HYPOTHESES ASSOCIATED WITH SUMS OF SQUARES CALCULATED BY STATISTICAL COMPUTER PACKAGES

BU-774-M

by

S. R. Searle

Biometrics Unit, Cornell University, Ithaca, New York

June, 1982

Abstract

Users of analysis of variance computer packages often face the question: when using a sum of squares from a computer package as the numerator of an F-statistic, exactly what is the hypothesis being tested? This question implies logic reverse to that customarily used for hypothesis testing, but it can arise whenever computer output fails to inform users exactly what hypothesis is associated with each computed sum of squares.

A general answer to the question is provided for a wide and useful class of sums of squares, and examples are given.

1. INTRODUCTION

The logic of hypothesis testing is to set up a hypothesis of interest and construct a test for it. With linear models this usually involves formulating a hypothesis about linear functions of parameters of the model and calculating an F-statistic which is the ratio of two independently χ^2-distributed mean squares. The denominator of that ratio is an estimate of the residual error variance and the numerator is determined by the form of the hypothesis.

Paper No. BU-774-M in the Biometrics Unit.

Unfortunately, in the presence of today's sophisticated computing packages, this logic is often (and all too easily) not followed. Data are often subjected to computer processing without sufficient thought for precisely what analyses are wanted. Instead of computer power being used to calculate mean squares and F-statistics whose need has been specifically planned, data are fed to a package and the ofttimes voluminous output that it produces often leads to the question "What does all this mean?". Answering this question can be difficult because computed mean squares often have labels that are not sufficiently explicit for users to be certain of knowing what hypotheses are being tested when those mean squares are used as numerators of F-statistics. A user can therefore be driven to the procedure of starting with a sum of squares and asking "What use is it?". Although this is not logically correct (insofar as hypothesis testing is concerned), it is nevertheless a question that occurs very naturally to anyone who has processed data through an analysis of variance computer package and is then faced with mean squares which, by the labels accompanying them, may not appear unequivocal in their meaning. Indeed they may appear to mean one thing but perhaps in reality mean something else.

One therefore needs to be able to derive hypotheses from given sums of squares in order to evaluate the applicability of computer-generated F-statistics to one's data. On occasion, despite the label attached to a sum of squares, the conclusion will be that by its very nature the associated hypothesis is not worth testing. For example, consider a row-by-column layout having a linear model

$$E(y_{ij}) = \mu + \alpha_i + \beta_j \;, \tag{1}$$

where y_{ij} is the observation in the i'th row and j'th column, μ represents a general mean and α_i and β_j represent effects, respectively, for the i'th row and j'th column. Then the sum of squares $R(\alpha|\mu) = \Sigma_i n_{i.}(\bar{y}_{i.} - \bar{y}_{..})^2$ can, for balanced data (with one observation per cell, wherein $n_{i.} = b$, the number of columns), be used to test the hypothesis $H: \alpha_i$ all equal. And for unbalanced data (wherein some cells are empty and some contain one observation and so $n_{i.}$ is the number of observations in the i'th row), the use of $R(\underset{\sim}{\alpha}|\mu)$ in the model $E(y_{ij}) = \mu + \alpha_i$ also tests $H: \alpha_i$ all equal. But in the model (1) it tests $H: \alpha_i + \Sigma_j n_{ij}\beta_j/n_{i.}$ all equal (Searle, 1971, p. 283), where n_{ij} is the number of observations (0 or 1) in the i'th row and j'th column. In brief

$$R(\underset{\sim}{\alpha}|\mu)_{\mu,\alpha} \quad \text{tests } H: \alpha_i \text{ all equal}$$

and

$$R(\underset{\sim}{\alpha}|\mu)_{\mu,\alpha,\beta} \text{ tests } H: \alpha_i + \Sigma n_{ij}\beta_j/n_{i.} \text{ all equal,} \tag{2}$$

where the subscripts on $R(\cdot|\cdot)$ are labels indicating the nature of the model, e.g., (2) is for the model (1). In point of fact this kind of labeling is seldom done; most computer output relies on context and brief text labeling for clarity, a reliance that may often be unwarranted by a package user's knowledge. Thus, without knowing the hypothesis (2), one could be led by the simplicity of a text label such as "rows" given to a computed $R(\underset{\sim}{\alpha}|\mu)$ into believing that it is of use when in fact it is not.

Deriving hypotheses such as (2) is not tantamount to promote their use; far from it. The sole purpose of the derivation is to have explicit hypotheses available as a means of understanding what certain sums of squares represent. If a hypothesis is applicable to one's data it can be used, but otherwise it can be ignored.

2. GENERAL FORMULATION

We consider our data to be a vector $\underset{\sim}{y}$ of order N, normally distributed with mean $\underset{\sim}{X}\underset{\sim}{\beta}$ and variance-covariance matrix $\sigma^2\underset{\sim}{I}$. The parameters of the model are elements of $\underset{\sim}{\beta}$, and $\underset{\sim}{X}$ is a known matrix that is often an incidence matrix of 0's and 1's, but may include columns of observed covariates. For a formulated hypothesis of interest

$$H : K'\beta = m \tag{3}$$

where K' has full rank and every element of $K'\beta$ is estimable, it is well known (e.g., Searle, 1971, p. 192) that a test-statistic for the hypothesis is

$$F_H = (K'\beta^0 - m)'(K'GK)^{-1}(K'\beta^0 - m)/\hat{\sigma}^2 r_K \tag{4}$$

where $X'XGX'X = X'X$, $\beta^0 = GX'y$, the rank of K is r_K and where

$$\hat{\sigma}^2 = y'(I - XGX')y/(N - r_X) \; . \tag{5}$$

This is how the correct logic of hypothesis testing is used: formulate one's hypothesis as (3) and calculate (4).

In contrast to the preceding development, the question we address here is the following: given a sum of squares $y'Ay$ for A symmetric of rank r_A, what is the hypothesis tested by

$$F_A = y'Ay/\hat{\sigma}^2 r_A \;\; ? \tag{6}$$

The question assumes that distributional requirements for F_A to have a non-central F-distribution are met: namely that A is symmetric and idempotent, that $y'Ay$ and $\hat{\sigma}^2$ are independent, and that $y'Ay/\hat{\sigma}^2$ has a non-central χ^2-density. Then F_A has a non-centrality parameter $\beta'X'AX\beta/2\sigma^2$ which is zero (using $A = A' = A^2$) if and only if $AX\beta = 0$. Therefore the hypothesis tested by F_A can be taken as

$$H : AX\beta = 0 \; . \tag{7}$$

3. FITTING CONSTANTS

A large class of sums of squares calculated by computer packages are those which arise from the method of fitting constants. Expressed as $y'Ay$ they satisfy the conditions of the preceding paragraph, and so the corresponding hypothesis is (7). Because, in general, $\underset{\sim}{A}$ can be expressed explicitly in terms of the model, a more useful expression of this hypothesis can be developed, as is now indicated.

Partition the model $E(\underset{\sim}{y}) = \underset{\sim}{X}\underset{\sim}{\beta}$ as

$$E(\underset{\sim}{y}) = \underset{\sim}{X}_1\underset{\sim}{\beta}_1 + \underset{\sim}{X}_2\underset{\sim}{\beta}_2 + \underset{\sim}{X}_3\underset{\sim}{\beta}_3 \tag{8}$$

and consider sums of squares such as that due to $\underset{\sim}{\beta}_2$ adjusted for $\underset{\sim}{\beta}_1$, namely $R(\underset{\sim}{\beta}_2|\underset{\sim}{\beta}_1) = R(\underset{\sim}{\beta}_1,\underset{\sim}{\beta}_2) - R(\underset{\sim}{\beta}_1)$, the sum of squares due to fitting $E(\underset{\sim}{y}) = \underset{\sim}{X}_1\underset{\sim}{\beta}_1 + \underset{\sim}{X}_2\underset{\sim}{\beta}_2$ minus that due to fitting $E(\underset{\sim}{y}) = \underset{\sim}{X}_1\underset{\sim}{\beta}_1$. Similar to (3.7) of Searle _et al._ (1981), this can be expressed as

$$R(\underset{\sim}{\beta}_2|\underset{\sim}{\beta}_1) = \underset{\sim}{y}'\underset{\sim}{A}\underset{\sim}{y} \tag{9}$$

for

$$\underset{\sim}{A} = \underset{\sim}{M}_1\underset{\sim}{X}_2(\underset{\sim}{X}_2'\underset{\sim}{M}_1\underset{\sim}{X}_2)^-\underset{\sim}{X}_2'\underset{\sim}{M}_1 \tag{10}$$

with

$$\underset{\sim}{M}_1 = \underset{\sim}{I} - \underset{\sim}{X}_1(\underset{\sim}{X}_1'\underset{\sim}{X}_1)^-\underset{\sim}{X}_1' = \underset{\sim}{I} - \underset{\sim}{X}_1\underset{\sim}{X}_1^+ \tag{11}$$

where $\underset{\sim}{X}_1'\underset{\sim}{X}_1(\underset{\sim}{X}_1'\underset{\sim}{X}_1)^-\underset{\sim}{X}_1'\underset{\sim}{X}_1 = \underset{\sim}{X}_1'\underset{\sim}{X}_1$ and $\underset{\sim}{X}_1^+$ is the Moore-Penrose inverse of $\underset{\sim}{X}_1$.

It is not difficult to show that $\underset{\sim}{A}$ of (10) satisfies the conditions of Section 2 with $r_{\underset{\sim}{A}} = r_{12} - r_1$ where r_{12} is the rank of $[\underset{\sim}{X}_1 \;\; \underset{\sim}{X}_2]$ and r_1 is that of $\underset{\sim}{X}_1$. Then on substituting (10) into (7) and using $\underset{\sim}{X} = [\underset{\sim}{X}_1 \;\; \underset{\sim}{X}_2 \;\; \underset{\sim}{X}_3]$ implicit in (8) we find that the hypothesis tested by

$$F = R(\underset{\sim}{\beta}_2|\underset{\sim}{\beta}_1)/\hat{\sigma}^2(r_{12} - r_1) \tag{12}$$

is

$$H : \underset{\sim}{M}_1\underset{\sim}{X}_2\underset{\sim}{\beta}_2 + \underset{\sim}{M}_1\underset{\sim}{X}_2(\underset{\sim}{X}_2'\underset{\sim}{M}_1\underset{\sim}{X}_2)^-\underset{\sim}{M}_1\underset{\sim}{X}_2\underset{\sim}{X}_3\underset{\sim}{\beta}_3 = \underset{\sim}{0}\,. \tag{13}$$

Thus a quite general form of hypothesis tested by using $R(\underset{\sim}{\beta}_2|\underset{\sim}{\beta}_1)$ as the numerator of an F-statistic in the model $E(\underset{\sim}{y}) = \underset{\sim}{X}_1\underset{\sim}{\beta}_1 + \underset{\sim}{X}_2\underset{\sim}{\beta}_2 + \underset{\sim}{X}_3\underset{\sim}{\beta}_3$ is (13). A distinctive feature of it is that it is for a model that includes not only $\underset{\sim}{\beta}_1$ and $\underset{\sim}{\beta}_2$ which are involved in the sum of squares, but also $\underset{\sim}{\beta}_3$ which is not. This is its generality, and it produces as special cases, all other combinations of model and hypothesis that are possible in this context, just five of them. All six are shown in the table, along with the values to be given to the general vector $[\beta_1' \;\; \beta_2' \;\; \beta_3']$ to yield the special cases. For example, using $[0 \;\; \beta_1' \;\; 0]$ reduces the model (8) to $E(y) = X_1\beta_1$, M_1 to I, M_1X_2 to X_1, $R(\beta_2|\beta_1)$ to $R(\beta_1)$ and the hypothesis to $H: X_1\beta_1 = 0$. This is line 1 of the table. Lines 2-4 and 6 are derived similarly.

[SHOW TABLE]

4. RANK CONSIDERATIONS

A noticeable feature of the hypothesis $H: AX\beta = 0$ and its special cases is that if written as $H: K'\beta = 0$, none of them has K' of full row rank as demanded of (3) in order for (4) to hold. Nevertheless, (4) does hold for $H: AX\beta = 0$, for the following reasons. The symmetry and idempotency of A mean that $A = LL'$ for some L such that $L'L = I$. Therefore, because $H: AX\beta = 0$ is equivalent to $H: L'X\beta = 0$, the numerator of F_H in (4) for the latter is $(L'X\beta^0)'(L'XGX'L)^{-1}L'X\beta^0$, which reduces very easily to $y'Ay$, as it should.

A practical consequence of the hypothesis formulated as $H: AX\beta = 0$ that is important in using (13) and its special cases is that all these hypotheses are statements about N linear combinations of elements of β_2 and β_3. But only r_A of them are linearly independent. But, by using Marsaglia and Styan (1974), we find that $r_A = r_{12} - r_1$. This number of linearly independent linear combinations of elements of β_2 and β_3 can therefore always be used as restatement of the hypothesis (13). This is illustrated in what follows.

5. APPLICATION

Application of (13) to any particular situation requires writing one's model partitioned as in (8), so that for a sum of squares describable as $R(\beta_2|\beta_1)$, one has β_2 being the effects of interest, so to speak, β_1 being the effects being adjusted for, and β_3 being all other effects in the model. After determining the respective X-matrices, X_1, X_2 and X_3, that go with the β's, they are used in (11) and (13) to derive the hypothesis. In that form it must then be reduced to full rank form, as discussed in the preceding section. We illustrate with two examples. In the first, (13) yields a result that is well known; in the second it gives one that is not so well known.

Example 1: Consider a completely randomized design of a groups, n observations per group, and a single covariate. We write the model as

$$E(y_{ij}) = \mu + \alpha_i + bx_{ij} \tag{14}$$

where y_{ij} is the response variable and x_{ij} the covariate on the j'th observation in the i'th group, for $i = 1, \cdots, a$ and $j = 1, \cdots, n$, μ is a general mean, α_i is the group effect and b is the coefficient of the covariate.

Notation: $\mathbf{1}_a$ represents an a $\times$ 1 vector of ones, J_a is an a $\times$ a matrix of ones, $\bar{J}_a = J_a/a$ and $D\{T\}$ is a block diagonal matrix of a matrices T. $N = an$.

With this notation (14) can be written as

$$E(y) = \mu\mathbf{1}_N + (D\{\mathbf{1}_n\})\alpha + bx ,$$

where $\alpha = \{\alpha_i\}$ for $i = 1, \cdots, a$ and y and x are, respectively, the vectors of elements y_{ij} and x_{ij}, in lexicon order.

We use (13) to derive the hypothesis corresponding to $R(\alpha|\mu)$. Thus $R(\beta_2|\beta_1)$ of the general notation is $R(\alpha|\mu)$ here. Therefore

$$\beta_2 = \alpha, \qquad \beta_1 = \mu \quad \text{and} \quad \beta_3 = b$$

and, correspondingly,

$$\mathbf{X}_2 = \mathbf{D}\{\mathbf{1}_n\}, \qquad \mathbf{X}_1 = \mathbf{1}_N \quad \text{and} \quad \mathbf{X}_3 = \mathbf{x} .$$

Therefore

$$\mathbf{M}_1 = \mathbf{I}_N - \mathbf{1}_N(\mathbf{1}_N'\mathbf{1}_N)^{-1}\mathbf{1}_N' = \mathbf{I} - \bar{\mathbf{J}}_N \quad \text{with} \quad \mathbf{M}_1\mathbf{X}_2 = \mathbf{D}\{\mathbf{1}_n\} - \frac{1}{a}\mathbf{J}_{N\times a} ,$$

and

$$\mathbf{X}_2'\mathbf{M}_1\mathbf{X}_2 = n(\mathbf{I}_a - \bar{\mathbf{J}}_a) \quad \text{and} \quad (\mathbf{X}_2'\mathbf{M}_1\mathbf{X}_2)^- = (1/n)(\mathbf{I}_a + \lambda\mathbf{J}_a) \quad \text{for any } \lambda .$$

It is easily shown that

$$\mathbf{M}_1\mathbf{X}_2(\mathbf{X}_2'\mathbf{M}_1\mathbf{X}_2)^-\mathbf{X}_2'\mathbf{M}_1 = \mathbf{D}\{\bar{\mathbf{J}}_n\} - \bar{\mathbf{J}}_N$$

and so from (13) the hypothesis associated with $R(\boldsymbol{\alpha}|\mu)$ is

$$H : (\mathbf{D}\{\mathbf{1}_n\} - \frac{1}{a}\mathbf{J}_{N\times a})\boldsymbol{\alpha} + (\mathbf{D}\{\bar{\mathbf{J}}_n\} - \bar{\mathbf{J}}_N)\mathbf{x}b = \mathbf{0} ,$$

i.e.,

$$H : \mathbf{D}\{\mathbf{1}_n\}\boldsymbol{\alpha} - \bar{\alpha}_{.}\mathbf{1}_N + b(\mathbf{D}\{\bar{x}_{i.}\mathbf{1}_n\} - \bar{x}_{..}\mathbf{1}_N) = \mathbf{0} . \tag{15}$$

This consists of N statements about the α_i's and b, but scrutiny reveals that they are n repetitions of the a statements

$$H : \alpha_i - \bar{\alpha}_{.} + b(\bar{x}_{i.} - \bar{x}_{..}) = 0 \;\forall\, i . \tag{16}$$

Since the left-hand sides of (16) sum to zero, it is equivalent to both $H : \alpha_i + b(\bar{x}_{i.} - \bar{x}_{..})$ equal $\forall\, i$, and to $H : \alpha_i + b\bar{x}_{i.}$ equal $\forall\, i$. Each of these is, of course, precisely as would be expected. The restatement of (15) in this form illustrates the restatement of (13) mentioned at the end of Section 4.

Example 2: The SAS GLM package includes in its output estimable functions that permit a user to formulate hypotheses corresponding to each output sum of squares. (See, for example, Searle, 1980.) But sometimes the origin of those estimable functions is difficult to determine. We show how (13) can be used for just such a determination.

Consider the "intra-class regression" form of (14):

$$E(y_{ij}) = \mu + \alpha_i + b_i x_{ij} \; . \tag{17}$$

Suppose for (14) that $R(b|\mu,\underset{\sim}{\alpha})$ has been calculated and we seek the hypothesis associated with it in (17). Since b does not explicitly occur in (17) we introduce it by writing

$$b = c, \quad b_i = c + c_i \quad \text{and} \quad E(y_{ij}) = \mu + \alpha_i + (c + c_i)x_{ij} \; . \tag{18}$$

With the model in (18), overparameterized as it is for the covariate, SAS GLM can be made to yield in a single run of the data both $R(b|\mu,\underset{\sim}{\alpha})$ and $R(\underset{\sim}{b}|\mu,\underset{\sim}{\alpha})$ and the difference between them, which provides a test of $H: b_i = b$. We show how (13) yields the hypothesis associated with $R(b|\mu,\underset{\sim}{\alpha})$ in (17). First write the model as

$$E(\underset{\sim}{y}) = \mu\underset{\sim}{1}_N + (\underset{\sim}{D}\{\underset{\sim}{1}_n\})\underset{\sim}{\alpha} + c\underset{\sim}{x} + \underset{\sim}{D}\{\underset{\sim}{x}_i\}\underset{\sim}{c} \tag{19}$$

where $\underset{\sim}{c}' = [c_1 \;\; c_2 \cdots c_a]$ and $D\{\underset{\sim}{x}_i\}$ is a block diagonal matrix of vectors $\underset{\sim}{x}_i = [x_{i1} \;\; x_{i2} \cdots x_{in}]'$. Then, on taking $R(\underset{\sim}{\beta}_2|\underset{\sim}{\beta}_1)$ of the general notation as $R(b|\mu,\underset{\sim}{\alpha}) \equiv R(c|\mu,\underset{\sim}{\alpha})$, we have

$$\underset{\sim}{\beta}_2 = c, \qquad \underset{\sim}{\beta}_1 = [\mu \;\; \underset{\sim}{\alpha}'] \quad \text{and} \quad \underset{\sim}{\beta}_3 = \underset{\sim}{c} \; ;$$

and correspondingly,

$$\underset{\sim}{X}_2 = \underset{\sim}{x}, \qquad \underset{\sim}{X}_1 = [\underset{\sim}{1}_N \;\; \underset{\sim}{D}\{\underset{\sim}{1}_n\}] \quad \text{and} \quad \underset{\sim}{X}_3 = \underset{\sim}{D}\{\underset{\sim}{x}_i\} \; .$$

Therefore

$$\underset{\sim}{M}_1 = \underset{\sim}{I} - \underset{\sim}{D}\{\bar{\underset{\sim}{J}}_n\} \quad \text{with} \quad \underset{\sim}{M}_1\underset{\sim}{X}_2 = \underset{\sim}{x} - \{\bar{x}_{i.}\underset{\sim}{1}_n\}$$

and

$$\underset{\sim}{X}_2'\underset{\sim}{M}_1\underset{\sim}{X}_2 = \sum_{i=1}^{a}\sum_{j=1}^{n}(x_{ij} - \bar{x}_{i.})^2 \quad \text{and} \quad \underset{\sim}{X}_2'\underset{\sim}{M}_1\underset{\sim}{X}_3\underset{\sim}{\beta}_3 = \sum_{i=1}^{a} c_i \sum_{j=1}^{n}(x_{ij} - \bar{x}_{i.})^2 \; .$$

Thus for

$$h_i = \sum_{j=1}^{n}(x_{ij} - \bar{x}_{i.})^2 / \sum_{i=1}^{a}\sum_{j=1}^{n}(x_{ij} - \bar{x}_{i.})^2 \quad \text{with} \quad \sum_{i=1}^{a} h_i = 1 , \tag{20}$$

the hypothesis in (13) is

$$H : (\underset{\sim}{x} - \{\bar{x}_{i.}\underset{\sim}{1}_n\})[c + \Sigma h_i c_i] = 0$$

which can be written equivalently as

$$H : \sum_{i=1}^{a} h_i b_i = 0 . \tag{21}$$

Thus, even though SAS GLM will include in its output numerical values of the h_i's of (20) and (21), the use of (13) in deriving (21) tells us the exact form of those h_i's, as shown in (20).

REFERENCES

Marsaglia, George and Styan, G. P. H. (1974). Rank conditions for generalized inverses of partitioned matrices. Sankhyā, 36, 437-442.

Searle, S. R. (1971). Linear Models. Wiley, New York.

Searle, S. R. (1980). Arbitrary hypotheses in linear models with unbalanced data. Communications in Statistics, A9, 181-200.

Searle, S. R., Speed, F. M. and Henderson, H. V. (1981). Some computational and model equivalences in analyses of variance of unequal-subclass-numbers data. The American Statistician, 35, 16-33.

Sums of Squares and Associated Hypotheses in Partitioned Linear Models.

[The general result (13) is line 5. Other lines are special cases of line 5.]

	Special value for the general vector $[\boldsymbol{\beta}_1' \ \boldsymbol{\beta}_2' \ \boldsymbol{\beta}_3']$	Model[1/] for $E(\mathbf{y})$	Sum of Squares[2/]	Associated Hypothesis[3/]
1.	$[\ \mathbf{0} \quad \boldsymbol{\beta}_1' \quad \mathbf{0}\ ]$	$\mathbf{X}_1\boldsymbol{\beta}_1$	$R(\boldsymbol{\beta}_1)$	H: $\mathbf{X}_1\boldsymbol{\beta}_1 = \mathbf{0}$
2.	$[\ \mathbf{0} \quad \boldsymbol{\beta}_1' \quad \boldsymbol{\beta}_2'\ ]$	$\mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2$	$R(\boldsymbol{\beta}_1)$	H: $\mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_1\mathbf{X}_1^+\mathbf{X}_2\boldsymbol{\beta}_2 = \mathbf{0}$
3.	$[\ \boldsymbol{\beta}_1' \quad \boldsymbol{\beta}_2' \quad \mathbf{0}\ ]$		$R(\boldsymbol{\beta}_2 \mid \boldsymbol{\beta}_1)$	H: $\mathbf{M}_1\mathbf{X}_2\boldsymbol{\beta}_2 = \mathbf{0}$
4.	$[\ \mathbf{0} \quad \boldsymbol{\beta}_1' \quad (\boldsymbol{\beta}_2' \ \boldsymbol{\beta}_3')\ ]$	$\mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \mathbf{X}_3\boldsymbol{\beta}_3$	$R(\boldsymbol{\beta}_1)$	H: $\mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_1\mathbf{X}_1^+(\mathbf{X}_2\boldsymbol{\beta}_2 + \mathbf{X}_3\boldsymbol{\beta}_3) = \mathbf{0}$
5.	$[\ \boldsymbol{\beta}_1' \quad \boldsymbol{\beta}_2' \quad \boldsymbol{\beta}_3'\ ]$		$R(\boldsymbol{\beta}_2 \mid \boldsymbol{\beta}_1)$	H: $\mathbf{M}_1\mathbf{X}_2\boldsymbol{\beta}_2 + \mathbf{M}_1\mathbf{X}_2(\mathbf{M}_1\mathbf{X}_2)^+\mathbf{X}_3\boldsymbol{\beta}_3 = \mathbf{0}$
6.	$[\ (\boldsymbol{\beta}_1' \ \boldsymbol{\beta}_2') \quad \mathbf{0} \quad \boldsymbol{\beta}_3'\ ]$		$R(\boldsymbol{\beta}_3 \mid \boldsymbol{\beta}_1, \boldsymbol{\beta}_2)$	H: $\mathbf{M}_{12}\mathbf{X}_3\boldsymbol{\beta}_3 = \mathbf{0}$

1/ In each model, $\hat{\sigma}^2 = \mathbf{y}'(\mathbf{I} - \mathbf{X}\mathbf{X}^+)\mathbf{y}/(N - r_{\mathbf{X}})$ where $\mathbf{X}$ is $\mathbf{X}_1$, $[\mathbf{X}_1 \mid \mathbf{X}_2]$ and $[\mathbf{X}_1 \mid \mathbf{X}_2 \mid \mathbf{X}_3]$, respectively.

2/ The F-statistic in each case is the sum of squares divided by $s\hat{\sigma}^2$ for s being the degrees of freedom of the sum of squares, which for $R(\boldsymbol{\beta}_2 \mid \boldsymbol{\beta}_1)$ is $r_{[\mathbf{X}_1 \mid \mathbf{X}_2]} - r_{\mathbf{X}_1}$.

3/ $\mathbf{M}_1 = \mathbf{I} - \mathbf{X}_1\mathbf{X}_1^+$ and $\mathbf{M}_{12} = \mathbf{I} - (\mathbf{X}_1 \mid \mathbf{X}_2)(\mathbf{X}_1 \mid \mathbf{X}_2)^+ = \mathbf{M}_1 - \mathbf{M}_1\mathbf{X}_2(\mathbf{M}_1\mathbf{X}_2)^+$.