

Chapter 4

Test of Hypotheses

C. R. Henderson

1984 - Guelph

Much of the statistical literature for many years dealt primarily with tests of hypotheses (or tests of significance). More recently increased emphasis has been placed, properly I think, on estimation and prediction. Nevertheless, many research workers and certainly most editors of scientific journals insist on tests of significance. Most tests involving linear models can be stated as follows. We wish to test the null hypothesis,

$$\mathbf{H}'_0\boldsymbol{\beta} = \mathbf{c}_0,$$

against some alternative hypothesis, most commonly the alternative that $\boldsymbol{\beta}$ can have any value in the parameter space. Another possibility is the general alternative hypothesis,

$$\mathbf{H}'_a\boldsymbol{\beta} = \mathbf{c}_a.$$

In both of these hypotheses there may be elements of $\boldsymbol{\beta}$ that are not determined by \mathbf{H} . These elements are assumed to have any values in the parameter space. \mathbf{H}'_0 and \mathbf{H}'_a are assumed to have full row rank with m and a rows respectively. Also $r \geq m > a$. Under the unrestricted hypothesis $a = 0$.

Two important restrictions are required logically for \mathbf{H}_0 and \mathbf{H}_a . First, both $\mathbf{H}'_0\boldsymbol{\beta}$ and $\mathbf{H}'_a\boldsymbol{\beta}$ must be estimable. It hardly seems logical that we could test hypotheses about functions of $\boldsymbol{\beta}$ unless we can estimate these functions. Second, the null hypothesis must be contained in the alternative hypothesis. That is, if the null is true, the alternative must be true. For this to be so we require that \mathbf{H}'_a can be written as $\mathbf{M}\mathbf{H}'_0$ and \mathbf{c}_a as $\mathbf{M}\mathbf{c}_0$ for some \mathbf{M} .

1 Equivalent Hypotheses

It should be recognized that there are an infinity of hypotheses that are equivalent to $\mathbf{H}'_0\boldsymbol{\beta} = \mathbf{c}$. Let \mathbf{P} be an $m \times m$, non-singular matrix. Then $\mathbf{P}\mathbf{H}'_0\boldsymbol{\beta} = \mathbf{P}\mathbf{c}_0$ is equivalent to

$\mathbf{H}'_0\boldsymbol{\beta} = \mathbf{c}$. For example, consider a fixed model

$$y_{ij} = \mu + t_i + e_{ij}, \quad i = 1, 2, 3.$$

A null hypothesis often tested is

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix} \mathbf{t} = \mathbf{0}.$$

An equivalent hypothesis is

$$\begin{pmatrix} 2/3 & -1/3 & -1/3 \\ -1/3 & 2/3 & -1/3 \end{pmatrix} \mathbf{t} = \mathbf{0}.$$

To convert the first to the second pre-multiply

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix} \text{ by } \begin{pmatrix} 2/3 & -1/3 \\ -1/3 & 2/3 \end{pmatrix}.$$

As an example of use of \mathbf{H}'_a consider a type of analysis sometimes recommended for a two way fixed model without interaction. Let the model be $y_{ijk} = \mu + a_i + b_j + e_{ijk}$, where $i = 1, 2, 3$ and $j = 1, 2, 3, 4$. The lines of the ANOVA table could be as follows.

Sum of Squares
Rows ignoring columns (column differences regarded as non-existent),
Columns with rows accounted for,
Residual.

The sum of these 3 sums of squares is equal to $(\mathbf{y}'\mathbf{y} - \text{correction factor})$. The first sum of squares is represented as testing the null hypothesis:

$$\begin{pmatrix} 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix} \boldsymbol{\beta} = \mathbf{0}.$$

and the alternative hypothesis:

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix} \boldsymbol{\beta} = \mathbf{0}.$$

The second sum of squares represents testing the null hypothesis:

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix} \boldsymbol{\beta} = \mathbf{0}.$$

and the alternative hypothesis: entire parameter space.

2 Test Criteria

2.1 Differences between residuals

Now it is assumed for purposes of testing hypotheses that \mathbf{y} has a multivariate normal distribution. Then it can be proved by the likelihood ratio method of testing hypotheses, Neyman and Pearson (1933), that under the null hypothesis the following quantity is distributed as χ^2 .

$$(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}_0)' \mathbf{V}^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}_0) - (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}_a)' \mathbf{V}^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}_a). \quad (1)$$

$\boldsymbol{\beta}_0$ is a solution to GLS equations subject to the restriction $\mathbf{H}'_0 \boldsymbol{\beta}_0 = \mathbf{c}_0$. $\boldsymbol{\beta}_0$ can be found by solving

$$\begin{pmatrix} \mathbf{X}'\mathbf{V}^{-1}\mathbf{X} & \mathbf{H}'_0 \\ \mathbf{H}'_0 & \mathbf{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_0 \\ \boldsymbol{\theta}_0 \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{V}^{-1}\mathbf{y} \\ \mathbf{c}_0 \end{pmatrix}$$

or by solving the comparable mixed model equations

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} & \mathbf{H}'_0 \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} & \mathbf{0} \\ \mathbf{H}'_0 & \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_0 \\ \mathbf{u}_0 \\ \boldsymbol{\theta}_0 \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{c}_0 \end{pmatrix}.$$

$\boldsymbol{\beta}_a$ is a solution to GLS or mixed model equations with restrictions, $\mathbf{H}'_a \boldsymbol{\beta}_a = \mathbf{c}_a$ rather than $\mathbf{H}'_0 \boldsymbol{\beta}_0 = \mathbf{c}_0$.

In case the alternative hypothesis is unrestricted ($\boldsymbol{\beta}$ can have any values), that is, $\boldsymbol{\beta}_a$ is a solution to the unrestricted GLS or mixed model equations. Under the null hypothesis (1) is distributed as χ^2 with $(m - a)$ degrees of freedom, m being the number of rows (independent) in \mathbf{H}'_0 , and a being the number of rows (independent) in \mathbf{H}'_a . If the alternative hypothesis is unrestricted, $a = 0$. Having computed (1) this value is compared with values of χ^2_{m-a} for the chosen level of significance.

Let us illustrate with a model

$$\begin{aligned} y &= \mu + t_i + e_{ij} \\ \mu, t_i \text{ fixed, } i &= 1, 2, 3 \\ \mathbf{R} &= \text{Var}(\mathbf{e}) = 5\mathbf{I}. \end{aligned}$$

Suppose that the number of observations on the levels of t_i are 4, 3, 2, and the treatment totals are 25, 15, 9 with individual observations, (6, 7, 8, 4, 4, 5, 6, 5, 4). We wish to test that the levels of t_i are equal, which can be expressed as

$$\begin{pmatrix} 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix} (\mu \ t_1 \ t_2 \ t_3)' = (0 \ 0)'$$

We use as the alternative hypothesis the unrestricted hypothesis. The GLS equations under the restriction are

$$.2 \begin{pmatrix} 9 & 4 & 3 & 2 & 0 & 0 \\ 4 & 4 & 0 & 0 & 1 & 0 \\ 3 & 0 & 3 & 0 & 0 & 1 \\ 2 & 0 & 0 & 2 & -1 & -1 \\ 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \theta_0 \end{pmatrix} = .2 \begin{pmatrix} 49 \\ 25 \\ 15 \\ 9 \\ 0 \\ 0 \end{pmatrix}.$$

A solution is

$$\beta'_o = (49 \ 0 \ 0 \ 0)/9, \theta'_o = (29 \ -12)/9.$$

The GLS equations with no restrictions are

$$.2 \begin{pmatrix} 9 & 4 & 3 & 2 \\ 4 & 4 & 0 & 0 \\ 3 & 0 & 3 & 0 \\ 2 & 0 & 0 & 2 \end{pmatrix} (\beta_a) = .2 \begin{pmatrix} 49 \\ 25 \\ 15 \\ 9 \end{pmatrix}.$$

A solution is $\beta_a = (0 \ 25 \ 20 \ 18)/4$.

$$\begin{aligned} (\mathbf{y} - \mathbf{X}\beta_o)' &= (5 \ 14 \ 23 \ -13 \ -13 \ -4 \ 5 \ -4 \ -13)/9. \\ (\mathbf{y} - \mathbf{X}\beta_o)'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\beta_o) &= 146/45. \\ (\mathbf{y} - \mathbf{X}\beta_a)' &= [-1, 3, 7, -9, -4, 0, 4, 2, -2]/4. \\ (\mathbf{y} - \mathbf{X}\beta_a)'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\beta_a) &= 9/4. \end{aligned}$$

The difference is $\frac{146}{45} - \frac{9}{4} = \frac{179}{180}$.

2.2 Differences between reductions

Two easier methods of computation that lead to the same result will now be presented. The first, described in Searle (1971b), is

$$\beta'_a \mathbf{X}' \mathbf{V}^{-1} \mathbf{y} + \theta'_a \mathbf{c}_a - \beta'_o \mathbf{X}' \mathbf{V}^{-1} \mathbf{y} - \theta'_o \mathbf{c}_o. \quad (2)$$

The first 2 terms are called reduction in sums of squares under the alternative hypothesis. The last two terms are the negative of the reduction in sum of squares under the null hypothesis. In our example

$$\begin{aligned} \beta'_a \mathbf{X}' \mathbf{V}^{-1} \mathbf{y} + \theta'_a \mathbf{c}_a &= 1087/20. \\ \beta'_o \mathbf{X}' \mathbf{V}^{-1} \mathbf{y} + \theta'_o \mathbf{c}_o &= 2401/45. \\ \frac{1087}{20} - \frac{2401}{45} &= \frac{179}{180} \text{ as before.} \end{aligned}$$

If the mixed model equations are used, (2) can be computed as

$$\boldsymbol{\beta}'_a \mathbf{X}' \mathbf{R}^{-1} \mathbf{y} + \mathbf{u}'_a \mathbf{Z}' \mathbf{R}^{-1} \mathbf{y} + \boldsymbol{\theta}'_a \mathbf{c}_a - \boldsymbol{\beta}'_o \mathbf{X}' \mathbf{R}^{-1} \mathbf{y} - \mathbf{u}'_o \mathbf{Z}' \mathbf{R}^{-1} \mathbf{y} - \boldsymbol{\theta}'_o \mathbf{c}_o. \quad (3)$$

2.3 Method based on variances of linear functions

A second easier method is

$$\begin{aligned} & (\mathbf{H}'_o \boldsymbol{\beta}^o - \mathbf{c}_o)' [\mathbf{H}'_o (\mathbf{X}' \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{H}_o]^{-1} (\mathbf{H}'_o \boldsymbol{\beta}^o - \mathbf{c}_o) \\ & - (\mathbf{H}'_a \boldsymbol{\beta}^o - \mathbf{c}_a)' [\mathbf{H}'_a (\mathbf{X}' \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{H}_a]^{-1} (\mathbf{H}'_a \boldsymbol{\beta}^o - \mathbf{c}_a). \end{aligned} \quad (4)$$

If $\mathbf{H}'_a \boldsymbol{\beta}$ is unrestricted the second term of (4) is set to 0. Remember that $\boldsymbol{\beta}^o$ is a solution in the unrestricted GLS equations. In place of $(\mathbf{X}' \mathbf{V}^{-1} \mathbf{X})^{-1}$ one can substitute the corresponding submatrix of a g-inverse of the mixed model coefficient matrix.

This is a convenient point to prove that an equivalent hypothesis, $\mathbf{P}(\mathbf{H}' \boldsymbol{\beta} - \mathbf{c}) = \mathbf{0}$ gives the same result as $\mathbf{H}' \boldsymbol{\beta} - \mathbf{c}$, remembering that \mathbf{P} is non-singular. The quantity corresponding to (4) for $\mathbf{P}(\mathbf{H}' \boldsymbol{\beta} - \mathbf{c})$ is

$$\begin{aligned} & (\mathbf{H}' \boldsymbol{\beta}^o - \mathbf{c})' \mathbf{P}' [\mathbf{P} \mathbf{H}' (\mathbf{X}' \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{H} \mathbf{P}']^{-1} \mathbf{P} (\mathbf{H}' \boldsymbol{\beta} - \mathbf{c}) \\ & = (\mathbf{H}' \boldsymbol{\beta}^o - \mathbf{c})' \mathbf{P}' (\mathbf{P}')^{-1} [\mathbf{H}' (\mathbf{X}' \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{H}]^{-1} \mathbf{P}^{-1} \mathbf{P} (\mathbf{H}' \boldsymbol{\beta}^o - \mathbf{c}) \\ & = (\mathbf{H}' \boldsymbol{\beta}^o - \mathbf{c})' [\mathbf{H}' (\mathbf{X}' \mathbf{V}^{-1} \mathbf{X}) \mathbf{H}]^{-1} (\mathbf{H}' \boldsymbol{\beta}^o - \mathbf{c}), \end{aligned}$$

which proves the equality of the two equivalent hypotheses.

Let us illustrate (3) with our example

$$\begin{pmatrix} 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix} \boldsymbol{\beta}^o = \begin{pmatrix} 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix} (0 \ 25 \ 20 \ 18)' / 4 = \begin{pmatrix} 7 \\ 2 \end{pmatrix} / 4.$$

A g-inverse of $\mathbf{X}' \mathbf{V}^{-1} \mathbf{X}$ is

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 15 & 0 & 0 \\ 0 & 0 & 20 & 0 \\ 0 & 0 & 0 & 30 \end{pmatrix} / 12.$$

$$\mathbf{H}'_0 (\mathbf{X}' \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{H}_0 = \begin{pmatrix} 45 & 30 \\ 30 & 50 \end{pmatrix} / 12.$$

The inverse of this is

$$\begin{pmatrix} 20 & -12 \\ -12 & 18 \end{pmatrix} /45.$$

Then

$$\frac{1}{4}(7 \ 2) \begin{pmatrix} 20 & -12 \\ -12 & 18 \end{pmatrix} \frac{1}{45} \begin{pmatrix} 7 \\ 2 \end{pmatrix} \frac{1}{4} = \frac{179}{180} \text{ as before.}$$

The d.f. for χ^2 are 2 because \mathbf{H}'_0 has 2 rows and the alternative hypothesis is unrestricted.

2.4 Comparison of reductions under reduced models

Another commonly used method is to compare reductions in sums of squares resulting from deletions of different subvectors of $\boldsymbol{\beta}$ from the reduction. The difficulty with this method is the determination of what hypothesis is tested by the difference between a pair of reductions. It is not true in general, as sometimes thought, that $Red(\boldsymbol{\beta}) - Red(\boldsymbol{\beta}_1)$ tests the hypothesis that $\boldsymbol{\beta}_2 = \mathbf{0}$, where $\boldsymbol{\beta}' = (\boldsymbol{\beta}'_1 \ \boldsymbol{\beta}'_2)$. In most designs, $\boldsymbol{\beta}_2$ is not estimable. We need to determine what $\mathbf{H}'\boldsymbol{\beta}$ imposed on a solution will give the same reduction in sum of squares as does $Red(\boldsymbol{\beta}_1)$.

In the latter case we solve

$$(\mathbf{X}'_1 \mathbf{V}^{-1} \mathbf{X}_1) \boldsymbol{\beta}_1^o = \mathbf{X}'_1 \mathbf{V}^{-1} \mathbf{y}$$

and then

$$\text{Reduction} = (\boldsymbol{\beta}_1^o)' \mathbf{X}'_1 \mathbf{V}^{-1} \mathbf{y}. \quad (5)$$

Consider a hypothesis, $\mathbf{H}'\boldsymbol{\beta}_2 = \mathbf{0}$. We could solve

$$\begin{pmatrix} \mathbf{X}'_1 \mathbf{V}^{-1} \mathbf{X}_1 & \mathbf{X}'_1 \mathbf{V}^{-1} \mathbf{X}_2 & \mathbf{0} \\ \mathbf{X}'_2 \mathbf{V}^{-1} \mathbf{X}_1 & \mathbf{X}'_2 \mathbf{V}^{-1} \mathbf{X}_2 & \mathbf{H} \\ \mathbf{0} & \mathbf{H}' & \mathbf{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_1^o \\ \boldsymbol{\beta}_2^o \\ \boldsymbol{\theta} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'_1 \mathbf{V}^{-1} \mathbf{y} \\ \mathbf{X}'_2 \mathbf{V}^{-1} \mathbf{y} \\ \mathbf{0} \end{pmatrix}. \quad (6)$$

Then

$$\text{Reduction} = (\boldsymbol{\beta}_1^o)' \mathbf{X}'_1 \mathbf{V}^{-1} \mathbf{y} + (\boldsymbol{\beta}_2^o)' \mathbf{X}'_2 \mathbf{V}^{-1} \mathbf{y}. \quad (7)$$

Clearly (7) is equal to (5) if a solution to (6) is $\boldsymbol{\beta}_2^o = \mathbf{0}$, for then

$$\boldsymbol{\beta}_1^o = (\mathbf{X}'_1 \mathbf{V}^{-1} \mathbf{X}_1)^{-1} \mathbf{X}'_1 \mathbf{V}^{-1} \mathbf{y}.$$

Consequently in order to determine what hypothesis is implied when $\boldsymbol{\beta}_2$ is deleted from the model, we need to find some $\mathbf{H}'\boldsymbol{\beta}_2 = \mathbf{0}$ such that a solution to (6) is $\boldsymbol{\beta}_2^o = \mathbf{0}$.

We illustrate with a two way fixed model with interaction. The numbers of observations per subclass are

$$\begin{pmatrix} 3 & 2 & 1 \\ 1 & 2 & 5 \end{pmatrix}.$$

The subclass totals are

$$\begin{pmatrix} 6 & 2 & 2 \\ 3 & 5 & 9 \end{pmatrix}.$$

An analysis sometimes suggested is

$$Red(\mu, r, c) - Red(\mu, c) \text{ to test rows.}$$

$$Red(full\ model) - Red(\mu, r, c) \text{ to test interaction.}$$

The least squares equations are

$$\begin{pmatrix} 14 & 6 & 8 & 4 & 4 & 6 & 3 & 2 & 1 & 1 & 2 & 5 \\ & 6 & 0 & 3 & 2 & 1 & 3 & 2 & 1 & 0 & 0 & 0 \\ & & 8 & 1 & 2 & 5 & 0 & 0 & 0 & 1 & 2 & 5 \\ & & & 4 & 0 & 0 & 3 & 0 & 0 & 1 & 0 & 0 \\ & & & & 4 & 0 & 0 & 2 & 0 & 0 & 2 & 0 \\ & & & & & 6 & 0 & 0 & 1 & 0 & 0 & 5 \\ & & & & & & 3 & 0 & 0 & 0 & 0 & 0 \\ & & & & & & & 2 & 0 & 0 & 0 & 0 \\ & & & & & & & & 1 & 0 & 0 & 0 \\ & & & & & & & & & 1 & 0 & 0 \\ & & & & & & & & & & 2 & 0 \\ & & & & & & & & & & & 5 \end{pmatrix} \beta^o = \begin{pmatrix} 27 \\ 10 \\ 17 \\ 9 \\ 7 \\ 11 \\ 6 \\ 2 \\ 2 \\ 3 \\ 5 \\ 9 \end{pmatrix}$$

A solution to these equations is

$$[0, 0, 0, 0, 0, 0, 2, 1, 2, 3, 2.5, 1.8],$$

which gives a reduction of 55.7, the full model reduction. A solution when interaction terms are deleted is

$$[1.9677, -.8065, 0, .8871, .1855, 0]$$

giving a reduction of 54.3468. This corresponds to an hypothesis,

$$\begin{pmatrix} 1 & 0 & -1 & -1 & 0 & 1 \\ 0 & 1 & -1 & 0 & -1 & 1 \end{pmatrix} \mathbf{rc} = \mathbf{0}.$$

When this is included as a Lagrange multiplier as in (6), a solution is

$$[1.9677, -.8065, 0, .8871, .1855, 0, 0, 0, 0, 0, 0, 0, -.1452, -.6935].$$

Note that $(\mathbf{rc})^o = \mathbf{0}$, proving that dropping \mathbf{rc} corresponds to the hypothesis stated above. The reduction again is 54.3468.

When \mathbf{r} and \mathbf{rc} are dropped from the equations, a solution is

$$[0, 2.25, 1.75, 1.8333]$$

giving a reduction of 52.6667. This corresponds to an hypothesis

$$\begin{pmatrix} 3 & -3 & 1 & 1 & 1 & -1 & -1 & -1 \\ 0 & 0 & 1 & 0 & -1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 & 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{r} \\ \mathbf{rc} \end{pmatrix} = \mathbf{0}.$$

When this is added as a Lagrange multiplier, a solution is

$$[2.25, 0, 0, 0, -0.5, -0.4167, 0, 0, 0, 0, 0, 0, -0.6944, -0.05556, -0.8056].$$

Note that \mathbf{r}^o and \mathbf{rc}^o are null, verifying the hypothesis. The reduction again is 52.6667. Then the tests are as follows:

Rows assuming \mathbf{rc} non-existent = 54.3468 - 52.6667.

Interaction = 55.7 - 54.3468.