to ensure that the MCEM algorithm has converged (Donohue and Xu, 2012). However, the ML estimator can be biased when the number of nuisance parameters increases with the sample size (Ha et al., 2010).

Furthermore, for the lognormal frailty the `coxph()` function uses the existing codes in linear mixed models so that it misses the $\partial \hat{v} / \partial \alpha$ term in solving the score equation $\partial p_{\beta,v}(h_p) / \partial \alpha = 0$; this can lead to an underestimation of the parameters, especially when the number of subjects q is large or censoring is high (Lee et al., 2006; Ha et al., 2010). To overcome this problem, in gamma frailty models Therneau and Grambsch (2000) develop the code for the ML estimator for α .

Fitting algorithm

Suppose that HL(0,1) is used. The fitting algorithm is as follows:

Step 1: Take (0,0,0.1) as initial estimates of components of (β, v, α) .

Step 2: Given $\hat{\alpha}$, new estimates $(\hat{\beta}, \hat{v})$ are obtained by solving the joint estimating equations $\partial h_p / \partial (\beta, v) = \{\partial h / \partial (\beta, v)\}|_{\lambda_0 = \hat{\lambda}_0} = 0$; then, given $(\hat{\beta}, \hat{v})$, new estimates $\hat{\alpha}$ are obtained by solving $\partial p_{\beta,v}(h_p) / \partial \alpha = 0$.

Step 3: Repeat Step 2 until the maximum absolute difference between the previous and current estimates for (β, v) and α is less than 10^{-6} .

After convergence, we compute the estimates of the standard errors of $\hat{\beta}$ and $\hat{\alpha}$.

Illustration: kidney infection data

To demonstrate differences of various estimation methods in small cluster size $n_i \equiv 2$, we use the kidney infection data (McGilchrist and Aisbett, 1991). The data consist of times until the first and second recurrences ($n_i \equiv 2$) of kidney infection in 38 ($q = 38$) patients using a portable dialysis machine. Each survival time (`time`) is the time until infection since the insertion of the catheter. The survival times for the same patient are likely to be related because of a shared frailty describing the common patient's effect. The catheter is later removed if infection occurs and can be removed for other reasons, which we regard as censoring; about 24% of the data were censored.

We fit frailty models with two covariates, the sex (1 = male; 2 = female) and age of each patient, using the functions (`frailtyHL()`, `coxph()`, `coxme()` and `phmm()`) in the four packages. The results are summarized in Table 2.

Table 2: Comparison of different estimation methods for the kidney infection data

Method	Sex $\hat{\beta}_1$ (SE)	Age $\hat{\beta}_2$ (SE)	Patient $\hat{\alpha}$ (SE)
lognormal model			
HL(0,1)	-1.380 (0.431)	0.005 (0.012)	0.535 (0.338)
HL(1,1)	-1.414 (0.432)	0.005 (0.012)	0.545 (0.340)
<code>coxph()</code> (Breslow)	-1.388 (0.441)	0.005 (0.012)	0.551 (-)
<code>coxph()</code> (Efron)	-1.411 (0.445)	0.005 (0.013)	0.569 (-)
<code>coxme()</code> (Breslow)	-1.332 (0.414)	0.005 (0.012)	0.440 (-)
<code>coxme()</code> (Efron)	-1.355 (0.417)	0.004 (0.012)	0.456 (-)
<code>phmm()</code>	-1.329 (0.452)	0.004 (0.012)	0.378 (-)
gamma model			
HL(0,2)	-1.691 (0.483)	0.007 (0.013)	0.561 (0.280)
HL(1,2)	-1.730 (0.485)	0.007 (0.013)	0.570 (0.281)
<code>coxph()</code> (Breslow)	-1.557 (0.456)	0.005 (0.012)	0.398 (-)
<code>coxph()</code> (Efron)	-1.587 (0.461)	0.005 (0.012)	0.412 (-)

In PPL procedures (`coxph()` and `coxme()`), the Breslow method provides slightly smaller estimate for α than the Efron method. In the lognormal frailty, REMPL procedures (`frailtyHL()` and `coxph()`) give larger estimates for α than ML (`phmm()`) and MPL (`coxme()`) procedures. However, both ML and MPL estimates from `phmm()` and `coxme()` are somewhat

different when cluster size is small, $n_i \equiv 2$ for all i . For the gamma frailty, `coxph()` uses the ML procedure, but it still gives smaller estimate for α than the REMPL (h-likelihood) procedures. Compared with the h-likelihood methods, PPL methods are computationally more efficient, but could have larger biases (Ha et al., 2010).

Potential future developments

The current version of the **frailtyHL** package allows multi-component (multilevel) frailties. Allowance for correlation (Ha et al., 2011) between random effects, for example correlations between random center effect and random treatment effect, is currently in progress. Other developments include dispersion frailty models based on double HGLMs (Lee and Nelder, 2006), which allow fixed and random effects in dispersion parts (i.e. variance of random effects) of the model as well as the hazard functions, and joint modelling (Ha et al, 2003) of HGLMs and frailty models.

Examples

Example 1: Lognormal and gamma frailty models on rat data

The data set presented by Mantel et al. (1977) is based on a tumorigenesis study of 50 ($q = 50$) litters of female rats. For each litter, one rat was selected to receive the drug and the other two rats were placebo-treated controls ($n_i \equiv 3$). Here each litter is treated as a cluster. The survival time (`time`) is the time to development of tumor, measured in weeks. Death before occurrence of tumor yields a right-censored observation; forty rats developed a tumor, leading to censoring of about 73%. The survival times for rats in a given litter may be correlated due to a random effect representing shared genetic or environmental effects.

We fit frailty models with one covariate, rx ($1 =$ drug; $0 =$ placebo), using `frailtyHL()`. Below, we present the code and results for the lognormal frailty model with HL(1,1). The output from the R code shows that the effect of rx is significant (t-value = 2.823 with p-value = 0.005). That is, the rx group has a significantly higher risk than in control group. Here, the variance estimate of the frailty is $\hat{\alpha} = 0.427$ (with SE = 0.423).

Note that although we report the SE of α , one should not use it for testing the absence of frailty $\alpha = 0$ (Vaida and Xu, 2000). Such a null hypothesis is on the boundary of the parameter space, so that the critical value of an asymptotic $(\chi_0^2 + \chi_1^2)/2$ distribution is 2.71 at 5% significant level (Lee et al., 2006; Ha et al., 2011). The difference in deviance (based on REMPL) $-2p_{\beta,v}(h_p)$ between the Cox model without

frailty and the lognormal frailty model is $364.15 - 362.56 = 1.59 (< 2.71)$, indicating that the frailty effect is non-significant, i.e. $\alpha = 0$. Here, the results from fitting the Cox model without frailty are available by adding the two arguments `varfixed = TRUE` and `varinit = 0` in the `frailtyHL()`: see below.

```
> library(survival)
> data(rats)
> frailtyHL(Surv(time,status) ~ rx + (1|litter),
+           data = rats,
+           varfixed = TRUE, varinit = 0)

iteration :
           4
convergence :
 4.801639e-09
[1] "converged"
[1] "Results from the Cox model"
[1] "Number of data : "
[1] 150
[1] "Number of event : "
[1] 40
[1] "Model for conditional hazard : "
Surv(time, status) ~ rx + (1 | litter)
[1] "Method : HL(0,1)"
[1] "Estimates from the mean model"
      Estimate Std. Error t-value p-value
rx  0.8982      0.3174    2.83 0.004655
[1] "Estimates from the dispersion model"
      Estimate Std. Error
litter "0"      "NULL"
      -2h0 -2*hp -2*p_b,v(hp)
[1,] 363.69 363.69  364.15
      cAIC  mAIC  rAIC
[1,] 365.69 365.69 364.15

> frailtyHL(Surv(time,status) ~ rx + (1|litter),
+           data = rats, RandDist = "Normal",
+           mord = 1, dord = 1)

iteration :
           87
convergence :
 9.97616e-07
[1] "converged"
[1] "Results from the log-normal frailty model"
[1] "Number of data : "
[1] 150
[1] "Number of event : "
[1] 40
[1] "Model for conditional hazard : "
Surv(time, status) ~ rx + (1 | litter)
[1] "Method : HL(1,1)"
[1] "Estimates from the mean model"
      Estimate Std. Error t-value p-value
rx  0.9107      0.3226    2.823 0.004754
[1] "Estimates from the dispersion model"
      Estimate Std. Error
litter 0.4272      0.4232
      -2h0 -2*hp -2*p_v(hp) -2*p_b,v(hp)
```

```
[1,] 335.97 397.36      362.14   362.56
      cAIC  mAIC  rAIC
[1,] 362.22 366.14 364.56
```

The code and results for the gamma frailty model with HL(1,2) are presented below. The output shows that these results are similar to those of the lognormal frailty, particularly for estimation of β . The deviance difference (based on REMPL) between the Cox model and gamma frailty model using the second-order approximation $-2s_{\beta,v}(h_p)$ is $364.15 - 362.12 = 2.03 (< 2.71)$, again indicating the absence of frailty effect (i.e. $\alpha = 0$) as evident in the lognormal frailty model.

```
> frailtyHL(Surv(time,status) ~ rx + (1|litter),
+           data = rats, RandDist = "Gamma",
+           mord = 1, dord = 2)

iteration :
      170
convergence :
  9.567765e-07
[1] "converged"
[1] "Results from the gamma frailty model"
[1] "Number of data : "
[1] 150
[1] "Number of event : "
[1] 40
[1] "Model for conditional hazard : "
Surv(time, status) ~ rx + (1 | litter)
[1] "Method : HL(1,2)"
[1] "Estimates from the mean model"
  Estimate Std. Error t-value p-value
rx  0.9126    0.3236    2.82 0.004806
[1] "Estimates from the dispersion model"
  Estimate Std. Error
litter  0.5757    0.5977
-2h0   -2*hp   -2*p_v(hp) -2*s_v(hp)
[1,] 331.60 413.85  365.35  361.71
-2*p_b,v(hp) -2*s_b,v(hp)
      365.77    362.12
      cAIC  mAIC  rAIC
[1,] 365.30 365.71 364.12
```

For the selection of a model among nested or non-nested models such as lognormal and gamma frailty models, we may use three Akaike information criteria (AIC) (Lee et al., 2006; Donohue et al., 2011; Ha et al., 2007) based on conditional likelihood, marginal likelihood and restricted likelihood, respectively, defined by

$$\begin{aligned} \text{cAIC} &= -2h_0 + 2df_c, \\ \text{mAIC} &= -2p_v(h_p) + 2df_m, \\ \text{rAIC} &= -2p_{\beta,v}(h_p) + 2df_r, \end{aligned}$$

where $h_0 = \ell_0^*$ in (4), and

$$df_c = \text{trace} \left\{ D^{-1}(h_p, (\beta, v)) D(h_0, (\beta, v)) \right\}$$

is an 'effective degrees of freedom adjustment' for estimating the fixed and random effects, computed using the Hessian matrices $D(h_p, (\beta, v)) = -\partial^2 h_p / \partial(\beta, v)^2$ and $D(h_0, (\beta, v)) = -\partial^2 h_0 / \partial(\beta, v)^2$. Note here that df_m is the number of fixed parameters and df_r is the number of dispersion parameters (Ha et al., 2007). For calculation of the mAIC and rAIC of gamma frailty model using HL(0,2) or HL(1,2), we use the corresponding second-order approximations, defined by $\text{mAIC} = -2s_v(h_p) + 2df_m$ and $\text{rAIC} = -2s_{\beta,v}(h_p) + 2df_r$. We select a model to minimize the AIC values among models. If the AIC difference is larger than 1 the choice can be made (Sakamoto et al., 1986). However, if the difference is less than 1 a simpler model can be selected by a parsimony principle (Donohue et al., 2011).

In the data set, in the Cox model $\text{cAIC}=365.69$, $\text{mAIC}=365.69$ and $\text{rAIC}=364.15$, and in the lognormal frailty model $\text{cAIC}=362.22$, $\text{mAIC}=366.14$ and $\text{rAIC}=364.56$, and in the gamma frailty model $\text{cAIC}=365.30$, $\text{mAIC}=365.71$ and $\text{rAIC}=364.12$. The likelihood tests based upon the REMPL showed the absence of frailty effect ($\alpha = 0$), so that mAIC and rAIC of all the three models are similar. Thus, we may choose the parsimonious Cox model. However, the cAIC selects the lognormal model, indicating that this model could give better prediction.

Example 2: Multilevel frailty models on CGD infection data

The CGD data set presented by Fleming and Harrington (1991) consists of a placebo-controlled randomized trial of gamma interferon (rIFN-g) in the treatment of chronic granulomatous disease (CGD). 128 patients (id) from 13 centers ($q_1 = 13, q_2 = 128$) were tracked for around 1 year. The number (i.e. cluster size) of patients per center ranged from 4 to 26. The survival times ($t_{\text{stop}} - t_{\text{start}}$) are the recurrent infection times of each patient from the different centers. Censoring occurred at the last observation for all patients, except one, who experienced a serious infection on the day he left the study; in the CGD study about 63% of the data were censored. The recurrent infection times for a given patient are likely to be correlated. However, each patient belongs to one of the 13 centers; hence, the correlation may also be attributed to a center effect.

Ignoring important random components may render invalid many of the traditional statistical analysis techniques. We fit a multilevel lognormal frailty model with two frailties and a single covariate, treat (rIFN-g, placebo), using `frailtyHL()`. Here, the two frailties are random center and patient effects. The code and results using HL(1,1) are provided below. The output shows that the effect of treatment is significant (t-value = -3.476 with p-value < 0.001), indicating that rIFN-g significantly reduces the rate of serious infection in CGD patients. The estimate of

variance of patient frailty ($\hat{\alpha}_2 = 1.002$) is considerably larger than variance of center frailty ($\hat{\alpha}_1 = 0.030$), indicating that the random-patient effect is more heterogeneous.

```
> library(survival)
> data(cgd)
> frailtyHL(Surv(tstop-tstart,status) ~ treat +
+           (1|center) + (1|id),
+           data = cgd,
+           RandDist = "Normal", mord = 1, dord = 1)

iteration :
           157
convergence :
           9.336249e-07
[1] "converged"
[1] "Results from the log-normal frailty model"
[1] "Number of data : "
[1] 203
[1] "Number of event : "
[1] 76
[1] "Model for conditional hazard : "
Surv(tstop-tstart,status)~treat+(1|center)+(1|id)
[1] "Method : HL(1,1)"
[1] "Estimates from the mean model"
           Estimate Std. Error t-value p-value
treatrIFN-g -1.184      0.3407  -3.476 0.0005085
[1] "Estimates from the dispersion model"
           Estimate Std. Error
center  0.02986      0.1572
id      1.00235      0.5089
      -2h0  -2*hp  -2*p_v(hp)  -2*p_b,v(hp)
[1,] 603.30 853.66  692.63    692.95
           cAIC  mAIC  rAIC
[1,] 684.92 698.63 696.95
```

For testing the absence of a random component ($\alpha_1 = 0$ or $\alpha_2 = 0$), we use the deviance based on REMPL, $-2p_{\beta,v}(h_p)$, and fit the following four models including the Cox model and three lognormal frailty models using HL(1,1) method,

M1: Cox model without frailty ($\alpha_1 = 0, \alpha_2 = 0$) has $-2p_{\beta,v}(h_p) = 707.48$,

M2: model without patient effect ($\alpha_1 > 0, \alpha_2 = 0$) has $-2p_{\beta,v}(h_p) = 703.66$,

M3: model without center effect ($\alpha_1 = 0, \alpha_2 > 0$) has $-2p_{\beta,v}(h_p) = 692.99$, and

M4: multilevel model above requiring both patient and center effects ($\alpha_1 > 0, \alpha_2 > 0$) has $-2p_{\beta,v}(h_p) = 692.95$.

The deviance difference between M3 and M4 is $692.99 - 692.95 = 0.04$, which is not significant at a 5% level ($\chi^2_{1,0.10} = 2.71$), indicating the absence of the random-center effects, i.e. $\alpha_1 = 0$. The deviance difference between M2 and M4 is $703.66 - 692.95 = 10.71$, indicating that the random-patient effects are necessary, i.e. $\alpha_2 > 0$. In addition, the deviance difference

between M1 and M3 is $707.48 - 692.99 = 14.49$, indicating that the random-patient effects are indeed necessary with or without random-center effects.

Let us choose a final model using information criteria. For M1 we have cAIC=708.68, mAIC=708.68 and rAIC=707.48; for M2 cAIC=702.96, mAIC=706.88 and rAIC=705.66; for M3 cAIC=684.84, mAIC=696.68 and rAIC=694.99; for M4 cAIC=684.92, mAIC=698.63 and rAIC=696.95. All of the three criteria choose M3 in the CGD data set.

Comparison of results with alternative procedures

Using the examples in the previous section, we compare the outcomes from `frailtyHL` and other packages. We consider the three functions (`coxph()`, `coxme()` and `phmm()`) for the lognormal frailty model and the `coxph()` function for the gamma frailty model.

Example 1: Rat data

The codes of `coxph()`, `coxme()` and `phmm()` for fitting lognormal frailty model are, respectively, as follows:

```
>coxph(Surv(time, status) ~ rx +
+       frailty(litter, dist = "gauss"),
+       method = "breslow", rats)

> coxme(Surv(time, status) ~ rx + (1|litter),
+       ties = "breslow", rats)

> phmm(Surv(time, status) ~ rx + (1|litter),
+       data = rats, Gbs = 2000, Gbsvar = 3000,
+       VARSTART = 1, NINIT = 10,
+       MAXSTEP = 200, CONVERG=90)
```

Table 3 summarizes the estimation results. Even though cluster size $n_i \equiv 3$ is not large, the results are similar. For example, MPL and ML estimates for α from `coxme()` and `phmm()` are somewhat different in Table 2 when $n_i \equiv 2$, while they become similar in Table 3.

Next, the code of `coxph()` for fitting the gamma frailty model is below:

```
> coxph(Surv(time, status) ~ rx +
+       frailty(litter, dist = "gamma"),
+       method = "breslow", rats)
```

The results of `frailtyHL()` (HL(0,2), HL(1,2)) and `coxph()` for gamma frailty are also presented in Table 3. For the estimation of β both results from `frailtyHL()` and `coxph()` are similar, but for α they are somewhat different. That is, our REMPL estimates from `frailtyHL()` ($\hat{\alpha} = 0.575$ with HL(0,2) and $\hat{\alpha} = 0.576$ with HL(1,2)) are somewhat larger than the ML estimates from `coxph()` ($\hat{\alpha} = 0.474$ with Breslow method and $\hat{\alpha} = 0.499$ with Efron method).

Table 3: Comparison of different estimation methods for the rat data

Method	Rx $\hat{\beta}$ (SE)	Litter $\hat{\alpha}$ (SE)
lognormal model		
HL(0,1)	0.906 (0.323)	0.427 (0.423)
HL(1,1)	0.911 (0.323)	0.427 (0.423)
coxph() (Breslow)	0.905 (0.322)	0.395 (-)
coxph() (Efron)	0.913 (0.323)	0.412 (-)
coxme() (Breslow)	0.905 (0.322)	0.406 (-)
coxme() (Efron)	0.913 (0.323)	0.426 (-)
phmm()	0.920 (0.326)	0.449 (-)
gamma model		
HL(0,2)	0.908 (0.324)	0.575 (0.598)
HL(1,2)	0.913 (0.324)	0.576 (0.598)
coxph() (Breslow)	0.906 (0.323)	0.474 (-)
coxph() (Efron)	0.914 (0.323)	0.499 (-)

Example 2: CGD data

The code of the `coxme()` function for fitting multi-level lognormal frailty model is as follows:

```
> coxme(Surv(tstop - tstart, status) ~
+       treat + (1|center) + (1|id),
+       ties = "breslow", cgd)
```

The results of `frailtyHL()` (HL(0,1), HL(1,1)) and `coxme()` are summarized in Table 4. The results from HL and PPL methods for frailty parameters become more similar because the cluster sizes (the number of patients from different centers) are somewhat large, ranging from 4 to 26.

Table 4: Comparison of different estimation methods for the CGD data

Method	Treat $\hat{\beta}$ (SE)	Center $\hat{\alpha}_1$ (SE)	Patient $\hat{\alpha}_2$ (SE)
lognormal model			
HL(0,1)	-1.074 (0.335)	0.026 (0.153)	0.982 (0.501)
HL(1,1)	-1.184 (0.341)	0.030 (0.157)	1.002 (0.509)
coxme() (Breslow)	-1.074 (0.333)	0.033 (-)	0.939 (-)
coxme() (Efron)	-1.074 (0.333)	0.032 (-)	0.947 (-)

Summary

The h-likelihood method offers novel possibilities to fit various models with random effects. The **frailtyHL** package for frailty models eliminates the nuisance parameters λ_0 in the h-likelihood (3) by profiling. Such models have important applications in multi-center clinical study (Vaida and Xu, 2000), meta analysis (Rondeau et al., 2008), and genetic analysis (Lee et al., 2006). Therefore, this package can be potentially adopted by statisticians in several different fields.

Acknowledgements

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (No. 2009-0088978, No. 2010-0021165 and No. 2010-0011372).

Bibliography

M. Alam, L. Ronnegard, and X. Shen. *hglm: hierarchical generalized linear models*, 2010. URL <http://CRAN.R-project.org/package=hglm>. R package version 1.1.1. [p28]

N. E. Breslow. Discussion of Professor Cox’s paper. *Journal of the Royal Statistical Society Series B*, 34: 216–217, 1972. [p29]

N. E. Breslow. Covariance analysis of censored survival data. *Biometrics*, 30:89–99, 1974. [p30]

D. R. Cox and N. Reid. Parameter orthogonality and approximate conditional inference (with discussion). *Journal of the Royal Statistical Society Series B*, 49:1–39, 1987. [p30]

L. Duchateau and P. Janssen. *The Frailty Model*. Springer-Verlag: New York, 2008. [p28]

M. Donohue, R. Overholser, R. Xu, and F. Vaida. Conditional Akaike information under generalized linear and proportional hazards mixed models. *Biometrika*, 98:685–700, 2011. [p33]

M. Donohue and R. Xu. *phmm: proportional hazards mixed-effects model*, 2012. URL <http://CRAN.R-project.org/package=phmm>. R package version 0.7-4. [p28, 30, 31]

T. R. Fleming and D. P. Harrington. *Counting Processes and Survival Analysis*. Wiley: New York, 1991. [p29, 33]

J. R. Gonzalez, V. Rondeau Y. Mazroui, A. Mauguen, and A. Diakit . *frailtypack: frailty models using a semi-parametrical penalized likelihood estimation or a*

- parametrical estimation*, 2012. URL <http://CRAN.R-project.org/package=frailtypack>. R package version 2.2-23. [p28]
- I. D. Ha and Y. Lee. Estimating frailty models via Poisson hierarchical generalized linear models. *Journal of Computational and Graphical Statistics*, 12: 663–681, 2003. [p28, 30]
- I. D. Ha and Y. Lee. Comparison of hierarchical likelihood versus orthodox best linear unbiased predictor approaches for frailty models. *Biometrika*, 92: 717–723, 2005. [p28, 30]
- I. D. Ha, Y. Lee and G. MacKenzie. Model selection for multi-component frailty models. *Statistics in Medicine*, 26:4790–4807, 2007. [p29, 33]
- I. D. Ha, Y. Lee and J. K. Song. Hierarchical likelihood approach for frailty models. *Biometrika*, 88: 233–243, 2001. [p28, 29, 30]
- I. D. Ha, M. Noh and Y. Lee. Bias reduction of likelihood estimators in semi-parametric frailty models. *Scandinavian Journal of Statistics*, 37:307–320, 2010. [p30, 31, 32]
- I. D. Ha, M. Noh and Y. Lee. *frailtyHL: frailty models using h-likelihood*, 2012. URL <http://CRAN.R-project.org/package=frailtyHL>. R package version 1.1. [p28, 30]
- I. D. Ha, T. Park and Y. Lee. Joint modelling of repeated measures and survival time data. *Biometrical Journal*, 45:647–658, 2003. [p32]
- I. D. Ha, R. Sylvester, C. Legrand and G. MacKenzie. Frailty modelling for survival data from multi-centre clinical trials. *Statistics in Medicine*, 30:2144–2159, 2011. [p31, 32]
- P. Hougaard. *Analysis of Multivariate Survival Data*. Springer-Verlag: New York, 2000. [p28, 29]
- Y. Lee and I. D. Ha. Orthodox BLUP versus h-likelihood methods for inferences about random effects in Tweedie mixed models. *Statistics and Computing*, 20:295–303, 2010. [p31]
- Y. Lee, M. Jang and W. Lee. Prediction interval for disease mapping using hierarchical likelihood. *Computational Statistics*, 26:159–179, 2011. [p31]
- Y. Lee and J. A. Nelder. Hierarchical generalized linear models (with discussion). *Journal of the Royal Statistical Society Series B*, 58:619–678, 1996. [p28, 30]
- Y. Lee and J. A. Nelder. Hierarchical generalised linear models: a synthesis of generalised linear models, random-effect models and structured dispersions. *Biometrika*, 88:987–1006, 2001. [p30]
- Y. Lee and J. A. Nelder. Double hierarchical generalised linear models (with discussion). *Applied Statistics*, 55:139–185, 2006. [p32]
- Y. Lee, J. A. Nelder and Y. Pawitan. *Generalised Linear Models with Random Effects: Unified Analysis via H-likelihood*. Chapman & Hall: London, 2006. [p28, 30, 31, 32, 33, 35]
- N. Mantel, N. R. Bohidar and J. L. Ciminera. Mantel-Haenszel analyses of litter-matched time-to-response data, with modifications for recovery of interlitter information. *Cancer Research*, 37:3863–3868, 1977. [p32]
- C. A. McGilchrist and C. W. Aisbett. Regression with frailty in survival analysis. *Biometrics*, 47:461–466, 1991. [p31]
- M. Molas. *HGLMMM: hierarchical generalized linear models*, 2010. URL <http://CRAN.R-project.org/package=HGLMMM>. R package version 0.1.1. [p28]
- M. Noh and Y. Lee. *dhglm: double hierarchical generalized linear models*, 2011. URL <http://CRAN.R-project.org/package=dhglm>. R package version 1.0. [p28]
- S. Ripatti and J. Palmgren. Estimation of multivariate frailty models using penalized partial likelihood. *Biometrics*, 56:1016–1022, 2000. [p30]
- V. Rondeau, S. Michiels, B. Liqueur and J. P. Pignon. Investigating trial and treatment heterogeneity in an individual patient data meta-analysis of survival data by means of the penalized maximum likelihood approach. *Statistics in Medicine*, 27:1894–910, 2008. [p35]
- Y. Sakamoto, M. Ishiguro and G. Kitagawa. *Akaike Information Criterion Statistics*. KTK Scientific Publisher: Tokyo, 1986. [p33]
- T. M. Therneau. *survival: survival analysis, including penalised likelihood*, 2010. URL <http://CRAN.R-project.org/package=survival>. R package version 2.36-2. [p28, 30, 31]
- T. M. Therneau. *coxme: Mixed Effects Cox Models*, 2011. URL <http://CRAN.R-project.org/package=coxme>. R package version 2.2-1. [p28, 30]
- T. M. Therneau and P. M. Grambsch. *Modelling Survival Data: Extending the Cox Model*. Springer-Verlag: New York, 2000. [p29, 30, 31]
- F. Vaida and R. Xu. Proportional hazards model with random effects. *Statistics in Medicine*, 19:3309–3324, 2000. [p32, 35]
- K. K. W. Yau. Multilevel models for survival analysis with random effects. *Biometrics*, 57:96–102, 2001. [p29]

Il Do Ha
Department of Asset Management
Daegu Haany University, Korea
idha@dhu.ac.kr

Maengseok Noh
Department of Statistics

Pukyong National University, Korea
msnoh@pknu.ac.kr

Youngjo Lee
Department of Statistics
Seoul National University, Korea
youngjo@snu.ac.kr