Contrast Analyses

Basic assumption of every inferential test:

\[
\text{DATA} = \text{MODEL} + \text{ERROR} ;
\]

With categorical variables that have three or more levels, it is suggested to test two conditions (Abelson & Prentice, 1997):

1. The MODEL (the hypothesis) explains a significant proportion of the DATA (of the variance), i.e., the contrast of interest is statistically significant.

2. The ERROR (the residual between-group variance) is so small that it explains only a non significant part of the DATA (of the variance), i.e., the “other” contrasts as a set are not statistically significant.

The different steps:

1. Determine the number of contrasts: If the categorical independent variable has \( m \) levels, you need \( m-1 \) contrasts.

2. Determine the contrast of interest: Find a contrast that best represents your operational hypothesis. You will need to find a contrast that is “centered,” i.e., a contrast for which the sum of the values is zero. For example, \( C_{1a} (-1, 0, +1) \), \( C_{1b} (+3, -1, -1, -1) \) et \( C_{1c} (-4, -1, +1, +2, +2) \) are centered contrasts, whereas \( C_{1d} (+1, 0, 0) \), \( C_{1e} (-1, 0, +1, +2) \) et \( C_{1f} (-1, -1, 0, 0, +1) \) are not.

3. Determine the “other” contrasts (sometimes called contrasts of disinterest): Find \( m-2 \) “other” contrasts that have no theoretical meaning and that have as a sole purpose test the residual (between-group) variance. You will need to find contrasts that are “centered” (see previous point) and “orthogonal” (the contrasts are not correlated with themselves or with the contrast of interest). Two contrasts are orthogonal if the sum of the products of the values is zero. See the appendix of Brauer and McClelland (2005) for more detailed explanations on how to find centered and orthogonal contrasts (pp. 302-305).

4. Test both conditions specified by Abelson and Prentice (1997): These tests depend on the design of the study (see below). In general, if the contrast of interest is significant and the “other” contrasts are not, one can conclude that the data confirm the hypothesis.

The case of a between-participant variable with three levels:

\[
Y = b_0 + b_1(C1) + b_2(C2) ;
\]

The regression coefficient \( b_1 \) tests the first condition; this condition is satisfied if \( b_1 \) is significant. The regression coefficient \( b_2 \) tests the second condition; this condition is satisfied if \( b_2 \) is non significant.
The case of a between-participant variable with more than 3 levels:

The example of an IV with four levels:

\[ Y = b_0 + b_1(C1) + b_2(C2) + b_3(C3) ; \]

The regression coefficient \( b_1 \) tests the first condition. The regression coefficients associated with the “other” contrasts (here \( b_2 \) and \( b_3 \)) test the second condition; it is necessary to test these “other” contrasts as a set (i.e., not individually) and with a one degree of freedom test. With equal \( N \), it is possible to simply add up the \( F \)s of the regression coefficients associated with the “other” contrasts. With unequal \( N \), one should use the following formula:

\[
F = \frac{SS_{\text{autres'' contrastes}}}{SS_{\text{error}} / df_{\text{error}}} \quad (\text{where } SS = \text{“sum of squares”})
\]

The case of an interaction between two between-participant variables:

The example of an IV with two levels (X1) and an IV with three levels (X2):

\[ Y = b_0 + b_1(X1) + b_2(C1) + b_3(C2) + b_4(X1\cdot C1) + b_5(X1\cdot C2) ; \]

The regression coefficient \( b_4 \) tests the first condition (the interaction hypothesis). The regression coefficients associated with the “other” interaction contrasts (here \( b_5 \)) test the second condition; if \( X2 \) has more than three levels, it is necessary to test the other interaction contrasts as a set and with one degree of freedom test.

The case of a within-participants variable:

The example of an IV with four levels (Y1, Y2, Y3, Y4):

It is necessary to compute \( m-1 \) centered, orthogonal contrasts that are each a linear combination of the measures that constitute the within-participant variable. For example:

\[
\begin{align*}
C_1 &= (-3)*Y1 + (+1)*Y2 + (+1)*Y3 + (+1)*Y4 ; \\
C_2 &= (0)*Y1 + (-2)*Y2 + (+1)*Y3 + (+1)*Y4 ; \\
C_3 &= (0)*Y1 + (0)*Y2 + (-1)*Y3 + (+1)*Y4 ;
\end{align*}
\]

Then get for each contrast \( SS_{IV} \), \( SS_{\text{error}} \), and \( df_{\text{error}} \). The first condition is tested with the formula \( F = (SS_{IV-C1}) / (SS_{\text{error-C1}} / df_{\text{error-C1}}) \). The second condition is tested with the following formula:

\[
F = \frac{\sum SS_{IV-\text{“other” contrast}}}{\sum SS_{\text{error-“other” contrast}} / \sum df_{\text{error-“other” contrast}}}
\]