Chapter 13

Fractional Factorials

13.1 Fractional replicates

A factorial design is a fractional replicate if not all possible combinations of the treatment factors occur. A fractional replicate can be useful if there are a large number of treatment factors to investigate and we can assume that some interactions are zero. In Chapter 9 we constructed some fractional replicate designs from Latin squares. Here we use characters to give us more types of fractional replicate.

Definition Let $\mathcal{T}'$ be any the subset of treatments in the design. Two characters F and G are *aliased* on $\mathcal{T}'$ if $F \equiv G$ on $\mathcal{T}'$. If F is aliased with I on $\mathcal{T}'$ (in other words, if F takes a single value throughout $\mathcal{T}'$) then F is a *defining contrast*.

If F and G are both defining contrasts then F and G both take constant values throughout $\mathcal{T}'$, so $F+G$ also take a constant value throughout $\mathcal{T}$. Thus the set of defining contrasts (including I) forms a subgroup of the group of all characters.

This suggests that we can use as a fractional replicate one block alone from a confounded block design.

Example 13.1 (Example 12.5 continued: Three factors with three levels) On any block from Example 12.5 we have

$$A+B+2C = \text{constant}.$$

So, on that block,

$$A = \text{constant} - (B+2C) = \text{constant} + C - B,$$

and thus

$$A \equiv 2B+C = C-B \equiv B+2C.$$

Hence we cannot distinguish between the effect of A and the effect of $B+2C$: we can estimate one only if the other is zero. But $B+2C$ is part of the B-by-C interaction, so we can estimate A if the B-by-C interaction is zero. Similarly, we can estimate B and C if the A-by-C and A-by-B interactions are zero. So this fraction consisting of nine treatments is a main-effects-only design.

Theorem 13.1 *Let $\mathcal{G}$ be a subgroup of the characters, and let $\mathcal{T}'$ be the set of treatments in any one block of the single-replicate block design which confounds $\mathcal{G}$. Then*

(i) two characters are aliased if and only if they are in the same coset of $\mathcal{G}$;

(ii) a character is a defining contrast if and only if it is in $\mathcal{G}$;

(iii) if two characters are neither defining contrasts nor aliased with each other then they are strictly orthogonal to each other on $\mathcal{T}'$.

We shall not prove this theorem here.

13.2 Choice of defining contrasts

Many choices of $\mathcal{G}$ are not suitable as defining-contrasts subgroups. If F is a defining contrast then its effect cannot be estimated. If F is aliased with G then the effect of one can be estimated only if the other can be assumed to be zero. Thus, in any coset of $\mathcal{G}$, *either* we must be able to assume that at most one character has a non-zero effect *or* we must accept that we cannot estimate the effect of anything in that coset. In this context, the cosets are known as *alias sets*.

Technique 13.2 (Fractional factorial designs) Given n treatment factors, each with p values, where p is prime, construct a fractional replicate design with p^{n-s} treatments as follows.

(i) Find a subgroup $\mathcal{G}$ of the set of all characters such that

- (a) every defining contrast (character in $\mathcal{G}$) *either* has zero effect *or* is not of interest;
- (b) of each pair of aliased characters (characters in the same coset of $\mathcal{G}$), *either* at most one has non-zero effect *or* neither is of interest;
- (c) $|\mathcal{G}| = p^s$.

(ii) Form the set $\mathcal{T}''$ consisting of those p^{n-s} treatments u satisfying

$$G(u) = 0 \quad \text{for all } G \text{ in } \mathcal{G}.$$

(iii) The desired fraction $\mathcal{T}'$ may be taken as $\mathcal{T}''$ or as any other coset of $\mathcal{T}''$ in $\mathcal{T}$.

Thus only a limited amount of information may be obtained from a fractional replicate, so the combination of too few plots with too many non-zero treatment effects may lead to an insoluble design problem.

The following concept is often useful in the seach for suitable groups of defining contrasts.

Definition The *weight* $w(F)$ of a character F is equal to m if F belongs to an m-factor interaction.

If there are no pseudofactors then $w(F)$ is just the number of non-zero coefficients in F. It must be defined more carefully if there are pseudofactors. In Example 12.10 the character $B+C+D_1+D_2$ has weight three, because only three letters are involved.

Example 13.2 (A two-level fraction) Suppose that A, B, C, D and E are factors with two levels, and that all interactions among them are zero except the A-by-B and C-by-D interaction. How small a fraction can we use to estimate all main effects and those two interactions?

Because all the treatment factors have two levels, all main effects and interactions have one degree of freedom. We want to estimate five main effects and two interactions, so we need at least seven degrees of freedom, and hence at least eight plots. A fraction with only eight treatments out of the possible 32 would be a quarter-replicate, so would require a defining-contrasts subgroup $\mathcal{G}$ of order four. Can we find such a subgroup $\mathcal{G}$ in which none of the seven characters

$$A, \quad B, \quad C, \quad D, \quad E, \quad A+B, \quad C+D$$

is aliased with another one?

If $\mathcal{G}$ contains A then A takes a constant value in the fraction (A is a defining contrast) so the main effect of A cannot be estimated. Thus $\mathcal{G}$ should contain no character of weight one. If $\mathcal{G}$ contains $A+B$ then

$$A = \text{constant} - B$$

so

$$A \equiv B$$

and so A and B are aliased and their main effects cannot be disentangled. (Another way to see this is to note that

$$A = (A+B)+B,$$

so that A and B are in the same coset of $\mathcal{G}$.) Thus $\mathcal{G}$ should contain no character of weight two. Thus $\mathcal{G}$ must contain (apart from I) only characters of weight three or more, apart from the single character I of weight zero.

Suppose that G and H are different characters in $\mathcal{G} \setminus \{I\}$. The weight of $G+H$ has the same parity as $w(G)+w(H)$ (because we are working modulo 2), and is at most five.

$$\begin{array}{l} \text{If } w(G)=5 \text{ and } w(H)=4 \text{ then } w(G+H)=1; \\ \text{if } w(G)=5 \text{ and } w(H)=3 \text{ then } w(G+H)=2; \\ \text{if } w(G)=4 \text{ and } w(H)=4 \text{ then } w(G+H)=2. \end{array}$$

Thus the only possibility for the three non-I characters in $\mathcal{G}$ is

$$\begin{array}{lllll} F_1 & +F_2 & +F_3 & & \\ F_1 & & & +F_4 & +F_5 \\ & F_2 & +F_3 & +F_4 & +F_5, \end{array}$$

where $(F_1,\ldots,F_5)$ is some reordering of (A,B,C,D,E).

If $\mathcal{G}$ contains $A+B+E$ then $A+B$ is aliased with E, because

$$A+B=(A+B+E)+E.$$

This is not allowed, because the A-by-B interaction is non-zero. Thus A and B are not both in $\{F_1,F_2,F_3\}$ or both in $\{F_1,F_4,F_5\}$. In particular, F_1 cannot be equal to A or to B. A similar argument shows that F_1 cannot be equal to C or to D, because the C-by-D interaction is non-zero. Hence $F_1=E$, and so

$$F_1+F_2+F_3+F_4=A+B+C+D.$$

But now $\mathcal{G}$ contains $A+B+C+D$, and

$$A+B=(A+B+C+D)+(C+D),$$

and so $A+B$ is aliased with $C+D$, which is not allowed if these two non-zero interactions are to be estimated. Hence no quarter-replicate will suffice.

We now try a fraction with 16 treatments, that is, a half-replicate. If we take $A+B+C+D+E$ as the single defining contrast (apart from I) then the alias sets are as shown in Table 13.1. The eight non-zero effects are in different alias sets and so are orthogonal to each other.

The design consists *either* of the 16 treatments u satisfying

$$(A+B+C+D+E)(u)=0$$

or of the 16 treatments u satisfying

$$(A+B+C+D+E)(u)=1.$$

The first set is shown in Table 13.2: it consists of those treatments which have an even number of coordinates equal to 1. The second set consists of those treatments with an odd number of coordinates equal to 1. For either fraction, the skeleton analysis of variance is shown in Table 13.3.

That example was given at some length, because it well demonstrates the kind of trial-and-error that is needed when specific effects are given as assumed non-zero. Several computer programs exist for carrying out such a search.

$$
\begin{array}{lcl}
\mathcal{G} & = & \{I, A+B+C+D+E\} \\
\mathcal{G}+A & = & \{A, B+C+D+E\} \\
\mathcal{G}+B & = & \{B, A+C+D+E\} \\
\mathcal{G}+C & = & \{C, A+B+D+E\} \\
\mathcal{G}+D & = & \{D, A+B+C+E\} \\
\mathcal{G}+E & = & \{E, A+B+C+D\} \\
\mathcal{G}+(A+B) & = & \{A+B, C+D+E\} \\
\mathcal{G}+(A+C) & = & \{A+C, B+D+E\} \\
\mathcal{G}+(A+D) & = & \{A+D, B+C+E\} \\
\mathcal{G}+(A+E) & = & \{A+E, B+C+D\} \\
\mathcal{G}+(B+C) & = & \{B+C, A+D+E\} \\
\mathcal{G}+(B+D) & = & \{B+D, A+C+E\} \\
\mathcal{G}+(B+E) & = & \{B+E, A+C+D\} \\
\mathcal{G}+(C+D) & = & \{C+D, A+B+E\} \\
\mathcal{G}+(C+E) & = & \{C+E, A+B+D\} \\
\mathcal{G}+(D+E) & = & \{D+E, A+B+C\}
\end{array}
$$

Table 13.1: Alias sets for the half-replicate in Example 13.2

00000,	11000,	10100,	10010,	10001,	01100,
01010,	01001,	00110,	00101,	00011,	11110,
11101,	11011,	10111,	01111		

Table 13.2: First half-replicate in Example 13.2

Stratum	Source	Degrees of freedom	
mean	mean		1
plots	A	1	
	B	1	
	C	1	
	D	1	
	E	1	
	$A \wedge B$	1	
	$C \wedge D$	1	
	residual	8	
	total		15
Total			16

Table 13.3: Analysis-of-variance table for the half-replicate in Example 13.2

13.3 Weight and resolution

In an experiment with several factors, sometimes the experimenter has no prior idea about which interactions are non-zero. Then it is common to assume that interactions involving more than a certain number of factors are all zero.

Definition A fraction has *resolution* M if every defining contrast (apart from I) has weight at least M.

Thus in a fraction of resolution $2m+1$, any character of weight m or less is aliased only with characters of weight at least $m+1$. Therefore all main effects and interactions between at most m factors can be estimated if all interactions between $m+1$ or more factors are assumed zero. In a fraction of resolution $2m$, any character of weight $m-1$ or less is aliased only with characters of weight $m+1$ or more, but some characters of weight m are aliased with each other. Therefore all main effects and interactions between at most $m-1$ factors can be estimated if all interactions between $m+1$ or more factors are assumed zero. Some authors take these properties as the definition of resolution, especially for fractions which are not subgroups of $\mathcal{T}$.

With either definition, a fraction of resolution M also has resolution $M-1$. Some authors limit the word *resolution* to the *maximum* value that we allow here for the resolution. Fractions with a given resolution are often tabulated.

13.4 Analysis of fractional replicates

The analysis of fractional replicates follows the same lines as the analysis of single replicates. If there are any alias sets in which all characters can be assumed to have zero effect then these can be used to estimate the error variance, which can in turn be used to test for the presence of effects. If there is an alias set in which only one character is assumed to have a non-zero effect then that effect can be estimated. If there are two or more potentially non-zero effects in an alias set then there is ambiguity: a large apparent effect may mean that one is large while the other is zero (but we cannot tell which is the large one!) or it may mean that both have small effects in the same direction. On the other hand, an apparently zero effect may mean that both are zero, or it may mean that they have effects of the same size but the opposite sign.

Example 13.3 (Chromatograph) Six factors were altered in an experiment on a chromatograph. The correspondence between the actual levels and the integers modulo 2 was as follows.

Factor	0	1
Flow rate (F)	0.5 ml/min	1.0 ml/min
Temperature (T)	25°C	45°C
Linear gradient time (L)	5 min	15 min
Injection volume (V)	0.5 ml	10 ml
Initial organic phase concentration (C)	0%	10 %
%TFA in mobile phase (M)	0.01%	0.2%

It was assumed that all interactions between three or more factors were zero, but that two-factor interactions might not be. Therefore a fraction with resolution 4 was required. Defining contrasts $T+L+V+C$, $F+T+C+M$ and $F+L+V+M$ were chosen.

The experiment consisted of the sixteen runs u for which

$$(T+L+V+C)(u) = 0 = (F+T+C+M)(u).$$

The chromatographic response function was measured on each run. The treatments and responses are shown in Table 13.4.

F	T	L	V	C	M	response
0	1	1	1	1	0	220
0	0	0	0	0	0	174
1	1	0	0	1	1	172
1	0	1	1	0	1	353
1	1	0	1	0	0	176
0	0	0	1	1	1	192
0	1	1	0	0	1	280
1	0	1	0	1	0	246
1	0	0	1	1	0	197
0	1	0	1	0	1	192
0	0	1	0	1	1	261
1	1	1	0	0	0	340
0	1	0	0	1	0	200
1	0	0	0	0	1	233
0	0	1	1	0	0	280
1	1	1	1	1	1	234

Table 13.4: Treatments and responses in Example 13.3

F	152	$F+T$	92	$F+T+L$	124
T	-122	$F+L$	-112	$F+T+V$	-68
L	678	$F+V$	0		
V	-62	$F+C$	200		
C	-306	$F+M$	18		
M	84	$T+L$	10		
		$T+V$	278		

Table 13.5: Difference between total response on level 1 and total response on level 0

stratum	source	sum of squares	degrees of freedom	mean square	variance ratio
mean	mean	878906.25	1	878906.25	–
runs	F	1444.00	1	1444.00	2.31
	T	930.25	1	930.25	1.49
	L	28730.25	1	28730.25	45.97
	V	240.25	1	240.25	0.38
	C	5852.25	1	5852.25	9.36
	M	441.00	1	441.00	0.71
	$F+T \equiv C+M$	529.00	1	529.00	0.85
	$F+L \equiv V+M$	784.00	1	784.00	1.25
	$F+V \equiv L+M$	0.00	1	0.00	0.00
	$F+C \equiv T+M$	2500.00	1	2500.00	4.00
	$F+M \equiv T+C \equiv L+V$	20.25	1	20.25	0.03
	$T+L \equiv V+C$	6.25	1	6.25	0.01
	$T+V \equiv L+C$	4830.25	1	4830.25	7.73
	residual	1250.00	2	625.00	–
	total	47557.75	15		
Total		926464.00	16		

Table 13.6: Analysis of variance in Example 13.3

Table 13.5 shows, for one character H in each alias set, the difference $\mathsf{sum}_{H=1} - \mathsf{sum}_{H=0}$. The sum of squares for this character is the square of this difference, divided by 16. Hence we obtain the analysis-of-variance in Table 13.6. The characters $F+T+L$ and $F+T+V$ both belong to alias sets with no characters of weight one or two, so these are both used for residual. Each main effect contributes one line to the analysis-of-variance table. Each of the remaining lines corresponds to two or three of the two-factor interactions.

The problem with such an analysis-of-variance table is that the F^1_2 distribution is very heavy-tailed, so it is hard to reject any null hypothesis of zero effect. The 95% point is 18.51 and the 90% point is 8.53. Thus the main effect of F is nonzero, and probably also the main effect of C and the effect caused by the aliased characters $C+L$ and $T+V$. Since neither the main effect of T nor the main effect of V seems to be nonzero, it is fairly safe to attribute this third nonzero effect to the interaction between C and L.

The quantile plot gives a similar conclusion. The mean squares for all fifteen characters are shown in Figure 13.1: the points do not lie on a straight line through the origin. However, removing the top three points gives the graph in Figure 13.2, which looks much more like such a line.

The conclusion from this experiment may well be that the experimenter should investigate factors L and C further. In this case, he might do well to include F in

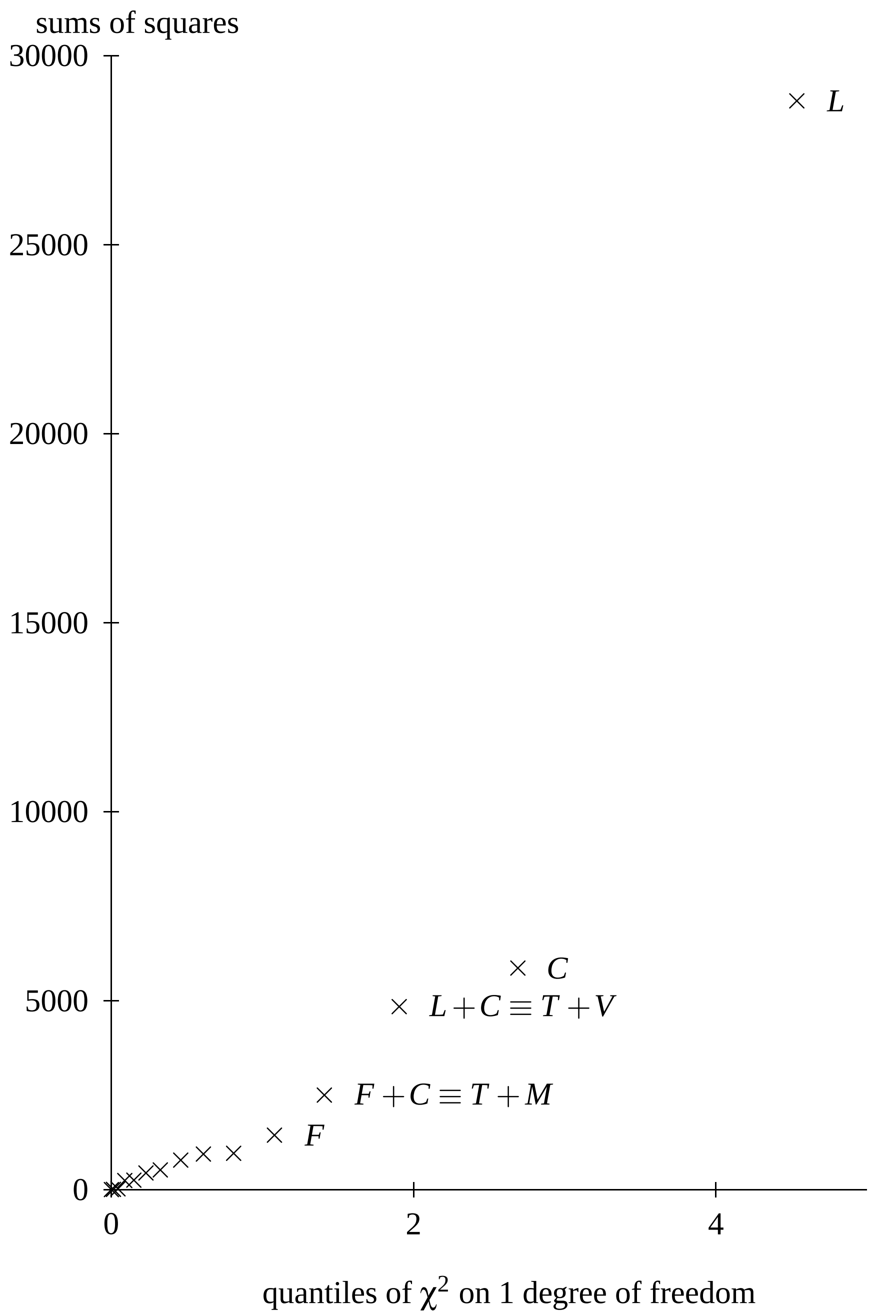

Figure 13.1: Plot of all sums of squares in Example 13.3

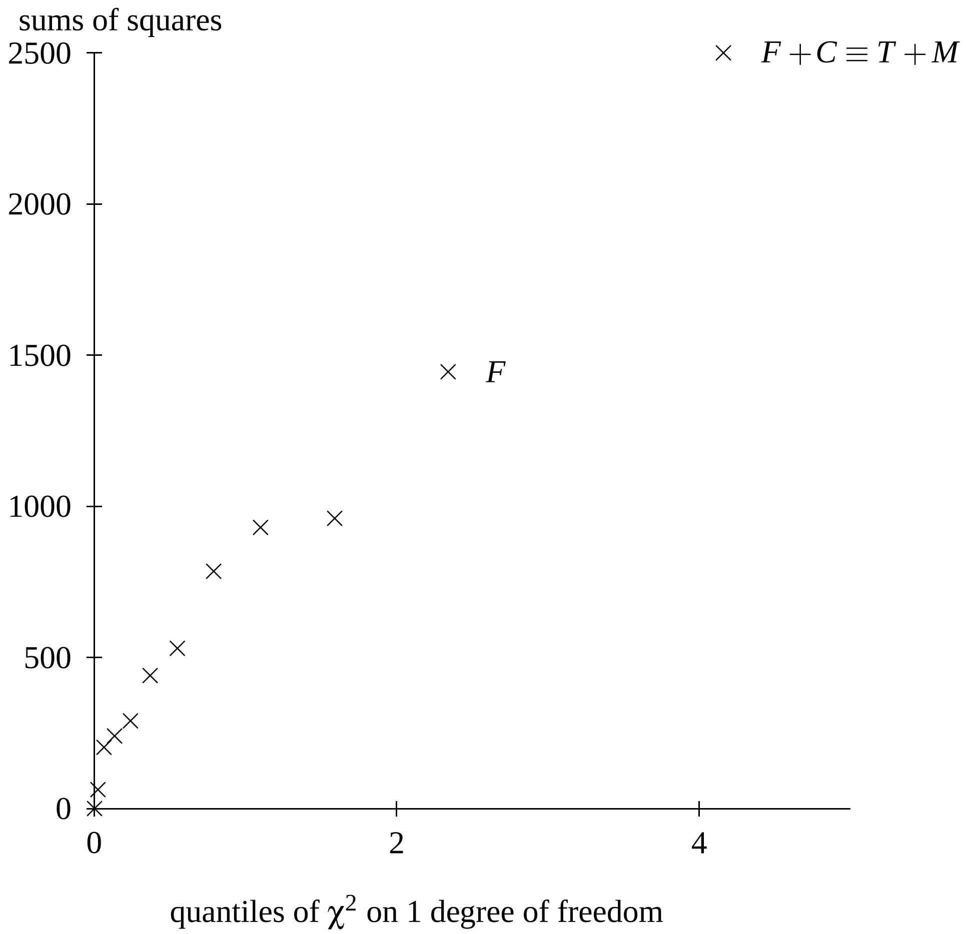

Figure 13.2: Plot of all sums of squares except the top three in Example 13.3

these further studies. The fouth largest effect is for $T+M$, which is aliased with $C+F$, and the fifth is for the main effect of F. These five effects are consistent with the model $V_{C\wedge L}+V_{C\wedge F}$.

Questions for Discussion

13.1 Five two-level factors A, B, C, D, E are to be tested for their effect on an industrial process. It is believed that the only non-zero interactions are the two-factor interactions A-by-B and A-by-C. Using the method of characters and aliasing, construct a fractional design which can be used to estimate all main effects and both non-zero two-factor interactions, and which is as small as possible.

13.2 Construct a quarter-replicate main-effects-only design for two two-level treatment factors and two four-level treatment factors.