Estimation of regression models with equi-correlated responses when some observations on the response variable are missing

H. Toutenburg¹ and Shalabh²

¹ Institut für Statistik, Universität München, 80799 München, Germany
² Department of Statistics, Panjab University, Chandigarh, India

Received: April 18, 2000; revised version: May 10, 2001

The date of receipt and acceptance will be inserted by the editor

Abstract: The present article deals with the problem of estimation of parameters in a linear regression model when some data on response variable is missing and the responses are equi-correlated. The ordinary least squares and optimal homogeneous predictors are employed to find the imputed values of missing observations. Their efficiency properties are analyzed using the small disturbances asymptotic theory. The estimation of regression co-
coefficients using these imputed values is also considered and a comparison of estimators is presented.

**Key words**  linear regression, missing data, equi-correlated response, comparison of estimators, optimal homogeneous predictors

1 Introduction

Let us consider the following linear regression model with equi-correlated disturbances:

\[ Y = X\beta + \sigma \epsilon \]  

(1.1)

where \( Y \) is a \( n \times 1 \) vector of \( n \) observations on the response variable, \( X \) is a \( n \times K \) full column rank matrix consisting of \( n \) observations on \( K \) explanatory variables, \( \beta \) is a \( K \times 1 \) vector of coefficients, \( \sigma \) is an unknown scalar and \( \epsilon \) is a \( n \times 1 \) vector of disturbances.

It is assumed that disturbances follow a multivariate normal distribution with mean 0, variances 1 and covariances or correlation coefficients \( \rho \) so that we can write

\[ E(\epsilon) = 0 \]

\[ E(\epsilon \epsilon') = \Sigma = (1 - \rho)I_n + \rho J_n J_n' \]  

(1.2)

where \( I_n \) is an identity matrix of order \( n \times n \), \( J_n \) denotes a column vector with all \( n \) elements unity and \( \rho \) is assumed to be known and different from zero.

Such a model provides an interesting framework for the analysis of data in many applications. For example, when data contain measurements on
symmetric organs like eyes of persons, the observations are found to be equi-
correlated; see, e.g. Múnoz, Rosner and Carey (1986) and Rosner (1984) for
details. Other instances relate to familial data and survey data arising from
cluster sampling; see, e.g., Christensen (1987), King and Evans (1986) and

For the estimation of model parameters, considerable attention has been
paid in the literature, see, e.g. Srivastava and Ng (1990) and the references
cited therein for a brief review of estimation procedures. A stringent assump-
tion made in all the procedures is that the data has no missing observation.
Such a specification may be violated in many practical situations and some
observations on the response variable may not be available for one reason or
the other. If $Y_{mis}$ denotes a $m \times 1$ vector of $m$ missing values in the response
variable and $X_*$ is a $m \times K$ matrix of $m$ observations on the $K$ explanatory
variables, we have

$$ Y_{mis} = X_* \beta + \sigma \epsilon_* $$ (1.3)

where $\epsilon_*$ is a $m \times 1$ vector of disturbances having same distributional prop-
erties as $\epsilon$, i.e.,

$$ E(\epsilon_* \epsilon'_*) = \Sigma_* = (1 - \varrho)I_m + \varrho J_m J'_m $$ (1.4)

$$ E(\epsilon_\epsilon') = \Sigma_* = \varrho J_m J'_m . $$

Estimation of $\beta$ on the basis of $(n + m)$ incomplete observations is the
subject matter of this article. In Section 2, we present three sets of imputed
values for the missing observations. Utilizing these imputed values, we repair
the incomplete data set and use it for the estimation of $\beta$. In Section 3, the efficiency properties of estimators of $\beta$ are analyzed. In the same way we discuss the properties of imputed values in Section 4.

2 Imputation of Missing Observations And Estimation of Coefficients

Let us assume, following Shalabh (1998), that the regression relationship contains no intercept term and the observations on explanatory variables are taken as deviations from their corresponding means so that $X'J_n$ and $X'J_m$ are null vectors. An interesting consequence of this specification is that ordinary least squares and generalized least squares estimators of $\beta$ from (1.1) are identically equal and given by

$$b = (X'X)^{-1}X'Y$$

(2.1)

which clearly does not utilize the $m$ incomplete observations at all; see Toutenburg and Trenkler (1998) for the case when intercept term is present.

A simple alternative for making use of $m$ incomplete observations is to find imputed values for missing observations on response variable and to substitute them in place of missing observations so that the thus repaired data set resembles the complete data set. Treating the problem of finding the imputed values as the problem of predicting the values of response variable outside the sample, Goldberger (1962) has presented the classical predictor of $Y_{mis}$ as

$$P_t = X_b$$

(2.2)
and the optimal homogeneous predictor in the class of unbiased predictors of $Y_{\text{mis}}$ as

$$P_2 = X_*b + \Sigma_* \Sigma^{-1}(Y - Xb)$$

$$= X_*b + \frac{\theta J'_n Y}{1 + (n - 1) \theta} J_m$$

which may be serve as the imputed values for missing observations on the response variable; see also Bibby and Toutenburg (1979).

If we relax the constraint of unbiasedness, the optimal homogeneous predictor is given by

$$P = \frac{\beta' X' \Sigma^{-1} Y}{\beta' X' \Sigma^{-1} X \beta + \sigma^2 X_* \beta}$$

$$+ \Sigma_* \Sigma^{-1} (Y - \frac{\beta' X' \Sigma^{-1} Y}{\beta' X' \Sigma^{-1} X \beta + \sigma^2 X \beta})$$

$$= \frac{\beta' X Y}{\beta' X' X \beta + \sigma^2 (1 - \theta)} X_* \beta + \frac{\theta J'_n Y}{1 + (n - 1) \theta} J_m$$

see, e.g., Rao and Toutenburg (1995, Sec. 6.5).

The predictor (2.4) has no practical utility due to involvement of unknown quantities $\sigma^2$ and $\beta$. A simple way to obtain a feasible version is to replace them by their unbiased estimators. Thus substituting $b$ in place of $\beta$ and

$$s^2 = \left( \frac{1}{n - K} \right) (Y - Xb)' \Sigma^{-1} (Y - Xb)$$

in place of $\sigma^2$ in (2.4), we find a feasible predictor for $Y_{\text{mis}}$ as follows:

$$\hat{P} = \frac{b' X' Y}{b' X' X b + s^2 (1 - \theta)} X_* b + \frac{\theta J'_n Y}{1 + (n - 1) \theta} J_m$$

$$= P_2 - \frac{s^2 (1 - \theta)}{b' X' X b + s^2 (1 - \theta)} P_1$$

(2.6)
which provides another set of imputed values for the missing observations on the response variable.

If we apply the method of ordinary least squares or generalized least squares to (1.1) and (1.3), we get the estimator of $\beta$ as

$$\hat{\beta}_* = (X'X + X'_*X_*)^{-1}(X'Y + X'_*Y_{mis}) \quad (2.7)$$

which has no utility due to presence of $Y_{mis}$. Replacing it by the vectors of imputed values, we obtain the following estimators of $\beta$:

$$\hat{\beta}_1 = (X'X + X'_*X_*)^{-1}(X'Y + X'_*P_1) \quad (2.8)$$

$$= b$$

$$\hat{\beta}_2 = (X'X + X'_*X_*)^{-1}(X'Y + X'_*P_2) \quad (2.9)$$

$$= b$$

$$\hat{\beta} = (X'X + X'_*X_*)^{-1}(X'Y + X'_*\hat{P}) \quad (2.10)$$

$$= \left[ I_k - \frac{s^2(1 - \varrho)}{b'X'Xb + s^2(1 - \varrho)}(X'X + X'_*X_*)^{-1}X'_*X_* \right] b.$$

Thus we observe that the estimators of $\beta$ employing the unbiased imputed values specified by (2.2) and (2.3) are identically equal to the least squares estimator $b$ which ignores the incomplete observations. In other words, the imputation procedure yielding unbiased imputed values does not serve any useful purpose. Such is, however, not the case when biased imputed values given by (2.6) are used. Here we obtain the estimator (2.10) that is clearly a shrunk estimator arising from $b$. 
3 Properties of Estimators of Coefficients

For the estimation of $\beta$, we have two distinct estimators $b$ and $\hat{\beta}$ specified by (2.1) and (2.10) respectively.

It is easy to see that $b$ is unbiased for $\beta$ while $\hat{\beta}$ is generally biased. The approximate expressions for analyzing the efficiency properties of $\hat{\beta}$ in relation to $b$ are obtained in Appendix and presented below:

**Theorem 1** The asymptotic approximation for the bias vector of $\hat{\beta}$ to order $O(\sigma^2)$ is given by

$$B(\hat{\beta}) = E(\hat{\beta} - \beta)$$

$$= -\sigma^2 \left( \frac{1 - \rho}{\beta' X' X \beta} \right) W \beta$$

while the difference matrix to order $O(\sigma^4)$ is

$$D(b; \hat{\beta}) = E(b - \beta)(b - \beta)' - E(\hat{\beta} - \beta)(\hat{\beta} - \beta)'$$

$$= \frac{\sigma^4 (1 - \rho)^2}{\beta' X' X \beta} \left[ W(X'X)^{-1} + (X'X)^{-1} W' \right.$$

$$- \frac{2}{\beta' X' X \beta} \{ W \beta \beta' + \beta \beta' W' + \left( \frac{n - K + 2}{2(n - K)} \right) W \beta \beta' W' \left. \right\}]$$

where

$$W = (X' X + X'_* X_*)^{-1} X'_* X_*.$$  \hspace{1cm} (3.3)

From (3.2), it is difficult to draw any clear inference about the superiority of $\hat{\beta}$ over $b$ or vice-versa according to the criterion of mean squared error matrix to the given order of approximation. We therefore consider trace of the mean squared error matrix as the performance criterion. With respect
to such a criterion, the estimator $\hat{\beta}$ is better than $b$ when

$$\text{tr} \left( W(X'X)^{-1} \right) > \left[ 2 \left( \frac{\beta'W\beta}{\beta'X'X\beta} \right) + \left( \frac{1}{2} + \frac{1}{n - K} \right) \frac{\beta'W'W\beta}{\beta'X'X\beta} \right]. \quad (3.4)$$

If $\lambda_{\text{max}}$ is the maximum eigenvalue of $W(X'X)^{-1}$, we have

$$\frac{\beta'W\beta}{\beta'X'X\beta} \leq \lambda_{\text{max}}$$

$$\frac{\beta'W'W\beta}{\beta'X'X\beta} \leq \frac{\beta'W\beta}{\beta'X'X\beta} \leq \lambda_{\text{max}}.$$

It may be noticed that $W(X'X)^{-1} = (X'X)^{-1} \left( X'X + X'_{\ast}X_{\ast} \right)^{-1}$ is a symmetric non-negative definite matrix so that all the eigenvalues are real and nonnegative. Further, if $T$ denotes the total of eigenvalues of $W(X'X)^{-1}$, the inequality (3.4) is satisfied as long as

$$T > \left( \frac{5}{2} + \frac{1}{n - K} \right) \lambda_{\text{max}} \quad \quad \quad (3.5)$$

which is a sufficient condition for the superiority of $\hat{\beta}$ over $b$ with respect to the criterion of trace of mean squared error matrix to order $O(\sigma^4)$. In other words, under the condition (3.5), use of imputation procedure providing biased imputed values for the missing observations of the response variable is worthwhile in comparison to the outright discard of incomplete observations so far as the estimation of coefficients in the model is concerned. An interesting aspect of the condition (3.5) is that it is easy to check in actual practice.
4 Properties Of Imputed Values

From (2.2), (2.3) and (2.6), it is easy to see that $P_1$ and $P_2$ are, but $\hat{P}$ is generally not, weakly unbiased for $Y_{mis}$ in the sense that

$$E(P_1 - Y_{mis}) = E(P_2 - Y_{mis}) = 0$$  \hspace{1cm} (4.1)

$$E(\hat{P} - Y_{mis}) \neq 0.$$  \hspace{1cm} (4.2)

Further, we have

$$V(P_1) = E(P_1 - Y_{mis})(P_1 - Y_{mis})'$$  \hspace{1cm} (4.3)

$$= \sigma^2 (1 - \rho) \left[ I_m + X_s(X'X)^{-1}X_s' + \frac{\theta}{1 + \rho} J_m J_m' \right]$$

$$V(P_2) = E(P_2 - Y_{mis})(P_2 - Y_{mis})'$$  \hspace{1cm} (4.4)

$$= \sigma^2 (1 - \rho) \left[ I_m + X_s(X'X)^{-1}X_s' + \frac{\theta}{1 + (n-1)\rho} J_m J_m' \right]$$

whence it is clearly seen that the variance covariance matrix of $P_1$ exceeds the variance covariance matrix of $P_2$ by a non-negative semi-definite matrix. This implies that $P_2$ is a better choice in comparison to $P_1$ for finding the imputed values of missing observations on the response variable.

Next, let us consider $\hat{P}$. The exact expression for the first and second order moments of $(\hat{P} - Y_{mis})$ can be derived but the resulting expressions will be sufficiently complex and it will be hard to draw any clear inference regarding the bias as well as the superiority of $\hat{P}$ over $P_1$ and $P_2$ and vice-versa. We therefore consider their approximations using the small disturbance asymptotic theory. Such results are derived in Appendix and presented below.
Theorem 2 The asymptotic approximation for the bias vector of $\hat{P}$ to order $O(\sigma^2)$ is given by

$$B(\hat{P}) = E(\hat{P} - Y_{mis}) = -\sigma^2 \left( \frac{1 - \varrho}{\beta'X'X\beta} \right) X_* \beta$$

(4.5)

while the difference matrix to order $O(\sigma^4)$ is

$$D(P_2; \hat{P}) = E(P_2 - Y_{mis})(P_2 - Y_{mis})' - E(\hat{P} - Y_{mis})(\hat{P} - Y_{mis})'$$

$$= \sigma^4 \frac{(1 - \varrho)^2}{\beta'X'X\beta} X_* Q X_*$$

(4.6)

where

$$Q = 2(X'X)^{-1} - \frac{5(n - K) + 2}{(n - K)\beta'X'X\beta} \beta\beta'.$$

Using Rao and Toutenburg (1995, Theorem A.7, p.303), it is observed that $Q$ cannot be a nonnegative definite matrix so that it follows from (4.6) that $\hat{P}$ does not dominate $P_2$ with respect to the criterion of mean squared error matrix to the given order of approximation. Similarly, using Rao and Toutenburg (1995, A.59, p. 304), we find that $(-Q)$ cannot be nonnegative definite except in the trivial case $K = 1$. This means that $P_2$ does not dominate $\hat{P}$. Thus $\hat{P}$ neither dominates $P_2$ nor is dominated by $P_2$ according to the mean squared error matrix criterion, at least to the order of our approximation.

Next, let us employ a weak criterion, viz., the trace of mean squared error matrix, for the comparison of $P_2$ and $\hat{P}$:

$$\text{tr} D(P_2; \hat{P}) = \sigma^4 \frac{2(1 - \varrho)^2}{\beta'X'X\beta} \left[ g - \left( \frac{5}{2} + \frac{1}{n - K} \right) \right]$$

(4.7)
where
\[ g = \left( \frac{\beta'X'X\beta}{\beta'X'_*X_*\beta} \right) \text{tr}(X'X)^{-1}X'_*X_* . \] (4.8)

Thus we observe that \( \hat{P} \) is better than \( P_2 \) when
\[ g > \left( \frac{5}{2} + \frac{1}{n-K} \right) \] (4.9)
while the opposite is true, i.e., \( P_2 \) is better than \( \hat{P} \) when
\[ g < \left( \frac{5}{2} + \frac{1}{n-K} \right) . \] (4.10)

The conditions (4.9) and (4.10) are not attractive as they are hard to be verified in practice owing to involvement of \( \beta \) which is unknown.

If \( \alpha_{\text{min}} \) and \( \alpha_{\text{max}} \) denote the minimum and maximum eigenvalues of \( X'_*X_* \) in the metric of \( X'X \) and \( S \) is the sum of all the eigenvalues, we observe that the condition (4.9) is satisfied as long as
\[ \frac{S}{\alpha_{\text{max}}} > \left( \frac{5}{2} + \frac{1}{n-K} \right) \] (4.11)
which is easy to verify in any given application.

Similarly, the condition (4.10) will hold true as long as
\[ \frac{S}{\alpha_{\text{min}}} < \left( \frac{5}{2} + \frac{1}{n-K} \right) \] (4.12)
that is easy to check.

It is interesting to note that conditions (4.9)-(4.12) are free from values of \( \varphi \).
Appendix

From (1.1), (2.1) and (2.5), we observe that

\[
\begin{align*}
    s^2 &= \left( \frac{\sigma^2}{n-K} \right) \epsilon' \left[ I_n - X (X'X)^{-1} X' \right] \\
    &\quad \cdot \left[ (1 - \theta) I_n + \rho J_n J_n' \right]^{-1} \left[ I_n - X (X'X)^{-1} X' \right] \epsilon \\
    &= \frac{\sigma^2}{(n-K)(1-\theta)} \epsilon' \left[ I_n - X (X'X)^{-1} X' \right] \left[ I_n - \frac{\theta}{1+(n-1)\rho} J_n J_n' \right] \\
    &\quad \cdot \left[ I_n - X (X'X)^{-1} X' \right] \epsilon \\
    &= \frac{\sigma^2}{(n-K)(1-\theta)} \epsilon' M \epsilon
\end{align*}
\]

so that

\[
\begin{align*}
    \frac{s^2(1-\theta)}{b'X'Xb + s^2(1-\theta)} \\
    &= \frac{\epsilon' M \epsilon}{(n-K)\beta' X' X \beta} \left[ 1 + 2\sigma \frac{\beta' X' \epsilon}{\beta' X' X \beta} + O_p(\sigma^2) \right]^{-1} \\
    &= \frac{\sigma^2}{(n-K)\beta' X' X \beta} \frac{\epsilon' M \epsilon}{(n-K)(\beta' X' X \beta)^2} - 2\sigma^3 \frac{\epsilon' M \epsilon \beta' X' \epsilon}{(n-K)(\beta' X' X \beta)^2} + O_p(\sigma^4)
\end{align*}
\]

where

\[
M = I_n - X (X'X)^{-1} X' - \frac{\theta}{1+(n-1)\rho} J_n J_n'.
\]

Let us first take up Theorem 2.

From (A.2), we have

\[
\begin{align*}
    (\hat{P} - Y_{mis}) &= (P_2 - Y_{mis}) - \frac{s^2(1-\theta)}{b'X'Xb + s^2(1-\theta)} X_* \\
    &= \frac{1}{\beta + \sigma (X'X)^{-1} X' \epsilon} \\
    &= \sigma \left[ X_* (X'X)^{-1} X' \epsilon + \frac{\theta J_n \epsilon}{1+(n-1)\rho} - \epsilon_* \right] \\
    &\quad - \frac{\epsilon' M \epsilon}{(n-K)(\beta' X' X \beta) X_* \beta} + O_p(\sigma^2)
\end{align*}
\]

(A.3)
whence it follows that

\[
E(\hat{P} - Y_{mis}) = -\sigma^2 \frac{\text{E}(\epsilon'M\epsilon)}{(n-K)\beta'X'X\beta} X_*\beta + O(\sigma^3) \\
= -\sigma^2 \frac{\text{tr} M \Sigma}{(n-K)\beta'X'X\beta} X_*\beta + O(\sigma^3)
\]

which leads to the result (4.5) of Theorem 2.

Similarly, we have

\[
D(P_2; \hat{P}) = E(P_2 - Y_{mis})(P_2 - Y_{mis})' - E(\hat{P} - Y_{mis})(\hat{P} - Y_{mis})' \\
= E \left[ \frac{s^2(1-\varrho)}{b'X'b + s^2(1-\varrho)} \left\{ X_* \beta (P_2 - Y_{mis})' + (P_2 - Y_{mis})b'X_*' \right\} \right] \\
- E \left[ \left\{ \frac{s^2(1-\varrho)}{b'X'b + s^2(1-\varrho)} \right\}^2 X_* \beta b'X_*' \right].
\]

Observing that

\[
X_* \beta (P_2 - Y_{mis})' = [\sigma X_*\beta + \sigma^2 X_* (X'X)^{-1} X' \epsilon] \\
= [\epsilon' X'(X'X)^{-1} X_*' + \frac{\varrho J_n^\epsilon}{1 + (n-1)\varrho} J_m' - \epsilon_*'] \\
X_* \beta b'X_*' = X_* \beta \beta' X_*' + O_p(\sigma)
\]

and using (A.2), we get

\[
D(P_2; \hat{P}) = \sigma^3 E(F + F') + \sigma^4 E(G + G' - H) \quad (A.4)
\]

where

\[
F = \frac{\epsilon'M\epsilon}{(n-K)\beta'X'X\beta} X_*\beta [\epsilon' X'(X'X)^{-1} X_*' + \frac{\varrho J_n^\epsilon}{1 + (n-1)\varrho} J_m' - \epsilon_*'] \\
G = \frac{\epsilon'M\epsilon}{(n-K)\beta'X'X\beta} X_* [(X'X)^{-1} - \frac{2}{\beta'X'X\beta} \beta \beta'] X' \epsilon \\
= [\epsilon' X(X'X)^{-1} X_*' + \frac{\varrho J_n^\epsilon}{1 + (n-1)\varrho} J_m' - \epsilon_*'] \\
H = \left[ \frac{\epsilon'M\epsilon}{(n-K)\beta'X'X\beta} \right]^2 X_* \beta \beta' X_*'.
\]
Next, by virtue of normality, we have
\[
E(\epsilon' M \epsilon \epsilon') = (\text{tr} M \Sigma) \Sigma + 2 \Sigma M \Sigma
\]
\[
= (n - K + 2)(1 - \rho)[(1 - \rho)J_n + \rho J_n J_n'] - 2(1 - \rho)^2 X(X'X)^{-1}X'.
\]

Further, \( \epsilon \) and \((\epsilon_* - \Sigma_* \Sigma_*^{-1} \epsilon)\) are stochastically independent so that
\[
E(\epsilon' M \epsilon \epsilon'_{*}) = E(\epsilon' M \epsilon \epsilon') \Sigma_*^{-1} \Sigma_*'
\]
\[
= (\text{tr} M \Sigma) \Sigma_*' + 2 \Sigma M \Sigma_*'.
\]

Utilizing these results, it is easy to see that
\[
E(F) = 0
\]
\[
E(G) = \frac{(1 - \rho)^2}{\beta'X'X\beta} X_*[(X'X)^{-1} - \frac{2}{\beta'X'X\beta} \beta'X_*']
\]
\[
E(H) = \frac{(1 - \rho)^2(n - K + 2)}{(n - K)(\beta'X'X\beta)^2} X_*\beta'X_*' .
\]

Substituting these results in (A.4), we obtain the expression (4.6) of Theorem 2.

For the results of Theorem 1, we observe from (2.10) and (A.2), that
\[
(\hat{\beta} - \beta) = \sigma(X'X)^{-1}X'\epsilon - \frac{\sigma^2 \epsilon M \epsilon}{(n - K)\beta'X'X\beta} W \beta + O_p(\sigma^3)
\]
with \( W = (X'X + X_*'X_*)^{-1}X_*'X_* \) so that the bias vector to order \( O(\sigma^2) \) is
\[
B(\hat{\beta}) = -\frac{\sigma^2 E(\epsilon' M \epsilon)}{(n - K)\beta'X'X\beta} W \beta
\]
\[
= -\sigma^2 \left( \frac{1 - \rho}{\beta'X'X\beta} \right) W \beta
\]
leading to the result (3.1).
Similarly, we have

$$D(b; \hat{\beta}) = E(b - \beta)(b - \beta)' - E(\hat{\beta} - \beta)(\hat{\beta} - \beta)'$$

$$= E \left[ \frac{s^2(1 - \varrho)}{b'X'Xb + s^2(1 - \varrho)} \{Wb(b - \beta)' + (b - \beta)b'W'\} \right]$$

$$- E \left[ \left\{ \frac{s^2(1 - \varrho)}{b'X'Xb + s^2(1 - \varrho)} \right\}^2 Wbb'W' \right]$$

$$= \sigma^3 E \left[ \frac{\epsilon'Me}{(n - K)\beta'X'X\beta} \{W\beta\epsilon'X(X'X)^{-1} + (X'X)^{-1}X'\epsilon\beta'W'\} \right]$$

$$+ \sigma^4 E \left[ \frac{\epsilon'Me}{(n - K)\beta'X'X\beta} W \left\{ (X'X)^{-1} - \frac{2}{\beta'X'X\beta} \beta' \right\} X\epsilon\epsilon'X(X'X)^{-1} \right]$$

$$+ \frac{\epsilon'Me}{(n - K)\beta'X'X\beta} (X'X)^{-1}X\epsilon\epsilon'X \left\{ (X'X)^{-1} - \frac{2}{\beta'X'X\beta} \beta' \right\}$$

$$- \frac{(\epsilon'Me)^2}{(n - K)^2(\beta'X'X\beta)^2} W\beta\beta'W' \right] + O(\sigma^5)$$

$$= \frac{\sigma^4(1 - \varrho)^2}{\beta'X'X\beta} \left[ W(X'X)^{-1} + (X'X)^{-1}W' \right]$$

$$- \frac{1}{\beta'X'X\beta} \left\{ 2(W\beta\beta' + \beta\beta'W') + \left( \frac{n - K + 2}{n - K} \right) W\beta\beta'W' \right\}$$

whence we get the result (3.2) of Theorem 1.

References


