CHAPTER 22

AN INTRODUCTION TO CONFOUNDING

22.1 The problem of block size

22.1.1 It has already been noted (§ 13.13.2) that a block size of 16 plots is looked upon as a reasonable maximum in field experiments, since the variability of plots within large blocks may result in inadequate control of error. Also in experiments where natural groupings are used as blocks (e.g. litters in small-animal experiments), there are obvious limitations on the maximum block size (§ 13.11.1). Where the number of treatments is large, therefore, an ordinary randomized blocks design may be either unsatisfactory or impossible, according to the type of experiment.

22.1.2 The problem of large treatment numbers is particularly acute in factorial experimentation. Even with factors at the minimum of two levels, the number of treatments with only 4 factors is already at the desirable maximum for a field experiment in randomized blocks. In exploratory experiments 6, 7, or 8 factors are quite commonly required, but in order to conduct such an experiment efficiently it is necessary to break through this barrier imposed by limitation of block size.

22.1.3 In general, the problem has been solved by the introduction of **incomplete block designs.** In these designs, as indicated by the name, each block does not contain a complete replication, but only a prescribed fraction of the treatments. Usually, but not always, the incomplete blocks can be grouped to form a complete replication, so that, if, for example, there are four blocks per replication, each block contains a quarter of the treatments. The particular manner in which the treatments are allocated to a block is prescribed by the design. This allocation is not usually the same from replication to replication.

22.2 Confounding

22.2.1 In any incomplete block design treatments and blocks are non-orthogonal, and there is said to be "confounding" (or confusing) of treatment and block effects—that is to say, they are not "capable of direct and separate estimation without any entanglement" (§ 13.6.3). The degree of entanglement encountered due to accidental non-orthogonality may be such as to necessitate an extremely complex type of analysis or else some form of rough analysis such as an approximate analysis or an analysis with certain data discarded.

In an incomplete block design there is premeditated non-orthogonality of such a nature that *any required disentanglement of treatment and block effects can be achieved without too great a departure from a standard analysis of variance procedure.* Otherwise such designs would not be useful in practice owing to the complicated analysis.

22.2.2 In factorial experiments in incomplete block designs the premeditated non-orthogonality is usually arranged by a device known as **confounding,** this term now being used in a specific sense to be explained below, as opposed to the wider sense in which it was used in § 22.2.1. In brief, the device consists in identifying in any replication one or more treatment D.F. with differences between blocks within that replication. In this way the mutual entanglement of block and treatment effects is restricted to certain treatment D.F. only, the remainder being still orthogonal to blocks and therefore unaffected. The treatment degrees of freedom identified with block differences are said to be confounded with blocks or confounded with block effects or differences, and information is lost partly or wholly on such D.F. This is the price to be paid for the smaller block size and probable reduced error, but the price may be negligible if the confounding is restricted, for example, to high order interactions, in which there may well be very little interest (cf. § 19.21). Designs employing the above device are known as **confounded factorial designs.**

22.2.3 The objective in confounding may therefore be stated as better control of error through smaller block size at the expense of the sacrifice of information on certain treatment effects (usually high order interactions). There is also the loss of some D.F. from error due to the greater number of D.F. devoted to local control in virtue of the increased number of blocks.

22.2.4 For the rest of this chapter we shall consider only the simplest possible case, viz. when each replication (or complete block) is divided into two blocks, a single treatment D.F. being confounded with the block difference within every replication.

22.3 A simple illustration of confounding

22.3.1 Consider a single replication of a 2^2 design with factors A and B in which it has been prescribed that there shall be two blocks, one block to contain the treatments (1) and ab and the other a and b, and that randomization of treatments to plots is to be performed separately within each block (§ 13.1.4). Blocks of this sort are sometimes called block pairs or sub-blocks, but it is only in the randomized blocks design that a block necessarily contains a complete replication, and so these terms are not really necessary.

22.3.2 Suppose that the following is an actual arrangement:

Block 1	Block 2
ab	a
(1)	b

In accordance with Table 21.2, and with γ_1 and γ_2 representing block effects

($\gamma_1 + \gamma_2 = 0$), we have the following expected yields for the treatments on these plots:

$$\begin{aligned}(1) &= \mu + \gamma_1 - \alpha - \beta + \alpha\beta \\ a &= \mu + \gamma_2 + \alpha - \beta - \alpha\beta \\ b &= \mu + \gamma_2 - \alpha + \beta - \alpha\beta \\ ab &= \mu + \gamma_1 + \alpha + \beta + \alpha\beta\end{aligned}$$

In terms of this model we have, applying the linear functions given in [21.4],

$$\begin{aligned}I &= 4\mu + 2(\gamma_1 + \gamma_2) = 4\mu \\ [A] &= 4\alpha \\ [B] &= 4\beta \\ [AB] &= 4\alpha\beta + 2(\gamma_1 - \gamma_2),\end{aligned} \qquad [22.1]$$

from which it is apparent that, whereas estimates of μ, α, and β will be clear of block effects, the estimate of $\alpha\beta$ will be confounded with block effects. The same is clear from an examination of the block totals in terms of the model, viz.

$$\left.\begin{aligned}B_1 &= (1) + ab = 2\mu + 2\gamma_1 + 2\alpha\beta \\ B_2 &= a + b = 2\mu + 2\gamma_2 - 2\alpha\beta,\end{aligned}\right\} \qquad [22.2]$$

whence

$$B_1 - B_2 = 2(\gamma_1 - \gamma_2) + 4\alpha\beta. \qquad [22.3]$$

Thus an estimate of the block difference will not be affected by the two main effects, but will be affected by the interaction.

22.3.3 The situation revealed by the above, then, is that two of the treatments D.F., corresponding to the two main effects, are orthogonal to blocks and are thus **unconfounded.** The third D.F. for treatments, however, is non-orthogonal to blocks, and we say that **the interaction AB is confounded.** This can be determined by inspection since

$$[AB] = \quad (1) - a - b + ab = \{(1) + ab\} - \{a + b\},$$

whereas

$$[A] = -(1) + a - b + ab = \{ab - (1)\} + \{a - b\},$$

and

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

In the case of each main effect the total effect is derived from two differences of plots within blocks, each such difference being obviously free of block effects; the total interaction, however, is the same as the block difference (see also [22.1] and [22.3]). In other words, the main effects are such that their linear functions have one positive sign and one negative sign applied to the treatments in any one block (represented by the curly brackets), thus causing a cancellation of block effects, whereas for $[AB]$ the signs are the same for any one block and no cancellation occurs.

22.3.4 If the block contents were arranged as follows

(1)
b

ab
a

,

it is clear that A would be confounded, B and AB unconfounded.

22.4 Complete and partial confounding

22.4.1 Confounding is said to be complete if the allocation of treatments between the two blocks of a replication is kept the same for all replications. There are then two types of block in the experiment, the one containing a particular half of the treatments and the other containing the other half. For example, if in every replication of a 2^2 design the block contents were kept the same as those of § 22.3.2, we would have a **completely confounded design** and **the interaction AB would be completely confounded.** This means that virtually all information on this interaction would be lost since it would be identified with the block difference in every replication. In the normal equations (cf. § 13.6.9) what happens is that the equations derived for the confounded interaction and for block effects are not independent and therefore cannot be solved. (Equations [22.1] and [22.2] show how this comes about. If we form the difference of the two equations [22.2] we get [22.3] which is identical with [22.1], so that, instead of having two simultaneous equations in two unknowns, viz. $\gamma_1 - \gamma_2$ and $\alpha\beta$, we have only a single equation, which cannot give a unique solution.) The reason why the loss of information is not absolutely complete will be discussed in § 22.9.

22.4.2 If, on the other hand, the block contents are varied from replication to replication, so that the treatment effects confounded are not the same in the different replications, we have what is called **partial confounding.** In this situation the partially confounded treatment effects are non-orthogonal to block effects in the replications in which they are confounded, but are orthogonal to the block effects in the other replications. *Hence they can be quite simply estimated from the replications in which they are not confounded,* and the loss of information on the confounded effects is only partial.

22.4.3 Complete confounding should be resorted to only when there is no interest whatsoever in the effect so confounded or if the effect is known, or can be assumed, to be negligible. Partial confounding offers a way of spreading the loss of information over more than one set of effects and is obviously valuable when there is some hesitation about complete confounding.

22.4.4 According to Yates: "If the gain in precision resulting from confounding is sufficiently great, even the partially confounded interactions may be more accurately determined than would be the case if the experiment were not confounded." Because of partial confounding the variance of an interaction estimated from only a fraction of the replications would be a greater fraction of the error variance than if it were estimated over all replications; yet, if the error variance is itself reduced sufficiently because of the smaller block size compared with an unconfounded design, the variance of the partially confounded interaction might actually be less than if there were no confounding. In addition, there can be no loss of information on unconfounded effects, and any reduction of the error variance can result only in increased precision on these effects.

22.5 Simple examples of complete confounding in 2^n designs

22.5.1 Complete confounding of the interaction of a 2^2 design as exemplified in § 22.3.2 is extremely unlikely to be used in a straightforward design because a block size of 4 plots is very small anyway and because information on a first order interaction is not lightly discarded.

*The arrangement considered in § 22.3.2 could, however, be of value in a split-plot design. Suppose it is desired to introduce two additional factors, A and B, each at two levels into an experiment already in progress (§ 20.9.2). If it is possible to divide each existing plot into not more than two sub-plots, then it can be arranged that half the plots of each treatment are split as | (1) *ab* | and half as | *a* *b* |. The two sub-plot main effects would then be orthogonal to whole-plot treatments. Similarly, if the whole-plot design is in randomized blocks, each block should contain equal numbers of the two types of whole-plot, so that the interaction AB is orthogonal to blocks. The loss of information on this interaction will be discussed in § 22.10.4.

22.5.2 Before considering more realistic examples of confounding, we first extend the findings of § 22.3.3 and state a general rule for determining the block contents in a replication divided into two blocks.

> *The treatment combinations corresponding to positive signs in the linear function of the treatment effect to be confounded are placed in one block and those corresponding to negative signs in the other.*

The 1 D.F. corresponding to the difference of the two blocks is then identical with the single treatment effect confounded, and by virtue of the orthogonality of the treatment effects all other treatment effects are unconfounded. The linear functions of unconfounded treatment effects contain an equal number of positive and negative signs corresponding to the treatment combinations of any one block. Conversely, if we know the block contents of any replication divided into two blocks, all that need be done to find which treatment effect is confounded is to find that effect for which all the treatment combinations of the same sign correspond to the contents of one of the blocks.

22.5.3 2^3 *design in blocks of* 4 *plots.* For factors A, B, C the second order interaction ABC might be confounded and the block size reduced to 4 plots. Contra-indications are that it might not be desired to sacrifice ABC and that the block size without confounding is not excessive. Inspection of the linear function for ABC in [21.5] shows that treatments (1), *ab*, *ac*, and *bc* have negative signs and treatments *a*, *b*, *c*, and *abc* have positive signs. The block contents with ABC confounded would therefore be:

(1)
ab
ac
bc

a
b
c
abc

With this arrangement the 3 main effects and the 3 first order interactions are unconfounded. For example, the linear function for AB may be written

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

which shows that block effects must cancel out.

22.5.4 2^4 *design in blocks of* 8 *plots and* 2^5 *design in blocks of* 16 *plots.* The natural candidates for confounding are the interactions of highest order in each case. The linear function corresponding to this interaction may be obtained by any of the methods explained in § 21.4 and the block contents determined according to the rule. The confounding will be worth while in a 2^4 design, since the block size would otherwise be large; a 2^5 design could hardly be considered for a field experiment without confounding.

22.6 Randomization and field plan of completely confounded design

22.6.1 The blocks are arranged in contiguous pairs, each pair making up a replication. The positions of the two blocks within each replication are randomized and, of course, the block contents are randomized within each separate block.

*22.6.2 There are two reasons for keeping the replications together rather than randomizing the two types of blocks over the whole experimental area. These are:

(i) This arrangement permits without the use of uniformity data an easy comparison with the ordinary randomized blocks design, to see whether the confounding has achieved its purpose by reducing error. (This comparison is, however, of limited value since it is possible only after the yields have been obtained and since the findings may not necessarily remain true for a subsequent experiment.)

(ii) The possibility of a block $\times$ treatment interaction is reduced. The presence of such an interaction could have serious consequences. To illustrate this, consider a 2^3 design with ABC confounded. Let X represent the block difference in any replication. Then $X = ABC$. By the symbolic multiplication rule (§ 21.4.3) $AX = BC$. Hence, if the main effect of A were to be different in one type of block from the other, this would show up as an interaction BC, causing a false interpretation. It is considered that this sort of possibility is rather remote if replications are kept together on the ground.

22.7 Statistical analysis of completely confounded design

22.7.1 *The effect of non-orthogonality of blocks and treatments.* The consequences of non-orthogonality have been previously discussed in § 13.6.6. It is always possible by the method of least squares (though it is in general difficult) to obtain a combined S.S. for block and treatment effects, called the Blocks + treatments S.S. Subtraction of this from the Total S.S. gives the Error S.S. Although it is impossible to subdivide the Blocks + treatments S.S. into a part representing purely block effects and a part representing purely treatment effects, if we calculate the S.S. for blocks ignoring treatments, this will "contain all the information" about block effects but also part of the information about treatment effects. Consequently, the remainder of the S.S. Blocks + treatments will reflect only treatment effects, but will not represent

all the information about treatments, part of which is locked up in the S.S. for blocks ignoring treatments. This remainder S.S. is called the S.S. for *treatments eliminating blocks*, and we have the following subdivision of the Total S.S., by means of which an over-all test of treatment effects may be made.

Table 22.1: Skeleton analysis of variance of a non-orthogonal design with b blocks, t treatments, and p plots

Source	D.F.
Blocks ignoring treatments	$b - 1$
Treatments eliminating blocks ..	$t - 1$
Blocks + treatments	$b + t - 2$
Error	$p - b - t + 1$
Total	$p - 1$

The S.S. for blocks ignoring treatments is that obtained from the block totals by the ordinary rules, ignoring non-orthogonality. The key item in Table 22.1 is therefore the S.S. for Treatments eliminating blocks. If this can be calculated, the Error S.S. is obtainable by subtraction.

22.7.2 In like manner, if we calculate the S.S. for treatments ignoring blocks, the remainder of the Blocks + treatments S.S. is called the S.S. for blocks eliminating treatments, which represents only block effects, but not all the information about blocks, part of which is locked up in the S.S. for treatments ignoring blocks. Only when blocks and treatments are orthogonal can the Blocks + treatments S.S. be subdivided into two parts which represent "the truth, the whole truth, and nothing but the truth" in respect of the two classes of effects.

22.7.3 In a completely confounded design with each replication divided into two blocks, the non-orthogonality is particularly simple and the analysis of variance offers no difficulty. The S.S. for treatments eliminating blocks can be calculated either as

(S.S. Treatments ignoring blocks) − (apparent S.S. for confounded interaction),

where the last term is obtained as if there were no confounding, or by adding the S.S.'s of the unconfounded treatment effects. The resulting Treatments S.S. is the S.S. for treatments eliminating blocks by virtue of the fact that the only degree of freedom for treatments non-orthogonal to blocks is removed from the S.S., which has in consequence one D.F. fewer than in an unconfounded design. The apparent S.S. for the confounded interaction is included in the S.S. for blocks ignoring treatments. This follows because the total effect of the confounded interaction worked out in the normal way is actually the difference of the totals of the two types of block, and this is automatically included in the S.S. for blocks ignoring treatments. The analysis follows Table 22.1 *except for the loss of* 1 *D.F. for treatments* (representing the total sacrifice of information on the confounded effect). Since this loss is premeditated and amounts to the assumption that the confounded effect is zero, the sources of variation are usually recorded simply as blocks, treatments,

and error, for, if the confounded effect is truly zero, the Blocks S.S. represents purely block effects. Nevertheless, the basis of the analysis is as summarized in Table 22.1.

22.7.4 In 2^n designs Yates's algorithm is usually employed as before to calculate the total effects, even though it is undiscriminating to the extent that it produces results for all effects, confounded as well as unconfounded. However, although the result for the confounded interaction can be discarded, it is useful for checking purposes if the Treatments S.S. is obtained by both the methods explained in § 22.7.3. Tests of significance of unconfounded treatment effects may be made in the way explained in § 21.5.1.

22.8 Presentation of results of completely confounded design

22.8.1 The purpose of the analysis of factorial experiments is to explain observed treatment differences in terms of main effects and interactions. The presentation of the results should therefore be concerned primarily with interaction tables, and not so much, when the design has many factors, with individual treatment means, since inspection of the latter is rarely worth while and will usually give a confused picture which it is the object of the analysis in terms of main effects and interactions to dissolve. Nevertheless, there is a popular desire to have the individual means available for inspection, as was acceded to in Examples 19.3 and 21.1. (N.B. An alternative, which is equivalent to the presentation of main effects and interaction tables is to calculate graduated means with high order interactions assumed zero—cf. § 21.7.5.)

22.8.2 In a confounded design it is most important that individual treatment means should not be presented as they stand. The reason is that the treatments do not all come from the same blocks; consequently differences between individual means may be due to block effects as much as treatment effects and may be misleading.

22.8.3 In the 2^3 design with *ABC* confounded, discussed in § 22.5.3, differences among the treatments (1), *ab*, *ac*, and *bc*, which occur in the same blocks throughout, are not subject to block differences. Nor are differences among the second group of treatments *a*, *b*, *c*, and *abc*. Differences between a treatment from the one group and a treatment from the other are, however, subject to block differences.

22.8.4 In any confounded design where certain means (not necessarily individual treatment means) are subject to block effects, the latter must be equalized or removed before the means are presented. Such means are called **adjusted treatment means** (adjusted for block effects).

22.8.5 In a 2^3 design with *ABC* confounded, since no main effect or first order interaction is confounded, all two-factor interaction tables are free from block effects and no adjustment is necessary to these. In drawing up any such interaction table it will be found that each entry in the table comes equally from all the blocks of the experiment. As already seen, however, this is not true of the individual treatment means, which must therefore be adjusted

if they are going to be presented. This adjustment is such that it does not affect the two-factor interaction tables, which could be computed either from the adjusted or from the unadjusted means. The adjustment is made by assuming that the confounded interaction (ABC) is zero. On this assumption the apparent value of $[ABC]$, which is the difference between the totals of the two types of blocks, viz.

$$d = \{[abc] + [a] + [b] + [c]\} - \{[(1)] + [ab] + [ac] + [bc]\},$$

where the square brackets indicate unadjusted treatment totals, reflects purely block effects. To adjust, we make d zero, thus removing block effects, which is done by subtracting $\frac{1}{8}d$ from each of the unadjusted totals in the first bracket and adding $\frac{1}{8}d$ to all the totals in the second bracket. In this way we get adjusted totals and hence adjusted means, as will be illustrated in Example 22.1.

22.8.6 Another reason why the presentation of adjusted treatment means is undertaken hesitantly is difficulties with S.E.'s. The adjusted means exemplified in § 22.8.5, calculated on the assumption that the second order interaction is zero, are exactly the same as those obtained on the same assumption but by a different method in § 21.7.5. As seen there, differences between adjusted means may have one of two variances, viz. $2\sigma^2/r$ or $\frac{3}{2}\sigma^2/r$ (r = no. of replications). The first of these applies to two adjusted means within the same group, where the assumption that the second order interaction is zero does not affect matters, and the second applies to two adjusted means in different groups. A single adjusted mean therefore has two S.E.'s according to the type of comparison envisaged. For a comparison involving different groups an effective S.E. (cf. § 16.18.6) is given by $\sqrt{\frac{3}{4}s^2/r}$, where s^2 = Error M.S. At first sight it is rather surprising that the variance of a difference of two adjusted means should be less for treatments in different groups than for treatments in the same group, but a close examination of § 21.7.5 should clear up the apparent anomaly.

*22.8.7 The S.E.'s of § 22.8.6 may be verified otherwise. Consider the comparison abc — (1), where the treatment combinations belong to different groups. The adjusted difference may be derived as a linear function of treatment totals as follows:

	[(1)]	[ab]	[ac]	[bc]	[a]	[b]	[c]	[abc]
$[abc]$	0	0	0	0	0	0	0	1
$\frac{1}{8}d$	$-\frac{1}{8}$	$-\frac{1}{8}$	$-\frac{1}{8}$	$-\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
$[abc] - \frac{1}{8}d$ = adjusted total	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$-\frac{1}{8}$	$-\frac{1}{8}$	$-\frac{1}{8}$	$\frac{7}{8}$
$[(1)]$	1	0	0	0	0	0	0	0
$[(1)] + \frac{1}{8}d$ = adjusted total	$\frac{7}{8}$	$-\frac{1}{8}$	$-\frac{1}{8}$	$-\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
$[abc]-[(1)]-\frac{1}{4}d$ = difference of adjusted totals	$-\frac{3}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$-\frac{1}{4}$	$-\frac{1}{4}$	$-\frac{1}{4}$	$\frac{3}{4}$

Hence the variance of the difference of two adjusted totals

$$= (2 \times \tfrac{9}{16} + 6 \times \tfrac{1}{16})r\sigma^2 = \tfrac{3}{2}r\sigma^2,$$

and that for adjusted means $= \frac{3}{2}\sigma^2/r$. This is for treatments belonging to different groups. For treatments belonging to the same group, the adjustment $\frac{1}{8}d$ cancels on subtraction and the variance is $2\sigma^2/r$ as usual. Notice how the coefficients of the linear function for the difference of adjusted totals sum to zero on either side of the dotted line, showing that the expression contains no block effects.

Example 22.1 (Data from Saunders and Rayner, *Statistical methods with special reference to field experiments*) In a 2^3 experiment on maize conducted in blocks of 4 plots 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*a*		Block 1*b*		Block 2*b*		Block 2*a*		Block 3*b*		Block 3*a*		Block 4*a*		Block 4*b*		Block 5*a*		Block 5*b*	
(1)	33	*g*	46	*p*	46	*ps*	59	*p*	48	*gs*	57	*pg*	44	*pgs*	64	*ps*	55	*s*	32
pg	47	*s*	45	*g*	43	*pg*	46	*s*	44	*pg*	45	*gs*	54	*g*	37	*gs*	50	*p*	39
ps	57	*p*	48	*pgs*	66	*gs*	55	*g*	46	(1)	34	(1)	31	*s*	36	*pg*	45	*g*	35
gs	55	*pgs*	65	*s*	42	(1)	31	*pgs*	70	*ps*	61	*ps*	58	*p*	37	(1)	29	*pgs*	61

Computation sheet

EFFECTS CONFOUNDED: *PGS* (all replications) (A)

	Treatments	Blocks 1*a*	2*a*	3*a*	4*a*	5*a*	Treatment totals
	(1)	33	31	34	31	29	158
	ps	57	59	61	58	55	290
(B)	*gs*	55	55	57	54	50	271
	pg	47	46	45	44	45	227
	Block totals (*a*)	192	191	197	187	179	946 (C)
		1*b*	2*b*	3*b*	4*b*	5*b*	
	p	48	46	48	37	39	218
	g	46	43	46	37	35	207
	s	45	42	44	36	32	199
	pgs	65	66	70	64	61	326
	Block totals (*b*)	204	197	208	174	167	950 (C)
(D)	Replication totals	396	388	405	361	346	1896 (C)

C.F. = 89,870·4 (E)

Total S.S. = 94,506·0
89,870·4
4,635·6 (E)

S.S. Blocks ignoring treatments = 90,254·5 (F)
89,870·4
384·1

S.S. Treatments ignoring blocks = 94,056·8 (E)
89,870·4
4,186·4

S.S. *PGS* ignoring blocks = 0·4 (G)

S.S. Treatments eliminating blocks = 4,186·0 (H)

Yates's method (E)

Treatment	Treatment totals	(1)	(2)	(3)	Effect
(1)	158	376	810	1896	G.T.
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*] (I)
	1896			2608 (J)	

S.S. Treatments eliminating blocks = 4,186·0 (K) (check).

Analysis of variance

Source	D.F.	S.S.	M.S.
Blocks	9	384·1	
Treatments	6 (L)	4,186·0	
Error	24	65·5	2·73 (M)
Total	39	4,635·6	

S.E. of single yield = 1·652

S.E. as % of mean = $\frac{1{\cdot}652}{1896} \times 40 \times 100 = 3{\cdot}49\%$

Estimate of variance of unconfounded total effect = 40 × 2·73 = 109·2

S.E. = 10·45

Least significant values for unconfounded total effects = 10·45 × *t* (24 D.F.)

$= 10{\cdot}45 \times \begin{cases} 2{\cdot}064 \\ 2{\cdot}797 \end{cases}$

= 21·6 (5%)
29·2 (1%)

S.E. of single treatment total: (N)

(i) for comparisons within a group = $\sqrt{5 \times 2{\cdot}73} = 3{\cdot}69$

(ii) for comparisons between groups (O) = $\sqrt{\frac{3}{4} \times 5 \times 2{\cdot}73} = 3{\cdot}20$.

Least significant differences (2 treatment totals):

(i) within a group = $\sqrt{10 \times 2{\cdot}73} \times \begin{cases} 2{\cdot}064 \\ 2{\cdot}797 \end{cases} = \begin{cases} 10{\cdot}78\ (5\%) \\ 14{\cdot}61\ (1\%) \end{cases}$

(ii) between groups = $\sqrt{\frac{3}{2} \times 5 \times 2{\cdot}73} \times \begin{cases} 2{\cdot}064 \\ 2{\cdot}797 \end{cases} = \begin{cases} 9{\cdot}34\ (5\%) \\ 12{\cdot}66\ (1\%) \end{cases}$

Conversion factors: As for Example 21.1

(P) Adjusted treatment totals: $\frac{1}{8}[PGS] = 0{\cdot}5$

	(1)	*ps*	*gs*	*pg*	Total
Group (*a*) (+0·5)	158·5	290·5	271·5	227·5	948·0 (Q)
	p	*g*	*s*	*pgs*	Total
Group (*b*) (−0·5)	217·5	206·5	198·5	325·5	948·0 (Q)

Presentation of results

TREATMENT MEANS IN BAGS PER MORGEN (ADJUSTED FOR BLOCK EFFECTS)

Group (*a*)		Group (*b*)	
(1)	15·8	*p*	21·8
ps	29·0	*g*	20·6
gs	27·2	*s*	19·8
pg	22·8	*pgs*	32·6

Mean 23·7

	For comparisons of treatments within a group	For comparisons of treatments in different groups
S.E.	±0·369	±0·320 (effective)
L.S.D. (5%)	1·08	0·93
L.S.D. (1%)	1·46	1·27

MEAN RESPONSES AND INTERACTIONS IN BAGS PER MORGEN

P	5·65**
G	4·15**
S	6·90**
PG	−1·90**
PS	1·65**
GS	1·25** (R)

S.E. = ±0·261

Least significant values:
0·54 (5%)
0·73 (1%)

Conclusions. As for Example 21.1, except for the adjustment of S.E.'s to suit the altered Error M.S. and the omission of the first sentence of the last paragraph.

Notes on the computations

The apparent familiarity of the example is no accident! The treatments and plot yields are identical with those of Example 21.1, but the design is different. This device of analysing the same figures in different ways in order to bring out the difference between the two analyses has been employed before (cf. §§ 9.8.4 and 20.3.9). The design was actually confounded as an Example 22.1.

(A) The block contents remain the same for all replications and, as explained in § 22.5.2 or by inspection of the example given in § 22.5.3, we identify *PGS* as the confounded interaction.

(B) The arrangement of treatments according to type of block, though not essential, is helpful for certain purposes.

(C) Checks down and across.

(D) Normally required for checking purposes only. See, however, § 22.9.

(E) As in Example 21.1.

(F) $\frac{1}{4}(192^2 + \cdots + 179^2 + 204^2 + \cdots + 167^2) - C.F.$

(G) [*PGS*] = Block totals (*b*) − Block totals (*a*) (see § 22.7.3) = 950 − 946 = 4.
S.S. *PGS* (ignoring blocks) = $4^2/40 = 0{\cdot}4$.
The sign of the difference is determined by the fact that the treatment combination containing all the letters (here *pgs*) always has a positive sign, as may be verified from § 21.4.

(H) The S.S. for treatments eliminating blocks is here calculated by the first method given in § 22.7.3. The actual difference between the S.S.'s for treatments ignoring and eliminating blocks is here very small, but this should not be regarded as typical. Frequently [*PGS*], consisting as it does of a difference of totals of two sets of blocks, is very large.

(I) The interaction *PGS* is bracketed to indicate that it is confounded (see § 22.12.3). Note that in the calculation of total effects, the confounding is otherwise ignored.

(J) This partial check is made in the same way as before (cf. § 21.5.4). The figure in square brackets is included in the total.

(K) This is here calculated as $\frac{1}{40}(226^2 + 166^2 + \cdots + 50^2)$. The check is strengthened by the fact that [*PGS*] has been independently checked.

(L) Blocks D.F. = 9 because there are 10 blocks. The extra 5 D.F. in comparison with Example 21.1 come from Error (4 D.F.) and Treatments (1 D.F., viz. *PGS*). There are 6 D.F. for treatments since 1 D.F. (*PGS*) is completely confounded. Error D.F. are obtained by subtraction.

(M) The Error M.S. is reduced in comparison with Example 21.1, but, since the previous analysis was really invalid because the design really was confounded, the difference is not an accurate reflection of the gain in precision due to confounding.

(N) Unnecessary if it is decided not to present individual means.

(O) This is the effective S.E. explained in § 22.8.6. Here it is calculated for an adjusted *total*, however.

(P) As explained in § 22.8.5.

(Q) These must be equal after the adjustments.

(R) Being confounded, *PGS* is, of course, omitted.

22.9 A test of significance for the completely confounded interaction. The recovery of inter-block information

22.9.1 In a completely confounded design there is a total loss of information on the confounded interaction *so far as an analysis based on the plots as unit is concerned.* It is, however, possible to salvage information on this interaction from an analysis based on the blocks as unit. This is known as **recovery of inter-block information** (inter = between); the analysis discussed hitherto has been concerned with information available from the variation of plots within blocks (**intra-block information**).

22.9.2 To illustrate the recovery of inter-block information in Example 22.1, consider an analysis of block totals only. The blocks are arranged in five pairs, each pair consisting of an (*a*) and a (*b*) block, located at random within each pair. Regarding the blocks as experimental units, then, we have the equivalent of a design with two treatments, *a* and *b*, and five replications, the block pairs constituting the blocks of a randomized blocks design. Moreover, the treatment contrast $b - a = pgs + p + g + s - pg - ps - gs - (1)$ is the second order interaction, *PGS*. We therefore have the following analysis of variance:

Table 22.2: Analysis of variance between blocks in Example 22.1

Source	D.F.	S.S.	M.S.
Between block pairs (replications) ..	4	307·4	
PGS	1	0·4	0·4
Within block pairs	4	76·3	19·1
Blocks	9	384·1	

This analysis has been calculated on a per-plot basis (i.e. divisors are based on the numbers of plots), so that the Total S.S. is the same as the Blocks S.S. in Example 22.1. The S.S. between block pairs is $\frac{1}{8}(396^2 + \cdots + 346^2) - C.F.$ Clearly *PGS* is non-significant. In any case the estimate obtained has very low precision, being based on an Error M.S. with only 4 D.F. (measuring random variation between blocks), and very much larger than the Error M.S. of Example 22.1 (measuring random intra-block variation). In view of this *the inter-block analysis as above is seldom performed.*

22.10 The split-plot design as an example of complete confounding

22.10.1 The above analysis is equivalent to a split-plot design. The blocks are the whole-plots and the plots sub-plots. The contrast *PGS* is the whole-plot treatment contrast and the other main effects and interactions the sub-plot treatment contrasts; *PGS* is confounded with respect to the plots of the experiment, but not with respect to the blocks. Similarly, in a split-plot design the whole-plot treatments are confounded with respect to sub-plots and must be tested against the whole-plot error and not the sub-plot error. In Table

22.2 the M.S. for within block pairs may be regarded as Error (*a*) for the testing of *PGS* and the Error M.S. as Error (*b*). The 4 D.F. for within block pairs are those transferred from the Error S.S. of the unconfounded analysis into the present Blocks S.S. (see Example 22.1, Note L).

22.10.2 To be specific, a split-plot design is a design with the main effects of one factor (the whole-plot factor) completely confounded. Thus the arrangement of § 22.3.4, if replicated, would be a 2×2 design with A as whole-plot factor and B as sub-plot factor. In general, of course, the whole-plot factor has more than two levels, each replication being divided into more than two sub-blocks (whole-plots). The idea of sacrificing information expressed in § 20.9.4 is exactly the same as the idea behind confounding (§ 22.2.3).

22.10.3 The difference between a split-plot design and the type of design considered in this chapter is that it is a main effect which is confounded, not a high order interaction. However, apart from the case considered in § 20.9.6 (1), the complete loss of information on a main effect is seldom acceptable to an experimenter. Hence in a split-plot design the recovery of inter-block information (i.e. the whole-plot analysis) is not ignored.

*22.10.4 The design mentioned in § 22.5.1 is an example of "split-plot confounding". The set-up is different from a standard split-plot design in that the contents of the whole-plots are not the same. It is easily seen from the above discussion that the interaction AB, which is confounded so far as the sub-plot error is concerned, will appear in the whole-plot analysis, where it can be tested against the whole-plot error.

22.11 Simple examples of partial confounding in 2^n designs

22.11.1 Partial confounding is most commonly employed in designs in which each replication is divided into more than two blocks. This means that more than one treatment D.F. is confounded in a replication; consequently, effects other than the highest order interaction will be involved and partial confounding becomes almost a necessity.

22.11.2 When consideration is restricted to cases where a replication is divided into two blocks only, there are not many designs of practical interest, but the principles are nevertheless worth study since they are applicable to the more complex designs. Disregarding 2^2 designs, where any confounding is unlikely, we come to the 2^3 design in blocks of 4 plots. If blocks of this size are preferred to an ordinary randomized blocks design, complete confounding of the second order interaction may not be desired. A possible solution is to confound each of the 4 interactions once each over 4 replications. This type of design is exemplified by Example 22.2, where with factors N, P, and K the interactions NPK, NP, NK, and PK are each confounded once. In the replication in which NPK is confounded, the block contents are the same as in § 22.5.3 (with interchange of letters) and all main effects and first order interactions are orthogonal to the block effects of this replication. In the other three replications the block contents are determined by the rule

given in § 22.5.2 and in these replications *NPK* is orthogonal to block effects. It is therefore possible to estimate *NPK* from only 3 out of the 4 replications (cf. § 22.4.2), and we say that there is "$\frac{3}{4}$ relative information on *NPK*" compared with an unconfounded design. Since this applies equally to all interactions, the design is said to be "balanced", there being equal relative information on all interactions. The use of 4 replications and a balanced design is not compulsory.

22.11.3 In a 2^4 design with blocks of 8 plots there may be partial confounding of second order interactions as well as the third order interaction in case it is desired not to confound the latter completely. In a 2^5 design with blocks of 16 plots there could scarcely be any reason for not completely confounding the fourth order interaction (cf. § 19.21.1).

22.11.4 As regards lay-out and randomization procedure, the directions given in § 22.6.1 apply equally to the partially confounded designs discussed here.

22.12 Statistical analysis of partially confounded 2^n designs

22.12.1 The analysis follows Table 22.1. In order to obtain the S.S. for treatments eliminating blocks it is necessary to estimate the partially confounded effects only from the replications in which they are not confounded (cf. § 22.4.2). Such estimates are free from block effects. The separate S.S.'s for the unconfounded effects and the partially confounded effects estimated in this manner are then added to give the S.S. for treatments eliminating blocks.

22.12.2 Consider the balanced 2^3 design introduced in § 22.11.2. The total effect [*NPK*], for example, over all replications, which is subject to block effects due to confounding, is made up of two parts—a part calculated from the replications where *NPK* is not confounded and a part from the replication in which it is confounded. This may be written

$$[NPK] = [NPK]' + [NPK]_c, \qquad [22.4]$$

where $[NPK]'$ signifies the total effect modified by subtraction of the block effects contained in [*NPK*] as estimated from the appropriate block totals. We may call this the adjusted total effect. Then

$$[NPK]' = [NPK] - [NPK]_c, \qquad [22.5]$$

and $[NPK]_c$ is the difference of the block totals in the replication in which *NPK* is confounded, or in other words $[NPK]'$ signifies here the total effect of *NPK* obtained from the replications in which *NPK* is not confounded. Since $[NPK]'$ is calculated from 24 plots, we have

$$NPK = \tfrac{1}{24}[NPK]',$$

i.e. $[NPK]'$ divided by the total number of plots from which it is calculated in accordance with the definition adopted in § 21.3. This estimate has variance $\frac{1}{24}\sigma^2$, where σ^2 = error variance. If there were no confounding, *NPK* would be taken as $\frac{1}{32}[NPK]$ with variance $\frac{1}{32}\sigma^2$, assuming the error variance to be

the same. On this assumption the relative information on *NPK*, which is defined to be the ratio of the variance of *NPK* with and without confounding $= \frac{1}{32}\sigma^2/\frac{1}{24}\sigma^2 = \frac{24}{32} = \frac{3}{4}$, as less formally obtained in § 22.11.2. Of course, a reduction in σ^2 in the confounded design due to smaller block size could offset this 25% loss of information (cf. § 22.4.4). All the work of this paragraph applies equally to all the partially confounded interactions in the balanced 2^3 design.

*Equation [22.5] may be neatly explained in terms of estimates of treatment and block effects. Let B_1 and B_2 be the block totals in the replication in which *NPK* is confounded such that

$$B_1 - B_2 = [NPK]_c.$$

Now

$$[NPK] = 32\,NPK + 4(b_1 - b_2), \qquad [22.6]$$

where b_1 and b_2 represent estimates of the components of yield received by plots in the blocks whose totals are B_1 and B_2 respectively. Also

$$B_1 - B_2 = 4(b_1 - b_2) + 8\,NPK, \qquad [22.7]$$

(illustrating the partial confounding of *NPK*), so that

$$[NPK]' = [NPK] - (B_1 - B_2) \qquad [22.8]$$

$$= 32\,NPK + 4(b_1 - b_2) - 4(b_1 - b_2) - 8\,NPK$$

$$= 24\,NPK. \qquad [22.9]$$

In seeking to cancel $4(b_1 - b_2)$ in the expression for $[NPK]$ by means of [22.7], we automatically (because of the confounding) bring in *NPK*, and the reduction of the coefficient from 32 in [22.6] to 24 in [22.9] is indicative of the loss of information incurred. Equations [22.6] and [22.7] are two of the normal equations for this design.

22.12.3 For any unconfounded effect the adjusted and unadjusted total effects are, of course, the same; as an example, for the main effect of *N* in the design quoted, $[N] = [N]'$. There is consequently a tendency to use this square-bracket notation only when dealing with (partially) confounded effects. The use of the square brackets to differentiate the confounded interaction *PGS* in Example 22.1 (cf. Note I) is an illustration of this tendency.

22.12.4 For the calculation of total effects Yates's method is again usually employed. It produces, of course, unadjusted total effects, but use of [22.5] or [22.8], for example, makes the adjusted effects easily obtainable. Care must be taken, however, with the sign of the block difference. A simple rule is that the treatment combination containing all the letters appears with a positive sign in the linear functions of all effects (cf. § 21.4) and consequently, for example, in $[NPK]_c$. Hence, in applying [22.5], the block containing the treatment combination with all the letters (here *npk*) must be the one which receives the negative sign.

22.12.5 It is also possible to calculate unconfounded effects directly by means of the appropriate linear functions as exemplified in § 21.5.3. For partially confounded effects, adjustments to the unadjusted total effects

obtained from treatment totals must be made as described in § 22.12.4, or else it is possible to follow literally the rule given in § 22.4.2 and apply the appropriate linear function to the *individual plot yields* of the replications in which the effect is not confounded. To illustrate this method from Example 22.2, we might calculate $[NP]'$, which is confounded in the third replication, as follows, using the linear function

$$NP = (1) - p - n + pn + k - pk - nk + pnk:$$

$$\begin{aligned}[NP]' &= 26 - 45 - 23 + 49 + 18 - 42 - 25 + 50 \text{ (Replication 1)} \\ &\quad + 26 - 49 - 30 + 50 + 26 - 44 - 26 + 47 \text{ (Replication 2)} \\ &\quad + 48 - 78 - 59 + 72 + 53 - 67 - 41 + 66 \text{ (Replication 4)} \\ &= 2.\end{aligned}$$

Unless there are only two replications, this method of working from individual yields is laborious and is not therefore recommended except perhaps as a check. If Yates's method is used, the unadjusted total effects can be checked as described in Chapter 21, which virtually limits sources of computing mistakes to the subtraction illustrated by [22.5] or [22.8], and if this is undertaken carefully the more laborious check is unnecessary. The points made in § 21.5.5 are still valid here.

22.12.6 In calculating the S.S. for treatments eliminating blocks, Formula [21.7] must be varied to the extent that the divisor for each partially confounded interaction is equal to the number of plots from which it is estimated. Similar variations are necessary in the calculation of S.E.'s.

22.13 Presentation of results of partially confounded 2^n design

22.13.1 The difficulties concerning S.E.'s with adjusted treatment means mentioned in § 22.8.6 are accentuated with partial confounding. For this reason it is best to avoid presentation of adjusted means for individual treatment combinations. This can also be supported on the general grounds discussed in § 22.8.1. If it is desired to present these means, they have to be calculated by the method explained in § 21.7. For example, in the 2^3 balanced design hitherto discussed

$$(1) = m - N - P - K + NP + NK + PK - NPK,$$

the conventional order being changed to keep unconfounded effects first. In practice this is calculated as

$$(1) = \tfrac{1}{32}(G.T. - [N] - [P] - [K]) + \tfrac{1}{24}([NP]' + [NK]' + [PK]' - [NPK]').$$

All the terms are orthogonal, so that this adjusted mean has variance $(\frac{4}{32} + \frac{4}{24})\sigma^2 = \frac{7}{24}\sigma^2$, since Var.$[P] = 32\sigma^2$ and Var.$[NP]' = 24\sigma^2$, etc.

22.13.2 All adjusted means have the same variance, viz. $\frac{7}{24}\sigma^2$, but the variance of the difference of any two such adjusted means is not in general double this. For example,

$$p = m - N + P - K - NP + NK - PK + NPK;$$

hence

$$p - (1) = 2P - 2NP - 2PK + 2NPK$$

with variance $(\frac{4}{32} + \frac{12}{24})\sigma^2 = \frac{5}{8}\sigma^2$, in this case slightly larger than $\frac{7}{12}\sigma^2$. In all

there are three possible variances according to which difference is taken, and in an unbalanced design the number would be greater. Actually, if $\frac{7}{12}\sigma^2$ were used as an average variance for the difference of two adjusted means, it would not be very far wrong.

22.13.3 *The difficulties with unadjusted means extend to first order interaction tables whenever the corresponding first order interaction is partially confounded.* In such a case the means in the body of the table require adjustment for block effects, though marginal means will remain unchanged if the corresponding main effects are unconfounded. Where adjustment is needed, two-factor interaction tables can be built up from the adjusted treatment means, but it is simpler to use the method given in Table 21.4. For example, to construct the $N \times P$ table in Example 22.2 we use

$$\begin{aligned} \dagger(1) &= m - N - P + NP = \tfrac{1}{32}(G.T. - [N] - [P]) + \tfrac{1}{24}[NP]' \\ n &= m + N - P - NP = \tfrac{1}{32}(G.T. + [N] - [P]) - \tfrac{1}{24}[NP]' \\ p &= m - N + P - NP = \tfrac{1}{32}(G.T. - [N] + [P]) - \tfrac{1}{24}[NP]' \\ np &= m + N + P + NP = \tfrac{1}{32}(G.T. + [N] + [P]) + \tfrac{1}{24}[NP]'. \end{aligned}$$

It will be seen that any marginal mean, e.g. $\frac{1}{2}(n + np) = m + N$, is unaltered by the adjustment.

22.13.4 Differences of adjusted means in the body of the table have different variances according as they belong to the same row or column or not. For two means in the same row or column, e.g. np and p,

$$\text{Var.}(np - p) = \text{Var.}(2N + 2NP) = (\tfrac{4}{32} + \tfrac{4}{24})\sigma^2 = \tfrac{7}{24}\sigma^2.$$

For two means not in the same row or column, e.g. n and p,

$$\text{Var.}(p - n) = \text{Var.}(2N + 2P) = \tfrac{1}{4}\sigma^2,$$

unaffected by adjustments. The variance of a single mean in the table $= (\frac{3}{32} + \frac{1}{24})\sigma^2 = \frac{13}{96}\sigma^2$, compared with $\frac{1}{8}\sigma^2$ if no adjustment had been necessary, but this variance is not useful since the variance of the difference of two means is not $\frac{13}{48}\sigma^2$.

22.13.5 Some final points. In the conversion to standard units different conversion factors must be used for unadjusted and adjusted total effects, based on the total number of plots from which each is estimated; for the adjusted means given in this section the conversion factor for a single plot (means to means) must be used. Of course, all the above formulae relate only to the balanced 2^3 design; for other partially confounded designs they would be different, but constructed on the same principles.

Example 22.2 (Data from Saunders and Rayner, *Statistical methods with special reference to field experiments*) The following are the plan and yields (in lb. per $\frac{1}{100}$ morgen plot) in a maize fertilizer experiment in which dressings of nitrogen (N), phosphate (P), and potash (K) were under test. Analyse the data presenting results in bags per morgen (1 bag = 200 lb.).

† Here used to denote the mean of all plots receiving neither N nor P, not merely the treatment combination (1). Similarly for n, p, and np.

1*a*		1*b*		2*a*		2*b*		3*a*		3*b*		4*a*		4*b*	
nk	25	*npk*	50	*np*	50	*nk*	26	*p*	64	(1)	48	*k*	53	(1)	48
(1)	26	*n*	23	*pk*	44	*npk*	47	*nk*	51	*np*	71	*np*	72	*n*	59
np	49	*k*	18	*k*	26	(1)	26	*pk*	64	*k*	49	*nk*	41	*pk*	67
pk	42	*p*	45	*n*	30	*p*	49	*n*	52	*npk*	61	*p*	78	*npk*	66

Computation sheet

(A) Interaction confounded	*NPK*		*NK*		*NP*		*PK*	
Block	1*a*	1*b*	2*a*	2*b*	3*a*	3*b*	4*a*	4*b*
(B) Treatments and yields	(1) 26 *np* 49 *pk* 42 *nk* 25	*p* 45 *n* 23 *k* 18 *npk* 50	*n* 30 *k* 26 *np* 50 *pk* 44	(1) 26 *p* 49 *nk* 26 *npk* 47	*p* 64 *n* 52 *pk* 64 *nk* 51	(1) 48 *k* 49 *np* 71 *npk* 61	*p* 78 *k* 53 *np* 72 *nk* 41	(1) 48 *n* 59 *pk* 67 *npk* 66
Block totals	142	136	150	148	231	229	244	240

G.T. = 1520 (C)

C.F. = 72,200·0

Total S.S. = 80,170·0 − 72,200·0 = 7,970·0

S.S. Blocks (ignoring treatments) (D) = 76,500·5 − 72,200·0 = 4,300·5

Yates's method

Treatment	Treatment totals	(1)	(2)	(3)	Effect (E)
(1)	148	312	790	1,520	G.T.
n	164	478	730	26	*N*
p	236	289	22	318**	*P*
np	242	441	4	0	[*NP*]
k	146	16	166	−60*	*K*
nk	143	6	152	−18	[*NK*]
pk	217	−3	−10	−14	[*PK*]
npk	224	7	10	20	[*NPK*]
Total	1,520 (C)			1,792 = 8 × 224	

S.S. Treatments ignoring blocks (F) = 75,522·5 − 72,200·0 = 3,322·5 = 3,293·8 + 28·7 = 3,322·5 (check)

Adjusted total effects: (G)

$[NP]' = 0 + 231 - 229 = 2$
$[NK]' = -18 + 150 - 140 = -16$
$[PK]' = -14 + 244 - 240 = -10$
$[NPK]' = 20 + 142 - 136 = 26$

S.S. Treatments eliminating blocks = 3,293·8 + 43·2 = 3,337·0 (H)

Analysis of variance

Source	D.F.	S.S.	M.S.
Blocks (ignoring treatments)	7	4,300·5	
Treatments (eliminating blocks)	7 (I)	3,337·0	
Error	17	332·5	19·56
Total	31	7,970·0	

S.E. of single yield = 4·42

S.E. as % of mean = $\frac{4{\cdot}42}{1520} \times 32 \times 100 = 9{\cdot}31\%$

Estimate of variance of unconfounded total effects $= 32 \times 19{\cdot}56 = 625{\cdot}92$

S.E. $= 25{\cdot}02$

Least significant values for unconfounded total effects $= 25{\cdot}02 \times t$ (17 D.F.)

$$= 25{\cdot}02 \times \begin{cases} 2{\cdot}110 \\ 2{\cdot}898 \end{cases}$$

$$= \begin{matrix} 52{\cdot}8\ (5\%) \\ 72{\cdot}5\ (1\%) \end{matrix}$$

Estimate of variance of adjusted total effect $= 24 \times 19{\cdot}56 = 469{\cdot}44$

S.E. $= 21{\cdot}67$

Least significant values for adjusted total effects $= 21{\cdot}67 \times \begin{cases} 2{\cdot}110 \\ 2{\cdot}898 \end{cases}$

$$= \begin{matrix} 45{\cdot}7\ (5\%) \\ 62{\cdot}8\ (1\%) \end{matrix} \text{ (J)}$$

Conversion factors:

(i) Main effects to mean responses in bags per morgen $= \frac{100}{200} \times \frac{1}{16} = 0{\cdot}03125$ (K)
(L)

(ii) Adjusted total effects $= \frac{100}{200} \times \frac{1}{12} = 0{\cdot}041667$
(L)

(iii) Single plot yield $\overset{\text{(M)}}{=} \frac{100}{200} = 0{\cdot}5$

Presentation of results

MEAN RESPONSES AND INTERACTIONS IN BAGS PER MORGEN

Mean responses		Interactions (adjusted for block effects)	
N	0·81	*NP*	0·08
P	9·94**	*NK*	−0·67
K	−1·88*	*PK*	−0·42
		NPK	1·08
S.E.	±0·782	S.E.	±0·903
Least significant values:	1·65 (5%)	Least significant values:	1·90 (5%)
	2·27 (1%)		2·62 (1%)

Mean yield = 23·75 bags per morgen (N)
C.V. = 9·31%

There was a moderate, non-significant response of 0·81 bags per morgen to the nitrogen application. The response to phosphate, however, of 9·94 bags per morgen was very marked and highly significant. On the other hand, potash gave a marked depressive effect of 1·88 bags per morgen (significant at the 5% level). These responses appear to be very consistent at all levels of the other two factors. (O)

Notes on the computations

(A) It is first noted that the block contents are different for each replication, which means that four different effects are partially confounded. It is also necessary to determine which effect is confounded in each replication, although actually this information is not used until the adjustment of total effects. The method given in § 22.5.2 is adequate here, for which purpose it is convenient to have the interaction matrix (as in [21.5], but with changed letters) available. For designs with more factors it is laborious to write out the whole matrix, but the linear functions for likely interactions (cf. § 22.11.3) can be separately obtained by the method of § 21.4.2 or § 21.4.5.

(B) In view of the different block contents, not much is possible in the way of presenting an orderly table. The order of treatments in each block has been derandomized but this is not really necessary.

(C) Neither Yates's algorithm nor the calculation of the S.S.'s should be proceeded with until the G.T. has been checked.

(D) $\frac{1}{4}(142^2 + 136^2 + \cdots + 240^2) - C.F.$

(E) Unadjusted total effects indicated by square brackets.

(F) This is done as a check as explained in § 22.12.5. The first calculation $= \frac{1}{4}(148^2 + 164^2 + \ldots + 224^2) - C.F.$ and the second is $\frac{1}{32}(26^2 + 318^2 + 60^2) + \frac{1}{32}(0^2 + 18^2 + 14^2 + 20^2)$. The latter is done in two parts since the first part, calculated for unconfounded effects, is free from block differences.

(G) Calculated as explained in § 22.12.4. The block total with the negative sign is the one containing *npk* in the replication concerned. The result for $[NP]'$ agrees with that obtained in § 22.12.5.

(H) The formula here is

$$\tfrac{1}{32}(26^2 + 318^2 + 60^2) + \tfrac{1}{24}(2^2 + 16^2 + 10^2 + 26^2),$$

cf. § 22.12.6. The first part was calculated above (see Note F).

(I) No loss of D.F. All treatment effects are estimated and no assumption is made that any interactions are negligible.

(J) The asterisks representing significance may now be inserted. Main effects *P* and *K* are significant, but no adjusted effects.

The calculation of the S.E. of a single treatment total which would usually appear at this stage is pointless in view of § 22.13.2. The intention is, in any case, not to present adjusted means of individual treatment combinations.

(K) In general, conversion factors are required for unadjusted (i.e. unconfounded) and adjusted total effects. In this example only main effects are unconfounded.

(L) The divisors are taken as half the divisors for the corresponding S.S.'s (cf. Note H) in accordance with Yates's definitions (§ 21.3.7).

(M) Required for adjusted interaction tables as explained in § 22.13.5.

(N) By giving the mean as well as the mean responses and interactions, it is possible for an interested person to work out adjusted means of individual treatment combinations. The method of § 22.13 could be used, but the mean responses and interactions must be divided by 2 since these are presented in accordance with Yates's definitions.

(O) The consistency of the responses is asserted because no interaction is very appreciable. This has the consequence that in this experiment it is hardly necessary to present any interaction tables. However, for the sake of the example, this will be illustrated for the $N \times K$ table.

	No *N*	*N*	Mean	Response to *N*
No *K*	23·9	25·4	24·7	1·5
K	22·7	22·9	22·8	0·2
Mean	23·3	24·2	23·8	0·9
Response to *K* ..	−1·2	−2·5	−1·9	—

For example, the mean for No *N*, No *K* (cf. § 22.13.3) is

$$\{\tfrac{1}{32}[1520 - 26 - (-60)] + \tfrac{1}{24}(-16)\}0{\cdot}5,$$

or alternatively

$$23{\cdot}75 + \tfrac{1}{2}[-0{\cdot}81 - (-1{\cdot}88) + (-0{\cdot}67)].$$

In the first expression 0·5 is the conversion factor; in the second the division by 2 is for the reason given in Note N. In repeated applications it would be desirable to work out the main effects and interactions as defined in § 21.3, i.e. half the values given in the presentation of results. S.E.'s could be evaluated as explained in § 22.13.4.

EXERCISES

22.1 The following are the field plan and yields (in cwt. per acre) of a 2^4 experiment on soybeans. The treatments are all combinations of

Dung : 10 tons per acre (*d*) or nil,
Nitrochalk : ½ cwt. per acre (*n*) or nil,
Superphosphate : ½ cwt. per acre (*p*) or nil.
Muriate of potash : 1 cwt. per acre (*k*) or nil.

Analyse the data.

Block 1A

nk	64·4	(1)	40·1
np	32·4	*dn*	63·4
dp	53·7	*pk*	44·7
dk	47·4	*dnpk*	65·2

Block 1B

p	33·6	*npk*	51·9
dnp	56·9	*n*	27·3
d	58·6	*k*	43·2
dnk	69·8	*dpk*	59·7

Block 2A

(1)	62·6	*dp*	72·8
nk	67·3	*dk*	77·2
np	49·6	*dnpk*	78·2
dn	74·7	*pk*	74·1

Block 2B

n	64·1	*k*	59·7
dnk	96·4	*dpk*	60·2
p	52·8	*dnp*	68·5
d	70·9	*npk*	52·4

22.2 In an experiment on potatoes conducted at Cedara the treatments were all combinations of the following:

Nitrogen

a = Single application of 150 lb. N per morgen as sulphate of ammonia
(1) = 150 lb. N per morgen as sulphate of ammonia, $\frac{1}{3}$ applied at planting and $\frac{2}{3}$ applied as a top-dressing after plant emergence

Potash

b = 336 lb. K_2O per morgen as potassium chloride
(1) = 336 lb. K_2O per morgen as potassium sulphate

Seed material

c = Whole seed potatoes
(1) = Cut seed potatoes

All plots received 500 lb. per morgen of superphosphate.

The field plan and tuber weights in lb. for each plot are given below. Analyse the data presenting mean responses and interactions in bags per morgen, given that the net plot size was 6 ft. × 45 ft., 1 bag = 200 lb., and 1 morgen = 10,244 sq. yds.

Replication I

Block 1		Block 2	
(1)	86·1	*ac*	98·6
ab	92·6	*a*	81·3
c	75·4	*b*	93·0
abc	63·6	*bc*	94·0

Replication II

Block 3		Block 4	
abc	77·4	*a*	81·5
(1)	91·1	*ab*	86·6
b	93·7	*bc*	92·3
ac	100·3	*c*	82·2

Replication III

Block 5		Block 6	
c	95·0	*ac*	119·7
a	82·7	(1)	104·8
b	87·3	*bc*	101·6
abc	76·0	*ab*	104·3

Replication IV

Block 7		Block 8	
abc	94·0	*ab*	82·9
c	104·6	(1)	96·3
a	79·9	*bc*	110·6
b	106·9	*ac*	59·7

Replication V

Block 9		Block 10	
bc	129·8	*c*	105·1
a	95·8	*ac*	98·2
abc	106·1	*ab*	91·4
(1)	106·0	*b*	105·8

Replication VI

Block 11		Block 12	
c	108·6	*ab*	104·0
b	94·6	*bc*	101·1
a	100·1	*ac*	91·5
abc	127·3	(1)	92·4

(Data from Department of Crop Science, Natal Region, Department of Agriculture Technical Services)

22.3 The following are the plan and yields of a 2^4 experiment with factors *L*, *P*, *K*, *N* on wheat. Each yield is in kilograms and represents 4 sample rectangles each 2 ft. × 35 in. Analyse, presenting results in bags per morgen given that 1 bag = 200 lb., 1 lb. = 0·4536 kg., 1 morgen = 10,244 sq. yd.

kn 1·20	*lpn* 1·25	*l* 1·08	(1) 1·16	*pn* 1·26	*pk* 1·13	*lpk* 1·15	*lkn* 1·29

n 1·19	*lpkn* 1·20	*p* 1·15	*lk* 1·40	*ln* 1·20	*k* 1·12	*lp* 1·35	*pkn* 1·24

p 1·10	(1) 1·07	*ln* 1·32	*lk* 1·26	*lpn* 1·12	*kn* 1·18	*pkn* 1·32	*lpk* 1·01

pk 1·19	*n* 1·26	*lpkn* 1·20	*k* 1·16	*lkn* 1·18	*pn* 1·15	*lp* 1·10	*l* 1·19

(Adapted from an experiment conducted by the Department of Agriculture, New Zealand)

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		read	
	0.17		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 - $\frac{148}{}$ = - 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.