COMMUNICATIONS

DE LA FACULTÉ DES SCIENCES
DE L'UNIVERSITÉ D'ANKARA

Série A_1: Mathématique

TOME 25 ANNÉE 1976

On the Analysis of Multiple Regression In T-Categories

by

M. ORUÇ and F. AKDENİZ

2

Faculté des Sciences de l'Université d'Ankara
Ankara, Turquie

Communications de la Faculté des Sciences de l'Université d'Ankara

---

La Revue "Communications de la Faculté des Sciences de l'Université d'Ankara" est un organe de publication englobant toutes les disciplines scientifiques représentées à la Faculté.

La Revue, jusqu'a 1975 à l'exception des tomes I, II, III, était composée de trois séries:

Série A: Mathématique, Physique et Astronomie.

Série B: Chimie.

Série C: Sciences naturelles.

A partir de 1975 la Revue comprend sept séries:

Série A_1: Mathématique
Série A_2: Physique
Série A_3: Astronomie
Série B : Chimie
Série C_1: Géologie
Série C_2: Botanique
Série C_3: Zoologie

En principe, la Revue est réservée aux mémoires originaux des membres de la Faculté. Elle accepte cependant, dans la mesure de la place disponible, les communications des auteurs étrangers. Les langues allemande, anglaise et française sont admises indifféremment. Les articles devront être accompagnés d'un bref sommaire en langue turque.

Adres: Fen Fakültesi Tebliğler Dergisi Fen Fakültesi, Ankara, Turquie.

On the Analysis of Multiple Regression In T-Categories

M. ORUÇ and F. AKDENİZ*

Dept. of Mathematics, Faculty of Science
(Ankara University, Ankara, TURKEY)

SUMMARY

This paper examines the formal procedures for testing the constancy of the entire set of coeffecients in a regression equation. Then we are interested in asking if β in equation $Y = \dot{X}\beta + U$ remains constant over T periods of time. Thus, the appropriate test statistic is provided for the null hypothesis. The distribution of the residuals and subsampling in the general linear model is discussed.

I. INTRODUCTION

Consider a system of T regression equations of which the typical j'th equation is

$$Y_j = X_j\beta_j + U_j \qquad (j = 1,2,\ldots, T) \tag{1.1}$$

where Y_j is a N_jx 1 vector of observations of the j'th dependent variable, X_j is a N_jxK non-stochastic matrix with rank K of observations of K independent variables, $(N_j > K)$, β_j is a Kx1 vector of unknown regression coeffecients to be estimated, and U_j is a N_jx1 vector of random disturbance terms with mean zero.

The system of which (1.1) is an equation may be written as

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \cdot \\ \cdot \\ \cdot \\ Y_T \end{bmatrix} = \begin{bmatrix} X_1 & 0 & \ldots & 0 \\ 0 & X_2 & \ldots & 0 \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ 0 & \cdot & \cdot & X_T \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \cdot \\ \beta_T \end{bmatrix} + \begin{bmatrix} U_1 \\ U_2 \\ \cdot \\ \cdot \\ \cdot \\ U_T \end{bmatrix} \tag{1.2}$$

or more compactly as

* This paper is a part of a thesis submitted to the Faculty of Science, University of Ankara in partial fulfillment of the requirements for the degree of Doctorate (Ph.D) by Fikri Akdeniz.

$$Y = \dot{X}\,\beta + U \tag{1.3}$$

where $Y_j = (Y'_{j1}, \ldots, Y'_{jN_j})'$, $\beta_j = (\beta'_{j1}, \ldots, \beta'_{jK})'$

$$U_j = (U_{j1}, \ldots, U_{jN_j})' \quad \text{and} \quad X_j = (x_{ijt}) = \begin{bmatrix} x_{1j1} x_{2j1} \cdots & x_{Kj1} \\ x_{1j2} \; x_{2j2} \cdots & x_{Kj2} \\ \vdots & \vdots \\ x_{1jN_j} \cdots & x_{KjN_j} \end{bmatrix}$$

and $\dot{X}$ represents the block-diagonal matrix on the right hand side of (1.3). The $(\sum_{j=1}^{T} N_j \times 1)$ dimensional disturbance vector in (1.2) and (1.3) is assumed to have the following distribution:

$$U \sim N(O, \sigma^2 I) \tag{1.4}$$

where I is a unit matrix of order $\sum_1^T N_j \times \sum_1^T N_j$. Consider the the two hypotheses, one of which is the null hypothesis:

H_o: $\beta_1 = \beta_2 = \cdots = \beta_T$

or equivalently

$\beta_j = \beta. = (\beta._1 = , \ldots, = \beta._K)' \qquad (j = 1,2,\ldots,T)$

and the alternative one is

H_1: $\beta_1 \neq \beta_2 \neq \ldots \neq \beta_T$.

Under the null hypothesis: $\beta_1 = \beta_2 = \ldots = \beta_T$ the system in (1.2) can be written as

$$(Y'_1, \ldots, Y'_T)' = (X'_1, \ldots, X'_T)'\beta. + (U'_1, \ldots, U'_T)' \tag{1.5}$$

on

$$Y = X\,\beta. + U \tag{1.6}$$

If the null hypothesis is true, the least squares (also maximum likelihood) estimator of $\beta.$ denoted by b_o becomes

$$b_o = \beta. + \left(\sum_{i=1}^{T} X'_i X_i\right)^{-1} \sum_{i=1}^{T} X'_i U_i \quad . \tag{1.7}$$

The residuals from the regression are

$$e = M U \tag{1.8}$$

$$\text{where } M = I - X (X'X)^{-1} X'. \tag{1.9}$$

The sum of squares of the residuals under H_o can be exspressed

$$Q_1 = \sum_{i=1}^{T} \sum_{t=1}^{N_i} y^2_{it} - \sum_{i=1}^{T} \sum_{t=1}^{N_i} \sum_{k=1}^{K} y_{it} x_{jit} b._j \tag{1.10}$$

According to the statistical theory, the best linear unbiased estimator of $\beta.$ is given by (1.7). If the alternative hypothesis is true we will go back to the model (1.2) and the least squares estimators of $\beta_1, \beta_2, \ldots, \beta_T$ are

$$\begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_T \end{bmatrix} = \begin{bmatrix} (X'_1 \; X_1)^{-1} \; X'_1 \; Y_1 \\ (X'_2 \; X_2)^{-1} \; X'_2 \; Y_2 \\ \cdot \\ \cdot \\ \cdot \\ (X'_T \; X_T)^{-1} \; X'_T \; Y_T \end{bmatrix}$$

or

$$b = (\dot{X}' \dot{X})^{-1} \dot{X}' Y. \tag{1.11}$$

The residuals under H_1 become

$$\begin{bmatrix} e_1 \\ \cdot \\ \cdot \\ \cdot \\ e_T \end{bmatrix} = \begin{bmatrix} [\, I - X_1(X'_1 X_1)^{-1} X'_1 \,]\, U_1 \\ \cdot \\ \cdot \\ \cdot \\ [\, I - X_T(X'_T X_T)^{-1} X'_T \,]\, U_T \end{bmatrix} = \begin{bmatrix} M_1 U_1 \\ \cdot \\ \cdot \\ \cdot \\ M_T U_T \end{bmatrix} \tag{1.12}$$

Similarly, the sum of squares of the residuals will be

$$Q_2 = \sum_{i=1}^{T} Y'_i M_i Y_i = \sum_{i=1}^{T} U'_i M_i U_i. \tag{1.13}$$

2. TESTS OF EQUALITY BETWEEN SETS OF COEFFECIENTS IN T LINEAR REGRESSIONS

Theorem 2.1 Under the hypothesis $H_0: \beta'_j = \beta'. = (\beta'._1, \ldots, \beta'._K)$ suppose $U_1, U_2, \ldots, U_T$ are independent and U_i's are distributed as $N(0, \sigma^2 I)$, where U_i is an N_i-component disturbance vector. Let us write $C = Y'X\ H^{-1}$, where $H = X'\ X$ and is nonsingular and symetric matrix. Therefore, $\sum_{i=1}^{T} Y'_i Y_i - CHC'$ is distributed as $\sum_{i=1}^{T-1} W'_i W_i + W^{*\prime}_T W^*_T$, where the W_i are independently disributed, each according to $N(0, \sigma^2 I)$.

Proof:

Let $Y'Y - CHC' = Q_1$ be, hence $Q_1 = Y'(I - X(X'X)^{-1}X')Y$ $Y = Y'MY$. Since $PMP' = D$ and $M = P'DP$, we have $(MU)'(MU) = U'M\ U = U'P'D\ P\ U$; where

$$D = \begin{bmatrix} I_{\sum_{i=1}^{T} N_i - K} & \vdots & 0 \\ \cdots\cdots & & \cdots\cdots \\ 0 & \vdots & 0_K \end{bmatrix} \tag{2.1}$$

and

$$PP' = P'P = I.$$

Now let us write

$$W = P\ U. \tag{2.2}$$

Clearly, W is

$$W = \begin{bmatrix} W_1 \\ \cdot \\ \cdot \\ \cdot \\ W_T \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{T} P_{N_1 N_i} U_i \\ \cdot \\ \cdot \\ \cdot \\ \sum_{i=1}^{T} P_{N_T N_i} U_i \end{bmatrix}$$

On the other hand, the orthogonal matrix P can be written as $P = [P_{N_j N_i}]$,

where

$$P_{N_j N_i} = \begin{bmatrix} P_{\sum_{k=1}^{i-1} N_k+1 \; j \sum_{k=1}^{j-1} N_k+1} & \cdots & P_{\sum_{k=1}^{i-1} N_k+N_i \; j \sum_{k=1}^{j-1} N_k+1} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ P_{\sum_{k=1}^{i-1} N_k+1 \; j \sum_{k=1}^{j-1} N_k+N_j} & & P_{\sum_{k=1}^{i-1} N_k+N_i \; j \sum_{k=1}^{j-1} N_k+N_j} \end{bmatrix}.$$

Furthermore, Q_1 can be clearly written as

$$Q_1 = \sum_{i=1}^{T-1} W_i' W_i + W'_T \begin{bmatrix} I_{N_T-K} & O_{N_T-K,K} \\ O_{K,N_T-K} & O_K \end{bmatrix} W_T \,. \quad (2.3)$$

Let W_T be partitioned as

$$W_T = \begin{bmatrix} W^*_T \\ W^{**}_T \end{bmatrix} (W_T^*:N_T-K \times 1,\ W_T^{**}: K \times 1) \quad (2.4)$$

then, we have

$$Q_1 = \sum_{i=1}^{T-1} W'_i W_i + W_T^{*\prime} W_T^* \,.$$

When both sides of this last expression are divided by σ^2, we get

$$Q_1/\sigma^2 = \sum_{j=1}^{T-1} \sum_{i=1}^{N_j} (W_{ji}/\sigma)^2 + \sum_{i=1}^{N_T-K} (W_{Ti}/\sigma)^2. \tag{2.5}$$

Now, let us consider $W = PU$ which is given by (2.2). W is a vector of normal variables, since each element is a linear combination of normal variables (see [1] pp. 51). Thus the W_i's $(i = 1,2,...,T)$ are independently distributed, each according to $N(0, \sigma^2 I)$. It is clear that

$$E(W) = 0,$$

and (2.6)

$$E(PUU'P') = P\,\sigma^2 I_N\, P' = \sigma^2 I_N.$$

Finally, (2.5) is evidently the sum of squares of independently distributed standart normal variables, then Q_1/σ^2 is distributed as $\chi^2(\sum_{i=1}^{T} N_i - K)$ with rank (M) degrees of freedom. This completes the proof.

Considering (1.13) again, since

$$E(U_i) = 0$$

and

$$E(U_i U'_j) = \begin{cases} \sigma^2 I_{N_i} & i=j \\ 0 & i \neq j \end{cases} \qquad (i,j=1,2,...,T),$$

Q_2 quadratic form has rank $\sum_{i=1}^{T} N_i - TK$. Under the null hypothesis, the sum of squares in (1.10) will also be written as

$$\begin{aligned} Q_1 &= (Y - X b_o)'(Y - X b_o) \\ &= [(Y - \dot{X} b) - \dot{X}(b - J_K b_o)]'[(Y - \dot{X} b) - \dot{X}(b - J_K b_o)] \\ &= (Y - \dot{X} b)'(Y - \dot{X} b) + (b - J_K b_o)'\dot{X}'\dot{X}(b - J_K b_o) \\ &= Q_2 + Q_3 \end{aligned} \tag{2.7}$$

where

$$J_K = I \otimes j_K$$

is a column of unit matrices, and rank $(Q_3) = (T-1)K$, $A \otimes B$ denotes the Kronecker product of matrices A and B (see Graybill [3]). Furthermore $M = I - X(X'X)^{-1}X'$, $\dot{M} = I - \dot{X}(\dot{X}'\dot{X})^{-1}\dot{X}'$ and $X = \dot{X} . J_K$. Hence, we have $M = I - \dot{X} J_K (J'_K \dot{X}'\dot{X} J_K)^{-1} J'_K \dot{X}'$. Note that $M - \dot{M}$ is symetric and idempotent, i.e., $M - \dot{M} = (M - \dot{M})'$ and $M - \dot{M} = (M - \dot{M})^2$.

If M and $\dot{M}$ are written in terms of $\dot{X}$ and J_K, then $M - \dot{M}$ becomes

$$M - \dot{M} = \dot{X} (\dot{X}'\dot{X})^{-1}\dot{X}' - \dot{X} J_K (J'_K \dot{X}' \dot{X} J_K)^{-1} J'_K \dot{X}' = (M - \dot{M})^2.$$

We now consider the product $\dot{M}$ M, then we obtain

$$\dot{M} M = \dot{M}. \qquad (2.8)$$

Recalling that $e = MU$, $\dot{e} = \dot{M}U$, $Q_1 = e'e$, $Q_2 = \dot{e}'\dot{e}$ and taking $v = (M - \dot{M})U$, we have $Q_3 = Q_1 - Q_2 = v'v$. The quantity Q_3/σ^2 is distributed as $\chi^2[(T-1)K]$, and the quantity Q_2/σ^2 is distributed as $\chi^2(\sum_{i=1}^{T} N_i - TK)$.

The preceding brings to mind another theorem in Searle [5] and Wallace [6]. Independence of two quadratic forms involving a multivariate normal vector mean zero and variance-covariance matrix $\sigma^2 I$ requires that the product of the matrix of one quadratic form times the matrix of the order quadratic form yields the null matrix. Thus, we have,

$$\dot{M} (M - \dot{M}) = 0. \qquad (2.9)$$

hence, the independence requirement is satisfied and it is provided that Q_2 and Q_3 are independent. Further,

$$E(v\,v') = \text{trace}\,(M - \dot{M})\,\sigma^2 = \sigma^2 (T-1) K,$$

$v'v/\sigma^2$ is distributed as χ^2 with $(T-1)K$ degrees of freedom, and $\dot{e}'\dot{e}/\sigma^2$ is distributed as χ^2 with $\sum_{i=1}^{T} N_i - TK$ degrees of freedom.

Then, the statistic

$$F = \frac{v'v}{\dot{e}'\dot{e}} \quad \frac{\sum_{i=1}^{T} N_i \ -TK}{(T-1)K}, \qquad (2.10)$$

$$= \frac{\sum_{i=1}^{T} \sum_{j=1}^{K} \sum_{t=1}^{N_i} y_{it} x_{jit}(b_{ij}-b._j)}{\sum_{i=1}^{T} \sum_{t=1}^{N_i} y^2_{it} - \sum_{i=1}^{T} \sum_{j=1}^{K} \sum_{t=1}^{N_i} y_{it} x_{jit} b_{ij}} \cdot \frac{\sum_{1}^{T} N_i - TK}{(T-1)\ K}$$

is distributed as F with (T–1) K and $\sum_{i=1}^{T} N_i$–TK degrees of freedom. Thus, we can test H_0 by the F ratio. When the null hypothesis is true, F has the analysis of variance distribution with tr(M) –tr ($\dot{M}$) and tr ($\dot{M}$) degrees of freedom; where tr (.) denotes the trace operator. (see Chow [2] and Kullback and Rosenbaltt [4]).

Theorem 2.2 Under the hypothesis H_1 the best linear unbiased estimator (BLUE) of β is $b = (\dot{X}'\dot{X})^{-1}\dot{X}'Y$. Under the null hypothesis H_0: $\beta_j = \beta. = (\beta'._1, \ldots, \beta'._K)'$the BLUE of $\beta.$ is $b_0 = (X'X)^{-1} X'Y$. Furthermore, let $Ab = \gamma$, $\dot{H} = \dot{X}'\dot{X}$, $M = I - X(X'X)^{-1}X'$ and $\dot{M} = I - \dot{X}(\dot{X}'\dot{X})^{-1}\dot{X}'$ then the statistic

$$\omega = \frac{\gamma'(A\ \dot{H}^{-1}A')^{-1}\gamma \ / \ [tr\ (M)\ -tr\ (\dot{M})\ ]}{Y'\dot{M}\ Y \ / \ tr\ (\dot{M})} \qquad (2.11)$$

is distributed as F with (T – 1) K and $\sum_{i=1}^{T}$ N_i–TK degrees of freedom.

Proof:

The restrictions given by the null hypothesis can be expressed as follows:

$A\ \beta = 0,$

where

$$A = \begin{bmatrix} I & -I & \dots & 0 \\ 0 & I & -I\dots & 0 \\ \cdot & & \cdot & \\ \cdot & & \cdot & \\ \cdot & & \cdot & \\ 0 & 0 & \dots I & -I \end{bmatrix} . \tag{2.12}$$

The unit and zero matrices in (2.12) are of order KxK. Let

$p = (T-1)\ K$, $q = TK$ and $N = \sum_{i=1}^{T} N_i$, we have seen that

$$\frac{e'e - \dot{e}'\dot{e}}{\dot{e}'\dot{e}} = \frac{p}{N-q}\ F\ (p,\ N-q) \ . \tag{2.13}$$

Furtheremore, γ is obtained as

$$\gamma = \begin{bmatrix} (X'_1X_1)^{-1}X'_1Y_1 - (X'_2X_2)^{-1}X'_2Y_2 \\ \cdot \\ \cdot \\ \cdot \\ (X'_{T-1}X_{T-1})^{-1}X'_{T-1}Y_{T-1} - (X'_TX_T)^{-1}X'_TY_T \end{bmatrix}$$

and we have $A\ J_K = A.\ (I \otimes j_K) = 0$. If we write

$$R = \dot{H}\ (I_q - J_K(\ J'_K\dot{H}\ J_K)^{-1}J'_K\dot{H}\ ) \tag{2.14}$$

then we get

$$\begin{aligned} Q_3 &= Y'(M - \dot{M})\ Y \\ &= b'\dot{H}\ (I_q - J_K(J'_K\dot{H}\ J_K)^{-1}\ J'_K\dot{H})\ b \\ &= b'R\ b. \end{aligned}$$

Finally, premultiplying both sides of the equation (2.14) by AH^{-1} we obtain

$$A\dot{H}^{-1}R = AI_q(I_q - J_K(J'_K\dot{H}\ J_K)^{-1}J'_K\dot{H}). \tag{2.15}$$

Next, postmultiplying both sides by J_K, we see that

$$A\dot{H}^{-1}R\ J_K = A\ I_q(I_qJ_K - J_KI_K).$$

This shows that

$$A\ \dot{H}^{-1}R = A.$$

Since we accept $A\ \dot{H}^{-1}A'(A\ \dot{H}^{-1}A')^{-1}A = A$, we get $R = A'(A\ \dot{H}^{-1}A')^{-1}A$. If R is substituted in (2.14) we obtain

$$Y'(M - \dot{M})\ Y = b'A'(A\ \dot{H}^{-1}\ A')^{-1}\ Ab$$
$$= \gamma'(A\ \dot{H}^{-1}A')^{-1}\gamma. \qquad (2.16)$$

The numerator in (2.11), when divided by σ^2 is a χ^2 variable divided by its degrees of freedom, the denominator of (2.11) when divided by σ^2 is distributed as χ^2 with (N-TK) degrees of freedom. Since the quadratic forms are independent, the statistic ω is disributed as F with p and N-q degrees of freedom.

3. EXAMINING RESIDUALS

Under the null hypothesis H_o, suppose $e = MY$, that is, a random variable, let rank $(X) = K$, it follows that M has rank $\sum_{i=1}^{T} N_i - K$. Therefore, there exists an orthogonal matrix P such that $P'P = PP' = I$ and

$$P\ M\ P' = \begin{bmatrix} I_{N-K} & 0 \\ 0 & 0_K \end{bmatrix},$$

where $N = \sum_{i=1}^{T} N_i$. Further, let P be partitioned as

$$P = \begin{bmatrix} P_1 \\ P_2 \end{bmatrix} \qquad (3.1)$$

so that

$P_1M\ P'_1 = I_{N-K},\ P_2\ M\ P'_2 = O_K,\ P_1P'_2 = 0_{N-K,K}\ P_2P'_1 = 0_{K,N-K}$. Where P_1 and P_2 are $(N-K \times N)$ and $(K \times N)$ matrices respectively. It is easy to show that $M = P'_1P_1$. If $w = P_1\ e$, then $w\ w' = P_1\ e\ e'\ P'_1$. But since $E\ (e) = 0$ and $E\ (e\ e') = \sigma^2 M$, it follows that, $E\ (w\ w') = \sigma^2 P_1 M\ P'_1 I_{N-K}$. Under H_o, w is distributed as

$$w \sim N\,(0,\ \sigma^2 I_{N-K}). \tag{3.2}$$

Considering Q_{11}, where Q_{11} is a quadratic form, and has rank N-K, such that,

$$Q_{11} = w'w = Y'M\ P'_1P_1M\ Y = U'M\ U \tag{3.3}$$

Now let θ be orthogonal matrix that is partitioned as follows:

$$\theta = \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} \tag{3.4}$$

such that $\theta_1\dot{M}\ \theta'_1 = I_{N-TK},\ \theta_2\dot{M}\ \theta'_2 = 0_{TK},\ \theta_1\theta'_2 = 0_{N-TK\,,\,TK}$ and $\dot{M} = \theta'_1\theta_1$, where $\dot{M} = I - \dot{X}\,(\dot{X}'\dot{X})^{-1}\dot{X}'$ and θ_1 and θ_2 are (N-TK x N) and TKxN matrices respectively. Also, if $\dot{w} = \theta_1\ \dot{e}$, then $\dot{w}\ \dot{w}' = \theta_1\ \dot{e}\ \dot{e}'\theta'_1$. From statistical theory, since $E\,(\dot{e}) = O$ and $E\,(\dot{e}\ \dot{e}') = \sigma^2\dot{M}$, we have $E\,(\dot{w}\ \dot{w}') = \sigma^2\theta_1\ \dot{M}\ \theta'_1 = \sigma^2 I_{N-TK}$. Hence we get

$$Q_{21} = \dot{w}'\dot{w} = \dot{e}'\theta'_1\theta_1\dot{e} = U'\dot{M}\ U. \quad (3.5$$

When Q_{11}-Q_{21} is written, we get

$$Q_{31} = Q_{11} - Q_{21} = U'\,(M-\dot{M})U \tag{3.6}$$

We know that, $\dot{M}\,(M-\dot{M}) = 0$, then Q_{31} and Q_{21} quadratic forms are independent. Consequenty, the statistic

$$F = \frac{Q_{31}\,/\ tr\ (M)-tr(\dot{M})}{Q_{21}\,/\ tr\ (\dot{M})}$$

is distributed as F with p and N-q degrees of freedom.

4. THE REGRESION ANALYSIS OF POOLING DATA

Unequal Subsample Size. Let us take the matrices $X_j = [x_{ijt}]$ ($i = 1,2,...,K,\ j = 1,2,...,T,\ t = 1,2,...,N_j$), and the vectors Y_j of model (1.1). In this section, we set up $T = K$ and consider the regressör matrix as follows:

$$X^* = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1K} \\ x_{21} & x_{22} & \dots & x_{2K} \\ \cdot & & \cdot & \\ \cdot & & \cdot & \\ \cdot & & \cdot & \\ x_{K1} & x_{K2} & \dots & x_{KK} \end{bmatrix} \tag{4.1}$$

for which rank $(X^*) = K$. We now consider the set of equations

$$Y_j = \bar{X}_j \beta_j + U_j \qquad j = 1,2,\dots,K \tag{4.2}$$

where

$$\bar{X}_j = \begin{bmatrix} x_{j1} & x_{j2} & \dots & x_{jK} \\ x_{j1} & x_{j2} & \dots & x_{jK} \\ \cdot & & \cdot & \\ \cdot & & \cdot & \\ \cdot & & \cdot & \\ x_{j1} & x_{j2} & \dots & x_{jK} \end{bmatrix} \tag{4.3}$$

The system of equations (4.2) can be written in the form

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \cdot \\ \cdot \\ \cdot \\ Y_K \end{bmatrix} = \begin{bmatrix} \bar{X}_1 & 0 & \dots & 0 \\ 0 & \bar{X}_2 & \dots & 0 \\ \cdot & & \cdot & \\ \cdot & & \cdot & \\ \cdot & & \cdot & \\ 0 & 0 & \dots & \bar{X}_K \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \cdot \\ \beta_K \end{bmatrix} + \begin{bmatrix} U_1 \\ U_2 \\ \cdot \\ \cdot \\ \cdot \\ U_K \end{bmatrix}$$

or more compactly

$$Y = \tilde{X} \beta + U. \tag{4.4}$$

where $\tilde{X}$ represents the block-diagonal matrix on the right hand side of (4.4). If the null hypothesis H_0 is true, then (4.4) can be written as

$$(Y'_1,\dots,Y'_K)' = (\bar{X}'_1,\dots,\bar{X}'_K)\,\beta_0 + (U'_1,\dots,U'_K)'. \tag{4.5}$$

This model shows that we have observed the response y_{ij} N_i times for each set of x_i' a row of design matrix or each "treatment combination" is sampled K times and included in the model. Let X^* be a KxK with rows x'_i in equation (4.1), $Y' = (Y'_1,\dots, Y'_K)$

be a $1 x \sum_{i=1}^{K} N_i$ vector of observations and $\beta'_o = (\beta_{\cdot 1}, \ldots, \beta_{\cdot K})$.

Then, the linear model (4.5) may be written as $Y = X^{**}\beta_o + U$ or clearly,

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \cdot \\ \cdot \\ \cdot \\ Y_K \end{bmatrix} = \begin{bmatrix} j_{N_1} & \otimes & x'_1 \\ j_{N_2} & \otimes & x'_2 \\ & \cdot & \\ & \cdot & \\ & \cdot & \\ j_{N_K} & \otimes & x'_K \end{bmatrix} \beta_o + U \tag{4.6}$$

where all N_i do not have to be different, and j_{N_i} is $(N_i \times 1)$ column vector of unit elements. The value of β_o that minimizes

$$W_1 = (Y - X^{**}\beta_o)'(Y - X^{**}\beta_o)$$

is given by the solution to $\frac{\partial}{\partial \beta_o}(W_1) = 0$. Under H_o, the least squares estimator b_o for β_o is given by

$$b_o = (X^{**\prime}X^{**})^{-1}X^{**\prime}Y. \tag{4.7}$$

The sum of squares of the residuals can be expressed by

$$\widetilde{W}_1 = \sum_{i=1}^{K}\sum_{j=1}^{N_i} y^2_{ij} - 2\sum_{i=1}^{K}\sum_{j=1}^{N_i}\sum_{k=1}^{K} x_{ik}y_{ij}b_{.k}$$

$$- \sum_{i=1}^{K}\sum_{k=1}^{K} N_i x^2_{ik} b^2_{.k} -$$

$$\sum_{i=1}^{K}\sum_{v=1}^{K}\sum_{k=1}^{K} N_i x_{ik}x_{iv}b_{.v}b_{.k} \quad .$$

If we take the derivative with respect to y_{ij}, and the result set equal to zero, we get

$$\begin{bmatrix} \sum_{j=1}^{N_i} \dfrac{\partial \widetilde{W}_1}{\partial y_{1j}} \\ \cdot \\ \cdot \\ \cdot \\ \sum_{j=1}^{N_K} \dfrac{\partial \widetilde{W}_1}{\partial y_{Kj}} \end{bmatrix} = 2 \begin{bmatrix} \sum_{j=1}^{N_1} y_{1j} - N_1 \sum_{k=1}^{K} x_{1k} b_{.k} \\ \cdot \\ \cdot \\ \cdot \\ \sum_{j=1}^{N_K} y_{Kj} - N_K \sum_{k=1}^{K} x_{Kk} b_{.k} \end{bmatrix} = 0 \quad (4.8)$$

or

$$\bar{Y} = \hat{Y} = X^* b_o.$$

Corollary. Under the hypothesis $A\ \beta = 0$, let us consider the general linear model $y_{ij} = \sum_{k=1}^{K} x_{ik}\beta_{.k} + u_{ij}$($j = 1,...,N_i,\ i = 1,...,K$).

Hence we can write

$$j'_{N_i} Y_i = N_i \sum_{k=1}^{K} x_{ik}\beta_{.k} + j'_{N_i} U_i$$

or

$$\bar{Y}_i = \sum_{k=1}^{K} x_{ik}\beta_{.k} + \frac{1}{N_i} \sum_{j=1}^{N_i} u_{ij}$$

where j_{N_i} is the unit column vector of N_i elements, and $\bar{Y}_i$ is the observed mean of $Y_{i.}$ The disturbance vector U is assumed to have the following distribution:

$$U \sim N\ (0,\ \sigma^2 I).$$

Then the solution of the system

$$\bar{Y}_i = \sum_{j=1}^{K} x_{ij} b_{.j} \qquad (i=1,...,K)$$

gives the least squares estimates $b_{.j}$ of the coefficients $\beta_{.j}$.

REFERENCES

[1] Anderson T.W (1958) An Introduction to Multivariate Statistical Analysis. Wiley, New York,

[2] Chow G.C (1960) "Test of equality between sets of coeffecients in two Linear Regressions". Econometrica 28 591-605.

[3] Graybill, F.A. (1969) Introduction to Matrices with Applications in Statistics, Wadworth Publishing Co.

[4] Kullback. S and Rosenbaltt H.M. (1957) "On the Analysis of Multiple Regression in k-categories". Biometrika 44. 67-83.

[5] Searle, S.R. (1971). Linear Models. Wiley: New York.

[6] Wallace T.D and Toro-Vizcarronda C.E (1969) "Table for the mean for square error test for exact linear restrictions in regression". Journal of the American Statist. Association 64, 1649-1663.

ÖZET

Bu çalışmada T zaman aralığında regresyon denklemindeki katsayıların tüm kümesisinin değişmezlik testi için yöntemler araştırıldı. Sıfır hipotezi için uygun test istatistiği verildi. Kalanların dağılımı ve genel doğrusal modelde alt örnekleme incelendi.

Prix de l'abonnement annuel

Turquie : 15 TL; Étranger: 30 TL.

Prix de ce numéro : 5 TL (pour la vente en Turquie).

Prière de s'adresser pour l'abonnement à : Fen Fakültesi Dekanlığı Ankara, Turquie.

Ankara Üniversitesi Basımevi. Ankara - 1976