On the Equivalence of Proportional Cell Frequencies and Orthogonality of Interaction Spaces in n-way ANOVA

by

Akimichi Takemura

Technical Report #83-49

Department of Statistics
Purdue University

Revised
March 1984

On the equivalence of proportional cell frequencies and orthogonality of interaction spaces in n-way ANOVA

Akimichi Takemura

Abstract

Consider a general n-way crossed classification in ANOVA with unequal cell frequencies. We prove that orthogonality of interaction spaces (suitably adjusted for lower order interaction spaces) holds if and only if the cell frequencies are proportional. Implications of proportional cell frequencies to the adjustment of all other effects are also studied.

Key words: ANOVA, proportional cell frequency, orthogonality, interaction.

1. Introduction.

In textbook treatments of twoway ANOVA under the additivity assumption, the case of proportional cell frequency is often mentioned as a special convenient case where row and column effects are orthogonal to each other and need no adjustments. See Section 4.4 of Scheffé (1959), Section 6.4 of Kempthorne (1952) for example. The converse of this is also true. See Section 4.2 of John (1971), for example.

For general n-way ANOVA precise statements on the implications of proportional cell frequencies are hard to find. A common conception about the case of proportional cell frequencies is that the same analysis can be carried out as in the usual case of equal cell frequencies. This is not exactly true as pointed out by Smith (1951) and Lewis and John (1976). For more detailed study see Mukerjee (1980).

For 3-way ANOVA a set of necessary and sufficient conditions is given in Theorems 4.5, 4.6 of Takeuchi, Yanai, and Mukherjee (1982). For general n-way ANOVA, Seber (1964) states a theorem closely related to our Theorem 2.1. However his proof is very sketchy. Jacobsen (1968, Lemma 12) states a lemma for n-way case essential for the proof of sufficiency, i.e., proportional cell frequency implies orthogonality. Other closely related works include Tjur (1982), Pukelsheim (1983).

In this article we prove that the proportional cell frequencies hold if and only if suitably defined interaction spaces are mutually orthogonal. For precise statements see Section 2. We also discuss implications of proportional cell frequencies to the adjustment of all other effects. In this respect, the case of proportional cell frequencies is not as convenient as the case of equal cell frequencies. For precise statements see Section 4.

In dealing with general n-way ANOVA rather elaborate notation is needed and the notational conventions of tensor analysis are found to be extremely useful. The n-way ANOVA with single observation per cell is summarized using tensor notation in Takemura (1982). In the sequel we freely use the notational conventions and results of Takemura (1982). In principle we follow the traditional notation of tensor analysis (see Chapter 2 of Sokolnikoff, 1964). For convenience we deviate from it at several occasions and these differences will be mentioned in remarks.

In Section 2 necessary definitions and notational conventions are introduced and main results are stated. Section 3 is devoted to the proof of the results in Section 2. In Section 4 we study implications of proportional cell frequency to the adjustment of all other effects.

2. Notation and statement of results

We begin with a quick remark on generalized least squares. Suppose that a random vector $X \in R^m$ has the covariance matrix V. For an observed x consider $(x-\hat{x})'V^{-1}(x-\hat{x}) = \min_y (x-y)'V^{-1}(x-y)$ where $\hat{x}$, y are restricted to a prescribed subspace L. Then under the assumption of normality $(X-\hat{X})'V^{-1}(X-\hat{X})$ and $\hat{X}V^{-1}\hat{X}$ are independently distributed according to certain noncentral χ^2-distributions. See Section 4a of Rao (1973) for detail. From geometric viewpoint, $\hat{X}$ is the orthogonal projection of X onto L when R^m is equipped with the inner product: $(x,y)_V = x'V^{-1}y$. This generalizes to the case of random tensor in an obvious way.

Consider a general $m_1 \times m_2 \times \ldots \times m_n$ n-way crossed classification. m_i is the number of levels of i-th factor, i=1,2,...,n. We use variables i, j, k to denote factors. Let S = {1,2,...,n}. I, J, K will be used to denote subsets of S. For example $I = \{i_1,\ldots,i_\ell\} \subset S$. Levels of i-th factor will be denoted by greek letters α_i, β_i, γ_i. Then an n-tuple $\alpha_S = (\alpha_1,\ldots,\alpha_n)$ denotes a particular combination of n factor levels. Here α_S stands for the multi-index $(\alpha_1,\ldots,\alpha_n)$. For a subset $I = \{i_1,\ldots i_\ell\}$ of S, $\alpha_I = \{\alpha_{i_1},\ldots,\alpha_{i_\ell}\}$ denotes a partial multi-index where only the levels of factors in I are of interest. Now let $r(\alpha_S) = r(\alpha_1,\ldots,\alpha_n)$ denote the number of observations at the combination of levels $\alpha_S = (\alpha_1,\ldots,\alpha_n)$. Throughout this article we assume that no cell is empty, i.e., $r(\alpha_S) > 0$ for all α_S. Let

$$R = \Sigma\, r(\alpha_1,\ldots,\alpha_n) = \sum_{\alpha_S} r(\alpha_S)$$

be the total number of observations. For convenience we will work with the relative frequency

(2.1) $$f(\alpha_S) = r(\alpha_S)/R.$$

Note that f can be regarded as probability function of n discrete random variables. Marginal frequency in the class $\alpha_I = (\alpha_{i_1},\ldots,\alpha_{i_\ell})$ is denoted by $f_I(\alpha_I)$, namely

(2.2) $$f_I(\alpha_I) = \sum_{\alpha_j,\ j\notin I} f(\alpha_1,\ldots,\alpha_n).$$

Writing $\alpha_S = \alpha_I \cup \alpha_{I^c}$ we express $f_I(\alpha_I)$ conveniently as

$$f_I(\alpha_I) = \sum_{\alpha_{I^c}} f(\alpha_I \cup \alpha_{I^c}).$$

Let $x^{\alpha_1\cdots\alpha_n} = x^{\alpha_S}$ denote the average of observations in the cell $\alpha_S = (\alpha_1,\ldots,\alpha_n)$ and let $x \in \bigotimes_{i=1}^{n} R^{m_i}$ be the random tensor with these elements

x^{α_S}. Since all interaction terms can be expressed in terms of cell averages, there is no loss of generality in considering the cell averages rather than individual observations. Now Var $(x^{\alpha_S}) = \sigma^2/r(\alpha_S)$ where σ^2 is the common error variance. According to the remark at the beginning of this section we equip $\bigotimes_{i=1}^{n} R^{m_i}$ with the inner product

$$(x,y)_r = \Sigma\,[r(\alpha_1,\ldots,\alpha_n)/\sigma^2]\; x^{\alpha_1\ldots\alpha_n}\; y^{\alpha_1\ldots\alpha_n}$$

$$= \Sigma\,[r(\alpha_S)/\sigma^2]\; x^{\alpha_S}\; y^{\alpha_S}$$

$$= (R/\sigma^2)\;\Sigma\; f(\alpha_S)\; x^{\alpha_S}\; y^{\alpha_S}.$$

In the sequel we ignore the constant factor R/σ^2 for convenience and work with the inner product

$$(x,y)_f = \Sigma\; f(\alpha_S)\; x^{\alpha_S}\; y^{\alpha_S}. \tag{2.3}$$

Remark 2.1. f is a metric tensor in the terminology of tensor analysis but to avoid double subscripts we do not use the usual notation for f.

Let $1_{(m_i)} \in R^{m_i}$ denote the m_i dimensional vector whose elements are all 1's. Here m_i is enclosed in parentheses to distinghish it from a covariant index. Let V_i^0 denote a subspace of R^{m_i} spanned by $1_{(m_i)}$. Let $V_i^1 = R^{m_i}$. Now for $I \subset S$ we define the I-effect space M_I as

Definition 2.1.

$$M_I = \bigotimes_{i=1}^{n} V_i^{\varepsilon_i}, \tag{2.4}$$

where

$$\varepsilon_i = \begin{cases} 1 & \text{if } i \in I, \\ 0 & \text{otherwise.} \end{cases}$$

Note that M_I is a tensor product of whole spaces and 1-dimensional subspaces corresponding to the mean. For a characterization of M_I see Lemma 3.1. Now define

$$\tilde{M}_I = \operatorname{span}\{M_J : J \subset I, J \neq I\}. \tag{2.5}$$

For $I \subset S$ we define L_I^{ℓ}, the I-interaction subspace adjusted for lower order interactions (contained in I), as the orthogonal complement of $\tilde{M}_I$ in M_I. Namely

Definition 2.2. L_I^{ℓ} is defined by the following requirements:

i) $\tilde{M}_I$ and L_I^{ℓ} span M_I,

ii) $\tilde{M}_I$ and L_I^{ℓ} are mutually orthogonal

(with respect to the inner product $(\, , \,)_f$).

For convenience we express this as

$$L_I^\ell = M_I/\tilde{M}_I. \tag{2.6}$$

Now we are ready to state the following theorem.

Theorem 2.1. L_I^ℓ and L_J^ℓ are mutually orthogonal (with respect to the inner product $(\ ,\ )_f$) for every $I \neq J$ if and only if

$$f(\alpha_1,\ldots,\alpha_n) = f_1(\alpha_1)\ \ldots\ f_n(\alpha_n), \tag{2.7}$$

i.e. the frequencies are proportional.

With a slight modification of proof we can generalize Theorem 2.1 as

Theorem 2.2. Let $K = (k_1,\ldots,k_\ell)$ be a fixed subset of $S = \{1,\ldots,n\}$. L_I^ℓ and L_J^ℓ are mutually orthogonal for every $I \neq J$, $I, J \subset K$ if and only if

$$f_K(\alpha_{k_1},\ldots,\alpha_{k_\ell}) = f_{k_1}(\alpha_{k_1})\ldots f_{k_\ell}(\alpha_{k_\ell}). \tag{2.8}$$

For interpretation of Theorem 2.1 see the discussion in Section 4.

Remark 2.2. If $I \subset J$ or $J \subset I$ then L_I^ℓ and L_J^ℓ are orthogonal by the construction of L^ℓ's. Therefore the assertions of the above theorem actually concern only those I, J such that $I \not\subset J$, $J \not\subset I$.

3. Proof.

The proof of sufficiency is fairly simple and this is given first.

3.1. Sufficiency.

Assume that the cell frequencies are proportional, namely (2.7) holds. We introduce an inner product $(\, , \,)_{f_i}$ in R^{m_i} using the marginal frequency f_i,

$$(3.1) \qquad (x,y)_{f_i} = \sum_{\alpha=1}^{m_i} f_i(\alpha)\, x^\alpha y^\alpha ,$$

where $x, y \in R^{m_i}$. Let $W_i^0 = V_i^0 = \text{span}\{1_{(m_i)}\}$ and $W_i^1 = R^{m_i}/W_i$ with respect to $(\, , \,)_{f_i}$. Let

$$(3.2) \qquad \tilde{L}_I = \bigotimes_{i=1}^{n} W_i^{\varepsilon_i} ,$$

where

$$\varepsilon_i = \begin{cases} 1 & \text{if } i \in I, \\ 0 & \text{otherwise.} \end{cases}$$

Then exactly as in Theorem 2.1 of Takemura (1982) we can show that $\tilde{L}_I$ and $\tilde{L}_J$ are mutually orthogonal for every $I \neq J$ with respect to $(\, , \,)_f$. Therefore it suffices to show that $\tilde{L}_I = L_I^\ell$ for all I. Note that $R^{m_i} = V_i^1 = W_i^0 + W_i^1$. Substituting this into (2.4) we see that $M_I = \text{span}\,\{\tilde{L}_J : J \subset I\}$. Using this relation in turn in the definition of $\tilde{M}_I$ (2.5) we obtain $\tilde{M}_I = \text{span}\,\{\tilde{L}_J : J \subset I, J \neq I\}$. This implies $L_I^\ell = M_I/\tilde{M}_I = \tilde{L}_I$. This proves the sufficiency part of Theorem 2.1. Q.E.D.

To prove the sufficiency part of Theorem 2.2 the following modifications are needed. We first note

Lemma 3.1. Let $K = (k_1, \ldots, k_\ell) \subset S$. $x \in M_K$ if and only if $x^{\alpha_1 \ldots \alpha_n} = x^{\alpha_S}$ depends only on $\alpha_K = (\alpha_{k_1}, \ldots, \alpha_{k_\ell})$.

Proof. Let M be the set of tensors which have the property stated in the lemma. Clearly M is a subspace. We want to show that $M = M_K$. Consider a decomposable element $y = y_{(1)}\otimes\cdots\otimes y_{(n)}$ of M_k. Then $y_{(j)} = 1_{(m_j)}$ if $j \notin K$ and $y_{(j)}$ for $j \in K$ are arbitrary. Hence $y^{\alpha_1 \cdots \alpha_n} = y_{(1)}^{\alpha_1} y_{(2)}^{\alpha_2} \cdots y_{(n)}^{\alpha_n}$ depends only on $(\alpha_{k_1}, \ldots, \alpha_{k_\ell})$. Since decomposable tensors of this form generate M_K we obtain $M_K \subset M$. However $\dim M_K = \pi_{i=1}^{\ell} m_{k_i} = \dim M$ hence $M_K = M$. Q.E.D.

By Lemma 3.1 if $x \in M_K$ then x^{α_S} depends only on α_K. We express this conveniently as

$$x^{\alpha_S} = x^{\alpha_K \cup \alpha_{K^c}} = x^{\alpha_K}.$$

Let $x, y \in M_K$. Then

$$\begin{aligned}(x,y)_f &= \Sigma\ f(\alpha_S)\ x^{\alpha_S}\ y^{\alpha_S} \\ &= \Sigma\ f(\alpha_K \cup \alpha_{K^c})\ x^{\alpha_K}\ y^{\alpha_K} \\ &= \sum_{\alpha_K} x^{\alpha_K}\ y^{\alpha_K} \sum_{\alpha_{K^c}} f(\alpha_K \cup \alpha_{K^c}) \\ &= \sum_{\alpha_K} f_K(\alpha_K)\ x^{\alpha_K}\ y^{\alpha_K}.\end{aligned}$$

Therefore we see that the inner product $(\ ,\ )_f$ restricted to M_K is given by the marginal frequencies $f_K(\alpha_K)$. Now in the proof of sufficiency of Theorem 2.1

above we consider M_K equipped with inner product given by f_K instead of the whole space $\otimes_{i=1}^{n} R^{m_i}$. Furthermore we only consider I,J which are subsets of K. Then exactly the same argument applies to the sufficiency part of Theorem 2.2. This completes the proof. Q.E.D.

3.2. Necessity.

We prove the necessity in a series of lemmas.

Lemma 3.2. L_I^ℓ and L_J^ℓ are mutually orthogonal for every $I \neq J$ if and only if $M_I/M_{I\cap J}$ and $M_J/M_{I\cap J}$ are mutually orthogonal for every $I \neq J$.

Proof. Assume that L_I^ℓ's are orthogonal. Now by construction of L_I^ℓ's we have $M_I = \text{span}\,\{L_{I'}^\ell: I' \subset I\}$, $M_J = \text{span}\,\{L_{I'}^\ell: I' \subset J\}$ and $M_{I\cap J} = \text{span}\,\{L_{I'}^\ell: I' \subset I\cap J\}$. Hence

$$M_I/M_{I\cap J} = \text{span}\,\{L_{I'}^\ell: I' \subset I\cap J^c\}$$

$$M_J/M_{I\cap J} = \text{span}\,\{L_{I'}^\ell: I' \subset J\cap I^c\}$$

Note that $(I\cap J^c) \cap (J\cap I^c) = \emptyset$. Therefore $M_I/M_{I\cap J}$ and $M_J/M_{I\cap J}$ are orthogonal. Conversely assume the latter condition. By Remark 2.2 we consider only I,J such that $I \not\subset J$, $J \not\subset I$. Then $I\cap J$ is a proper subset of I as well as of J. Hence $M_{I\cap J} \subset \tilde{M}_I$, $M_{I\cap J} \subset \tilde{M}_J$. Then $L_I^\ell = M_I/\tilde{M}_I \subset M_I/M_{I\cap J}$ and similarly $L_J^\ell \subset M_J/M_{I\cap J}$. This implies that L_I^ℓ and L_J^ℓ are orthogonal. Q.E.D.

Lemma 3.3. $M_I \cap M_J = M_{I\cap J}$.

This is obvious from the characterization of M_I in Lemma 3.1.

Now we present two lemmas concerning orthogonal projectors in an inner product space V. Let A, B be subspaces of V. A linear mapping P_A from V to itself is called the orthogonal projector onto A if $P_A x=x$ for all $x \in A$ and $P_A x=0$ for all $x \in V/A$ where V/A denotes the orthogonal complement of A in V.

Lemma 3.4. A, B are mutually orthogonal if and only if $P_A P_B=0$.

See Theorem 1.21 of Takeuchi et. al. (1982).

Lemma 3.5. $A/A\cap B$ and $B/A\cap B$ are mutually orthogonal if and only if $P_A P_B = P_{A\cap B}$.

Proof of this is fairly easy using the previous lemma and is omitted.

Combining Lemma 3.2 - Lemma 3.5 we have

Lemma 3.6. L_I^ℓ and L_J^ℓ are mutually orthogonal for every $I\neq J$ if and only if

$$P_{M_I} P_{M_J} = P_{M_{I\cap J}} \qquad \text{for every } I\neq J.$$

The essential step of our proof now is to give the explicit form of the projector P_{M_I}. To do this we first characterize $\otimes R^{m_i}/M_I$ and introduce a notation on Kronecker delta.

Lemma 3.7. $y \in \otimes_{i=1}^{n} R^{m_i}/M_I$ if and only if

$$\sum_{\alpha_{I^c}} f(\alpha_I \cup \alpha_{I^c})\, y^{\alpha_I \cup \alpha_{I^c}} = 0$$

for every α_I.

Proof. Let $x \in M_I$ then $x^{\alpha_S} = x^{\alpha_I}$ by Lemma 3.1. Now

$$(x,y)_f = \Sigma\, f(\alpha_I \cup \alpha_{I^c})\, x^{\alpha_I}\, y^{\alpha_I \cup \alpha_{I^c}}$$

$$= \sum_{\alpha} x^{\alpha_I} \sum_{\alpha_{I^c}} f(\alpha_I \cup \alpha_{I^c})\, y^{\alpha_I \cup \alpha_{I^c}} .$$

Hence $y \in \otimes R^{m_i}/M_I$ iff the right hand side is zero for all x^{α_I}. This holds iff the coefficient of x^{α_I} is zero for all α_I. Q.E.D.

Let δ^{α}_{β} be Kronecker delta. For $I = \{i_1,\ldots,i_\ell\}$ we define

(3.3) $$\delta^{\alpha_I}_{\beta_I} = \delta^{\alpha_{i_1}}_{\beta_{i_1}} \cdot \delta^{\alpha_{i_2}}_{\beta_{i_2}} \cdots \delta^{\alpha_{i_\ell}}_{\beta_{i_\ell}} .$$

Remark 3.1. This differs from the usual definition of generalized Kronecker delta (see Section 40 of Sokolnikoff, 1964).

Now we give the explicit form of the projector P_{M_I}. Being a linear map from $\otimes R^{m_i}$ to itself, P_{M_I} has n contravariant indices and n covariant indices:

$$P_{M_I}{}^{\alpha_S}_{\beta_S} = P_{M_I}{}^{\alpha_1 \ldots \alpha_n}_{\beta_1 \ldots \beta_n}$$

P_{M_I} is explicitly specified if the $(\alpha_S,\beta_S) = ((\alpha_1,\ldots,\alpha_n), (\beta_1,\ldots,\beta_n))$ element of P_{M_I} is specified for all (α_S,β_S).

Remark 3.2. If multi-indices α_S, β_S are ordered lexicographically, then P_{M_I} becomes a $(\pi m_i) \times (\pi m_i)$ matrix. For specifying the entries the ordering is irrelevant.

Lemma 3.8.

(3.4) $$P_{M_I}{}^{\alpha_S}_{\beta_S} = \delta^{\alpha_I}_{\beta_I} \frac{f(\alpha_I \cup \beta_{I^c})}{f_I(\alpha_I)}$$

For example if I = {1,2} then

$$P_{M_{\{1,2\}}}{}^{\alpha_1\ldots\alpha_n}_{\beta_1\ldots\beta_n} = \delta^{\alpha_1}_{\beta_1}\,\delta^{\alpha_2}_{\beta_2}\,\frac{f(\alpha_1,\alpha_2,\beta_3,\ldots,\beta_n)}{f_{12}(\alpha_1,\alpha_2)}$$

Proof. We show that if $x \in M_I$ then $P_{M_I}x=x$ and if $x \in \otimes R^{m_i}/M_I$ then $P_{M_I}x=0$. Let $x \in M_I$. Then $x^{\beta_S} = x^{\beta_I}$. Therefore the α_S-element of $P_{M_I}x$ is

$$(P_{M_I}x)^{\alpha_S} = \sum_{\beta_S} P_{M_I}{}^{\alpha_S}_{\beta_S}\, x^{\beta_S}$$

$$= \sum_{\beta_S} \delta^{\alpha_I}_{\beta_I} \frac{f(\alpha_I \cup \beta_{I^c})}{f_I(\alpha_I)}\, x^{\beta_I}$$

$$= \sum_{\beta_I} \delta^{\alpha_I}_{\beta_I}\, x^{\beta_I} \left[\sum_{\beta_{I^c}} \frac{f(\alpha_I \cup \beta_{I^c})}{f_I(\alpha_I)}\right]$$

$$= \sum_{\beta_I} \delta^{\alpha_I}_{\beta_I}\, x^{\beta_I} = x^{\alpha_I}.$$

This shows $P_{M_I}x=x$.

Now let $x \in \otimes R^{m_i}/M_I$. Then α_S-element of P_{M_I} is

$$(P_{M_I}x)^{\alpha_S} = \sum_{\beta_S} \delta_{\beta_I}^{\alpha_I} \frac{f(\alpha_I \cup \beta_{I^c})}{f_I(\alpha_I)} x^{\beta_I \cup \beta_{I^c}}$$

$$= \frac{1}{f_I(\alpha_I)} \sum_{\beta_{I^c}} f(\alpha_I \cup \beta_{I^c}) \left[\sum_{\beta_I} \delta_{\beta_I}^{\alpha_I} x^{\beta_I \cup \beta_{I^c}} \right]$$

$$= \frac{1}{f_I(\alpha_I)} \sum_{\beta_{I^c}} f(\alpha_I \cup \beta_{I^c}) x^{\alpha_I \cup \beta_{I^c}}$$

$$= 0$$

by Lemma 3.7. Hence $P_{M_I} x = 0$. Q.E.D.

Now the last lemma is the following.

Lemma 3.9. Let $I \cap J = \emptyset$. $P_{M_I} P_{M_J} = P_{M_\emptyset}$ implies

$$f_{I \cup J} (\alpha_I \cup \alpha_J) = f_I(\alpha_I) \cdot f_J(\alpha_J)$$

for all α_I, α_J.

Proof. If $P_{M_I} P_{M_J} = P_{M_\emptyset}$ then for all $(\alpha_S, \varepsilon_S)$ we have

$$(3.5) \qquad \sum_{\gamma_S} (P_{M_I} P_{M_J})_{\gamma_S}^{\alpha_S} \delta_{\varepsilon_J}^{\gamma_J} = \sum_{\gamma_S} P_{M_\emptyset}{}_{\gamma_S}^{\alpha_S} \delta_{\varepsilon_J}^{\gamma_J}$$

Now by Lemma 3.8

$$P_{M_\emptyset}{}^{\alpha_S}_{\gamma_S} = f(\gamma_S).$$

Hence the right hand side of (3.5) is

$$(3.6) \qquad \sum_{\gamma_S} f(\gamma_S)\ \delta^{\gamma_J}_{\varepsilon_J} = \sum_{\gamma_J} \delta^{\gamma_J}_{\varepsilon_J} \sum_{\gamma_{J^c}} f(\gamma_J \cup \gamma_{J^c})$$

$$= \sum_{\gamma_J} \delta^{\gamma_J}_{\varepsilon_J}\ f_J(\gamma_J)$$

$$= f_J(\varepsilon_J).$$

Now the left hand side of (3.5) is

$$(3.7) \qquad \sum_{\gamma_S, \beta_S} P_{M_I}{}^{\alpha_S}_{\beta_S}\ P_{M_J}{}^{\beta_S}_{\gamma_S}\ \delta^{\gamma_J}_{\varepsilon_J}\ .$$

Now

$$\sum_{\gamma_S} P_{M_J}{}^{\beta_S}_{\gamma_S}\ \delta^{\gamma_J}_{\varepsilon_J} = \sum_{\gamma_S} \delta^{\beta_J}_{\gamma_J} \frac{f(\beta_J \cup \gamma_{J^c})}{f_J(\beta_J)}\ \delta^{\gamma_J}_{\varepsilon_J}$$

$$= \sum_{\gamma_J} \delta^{\beta_J}_{\gamma_J}\ \delta^{\gamma_J}_{\varepsilon_J} \left[\sum_{\gamma_{J^c}} \frac{f(\beta_J \cup \gamma_{J^c})}{f_J(\beta_J)} \right]$$

$$= \sum_{\gamma_J} \delta^{\beta_J}_{\gamma_J}\ \delta^{\gamma_J}_{\varepsilon_J} = \delta^{\beta_J}_{\varepsilon_J}\ .$$

Hence (3.7) is

$$(3.8) \qquad \sum_{\beta_S} P_{M_I}{}^{\alpha_S}_{\beta_S} \, \delta^{\beta_J}_{\varepsilon_J} = \sum_{\beta_S} \delta^{\alpha_I}_{\beta_I} \frac{f(\alpha_I \cup \beta_{I^c})}{f_I(\alpha_I)} \delta^{\beta_J}_{\varepsilon_J}$$

$$= \frac{1}{f_I(\alpha_I)} \sum_{\beta_{I^c}} f(\alpha_I \cup \beta_{I^c}) \, \delta^{\beta_J}_{\varepsilon_J} \left[\sum_{\beta_I} \delta^{\alpha_I}_{\beta_I} \right]$$

$$= \frac{1}{f_I(\alpha_I)} \sum_{\beta_J, \beta_{(I \cup J)^c}} f(\alpha_I \cup \beta_J \cup \beta_{(I \cup J)^c}) \, \delta^{\beta_J}_{\varepsilon_J}$$

$$= \frac{1}{f_I(\alpha_I)} \sum_{\beta_{(I \cup J)^c}} f(\alpha_I \cup \varepsilon_J \cup \beta_{(I \cup J)^c})$$

$$= \frac{1}{f_I(\alpha_I)} \, f_{I \cup J}(\alpha_I \cup \varepsilon_J).$$

Equating (3.6) and (3.8) we have

$$f_J(\varepsilon_J) = f_{I \cup J}(\alpha_I \cup \varepsilon_J)/f_I(\alpha_I)$$

for all α_I and ε_J. This proves the lemma. Q.E.D.

Now the necessity part of Theorem 2.1 easily follows from Lemma 3.6 and Lemma 3.9. Except for considering M_K instead of $\otimes R^{m_i}$ no modification is needed in the above argument to prove the necessity part of Theorem 2.2. This completes the entire proof.

4. Relation to adjustment for all other interactions.

In Theorem 2.1 the I-interaction term is only adjusted for lower order interactions contained in I. Often adjustments for all other interactions are desired in practical applications of ANOVA. Here we discuss the implication of proportional cell frequency for adjustment of all other effects. It turns out that when the usual parametrization of mean tensor μ is employed, then the case of proportional cell frequency is convenient for "maximal" interaction terms. For other interaction terms the situation is not as simple. Precise statements will be given in Theorem 4.1 and Corollary 4.1.

To discuss the adjustment for other effects we have to first specify what effects are included in a model. This can be done by specifying a subspace $\mathcal{M}$ where the mean tensor $\mu = E(x)$ is assumed to lie. Usually μ is written as sum of overall mean, main effects, two-factor interactions, etc., where the parameters are under the usual linear restrictions. This amounts to decomposing μ into various subspaces. Now the point is that the subspaces in question are not mutually orthogonal with respect to the inner product $(\, , \,)_f$. Instead they are mutually orthogonal with respect to the natural inner product of $\otimes R^{m_i}$ (see (2.5) of Takemura, 1982). Therefore in the case of unequal cell frequencies two inner products are considered at once: the natural inner product for decomposing μ and the inner product $(\, , \,)_f$ for decomposing x. This is the reason why the case of unequal cell frequency is difficult.

Let $I = \{i_1,\ldots,i_k\}$ and consider the I-interaction space L_I defined in (3.1) of Takemura (1982). Let P_I be given by (3.3) of Takemura (1982) which is the orthogonal projector onto L_I with respect to the natural inner product. Let

$$\mu_I = P_I\mu,$$

where $\mu = E(x)$.

Then

$$\mu = (\sum_I P_I)\ \mu$$

$$= \sum_I P_I\ \mu = \sum_I \mu_I\ ,$$

where I ranges over all subsets of S. Let $|I|$ denote the number of elements of I. Writing

$$\mu = \mu_\emptyset + \sum_{|I|=1} \mu_I + \sum_{|I|=2} \mu_I + \ldots$$

We obtain the usual parametric expression of μ. Following the notation in Chapter 4 of Scheffé (1959), the hypothesis H_I of no I-interaction is expressed as

$$H_I\colon\ \mu_I = 0.$$

Under a specific model certain μ_I's are assumed to be zeros. Let

$$\mathcal{I} = \{I\colon\ \mu_I \neq 0\}$$

and let

$$\mathcal{M} = \text{span } \{L_I: I \in \mathcal{J}\}.$$

$\mathcal{M}$ may be simply called as model. Now for $I \in \mathcal{J}$ we define L_I^a, the I-interaction space adjusted for all other effects in the model $\mathcal{M}$, as follows:

Definition 4.1. Let $\tilde{\mathcal{M}}_I = \text{span } \{L_J: J \neq I, J \in \mathcal{J}\}$. Then

$$L_I^a = \mathcal{M} / \tilde{\mathcal{M}}_I$$

(with respect to $(\, , \,)_f$).

To proceed further we assume that $\mathcal{J}$ satisfies the following condition (A):

(A): If $I \in \mathcal{J}$ and $J \subset I$, then $J \in \mathcal{J}$.

This means that if I-interaction is assumed to be present, then lower order interactions contained in I have to be assumed as well. When condition (A) is satisfied, $I \in \mathcal{J}$ is called maximal if $J \supset I$, $J \in \mathcal{J}$ imply $J = I$. Namely, I is maximal if I-interaction is the highest order interaction in the model $\mathcal{M}$ with respect to the partial ordering of inclusion.

Now we can state the following theorem.

Theorem 4.1. Assume that cell frequencies are proportional and $\mathcal{J}$ satisfies the condition (A). Then for maximal I we have $L_I^\ell = L_I^a$.

Let $\tilde{P}_I$ denote the orthogonal projector onto L_I^ℓ (with respect to $(\, , \,)_f$). Let $m_I = \dim L_I^\ell = \pi_{i \in I} (m_i - 1)$. Then as a corollary to Theorem 4.1 we have

Corollary 4.1. Assume that cell frequencies are proportional and $\mathcal{J}$ satisfies condition (A). Let I be maximal. Under the assumption of normality, $(R/\sigma^2)\ (\tilde{P}_I x, \tilde{P}_I x)_f$ has the noncentral x^2-distribution with m_I degrees of freedom and the noncentrality parameter $(R/\sigma^2)\ (\tilde{P}_I \mu_I, \tilde{P}_I \mu_I)_f$.

Proof of Theorem 4.1. Note that M_I = span $\{L_J: J \subset I\}$. Therefore under the condition (A), $M_I \subset \mathcal{M}$. It follows that $\mathcal{M}$ = span $\{M_I: I \in \mathcal{J}\}$. However M_I = span $\{L_J^{\ell}: J \subset I\}$ as well. Hence $\mathcal{M}$ = span $\{L_J^{\ell}: J \in \mathcal{J}\}$. If I is maximal the same argument yields

$$\tilde{\mathcal{M}}_I = \text{span } \{M_J: J \neq I,\ J \in \mathcal{J}\}$$

$$= \text{span } \{L_J^{\ell}: J \neq I,\ J \in \mathcal{J}\}.$$

Hence

$$L_I^{\ell} = \mathcal{M}/\tilde{\mathcal{M}}_I = L_I^{a}. \qquad \text{Q.E.D.}$$

Proof of Corollary 4.1. It suffices to check that $\tilde{P}_I \mu = \tilde{P}_I \mu_I$. For $J \neq I$, $\mu_J \in L_J \subset M_J \subset \tilde{\mathcal{M}}_I$. Since $\tilde{\mathcal{M}}_I$ and L_I^{ℓ} are mutually orthogonal we have $\tilde{P}_I \mu_J = 0$. Hence $\tilde{P}_I \mu = \tilde{P}_I(\Sigma_J \mu_J) = \tilde{P}_I \mu_I$. Q.E.D.

Acknowledgement. The author wishes to thank Prof. H. Yanai for his valuable suggestions.

References

Jacobsen, R. L. (1968). Linear algebra and ANOVA. Thesis. Cornell University.

John, P. W. M. (1971). Statistical Design and Analysis of Experiments. Macmillan, New York.

Kempthorne, O. (1952). The Design and Analysis of Experiments. Wiley, New York.

Lewis, S. M. and John, J. A. (1976). Testing main effects in fractions of asymmetrical factorial experiments. Biometrika, 63, 678-680.

Mukerjee, R. (1980). Orthogonal fractional factorial plans. Calcutta Statist. Assoc. Bull., 29, 143-160.

Pukelsheim, F. (1983). On optimality properties of simple block designs in the approximate design theory. J. Statistical Planning and Inference, 8, 193-208.

Rao, C. R. (1973). Linear Statistical Inference and Its Applications, 2nd ed. Wiley, New York.

Scheffé, H. (1959). The Analysis of Variance. Wiley, New York.

Seber, G. A. F. (1964). Orthogonality in analysis of variance. Ann. Math. Statist., 35, 705-710.

Smith, H. F. (1951). Analysis of variance with unequal but proportionate numbers of observations in the sub-classes of a two-way classification. Biometrics, 7, 70-74.

Sokolnikoff, I. S. (1964). Tensor analysis, theory and applications to geometry and mechanics of continua, 2nd ed. Wiley, New York.

Takemura, A. (1983). Tensor analysis of ANOVA decomposition. J. Amer. Stat. Assoc., 78, 894-900.

Takeuchi, K., Yanai, H., and Mukherjee, B. N. (1982). The Foundations of Multivariate Analysis. Wiley Eastern, New Delhi.

Tjur, T. (1982). Variance component models in orthogonal designs. Preprint 3, Institute of Mathematical Statistics, University of Copenhagen.