A METHOD OF ANALYSIS FOR UNEQUAL NUMBERS OF REPLICATES IN A FACTORIAL EXPERIMENT

BU-221-M U. B. Paik July, 1966

Abstract

A method of analysis for unequal numbers of replicates in a factorial experiment is presented using Searle's [1966] method. A numerical example is included to compare with the results of Federer and Zelen [1966].

Biometrics Unit, Plant Breeding Department, Cornell University.

A METHOD OF ANALYSIS FOR UNEQUAL NUMBERS OF REPLICATES IN A FACTORIAL EXPERIMENT

BU-221-M U. B. Paik July, 1966

In a 4 X 3 X 2 factorial experiment in which the factors are A consisting of four levels, B of three levels and C of two levels, and if we want to test

$$H_0: A_L = 0 \quad , \quad H_a: A_L \neq 0$$

$$H_0: A_Q = 0 \quad , \quad H_a: A_Q \neq 0$$

$$\vdots$$

$$H_0: A_C B_Q C = 0 \quad , \quad H_a: A_C B_Q C \neq 0$$

or

$$H_0: A_L = A_Q = A_C = 0 \quad , \quad H_a: \text{not } H_0$$

$$H_0: B_L = B_Q = 0 \quad , \quad H_a: \text{not } H_0$$

$$\vdots$$

$$H_0: A_L B_L C = A_L B_Q C = A_Q B_L C = A_Q B_Q C = A_C B_L C = A_C B_Q C = 0 \ , \ H_a: \text{not } H_0,$$

then we will be able to use the following method.

Let

$$Y_{ijkh} = n_{ijkh} (\mu + \tau_{ijk} + \epsilon_{ijkh}) \qquad (1)$$

where

$$E(Y_{ijkh}) = \mu + \tau_{ijk}$$

$$E(\epsilon\epsilon') = \sigma^2 I$$

and

$$i = 0, 1, 2, 3; \quad j = 0, 1, 2; \quad k = 0, 1; \quad h = 0, 1, \cdots, (n_{ijk\cdot} - 1)$$

Biometrics Unit, Plant Breeding Department, Cornell University.

n_{ijkh} = 1 if the treatment (ijk) is replicated at the h^{th} time and otherwise zero.

Let X be the design matrix,

then

$$X'X = \begin{bmatrix} n_{....} & n_{000\cdot} & n_{001\cdot} & \cdots & n_{321\cdot} \\ n_{000\cdot} & n_{000\cdot} & 0 & \cdots & 0 \\ n_{001\cdot} & 0 & n_{001\cdot} & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ n_{321\cdot} & 0 & 0 & \cdots & n_{321\cdot} \end{bmatrix} \tag{2}$$

Generalized inverse of $X'X$ is

$$G = (X'X)^{-} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & \frac{1}{n_{000\cdot}} & 0 & \cdots & 0 \\ 0 & 0 & \frac{1}{n_{001\cdot}} & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \frac{1}{n_{321\cdot}} \end{bmatrix} \tag{3}$$

$$\text{for which } H = GX'X = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \tag{4}$$

Then

$$\underline{\hat{\tau}} = GX'Y + (H - I)Z \qquad (5)$$

where Z is arbitrary.

Let Z = 0, then we have

$$_0\underline{\hat{\tau}} = \begin{bmatrix} _0\hat{\mu} \\ _0\hat{\tau}_{000} \\ _0\hat{\tau}_{001} \\ _0\hat{\tau}_{010} \\ \vdots \\ _0\hat{\tau}_{321} \end{bmatrix} = \begin{bmatrix} 0 \\ Y_{000\cdot}/n_{000\cdot} \\ Y_{001\cdot}/n_{001\cdot} \\ Y_{010\cdot}/n_{010\cdot} \\ \vdots \\ Y_{321\cdot}/n_{321\cdot} \end{bmatrix} \qquad (6)$$

Let's consider the following linear comparisons:

$$
\begin{bmatrix} A_L \\ A_Q \\ A_C \\ B_L \\ B_Q \\ C \\ A_LB_L \\ A_LB_Q \\ A_QB_L \\ A_QB_Q \\ A_CB_L \\ A_CB_Q \\ A_LC \\ A_QC \\ A_CC \\ B_LC \\ B_QC \\ A_LB_LC \\ A_LB_QC \\ A_QB_LC \\ A_QB_QC \\ A_CB_LC \\ A_CB_QC \end{bmatrix}
=
\begin{bmatrix}
0 & -3 & -3 & -3 & -3 & -3 & -3 & -1 & -1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & 1 & 1 & 3 & 3 & 3 & 3 & 3 & 3 \\
0 & 1 & 1 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & 1 & 1 \\
0 & -1 & -1 & -1 & -1 & -1 & -1 & 3 & 3 & 3 & 3 & 3 & 3 & -3 & -3 & -3 & -3 & -3 & -3 & 1 & 1 & 1 & 1 & 1 & 1 \\
0 & -1 & -1 & 0 & 0 & 1 & 1 & -1 & -1 & 0 & 0 & 1 & 1 & -1 & -1 & 0 & 0 & 1 & 1 & -1 & -1 & 0 & 0 & 1 & 1 \\
0 & 1 & 1 & -2 & -2 & 1 & 1 & 1 & 1 & -2 & -2 & 1 & 1 & 1 & 1 & -2 & -2 & 1 & 1 & 1 & 1 & -2 & -2 & 1 & 1 \\
0 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 \\
0 & 3 & 3 & 0 & 0 & -3 & -3 & 1 & 1 & 0 & 0 & -1 & -1 & -1 & -1 & 0 & 0 & 1 & 1 & -3 & -3 & 0 & 0 & 3 & 3 \\
0 & -3 & -3 & 6 & 6 & -3 & -3 & -1 & -1 & 2 & 2 & -1 & -1 & 1 & 1 & -2 & -2 & 1 & 1 & 3 & 3 & -6 & -6 & 3 & 3 \\
0 & -1 & -1 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & -1 & -1 & 1 & 1 & 0 & 0 & -1 & -1 & -1 & -1 & 0 & 0 & 1 & 1 \\
0 & 1 & 1 & -2 & -2 & 1 & 1 & -1 & -1 & 2 & 2 & -1 & -1 & -1 & -1 & 2 & 2 & -1 & -1 & 1 & 1 & -2 & -2 & 1 & 1 \\
0 & 1 & 1 & 0 & 0 & -1 & -1 & -3 & -3 & 0 & 0 & 3 & 3 & 3 & 3 & 0 & 0 & -3 & -3 & -1 & -1 & 0 & 0 & 1 & 1 \\
0 & -1 & -1 & 2 & 2 & -1 & -1 & 3 & 3 & -6 & -6 & 3 & 3 & -3 & -3 & 6 & 6 & -3 & -3 & 1 & 1 & -2 & -2 & 1 & 1 \\
0 & 3 & -3 & 3 & -3 & 3 & -3 & 1 & -1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & -1 & 1 & -3 & 3 & -3 & 3 & -3 & 3 \\
0 & -1 & 1 & -1 & 1 & -1 & 1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & -1 & 1 \\
0 & 1 & -1 & 1 & -1 & 1 & -1 & -3 & 3 & -3 & 3 & -3 & 3 & 3 & -3 & 3 & -3 & 3 & -3 & -1 & 1 & -1 & 1 & -1 & 1 \\
0 & 1 & -1 & 0 & 0 & -1 & 1 & 1 & -1 & 0 & 0 & -1 & 1 & 1 & -1 & 0 & 0 & -1 & 1 & 1 & -1 & 0 & 0 & -1 & 1 \\
0 & -1 & 1 & 2 & -2 & -1 & 1 & -1 & 1 & 2 & -2 & -1 & 1 & -1 & 1 & 2 & -2 & -1 & 1 & -1 & 1 & 2 & -2 & -1 & 1 \\
0 & -3 & 3 & 0 & 0 & 3 & -3 & -1 & 1 & 0 & 0 & 1 & -1 & 1 & -1 & 0 & 0 & -1 & 1 & 3 & -3 & 0 & 0 & -3 & 3 \\
0 & 3 & -3 & -6 & 6 & 3 & -3 & 1 & -1 & -2 & 2 & 1 & -1 & -1 & 1 & 2 & -2 & -1 & 1 & -3 & 3 & 6 & -6 & -3 & 3 \\
0 & 1 & -1 & 0 & 0 & -1 & 1 & -1 & 1 & 0 & 0 & 1 & -1 & -1 & 1 & 0 & 0 & 1 & -1 & 1 & -1 & 0 & 0 & -1 & 1 \\
0 & -1 & 1 & 2 & -2 & -1 & 1 & 1 & -1 & -2 & 2 & 1 & -1 & 1 & -1 & -2 & 2 & 1 & -1 & -1 & 1 & 2 & -2 & -1 & 1 \\
0 & -1 & 1 & 0 & 0 & 1 & -1 & 3 & -3 & 0 & 0 & -3 & 3 & -3 & 3 & 0 & 0 & 3 & -3 & 1 & -1 & 0 & 0 & -1 & 1 \\
0 & 1 & -1 & -2 & 2 & 1 & -1 & -3 & 3 & 6 & -6 & -3 & 3 & 3 & -3 & -6 & 6 & 3 & -3 & -1 & 1 & 2 & -2 & -1 & 1
\end{bmatrix}
\begin{bmatrix} \mu \\ \tau_{000} \\ \tau_{001} \\ \tau_{010} \\ \tau_{011} \\ \tau_{020} \\ \tau_{021} \\ \tau_{100} \\ \tau_{101} \\ \tau_{110} \\ \tau_{111} \\ \tau_{120} \\ \tau_{121} \\ \tau_{200} \\ \tau_{201} \\ \tau_{210} \\ \tau_{211} \\ \tau_{220} \\ \tau_{221} \\ \tau_{300} \\ \tau_{301} \\ \tau_{310} \\ \tau_{311} \\ \tau_{320} \\ \tau_{321} \end{bmatrix}
\quad (7)
$$

We can rewrite (7) as follows:

$$\underline{b} = L'\underline{\tau} \tag{8}$$

Clearly

$$L'H = L' \tag{9}$$

then the following linear combinations are estimable

$$L'\underline{\tau} \tag{10}$$

and their best unbiased linear estimators are

$$\hat{\underline{b}} = L'_0\hat{\underline{\tau}} \; ; \tag{11}$$

$$\text{var}[L'_0\hat{\underline{\tau}}] = L'GL\sigma^2 \; . \tag{12}$$

Now, using the method used by S. R. Searle [1966] in the class on linear hypotheses, the following testing of hypothesis will be possible. For example,

(I) H_0: $A_L = 0$, H_a: $A_L \neq 0$

$$SS_{A_L} = [Q'_0\hat{\underline{\tau}}]'(Q'GQ)^{-1}[Q'_0\hat{\underline{\tau}}] \tag{13}$$

where

$$Q' = (0\ -3\ -3\ -3\ -3\ -3\ -3\ -1\ -1\ -1\ -1\ -1\ -1\ 1\ 1\ 1\ 1\ 1\ 1\ 3\ 3\ 3\ 3\ 3\ 3)$$

and

$$F = \frac{SS_{A_L}}{\hat{\sigma}^2} \; , \quad \text{degrees of freedom: } (1, 19) \tag{14}$$

I conjecture that

$$E[SS_{A_L}] = (Q'GQ)^{-1}\,\sigma^2_{A_L} + \sigma^2 \tag{15}$$

(II) $\qquad H_0: \ A_L = A_Q = A_C = 0 \ , \ H_a: \ \text{not } H_0$

$$SS_A = [Q'_0\hat{\tau}]'(Q'GQ)^{-1}[Q'_0\hat{\tau}] \tag{16}$$

where

$$Q' = \begin{bmatrix} 0 & -3 & -3 & -3 & -3 & -3 & -3 & -1 & -1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & 1 & 1 & 3 & 3 & 3 & 3 & 3 & 3 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & -1 & -1 & -1 & -1 & -1 & -1 & 3 & 3 & 3 & 3 & 3 & 3 & -3 & -3 & -3 & -3 & -3 & -3 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

and

$$F = \frac{SS_A}{3\hat{\sigma}^2} \ , \quad \text{degrees of freedom: } (3, 19) \tag{17}$$

I conjecture that

$$E[SS_A] = E\begin{bmatrix} A_L \\ A_Q \\ A_C \end{bmatrix}' (Q'GQ)^{-1} \begin{bmatrix} A_L \\ A_Q \\ A_C \end{bmatrix} + 3\sigma^2 \tag{18}$$

$$= \alpha_{11}\sigma^2_{A_L} + \alpha_{22}\sigma^2_{A_Q} + \alpha_{33}\sigma^2_{A_C} + 3\sigma^2$$

where

$$(Q'GQ)^{-1} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix}$$

The computing result above should not be different from the Federer and Zelen [1966] computing result, except for rounding errors. The latter method inverts a much smaller matrix than the method described here. The Cornell University Computing Center CUSTAT program called MINT (matrix inversion) was used to obtain the results in the following example.

Example:

	B_0		B_1		B_2	
	C_0	C_1	C_0	C_1	C_0	C_1
A_0	(000) 1.11 0.97	(001) 1.52 1.45	(010) 1.09 0.99	(011) 1.27 1.22	(020) 1.21	(021) 1.24
A_1	(100) 1.30 1.00	(101) 1.55 1.53	(110) 1.03 1.21	(111) 1.24 1.34	(120) 1.12 0.96	(121) 1.27
A_2	(200) 1.22 1.13	(201) 1.38 1.08	(210) 1.34 1.41	(211) 1.40 1.21	(220) 1.34 1.19	(221) 1.46 1.39
A_3	(300) 1.19 1.03	(301) 1.29	(310) 1.36 1.16	(311) 1.42 1.39	(320) 1.46 1.03	(321) 1.62

Analysis of Variance

Source of variation	d.f.	SS	Null hypothesis for F test
Total	43	69.3586	
Correction for mean	1	68.115684	
Among groups	23	0.936266	$A_L = A_Q = A_C = B_C = B_Q = C = \cdots = A_C B_Q C = 0$ or $\tau_{000} = \tau_{001} = \cdots = \tau_{321}$
Within groups	19	0.306650	
A	3	0.076645	$A_L = A_Q = A_C = 0$
B	2	0.010012	$B_L = B_Q = 0$
C	1	0.369602	$C = 0$
AB	6	0.212323	$A_L B_L = A_L B_Q = A_Q B_L = A_Q B_Q = A_C B_L = A_C B_Q = 0$
AC	3	0.080971	$A_L C = A_Q C = A_C C = 0$
BC	2	0.045573	$B_L C = B_Q C = 0$
ABC	6	0.083617	$A_C B_L C = A_L B_Q C = A_Q B_L C = A_Q B_Q C = A_C B_L C = A_C B_Q C = 0$

References

Federer, W. T. and Zelen, M. [1966] "Analysis of Multifactor Classifications with Unequal Numbers of Observations." Biometrics, September.

Searle, S. R. [1966] Matrix Algebra for the Biological Sciences. John Wiley and Sons, New York. pp. 262-279.