ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
20, 3 (1990), pp. 367–375

M. KOZŁOWSKA and R. WALKOWIAK (Poznań)

INCOMPLETE HOMOGENEOUS MULTIRESPONSE MODELS: TESTING OF HYPOTHESES

1. Preliminaries. In our previous paper [2], a method of estimation of treatment contrasts in the incomplete homogeneous multiresponse model was presented. In this paper we consider the problem of testing the hypotheses concerning certain linear functions of treatment parameters in this model.

As in the previous paper, we assume here the following model of observations:

$$\left(y, [U', \Delta']\begin{bmatrix}\xi\\ \gamma\end{bmatrix}, \Sigma^*\right), \tag{1.1}$$

where y denotes the vector of observations, $[U', \Delta']$ is the design matrix in which U' is the matrix for all not interesting classifications connected with the vector ξ of unknown nuisance parameters, the matrix Δ' corresponds to the vector γ of unknown treatment parameters, and Σ^* is an unknown covariance matrix. This model is not an ordinary linear model because all n experimental units are grouped in u disjoint groups with respect to the observed variables. Then we can write

$$y = \begin{bmatrix}\operatorname{vec} Y_1\\ \dots\\ \operatorname{vec} Y_u\end{bmatrix}, \quad U = \begin{bmatrix} I\otimes U'_1 & \dots & 0\\ \dots & \dots & \dots\\ 0 & \dots & I\otimes U'_u\end{bmatrix},$$

$$\Delta' = \begin{bmatrix} M'_1\otimes \Delta'_1\\ \dots\\ M'_u\otimes \Delta'_u\end{bmatrix}, \quad \xi = \begin{bmatrix}\operatorname{vec}\Xi_1\\ \dots\\ \operatorname{vec}\Xi_u\end{bmatrix}, \quad \gamma = \operatorname{vec}\Gamma,$$

$$\Sigma^* = \begin{bmatrix} M'_1\Sigma M_1\otimes I & \dots & 0\\ \dots & \dots & \dots\\ 0 & \dots & M'_u\Sigma M_u\otimes I\end{bmatrix},$$

where Y_i, $i = 1, 2, \ldots, u$, are $(n_i \times t_i)$-matrices of observations, U_i' are $n_i \times s_i$ design matrices corresponding to the $(s_i \times t_i)$-matrices Ξ_i of unknown parameters, Γ is the $(v \times t)$-matrix of unknown treatment parameters, Δ_i' are the $n_i \times v$ design matrices corresponding to the $(v \times t_i)$-matrices ΓM_i, while M_i are the $(t \times t_i)$-matrices obtained from the identity matrix I through the elimination of columns corresponding to variables which are not observed in the i-th group, Σ is the $t \times t$ unknown covariance matrix. Details of the incomplete homogeneous multiresponse model are described by Kozłowska and Walkowiak [2]. Particularly, Definition 1.1 from that paper defined the property of model (1.1) called *homogeneity*. The necessary and sufficient condition for homogeneity is that for every i, i', $i \neq i' = 1, 2, \ldots, u$, the matrix $C_i X^{-1} C_{i'}$ was a symmetric one, where X is a positive definite symmetric matrix, and

$$C_i = \Delta_i \Phi_i \Delta_i', \qquad \Phi_i = I - U_i'(U_i U_i')^- U_i.$$

2. Results. Suppose that we wish to test the following hypothesis:

$$H_0\colon\ Z = 0, \tag{2.1}$$

where $Z = (d' \otimes c')\gamma$ is a certain estimable linear function of the treatment parameters, the vector c defines the contrast between treatments, and the vector d defines the linear combination of variables. It is known (see, e.g., [3]) that each function Z if it is estimable, may be written as a linear combination of estimable basic contrasts. According to the main results from [2], the least square estimator of the function Z takes the following form:

$$\begin{aligned} \hat{Z} &= \sum_{j \in T} \sum_{l \in L_j} g_{jl} \hat{Z}_{jl} \\ &= \sum_{j \in T} \sum_{l \in L_j} g_{jl} \Big(\sum_{i=1}^{u} m_{il} \lambda_{ij}\Big)^{-1} (e_l' \otimes w_j') \Delta \Phi y. \end{aligned} \tag{2.2}$$

Here g_{jl} are any numbers, $Z_{jl} = (e_l' \otimes w_j' X)\gamma$, e_l is a $(t \times 1)$-vector with the l-th element equal to 1 and the remaining elements equal to 0, w_j $(j = 1, 2, \ldots, v)$ are orthonormal common eigenvectors of all matrices C_i with respect to X corresponding to eigenvalues λ_{ij} $(i = 1, 2, \ldots, u)$, m_{il} is the (l, l)-th element of $M_i M_i'$,

$$\Phi = \begin{bmatrix} I \otimes \Phi_1 & \ldots & 0 \\ \ldots & \ldots & \ldots \\ 0 & \ldots & I \otimes \Phi_u \end{bmatrix},$$

and

$$T = \{j\colon \underset{i}{\exists}\, \lambda_{ij} > 0\}, \qquad L_j = \{l\colon \underset{i}{\exists}\, m_{il}\lambda_{ij} \neq 0\}.$$

It is easy to see that the variance of the estimator of the function Z is a linear combination of variances of estimators of basic contrasts:

$$\operatorname{Var}(\hat{Z}) = \sum_{j \in T} \sum_{l \in L_j} g_{jl}^2 \operatorname{Var}(\hat{Z}_{jl}).$$

Applying (2.2) we have

$$\mathrm{Var}(\hat{Z}_{jl}) = \Big(\sum_{i=1}^{u} m_{il}\lambda_{ij}\Big)^{-2} \sum_{i=1}^{u} \lambda_{ij}\sigma_{il},$$

where $\sigma_{il} = e_l' M_i M_i' \Sigma M_i M_i' e_l$. The estimator of the above variance of a basic contrast estimator is of the form

$$\hat{\mathrm{Var}}(\hat{Z}_{jl}) = \Big(\sum_{i=1}^{u} m_{il}\lambda_{ij}\Big)^{-2} \sum_{i=1}^{u} \lambda_{ij}\hat{\sigma}_{il},$$

where

$$\hat{\sigma}_{il} = e_l' M_i \hat{\Sigma}_i M_i' e_l. \tag{2.3}$$

The matrix $\Sigma_i = M_i'\Sigma M_i$ is the variance matrix in the ordinary multivariate model, $\hat{\Sigma}_i$ is its estimator and takes the form

$$\hat{\Sigma}_i = \frac{1}{v_i} Y_i' \Phi_i [I - \Delta_i' C_i^- \Delta_i] \Phi_i Y_i, \tag{2.4}$$

where $v_i = n_i - r(C_i) - r(U_i)$, $r(\cdot)$ denotes the rank of the matrix $(\cdot)$, and C_i^- is any generalized inverse of the matrix C_i. Hence

$$\hat{\mathrm{Var}}(\hat{Z}) = \sum_{j\in T} \sum_{l\in L_j} g_{jl}^2 \hat{\mathrm{Var}}(\hat{Z}_{jl}) = \sum_{j\in T} \sum_{l\in L_j} a_{jl}^2 \Big(\sum_{i=1}^{u} \lambda_{ij}\hat{\sigma}_{il}\Big), \tag{2.5}$$

where

$$a_{jl} = g_{jl}\Big(\sum_{i=1}^{u} m_{il}\lambda_{ij}\Big)^{-1}.$$

Now we assume that the vector of observations y has the $\big(\sum_{i=1}^{u} n_i t_i\big)$-dimensional normal distribution with expected value and dispersion matrix as in model (1.1), i.e.,

$$y \sim N\left([U', \Delta']\begin{bmatrix}\xi\\ \gamma\end{bmatrix}, \Sigma^*\right).$$

Hence the estimator of the function Z described in (2.2) has the distribution

$$\hat{Z} \sim N\Big(Z, \sqrt{\sum_{j\in T}\sum_{l\in L_j} a_{jl}^2 \Big(\sum_{i=1}^{u} \lambda_{ij}\sigma_{il}\Big)}\Big).$$

Therefore

$$\xi^2 = \frac{\Big(\sum_{j\in T}\sum_{l\in L_j} a_{jl}(e_l' \otimes w_j')\Delta\Phi y\Big)^2}{\sum_{j\in T}\sum_{l\in L_j} a_{jl}^2 \Big(\sum_{i=1}^{u} \lambda_{ij}\sigma_{il}\Big)}$$

has the noncentral χ^2-distribution with 1 degree of freedom and noncentrality parameter

$$\frac{Z}{\sum_{j\in T}\sum_{l\in L_j} a_{jl}^2\left(\sum_{i=1}^{u}\lambda_{ij}\sigma_{il}\right)}.$$

If the hypothesis H_0 is true, ξ^2 has the central χ^2-distribution.

Now, let us consider the distribution of estimator (2.5). It is known (see, e.g., [4]) that estimator (2.4) has the central Wishart distribution $W(M_i'\Sigma M_i, v_i)$ with v_i degrees of freedom. Then $\hat{\sigma}_{il}$ $(i = 1, 2, \ldots, u;\ l = 1, 2, \ldots, t)$ have independent distributions

$$\hat{\sigma}_{il} \sim \sigma_{il}\chi_{v_i}^2/v_i.$$

Hence

$$\hat{\mathrm{Var}}(\hat{Z}) \sim \sum_{j\in T}\sum_{l\in L_j} a_{jl}^2 \sum_{i=1}^{u} (\lambda_{ij}\sigma_{il}/v_i)\chi_{v_i}^2.$$

Since the distribution of the estimator of the variance of the contrast estimator Z is a linear combination of χ^2 variables, we cannot use the exact F-test for testing the hypothesis (2.1). Now we try to find an approximate F-test (see, e.g., [5]). Let us approximate

$$\sum_{j\in T}\sum_{l\in L_j} a_{jl}^2 \sum_{i=1}^{u} (\lambda_{ij}\sigma_{il}/v_i)\chi_{v_i}^2 \tag{2.6}$$

with a random variable of the form $a\chi_v^2$, where a and v are determined so that (2.6) and $a\chi_v^2$ have the same expected value and variance. This gives

$$\sum_{j\in T}\sum_{l\in L_j} a_{jl}^2\left(\sum_{i=1}^{u}\lambda_{ij}\sigma_{il}\right) = av,$$

$$\sum_{j\in T}\sum_{l\in L_j} 2a_{jl}^4\left(\sum_{i=1}^{u}\lambda_{ij}^2\sigma_{il}^2/v_i\right) = 2a^2 v.$$

Hence

$$a = \frac{1}{v}\sum_{j\in T}\sum_{l\in L_j} a_{jl}^2\left(\sum_{i=1}^{u}\lambda_{ij}\sigma_{il}\right),$$

$$v = \frac{\left(\sum_{j\in T}\sum_{l\in L_j} a_{jl}^2\sum_{i=1}^{u}\lambda_{ij}\sigma_{il}\right)^2}{\sum_{j\in T}\sum_{l\in L_j} a_{jl}^4\sum_{i=1}^{u}(\lambda_{ij}^2\sigma_{il}^2/v_i)}. \tag{2.7}$$

Now we may state that $\hat{\mathrm{Var}}(\hat{Z})$ has the following approximate distribution:

$$\hat{\mathrm{Var}}(\hat{Z}) \dot{\sim} \sum_{j\in T}\sum_{l\in L_j} a_{jl}^2\left(\sum_{i=1}^{u}\lambda_{ij}\sigma_{il}\right)\chi_v^2/v;$$

therefore

$$\hat{\sigma} = \frac{\hat{\mathrm{Var}}(\hat{Z})}{\sum_{j\in T}\sum_{l\in L_j} a_{jl}^2\left(\sum_{i=1}^{u}\lambda_{ij}\sigma_{il}\right)} \mathrel{\dot{\sim}} \chi_v^2/v.$$

The approximate noncentral F-test for testing the hypothesis (2.1) is then obtained by using the ratio

$$\frac{\xi^2}{\hat{\sigma}} = \frac{\left(\sum_{j\in T}\sum_{l\in L_j} a_{jl}(e_l'\otimes w_j')\Delta\Phi y\right)^2}{\sum_{j\in T}\sum_{l\in L_j} a_{jl}^2\left(\sum_{i=1}^{u}\lambda_{ij}\hat{\sigma}_{il}\right)}, \tag{2.8}$$

where $\hat{\sigma}_{il}$ is described in (2.3), and by proceeding as though (2.8) was distributed as

$$\frac{v}{\chi_v^2}\chi_A^2, \quad \text{where } A = \left(1, \frac{Z}{\sum_{j\in T}\sum_{l\in L_j} a_{jl}^2\left(\sum_{i=1}^{u}\lambda_{ij}\sigma_{il}\right)}\right).$$

It is obvious that (2.8) has the approximate noncentral distribution F with 1 and v degrees of freedom and with noncentrality parameter

$$Z^2/\sum_{j\in T}\sum_{l\in L_j} a_{il}^2\left(\sum_{i=1}^{u}\lambda_{ij}\sigma_{il}\right).$$

If the hypothesis H_0 is true, it is the central approximate F-distribution. Hence we reject the hypothesis H_0 when

$$\xi^2/\hat{\sigma} > F_{\alpha,1,v},$$

where α is the level of significance. Unfortunately, the constant v is presented in terms of unknown parameters. In this situation we must estimate it. We obtain the required estimator $\hat{v}$ replacing σ_{il} by $\hat{\sigma}_{il}$ in (2.7). Notice that the constant v satisfies the following relation:

$$\min_i v_i \leqslant v \leqslant \sum_{i=1}^{u} v_i.$$

When the value of the ratio (2.8) is greater than $F_{\alpha,1,d}$, where

$$d = \min_i v_i$$

or (2.8) is not greater than $F_{\alpha,1,g}$, where

$$g = \sum_{i=1}^{u} v_i,$$

then we need not estimate v. In the first case we reject the hypothesis H_0, and in the other one we do not reject it. In every other case we must calculate $\hat{v}$.

A particularly interesting case of hypothesis (2.1) is

$$H_0\colon\ Z_{jl} = 0, \quad \text{where } l \in L_j,\ j \in T.$$

The ratio (2.8) takes the form

$$\frac{\xi^2}{\hat{\sigma}} = \frac{((e_l' \otimes w_j')\Delta\Phi y)^2}{\sum_{i=1}^{u} \lambda_{ij}\hat{\sigma}_{il}},$$

where $\hat{\sigma}_{il}$ is described in (2.3). Now the constant $\hat{v}$ takes the form

$$\hat{v} = \frac{\left(\sum_{i=1}^{u} \lambda_{ij}\hat{\sigma}_{il}\right)^2}{\sum_{i=1}^{u} \lambda_{ij}^2\hat{\sigma}_{il}^2/v_i}.$$

3. Example. Now we want to explain the above theory by an example. We will use the experiment described by Kozłowska [1].

The considered experiment was carried out in an incomplete multiresponse block design homogeneous with respect to the identity matrix ($X = I$). Six treatments are applied to $n = 30$ experimental units which are grouped in two sets, $u = 2$, $n_1 = 12$, $n_2 = 18$. Five variables are observed: the first, second and third ones on units from the first set, and the first, second, fourth and fifth on the units from the second one. Hence the matrices M_i ($i = 1, 2$) are as follows:

$$M_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad M_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

The scheme of this experiment may be described by the incidence matrices

$$N_1 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \quad N_2 = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

We are interested in testing the hypothesis concerning the contrast between third and sixth treatments for the sum of the first and third variables. It is

$$H_0\colon\ ([1\ \ 0\ \ 1\ \ 0\ \ 0] \otimes [0\ \ 0\ \ 1\ \ 0\ \ 0\ \ -1])\gamma = 0. \tag{3.1}$$

This contrast is estimable and it may be written in the form of a linear combination of basic contrasts $(e_i' \otimes w_j')\gamma$. The vectors $w_j, j \in T = \{1, 2, 3, 4, 5\}$, are common eigenvectors of the matrices

$$C_1 = \frac{1}{3}\begin{bmatrix} 4 & -1 & -1 & -1 & -1 & 0 \\ -1 & 4 & 0 & -1 & -1 & -1 \\ -1 & 0 & 4 & -1 & -1 & -1 \\ -1 & -1 & -1 & 4 & 0 & -1 \\ -1 & -1 & -1 & 0 & 4 & -1 \\ 0 & -1 & -1 & -1 & -1 & 4 \end{bmatrix},$$

$$C_2 = \frac{1}{3}\begin{bmatrix} 6 & -1 & -1 & -1 & -1 & -2 \\ -1 & 6 & -2 & -1 & -1 & -1 \\ -1 & -2 & 6 & -1 & -1 & -1 \\ -1 & -1 & -1 & 6 & -2 & -1 \\ -1 & -1 & -1 & -2 & 6 & -1 \\ -2 & -1 & -1 & -1 & -1 & 6 \end{bmatrix}$$

and take the following forms:

$$w_1 = (1/\sqrt{12})[-2 \quad 1 \quad 1 \quad 1 \quad 1 \quad -2]',$$

$$w_2 = (1/2)[0 \quad 1 \quad 1 \quad -1 \quad -1 \quad 0]',$$

$$w_3 = (1/\sqrt{6})[-1 \quad -1 \quad 1 \quad 1 \quad -1 \quad 1]',$$

$$w_4 = (1/\sqrt{12})[2 \quad -1 \quad 1 \quad 1 \quad -1 \quad -2]',$$

$$w_5 = (1/2)[0 \quad -1 \quad 1 \quad -1 \quad 1 \quad 0]'.$$

These vectors correspond to the eigenvalues

$$\lambda_{11} = \lambda_{12} = 2, \qquad \lambda_{13} = \lambda_{14} = \lambda_{15} = 4/3$$

and

$$\lambda_{21} = \lambda_{22} = 2, \qquad \lambda_{23} = \lambda_{24} = \lambda_{25} = 8/3.$$

We can write

$$\begin{aligned} Z &= ([1 \ 0 \ 1 \ 0 \ 0] \otimes [0 \ 0 \ 1 \ 0 \ 0 \ -1])\gamma \\ &= (\sqrt{3}/2)(e_1' \otimes w_1')\gamma + (1/2)(e_1' \otimes w_2')\gamma + (\sqrt{3}/2)(e_1' \otimes w_4')\gamma \\ &\quad + (1/2)(e_1' \otimes w_5')\gamma + (\sqrt{3}/2)(e_3' \otimes w_1')\gamma + (1/2)(e_3' \otimes w_2')\gamma \\ &\quad + (\sqrt{3}/2)(e_3' \otimes w_4')\gamma + (1/2)(e_3' \otimes w_5')\gamma. \end{aligned}$$

Now we estimate the considered contrast and variance of its estimator applying formulas (2.2), (2.5) and the values of observations from Table 1.

TABLE 1. Observations of variables from experimental units

No. of unit from the first set	1	2	3	4	5	6
	7	8	9	10	11	12
No. of block	1	1	1	2	2	2
	3	3	3	4	4	4
No. of treatment	1	2	5	1	3	4
	2	4	6	3	5	6
First variable	112.1	94.2	83.7	99.6	98.6	64.3
	86.9	48.9	58.4	90.3	70.1	54.6
Second variable	146.6	112.4	101.1	130.5	104.6	73.4
	94.3	68.2	76.4	100.1	99.3	77.2
Third variable	23.6	22.9	13.9	21.8	16.8	17.6
	19.6	12.6	15.1	12.5	11.5	9.4
No. of unit from the second set	1	2	3	4	5	6
	7	8	9	10	11	12
	13	14	15	16	17	18
No. of block	1	1	1	2	2	2
	3	3	3	4	4	4
	5	5	5	6	6	6
No. of treatment	1	2	6	1	3	6
	1	4	5	2	3	4
	2	3	5	4	5	6
First variable	96.8	99.3	68.2	80.6	69.9	61.1
	106.4	66.2	76.5	80.1	82.3	71.3
	68.4	78.9	76.6	52.4	90.0	49.1
Second variable	120.1	106.7	79.4	108.6	94.5	80.1
	136.1	78.2	90.5	99.4	102.1	94.3
	81.5	87.8	92.4	69.4	105.1	62.5
Fourth variable	7.84	7.12	6.51	8.04	7.60	6.83
	7.95	6.68	7.72	7.30	7.45	6.92
	7.15	7.38	7.85	6.90	7.95	6.75
Fifth variable	9.6	9.2	8.7	10.8	9.4	8.3
	10.4	8.9	9.2	8.9	9.6	8.6
	9.3	9.0	9.0	9.0	9.4	9.1

(Notice that, for $j = 1, 2, 3, 4, 5$, $L_j = \{1, 2, 3, 4, 5\}$.) From (2.2) we see that the estimator of the considered contrast is equal to 137.33. The estimators of covariance matrices (2.4) are

$$\hat{\Sigma}_1 = \begin{bmatrix} 637.979 & 34.550 & 159.833 \\ 34.550 & 25.164 & 38.004 \\ 159.833 & 38.004 & 115.063 \end{bmatrix},$$

$$\hat{\Sigma}_2 = \begin{bmatrix} 2597.150 & 512.828 & 1.705 & -3.041 \\ 512.828 & 115.101 & 0.465 & -0.788 \\ 1.705 & 0.465 & 0.003 & -0.010 \\ -3.041 & -0.788 & -0.010 & 0.138 \end{bmatrix}$$

and from (2.3) and (2.5) we get

$$\hat{\mathrm{Var}}(\hat{Z}) = 1034.243.$$

Hence and from (2.8) we obtain

$$\frac{\xi^2}{\hat{\sigma}} = \frac{(137.33)^2}{1034.243} = 18.235.$$

Since the ratio (2.8) is greater than $F_{0.05,1,3} = 10.1$, we reject the hypothesis (3.1) without estimating the degrees of freedom v.

References

[1] M. Kozłowska, *Analiza doświadczenia wielocechowego założonego w układzie blokowym wieloreakcjowym niekompletnym*, Dziesiąte Colloquium Metodologiczne z Agro-Biometrii, 1979, pp. 254–275.

[2] – and R. Walkowiak, *Incomplete homogeneous multiresponse models: Estimation*, Zastos. Mat. 20 (1988), pp. 211–217.

[3] S. C. Pearce, T. Caliński and T. F. de C. Marshall, *The basic contrasts of an experimental design with special reference to the analysis of data*, Biometrika 61 (1974), pp. 449–460.

[4] C. R. Rao, *Linear Statistical Inference and its Applications*, J. Wiley, New York 1965.

[5] H. Scheffé, *The Analysis of Variance*, J. Wiley, New York 1959.

MARIA KOZŁOWSKA, RYSZARD WALKOWIAK
DEPARTMENT OF MATHEMATICAL AND STATISTICAL METHODS
ACADEMY OF AGRICULTURE
60-637 POZNAŃ

Received on 1987.08.11;
revised version on 1988.04.22