Proceedings of the 8th WSEAS International Conference on APPLIED MATHEMATICS, Tenerife, Spain, December 16-18, 2005 (pp124-138)

Influence Diagnostics on Competing Risks Using Cox's Model with Censored Data

F. A. M. Elfaki[1], Bin Daud. I[2], N. A. Ibrahim[2], M. Y. Abdullah[2] and M. Usman[1]

[1]Department of Science in Engineering, Kulliyyah of Engineering, IIUM
Jalan Gombak, 53100, Kuala Lumpur, Malaysia.

[2]Department of Mathematics, Faculty of Science and Environmental Studies
University Putra Malaysia, 43400, Serdang, Selangor.

Abstract

This paper studies influence diagnostics (Cook's Distance and Likelihood Distance) on competing risk models when true covariates are observable. The one-step EM and one-step ML methods are used to detect influential observations for Cox's model. It was found that results generated from data analysis using one-step EM algorithm are better than results obtained by using one-step ML method. Moreover, Cook's distance show better results compare to likelihood distance.

Key word: Censored data; influence measurement; one-step methods; Cox's model; Competing risks.

1. Introduction

Diagnostics are used to assess the adequacy of assumptions underlying the modeling process and to identify unexpected characteristics of the data that may seriously influence conclusions or require special attention. A variety of graphical and nongraphical methods are available to aid one in linear regression analyses (Cook and Weisberg 1982) but most of these methods require the a priori specification of a model. Outliers and influential observations, for example, are always judged relative to some model, either implicit or explicit.

The detection of influential observations, that is observations whose deletion, either singly or multiply, result in substantial changes in parameter estimates, fitted values or tests of hypothesis, has received considerable attention in recent years. Several methods have been proposed for studying the impact of deletion of observations on parameter estimates obtained from the linear model (Belsley, Kuh & Welsch, 1980; Cook &Weisberg, 1982), the logistic regression model (Pregibon, 1981; Johnson, 1985), the Weibull model for censored data (Pregibon, 1981) and the proportional hazards model (Reid & Crepeau, 1985, Bin Daud 1987, Noor Akma 1994, Elfaki 2000).
The focus of this paper is use of diagnostic technique by performing it on independent competing risks with case deletion using proportional hazards model based on censored data.

2. The Proportional hazards Regression Model

We assume that we have n cases and for each we observe the vector (t, δ, z) where t is the time until failure, if the case is uncensored $(\delta = 1)$, or it is the time until removal or censoring $(\delta = 0)$. For each case there is a $p \times 1$ vector z of the explanatory or regression variables. The proportional hazards regression model is defined as follows; we let the hazard function of the failure time to a time constant vector explanatory variable z by taking

$$h(t / z) = \lambda(t) \exp(z\beta) \tag{1}$$

where $\lambda(t)$ is an unspecified base-line hazard function corresponding to the case $z = 0$ and β is a $p \times 1$ regression parameter. Cox (1972; 1975) proposed that we may use the partial likelihood

$$L(\beta) = \prod_{i=1}^{n} \frac{\exp(\beta^T z)}{\sum_{i \in R_i} \exp(\beta^T z)} \tag{2}$$

for inference concerning β, where the product is evaluated over all observed uncensored failure times, and R_i is the "risk set" for ith observed failure, that is, the set of individual surviving and uncensored at t.

To find $\hat{\beta}$ usually $U(\beta) = 0$ is solved where $U(\beta)$ is the $p \times 1$ score vector of derivatives of $L(\beta)$, the log likelihood is

$$LogL(\beta) = \sum_{i=1}^{n} (\beta^T z) - \log\left[\sum_{i \in R_i} \exp(\beta^T z)\right] \tag{3}$$

From equation (3) we obtained

$$U(\beta) = \frac{\partial \log L(\beta)}{\partial \beta_i} = \sum_{i \in R_i} \left[z - \frac{\sum_{i \in R_i} z \exp(\beta^T z)}{\sum_{i \in R_i} \exp(\beta^T z)} \right] \tag{4}$$

By taking the second derivative of equation (3), an expression is obtained which has the form of a variance. For example, the derivative of (4) with respect to β_p is:

$$\frac{\partial^2 \log L_1}{\partial \beta_p^2} = -\sum_{i \in R_i} \left[\frac{\sum_{i \in R_i} z^2 \exp(\beta^T z)}{\sum_{i \in R_i} \exp(\beta^T z)} - \left(\frac{\sum_{i \in R_i} z^2 \exp(\beta^T z)}{\sum_{i \in R_i} \exp(\beta^T z)} \right)^2 \right] \tag{5}$$

Maximum-likelihood estimates of β can be obtained by iterative (EM algorithm or Newton-Raphson methods) use of (4) and (5) in usual way.

3. Influence Measurement

A general approach to influence is given in Cook and Weisberg (1982). We shall confine our study mainly by adopting the case deletion approach. In linear regression, Cook (1977), Cook and Weisberg (1982) and others, suggested a suitably weighted combination of the changes $\hat{\beta}_i - \hat{\beta}$ as measures of influence. They all gave versions of the Cook's (1977) distance define as;

$$D_i = [(\hat{\beta}_{(i)} - \hat{\beta})' X' X (\hat{\beta}_{(i)} - \hat{\beta})]/(s\sigma^2), \quad i = 1,\ldots,n \tag{6}$$

where $\hat{\beta}$ indicates an estimate for β with full data. Full data in this context refers to the failure time[1] for all observations that can be obtained until the study is completed, while $\hat{\beta}_{(i)}$ indicates estimate for β by deleting data point i, $X'X$ is a positive (semi-) definite matrix, s is the parameter number, and σ^2 is the variance. Likewise, equation (6) becomes the basis for most distance measurements in detecting the influence of an observation or a case.

Influence measurements for the ordinary least square are generally based on the change in parameter estimate when the i observation is deleted, that is, $\hat{\beta}_i - \hat{\beta}$, where $\hat{\beta}_i$ is the estimate of β_i when the ith observation is deleted (Cook & Weisberg, 1980; Belsley et al., 1980). This difference in measurement has been applied in other computationally more complex settings by using a one-step estimate of $\hat{\beta}_i$ (Cook & Wang, 1983), and can be implemented for either the EM algorithm or Newton-Raphson method that will be discussed in a later section in this paper. For a single deletion with the ith case omitted from data, the change is given by

$$\hat{\beta}_i - \hat{\beta} = \frac{(X'X)^{-1} X_i r_i}{m_{ii}} \tag{7}$$

where X_i is the ith row of the design matrices X, r_i is the ith residual and m_{ii} is the ith diagonal element of the projection matrices $m = 1 - H$ with $H = X(X'X)^{-1}X'$.

4. Cook's Distance

To get the influence in equation (7) in a quantitative form, which is more meaningful, we used Cook's distance (Cook, 1977, 1979; Cook and Weisberg, 1982), which was defined earlier in this paper by equation (6). The usefulness of this is the availability of several ways to measure the scale of change vector involved in perturbation like equation (7), as suggested. For normal linear regression model with the least square, Cook (1977) recommended a form of scale measurement known as Cook's distance, which is given as follows,

$$D_i = \frac{(\hat{\beta}_i - \hat{\beta})' X'X (\hat{\beta}_i - \hat{\beta})}{q\hat{\sigma}^2} \tag{8}$$

where $\hat{\beta}$ is the estimate of the least square for full data, $\hat{\beta}_i$ is the estimate of β_i when the i observation is deleted, q is the parameter number and $\hat{\sigma}^2$ is the estimate of variance.

[1] The time observed on individual or object from one original point to the time an anticipated event occurs.

In line with the least square technique, Pregibon (1981) has produced a matrix which has a role like *H.* Cook's distance will be considered based on Cook and Weisberg (1982), that is,

$$D_i = \frac{(\hat{\beta}_i - \hat{\beta})' M (\hat{\beta}_i - \hat{\beta})}{g} \tag{9}$$

where M is a semi positive exact symmetry matrices, and g a positive scale factor.

5. Likelihood Distance

A more general method to get measurement of influence is by using contour measurement for the log likelihood function. Let $L(\beta)$ be the log likelihood on parameter β based on full data. Likelihood distance (Cook and Weisberg, 1982) is defined as

$$LD_i = 2\left[L(\beta) - L(\beta_i)\right] \tag{10}$$

where $L(\beta_i)$ is the log likelihood function of β_i when the i observation is deleted. LD_i measurement can also be interpreted in terms of the asymptotic confidence region (see Cox and Hinkley, 1974)

$$\beta : 2\{L(\beta) - L(\beta_i)\} \leq \chi^2(\alpha; s) \tag{11}$$

where $\chi^2(\alpha; s)$ is the upper α point from chi-squared distribution with s degree of freedom.

If the log likelihood contours are approximately elliptical, LD_i is quadratic, and can be drawn closer to Taylor's expansion around $\hat{\beta}$, as follows,

$$L(\hat{\beta}_i) \cong L(\beta) + (\hat{\beta}_i - \hat{\beta})' \frac{\partial L(\beta)}{\partial \beta} + \frac{1}{2}(\hat{\beta}_i - \hat{\beta})' \frac{\partial^2 L(\beta)}{\partial \beta \partial \beta_i}, \tag{12}$$

since $\dfrac{\partial L(\beta)}{\partial \beta} = 0$ at maximum likelihood,. Then,

$$LD_i \cong (\hat{\beta}_i - \hat{\beta})' \left(- \frac{\partial^2 L(\beta)}{\partial \beta \partial \beta_i} \right) (\hat{\beta}_i - \hat{\beta}) \tag{13}$$

where $-\dfrac{\partial^2 L(\beta)}{\partial \beta \partial \beta_i}$, is the observed information matrix. This approach can be confusing if log likelihood contour is nonelliptical. Hence, the approach of Taylor's expansion for log likelihood, which is not elliptic, needs a better development, which will involve a complex expression.

6. Methods for Estimation

We turn now to a discussion of methods that can be used to detect influential cases under competing risks model that is fitted to censored data. The $\beta' s$ are found by iterative numerical techniques. This is a cumbersome task, since, apart from the computation needed to obtain $\hat{\beta}$, we are still required to calculate for each β_i, $i = 1,2,...,n$. To reduce computation, approximations to β_i are of great value. An obvious choice in such a situation is the one-step technique, that is, to compute from the maximum likelihood estimate $\hat{\beta}$ the first step of an iterative process to find $\hat{\beta}_i$.

Cain and Lange (1984) considered the case-weight perturbation scheme in proportional hazard regression model by

$$l(\beta / w) = \prod_{i=1}^{n} \frac{\exp(\beta^T z_i \delta_i w_i)}{\left[\sum_{i \in R_i} w_i \exp(\beta^T z_i)\right]^{w_i}} \tag{14}$$

where z_i, δ_i, and β are defined in earlier in this paper, and w_i is a vector of ones for the complete data set-up. If $w_i = 0$ and the other $w_i{}'s$ are one, then $l(\beta / w)$ corresponds to the partial likelihood with the ith case omitted, and $\hat{\beta}(w)$ corresponds to $\hat{\beta}_i$. Using the perturbation scheme as in equation (14), we have $l(\beta / w_i = 0)$ equivalent to the partial likelihood with ith case omitted from the data set. Our attempt to change the parameter involves one-step iterative for maximizing $l(\beta)$ applied to $l(\beta / w_i = 0)$ starting at $\hat{\beta}$.

The one-step iterative method for competing risks model is complex, due to the expression to log likelihood. The one-step methods for both EM algorithm and Newton-Raphson will be discussed in the following section.

6.1. One-step EM

The general arguments that follow are applicable in the scale Cox's finite, and continuous models, only by using the EM and ML methods. The one-step techniques, as mentioned earlier in this paper, which are defined to be the maximum likelihood estimate of β with the ith case deleted, are formed by taking one-step of the iterative process for finding $\hat{\beta}_i$ starting at $\hat{\beta}$. The EM approach with exponential distribution is the best suited for adapting to cases that are deleted from the data set. We can define the resulting iterative weighted least square scheme for the full data by

$$\beta_{j+1} = \beta_j + (z^T z)^{-1} z^T b \tag{15}$$

where z is the $n \times p$ design matrices of the covariates, and b has a component

$$b_i = 1 - \lambda_i \left\{ \frac{1-\delta_i}{\lambda_i} + \sum_{t_j \le t_i}^{n} \left[\frac{\delta_i}{\sum_{r \in R_j} \lambda_r} \right] \right\}$$

$$= \delta_i - \lambda_i \sum_{t_j \le t_i}^{n} \frac{\delta_j}{\sum_{r \in R_j} \lambda_r} \tag{16}$$

where $\lambda_i = \exp(\beta^T z)$.

The covariate matrices z are column centered, that is, $\sum_{i=1}^{n} z_i = 0$.

The change in estimates, after deletion of the ith observation, based on the one-step EM algorithm, is given by

$$\hat{\beta}_i - \hat{\beta} = \left[z_i^T z_i\right]^{-1} z_i^T \qquad (17)$$

where z_i is the covariate matrices obtained from z with the ith row omitted, starting from $\hat{\beta} = \beta$.

6.2. One-step ML

The iterative scheme for the Newton-Raphson is given by

$$\hat{\beta} = \beta^* + U(\beta^*) I^{-1}(\beta^*) \qquad (18)$$

where $U(\beta^*)$ and $I^{-1}(\beta^*)$ are defined in equation (4) and (5). Using the same arguments as in the preceding section, the one-step change from β, for a Newton-Raphson scheme such as (18), is written as

$$\hat{\beta}_i - \hat{\beta} = U_i(\beta) I_i^{-1}(\beta) \qquad (19)$$

where U_i is the ith element of $U(\beta)$ and I_i is the ith element of I. Equation (19) is equivalent to that of the robust method of Cook and Weisberg, (1982).

The one-step estimate can be computed directly at the final iteration for full data, thus a one-step influence measures may be obtained. SAS software may easily be used to carry out the computation of (17) and (19). However, for ease of manipulation of matrices, the computation in our analysis is programmed in SAS. Results from some simulation data sets are given in the next section.

7. Simulation Data

The two simulations were performed according to two different sample sizes with varying percentage of censored observation (simulation 1 and 2 for sample size 15 with 50 percent censoring, sample size 30 with 33 percent censoring, respectively, generated 1000 times). We took the average for the failure time and covariate to make the calculation simple. To generate failure time, the value of $\lambda = .45$ for the first type of failure and $\lambda = .057$ for the second type of failure were used. In this simulation, the objective is to find Cook's distance based on equation (6) and likelihood distance based on equation (10) from the first risk and the second risk for every sample size. Both distances were calculated by using one-step EM algorithm based on equation (17) and one-step ML based on equation (19) under competing risks model.

Simulation 1 (sample size 15 with 50 percent censoring), in Table 1, shows the Cook's distance and likelihood distance obtained by the one-step EM algorithm and one-step ML under competing risk model based on equation (1).

From the first risk $(k = 1)$, based on a large value for Cook's distance from both methods, and the five observations, that is, number 5, 7, 9, 11, and 14, show a large influence on parameter estimates compared to other observations. This can clearly be seen in Figures 1 and 2. Also, based on a large value for likelihood distance, the same observations show to have a large influence on parameter estimates, compared with other observations, as it can clearly be seen in Figures 3 and 4. But, Cook's distance obtained by both methods was found to have smaller value compared with likelihood distance.

From the second risk $(k = 2)$, the observations number 5, 7, 9, 11 and 14 showed to have large value from Cook's distance and likelihood distance obtained by one-step EM. The same observations and observations number 13 and 15 have large value from likelihood distance obtained by one-step ML method. However, Cook's distance obtained by both methods has smaller value compared with likelihood distance. Finally, the Cook's distance and likelihood distance obtained by one-step EM algorithm showed to be smaller than the others obtained by one-step ML methods from both risks.

Plots of Cook's distance (D_i) and Likelihood distance (LD_i) one-step EM (first risk) against case number were given in Figure 1 and 3, respectively. Inspecting the $D_i's$, the following can be inferred: case 7 has $D_7 = .33$ and $LD_7 = 0.63$ for D_i and LD_i, respectively, suggesting that case 7 for LD_i, $LD_7 = 0.63$, may have a large enough influence to induce the anomaly. Similarly, for D_i and LD_i one-step ML (first risk), Figure 2 and 4, respectively, in cases 5 and 14, that is, $D_5 = D_{14} = 0.198$ for D_i, while $LD_7 = 0.58$ for LD_i for case 7.

However, considering Figures 5 and 7 for D_i and LD_i, respectively, similar results with a possible cause: case 7 with $D_7 = 0.32$ for D_i and case 11 with $LD_{11} = 0.37$ for LD_i are obtained by one-step EM from the second risk, while D_i and LD_i for one-step ML from second risk (Figure 6 and 8, respectively) are obtained in case 14 with $D_{14} = 0.214$ and case 7 with $LD_7 = 0.58$. Despite the fact that $LD_i's$ have higher values for both one-step EM and ML from the second risk, it can clearly be seen that the ranges between the highest two values is too small compared to the highest two values of $D_i's$.

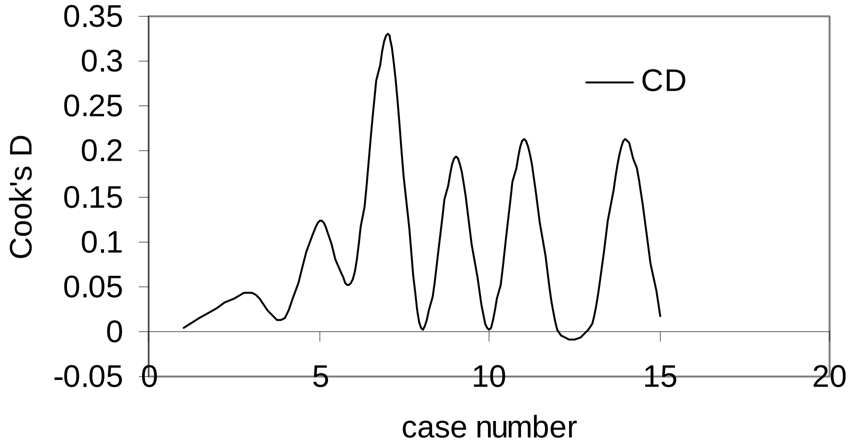

Figure 1: Cook's Distance from simulation 1 obtained by one-step EM (first risk)

Table 1: Cook's distance and Likelihood distance obtained by one-step EM and one-step ML under competing risks model from simulation 1 (sample size 15)

First Risk

	One-step EM		One-step-ML	
case	Cook's D	Likelihood D	Cook's D	Likelihood D
1	0.003	0.055	0.054	0.062
2	0.025	0.028	0.045	0.061
3	0.044	0.050	0.044	0.056
4	0.015	0.016	0.034	0.024
5	0.122	0.16	0.198	0.30
6	0.058	0.08	0.058	0.09
7	0.33	0.63	0.057	0.58
8	0.003	0.005	0.043	0.025
9	0.195	0.30	0.088	0.12
10	0.001	0.002	0.045	0.06
11	0.214	0.30	0.088	0.14
12	0.001	0.001	0.054	0.04
13	0.009	0.012	0.089	0.12
14	0.214	0.31	0.198	0.34
15	0.017	0.017	0.054	0.061

Second Risk

	One-step EM		One-step-ML	
case	Cook's D	Likelihood D	Cook's D	Likelihood D
1	0.004	-0.065	0.005	0.062
2	0.025	0.028	0.045	0.061
3	0.045	0.055	0.046	0.067
4	0.016	0.036	0.034	0.044
5	0.152	0.36	0.198	0.45
6	0.068	0.08	0.088	0.095
7	0.327	0.34	0.078	0.58
8	0.003	0.097	0.045	0.097
9	0.213	0.36	0.088	0.253
10	0.003	0.042	0.065	0.06
11	0.214	0.37	0.088	0.41
12	0.002	0.052	0.056	0.063
13	0.009	0.195	0.089	0.21
14	0.214	0.34	0.214	0.44
15	0.037	0.223	0.055	0.261

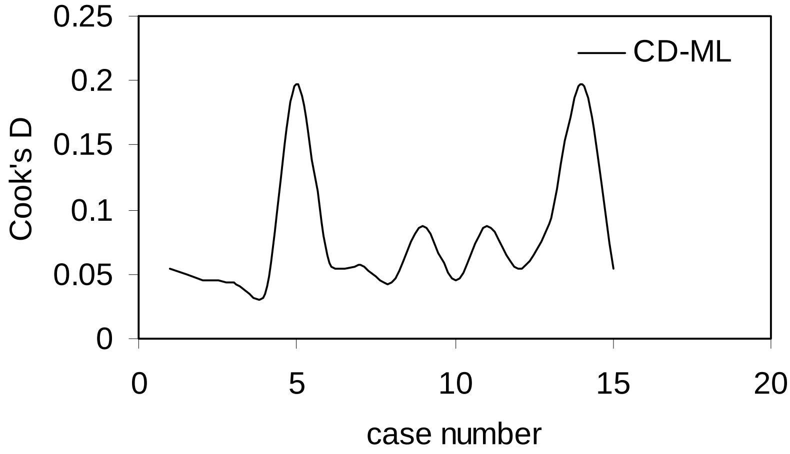

Figure 2: Cook's Distance from simulation 1 obtained by one-step ML (first risk)

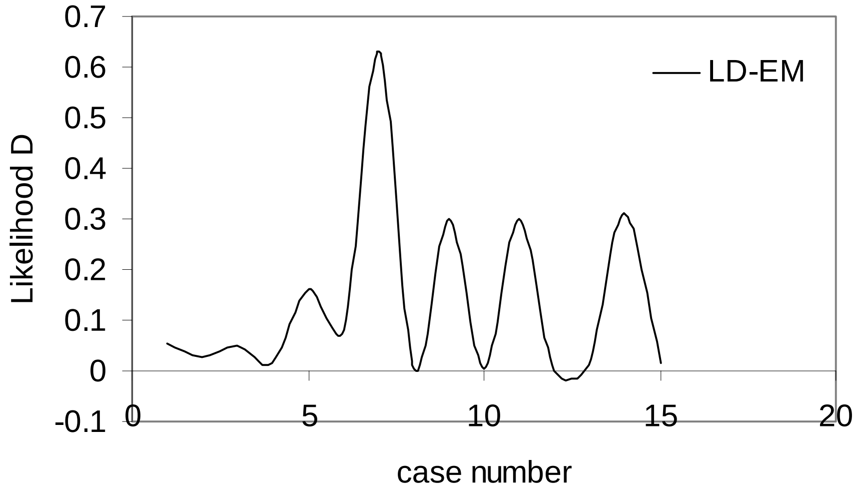

Figure 3: Likelihood distance from simulation 1 obtained by one-step EM (first risk)

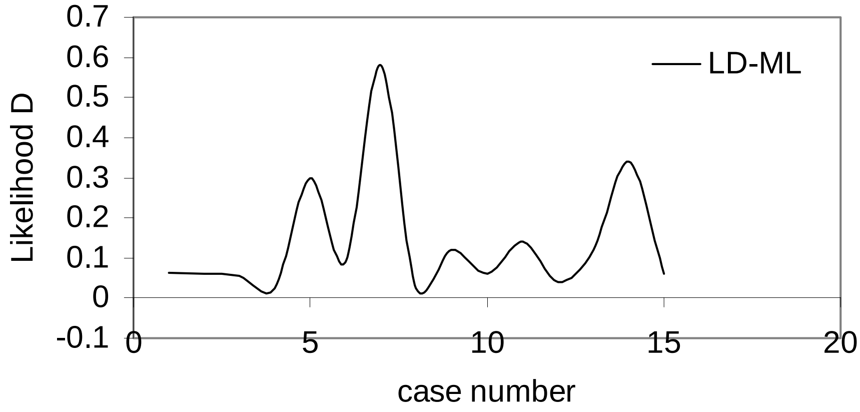

Figure 4: Likelihood distance from simulation 1 obtained by one-step ML (first risk)

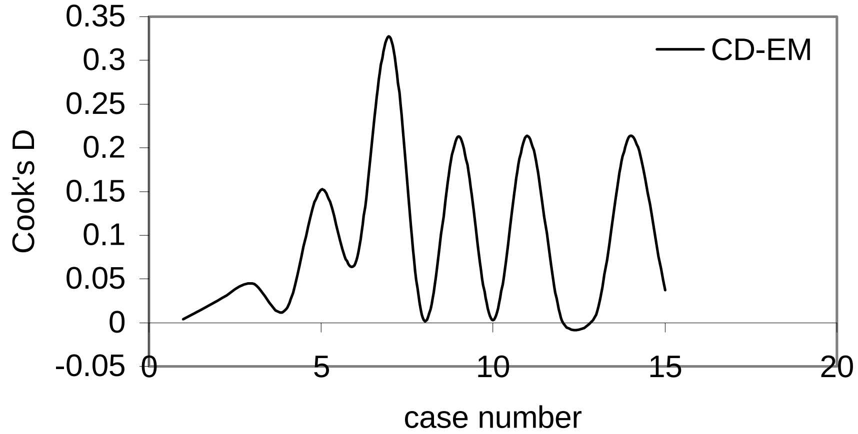

Figure 5: Cook's distance from simulation 1 obtained by one-step EM (second risk)

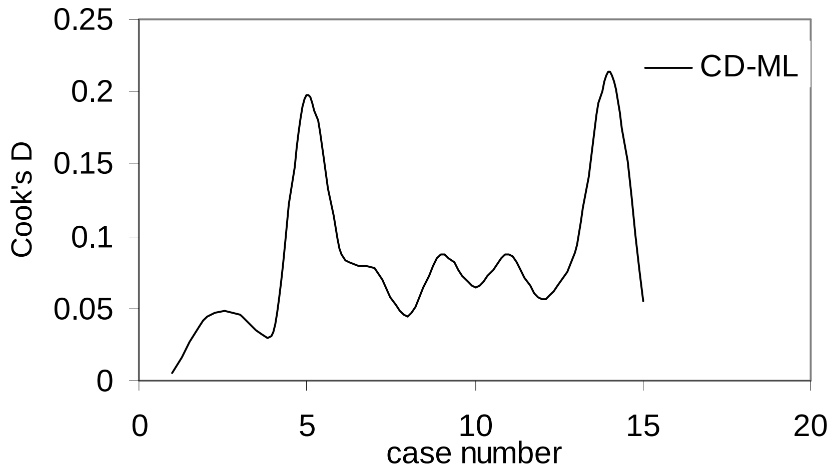

Figure 6: Cook's distance from simulation 1 obtained by one-step ML (second risk)

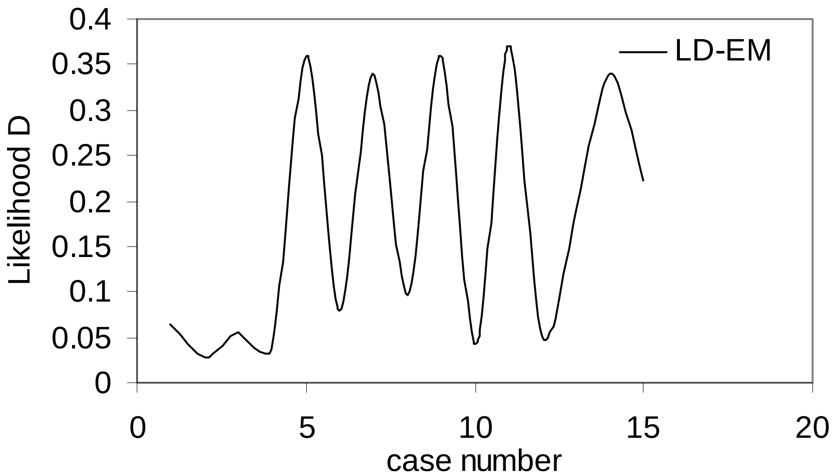

Figure 7: Likelihood distance from simulation 1 obtained by one-step EM (second risk)

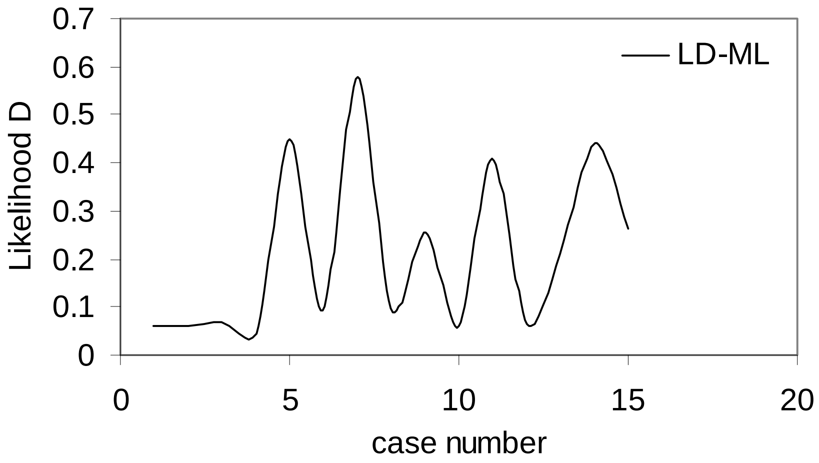

Figure 8: Likelihood distance from simulation 1 obtained by one-step ML (second risk)

Simulation 2 (sample size 30, with 33 percent censoring), Tables 2 and 3 show Cook's distance and likelihood distance obtained by one-step EM algorithm method and one-step ML method under competing risks model. Table 2 shows the first risk $(k = 1)$, and the value of Cook's distance for case number 30 obtained by one-step EM has the largest potential and largest influence. Also, the value of Cook's distance for cases 7, 18 and 24 obtained by one-step ML have largest potential and largest influence and the second largest value of Cook's distance, $D_1 = D_{13} = D_{15} = D_{27} = 0.105$. Only case number 1 is individually influential, compared with other cases of the value of likelihood distance obtained by one-step ML method. However, Cook's distance obtained by both methods have smaller values than likelihood distance.

Table 3 shows the second risk $(k = 2)$, and the value of Cook's distance of case number 30 only has the largest influence, compared with other cases obtained by one-step EM. Also, $D_1 = D_{18} = D_{24} = 0.145$ obtained by one-step ML method have the largest potential, and the second largest value of Cook's distance obtained by this method is from case numbers 2, 5, 8, 12, 14, 17, 21, 22, 25 and 29. However, based on likelihood distance, case number 10 and 30 obtained by one-step EM and one-step ML, respectively, have the largest value, compared with other cases from both methods. Finally, the value of Cook's distance and likelihood distance obtained by one-step EM show to be smaller than the others obtained by one-step ML. Moreover, Cook's distances show to have a smaller value than likelihood distance.

Table 2: Cook's distance and Likelihood distance obtained by one-step EM and one-step ML under competing risks model from simulation 2 (first risk)

First Risk				
	One-step EM		One-step-ML	
case	Cook's D	Likelihood D	Cook's D	Likelihood D
1	0.000	0.001	0.105	0.423
2	0.000	0.004	0.012	0.168
3	0.008	0.141	0.010	0.014
4	0.000	0.008	0.012	0.047
5	0.000	0.0003	0.012	0.0167
6	0.007	0.136	0.010	0.191
7	0.006	0.1225	0.134	0.214
8	0.000	0.003	0.012	0.024
9	0.000	0.007	0.012	0.034
10	0.007	0.164	0.010	0.097
11	0.006	0.120	0.010	0.095
12	0.000	0.002	0.012	0.123
13	0.000	0.006	0.105	0.034
14	0.000	0.0022	0.012	0.043
15	0.000	0.0045	0.105	0.054
16	0.006	0.1272	0.010	0.024
17	0.000	0.001	0.012	0.165
18	0.006	0.1188	0.134	0.054
19	0.007	0.1293	0.010	0.054
20	0.006	0.1275	0.010	0.077
21	0.000	0.003	0.012	0.168
22	0.000	0.004	0.012	0.045
23	0.007	0.137	0.010	0.097
24	0.005	0.115	0.134	0.123
25	0.000	0.001	0.012	0.021
26	0.006	0.121	0.010	0.0158
27	0.000	0.0045	0.105	0.043
28	0.006	0.124	0.010	0.158
29	0.000	0.008	0.012	0.168
30	1.077	0.000	0.010	0.158

Table 3: Cook's distance and Likelihood distance obtained by one-step EM and one-step ML under competing risks model from simulation 2 (second risk)

	Second Risk			
	One-step EM		One-step-ML	
case	Cook's D	Likelihood D	Cook's D	Likelihood D
1	0.001	0.001	0.005	0.003
2	0.001	0.004	0.012	0.168
3	0.010	0.141	0.010	0.014
4	0.002	0.008	0.010	0.075
5	0.003	0.0003	0.012	0.017
6	0.007	0.136	0.010	0.214
7	0.006	0.1225	0.145	0.214
8	0.001	0.003	0.012	0.024
9	0.002	0.007	0.010	0.034
10	0.007	0.164	0.010	0.097
11	0.006	0.120	0.010	0.095
12	0.001	0.002	0.012	0.173
13	0.000	0.006	0.005	0.034
14	0.000	0.0022	0.012	0.043
15	0.004	0.0045	0.005	0.054
16	0.006	0.1272	0.010	0.024
17	0.000	0.001	0.012	0.156
18	0.008	0.1188	0.145	0.054
19	0.007	0.1293	0.010	0.054
20	0.007	0.1275	0.010	0.077
21	0.000	0.003	0.012	0.168
22	0.003	0.004	0.012	0.045
23	0.007	0.137	0.010	0.097
24	0.008	0.115	0.145	0.123
25	0.000	0.001	0.012	0.021
26	0.009	0.121	0.010	0.0158
27	0.000	0.0045	0.005	0.043
28	0.005	0.124	0.010	0.058
29	0.000	0.008	0.012	0.168
30	1.230	0.000	0.010	0.580

8. Conclusion

The influence measurements are based on case deletion. Within this technique, a vector needs to be introduced so that an assessment on the case can be used to obtain Cook's distance and likelihood distance. The diagnostic techniques and influence used identify cases of influence constructed for competing risk model were successful, and these techniques were able to identify odd cases using the simulation data sets(Elfaki, 2000).

One-step technique was used to derive the estimate of parameter, followed by distance measurement. These techniques are one-step EM and one-step ML. they were introduced to reduce the iterative steps in computing the influence measurements. It was found that from both simulations the Cook's distance was significant compare to likelihood distance because it can show clearly the outlier observation. Also the results obtained by one-step EM is significant compare to the one obtained by one-step ML.

Reference

Andersen, P. K., Borgan. Q., Gill. R. D. and Keiding, N. (1993). *Statistical Models Based on Counting Processes*. Springer-Verlag.

Belsley, D. A., Kuh, E., and Welsch, R. E. (1980). Regression Diagnostics: *Identifying Influential Data and Sources of Collinearity*. New York. Wiley.

Bin Daud, I. (1987). *Influence Diagnostics in Regression with Censored Data*. Ph.D., thesis, Loughborough University.

Cain, K. C. and Lange, N. T. (1984). "*Approximate Case Influence for the Proportional hazards regression Model with Censored data."* Biometrics. 40: 493-499.

Cook, R. D. (1977). "*Detection of Influence Observation in Linear Regression.*" Technometrics. 19: 15-18

Cook, R. D. (1979). "*Influence Observations in Linear Regression.* " Journal of American Statistical Association.74: 169-174.

Cook, R. D. (1986). "*Assessment of Local Influence (with Discussion).*" *Journal of Royal Statistical Society*. 48:133-169.

Cook, R. D. and Wange, P. C. (1983). "*Transformations and Influential Cases in Regression.*" Technometrics. Vol. 25. No. 4: 337-343.

Cook, R. D. and Weisberg. S. (1980). "*Characterizations of an Empirical Influence Function for Detecting influential Cases in Regression.*" Technometrics. Vol. 22. No. 4: 495-507.

Cook, R. D. and Weisberg. S. (1982). *Residuals and Influence in Regression*. New York and London; Chapman and Hall.

Cox, D. R. (1972)."*Regression Models and Life Tables (with discussion)*". J. R. . Statistic. Soc. 34: 187-220.

Cox, D. R. (1975). "*Partial likelihood*". Biometrika. 62: 269-276.

Cox, D. R. and. Hinkely, D. V (1974). *Theoretical Statistics*. London: Chapman and Hall

Cox, D. R. and D. Oakes (1984) *Analysis of Survival Data*. London: Chapman and Hall.

Elfaki, F. A. M. (2000) *EM Approach on Influence Measures in Competing Risks Via Proportional Hazard Regression Model*. M.Sc thesis University Putra Malaysia.

Noor, Akma. Ibrahim (1994). *Influence Diagnostics in Competing Risks*. Ph. D., thesis. Universiti Pertanian Malaysia.

Pregibon, D. (1981). "*Logistic Regression Diagnostics.*" Annals of Statistics. 9: 705-724.

Reid, N. and Crepeau, H. (1985). "Influence Function for Proportional Hazard *Regression.*" Biometrika. 72: 1-9.