Thoughts on Interaction

Ronald Christensen

Department of Mathematics and Statistics

University of New Mexico

November 16, 2016

Abstract

KEY WORDS:

The first section examines interactions in an unbalanced two-way ANOVA. The second section uses the first to establish that the two-way ANOVA displays orthogonality if and only if the unbalanced numbers are proportional. The third section looks at Boik's work on interaction testing for proportional numbers.

1. Characterizing the Interaction Space

In a two-way ANOVA with interaction and unbalanced numbers, the cell-means parameterization is

$$y_{ijk} = \mu_{ij} + \varepsilon_{ijk}, \qquad i = 1, \ldots, a, \quad j = 1, \ldots, b, \quad k = 1, \ldots, N_{ij},$$

which is really just a one-way ANOVA with unequal numbers and the pair of subscripts ij identifying the ab different groups. The traditional parameterization is $\mu_{ij} = \mu + \alpha_i + \eta_j + \gamma_{ij}$. Write the cell-means model in matrix form as $Y = X\boldsymbol{\mu} + e$. X has ab columns and the $i'j'$ column has the form

$$X_{i'j'} = [t_{ijk}], \quad \text{with} \quad t_{ijk} = \delta_{(i,j)(i',j')}$$

where for any two symbols a and b, the kronecker delta has

$$\delta_{ab} = \begin{cases} 1 & \text{if } a = b \\ 0 & \text{if } a \neq b. \end{cases}$$

Note that

$$C(X) = \{v | v = [v_{ijk}], \quad \text{with} \quad v_{ijk} = \mu_{ij} \text{ for some } \mu_{ij}\}.$$

Similar to the balanced case, we will see that the interaction space is the set of vectors

$$T = [t_{ijk}], \quad \text{with} \quad t_{ijk} = q_{ij}/N_{ij} \qquad \text{where} \quad q_{i\cdot} = 0 = q_{\cdot j}, \quad \text{for all } i, j.$$

An interaction contrast is $T'X\boldsymbol{\mu} = \sum_{ij} q_{ij}\mu_{ij} = \sum_{ij} q_{ij}\gamma_{ij}$. Clearly, $T \in C(X)$, so $T'MY = T'Y$ where M is the perpendicular projection operator (ppo) onto $C(X)$. It follows that the least squares estimate of $T'X\boldsymbol{\mu}$ is $T'Y = \sum_{ij} q_{ij}\bar{y}_{ij\cdot}$. These are essentially the same results as for balanced ANOVA.

To see that vectors T characterize the interaction space, write the corresponding main effects model

$$Y = J\mu + X_\alpha \alpha + X_\eta \eta + e.$$

The matrix X_α has columns

$$X_{i'} = [t_{ijk}], \quad t_{ijk} = \delta_{ii'} \qquad i' = 1, \ldots, a$$

The matrix X_η has columns

$$X_{a+j'} = [t_{ijk}], \quad t_{ijk} = \delta_{jj'} \qquad j' = 1, \ldots, b$$

By definition, the interaction space is $C(J, X_\alpha, X_\eta)^{\perp}_{C(X)}$. Thus, the characterization of the interaction space results from vectors of the form T spanning a space of sufficient rank [$(a-1)(b-1)$ when $N_{ij} > 0$ for all ij], having $T \in C(X)$, and $X_h'T = 0$, $h = 1, \ldots, a+b$ [which follows from the definitions and arithmetic]. Note that to have orthogonal interaction contrasts we need the T vectors corresponding to the interaction contrasts to be orthogonal.

Note also that all interaction contrasts are contrasts in the μ_{ij} one-way model, but not all μ_{ij} contrasts are interaction contrasts. As in PA Chapter 4, an arbitrary element of the μ_{ij} contrast space is

$$S = [s_{ijk}], \quad \text{with} \quad s_{ijk} = s_{ij}/N_{ij} \qquad \text{where} \quad s_{\cdot\cdot} = 0.$$

In addition to the interaction space being a subset of the contrast space, there is an $a-1$ dimensional subspace of the contrast space consisting of vectors

$$T_\alpha = [t_{ijk}], \quad \text{with} \quad t_{ijk} = c_i/bN_{ij} \qquad \text{where} \quad c_\cdot = 0$$

and a $b-1$ dimensional subspace of vectors

$$T_\eta = [t_{ijk}], \quad \text{with} \quad t_{ijk} = d_j/aN_{ij} \qquad \text{where} \quad d_\cdot = 0.$$

These define contrasts in the interaction model of

$$T_\alpha' X\boldsymbol{\mu} = \sum_{ij} c_i \mu_{ij}/b = \sum_{ij} c_i \bar{\mu}_{i\cdot} = \sum_i c_i(\alpha_i + \bar{\gamma}_{i\cdot})$$

with estimate $T_\alpha' Y = \sum_i c_i \bar{y}_{i\cdot}$ and

$$T_\eta' X\boldsymbol{\mu} = \sum_{ij} d_j \mu_{ij}/a = \sum_{ij} d_j \bar{\mu}_{\cdot j} = \sum_j d_j(\eta_j + \bar{\gamma}_{\cdot j})$$

with estimate $T\eta' Y = \sum_j d_j \bar{y}_{\cdot j}$. Under proportional numbers, these three spaces are orthogonal but in general they intersect in the zero vector.

For the main effects model, the vectors T_α and T_η typically are not in $C(J, X_\alpha, X_\eta)$, so although, for example, $T_\alpha'[J\mu + X_\alpha\alpha + X_\eta\eta] = \sum_i c_i\alpha_i$, the estimate is *not* $T_\alpha' Y$, it is $T_\alpha' M_0 Y$ where M_0 is the ppo onto $C(J, X_\alpha, X_\eta)$.

2. Orthogonal Main Effects iff Proportional Numbers

An interesting sidelight from writing the unbalanced model this way is that correcting for the grand mean gives

$$[I - (1/n)JJ']X_{i'} = [t_{ijk}], \quad t_{ijk} = \delta_{ii'} - \frac{N_{i'\cdot}}{n}$$

and

$$[I - (1/n)JJ']X_{a+j'} = [t_{ijk}], \quad t_{ijk} = \delta_{jj'} - \frac{N_{\cdot j'}}{n}$$

which implies that main effects are orthogonal iff the data have proportional numbers. To see this,

$$X_{i'}'[I-(1/n)JJ']X_{a+j'} = 0 \iff 0 = N_{i'j'} - \frac{N_{i'\cdot}N_{\cdot j'}}{n} = \sum_i \sum_j \sum_k \left(\delta_{ii'} - \frac{N_{i'\cdot}}{n}\right)\left(\delta_{jj'} - \frac{N_{\cdot j'}}{n}\right)$$

I don't remember if orthogonality implies proportional numbers is proven in the PA Chap 7. The reverse certainly is proven.

3. Thoughts on Testing Interaction

This requires some knowledge of Multivariate Analysis. We know how to test interaction, just test the μ_{ij} model against the additive model. As discussed above and in PA (Chap. 7), in the balanced case and now in the unbalanced case, we can characterize the interaction space. The interaction space involves contrasts $\sum_i \sum_j q_{ij}\mu_{ij}$ with q_{ij}s defined as above. But the most interpretable interaction contrasts have a special form (product interactions), built by combining a contrast in the α_is with a contrast in the η_j, see below. Boik developed a test for interactions of this specific form. However, it seems that the distribution theory was ugly. (Admittedly, I think that all distribution theory is ugly.) At least in some special cases, one can rewrite Boik's approach so that standard results from multivariate ANOVA equate to Boik's results. Boik's results apply when one has proportional numbers. The results below are more restrictive than proportional numbers.

1. Boik, R. J. (1986), Testing the Rank of a Matrix with Applications to the Analysis of Interaction in ANOVA, *Journal of the American Statistical Association*, **81**, 243-248.

2. Boik, R. J. (1993), The Analysis of Two-Factor Interactions in Fixed Effects Linear Models, *Journal of Educational and Behavioral Statistics*, **18**, 1-40.

Boik's is a test of *product interactions*, i.e., $q_{ij} \equiv c_i d_j$ where $c_\cdot = 0 = d_\cdot$. Note that these provide a basis for the interaction space by picking $a-1$ linearly independent α contrasts and $b-1$ linearly independent η contrasts. But vectors T with this form generate the entire interaction space, not just vectors of product form. Product form is not closed under vector addition.

Write

$$y_{ijk} = \mu + \alpha_i + \eta_j + \gamma_{ij} + \varepsilon_{ijk}, \qquad i = 1, \ldots, a, \qquad j = 1, \ldots, b, \quad k = 1, \ldots, N_{ij}.$$

Fix j and write a series of one-way ANOVA models

$$y_{ijk} = (\mu + \eta_j) + (\alpha_i + \gamma_{ij}) + \varepsilon_{ijk}, \qquad i = 1, \ldots, a, \quad k = 1, \ldots, N_{ij}.$$

When $N_{ij} = N_{ij'} = N_{i\cdot}/b$ for all i, j, j', we can treat this as a multivariate one-way model

$$[Y_1, \cdots, Y_b] = \boldsymbol{X}B + e.$$

The condition on the N_{ij} is a special case of proportional numbers. Note that X has $n \equiv N_{\cdot\cdot}$ rows, whereas $\boldsymbol{X}$ has $N_{\cdot\cdot}/b$ rows.

Let Z be a matrix whose columns are $b-1$ linearly independent sets of contrast coefficients. Think of fitting the model

$$YZ = \boldsymbol{X}(BZ) + eZ.$$

Our test for product interactions is based on the test for group ("α") effects in this one-way model. The test statistic is

$$H = Z'Y'M_{\alpha}YZ$$

where M_{α} is the ppo for group effects in the one-way. We could just use a standard multivariate linear model test, like Roy's, or we could try to incorporate the fact that the columns in this multivariate linear model are independent and homoscedastic to develop an even better test. Roy's maximum F test should be very similar to what is in Boik, especially the '93 paper. Boik showed that his results hold for arbitrary sets of proportional numbers, not just the special case considered here. Incidentally, if we reverse the roles of i and j we should get the same results provided $N_{ij} = N_{\cdot j}/a$, giving us two ways to achieve the needed balance.

Acknowledgement

My thanks to Robert for discussing his results with me. If I have misrepresented his work, the responsibility is entirely mine.