J. Sci. Hiroshima Univ. Ser. A-I
**32** (1968) 357-370

# *Two-way Classification Designs with Unequal Cell Frequencies*

Sumiyasu Yamamoto and Yasunori Fujikoshi

(Received September 20, 1968)

## 1. Introduction and Summary

In a two-way classification design on two factors, say $A$ and $B$, we apply each factor on varying levels to various experimental units. We assume that this application yields for each unit a quantity which we call the yield of this unit. We denote by $\eta(I, J)$ the mean value of the yield obtained when the factor $A$ is applied at level $I$ and the factor $B$ at level $J$. These levels may be qualitative or quantitative and could assume discrete or continuous values. Usually they are chosen deterministically by the experimenter. In some cases, however, they are selected randomly according to a probability scheme. Even when the levels vary continuously, the experimenter can calibrate or can group them into a finite number of discrete values. We, therefore, assume that $I$ and $J$ can take the values $1, 2, \cdots, r$ and $1, 2, \cdots, s$, respectively.

The object of a two-way classification design is to make some inferences on the behavior of the mean yield function $\eta(I, J)$. For such purpose, the function $\eta(I, J)$ is usually broken up into a general mean $\mu$, a main effect $\alpha(I)$ of the factor $A$, a main effect $\beta(J)$ of the factor $B$, and an interaction effect $\gamma(I, J)$ ascribed to the combination of level $I$ of the factor $A$ with level $J$ of the factor $B$, i.e.,

$$\eta(I, J)=\mu+\alpha(I)+\beta(J)+\gamma(I, J). \tag{1.1}$$

If nothing more is stated about the decomposition, these components of the decomposition are not uniquely defined. It is for this reason to impose some constraints among these components.

In order to seek for a set of reasonable and intutively acceptable constraints, we introduce a non-negative weight function $W(I, J)$ associated to a pattern of the yield function $\eta(I, J)$. A purpose of introducing such a weight function is to develop a unified treatment of the identification problem in the decomposition (1.1) of the yield function.

The weight function $W(I, J)$ might be considered as an apriori probability function over the combination of the levels of the experiment in the wide sense (the total mass of the distribution of $W(I, J)$ may not be necessarily unity).

Upon introducing the weight function $W(I, J)$, the general mean $\mu$ is defined as an overall expectation of $\eta(I, J)$ with respect to the weight function

$W(I, J)$ by

$$\mu=\frac{\sum_{i=1}^{r}\sum_{j=1}^{s}W(i, j)\eta(i, j)}{\sum_{i=1}^{r}\sum_{j=1}^{s}W(i, j)}. \tag{1.2}$$

The following constraints imposed on the other components might be reasonable and intutively acceptable, i.e.,

$$\begin{aligned}&\sum_{i=1}^{r}W(i, \cdot)\alpha(i)=0, \quad \sum_{j=1}^{s}W(\cdot, j)\beta(j)=0\\ &\sum_{i=1}^{r}W(i, J)\gamma(i, J)=0 \qquad \text{for all } J\\ &\sum_{j=1}^{s}W(I, j)\gamma(I, j)=0 \qquad \text{for all } I\end{aligned} \tag{1.3}$$

where $W(I, \cdot)=\sum_{j=1}^{s}W(I, j)$ and $W(\cdot, J)=\sum_{i=1}^{r}W(i, J)$.

It is easy to see that the constraints introduced in a usual two-way classification design are based on a special weight function with $W(I, J)=1$ for all $I$ and $J$.

Scheffé [6] introduced a system of non-negative weights $U(I)$ and $V(J)$ with $\sum_{i=1}^{r}U(i)=1$ and $\sum_{j=1}^{s}V(j)=1$, and imposed a set of constraints on the components for the identification of the decomposition;

$$\begin{aligned}&\sum_{i=1}^{r}U(i)\alpha(i)=0, \quad \sum_{j=1}^{s}V(j)\beta(j)=0\\ &\sum_{i=1}^{r}U(i)\gamma(i, J)=0 \qquad \text{for all } J\\ &\sum_{j=1}^{s}V(j)\gamma(I, j)=0 \qquad \text{for all } I.\end{aligned} \tag{1.4}$$

This set of constraints is based on a special weight function with $W(I, J)=U(I)V(J)$ for all $I$ and $J$, provided all $U(I)$ and $V(J)$ are positive. In order to arrive at Scheffé's constraints when $U(I)=0$ for some $I$ (and/or $V(J)=0$ for some $J$), all we have to do is to impose one additional constraint $\sum_{j=1}^{s}V(j)\gamma(I, j)=0$ (and/or $\sum_{i=1}^{r}U(i)\gamma(i, J)=0$). Therefore no special mention will be made of such cases.

In a two-way classification design with unequal cell frequencies, it can be seen in Section 4 that the use of the weights $W(I, J)$ proportional to the cell frequencies $n(I, J)$ is necessary to explain the appropriateness of the analysis of variances described without introducing the definition of the components in (1.1) by Rao [4]. Such a weight function is neither a usual one nor of Scheffé's type.

In section 2 we show the consistency of our constraints (1.3) and derive the necessary and sufficient conditions for a set of weights to be identifiable. These conditions show that the weight function $W(I, J)$ may assume zero without violating the identifiability. A balanced incomplete block design and a partially balanced incomplete block design with replication may be considered as typical examples of such a two-way classification design. In section 3 we investigate some properties of the components. In section 4 we treat the analysis of variances of the two-way classification design under our general model. An important special case of interest will also be considered there. In section 5 we treat the orthogonality property in the analysis of variances.

The considerations given in this paper may be generalized to the multi-way classification design*. Similar considerations may also be possible for the multivariate analysis of variances.

For simplicity, we denote $\alpha(I)$, $\beta(J)$, $\gamma(I, J)$, $W(I, J)$, $W(I, \cdot)$, $W(\cdot, J)$, $\eta(I, J)$ and $n(I, J)$ by $\alpha_i$, $\beta_j$, $\gamma_{ij}$, $w_{ij}$, $w_{i\cdot}$, $w_{\cdot j}$, $\eta_{ij}$ and $n_{ij}$, respectively in the following discussion.

## 2. Identifiability

We shall first prove the consistency of our constraints (1.3) in defining the components $\mu$, $\alpha_i$, $\beta_j$ and $\gamma_{ij}$.

(a) Consistency

Let $\boldsymbol{\eta}'=(\eta_{11}, \eta_{12}, \cdots, \eta_{rs-1}, \eta_{rs})$, $\boldsymbol{\alpha}'=(\alpha_1, \cdots, \alpha_r)$, $\boldsymbol{\beta}'=(\beta_1, \cdots, \beta_s)$, and $\boldsymbol{\gamma}'=(\gamma_{11}, \gamma_{12}, \cdots, \gamma_{rs-1}, \gamma_{rs})$. Then the decomposition (1.1) and the constraints (1.3) can be expressed in matrix notation as follows;

$$\boldsymbol{\eta}=\boldsymbol{j}_{rs}\mu+[I_r\otimes\boldsymbol{j}_s]\boldsymbol{\alpha}+[\boldsymbol{j}_r\otimes I_s]\boldsymbol{\beta}+[I_r\otimes I_s]\boldsymbol{\gamma} \tag{2.1}$$

$$\boldsymbol{j}'_{rs}D_w[I_r\otimes\boldsymbol{j}_s]\boldsymbol{\alpha}=0, \quad \boldsymbol{j}'_{rs}D_w[\boldsymbol{j}_r\otimes I_s]\boldsymbol{\beta}=0$$

$$\begin{bmatrix} I_r\otimes\boldsymbol{j}'_s \\ \boldsymbol{j}'_r\otimes I_s \end{bmatrix} D_w\boldsymbol{\gamma}=0 \tag{2.2}$$

where $\boldsymbol{j}'_p=(1, \cdots, 1)$, $I_p$ denotes the unit matrix of order $p$, $X\otimes Y$ denotes the Kronecker product of $X$ and $Y$, and $D_w=\text{diag}(w_{11}, w_{12}, \cdots, w_{rs-1}, w_{rs})$.

The problem is to prove the consistency of the constraints, i.e., the matrix equation

* (c.f. Yamamoto and Fujikoshi (1967). Models and ANOVA in factorial designs (in Japanese). *Res. Memoir of Res. Inst. Math. Sci. (Kyoto).* No. 25, p. 240-260.)

$$(2.3)\quad \begin{pmatrix} \boldsymbol{j}_r\otimes\boldsymbol{j}_s & I_r\otimes\boldsymbol{j}_s & \boldsymbol{j}_r\otimes I_s & I_r\otimes I_s \\ 0 & \boldsymbol{j}'_{rs}D_w[I_r\otimes\boldsymbol{j}_s] & 0 & 0 \\ 0 & 0 & \boldsymbol{j}'_{rs}D_w[\boldsymbol{j}_r\otimes I_s] & 0 \\ 0 & 0 & 0 & \begin{bmatrix} I_r\otimes\boldsymbol{j}'_s \\ \boldsymbol{j}'_r\otimes I_s \end{bmatrix} D_w \end{pmatrix} \begin{pmatrix} \mu \\ \boldsymbol{\alpha} \\ \boldsymbol{\beta} \\ \boldsymbol{\gamma} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\eta} \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

has at least a solution with respect to $\mu$, $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ for every $\boldsymbol{\eta}$ in the $rs$ dimensional Euclidean space $E_{rs}$.

THEOREM 1. *The equation* (2.3) *is consistent.*

PROOF. It is sufficient to show that

$$(2.4)\quad \mathcal{R}\begin{pmatrix} \boldsymbol{j}'_r\otimes\boldsymbol{j}'_s \\ I_r\otimes\boldsymbol{j}'_s \\ \boldsymbol{j}'_r\otimes I_s \\ I_r\otimes I_s \end{pmatrix} \cap \mathcal{R}\begin{pmatrix} 0 & 0 & 0 \\ [I_r\otimes\boldsymbol{j}'_s]D_w\boldsymbol{j}_{rs} & 0 & 0 \\ 0 & [\boldsymbol{j}'_r\otimes I_s D_w]\boldsymbol{j}_{rs} & 0 \\ 0 & 0 & D_w[I_r\otimes\boldsymbol{j}_s, \boldsymbol{j}_r\otimes I_s] \end{pmatrix} = \{0\}$$

(c.f. Yamamoto and Fujikoshi [7], p. 213), where $\mathcal{R}[X]$ denotes the space spanned by the column vectors of $X$. Suppose that

$$\boldsymbol{\xi} = \begin{pmatrix} \boldsymbol{j}'_r\otimes\boldsymbol{j}'_s \\ I_r\otimes\boldsymbol{j}'_s \\ \boldsymbol{j}'_r\otimes I_s \\ I_r\otimes I_s \end{pmatrix} [\boldsymbol{u}] = \begin{pmatrix} 0 & 0 & 0 \\ [I_r\otimes\boldsymbol{j}'_s D_w]\boldsymbol{j}_{rs} & 0 & 0 \\ 0 & [\boldsymbol{j}'_r\otimes I_s]D_w\boldsymbol{j}_{rs} & 0 \\ 0 & 0 & D_w[I_r\otimes\boldsymbol{j}_s, \boldsymbol{j}_r\otimes I_s] \end{pmatrix} \begin{pmatrix} a \\ b \\ \boldsymbol{c} \\ \boldsymbol{d} \end{pmatrix}$$

holds for some vectors $\boldsymbol{u}'=(u_{11}, u_{12}, \cdots, u_{rs})$ and $(a, b, \boldsymbol{c}', \boldsymbol{d}')=(a, b, c_1, \cdots, c_r, d_1, \cdots, d_s)$. Then we have

$$\begin{aligned} \boldsymbol{\xi}' &= (u_{..}, u_{1.}, \cdots, u_{r.}, u_{.1}, \cdots, u_{.s}, u_{11}, u_{12}, \cdots, u_{rs}) \\ &= \big(0, aw_{1.}, \cdots, aw_{r.}, bw_{.1}, \cdots, bw_{.s}, (c_1+d_1)w_{11}, (c_1+d_2)w_{12}, \cdots, (c_r+d_s)w_{rs}\big) \end{aligned}$$

where $u_{..}=\sum_{i=1}^{r}\sum_{j=1}^{s}u_{ij}$, $u_{i.}=\sum_{j=1}^{s}u_{ij}$, and $u_{.j}=\sum_{i=1}^{r}u_{ij}$. If $w_{ij}=0$ for all $i$ and $j$, we easily get $\boldsymbol{\xi}=0$. Assume that $w_{..}=\sum_{i=1}^{r}\sum_{j=1}^{s}w_{ij}>0$. Since $aw_{..}=u_{..}$, $bw_{..}=u_{..}$ and $u_{..}=0$, we have $a=b=0$. Thus we have $u_{i.}=0$, $u_{.j}=0$ for all $i$ and $j$. Applying these conditions for $u_{ij}=(c_i+d_j)w_{ij}$, we have

$$\begin{bmatrix} I_r\otimes\boldsymbol{j}'_s \\ \boldsymbol{j}'_r\otimes I_s \end{bmatrix} D_w[I_r\otimes\boldsymbol{j}_s, \boldsymbol{j}_r\otimes I_s] \begin{bmatrix} \boldsymbol{c} \\ \boldsymbol{d} \end{bmatrix} = 0.$$

Since $D_w$ is positive semi-definite, we get $D_w[I_r\otimes \boldsymbol{j}_s, \boldsymbol{j}_r\otimes I_s][\boldsymbol{c}', \boldsymbol{d}']'=0$, i.e., $(c_i+d_j)w_{ij}=0$ for all $i$ and $j$. Thus we have $\boldsymbol{\xi}=0$. This completes the proof.

(b) Identifiability of a set of weights

Definition; A set of weights $\{w_{ij}\}$ is called identifiable if and only if the solution ($\mu$, $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$) of the equation (2.3) is uniquely determined for every $\boldsymbol{\eta}$ in $\mathrm{E}_{rs}$.

We have the following

THEOREM 2. *A set of weights $\{w_{ij}\}$ is identifiable if and only if one of the following conditions (i)$\sim$(vi) holds;*

(i) rank $D_w[I_r\otimes \boldsymbol{j}_s, \boldsymbol{j}_r\otimes I_s]=r+s-1$.

(ii) rank $\begin{bmatrix} D_r & W \\ W' & D_s \end{bmatrix} = \operatorname{rank} D_w(I_{rs}-A^{\#}_{11})D_w = r+s-1$.

(iii) rank $[D_r - WD_s^{-1}W'] = r-1$.

(iv) *The maximum latent root of $D_r^{-\frac{1}{2}}WD_s^{-1}W'D_r^{-\frac{1}{2}}$ is simple.*

(v) *There exists a positive integer $l$ such that each element of $(D_r^{-\frac{1}{2}}WD_s^{-1}W'D_r^{-\frac{1}{2}})^l$ is positive.*

(vi) *For any $i$, $i'=1, \cdots, r$ there exists a chain $i, j_1, i_1, j_2, \cdots, i_{l-1}, j_l, i'$ such that $w_{ij_1}w_{i_1j_1}w_{i_1j_2}\cdots w_{i_{l-1}j_l}w_{i'j_l}>0$.*

*Where* $A^{\#}_{11}=\left(I_r-\frac{1}{r}G_r\right)\otimes\left(I_s-\frac{1}{s}G_s\right)$, $G_p=\boldsymbol{j}_p\boldsymbol{j}_p'$, $D_r=\operatorname{diag}(w_{1.}, \cdots, w_{r.})$, $D_s=\operatorname{diag}(w_{.1}, \cdots, w_{.s})$, and $W=\|w_{ij}\|$.

PROOF. The proof of Theorem 1 shows that a set of weights $\{w_{ij}\}$ is identifiable if and only if the rank of the coefficient matrix of the equation (2.3) is $rs+r+s+1$. From (2.4) it can be easily seen that the rank of the coefficient matrix of the equation (2.3) equals to $rs+2+\operatorname{rank} D_w[I_r\otimes \boldsymbol{j}_s, \boldsymbol{j}_r\otimes I_s]$. This shows that a set of weights $\{w_{ij}\}$ is identifiable if and only if (i) holds.

Since $D_w$ is positive semi-definite and $\operatorname{rank} X=\operatorname{rank} X'X=\operatorname{rank} XX'$, we have

$$(2.5)\qquad \operatorname{rank} D_w[I_r\otimes \boldsymbol{j}_s, \boldsymbol{j}_r\otimes I_s]=\operatorname{rank} D_w^{\frac{1}{2}}[I_r\otimes \boldsymbol{j}_s, \boldsymbol{j}_r\otimes I_s]$$
$$=\operatorname{rank}\begin{bmatrix} I_r\otimes \boldsymbol{j}_s' \\ \boldsymbol{j}_r'\otimes I_s \end{bmatrix} D_w[I_r\otimes \boldsymbol{j}_s, \boldsymbol{j}_r\otimes I_s]=\operatorname{rank}\begin{bmatrix} D_r & W \\ W' & D_s \end{bmatrix}.$$

Furthermore,

$$(2.6)\quad \operatorname{rank} D_w[I_r\otimes \boldsymbol{j}_s, \boldsymbol{j}_r\otimes I_s]=\operatorname{rank} D_w[I_r\otimes \boldsymbol{j}_s, \boldsymbol{j}_r\otimes I_s][I_r\otimes \boldsymbol{j}_s, \boldsymbol{j}_r\otimes I_s]'D_w$$
$$=\operatorname{rank} D_w\left(sA^{\#}_{10}+rA^{\#}_{01}+(r+s)A^{\#}_{00}\right)D_w=\operatorname{rank}(\sqrt{s}\,A^{\#}_{10}+\sqrt{r}\,A^{\#}_{01}+\sqrt{r+s}\,A^{\#}_{00})D_w$$
$$=\operatorname{rank}(A^{\#}_{10}+A^{\#}_{01}+A^{\#}_{00})D_w=\operatorname{rank}(I_{rs}-A^{\#}_{11})D_w=\operatorname{rank} D_w(I_{rs}-A^{\#}_{11})D_w$$

where $A^{\#}_{10}\equiv\left(I_r-\frac{1}{r}G_r\right)\otimes\frac{1}{s}G_s$, $A^{\#}_{01}\equiv\frac{1}{r}G_r\otimes\left(I_s-\frac{1}{s}G_s\right)$, $A^{\#}_{00}\equiv\frac{1}{r}G_r\otimes\frac{1}{s}G_s$ and $A^{\#}_{11}$ defined above are mutually orthogonal symmetric idempotent matrices. Thus the conditions (i) and (ii) are equivalent.

The condition (ii) shows that $w_{i\cdot}$ and $w_{\cdot j}$ must be positive for all $i$ and $j$. So we have

$$\text{(2.7)}\qquad \operatorname{rank}\begin{bmatrix} D_r & W \\ W' & D_s \end{bmatrix}=\operatorname{rank}\begin{bmatrix} I_r & -WD_s^{-1} \\ 0 & I_s \end{bmatrix}\begin{bmatrix} D_r & W \\ W' & D_s \end{bmatrix}$$

$$=\operatorname{rank}\begin{bmatrix} D_r-WD_s^{-1}W' & 0 \\ W' & D_s \end{bmatrix}=\operatorname{rank}[D_r-WD_s^{-1}W']+s.$$

Hence we have (iii). The converse is obvious.

Since each element of the matrix $D_r^{-1}WD_s^{-1}W'$ is non-negative and $D_r^{-1}WD_s^{-1}W'\boldsymbol{j}_r=\boldsymbol{j}_r$ holds, the matrix $D_r^{-1}WD_s^{-1}W'$ is a stochastic matrix. By the well known results concerning the bound of latent roots of a stochastic matrix (c.f. Bartlett [1], p. 52), we have $|\rho|\leq 1$, where $\rho$ denotes a latent root of the matrix $D_r^{-1}WD_s^{-1}W'$. We, moreover, can show that $0\leq\rho\leq 1$, since the latent roots of $D_r^{-1}WD_s^{-1}W'$ coincide with those of the symmetric positive semi-definite matrix $D_r^{-\frac{1}{2}}WD_s^{-1}W'D_r^{-\frac{1}{2}}=(D_r^{-\frac{1}{2}}WD_s^{-\frac{1}{2}})(D_r^{-\frac{1}{2}}WD_s^{-\frac{1}{2}})'$. The condition rank $(D_r-WD_s^{-1}W')=r-1$ is equivalent to the condition $\operatorname{rank}(I_r-D_r^{-1}WD_s^{-1}W')=r-1$. The latter is necessary and sufficient for that $\rho=1$ is simple. Thus $\rho=1$ is simple and the maximum latent root of the matrix $D_r^{-1}WD_s^{-1}W'$. Hence (iii) and (iv) are equivalent.

Suppose that (iv) holds, then we have the spectral resolution $D_r^{-\frac{1}{2}}WD_s^{-1}W'D_r^{-\frac{1}{2}}=B^{\#}_1+\rho_2B^{\#}_2+\cdots+\rho_qB^{\#}_q$, where $0\leq\rho_i<1$, $i=2,\dots,q$ and $B^{\#}_1=\frac{1}{w_{\cdot\cdot}}D_r^{\frac{1}{2}}G_rD_r^{\frac{1}{2}}$, $B^{\#}_2,\dots,B^{\#}_q$ are mutually orthogonal idempotent matrices. Since each element of $B^{\#}_1$ is positive and $(D_r^{-\frac{1}{2}}WD_s^{-1}W'D_r^{-\frac{1}{2}})^l=B^{\#}_1+\rho_2^lB^{\#}_2+\cdots+\rho_q^lB^{\#}_q$, we have (v). Conversely, suppose that (v) holds. From Perron's theorem (c.f. Bellmann [2], p. 278) the maximum latent root of $(D_r^{-\frac{1}{2}}WD_s^{-1}W'D_r^{-\frac{1}{2}})^l$ is simple. Since $0\leq\rho\leq 1$, we have (iv).

The equivalence of (v) and (vi) can be easily seen by using the facts that $w_{ij}\geq 0$ for all $i$, $j$ and $(i, i')$ element of $(D_r^{-\frac{1}{2}}WD_s^{-1}W'D_r^{-\frac{1}{2}})^l$ can be expressed as

$$\sum_{j_1,i_1,\dots,i_{l-1}j_l}\frac{w_{ij_1}}{\sqrt{w_{i\cdot}w_{\cdot j_1}}}\frac{w_{i_1j_1}}{\sqrt{w_{i_1\cdot}w_{\cdot j_1}}}\frac{w_{i_1j_2}}{\sqrt{w_{i_1\cdot}w_{\cdot j_2}}}\frac{w_{i_2j_2}}{\sqrt{w_{i_2\cdot}w_{\cdot j_2}}}\cdots\frac{w_{i_{l-1}j_l}}{\sqrt{w_{i_{l-1}\cdot}w_{\cdot j_l}}}\frac{w_{i'j_l}}{\sqrt{w_{i'\cdot}w_{\cdot j_l}}}.$$

It is worthwhile to note that a set of weights $\{w_{ij}\}$ is identifiable when non of $w_{ij}$ is zero. Some of the weights $w_{ij}$, however, can assume zero without violating the identifiability property.

## 3. Interaction effects

In this section we investigate some properties of the components $\mu$, $\alpha_i$, $\beta_j$ and $\gamma_{ij}$ defined by our general set of identifiable weights $\{w_{ij}\}$, i.e., a solution of (2.3).

First, we show one of the extremal properties of the interaction effects. Multiplying (2.1) on the left by $\begin{bmatrix} I_r \otimes \boldsymbol{j}'_s \\ \boldsymbol{j}'_r \otimes I_s \end{bmatrix} D_w$ and using (2.2), we have

$$(3.1)\qquad \begin{bmatrix} I_r \otimes \boldsymbol{j}'_s \\ \boldsymbol{j}'_r \otimes I_s \end{bmatrix} D_w [\boldsymbol{j}_{rs},\ I_r \otimes \boldsymbol{j}_s,\ \boldsymbol{j}_r \otimes I_s] \begin{bmatrix} \mu \\ \boldsymbol{\alpha} \\ \boldsymbol{\beta} \end{bmatrix} = \begin{bmatrix} I_r \otimes \boldsymbol{j}'_s \\ \boldsymbol{j}'_r \otimes I_s \end{bmatrix} D_w \boldsymbol{\eta}.$$

This shows that the components $\mu$, $\alpha_i$, and $\beta_j$ are so determined as to minimize the quadratic form

$$\|D_w^{\frac{1}{2}}(\boldsymbol{\eta} - \boldsymbol{j}_{rs}\mu - [I_r \otimes \boldsymbol{j}_s]\boldsymbol{\alpha} - [\boldsymbol{j}_r \otimes I_s]\boldsymbol{\beta})\|^2$$
$$= \sum_{i=1}^{r} \sum_{j=1}^{s} w_{ij}(\eta_{ij} - \mu - \alpha_i - \beta_j)^2 = \sum_{i=1}^{r} \sum_{j=1}^{s}{}' w_{ij} \gamma_{ij}^2$$

where $\|\boldsymbol{x}\|$ denotes the modulus of the vector $\boldsymbol{x}$. Our set of constraints is so contemplated as to minimize the weighted dispersion of the interaction effects.

Next we show that a theorem due to Scheffé ([6] p. 93) still holds for our components defined by a set of identifiable weights.

THEOREM 3. *If the interaction effects $\{\gamma_{ij}^{(0)}\}$ are all zero for some set of identifiable weights $\{w_{ij}^{(0)}\}$, they are all zero for every set of identifiable weights. In that case every contrast in main effects $\{\alpha_i\}$ or $\{\beta_j\}$ has a value that does not depend on any set of identifiable weights $\{w_{ij}\}$.*

PROOF. Suppose that the components defined for a particular set of identifiable weights $\{w_{ij}^{(0)}\}$ are $\mu^{(0)}$, $\alpha_i^{(0)}$, $\beta_j^{(0)}$, $\gamma_{ij}^{(0)}$, and suppose that all $\gamma_{ij}^{(0)} = 0$. Then

$$(3.2)\qquad \boldsymbol{\eta} = \boldsymbol{j}_{rs}\mu^{(0)} + [I_r \otimes \boldsymbol{j}_s]\boldsymbol{\alpha}^{(0)} + [\boldsymbol{j}_r \otimes I_s]\boldsymbol{\beta}^{(0)}$$

where $\boldsymbol{\alpha}^{(0)\prime} = (\alpha_1^{(0)}, \cdots, \alpha_r^{(0)})$ and $\boldsymbol{\beta}^{(0)\prime} = (\beta_1^{(0)}, \cdots, \beta_s^{(0)})$. After substituting (3.2) into (3.1) we can write

$$(3.3)\qquad \begin{bmatrix} \boldsymbol{w}_r & D_r & W \\ \boldsymbol{w}_s & W' & D_s \end{bmatrix} \begin{bmatrix} \mu - \mu^{(0)} \\ \boldsymbol{\alpha} - \boldsymbol{\alpha}^{(0)} \\ \boldsymbol{\beta} - \boldsymbol{\beta}^{(0)} \end{bmatrix} = 0$$

where $\boldsymbol{w}_r'=(w_{1\cdot}, \cdots, w_{r\cdot})$ and $\boldsymbol{w}_s'=(w_{\cdot 1}, \cdots, w_{\cdot s})$. By using Theorem 2 we have $\text{rank}\begin{bmatrix} \boldsymbol{w}_r & D_r & W \\ \boldsymbol{w}_s & W' & D_s \end{bmatrix} = \text{rank}\begin{bmatrix} D_r & W \\ W' & D_s \end{bmatrix} = r+s-1$. It can be seen that two independent vectors $(-1, \boldsymbol{j}_r', \boldsymbol{O}')$ and $(-1, \boldsymbol{O}', \boldsymbol{j}_s')$ are orthogonal to every row vector of the coefficient matrix of (3.3). The general solution of (3.3) is, therefore,

$$\begin{bmatrix} \mu-\mu^{(0)} \\ \boldsymbol{\alpha}-\boldsymbol{\alpha}^{(0)} \\ \boldsymbol{\beta}-\boldsymbol{\beta}^{(0)} \end{bmatrix} = b_1\begin{bmatrix} -1 \\ \boldsymbol{j}_r \\ 0 \end{bmatrix} + b_2\begin{bmatrix} -1 \\ 0 \\ \boldsymbol{j}_s \end{bmatrix} \tag{3.4}$$

or $\mu-\mu^{(0)}=-b_1-b_2$, $\alpha_i-\alpha_i^{(0)}=b_1$ for all $i$, and $\beta_j-\beta_j^{(0)}=b_2$ for all $j$, where $b_1$ and $b_2$ are arbitrary constants. From (3.2) and (3.4) we get $\mu+\alpha_i+\beta_j=\mu^{(0)}+\alpha_i^{(0)}+\beta_j^{(0)}=\eta_{ij}$. This implies that $\gamma_{ij}=0$ for all $i$ and $j$. We, moreover, can easily get $\sum_{i=1}^{r} c_i\alpha_i=\sum_{i=1}^{r} c_i\alpha_i^{(0)}$ and $\sum_{j=1}^{s} d_j\beta_j=\sum_{j=1}^{s} d_j\beta_j^{(0)}$ for any contrast $\sum_{i=1}^{r} c_i\alpha_i$ and $\sum_{j=1}^{s} d_j\beta_j$, respectively.

THEOREM 4. *If the main effects $\{\alpha_i^{(0)}\}$ (or $\{\beta_j^{(0)}\}$) and the interaction effects $\{\gamma_{ij}^{(0)}\}$ are all zero for some set of identifiable weights $\{w_{ij}^{(0)}\}$, they are all zero for every set of identifiable weights.*

PROOF. Suppose that $\alpha_i^{(0)}=0$ for all $\alpha_i^{(0)}=0$ and all $\gamma_{ij}^{(0)}=0$. From Theorem 3 we have that all $\gamma_{ij}=0$ and $\sum_{i=1}^{r} c_i\alpha_i=0$ for any $\{c_i\}$ such that $\sum_{i=1}^{r} c_i=0$. The main effects $\{\alpha_i\}$ satisfy the constraint $\sum_{i=1}^{r} w_{i\cdot}\alpha_i=0$. Thus we have $\alpha_i=0$ for all $i$.

## 4. Analysis of variances

In this section we discuss the analysis of variances of a two-way classification design with unequal cell frequencies. The components of the yield function $\eta_{ij}$ are assumed to be parametric and they are assumed to be determined by a set of identifiable weights $\{w_{ij}\}$.

Let $y_{ijk}$ be the $k$th yield in the $(i, j)$ cell, i.e., the value of the $k$th observation obtained when the factor $A$ is applied at level $i$ and the factor $B$ at level $j$. Let $\boldsymbol{y}$ be a column vector in which those yields are arranged in the lexicographical order with respect to the factors $A$, $B$ and replication.

Let $n_{ij}$ be the number of replications in the $(i, j)$ cell. We only assume that there is no level that has no observation at all in the marginals, i.e., $n_{i\cdot}=\sum_{j=1}^{s} n_{ij}>0$ and $n_{\cdot j}=\sum_{i=1}^{r} n_{ij}>0$ for all $i$ and $j$. Nevertheless, the number $n_{ij}$ may assume zero for some pair $(i, j)$.

Our basic model under which all the hypotheses are to be tested is as follows in matrix notation,

$$(4.1)\qquad G_W: \begin{cases} \boldsymbol{y} \text{ is distributed as } N[\boldsymbol{\theta}, \sigma^2 I_n], \text{ i.e., the multivariate normal} \\ \text{distribution with mean } \boldsymbol{\theta} \text{ and dispersion matrix } \sigma^2 I_n \\ \boldsymbol{\theta} = \boldsymbol{j}_n\mu + \Phi\boldsymbol{\alpha} + \Psi\boldsymbol{\beta} + \Gamma\boldsymbol{\gamma} \\ \boldsymbol{j}'_{rs} D_w [I_r \otimes \boldsymbol{j}'_s]\boldsymbol{\alpha} = 0, \quad \begin{bmatrix} I_r \otimes \boldsymbol{j}'_s \\ \boldsymbol{j}'_r \otimes I_s \end{bmatrix} D_w \boldsymbol{\gamma} = 0 \\ \boldsymbol{j}'_{rs} D_w [\boldsymbol{j}'_r \otimes I_s]\boldsymbol{\beta} = 0, \end{cases}$$

where $n = \sum_{i=1}^{r} \sum_{j=1}^{s} n_{ij}$,

$\Gamma(n \times rs) = ||\gamma_{gl}||, \quad l = (i-1)s + j,$

$$\gamma_{gl} = \begin{cases} 1, & \text{if the combination of } i\text{th level of the factor } A \text{ with} \\ & j\text{th level of the factor } B \text{ is applied to } g\text{th yield.} \\ 0, & \text{otherwise.} \end{cases}$$

$\Phi(n \times r) = \Gamma[I_r \otimes \boldsymbol{j}_s]$ and $\Psi(n \times s) = \Gamma[\boldsymbol{j}_r \otimes I_s]$.

We now consider the tests of the following hypotheses $H_1$, $H_2$, $H_3$ and $H_4$ under $G_W$ assuming that some of the $n_{ij} \geq 2$, where

$H_1: \boldsymbol{\gamma} = 0$ (interaction effects zero)

$H_2: \boldsymbol{\alpha} = 0$ (main effects of $A$ zero)

$H_3: \boldsymbol{\beta} = 0$ (main effects of $B$ zero)

$H_4: \mu = 0$ (general mean zero).

The underlying model $G_W$ can be specified by $\boldsymbol{\theta} \in \Omega$, a linear subspace in $E_n$. Each hypothesis $H_t$ is a linear hypothesis and is specified by $\boldsymbol{\theta} \in \omega_t$, a subspace of $\Omega$. Let $P_\Omega$ and $P_{\omega_t}$ be projection operators to $\Omega$ and $\omega_t$, respectively, then the likelihood ratio criterion for testing $H_t$ is equivalent to

$$(4.2)\qquad F_t = \frac{f_e}{f_t} \frac{\boldsymbol{y}'(P_\Omega - P_{w_t})\boldsymbol{y}}{\boldsymbol{y}'(I_n - P_\Omega)\boldsymbol{y}}$$

where $f_e = n$-dimension $(\Omega) = n$-(number of non-empty cells) and $f_t$=dimension $(\Omega)$-dimension $(w_t)$. $F_t$ has a central $F$ distribution under the hypothesis $H_t$ and a non-central $F$ distribution under the alternative with $f_t$, $f_e$ degrees of freedom.

In the following we give the explicit expressions of those projection operators and discuss some of the properties of the likelihood ratio criterions.

Under $G_W$, since $\mathcal{E}(\boldsymbol{y})=\boldsymbol{j}_n\mu+\boldsymbol{\Phi}\boldsymbol{\alpha}+\boldsymbol{\Psi}\boldsymbol{\beta}+\Gamma\boldsymbol{\gamma}=\Gamma\boldsymbol{\eta}$, $\boldsymbol{\eta}\in\mathrm{E}_{rs}$, we have $\Omega=\mathcal{R}[\Gamma]$, and consequently, we have $\mathrm{P}_\Omega=\mathcal{P}[\Gamma]$, where $\mathcal{P}[X]$ denotes the projection operator to $\mathcal{R}[X]$.

The space $\omega_1$ can be expressed as

$$\omega_1=\left\{\boldsymbol{\theta}\,|\,\boldsymbol{\theta}=[\boldsymbol{j}_n,\boldsymbol{\Phi},\boldsymbol{\Psi}]\begin{bmatrix}\mu\\ \boldsymbol{\alpha}\\ \boldsymbol{\beta}\end{bmatrix},\ \begin{bmatrix}0 & \boldsymbol{j}'_{rs}D_w[I_r\otimes\boldsymbol{j}_s] & 0\\ 0 & 0 & \boldsymbol{j}'_{rs}D_w[\boldsymbol{j}_r\otimes I_s]\end{bmatrix}\begin{bmatrix}\mu\\ \boldsymbol{\alpha}\\ \boldsymbol{\beta}\end{bmatrix}=0\right\}$$

$$=\mathcal{R}[\boldsymbol{\Phi},\boldsymbol{\Psi}].$$

The last equality follows from the arguments similar to those in the proof of Theorem 1. The likelihood ratio criterion for testing $H_1$ under $G_w$ is, therefore, equivalent to

$$F_1=\frac{f_e}{f_1}\frac{\boldsymbol{y}'(\mathcal{P}[\Gamma]-\mathcal{P}[\boldsymbol{\Phi},\boldsymbol{\Psi}])\boldsymbol{y}}{\boldsymbol{y}'(I_n-\mathcal{P}[\Gamma])\boldsymbol{y}}. \tag{4.3}$$

The test statistic (4.3) shows that its value does not affected by a set of identifiable weights $\{w_{ij}\}$. Since our set of constraints is consistent, it is interesting to note that the condition of the identifiability of the set of weights is not necessary in deriving the test statistic (4.3).

Although each of the spaces $\omega_2$, $\omega_3$ and $\omega_4$ is defined in the form $\{\boldsymbol{\theta}\,|\,\boldsymbol{\theta}=A\boldsymbol{\tau}.\ B\boldsymbol{\tau}=0\}$, neither $\mathcal{R}[A']\supset\mathcal{R}[B']$ nor $\mathcal{R}[A']\cap\mathcal{R}[B']=\{0\}$ is in general true (see Yamamoto and Fujikoshi [7]). The explicit expressions of $P_{\omega_2}$, $P_{\omega_3}$ and $P_{\omega_4}$ are very complicated. For example we give an explicit expression of $P_{\omega_2}$ in the following.

$$\omega_2=\left\{\boldsymbol{\theta}\,|\,\boldsymbol{\theta}=[\boldsymbol{j}_n,\boldsymbol{\Psi},\Gamma]\begin{bmatrix}\mu\\ \boldsymbol{\beta}\\ \boldsymbol{\gamma}\end{bmatrix},\ \begin{pmatrix}0 & \boldsymbol{j}'_{rs}D_w[\boldsymbol{j}_r\otimes I_s] & 0\\ 0 & 0 & \begin{bmatrix}I_r\otimes\boldsymbol{j}'_s\\ \boldsymbol{j}'_r\otimes I_s\end{bmatrix}D_w\end{pmatrix}\begin{bmatrix}\mu\\ \boldsymbol{\beta}\\ \boldsymbol{\gamma}\end{bmatrix}=0\right\}$$

$$=\mathcal{R}[\boldsymbol{\Psi}]+\left\{\boldsymbol{\theta}\,|\,\boldsymbol{\theta}=\Gamma\boldsymbol{\gamma},\ \begin{bmatrix}I_r\otimes\boldsymbol{j}'_s\\ \boldsymbol{j}'_r\otimes I_s\end{bmatrix}D_w\gamma=0\right\}$$

where $V_1+V_2$ denotes the sum of vector spaces $V_1$ and $V_2$. Using the theorem in Yamamoto and Fujikoshi [7], we get $\omega_2=\mathcal{R}[\boldsymbol{\Psi}]+\mathcal{R}[Q]$, where $Q=\Gamma(\Gamma'\Gamma)^g\Gamma'-\Gamma(\Gamma'\Gamma)^gS'\{S(\Gamma'\Gamma)^gS'\}^gS(\Gamma'\Gamma)^g\Gamma'$, $S'=B'_1(I_{r+s}-\mathcal{P}[B_2])$, $B'_1=\mathcal{P}[\Gamma']D_w[I_r\otimes\boldsymbol{j}_s,\boldsymbol{j}_r\otimes I_s]$, $B'_2=(I_{rs}-\mathcal{P}[\Gamma'])D_w[I_r\otimes\boldsymbol{j}_s,\boldsymbol{j}_r\otimes I_s]$ and the symbol $X^g$ denotes a generalized inverse of a matrix $X$. Therefore we can express $P_{\omega_2}$ as

$$\mathrm{P}_{\omega_2}=\mathcal{P}[\boldsymbol{\Psi}]+\mathcal{P}[(I_n-\mathcal{P}[\boldsymbol{\Psi}])Q]. \tag{4.4}$$

Similarly we can obtain the explicit expressions of $P_{\omega_3}$ and $P_{\omega_4}$.

In general, the likelihood ratio criteria for testing $H_2$, $H_3$ and $H_4$ depend on the set of weights $\{w_{ij}\}$. This means that the sums of squares appropriate for testing $H_t$ $(t=2, 3, 4)$ depend on how the components $\mu$, $\alpha_i$, $\beta_j$ and $\gamma_{ij}$ are defined.

Let us consider an important special case in which the set of cell frequencies $\{n_{ij}\}$ is chosen as a set of weights $\{w_{ij}\}$. Choice of such a set of weights is plausible not only in a usual two-way classification design with equal frequencies but also in a fairly general class of designs with unequal cell frequencies in which the unequality comes from the stochastic structure of data collection. The former applies to the case where the yield function is considered to be a deterministic function of the levels. The latter applies to the case where $n_{ij}$ may be considered as aposteriori estimates of apriori weights or probabilities $w_{ij}$.

We further assume that $\{n_{ij}\}$ satisfies the identifiability conditions. It is easy to see that the identifiability of the set of weights $\{n_{ij}\}$ is equivalent to the connectedness of the two-way classification design.

The constraints (2.2) can be expressed as

$$(4.5)\qquad \begin{aligned} &\boldsymbol{j}_n'\boldsymbol{\Phi\alpha}=0 &&\text{or}\quad \sum_{i=1}^{r} n_{i\cdot}\alpha_i=0 \\ &\boldsymbol{j}_n'\boldsymbol{\Psi\beta}=0 &&\text{or}\quad \sum_{j=1}^{s} n_{\cdot j}\beta_j=0 \\ &\begin{bmatrix}\boldsymbol{\Phi}'\\ \boldsymbol{\Psi}'\end{bmatrix}\boldsymbol{\Gamma\gamma}=0 &&\text{or}\quad \begin{cases}\sum_{j=1}^{s} n_{ij}\gamma_{ij}=0 & \text{for all } i\\ \sum_{i=1}^{r} n_{ij}\gamma_{ij}=0 & \text{for all } j,\end{cases} \end{aligned}$$

and our basic model can be expressed as

$$(4.6)\qquad G_N:\begin{cases} \boldsymbol{y} \text{ is distributed as } N[\boldsymbol{\theta}, \sigma^2 I_n] \\ \boldsymbol{\theta}=\boldsymbol{j}_n\mu+\boldsymbol{\Phi\alpha}+\boldsymbol{\Psi\beta}+\boldsymbol{\Gamma\gamma} \\ \boldsymbol{j}_n'\boldsymbol{\Phi}\alpha=0, \qquad \begin{bmatrix}\boldsymbol{\Phi}'\\ \boldsymbol{\Psi}'\end{bmatrix}\boldsymbol{\Gamma\gamma}=0. \\ \boldsymbol{j}_n'\boldsymbol{\Psi}\beta=0, \end{cases}$$

We already know that the statistic for testing $H_1$ is not affected by the set of weights and is given by (4.3). Noting that $B_1'=\mathcal{P}[\Gamma']\Gamma'[\boldsymbol{\Phi}, \boldsymbol{\Psi}]=\Gamma'[\boldsymbol{\Phi}, \boldsymbol{\Psi}]$, $B_2'=0$ and $\mathcal{P}[\Gamma][\boldsymbol{\Phi}, \boldsymbol{\Psi}]=[\boldsymbol{\Phi}, \boldsymbol{\Psi}]$, we can express $P_{\omega_2}$ in (4.4) as

$$(4.7)\qquad \begin{aligned} \mathrm{P}_{\omega_2} &= \mathcal{P}[\boldsymbol{\Psi}]+\mathcal{P}[(I_n-\mathcal{P}[\boldsymbol{\Psi}])(\mathcal{P}[\Gamma]-\mathcal{P}[\boldsymbol{\Phi}, \boldsymbol{\Psi}])] \\ &= \mathcal{P}[\Gamma]-\mathcal{P}[\boldsymbol{\Phi}, \boldsymbol{\Psi}]+\mathcal{P}[\boldsymbol{\Psi}]. \end{aligned}$$

Hence the likelihood ratio criterion for testing $H_2$ is equivalent to

(4.8) $$F_2=\frac{f_e}{r-1}\frac{\boldsymbol{y}'(\mathcal{P}[\Phi,\Psi]-\mathcal{P}[\Psi])\boldsymbol{y}}{\boldsymbol{y}'(I_n-\mathcal{P}[\Gamma])\boldsymbol{y}}.$$

Similarly the likelihood ratio criterion for testing $H_3$ is equivalent to

(4.9) $$F_3=\frac{f_e}{s-1}\frac{\boldsymbol{y}'(\mathcal{P}[\Phi,\Psi]-\mathcal{P}[\Phi])\boldsymbol{y}}{\boldsymbol{y}'(I_n-\mathcal{P}[\Gamma])\boldsymbol{y}}.$$

Since $\omega_4$ and $\mathcal{R}[\boldsymbol{j}_n]$ are mutually orthogonal and $\omega_4+\mathcal{R}[\boldsymbol{j}_n]=\mathcal{R}[\Gamma]$, we get $\mathrm{P}_{\omega_4}=\mathcal{P}[\Gamma]-\mathcal{P}[\boldsymbol{j}_n]$. Hence the likelihood ratio criterion for testing $H_4$ is equivalent to

(4.10) $$F_4=f_e\frac{\boldsymbol{y}'\mathcal{P}[\boldsymbol{j}_n]\boldsymbol{y}}{\boldsymbol{y}'(I_n-\mathcal{P}[\Gamma])\boldsymbol{y}}.$$

From (4.3), (4.8), (4.9) and (4.10) we have two types of ANOVA (the analysis of variances);

(4.11) $$\begin{aligned}\boldsymbol{y}'\boldsymbol{y}&=\boldsymbol{y}'(I_n-\mathcal{P}[\Gamma])\boldsymbol{y}+\boldsymbol{y}'(\mathcal{P}[\Gamma]-\mathcal{P}[\Phi,\Psi])\boldsymbol{y}+\boldsymbol{y}'(\mathcal{P}[\Phi,\Psi]\\&\quad-\mathcal{P}[\Psi])\boldsymbol{y}+\boldsymbol{y}'(\mathcal{P}[\Psi]-\mathcal{P}[\boldsymbol{j}_n])\boldsymbol{y}+\boldsymbol{y}'\mathcal{P}[\boldsymbol{j}_n]\boldsymbol{y}\\&=\boldsymbol{y}'(I_n-\mathcal{P}[\Gamma])\boldsymbol{y}+\boldsymbol{y}'(\mathcal{P}[\Gamma]-\mathcal{P}[\Phi,\Psi])\boldsymbol{y}+\boldsymbol{y}'(\mathcal{P}[\Phi,\Psi]\\&\quad-\mathcal{P}[\Phi])\boldsymbol{y}+\boldsymbol{y}'(\mathcal{P}[\Phi]-\mathcal{P}[\boldsymbol{j}_n])\boldsymbol{y}+\boldsymbol{y}'\mathcal{P}[\boldsymbol{j}_n]\boldsymbol{y}.\end{aligned}$$

$\boldsymbol{y}'(\mathcal{P}[\Phi,\Psi]-\mathcal{P}[\Psi])\boldsymbol{y}$ is called the sum of squares due to $\boldsymbol{\alpha}$ eliminating $\boldsymbol{\beta}$ and its noncentrality parameter is given by $\frac{1}{2\sigma^2}\mathcal{E}(\boldsymbol{y})'(\mathcal{P}[\Phi,\Psi]-\mathcal{P}[\Psi])\mathcal{E}(\boldsymbol{y})=\frac{1}{2\sigma^2}\boldsymbol{\alpha}'\Phi'(I_n-\mathcal{P}[\Psi])\Phi\boldsymbol{\alpha}$. Similarly $\boldsymbol{y}'(\mathcal{P}[\Phi,\Psi]-\mathcal{P}[\Phi])\boldsymbol{y}$ is called the sum of squares due to $\boldsymbol{\beta}$ eliminating $\boldsymbol{\alpha}$ and its noncentrality parameter is given by $\frac{1}{2\sigma^2}\boldsymbol{\beta}'\Psi'(I_n-\mathcal{P}[\Phi])\Psi\boldsymbol{\beta}$. If all the $n_{ij}$ are positive, we can easily recognize the agreement between the ANOVA (4.11) and the ANOVA being described by Rao [4] under the heading "Two-way classifications with unequal numbers in cells" without introducing the definition of the parameters.

In connection with the ANOVA given in (4.11) we have the following theorem.

THEOREM 5. *A necessary and sufficient condition that the noncentrality parameters of the component sum of squares* $\boldsymbol{y}'\mathcal{P}[\boldsymbol{j}_n]\boldsymbol{y}$, $\boldsymbol{y}'(\mathcal{P}[\Phi,\Psi]-\mathcal{P}[\Psi])\boldsymbol{y}$ *and* $\boldsymbol{y}'(\mathcal{P}[\Phi,\Psi]-\mathcal{P}[\Phi])\boldsymbol{y}$ *depend only on general mean* $\mu$, *main effects* $\{\alpha_i\}$ *and main effects* $\{\beta_j\}$, *respectively, are that these parameters are so identified by the set of weights* $\{n_{ij}\}$.

PROOF. Sufficiency of the theorem has been shown above. Necessity can be proved as follows. From the conditions we have $\boldsymbol{j}_n'\Phi\boldsymbol{\alpha}=0$, $\boldsymbol{j}_n'\Psi\boldsymbol{\beta}=0$,

$j_n'\Gamma\gamma=0$, $(\mathcal{P}[\Phi, \Psi]-\mathcal{P}[\Psi])\Gamma\gamma=0$ and $(\mathcal{P}[\Phi, \Psi]-\mathcal{P}[\Phi])\Gamma\gamma=0$. Since $\mathcal{R}[\Phi, \Psi]=\mathcal{R}[j_n]+\mathcal{R}[\mathcal{P}[\Phi, \Psi]-\mathcal{P}[\Psi]]+\mathcal{R}[\mathcal{P}[\Phi, \Psi]-\mathcal{P}[\Phi]]$, we have $\Phi'\Gamma\gamma=0$ and $\Psi'\Gamma\gamma=0$. This completes the proof.

This theorem shows that the appropriateness of the ANOVA described by Rao [4] can not be explained by any set of weights given by Scheffé [6].

## 5. Orthogonality

In this section we derive the necessary and sufficient conditions for the hypotheses $H_1$, $H_2$, $H_3$ and $H_4$ to be orthogonal with respect to the basic model. The orthogonality of the hypotheses implies that the likelihood ratio criteria for testing those hypotheses are quasi-independent (Roy and Gnanadesikan [5]), i.e., the projection operators $P_\Omega-P_{\omega_1}$, $P_\Omega-P_{\omega_2}$, $P_\Omega-P_{\omega_3}$ and $P_\Omega-P_{\omega_4}$ are mutually orthogonal, or equivalently, the linear spaces $\Omega\cap\omega_1^\perp$, $\Omega\cap\omega_2^\perp$, $\Omega\cap\omega_3^\perp$ and $\Omega\cap\omega_4^\perp$ are mutually orthogonal (Darroch and Silvey [3]), where $\omega^\perp$ denotes the orthogonal complement of the linear space $\omega$.

At first, we consider the case where the basic model is $G_N$. From (4.3), (4.8), (4.9) and (4.10) we have

$$
\begin{aligned}
\Omega\cap\omega_1^\perp &= \mathcal{R}[\Gamma]\cap(\mathcal{R}[\Phi, \Psi])^\perp \\
\Omega\cap\omega_2^\perp &= \mathcal{R}[\Phi, \Psi]\cap(\mathcal{R}[\Psi])^\perp = \mathcal{R}[(I_n-\mathcal{P}[\Psi])\Phi] \\
\Omega\cap\omega_3^\perp &= \mathcal{R}[\Phi, \Psi]\cap(\mathcal{R}[\Phi])^\perp = \mathcal{R}[(I_n-\mathcal{P}[\Phi])\Psi] \\
\Omega\cap\omega_4^\perp &= \mathcal{R}[j_n].
\end{aligned}
\tag{5.1}
$$

THEOREM 6. *The hypotheses $H_1$, $H_2$, $H_3$ and $H_4$ are orthogonal with respect to $G_N$ if and only if $n_{ij}=n_{i\cdot}n_{\cdot j}/n$.*

PROOF. Since $\mathcal{R}[\Gamma]\cap(\mathcal{R}[\Phi, \Psi])^\perp$, $\mathcal{R}[\Phi, \Psi]\cap(\mathcal{R}[\Psi])^\perp$, $\mathcal{R}[j_n]$ (or $\mathcal{R}[\Gamma]\cap(\mathcal{R}[\Phi, \Psi])^\perp$, $\mathcal{R}[\Phi, \Psi]\cap(\mathcal{R}[\Phi])^\perp$, $\mathcal{R}[j_n]$) are mutually orthogonal, it is sufficient to show that $\mathcal{R}[(I_n-\mathcal{P}[\Psi])\Phi]$ and $\mathcal{R}[(I_n-\mathcal{P}[\Phi])\Psi]$ are mutually orthogonal if and only if $n_{ij}=n_{i\cdot}n_{\cdot j}/n$. The condition $\Phi'(I_n-\mathcal{P}[\Psi])'(I_n-\mathcal{P}[\Phi])\Psi=0$ is equivalent to that $\{n_{ij}\}$ satisfies the condition, $(I_r-N\Delta_s^{-1}N'\Delta_r^{-1})N=0$, where $N=\Phi'\Psi=\|n_{ij}\|$, $\Delta_r=\Phi'\Phi=\mathrm{diag}(n_{1\cdot}, \cdots, n_{r\cdot})$ and $\Delta_s=\Psi'\Psi=\mathrm{diag}(n_{\cdot 1}, \cdots, n_{\cdot s})$. From Theorem 2 we have rank $(I_r-N\Delta_s^{-1}N'\Delta_r^{-1})=\mathrm{rank}(\Delta_r-N\Delta_s^{-1}N')=r-1$. Hence we have that $(I_r-N\Delta_s^{-1}N'\Delta_r^{-1})N=0$ holds if and only if rank $N=1$, i.e., $n_{ij}=n_{i\cdot}n_{\cdot j}/n$.

Next we consider the case where the basic model is $G_W$.

THEOREM 7. *Assume that all the $n_{ij}$ are positive. Then the hypotheses $H_1$, $H_2$, $H_3$ and $H_4$ are orthogonal with respect to $G_W$ if and only if $n_{ij}=n_{i\cdot}n_{\cdot j}/n$ and $w_{ij}=pn_{ij}$ for all $i$ and $j$, where $p$ is a positive constant.*

PROOF. Sufficiency of the theorem was proved in Theorem 6. Necessity can be derived as follows. Consider the hypothesis $H$: $\mu=0$, $\boldsymbol{\alpha}=0$, $\boldsymbol{\beta}=0$. We can express the hypothesis $H$ in the form $\mathcal{E}(\boldsymbol{y}) \in \omega$, a subspace of $\Omega$. Since

$$\omega = \{\boldsymbol{\theta} \mid \boldsymbol{\theta} = \Gamma\boldsymbol{\gamma}, \quad \begin{bmatrix} I_r \otimes \boldsymbol{j}_s' \\ \boldsymbol{j}_r' \otimes I_s \end{bmatrix} D_w \boldsymbol{\gamma} = 0\},$$

we have

$$\Omega \cap \omega^\perp = \{\boldsymbol{\theta} \mid \boldsymbol{\theta} = \Gamma\boldsymbol{\zeta}, \ \boldsymbol{\zeta}' D_n \boldsymbol{\gamma} = 0 \quad \text{for any } \boldsymbol{\gamma} \text{ such that } \begin{bmatrix} I_r \otimes \boldsymbol{j}_s' \\ \boldsymbol{j}_r' \otimes I_s \end{bmatrix} D_W \boldsymbol{\gamma} = 0\}$$

$$= \mathcal{R}[\Gamma D_n^{-1} D_W [I_r \otimes \boldsymbol{j}_s, \boldsymbol{j}_r \otimes I_s]]$$

where $D_n = \mathrm{diag}(n_{11}, n_{12}, \cdots, n_{rs-1}, n_{rs})$. From (4.3) we have

$$\Omega \cap \omega_1^\perp = \mathcal{R}[\Gamma] \cap \mathcal{R}[\boldsymbol{\Phi}, \boldsymbol{\Psi}]^\perp$$

$$= \{\boldsymbol{\theta} \mid \boldsymbol{\theta} = \Gamma\boldsymbol{\xi}, \quad \begin{bmatrix} I_r \otimes \boldsymbol{j}_s' \\ \boldsymbol{j}_r' \otimes I_s \end{bmatrix} D_n \boldsymbol{\xi} = 0\}.$$

Since $\Omega \cap \omega^\perp$ and $\Omega \cap \omega_1^\perp$ are mutually orthogonal, we have

$$\begin{bmatrix} I_r \otimes \boldsymbol{j}_s' \\ \boldsymbol{j}_r' \otimes I_s \end{bmatrix} D_W \boldsymbol{\xi} = 0 \quad \text{for any } \boldsymbol{\xi} \text{ such that } \begin{bmatrix} I_r \otimes \boldsymbol{j}_s' \\ \boldsymbol{j}_r' \otimes I_s \end{bmatrix} D_n \boldsymbol{\xi} = 0.$$

Therefore there exist matrices $U(\overline{r+s} \times r)$ and $V(\overline{r+s} \times s)$ such that

$$D_W [I_r \otimes \boldsymbol{j}_s, \boldsymbol{j}_r \otimes I_s] = D_n [I_r \otimes \boldsymbol{j}_s, \boldsymbol{j}_r \otimes I_s][U, V].$$

Comparing the elements of both matrices and noting that all the $n_{ij}$ are positive, we can derive that $w_{ij}/n_{ij}=p$, a positive constant, for all $i$ and $j$. Applying Theorem 6 we get $n_{ij}=n_{i.}n_{.j}/n$. This completes the proof.

## References


[1] Bartlett, M. S. (1956). *An introduction to stochastic processes.* The Cambridge University Press, London.
[2] Bellman, R. (1960). *Introduction to matrix analysis.* Mcgraw-Hill, New York.
[3] Darroch, J. N. and Silvey, S. D. (1963). On testing one or more hypothesis. *Ann. Math. Statist.* **34** 555-567.
[4] Rao, C. R. (1952). *Advanced statistical methods in biometric research.* Wiley, New York.
[5] Roy, S. N. and Gnanadesikan, R. (1959). Some contributions to ANOVA in one or more dimensions: I. *Ann. Math. Statist.* **30** 304-317.
[6] Scheffé, H. (1959). *The analysis of variance.* Wiley, New York.
[7] Yamamoto, S. and Fujikoshi, Y. (1967). The linear hypotheses and constraints. *J. Sci. Hiroshima Univ.* Ser. A-1. **31** 211-219.



*Department of Mathematics*
*Faculty of Science*
*Hiroshima University*