Supplemental Instructions for Robust Contrasts

This document provides orienting information for you to use when conducting specialized contrast analyses for the Robust Contrasts program. I focus first on principles for means and then later consider other parameters, like median, M estimates, trimmed means, and quantiles.

Mean Contrasts

Many programs that that focus on means rely on omnibus tests of mean differences. However, when it is all said and done, we almost always move to the analysis of single degree of freedom contrasts that underlie those omnibus tests. In this section, I describe the framework I use to conceptualize one way and factorial between-subject designs vis-a-vis contrast analysis.

I make use of a general formula to calculate the estimate of a contrast parameter value in a population based on means. I illustrate the formula here using a 2X2 factorial design. In practice, it can be applied to any size design, from one factor designs to five or more factor designs. In this example, suppose I examine the effects of grade (8th grade versus 9th grade) and sex (male versus female) on how satisfied adolescents are with their relationship with their parents (measured on a scale ranging from 0 to 100, with higher scores indicating more satisfaction). Let

M_{8M} = the mean for 8th grade males

M_{8F} = the mean for 8th grade females

M_{9M} = the mean for 9th grade males

M_{9F} = the mean for 9th grade females

I begin by forming an equation that weights (i.e., multiplies) each of the means by a "contrast coefficient," as follows:

$$PE_1 = c_1 M_{8M} + c_2 M_{8F} + c_3 M_{9M} + c_4 M_{9F} \qquad [1]$$

where PE_1 stands for "parameter estimate" and is the estimated contrast parameter value of interest. I assign values to the contrast coefficients to isolate comparisons I am interested in. For example, suppose I want to estimate the population mean difference between 8th grade males and 8th grade females. This is $M_{8M} – M_{8F}$. I can express this difference using Equation 1 by assigning the values $c_1 = 1$, $c_2= -1$, $c_3 = 0$, and $c_4 = 0$. This yields

PE_1 = (1) M_{8M} + (-1) M_{8F} + (0) M_{9M} + (0) M_{9F}

= (1) M_{8M} + (-1) M_{8F}

= $M_{8M} – M_{8F}$

Contrast analysis involves estimating the above parameter (i.e., the group mean difference), calculating a confidence interval and margin of error for it, and testing it for "statistical significance" in a null hypothesis sense. If I want to compare ninth grade males with ninth grade females, I use the coefficients c1 = 0, c2= 0, c3 = 1, and c4 = -1. This yields

$PE_2 = (0)\ M_{8M} + (0)\ M_{8F} + (1)\ M_{9M} + (-1)\ M_{9F}$

$= (1)\ M_{9M} + (-1)\ M_{9F}$

$= M_{9M} – M_{9F}$

Suppose I want to estimate the mean difference between 8th graders and 9th graders, collapsing across biological sex. In this case, the mean for 8th graders is $(M_{8M} + M_{8F})/2$ and for 9th graders it is $(M_{9M} + M_{9F})/2$. The difference between these is $[(M_{8M} + M_{8F})/2] – [(M_{9M} + M_{9F})/2]$. I can isolate this parameter with contrast coefficients. Focus first on the mean for 8th graders, $(M_{8M} + M_{8F})/2$. This can be re-written as

$(M_{8M} + M_{8F})/2 = ½\ (M_{8M} + M_{8F})$

$= 0.5\ (M_{8M} + M_{8F})$

$= 0.5\ M_{8M} + 0.5\ M_{8F}$

Using similar logic, the mean for 9th graders $(M_{9M} + M_{9F})/2$ can be rewritten as

$(M_{9M} + M_{9F})/2 = ½\ (M_{9M} + M_{9F})$

$= 0.5\ (M_{9M} + M_{9F})$

$= 0.5\ M_{9M} + 0.5\ M_{9F}$

and I can isolate the difference by assigning the values $c_1 = .5$, $c_2= .5$, $c_3 = -.5$, and $c_4 = -.5$. This yields

$PE_3 = (.5)\ M_{8M} + (.5)\ M_{8F} + (-.5)\ M_{9M} + (-.5)\ M_{9F}$

If we are creative, we can determine a set of coefficients that will isolate most any focused contrast of interest. For example, suppose I want to know if the male-female difference for 8th graders is the same as the male-female difference for 9th graders. The male-female difference for 8th graders is

$(1)\ M8_M + (-1)\ M_{8F}$

and the male-female difference for 9th graders is

$(1)\ M_{9M} + (-1)\ M_{9F}$

and the difference between these two differences is found by

$PE_4 = (1)\ M_{8M} + (-1)\ M_{8F} + (-1)\ M_{9M} + (1)\ M_{9F}$

In factorial designs, the three contrasts I just specified are called (1) a simple effect contrast (PE1), (2) a main effect contrast (PE3) and (3) an interaction contrast (PE4). The important point for now is that I can use the contrast formula approach to isolate estimates of a wide range of population contrasts that may be of interest. For factorial designs, we lay the means out for every cell in the design as if it were a one way design and then apply contrasts coefficients to the different cell means to isolate the contrast we are interested in. This approach to contrast analysis is very flexible and useful.

Note that PE_1 and PE_2 correspond to simple effects in a factorial design, PE_3 is a main effect and PE_4 is an interaction/moderation effect. In this sense, I can express any between-subject factorial design as a one way design and use carefully chosen contrast coefficients to isolate the contrasts I am interested in. I can readily use this framework for ANCOVA by conceptualizing the means as adjusted means.

Most researchers use contrast coefficients that preserve the metric of the outcome variable but doing so is not necessary for purposes of significance tests. For example, if for three groups one tests the difference between the first group and the average of the second two groups, one could use either of the following set of coefficients:

1, -.5, -.5

2, -1, -1

Both will yield the same result for a significance test, but the parameter estimate for the latter will be twice the size of the parameter estimate for the former because the metric of the outcome has been doubled. I personally prefer to define coefficients so that I retain the original metric.

An issue also arises when there are unequal n in the groups in the population. Suppose I want to compare the mean of the first group with the average of the second two groups where the means and sample sizes are as follows:

Group 1: Mean (M1) = 100 n1 = 100

Group 2: Mean (M2) = 50 n2 = 50

Group 3: Mean (M3) = 30 n3 = 25

There are two ways we can combine groups 2 and 3. First, I can average their two means, yielding (50 + 30)/2 = 40. This method ignores the sample size differences and yields what is called an unweighted mean. Second, I can calculate a weighted mean that takes into account the sample sizes (which would be the same as merging the respondents in the two groups together and calculating the mean on this combined sample size). In this case, the weighted mean equals 43.333. The most common approach to contrast analysis in the social sciences uses the unweighted mean method.

The contrast coefficients in the above example for unweighted means analysis are 1, -.5, and -.5. If I want to work with weighted means, I have to define the coefficients differently. For the first group, the coefficient would still be 1, but for the second two groups, the coefficients would take into account the respective expected sample sizes. The sum of the sample sizes for the two groups is 50 + 25 = 75. The sample size for the first group divided by this value (50/75) yields a coefficient of = .66667 and for the second group, it is 25/75 = 0.3333. So instead of using 1, -.5, -5 per the unweighted case, I would use for the weighted case 1, -.6667, -.3333. The choice of which method to use depends on your conceptual questions.

A requirement of contrast analysis as used in my program is that the coefficients must sum to 0. Note that such is the case for all the above examples.

Contrasts for Other Types of Parameters

The above strategy for means can be applied to medians, M estimates, trimmed means, and quantiles. The primary caution with these other parameter types is that when you work with values that combine two or more of the groups into a single composite, such as comparing group 1 versus the average of groups 2 and 3, the composite parameter (e.g., the average median for groups 2 and 3) may not equal the median calculated on the raw scores of groups 2 and 3 combined. Be sure to carefully phrase your results to conform to what you have done in the contrast analyses, in this case, averaged the medians of groups 2 and 3.