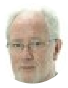
Likelihood ratio and score tests for a shared frailty model: a non-standard problem



Paul Janssen
Hasselt University

T E C H N I C A L

R E P O R T

0309

LIKELIHOOD RATIO AND SCORE TESTS FOR A SHARED FRAILTY MODEL: A NON-STANDARD PROBLEM.

R. NGUTI, G. CLAESKENS and P. JANSSEN

I A P S T A T I S T I C S
N E T W O R K

INTERUNIVERSITY ATTRACTION POLE

http://www.stat.ucl.ac.be/IAP

Likelihood ratio and score tests for a shared frailty model: a non-standard problem

Rosemary Nguti[1,3], Gerda Claeskens[2] and Paul Janssen[3]

[1] *University of Nairobi and International Livestock Research Institute*

[2] *Texas A&M University*

[3] *Limburgs Universitair Centrum*

Abstract

Tests for the presence of heterogeneity in frailty models use an alternative hypothesis in which the heterogeneity parameter is subject to an inequality constraint. As a result the classical likelihood ratio asymptotic χ^2-distribution theory is no longer valid. Based on simulations and inspired by results from mixed models theory it has been conjectured that the likelihood ratio test and the score test for testing heterogeneity has an asymptotic distribution that is a 50:50 mixture of a χ^2_0 and a χ^2_1 distribution. We proof this conjecture for the bivariate shared gamma frailty model with Weibull baseline hazard. We consider the likelihood ratio and the score test. Our theorems provide a contribution to the theme of "statistical inference under inequality constraints" (Sen and Silvapulle, 2002) for bivariate survival data; and in that sense our results contribute to a "better understanding of the asymptotic theory of frailty models" (Bjarnason and Hougaard, 2000).

Key Words: Inference under inequality constraints, Frailty models, Likelihood ratio test, Mixture of χ^2-distributions, Score test, Survival data.

Running headline: Tests for a shared frailty model.

1 Introduction

The proportional hazards model with random effects, an extension of the Cox model, has become very popular for analysing multivariate survival data. The model is also called the mixed proportional hazards model or the frailty model. A rather general description of frailty models has been given by Vaida and Xu (2000). Frailty models for survival data are the counterpart for mixed effects models for normal data; the main idea is to have a model that can handle survival times that are dependent within clusters. There are indeed many situations where the survival data are dependent: in a multi-centre trial survival times within a centre can be dependent (heterogeneity between centres; Duchateau et al. (2002)), survival times of lambs from the same sire can be dependent (heterogeneity between sires; Nguti et al. (2003)), recurrent event times for a patient can be dependent (heterogeneity between patients; Wassell et al. (1999)). See Section 2 for a further discussion on how dependence within a cluster is generated by heterogeneity in the conditional model.

In spite of the fact that mixed proportional hazards are very useful to describe and to model multivariate survival data, the inferential properties are not yet well examined. The main reason is that, due to the complexity of the modelling, it is very hard to derive statistical properties (e.g. asymptotic properties) for frailty models in general, see e.g. Murphy (1995) and Parner (1998). The complexity lies in the fact that the likelihood expression needed for the inference is implicit and difficult, so that in many situations numerical algorithms are needed to obtain estimates and standard errors.

One of the important methodological questions is to provide information on the asymptotic distributional behaviour of the likelihood ratio test for heterogeneity. To test for heterogeneity (to test within cluster correlation) we consider the following hypotheses testing problem. Assume that the random effect, present in the mixed proportional hazards model, has variance θ. The relevant hypotheses testing problem is:

$$H_0 : \theta = 0 \text{ versus } H_a : \theta > 0. \tag{1}$$

From the theory of mixed effects models we know that the asymptotic distribution theory for the likelihood ratio statistic for such hypotheses testing problems does not follow the classical chi-square limit theory. The reason is that, under the null hypothesis, the parameter of interest is at the boundary of the parameter space (in the alternative hypothesis the heterogeneity parameter is subject to an inequality constraint). A consequence of this is that the classical conditions needed for the likelihood ratio theory are not satisfied. We therefore need to develop “likelihood ratio theory under non-standard conditions”. This phenomenon has been recognized in the literature on frailty models. Vaida and Xu (2000, p. 3322) write that "for the likelihood ratio test a correction for the null distribution, which is no longer a chi-square distribution, is needed as discussed in similar set-ups by Stram and Lee (1994) and Self and Liang (1987) in the context of mixed effects models". Duchateau et al. (2002) simulate the limit distribution of the likelihood ratio test and conjecture that the simulated distribution is a 50:50 mixture of a χ^2_0 and a χ^2_1 distribution.

In this manuscript we give an explicit proof for this conjecture for the shared gamma frailty with a Weibull baseline hazard (which includes the exponential baseline hazard as a special case). We further assume that the survival data are complete (censoring makes the formulas much more complicated) and that there are no covariates. A final simplification is that we assume that each cluster contains two observations. Bjarnason and Hougaard (2000) use this model to study the Fisher information. The idea behind the simplification is to fully understand a statistical property for a simple, but relevant, model. In our case we want to derive the asymptotic null distribution for the likelihood ratio test and the score test for heterogeneity. The formal model and the main results are given in Section 2. The proofs are given in Section 3, key references are Vu and Zhou (1997) and Silvapulle and Silvapulle (1995). A discussion on possible extensions is given in Section 4.

2 The model and the main results

We observe a set of n independent and identically distributed random vectors $T_i = (T_{i1}, T_{i2})$, $i = 1, \ldots, n$. Each vector is considered as a cluster of size two. We assume that, conditional on the frailty variables Z_i, the lifetimes T_{i1} and T_{i2} are independent with (for $Z_i = z$) a Weibull$(z\lambda, \gamma)$ distribution, i.e., the conditional hazard is

$$h(t \mid z) = z\lambda\gamma t^{\gamma-1}$$

with $\lambda > 0$ and $\gamma > 0$ and where Z_i has the gamma density

$$f_{Z_i}(z) = z^{\frac{1}{\theta}-1} \exp\left(-\frac{z}{\theta}\right) / \{\Gamma\left(1/\theta\right) \theta^{\frac{1}{\theta}}\}.$$

The key idea is that within cluster dependence is caused by the frailty variables $Z_1, \ldots, Z_n$ representing unobserved common risk factors. The frailty variables are assumed to be independent. Also note that $\mathrm{Var}(Z_i) = \theta$. Given $Z_i = z$, the conditional survival function of (T_{i1}, T_{i2}) is

$$S(t_1, t_2 \mid z) \;\; = \;\; P(T_{i1} > t_1, T_{i2} > t_2 \mid Z = z)$$

$$= \exp\{-z\lambda(t_1^\gamma + t_2^\gamma)\}.$$

The (unconditional) survival function is

$$\begin{aligned} S(t_1, t_2) &= E\left[\exp\{-Z\lambda(t_1^\gamma + t_2^\gamma)\}\right] \\ &= \{1 + \theta\lambda(t_1^\gamma + t_2^\gamma)\}^{-\frac{1}{\theta}}. \end{aligned}$$

The corresponding joint density is

$$f(t_1, t_2) = \frac{(1+\theta)\lambda^2\gamma^2 t_1^{\gamma-1} t_2^{\gamma-1}}{\{1 + \theta\lambda(t_1^\gamma + t_2^\gamma)\}^{\frac{1}{\theta}+2}}.$$

For $\theta > 0$ (heterogeneity between clusters) the components of the vector (T_{i1}, T_{i2}) are correlated (within cluster correlation). To quantify the within cluster dependence we can use Kendall's coefficient of concordance which, in terms of the joint density and survival function, is given by

$$4\int_0^\infty \int_0^\infty f(t_1, t_2) S(t_1, t_2) dt_1 dt_2 - 1,$$

see Hougaard (2000, p. 132) and Bjarnason and Hougaard (2000). For our model Kendall's coefficient of concordance is $\theta/(2+\theta)$ which is zero for $\theta = 0$ (homogeneity between clusters). Moreover we easily obtain that

$$\lim_{\substack{\theta \to 0 \\ >}} f(t_1, t_2) = (\lambda\gamma t_1^{\gamma-1} e^{-\lambda t_1^\gamma})(\lambda\gamma t_2^{\gamma-1} e^{-\lambda t_2^\gamma}),$$

i.e., T_{i1} and T_{i2} are independent Weibull distributed random variables.

The likelihood for the data is given by

$$\prod_{i=1}^n \frac{(1+\theta)\lambda^2\gamma^2 T_{i1}^{\gamma-1} T_{i2}^{\gamma-1}}{\{1 + \theta\lambda(T_{i1}^\gamma + T_{i2}^\gamma)\}^{\frac{1}{\theta}+2}}$$

with corresponding loglikelihood

$$\begin{aligned} L_n &= \sum_{i=1}^n \{2\ln\lambda + 2\ln\gamma + \ln(1+\theta) + (\gamma - 1)(\ln T_{i1} + \ln T_{i2}) \\ &\qquad - \left(\frac{1}{\theta} + 2\right) \ln(1 + \theta\lambda(T_{i1}^\gamma + T_{i2}^\gamma))\}. \end{aligned}$$

To test the within cluster dependence we consider the testing problem $H_0 : \theta = 0$ against $H_a : \theta > 0$. For the further discussion it is convenient to work with the following transformed

Weibull parameters: $\eta = -\ln\lambda$ and $\alpha = -\ln\gamma$; or $\lambda = \exp(-\eta)$ and $\gamma = \exp(-\alpha)$. Further we use τ as shorthand notation for the set of model parameters (θ, η, α) and $\nu = (\eta, \alpha)$ for the set of nuisance parameters. In terms of τ the parameter space is $\Theta = [0, \infty) \times \mathbb{R} \times \mathbb{R}$ and the testing problem can be written as

$$H_0 : \tau \in \Theta_0 = \{0\} \times \mathbb{R} \times \mathbb{R} \text{ against } H_a : \tau \in \Theta_1 = (0, \infty) \times \mathbb{R} \times \mathbb{R}$$

with as corresponding likelihood ratio statistic

$$\mathcal{L}_n = 2\left\{\sup_{\tau\in\Theta} L_n(\tau) - \sup_{\tau\in\Theta_0} L_n(\tau)\right\}$$

where we use the notation $L_n \equiv L_n(\tau)$.
Under the null hypothesis the parameter vector of interest is at the boundary of the parameter space. Therefore the standard asymptotic distribution theory for likelihood ratio tests does not work. Instead we have the following theorem.

Theorem 1
The likelihood ratio statistic $\mathcal{L}_n$ for testing the heterogeneity hypothesis (1.1) has an asymptotic null distribution which is an equal mixture of a point mass at zero and a chi-square distribution with one degree of freedom, abbreviated as $\frac{1}{2}\chi^2_0 + \frac{1}{2}\chi^2_1$.

Remark. This result is important since it has an immediate impact on how to determine (asymptotic) critical values and P-values for likelihood ratio test for heterogeneity. Erroneously relying on classical χ^2- distribution theory will lead to P-values that are too big (or critical values that are too large), which means that not using the appropriate statistical inference leads to a conservative strategy in rejecting the null hypothesis of independence.

Under standard conditions, that is, when parameters constrained under the null hypothesis belong to the interior of the parameter space, it is well known that likelihood ratio, Wald and score statistics have asymptotically the same distribution under the null hypothesis. Under inequality constraints in the alternative hypothesis, a score statistic is no longer uniquely defined, see Silvapulle and Silvapulle (1995). Robertson, Wright and Dykstra (1988, pp. 320–321) propose a Wald and score statistic the latter of which has the disadvantage of requiring estimation of model parameters both under the null and alternative hypothesis. Silvapulle and Silvapulle (1995) propose a different score-type statistic which only requires estimation under the null hypothesis. Under mild regularity conditions, they obtain that under the null hypothesis asymptotically the score statistic follows the same mixture distribution as the likelihood ratio statistic.

Before we further explain the basic idea in Silvapulle and Silvapulla (1995) we introduce the score vector for the shared frailty model in Section 2. The explicit expressions for the components of the score vector $S_n(\tau) = (S_{n,\theta}(\tau), S_{n,\nu}(\tau))$ with $S_{n,\theta}(\tau) = \partial L_n(\tau)/\partial\theta$ and $S_{n,\nu}(\tau) = (\partial L_n(\tau)/\partial\eta, \partial L_n(\tau)/\partial\alpha)^T$ are given in Section 5.

Via a Taylor series expansion Silvapulle and Silvapulle (1995) rewrite the likelihood ratio statistic as the difference of the minimum of two quadratic forms, of which the minimisation of the first one under the null hypothesis can be performed exactly. We state the resulting score statistic in the following theorem. Let $\hat{\nu}$ be the maximum likelihood estimator of the nuisance parameters under the null hypothesis and let $S_{n,\theta}(0,\hat{\nu})$ denote the score vector evaluated at $(0,\hat{\nu})$.

Theorem 2

(i) For a shared gamma frailty model with exponential baseline hazard a score statistic for testing the heterogeneity hypothesis (1.1) is given by

$$\mathcal{S}_n = \frac{1}{3n^2}\{S_{n,\theta}(0,\hat{\eta})\}^2 - 3n \inf_{b\geq 0}\left\{\left(\frac{1}{3n^{3/2}}S_{n,\theta}(0,\hat{\eta}) - b\right)^2\right\}.$$

(ii) For a Weibull distribution as the baseline hazard function a score statistic for testing the heterogeneity hypothesis (1.1) is given by

$$\mathcal{S}_n = \frac{1}{3n^2}\frac{\pi^2}{\pi^2-4}\{S_{n,\theta}(0,\hat{\nu})\}^2 - 3n\left(1-\frac{4}{\pi^2}\right)\inf_{b\geq 0}\left\{\left(\frac{1}{3n^{3/2}}\frac{\pi^2}{\pi^2-4}S_{n,\theta}(0,\hat{\nu}) - b\right)^2\right\}.$$

For both models the corresponding score statistic has, under the null hypothesis, asymptotic distribution $\frac{1}{2}\chi^2_0 + \frac{1}{2}\chi^2_1$.

A proof is given in the next section. Note that in the exponential baseline hazard model when $S_{n,\theta}(0,\hat{\eta}) \geq 0$, the expression of the score statistic simplifies to

$$\mathcal{S}_n = \frac{1}{3n^2}\{S_{n,\theta}(0,\hat{\eta})\}^2.$$

A similar simplification holds for the Weibull baseline hazard model.

Wald-type test statistics for testing hypothesis (1.1) may be employed as well. Robertson, Wright and Dykstra (1988) construct a Wald statistic for the situation where the alternative hypothesis is described by inequalities. Their test statistic requires estimation of model parameters under both the null and alternative hypothesis. Sen and Silvapulle (2002) state a Wald test statistic as a difference of the minimum of two quadratic forms which has under the null hypothesis the same asymptotic distribution as the score and likelihood ratio statistic. For more details, see the recent review paper by Sen and Silvapulle (2002).

3 Proofs

Vu and Zhou (1997) give a set of conditions under which a general result holds on the asymptotic behaviour of likelihood ratio tests where, under the null hypothesis, the true values are allowed to be on the boundary of the parameter space. For the model specified in Section 2, we will show that their set of conditions is satisfied.

First define $\boldsymbol{D}_n(\nu) = E[S_n^T(0,\nu)S_n(0,\nu)]$ and $\boldsymbol{G}_n(\nu) = E[\boldsymbol{F}_n(0,\nu)]$ with $\boldsymbol{F}_n(\tau)$ the matrix of the negative of the second derivatives of $L_n(\tau)$. As derived in Section 5 we have for the shared frailty model with Weibull baseline hazard that

$$\boldsymbol{G}_n(\nu) = n\boldsymbol{G}(\eta) = n \begin{pmatrix} 5 & 2 & 2(2-\gamma_e+\eta) \\ 2 & 2 & 2(1-\gamma_e+\eta) \\ 2(2-\gamma_e+\eta) & 2(1-\gamma_e+\eta) & \pi^2/3 + 2(1-\gamma_e+\eta)^2 \end{pmatrix} \tag{2}$$

with γ_e the Euler constant. From Property 2 in Section 5 we know that $\boldsymbol{D}_n(\nu) = \boldsymbol{G}_n(\nu)$, a property that we expect since our likelihood is based on a sample of independent and identically distributed vectors.

3.1 Proof of Theorem 1.

The main issue is to check the conditions (A1)-(A3) and (B1)-(B5) of Vu and Zhou (1997) needed for the validity of their Theorem 2.2.

(A1) The likelihood function, score vector and components of the matrix of second derivatives of the log likelihood (see Section 5 for explicit expressions) are continuous and finite on a neighborhood of the true parameter value $(0, \nu_0)$. For the score component and second derivatives of the log likelihood with respect to θ the boundedness can be shown by an expansion of the logarithmic function in the second term of equation (12) in Section 5. For the other derivatives the result is straightforward to obtain.

(A2) Let $\beta_{\min}(\boldsymbol{A})$ and $\beta_{\max}(\boldsymbol{A})$ denote the smallest and the largest eigenvalue of a symmetric positive definite matrix $\boldsymbol{A}$. From Property 1 in Section 5 we have that

$$\liminf_{n\to\infty} \frac{\beta_{\min}(\boldsymbol{G}_n(\nu))}{\beta_{\max}(\boldsymbol{G}_n(\nu))} = \frac{\beta_{\min}(\boldsymbol{G}(\eta))}{\beta_{\max}(\boldsymbol{G}(\eta))} > 0.$$

It therefore suffices to show the Chernoff regularity, which is satisfied since the parameter space $\Theta_1 = (0,\infty) \times \mathbb{R} \times \mathbb{R}$ is convex (Geyer, 1994).

(A3) The approximating cones for Θ_0 and Θ_1 are $C_{\Theta_0} \equiv \Theta_0$ and $C_{\Theta_1} \equiv \Theta$. The transformed cones, used to obtain the asymptotic distribution of the likelihood ratio test, are for $j = 0, 1$

$$\widetilde{C}_{n,\Theta_j} = \left\{ (\widetilde{\theta}, \widetilde{\eta}, \widetilde{\alpha}) = \boldsymbol{G}_n^{T/2}(\nu) \begin{pmatrix} \theta \\ \eta \\ \alpha \end{pmatrix} \text{ with } \begin{pmatrix} \theta \\ \eta \\ \alpha \end{pmatrix} \in C_{\Theta_j} \right\}$$

with $\boldsymbol{G}_n^{1/2}(\nu)$ and $\boldsymbol{G}_n^{T/2}(\nu)$ the left and the corresponding right Cholesky square root of $\boldsymbol{G}_n(\nu)$. A direct calculation shows that $\boldsymbol{G}_n^{T/2}(\nu) = n^{1/2}\boldsymbol{G}^{T/2}(\eta)$ with

$$\boldsymbol{G}^{T/2}(\eta) = \begin{bmatrix} \sqrt{5} & 2/\sqrt{5} & 2(2-\gamma_e+\eta)/\sqrt{5} \\ 0 & \sqrt{6/5} & \sqrt{2}(1-3\gamma_e+3\eta)/\sqrt{15} \\ 0 & 0 & \sqrt{(\pi^2-4)/3} \end{bmatrix}$$

We therefore have

$$\begin{aligned} \widetilde{C}_{n,\Theta_0} &= \left\{ (\widetilde{\theta}, \widetilde{\eta}, \widetilde{\alpha}) : \widetilde{\theta} - \sqrt{\frac{2}{3}}\widetilde{\eta} - \frac{2\sqrt{5}}{\sqrt{3(\pi^2-4)}}\widetilde{\alpha} = 0 \right\} \equiv \widetilde{C}_{\Theta_0} \\ \widetilde{C}_{n,\Theta_1} &= \left\{ (\widetilde{\theta}, \widetilde{\eta}, \widetilde{\alpha}) : \widetilde{\theta} - \sqrt{\frac{2}{3}}\widetilde{\eta} - \frac{2\sqrt{5}}{\sqrt{3(\pi^2-4)}}\widetilde{\alpha} \geq 0 \right\} \equiv \widetilde{C}_{\Theta_1}. \end{aligned}$$

Since $\widetilde{C}_{n,\Theta_j} \equiv \widetilde{C}_{\Theta_j}$, condition (A3) holds.

(B1) Based on the expressions in Section 5 we can easily show that $E[S_n(0,\nu)] = 0$ and we know that $\boldsymbol{D}_n(\nu) = \boldsymbol{G}_n(\nu)$ are finite matrices.

(B2) $\beta_{\min}(\boldsymbol{G}_n(\nu)) = n\beta_{\min}(\boldsymbol{G}(\eta)) \to \infty$, $n \to \infty$. This follows from Property 1 in Section 5.

(B3) Let $\|\boldsymbol{W}\|_1$ denote the sum of the absolute values of the elements of a matrix $\boldsymbol{W}$.

For $(0,\nu_0)$, the true parameter value, define

$$N_n(A)=\left\{\tau=(\theta,\eta,\alpha):(\theta,\eta-\eta_0,\alpha-\alpha_0)\boldsymbol{G}_n(\nu_0)\begin{pmatrix}\theta\\ \eta-\eta_0\\ \alpha-\alpha_0\end{pmatrix}\leq A^2,\tau\in\Theta\right\}.$$

To prove (B3) we need to show that

$$\sup_{\tau\in N_n(A)}\|\boldsymbol{G}_n^{-1/2}(\nu_0)\boldsymbol{F}_n(\tau)\boldsymbol{G}_n^{-T/2}(\nu_0)-\boldsymbol{I}_3\|_1=o_P(1). \tag{3}$$

with P a shorthand notation for P_{τ_0} ($\tau_0=(0,\nu_0)\in\Theta_0$, the true value of the parameter under the null hypothesis). Note that (recall equation (2))

$$\begin{aligned}&\boldsymbol{G}_n^{-1/2}(\nu_0)\boldsymbol{F}_n(\tau)\boldsymbol{G}_n^{-T/2}(\nu_0)-\boldsymbol{I}_3\\ &\quad=\boldsymbol{G}^{-1/2}(\eta_0)\left(\frac{\boldsymbol{F}_n(0,\nu_0)-\boldsymbol{G}_n(\nu_0)}{n}\right)\boldsymbol{G}^{-T/2}(\eta_0)\\ &\quad\quad+\boldsymbol{G}^{-1/2}(\eta_0)\left(\frac{\boldsymbol{F}_n(\tau)-\boldsymbol{F}_n(0,\nu_0)}{n}\right)\boldsymbol{G}^{-T/2}(\eta_0).\end{aligned} \tag{4}$$

Note that, for matrices $\boldsymbol{W}_1$ and $\boldsymbol{W}_2$, $\|\boldsymbol{W}_1\boldsymbol{W}_2\|_1\leq\|\boldsymbol{W}_1\|_1\|\boldsymbol{W}_2\|_1$. Since $\|\boldsymbol{G}^{-1/2}(\eta_0)\|_1=\|\boldsymbol{G}^{-T/2}(\eta_0)\|_1\leq C(\eta_0)$, with $0<C(\eta_0)<\infty$, (3) follows by showing that

$$\left\|\frac{\boldsymbol{F}_n(0,\nu_0)}{n}-\boldsymbol{G}(\eta_0)\right\|_1=o_P(1) \tag{5}$$

and

$$\sup_{\tau\in N_n(A)}\left\|\frac{\boldsymbol{F}_n(\tau)-\boldsymbol{F}_n(0,\nu_0)}{n}\right\|_1=o_P(1). \tag{6}$$

To establish the validity of (5) we need the entries of $\boldsymbol{F}_n(0,\nu_0)$ given in Section 5. For each entry we apply the law of large numbers to obtain

$$\left|\left(\frac{\boldsymbol{F}_n(0,\nu_0)}{n}\right)_{[i,j]}-(\boldsymbol{G}(\eta_0))_{[i,j]}\right|=o_P(1).$$

Hence (5) is valid.

To establish (6) we need the entries of $\boldsymbol{F}_n(\tau)$, the negative of the second derivatives of L_n, which are also given in Section 5. For each entry we need to show that

$$\sup_{\tau\in N_n(A)}\left|\left(\frac{\boldsymbol{F}_n(\tau)-\boldsymbol{F}_n(0,\nu_0)}{n}\right)_{[i,j]}\right|=o_P(1). \tag{7}$$

We show how to prove (7) for the $[2,2]$-entry.

$$\left(\frac{\boldsymbol{F}_n(\tau)-\boldsymbol{F}_n(0,\nu_0)}{n}\right)_{[2,2]} = \frac{1}{n}\sum_{i=1}^{n} H_{22}(T_i,\tau)$$

where, with $U_i = T_{i1}^{e^{-\alpha}} + T_{i2}^{e^{-\alpha}}$ (as defined in Section 5),

$$H_{22}(T_i,\tau) = -\theta e^{-2\eta}(1+2\theta)\frac{U_i^2}{(1+\theta e^{-\eta}U_i)^2} + e^{-\eta}(1+2\theta)\frac{U_i}{1+\theta e^{-\eta}U_i} - e^{-\eta_0}U_i.$$

Note that $H_{22}(T_i,\tau_0)\equiv 0$. There exists a fixed positive integer n_0 such that for all $n\geq n_0$

$$\sup_{\tau\in N_n(A)} e^{-\alpha} \leq K \equiv 2(e^{-\alpha_0}+1)$$

and, for some constant $D>0$,

$$|H_{22}(T_i,\tau)| < D(T_{i1}^{2K}+T_{i2}^{2K}).$$

With $\mu(\tau) = E_{\tau_0}H(T_i,\tau)$ we have by the dominated convergence theorem that $\lim_{\tau\to\tau_0}\mu(\tau) = \mu(\tau_0)\equiv 0$. Now the proof of (6) follows since

$$\begin{aligned}
&\sup_{\tau\in N_n(A)}\left|\frac{1}{n}\sum_{i=1}^{n} H_{22}(T_i,\tau)\right| \\
&\quad\leq \sup_{\tau\in N_n(A)}\left|\frac{1}{n}\sum_{i=1}^{n} H_{22}(T_i,\tau)-\mu(\tau)\right| + \sup_{\tau\in N_n(A)}|\mu(\tau)| = o_P(1). \qquad (8)
\end{aligned}$$

An application of Theorem 16(a) in Ferguson (1996), p. 108 implies indeed that the first term in the right-hand side of (8) is $o_P(1)$ (uniform law of large numbers); elementary analysis implies that the second term in the right-hand side of (8) is $o(1)$.
Similar proofs hold for all the other entries of $(\boldsymbol{F}_n(\tau)-\boldsymbol{F}_n(0,\nu_0))/n$.

(B4) The matrix $\boldsymbol{V}=\boldsymbol{I}_3$ and (B4) holds since $\boldsymbol{G}_n(\nu)=\boldsymbol{D}_n(\nu)$.

(B5) Since (T_{i1},T_{i2}), $i=1,\ldots,n$, are independent and identically distributed vectors (B5) follows from classical multivariate central limit theory.

Since the Vu and Zhou (1997) conditions (A1)-(A3) and (B1)-(B5) are valid, an application of their Theorem 2.2 gives that the asymptotic null distribution of $\mathcal{L}_n$, the likelihood ratio statistic, is the same as the distribution of

$$\inf_{\widetilde{\tau}\in\widetilde{C}_{\Theta_0}} |N-\widetilde{\tau}|^2 - \inf_{\widetilde{\tau}\in\widetilde{C}_{\Theta_1}} |N-\widetilde{\tau}|^2 \qquad (9)$$

where $\widetilde{\tau} = (\widetilde{\theta}, \widetilde{\eta}, \widetilde{\alpha})$ and $N = (N_1, N_2, N_3)$ is multivariate normal with mean vector zero and covariance matrix $\boldsymbol{I}_3$.

From the definitions $\widetilde{C}_{\Theta_0}$ and $\widetilde{C}_{\Theta_1}$ we have

$$\inf_{\widetilde{\tau} \in \widetilde{C}_{\Theta_0}} |N - \widetilde{\tau}|^2 = \left(\frac{N_1 + aN_2 + bN_3}{\sqrt{1 + a^2 + b^2}} \right)^2 \tag{10}$$

with $a = -\sqrt{\frac{2}{3}}$ and $b = -\frac{2\sqrt{5}}{\sqrt{3(\pi^2-4)}}$ (see the proof of condition (A3)), i.e., the random variable in (10) has a χ_1^2 distribution. We further have

$$\inf_{\widetilde{\tau} \in \widetilde{C}_{\Theta_1}} |N - \widetilde{\tau}|^2 = \begin{cases} 0 & N \in \widetilde{C}_{\Theta_1} \\ \left(\frac{N_1 + aN_2 + bN_2}{\sqrt{1 + a^2 + b^2}} \right)^2 & N \notin \widetilde{C}_{\Theta_1} \end{cases} \tag{11}$$

Moreover we have $P(N \in \widetilde{C}_{\Theta_1}) = 0.5$. This, together with (9) - (11) implies that the asymptotic distribution of the likelihood ratio test is $0.5\chi_0^2 + 0.5\chi_1^2$. ■

3.2 Proof of Theorem 2.

We first state the general form of the score statistic to test the heterogeneity hypothesis. Partition the Fisher information matrix $\boldsymbol{G}_n(\nu)$ such that the upper left block corresponds to the parameter θ constrained to zero under the null hypothesis and the lower right block is defined by the nuisance parameters ν. Specifically,

$$\boldsymbol{G}_n(\nu) = \begin{pmatrix} G_{n,00}(\nu) & G_{n,01}(\nu) \\ G_{n,01}^T(\nu) & G_{n,11}(\nu) \end{pmatrix}.$$

Further, define $G_n^{00}(\nu) = (\boldsymbol{G}_n^{-1}(\nu))_{00} = (G_{n,00}(\nu) - G_{n,01}(\nu) G_{n,11}^{-1}(\nu) G_{n,01}^T(\nu))^{-1}$, let $\hat{\nu}$ be the maximum likelihood estimator of the nuisance parameters under the null hypothesis and let $S_{n,\theta}(0, \hat{\nu})$ denote the score vector evaluated at $(0, \widehat{\nu})$.

Under the assumption of the existence of a matrix $\boldsymbol{H}$ such that for any $a > 0$

$$\sup_{\|h\| \le a} \left[n^{-1/2} \{ S_n(\tau + n^{-1/2} h) - S_n(\tau) \} + \boldsymbol{H}(\tau) h \right] = o_p(1),$$

Silvapulle and Silvapulle (1995) define the following score statistic for testing the heterogeneity hypothesis (1.1):

$$\begin{aligned} \mathcal{S}_n &= n^{-1} S_{n,\theta}^T(0, \hat{\nu}) G_n^{00}(0, \hat{\nu}) S_{n,\theta}(0, \hat{\nu}) \\ &\quad - \inf_{b \ge 0} \left\{ \left(n^{-1/2} G_n^{00}(0, \hat{\nu}) S_{n,\theta}(0, \hat{\nu}) - b \right)^T \{ G_n^{00}(0, \hat{\nu}) \}^{-1} \left(n^{-1/2} G_n^{00}(0, \hat{\nu}) S_{n,\theta}(0, \hat{\nu}) - b \right) \right\}. \end{aligned}$$

It is also shown that this statistic has asymptotically the same distribution as the likelihood ratio statistic. We take $\boldsymbol{G}_n(\tau)$ as the matrix $\boldsymbol{H}$ from the theorem.

For the case of a shared gamma frailty model with an exponential baseline hazard there is only the nuisance parameter η (or $\lambda = \exp(-\eta)$). For this special case, with Fisher information matrix

$$\boldsymbol{G}_n = n \begin{pmatrix} 5 & 2 \\ 2 & 2 \end{pmatrix},$$

we have that $G_n^{00} = (3n)^{-1}$, not dependent on any nuisance parameters, and hence we obtain the following score statistic

$$\mathcal{S}_n = \frac{1}{3n^2}\{S_{n,\theta}(0,\hat{\eta})\}^2 - 3n \inf_{b \geq 0} \left\{ \left(\frac{1}{3n^{3/2}} S_{n,\theta}(0,\hat{\eta}) - b \right)^2 \right\}$$

For the Weibull baseline hazard the nuisance parameter is $\nu = (\eta, \alpha)$ and the Fisher information matrix is $\boldsymbol{G}_n(\nu)$ as given in (2), from which it is deduced that $G_n^{00} = \pi^2/(3n(\pi^2 - 4))$. Hence the resulting score statistic is obtained as given in Theorem 2. ∎

4 Discussion

For the shared gamma frailty model with Weibull baseline hazard, as defined in Section 2, we give explicit results on the asymptotic distributional behaviour of the likelihood ratio test and the score test for heterogeneity. These results are an important first step towards a better understanding of the asymptotic properties on test statistics for testing heterogeneity in more general frailty models. A further challenge is indeed to extend these results to more complex models and data settings. More complex families of models should include frailty models with covariates present and frailty models with unspecified baseline hazard. Regarding the type of data there is the need to obtain explicit results for censored data. The likelihood expressions in the papers by Murphy (1995) and Murphy and van der Vaart (1997, 2000) will provide useful information to tackle these problems.

A second interesting theme for further work is to study the distributional behaviour of the likelihood ratio and the score test for heterogeneity under local alternatives converging to the null hypothesis at the rate $n^{-1/2}$. As in two-sided testing problems, it is expected that the test statistics will have the same power characteristics under these local circumstances.

A third relevant issue for further study is to provide information on good finite sample approximations of the mixing properties, i.e., can we improve the asymptotic 50:50 mixture of the χ_0^2 and the χ_1^2 by finding mixing proportions that depend on the information of the sample size? In a setting of regression spline mixed models, Claeskens (2002) calculates finite sample approximations to the mixing probabilities. In the frailty models currently under consideration the situation is more complex by the presence of nuisance parameters under the null hypothesis. Bootstrapping the distribution of the test statistic can provide another alternative to the asymptotic distribution.

5 Technical results

Define the following random variables: $U_i = T_{i1}^{e^{-\alpha}} + T_{i2}^{e^{-\alpha}}$, $V_i = \ln(T_{i1}) + \ln(T_{i2})$ and $W_i = T_{i1}^{e^{-\alpha}} \ln(T_{i1}) + T_{i2}^{e^{-\alpha}} \ln(T_{i2})$.

Calculation of the three-dimensional score vector $S_n(\theta, \eta, \alpha)$:

$$\begin{aligned}
\frac{\partial}{\partial\theta} L_n(\tau) &= \frac{n}{(1+\theta)} + \frac{1}{\theta^2}\sum_{i=1}^{n} \ln(1+\theta e^{-\eta}U_i) - (\frac{1}{\theta}+2)e^{-\eta}\sum_{i=1}^{n}\frac{U_i}{1+\theta e^{-\eta}U_i} \qquad (12)\\
\frac{\partial}{\partial\eta} L_n(\tau) &= -2n + (1+2\theta)e^{-\eta}\sum_{i=1}^{n}\frac{U_i}{(1+\theta e^{-\eta}U_i)}\\
\frac{\partial}{\partial\alpha} L_n(\tau) &= -2n - e^{-\alpha}\sum_{i=1}^{n} V_i + (1+2\theta)e^{-(\eta+\alpha)}\sum_{i=1}^{n}\frac{W_i}{1+\theta e^{-\eta}U_i}.
\end{aligned}$$

Under the null hypothesis, using an expansion of the logarithm in the second term of (12), it follows that:

$$\begin{aligned}
\frac{\partial}{\partial\theta} L_n(0,\nu) &= n - 2e^{-\eta}\sum_{i=1}^{n} U_i + \frac{1}{2}e^{-2\eta}\sum_{i=1}^{n} U_i^2\\
\frac{\partial}{\partial\eta} L_n(0,\nu) &= -2n + e^{-\eta}\sum_{i=1}^{n} U_i\\
\frac{\partial}{\partial\alpha} L_n(0,\nu) &= -2n - e^{-\alpha}\sum_{i=1}^{n} V_i + e^{-(\eta+\alpha)}\sum_{i=1}^{n} W_i.
\end{aligned}$$

The components needed in the calculation of the Fisher information matrix $\boldsymbol{F}_n(\tau)$ are :

$$[1,1]: \quad \frac{\partial^2}{\partial\theta^2} L_n(\tau) = -\frac{n}{(1+\theta)^2} - \frac{2}{\theta^3}\sum_{i=1}^n \ln(1+\theta e^{-\eta}U_i) + \frac{2e^{-\eta}}{\theta^2}\sum_{i=1}^n \frac{U_i}{1+\theta e^{-\eta}U_i} + (\frac{1}{\theta}+2)e^{-2\eta}\sum_{i=1}^n \frac{U_i^2}{(1+\theta e^{-\eta}U_i)^2}$$

$$[2,2]: \quad \frac{\partial^2}{\partial\eta^2} L_n(\tau) = \theta e^{-2\eta}(1+2\theta)\sum_{i=1}^n \frac{U_i^2}{(1+\theta e^{-\eta}U_i)^2} - e^{-\eta}(1+2\theta)\sum_{i=1}^n \frac{U_i}{1+\theta e^{-\eta}U_i}$$

$$[3,3]: \quad \frac{\partial^2}{\partial\alpha^2} L_n(\tau) = e^{-\alpha}\sum_{i=1}^n V_i - (1+2\theta)e^{-(\eta+\alpha)}\sum_{i=1}^n \frac{W_i}{1+\theta e^{-\eta}U_i} - (1+2\theta)e^{-(\eta+2\alpha)}\sum_{i=1}^n \frac{T_{i1}^{e^{-\alpha}}(\ln(T_{i1}))^2 + T_{i2}^{e^{-\alpha}}(\ln(T_{i1}))^2}{1+\theta e^{-\eta}U_i} + \theta(1+2\theta)e^{-2(\eta+\alpha)}\sum_{i=1}^n \frac{W_i^2}{(1+\theta e^{-\eta}U_i)^2}$$

$$[1,2]: \quad \frac{\partial^2}{\partial\theta\partial\eta} L_n(\tau) = -e^{-2\eta}(1+2\theta)\sum_{i=1}^n \frac{U_i^2}{(1+\theta e^{-\eta}U_i)^2} + 2e^{-\eta}\sum_{i=1}^n \frac{U_i}{1+\theta e^{-\eta}U_i}$$

$$[1,3]: \quad \frac{\partial^2}{\partial\theta\partial\alpha} L_n(\tau) = 2e^{-(\eta+\alpha)}\sum_{i=1}^n \frac{W_i}{1+\theta e^{-\eta}U_i} - (1+2\theta)e^{-(2\eta+\alpha)}\sum_{i=1}^n \frac{U_i W_i}{(1+\theta e^{-\eta}U_i)^2}$$

$$[2,3]: \quad \frac{\partial^2}{\partial\eta\partial\alpha} L_n(\tau) = -(1+2\theta)e^{-(\eta+\alpha)}\sum_{i=1}^n \frac{W_i}{(1+\theta e^{-\eta}U_i)^2}.$$

Under the null hypothesis, and using an expansion of the logarithm in the second term of the second derivative with respect to θ, it follows that:

$$[1,1]: \quad \frac{\partial^2}{\partial\theta^2}L_n(0,\nu) = -n + 2e^{-2\eta}\sum_{i=1}^{n}U_i^2 - \frac{2}{3}e^{-3\eta}\sum_{i=1}^{n}U_i^3$$

$$[2,2]: \quad \frac{\partial^2}{\partial\eta^2}L_n(0,\nu) = -e^{-\eta}\sum_{i=1}^{n}U_i$$

$$\begin{aligned}[3,3]: \quad \frac{\partial^2}{\partial\alpha^2}L_n(0,\nu) &= e^{-\alpha}\sum_{i=1}^{n}V_i - e^{-(\eta+\alpha)}\sum_{i=1}^{n}W_i \\ &-e^{-(\eta+2\alpha)}\sum_{i=1}^{n}\{T_{i1}^{e^{-\alpha}}(\ln(T_{i1}))^2 + T_{i2}^{e^{-\alpha}}(\ln(T_{i1}))^2\}\end{aligned}$$

$$[1,2]: \quad \frac{\partial^2}{\partial\theta\partial\eta}L_0(0,\nu) = -e^{-2\eta}\sum_{i=1}^{n}U_i^2 + 2e^{-\eta}\sum_{i=1}^{n}U_i$$

$$[1,3]: \quad \frac{\partial^2}{\partial\theta\partial\alpha}L_0(0,\nu) = 2e^{-(\eta+\alpha)}\sum_{i=1}^{n}W_i - e^{-(2\eta+\alpha)}\sum_{i=1}^{n}U_iW_i$$

$$[2,3]: \quad \frac{\partial^2}{\partial\eta\partial\alpha}L_0(0,\nu) = -e^{-(\eta+\alpha)}\sum_{i=1}^{n}W_i.$$

If we let ψ be the digamma function and define the function $h(\theta) = (\eta - \psi(\frac{1}{\theta}) - \ln(\theta))$ and $\zeta(2,q) = \int_0^\infty \frac{te^{-qt}}{1-e^{-t}}dt$, it then follows that the expected values are

$$\begin{aligned}
E[-\frac{\partial^2}{\partial\theta^2}L_n(\tau)] &= \frac{(5+9\theta+6\theta^2)}{(1+\theta)^2(1+2\theta)(1+3\theta)} \\
E[-\frac{\partial^2}{\partial\eta^2}L_n(\tau)] &= \frac{2}{(1+3\theta)} \\
E[-\frac{\partial^2}{\partial\alpha^2}L_n(\tau)] &= 2n + 2n\zeta(2,2) + \frac{2n}{(1+3\theta)}[(\psi(2)+h(\theta))^2 + \zeta(2,\frac{1}{\theta}) \\
&\quad -2\theta[\psi(3)^2 - \psi(2)^2 + 2(\psi(3)-\psi(2))h(\theta) + \zeta(2,3)]] \\
E[-\frac{\partial^2}{\partial\theta\partial\eta}L_n(\tau)] &= \frac{2}{(1+3\theta)(1+2\theta)} \\
E[-\frac{\partial^2}{\partial\theta\partial\alpha}L_n(\tau)] &= \frac{4n\psi(3)}{(1+3\theta)} + \frac{2nh(\theta)}{(1+2\theta)(1+3\theta)} - \frac{n(2+8\theta)\psi(2)}{(1+2\theta)(1+3\theta)} \\
E[-\frac{\partial^2}{\partial\eta\partial\alpha}L_n(\tau)] &= \frac{2n}{(1+3\theta)}[\psi(2) + \eta - \psi(\frac{1}{\theta}+1) - \ln(\theta)]
\end{aligned}$$

To obtain the matrix $G_n(\nu)$ we need these expected values under H_0. This yields

$$\boldsymbol{G}_n(\nu) = n\boldsymbol{G}(\eta) = n \begin{pmatrix} 5 & 2 & 4\psi(3)+2\eta-2\psi(2) \\ 2 & 2 & 2(\psi(2)+\eta) \\ 4\psi(3)+2\eta-2\psi(2) & 2(\psi(2)+\eta) & 2(1+\zeta(2,2)+(\psi(2)+\eta)^2) \end{pmatrix}.$$

Note that $\det\{\boldsymbol{G}(\eta)\} = 2\eta^2 - 8$. Since the submatrices (5) and $\begin{pmatrix} 5 & 2 \\ 2 & 2 \end{pmatrix}$ have positive determinants it follows that $\boldsymbol{G}(\eta)$ is positive definite (see e.g. Martin (1991), p. 242). Moreover since $\boldsymbol{G}(\eta)$ is symmetric its eigenvalues are real . For β an eigenvalue of $\boldsymbol{G}(\eta)$ and x_β the corresponding eigenvector we have

$$x_\beta^T G(\eta) x_\beta = \beta x_\beta^T x_\beta > 0$$

which implies that the eigenvalues of $\boldsymbol{G}(\eta)$ are strictly positive.

Property 1. *The symmetric matrix $\boldsymbol{G}(\eta)$ is positive definite and therefore has for every fixed value of η, three positive eigenvalues.*

Now

$$\begin{aligned}
E[\frac{\partial}{\partial\theta}L_n(0,\nu)\frac{\partial}{\partial\theta}L_n(0,\nu)] &= 5n \\
E[\frac{\partial}{\partial\eta}L_n(0,\nu)\frac{\partial}{\partial\eta}L_n(\tau)] &= 2n \\
E[\frac{\partial}{\partial\alpha}L_n(0,\nu)\frac{\partial}{\partial\alpha}L_n(0,\nu)] &= n(4(\psi(3)+\eta)^2 - 6(\psi(2)+\eta)^2 \\
&\quad +4(\psi(2)+\eta)(\psi(1)+\eta) + 2\zeta(2,1) + 4\zeta(2,3) - 4\zeta(2,2)) \\
E[\frac{\partial}{\partial\theta}L_n(0,\nu)\frac{\partial}{\partial\eta}\ L_n(0,\nu)] &= 2n \\
E[\frac{\partial}{\partial\theta}L_n(0,\nu)\frac{\partial}{\partial\alpha}\ L_n(0,\nu)] &= n(6\psi(4) - 6\psi(3) + 2\psi(2) + 2\eta) \\
E[\frac{\partial}{\partial\eta}L_n(0,\nu)\frac{\partial}{\partial\alpha}\ L_n(0,\nu)] &= n(4\psi(3) - 4\psi(2) + 2\psi(1) + 2\eta)
\end{aligned}$$

Property 2. $\boldsymbol{G}_n(\nu) = \boldsymbol{D}_n(\nu) = n\boldsymbol{G}(\eta)$.

Proof. Using the recursive property of the digamma function, $\psi(\nu+1) = \psi(\nu) + \frac{1}{\nu}$, we have $\psi(3) = \psi(2) + \frac{1}{2}$ and $\psi(1) = \psi(2) - 1$. From this we obtain that

$$4(\psi(3)+\eta)^2 - 6(\psi(2)+\eta)^2 + 4(\psi(2)+\eta)(\psi(1)+\eta) = 2(\psi(2)+\eta)^2 + 1 \tag{13}$$

A direct calculation also shows that

$$2\zeta(2,1) + 4\zeta(2,3) - 4\zeta(2,2) \;=\; 2\zeta(2,2) + 1 \tag{14}$$

From (13) and (14) we obtain $(\boldsymbol{G}_n(\eta))_{33} = (\boldsymbol{D}_n(\eta))_{33}$ where $\boldsymbol{D}_n(\eta) = \boldsymbol{D}_n(\eta, \alpha)$. Equality of the other entries can be showed in a similar (more easy) way. Direct calculation of the matrix elements yields the simplified expression for $\boldsymbol{G}_n$ in (2).

Acknowledgements

Rosemary Nguti acknowledges partial support from the Belgian Technological Cooperation (BTC). Paul Janssen acknowledges partial support by the Ministry of the Flemish Community (Project BIL00/28, International Scientific and Technological Cooperation) and by the Interuniversity Attraction Pole Programme P5/24 - Belgian State - Federal Office for Scientific, Technical and Cultural Affairs.

References

Artin, M. (1991). *Algebra.* Prentice Hall, New Jersey.

Bjarnason, H. and Hougaard, P. (2000). Fisher information for two gamma frailty bivariate Weibull models. *Lifetime Data Anal.* **6**, 59–71.

Claeskens, G. (2002). Restricted likelihood ratio lack of fit tests using mixed spline models. Submitted.

Duchateau, L., Janssen, P., Lindsey, P., Legrand, C., Nguti, R. and Sylvester, R. (2002). The shared frailty model and the power for heterogeneity tests in multicenter trials. *Comput. Statist. Data Anal.* **40**, 603–620.

Ferguson, T. (1996). *A course in large sample theory.* Chapman and Hall, London.

Geyer, C.J. (1994). On the asymptotics of constrained M-estimation. *Ann. Statist.* **22**, 1993–2010.

Hougaard, P. (2000). *Analysis of multivariate survival data.* Springer, New York.

Nguti, R., Janssen, P., Rowlands, J., Audho, J., Thorpe, W. and Baker, L. (2003). Mortality of red maasai, dorper and crossbreed lambs in the sub-humid tropics: a survival analysis approach. *British J. Animal Science* (to appear).

Murphy, S.A. (1995). Asymptotic theory for the frailty model. *Ann. Statist.* **23**, 182–198.

Murphy, S.A. and van der Vaart, A.W. (1997). Semiparametric likelihood ratio inference. *Ann. Statist.* **25**, 1471–1509.

Murphy, S.A. and van der Vaart, A.W. (2000). On profile likelihood. *J. Amer. Statist. Assoc.* **95**, 449–485.

Parner, E. (1998). Asymptotic theory for the correlated gamma-frailty model. *Ann. Statist.* **26**, 183-214.

Robertson, T., Wright, F.T. and Dykstra, R.L. (1988). *Order-restricted statistical inference.* Wiley, New York.

Self, S.G. and Liang, K.-Y. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under non-standard conditions. *J. Amer. Statist. Assoc.* **82**, 605–610.

Sen, P.K. and Silvapulle, M.J. (2002). An appraisal of some aspects of statistical inference under inequality constraints. *J. Statist. Planning Inference* **107**, 3–43.

Silvapulle, M.J. and Silvapulle, P. (1995). A score test against one-sided alternatives. *J. Amer. Statist. Assoc.* **90**, 342–349.

Stram, D.O. and Lee, J.W. (1994). Variance components testing in the longitudinal mixed effects models. *Biometrics* **50**, 1171–1177.

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

Vu, H.T.V. and Zhou, S. (1997). Generalization of likelihood ratio tests under nonstandard conditions. *Ann. Statist.* **25**, 897–916.

Wassell, J.T., Wojciechowski, W.C. and Landen, D.D. (1999). Recurrent injury event-time analysis. *Statist. in Medicine* **18**, 3355–3363.

Address of corresponding author:

P. Janssen, Center for Statistics, Limburgs Universitair Centrum, Universitaire Campus, B-3590 Diepenbeek, Belgium. paul.janssen@luc.ac.be