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

U. B. Paik

BU-221-M

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

\[
H_0: \alpha_L = 0, \quad H_a: \alpha_L \neq 0
\]

\[
H_0: \alpha_Q = 0, \quad H_a: \alpha_Q \neq 0
\]

\[
H_0: \alpha_{BQ} = 0, \quad H_a: \alpha_{BQ} \neq 0
\]

or

\[
H_0: \alpha_L = \alpha_Q = \alpha_C = 0, \quad H_a: \text{ not } H_0
\]

\[
H_0: \beta_L = \beta_Q = 0, \quad H_a: \text{ not } H_0
\]

\[
H_0: \alpha_{LQ} = \alpha_{LQ} = \alpha_Q \beta_C = \alpha_Q \beta_C = \alpha_C \beta_Q = \alpha_C \beta_Q = 0, \quad H_a: \text{ not } H_0
\]

then we will be able to use the following method.

Let

\[
y_{ijth} = n_{ijth} (\mu + \tau_{ijk} + \epsilon_{ijth}) \tag{1}
\]

where

\[
E(y_{ijth}) = \mu + \tau_{ijk}
\]

\[
E(\epsilon \epsilon') = \sigma^2 I
\]

and

\[
i = 0, 1, 2, 3; \quad j = 0, 1, 2; \quad k = 0, 1; \quad h = 0, 1, \ldots, (n_{ijk} - 1)
\]

Biometrics Unit, Plant Breeding Department, Cornell University.
\[ n_{ijkh} = 1 \text{ if the treatment (ijk) is replicated at the } h^{th} \text{ time} \]

and otherwise zero.

Let \( X \) be the design matrix, then

\[
X'X = \begin{bmatrix}
  n_{000} & n_{001} & \cdots & n_{321} \\
  n_{000} & 0 & \cdots & 0 \\
  n_{001} & 0 & \cdots & 0 \\
  \vdots & \vdots & \ddots & \vdots \\
  n_{321} & 0 & 0 & \cdots & n_{321}
\end{bmatrix}
\]  

(2)

Generalized inverse of \( X'X \) is

\[
G = (X'X)^{-1} = \begin{bmatrix}
  0 & 0 & 0 & \cdots & 0 \\
  0 & \frac{1}{n_{000}} & 0 & \cdots & 0 \\
  0 & 0 & \frac{1}{n_{001}} & \cdots & 0 \\
  \vdots & \vdots & \vdots & \ddots & \vdots \\
  0 & 0 & 0 & \cdots & \frac{1}{n_{321}}
\end{bmatrix}
\]  

(3)

for which \( H = GX'X = \)

\[
\begin{bmatrix}
  0 & 0 & 0 & \cdots & 0 \\
  1 & 1 & 0 & \cdots & 0 \\
  1 & 0 & 1 & \cdots & 0 \\
  \vdots & \vdots & \vdots & \ddots & \vdots \\
  1 & 0 & 0 & 0 & \cdots & 1
\end{bmatrix}
\]  

(4)
Then

\[ \hat{\tau} = GX'Y + (H - I)Z \]

where \( Z \) is arbitrary.

Let \( Z = 0 \), then we have

\[
\begin{bmatrix}
\hat{\tau}^\mu \\
\hat{\tau}^{000} \\
\hat{\tau}^{001} \\
\hat{\tau}^{010} \\
\vdots \\
\hat{\tau}^{321}
\end{bmatrix}
= 
\begin{bmatrix}
0 \\
y_{000}/n_{000} \\
y_{001}/n_{001} \\
y_{010}/n_{010} \\
\vdots \\
y_{321}/n_{321}
\end{bmatrix}
\]

Let's consider the following linear comparisons:
We can rewrite (7) as follows:
\[ \hat{y} = L' \hat{\tau} \] (8)

Clearly
\[ L' \Sigma L = L' \] (9)

then the following linear combinations are estimable
\[ L' \hat{\tau} \] (10)

and their best unbiased linear estimators are
\[ \hat{\tau} = L' \hat{\hat{\tau}} \] ; (11)

\[ \text{var}[L' \hat{\hat{\tau}}] = L' \Sigma L \sigma^2 \] . (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) \[ H_0: \; A_L = 0 \; , \; H_\alpha: \; A_L \neq 0 \]

\[ SS_{A_L} = [Q' \hat{\hat{\tau}}] (Q'Q)^{-1} [Q' \hat{\hat{\tau}}] \] (13)

where,
\[ Q' = (0 -3 -3 -3 -3 -3 -1 -1 -1 -1 1 1 1 1 1 1 1 1 1 1 1 1 3 3 3 3 3 3 3 3 3) \]

and
\[ F = \frac{SS_{A_L}}{\hat{\hat{\tau}}} \] , degrees of freedom: (1, 19) (14)

I conjecture that
\[ E[SS_{A_L}] = (Q'Q)^{-1} \sigma^2_{A_L} + \sigma^2 \] (15)
(II)  \[ H_0: A_L = A_Q = A_C = 0 \quad \text{and} \quad H_a: \text{not } H_0 \]

\[ SS_A = [Q'Q']^{-1} (Q'Q) [Q'Q']^{-1} \]

where

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

1 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 1 & 1 & 1 & 1 & 1

-1 & -1 & -1 & -1 & -1 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 1 & 1 & 1 & 1 & 1
\end{bmatrix} \]

and

\[ F = \frac{SS_A}{3\sigma^2} \quad \text{degrees of freedom: } (3, 19) \]

I conjecture that

\[ E[SS_A] = E \left[ \begin{bmatrix} A_L & A_Q & A_C \end{bmatrix}' \right] (Q'Q) \left[ \begin{bmatrix} A_L & A_Q & A_C \end{bmatrix} \right] + 3\sigma^2 \]

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

where

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

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:

<table>
<thead>
<tr>
<th></th>
<th>$B_0$</th>
<th></th>
<th>$B_1$</th>
<th></th>
<th>$B_2$</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$c_0$</td>
<td>$c_1$</td>
<td>$c_0$</td>
<td>$c_1$</td>
<td>$c_0$</td>
<td>$c_1$</td>
</tr>
<tr>
<td>$A_0$</td>
<td>(000)</td>
<td>(001)</td>
<td>(010)</td>
<td>(011)</td>
<td>(020)</td>
<td>(021)</td>
</tr>
<tr>
<td></td>
<td>1.11</td>
<td>1.52</td>
<td>1.09</td>
<td>1.27</td>
<td>1.21</td>
<td>1.24</td>
</tr>
<tr>
<td></td>
<td>0.97</td>
<td>1.45</td>
<td>0.99</td>
<td>1.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$A_1$</td>
<td>(100)</td>
<td>(101)</td>
<td>(110)</td>
<td>(111)</td>
<td>(120)</td>
<td>(121)</td>
</tr>
<tr>
<td></td>
<td>1.30</td>
<td>1.55</td>
<td>1.03</td>
<td>1.24</td>
<td>1.12</td>
<td>1.27</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>1.53</td>
<td>1.21</td>
<td>1.34</td>
<td>0.96</td>
<td></td>
</tr>
<tr>
<td>$A_2$</td>
<td>(200)</td>
<td>(201)</td>
<td>(210)</td>
<td>(211)</td>
<td>(220)</td>
<td>(221)</td>
</tr>
<tr>
<td></td>
<td>1.22</td>
<td>1.38</td>
<td>1.34</td>
<td>1.40</td>
<td>1.34</td>
<td>1.46</td>
</tr>
<tr>
<td></td>
<td>1.13</td>
<td>1.08</td>
<td>1.41</td>
<td>1.21</td>
<td>1.19</td>
<td>1.39</td>
</tr>
<tr>
<td>$A_3$</td>
<td>(300)</td>
<td>(301)</td>
<td>(310)</td>
<td>(311)</td>
<td>(320)</td>
<td>(321)</td>
</tr>
<tr>
<td></td>
<td>1.19</td>
<td>1.29</td>
<td>1.36</td>
<td>1.42</td>
<td>1.46</td>
<td>1.62</td>
</tr>
<tr>
<td></td>
<td>1.03</td>
<td>1.16</td>
<td>1.39</td>
<td>1.03</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Analysis of Variance

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>d.f.</th>
<th>SS</th>
<th>Null hypothesis for F test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>43</td>
<td>69.3596</td>
<td></td>
</tr>
<tr>
<td>Correction for mean</td>
<td>1</td>
<td>68.115684</td>
<td></td>
</tr>
<tr>
<td>Among groups</td>
<td>23</td>
<td>0.936266</td>
<td>( A_L = A_Q = A_C = B_C = B_Q = C = \cdots = A_C B_Q C = 0 )</td>
</tr>
<tr>
<td>Within groups</td>
<td>19</td>
<td>0.306550</td>
<td>( \tau_{000} = \tau_{001} = \cdots = \tau_{321} )</td>
</tr>
<tr>
<td>A</td>
<td>3</td>
<td>0.076645</td>
<td>( A_L = A_Q = A_C = 0 )</td>
</tr>
<tr>
<td>B</td>
<td>2</td>
<td>0.010012</td>
<td>( B_L = B_Q = 0 )</td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>0.369602</td>
<td>( C = 0 )</td>
</tr>
<tr>
<td>AB</td>
<td>6</td>
<td>0.212323</td>
<td>( 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 )</td>
</tr>
<tr>
<td>AC</td>
<td>3</td>
<td>0.080971</td>
<td>( A_L C = A_Q C = A_C C = 0 )</td>
</tr>
<tr>
<td>BC</td>
<td>2</td>
<td>0.045573</td>
<td>( B_L C = B_Q C = 0 )</td>
</tr>
<tr>
<td>ABC</td>
<td>6</td>
<td>0.083617</td>
<td>( A_L B_C = A_B C = A_Q B_C = A_Q B_Q C = A_C B_C = A_C B_Q C = 0 )</td>
</tr>
</tbody>
</table>

References
