Sankhyā: The Indian Journal of Statistics
1995, Volume 57, Series A, Pt. 3, pp. 393-409

A NEW APPROACH TO INFLUENCE ANALYSIS IN LINEAR MODELS

By J. MUÑOZ-PICHARDO
J. MUÑOZ-GARCIA
J.L. MORENO-REBOLLO
and
R. PINO-MEJIAS
Universidad de Sevilla

SUMMARY. We propose a new approach to the study of influence in the General Linear Model based on conditional bias. This approach enables us to apply such an analysis to all particular cases of this model. The theoretical foundation, on which this approach is based, does not presuppose a particular hypothesis on the distribution of the variables. Applying the results obtained to the Multiple Linear Regression Model, measures of influence are obtained as already proposed by other authors. Finally we carry out an application of the results on the analysis of covariance.

1. INTRODUCTION

Influence Analysis arises when one is faced with the need to analyze the sensitivity of the results of a statistical analysis. This need is due to the fact that the conclusions of any analysis are drawn from methods based on sample observations and on asumptions on the underlying models of an experimentation.

On the whole, in order to measure the influence that one or a set of observations has on the statistical analysis, slight disturbances are introduced into the model and the changes produced are quantified. In the literature, the most common disturbance pattern is that of the omission of the observations whose influence is to be studied.

Cook (1987) tried to unify the problem under a general formulation which is valid for the different approaches and methods carried out on it. In this formulation, the key questions are the choice of the perturbation pattern and the choice of the model of comparison of the results obtained under the postulated model and the disturbed model.

Paper received. August 1993; revised February 1994.
AMS (1980) (revised 1985) *subject classification.* 62J99.
Key words and phrases. Conditional bias, general linear model, influence observations.

On the other hand, Influence Analysis has developed mainly in the Multiple Linear Regression Model (Belsley, Kuh, Welsch, 1980; Cook and Weisberg, 1982; Cook 1986, 1987; Chatterjee and Hadi, 1986, 1988) and, by extension, in the Multivariate Regression Model (Caroni, 1987; Barret and Ling, 1992).

In this paper, we propose a new approach to study the influence on the General Linear Model (Kshirsagar, 1983) based on **conditional bias**. This approach permits us to apply such an analysis to all particular cases of this model. The theoretical foundation, on which this approach is based, does not presuppose a particular hypothesis on the distribution of the variables.

As a particular case, by moving the results obtained from the General Linear Model to the Multiple Linear Regression Model, we obtain measures of influence that have been already proposed by other authors (Belsley *et al.* (1980); Atkinson (1982); Cook and Weisberg (1982);). Finally we carry out an application of the results on the analysis of covariance.

2. CONDITIONAL BIAS IN THE GENERAL LINEAR MODEL

2.1 *The general linear model : Notation.* The study of the General Linear Model is of great importance within statistics mainly due to its wide range of applications. It is used as a model for the development of other models, such as the multiple regression, experimental designs, analysis of covariance, etc.

The general linear model is defined by

$$\mathbf{Y} = \mathbf{X}\beta + \epsilon, \quad E[\boldsymbol{\epsilon}] = 0, \; Var(\epsilon) = \sigma^2 I_n \quad \text{(GLM)}$$

where $\mathbf{Y}$ is a random n-vector; $\mathbf{X}$ is a known $n \times p$ matrix with rank r ($r \leq p < n$); $\boldsymbol{\beta}$ is a p-vector of unknown parameters and $\boldsymbol{\epsilon}$ is an n-vector which represents the non-observable random perturbations.

The fitted values vector is denoted by $\widehat{\mathbf{Y}} = \mathbf{V}\mathbf{Y}$, with $\mathbf{V} = \mathbf{X}\mathbf{S}^{-}\mathbf{X}' = ((v_{ij}))_{i,j=1,\ldots,n}$ the prediction matrix, which is symmetric, idempotent, with rank r and unique for any $\mathbf{S}^{-}$ generalized inverse of $\mathbf{S} = \mathbf{X}'\mathbf{X}$ (Kshirsagar, 1983). The residuals vector is represented by $\mathbf{e} = \mathbf{Y} - \widehat{\mathbf{Y}} = \mathbf{M}\epsilon$, where $\mathbf{M} = \mathbf{I}_n - \mathbf{V}$, and the least squares estimators of $\boldsymbol{\beta}$ and σ^2 are denoted by $\widehat{\boldsymbol{\beta}}$ and $\widehat{\sigma}^2$, respectively. On the other hand, the estimable linear parametric functions of β and $\Lambda\beta$, where Λ is a $q \times p$ matrix, with rank $(\Lambda) = q$, so that $\Lambda\boldsymbol{\beta}$ is the BLUE vector of the components of $\Lambda\boldsymbol{\beta}$.

Finally, given a collection of subindexes $I = \{i_1, \ldots, i_m\} \subset \{1, \ldots, n\}$, the m-vector composed of the components of $\mathbf{Y}$ subindicated by I is denoted by $\mathbf{Y}_I$. Similarly, the matrix formed by the rows of $\mathbf{X}$ corresponding to the collection I is denoted by $\mathbf{X}_I$, and the submatrix corresponding to $\mathbf{V}$ is represented by $\mathbf{V}_I = \mathbf{X}_I\mathbf{S}^{-}\mathbf{X}_I'$. Likewise, the omission of the observations indexed by I is indicated by the subindex (I).

2.2. *Conditional bias.* The Decomposition Lemma from Efron and Stein (1981) expresses a statistic T defined on a simple random sample $Y_1, \ldots, Y_n$ as

a finite sum, whose terms are given according to the conditional expectations of the statistic T given the sample observations, which can be considered either individually or together. Taking this lemma as a reference, the following definitions are given.

Definition 2.1. Let $Y_1, \ldots, Y_n$ be a random sample of the random variable Y, let $T = T(Y_1, \ldots, Y_n)$ be a statistic defined on the sample, and let $y_1, \ldots y_n$ be a realization of the sample. The **conditional bias of T given the i-th observations** is defined as

$$\mathcal{J}[y_1; T] = E[T \mid Y_i = y_i] - E[T] \qquad \ldots(2.1)$$

The conditional bias $\mathcal{J}[y_i; T]$ can be interpreted as the average effect that the realization of the i-th sample observation has on the statistic T. Therefore, it can be considered as a measure of the influence that such a sample realization has on T.

By generalizing this, we can define the conditional bias given a set of observations.

Definition 2.2. Under the conditions of Defintion 2.1, the **conditional bias of T given the set of observations** $\{y_{i_1}, \ldots, y_{i_m}\}$ is defined as

$$\mathcal{J}[y_{i_1}, \ldots, y_{i_m}; T] = E[T \mid y_{i_1} = y_{i_1}, \ldots, y_{i_m} = y_{i_m}] - E[T] \qquad \ldots(2.2)$$

Therefore, it is considered to be a measure of the joint influence of observations $\{y_{i_1}, \ldots, y_{i_m}\}$ on T.

2.3. *Conditional bias of the general linear model estimators.* We shall now calculate the conditional bias for the BLUE of an estimate linear function of the parameter vector and for the estimator $\widehat{\sigma}^2$.

For $\mathbf{Y}_I$, the conditional bias is denoted by

$$\mathcal{J}[y_{i_1}, \ldots, y_{i_m} : T] = \mathcal{J}[y_I; T] = E[T \mid \mathbf{Y}_I = \mathbf{y}_I] - E[T] \qquad \ldots(2.3)$$

Without loss of generality, we assume that these m observations are the last observations. Thus, the following decompositions are obtained

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_{(I)} \\ \mathbf{X}_I \end{bmatrix}; \mathbf{Y} = \begin{bmatrix} \mathbf{Y}_{(I)} \\ \mathbf{Y}_I \end{bmatrix}; \mathbf{V} = \begin{bmatrix} \mathbf{V}_{(I)} & \mathbf{V}_0 \\ \mathbf{V}_0' & \mathbf{V}_I \end{bmatrix}; \mathbf{I}_n - \mathbf{V} = \begin{bmatrix} \mathbf{M}_{(I)} & \mathbf{M}_0 \\ \mathbf{M}_0' & \mathbf{M}_i \end{bmatrix} \qquad \ldots(2.4)$$

where $\mathbf{V}_0 = \mathbf{X}_{(I)}\mathbf{S}^{-}\mathbf{X}_I'$, $\mathbf{M}_{(I)} = \mathbf{I}_{n-m} - \mathbf{V}_{(I)}$, $\mathbf{M}_I = \mathbf{I}_m - \mathbf{V}_I$ and $\mathbf{M}_0 = -\mathbf{V}_0$. We assume that $\mathbf{V}_I$ and $\mathbf{M}_I$ are non-singular and that $n > r + m$.

In the following theorem, we obtain the conditional bias for the BLUE of an estimable linear function.

Theorem 2.3. *Let $\Lambda\boldsymbol{\beta}$ be an estimable linear function, with Λ being a $q \times p$ matrix with rank $(\Lambda) = q$. Then*

$$\mathcal{J}\left[\mathbf{y}_I; \Lambda\widehat{\boldsymbol{\beta}}\right] = \Lambda\mathbf{S}^{-}\mathbf{X}_I'\left[\mathbf{y}_I - \mathbf{X}_I\beta\right] \qquad \ldots(2.5)$$

Proof. $E\left[\Lambda\widehat{\boldsymbol{\beta}} \mid \mathbf{Y}_I = \mathbf{y}_I\right] - E\left[\Lambda\widehat{\boldsymbol{\beta}}\right]$

$= \Lambda\mathbf{S}^{-}\left\{E\left[\mathbf{X}_I'\mathbf{Y}_I \mid \mathbf{Y}_I = \mathbf{y}_I\right] + E\left[\mathbf{X}_{(I)}'\mathbf{Y}_{(I)} \mid \mathbf{Y}_I - \mathbf{y}_I\right]\right\} - \Lambda\boldsymbol{\beta}$

$= \Lambda\mathbf{S}^{-}\left[\mathbf{X}_I'\mathbf{y}_I + \mathbf{X}_{(I)}'\mathbf{X}_{(I)}\boldsymbol{\beta}\right] - \Lambda\boldsymbol{\beta}$

$= \Lambda\mathbf{S}^{-}\left[\mathbf{X}'\mathbf{X}\boldsymbol{\beta} + \mathbf{X}_I'(\mathbf{y}_I - \mathbf{X}_I\boldsymbol{\beta})\right] - \Lambda\boldsymbol{\beta} = \Lambda\mathbf{S}^{-}\mathbf{X}_I'(\mathbf{y}_I - \mathbf{X}_I\boldsymbol{\beta})$ ▱

Corollary 2.4. *For $m = 1$ and $I = \{i\}$,*

$$\mathcal{J}\left[y_i; \Lambda\boldsymbol{\beta}\right] = \Lambda\mathbf{S}^{-}\mathbf{x}_i'\left[\mathbf{y}_i - \mathbf{x}_i\beta\right] \qquad \ldots(2.6)$$

where $\mathbf{x}_i$ is the i-th row of $\mathbf{X}$.

The next lemma is useful to obtain the conditional bias for $\widehat{\sigma}^2$.

Lemma 2.5. *Given the partitioned matrices (2.4), we obtain*

$$\mathbf{M}_0'\mathbf{X}_{(I)} = -[\mathbf{I}_m - \mathbf{V}_I]\mathbf{X}_I \qquad \ldots(2.7)$$

Proof. As $\mathbf{V}\,\mathbf{X} = \mathbf{X}$, then $\mathbf{M}\,\mathbf{X} = \Theta$, with Θ being the null matrix. Therefore $\mathbf{M}_0'\mathbf{X}_{(I)} = -\mathbf{M}_I\mathbf{X}_I = -[\mathbf{I}_m - \mathbf{V}_I]\mathbf{X}_I$

▱

Theorem 2.6. *The conditional bias of $\widehat{\sigma}^2$ given the set of sample observations indexed by I is*

$$\mathcal{J}\left[\mathbf{y}_I; \widehat{\sigma}^2\right] = \frac{1}{n-r}\left\{\left[\mathbf{y}_I - \mathbf{X}_I\boldsymbol{\beta}\right]'\left[\mathbf{I}_m - \mathbf{V}_I\right]\left[\mathbf{y}_I - \mathbf{X}_I\boldsymbol{\beta}\right] - \sigma^2 tr[\mathbf{I}_m - \mathbf{V}_I]\right\} \qquad \ldots(2.8)$$

Proof. The unbiased estimator of σ^2 can be expressed in the following way

$$\widehat{\sigma}^2 = \frac{1}{n-r}\left[\mathbf{Y}_{(I)}'\mathbf{M}_{(I)}\mathbf{Y}_{(I)} + 2\mathbf{Y}_I'\mathbf{M}_0'\mathbf{Y}_{(I)} + Y_I'\mathbf{M}_I Y_I\right] \qquad \ldots(2.9)$$

Besides

$E\left[\mathbf{Y}_I'\mathbf{M}_I\mathbf{Y}_I \mid \mathbf{Y}_I = \mathbf{y}_I\right] = \mathbf{y}_I'\left[\mathbf{I}_m - \mathbf{V}_I\right]y_I,$

$E\left[\mathbf{Y}_I'\mathbf{M}_0'\mathbf{Y}_{(I)} \mid \mathbf{Y}_I = \mathbf{y}_I\right] = \mathbf{y}_I'\mathbf{M}_0'\mathbf{X}_{(I)}\boldsymbol{\beta} = -\mathbf{y}_I'[\mathbf{I}_m - \mathbf{V}_I]\mathbf{X}_I\boldsymbol{\beta}$

and

$$E\left[\mathbf{Y}_{(I)}'\mathbf{M}_{(I)}\mathbf{Y}_{(I)} \mid \mathbf{Y}_I = \mathbf{y}_I\right] = E\left[\mathbf{Y}_{(I)}'\mathbf{M}_{(I)}\mathbf{Y}_{(I)}\right]$$

$$= \sum_{k\notin I}(1 - v_{kk})E[Y_k^2] - \sum_{k\notin I}\sum_{s\notin I\ k\neq s} v_{ks}E[Y_kY_s] = \sigma^2\left[\sum_{k\notin I}(1 - v_{kk})\right]$$

$$+ \sum_{k\notin I}(1 - v_{kk}))\mathbf{x}_k\boldsymbol{\beta})^2 - \sum_{k\notin I}\left[\sum_{s\notin I\ s\neq k} v_{ks}(\mathbf{x}_s\beta)\right](\mathbf{x}_k\boldsymbol{\beta})$$

$$= \sigma^2 \{n - r - tr [\mathbf{I}_m - \mathbf{V}_I]\} + \sum_{k \notin I} (1 - v_{kk})(\mathbf{x}_k \boldsymbol{\beta})^2$$

$$- \sum_{k \notin I} \left[(1 - v_{kk})(\mathbf{x}_k \boldsymbol{\beta}) - \sum_{j \in I} v_{kj} (\mathbf{x}_j \boldsymbol{\beta}) \right] (\mathbf{x}_k \boldsymbol{\beta})$$

$$= \sigma^2 \{n - r - tr [\mathbf{I}_m - \mathbf{V}_I]\} + \sum_{j \in I} \left[\sum_{k \notin I} v_{jk} (\mathbf{x}_k \boldsymbol{\beta}) \right] (\mathbf{x}_j \boldsymbol{\beta})$$

$$= \sigma^2 \{n - r - tr [\mathbf{I}_m - \mathbf{V}_I]\} + \sum_{j \in I} \left[(1 - v_{jj})(\mathbf{x}_j \boldsymbol{\beta}) + \sum_{t \in I \; s \neq j} v_{jt} (\mathbf{x}_t \beta) \right] (\mathbf{x}_j \boldsymbol{\beta})$$

$$= \sigma^2 \{n - r - tr [\mathbf{I}_m - \mathbf{V}_I]\} + [\mathbf{X}_I \boldsymbol{\beta}]' [\mathbf{I}_m - \mathbf{V}_I] [\mathbf{X}_I \beta]$$

Therefore, the following is obtained

$$E [\widehat{\sigma}^2 \mid \mathbf{Y}_I = \mathbf{y}_I] = \sigma^2 \left\{ 1 - \frac{1}{n - r} tr [\mathbf{I}_m - \mathbf{V}_I] \right\}$$

$$+ \frac{1}{n - r} [\mathbf{y}_I - \mathbf{X}_I \boldsymbol{\beta}]' [\mathbf{I}_m - \mathbf{V}_I] [\mathbf{y}_I - \mathbf{X}_I \boldsymbol{\beta}]$$

from which we can deduce the proposed result.

Corollary 2.7. *For $m = 1$ and $I = \{i\}$,*

$$\mathcal{J} [y_I; \widehat{\sigma}^2] = \frac{1}{n - r} [1 - v_{ii}] \left\{ [y_i - \mathbf{x}_i \boldsymbol{\beta}]^2 - \sigma^2 \right\} \qquad \ldots (2.10)$$

2.4. *An estimation of the conditional BIAS.* In the results previously obtained we observe that the conditional bias depends on the unknown parameter $\boldsymbol{\beta}$ and σ^2. The following theorem offers an estimation of the conditional bias of an estimator $\widehat{\mathcal{V}} = \widehat{\mathcal{V}}(Y_1, \ldots, Y_n)$ of a vector θ of parameters.

Theorem 2.8. *Let $Y_1, \ldots, Y_n$ be a random sample of the random variable Y, whose distribution depends on an unknown parameter $\theta \in \Omega \subseteq \Re^p$. Let $\widehat{\mathcal{V}} = \widehat{\mathcal{V}}(Y_1, \ldots, Y_n)$ be an unbiased estimator of θ and let $\widehat{\mathcal{V}}(i_1, \ldots, i_m)$ be the unbiased estimator obtained from the sample with the omission of the observations $\{Y_k \mid k \in \{i_1, \ldots, i_m\}\}$. Then*

$$E \left[\widehat{\mathcal{V}} - \widehat{\mathcal{V}}_{(i_1, \ldots, i_m)} \mid Y_{i_1} = y_{i_1}, \ldots, Y_{i_m} = y_{i_m} \right] = \mathcal{J} \left[y_{i_1}, \ldots, y_{i_m}; \widehat{\mathcal{V}} \right] \qquad \ldots (2.11)$$

The result follows directly from the fact that $\widehat{\mathcal{V}}_{(i_1, \ldots, i_m)}$ is unbiased and does not depend on $Y_{i_1}, \ldots, Y_{i_m}$. Hence, we denote $\widehat{\mathcal{V}} - \widehat{\mathcal{V}}_{(i_1, \ldots, i_m)}$ by $\widehat{\mathcal{J}} \left[y_{i_1}, \ldots, y_{i_m}; \widehat{\mathcal{V}} \right]$.

In order to estimate the conditional bias either of the BLUE of an estimable linear parametric function for the model (GLM), or of $\widehat{\sigma}^2$, we must obtain the estimators corresponding to the model in which the observations indicated by I are omitted.

The resulting linear model with such observations omitted is

$$\mathbf{Y}_{(I)} = \mathbf{X}_{(I)}\boldsymbol{\beta} + \boldsymbol{\epsilon}_{(I)} \qquad [\mathrm{GLM}(I)]$$

where $\mathbf{Y}_{(I)}, \mathbf{X}_{(I)}, \boldsymbol{\epsilon}_{(I)}$ are the corresponding subvectors or submatrices associated with the set of subindexes I. For the model, the least squares estimators are expressed as

$$\widehat{\boldsymbol{\beta}}_{(I)} = \mathbf{S}_{(I)}^{-}\mathbf{X}'_{(I)}\mathbf{Y}_{(I)} + [\mathbf{I}_k - \mathbf{H}_{(I)}]\boldsymbol{\nu}, \quad \boldsymbol{\nu}\epsilon\Re^p \quad \text{arbitrary} \qquad \ldots(2.12)$$

with $\mathbf{S}_{(I)} = \mathbf{X}'_{(I)}\mathbf{X}_{(I)}$ and $\mathbf{H}_{(I)} = \mathbf{S}_{(I)}^{-}\mathbf{S}_{(I)}$,

Theorem 2.9. *If $\Lambda\boldsymbol{\beta}$ is an estimable linear function for (GLM), it is also for the model* [GLM(I)]

Proof. We consider the following decomposition

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_I \\ \mathbf{X}_{(I)} \end{bmatrix} \quad \mathbf{V} = \begin{bmatrix} \mathbf{V}_I & \mathbf{V}_0 \\ \mathbf{V}_0' & \mathbf{V}_{(I)} \end{bmatrix}$$

As $\mathbf{VX} = \mathbf{X}$, then we obtain $\mathbf{V}_I\mathbf{X}_I + \mathbf{V}_0\mathbf{X}_{(I)} = \mathbf{X}_I$. Likewise, if $\Lambda\boldsymbol{\beta}$ is estimable for the model (GLM), there exists a $q \times n$ matrix $\mathbf{A}$, such that

$$\Lambda = \mathbf{A}\begin{bmatrix} \mathbf{M}_I^{-1}\mathbf{V}_0 \\ \mathbf{I}_m \end{bmatrix}\mathbf{X}_{(I)} \Rightarrow \Lambda = \mathbf{B}\mathbf{X}_{(I)} \text{ where } \mathbf{B} = \mathbf{A}\begin{bmatrix} \mathbf{M}_I^{-1}\mathbf{V}_0 \\ \mathbf{I}_m \end{bmatrix}$$

Then $\Lambda\boldsymbol{\beta}$ is an estimable linear function for the model [GLM(I)].

As a consequence, if $\Lambda\boldsymbol{\beta}$ is an estimable linear function for (GLM), then the BLUE for [GLM(I)] is $\Lambda\widehat{\boldsymbol{\beta}}_{(I)}$. The following results, Lemma 2.10 and Theorem 2.11, allow us to obtain the estimation of the conditional bias of $\Lambda\widehat{\boldsymbol{\beta}}$.

Lemma 2.10. *Let* $\mathbf{A}$ *be a* $p \times q$ *matrix and* L *be a* $m \times p$ *matrix* ($m \leq p$), *such that* L *belongs to the row space generated by the matrix* $\mathbf{A}$, *and the inverse of the matrix* [$\mathbf{I}$ - $\mathbf{L}$ $\mathbf{A}^{-}\mathbf{L}'$] *exists. Therefore, the matrices*

$$\left\{\mathbf{A}^{-} - \mathbf{A}^{-}\mathbf{L}'[\mathbf{I} - \mathbf{L}\mathbf{A}^{-}\mathbf{L}']^{-1}\mathbf{L}\mathbf{A}^{-}\right\}$$

are generalized inverses of [$\mathbf{A}$ - $\mathbf{L}$ $\mathbf{L}'$].

Proof. It suffices to base the proof on Pringle and Rayner (1971), where the equalities $\mathbf{L}\mathbf{A}^{-}\mathbf{A} = \mathbf{L}$ and $\mathbf{A}\mathbf{A}'\mathbf{L}' = \mathbf{L}'$ verify if $\mathbf{L}$ belongs to the row space of $\mathbf{A}$.

Theorem 2.11. *The least squares estimators of* $\boldsymbol{\beta}$ *in [GLM(I)] can be expressed as*

$$\widehat{\boldsymbol{\beta}}_{(I)} = \mathbf{S}^{-}\mathbf{X}'\mathbf{Y} - \mathbf{S}^{-}\mathbf{X}'_I\mathbf{M}_I^{-1}\mathbf{e}_I + (\mathbf{I}_k - \mathbf{H})\boldsymbol{\nu}, \quad \boldsymbol{\nu}\epsilon\Re^p \text{ arbitrary.} \qquad \ldots(2.13)$$

Proof. From Lemma 2.10, we obtain

$$\mathbf{S}_{(I)}^{-} = (\mathbf{S} - \mathbf{X}'_I\mathbf{X}_I)^{-} = \mathbf{S}^{-} + \mathbf{S}^{-}\mathbf{X}'_I\mathbf{M}_I^{-1}\mathbf{X}_I\mathbf{S}^{-}$$

on the other hand

$$\mathbf{H}_{(I)} = \mathbf{H} - \mathbf{S}^{-}\mathbf{X}_I'\mathbf{X}_I + \mathbf{S}^{-}\mathbf{X}_I'\mathbf{M}_I^{-1}\mathbf{X}_I - \mathbf{S}^{-}\mathbf{X}_I'\mathbf{M}_I^{-1}\mathbf{V}_I\mathbf{X}_I = \mathbf{H}$$

Therefore, based on (2.12), we obtain

$$\begin{aligned}\boldsymbol{\beta}_{(I)} &= (\mathbf{S}^{-} + \mathbf{S}^{-}\mathbf{X}_I'\mathbf{M}_I^{-1}\mathbf{X}_I\mathbf{S}^{-})(\mathbf{X}'\mathbf{Y} - \mathbf{X}_I'\mathbf{Y}_I) + (\mathbf{I}_k - \mathbf{H})\boldsymbol{\nu} \\ &= \mathbf{S}^{-}\mathbf{X}'\mathbf{Y} - \mathbf{S}^{-}\mathbf{X}_I'\mathbf{M}_I^{-1}(\mathbf{Y}_I - \mathbf{X}_I\mathbf{S}^{-}\mathbf{X}'\mathbf{Y}) + (\mathbf{I}_k - \mathbf{H})\boldsymbol{\nu} \\ &= \mathbf{S}^{-}\mathbf{X}'\mathbf{Y} - \mathbf{S}^{-}\mathbf{X}_I'\mathbf{M}_I^{-1}\mathbf{e}_I + (\mathbf{I}_k - \mathbf{H})\boldsymbol{\nu}\end{aligned}$$

□

The estimation of the conditional bias of $\Lambda\widehat{\boldsymbol{\beta}}$ is obtained in the following.

Corollary 2.12.

$$\widehat{\mathcal{J}}[\mathbf{y}_I; \Lambda\widehat{\boldsymbol{\beta}}] = \Lambda\mathbf{S}^{-}\mathbf{X}_I'\mathbf{M}_I^{-1}\mathbf{e}_I \qquad \ldots(2.14)$$

In relation to the unbiased estimator of the parameter σ^2, we obtain

$$\widehat{\sigma}^2_{(I)} = \frac{1}{n-r-m}\left[\mathbf{Y}_{(I)} - \mathbf{X}_{(I)}\widehat{\boldsymbol{\beta}}_{(I)}\right]'\left[\mathbf{Y}_{(I)} - \mathbf{X}_{(I)}\widehat{\boldsymbol{\beta}}_{(I)}\right] \qquad \ldots(2.15)$$

The relation between (2.9) and (2.15) is stated below,

Theorem 2.13. *In the model [GLM(I)]*

$$\widehat{\sigma}^2_{(I)} = \frac{1}{n-r-m}\left[(n-r)\widehat{\sigma}^2 - \mathbf{e}_I'\mathbf{M}_I^{-1}\mathbf{e}_I\right] \qquad \ldots(2.16)$$

Proof. $\left[\mathbf{Y}_{(I)} - \mathbf{X}_{(I)}\widehat{\boldsymbol{\beta}}_{(I)}\right]'\left[\mathbf{Y}_{(I)} - \mathbf{X}_{(I)}\widehat{\boldsymbol{\beta}}_{(I)}\right] = \mathbf{Y}_{(I)}'\mathbf{Y}_{(I)} - \mathbf{Y}_{(I)}'\mathbf{X}_{(I)}\widehat{\boldsymbol{\beta}}_{(I)}$

$$\begin{aligned}&= [\mathbf{Y}'\mathbf{Y} - \mathbf{Y}_I'\mathbf{Y}_I] - [\mathbf{Y}'\mathbf{X} - \mathbf{Y}_I'\mathbf{X}_I]\left[\mathbf{S}^{-}\mathbf{X}'\mathbf{Y} - \mathbf{S}^{-}\mathbf{X}_I'\mathbf{M}_I^{-1}\mathbf{e}_I + (\mathbf{I} - \mathbf{H})\boldsymbol{\nu}\right] \\ &= \mathbf{Y}'(\mathbf{I} - \mathbf{V})\mathbf{Y} - \mathbf{Y}_I'\mathbf{Y}_I + \boldsymbol{\beta}\mathbf{X}_I'\mathbf{M}_I^{-1}\mathbf{e}_I + \mathbf{Y}_I'\mathbf{X}_I\boldsymbol{\beta} - \mathbf{Y}_I'\mathbf{V}_I\mathbf{M}_I^{-1}\mathbf{e}_I \\ &= \mathbf{Y}'(\mathbf{I} - \mathbf{V})\mathbf{Y} - \mathbf{Y}_I'\mathbf{e}_I + \boldsymbol{\beta}\mathbf{X}_I'\mathbf{M}_I^{-1}\mathbf{e}_I - \mathbf{Y}_I'\mathbf{V}_I\mathbf{M}_I^{-1}\mathbf{e}_I \\ &= \mathbf{Y}'(\mathbf{I} - \mathbf{V})\mathbf{Y} - \left[\mathbf{Y}_I'[\mathbf{I} + \mathbf{V}_I\mathbf{M}_I^{-1}] - \boldsymbol{\beta}\mathbf{X}_I'\mathbf{M}_I^{-1}\right]\mathbf{e}_I \\ &= \mathbf{Y}'(\mathbf{I} - \mathbf{V})\mathbf{Y} - \left[\mathbf{Y}_I'\mathbf{M}_I^{-1} - \boldsymbol{\beta}\mathbf{X}_I'\mathbf{M}_I^{-1}\right]\mathbf{e}_I = \mathbf{Y}'(\mathbf{I} - \mathbf{V})\mathbf{Y} - \mathbf{e}_I'\mathbf{M}_I^{-1}\mathbf{e}_I \\ &\quad (n-r)\widehat{\sigma}^2 - \mathbf{e}_I'\mathbf{M}_I^{-1}\mathbf{e}_I\end{aligned}$$

from which we can deduce what was previously stated.

□

Consequently, we obtain the estimation of the conditional bias of $\widehat{\sigma}^2$.

Corollary 2.14.

$$\widehat{\mathcal{J}}[y_I; \widehat{\sigma}^2] = \frac{1}{n-r-m}\mathbf{e}_I'M_I^{-1}\mathbf{e}_I - \frac{m}{n-r-m}\widehat{\sigma}^2 \qquad \ldots(2.17)$$

3. INFLUENCE MEASURES

Corollary 2.12 and 2.14 enable us to propose the following influence measures for the estimators $\Lambda\widehat{\beta}$ and $\widehat{\sigma}^2$ in the (GLM).

3.1 *Influence measures on* $\Lambda\widehat{\beta}$. Based on Corollary 2.12,

$$\widehat{\mathcal{J}}[y_I; \Lambda\widehat{\beta}] = \Lambda\mathbf{S}^{-}\mathbf{X}_I'\mathbf{M}_I^{-1}\mathbf{e}_I \in \Re^q$$

can be interpreted as a measure of the influence that the set of observations, subindicated by I, has on $\Lambda\widehat{\beta}$. And,

$$\widehat{\mathcal{J}}[y_i; \Lambda\widehat{\beta}] = \frac{e_i}{1 - V_{ii}}\Lambda\mathbf{S}^{-}x_i' \in \Re^q$$

can be interpreted as a measure of the influence that the i-th observation has on $\Lambda\widehat{\beta}$.

These influence measures on the BLUE of an estimable linear function belong to the q-dimensional real space. So, in order to characterize the influence, a metric of generalized distance type will be applied, according to the characterization given by Barnett (1976).

Hence, given a symmetric, positive definite matrix $\mathbf{Q}$ and a positive scalar c, we define the $(\mathbf{Q}, c)$-norm of a vector X as

$$\|X\|_{(\mathbf{Q},c)} = (1/c)X'\mathbf{Q}X$$

Considering the matrix $\mathbf{Q} = (\Lambda\mathbf{S}^{-}\Lambda')^{-1}$, with Λ a $q \times p$ matrix such that $\Lambda\beta$ is linearly estimable, and an adequate choice of the scalar c, we can establish the following norms.

(I) For $c_1 = q\widehat{\sigma}^2$,

$D_I^*[\Lambda\widehat{\beta}] = \|\mathcal{J}[y_I : \Lambda\widehat{\beta}]\|_{(\mathbf{Q},c_1)}$, called D_I^*-DISTANCE associated with the set of observations y_I.

$D_I[\Lambda\widehat{\beta}] = \|\widehat{\mathcal{J}}[y_I : \Lambda\widehat{\beta}]\|_{(\mathbf{Q},c_1)}$, called D_I-DISTANCE associated with the set of observations y_I.

Easily, we obtain the following expressions for $D_I^*[\Lambda\widehat{\beta}]$ and $D_I[\Lambda\widehat{\beta}]$:

$$D_I^*[\Lambda\widehat{\beta}] = \frac{1}{q\widehat{\sigma}^2}\left[y_I - \mathbf{X}_I\beta\right]'\mathbf{X}_I\mathbf{S}^{-}\Lambda'\left[\Lambda\mathbf{S}^{-}\Lambda'\right]^{-1}\Lambda\mathbf{S}^{-}\mathbf{X}_I'\left[y_I - \mathbf{X}_I\beta\right] \quad \ldots(3.1)$$

$$D_I[\Lambda\widehat{\beta}] = \frac{1}{q\widehat{\sigma}^2}\mathbf{e}_I'\mathbf{M}_I^{-1}\mathbf{X}_I\mathbf{X}^{-}\Lambda'\left[\Lambda\mathbf{S}^{-}\Lambda'\right]^{-1}\Lambda\mathbf{S}^{-}\mathbf{X}_I'\mathbf{M}_I^{-1}\mathbf{e}_I \quad \ldots(3.2)$$

In particular, in the study of the influence of a single observation, that is to say $I = \{i\}$, the D_i^*-distance and D_i-distance associated with y_i are, respectively,

$$D_i^*[\Lambda\widehat{\beta}] = \frac{1}{q\widehat{\sigma}^2}\mathbf{x}_i\mathbf{S}^{-}\Lambda'\left[\Lambda\mathbf{S}^{-}\Lambda'\right]^{-1}\Lambda\mathbf{S}^{-}\mathbf{x}_i'\left[y_i - \mathbf{x}_i\beta\right]^2 \quad \ldots(3.3)$$

and

$$D_i[\Lambda\widehat{\beta}] = \frac{1}{q(1-v_{ii})}\mathbf{x}_i\mathbf{S}^-\Lambda'\left[\Lambda\mathbf{S}^-\Lambda'\right]^{-1}\Lambda\mathbf{S}^-\mathbf{x}_i'r_i^2 \qquad \ldots(3.4)$$

where $r_i^2 = e_i^2/[\widehat{\sigma}^2(1-v_{ii})]$

Thus, we propose $D_i[\Lambda\widehat{\beta}]$ and $D_i[\Lambda\widehat{\beta}]$ as an influence measure of y_I and y_i on $\Lambda\widehat{\beta}$, respectively. In the multiple Regression Model, $D_I[\widehat{\beta}]$ coincides with Cook's distance (Cook and Weisberg (1982)).

(II) For $c_2 = \widehat{\sigma}^2_{(I)}$,

$W_i^*[\Lambda\widehat{\beta}] = \|\mathcal{J}[\mathbf{y}_I;\Lambda\widehat{\beta}]\|_{(\mathbf{Q},c_2)}$, called W_I^*-DISTANCE associated with the set of observations $\mathbf{y}_I$.

$W_i[\Lambda\widehat{\beta}] = \|\widehat{\mathcal{J}}[y_I;\Lambda\widehat{\beta}]\|_{(\mathbf{Q},c_2)}$, called W_I-DISTANCE associated with the set of observations y_I.

Easily, we obtain the following expressions for these norms :

$$W_I^*[\Lambda\widehat{\beta}] = \frac{1}{\widehat{\sigma}^2_{(I)}}[y_I - \mathbf{X}_I\beta]'\mathbf{X}_I\mathbf{S}^-\Lambda'\left[\Lambda\mathbf{S}^-\Lambda'\right]^{-1}\Lambda\mathbf{S}^-\mathbf{X}_I'[y_I - \mathbf{X}_I\beta] \qquad \ldots(3.5)$$

$$W_I[\Lambda\widehat{\beta}] = \frac{1}{\widehat{\sigma}^2_{(I)}}\mathbf{e}_I'\mathbf{M}_I^{-1}\mathbf{X}_I\mathbf{S}^-\Lambda'\left[\Lambda\mathbf{S}^-\Lambda'\right]^{-1}\Lambda\mathbf{S}^-\mathbf{X}_I'\mathbf{M}_I^{-1}\mathbf{e}_I \qquad \ldots(3.6)$$

In particular, if $I = \{i\}$ the W_i^* and W_i distances associated with the single observation are, respectively.

$$W_i^*[\Lambda\widehat{\beta}] = \frac{1}{\widehat{\sigma}^2_{(I)}}\mathbf{x}_i\mathbf{S}^-\Lambda'[\Lambda\mathbf{S}^-\Lambda']^{-1}\Lambda\mathbf{x}_i'[y_i - \mathbf{x}_i\beta]^2 \qquad \ldots(3.7)$$

$$W_i[\Lambda\widehat{\beta}] = \frac{1}{(1-v_{ii})}\mathbf{x}_i\mathbf{S}^-\Lambda'[\Lambda\mathbf{S}^-\Lambda']^{-1}\Lambda\mathbf{x}_i't_i^2 \qquad \ldots(3.8)$$

where $t_i^2 = e_i^2/[\widehat{\sigma}^2_{(i)}(1-v_{ii})]$.

Likewise, we propose $W_I[\Lambda\widehat{\beta}]$ and $W_i[\Lambda\widehat{\beta}]$ as an influence measure of y_I and y_i on $\Lambda\boldsymbol{\beta}$, respectively. In the Multiple Regression Model, $W_I[\widehat{\beta}]$ coincides with the squared Welsch-Kuh's distance (Belsley *et al.* (1980)).

(III) For $c_3 = [r/(n-r)]\widehat{\sigma}^2_{(I)}$,

$C_I^*[\Lambda\widehat{\beta}] = \|\mathcal{J}[y_I;\Lambda\widehat{\beta}]\|_{(\mathbf{Q},c_3)}$, called C_I^*-DISTANCE associated with the set of observations y_I.

$C_I[\Lambda\widehat{\beta}] = \|\widehat{\mathcal{J}}[y_I;\Lambda\widehat{\beta}]\|_{(\mathbf{Q},c_3)}$ is called C_I^*-DISTANCE associated with the set of observations y_I.

The expressions of these norms are given by

$$C_I^*[\Lambda\widehat{\beta}] = \frac{1}{\frac{r}{n-r}\widehat{\sigma}^2_{(I)}}[y_I - \mathbf{X}_I\beta]'\mathbf{X}_I\mathbf{S}^-\Lambda'\left[\Lambda\mathbf{S}^-\Lambda'\right]^{-1}\Lambda\mathbf{S}^-\mathbf{X}_I'[y_I - \mathbf{X}_I\beta] \qquad \ldots(3.9)$$

$$C_I[\Lambda\widehat{\beta}] = \frac{1}{\frac{r}{n-r}\widehat{\sigma}^2_{(I)}} \mathbf{e}_I^* \mathbf{M}_I^{-1} \mathbf{X}_I \mathbf{S}^- \Lambda' \left[\Lambda \mathbf{S}^- \Lambda'\right]^{-1} \Lambda \mathbf{S}^- \mathbf{X}_I' \mathbf{M}_I^{-1} \mathbf{e}_I \qquad \ldots (3.10)$$

In particular, in the study of the influence of a single observation, the C_i^*-distance and C_i-distance associated with y_i are, respectively,

$$C_i^*[\Lambda\widehat{\beta}] = \frac{1}{\frac{r}{n-r}\widehat{\sigma}^2_{(I)}} \mathbf{x}_i \mathbf{S}^- \Lambda' \left[\Lambda \mathbf{S}^- \Lambda'\right]^{-1} \Lambda \mathbf{S}^- \mathbf{x}_i' [y_i - \mathbf{x}_i \beta]^2 \qquad \ldots (3.11)$$

$$C_i[\Lambda\widehat{\beta}] = \frac{n-r}{r} \frac{1}{(1-v_{ii})} \mathbf{x}_i \mathbf{S}^- \Lambda' \left[\Lambda \mathbf{S}^- \Lambda'\right]^{-1} \Lambda \mathbf{S}^- x_i' t_i^2 \qquad \ldots (3.12)$$

Likewise, we propose $C_I[\Lambda\widehat{\beta}]$ and $C_i[\Lambda\widehat{\beta}]$ as influence measure of y_I and y_i on $\Lambda\widehat{\beta}$, respectively. In the Multiple Regression Model, $C_I[\widehat{\beta}]$ coincides with the square modified Cook's distance (Atkinson (1982)).

3.2. *Influence measures on* $\widehat{\sigma}^2$. Based on Corollary 2.14 we obtain the following influence measures on the unbiased estimator of the variance σ^2.

1. Measure of influence that the set of observations, indexed by I, has on $\widehat{\sigma}^2$

$$SMED_I = \frac{1}{n-r-m} \mathbf{e}_i' \mathbf{M}_I^{-1} \mathbf{e}_I - \frac{m}{n-r-m} \widehat{\sigma}^2 \in \Re$$

2. Measure of influence that the i-th observation has on $\widehat{\sigma}^2$

$$SMED_i = \frac{1}{n-r-1} \frac{e_i^2}{1-v_{ii}} - \frac{1}{n-r-1} \widehat{\sigma}^2 = \frac{1}{n-r-1} \widehat{\sigma}^2 \left[r_i^2 - 1\right] \in \Re.$$

4. APPLICATION TO THE ANALYSIS OF COVARIANCE

In this section the results previously obtained are applied to the analysis of Covariance model with one factor and two covariates using a data set taken from BMDP STATISTICAL SOFTWARE in the file EXERCISE.DAT (Frane, Jennrich and Sampson (1988)). The data are measurements corresponding to 40 people whose pulses were taken before and after running a mile. The following variables are measured :

ID (Identification number); SEX (Male, Female); SMOKER (Yes, No)

AGE (Subject's age in years); PULSE-1 (Pre-exercise pulse rate)

PULSE-2 (Post-exercise pulse rate)

The Analysis of Covariance of the variable PULSE-2 regarding the SEX factor and the covariates AGE and PULSE-1 can be formulated by the following linear model

$$PULSE_{2_{ij}} = \mu + \alpha_i + \gamma_1 AGE_{ij} + \gamma_2 PULSE_{1_{ij}} + \epsilon_{ij}.\epsilon_{ij} = N(0, \sigma^2)$$

where the subindex ij represents the value of the corresponding variable on the j-th individual of the i-th group (SEX). We obtain that the BLUE's vector of the estimate linear function

$$\begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \beta = \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix}$$

where $\beta' = [\mu\alpha_1\alpha_2\gamma_1\gamma_2]$ is $[\widehat{\gamma}_1\widehat{\gamma}_2]' = [-0.2639\ 0.9977]'$. The influence analysis of each observation, taken individually, on the BLUE's vector applying the influence measures obtained from the conditional bias concept leads to the results recorded in Table 1, which are represented in Figures 1-3.

Table 1. INFLUENCE MEASURES $[\widehat{\gamma}_1\widehat{\gamma}_2]'$

ID	D-DIST	W-DIST	C-DIST	ID	D-DIST	W-DIST	C-DIST
1	0.003	0.003	0.053	21	0.001	0.001	0.023
2	0.000	0.000	0.003	22	0.001	0.001	0.014
3	0.001	0.001	0.023	23	0.000	0.000	0.001
4	0.031	0.031	0.553	24	0.000	0.000	0.002
5	0.000	0.000	0.007	25	0.009	0.008	0.152
6	1.552	8.013	144.235	26	0.001	0.000	0.009
7	0.003	0.003	0.045	27	0.002	0.002	0.040
8	0.002	0.001	0.027	28	0.002	0.002	0.035
9	0.000	0.000	0.002	29	0.000	0.000	0.001
10	0.004	0.003	0.063	30	0.011	0.011	0.204
11	0.008	0.008	0.139	31	0.009	0.009	0.168
12	0.000	0.000	0.000	32	0.000	0.000	0.000
13	0.072	0.071	1.279	33	0.013	0.013	0.237
14	0.000	0.000	0.001	34	0.001	0.001	0.013
15	0.082	0.082	1.473	35	0.007	0.007	0.129
16	0.024	0.023	0.421	36	0.016	0.016	0.289
17	0.005	0.006	0.099	37	0.016	0.016	0.283
18	0.007	0.007	0.127	38	0.029	0.028	0.509
19	0.002	0.002	0.029	39	0.017	0.017	0.307
20	0.003	0.003	0.052	40	0.000	0.000	0.004

We can use that the observation number 6 has a considerable influence on the BLUE's vector of the estimable linear function under consideration, which could mask the behaviour of the other observations. For this reason, Fig. 4, Fig. 5 and Fig. 6 represent the data from Table 1, omitting observation number 5.

In these plots we can see that observations number 13 and 15 has greater influence than the rest of the observations, and for this reason, along with number 6, they should undergo a more detailed study.

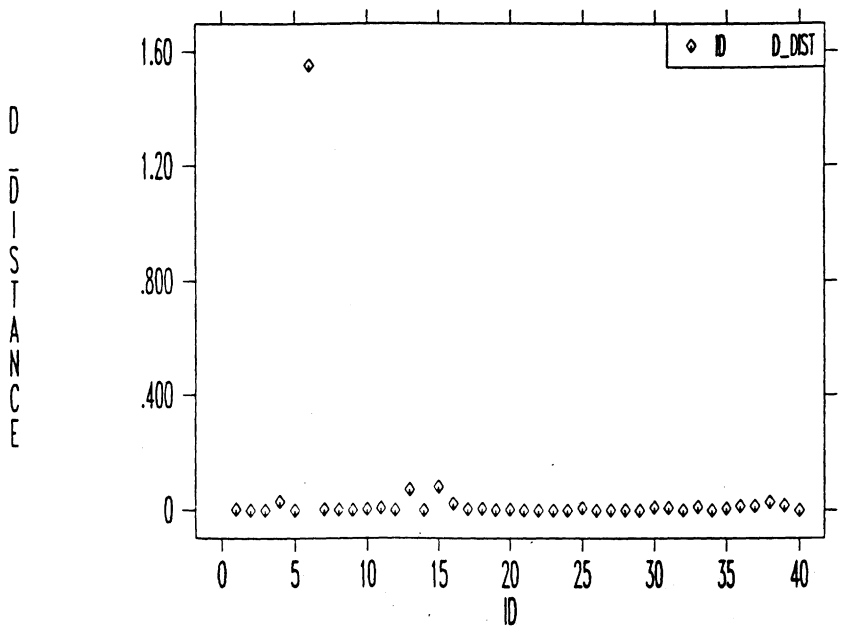

Fig. 1. Plot of D-distances

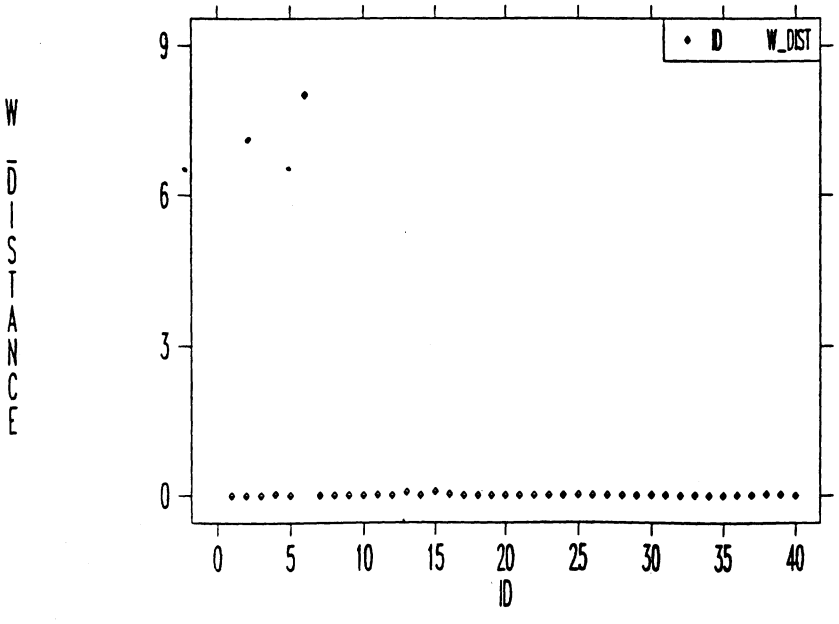

Fig. 2. Plot of W-distances

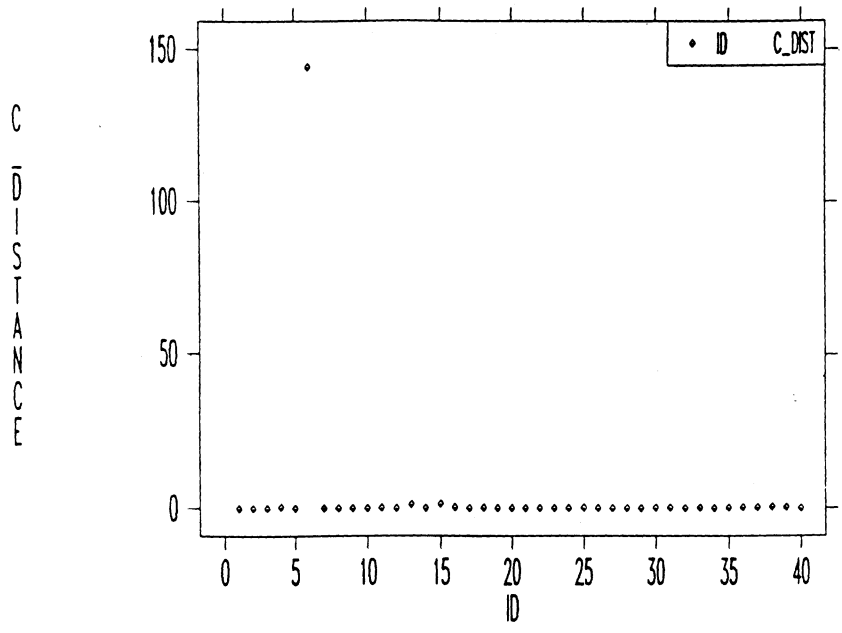

Fig. 3. Plot of C-distances

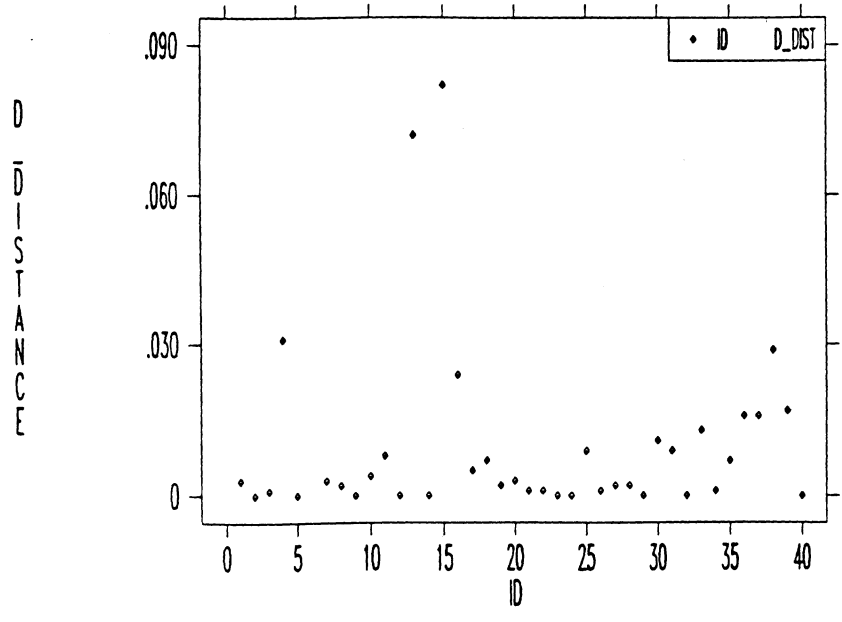

Fig. 4. Plot of D-distances (deleted case 6)

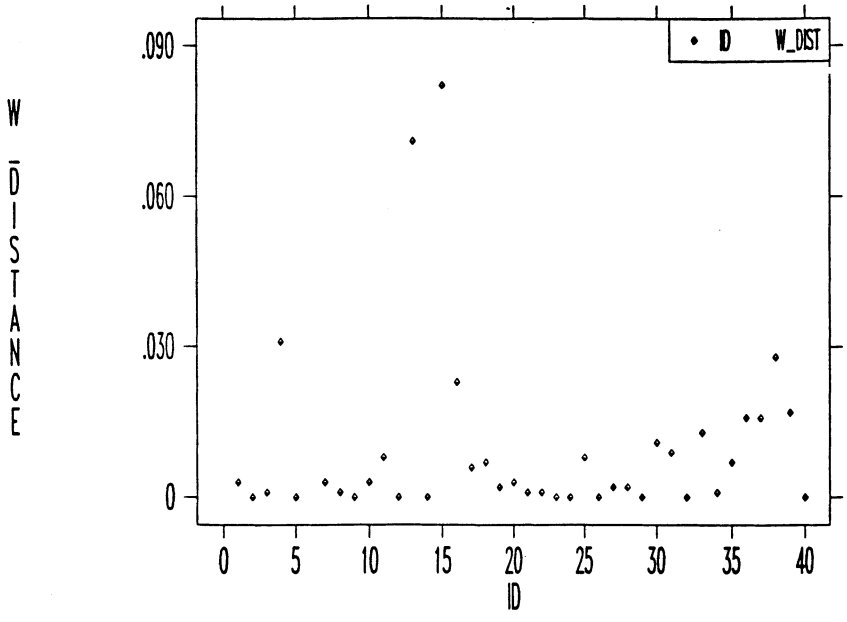

Fig. 5. Plot of W-distances (deleted case 6)

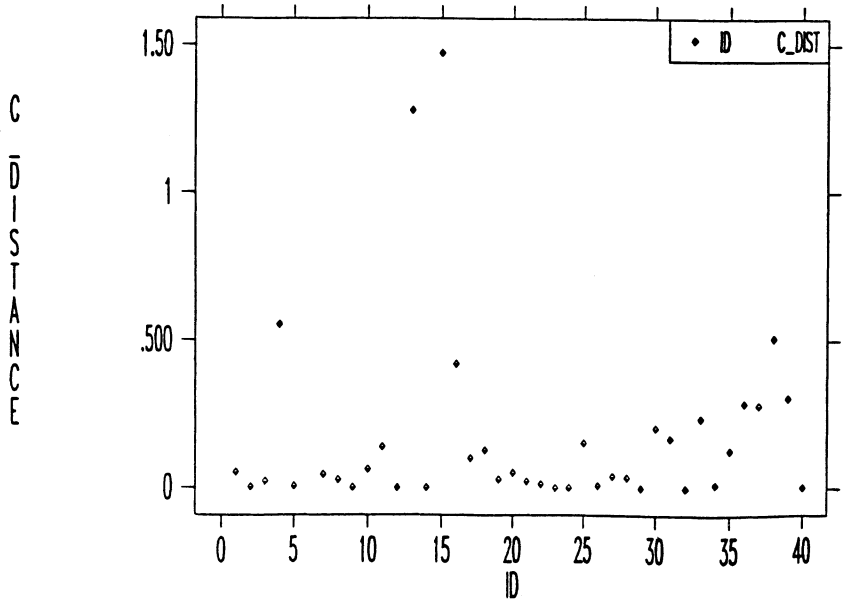

Fig. 6. Plot of C-distances (deleted case 6)

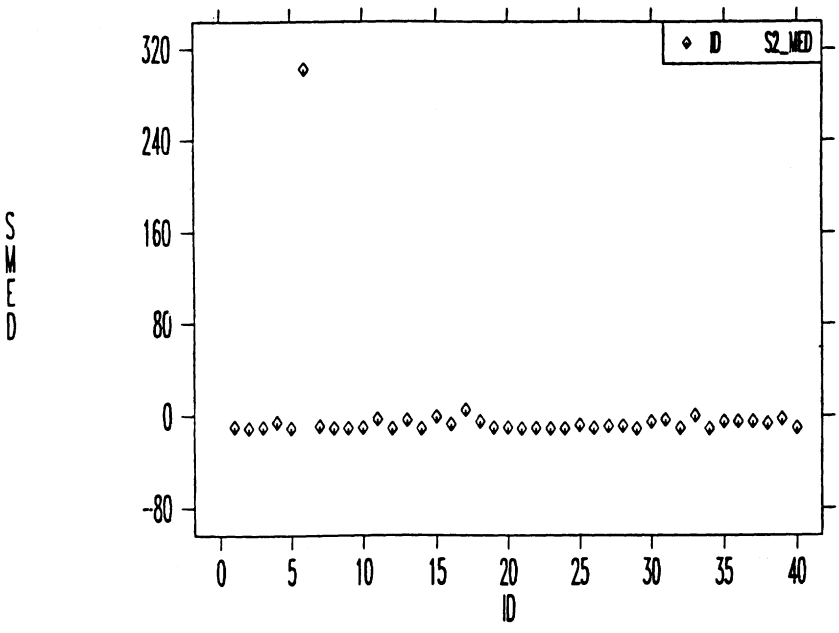

Fig. 7. Plot of SMED

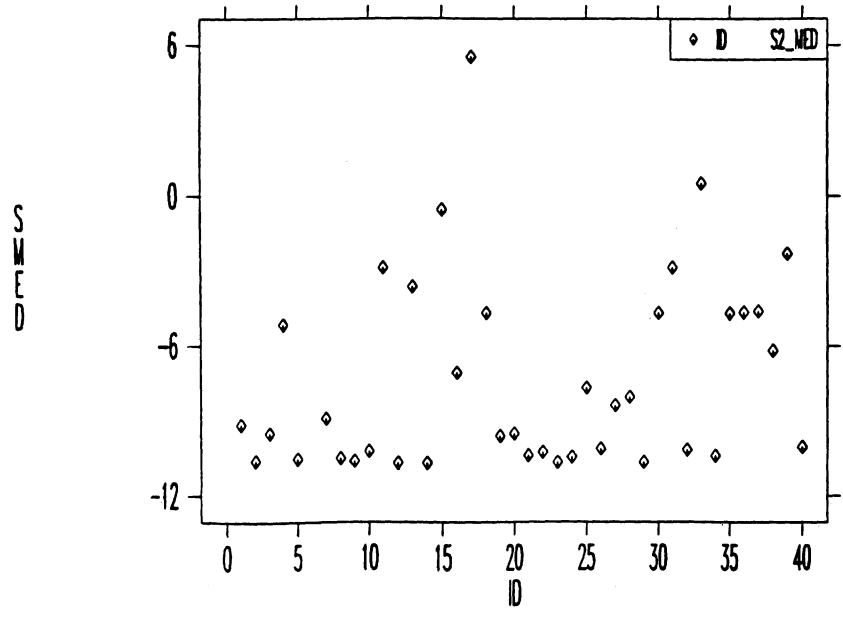

Fig. 8. Plot SMED (deleted case 6)

The influence measures, $SMED_i$, on the unbiased estimator of the variance $\widehat{\sigma}^2$, Table 2, are represented in FIG. 7 and FIG. 8. From these we can draw the conclusion that observation number 6 has a considerable greater influence on this estimator than the rest of the observations. However, observations 13 and 15 have very little effect on this estimation, as opposed to that obtained for the influence analysis on $[\widehat{\gamma}_1\widehat{\gamma}_2]'$.

Table 2. INFLUENCE MEASURE ON $\widehat{\sigma}^2$

1	-9.164	11	-2.864	21	-10.394	31	-2.843
2	-10.605	12	-10.693	22	-10.252	32	-10.163
3	-9.511	13	-3.650	23	-10.661	33	0.514
4	-5.167	14	-10.699	24	-10.454	34	-10.391
5	-10.527	15	-0.535	25	-7.676	35	-4.702
6	302.158	16	-7.076	26	-10.130	36	-4.653
7	-8.926	17	5.536	27	-8.400	37	-4.614
8	-10.505	18	-4.695	28	-8.062	38	-6.174
9	-10.623	19	-9.614	29	-10.653	39	-2.303
10	-10.231	20	-9.519	30	-4.703	40	-10.016

5. CONCLUSION

In conclusion, in this paper, we propose :

A new theoretical approach to the study of influence, that it can be extended to different statistical models. This approach is based on conditional bias of a given statistic T and does not presuppose hypothesis on the underlying distribution.

Influence measures for the statistics of interest in the general linear model. So, we propose D_I-distance W_I-distance and C_I-distance to quantify the influence of observations indexed by I on $\Lambda\widehat{\beta}$. For the linear regression model, the distances coincides with Cook's distance, squared Welsch-Kuh distance and modified Cook's distance, respectively. Nevertheless, these distances can be applied to all particular cases of (GLM). Finally, we propose a measure of the influence on $\widehat{\sigma}^2$, $SMED_I$.

REFERENCES

ATKINSON, A.C. (1982). Regression Diagnostics. Transformations and Constructed Variables. *Jour. Roy. Statist. Soc.*, Ser. B, **44**, No. 1, 1-36.

BARNETT, V. (1976). The ordering of multivariate data. *Jour. Roy. Statist. Soc.*, Ser. A, **139**, Part. 3, 318-344.

BARRETT, B.E. and LING, R.F. (1992). General Classes of Influence Measures for Multivariate Regression. *Jour. Amer. Statist. Asssoc.*, **87**, No. 417, 184-191.

BELSEY, D.A., KUH, E. and WELSCH, R.E. (1980). *Regression Diagnostics : Identifying Influential Data an Sources of Collinerity*, New York : John Wiley & Sons.

CARONI, C. (1987). Residuals and influence in the multivariate linear model. *The Statistician*, **36**, 365-370.

CHATTERJEE, S. and HADI, A.S. (1986). Influential observations, high leverage points and outliers in linear regression. *Statistical Science*, **1**, No. 3, 379-416.

– – – – (1988). *Sensitivity Analysis in Linear Regression*, New York : John Wiley and Sons.

COOK, R.D. (1987). Influence assessment. *Jour. Appl. Statist.*, **14**, No. 2, 117-132.

COOK, R.D. and WEISBERG, S. (1982). *Residuals and Influence in Regression*, New York : Chapman and Hall.

EFRON, B. and STEIN, C. (1981). The jackknife estimate of variance. *Ann. Statist.*, **9**, No. 3, 586-596.

FRANE, J., JENNRICH, R. and SAMPSON, P. (1990). *BMDP Statistical Software Manual*, Ed. Dixon, W.J. University of California Press, Berkeley.

KSHIRSAGAR, A.M. (1983). *A Course in Linear Models*, New York : Marcel Dekker, Inc.

PRINGLE, R.M. and RAYNER, A.A. (1971). *Generalized Inverse Matrices with Applications to Statistics*, London : Charles Griffin and Co.

DEPARTAMENTO DE ESTADÍSTICA E INVESTIGACIÓN OPERATIVA
FACULTAD DE MATEMÁTICAS. UNIVERSIDAD DE SEVILLA
C/. TARFIA S/N.
41012 - SEVILLA
ESPANA.