BIOMETRICS 49, 221–225
March 1993

SHORTER COMMUNICATIONS

EDITOR:
BYRON J. T. MORGAN

REML Estimation for Survival Models with Frailty

C. A. McGilchrist*

Fred Hutchinson Cancer Research Center, 1124 Columbia Street MP702, Seattle, Washington 98107, U.S.A.

SUMMARY

A method of estimation for generalised mixed models is applied to the estimation of regression parameters in proportional hazards models for failure times when there are repeated observations of failure on each subject. The subject effect is incorporated into the model as a random frailty term. Best linear unbiased predictors are used as an initial step in the computation of maximum likelihood and residual maximum likelihood estimates.

1. Introduction

Multivariate failure times consisting of repeated observations of failure on the same subject are correlated over time by the presence of a common subject or frailty effect taken to be randomly distributed over subjects. Proportional hazards models for fixed-effect regression models were popularised by Cox (1972) in his partial likelihood estimation method, in which the unknown baseline hazard function cancelled out and was thus eliminated from the estimation process for the regression parameters. When random effects (frailties) are included, the failure time distribution obtained by integration over the distribution of frailties loses this simple cancellation property. Some models for such failure time distributions have been developed by Hougaard (1986a, 1986b).

Methods that stratify on the number of previous failures were introduced by Prentice, Williams, and Peterson (1981) and this approach has been extended using a semiparametric method of combination of results by Wei, Lin, and Weissfeld (1989).

The current approach is an extension of the best linear unbiased predictor (BLUP) method of Henderson (1975) as described in McGilchrist and Aisbett (1991). The method is now extended to using the BLUP approach as an initial step in the computation of both maximum likelihood (ML) and residual maximum likelihood (REML) estimates. The basic method is described for generalised mixed models in work as yet unpublished by McGilchrist, and developed in detail here for multivariate survival data.

The problem that motivates this work is the time to infection for a kidney patient using portable dialysis. At entry to the study, a catheter is inserted and remains in place until infection occurs at the point of insertion. The catheter is then removed and is reinserted only after the infection is cleared up, usually some weeks later. Time is recorded from the point of reinsertion until the next infection occurs. During the study several infections may be observed on each patient. The aim is to relate the incidence rate of infection to risk variables (age, sex, type of kidney disorder) and to explore the extent of variability among patients as modelled by a patient effect additional to that included in the recorded risk variables. It is assumed that the recovery interval is sufficiently long to make negligible any carry-over effects from one failure recurrence interval to the next. The only correlation between recurrence intervals for failures is due to a common patient effect, which is assumed not to vary with time. The BLUP estimation for this problem is described in McGilchrist and Aisbett (1991).

---

* *Permanent address*: Department of Statistics, University of New South Wales, P.O. Box 1, Kensington NSW 2033, Australia.

Key words: Frailty; Maximum likelihood; Multivariate survival times; Residual maximum likelihood.

The current work also includes a small simulation comparing BLUP, ML, and REML estimation, allowing the variance of frailty effects to increase and varying the number of regression variables.

2. Model and Estimation

A total of M subjects are followed over time and T_{ij} is the jth time to failure or censoring for subject i ($j = 1, 2, \ldots, n_i$). Associated risk variables for the jth failure/censoring time interval of subject i are collected into a vector $\mathbf{x}_{ij}$. The proportional hazards model is

$$h(t; i, j) = \lambda(t)g(\eta_{ij}), \quad \eta_{ij} = \mathbf{x}_{ij}'\boldsymbol{\beta} + U_i,$$

where $g(\cdot) = \exp(\cdot)$ for the Cox hazard function and U_i is the random frailty effect for subject i. The random frailty terms are taken to be independent $N(0, \theta)$ and are constant over time.

The estimation procedure, adapted from McGilchrist (unpublished manuscript), is a generalisation of Schall (1991) and follows from the work of Thompson (1980) and Fellner (1986, 1987). A parallel development of generalised mixed models is given by Breslow and Clayton (Technical Report 106, Biostatistics, University of Washington, 1991). It consists of first finding the best linear unbiased predictors (BLUP) of the fixed and random components and then using these results to find both maximum likelihood (ML) and residual maximum likelihood (REML) estimators.

The BLUP estimation of the hazard model consists of choosing estimates to maximise the sum of two components:

$$l_1 = \text{Partial log-likelihood of failure times taking } U \text{ fixed},$$

$$l_2 = -\left(\frac{1}{2}\right)\left[M \ln 2\pi\theta + \sum_{i=1}^{M} \frac{U_i^2}{\theta}\right].$$

To write down the partial log-likelihood, arrange the failure/censoring times in increasing order; the nth such time is denoted by T_i ($i = 1, 2, \ldots, N$) and a corresponding variable D_i takes value 1 if a failure occurs at T_i and 0 if censoring occurs there. The value of η for the subject failing or being censored at T_i is denoted by η_i. Using this notation, the partial log-likelihood for the Cox proportional hazards model, conditional on U fixed, is

$$l_1 = \sum_{i=1}^{N} D_i\left[\eta_i - \ln \sum_{j=1}^{N} \exp(\eta_j)\right].$$

Letting

$$w_k = \exp \eta_k, \quad a_k = D_k \Big/ \sum_{j=k}^{N} w_j, \quad b_k = \sum_{j=1}^{k} a_j, \quad \mathbf{d}' = [D_1, D_2, \ldots, D_N],$$

$$\mathbf{W} = \operatorname{diag}(w_1, w_2, \ldots, w_N), \quad \mathbf{A} = \operatorname{diag}(a_1, a_2, \ldots, a_N),$$

$$\mathbf{M} = \text{Lower-triangular matrix with ones on/below the principal diagonal},$$

$$\mathbf{B} = \operatorname{diag}(b_1, b_2, \ldots, b_N) = \operatorname{diag}(\mathbf{MA1}), \quad \mathbf{1} = \text{Vector of ones},$$

$$\boldsymbol{\eta}' = [\eta_1, \eta_2, \ldots, \eta_N],$$

then

$$\frac{dl_1}{d\boldsymbol{\eta}} = \mathbf{d} - \mathbf{WMA1},$$

$$-\frac{d^2 l_1}{d\boldsymbol{\eta}\, d\boldsymbol{\eta}'} = \mathbf{WB} - \mathbf{WMA^2M'W}.$$

Note that $\boldsymbol{\eta}$ has components that are a reordering of the η_{ij} given in the original model. For the reordered model let

$$\boldsymbol{\eta} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Zu}, \quad \mathbf{u}' = [U_1, U_2, \ldots, U_M].$$

The appropriate ordering of risk variable values makes up the $\mathbf{X}$ matrix and a similar ordering of subject effects gives the matrix $\mathbf{Z}$.

From a set of initial values $\boldsymbol{\beta}_0$, $\mathbf{u}_0$ of $\boldsymbol{\beta}$, $\mathbf{u}$, the Newton–Raphson iterative procedure for maximising $l_1 + l_2$ to obtain BLUP estimators $\tilde{\boldsymbol{\beta}}$, $\tilde{\mathbf{u}}$ is

$$\begin{bmatrix} \tilde{\beta} \\ \tilde{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\beta}_0 \\ \mathbf{u}_0 \end{bmatrix} - \mathbf{V}^{-1}\begin{bmatrix} \mathbf{0} \\ \theta^{-1}\mathbf{u}_0 \end{bmatrix} + \mathbf{V}^{-1}[\mathbf{X}\ \mathbf{Z}]\frac{dl_1}{d\boldsymbol{\eta}},$$

where

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{X}' \\ \mathbf{Z}' \end{bmatrix} \left(\frac{-d^2 l_1}{d\boldsymbol{\eta}\, d\boldsymbol{\eta}'} \right) [\mathbf{X}\ \mathbf{Z}] + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \theta^{-1}\mathbf{I} \end{bmatrix},$$

$$\mathbf{V}^{-1} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}.$$

The variance matrix for the estimate of $\boldsymbol{\beta}$ is given by $\mathbf{A}_{11}$.

The maximum likelihood and REML estimators of $\boldsymbol{\beta}$ are the same as the BLUP estimator for any given estimate of θ. However, the estimators for θ are different in the three methods, hence giving rise to different numerical values for the $\boldsymbol{\beta}$ estimators. The maximum likelihood estimator of θ is

$$\hat{\theta}_{\text{ML}} = \tilde{\mathbf{u}}'\tilde{\mathbf{u}}/(M - r^*), \quad r^* = \hat{\theta}_{\text{ML}}^{-1} \operatorname{tr} \mathbf{V}_{22}^{-1}.$$

The asymptotic variance of $\hat{\theta}_{\text{ML}}$ is $2\theta^2[M - 2r^* + \theta^{-2} \operatorname{tr}(\mathbf{V}_{22}^{-2})]^{-1}$.

The REML estimator of θ has a similar expression to the above:

$$\hat{\theta}_{\text{REML}} = \tilde{\mathbf{u}}'\tilde{\mathbf{u}}/(M - r), \quad r = \hat{\theta}_{\text{REML}}^{-1} \operatorname{tr}(\mathbf{A}_{22}).$$

The asymptotic variance of $\hat{\theta}_{\text{REML}}$ is $2\theta^2[M - 2r + \theta^{-2}\operatorname{tr}(\mathbf{A}_{22}^2)]^{-1}$.

3. Application and Simulation

The method is now applied to the kidney data used in McGilchrist and Aisbett (1991), where a full description is given. The technique described here for BLUP estimates is equivalent to that given in the previous paper. Both ML and REML estimates of regression coefficients $\boldsymbol{\beta}$ are given below with values that are not greatly different from the BLUP estimates given previously. In the table GN, AN, and PKD refer to three different classifications of kidney disease: GN = glomerulo nephritis, AN = acute nephritis, PKD = polycystic kidney disease.

Estimates of regression coefficients (S.E. *in parentheses*)

Variable	Age	Sex	GN	AN	PKD
ML estimates	.0037	−1.6051	.1317	.3573	−1.2946
	(.0126)	(.4066)	(.4610)	(.4583)	(.7244)
REML estimates	.0046	−1.7399	.1860	.3918	−1.1428
	(.0152)	(.4715)	(.5516)	(.5533)	(.8289)
	$\hat{\theta}_{\text{ML}} = .1793$ (.1204),	$\hat{\theta}_{\text{REML}} = .5460$ (.3099)			

As in the previous analysis, the only significant regression variable is sex, with females having a lower infection rate than males.

It is useful to compare BLUP, ML, and REML estimates when the failure rate distribution is known. To this end a hazard rate model was chosen as

$$h(t; i, j) = \lambda(t) \exp(\mathbf{x}_{ij}'\boldsymbol{\beta} + U_i),$$

with $\lambda(t) = .1$, components of $\mathbf{x}_{ij}$ selected randomly as 0 or 1, and U_i as independent $N(0, \theta)$. For each of $M = 30$ subjects, three failure times are generated and there is no censoring. Table 1 gives the results for $\theta = 1$ or 4, $\mathbf{x}_{ij}$ a one-component or a four-component vector. It is expected that the increase in the number of risk variables from 1 to 4 should indicate the type of differences that are likely to occur between ML and REML estimators. For most variance component problems, the bias of REML estimators of the variance component parameters (in this case, θ) does not change appreciably as more regression variables are fitted, whereas the bias often increases for ML estimators. The value of $\theta = 1$ gives a not uncommon spread of frailty effects, whereas $\theta = 4$ gives a range of U from, say, -4 to $+4$. The corresponding variation in relative hazard rate is $e^8 \approx 3{,}000$, which would be a fairly extreme variation of hazard rates due to a subject effect only.

The results of the simulation show that the BLUP estimates of the regression coefficient β tend to be shrunk slightly towards the origin, but ML and REML estimates are never appreciably biased. BLUP estimates of θ are severely negatively biased and ML estimates are slightly negatively biased, with the bias increasing to significance when the number of regression variables fitted increases from 1 to 4. REML estimates of θ tend to be positively biased but only slightly.

Table 1
Estimated biases and standard errors for 100 *simulations of BLUP, ML, and REML estimation*

Parameter	True value	Average bias (S.E. in parentheses)			Average of S.E.[a] of estimates			S.E. of estimates[b] over simulations		
		BLUP	ML	REML	BLUP	ML	REML	BLUP	ML	REML
Simulation 1: $\theta = 1$, x one component										
β	.5	−.109	.064	.071	.230	.426	.439	.330	.460	.465
		(.033)	(.046)	(.046)						
θ	1	−.975	−.035	.051	—	.351	.380	.182	.546	.579
		(.018)	(.055)	(.058)						
Mean squared error in estimation of frailties								.287	.110	.109
Simulation 2: $\theta = 4$, x one component										
β	.5	−.095	−.023	−.016	.611	.761	.785	.694	.769	.778
		(.069)	(.077)	(.078)						
θ	4	−1.457	−.128	.139	—	1.125	1.217	1.615	1.481	1.576
		(.161)	(.148)	(.158)						
Mean squared error in estimation of frailties								.324	.210	.209
Simulation 3: $\theta = 1$, x four components										
β_1	.5	−.184	−.105	−.087	.248	.430	.487	.360	.441	.462
		(.036)	(.044)	(.046)						
β_2	−.5	.177	.050	.027	.247	.431	.487	.358	.460	.483
		(.036)	(.046)	(.048)						
β_3	.8	−.231	−.039	−.001	.249	.433	.490	.337	.443	.469
		(.034)	(.044)	(.047)						
β_4	−.8	.187	−.021	−.064	.253	.436	.493	.319	.437	.464
		(.032)	(.044)	(.046)						
θ	1	−.985	−.177	.168	—	.314	.437	.138	.427	.545
		(.014)	(.043)	(.054)						
Mean squared error in estimation of frailties								.300	.131	.130
Simulation 4: $\theta = 4$, x four components										
β_1	.5	−.028	.071	.099	.576	.759	.858	.776	.885	.930
		(.078)	(.088)	(.093)						
β_2	−.5	.156	.078	.076	.573	.754	.853	.699	.841	.869
		(.070)	(.084)	(.087)						
β_3	.8	−.114	.009	.048	.572	.754	.852	.684	.779	.819
		(.068)	(.078)	(.082)						
β_4	−.8	.099	−.030	−.097	.571	.755	.853	.665	.831	.841
		(.066)	(.083)	(.084)						
θ	4	−2.087	−.694	.342	—	.977	1.348	1.421	1.235	1.585
		(.142)	(.123)	(.158)						
Mean squared error in estimation of frailties								.457	.311	.309

[a] Each simulation gives a S.E. of estimate. Tabulated value is the average S.E.
[b] Tabulated value is the S.E. of the 100 simulated estimates.

The average standard error of estimates is always less than the standard error obtained over simulations for BLUP and ML, but alternates for REML estimators. The result indicates that the standard error usually reported for REML estimators is, on average, about right, whereas the standard errors for BLUP and ML will usually be underreported.

Finally the BLUP, ML, and REML estimates of the frailty realisations are compared to the true generated values. It is seen from the mean squared errors reported in Table 1 that REML and ML perform equally well and both do substantially better than BLUP.

ACKNOWLEDGEMENT

This research was partially supported by the Australian Research Council.

Résumé

Une méthode d'estimation pour les modèles mixtes généralisés est appliquée à l'estimation des paramètres de régression dans les modèles à risques instantanés proportionnels pour les données de survie en présence d'observations répétes de l'événement d'intérêt pour chaque sujet. L'effet sujet est inclus dans le modèle sous forme d'un terme aléatoire de fragilité. Les meilleurs prédicteurs linéaires sans biais sont utilisés comme valeur initiale dans le calcul des estimateurs du maximum de vraisemblance et du maximum de vraisemblance résiduelle.

References

Cox, D. R. (1972). Regression models and life tables (with Discussion). *Journal of the Royal Statistical Society, Series B* **45,** 311–354.

Fellner, W. H. (1986). Robust estimation of variance components. *Technometrics* **28,** 51–60.

Fellner, W. H. (1987). Sparse matrices, and the estimation of variance components by likelihood methods. *Communications in Statistics—Simulation* **16,** 439–463.

Henderson, C. R. (1975). Best linear unbiased estimation and prediction under a selection model. *Biometrics* **31,** 423–447.

Hougaard, P. (1986a). Survival models for heterogeneous populations derived from stable distributions. *Biometrika* **73,** 387–396.

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

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

Prentice, R. L., Williams, B., and Peterson, A. V. (1981). On the regression analysis of multivariate failure time data. *Biometrika* **68,** 373–379.

Schall, R. (1991). Estimation in generalised linear models with random effects. *Biometrika* **78,** 719–727.

Thompson, R. (1980). Maximum likelihood estimation of variance components. *Math. Operationsforsch. Statist., Ser. Statistics* **11,** 545–561.

Wei, L. J., Lin, D. Y., and Weissfeld, L. (1989). Regression analysis of multivariate incomplete failure time data by modelling marginal distributions. *Journal of the American Statistical Association* **84,** 1065–1073.

Received April 1992; *revised September* 1992; *accepted October* 1992.