ON A GENERAL RESULT IN THE METHOD OF FITTING CONSTANTS

S. R. Searle*

Texas A and M University

Abstract

Two conditions are given for using a general result concerning expectations of sums of squares in the method of fitting constants. Application is illustrated for a split-plot design.

Paper No. BU-264-M of the Biometrics Unit of the Department of Plant Breeding and Biometry, Cornell University, Ithaca, New York 14850.

*On leave from Cornell University, 1968-1969.

ON A GENERAL RESULT IN THE METHOD OF FITTING CONSTANTS

S. R. Searle*

Texas A&M University

Summary

Two conditions are given for using a general result concerning expectations of sums of squares in the method of fitting constants. Application is illustrated for a split-plot design.

Introduction

Consider the familiar linear model $\underline{y} = \underline{X}\underline{\beta} + \underline{e}$ written, by partitioning $\underline{X}$ and $\underline{\beta}$, as

$$\underline{y} = \underline{X}_1\underline{\beta}_1 + \underline{X}_2\underline{\beta}_2 + \underline{e} \;. \tag{1}$$

Let $R(\beta_1, \beta_2)$ be the reduction in sum of squares due to fitting (1). Similarly, for fitting the sub-model

$$\underline{y} = \underline{X}_1\underline{\beta}_1 + \underline{e} \;, \tag{2}$$

let $R(\beta_1)$ be the reduction in sum of squares. Then it is shown in Searle [1968, equation (28)] that for

$$R(\beta_2|\beta_1) = R(\beta_1, \beta_2) - R(\beta_1) \tag{3}$$

$$E\, R(\beta_2|\beta_1) = tr\left\{\underline{X}_2' \,[I - \underline{X}_1(\underline{X}_1'\underline{X}_1)^-\underline{X}_1']\underline{X}_2 \; E(\underline{\beta}_2\underline{\beta}_2')\right\} + \sigma_e^2[r(\underline{X}_1 \;\; \underline{X}_2) - r(\underline{X}_1)] \tag{4}$$

In this result E represents expectation over the full model, model (1); 'tr' represents the trace of a matrix (the sum of its diagonal elements), and $(\underline{X}_1'\underline{X}_1)^-$

---

* On leave from Cornell University, 1968-69.

is any generalized inverse of $X_1'X_1$. $[X_1(X_1'X_1)^-X_1'$ is invariant to whichever generalized inverse is used for $(X_1'X_1)^-$.]

Result (4) is true quite generally. It holds for any combination of fixed and/or random effects that occur in β_1 and/or β_2, and no matter what random effects there are in either β_1 or β_2 (4) is true for whatever value $E(\beta_2\beta_2')$ has in the model. If part of β_2 contains fixed effects the corresponding part of $E(\beta_2\beta_2')$ will be squares and products of those fixed effects; and if part of β_2 contains random effects (with zero means) the corresponding part of $E(\beta_2\beta_2')$ will be the variance-covariance matrix of those random effects, whatever its form may be. Furthermore, $E(\beta_1\beta_1')$ does not arise in (4) and neither does $E(\beta_2\beta_1')$; (4) depends solely on $E(\beta_2\beta_2')$. No matter what fixed effects or random effects occur in β_1, they do not occur in (4).

Conditions for using the general result

In view of the generality of (4) it can be called the general result of the fitting constants method. Implications of its use in estimating variance components are discussed in Searle (1968). But, as has been emphasized, the result applies for all linear models, fixed, random, and mixed. However, two conditions pertaining to its use need to be stated: (i) Every expression of the form $R(\beta_2|\beta_1)$ is the reduction for fitting (1), the full model [under which expectation in (4) is being taken], minus the reduction for fitting some sub-model, represented by (2). (ii) The models used as sub-models in (2) and (3) must, so far as sums of squares are concerned, be different from each other and from the full model.

At first reading, the necessity of these two conditions may seem obvious. Yet, overlooking them can be all too easy and leads, of course, to erroneous results. Suppose we have a 2-way classification with interaction, with its model being

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ijk} . \tag{5}$$

Then the reduction in sum of squares for fitting (5) will be denoted by $R(\mu, \alpha, \beta, \alpha\beta)$. And for fitting the sub-model

$$y_{ijk} = \mu + \alpha_i + \beta_j + e_{ijk} \tag{6}$$

the reduction in sum of squares is similarly denoted $R(\mu, \alpha, \beta)$. Thus

$$R(\alpha\beta \mid \mu, \alpha, \beta) = R(\mu, \alpha, \beta, \alpha\beta) - R(\mu, \alpha, \beta)$$

is an example of (3) to which (4) applies. But

$$R(\beta \mid \mu, \alpha) = R(\mu, \alpha, \beta) - R(\mu, \alpha), \tag{7}$$

although an example of (3), cannot be used in (4) when expectation is under the model (5), because the first term of (7) is not the reduction due to fitting (5); i.e. in taking the expectations of (7) under (5), condition (i) is not upheld, and so the expected value of (7) cannot be derived from (4).

A consequence of (i) can be expressed in terms of the symbolism $R(.|.)$: every expression of this form, in order to be used in (4) must have in it all symbols of the full model. These can appear on either side of the bar in $R(.|.)$, but they must all be there. Thus $R(\beta|\mu, \alpha)$ of (7) does not qualify because it lacks the symbol $\alpha\beta$ of the full model; but $R(\mu, \alpha|\beta, \alpha\beta)$ and $R(\mu, \beta|\alpha, \alpha\beta)$ do qualify.

Of course, condition (i) does not preclude using (4) to obtain the expected value of (7) under (5); it only precludes using it directly. But it can be used indirectly by writing

$$R(\beta|\mu, \alpha) = R(\mu, \alpha, \beta) - R(\mu, \alpha)$$

$$= R(\mu, \alpha, \beta, \alpha\beta) - R(\mu, \alpha) - [R(\mu, \alpha, \beta, \alpha\beta) - R(\mu, \alpha, \beta)]$$

$$= R(\beta, \alpha\beta|\mu, \alpha) - R(\alpha\beta|\mu, \alpha, \beta)$$

and for each of these last two terms (4) can be used.

Having emphasized that in $R(.|.)$ every symbol of a model must appear in order to be able to use (4), it is clear from (4) that $E\ R(.|.)$ involves σ_e^2 and only those elements of the model indicated on the left of the bar in the symbol $R(.|.)$. For example,

$$E\ R(\beta, \alpha\beta|\mu, \alpha) \text{ involves } \sigma_e^2, \beta\text{'s and } \alpha\beta\text{'s,}$$

and (8)

$$E\ R(\alpha\beta|\mu, \alpha, \beta) \text{ involves } \sigma_e^2 \text{ and } \alpha\beta\text{'s.}$$

In random models this involvement of elements other than σ_e^2 is always in terms of variance components, but in mixed models it can be in terms of quadratic functions of the fixed effects just as it is in fixed models.

Condition (ii) ensures that the $R(.|.)$ symbols are not used blindly. For example, in the 2-way classification, one <u>cannot</u> <u>use</u>

$$R(\beta|\mu, \alpha, \alpha\beta) = R(\mu, \alpha, \beta, \alpha\beta) - R(\mu, \alpha, \alpha\beta). \tag{9}$$

By condition (i) this term is permissible because all symbols of the full model are included in the symbol $R(\beta|\mu, \alpha, \alpha\beta)$, but on the right-hand side of (9) the two terms are identical. This is so because $R(\mu, \alpha, \alpha\beta)$ comes from fitting the model

$$y_{ijk} = \mu + \alpha_i + (\alpha\beta)_{ij} + e_{ijk} ,$$

and so far as reduction in sum of squares is concerned this is indistinguishable from the full model (5). Hence $R(\mu, \alpha, \beta, \alpha\beta)$ and $R(\mu, \alpha, \alpha\beta)$ are identical, equal to $\sum_i \sum_j y_{ij.}^2/n_{ij}$, and so (9) is identically zero. Similarly $R(\mu, \alpha\beta) = R(\mu, \alpha, \beta, \alpha\beta)$ and $R(\alpha, \beta | \mu, \alpha\beta)$ is identically zero and can never be used as part of the fitting constants method.

A split-plot design

Consider a split-plot experiment where the main plots form a randomized complete blocks design. Let

$$y_{ijk} = \mu + t_i + b_j + (tb)_{ij} + p_k + (tp)_{jk} + e_{ijk} \tag{10}$$

be the model for the observation on the k'th sub-plot in the j'th block receiving the i'th treatment. If there are T treatments, B blocks and P sub-plots per block, the sum of squares due to treatments is

$$SST = BP \sum_{i=1}^{T} (\bar{y}_{i..} - \bar{y}_{...})^2 = BP \sum_{i=1}^{T} \bar{y}_{i..}^2 - BPT\bar{y}_{...}^2 \, . \tag{11}$$

Now suppose the block effects and the treatment-block interaction effects are random, with zero means, and variances σ_b^2 and σ_{tb}^2 respectively. Then, on substituting (10) into (11) and taking expected values it can be shown in the usual way that

$$E(SST) = (T - 1)\left\{ \frac{B\,P}{T-1} \sum_{i=1}^{T} [t_i - \bar{t}_i + (\bar{t}p)_{i.} - (\bar{t}p)_{..}]^2 + P\sigma_{tb}^2 + \sigma_e^2 \right\} \tag{12}$$

where $\bar{t}_{.} = \sum_{i=1}^{T} t_i/T$, $(\bar{t}p)_{i.} = \sum_{j=1}^{B} (\bar{t}p)_{ij}/B$ and $(\bar{t}p)_{..} = \sum_{i=1}^{T} \sum_{j=1}^{B} (tp)_{ij}/TB$. With the customary restrictions

$$\sum_{i=1}^{T} t_i = 0, \quad \text{and} \quad \sum_{j=1}^{B} (tp)_{ij} = 0 \text{ for all } i \tag{13}$$

as part of the model, $\bar{t}_{.} = (\bar{t}p)_{i.} = (\bar{t}p)_{..} = 0$ and (12) takes its familiar form

$$E(SST) = (T - 1)\left[BP \frac{\sum_{i=1}^{T} t_i^2}{T - 1} + P\,\sigma_{tb}^2 + \sigma_e^2\right] , \tag{14}$$

as found in Steel and Torrie (1960), for example. We show that this is in complete agreement with the general result (4) for fitting constants.

Because the data are balanced

$$SST = R(\text{full model}) - R(\text{model with no t-effects}) \tag{15}$$

which, in view of the model (10) one would be tempted to write as

$$SST = R(\mu, t, b, tb, p, tp) - R(\mu, b, tb, p, tp) . \tag{16}$$

However, for the reasons illustrated in discussing condition (ii) above, the two terms on the right of (16) are identically equal. The error in suggesting (16) as the form of (15) is that in the second term of (15) one must omit not only the t-effects but also the interactions involving treatments. Hence in (15) SST is

$$SST = R(\mu, t, b, tb, p, tp) - R(\mu, b, p)$$

$$= R(t, tb, tp|\mu, b, p) .$$

To this (4) applies, giving

$$\begin{aligned} &E(SST) \text{ involving t's, tb's, tp's and } \sigma_e^2 , \\ \text{i.e.} \quad &E(SST) \text{ involves t's, } \sigma_{tb}^2\text{, tp's and } \sigma_e^2 . \end{aligned} \tag{17}$$

And in (12) we see that this is exactly what E(SST) does involve. It is only through the further use of (13) that E(SST) reduces to its more familiar form (14), which shows no terms in the tp's, because their occurrence in (12) is such that (13) makes them vanish.

The use of (4) of course is not to just indicate what terms are involved in E(SST) as in (17) but to evaluate E(SST) specifically. Carrying out the details of this yields (12) which can be modified to (14) as indicated.

References

Searle, S. R. (1968). Another look at Henderson's methods of estimating variance components. Biometrics, 24, 749-788.

Steele, R. G. D., and Torrie, J. H. (1960). "Principles and Procedures of Statistics". McGraw-Hill, New York.