Experiments With Mixtures

By Henry Scheffé*

University of California, Berkeley

[Received August, 1957]

Summary

A theory is developed for experiments with mixtures of q components whose purpose is the empirical prediction of the response to any mixture of the components, when the response depends only on the proportion of the components and not on the total amount.

1. Introduction

We consider experiments with mixtures, of which the property studied depends on the proportions of the components present, but not on the amount of the mixture. Examples are the tensile strength of an alloy of different metals, the wear-resistance, suitably defined, of a mixture of different kinds of rubber, and the octane rating of a blend of different gasoline stocks. The effect of a fertilizer which is a mixture of certain components, on the yield of a crop would not be an example, because this yield would depend not only on the proportions used but on the total amount.

In a q-component mixture ($q \geqq 3$) let x_i be the proportion (by volume, or by weight, or by moles, etc.) of the i^{th} component in the mixture, so that

$$x_i \geqq 0 \ (i = 1, 2, \ \ldots \ , q), \qquad x_1 + x_2 + \ \ldots \ + x_q = 1. \tag{1.1}$$

The factor space is thus a regular $(q - 1)$-dimensional simplex (triangle for $q = 3$, tetrahedron for $q = 4$). The customary complete or fractional factorial experimental designs are possible if we are interested only in mixtures in which one of the components, possibly an inert one or a diluent, is present in a proportion near unity, and the other components only in proportions near zero. Then the large component may be ignored in the analysis, and all combinations of various small proportions chosen for the other components are possible. An example might be an experiment on the tensile strength of steel with small proportions of sulphur, manganese, and nickel in the iron. The usual complete factorial design then corresponds to the choice of a lattice of points in the form of a parallelopiped fitted into the corner of the simplex where the large component takes on the value unity. This design is not possible if the sum of the largest proportions to be used for each of the $q - 1$ small components exceeds unity.

That the factor space for experiments with mixtures is a simplex is noted by Claringbold (1955) in a paper on the joint action of hormones. The term *simplex design* is there used for any set of N experimental points on the simplex. Designs are briefly considered for

* This paper was prepared with the partial support of the Office of Naval Research.

experiments where the response depends also on the total amount, different amounts being used as different levels of an extra factor, and possibly other factors are varied.

2. Simplex Lattices

If it is desired to explore the whole factor space a design to which one is naturally led, and which gives a uniformly spaced distribution of points over the factor space, is what we shall call a $\{q, m\}$ *lattice*: In this the proportions used for each factor have the $m + 1$ equally spaced values from 0 to 1, $x_i = 0, 1/m, 2/m, \ldots, 1$, and all possible

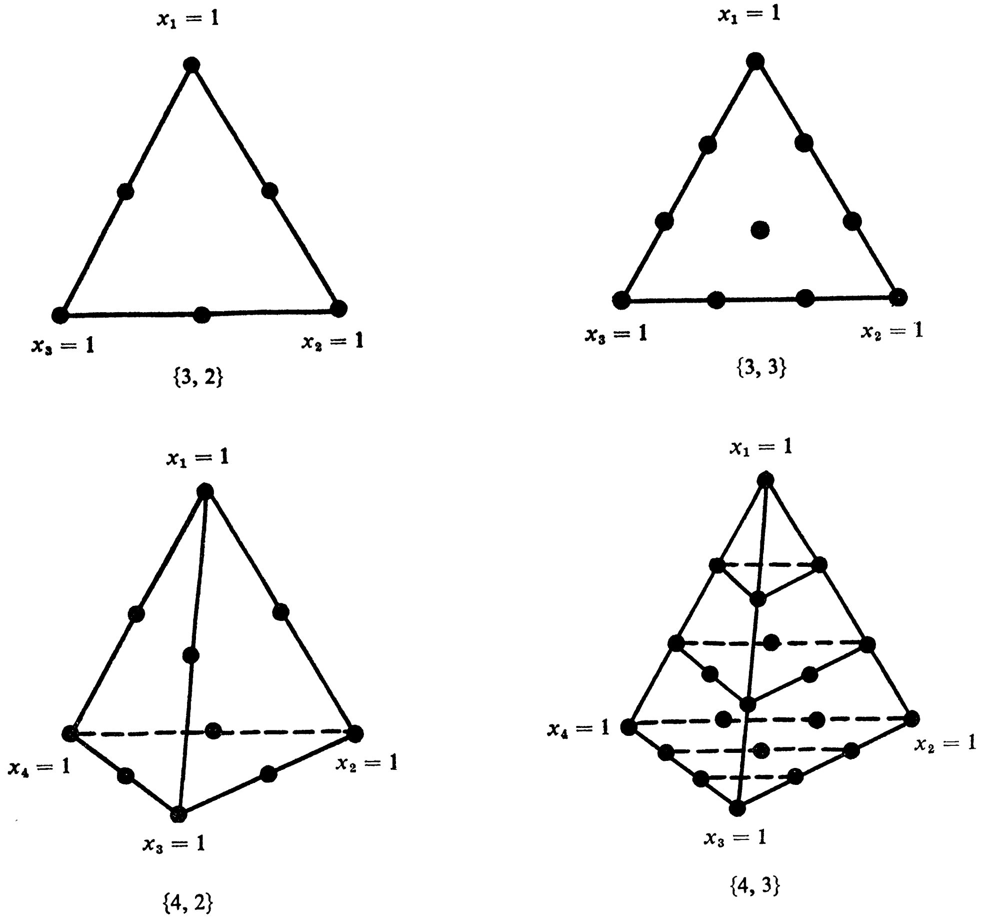

FIG. 1.—Some $\{q, m\}$ lattices.

mixtures with these proportions for each factor are used. Some $\{q, m\}$ lattices are pictured in Fig. 1. It is shown in sec. 9.1 of the appendix that the number of points in this design is

$$\binom{m+q-1}{m} = \frac{q(q+1) \ \ldots \ (q+m-1)}{1 \, . \, 2 \ \ldots \ m}. \tag{2.1}$$

It is necessary to modify this design if for some reason one or more of the components can be used only in small proportions; there is some discussion of such modifications in sec. 8.

The property studied, assumed to be a real-valued function on the simplex, will be called the *response*. It does not seem possible to extend the customary resolution of the response into general mean, main effect, and interactions of various orders, to this kind of experiment. We shall introduce instead an appropriate form of polynomial regression. In practice the polynomial will have to be of low degree, because of the large number of coefficients otherwise present in a polynomial in several variables. This approach will then not be useful if the response function has discontinuities or other peculiarities which a low degree polynomial cannot emulate.

3. Polynomials on a Simplex

A *polynomial* function of degree n in the q variables $x_1, x_2, \ldots, x_q$ subject to (1.1) will be called a $\{q, n\}$ *polynomial*. It will be of the form*

$$\begin{aligned}\eta = \beta_0 + \sum_{1 \leq i \leq q} \beta_i x_i + \sum_{1 \leq i \leq j \leq q} \beta_{ij} x_i x_j \\ + \sum_{1 \leq i \leq j \leq k \leq q} \beta_{ijk} x_i x_j x_k + \ldots \\ \ldots + \sum_{1 \leq i_1 \leq i_2 \leq \ldots \leq i_n \leq q} \beta_{i_1 i_2 \ldots i_n} x_{i_1} x_{i_2} \ldots x_{i_n},\end{aligned} \tag{3.1}$$

where the coefficients β are constants. It is shown in sec. 9.2 that the number of coefficients in (3.1) is $\binom{n+q}{n}$. Because the function (3.1) has meaning for us only subject to the restriction (1.1), the coefficients β are not unique. We might for example substitute

$$x_q = 1 - \sum_{i=1}^{q-1} x_i \tag{3.2}$$

in (3.1), and the same function would appear as a polynomial in the $q-1$ variables $x_1, x_2, \ldots, x_{q-1}$. This then involves only $\binom{n+q-1}{n}$ coefficients, the number being the same as the number of points in a $\{q, n\}$ lattice. The equality of these numbers suggests that polynomial regression methods may be especially well adapted to the simplex lattice design and its modifications. The writer conjectures that the values of a $\{q, n\}$ polynomial can be assigned arbitrarily on a $\{q, n\}$ lattice and that its values on the simplex (1.1) are then uniquely determined; we shall verify this below for $n = 1, 2$, and 3.

The special form of (3.1) resulting from the substitution (3.2), namely that in which all coefficients β for which q appears among the subscripts are zero, is not a nice canonical form because of its asymmetry, x_q having the distinguished role of absence. Furthermore, the coefficients β in this form do not have simple interpretations as in the canonical forms we will now introduce for $n = 1, 2, 3$.

* We do not require that in a $\{q, n\}$ polynomial some coefficient in the last sum in (3.1) be non-zero: Thus a $\{q, n\}$ polynomial is also a $\{q, n'\}$ polynomial for $n' > n$. If we did impose the requirement, the degree would still not be unique, since if the polynomial is not identically zero the degree can be raised by multiplying the polynomial by $x_1 + x_2 + \ldots + x_q$.

It is shown in sec. 9.3 that the most general polynomial of degree n in q variables subject to (1.1) may be written

$$\eta = \sum_{1 \leqq i \leqq q} \beta_i x_i \tag{3.3$_1$}$$

if $n = 1$,

$$\eta = \sum_{1 \leqq i \leqq q} \beta_i x_i + \sum_{1 \leqq i < j \leqq q} \beta_{ij} x_i x_j \tag{3.3$_2$}$$

if $n = 2$, and if $n = 3$

$$\eta = \sum_{1 \leqq i \leqq q} \beta_i x_i + \sum_{1 \leqq i < j \leqq q} \beta_{ij} x_i x_j + \sum_{1 \leqq i < j \leqq q} \gamma_{ij} x_i x_j (x_i - x_j) + \sum_{1 \leqq i < j < k \leqq q} \beta_{ijk} x_i x_j x_k. \tag{3.3$_3$}$$

(The notation in (3.3$_3$) might be made more symmetrical by writing $\gamma_{ij} x_i x_j (x_i - x_j)$ in the form $\beta_{iij} x_i^2 x_j + \beta_{ijj} x_i x_j^2$ with $\beta_{iij} + \beta_{ijj} = 0$.) In (3.3$_2$) and (3.3$_3$) the number of coefficients β_{ij} is $q(q-1)/2$, in (3.3$_3$) the number of coefficients γ_{ij} is the same, and the number of coefficients β_{ijk} in (3.3$_3$) is $q(q-1)(q-2)/6$.

4. DETERMINATION OF THE COEFFICIENTS IN THE POLYNOMIALS

To interpret the coefficients in these polynomials suppose the response function is denoted by η, and that η has any of the forms (3.3). Then the coefficients with a single subscript i may be interpreted in terms of the response to pure component i, those with the subscripts i and j, in terms of the response to binary mixtures of components i and j, and those with the subscripts i, j, and k in terms of the response to ternary mixtures of components i, j, and k.

We shall denote the response to pure component i by η_i. Putting $x_i = 1$ (hence $x_j = 0$ for $j \neq i$) and $\eta = \eta_i$ in any of the three equations (3.3) we find

$$\beta_i = \eta_i, \tag{4.1}$$

and so the coefficients β_i are the responses to the pure components. If the linear formula (3.3$_1$) were valid the response would then be given by

$$\sum_{i=1}^{q} \beta_i x_i = \sum_{i=1}^{q} x_i \eta_i.$$

The excess of the response η over this linear mixing or linear blending value calculated from the proportions x_i and the responses η_i to the pure components is called in various fields (pharmacology, petroleum refining) the *synergism* of the mixture; its negative is also called the *antagonism*.

The response to a binary mixture of components i and j in proportions x_i and x_j ($x_i + x_j = 1$, $i < j$) is given by the quadratic formula (3.3$_2$) as $\beta_i x_i + \beta_j x_j + \beta_{ij} x_i x_j$. The synergism of the binary mixture is thus $\beta_{ij} x_i x_j$, and the coefficient β_{ij} may therefore be called the quadratic coefficient of the binary synergism of components i and j. If instead the cubic formula (3.3$_2$) were valid the synergism of the binary mixture would have the additional term $\gamma_{ij} x_i x_j (x_i - x_j)$, and so γ_{ij} may be called the cubic coefficient of the binary synergism of components i and j.

If we assume the validity of the cubic formula (3.3$_3$) we find that the synergism of a ternary mixture of components i, j and k in proportions x_i, x_j and x_k ($x_i + x_j + x_k = 1$, $i < j < k$) is

$$[\beta_{ij}x_ix_j + \gamma_{ij}x_ix_j(x_i - x_j)] + [\beta_{ik}x_ix_k + \gamma_{ik}x_ix_k(x_i - x_k)] \\ + [\beta_{jk}x_jx_k + \gamma_{jk}x_jx_k(x_j - x_k)] + \beta_{ijk}x_ix_jx_k.$$

If we extend the terminology to call $\beta_{ij}x_ix_j + \gamma_{ij}x_ix_j(x_i - x_j)$ the binary synergism of components i and j even if $x_i + x_j < 1$, and similarly call the other terms in square brackets binary synergisms in the ternary mixture, then we are led to call the last term the ternary synergism, and β_{ijk}, the cubic coefficient of the ternary synergism of components i, j and k.

To evaluate the coefficients β and γ in terms of the responses at the points of a $\{q, 2\}$ or $\{q, 3\}$ lattice it will be convenient to introduce the notation for these responses indicated in Figs. 2 and 3: The response to pure component i is denoted by η_i, the response to a

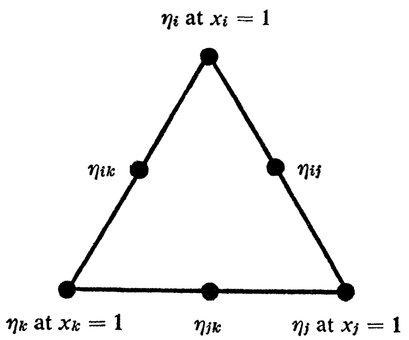

FIG. 2.—Notation for responses on $\{q, 2\}$ lattice.

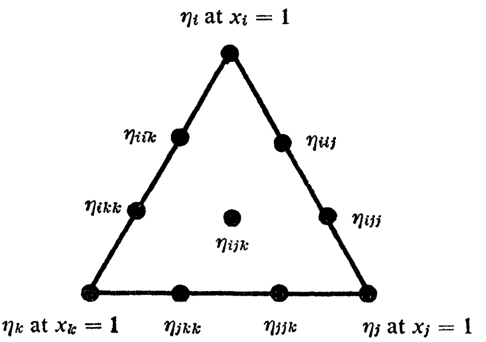

FIG. 3.—Notation for responses on $\{q, 3\}$ lattice.

1 : 1 binary mixture of components i and j by η_{ij} ($i < j$), the response to a 1 : 1 : 1 ternary mixture of components i, j, k by η_{ijk} ($i < j < k$), and the responses to 2 : 1 and 1 : 2 binary mixtures of components i and j, respectively by η_{iij} and η_{ijj} ($i < j$).

To evaluate β_{ij} in terms of the responses on the $\{q, 2\}$ lattice, put $x_i = x_j = 1/2$ and $\eta = \eta_{ij}$ in (3.3_2) or (3.3_3) to get $\eta_{ij} = \frac{1}{2}\beta_i + \frac{1}{2}\beta_j + \frac{1}{4}\beta_{ij}$, which together with (4.1) gives

$$\beta_{ij} = 4\eta_{ij} - 2\eta_i - 2\eta_j. \tag{4.2}$$

To evaluate it on the $\{q, 3\}$ lattice put $x_i = \frac{2}{3}$, $x_j = \frac{1}{3}$, $\eta = \eta_{iij}$ ($i < j$) in (3.3_3) to get

$$\eta_{iij} = \frac{2}{3}\beta_i + \frac{1}{3}\beta_j + \frac{2}{9}\beta_{ij} + \frac{2}{27}\gamma_{ij}, \tag{4.3}$$

and put $x_i = \frac{1}{3}$, $x_j = \frac{2}{3}$, $\eta = \eta_{ijj}$ to get

$$\eta_{ijj} = \frac{1}{3}\beta_i + \frac{2}{3}\beta_j + \frac{2}{9}\beta_{ij} - \frac{2}{27}\gamma_{ij}. \tag{4.4}$$

Adding these equations, we find

$$\beta_{ij} = \frac{9}{4}(\eta_{iij} + \eta_{ijj} - \eta_i - \eta_j). \tag{4.5}$$

If we subtract (4.4) from (4.3) we find

$$\gamma_{ij} = \frac{9}{4}(3\eta_{iij} - 3\eta_{ijj} - \eta_i + \eta_j). \tag{4.6}$$

Those accustomed to using orthogonal polynomials in one variable on an equally spaced set of points will recognize (4.2) as the quadratic contrast on three points, and (4.5) and (4.6) as the quadratic and cubic contrasts, respectively, on four points. The three points involved in (4.2) lie on one edge of the $\{q, 2\}$ lattice as shown in Fig. 2, the four points in (4.5) and (4.6), on one edge of the $\{q, 3\}$ lattice as in Fig. 3.

To evaluate β_{ijk} we put $x_i = x_j = x_k = \frac{1}{3}$ and $\eta = \eta_{ijk}$ $(i < j < k)$ in (3.3_3) to get

$$\beta_{ijk} = 27\eta_{ijk} - 9(\eta_i + \eta_j + \eta_k) - 3(\beta_{ij} + \beta_{ik} + \beta_{jk}), \tag{4.7}$$

and then with the help of (4.2),

$$\beta_{ijk} = 27\eta_{ijk} - \frac{4}{3}(\eta_{ij} + \eta_{ik} + \eta_{jk}) - 3(\eta_i + \eta_j + \eta_k). \tag{4.8}$$

This evaluation will be useful on the augmented $\{q, 2\}$ lattice to be introduced in sec. 6. To express β_{ijk} in terms of the responses on the $\{q, 3\}$ lattice we substitute (4.5) into (4.7); the result is

$$\beta_{ijk} = 27\eta_{ijk} - \frac{27}{4}(\eta_{iij} + \eta_{ijj} + \eta_{iik} + \eta_{ikk} + \eta_{jjk} + \eta_{jkk}) + \frac{9}{2}(\eta_i + \eta_j + \eta_k). \tag{4.9}$$

It is now clear from the canonical forms (3.3) and the formulae (4.1), (4.2), (4.5), (4.6), (4.9) for the coefficients in these forms that for $n = 1, 2, 3$ the values of the general $\{q, n\}$ polynomial on the simplex (1.1) are uniquely determined by those on the $\{q, n\}$ lattice and that the latter may be assigned arbitrarily.

5. DESIGN OF EXPERIMENTS TO FIT POLYNOMIAL REGRESSION FUNCTIONS

In designing an experiment to fit a quadratic or cubic polynomial to the response to a q-component mixture we may take our observations at the points of a $\{q, m\}$ lattice, possibly modified as indicated elsewhere. If n has been chosen as 2 or 3, several paths are then open, including the following two: (i) We may use the $\{q, n\}$ lattice with possibly more than one observation per point, or (ii) we may use a $\{q, m\}$ lattice with $m > n$ and one observation per point. Under design (i) we shall also include the augmented $\{q, 2\}$ lattice defined in sec. 6 for fitting the special cubic introduced there.

We shall consider mainly design (i), which has the advantage that least-squares estimates $\hat{\beta}$ and $\hat{\gamma}$ of the regression coefficients β and γ are easily calculable from the means $\hat{\eta}$ of the observations at the points of the lattice, by replacing the β, γ, η by $\hat{\beta}$, $\hat{\gamma}$, $\hat{\eta}$ in (4.1), (4.2), (4.5), (4.6), (4.8), and (4.9).

With design (ii) the regression coefficients would be calculated by least squares.* It has the advantage that since the true regression function is generally not actually a polynomial of the degree chosen we would expect the maximum deviation of the polynomial approximation from the true function to be usually smaller if instead of forcing the polynomial function through the observed means in design (i) we fit it by least squares on the less distantly spaced points of the lattice in (ii). However, design (ii) has two disadvantages which increase with q.

The number of normal equations which must be constructed and solved is

$$\binom{n+q-1}{n}.$$

If $n = 2$, then already for $q = 3$, 4, 5, this number is 6, 10, 15, respectively. Even with some simplification of the solution permitted by the symmetry of the design, the numerical problem rapidly becomes formidable with increasing q. The solution for $n = 2$, $q = 3$ is considered further in sec. 9.7.

The second disadvantage is on intuitive grounds. In some cases such as gasoline blending, the number q of components in the experiment may be large, while each mixture for which it is desired to predict the response in practice may consist of a relatively small number q' of the q components. In calculations made from design (i) the prediction is based only on observations on mixtures formed from the q' components actually present, whereas for those made from (ii), the prediction is affected by observations on mixtures containing other components than the q' actually present, including mixtures containing none of the q' components.

6. Testing the Fit of a Second Degree Polynomial

If the response function is a quadratic polynomial (3.3_2), the response to any ternary mixture of the components i, j, k $(i < j < k)$ is determined by the three responses η_i, η_j, η_k to the pure components and the three responses η_{ij}, η_{ik}, η_{jk} to their 1 : 1 binary mixtures by equations (3.3_2), (4.1), and (4.2): In particular, by putting $x_i = x_j = x_k = \frac{1}{3}$, $\eta = \eta_{ijk}$ in these equations, we find for the response η_{ijk} to the 1:1:1 ternary mixture.

$$\eta_{ijk} = \frac{4}{9}(\eta_{ij} + \eta_{ik} + \eta_{jk}) - \frac{1}{9}(\eta_i + \eta_j + \eta_k). \tag{6.1}$$

We may then consider how well the observations on these seven responses agree with the formula (6.1) implied by the quadratic regression function (3.3_2). For this purpose, and other tests of fit, we would of course select mixtures suspected of fitting most poorly in the light of whatever prior knowledge or conjecture is available.

Suppose we make r_1 observations of each of the responses η_i, η_j, η_k, r_2 of η_{ij}, η_{ik}, η_{jk}, and r_3 of η_{ijk}, and denote the respective means of the observed responses by $\hat{\eta}_i$, $\hat{\eta}_j$, $\hat{\eta}_k$, $\hat{\eta}_{ij}$, $\hat{\eta}_{ik}$, $\hat{\eta}_{jk}$, $\hat{\eta}_{ijk}$. Then the mean value $\hat{\eta}_{ijk}$ observed for the response η_{ijk} differs from

* The easily calculated estimates used in design (i) are also least squares estimates.

that predicted from (6.1) by

$$d_{ijk} = \hat{\eta}_{ijk} - \frac{4}{9}(\hat{\eta}_{ij} + \hat{\eta}_{ik} + \hat{\eta}_{jk}) + \frac{1}{9}(\hat{\eta}_i + \hat{\eta}_j + \hat{\eta}_k).$$

If the observations are independent with equal variance σ^2 then the variance of d_{ijk} is

$$\sigma^2{}_d = \sigma^2\left(\frac{1}{r_3} + \frac{16}{27r_2} + \frac{1}{27r_1}\right). \tag{6.2}$$

If the total number of observations involved in d_{ijk} is fixed,

$$r_3 + 3r_2 + 3r_1 = \text{constant}, \tag{6.3}$$

then (6.2) is minimized if the numbers of observations are in the proportion

$$r_1 : r_2 : r_3 = 1 : 4 : 9.$$

For the purpose of testing the fit in this way (allocation of measurements for the purpose of estimating the response after the type of polynomial has been decided on is considered in sec. 7) it is then most efficient to take the observations as nearly as possible in these proportions; for example, if we wish to take about 10–15 observations to test a difference d_{ijk}, then we should take $r_1 = 1$, $r_2 = 2$, $r_3 = 4$ or 5. In formulating the side condition (6.3) for this calculation we have ignored the fact that the same $\hat{\eta}_i$ or $\hat{\eta}_{ij}$ may enter into different d_{ijk} tested. If the observations are unbiased (i.e., the expected value of an observation is the true response) then the expected value of d_{ijk} will be zero if a quadratic response formula (3.3_2) is valid. If the observations are assumed to be normal, an independent estimate of σ^2 is available from the repeated observations, which may be employed in (6.2) to construct a t-test for d_{ijk}.

Let δ_{ijk} denote the difference of the true value of the 1 : 1 : 1 ternary response η_{ijk} from the value (6.1) of η_{ijk} implied by the quadratic response formula (3.3_2) and determined from the true values of $\eta_i, \eta_j, \eta_k, \eta_{ij}, \eta_{ik}, \eta_{jk}$; then

$$\delta_{ijk} = \eta_{ijk} - \frac{4}{9}(\eta_{ij} + \eta_{ik} + \eta_{jk}) + \frac{1}{9}(\eta_i + \eta_j + \eta_k).$$

It is easy to get a confidence interval for δ_{ijk} or to calculate the power of the above test of the hypothesis that $\delta_{ijk} = 0$, and either of these techniques will aid in deciding whether a quadratic response function is adequate for the purpose of predicting the response.

An F-statistic for simultaneously testing several, say r, of the d_{ijk} may be constructed. This is easy if the d_{ijk} have no subscripts in common, for then they will be independent, and their sum of squares will be distributed as σ_d^2 times χ^2 with r d.f. If no two d_{ijk} share more than one subscript the correlations will be very small and the same F-test will still be approximately valid. If some of the d_{ijk} share two subscripts it may be desirable to calculate the F-statistic by the exact method given in sec. 9.6. In any case, the denominator of the F-statistic is the estimate of σ^2 constructed from the repeated observations. The covariance of two d_{ijk} which share a single subscript is

$$\frac{\sigma^2}{81r_1}, \tag{6.4}$$

of two d_{ijk} which share two subscripts,

$$\frac{\sigma^2}{81}\left[\frac{16}{r_2}+\frac{2}{r_1}\right]. \tag{6.5}$$

If the quadratic formula is deemed inadequate for graduating the response the simplest remedy would be to go on to the special cubic obtained by taking all the γ_{ij} in (3.3$_3$) to be zero. Estimates $\hat{\beta}_{ijk}$ of the β_{ijk} could be obtained by adding to the $\{q, 2\}$ lattice the $q(q-1)(q-2)/6$ points for the $1:1:1$ ternary mixtures, and using (4.8) with β and the η replaced by $\hat{\beta}$ and $\hat{\eta}$. This forces the fitted response function through all the mean responses observed on the augmented lattice.

We remark that the above test of the hypothesis that $\delta_{ijk}=0$ may also be regarded as a test of the hypothesis that $\beta_{ijk}=0$ in the special or general cubic since $\beta_{ijk}=27\delta_{ijk}$.

Before launching a large-scale programme to determine coefficients for any particular kind of regression formula, such as the quadratic, the above special cubic, or the general cubic, it would usually be wise to investigate, by methods similar to that discussed above for the $1:1:1$ ternary mixture, the fit for some mixtures other than those used to calculate the coefficients in the formula; we have seen this is possible without first completing the programme of determining all the coefficients.

If it is decided that the special cubic is also inadequate we may go on to the general cubic.

7. Variance of Predicted Response. Allocation of Measurements

We consider fitting a second or third degree polynomial formula to the response. After we have decided whether to use the quadratic formula

$$\eta=\sum_{1\leq i\leq q}\beta_i x_i+\sum_{1\leq i<j\leq q}\beta_{ij}x_i x_j, \tag{7.1$_1$}$$

or the special cubic

$$\eta=\sum_{1\leq i\leq q}\beta_i x_i+\sum_{1\leq i<j\leq q}\beta_{ij}x_i x_j+\sum_{1\leq i<j<k\leq q}\beta_{ijk}x_i x_j x_k, \tag{7.1$_2$}$$

or the general cubic

$$\eta=\sum_{1\leq i\leq q}\beta_i x_i+\sum_{1\leq i<j\leq q}\beta_{ij}x_i x_j + \sum_{1\leq i<j\leq q}\gamma_{ij}x_i x_j(x_i-x_j)+\sum_{1\leq i<j<k\leq q}\beta_{ijk}x_i x_j x_k, \tag{7.1$_3$}$$

we adopt the $\{q, 2\}$ lattice design if we have chosen (7.1$_1$), the $\{q, 2\}$ lattice augmented by the points corresponding to all the $1:1:1$ ternary mixtures if we have chosen (7.1$_2$), and the $\{q, 3\}$ lattice if (7.1$_3$). The estimates of the coefficients in (7.1) are then linear combinations of the mean responses $\hat{\eta}$ observed on the lattice used, the estimates $\hat{\beta}$ and $\hat{\gamma}$ being the same linear combinations of the $\hat{\eta}$ as the β and γ are of the η in (4.1), (4.2) for polynomial (7.1$_1$), in (4.1), (4.2), (4.8) for (7.1$_2$), and in (4.1), (4.5), (4.6), (4.9) for (7.1$_3$). The variances and covariances of these estimated coefficients are thus easily calculable since the $\hat{\eta}$ are independent. Of greater interest usually will be the variance of the predicted response calculated with these estimated coefficients.

If in the chosen formula (7.1) we replace β and γ by their estimates $\hat{\beta}$ and $\hat{\gamma}$ evaluated in terms of the mean responses $\hat{\eta}$ as just described, we may collect coefficients on the $\hat{\eta}$

and we then obtain

$$\tilde{\eta} = \sum_{1 \leq i \leq q} a_i \hat{\eta}_i + \sum_{1 \leq i < j \leq q} a_{ij} \hat{\eta}_{ij} \tag{7.2$_1$}$$

for the quadratic, where $\tilde{\eta}$ denotes the predicted response to a mixture with proportions $x_1, x_2, \ldots, x_q$, the a_i and a_{ij} are functions of $(x_1, x_2, \ldots, x_q)$, and $\hat{\eta}_i$ and $\hat{\eta}_{ij}$ are the observed mean responses on the $\{q, 2\}$ lattice. Likewise for the special cubic

$$\tilde{\eta} = \sum_{1 \leq i \leq q} b_i \hat{\eta}_i + \sum_{1 \leq i < j \leq q} b_{ij} \hat{\eta}_{ij} + \sum_{1 \leq i < j < k \leq q} b_{ijk} \hat{\eta}_{ijk}, \tag{7.2$_2$}$$

and for the general cubic

$$\tilde{\eta} = \sum_{1 \leq i \leq q} c_i \hat{\eta}_i + \sum_{1 \leq i < j \leq q} (c_{iij} \hat{\eta}_{iij} + c_{ijj} \hat{\eta}_{ijj}) + \sum_{1 \leq i < j < k \leq q} c_{ijk} \hat{\eta}_{ijk}. \tag{7.2$_3$}$$

The formulae for the coefficients a, b, c in these equations will be given below; they are useful for calculating the variance of the predicted response; for example, for the quadratic,

$$\text{Var}\,(\tilde{\eta}) = \sigma^2 \left(\sum_{1 \leq i \leq q} \frac{a^2_i}{r_i} + \sum_{1 \leq i < j \leq q} \frac{a^2_{ij}}{r_{ij}} \right),$$

where r_i and r_{ij} are respectively the numbers of observations on η_i and η_{ij}, and the observations are assumed to be independent with equal variance σ^2.

We shall now propose a certain formulation of the problem of allocation of the measurements to the points of the lattice, which we will be able to solve, and for which the solution turns out to be (with a qualification of no practical importance) to take equal numbers of observations on each of the lattice points. The criterion* is to take the number of observations at each lattice point proportional to the maximum, over the whole factor space (1.1), of the square of the coefficient in (7.2) of the $\hat{\eta}_i$, $\hat{\eta}_{ij}$, $\hat{\eta}_{iij}$, $\hat{\eta}_{ijj}$, or $\hat{\eta}_{ijk}$ observed at the lattice point. This allocates to each observed mean such a number of observations that the maximum contribution to the variance of the predicted response is the same from each of the observed means entering into the formula (7.2) for the response.

It is evident that each coefficient in (7.2) takes on the value unity, namely at the point of the lattice where the $\hat{\eta}$ it multiplies is measured, and that the other coefficients in the same formula (7.2) must be zero at this point, since the predicted response $\tilde{\eta}$ equals the observed mean response at this point no matter what the values of the other mean responses. It is shown in sec. 9.5 that the maximum of the square of each of the coefficients is unity, with the exception of c_{iij} and c_{ijj} for which it is 1.12.

The formulae for the coefficients are

$$a_i = x_i(2x_i - 1), \quad a_{ij} = 4x_i x_j,$$

$$b_i = \frac{1}{2} x_i (6x^2_i - 2x_i + 1 - 3 \sum_{j=1}^{q} x^2_j), \quad b_{ij} = 4x_i x_j (3x_i + 3x_j - 2), \quad b_{ijk} = 27 x_i x_j x_k,$$

$$c_i = \frac{1}{2} x_i (3x_i - 1)(3x_i - 2), \quad c_{iij} = \frac{9}{2} x_i x_j (3x_i - 1), \quad c_{ijj} = \frac{9}{2} x_i x_j (3x_j - 1), \quad c_{ijk} = 27 x_i x_j x_k.$$

* A natural minimax criterion might be some variation of the following, which has unfortunately not been tractable for me: First maximize Var $(\tilde{\eta})$ over the simplex for fixed numbers of observations, then allocate the numbers to minimize this maximum variance subject to a fixed total number.

Their derivation is indicated in sec. 9.4. Because of the properties stated in the preceding paragraph these formulae give, for each of the three lattices considered, a set of polynomials orthogonal on the lattice in the sense that each equals unity at the lattice point associated with it, that is, at the point where the mean response is observed that multiplies the polynomial in (7.2), and equals zero at the other lattice points. If required these orthogonal polynomials could be expressed in our canonical forms (3.3) by substituting the expressions of sec. 9.3, or by deriving them as explained above.

8. Modifications of Simplex Lattice Designs

If it is necessary to keep the proportions of one or more of the components below certain bounds, so that the $\{q, m\}$ lattice design is ruled out, it may yet be possible to satisfy these conditions without enormously complicating the numerical solution for the regression coefficients, by a judicious modification of the $\{q, m\}$ lattice; this holds also for the augmented $\{q, 2\}$ lattice.

Suppose that only the first component has to be restricted and is subject to the condition $x_1 \leqq h$. The modification which is computationally simplest is to replace the first component throughout the simplex lattice designs by a "pseudocomponent" which is a mixture of a proportion h of the first component and proportions p_i of the other components

$$(i > 1, \sum_{i=2}^{q} p_i = 1 - h),$$

the p_i being selected by the experimenter. If x'_1 is the proportion of the pseudocomponent and $x'_2, \ldots, x'_q$ are the proportions of the other components added to the pseudocomponent (for $i > 1$, x'_1 does not include the part of component i in the pseudocomponent), then we can use the methods discussed thus far to establish a regression equation for the response as a polynomial in $x'_1, x'_2, \ldots, x'_q$, and by substituting in this

$$x'_1 = h^{-1} x_1, \; x'_i = x_i - h^{-1} p_i x_1 \qquad (i > 1),$$

get a polynomial of the same degree in the actual proportions $x_1, x_2, \ldots, x_q$ of the components. The last polynomial will not be in canonical form but can be brought to this by the substitutions of sec. 9.3.

This method can be generalized to the case of several restricted components by introducing a pseudocomponent for each. A shortcoming of the method is that it does not distribute the experimental points very evenly over the part of the factor space satisfying the bounds imposed. Thus, if only one component is restricted, say the first, and $x_1 \leqq h$, then the available factor space is the frustrum of the simplex (1.1) satisfying $x_1 \leqq h$, and the lattice of points used lies in a simplex which contains about $1/(q-1)$ of the volume of the available factor space for small h; the exact proportion is in fact $h/[1-(1-h)^{q-1}]$.

A modification of the $\{q, 2\}$ lattice design for fitting a $\{q, 2\}$ polynomial, which spreads the experimental points out into the corners of the available factor space if there is only one small component, say the first, with $x_1 \leqq h$, is pictured in Fig. 4 for $q = 4$, and defined in general by taking observations on the responses η_i to the other pure components ($i > 1$), the responses η_{ij} to their 1 : 1 binary mixtures ($j > i > 1$), the responses η'_{ij} of the binary mixtures for which $x_1 = h$, $x_j = 1 - h$ ($j > 1$), and the response η'_1 of the mixture with $x_1 = \frac{1}{2}h$, $x_2 = x_3 = \ldots = x_q = (1 - \frac{1}{2}h)/(q-1)$. Then the coefficients β_i for

$i > 1$ and β_{ij} for $j > i > 1$ are still given by (4.1) and (4.2), and β_1 and the β_{1j} are easily found from the following equations obtained from (3.3_2):

$$h\beta_1 + h(1-h)\,\beta_{1j} = \eta'_{1j} - (1-h)\,\beta_j \qquad (j = 2, \quad . \; . \; . \quad , q), \tag{8.1}$$

$$\frac{1}{2}h\beta_1 + \frac{\frac{1}{2}h\left(1-\frac{1}{2}h\right)}{q-1} \sum_{2\leq j\leq q} \beta_{1j} = \eta'_1 - \frac{1-\frac{1}{2}h}{q-1} \sum_{2\leq j\leq q} \beta_j - \frac{\left(1-\frac{1}{2}h\right)^2}{(q-1)^2} \sum_{2\leq i<j\leq q} \beta_{ij}. \tag{8.2}$$

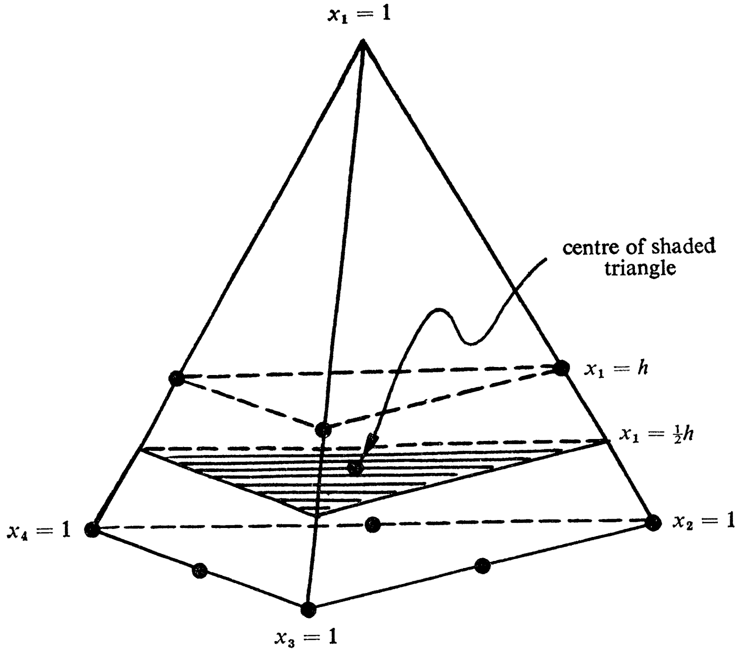

FIG. 4.—Modified {4, 2} design for the case of one small component.

The right sides may be regarded as known, and if we solve (8.1) for β_{1j} and substitute this in (8.2) we get an equation in β_1 only, whose solution substituted back into (8.1) then gives β_{1j}. Similar useful modifications of the $\{q, m\}$ lattice for $m > 2$ or for the case of more than one small component remain to be explored.

9. MATHEMATICAL APPENDIX

9.1. *Number of Points in a $\{q, m\}$ Lattice*

Let us term $\{q, m\}'$ lattice the result of multiplying by m the co-ordinates of the points of the $\{q, m\}$ lattice; it consists of the points $(x_1, x_2, \; . \; . \; . \; , x_q)$ with

$$x_i = 0, 1, \quad . \; . \; . \quad , m; \quad x_1 + x_2 + \quad . \; . \; . \quad + x_q = m. \tag{9.1.1}$$

To derive the number of points in a $\{q, m\}'$ lattice we may assign a code number consisting of zeros and ones to each point, the point $(x_1, x_2, \ldots, x_q)$ having the code number whose successive digits are x_1 ones, a zero, x_2 ones, a zero, $\ldots$, x_{q-1} ones, a zero, x_q ones. There is thus established a 1 : 1 correspondence between the lattice points and the $(m + q - 1)$-digit numbers consisting of $q - 1$ zeros and m ones. But the number of such code numbers is the number (2.1) of combinations of $m + q - 1$ things taken m at a time, since the code number is uniquely determined by the m positions we choose for the ones out of the $m + q - 1$ possible positions.

9.2. *Number of Coefficients in a $\{q, n\}$ Polynomial*

The number of coefficients in the $\{q, n\}$ polynomial written in the general form (3.1) is evidently

$$N_{q0} + N_{q1} + \ldots + N_{qn}, \tag{9.2.1}$$

where N_{qm} is the number in the homogeneous polynomial of degree m in the q variables. But the N_{qm} coefficients in the homogeneous polynomial can be placed in 1 : 1 correspondence with the points of the $\{q, m\}'$ lattice, the coefficient of $x_1^{p_1} x_2^{p_2} \ldots x_q^{p_q}$, where $p_1 + p_2 + \ldots + p_q = m$, corresponding to the lattice point with co-ordinates $(p_1, p_2, \ldots, p_q)$. Thus N_{qm} equals the number of points in the $\{q, m\}'$ lattice. Next we note that if the $N_{q+1,\, n}$ points of the $\{q + 1, n\}'$ lattice are divided into $n + 1$ sets $S_0, S_1, \ldots, S_n$, such that $x_{q+1} = i$ in S_i, then the first q of the co-ordinates $(x_1, x_2, \ldots, x_q, i)$ of the points in S_i satisfy the conditions (9.1.1) with $m = n - i$, and so the number of points in S_i is the same as that in a $\{q, n - i\}'$ lattice, namely $N_{q,\, n-i}$. This proves that the sum (9.2.1) equals $N_{q+1,\, n}$, that is, that the number of coefficients in the general form of the $\{q, n\}$ polynomial equals the number of points in a $\{q + 1, n\}$ lattice.

9.3. *Canonical Form of the $\{q, n\}$ Polynomial for $n = 1, 2, 3$*

The canonical form (3.3_1) of the $\{q, 1\}$ polynomial is obtained by replacing β_0 by $\beta_0 x_1 + \beta_0 x_2 + \ldots + \beta_0 x_q$. The canonical form (3.3_2) of the $\{q, 2\}$ polynomial is obtained by also replacing $x^2{}_i$ in $\beta_{ii} x^2{}_i$ by

$$x_i - \sum_{\substack{j=1 \\ j \neq i}}^{q} x_i x_j.$$

To obtain the canonical form (3.3_3) of the $\{q, 3\}$ polynomial we also replace $x^3{}_i$ by

$$x_i - \frac{1}{2} \sum_{\substack{j=1 \\ j \neq i}}^{q} [3x_i x_j + x_i x_j (x_i - x_j) - \sum_{\substack{k=1 \\ k \neq i, j}}^{q} x_i x_j x_k], \tag{9.3.1}$$

and substitute for $x_s{}^2 x_t$ (this takes care of $\beta_{iij} x^2{}_i x_j$ and $\beta_{ijj} x_i x^2{}_j$ for $i < j$)

$$\frac{1}{2} [x_s x_t + x_s x_t (x_s - x_t) - \sum_{\substack{k=1 \\ k \neq s, t}}^{q} x_s x_t x_k] \qquad (s \neq t). \tag{9.3.2}$$

The validity of the substitutions (9.3.1) and (9.3.2) is easily verified by factoring $x_i x_j$ out of the terms in square brackets in (9.3.1), and $x_s x_t$, in (9.3.2).

9.4. *Calculation of the Orthogonal Polynomials a, b, c*

As an example of the calculation we shall derive the form of b_1 given at the end of sec. 7. By the procedure stated above (7.2$_1$) we find

$$b_1 = x_1 - 2 \sum_{j=2}^{q} x_1 x_j + 3 \sum_{2 \leqq j < k \leqq q} x_1 x_j x_k$$

$$= x_1(1 - 2 \sum_{j=2}^{q} x_j + 3 \sum_{2 \leqq j < k \leqq q} x_j x_k).$$

If in this we first substitute

$$\sum_{2 \leqq j < k \leqq q} x_j x_k = \frac{1}{2} [(\sum_{j=2}^{q} x_j)^2 - \sum_{j=2}^{q} x^2{}_j],$$

and then

$$\sum_{j=2}^{q} x_j = 1 - x_1,$$

we get

$$b_1 = \frac{1}{2} x_1(3x^2{}_1 - 2x_1 + 1 - 3 \sum_{j=2}^{q} x^2{}_j), \tag{9.4.1}$$

which may then be written as in sec. 7.

9.5. *Calculation of the Maxima of a^2, b^2, c^2*

Let C denote any of the coefficients a, b or c, except c_{iij} or c_{ijj}. We may prove $C^2 \leqq 1$ by showing $-1 \leqq C \leqq 1$. This is elementary, except perhaps for b_i and b_{ij}, which are treated below. Since we know the value $C = 1$ is attained, this then proves max $C^2 = 1$.

To prove $-1 \leqq b_i \leqq 1$ let us take $i = 1$ and use the form (9.4.1) so

$$b_1 = \frac{1}{2} x_1(3x^2{}_1 - 2x_1 + 1 - 3z),$$

where

$$z = \sum_{j=2}^{q} x^2{}_j.$$

If at first we hold x_1 fixed, then z is the square of the radius from the origin of the space of $(x_2, \ldots, x_q)$ to the point $(x_2, \ldots, x_q)$ in the $(q - 2)$-dimensional simplex

$$x_j \geqq 0 \, (j = 2, \ldots, q), \quad \sum_{=2}^{q} x_j = 1 - x_1.$$

The maximum of z is thus the square of the radius to a corner of the simplex, namely $(1 - x_1)^2$. We shall need only the bounds $0 \leqq z \leqq (1 - x_1)^2$. Using $z \geqq 0$, we get $b_1 \leqq f(x_1)$, where

$$f(x_1) = \frac{1}{2} x_1(3x^2{}_1 - 2x_1 + 1).$$

The two roots of $f'(x_1)$ are imaginary, hence $f(x_1)$ takes on its maximum value at one of the endpoints of the interval $0 \leqq x_1 \leqq 1$. Thus $b_1 \leqq \max\ [f(0), f(1)] = 1$. Using

$z \leqq (1 - x_1)^2$, we get

$$b_1 \geqq \frac{1}{2} x_1[3x^2{}_1 - 2x_1 + 1 - 3(1 - x_1)^2] = x_1(2x_1 - 1) > -1,$$

the last inequality being proved by elementary methods. Likewise, $-1 \leqq b_i \leqq 1$.

To treat b_{ij} note that it is a function of x_i and x_j only, and that (x_i, x_j) must lie in the triangle

$$x_i \geqq 0,\ x_j \geqq 0,\ x_i + x_j \leqq 1. \tag{9.5.1}$$

If we equate to zero the first partial derivatives of b_{ij} with respect to x_i and x_j we obtain

$$x_j(6x_i + 3x_j - 2) = 0,\ x_i(3x_i + 6x_j - 2) = 0. \qquad (9.5.2.)$$

We shall investigate the behaviour of b_{ij} on the boundary of the triangle later. Inside the triangle the equations (9.5.2) can be satisfied only if $6x_i + 3x_j = 3x_i + 6x_j = 2$, that if, is $x_i = x_j = 2/9$. Then $b_{ij} = -32/243$, and this must be the minimum of b_{ij} since $b_{ij} \geqq 0$ on the boundary of the triangle (9.5.1). Since there exists only one solution of (9.5.2) inside the triangle, b_{ij} must attain its maximum on the boundary. On the part of the boundary where $b_{ij} > 0$, $x_i + x_j = 1$, hence $b_{ij} = 4x_ix_j = 4x_i(1 - x_i)$, and therefore max $b_{ij} = 1$. Thus $-1 \leqq b_{ij} \leqq 1$.

The maximum of

$$c^2{}_{iij} = \frac{81}{4} x^2{}_i x^2{}_j (3x_i - 1)^2$$

will be attained for values x_i, x_j satisfying $x_i + x_j = 1$, else x_j could be increased, holding x_i fixed, and $c^2{}_{iij}$ would increase. Under the restriction $x_i + x_j = 1$,

$$c_{iij} = \frac{9}{2} x_i(1 - x_i)(3x_i - 1). \tag{9.5.3}$$

Elementary methods show that the minimum of this cubic on the interval $0 \leqq x_i \leqq 1$ is attained at $x_i = (4 - \sqrt{7})/9$, where $c_{iij} > -1$, and the maximum at $x_i = (4 + \sqrt{7})/9$, where $c_{iij} = (10 + 7\sqrt{7})/27$. Hence max $c^2{}_{iij} = (10 + 7\sqrt{7})^2/729 = 1{\cdot}116$. The same bound holds for $c^2{}_{ijj}$.

9.6. *Statistic for Testing Several Correlated d_{ijk}*

Let us shift to a single-subscript notation for the r differences d_{ijk} discussed at the end of sec. 6, denoting them by $d_\nu(\nu = 1, \ . \ . \ . \ , r)$, and the covariance of d_ν and $d_{\nu'}$ by $u_{\nu\nu'}\sigma^2$. Then $u_{\nu\nu'}\sigma^2$ is given by (6.2) if $\nu = \nu'$, by (6.5) if the d_{ijk} corresponding to d_ν and $d_{\nu'}$ have two subscripts in common, by (6.4) if they have one subscript in common, and $u_{\nu\nu'} = 0$ if they have none in common. The numerator of the F-statistic is $\Sigma_{\nu\nu'} u^{\nu\nu'} d_\nu d_{\nu'}/r$, where $(u^{\nu\nu'})$ is the inverse of the matrix $(u_{\nu\nu'})$, and may be calculated as

$$\frac{1}{r}\left[\frac{|\,u_{\nu\nu'} + d_\nu d_{\nu'}\,|}{|\,u_{\nu\nu'}\,|} - 1\right]$$

without actually inverting the matrix but by merely evaluating two r^{th} order determinants (Rao, 1948, p. 60).

9.7. Estimation of the Regression Coefficients with Design (ii)

For design (ii) of sec. 5 we shall consider only the simple case where a $\{3, 2\}$ polynomial is fitted to the observations on a $\{3, m\}$ lattice with $m > 2$. Suppose the points of the lattice are ordered by a subscript ν, where $\nu = 1, 2, \ldots, (m+1)(m+2)/2$, the co-ordinates of the ν^{th} point being denoted by $(x_{1\nu}, x_{2\nu}, x_{3\nu})$, and the observed response there by y_ν. The method of least squares estimates the coefficients β as the values which minimize

$$\sum_\nu (\beta_1 x_{1\nu} + \beta_2 x_{2\nu} + \beta_3 x_{3\nu} + \beta_{23} x_{2\nu} x_{3\nu} + \beta_{13} x_{1\nu} x_{3\nu} + \beta_{12} x_{1\nu} x_{2\nu} - y_\nu)^2.$$

The normal equations are, in matrix form,

$$\begin{bmatrix} A & B & B & C & D & D \\ B & A & B & D & C & D \\ B & B & A & D & D & C \\ C & D & D & E & F & F \\ D & C & D & F & E & F \\ D & D & C & F & F & E \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_{23} \\ \beta_{13} \\ \beta_{12} \end{bmatrix} = \begin{bmatrix} \Sigma_\nu x_{1\nu} y_\nu \\ \Sigma_\nu x_{2\nu} y_\nu \\ \Sigma_\nu x_{3\nu} y_\nu \\ \Sigma_\nu x_{2\nu} x_{3\nu} y_\nu \\ \Sigma_\nu x_{1\nu} x_{3\nu} y_\nu \\ \Sigma_\nu x_{1\nu} x_{2\nu} y_\nu \end{bmatrix}, \tag{9.7.1}$$

where

$$A = \sum_\nu x^2_{1\nu}, \qquad B = \sum_\nu x_{1\nu} x_{2\nu}, \qquad C = \sum_\nu x_{1\nu} x_{2\nu} x_{3\nu},$$

$$D = \sum_\nu x^2_{1\nu} x_{2\nu}, \qquad E = \sum_\nu x^2_{1\nu} x^2_{2\nu}, \qquad F = \sum_\nu x^2_{1\nu} x_{2\nu} x_{3\nu},$$

and we have made use of the symmetry of the design in substituting

$$\Sigma_\nu x^2_{2\nu} = \Sigma_\nu x^2_{1\nu}, \; \Sigma_\nu x^2_{2\nu} x_{3\nu} = \Sigma_\nu x^2_{1\nu} x_{2\nu}, \text{ etc.}$$

The matrix in (9.7.1) is of the composite form

$$\begin{bmatrix} \mathbf{U} & \mathbf{V} \\ \mathbf{V} & \mathbf{W} \end{bmatrix}, \tag{9.7.2}$$

where each of the matrices **U**, **V**, **W** is of the type

$$\begin{bmatrix} f & g & g \\ g & f & g \\ g & g & f \end{bmatrix}. \tag{9.7.3}$$

Multiplication of matrices of the type (9.7.3) is commutative. Products and differences are of the same type. If $\Delta = (f - g)(f + 2g)$ and $\Delta \neq 0$, then (9.7.3) has an inverse, which is of the same type, with diagonal elements $(f + g)/\Delta$, and off-diagonal elements

$-g/\Delta$. Products, differences, and inverses are thus easily computed, requiring the evaluation of only two matrix elements. The following four matrices are defined by operations of the type just discussed:

$$\mathbf{M} = (\mathbf{UW} - \mathbf{V}^2)^{-1}, \ \mathbf{X} = \mathbf{WM}, \ \mathbf{Y} = -\mathbf{VM}, \ \mathbf{Z} = \mathbf{UM}.$$

The inverse of (9.7.2), from which the least squares estimates of the β and their covariance matrix are easily calculated, is

$$\begin{bmatrix} \mathbf{X} & \mathbf{Y} \\ \mathbf{Y} & \mathbf{Z} \end{bmatrix}.$$

REFERENCES

CLARINGBOLD, P. J. (1955), "Use of the simplex design in the study of joint action of related hormones", *Biometrics*, **11**, 174–185.
RAO, C. R. (1948), "Tests of significance in multivariate analysis", *Biometrika*, **35**, 58–79.