QBS
Vol. 36, No. 1, pp. 79-84 (2017)
https://doi.org/10.22283/qbs.2017.36.1.79
pISSN 2288-1344 eISSN 2508-7185 **Original Article**

Quantitative Bio-Science
Official Journal of the Institute of Natural Science at Keimyung University
Available Online at https://qbs.kmu.ac.kr:442/

Variable Selection Procedures in Competing-Risks Models Using Penalized Likelihood

Il Do Ha*

Department of Statistics, Pukyong National University, Busan 48513, Korea
(Received May 3, 2017; Revised May 11, 2017; Accepted May 13, 2017)

ABSTRACT

Competing risks data commonly arise when an occurrence of an event precludes other type of events from being observed. Recently, competing-risks models with frailty have been studied for clustered competing-risks data that may be correlated. In this paper, we propose a variable selection procedure for fixed effects in the cause-specific hazard frailty model for the clustered competing-risks data using a penalized likelihood. Here we consider two popular penalty functions, least absolute shrinkage and selection operator (LASSO) and smoothly clipped absolute deviation (SCAD). We derive simple matrix forms for the variable selection procedure. The usefulness of the proposed method is illustrated using a practical example data set.

Key words : Competing risks, Frailty models, H-likelihood, Penalized likelihood, Variable selection

1. Introduction

Two broad classes of competing-risks regression models are cause-specific hazard model [1] and subdistribution hazard model [2]. The former is used to model the cause-specific hazard of each event type separately, while the latter is used to model the hazard function of subdistribution (subhazard) for a particular event of interest. Recently, these models have been extended to clustered competing risks data via frailty [3,4], i.e. semi-parametric frailty models have been applied to these correlated competing risks data.

It is important to determine relevant variables in the analysis of regression models. In classical regression models, there are several techniques for variable selection (e.g., forward selection, backward elimination and stepwise selection). However, these traditional methods may be computationally intensive for many covariates and often suffer from high variability [5,6]. Recently, variable selection methods using a penalized likelihood allowing for various penalty functions have been widely developed in linear models, generalized linear models and Cox's proportional hazards (PH) models [6]. The main advantage of this method is to select important covariates and to estimate the regression coefficients, simultaneously; i.e. it deletes insignificant variables by estimating their coefficients as zero.

Ha et al. [7] have developed a variable selection procedure for fixed effects in subhazard competing-risks frailty models using a penalized hierarchical likelihood (h-likelihood) [8] with a penalty function (e.g. LASSO [9], SCAD [6]). The h-likelihood [10,11] avoids the integration of random effects itself, whereas the marginal likelihood approach often involves intractable integrations when eliminating the frailties [12]. In this paper, we extend this penalized variable-selection procedure [8] to a cause-specific hazard frailty model.

In Section 2 we describe the cause-specific competing-risks models with frailty and outline the h-likelihood procedure. In Section 3, we propose a variable selection procedure for the

* Correspondence should be addressed to Dr. Il Do Ha, Department of Statistics, Pukyong National University, Busan 48513, Korea. Tel: +82-51-629-5536, Fax: +82-51-629-5544, E-mail: idha1353@pknu.ac.kr

cause-specific hazard frailty models using a penalized h-likelihood. The proposed method is illustrated with a competing-risks data set from multi-center bladder cancer trial in Section 4. Finally, some discussions are provided in Section 5.

2. H-likelihood for Cause-Specific Hazard Frailty Model

2.1 The model

Suppose there are $i=1,\ldots,q$ centers where each center (cluster) has $j=1,\ldots,n_i$ observations, so that the total sample size is $n=\sum_{i=1}^{q} n_i$. For a patient j in center i, let T_{ijk} be time to event from cause k and let $\varepsilon_{ij}\in\{1,2,\ldots,K\}$ be the corresponding cause of event. Let C_{ij} denote independent censoring time. The log-frailty for center i is denoted by V_i. Assume that given $V_i=v_i$, C_{ij} is conditionally independent and noninformative of $(T_{ij}, \varepsilon_{ij})$ for $j=1,\ldots,n_i$. For simplicity, in this paper we consider two types of events ($k=1,2$); let T_{ij1} be an event of interest (Type 1) and T_{ij2} be a competing event (Type 2).

The cause-specific hazard function conditional on the common log-frailty v_i for the jth observation in the patient i who failed from cause k ($k=1,2$) is described as [4]

$$\lambda_{ijk}(t\,|\,v_i)=\lambda_{0k}(t)\exp(x_{ij}^T\beta_k+v_i), \tag{1}$$

where v_i is an unobserved random variable from a univariate distribution with parameter θ, $\lambda_{0k}(t)$ is the unspecified baseline hazard function for event type k and $\beta_k=(\beta_{k1},\beta_{k2},\ldots,\beta_{kp})^T$ is a $p\times 1$ vector of regression parameters for event type k, and $x_{ij}=(x_{ij1},\ldots,x_{ijp})^T$ is a $p\times 1$ vector of fixed covariates corresponding to β_k. Let $\beta=(\beta_1^T,\beta_2^T)^T$ be a $2p\times 1$ vector of all the regression coefficients for all event types. Similarly, let $\lambda_0=(\lambda_{01}(\cdot),\lambda_{02}(\cdot))^T$ denote the collection of all baseline hazard functions. If there is only one event type $K=1$, then the cause-specific frailty model (1) simply reduces to the standard univariate frailty model [11]. For the distribution of v_i we assume a normal distribution, i.e. $v_i\sim$ i.i.d. $N(0,\theta)$, which is useful for modelling various structures of frailties [13].

Note that the model (1) can be expressed as a joint model with a common frailty v_i as follows. The event times (T_{ij1}) from cause 1 would follow a cause-specific PH model

$$\lambda_{ij1}(t\,|\,v_i)=\lambda_{01}(t)\exp(x_{ij}^T\beta_1+v_i),$$

and similarly event times (T_{ij2}) from cause 2 would follow a model

$$\lambda_{ij2}(t\,|\,v_i)=\lambda_{02}(t)\exp(x_{ij}^T\beta_2+v_i).$$

In the traditional cause-specific analysis, patients who failed from cause 2 are treated as censored for the analysis of Type 1 events, which ignores a potential correlation between the two event times.

2.2 H-likelihood procedure

The observed event time and the event indicator are, respectively, defined by

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

where $T_{ij}=\min(T_{ij1},T_{ij2})$ is the time to the first event, $\delta_{ijk}=1$ if Type k event occurs first (i.e. $Y_{ij}=T_{ijk}$) and otherwise is 0. Note that δ_{ijk} is often referred to as a cause-specific event indicator and that it can also be expressed as

$$\delta_{ijk}=I(T_{ij}\le C_{ij})I(\varepsilon_{ij}=k).$$

Now, we outline the h-likelihood for the cause-specific frailty model (1). Following Ha et al. [11] and Christian et al. [4], the h-likelihood for the joint model above (1) is defined by

$$h=h(\beta,\lambda_0,v,\theta)=\sum_{ijk} l_{ijk}(\beta_k,\lambda_{0k};y_{ij},\delta_{ijk}\,|\,v_i)+\sum_i l_i(\theta;v_i),$$

where

$$\begin{aligned} l_{ijk}&=l_{ijk}(\beta_k,\lambda_{0k};y_{ij},\delta_{ijk}\,|\,v_i)\\ &=\delta_{ijk}(\log\lambda_{0k}(y_{ij})+\eta_{ijk})-\Lambda_{0k}(y_{ij})\exp(\eta_{ijk}) \end{aligned}$$

and

$$l_i=l_i(\theta;v_i)=-\log(2\pi\theta)/2-v_i^2/(2\theta).$$

Here $\lambda_0=(\lambda_{01},\lambda_{02})$ is a collection of all the baseline hazards, $v=(v_1,\ldots,v_q)^T$, $\Lambda_{0k}(\cdot)$ is the baseline cumulative hazard function for cause k, and $\eta_{ijk}=x_{ij}^T\beta_k+v_i$.

Note that the h-likelihood above h has unknown parameters λ_{0k} that increases with the number of events. Firstly, to eliminate the nuisance parameters λ_{0k} we use the corresponding profile h-likelihood h_p [11,13], defined by

$$\begin{aligned} h_p(\beta,v,\theta)=&\sum_{ijk}\delta_{ijk}(x_{ij}^T\beta_k+v_i)-\sum_{kr}d_{(kr)}\log\Big\{\sum_{ij\in R_{(kr)}}\exp(x_{ij}^T\beta_k+v_i)\Big\}\\ &+\sum_{i=1}^{q} l_{2i}(\theta;v_i), \end{aligned}$$

where $R_{(kr)}=\{ij:y_{ij}\ge y_{(kr)}\}$ is the risk set at time $y_{(kr)}$, i.e. the set of all individuals who are still at risk of experiencing an event. Here $y_{(k1)}<y_{(k2)}<\ldots<y_{(kD_k)}$ are the D_k are the ordered

unique event times for type k events among all of the y_{ij}'s, and d_{kr} is the number of events that occur at time $y_{(kr)}$. Then we use the partial h-likelihood (PHL) h_p to estimate the fixed and random effects (β,v), and the the partial restricted likelihood $p_{\beta,v}(h_p)$ [11,13] to estimate the frailty parameter θ, given by

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

where

$$H_p=-\partial^2 h_p/\partial(\beta,v)^2$$

is the observed information matrix from the PHL h_p.

3. Variable Selection Procedure

Following Ha et al. [7,8], we consider variable selection of fixed effects β in the joint competing-risks model (1) via maximization of a penalized profile h-likelihood, l_p, using h_p and a penalty; it is defined by

$$l_p(\beta,v,\alpha)=h_p-n\sum_{j=1}^{p^*}J_\gamma(|\beta_j|), \tag{2}$$

where $p^*=Kp$ with $K=2$, and $J_\gamma(|\cdot|)$ is a penalty function that controls model complexity using the tuning parameter γ. Note that we do not impose any penalty on the frailty parameter θ. Typically, the larger value of γ tends to choose the simple model, whereas the smaller value of γ inclines towards the complex model [8]. To choose an optimal value of γ, we use a BIC-type criterion [8]; the generalized cross-validation cannot select the tuning parameter satisfactorily [14]. For details of variable selection procedures using (2), refer to Ha et al. [8].

Various penalty functions have been used in the literature on variable selection in statistical models including the Cox PH model. Here, we consider the following two penalty functions (LASSO and SCAD), but our results can also be applied to other penalty functions (e.g. h-likelihood penalty [15]).

- LASSO [9]: $J_\gamma(|\beta|)=\gamma|\beta|$,
- SCAD [6]:

$$J'_\gamma(|\beta|)=\gamma I(|\beta|\le\gamma)+\frac{(a\gamma-|\beta|)_+}{a-1}I(|\beta|>\gamma),$$

where x_+ is x if $x>0$, 0 otherwise, and $a=3.7$.

3.1 Fitting procedure

In this subsection, we present the gradient vector of $\tau=(\beta,v)$ given θ and the observed information matrix from the partial h-likelihood h_p. In particular, we further derive the ILS equations and useful matrix forms for estimating (v,β,θ).

The matrix forms for estimating $\beta=(\beta_1^T,\beta_2^T)^T$ and v are as follows:

$$\frac{\partial l_p}{\partial\beta_j}=\frac{\partial h_p}{\partial\beta_j}-n\sum_{j=1}^{p^*}[J_\gamma(|\beta_j|)]'=0,\ (j=1,\ldots,p^*)$$

where $\partial h_p/\partial\beta_k=X^T(\delta_k-\mu_k)$ $(k=1,2)$, and

$$\frac{\partial l_p}{\partial v}=Z^T(\delta_1-\mu_1)+Z^T(\delta_2-\mu_2)-\theta^{-1}v.$$

Here Z is a $n\times q$ model matrix of v, δ_k is an $n\times 1$ Type k event indicator vector with ij^{th} element δ_{ijk}, and $\mu_k=\hat{\Lambda}_{0k}\exp(\eta_k)$.

Let $\boldsymbol{X}=\begin{pmatrix}X & \mathbf{0}\\ \mathbf{0} & X\end{pmatrix}$, $\boldsymbol{Z}=\begin{pmatrix}Z\\ Z\end{pmatrix}$ and $\boldsymbol{W}^*=\begin{pmatrix}W_1^* & \mathbf{0}\\ \mathbf{0} & W_2^*\end{pmatrix}$.

Here X is a $n\times p$ model matrix of β_k, and $W_k^*=-\partial^2 h_p/\partial\eta_k\partial\eta_k^T$ with $\eta_k=X\beta_k+Zv$. Along the lines of Ha et al. [8] and local quadratic approximation (LQA) [6], we can show that the iterative least squares (ILS) equations for (β,v) with $\beta=(\beta_1^T,\beta_2^T)^T$ are given by

$$\begin{pmatrix}\boldsymbol{X}^T\boldsymbol{W}^*\boldsymbol{X}+n\Sigma_\gamma & \boldsymbol{X}^T\boldsymbol{W}^*\boldsymbol{Z}\\ \boldsymbol{Z}^T\boldsymbol{W}^*\boldsymbol{X} & \boldsymbol{Z}^T\boldsymbol{W}^*\boldsymbol{Z}+Q\end{pmatrix}\begin{pmatrix}\hat{\beta}\\ \hat{v}\end{pmatrix}=\begin{pmatrix}\boldsymbol{X}^T\boldsymbol{w}^*\\ \boldsymbol{Z}^T\boldsymbol{w}^*\end{pmatrix},$$

where $\Sigma_\gamma=\mathrm{diag}\{J'_\gamma(|\beta_j|)/|\beta_j|\}$, $Q=-\partial^2 l_2/\partial v^2=\theta^{-1}I_q$ is a $q\times q$ identity matrix, and $\boldsymbol{w}^*=(w_1^{*T},w_2^{*T})^T$ with $w_k^*=W_k^*\,\eta_k+(\delta_k-\mu_k)$. For estimation of the frailty parameter α, we use an penalized partial restricted h-likelihood $p_{\beta,v}(l_p)$. Note that for the cause-specific hazard model without frailty v, the ILS equations reduce to

$$(\boldsymbol{X}^T\boldsymbol{W}^*\boldsymbol{X}+n\Sigma_\gamma)\hat{\beta}=\boldsymbol{X}^T\boldsymbol{w}^*.$$

Accordingly, we see that the variable-selection procedure mentioned above is easily implemented for competing-risks data via a slight modification to the existing PHL procedures [8,16].

3.2 Standard error and selection of tuning parameter

Following Ha et al. [7,8], an approximated standard error

(SE) of $\hat{\beta}$ is obtained from a sandwich formula based on l_p:

$$\mathrm{cov}(\hat{\beta}) = (H_{\beta\beta} + n\Sigma_\gamma)^{-1} H_{\beta\beta} (H_{\beta\beta} + n\Sigma_\gamma)^{-1}, \quad (3)$$

where $H_{\beta\beta} = \{(X^T W^* X) - (X^T W^* Z)(Z^T W^* Z + Q)^{-1}(Z^T W^* X)\}|_{v=\hat{v}}$. For the choice of tuning parameter γ, Ha et al. [8] used a BIC-type criterion based on the penalized PHL, defined by

$$\mathrm{BIC}(\gamma) = -2p_v(h_p) + e(\gamma)\log(n), \quad (4)$$

where $p_v(h_p)$ is the first-order Laplace approximation to the marginal partial likelihood [7,8] and $e(\gamma) = \mathrm{tr}[\{H_{\beta\beta} + n\Sigma_\gamma\}^{-1} H_{\beta\beta}]$ is the effective number of parameters [10,13].

Our variable-selection algorithm consists of two loops as follows.

1. Inner loop: we maximize l_p for $\tau = (\beta^T, v^T)^T$ (i.e., we solve the ILS equations above for (β, v) and $p_\tau(l_p)$ for θ, respectively.
2. Outer loop: we select γ that minimizes BIC(γ) in (4).

After convergence, we compute the estimates of the SEs for $\hat{\beta}$ using (3).

4. An Illustrative Example

We illustrate the proposed procedure using a multicenter clinical dataset [17] from a bladder cancer trial conducted by the European Organization for Research and Treatment of Cancer (EORTC). We consider 396 patients with stage Ta and T1 bladder cancer from 21 centers, where the numbers of patients per center varied from 3 to 78, with mean 18.9 and median 14. Two competing endpoints are time to first bladder recurrence (an event of interest; Type 1 event) and time to death prior to recurrence (competing event; Type 2 event). Of 396 patients, 200 (50.51%) had recurrence of bladder cancer and 81 (20.45%) died prior to recurrence. The number of patients who were still alive without recurrence were 115 (29.04%) and they censored at the date of the last available follow-up.

We consider the following 12 categorical covariates (x) of interest:

- Chemotherapy as the main covariate (CHEMO; no = 0, yes = 1),
- Age (0 if Age $\leq$ 65 years, 1 if Age $>$ 65 years),
- Sex (male = 0, female = 1),
- Prior recurrent rate (PRIORREC; primary, $\leq$ 1/yr, $>$ 1/yr);
  PRIORREC1 = I (PRIORREC $\leq$ 1/yr),
  PRIORREC2 = I (PRIORREC $>$ 1/yr)
- Number of tumors (NOTUM; single, 2$\sim$7 tumors, $\geq$ 8 tumors);
  NOTUM1 = I (NOTUM = 2$\sim$7 tumors),
  NOTUM2 = I (NOTUM $\geq$ 8 tumors);
- Tumor size (TUM3CM; 0 if Tumor size $<$ 3 cm, 1 if Tumor size $\geq$ 3 cm),
- T category (TLOCC; Ta = 0, T1 = 1),
- Carcinoma in situ (CIS; no = 0, yes = 1),
- G grade (GLOCAL; G1, G2, G3);
  GLOCAL1 = I (GLOCAL = G2),
  GLOCAL2 = I (GLOCAL = G3).

Ha et al. [7] also used the same data set as above for the variable selection using the subhazard frailty model. Now, we fitted the cause-specific hazard model (1) with a common frailty using the penalized h-likelihood procedure proposed in Section 3. The selected values of the tuning parameters γ by the BIC were 0.013 and 0.067 for LASSO and SCAD, respectively. The estimates of the frailty parameter θ for no-penalty, LASSO and SCAD are 0.066, 0.061 and 0.079, respectively. The estimated coefficients and their standard errors for Type 1 (i.e. bladder cancer recurrence) are given in Table 1. The main covariate, CHEMO (x_1), is very significant in all the three methods (i.e. no-penalty, LASSO and SCAD). LASSO chooses 9 covariates (x_1, x_2, x_5, x_6, x_7, x_8, x_9, x_{11}, x_{12}) among 12 covariates, while SCAD chooses 6 covariates (x_1, x_5, x_6, x_7, x_{11}, x_{12}), respectively. Notice here that LASSO chooses the two covariates (x_8 and x_9) which are not significant under no-penalty. In particular, we find that the trends of the variable selection for Type 1 are similar to those in subhazard frailty model by Ha et al. [7]. For Type 2 (i.e. death prior to recurrence), the estimated coefficients and their standard errors are also presented in Table 1. Here, LASSO selects 2 covariates (x_2 and x_{11}) where x_{11} is not significant under no-penalty, while SCAD chooses only one covariate (x_2).

We observe that the non-zero estimates by SCAD are overall similar to the corresponding estimates without penalty ($\gamma = 0$). As expected by Ha et al. [7,8], LASSO selects more covariates as compared to SCAD. A possible reason may be that LASSO selects unimportant variables much more than SCAD method. These findings indicate that LASSO may not properly identify important variables in the cause-specific hazard frailty model (1), that confirms the simulation results of Ha et al. [7,8].

Table 1. Estimated coefficients and standard errors (in parentheses) in the cause-specific hazard frailty model for two types of events in the multi-center bladder cancer data

Event	Variable	No-penalty	LASSO	SCAD
Type 1	x_1: CHEMO	−0.862 (0.186)	−0.598 (0.142)	−0.865 (0.181)
	x_2: Age	−0.271 (0.146)	−0.131 (0.079)	0 (0)
	x_3: Sex	−0.011 (0.209)	0 (0)	0 (0)
	x_4: PRIORREC1	0.269 (0.250)	0 (0)	0 (0)
	x_5: PRIORREC2	0.518 (0.200)	0.337 (0.119)	0.423 (0.177)
	x_6: NOTUM1	0.669 (0.166)	0.455 (0.118)	0.664 (0.163)
	x_7: NOTUM2	1.168 (0.280)	0.693 (0.171)	1.196 (0.267)
	x_8: TUM3CM	0.155 (0.174)	0.002 (0.002)	0 (0)
	x_9: TLOCC	0.239 (0.170)	0.183 (0.091)	0 (0)
	x_{10}: CIS	0.222 (0.275)	0 (0)	0 (0)
	x_{11}: GLOCAL1	0.517 (0.165)	0.269 (0.103)	0.540 (0.158)
	x_{12}: GLOCAL2	0.762 (0.271)	0.276 (0.111)	0.906 (0.248)
Type 2	x_1: CHEMO	0.346 (0.393)	0 (0)	0 (0)
	x_2: Age	0.906 (0.288)	0.357 (0.131)	0.712 (0.264)
	x_3: Sex	−0.508 (0.357)	0 (0)	0 (0)
	x_4: PRIORREC1	0.152 (0.401)	0 (0)	0 (0)
	x_5: PRIORREC2	0.463 (0.327)	0 (0)	0 (0)
	x_6: NOTUM1	−0.503 (0.261)	0 (0)	0 (0)
	x_7: NOTUM2	−1.335 (0.560)	0 (0)	0 (0)
	x_8: TUM3CM	−0.154 (0.275)	0 (0)	0 (0)
	x_9: TLOCC	−0.095 (0.278)	0 (0)	0 (0)
	x_{10}: CIS	0.481 (0.502)	0 (0)	0 (0)
	x_{11}: GLOCAL1	0.268 (0.264)	0.010 (0.008)	0 (0)
	x_{12}: GLOCAL2	−0.176 (0.480)	0 (0)	0 (0)

5. Discussion

The h-likelihood methods are useful for variable selection in the competing-risks frailty models. In Section 4, we have illustrated how to select jointly important variables of fixed effects in the cause-specific hazard frailty models with two types of events via a penalized h-likelihood procedure. Ha et al. [7,8] have demonstrated via simulation studies and data analysis that the penalized h-likelihood procedure with SCAD penalty performs well, that confirm the results of Table 1. In Section 3, we have shown that the proposed variable selection method can be easily implemented by a slight modification of the standard h-likelihood estimation procedures for competing risks frailty models [4], leading to a fast and efficient procedure. Thus the h-likelihood method would be easily applied to variable selection in various random-effect models such as hierarchical generalized linear models [10,12].

In this paper, the proposed variable-selection method was applied to only a real-data analysis. Thus, the simulation study would be a necessity for future work. We proposed a variable selection for the cause-specific frailty model with a univariate frailty. Development of variable selection to cause-specific models with correlated frailties [4] or high dimensional case having $p > n$ [18] would also be further work of interest.

In this paper we also considered selecting individual variables only. In many regression problems, covariates often possess a natural group structure; for example, categorical variables are often represented by a group of indicator variables. In these situations, the problem of selecting relevant variables is that of selecting groups rather than selecting individual variables [19,20]. Extension of the group penalization methods such as group LASSO [19] to the competing risks models would also be an interesting topic.

Acknowledgements

This research was supported by a Research Grant of Pukyong National University (2016 year).

References

1. Prentice R, Kalbfleisch JD, Peterson AV, Flournoy N, Farewell VT, Breslow NE. The analysis of failure times in the presence

of competing risks. Biometrics 1978;34:541-554.
2. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc 1999;94:496-509.
3. Ha ID, Christian NJ, Jeong JH, Park J, Lee Y. Analysis of clustered competing risks data using subdistribution hazards models with multivariate frailties. Stat Methods Med Res 2016;25:2488-2505.
4. Christian NJ, Ha ID, Jeong, JH. Hierarchical likelihood inference on clustered competing risks data. Stat Med 2016;35:251-267.
5. Breiman L. Heuristics of instability and stabilization in model selection. Ann Stat 1996;24:2350-2383.
6. Fan J, Li R. Variable selection via nonconcave penalized likelihood and its oracle properties. J Am Stat Assoc 2001:96:1348-1360.
7. Ha ID, Lee M, Oh S, Jeong JH, Sylvester R, Lee Y. Variable selection in subdistribution hazard frailty models with competing risks data. Stat Med 2014;33:4590-4604.
8. Ha ID, Pan J, Oh S, Lee Y. Variable selection in general frailty models using penalized h-likelihood. J Comput Graph Stat 2014;23:1044-1060.
9. Tibshirani R. Regression shrinkage and selection via the Lasso. J Roy Stat Soc B 1996;58:267-288.
10. Lee Y, Nelder JA. Hierarchical generalized linear models (with discussion). J Roy Stat Soc B 1996;58:619-678.
11. Ha ID, Lee Y, Song JK. Hierarchical likelihood approach for frailty models. Biometrika 2001;88:233-243.
12. Lee Y, Nelder JA, Pawitan Y. Generalised linear models with random effects: unified analysis via h-likelihood. London: Chapman and Hall; 2006.
13. Ha ID, Lee Y, MacKenzie G. Model selection for multi-component frailty models. Stat Med 2007;26:4790-4807.
14. Wang H, Li R, Tsai CL. Tuning parameter selectors for the smoothly clipped absolute deviation method. Biometrika 2007;94:553-568.
15. Lee Y, Oh HS. A new sparse variable selection via random-effect model. J Multivariate Anal 2014;125:89-99.
16. Ha ID, Lee Y. Estimating frailty models via Poisson hierarchical generalized linear models. J Comput Graph Stat 2003;12:663-681.
17. Sylvester R, van der Meijden APM, Oosterlinck W, Witjes J, Bouffioux C, Denis L, et al. Predicting recurrence and progression in individual patients with stage Ta T1 bladder cancer using EORTC risk tables: a combined analysis of 2596 patients from seven EORTC trials. Eur Urol 2006;49:466-477.
18. Fan J, Lv J. A selective overview of variable selection in high dimensional feature space. Stat Sinica 2010;20:101-148.
19. Yuan M, Lin Y. Model selection and estimation in regression with grouped variables. J Roy Stat Soc B 2006;68:49-67.
20. Huang J, Breheny P, Ma SG. A selective review of group selection in high-dimensional models. Stat Sci 2012;27:48199.