Journal of Reliability and Statistical Studies; ISSN (Print): 0974-8024, (Online):2229-5666
Vol. 5, Issue Special (2012): 97-106

ON THE UNIFORM FRAILTY MODEL WITH PENALIZED LIKELIHOOD AND CLUSTERED DATA

Emmanouil Androulakis, Christos Koukouvinos and Filia Vonta

Department of Mathematics, National Technical University of Athens,
Zografou 15773, Athens, Greece
Email: manolis.androulakis@gmail.com; ckoukouv@math.ntua.gr;
vonta@math.ntua.gr

Abstract

In the field of survival analysis, when heterogeneity is suspected across study subjects, a model that can account for that variability is preferred. Moreover, an important and challenging task in that field is to efficiently select a subset of significant variables upon which the hazard function depends. To this end, frailty models along with the penalized likelihood methodology can be applied. In this paper, we extend the Gamma frailty model methodology of Fan and Li (2002) to the Uniform frailty model, based on the results of our previous work (Androulakis et al., 2012). Theoretical findings are illustrated via a thorough simulation study.

Key words: Frailty model; Uniform distribution; Penalized likelihood; Variable selection.

1. Introduction

The problem of analyzing time to event data arises in a number of applied fields, such as medicine, biology, public health, epidemiology, engineering, economics and demography. Due to the presence of censored data, the traditional statistical analysis tools cannot be used. As a result, survival analysis methods have emerged and the Cox proportional hazards model is still the most popular approach.

However, the survival times of subjects are assumed to be independent when the Cox model is used. In the cases when this assumption is violated, frailty models are used, either univariate or shared, which allow for such dependence. A frailty random effect is imposed on the model, following a specific distribution, so as to describe heterogeneity among individuals or dependence within a cluster of individuals.

Hougaard (1984, 1995, 2000) discussed the choice of frailty distribution, especially for the shared frailty case. For example, the Gamma frailty model describes high late dependence while the Inverse Gaussian frailty model describes moderate dependence along the whole range of times. Moreover, the Uniform frailty case could be considered for example when one analyzes data coming from the reliability of two components of a system, exposed to the same environmental random stress factor. When the operating environment of the system differs from the laboratory environment, in the case for example when it is more severe, it makes sense to restrict the support of the random stress factor to a fixed interval $[\alpha, b]$ (Lee and Klein, 1988). When both α and b are below the value 1, this corresponds to an operating environment which is less severe than the test environment, while if α and b are greater than 1, the operating environment is more severe. Also when $[\alpha, b]$ contains 1, this corresponds to an environment which incurs the possibility of no differential effect from that found in the laboratory.

Among other distributions, the Gamma is typically used to fit the frailty term, due to mathematical reasons and its simplicity. There are no known biological reasons motivating that choice. The same distribution was also used in the paper of Fan and Li (2002), in which the Gamma frailty model with penalized likelihood was analyzed. Their methodology was extended in Androulakis et al. (2012) to a whole class of frailty models, in which a generalized form of the penalized likelihood function designed for clusters is given (see also Hougaard (2000) for the non-penalized case). This approach allows the direct use of many continuous distributions for the frailty parameter. Two theoretical examples were analyzed, concerning the cases of the Gamma frailty (which slightly corrects the corresponding likelihood of Fan and Li, 2002) and the Inverse Gaussian frailty. In this work, we focus on the Uniform frailty model with penalized likelihood and clustered data.

This paper is organized as follows. In Section 2, we provide a brief description of our previous work. In Section 3, the Uniform frailty model with penalized likelihood is analyzed. In the next Section 4, we provide simulations to illustrate our methodology for various sample sizes and cluster sizes. The results are discussed and some comments are given. In the last Section 5, some general conclusions are presented.

2. Frailty models with penalized likelihood for clustered data

In this section, we provide a brief description of our previous work (Androulakis et al., 2012). Assume that the data are described by $(z_{ij}, \delta_{ij}, \boldsymbol{x}_{ij})$ where $z_{ij} = \min\{T_{ij}, C_{ij}\}$, T_{ij} is the survival time, C_{ij} the censoring time, δ_{ij} the censoring indicator and $\boldsymbol{x}_{ij}$ the $d-$dimensional vector of covariates of the j^{th} individual belonging to the cluster i. We assume that the hazard rate of the j^{th} subject in the i^{th} subgroup-cluster is given conditionally on the covariates $\boldsymbol{x}_{ij}$ and the shared frailty u_i as

$$h_{ij}(t \mid \boldsymbol{x}_{ij}, u_i) = u_i h_0(t) e^{\boldsymbol{x}_{ij}^T \boldsymbol{\beta}} \tag{1}$$

where $\boldsymbol{\beta} \in R^d$ is the parameter of interest and $h_0(t)$ the baseline hazard function.

Let n be the number of clusters, J_i the size of the cluster i, $i = 1, \dots, n$ and F_{u_i} is the c.d.f. of the shared frailty u_i. The generalized penalized loglikelihood has the following form:

$$\sum_{i=1}^{n}\sum_{j=1}^{J_i}\delta_{ij}\ln(h_0(z_{ij})) + \sum_{i=1}^{n}\sum_{j=1}^{J_i}\delta_{ij}\boldsymbol{x}_{ij}^T\boldsymbol{\beta} + \sum_{i=1}^{n}\ln\left(\left|L^{(A_i)}\left(\sum_{j=1}^{J_i}e^{\boldsymbol{x}_{ij}^T\boldsymbol{\beta}}H_0(z_{ij})\right)\right|\right) - n\sum_{j=1}^{d}p_\lambda(|\beta_j|) \tag{2}$$

where

$$L^{(A_i)}(x) = \int_0^\infty e^{-ux}(-1)^{A_i}u^{A_i}\,dF_u(u) \tag{3}$$

denotes the $A_i th$ derivative of the Laplace transform $L(x)=\int_0^\infty e^{-ux}\,dF_u(u)$ of the distribution function F_u of the frailty, $p_\lambda(.)$ is an appropriate penalty function and λ is the tuning parameter which can be chosen by data-driven approaches, such as cross-validation (CV) and generalized cross-validation (GCV) (Craven and Wahba, 1979). We should point out here that different continuous distributions of the frailty parameter give rise to different Laplace transforms and therefore a class of frailty models arises with corresponding penalized loglikelihood given in (2).

The most used penalties, these are the L_1 penalty $p_\lambda(|\beta|)=\lambda|\beta|$, for $\beta\in R$, which results in the Least Absolute Shrinkage and Selection Operator method (LASSO) (Tibshirani, 1996), and the hard thresholding penalty (Antoniadis, 1997), $p_\lambda(|\beta|)=\lambda^2-(|\beta|-\lambda)^2 I(|\beta|<\lambda)$, where $I(\cdot)$ is an indicator function. However, these penalties do not simultaneously satisfy the necessary mathematical conditions for unbiasedness, sparsity and continuity. Therefore, Fan and Li (2001) proposed the continuous differentiable penalty function, known as Smoothly Clipped Absolute Deviation penalty (SCAD), the first derivative of which is defined by

$p'_\lambda(\beta)=\lambda\left\{I(\beta\le\lambda)+\frac{(\alpha\lambda-\beta)_+}{(\alpha-1)\lambda}I(\beta>\lambda)\right\}$, for some $\beta>0$ and $\alpha>2$,

with $p_\lambda(0)=0$. For the choice of α, according to the relevant literature (Fan and Li, 2001 and 2002), the value $\alpha\approx 3.7$ appears to perform quite satisfactorily in numerous variable selection problems.

We also considered the 'least informative' nonparametric modeling for $H_0(\cdot)$ in which $H_0(t)$ has a possible jump of size μ_l at the observed failure time z_l. Then,

$$H_0(t)=\sum_{l=1}^{N}\mu_l I(z_l\le t), \tag{4}$$

where $z_1,\ldots,z_N$ are pooled observed failure times. Substituting (4) into (2) we obtain the following penalized profile likelihood

$$\sum_{i=1}^{n}\sum_{j=1}^{J_i}\delta_{ij}\mathbf{x}_{ij}^T\boldsymbol{\beta}+\sum_{l=1}^{N}\ln\mu_l+\sum_{i=1}^{n}\ln\left(\int_0^\infty \exp\left(-u_i\sum_{j=1}^{J_i}e^{\mathbf{x}_{ij}^T\boldsymbol{\beta}}\sum_{l=1}^{N}\mu_l I(z_l\le z_{ij})\right)u_i^{A_i}\,dF_{u_i}(u_i)\right)-n\sum_{j=1}^{d}p_\lambda(|\beta_j|). \tag{5}$$

The derivative of the above loglikelihood without the penalty term and with respect to μ_l, $l=1,\ldots,N$ gives

$$\frac{1}{\mu_l}=\sum_{i=1}^{n}\frac{\left|L^{(A_i+1)}\left(\sum_{j=1}^{J_i}e^{\mathbf{x}_{ij}^T\boldsymbol{\beta}}\sum_{l=1}^{N}\mu_l I(z_l\le z_{ij})\right)\right|}{\left|L^{(A_i)}\left(\sum_{j=1}^{J_i}e^{\mathbf{x}_{ij}^T\boldsymbol{\beta}}\sum_{l=1}^{N}\mu_l I(z_l\le z_{ij})\right)\right|}\sum_{j=1}^{J_i}e^{\mathbf{x}_{ij}^T\boldsymbol{\beta}}I(z_l\le z_{ij}). \tag{6}$$

Moreover, the first and second gradient derivatives of (5) without the penalty term, with respect to $\boldsymbol{\beta}=(\beta_1,\ldots,\beta_d)^T$ are given below. More specifically, the derivative with respect to β_{k_1}, $k_1=1,\ldots,d$, is given as

$$\frac{\vartheta l(\boldsymbol{\beta})}{\vartheta\beta_{k_1}}=\sum_{i=1}^{n}\sum_{j=1}^{J_i}\delta_{ij}x_{ijk_1}-\sum_{i=1}^{n}\frac{|L^{(A_i+1)}(x)|}{|L^{(A_i)}(x)|}\sum_{j=1}^{J_i}e^{x_{ij}^T\boldsymbol{\beta}}\sum_{l=1}^{N}\mu_l I(z_l\le z_{ij})\,x_{ijk_1} \tag{7}$$

while for $\dfrac{\vartheta^2 l(\boldsymbol{\beta})}{\vartheta\beta_{k_1}\vartheta\beta_{k_2}}$, $k_1=1,\ldots,d$, $k_2=1,\ldots,d$, we have

$$\sum_{i=1}^{n}\left\{\frac{|L^{(A_i+2)}(x)|}{|L^{(A_i)}(x)|}-\left(\frac{|L^{(A_i+1)}(x)|}{|L^{(A_i)}(x)|}\right)^2\right\}\sum_{j=1}^{J_i}e^{x_{ij}^T\boldsymbol{\beta}}\sum_{l=1}^{N}\mu_l I(z_l\le z_{ij})x_{ijk_1}\sum_{j=1}^{J_i}e^{x_{ij}^T\boldsymbol{\beta}}\sum_{l=1}^{N}\mu_l I(z_l\le z_{ij})x_{ijk_2}$$

$$-\sum_{i=1}^{n}\frac{|L^{(A_i+1)}(x)|}{|L^{(A_i)}(x)|}\sum_{j=1}^{J_i}e^{x_{ij}^T\boldsymbol{\beta}}\sum_{l=1}^{N}\mu_l I(z_l\le z_{ij})x_{ijk_1}x_{ijk_2} \tag{8}$$

where $x=\sum_{j=1}^{J_i}e^{x_{ij}^T\boldsymbol{\beta}}\sum_{l=1}^{N}\mu_l I(z_l\le z_{ij})$, which actually depends on i.

3. The Uniform frailty model with penalized likelihood

Consider a uniformly distributed frailty on [b, α] where $a>b>0$. Then $f_u(u)=1/(a-b),\ \ b\le u\le a$ and the Laplace transform is given as

$$L(x)=\frac{1}{(a-b)x}(e^{-bx}-e^{-ax}).$$

(9)

From (3) by a series of integration by parts we get that $L^{(A_i)}(x)$ is equal to

$$(-1)^{A_i}\left\{e^{-bx}\left(\frac{b^{A_i}}{(a-b)x}+\frac{A_ib^{A_i-1}}{(a-b)x^2}+\frac{A_i(A_i-1)b^{A_i-2}}{(a-b)x^3}+\ldots+\frac{A_i(A_i-1)\ldots2b^1}{(a-b)x^{A_i}}+\frac{A_i(A_i-1)\ldots2}{(a-b)x^{A_i+1}}\right)\right.$$

$$\left.-e^{-ax}\left(\frac{a^{A_i}}{(a-b)x}+\frac{A_ia^{A_i-1}}{(a-b)x^2}+\frac{A_i(A_i-1)a^{A_i-2}}{(a-b)x^3}+\ldots+\frac{A_i(A_i-1)\ldots2a^1}{(a-b)x^{A_i}}+\frac{A_i(A_i-1)\ldots2}{(a-b)x^{A_i+1}}\right)\right\}$$

$$=(-1)^{A_i}\left\{\frac{e^{-bx}}{(a-b)x}\sum_{j=0}^{A_i}j!\binom{A_i}{j}\left(\frac{1}{x}\right)^j b^{A_i-j}-\frac{e^{-ax}}{(a-b)x}\sum_{j=0}^{A_i}j!\binom{A_i}{j}\left(\frac{1}{x}\right)^j a^{A_i-j}\right\}$$

$$=\frac{(-1)^{A_i}}{(a-b)x}\left\{e^{-bx}\frac{e^{bx}\left(2x^{-A_i}-x^{-A_i+1}\right)\Gamma\left(A_i+1,bx\right)}{-2+x}-e^{-ax}\frac{e^{ax}\left(2x^{-A_i}-x^{-A_i+1}\right)\Gamma\left(A_i+1,ax\right)}{-2+x}\right\}$$

which simplifies to

$$\frac{(-1)^{A_i}}{(a-b)x}\frac{2x^{-A_i}-x^{-A_i+1}}{-2+x}\{\Gamma\left(A_i+1,bx\right)-\Gamma\left(A_i+1,ax\right)\}=$$

$$\frac{(-1)^{A_i+1}}{(a-b)x^{A_i+1}}\{\Gamma\left(A_i+1,bx\right)-\Gamma\left(A_i+1,ax\right)\} \tag{10}$$

where $\Gamma(a,x)$ is the incomplete Gamma function defined as $\int_x^{\infty} t^{a-1}e^{-t}dt$. The loglikelihood (5) for the appropriate penalty function, becomes

$$\sum_{l=1}^{N}\ln\mu_l+\sum_{i=1}^{n}\sum_{j=1}^{J_i}\delta_{ij}\boldsymbol{x}_{ij}^T\boldsymbol{\beta}+\sum_{l=1}^{n}\ln\left(\Gamma(A_i+1,bx)-\Gamma(A_i+1,ax)\right)-$$
$$\sum_{l=1}^{n}\ln(a-b)-\sum_{l=1}^{n}(A_i+1)\ln x-n\sum_{j=1}^{d}p_{\lambda}(|\beta_j|). \tag{11}$$

The derivative with respect to μ_l, $l=1,\ldots,N$ by (6) and (10) becomes

$$\frac{1}{\mu_l}=\sum_{l=1}^{n}\left\{\frac{\Gamma\left(A_i+2,bx\right)-\Gamma\left(A_i+2,ax\right)}{\left(\Gamma\left(A_i+1,bx\right)-\Gamma\left(A_i+1,ax\right)\right)x}\right\}\sum_{j=1}^{J_i}e^{x_{ij}^T\beta}I(z_l\le z_{ij}). \tag{12}$$

Analogous expressions hold for the first and second derivatives with respect to β.

4. Simulations

We perform in this section a number of simulations in order to demonstrate our methodology. We were based on the simulation scheme of Fan and Li (2002), therefore 100 datasets were simulated, consisting of n groups and J subjects in each group from the exponential hazard frailty model

$$h(t\,|\,\mathbf{x},u)=u\exp(\mathbf{x}^T\boldsymbol{\beta}),$$

where the true parameter $\beta_0=(0.8,0,0,1,0,0,0.6,0)$ and the x_i are marginally standard normal and the correlation between x_i and x_j is $\rho^{|i-j|}$ with $\rho=0.5$. The distribution of the censoring time is an exponential distribution with mean $U\exp(\boldsymbol{x}^T\boldsymbol{\beta}_0)$, where U is randomly generated from the Uniform distribution over [1,3] for each simulated dataset so that about 30% data are censored. Here β_0 is regarded as a known constant so that the censoring scheme is noninformative. Concerning the frailty u, we present three cases. First, we consider a Uniform frailty with $b=0.25$ and $a=0.75$. Secondly, we took $b=1.25$ and $a=1.75$ while in the last case, the values of b and a were 0.75 and 1.25, respectively.

The performance of the proposed generalized penalized likelihood is examined in terms of its model errors, model complexity and accuracy. Model errors of our procedure are compared to those of the maximum profile likelihood estimators. Specifically, we present the Median of Relative Model Errors (MRME) over the 100 simulated data sets with some combinations of *n* and *J*. Moreover, the average number of zero coefficients are also reported in the following Tables, in which the column labeled "correct" presents the average restricted only to the true zero coefficients, while the column labeled "incorrect" depicts the average of coefficients erroneously set to 0. We also test the accuracy of the standard error formula proposed in Fan and Li (2002). The median absolute deviation divided by 0.6745, denoted by SD, of the 100 estimated coefficients in the 100 simulations can be regarded as the true standard error except the Monte Carlo error. The median of the 100 estimated SDs, resulting from 100 simulations, denoted by SD_m, and the median absolute deviation error of the 100 estimated standard errors divided by 0.6745, denoted by SD_{mad}, gauge the overall performance of the standard error formula. In our simulations, the standard errors of estimated coefficients were set to be 0, if they were excluded from the selected model, therefore, only the results for the non-zero coefficients are shown. We used the penalized likelihood approach with the SCAD penalty (SCAD), the L_1 penalty (LASSO) and the hard thresholding penalty (Hard), along with the generalized cross-validation as a tuning parameter selection method. Moreover we present the results for the Oracle estimator, obtained by fitting the ideal model consisting only of the variables X_1, X_4 and X_7. All computations were conducted using Matlab codes.

From Tables 1-6, we extract the following conclusions:

• In model identifications, the SCAD and Hard methods have a high chance of correctly setting the non-significant coefficients to zero. However, the LASSO method performs also well in many cases.

• The SCAD and Hard outperform the LASSO in general and perform as well as the oracle estimator in terms of MRME.

• The standard error formula proposed by Fan and Li (2002) performs also well.

• In general, based on our results, the proposed methodology can be considered quite reliable and effective in terms of estimation and variable selection.

5. Concluding remarks

In this paper, we extend the Gamma frailty model with penalized likelihood, proposed by Fan and Li (2002), to the Uniform frailty model. The theory presented is supported by a thorough simulation study, where the performance of our methodology is examined in terms of model errors, model complexity and accuracy. Based on our results, we can conclude that our proposed methodology can be considered very effective for parameter estimation and variable selection, when a Uniform frailty model is considered.

Acknowledgment

We would like to thank Professors Dennis K. J. Lin and Runze Li for sending the Matlab codes for the procedures proposed in their papers.

CASE 1: b=0.25 and α=0.75

Table 1: MRME and average no. of 0 coefficients

Method	**MRME(%)**	**Aver. no. of 0 coeff.**	
		correct	incorrect
n=50, J=2			
SCAD	0.2723	4.4100	0.1300
LASSO	0.5858	4.0400	0.0200
Hard	0.4425	4.5700	0.1800
Oracle	0.3114	5	0
n=50, J=5			
SCAD	0.6909	4.8000	0
LASSO	1.1094	4.0200	0
Hard	0.6308	4.8400	0
Oracle	0.6318	5	0
n=100, J=2			
SCAD	0.6193	4.7200	0
LASSO	1.1136	3.8900	0
Hard	0.6737	4.8700	0.0200
Oracle	0.5720	5	0
n=100, J=5			
SCAD	0.8584	4.4600	0
LASSO	1.2529	3.9700	0
Hard	0.8928	4.8700	0
Oracle	0.8763	5	0

Table 2: Standard deviations

	$\hat{\beta}_1$			$\hat{\beta}_4$			$\hat{\beta}_7$		
Method	SD	SD_m	SD_{mad}	SD	SD_m	SD_{mad}	SD	SD_m	SD_{mad}
n=50, J=2									
SCAD	0.1346	0.1188	0.0066	0.1275	0.1040	0.0084	0.1149	0.0916	0.0106
LASSO	0.1106	0.0821	0.0120	0.0985	0.0659	0.0127	0.1018	0.0725	0.0082
Hard	0.1359	0.1062	0.0084	0.1365	0.1145	0.0081	0.1158	0.0902	0.0095
Oracle	0.1256	0.1091	0.0078	0.1298	0.0901	0.0085	0.1002	0.0809	0.0068
n=50, J=5									
SCAD	0.0597	0.0568	0.0043	0.0731	0.0561	0.0050	0.0595	0.0537	0.0041
LASSO	0.0632	0.0402	0.0034	0.0612	0.0445	0.0036	0.0615	0.0454	0.0043
Hard	0.0632	0.0570	0.0042	0.0729	0.0563	0.0051	0.0599	0.0550	0.0047
Oracle	0.0635	0.0569	0.0041	0.0756	0.0567	0.0051	0.0581	0.0552	0.0045
n=100, J=2									
SCAD	0.0744	0.0603	0.0043	0.0810	0.0712	0.0012	0.0730	0.0657	0.0057
LASSO	0.0810	0.0479	0.0050	0.0963	0.0506	0.0060	0.0996	0.0446	0.0065
Hard	0.0649	0.0560	0.0040	0.0815	0.0575	0.0041	0.0715	0.0549	0.0044
Oracle	0.0683	0.0565	0.0038	0.0785	0.0582	0.0036	0.0685	0.0552	0.0040
n=100, J=5									
SCAD	0.0435	0.0411	0.0020	0.0531	0.0404	0.0020	0.0445	0.0401	0.0027
LASSO	0.0480	0.0339	0.0018	0.0584	0.0345	0.0028	0.0405	0.0309	0.0025
Hard	0.0428	0.0411	0.0020	0.0535	0.0404	0.0023	0.0411	0.0406	0.0026
Oracle	0.0412	0.0412	0.0019	0.0504	0.0405	0.0020	0.0398	0.0406	0.0026

CASE 2: b=1.25 and α=1.75

Table 3: MRME and average no. of 0 coefficients

Method	MRME(%)	Aver. no. of 0 coeff.	
		correct	incorrect
n=50, J=2			
SCAD	0.4762	4.4100	0.1500
LASSO	1.0312	3.9600	0.0600
Hard	0.4494	4.8700	0.1900
Oracle	0.3112	5	0
n=50, J=5			
SCAD	0.5304	4.2800	0.0300
LASSO	0.8233	4.0600	0
Hard	0.6513	4.9300	0
Oracle	0.5767	5	0
n=100, J=2			
SCAD	0.5435	4.5700	0.0100
LASSO	1.0674	3.8200	0.0100
Hard	0.4759	4.7400	0.0600
Oracle	0.4614	5	0
n=100, J=5			
SCAD	0.5802	4.6500	0
LASSO	0.8838	3.9900	0
Hard	0.5738	4.4800	0
Oracle	0.5242	5	0

Table 4: Standard deviations

	$\hat{\beta}_1$			$\hat{\beta}_4$			$\hat{\beta}_7$		
Method	SD	SD_m	SD_{mad}	SD	SD_m	SD_{mad}	SD	SD_m	SD_{mad}
n=50, J=2									
SCAD	0.1344	0.1046	0.0086	0.1261	0.0859	0.0126	0.1215	0.1013	0.0113
LASSO	0.1144	0.0721	0.0093	0.1038	0.0755	0.0180	0.1497	0.0816	0.0137
Hard	0.1383	0.1045	0.0119	0.1127	0.0861	0.0061	0.1355	0.1071	0.0194
Oracle	0.1267	0.0972	0.0089	0.1234	0.0965	0.0076	0.1289	0.0972	0.0093
n=50, J=5									
SCAD	0.0968	0.0896	0.0103	0.1096	0.0684	0.0089	0.0810	0.0777	0.0077
LASSO	0.0858	0.0606	0.0138	0.0954	0.0630	0.0093	0.0889	0.0560	0.0123
Hard	0.0711	0.0722	0.0092	0.0890	0.0717	0.0081	0.0754	0.0716	0.0073
Oracle	0.0720	0.0741	0.0063	0.0817	0.0738	0.0051	0.0673	0.0731	0.0049
n=100, J=2									
SCAD	0.0774	0.0702	0.0096	0.0853	0.0701	0.0083	0.0702	0.0671	0.0092
LASSO	0.0687	0.0612	0.0045	0.0774	0.0636	0.0047	0.0805	0.0552	0.0048
Hard	0.0717	0.0715	0.0077	0.0853	0.0717	0.0074	0.0785	0.0711	0.0049
Oracle	0.0727	0.0748	0.0051	0.0767	0.0737	0.0059	0.0757	0.0733	0.0058
n=100, J=5									
SCAD	0.0479	0.0475	0.0050	0.0632	0.0483	0.0043	0.0752	0.0466	0.0059
LASSO	0.0491	0.0440	0.0021	0.0598	0.0459	0.0022	0.0527	0.0415	0.0020
Hard	0.0498	0.0498	0.0027	0.0553	0.0505	0.0028	0.0477	0.0495	0.0030
Oracle	0.0478	0.0500	0.0026	0.0558	0.0507	0.0026	0.0482	0.0496	0.0027

CASE 3: b=0.75 and α=1.25

Table 5: MRME and average no. of 0 coefficients

Method	MRME(%)	Aver. no. of 0 coeff.	
		correct	incorrect
n=50, J=2			
SCAD	0.3542	4.6100	0.0900
LASSO	1.0597	3.8600	0.0400
Hard	0.3773	4.3300	0.1500
Oracle	0.2797	5	0
n=50, J=5			
SCAD	0.6394	4.1900	0
LASSO	1.0649	4.0500	0
Hard	0.7948	4.1300	0
Oracle	0.5157	5	0
n=100, J=2			
SCAD	0.5767	4.9800	0.0100
LASSO	0.8588	4.1400	0
Hard	0.6065	4.9200	0.0100
Oracle	0.5662	5	0
n=100, J=5			
SCAD	0.6835	4.7100	0
LASSO	0.8057	4.0600	0
Hard	0.7770	4.1800	0
Oracle	0.6558	5	0

Table 6: Standard deviations

	$\hat{\beta}_1$			$\hat{\beta}_4$			$\hat{\beta}_7$		
Method	SD	SD_m	SD_{mad}	SD	SD_m	SD_{mad}	SD	SD_m	SD_{mad}
n=50, J=2									
SCAD	0.1050	0.0801	0.0074	0.1136	0.0904	0.0123	0.1117	0.0981	0.0123
LASSO	0.1245	0.0755	0.0075	0.1143	0.0804	0.0062	0.1263	0.0885	0.0101
Hard	0.1080	0.0771	0.0091	0.1093	0.0884	0.0120	0.1152	0.0895	0.0115
Oracle	0.1061	0.0855	0.0072	0.1119	0.0952	0.0058	0.1053	0.0958	0.0088
n=50, J=5									
SCAD	0.0643	0.0626	0.0057	0.0841	0.0608	0.0055	0.0811	0.0612	0.0066
LASSO	0.0569	0.0524	0.0031	0.0750	0.0524	0.0037	0.0752	0.0687	0.0035
Hard	0.0677	0.0666	0.0058	0.0864	0.0644	0.0067	0.0773	0.0675	0.0063
Oracle	0.0618	0.0645	0.0059	0.0845	0.0616	0.0053	0.0725	0.0647	0.0048
n=100, J=2									
SCAD	0.0831	0.0600	0.0038	0.0847	0.0616	0.0038	0.0935	0.0613	0.0088
LASSO	0.0800	0.0557	0.0040	0.0858	0.0592	0.0034	0.0993	0.0617	0.0048
Hard	0.0709	0.0658	0.0057	0.0867	0.0660	0.0055	0.0933	0.0639	0.0051
Oracle	0.0702	0.0673	0.0047	0.0830	0.0670	0.0049	0.0914	0.0648	0.0043
n=100, J=5									
SCAD	0.0623	0.0454	0.0026	0.0575	0.0456	0.0027	0.0503	0.0459	0.0026
LASSO	0.0490	0.0397	0.0020	0.0520	0.0410	0.0021	0.0472	0.0383	0.0021
Hard	0.0645	0.0456	0.0024	0.0600	0.0476	0.0036	0.0554	0.0476	0.0035
Oracle	0.0612	0.0453	0.0026	0.0583	0.0456	0.0027	0.0480	0.0459	0.0026

References

1. Androulakis, E., Koukouvinos, C. and Vonta, F. (2012). Estimation and variable selection via frailty models with penalized likelihood, Statistics in Medicine, DOI: 10.1002/sim.5325.
2. Antoniadis, A. (1997). Wavelets in statistics: a review (with discussion), J. Italian Statist. Assos., 6, p. 97-144.
3. Craven, P. and Wahba, G. (1979). Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalized cross-validation, Numer. Math., 31, p. 377-403.
4. Fan, J. and Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties, J. Amer. Statist. Assos., 96, p. 1348-1360.
5. Fan, J. and Li, R. (2002). Variable selection for Cox's proportional hazards model and frailty model, The Annals of Statistics, 30, p. 74-99.
6. Hougaard, P. (1984). Life table methods for heterogeneous populations: distributions describing the heterogeneity, Biometrika, 71, p. 75-83.
7. Hougaard, P. (1995). Frailty models for survival data, Lifetime Data Analysis, 1, p. 255-273.
8. Hougaard, P. (2000). Analysis of multivariate survival data, Springer, New York.
9. Lee, S. and Klein, J. P. (1988). Bivariate models with a random environmental factor. Industrial J. Productivity, Reliability, and Quality Control, 13, p. 1-18.
10. Tibshirani, R. J. (1996). Regression shrinkage and selection via the LASSO, J. Roy. Statist. Soc. Ser. B, 58, p. 267-288.