20 : CHAPTER

The Analysis of Variance

The Two-Way Model

We shall begin our discussion of the analysis of variance by considering a model in the form of

$$\begin{aligned} &y_{tj} = \mu + \gamma_t + \delta_j + \varepsilon_{tj}; \\ &\text{where} \quad t = 1, \ldots, T \quad \text{and} \quad j = 1, \ldots, M. \end{aligned} \tag{1}$$

It is assumed that the disturbances ε_{tj} are independently and identically distributed with $E(\varepsilon_{tj}) = 0$ and $V(\varepsilon_{tj}) = \sigma^2$ for all t, j.

The parameters γ_t and δ_j are subject to the restrictions

$$\sum_t \gamma_t = 0 \quad \text{and} \quad \sum_j \delta_j = 0; \tag{2}$$

and therefore $\mu = E(\sum_t \sum_j y_{tj}/MT)$ represents the expected value of the overall average of y_{tj}.

For the sake of a concrete interpretation, let us imagine that y_{tj} is an observation taken at time t in the jth region. Then the parameter γ_t represents an effect which is common to all observations taken at time t, whilst the parameter δ_j represents a characteristic of the jth region which prevails through time.

In order to apply our existing techniques to the problem of estimating the parameters of this model, we must first find an appropriate way of formatting the TM observations. Let us begin by using the index notation to represent the equations of (1) by

$$(y_{tj} e_t^j) = \mu(e_t^j) + (\gamma_t e_t^j) + (\delta_j e_t^j) + (\varepsilon_{tj} e_t^j). \tag{3}$$

In ordinary matrix notation, this becomes

$$Y = \mu \iota_T \iota'_M + \gamma \iota'_M + \iota_T \delta' + \mathcal{E}, \tag{4}$$

where $Y = [y_{tj}]$ and $\mathcal{E} = [\varepsilon_{tj}]$ are matrices of order $T \times M$ and $\gamma = [\gamma_1, \ldots, \gamma_T]'$ and $\delta = [\delta_1, \ldots, \delta_M]'$ are vectors of orders T and M respectively. It can be seen that $(e_t^j) = (e_t) \otimes (e^j) = \iota_T \otimes \iota_M'$ where ι_T and ι_M are vectors of units whose orders are indicated by their subscripts.

As an illustration, we may consider the case where $T = M = 3$. Then equation (3) assumes the following form:

$$
\begin{aligned}
\begin{bmatrix} y_{11} & y_{12} & y_{13} \\ y_{21} & y_{22} & y_{23} \\ y_{31} & y_{32} & y_{33} \end{bmatrix} =& \mu \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} + \begin{bmatrix} \gamma_1 & \gamma_1 & \gamma_1 \\ \gamma_2 & \gamma_2 & \gamma_2 \\ \gamma_3 & \gamma_3 & \gamma_3 \end{bmatrix} \\
&+ \begin{bmatrix} \delta_1 & \delta_2 & \delta_3 \\ \delta_1 & \delta_2 & \delta_3 \\ \delta_1 & \delta_2 & \delta_3 \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} & \varepsilon_{13} \\ \varepsilon_{21} & \varepsilon_{22} & \varepsilon_{23} \\ \varepsilon_{31} & \varepsilon_{32} & \varepsilon_{33} \end{bmatrix}.
\end{aligned} \tag{5}
$$

To put our equation into a form which is amenable to estimation by least squares regression, we must vectorise it. In terms of the notation of (3), the vectorised equation is

$$
(y_{tj} e_{jt}) = \mu(e_{jt}) + (e_{jt}^{t})(\gamma_t e_t) + (e_{jt}^{j})(\delta_j e_j) + (\varepsilon_{tj} e_{jt}). \tag{6}
$$

Using the notation of the Kronecker product, this can be rendered as

$$
\begin{aligned}
Y^c &= \mu(\iota_M \otimes \iota_T) + (\iota_M \otimes I_T)\gamma + (I_M \otimes \iota_T)\delta + \mathcal{E}^c \\
&= X\beta + \mathcal{E}^c.
\end{aligned} \tag{7}
$$

In comparing (6) and (7), we see, for example, that $(e_{jt}^{t}) = (e_j) \otimes (e_t^t) = \iota_M \otimes I_T$. We recognise that (e_t^t) is the sum over the index t of the matrices of order T which have a unit in the tth diagonal position and zeros elsewhere; and this sum amounts, of course, to the unit matrix of order T.

By vectorising our example under (5), we get

$$
\begin{bmatrix} y_{11} \\ y_{21} \\ y_{31} \\ y_{12} \\ y_{22} \\ y_{32} \\ y_{13} \\ y_{23} \\ y_{33} \end{bmatrix} = \left[\begin{array}{c|ccc|ccc} 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ \hline 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ \hline 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \left[\begin{array}{c} \mu \\ \hline \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \hline \delta_1 \\ \delta_2 \\ \delta_3 \end{array}\right] + \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{21} \\ \varepsilon_{31} \\ \varepsilon_{12} \\ \varepsilon_{22} \\ \varepsilon_{32} \\ \varepsilon_{13} \\ \varepsilon_{23} \\ \varepsilon_{33} \end{bmatrix}. \tag{8}
$$

The matrix X consisting of zeros and ones is described as the design matrix.

So far, in our description of the two-way model, we have assumed that there is a single observation in every cell. That is to say, only a single observation is taken at the time t in the jth region. Let us now imagine that a total of N_{tj} observations are taken. Then, in place of the element y_{tj} within the vector Y^c, we have a subvector containing the elements $y_{tij}; i = 1, \ldots, N_{tj}$. Corresponding to these N_{tj} observations, there will be N_{tj} replications of the row of the original design matrix which was associated with y_{tj}. Although this extension of the model is easily accomplished, it somewhat spoils the neatness of the algebra unless it can be assumed that $N_{tj} = N$ for all t, j. If this condition does prevail, then the equations of the model can be written as

$$
\begin{aligned}
y_{tij} &= \mu + \gamma_t + \delta_j + \varepsilon_{tij}; \\
\text{where} \quad & t = 1, \ldots, T, \\
& i = 1, \ldots, N \quad \text{and} \\
& j = 1, \ldots, M.
\end{aligned} \tag{8}
$$

The equations can be put in a vectorised format which is represented in our index notation by

$$(y_{tij}e_{jit}) = \mu(e_{jit}) + (e_{jit}^{t})(\gamma_t e_t) + (e_{jit}^{j})(\delta_j e_j) + (\varepsilon_{tij}e_{jit}). \tag{10}$$

Using the notation of the Kronecker product, this can be rendered as

$$
\begin{aligned}
y &= \mu(\iota_M \otimes \iota_N \otimes \iota_T) + (\iota_M \otimes \iota_N \otimes I_T)\gamma + (I_M \otimes \iota_N \otimes \iota_T)\delta + \varepsilon \\
&= X\beta + \varepsilon,
\end{aligned} \tag{11}
$$

where y and ε are long vectors whose elements are arrayed according to a reversed lexicographic ordering of their indices.

A cursory inspection of the example under (8) reveals that the design matrix is singular; for the first column can be obtained either as the sum of the columns 2, 3 and 4 or as the sum of the columns 5, 6 and 7. Therefore the matrix has a rank of 5; and it does not seem possible to estimate the parameters within β. However, it is always possible to estimate the mean vector $X\beta$. If we calculate the Q–R decomposition of $X = Q_rR$, then this estimate is available in the form of $X\hat{\beta} = PY^c = Q_rQ_r'Y^c$ where $P = Q_rQ_r'$ is the orthogonal projector onto the manifold of the design matrix.

Estimation with Restrictions

There are two ways in which we might overcome the problems posed by the singularity of the design matrix. The first way is to confine our attention to functions of the parameters in β which can be estimated from the information contained in the observations on y_{tj} and in the corresponding rows of the design matrix. We shall examine this approach in the next section

The alternative approach is to overcome the singularity of the design matrix by supplementing the equations $Y^c = X\beta + \mathcal{E}^c$ of (7) by a further set of equations which are derived from the restrictions under (2). The supplementary equations can be written as

$$\iota_T'\gamma = 0 \quad \text{and} \quad \iota_M'\delta = 0, \tag{12}$$

or as

$$\begin{aligned} &H'\beta = 0, \qquad \text{where} \\ &H' = \begin{bmatrix} 0 & \iota_T' & 0 \\ 0 & 0 & \iota_M' \end{bmatrix} \qquad \text{and} \quad \beta' = [\,\mu \quad \gamma' \quad \delta'\,]. \end{aligned} \tag{13}$$

In the example under (8), the singularity of the design matrix is overcome by adding to it the rows $[0,1,1,1,0,0,0]$ and $[0,0,0,0,1,1,1]$ which relate to the restrictions $\gamma_1 + \gamma_2 + \gamma_3 = 0$ and $\delta_1 + \delta_2 + \delta_3 = 0$ respectively. At the same time, the vectors Y^c and $\mathcal{E}^c$ must be lengthened by adding to each of them a pair of zeros.

It can be shown that this extra information, when it is added to the equations, in no way conflicts with the pre-existing observational information. The consequence is that, when we use the ordinary least-squares regression to estimate the parameters from the combined equations, we find that the estimates obey the restrictions of (13) precisely.

By combining the equations of (13) with the equations of (7), we get the system

$$\begin{bmatrix} Y^c \\ 0 \end{bmatrix} = \begin{bmatrix} X \\ H' \end{bmatrix} \beta + \begin{bmatrix} \mathcal{E}^c \\ 0 \end{bmatrix} \tag{14}$$

and, when ordinary least-squares regression is applied, we obtain the estimator

$$\hat{\beta} = (X'X + HH')^{-1}X'Y^c. \tag{15}$$

Exactly the same estimator is derived if we consider joining the equations $X\beta = PY^c$ to the equations $H'\beta = 0$ to derive the system

$$\begin{bmatrix} X \\ H' \end{bmatrix} \beta = \begin{bmatrix} PY^c \\ 0 \end{bmatrix}. \tag{16}$$

On premultiplying this system by the matrix $[X', H]$ and using the result that $X'P = X$, we get $(X'X + HH')\beta = X'PY^c = X'Y^c$; and the solution is provided, once more, by (15). To demonstrate that the estimate $\hat{\beta}$ fulfils the restrictions, we need only prove the following proposition:

(17) The equations $PY^c = X\beta$ and $H'\beta = 0$ have a unique solution for all Y^c if and only if the rows of X and H' are linearly independent and the matrix $[X', H]$ has full rank.

Proof. Assume to the contrary that the rows of X and H' are linearly dependent. Then there exists a vector $v' = [v_1', v_2']$ such that $v_1'X + v_2'H' = 0$ with $v_1'X, v_2'H' \neq 0$. Since the vector Y^c is unconstrained, there must be a value for Y^c such that $v_1'PY^c = v_1'X\beta \neq 0$. But this conflicts with the result that $(v_1'X + v_2'H')\beta = v_1'X\beta = 0$. Therefore, if there are to be solutions for all values of Y^c, the rows of X and H' cannot be linearly dependent.

Now assume that the rows of X and H' are linearly independent. Then $v_1'X + v_2'H' = 0$ implies $v_1'X, v_2'H' = 0$. Since $PY^c = X\beta$, it follows that, if $v_1'X + v_2'H' = 0$, then $v_1'PY^c + v_2'0 = 0$. Thus every vector which is orthogonal to the matrix $[X', H]'$ is also orthogonal to the vector $[(PY^c)', 0]'$. Therefore, the vector $[(PY^c)', 0]'$ must be a linear combination of the columns of the matrix, which implies the existence of a vector β such that $X\beta = PY^c$ and $H'\beta = 0$; and this is the solution we are seeking. Finally, the condition that $[X', H]$ has full rank is necessary and sufficient for the uniqueness of this solution.

If the order $T + M + 1$ of the matrix $X'X + HH'$ is great, then it may be impractical to estimate the elements of β in the manner suggested by equation (15). In that case, we may be inclined to approach the problem of estimation by way of the estimable parametric functions of β.

Estimable Parametric Functions

The alternative way of overcoming the problem of the singularity of the design matrix is to reparametrise the model so that its new parameters are estimable functions of the old parameters. We say that a parametric function $p'\beta$ is estimable if and only if there exists a vector q such that $E(q'Y^c) = p'\beta$. It is easy to prove that

(18) A parametric function $p'\beta$ is estimable if and only if there exists a vector q such that $p' = q'X$; that is to say, if and only if p' may be expressed as a linear combination of the rows of X.

Proof. If $p'\beta$ is estimable, then there exists, by definition, a vector q such that $E(q'Y^c) = q'X\beta = p'\beta$ for all values of β; whence it follows that $p' = q'X$. Conversely, if $p' = q'X$, then $p'\beta = q'X\beta = E(q'Y^c)$ and $p'\beta$ is an estimable parametric function for all β.

Example. We should note that, according to this criterion of estimability, the fact that we are unable to construct the identity matrix of order 7 from linear combinations of the rows of the design matrix in the example under (8) signifies that the seven parameters in $\beta = [\mu, \gamma_1, \gamma_2, \gamma_3, \delta_1, \delta_2, \delta_3]'$ cannot be estimated individually unless we can somehow impose the restrictions $\sum_t \gamma_t = 0$ and $\sum_j \delta_j = 0$. However, it is easy to see that the functions $\mu_{*j} = \mu + \delta_j$ and $\mu_{t*} = \mu + \gamma_t$ are estimable for all j and t.

Consider, for example, $\mu_{*3} = \mu + \delta_3$. Since $\sum_t \gamma_t = 0$, this can also be written as $\mu_{*3} = \mu + \frac{1}{3}\sum_t \gamma_t + \delta_3 = p'\beta$ where $p' = [1, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, 0, 1]$. We can see that $p'\beta$ is an estimable function by virtue of the fact that $p' = q'X$, where $q' = [0, 0, 0, 0, 0, 0, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}]$ and where X is the design matrix of equation (8). It follows that we can estimate $p'\beta$ by $q'y = \frac{1}{3}\sum_t y_{t3} = \bar{y}_{*3}$.

We can approach the problem of constructing estimable parametric functions in a wide variety of ways. One way, which is particularly simple, is to partition the columns of the design matrix to give $X = [X_1, X_2]$, wherein X_2, which has $\text{Rank}(X_2) = \text{Rank}(X)$, serves as a basis for X. It follows that $X_1 = P_2 X_1$, where $P_2 = X_2(X_2'X_2)^{-1}X_2'$ is the orthogonal projector onto the manifold of X_2.

To enable such a partitioning, we must allow the order of the columns of X to be chosen at will. However, in our example under (8), the columns of the design matrix are already in an appropriate order; for we can form the matrix X_1 from the first two and the matrix X_2 from the remaining five columns.

From the way in which we partition the design matrix, it follows that, if $p'\beta = p_1'\beta_1 + p_2'\beta_2$ is estimable by virtue of the condition $p' = q'X$, then

$$\begin{aligned} p' &= q'[X_1,\ X_2] \\ &= q'X_2[(X_2'X_2)^{-1}X_2'X_1,\ I] \\ &= p_2'[(X_2'X_2)^{-1}X_2'X_1,\ I]. \end{aligned} \tag{19}$$

We can use this result in either of two ways. On the one hand, we can ascertain whether a proposed parametric function $p_1'\beta_1 + p_2'\beta_2$ does satisfy the condition of estimability by seeking to confirm the condition

$$p_1' = p_2'(X_2'X_2)^{-1}X_2'X_1. \tag{20}$$

On the other hand, we can generate estimable parametric functions by freely choosing the vector p_2 and then using the formula of (20) to generate the value of p_1.

Hypothesis Testing and the Decomposition of the Variance

In the analysis of variance, our principle aim is to ascertain whether the classificatory scheme which gives rise to the design matrix has any power in explaining the variability of the dependent variable. Thus, if the locality in which the observations are taken is a factor in determining their values, then we would expect the values of δ_j to be significantly different from zero; and to test their significance we would wish to evaluate a null hypothesis maintaining that $\delta_1 = \delta_2 = \cdots = \delta_M = 0$. Unfortunately these parameters are not estimable unless the design matrix is supplemented by the necessary restrictions. Nevertheless, the values $\mu_{*j} = \mu + \delta_j$ will constitute estimable parametric functions in the context of our two-way model; and a reasonable index of their effect would be the sum of squares $\sum_j (\bar{y}_{*j} - \bar{y})^2$ wherein $\bar{y}$ is the average of all values of y_{tj} and $\bar{y}_{*j}$ is the average over time of all the observations taken in the jth region. Thus it transpires that, to test an hypothesis, it is not always necessary to estimate the parameters in question.

Hypothesis Testing in the One-Way Model

In order to demonstrate the derivation of the standard hypothesis tests, let us begin with a simplified one-way model which can be obtained by setting $\gamma_1 = \gamma_2 = \cdots = \gamma_T = 0$ in the equations of (1). This gives the equations

$$\begin{aligned} y_{tj} &= \mu + \delta_j + \varepsilon_{tj} \\ &= \mu_{*j} + \varepsilon_{tj}; \\ \text{where} \quad t &= 1, \ldots, T \quad \text{and} \quad j = 1, \ldots, M. \end{aligned} \tag{21}$$

Now consider the sum of squares of the deviations of the observations y_{tj} about their overall sample mean: $\sum_t \sum_j (y_{tj} - \bar{y})^2$. This can be decomposed into

$$\begin{aligned} \sum_t \sum_j (y_{tj} - \bar{y})^2 &= \sum_t \sum_j \left\{ (y_{tj} - \bar{y}_{*j}) + (\bar{y}_{*j} - \bar{y}) \right\}^2 \\ &= \sum_t \sum_j (y_{tj} - \bar{y}_{*j})^2 + \sum_t \sum_j (\bar{y}_{*j} - \bar{y})^2 \\ &\quad + 2 \sum_t \sum_j (y_{tj} - \bar{y}_{*j})(\bar{y}_{*j} - \bar{y}). \end{aligned} \tag{22}$$

However, the final term on the RHS is zero since $\sum_t (y_{tj} - \bar{y}_{*j}) = 0$ for all j. Therefore we find that

$$\sum_t \sum_j (y_{tj} - \bar{y})^2 = \sum_t \sum_j (y_{tj} - \bar{y}_{*j})^2 + T \sum_j (\bar{y}_{*j} - \bar{y})^2. \tag{23}$$

The first term on the RHS may be described as the sum of squares of the local variations and the second term may described as the sum of squares of the inter-regional variations. The local variation provides a standard by which to measure the inter-regional variation. Therefore it seems reasonable to adopt the following ratio as a statistic for testing the null hypothesis of no significant inter-regional variations:

$$F = \left\{ \frac{T\sum_j(\bar{y}_{*j} - \bar{y})^2/(M-1)}{\sum_t\sum_j(\bar{y}_{tj} - \bar{y}_{*j})^2/M(T-1)} \right\}. \tag{24}$$

To analyse this statistic and to find its distribution, let us use the Kronecker notation to express the vectorised equations as

$$\begin{aligned} Y^c &= \mu(\iota_M \otimes \iota_T) + (I_M \otimes i_T)\delta + \mathcal{E}^c \\ &= X\beta + \mathcal{E}^c. \end{aligned} \tag{25}$$

This is just a specialisation of equation (7) which comes from setting $\gamma = 0$. Let us also define the following projectors:

$$\begin{aligned} P_M &= \iota_M(\iota_M'\iota_M)^{-1}\iota_M', \qquad & P_T &= \iota_T(\iota_T'\iota_T)^{-1}\iota_T', \\ P_{*j} &= I_M \otimes P_T \qquad \text{and} & P &= P_M \otimes P_T. \end{aligned} \tag{26}$$

These projectors represent various averaging operators. To illustrate their effect, let us consider the vector $y_{*j} = [y_{1j}, \ldots y_{Tj}]'$ of order T. Then, since $(\iota_T'\iota_T)^{-1} = T^{-1}$ and $\iota_T' y_{*j} = \sum_t y_{tj}$, it follows that

$$\begin{aligned} P_T y_{*j} &= \iota_T(\iota_T'\iota_T)^{-1}\iota_T' y_{*j} \\ &= \bar{y}_{*j}\iota_T, \end{aligned} \tag{27}$$

which is just the vector $[\bar{y}_{*j}, \ldots \bar{y}_{*j}]'$ of order T which has the mean of the jth region as its repeated element. We can also see that $P_{*j} = I_M \otimes P_T$ is the operator which averages over time throughout the M regions. The operator P is a total averaging operator which extends over space and time. We can easily verify that

$$PP_{*j} = P_{*j}P = P \quad \text{and} \quad (P_{*j} - P)(I_{MT} - P_{*j}) = 0. \tag{28}$$

Now consider the following identity:

$$\mathcal{E}^{c\prime}(I_{MT} - P)\mathcal{E}^c = \mathcal{E}^{c\prime}(P_{*j} - P)\mathcal{E}^c + \mathcal{E}^{c\prime}(I_{MT} - P_{*j})\mathcal{E}^c. \tag{29}$$

Since the elements ε_{tj} of $\mathcal{E}^c$ are independently and identically distributed normal variates with $E(\varepsilon_{tj}) = 0$ and $V(\varepsilon_{tj}) = \sigma^2$ for all t,j, it follows immediately from Cochran's Theorem that

$$\begin{aligned} &1. \quad \frac{1}{\sigma^2}\mathcal{E}^{c\prime}(I_{MT} - P)\mathcal{E}^c \sim \chi^2(MT-1), \\ &2. \quad \frac{1}{\sigma^2}\mathcal{E}^{c\prime}(P_{*j} - P)\mathcal{E}^c \sim \chi^2(M-1), \\ &3. \quad \frac{1}{\sigma^2}\mathcal{E}^{c\prime}(I_{MT} - P_{*j})\mathcal{E}^c \sim \chi^2(MT-M), \end{aligned} \tag{30}$$

where the expressions under 2 and 3 are for mutually independent variates. Since $\mathcal{E}^c = (Y^c - X\beta)$ and since $(I_{MT} - P_{*j})(Y^c - X\beta) = (I_{MT} - P_{*j})Y^c$, it follows that the ratio of these expressions divided by their respective degrees of freedom can be written

$$F = \left\{ \frac{(Y^c - X\beta)'(P_{*j} - P)(Y^c - X\beta)/(M-1)}{Y^{c\prime}(I_{MT} - P_{*j})Y^c/M(T-1)} \right\} \tag{31}$$

and this is distributed as an $F\{M-1, M(T-1)\}$ variate. Under the null hypothesis that $\delta = 0$ in equation (25), we have

$$\begin{aligned}(P_{*j} - P)(Y^c - X\beta) &= (P_{*j} - P)\{Y^c - \mu(\iota_M \otimes \iota_T)\} \\ &= (P_{*j} - P)Y^c.\end{aligned} \tag{32}$$

Therefore, on the hypothesis that $\delta_1 = \delta_2 = \cdots = \delta_M$, the ratio in (31) becomes

$$F = \left\{ \frac{Y^{c\prime}(P_{*j} - P)Y^c/(M-1)}{Y^{c\prime}(I_{MT} - P_{*j})Y^c/M(T-1)} \right\} \tag{33}$$

It can be seen that this is precisely the statistic which appears, in alternative notation, under (24).

Hypothesis Testing in the Two-Way Model

Now let us develop tests for the standard hypotheses in the context of the two-way model. The model is expressed in the equations

$$\begin{aligned} y_{tj} &= \mu + \gamma_t + \delta_j + \varepsilon_{tj} \\ &= \mu_{t*} + \mu_{*j} - \mu + \varepsilon_{tj}, \\ \text{where} \quad t &= 1, \ldots, T \quad \text{and} \quad j = 1, \ldots, M. \end{aligned} \tag{34}$$

Here we have defined $\mu_{t*} = \mu + \gamma_t$ and $\mu_{*j} = \mu + \gamma_j$ which are respectively a time-specific and a regional mean. The model is also represented by equations (6) and (7). The sum of squares of the deviations of the observations about the overall sample average $\bar{y}$ may be decomposed as follows:

$$\begin{aligned} \sum_t \sum_j (y_{tj} - \bar{y})^2 = {} & \sum_t \sum_j (y_{tj} - \bar{y}_{t*} - \bar{y}_{*j} + \bar{y})^2 \\ & + M \sum_t (\bar{y}_{t*} - \bar{y})^2 + T \sum_j (\bar{y}_{*j} - \bar{y})^2. \end{aligned} \tag{35}$$

Using the projectors defined in (26), together with the new projector

$$P_{t*} = P_M \otimes I_T, \tag{36}$$

we can write this decomposition in the equivalent form of

$$(37)\qquad \begin{aligned} Y^{c\prime}(I_{MT}-P)Y^c &= Y^{c\prime}(I_{MT}-P_{t*}-P_{*j}+P)Y^c\\ &\quad + Y^{c\prime}(P_{t*}-P)Y^c + Y^{c\prime}(P_{*j}-P)Y^c. \end{aligned}$$

Also, if $\mathcal{E}^c = Y^c - X\beta = Y^c - \mu(\iota_M\otimes\iota_T) - (\iota_M\otimes I_T)\gamma - (I_M\otimes\iota_T)\delta$ is the normally distributed disturbance vector from equation (7), then it follows immediately from Cochran's Theorem that

$$(38)\qquad \begin{aligned} &1.\quad \frac{1}{\sigma^2}\mathcal{E}^{c\prime}(I_{MT}-P)\mathcal{E}^c \sim \chi^2(MT-1),\\ &2.\quad \frac{1}{\sigma^2}\mathcal{E}^{c\prime}(I_{MT}-P_{t*}-P_{*j}+P)\mathcal{E}^c \sim \chi^2(MT-M-T+1),\\ &3.\quad \frac{1}{\sigma^2}\mathcal{E}^{c\prime}(P_{t*}-P)\mathcal{E}^c \sim \chi^2(T-1),\\ &4.\quad \frac{1}{\sigma^2}\mathcal{E}^{c\prime}(P_{*j}-P)\mathcal{E}^c \sim \chi^2(M-1), \end{aligned}$$

where the expressions under 2, 3 and 4 are for mutually independent variates. The ratio of any two of these, divided by their respective degrees of freedom gives a chi-square variate.

Let us imagine that we wish to test the hypothesis that $\delta = 0$ which signifies that there is no inter-regional variation. This is the hypothesis for which we constructed a test in the context of the one-way model. In the two-way model, the hypothesis has the implication that

$$(39)\qquad \begin{aligned} (P_{*j}-P)\mathcal{E}^c &= (P_{*j}-P)\{Y^c - \mu(\iota_M\otimes\iota_T) - (\iota_M\otimes I_T)\gamma\}\\ &= (P_{*j}-P)Y^c. \end{aligned}$$

We also have the result that

$$(40)\qquad \begin{aligned} (I_{MT}-P_{t*}-P_{*j}+P)\mathcal{E}^c &= (I_{MT}-P_{t*}-P_{*j}+P)\\ &\qquad\times\{Y^c - \mu(\iota_M\otimes\iota_T) - (\iota_M\otimes I_T)\gamma\}\\ &= (I_{MT}-P_{t*}-P_{*j}+P)Y^c. \end{aligned}$$

Therefore, under the conditions of the hypothesis, the following statistic

$$(41)\qquad F = \left\{\frac{Y^{c\prime}(P_{*j}-P)Y^c/(M-1)}{Y^{c\prime}(I_{MT}-P_{t*}-P_{*j}+P)Y^c/(M-1)(T-1)}\right\}$$

has an $F(M-1, MT-M-T+1)$ distribution; and we can test the hypothesis by testing the validity of this implication.