MULTIPLE LINEAR REGRESSION VIEWPOINTS
VOLUME 18, NUMBER 1, FALL 1991

The Case for Non-Zero Restrictions in Statistical Analysis

Keith McNeil
New Mexico State University

One of the many advantages of MLR is its versatility and its ability to answer a vast array of questions. Unfortunately, most researchers fall into the habit of asking a small subset of very similar questions. The question being tested should be stated first, but can be identified from the full model and the restriction(s) placed on that full model. While the restrictions can take on any numerical value, almost all applications use the "default " value of zero:

1. $a_1 = a_2$ or $(a_1 - a_2 = 0)$ (t-test)
2. $a_1 = 0$ (Correlation)
3. $a_1 = a_2 = a_3 = \ldots a_i$ or $(a_1 - a_2 = a_2 - a_3 = a_3 - a_2 = \ldots 0)$ (F-test)
4. $(a_1 - a_2) = (a_3 - a_4)$ or $((a_1 - a_2) - (a_3 - a_4) = 0)$ (interaction)

The focus of this paper will be on the utility of making a non-zero restriction. Why the zero restriction occurs so frequently will be questioned and hopefully researchers and statisticians will see how the zero restriction limits the conclusions of the research. The argument will be made for making non-zero restrictions, resulting of course, from "non-zero" research hypotheses. The argument will be made for each of these statistical procedures: two group t test, Pearson correlation, single population mean, one-way analysis of variance, and interaction.

<u>Two Group t Test</u>

Perhaps the most widely used design compares the performance of two groups. The research hypothesis takes the following form:
Research Hypothesis 1: For a given population, the New treatment is better than the Traditional treatment on Y. (See Note 1 for discussion of directional hypothesis testing.)
Full Model: $Y = a_1N + a_2T + E_1$
Where Y = criterion of interest,
N = 1 if subject in New treatment; 0 otherwise, and
T = 1 if subject in Traditional treatment; 0 otherwise.
The research hypothesis implies that the sample mean for N should be greater than the sample mean for T, or $a_1 > a_2$, or $a_1 - a_2 > 0$.
Restriction: $a_1 = a_2$, or $(a_1 - a_2 = 0)$
Forcing the restriction into the full model results in:
Restricted Model: $Y = a_1N + a_1T + E_2$
But since the two vectors (N and T) are multiplied by the same weights, the vectors can be added first. But N + T equals the Unit vector (or everyone). Therefore:
Restricted Model: $Y = a_1U + E_2$

There are two linearly independent pieces of information in the full model. Forcing the one restriction on the full model results in one linearly independent piece of information in the restricted model. (See Note 2 for test of significance.)

A significant drop in the R^2 from the full model to the restricted model results in a significant F. If the sample means are in accord with the anticipated result, then Research Hypothesis 1 can be held as tenable and the conclusion would be: For the given population, the New treatment is better than the Traditional treatment on Y. But all that has been said is that the New treatment is better than the Traditional treatment. We do not know how much better; all we know is that the two treatments are not equally effective.

But what if the cost of the two treatments is not the same? The Traditional treatment has surely been somewhat effective in the past. The New treatment will surely require some additional cost in the form of special inservice, purchase of new materials, acceptance by teachers, students, and community, etc. Before the Traditional treatment is replaced by the New treatment, perhaps the researcher should demonstrate that there is, say, more than a five-point superiority of the New treatment over the Traditional treatment.

When a non-zero research hypothesis is proposed, other researchers and statisticians often ask for the justification for the actual non-zero value chosen, as they should. But why should more justification be required for a non-zero value than for a zero value? Or looking at the issue form the other side, why are researchers allowed to test a zero value with little or no justification. When one realizes that zero is only one of an infinite number of values, then one realizes that the same amount of justification should be required of a zero value as of a non-zero value. Furthermore, when one attempts to justify the zero value restriction, one may realize that zero is not the value of interest. Those researchers who have been defaulting with zero should know how to choose a value, but may not. It is not the intent of this paper to illustrate how one determines the magnitude of the value tested in the research hypothesis, although a few suggestions will be provided.

In the case where there was an expectation of a five-point superiority, the research hypothesis would be:

Research Hypothesis 2: For a given population, the New treatment is more than five points better than the Traditional treatment on Y.

Full Model: $Y = a_1N + a_2T + E_3$

The research hypothesis implies that the sample mean for the New treatment is more than five units greater than the sample mean for the treatment or, a_1 greater than $(a_2 + 5)$ or $(a_1 - a_2 > 5)$

Restriction: $a_1 = a_2 + 5$, or $(a_1 - a_2 = 5)$ or $(a_2 = a_1 - 5)$

Restricted Model:
$$Y = a_1N + (a_1 - 5)T + E_4$$
$$Y = a_1N + a_1T - 5T + E_4$$
$$(Y + 5T) = a_1(N + T) + E_4$$
$$(Y + 5T) = a_1U + E_4$$

There are two linearly independent pieces of information in the full model. Forcing the one restriction on the full model results in one linearly independent piece of information in the restricted model. (See Note 1 for test of significance.)

Notice that the full model in Research Hypothesis 1 is exactly the same as the full model in Research Hypothesis 2. The number of restrictions is also the same, resulting in the same number of degrees of freedom. What is different, though, is the nature of the restriction and hence the restricted models are different. The two research hypotheses are both "correct" and equally "valid" - they just test two different hypotheses. Research Hypothesis 2 provides a more definitive conclusion.

The actual "cost" of any treatment may be difficult to determine. But one must remember that Research Hypothesis 1 reduces to the default assumption that the "costs" are equal. The choice of a research hypothesis leading to a restriction of ($a_1 - a_2 = 0$) should be defended as much as a research hypothesis leading to a restriction of ($a_1 - a_2$ = aome non-zero value). The restriction ($a_1 - a_2 = 0$) has become a widely used default value, but we must realize that it is only one of an infinite number of values.

Pearson Correlation

The usual application of the Pearson correlation hypothesis is:

Research Hypothesis 3: For a given population, the linear correlation between X and Y is greater than zero.

Full Model: $Y = a_0U + a_1X + E_6$

The research hypothesis implies that the slope of the line of beat fit in the sample is positive, or $a_1 > 0$.

Reatriction: $a_1 = 0$

Restricted Model: $Y = a_0U + 0X + E_6$

$Y = a_0U + E_6$

There are two linearly independent pieces of information in the full model. Forcing the one restriction on the full model results in one linearly independent piece of information in the restricted model.

If the F test is significant, then one concludes that the research hypothesis is tenable, that the linear correlation between X and Y is greater than 0, or that the change in Y per unit change in X is greater than 0; but we do not know how much greater than 0. There may be reasons for wanting to know if the correlation is greater than a particular value. For instance, if the correlation under consideration is either a validity coefficient or a reliability coefficient, then we would definitely want a correlation coefficient above some specified value, such as:

Research Hypothesis 4: For a given population, the linear correlation between Y and the Retest of Y is greater than .80.

Full Model: $Y = a_0U + a_1R + E_7$

Restriction: Restricted Model $R2 = .64$

The research hypothesis implies that the restricted model R^2 will be $(.80)^2$ or .64. The formula in Note 2 can be used when testing this hypothesis for significance.

Consider Research Hypothesis 5: For a given population, there is more than a .6 unit change in Y for every unit change in X. In this case the models would be:

Full Model: $Y = a_0U + a_1X + E_8$

Restriction: $a_1 = .6$

Restricted Model: $Y = a_0U + .6X + E_9$

$(Y - .6X) = a_0U + E_9$

Notice that the full model in Research Hypotheses 3 and 4 is exactly the same as Research Hypothesis 5. The number of restrictions is also the same; resulting in the same number of degrees of freedom. What is different, though, is the nature of the restriction and hence the restricted models are different. The three research hypotheses are all "correct" and equally "valid" - they just test three different hypotheses. Research Hypotheses 4 and 5 provide more definitive conclusions.

The desired correlation (reliability, validity, etc.) may be difficult to determine, but should be no more difficult to justify than justifying the default value of 0. Just because $a_1 = 0$ has been used in the past does not justify its use, particularly with hypotheses about reliability and validity.

Single Population Mean

The usual application of the single population mean hypothesis is:

Research Hypothesis 6: For a given population, the population mean is greater than a particular value, S.

Here S is some meaningful value, depending on the given circumstances. Maybe the researcher wants to establish that the population mean height is greater than 72 inches. Or possibly the researcher is concerned that a four-choice, 100 item multiple choice test score is greater than a chance score of 25. Note that in these two examples (and in most hypotheses regarding a single population mean), the value of zero makes no sense. Suppose that a researcher wanted to establish that the population of freshman at a particular University had a mean College Board Score above the national average of 450:

Research Hypothesis 7: The population of freshmen at University X has a mean College Board Score greater than the national mean of 450.

Full Model: College Board Scores = $a_0U + E_{10}$

The research hypothesis implies that the sample mean is greater than 450, or $a_0 > 450$

Restriction: $a_0 = 450$

Restricted Model: (College Board Scores) = $450\ U + E_{11}$, or

(College Board Scores - 450) = E_{11}

(See bottom of Note 2 for test of significance and McNeil,1973 and McNeil, et al., 1975, p 315 for further details.)

The desired mean may be difficult to determine (i.e., it may require some thought or knowledge of the phenomenon under consideration), but no more difficult than justifying the default mean of 0. Indeed, using a mean of 0 in this example makes absolutely no sense at all, and that is why it doesn't appear in the literature.

One-way Analysis of Variance

The usual application of the multiple group F test (one-way ANOVA) is:

Research Hypothesis 8: There is at least one difference in the means on Y between the i populations.

Full Model: $Y = a_1G_1 + a_2G_2 + \ldots a_iG_i + E_{12}$

The research hypothesis implies that not all the sample means are equal, or that a_1 not equal a_2 not equal $\ldots a_i$ for at least one pair of means, or ($a_1 - a_i$ not equal 0; $a_2 - a_i$ not equal 0; $\ldots a_{i-1} - a_i$ not equal 0 for at least one pair of means.)

Restriction: $a_1 = a_2 = \ldots a_i$; or

$a_1 - a_i = 0;\ a_2 - a_i = 0;\ \ldots\ a_{i-1} - a_i = 0$

By replacing all the coefficients with a common coefficient, a_0, we arrive at the following restricted model:

Restricted Model: $Y = a_0G_1 + a_0G_2 + \ldots a_0G_i + E_{13}$

Restricted Model: $Y = a_0 * (G_1 + G_2 + \ldots G_i) + E_{13}$

Restricted Model: $Y = a_0U + E_{13}$

When the F test is significant then the restriction is rejected and the research hypothesis is accepted as tenable. But the research hypothesis just indicates that the i means are not all equal. Since most researchers are not satisfied with that information (confirming that the research hypothesis wasn't very interesting in the first place), most researchers turn to post-hoc comparisons to find out where the differences lie. These post-hoc comparisons are basically t-test comparisons and are thus like Research Hypothesis 1. (See Williams 1974). The suggestion here is to avoid asking a research hypothesis that you aren't interested in, and to go directly to non-zero research hypotheses that will yield satisfying information.

Interaction

Interaction is usually viewed only as a potentially contaminating factor when trying to explain main effects. That is, most researchers hope that there is no interaction so that they can proceed with interpreting main effects. But the interaction research hypothesis may be important in and of itself. Indeed, whenever an F has been computed for the interaction, the interaction research hypothesis has been tested. The usual interaction research hypothesis in a 2x2 design is as follows:

Research Hypothesis 9: For a given population, the difference on Y between Treatment 1 and Treatment 2 is not the same on Level 1 as on Level 2.

Full Model: $Y = a_1(T_1*L_1) + a_2(T_1*L_2) + a_3(T_2*L_1) + a_4(T_2*L_2) + E_{14}$

$m_{2)} = (R^2_F - R^2_R) / (11_F - 11_R)$

Where T_1 = 1 if in Treatment 1; 0 otherwise,

T_2 = 1 if in Treatment 2; 0 otherwise,

L_1 = 1 if in Level 1; 0 otherwise,

L_2 = 1 if in Level 2; 0 otherwise,

$(T_1 * L_1) = 1$ if in Treatment 1 <u>and</u> Level 1, etc.
The research hypothesis implies that the two differences are not the same, and that in the sample $(a_1 - a_3)$ not equal $(a_2 - a_4)$, or $[(a_1 - a_3) - (a_2 - a_4)$ not equal $0]$.
Restriction: $(a_1 - a_3) = (a_2 - a_4)$, or $[(a_1 - a_3) - (a_2 - a_4) = 0]$.
By placing the one restriction on the full model, one arrives at the following restricted model (See Note 3 and McNeil, et al., 1975):
Restricted Model: $Y = b_1T_1 + b_2T_2 + b_3L_1 + b_4L_2 + E_{15}$

Acceptance of the non-directional research hypothesis leads to a non-directional statement. All that can be concluded is that the differences are not the same. Hence we don't even know if the differences are greater at Level 1 or Level 2, let alone the magnitude of the difference of the differences. We have just conducted a non-directional test of interaction; a directional test of interaction is reflected in the following:
Research Hypothesis 10: For a given population, the difference on Y between Treatment 1 and Treatment 2 is greater at Level 1 than at Level 2.
Full Model: $Y = a_1(T_1*L_1) + a_2(T_1*L_2) + a_3(T_2*L_1) + a_4(T_2*L_2) + E_{16}$
The research hypothesis implies that the difference between T_1 and T_2 is greater at Level 1 than at Level 2, or in the sample $(a_1 - a_3)$ higher than $a_2 - a_4)$ or $[(a_1 - a_3) - (a_2 - a_4) > 0]$.
Restriction: $(a_1 - a_3) = (a_2 - a_4)$ or $[(a_1 - a_3) - (a_2 - a_4) = 0]$
Restricted Model: $Y = b_1T_1 + b_2T_2 + b_3L_1 + b_4L_2 + E_{17}$
A significant F for Research Hypothesis 10 provides more insight than would one for Research Hypothesis 9. We know that the differences are greater at Level 1, but again we do not know how much greater. If cost or theory dictate, say, a difference greater than six before a decision is made, the following Research Hypothesis would be appropriate:
Research Hypothesis 11: For a given population, the difference on Y between Treatment 1 and Treatment 2 is more than 6 units at Level 1 than at Level 2.
Full Model: $Y = a_1(T_1*L_1) + a_2(T_1*L_2) + a_3(T_2*L_1) + a_4(T_2*L_2) + E_{18}$
The research hypothesis implies that the difference between T_1 and T_2 is greater at Level 1 than at Level 2 by more than 6 units, or in the sample $(a_1 - a_3)$ higher than $(a_2 - a_4 + 6)$ or $[(a_1 - a_3) - (a_2 - a_4) > 6]$.
Restriction: $(a_1 - a_3) = (a_2 - a_4) + 6$; or
$(a_1 - a_3) - (a_2 - a_4) > 6$
Restricted Model: $(Y - 6) = b_1T_1 + b_2T_2 + b_3L_1 + b_4L_2 + E_{19}$

Research Hypotheses 9, 10, and 11 all test an interaction question, but in slightly different ways. In all three hypotheses, there are four linearly independent pieces of information in the full model. Forcing the one restriction on the full model results in three linearly independent pieces of information in the restricted model. Notice that the full models in Research Hypotheses 9, 10, and 11 are exactly the same. The number of restrictions is also the same; resulting in the same number of degrees of freedom. What is different, though, is the nature of the restriction and hence the restricted models are different. The three research hypotheses are all "correct" and equally "valid" - they just test three different hypotheses. Research Hypothesis 11, though, provides a more definitive conclusion, because as in the previous examples, a non-zero restriction was made.

Note 1. All the Research Hypotheses in this paper (except the one-way ANOVA) are directional Research Hypotheses. This follows the author's contention that a directional Research Hypothesis provides conclusive information whereas a non-directional Research Hypothesis provides no conclusive information. The Full and Restricted models are the same for the directional and non-directional hypotheses. The non-directional Research Hypothesis allows the researcher to conclude that a_1 does not equal 0, while the directional Research Hypothesis allows the researcher to conclude that $a_1 > 0$ (McNeil & Beggs, 1971). With reference to the non-zero restriction, of, say 6, the non-directional Research Hypothesis allows the conclusion that a_1 not equal to 6, while the directional Research Hypothesis allows the conclusion that $a_1 > 6$. The directional Research Hypothesis allows a more definitive conclusion using the same data and the same degrees of freedom.

Note 2. The general F test for testing two regression models is

$$F(m_1, m_2) = \frac{(R^2_F - R^2_R) / (li_F - li_R)}{(1 - R^2_F) / (N - li_F)}$$

Where: R^2_F = R^2 of the full model,
R^2_R = R^2 of the restricted model,
li_F = pieces of linearly independent information in the full model,
li_R = pieces of linearly independent information in the restricted model,
m_1 = $(li_F - li_R)$, and
m_2 = $(N - li_F)$.

This test cannot be used when either the restricted model has no predictors, when the criterion variable is different in the two models, or when the Unit vector is not in the restricted Model. In these cases, the F test must rely upon the sum of the squared scores in the error vector, E in both the full model (ESS_F) and the restricted model (ESS_R):

$$F = \frac{(ESS_R - ESS_F) / (li_F - li_R)}{(ESS_F) / (N - li_R)}$$

Note 3. The interaction examples all assumed equal N. The concepts still apply to the unequal N situation, although the restricted models will be different. (See Williams, 1972.)

SUMMARY

SIGNIFICANCE TEST	USUAL RESTRICTION	SUGGESTION
Pearson Correlation	zero	non-zero based on theory or cost
difference between two means	zero	non-zero based on theory or cost
difference between means (one-way f)	only zero	ignore omnibus F, go with planned comparisons based on theory or cost
interaction	almost always zero	non-zero based on theory or cost
single population mean	always non-zero	use more often

Bibliography

McNeil, K. A. Testing an hypothesis about a single population mean with multiple linear regression. Multiple Linear Regression Viewpoints, 1973, 4(1), 7-14.

McNeil, K. A., & Beggs, D. L. Directional hypotheses with the multiple linear regression approach. Paper presented at the meeting of the American Educational Research Association, New York, February 1971.

McNeil, K. A., Kelly, F. J., & McNeil, J. T. Testing Research Hypotheses Using Multiple Linear Regression. Cardondale: Southern Illinois University Press, 1975.

Williams, J. D. Two way fixed effects analysis of variance with disproportionate cell frequencies. Multivariate Behavioral Research, 1972, 7, 67-83.

Williams, J. D. Regression Analysis In Educational Research. New York: MSS Information Corporation, 1974.

C:CASENON