Journal of Statistical Research *ISSN 0256 - 422 X*
2002, Vol. 36, No. 1, pp. 85-97
Bangladesh

TESTING THE SIGNIFICANCE OF LOCAL INFLUENCE

Monzur Hossain[1] and M. Ataharul Islam[2]

SUMMARY

The motivation behind the influence analysis is to increase the adequacy of the fitted model. Several local influence diagnostics have been proposed for different models such as linear regression, generalized linear, Weibull regression, proportional hazards etc. models by different authors on the basis of Cook's (1986) local influence method proposed for linear regression. As testing the significance of local influence is a motivating problem, in this paper we develop a likelihood ratio based test procedure for testing the significance of local influence on the parameter estimates and application is shown by fitting a logistic regression model to the Framingham Heart Study data set. This test procedure is based on the ideology of testing the equality of parameters of postulated and perturbed model. The proposed test procedure can be extended to the model having smooth and well-behaved likelihood and perturbation function.

Keywords: Local influence, Diagnostics, Perturbation function, Logistic regression model.

1 Introduction

Influence emerges from the interaction between the model and the bad elements of the data for which valid conclusion from a fitted model can not be drawn. Local influence is the minor perturbations of a model and assessment of local influence is necessary for the best fit of a model. Several diagnostics have been developed for assessing the local influence for the perturbations of case-weights, explanatory variables and for assessing effect of specific perturbations on the estimates by different authors for different models. Considering local influence diagnostics, index plot of U_{max} (usually eigen vector components of the influence matrix corresponding to maximum eigen value) can detect the influential cases and maximum curvature diagnostic (C_{max}) gives a gross idea about the presence of local influence on

[1]Monzur Hossain is an Assistant Director of Bangladesh Bank (the central bank of Bangladesh). When the paper was submitted, he was working as a researcher (data analyst) at Bangladesh Institute of Research for Promotion of Essential and Reproductive Health and Technologies (BIRPERHT).
Email: monzur_ h@yahoo.com
[2]M. Ataharul Islam Ph.D., Professor, Department of Statistics, University of Dhaka, Bangladesh . Now on leave as a visiting faculty in the Department of Mathematical Sciences, University Sains Malaysia, Malaysia.

the parameter estimates. The usual local influence diagnostics are not able to indicate the presence of local influence on single parameter estimates. However, no successful attempt has been made so far to test the significance of local influence on the estimates of parameters. The focus of this paper is to develop a test procedure to test the significance of local influence on the estimates of parameters.

On the basis of Cook's (1986) local influence method for linear regression model, Thomas and Cook (1989) apply local influence methods to the generalized linear model, Weissfeld (1990) do the same for the proportional hazards model while Weissfeld and Schneider (1990 a,b) extend local influence diagnostics for normal linear model and for Weibull regression model with censored data respectively. Various aspects of local influence techniques and likelihood displacement have also been discussed by Escobar and Meeker (1992). But most of these are concerned with the identification of influential elements and measurement of influence on the parameter estimates by the curvature diagnostics. Besides this, Pregibon (1979, 1981) shows that if log-likelihood displacement $D(\omega) > {\chi^2}_{(1-\alpha,p)}$, the perturbation ω results in a $\hat{\theta}_\omega$ that lies outside of the null perturbation approximate likelihood-ratio-based $100(1-\alpha)\%$ confidence region for θ. This is the generalization of the F-calibration of Cook's D statistics.

After assessing locally influential elements and local influence on the parameter estimates, it is important and necessary to test the significance of local influence on the parameter estimates. Escobar and Meeker (1992) shows for perturbing a single case that $\frac{1}{2}H_{ii} > \chi^2_{(0.5,p)}$ provides a warning signal that the perturbation could be influential. Also the classical test procedures do not seem adequate to test the influence on the parameter estimates. Here a test procedure for testing the significance of local influence on the parameter estimates is developed. This test procedure is mainly based on the test procedure developed by Islam (1994) to test the equality of parameters for repeated transitions for proportional hazards (PH) models. In his test procedure Islam compares the parameters of two types of PH models for two time periods. This concept is extended in our test procedure for testing whether the parameters of postulated and the perturbed model are equal or not, i.e., whether perturbation scheme ω is influential on the parameter estimates. For illustration we apply the local influence diagnostics of logistic regression model (Hossain, 1997) for different perturbation schemes to the Framingham Heart Study (FHS) data set (see Kahn and Sempos, 1989). The proposed test procedure is then applied to test the significance of local influence on the parameter estimates.

2 Local Influence Techniques

We discuss some facts about local influence methods in this section. Let $\beta = (\beta_0, \beta_1, \ldots, \beta_p)^T$ be the $(p+1) \times 1$ vector of regression coefficients of a particular model which follows the assumptions of local influence techniques (see Cook, 1986). In local influence analysis an open region Ω of small perturbation ω is considered and it is thought out that likelihood displacement $D(\omega)$ contains essential information on β for the effect of perturbation scheme

ω which is defined as

$$D(\omega) = 2\{L(\hat{\beta}) - L(\hat{\beta} \mid \omega)\}$$

where $L(\hat{\beta})$ and $L(\hat{\beta} \mid \omega)$ denote the log-likelihood of the postulated and perturbed model respectively. For the specific perturbation scheme there is a point ω in Ω that represents null perturbation of the data i.e., $L(\beta \mid \omega) = L(\beta)$. Cook (1986) argued that a direction of large change at $D(\omega)$ can be used to identify the locally influential elements. Since $D(\omega)$ achieves a local minimum at ω_0, a simple representation of normal curvature can be obtained as

$$C_u = \mid u' H u \mid$$

where U is a direction vector of unit length and $H = \delta^2 L(\beta \mid \omega)/\delta\omega\omega'$. In this form computation of H is very difficult so that Cook (1986) recalculate it as

$$H = \Delta' I^{-1} \Delta$$

where I^{-1} is the variance-covariance matrix of the model evaluated at $\omega = \omega_0$ and $\Delta = \delta^2 L(\beta \mid \omega)/\delta\beta_i\omega_j$ evaluated at $\beta = \hat{\beta_\omega}$. Hence direction vector of unit length U_{max} of C_{max} can be obtained by considering the eigen vector of the influence matrix H corresponding to the largest eigen value which produces greatest local change in the regression coefficients. From the above discussion one thing is clear that ω results through the perturbed parameter estimates $\hat{\beta_\omega}$ and then the local influence diagnostics are used to identify the locally influential elements. Hence it is important to investigate the local change on $\hat{\beta}$ for the perturbation scheme ω and this can be done by comparing perturbed estimates with the estimates evaluated at $\omega = \omega_0$.

3 Logistic Regression Model and Perturbations

It is noted that logistic regression model has found pretty well in dealing with categorical outcome variable (Hosmer and Lemeshow, 1989) and it involves non-linear estimation procedure. Let us consider the logistic model of the form

$$\pi_i = P(Y_i = 1 \mid X_i = x) = \frac{e^{X_i\beta}}{1 + e^{X_i\beta}}$$

and

$$1 - \pi_i = P(Y_i = 0 \mid X_i = x) = \frac{1}{1 + e^{X_i\beta}}$$

where β denote $(p+1)\times 1$ vector of regression coefficients and X denote explanatory variables in which X_i is the i^{th} row of the design matrix X. Here Y_i denote the dichotomous outcome variables where 1 indicates the i^{th} individual has experienced an event and 0 otherwise.

The perturbed estimates can be obtained by considering likelihood functions for different perturbation schemes such as for the perturbations of (i) case-weights, (ii) explanatory variable (one or more), (iii) individual coefficient etc. To assess the local influence or to obtain perturbed estimates of logistic regression model, the diagnostics and likelihood functions proposed by Thomas and Cook (1989) for generalized linear model can be used by taking link functions for logit distributions. The perturbed estimates can also be obtained by extending local influence diagnostics directly for logistic regression model (Hossain, 1997) on the basis of Cook's (1986) method. The methods are briefly discussed below.

3.1 Perturbations of Case-weights

Let $\omega' = (= \omega_1, \ldots, \omega_n)$ be a vector of weights and $\omega_0' = (1, \ldots, 1)$ regresents $n \times 1$ vector of null perturbation. To assess the influence for the case-weight perturbations, the perturbed log-likelihood is defined by

$$L(\beta \mid \omega) = \sum_{i=1}^{n} \omega_i \left[Y_i X_i \beta - \log(1 + e^{X_i \beta}) \right].$$

The $(i,j)^{th}$ elements of Δ matrix of order $(p+1) \times n$ is given by

$$\Delta_{ij} = \frac{\delta^2 L(\beta \mid \omega)}{\delta \beta_j \delta \omega_i} = \left[y_i x_{ij} - \frac{e^{x_{ij}\beta_j}}{1 + e^{x_{ij}\beta_j}} x_{ij} \right] = (y_i - \pi_i) x_{ij} \quad ; j = 0, 1, \ldots, p \quad ; i = 1, 2, \ldots, n$$

evaluated at ω_0 and $\hat{\beta}$. The influence matrix H of order $n \times n$ can be obtained by following the relation $H = \Delta' I^{-1} \Delta$ under the perturbation scheme. Thus the maximum curvature C_{max} can be easily computed as defined in section 2.

3.2 Perturbations of Explanatory Variables

Here we consider a general method for perturbing the whole design matrix X i.e., for the modification of all explanatory variables. Let β be a vector of parameters and the perturbed log-likelihood $L(\beta \mid \omega)$ can be obtained by replacing explanatory variables X with Z i.e.,

$$Z = X + WV$$

where $W = (\omega_{ij})$ is a $n \times (p+1)$ matrix of perturbations and $V = \text{diag}(v_1, v_2, \ldots, v_{p+1})$ be the scaling factor which is used to convert the perturbations ω_{ij} to the appropriate size and units so that $\omega_{ij} v_j$ is consistent with the $(ij)^{th}$ element of X. Under this perturbation scheme, the perturbed log-likelihood will take the form

$$L(\beta \mid \omega) = \sum_i \left[y_i (x_{ij} + \omega_{ij} v_j) \beta_j - \ln(1 + e^{(x_{ij} + \omega_{ij} v_j)\beta_j}) \right];$$

i=1,2, ...,n and j=0,1,...,p

To compute the curvature diagnostics, we partition the Δ matrix of order $(p+1) \times n(p+1)$ into $p+1$ sub-matrices as

$$\Delta = (\Delta_1, \Delta_2, \ldots, \Delta_{p+1})$$

where k_{th} sub-matrix Δ_k of order $(p+1)\times n$ is defined by

$$\begin{aligned}\Delta_k &= \frac{\delta^2 L(\beta \mid \omega)}{\delta\beta_j\delta\omega_{ik}}\\ &= \frac{\delta}{\delta\omega_{ik}}\left[\sum_i (y_i-\pi_{\omega i})(x_{ij}+\omega_{ij}v_j)\right]\\ &= \sum_i \left((y_i-\pi_{\omega i})-\pi_{\omega i}(1-\pi_{\omega i})z_i\beta_j\right)v_j; i=1,2,\ldots,n \quad \text{and} \quad j=0,1,\ldots,p\end{aligned}$$

evaluated at ω_0 and $\hat{\beta}$. Thus the influence matrix H can be obtained through the relation $H = \Delta' I^{-1}\Delta$. These results can be extended to the situations where only one or more explanatory variables is of interest by setting $v_j = 0$ for the unperturbed variables.

3.3 Perturbations of Individual Coefficient

For examining the sensitivity of the i^{th} coefficient to each of the perturbation scheme discussed above, a curvature diagnostic is developed as suggested by Cook (1986). First we rearrange the columns of X as $X = (X^{(1)}, X^{(2)})$ so that the first column $X^{(1)}$ corresponds to the coefficient β_1 of interest. The curvature diagnostic for an individual coefficient is defined by

$$C_u(\beta_1) = 2 \mid u'\Delta'(I^{-1}-\ell)\Delta u \mid$$

where u is the eigen vector of $\Delta'(I^{-1}-\ell)\Delta$ corresponding to the largest eigenvalue, ℓ is defined by

$$\ell = \begin{pmatrix} 0 & 0 \\ 0 & P^{-1}\end{pmatrix}$$

and P is obtained from the partition of

$$I = \begin{pmatrix} M & N \\ O & P\end{pmatrix}.$$

Under specific perturbation scheme, the sensitivity of the coefficients can be investigated by this procedure.

4 Testing the Significance of Local Influence on the Parameter Estimates

Let β be the $(p+1)\times 1$ vector of unknown model parameters evaluated at $\omega = \omega_0$ and θ denote the $(p+1)\times 1$ vector of parameters of the model evaluated at ω. Consider the $(p+1)\times 1$ vector of parameters $\theta' = (\theta_1, \theta_2)$ of the another model evaluated at ω where θ_1 is the vector of parameters (order $\leq p$) of the model evaluated at $\omega = \omega_0$ and is a sub-set of β. Then influence of ω on θ_1 can be tested by the likelihood ratio test statistic

$$\lambda = -2[L(\hat{\theta}_1, \hat{\theta}_2) - L(\hat{\beta})]$$

which follows χ^2 distribution with d.f. equal to order of θ_1. Here it is important to note that the first or second log-likelihood term of right hand side can be interchanged since both the terms contain equal number of parameters (see Cook and Weissberg, 1982).

To test the local influence on the parameter estimates $\hat{\theta}_1$, our interest will mainly center on the total parameter vector of order p, an individual parameter or a subset of θ_1.

In the following section, the proposed test procedure is applied to test the significance of local influence on the parameter estimates of the logistic regression model as a specific example.

5 Application of the Test Procedure

For testing the significance of local influence on the j^{th} $(j = 1, 2, \ldots, p)$ parameter estimates $\hat{\beta}_j$, we consider the null hypothesis

$$H_0 \quad : \quad \beta_j = \beta_{\omega j}$$

against

$$H_1 \quad : \quad \beta_j \neq \beta_{\omega j}$$

where ω denote the particular perturbation scheme. Here we want to test the effect of small perturbation ω on the j^{th} parameter estimate $\hat{\beta}_j$.

To test the above null hypothesis, according to the theory stated above we consider the following three logistic models

$$\begin{aligned} \pi_i(1) &= \frac{e^{X_i\beta}}{1+e^{X_i\beta}} \\ \pi_{\omega i}(2) &= \frac{e^{X_i\beta_\omega}}{1+e^{X_i\beta_\omega}} \\ \text{and } \pi_{\omega i}(3) &= \frac{e^{X_i\theta}}{1+e^{X_i\theta}} \,. \end{aligned}$$

In the model (3), $\theta = (\hat{\theta}_1, \theta_2)$ where $\hat{\theta}_1$ is the estimates of the model (1) evaluated at $\omega = \omega_0$ and replacing it in the model (2), we will get the estimates $\hat{\theta}_2$ of θ at ω. In simple form θ can be expressed as follows:

(a) for testing local influence on total parameter vector, $\hat{\theta}_2 = \hat{\theta}_0$ evaluated at ω and $\hat{\theta}_1^T = (\hat{\beta})_{p\times 1}$ evaluated at $\omega = \omega_0$ i.e., $\hat{\theta}^T = (\hat{\theta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_p)$ and

(b) for testing local influence on single parameter, $\hat{\theta}^T = (\hat{\theta}_0, \hat{\beta}_j, \ldots, \hat{\theta}_p)$ evaluated at ω.

The concept of the proposed test procedure is that we replace the estimates of the postulated model to the perturbed model to see whether these estimates are equal or not on the basis of likelihood ratio test statistic. The likelihood ratio test statistic for testing the significance of influence on single parameter estimate $\hat{\theta}_1$ is then can be defined as

$$\lambda = -2 \ln \left[\frac{\ell(\hat{\theta}_1, \hat{\theta}_2)}{\ell(\hat{\beta})} \right] \sim \chi^2 \text{with 1 d.f.}$$

6 Illustration

For testing purpose, we fit a logistic regression model to the Framingham Heart Study data set. A random sample of 200 out of 669 men have been taken for the computational convenience and Fortran programming language has been used for analyzing the data. In our study we use the variables age, systolic blood pressure (SBP), diastolic blood pressure (DBP), serum cholesterol (CHOL), Farmingham relative weight (FRW) and cigarette (CIG) as independent variables and coronary heart disease (CHD) as outcome variable. First we fit a logistic regression model and the estimated coefficients are given below:

Variables	Intercept	AGE(β_1)	SBP(β_2)	DBP(β_3)	CHOL (β_4)	FRW (β_5)	CIG (β_6)
Estimated Coefficients	-7.9116	0.0113	0.00108	0.0389	0.011	0.0013	-0.0102

Then we fit logistic model for different perturbation schemes and local influence diagnostics have been used to assess local influence and to detect the influential elements. In the local influence analysis the case 52 and explanatory variable CIG has been found very influential.

For testing the significance of local influence on the parameter estimates, here we consider the estimates for the perturbation of case-weights ($\omega_{52} = 0.8$), explanatory variable cigarette (CIG) and for the deletion of case 52. In Table 1, 2 and 3, the null hypotheses are replaced in the perturbed model. Then the estimates of the 3rd model (see section 5) obtained are listed in column 3. Column 4 presents the value of conditional log-likelihood and column 5 presents corresponding χ^2 value. For the χ^2 value, the unconditional log-likelihood is calculated $\text{Ł}(\hat{\beta}) = -103.64$. Finally column 6 presents corresponding p-value.

6.1 Case-weight Perturbations

To test the significance of local influence on the parameter estimates for case-weight perturbations, here we consider only the minor modification of the case 52 i.e., the case 52 has given weight 0.8 and the local influence on the parameter estimates are assessed (Table-1).

The null hypotheses 1, 2, 3, 4 and 5 in Table-1 may be rejected since the p-values are moderately small and we may conclude that there are significant influences on the estimates of the parameters β_1, β_2, β_3, β_4 and β_5 for the perturbations of case 52. Under the null hypothesis of equality of parameters, the chi-square values have found negative for testing null hypothesis regarding β_6 and overall influence. The absolute value is considered in such cases because the log-likelihood values can be interchanged. Since the p-value is high (0.599), the null hypothesis 6 may not be rejected and we may conclude that there is no significant influence on β_6. Similarly we may conclude that there is no overall significance of local influence on the estimates for the case-weight perturbation introduced. However, although $\hat{\beta}$ (entire parameter vector) is insensitive to ω, one can begin to check for influences on different subsets of $\hat{\beta}$.

6.2 Perturbations of Explanatory Variable

In Table-2 we consider the testing procedure of the significance of influence on the parameter estimates for the perturbation of explanatory variable CIG. For perturbing the explanatory variable CIG, the elements of the perturbation vector ω of the cases 41, 47, 52, 54, 64, 98, 117, 125, 171, 190, 193, 197 are so chosen that the value of CIG of the above cases have been transformed into 5. Thus we assess the local influence and effect of perturbation is tested (Table-2).

In Table-2, all the null hypotheses may be accepted and so thus there is no significant influence on the parameter estimates for the modification of explanatory variable CIG which is also indicated in the local influence analysis of Hossain (1997) where all influential elements have lost their influential nature for such modifications. Hence the test procedure successfully tests the significance of local influence on the parameter estimates for the perturbation of explanatory variable CIG.

6.3 Case-deletion Effect

Since we see that the case 52 has had greater influence for the minor modifications of the data, in Table-3 we make an attempt to observe the effect for the deletion of case 52 on the parameter estimates. The p-values corresponding to the null hypotheses 1 to 6 are very much small and so the null hypotheses are rejected. Hence the deletion of case 52 makes a substantial changes on the parameter estimates. This is also illustrated in the local influence analysis of Hossain (1997) and our proposed test procedure absolutely support his result. But for testing the overall significance, the test procedure implies that there is no overall significant influences on the parameter estimates for the deletion of the case 52.

7 Discussion and Conclusion

In this paper we propose a simple likelihood ratio based test procedure for testing the significance of local influence on the parameter estimates which is developed on the ideology of testing the equality of parameters of different models. As a particular case, this test procedure is used for testing the significance of local influence on the parameter estimates of logistic regression model and it can be extended to the model having smooth and well-behaved likelihood and perturbation functions. Here it is investigated that although the whole parameter vector is insensitive to the perturbation scheme introduced, it might be sensitive to a single parameter or a sub-set of parameter vector.

The proposed test procedure gives a measure of local influence on the parameter estimates which could not be possible by the usual local influence diagnostics or test procedures. The test procedure is illustrated by fitting the logistic regression models to the FHS data set under specific perturbation schemes.

Table 1: Newly estimated parameters, χ^2 value and corresponding p-value for testing the null hypotheses $\beta_j = \beta_{\omega j}$ when case-weights perturbations ($\omega_{52} = 0.8$) are considered

Serial No.	Null Hypotheses	$\hat{\theta}_0$	Newly estimated parameters of the perturbed model $\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\theta}_4$	$\hat{\theta}_5$	$\hat{\theta}_6$	Conditional log-likelihood $L(\hat{\theta})$	χ^2 value (λ)	p-value
1	$H_0 : \beta_1 = \beta_{\omega 1}$ ($\hat{\beta}_1 = 0.0113$)	-7.63	0.0113*	0.0012	0.037	0.0109	0.0011	-0.031	-105.054	2.82	0.093
2	$H_0 : \beta_2 = \beta_{\omega 2}$ ($\hat{\beta}_2 = 0.001$)	-7.16	0.0031	0.0010*	0.037	0.0108	0.0014	-0.032	-105.027	2.76	0.096
3	$H_0 : \beta_3 = \beta_{\omega 3}$ ($\hat{\beta}_3 = 0.038$)	-7.22	0.003	0.0014	0.038*	0.0108	0.0012	-0.032	-105.015	2.75	0.097
4	$H_0 : \beta_4 = \beta_{\omega 4}$ ($\hat{\beta}_4 = 0.0109$)	-7.14	0.0017	0.0021	0.036	0.0109*	0.0015	-0.032	-105.036	2.79	0.094
5	$H_0 : \beta_5 = \beta_{\omega 5}$ ($\hat{\beta}_5 = 0.0013$)	-7.12	0.0018	0.0021	0.036	0.0108	0.0013*	-0.032	-105.0366	2.79	0.094
6	$H_0 : \beta_6 = \beta_{\omega 6}$ ($\hat{\beta}_6 = -0.0102$)	-7.88	0.0139	0.0011	0.038	0.0105	0.0009	-0.0102*	-103.505	-0.28 (0.28)	0.596
	For testing overall significance $\hat{\theta}_0 = -7.83$								-103.498	-0.28 (0.28)	0.999

$*$ Replaced estimates.

Table 2: Newly estimated parameters, χ^2 value and corresponding p-value for testing the null hypotheses $\beta_j = \beta_{\omega j}$ when perturbations of CIG are considered

Serial No.	Null Hypotheses	$\hat{\theta}_0$	Newly estimated parameters of the perturbed model $\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\theta}_4$	$\hat{\theta}_5$	$\hat{\theta}_6$	Conditional log-likelihood $L(\hat{\theta})$	χ^2 value (λ)	p-value
1	$H_0 : \beta_1 = \beta_{\omega 1}$ ($\hat{\beta}_1 = 0.0113$)	-7.86	0.0113*	0.0013	0.038	0.011	0.0007	-0.011	-103.58	-0.12	0.72
2	$H_0 : \beta_2 = \beta_{\omega 2}$ ($\hat{\beta}_2 = 0.001$)	-7.87	0.0117	0.0010*	0.038	0.011	0.0006	-0.0112	-103.64	0.02	0.88
3	$H_0 : \beta_3 = \beta_{\omega 3}$ ($\hat{\beta}_3 = 0.038$)	-7.84	0.011	0.0015	0.038*	0.011	0.0008	-0.0113	-103.50	-0.28 (0.28)	0.59
4	$H_0 : \beta_4 = \beta_{\omega 4}$ ($\hat{\beta}_4 = 0.0109$)	-7.83	0.0114	0.0012	0.038	0.0109*	0.0006	-0.011	-103.49	-0.29 (0.29)	0.59
5	$H_0 : \beta_5 = \beta_{\omega 5}$ ($\hat{\beta}_5 = 0.0013$)	-7.91	0.0111	0.0013	0.038	0.011	0.0013*	-0.011	-103.50	-0.28 (0.28)	0.59
6	$H_0 : \beta_6 = \beta_{\omega 6}$ ($\hat{\beta}_6 = -0.0102$)	-7.91	0.0123	0.0012	0.038	0.0109	0.0006	-0.0102*	-103.505	-0.28 (0.28)	0.59
	For testing overall significance $\hat{\theta}_0 = -7.82$								-103.49	-0.28 (0.28)	0.99

$*$ Replaced estimates.

Table 3: Newly estimated parameters, χ^2 value and corresponding p-value for testing the null hypotheses $\beta_j = \beta_{\omega j}$ when the most influential case is deleted

Serial No.	Null Hypotheses	Newly estimated parameters of the perturbed model $\hat{\theta}_0$	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\theta}_4$	$\hat{\theta}_5$	$\hat{\theta}_6$	Conditional log-likelihood $L(\hat{\theta})$	χ^2 value (λ)	p-value
1	$H_0 : \beta_1 = \beta_{\omega 1}$ ($\hat{\beta}_1 = 0.0113$)	-5.95	0.0113*	0.0072	0.05	0.006	-0.0062	-0.0124	-328.104	448.9	1.25e-99
2	$H_0 : \beta_2 = \beta_{\omega 2}$ ($\hat{\beta}_2 = 0.001$)	-6.74	0.018	0.0010*	0.039	0.007	-0.0056	-0.0112	-129.68	52.04	5.32e-13
3	$H_0 : \beta_3 = \beta_{\omega 3}$ ($\hat{\beta}_3 = 0.038$)	-6.28	0.0208	-0.0028	0.038*	0.007	-0.0038	-0.0111	-186.423	165.56	6.9e-38
4	$H_0 : \beta_4 = \beta_{\omega 4}$ ($\hat{\beta}_4 = 0.0109$)	-8.11	0.0307	-0.0085	0.052	0.0109*	-0.0049	-0.0125	-339.09	470.9	2.04e-104
5	$H_0 : \beta_5 = \beta_{\omega 5}$ ($\hat{\beta}_5 = 0.0013$)	-7.49	0.0269	-0.008	0.049	0.007	0.0013*	-0.0113	-213.63	219.99	9.09e-50
6	$H_0 : \beta_6 = \beta_{\omega 6}$ ($\hat{\beta}_6 = -0.0102$)	-6.95	0.0312	-0.009	0.053	0.0068	0.0072	-0.0102*	-421.18	635.08	3.92e-140
	For testing overall significance $\hat{\theta}_0 = -7.83$								-103.49	0.28	0.999

$*$ Replaced estimates.

References

[1] Cook, R.D. (1986). Assessment of local influence. *Journal of Royal Statistical Society*, Series B, 48, 133-169.

[2] Cook, R. D. and Weissberg, S. (1982). Residual and Influence in Regression. New York & London. Chapman and Hall.

[3] Escobar, L. A. and Meeker, W. Q. (1992). Assessing influence in regression analysis with censored data. *Biometrics*, 48, 507-528.

[4] Hosmer, D. W. and Lemeshow, S. (1989). Applied logistic regression. John Wiley and Sons, New York.

[5] Hossain, M. (1997). Extension of diagnostics for assessing local influence in logistic regression models and some test procedures. Unpublished M.Sc. thesis, Department. of Statistics, University of Dhaka.

[6] Islam, M. A. (1994). Multistate survival models for transitions and reverse transitions: An application to contraceptive use data. *Journal of Royal Statistical Society*, Series A, 157, Part 3, 441-455.

[7] Kahn, H. A. and Sempos, C. D. (1989). Statistical methods in epidemiology. Oxford University Press, New York.

[8] Pregibon, D. (1979). Data analytic methods for generalized linear models. Unpublished Ph.D. Thesis, University of Toronto.

[9] Pregibon, D. (1981). Logistic regression diagnostics. *Annals of Statistics*, 9, 705-724.

[10] Thomas, W. and Cook, R. D. (1989). Assessing influence on regression coefficients in generalized linear model. *Biometrika*, 76, 741-749.

[11] Weissfeld, L. A. and Schneider, H. (1990a). Influence diagnostics for normal linear model with censored data. *Australian Journal of Statistics*, 32, 11-20.

[12] Weissfeld, L. A. and Schneider, H. (1990b). Influence diagnostics for the Weibull model fit to censored data. *Statistics and Probability Letters*, 9, 67-73.

[13] Weissfeld, L. A. (1990). Influence diagnostics for the proportional hazards models. *Statistics and Probability Letters*, 10, 411-417.