CHAPTER 21

FACTORS AT TWO LEVELS ONLY

21.1 The 2^n design

21.1.1 Factorial treatment arrangements with all factors at two levels, commonly called (despite the remark of § 19.3.1) 2^n factorial designs, where n here (and throughout Chapters 21 and 22) indicates the number of factors, are a special case of the designs discussed in Chapter 19. As such, they will not involve any essentially new principles, and, in fact, any experiment of this type could be analysed by the general methods of Chapter 19.

21.1.2 In § 19.11, however, we have already made acquaintance with special computational methods which are available when one or more factors are at two levels. These methods can be extended and rationalized to such an extent that the general method of analysis seems to be something quite different, the purpose being, of course, the greater speed of computation made possible by the special methods.

21.1.3 An additional reason for the separate chapter devoted to these designs is that a special notation suitable for algebraic treatment has become associated with them, and it is necessary to make acquaintance with this notation and the type of algebraic manipulation that it makes possible.

21.1.4 The particular property of 2^n designs which is the underlying reason for this special treatment is that each main effect and interaction has 1 D.F. and therefore represents a single comparison among the treatments. To make these comparisons we form linear functions of the treatment totals (or means) as explained in § 11.2, and from these can be calculated the S.S.'s for each main effect and interaction. Since main effects and interactions are orthogonal (§ 19.9), it follows that in a 2^n design the main effects and interactions constitute a complete orthogonal set (§ 11.3.5) and that their linear functions must obey the condition [11.6] that products of corresponding coefficients sum to zero.

21.2 Notation

21.2.1 The two levels of a factor A could be denoted as a_0 and a_1 as before, but it is convenient to avoid suffices by regarding the two levels as an upper level and a lower level, the former being denoted by the symbol a and the latter by the absence of this symbol. The two levels may then be thought of as "A" and "No A" respectively, e.g. phosphate and no phosphate. Thus in a 2^3 experiment with factors A, B, and C the symbol ac would stand for that treatment combination containing A and C, but no B (cf. § 19.2.3). This enables 7 of the 8 treatment combinations to be written down. The eighth, "absence" of all three factors, commonly called the "control", is given the symbol unity instead of zero (as might have been expected), but the figure 1 is conventionally bracketed, presumably because on a typewriter the lower-case L and the figure 1 are identical; otherwise there would be confusion in experiments where lime (usually given the symbol L) was a factor, and fertilizer experiments with factors $LPKN$, the four basic nutrient elements, are very common. The eight treatment combinations in a 2^3 experiment with factors A, B, C could therefore be written as

$$(1) \quad a \quad b \quad ab \quad c \quad ac \quad bc \quad abc$$

The convenience of this notation and the reason for the particular order of writing down the combinations will emerge presently.

21.2.2 The lower levels of the factors may actually be zero applications of a quantitative factor, in which case the treatment (1) is actually a control treatment. However, the lower level need not be, and often is not, a zero application, in which case the two levels would correspond to "some" and "more", or "a little" and "a lot". Nevertheless it is still convenient to use the above notation and to think of the levels as "nil" and "some". The same applies even to a qualitative factor, such as two varieties. One variety is arbitrarily given the symbol v (say) and the other is denoted by the absence of the symbol.

21.2.3 While the above is convenient theoretically, it might lead to mistakes in laying out the design. Consequently, for guidance of field-workers it may be desirable to use on the field plans symbols such as v_1 and v_2 for two varieties, and similarly for a quantitative factor where the lower level is not zero.

21.2.4 The suggested retention of lower-case letters for treatment combinations (§ 19.2.1) must be rigidly observed in 2^n designs, where the use of lower-case letters and capitals is the means of distinguishing between treatment combinations and treatment effects. Without this convention confusion would arise, as will be realized from the following examples for a 2^3 design with factors A, B, C:

a = treatment combination with A only—no B or C, i.e. the combination with the upper level of A and the lower levels of B and C
A = main effect of factor A
ab = treatment combination with the upper levels of A and B and the lower level of C
AB = interaction of factors A and B.

It is not uncommon for capitals to be used for treatment combinations on field plans, but it is better to adhere to the convention.

21.3 Definitions of main effects and interactions in 2^n designs

21.3.1 In a 2^2 design with factors A and B, the comparison between the upper and lower levels of A is made by the difference of treatments with and without A, viz.

$$\{ab + a\} - \{b + (1)\}. \qquad [21.1]$$

If the treatment symbols represent treatment totals, this is the total response to A, which we may denote as $[A]$. This total response depends on both the number of replications (r) as well as n, and so it is necessary to convert $[A]$ to a per-plot basis with conversion to standard units also where necessary. Disregarding the latter, we define [21.1] divided by $2r$ as the **mean response** to A. Since

$$\frac{1}{2r}[A] = \frac{1}{2r}\{ab + a\} - \frac{1}{2r}\{b + (1)\}$$
$$= \text{(Mean of plots with } A) - \text{(Mean of plots without } A),$$

we see the justification for this. Similarly the mean response to B is $\frac{1}{2r}[B]$, where

$$[B] = \{ab + b\} - \{a + (1)\}. \qquad [21.2]$$

If we now consider a 2^3 design with an additional factor C, we define the mean responses to the three factors A, B, and C respectively as

$$\frac{1}{4r}\{abc + ab + ac + a\} - \frac{1}{4r}\{bc + b + c + (1)\} = \frac{1}{4r}[A]$$
$$\frac{1}{4r}\{abc + ab + bc + b\} - \frac{1}{4r}\{ac + a + c + (1)\} = \frac{1}{4r}[B]$$
$$\frac{1}{4r}\{abc + ac + bc + c\} - \frac{1}{4r}\{ab + a + b + (1)\} = \frac{1}{4r}[C].$$

These represent

(Mean of plots with A) — (Mean of plots without A)

etc., as before.

21.3.2 From the above it is easy to see the generalization to a 2^n design. The total response to any factor is the difference between the total yield of all treatment combinations with the upper level of the factor and the total yield of all combinations with the lower level of the factor. To get the mean response per plot this is divided by $2^{n-1}r$ = half the total number of plots in the experiment.

21.3.3 Yates originally defined the main effect of a factor as the total response to that factor, but soon changed this to the mean response. Actually, it is not uncommon still to refer to undivided linear functions such as [21.1] or [21.2] as main effects, but clearly some divisor is necessary for presentation of results. Equally clearly, the mean response is ideal for the latter purpose. Nevertheless, as a measure of the main effect any fraction of the total response which is independent of r and n would do, and likewise any multiple of the total response provides the same test of significance (cf. § 11.3.9). The author believes that there is good reason for differentiating between mean response and main effect. Applying the definition of § 19.6.2 to a factor at two levels, we would write the main effects of these levels as (say) $y_{10} - \bar{y}$ and $y_{20} - \bar{y}$. These two quantities are, of course, equal in magnitude but of opposite sign, so that

$$y_{20} - \bar{y} = -(y_{10} - \bar{y}) = \tfrac{1}{2}(y_{20} - y_{10}).$$

The "main effects" of the two levels reduce effectively to one since their sum is zero, and it is therefore seen to be reasonable to refer to $\frac{1}{2}(y_{20} - y_{10}) = \frac{1}{2}$(mean response) as the main effect of the factor—in other words to use the total number of plots in dividing the total response instead of half the total number of plots. This is also the divisor for the S.S. of the main effect of the factor if we square the total response, as is usual (cf. § 17.17.13 and Example 19.3, Note H). There are, in fact, distinct advantages in defining the main effect as "Total response divided by total number of plots", and there is no reason why these advantages should be sacrificed simply because of the fact that the mean response is required for the presentation of results. *We therefore define the main effect of a factor A as $\frac{1}{2^n r}[A]$ and symbolize this as A. The mean response to A, under this definition, is then $2A$.*

21.3.4 In a 2^2 experiment with factors A and B, the linear function

$$ab - a - b + (1)$$

is a measure of the interaction AB (cf. § 19.11.3). This can be seen by arranging the function as

$$\{ab - a\} - \{b - (1)\},$$

i.e. the difference between the response to B in the presence and absence of A, or as

$$\{ab - b\} - \{a - (1)\}$$

i.e. the difference between the response to A in the presence and absence of B (cf. § 19.7.3). The mean interaction (with treatment symbols representing treatment totals, as before) would be $\frac{1}{r}\{ab - a - b + (1)\}$, but it is theoretically and practically desirable to divide the linear function by the same quantity as used in the definition of a main effect, viz. 2^2r. For one thing, there is the obvious advantage that the main effects and interaction all have the same variance $\left(= \frac{1}{2^2 r}\sigma^2\text{, where } \sigma^2 = \text{error variance}\right)$, which saves time in tests of significance. There are also theoretical advantages, as will be explained later.

21.3.5 In a 2^3 experiment with factors A, B, and C we have measures of the interaction AB in the presence and absence of C, viz.

$$\frac{1}{r}\{ab - a - b + (1)\}$$

and

$$\frac{1}{r}\{abc - ac - bc + c\}.$$

The mean interaction is therefore

$$\frac{1}{2r}\{abc - ac - bc + c + ab - a - b + (1)\},$$

or in general

$$\frac{1}{2^{n-2}r}\,[AB],$$

where $[AB]$ represents the undivided linear function of treatment totals. Again it is convenient to have the divisor the same as for main effects and so *we define the interaction AB as* $\frac{1}{2^n r}[AB]$, *symbolized as AB*. Similarly for any first order interaction, Yates defines AB as $\frac{1}{2^{n-1} r}[AB]$, retaining for the same reasons as discussed above the same divisor as for the mean response (or main effect as he defines it). However, it will be noticed that his definition does not give the mean interaction in the same way as his definition of a main effect gives the mean response.

21.3.6 A measure of the triple interaction ABC is given by

$$\begin{aligned}[ABC] &= \{abc - ac - bc + c\} - \{ab - a - b + (1)\} \\ &= \text{difference in } [AB] \text{ in the presence and absence of } C.\end{aligned}$$

The same linear function is obtained by considering the difference in $[AC]$ in the presence and absence of B or the difference in $[BC]$ in the presence and absence of A (cf. § 19.18.2). For the same reasons as before we take as definition of the triple interaction

$$ABC = \frac{1}{2^n r}[ABC].$$

The generalization to interactions of higher order is obvious.

21.3.7 The only difference between Yates's definitions for main effects and interactions and those proposed here is the division of the linear function by the total number of plots instead of half the total number of plots; i.e. Yates's definitions are double those adopted here. The reason for changing Yates's definitions is convenience in theoretical work. Nevertheless, as already noted, the mean response (Yates's "main effect") will be used in presenting results and it has become conventional to present "main effects and interactions" according to Yates's definitions. There is no reason to depart from this convention, but we shall prefer to allude to "mean responses and interactions". Here and here only (for lack of an alternative term) will there be any departure from the definitions of interactions adopted in §§ 21.3.5 and 21.3.6. Since there is no intermingling of theoretical work and presentation of results, this will cause no difficulty.

21.4 Methods of obtaining symbolic linear functions for main effects and interactions in 2^n designs

21.4.1 Main effects and interactions have been defined as certain linear functions of treatment totals (in which the coefficients are all $+1$ or -1) divided by the total number of plots in the experiment. These linear functions can be written down in accordance with principles followed in § 21.3, but it is convenient to have a mechanical method.

21.4.2 The 3 D.F. for treatments in a 2^2 design with factors A and B may be represented by the following linear functions, written in vector form in the manner of Chapter 11:

	(1)	a	b	ab
$A =$	-1	1	-1	1
$B =$	-1	-1	1	1
$AB =$	1	-1	-1	1

It will be immediately noticed that the multiplication rule of corresponding coefficients given in § 19.14.15 could have been used to obtain the coefficients of AB, and we have symbolically

$$A \times B = AB.$$

(N.B.: Throughout § 21.4 the square-bracket notation is dropped for convenience and capital letters stand for undivided linear functions.) This rule may be applied to obtain any required linear functions very easily. Those for main effects can be written down directly; for example, that for the main effect of A is given by allotting treatment combinations a positive or negative sign accordingly as they do or do not contain the letter a. The linear function for any interaction is then obtainable by multiplying corresponding coefficients in the linear functions of the main effects of the factors concerned in the interaction. Thus to obtain the linear function for ABC in a 2^3 design we work on $A \times B \times C = ABC$ and write down the functions for A, B, and C, viz.

	(1)	a	b	ab	c	ac	bc	abc
$A =$	-1	1	-1	1	-1	1	-1	1
$B =$	-1	-1	1	1	-1	-1	1	1
$C =$	-1	-1	-1	-1	1	1	1	1,

whence the linear function for ABC is obtained by multiplying all three corresponding coefficients, column by column, as

$$ABC = -1 \quad 1 \quad 1 \quad -1 \quad 1 \quad -1 \quad -1 \quad 1\,.$$

Alternatively, it could have been obtained in two stages, viz. $A \times B = AB$, $AB \times C = ABC$, in case the linear function for AB was also required.

21.4.3 The symbolic products of linear functions of treatment effects in 2^n designs possess a reversible property which does not apply in general. This is, for example, $AB \times A = B$, as may be confirmed for the 2^2 case set out in § 21.4.2. On the other hand in §§ 19.14.14 and 19.14.15, although $P' \times K' = P'K'$, we do not find $P' \times P'K' = K'$. If $AB \times A = B$, then symbolically $A \times A =$ unity, where unity is represented by a linear function with all coefficients $+1$, denoted by the symbol I and called the "identity". Thus, when we symbolically multiply two effects in a 2^n design, common letters disappear, e.g. $ABC \times AB = C$.

21.4.4 Algebraically it is convenient to introduce a similar system of symbolic multiplication for the treatment combinations. Thus we may consider $ab \times c = abc$, $abc \times ab = c$, etc., though we shall have use here only for the former type of product. The symbolic multiplication of treatment combinations seems even more abstract than that of treatment effects, because at least in the latter case there are linear functions to "multiply", but its introduction is justified by its usefulness. Notice that $a \times (1) = a$, showing one reason why the control is symbolized as unity.

21.4.5 We may make use of the above as follows: Consider a single factor A; then symbolically we have

$$\begin{array}{rcc} & (1) & a \\ \hline I = & 1 & 1 \\ A = & -1 & 1 \end{array} \qquad [21.3]$$

or $I_a = (a + 1)$, $A = (a - 1)$. Similarly, for a factor B we have $I_b = (b + 1)$, $B = (b - 1)$. Combining these, we obtain the effects for a 2^2 design, viz.

$$\begin{aligned} I &= I_a \times I_b = (a + 1)(b + 1) = ab + b + a + (1) \\ A &= A \times I_b = (a - 1)(b + 1) = ab - b + a - (1) \\ B &= I_a \times B = (a + 1)(b - 1) = ab + b - a - (1) \\ AB &= A \times B = (a - 1)(b - 1) = ab - b - a + (1) \end{aligned}$$

It will be seen that where a letter appears in the effect concerned a negative sign occurs in the bracket containing that letter in the factorized linear function. This gives a rule for writing down any required linear function. For example, in a 2^3 design

$$\begin{aligned} ABC &= (a - 1)(b - 1)(c - 1) \\ &= abc - ab - ac - bc + a + b + c - (1), \end{aligned}$$

confirming the result of § 21.4.2. This rule is particularly useful for writing down single linear functions.

21.4.6 As exemplified by [21.3], it is also algebraically convenient to include the identity as a treatment effect. For the 2^2 case we have:

$$\begin{array}{rcccc} & (1) & a & b & ab \\ \hline I = & 1 & 1 & 1 & 1 \\ A = & -1 & 1 & -1 & 1 \\ B = & -1 & -1 & 1 & 1 \\ AB = & 1 & -1 & -1 & 1 \end{array} \qquad [21.4]$$

To obtain the main effects and interaction the vectors for A, B, and AB are applied to the vector of treatment totals and the results are divided by the total number of plots in the experiment. If we do the same for I we get the grand total divided by the total number of plots, i.e. the general mean. Thus I represents the fourth degree of freedom for treatments, that taken up by the correction factor.

21.4.7 An arrangement of vectors one under the other so that the coefficients form columns as well as rows is called a **matrix**. A matrix such as [21.4] which when applied to a vector of treatment totals (or means) produces (after suitable division) the main effects and interactions may be termed an **interaction matrix.** We may call $[(1) \ \ a \ \ b \ \ ab]$ the treatments vector, and the corresponding vector of capitals the effects vector. The coefficients of these vectors, being symbolic, may perhaps be better termed "elements".

21.4.8 It is possible to construct the interaction matrix for any number of factors in a systematic way. Let us give the matrix of [21.3] (the interaction matrix for $n = 1$) the symbol M_1, i.e.

$$M_1 = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.$$

It will be seen that M_2, the interaction matrix for $n = 2$, as given in [21.4] is made up as follows:

$$M_2 = \begin{bmatrix} M_1 & M_1 \\ -M_1 & M_1 \end{bmatrix}$$

The effects and treatments vectors must, however, correspond to the particular order shown in [21.4]. If A is the first factor, then the treatments vector $[(1) \quad a]$ is extended by the addition of two more terms which are products of the original elements (in order) with b. This gives $[(1) \quad a \quad b \quad ab]$, and the same, of course, holds for the effects vector.

21.4.9 The addition of further factors is governed by similar rules. Thus, if a third factor C, also at 2 levels, is introduced, making a 2^3 experiment, the interaction matrix is

$$M_3 = \begin{bmatrix} M_2 & M_2 \\ -M_2 & M_2 \end{bmatrix},$$

and the effects and treatments vectors are both extended in a similar manner, the latter by the addition of the former treatment combinations (in the same order) multiplied by c. We thus have the following complete orthogonal set for a 2^3 design:

	(1)	a	b	ab	c	ac	bc	abc	
$I =$	1	1	1	1	1	1	1	1	
$A =$	−1	1	−1	1	−1	1	−1	1	
$B =$	−1	−1	1	1	−1	−1	1	1	
$AB =$	1	−1	−1	1	1	−1	−1	1	
$C =$	−1	−1	−1	−1	1	1	1	1	[21.5]
$AC =$	1	−1	1	−1	−1	1	−1	1	
$BC =$	1	1	−1	−1	−1	−1	1	1	
$ABC =$	−1	1	1	−1	1	−1	−1	1	

The linear functions previously obtained by the other methods discussed above may now be checked. This last method is particularly suitable if *all* linear functions are required (cf. § 21.4.5). *The order of treatment combinations in the treatments vector and of effects in the effects vector as given by the above rules is conventional and is known as the standard order.*

21.5 Statistical analysis of 2^n design

21.5.1 The analysis varies little from that ordinarily performed on a design with a complete orthogonal set of predetermined comparisons (cf. §§ 11.3.10, 11.5.3). The total effects ([A], [B], [AB], etc.) may be evaluated as linear functions of the treatment totals in the manner symbolically exemplified for a 2^3 design by [21.5]. These total effects have variance $2^n r\sigma^2$ and may be tested either by the t-test or by the F-test. An advantage suggested earlier for the F-test is the orthogonality check supplied by the exact summation of the S.S.'s $\frac{[A]^2}{2^n r}, \frac{[B]^2}{2^n r}$, etc. Here, however, with the variances of all effects the same, it is more convenient to calculate least significant values as

$$\sqrt{2^n r\,(\text{Error M.S.})} \times t. \qquad [21.6]$$

Tests of significance can then be made by direct comparison of the total effects with these. The orthogonality check should still be made, but without finding the S.S. for each individual effect, i.e. by alternatively calculating the Treatments S.S. as

$$\frac{1}{2^n r}\{[A]^2 + [B]^2 + [AB]^2 + \cdots\}. \qquad [21.7]$$

21.5.2 *Yates's systematic method for computing the total effects.* Especially when n is fairly large, the computation of the total effects as just described becomes laborious and liable to error. Fortunately, a systematic method due to Yates enables these to be computed in a more felicitous manner. The method will be illustrated with the use of the data from the 2^3 experiment with factors denoted by P, G, and S analysed in Example 21.1.

The treatment totals are arranged in the *standard order* (§ 21.4.9), each factor being introduced in turn followed by all combinations of itself and the combinations previously introduced. The computation is then set out as follows:

Treatment combination	Treatment totals	(1)	(2)	(3)	Effect
(1)	158	376	810	1896	I
p	218	434	1086	226	P
g	207	489	80	166	G
pg	227	597	146	−76	PG
s	199	60	58	276	S
ps	290	20	108	66	PS
gs	271	91	−40	50	GS
pgs	326	55	−36	4	PGS

Column (1) is formed from the treatment totals by adding them in pairs to form the first 4 values and then finding the differences of these pairs, the *upper number being subtracted from the lower* in each case. Thus $376 = 158 + 218$ and $55 = 326 - 271$. Column (2) is found from column (1) in an identical manner and the process is carried out as many times as there are factors, i.e. in general n times, and here 3 times. The final column gives the total effects [P], [G], [PG], etc. in the standard order, headed by the grand total of yields.

21.5.3 For comparison and confirmation the direct calculation of [P] and [PG] is set out below:

$$[P] = -158 + 218 - 207 + 227 - 199 + 290 - 271 + 326 = 226$$

$[PG] = \quad 158 - 218 - 207 + 227 + 199 - 290 - 271 + 326 = -76$

The signs are obtained from the matrix of [21.5], viz. the second and fourth rows respectively.

21.5.4 The student is warned that even with Yates's method it is fatally easy to make arithmetical errors, particularly with the subtractions and when the quantities subtracted are not both positive. Consequently, it is advisable to check the work. The appearance of the G.T. at the top of the final column is a check, but only a partial one, since it checks only figures near the top of the table where mistakes are least likely to occur. The check provided by [21.7] is not absolute, since it does not guard against the attachment of an incorrect sign to any addition or subtraction. A third check, also not infallible, is as follows:

$$\text{Total of final column} = 2^n \text{ (last total in column of treatment totals)}$$

In the Example above

$$1896 + 226 + 166 - 76 + 276 + 66 + 50 + 4 = 2608 = 8 \times 326.$$

Since very litte extra work is involved, all three checks should be made. Should the failure of either of the last two checks necessitate recomputation, it is convenient (for $n > 3$) to be able to locate the first column in which an error occurs. A method of doing this will be given in conjunction with Example 21.1 (Note D).

*21.5.5 In experiments with a large number of factors it may be necessary or desirable to use higher order interactions for estimation of error (cf. § 19.21). If, to save computation, the Error S.S. (including higher order interactions) is to be obtained by subtraction, it will not be necessary to compute all the treatment effects. In this case the method of directly calculating the required effects as linear functions of the treatment totals may be preferred to Yates's method, which will generate unwanted information. Nevertheless, Yates's method may still be preferred for its convenience. Even if it is, however, it may not be desired to compute the Treatments S.S. from treatment totals, since the saving in computation comes from calculating a Treatments S.S. from main effects and low order interactions only. In these circumstances the check given by [21.7] is not available.

21.5.6 *Yates's method is applicable only with factors at two levels.* The student is warned against attempting to apply it to the general types of examples discussed in Chapter 19.

Example 21.1 (Data from Saunders and Rayner, *Statistical methods with special reference to field experiments*) In a 2^3 experiment on maize conducted as a randomized blocks design the factors were:

Phosphate (P): phosphate (p) v. no phosphate
Greenmanure (G): cowpeas ploughed in (g) v. no greenmanure
Spacing in rows (S): 18 in. spacing (s) v. 36 in. spacing.

There were five replications, the plan and yields in lb. per $\frac{1}{100}$ morgen plot being shown below. Analyse the data.

Block 1

(1) 33	*pg* 47	*g* 46	*s* 45
p 48	*ps* 57	*pgs* 65	*gs* 55

Block 2

p 46	*ps* 59	*pg* 46	*gs* 55
g 43	*pgs* 66	*s* 42	(1) 31

Block 3

p 48	*s* 44	*gs* 57	*g* 46
pgs 70	*pg* 45	(1) 34	*ps* 61

Block 4

pg 44	*pgs* 64	*gs* 54	*g* 37
(1) 31	*s* 36	*p* 37	*ps* 58

Block 5

ps 55	*gs* 50	*s* 32	*pg* 45
p 39	*g* 35	(1) 29	*pgs* 61

Computation sheet

Treatments	Blocks 1	2	3	4	5	Treatment totals
(1)	33	31	34	31	29	158
p	48	46	48	37	39	218
g	46	43	46	37	35	207
pg	47	46	45	44	45	227
s	45	42	44	36	32	199
ps	57	59	61	58	55	290
gs	55	55	57	54	50	271
pgs	65	66	70	64	61	326
Block totals	396	388	405	361	346	1,896

C.F. = 89,870·4

Blocks S.S. = 90,177·8 − 89,870·4 = 307·4

Total S.S. = 94,506·0 − 89,870·4 = 4,635·6

Treatments S.S. = 94,056·8 − 89,870·4 = 4,186·4

Analysis of variance

Source	D.F.	S.S.	M.S.
Blocks	4	307·4	
Treatments	7	4,186·4	
Error	28	141·8	5·064
Total	39	4,635·6	

S.E. of single yield = 2·250

$$\text{C.V.} = \frac{2{\cdot}250}{1896} \times 40 \times 100 = 4{\cdot}75\% \text{ (A)}$$

Yates's method **(B)**

Treatment	Treatment totals	(1)	(2)	(3)		Effect
(1)	158	376	810	1896 (D)		G.T.
p	218	434	1086	226**		*P*
g	207	489	80	166**		*G*
pg	227	597	146	−76**	(H)	*PG*
s	199	60	58	276**		*S*
ps	290	20	108	66**		*PS*
gs	271	91	−40	50**		*GS*
pgs	326	55	−36	4		*PGS*
	1896 (C)			2608 (D)		

$$\text{Treatments S.S. (check)} = 4{,}186{\cdot}4 \text{ (D) (E)}$$

$$\text{Estimate of variance of total effect} = 40 \times 5{\cdot}064^{\text{(F)}} = 202{\cdot}56$$

$$\text{S.E.} = 14{\cdot}23$$

$$\text{Least significant values for total effects} = 14{\cdot}23 \times t \text{ (28 D.F.) (G)}$$

$$= 14{\cdot}23 \times \begin{cases} 2{\cdot}048 \\ 2{\cdot}763 \end{cases}$$

$$= 29{\cdot}1 \ (5\%)$$
$$39{\cdot}3 \ (1\%).$$

$$\text{S.E. of single treatment total} = \sqrt{5 \times 5{\cdot}064} = 5{\cdot}032$$

$$\text{Least significant differences (2 treatment totals)} = \sqrt{10 \times 5{\cdot}064} \times t$$

$$= 7{\cdot}116 \times \begin{cases} 2{\cdot}048 \\ 2{\cdot}763 \end{cases}$$

$$= 14{\cdot}57 \ (5\%)$$
$$19{\cdot}66 \ (1\%)$$

Conversion factors:

(1) Treatment totals to means in bags per morgen (1 bag = 200 lb.) $= \frac{100}{5} \times \frac{1}{200} = 0{\cdot}1$

(2) Total effects to mean responses and interactions (Yates's definition) in bags per morgen

$$= \frac{100}{200} \times \frac{1}{20}^{\text{(I)}} = 0{\cdot}025.$$

Presentation of results

TREATMENT MEANS IN BAGS PER MORGEN

(1)	15·8
p	21·8
g	20·7
pg	22·7
s	19·9
ps	29·0
gs	27·1
pgs	32·6
Mean	23·7

S.E. ±0·503

L.S.D.'s { 1·46 (5%) / 1·97 (1%)

MEAN RESPONSES AND INTERACTIONS (J) IN BAGS PER MORGEN

P	5·65**
G	4·15**
S	6·90**
PG	−1·90**
PS	1·65**
GS	1·25**
PGS	0·10

S.E. ±0·356 (K)

Least significant values { 0·72 (5%) / 0·98 (1%) (L)

Highly significant average responses were obtained to both the phosphate and greenmanure applications of 5·65 and 4·15 bags per morgen respectively. In addition a highly significant average difference of 6·90 bags per morgen was recorded in favour of the 18 in. spacing over the 36 in. spacing.(M) These average effects were not, however, consistent over the levels of the other factors, as the accompanying interaction tables (of means yields in bags per morgen) reveal.

Although the response to phosphate is highly significant, both in the presence and absence of greenmanure, the response in the presence of greenmanure (3·75 bags/morg.) is less than in the absence of greenmanure (7·55 bags/morg.), the difference being highly significant. (V) Similarly, the response to phosphate is highly significantly greater with the 18 in. spacing (7·30 bags/morg.) than with the 36 in. spacing (4·00 bags/morg.), both responses being themselves highly significant.

The response to greenmanure is highly significant at both phosphate levels, but the response is highly significantly greater in the absence of phosphate (6·05 bags/morg.) than in the presence of phosphate (2·25 bags/morg.). Responses at both spacings are highly significant, but there is a highly significant difference between that at the 18 in. spacing (5·40 bags/morg.) and that at the 36 in. spacing (2·90 bags/morg.).

The 18 in. spacing is highly significantly better than the 36 in. spacing whether phosphate or greenmanure is present or absent, but there is a highly significant enhancement of the difference when phosphate is present (8·55 compared with 5·25 bags/morg.) and when greenmanure is present (8·15 compared with 5·65 bags/morg.).

There is no significant evidence that any of the responses discussed in the last three paragraphs varies over the levels of the third factor. (W) The recommended treatment combination would be the application of phosphate and greenmanure at the 18 in. spacing. The response to the combined fertilizer-manure application at this spacing is estimated as

Phosphate × greenmanure (**N**)

	No phosphate	Phosphate	Mean	Response to phosphate
No greenmanure	17·85 (**O**)	25·40	21·62	7·55**
Greenmanure	23·90	27·65	25·78	3·75**
Mean	20·88	26·52	23·70	5·65 (**P**)
Response to greenmanure	6·05**	2·25**	4·15 (**Q**)	— (**R**)

Phosphate × spacing

	No phosphate	Phosphate	Mean	Response to phosphate
18 in.	23·50	30·80	27·15	7·30**
36 in.	18·25	22·25	20·25	4·00**
Mean	20·88	26·52	23·70	5·65
Difference in favour of 18 in.	5·25**	8·55**	6·90	—

Greenmanure × spacing

	18 in.	36 in.	Mean	Difference in favour of 18 in.
No greenmanure	24·45	18·80	21·62	5·65**
Greenmanure	29·85	21·70	25·78	8·15**
Mean	27·15	20·25	23·70	6·90
Response to greenmanure	5·40**	2·90**	4·15	—

S.E.'s for above tables:

Body of table ±0·356 (**S**)
Marginal means ±0·252 (**T**)
Marginal responses ±0·503 (**U**)

Least significant values for marginal responses $\begin{cases} 1\cdot03\ (5\%) \\ 1\cdot39\ (1\%) \end{cases}$

12·7 ± 0·71 bags/morg. and the gain due to the narrower spacing with both phosphate and greenmanure applied is 9·8 ± 0·62 bags/morg. (**X**) The combined response to this treatment over no fertilizer and no manure at the wider spacing is 16·7 ± 0·62 bags/morg. (**X**)

Notes on the computations

(A) Down to this point the analysis follows that of an ordinary randomized blocks design.

(B) The method is explained in § 21.5.2. In small examples like the present one, the works ing of Yates's method could be done alongside the treatment totals in the data table. This would save recopying the treatment totals, but would not be possible for designs with n large except on large computing sheets. If desired, Yates's method could precede the analysi- of variance.

(C) A check here saves the annoyance of afterwards discovering a copying error, usually after reaching the top of the final column.

(D) Partial checks. See § 21.5.4.

Yates gives the following check on the working of any column:
Add up

(a) the numbers appearing 1st, 3rd, 5th, etc. (i.e. in odd order) in the top half of the column;
(b) the numbers appearing in even order in the top half of the column;
(c) the numbers appearing in odd order in the lower half of the column;
(d) the numbers appearing in even order in the lower half of the column.

For the k^{th} column, let these totals be a_k, b_k, c_k, d_k. Do the same to the previous column in order to obtain a_{k-1}, b_{k-1}, c_{k-1}, d_{k-1} (for checking the first column, the column of treatment totals counts as the previous column). Then check that

$$a_k + b_k = a_{k-1} + b_{k-1} + c_{k-1} + d_{k-1}$$
$$c_k + d_k = -a_{k-1} + b_{k-1} - c_{k-1} + d_{k-1}.$$

In this analysis, for example, suppose we wish to check column (2):

$$a_2 = 810 + 80 = 890 \qquad a_1 = 376 + 489 = 865$$
$$b_2 = 1086 + 146 = 1232 \qquad b_1 = 434 + 597 = 1031$$
$$c_2 = 58 - 40 = 18 \qquad c_1 = 60 + 91 = 151$$
$$d_2 = 108 - 36 = 72 \qquad d_1 = 20 + 55 = 75$$

Checks
$$\left.\begin{array}{r} a_2 + b_2 = 2122 \\ a_1 + b_1 + c_1 + d_1 = 2122 \end{array}\right\} \qquad \left.\begin{array}{r} c_2 + d_2 = 90 \\ -a_1 + b_1 - c_1 + d_1 = 90 \end{array}\right\}$$

(E) Formula [21.7]: $\frac{1}{40}(226^2 + 166^2 + \cdots + 4^2)$.

(F) Each total effect is a linear function involving every plot in the experiment with coefficient ± 1. Hence variance equals $40\sigma^2$ (40 = total number of plots in the experiment).

(G) Formula [21.6].

(H) All total effects except *PGS* are greater than the least significant value at the 1% level. This means that all main effects and interactions except *PGS* are significant at the 1% level.

Alternatively, the check by means of the Treatments S.S. and the tests of significance could be done as follows, but it is a little longer:

Analysis of variance

Source	D.F.	S.S.		M.S.	F
Blocks	4		307·4		
Treatments					
P	1	1276·9		1276·9	252·2**
G	1	688·9		688·9	136·0**
S	1	1904·4		1904·4	376·1**
PG	1	144·4		144·4	28·5**
PS	1	108·9		108·9	21·5**
GS	1	62·5		62·5	12·3**
PGS	1	0·4		0·4	0·1 N.S.
	7		4,186·4		
Error	28		141·8	5·064	
Total	39		4,635·6		

(I) The conversion factor here is that applicable for a single plot divided by $\frac{1}{2}$(total number of plots).

(J) Yates's definition.

(K) $0{\cdot}356 = 14{\cdot}23 \times 0{\cdot}025$.

(L) $0{\cdot}72 = 29{\cdot}1 \times 0{\cdot}025$
$0{\cdot}98 = 39{\cdot}3 \times 0{\cdot}025$.

(M) The narrow spacing was decided upon as the higher level of the factor (see introduction to this example). Thus a positive response means that the narrow spacing is the better.

(N) The principles followed in presenting interaction tables have already been discussed in § 19.19.4. Many workers set out all two-factor tables whether the relevant two-factor interactions are significant or not. Others may give the tables only for those factor-pairs whose interaction is significant. Had the three-factor interaction proved significant, a three-factor table such as that employed in Example 19.3 would have been desirable. An alternative mode of presentation will be given later (§ 21.7.6).

(O) $17{\cdot}85 = \frac{1}{2}(15{\cdot}8 + 19{\cdot}9)$ from table of treatment means in bags per morgen.

(P) This is, of course, the mean response to phosphate in bags per morgen.

(Q) Similarly this is the mean response to greenmanure.

(R) If this figure were entered it would be

$$3{\cdot}75 - 7{\cdot}55 = -3{\cdot}80,$$
or
$$2{\cdot}25 - 6{\cdot}05 = -3{\cdot}80.$$

This is the difference of responses to P in the presence and absence of G, or *vice versa*, and is a measure of the interaction *PG*. The interaction according to Yates's definition is half this ($-1{\cdot}90$).

(S) $0{\cdot}356 = \frac{1}{\sqrt{2}}(0{\cdot}503)$, since the variance of the mean of two treatment means is half that of a single mean. Here this variance is $\frac{\sigma^2}{10}$ (σ^2 = error variance). This is the same as the variance of a mean response, viz. $\frac{40\sigma^2}{(20)^2} = \frac{\sigma^2}{10}$.

(T) A marginal mean is a mean of 4 treatment means and so its S.E. is half that of a single mean.

(U) A marginal, or differential, response is the difference of two means in the body of the table and therefore has S.E. $= \sqrt{2}(0\cdot356) = 0\cdot503$. Otherwise, its variance $= \frac{20\sigma^2}{10^2}$ (cf. Note S) $= \frac{\sigma^2}{5}$ = variance of a single treatment mean.

(V) Because the interaction *PG* is significant at the 1% level.

(W) The interaction *PGS* is not significant.

(X) These statements will be explained in § 21.7.5.

21.6 Statistical model for treatment effects in a 2^n design

21.6.1 Consider first the case of two factors only, *A* and *B*, each at two levels denoted by suffices 1 and 2. In accordance with Model [19.3], the mean yields of treatment means arranged in an interaction table and of the marginal means are as follows:

Table 21.1: Mean yields in a 2 × 2 experiment in terms of Model [19.3]

	b_1	b_2	Means
a_1	$\mu + a_1 + \beta_1 + (a\beta)_{11}$	$\mu + a_1 + \beta_2 + (a\beta)_{12}$	$\mu + a_1$
a_2	$\mu + a_2 + \beta_1 + (a\beta)_{21}$	$\mu + a_2 + \beta_2 + (a\beta)_{22}$	$\mu + a_2$
Means	$\mu + \beta_1$	$\mu + \beta_2$	μ

It will be remembered that we choose $\sum_i a_i = 0$, and $\sum_j \beta_j = 0$. Here this means $a_1 = -a_2$, and $\beta_1 = -\beta_2$. Also that we choose $\sum_j (a\beta)_{ij} = \sum_i (a\beta)_{ij} = 0$. This means here that $(a\beta)_{11} = -(a\beta)_{12}$, $(a\beta)_{11} = -(a\beta)_{21}$, $(a\beta)_{12} = -(a\beta)_{22}$, and $(a\beta)_{21} = -(a\beta)_{22}$, or, in short, $(a\beta)_{11} = -(a\beta)_{12} = -(a\beta)_{21} = (a\beta)_{22}$. From this it is clear that we may replace the 8 parameters of Table 21.1 (apart from μ) by 3 linearly independent parameters, corresponding to the 3 D.F. for treatments, as follows:

$$a = -a_1 = a_2$$
$$\beta = -\beta_1 = \beta_2$$
$$a\beta = (a\beta)_{11} = -(a\beta)_{12} = -(a\beta)_{21} = (a\beta)_{22}.$$

The symbol $a\beta$ represents a distinct parameter and is *not* the product of a and β.

Table 21.1 may now be rewritten as:

Table 21.2: Revised model of true mean yields in a 2 × 2 experiment

	No *B*	*B*	Means
No *A*	$\mu - a - \beta + a\beta$	$\mu - a + \beta - a\beta$	$\mu - a$
A	$\mu + a - \beta - a\beta$	$\mu + a + \beta + a\beta$	$\mu + a$
Means	$\mu - \beta$	$\mu + \beta$	μ

The parameters a, β, and $a\beta$ represent the *true* main effect of *A*, main effect of *B*, and interaction *AB*, corresponding to the definitions of main effects and interactions adopted in §§ 21.3.3 and 21.3.5. The *true* mean response to *A* is $2a$ (cf. § 21.3.3). Adopting, as is conventional, the symbols *A*, *B*, and *AB* to stand for *estimates* of α, β, and $\alpha\beta$, we have in the notation of Chapter 19

$$A = y_{20} - \bar{y} = -(y_{10} - \bar{y}) = \tfrac{1}{2}(y_{20} - y_{10}) = \text{main effect of } A, \quad [21.8]$$

$$2A = y_{20} - y_{10} = \text{mean response to } A,$$

and $AB = y_{11} - y_{10} - y_{01} + \bar{y}$ (cf. § 19.8.6) [21.9]

$$= y_{11} - \tfrac{1}{2}(y_{11} + y_{12}) - \tfrac{1}{2}(y_{11} + y_{21}) + \tfrac{1}{4}(y_{11} + y_{12} + y_{21} + y_{22})$$

$$= \tfrac{1}{4}(y_{22} - y_{12} - y_{21} + y_{22})$$

$$= \frac{1}{4r}(Y_{11} - Y_{12} - Y_{21} + Y_{22})$$

$$= \frac{1}{4r}[AB]$$

in accord with the definition of § 21.3.5.

21.6.2 The model for treatment means given in Table 21.2 may be extended by noticing (i) that the sign of the main-effect component of a factor is negative for the lower level of the factor, positive for the upper level, and (ii) that the sign of $\alpha\beta$ is the product of the signs of α and β. Introducing a third factor C with true main effect γ, we may write down the following model for the true means of the various treatments:

Table 21.3: Model for the true mean yields of treatments in a 2^3 experiment

Treatment combination	True means
(1)	$\mu - \alpha - \beta + \alpha\beta - \gamma + \alpha\gamma + \beta\gamma - \alpha\beta\gamma$
a	$\mu + \alpha - \beta - \alpha\beta - \gamma - \alpha\gamma + \beta\gamma + \alpha\beta\gamma$
b	$\mu - \alpha + \beta - \alpha\beta - \gamma + \alpha\gamma - \beta\gamma + \alpha\beta\gamma$
ab	$\mu + \alpha + \beta + \alpha\beta - \gamma - \alpha\gamma - \beta\gamma - \alpha\beta\gamma$
c	$\mu - \alpha - \beta + \alpha\beta + \gamma - \alpha\gamma - \beta\gamma + \alpha\beta\gamma$
ac	$\mu + \alpha - \beta - \alpha\beta + \gamma + \alpha\gamma - \beta\gamma - \alpha\beta\gamma$
bc	$\mu - \alpha + \beta - \alpha\beta + \gamma - \alpha\gamma + \beta\gamma - \alpha\beta\gamma$
abc	$\mu + \alpha + \beta + \alpha\beta + \gamma + \alpha\gamma + \beta\gamma + \alpha\beta\gamma$

The sign of $\alpha\beta\gamma$ is the product of the signs of α, β, and γ, or the product of the signs of $\alpha\beta$ and γ, etc. It will be seen that *the signs in the above expressions are, in order, the signs of the columns of the matrix of* [*21.5*].

21.6.3 Extension of the model to cover more than three factors follows the above rules. If it is considered that a certain high order interaction is likely to be negligible, it can be ignored by pooling its S.S. with error, which is equivalent to dropping this interaction term from the model, i.e. equating it to zero.

21.6.4 Yates's definition of main effects corresponds to the adoption of $\mu - \tfrac{1}{2}\alpha$ and $\mu + \tfrac{1}{2}\alpha$, for example, as the true means for the lower and upper levels of A, respectively. This, however, is not usual in theoretical work (cf. § 9.7.7) since the fractions are rather a nuisance. For example, in Table 21.3 a factor $\tfrac{1}{2}$ would have to be introduced for all components except μ. Under Yates's definition the true response to A is α (not 2α), which is thus the same as his true main effect.

21.6.5 From the models any true response can be represented algebraically. Thus from Table 21.2,

$$\text{True response to } A \text{ in absence of } B = a - (1) = 2(\alpha - \alpha\beta)$$
$$\text{True response to } A \text{ in presence of } B = ab - b = 2(\alpha + \alpha\beta)$$

(Only if $\alpha\beta = 0$ are these equal and the factors additive.)

$$\text{True response to } A \text{ and } B \text{ together} = ab - (1) = 2(\alpha + \beta)$$

(Notice that this does not contain $\alpha\beta$, but that this does not mean that the factors are additive.)

21.7 Expression of treatment means in terms of main effects and interactions

21.7.1 Since from [21.8] $A = y_{20} - \bar{y} = -(y_{10} - \bar{y})$ and since $\bar{y} = m =$ estimate of μ, we have

$$y_{10} = m - A,$$
$$y_{20} = m + A.$$

Similarly,

$$y_{01} = m - B,$$
$$y_{02} = m + B.$$

Also, since

$$\begin{aligned} AB &= y_{11} - y_{10} - y_{01} + \bar{y} \quad \text{(from [21.9])} \\ &= y_{11} - (y_{10} - \bar{y}) - (y_{01} - \bar{y}) - \bar{y} \\ &= y_{11} + A + B - m, \end{aligned}$$

we have

$$y_{11} = m - A - B + AB.$$

Similarly,

$$\begin{aligned} y_{12} &= m - A + B - AB, \\ y_{21} &= m + A - B - AB, \\ y_{22} &= m + A + B + AB. \end{aligned}$$

The following table of treatment means in terms of main effects and interaction may therefore be drawn up:

Table 21.4: Treatment means in a 2 × 2 experiment in terms of main effects and interaction

	No B	B	Means
No A	$m - A - B + AB$	$m - A + B - AB$	$m - A$
A	$m + A - B - AB$	$m + A + B + AB$	$m + A$
Means	$m - B$	$m + B$	m

The correspondence between Tables 21.2 and 21.4 will be noted. Table 21.2 refers to true means, Table 21.4 to observed means, i.e. Table 21.2 gives the true means exactly in terms of the model, while Table 21.4 gives the observed means exactly in terms of the main effects and interaction. The addition of error components (ϵ_{ij}) as in [19.3] would be necessary to express the observed means exactly in terms of the hypothetical parameters, but no corresponding

e_{ij} are necessary to express them in terms of the estimates of these parameters, i.e. Table 21.4 is an exact representation of observed treatment means in terms of estimated parameters (main effects and interaction). This is always possible when the model for treatment effects corresponds to a complete orthogonal subdivision of the Treatments S.S., all available D.F. having been used for the estimation of parameters so that no deviations of observed means from estimates in terms of the model are possible (cf. § 17.16.2).

21.7.2 The extension of the above to the case of 3 or more factors follows §§ 21.6.2 and 21.6.3. In fact, Table 21.3 with A, B, C, . . . replacing α, β, γ, . . . etc. becomes an exact expression of treatment means in terms of main effects and interactions. The signs of the terms in these expressions may be determined by the rule given in § 21.6.2, or are available from the columns of the appropriate interaction matrix.

21.7.3 To exemplify numerically, we turn to Example 21.1 and work out the treatment mean for *pgs*. The matter is complicated by the adoption of the conventional method of presenting results, since if we were to work from mean responses and interactions, a division by 2 would be necessary. Rather we work from the total effects and convert directly, e.g.

$$\begin{aligned} pgs &= (1896 + 226 + 166 - 76 + 276 + 66 + 50 + 4)\,\frac{100}{200} \times \frac{1}{40} \\ &= (2608)\,\frac{1}{80} \\ &= 32\cdot 6 \text{ bags/morg.} \end{aligned}$$

Of course, it is pointless to work out treatment means in this long-winded way when they are otherwise available, but the method will find a useful application in Chapter 22. Yates has presented a systematic method of calculating treatment means from main effects and interactions which is rather like the method of § 21.5.2 in reverse. This is not likely to be of much practical use, however.

21.7.4 A further usefulness of the expressions discussed in this section is that they permit easy calculation of certain estimated responses which may be required. Thus following § 21.6.5, we may calculate the following estimated responses for Example 21.1:

Mean response to P in absence of G

$$= p - (1)\dagger = 2(P - PG) = 5\cdot 65 + 1\cdot 90 = 7\cdot 55 \text{ bags/morg.}$$

(N.B.: The multiplication by 2 is in effect already performed by the adoption of Yates's definition in the presentation of results.)

Mean response to P in the presence of G

$$= pg - g\dagger = 2(P + PG) = 5\cdot 65 - 1\cdot 90 = 3\cdot 75 \text{ bags/morg.}$$

And so on.

It is true that the above results are readily available from the two-way interaction tables, but in Chapter 22 we shall come across cases where they are not. See also § 21.7.6 below.

† These symbols here represent means over both levels of S, in line with the description "*mean* response" (over both levels of S).

21.7.5 However, even in the present example the method has proved useful. In the final paragraph of the summary of results, the following estimated responses are quoted:

(i) Response to combined fertiliser-manure application at the 18 in. spacing = 12·7 bags./morg.

(ii) Gain due to narrower spacing with both phosphate and greenmanure applied = 9·8 bags/morg.

(iii) Combined response to fertilizer + manure at 18 in. spacing over no fertilizer or manure at 36 in. spacing = 16·7 bags/morg.

These responses have been estimated *on the assumption that* $PGS = 0$, as is justified by the fact that it is small and non-significant. They are therefore graduated responses calculated from graduated means (cf. § 17.17.18) and are in general slightly different from the estimates derived directly from the treatment means. Although the difference is trivial here, the method is nevertheless of importance as a general principle.

(i)
$$\begin{aligned} pgs &= m + P + G + PG + S + PS + GS \\ s &= m - P - G + PG + S - PS - GS \\ pgs - s &= 2(P + G + PS + GS) \\ &= 5{\cdot}65 + 4{\cdot}15 + 1{\cdot}65 + 1{\cdot}25 \\ &= 12{\cdot}70 \end{aligned}$$

(Here the result is the same as derived from the treatment means, viz. 32·6—19·9, because PGS would have cancelled even if included.)

(ii)
$$\begin{aligned} pg &= m + P + G + PG - S - PS - GS \\ pgs - pg &= 2(S + PS + GS) \\ &= 6{\cdot}90 + 1{\cdot}65 + 1{\cdot}25 \\ &= 9{\cdot}80 \end{aligned}$$

(iii)
$$\begin{aligned} (1) &= m - P - G + PG - S + PS + GS \\ pgs - (1) &= 2(P + G + S) \\ &= 5{\cdot}65 + 4{\cdot}15 + 6{\cdot}90 \\ &= 16{\cdot}70 \end{aligned}$$

It must be noted that this response is $2(P + G + S + PGS)$ if PGS is not assumed zero, i.e. it is not in general the simple sum of the separate mean responses to the factors.

As regards the S.E.'s of these quantities, it must be remembered that the main effects and interactions are orthogonal and that their variances are therefore additive. Consequently, $\text{Var.}\{2(P + G + PS + GS)\} = 4 \times \frac{\sigma^2}{10}$, since $\frac{\sigma^2}{10}$ is the variance of a mean response or interaction (Yates's definition), as seen in Note S of Example 21.1. This is $\frac{2\sigma^2}{5}$, here equal to the variance of the difference of two treatment means, as must be correct in view of the bracketed note just above. The S.E. is therefore $\sqrt{2} \times 0{\cdot}503 = 2 \times 0{\cdot}356 = 0{\cdot}71$. However, for the response to pg at the 18 in. spacing, because of the assumption that $PGS = 0$, the required variance is that of $2(S + PS + GS)$, viz. $\frac{3\sigma^2}{10}$. The S.E. is therefore $\sqrt{3} \times 0{\cdot}356 = 0{\cdot}62$. The same applies to the third response estimated.

*21.7.6 *Alternative method of presenting results.* A method of presenting the results of a 2^n design, which is equivalent to the presentation of all the two-factor interaction tables, has the advantage of a neat tabular form. In Example 21.1 it would be applied as follows:

	Mean response	Differential responses					
		Phosphate		Greenmanure		Spacing	
		Absent	Present	Absent	Present	36 in.	18 in.
Phosphate	5·65	—	—	7·75	3·75	4·00	7·30
Greenmanure	4·15	6·05	2·25	—	—	2·90	6·05
18 in. spacing over 36 in.	6·90	5·25	8·55	5·65	8·15	—	—

The differential responses can be calculated without forming the interaction tables by the method shown in § 21.7.4. As far as the S.E.'s of the differential responses are concerned, they are always $\sqrt{2}$ times the S.E. for a mean response in view of the orthogonality of main effects and interactions. Significance can be indicated by asterisks in the usual way, or by printing in bold type (which saves space). The table is completed by addition of the general mean, which then permits the construction of any two-factor interaction table.

EXERCISES

21.1 In a variety-fertilizer trial with sugar beet, the yields of unwashed roots in lb. per $\frac{1}{60}$ acre plot are given below. Analyse, presenting results in tons per acre (1 ton = 2,240 lb.). The treatment symbols are as follows:

A, B } Varieties of sugar beet

P = superphosphate
K = potash
O = no fertilizer

(Rothamsted Report, 1932, page 200)

AP 542	*AK* 554	*BPK* 501	*APK* 505	*AO* 524	*BK* 484	*BP* 506	*BO* 512
BK 455	*AK* 454	*APK* 497	*AO* 473	*BP* 503	*BPK* 514	*BO* 465	*AP* 506
AO 517	*AK* 568	*APK* 564	*BP* 458	*AP* 493	*BPK* 465	*BK* 452	*BO* 429
BP 497	*BPK* 496	*BO* 432	*AO* 492	*APK* 514	*AK* 467	*BK* 352	*AP* 449

(Hint: First convert to standard 2^n notation.)

21.2 The following are the field plan and yields of dry hay (in lb.) in a veld fertilizer experiment. The treatments were all combinations of:

Superphosphate: 400 lb./morg. (p) or 200 lb./morg.
Nitramoncal: 200 lb./morg. (n) or 100 lb./morg.
Chloride of potash: 100 lb./morg. (k) or nil.
Lime: 1,000 lb./morg. (l) or nil.

pk 32	*n* 44	*p* 50	*k* 47	*pl* 161	*nk* 185	*pnkl* 96	*nl* 38
l 35	*pn* 127	*nkl* 34	*kl* 36	*pnk* 79	*pkl* 38	(1) 56	*pnl* 35

nkl 39	*pnk* 152	*p* 52	(1) 45	*pnkl* 57	*pk* 77	*pn* 62	*pl* 108
nl 155	*kl* 66	*nk* 70	*l* 187	*k* 50	*pkl* 107	*n* 53	*pnl* 39

n 83	*pkl* 168	*kl* 41	*nkl* 158	*nk* 106	*pl* 38	*l* 115	*p* 42
pk 42	*pn* 54	*pnkl* 32	*nl* 62	(1) 52	*pnl* 93	*pnk* 77	*k* 81

The plot size (net) is 20 yds. × 12½ yds. Analyse the data evaluating all main effects and interactions and presenting the results in tons per morgen (1 ton = 2,000 lb., 1 morgen = 10,244 sq. yds.)

(Adapted from data supplied by D. P. le Roux, Highveld Region, Department of Agricultural Technical Services)

A First Course in Biometry for Agriculture Students

ERRATA

Title page and following page: The date of publication should be 1969, not 1967.

Page 6 (4 lines from foot of page): For "preceeding" read "preceding".

Page 11: The two diagrams in § 1.9.6. should be labelled (1) and (2).

Page 62: Formula 5.3 should read:

$$\bar{x} = \sum_j x_j \left(\frac{n_j}{n}\right)$$

Page 344, second line of Note K: For $\Sigma\xi_1 n$ * read $\Sigma\xi_1 n_1$*.

Page 376, Table 17.2: The last M.S. should be $s^2_{y,x}$.

Page 408: Value of x for Treatment D, Block 4, should be 62.2, not 62.3.

Page 486, line 13: For "D.F. of s^2" read "D.F. of $s_a{}^2$".

Page 507, end of first line of Note B: For "works" read "work-"; end of line 4 of same paragraph: "analysi-" to read "analysis".

Page 537: Data acknowledgement at end of Exercise 22.2 should read "Department of Agricultural Technical Services".

Page 570, line 11 of § 25.2.1: For "hervesting" read "harvesting".

Page 577, Table 25.6: Heading of last column should read "Estimated C.V.".

A FIRST COURSE IN BIOMETRY FOR AGRICULTURE STUDENTS.

A.A. RAYNER.

The following errata in addition to those on the printed slip, have been noticed during one year's use of the book as a class text:-

Page 115: In Table 7.1 the entry for N = 4, x = 3 should be $\frac{4}{16}$ ($= \frac{1}{4}$) not $\frac{1}{8}$.

Page 119: In Formula 7.2 the last term should be $x^3/3!$, not $x^2/3!$.

Page 197: Note H: For 1025.3/64 read 1052.3/64.

Page 248: (4 lines under the analysis of variance table): For 225.2 read 225.5

Page 293: (6 lines from foot of page): For "than" read "that".

Page 299: (first line of Example 15.5): For "Example 9.4" read "Example 15.4".

Page 300: The second P(10) should be P(11).

Page 308: Example 15.10: Some rounding errors occur in the last 6 entries in the χ^2 column:

For	0.17	read	0.18
	0.05		0.06
	6.54		6.56
	3.08		3.10

Page 309: Line 2: The P value for Heterogeneity should read: "$0.98 > P > 0.95$".

Page 315: Note H: For "souce" read "source".

Page 340: Note B: Delete the second and third lines of this note and replace by the following: "The C.F. for the S.P. (101.3256) must lie between the C.F.'s for the two S.S.'s (256.2560 and 40.0649). Also the uncorrected S.P. (103.5921) usually (i.e. when the variate-values are all positive and $\bar{x}_1$ and $\bar{x}_2$ are are reasonably different) lies between the two uncorrected S.S.'s (262.5634 and 40.9593), as here. Gross errors ......".

Page 342: The entries in the last five lines of the table have in the printing got out of line with the columns higher up. The "0" in the ξ_1 line should be under the "141" in the n_1^* line, etc.

Page 367: (lines 10, 12 and 13): For 0.25919 read 0.25906.
(line 10): For 7751.29 read 7755.29.
(line 13): For 44.887 read 44.896

Page 394: Figure 17.6: x-axis should be calibrated 0, 50, 100, 150, 200, 250.

Page 411: (Example 18.2): In the table Linear effect of phosphate, the totals for X_{i0} and Y_{i0} should be interchanged, viz. 1442.5 for X_{i0}, and 1047.5 for Y_{i0}.

Page 504: Line 16: For "litte" read "little".

Page 534: Just above analysis of variance table: The calculation for [NK]' should read "$-18 + 150 - 148 = -16$".

Page 541: Formula 23.7: The denominator of the second term in the curly brackets should be $r(r - 1)(t - 1)$.

I am grateful to the many sharp-eyed students this year who detected errors. I should also be grateful to any reader who detects any further errors.

A.A. RAYNER.
29/10/1969.