World Applied Sciences Journal 7 (11): 1382-1387, 2009
ISSN 1818-4952


# Comparison of Some Smoothing Parameter Selection Methods in Generalized Estimating Equation-Smoothing Spline

*[1]Suliadi, [2]Noor Akma Ibrahim, [3]Isthrinayagy S. Krishnarajah and [4]Isa Daud*

[123]Institute for Mathematical Research, Universiti Putera Malaysia, Malaysia
[1]Department of Statistics, Bandung Islamic University, Indonesia
[234]Department of Mathematics, Faculty of Science, Universiti Putra Malaysia, Malaysia

**Abstract:** This paper considers performance of some smoothing parameter selection methods in Generalized Estimating Equation-Smoothing Spline for nonparametric regression with binary data. We evaluated eight methods, GCV given by Green and Silverman, GCV and AIC given by Ruppert *et al*, ACV and GACV given by Xiang and Wahba, AIC given by Chiou and Tsai, SCVD given by WU and Zhang and the last method is AIC*, modification of AIC given by Chiou and Tsai. Using simulation we found that for nonlinear systematic component (sinusoidal) AIC and AIC* of Chiou and Tsai are the best methods and the worst method is GCV of Green and Silverman. For linear systematic component, GCV of Green and Silverman is the best, while AIC and AIC* are the worst. Since in practical situation we do not know the form of the systematic component, hence we suggest the use of ACV and GACV of Xiang and Wahba or AIC of Ruppert *et al*, which give moderate results.

**Key words:** Smoothing parameter · Smoothing spline · GEE · Nonparametric Regression · Longitudinal data · Correlated data

## INTRODUCTION

Recently, nonparametric regression has become a popular method to analyze relationship between dependent and independent variables. In nonparametric regression, one relaxes some of the assumptions of the parametric regression. One of the nonparametric regression methods is smoothing spline. This method is popular since the estimation can be made using simple linear algebra. The roughness penalty can be stated as multiplication of matrix and vector [3].

Performance of the smoothing spline estimator depends on the smoothing parameter ($\lambda$). It represents the rate of exchange (*trade-off*) between goodness-of-fit of data and the smoothness of the curve $f$. If $\lambda$ is too small then the curve will be too rough but the variance will be low and vice versa. Thus this value is very important in smoothing spline estimation. For continuous data, there exist methods for smoothing parameter selection. Some examples are: Cross Validation (CV), Generalized Cross Validation (GCV), Akaike Information Criteria (AIC), Mallow's Cp, Leave-one-subject-out Cross Validated Deviance (SCVD), Leave-one-point out Cross Validation (PCV). See [4,7,9,16 and 18].

For non-Gaussian data, selection of smoothing parameter is more complicated since the relation between responses and covariates are not linear. Some authors have proposed smoothing parameter selection method for model in the class of generalized linear model for independent data. GCV and AIC using deviance have been proposed [16], other methods that have also been proposed are approximate of cross validation and generalized approximate cross validation (ACV and GACV) [19]. Motivated from Gaussian data, by applying working responses instead of the real data, modified cross validation has also been proposed [4]. All these methods are developed for independent non-Gaussian data. The question is which method is the best to be applied for correlated data. Correlated data are common in

**Corresponding Author:** Suliadi, Institute for Mathematical Research, Universiti Putra Malaysia, UPM Serdang, Selangor Darul Ehsan, 43400, Malaysia, Tel: +60173985509, E-mail: suliadi@gmail.com

many areas of study, such as medical and clinical trial, biology, and economic. The source of correlation comes from repeated observations from the same subject, also called as longitudinal study. Several examples of study that result correlated data given below. Manjunath & Telles [12] studied the effect of a cold chest pack for 30 minutes daily over period time to the bronchial asthmatics patients, where the response measured at day and day 21. Arora *et al.* [1] studied the characteristics of diabetes mellitus induced by different doses of streptozotocin in mice. The response is the blood glucose concentration and this response was measured once a week from week 0 until week 5 for each mice. Other examples of longitudinal data are conducted by Hashemi *et al.* [6] and Oduola *et al.* [15]. Statistically, the important point of their studies is each subject is measured repeatedly and this results correlated data.

Generalized estimating equation (GEE) is a parametric method usually used to analyze longitudinal (correlated) data, based on marginal model. GEE was initially proposed by Liang and Zeger [11]. Suliadi, *et al.* [17] extended the parametric GEE into nonparametric GEE using smoothing spline, called GEE-Smoothing Spline. Smoothing parameter selection for GEE-Smoothing Spline is rather complicated, since GEE do not assume the distribution of the response. As known, GEE is the extension of the quasi-likelihood for correlated data [3,11] by introducing working correlation into the estimation. Thus it does not have real deviance as in GLM. Chiou and Tsai [1] modified the improved AIC given by Naik and Tsai [11] and applied it to quasi-likelihood models for Local Polynomial Kernel (LPK)-GEE. Another method that has been proposed is SCVD. This method was proposed by Wu and Zhang [18] (p. 326) and has been applied to the LPK-GEE.

The purpose of this paper is to evaluate the performance of smoothing parameter selection methods in the GEE-Smoothing Spline of Suliadi *et al.* [17]. We evaluate the performances by simulation.

## Smoothing Parameter Selection

**For Non-Gaussian Data:** Ruppert *et al.* [16] (p.220) proposed to use GCV and AIC in selection of smoothing parameter for non-Gaussian data. Let $A$ be the smoother matrix (“*hat matrix*”), $D(y,\hat{y})$ be the deviance of the model and $n$ be the sample size. Then GCV based on deviance is defined by

$$GCV_R(\lambda) = \frac{n^{-1}D(y,\hat{y})}{[1-n^{-1}tr(A)]^2} \qquad (1)$$

Another selection method is AIC criterion, which is defined as

$$AIC_R(\lambda) = n^{-1}[D(y,\hat{y}) + 2tr(A)\phi] \qquad (2)$$

where $\phi$ is the scale parameter. The optimal $\lambda$ is obtained by minimizing GCV($\lambda$) or AIC($\lambda$).

Xiang and Wahba [15] also gave the approximation of CV and GCV. Suppose the responses *y's* are from exponential family with canonical parameter $\theta = \eta(x)$. Thus $E(y_i) = b'(\eta(x_i))$ and $var(y_i) = b''(\eta(x_i)).a(\phi)$. Let $W = Diag(b''(\eta(x_1), \ldots, b''\eta(x_n)) = Diag(w_1, \ldots, w_n)$ and $H$ is the inverse of Hessian matrix with $h_{ii}$ is the *i-th* diagonal element of $H$. The approximate of cross validation and generalized approximate cross validation (ACV and GACV) are defined as

$$ACV_{XW}(\lambda) = \frac{1}{n}\sum_{i=1}^{n}[-y\eta(x_i) + b(\eta(x_i))] + \frac{1}{n}\sum_{i=1}^{n}\left[\frac{h_{ii}y_i[y_i - \mu(x_i)]}{1-h_{ii}b''(\eta(x_i))}\right] \qquad (3)$$

and

$$GACV_{XW}(\lambda) = \frac{1}{n}\sum_{i=1}^{n}[-y\eta(x_i) + b(\eta(x_i))] + \left[\frac{tr(H)}{n}\right]\frac{\sum_{i=1}^{n} y_i[y_i - \mu(x_i)]}{n - tr(W^{1/2}HW^{1/2})} \qquad (4)$$

The forms of (3) and (4) were designed with the assumption that covariates have different values (different design time points). Little modification is needed if there are some points of the covariates that have the same design time points. One selects $\lambda$ by minimizing ACV($\lambda$) or GACV($\lambda$).

Green and Silverman [3] proposed the approximated Cross Validation that is applied to the working response instead of response data. Suliadi *et al.* [17] applied this method to the GEE-Smoothing Spline for correlated binary data and found that this method had poor performance. It produces over smooth curve. The approximate generalized cross validation is motivated from GCV for Gaussian data, defined as

$$GCV_{GS}(\lambda) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{z_i - x_i^T\hat{f}}{(1-n^{-1}tr(A))^2}\right)^2 \qquad (5)$$

where the working response $z_i = (y - \hat{\mu})g'(\mu_i) + x_i^T\hat{f}$ and $A$ is the “hat matrix”.

Chiou and Tsai [1] proposed smoothing parameter selection using Akaike Information Criteria (AIC) by extending the improved version of AIC given by Naik and Tsai [11] and applied it to the quasi-likelihood nonparametric models. The extension is based on the extended quasi-likelihood. The method can be described as follows. Let $H_{np}$ be a smoother matrix such that $\eta(X\beta)=H_{np}Y$, $H_p=M(M^TVM)^{-1}M^T$ where $M=\partial\eta/\partial\beta$, $H=H^*_{np}+H^*_p-H^*_{np}H^*_p$ with $H^*_{(\bullet)}=V^{1/2}H_{(\bullet)}V^{1/2}$ and $V=b''(\hat{\theta})=V(\hat{\mu})$. The improved AIC based on quasi-likelihood models is defined by

$$AIC_{CT}(\lambda)=\log(\hat{\sigma}^2)+\frac{1+tr(H)/n}{1+[tr(H)+2]/n} \tag{6}$$

where $\hat{\sigma}^2=(y-\hat{\mu})^TV^{-1}(y-\hat{\mu})/n$.

Smoothing parameter selection methods which are discussed above are proposed for individual observation. Wu and Zhang [18] (p.326) proposed smoothing parameter selection method for discrete longitudinal data and applied it to the local polynomial kernel (LPK)-GEE. This method is called *leave-one-subject-out cross validated deviance* (SCVD) and defined by

$$SCVD_{WZ}(\lambda)=\sum_{i=1}^{n}\sum_{j=1}^{n_i}d(y_{ij},\hat{\mu}_{ij}^{(-i)})/nn_i \tag{7}$$

The $\hat{\mu}_{ij}^{(-i)}$ is the estimate value for the *i-th* subject and the *j-th* time observation when the estimation of $f$ is obtained without the *i-th* observation, $f^{(-i)}$. Let $H$ and $U$ be the Hessian matrix and the estimating equation matrix from the final iteration respectively. Given a smoothing parameter $\lambda$, the approximate of $f^{(-i)}$ is given by

$$\hat{f}^{(-i)}\approx\left\{H^{(-i)}\right\}^{-1}U^{(-i)},$$

where $H^{(-i)}$ and $U^{(-i)}$ are obtained from $H$ and $U$ by removing the *i-th* observation. Computation of (7) is time consuming since one must compute $\hat{f}^{(-i)}$ as many as the number of subjects there are.

**Smoothing Spline for Longitudinal Categorical Data:** Suliadi *et al.* [17] proposed GEE-Smoothing Spline to analyze longitudinal binary data nonparametrically by combining natural cubic spline [3] and generalized estimating equation given by Liang and Zeger [9].

Suppose there are $n$ subjects and for the *i-th* subject, $n_i$ measurement is taken. Let realization of the outcome and time measurement for the *i-th* subject and the *j-th* time measurement be $y_{ij}$ and $t_{ij}$, respectively. Observations from the same subject are correlated following a specific form, but observations from different subject are independent. Consider the nonparametric population average model. In this model, the systematic component relates to the mean of response in the form of

$$g(\mu_{ij})=\eta_{ij}=f(t_{ij}),\quad i=1,2,\dots,n;\ j=1,2,\dots,n_i.$$

where $f$ is an unknown smooth function and $g(\mu)$ is the link function. In this paper we use the canonical link function.

Let $t_{(1)}, t_{(2)}, \dots, t_{(q)}$ be the different time points for all $t_i$'s such that $t_{(1)} < t_{(2)} < \dots < t_{(q)}$ and let $X_i$ be the $n_i\times q$ incidence matrix for all $t_i$'s. The relation $t_{()}$'s to $x_{ijk}$ are $x_{ijk}=1$, if $t_{ij}=t_{(k)}$ and $x_{ijk}=0$ otherwise, for $k=1,\dots,q$. Let $x_{ij}=(x_{ij1},\dots,x_{ijq})^T$ and vector of the unknown function at different design point is $f=[f(t_{(1)}), f(t_{(2)}),\dots,f(t_{(q)})]^T$. Thus $f(t_{ij})=x_{ij}^Tf$. Set $X_i=(x_{i1},\dots,x_{in_i})^T$, $y_i=(y_{i1},\dots,y_{in_i})^T$, $\eta_i=(\eta_{i1},\dots,\eta_{in_i})^T=X_if$. Following Green and Silverman [3], the roughness penalty can be stated as $\int[f''(t)]^2dt=f^TKf$. The matrix $K$ is obtained as follows (see [4] p.12).

Let $h_i=t_{(i+1)}-t_{(i)}$, for $i=1,2,\dots,q-1$. Let $G$ be the $q$ x $(q-2)$ matrix with elements $g_{ij}$, $i=1,\dots,q$ and $j=2,\dots,q-1$, given by

$$g_{j-1,j}=h_{j-1}^{-1};\ g_{j,j}=h_{j-1}^{-1}-h_j^{-1};\ g_{j+1,j}=h_j^{-1}.$$

The $P$ matrix is defined as follows. The symmetric matrix $P$ is $(q-2)$ x $(q-2)$ with elements $p_{ij}$, for $i$ and $j$ running from $2$ to $(q-1)$, given by

$$p_{ii}=(h_{i-1}+h_i)/3,\quad \text{for i = 2, \&, q-1}$$
$$p_{i,i+1}=p_{i+1,i}=h_i/6,\quad \text{for i = 2, \&, q-2}$$

Matrix $P$ and $G$ are numbered in non standard way. The matrix $P$ is strictly diagonal dominant, in which $|p_{ii}|>\sum_{i\neq j}|p_{ij}|$. Thus $P$ is strictly positive-definite, hence $P^{-1}$ exist. Defined a matrix $K$ by

$$K=GP^{-1}G^T$$

The generalized estimating equation-smoothing spline is defined as

$$\begin{aligned}U(f)&=\sum_{i=1}^{n}D_i^TV_i^{-1}S_i-\frac{\partial}{\partial f}\left[\frac{1}{2}\lambda\int[f''(t)]^2dt\right]\\&=\sum_{i=1}^{n}D_i^TV_i^{-1}S_i-\lambda Kf\end{aligned} \tag{8}$$

where $S_i(f)$, $y_i = \mu_i(f)$ $D_i(f) = \partial(\mu_i(f))/\partial f = W_i(f)\Delta_i X_i = W_i(f) X_i$ (where $\Delta_i = I_{ni}\,\partial\theta/\partial\eta$ since for the canonical link function $\theta = \eta$), $W_i = diag\{\mu_{ij}(1-\mu_{ij})\}$ and $V_i = W_i^{1/2} R(\hat{\alpha}) W_i^{1/2}$. $R(\hat{\alpha})$ is working covariance matrix and $\alpha$ is the association parameter. The iterative procedure using modified Fisher scoring for $f$ for a given current estimate of $\hat{\alpha}$ is

$$\hat{f}_{s+1} = \hat{f}_s + \left[\sum_{i=1}^{n} X_i^T W_i V_i^{-1} W_i X_i + \lambda K\right]^{-1} \left[\sum_{i=1}^{n} X_i^T W_i V_i^{-1} S_i - \lambda K \hat{f}_s\right] \tag{9}$$

where $W$, $V$ and $S$ are evaluated at $\hat{f}_s$. Expression (9) can be written as an iteratively re-weighted least square (IRWLS). Using the IRWLS form, we can obtain the "hat" or smoother matrix. Defined $X = (X_1^T,\ldots,X_n^T)^T$, $y = (y_1^T,\ldots,y_n^T)^T$, $W = diag(W_1,\ldots,W_n)$ $V = diag(V_1,\ldots,V_n)$, and $z = Xf + W^{-1}(y-\mu)$, then (9) can be written as

$$\hat{f}_{s+1} = (X^T W V^{-1} W X + \lambda K)^{-1} X^T W V^{-1} W z \tag{10}$$

The smoothing parameter selection methods (1) – (7), need adjustment before they can be applied to (10).

The first adjustment is the deviance part in (1), (2) and (7). It is known that there is no deviance in GEE. Since GEE is based on quasi-likelihood thus the deviance part in (1), (2) and (7) is based on quasi-likelihood (see: [5] Ch. 4; [13] Ch. 9).

Another adjustment is the hat or smoother matrix. The hat matrix for a linear model is defined as matrix $A$, such that $\hat{y} = Ay$. This definition is no longer applicable to the class of all exponential family distributions, except for normal distribution. Following Green and Silverman [4] (Ch. 5) and Ruppert *et al.* [16] (p.212), the hat or smoother matrix for (10) is $A = X(X^T W V^{-1} W X + \lambda K)^{-1} X^T W V^{-1} W$.

Computation for $\hat{\mu}^{(-i)}$ in (7) is based on $\hat{f}^{(-i)}$ where $\hat{f}^{(-i)}$ is the estimate of $f$ without the *i-th* subject. Let $W$ and $V$ be obtained from the final estimate of $f$. Then $\tilde{W}^{(-i)} = W^{(-i)}(\hat{f})$ and $\tilde{V}^{(-i)} = V^{(-i)}(\hat{f})$ are $W$ and $V$ respectively, after the deletion of the *i-th* subject. Matrices $X^{(-i)}$ and $z^{(-i)}$ are obtained using the same way. Then,

$$\hat{f}^{(-i)} \approx (X^{(-i)T}\tilde{W}^{(-i)}\tilde{V}^{(-i)-1}\tilde{W}^{(-i)}X^{(-i)} + \lambda K)^{-1} X^{(-i)T}\tilde{W}^{(-i)}\tilde{V}^{(-i)-1}\tilde{W}^{(-i)} z^{(-i)}$$

**Simulation Study:** We run simulations to evaluate the performance of these smoothing parameter selection methods, by using correlated binary data. The data were generated using R Language 2.5.0 [8]. The simulation was done in two conditions of systematic component, the sinus curve and linear curvature. We evaluate the eight methods, $GCV_{GS}$ given by Green and Silverman (5), $GCV_R$ and $AIC_R$ given by Ruppert *et al*, (1) and (2) respectively, $ACV_{XW}$ and $GACV_{XW}$ given by Xiang and Wahba, (3) and (4) respectively, $AIC_{CT}$ given by Chiou and Tsai (6), $SCVD_{WZ}$ given by WU and Zhang (7). The last method is $AIC^*_{CT}$, the AIC given by Chiou and Tsai (6) by replacing $V$ in $\hat{\sigma}^2$ with $V^{1/2} R V^{1/2}$. Here we introduced the within subject correlation into $\hat{\sigma}^2$.

The sinus curve is formed with systematic component $\eta(t) = \sin(\pi\, t/90)$, $t = 9, 18, \ldots, 180$. Number of (independent) subject is 10 with the structure of the correlation taken as exchangeable with $\alpha = \rho_{ij} = 0.1$, $i \neq j$. The second data structure has systematic component $\eta(t) = -1.1818+(2/99)t$, $t = 9, 18, \ldots, 108$. The number of subject is 10, with exchangeable correlation structure with $\alpha = r_{ij} = 0.35$, $i \neq j$. These two models were replicated 100 times. To assess the performance of these methods, we considered the square root of the average deviation of the mean (SRAD) and the average of absolute deviation of mean (AAD), both are defined as follows:

$$SRAD = \left[\sum_{i=1}^{n}\sum_{j=1}^{n_i}[\hat{\mu}(x_{ij}) - \mu(x_{ij})]^2 / (\sum_i n_i)\right]^{1/2}$$

$$AAD = \sum_{i=1}^{n}\sum_{j=1}^{n_i}\left|\hat{\mu}(x_{ij}) - \mu(x_{ij})\right| / (\sum_i n_i)$$

Method with lower SRAD or AAD is better than method with higher SRAD or AAD.

## RESULTS AND DISCUSSION

Table 1 shows the statistical descriptive of the SRAD and AAD for model 1. Result of SRAD and AAD are similar. These results show that for nonlinear (sinus) curve, AIC of Chiou and Tsai [1] and its modified are the best methods. It gives lowest mean and median of SRAD and AAD. Whilst GCV gives the worst method since it gives the highest mean and median of SRAD and AAD. Others have almost the same performance, which can be seen from their means and medians of SRAD and AAD.

Table 1: Statistical Descriptive of Model 1

| | SRAD | | AAD | |
|---|---|---|---|---|
| Method | Mean | Median | Mean | Median |
| $GCV_{GS}$ | 0.1085 | 0.1069 | 0.0914 | 0.0907 |
| $GCV_R$ | 0.0909 | 0.0916 | 0.0762 | 0.0780 |
| $AIC_R$ | 0.0870 | 0.0850 | 0.0722 | 0.0690 |
| $ACV_{XW}$ | 0.0861 | 0.0823 | 0.0712 | 0.0667 |
| $GACV_{XW}$ | 0.0854 | 0.0829 | 0.0708 | 0.0680 |
| $AIC_{CT}$ | 0.0817 | 0.0776 | 0.0687 | 0.0628 |
| $SCVD_{WZ}$ | 0.0879 | 0.0860 | 0.0729 | 0.0670 |
| $AIC^*_{CT}$ | 0.0816 | 0.0776 | 0.0686 | 0.0628 |

Table 2: Statistical Descriptive of Model 2

| | SRAD | | AAD | |
|---|---|---|---|---|
| Method | Mean | Median | Mean | Median |
| $GCV_{GS}$ | 0.1194 | 0.1071 | 0.0797 | 0.0720 |
| $GCV_R$ | 0.1207 | 0.1071 | 0.0818 | 0.0741 |
| $AIC_R$ | 0.1207 | 0.1071 | 0.0818 | 0.0736 |
| $ACV_{XW}$ | 0.1204 | 0.1070 | 0.0820 | 0.0734 |
| $GACV_{XW}$ | 0.1203 | 0.1070 | 0.0816 | 0.0728 |
| $AIC_{CT}$ | 0.1227 | 0.1081 | 0.0850 | 0.0751 |
| $SCVD_{WZ}$ | 0.1250 | 0.1113 | 0.0882 | 0.0805 |
| $AIC^*_{CT}$ | 0.1325 | 0.1247 | 0.1173 | 0.1079 |

The second model is model with linear systematic component. From Table 2, $GCV_{GS}$ gives minimum value of mean, but for the value of median, the performance of this method is no different with $GCV_R$, $AIC_R$, $ACV_{XW}$ and $GACV_{XW}$. Median values of these five methods are almost the same. Whilst, the last three methods, $AIC_{CT}$, $SCVD_{WZ}$ and $AIC^*_{CT}$, are the worst. They give highest value of the mean and median of SRAD and AAD.

## CONCLUSION

According to the results of Model 1 and Model 2, the methods seemed to have opposing performances. A method might have a good performance in nonlinear curve, but bad performance for linear case.

Since in practice we actually do not know the form of curvature (function), we recommend using ACV or GCV of Xiang and Wahba [15] or AIC of Ruppert *et al.* [12] for GEE-Smoothing Spline. These methods are not the best for nonlinear curve, but they do not give the worst performance. They have moderate performances.

## ACKNOWLEGEMENT

This paper is supported by Science Fund grant Vote 5450434 from Ministry of Science, Technology and Innovation Malaysia.

## REFERENCES

1. Arora, S., S.K. Ojha and D. Vohora, 2009. Characterization of Streptozotocin Induced Diabetes Mellitus in Swiss Albino Mice. Global Journal of Pharmacology, 3: 81-84.
2. Chiou, Jeng-Min and Chih-Ling Tsai, 2006. Smoothing Parameter Selection in Quasi-Likelihood Models. Journal of Nonpara-metric Statistics, 18: 307-314.
3. Davis, Charles, S., 2002. Statistical Methods for the Analysis of Repeated Measurements. Springer-Verlag. New York. USA.
4. Green, P.J. and B.W. Silverman, 1994. Nonparametric Regression and Generalized Linear Models. A Roughness Penalty Approach. Chapman and Hall/CRC. Washing-ton, USA.
5. Hardin, James W. and Joseph M. Hilbe, 2003. Generalized Estimating Equation. Chapman and Hall/CRC. Washington. USA.
6. Hashemi, M., Y. Dotsar, S.R. Rohani, A.R.A. Saraji, and M. Bayat, 2009. Influence of Aloxanes on the Apoptosis of Pancreas B-Cells of Rat. World Journal of Medical Sciences, 4: 70-73.
7. Hoover, D.R., John A. Rice, Colin O. Wu, Li-Ping Yang, 1998. Nonparametric Smoothing Estimates of Time-Varying Coefficient Models with Longitudinal Data. Biometrika, 85: 809-822.
8. Hurvich, C.M., J.S. Simonoff and C.L. Tsai, 1998. Smoothing Parameter Selection in Nonparametric Regression Using An Improved Akaike Information Criterion. Journal of the Royal Statistical Society. Series B, 60: 271-293.
9. Lee, Thomas, C.M, 2003. Smoothing Parameter Selection for Smoothing Spline: A Simulation Study. Computational Statistics and Data Analysis, 42: 139-148.
10. Leisch, Friedrich andreas Weingessel and Kurt Hornik, 1998. On the Generation of Correlated Artificial Binary Data. Working Paper. No 13. SFB. Adaptive Information System and Modeling in Economics and Management Science. Vienna University of Economics and Business Administration. Austria.
11. Liang, K.Y. and S.L. Zeger, 1986. Longitudinal Data Analysis using Generalized Linear Models. Biometrika, 73: 13-22.
12. Manjunath, N.K. and S. Telles, 2006. Therapeutic Application of a Cold Chest Pack in Bronchial Asthma. World Journal of Medical Sciences, 1: 18-20.
13. McCullagh, P. and J.A. Nelder, 1989. Generalized Linear Models. 2nd Edition. Chapman and Hall. London. UK.
14. Naik, Prasad A. and Chih-Ling Tsai, 2001. Single-Index Model Selections. Biometrika, 88: 821-832.
15. Oduola, T., A.A. Akinbolade, L.O. Oladokun, O.G. Adeosun, I.S. Bello and T.I. Ipadeola, 2009. Lipid Profiles in People Living with HIV/AIDS on ARV Therapy in an Urban Area of Osun State, Nigeria. World Journal of Medical Sciences, 4: 18-21.

16. Ruppert, David, M.P. Wand and R.J. Carrol, 2003. Semiparametric Regression. Cambridge University Press, Cambridge, UK.
17. Suliadi, Noor Akma Ibrahim, Isthrinayagy S. Krishnarajah and Isa Daud, 2007. Analyzing Longitudinal Binary Data using Cubic Spline. Paper was presented in The Islamic Countries Conference on Statistical Sciences (ICCS-IX).
18. Wu, Hulin and Jinting Zhang, 2006. Nonparametric Regression Methods for Longitudinal Data Analysis: Mixed-Effects Modeling Approaches, John Wiley and Sons, New York.
19. Xiang, Dong and Grace Wahba, 1996. A Generalized Approximate Cross Validation For Smoothing Splines With Non-Gaussian Data. Statistica Sinica, 6: 675-692.