CHAPTER 10

Factorial Designs at Two Levels

A class of designs is now discussed that are of great practical importance—two-level factorial designs. In this chapter the uses, generation, and analysis of these designs are considered.

10.1. GENERAL FACTORIAL DESIGNS AND DESIGNS AT TWO LEVELS

To perform a general factorial design, an investigator selects a fixed number of "levels" (or "versions") for each of a number of variables (factors) and then runs experiments with all possible combinations. If there are l_1 levels for the first variable, l_2 for the second, ..., and l_k for the kth, the complete arrangement of $l_1 \times l_2 \times \cdots \times l_k$ experimental runs is called an $l_1 \times l_2 \times \cdots \times l_k$ factorial design. For example, a $2 \times 3 \times 5$ factorial design requires $2 \times 3 \times 5 = 30$ runs, and a $2 \times 2 \times 2 = 2^3$ factorial design 8 runs. An application and analysis of a 3×4 factorial design in four replications was given in Section 7.7. In the present chapter we discuss designs in which each variable occurs at only two levels. These designs are of importance for a number of reasons.

1. They require relatively few runs per factor studied; and although they are unable to explore fully a wide region in the factor space, they can indicate major trends and so determine a promising direction for further experimentation.
2. We see in Chapter 15 that, when a more thorough local exploration is needed, they can be suitably augmented to form composite designs.

3. In Chapter 12 we see that they form the basis for two-level *fractional* factorial designs. These fractional designs are often of great value at an early stage of an investigation, when it is frequently good practice to use a preliminary experimental effort to look at a large number of factors superficially rather than a small number (which may or may not include the important ones) thoroughly.
4. These designs and the corresponding fractional designs may be used as building blocks so that the degree of complexity of the finally constructed design can match the sophistication of the problem.
5. The interpretation of the observations produced by the designs can proceed largely by using common sense and elementary arithmetic.

In all these applications the designs fit naturally into the sequential strategy discussed in Chapter 1, which is an essential feature of the scientific process.

10.2. AN EXAMPLE OF A 2^3 FACTORIAL DESIGN: PILOT PLANT INVESTIGATION

Table 10.1*a* shows a 2^3 factorial experiment in which there are two quantitative variables—temperature and concentration—and a single qualitative variable—catalyst. The response is the chemical yield. The original data were part of a pilot plant investigation of a process and have been simplified somewhat for illustrative purposes. Table 10.1*b* shows the recorded data with the levels coded so that for the quantitative variables a minus sign represents the low level and a plus sign the high level. For a qualitative variable the two versions or "levels" can also be conveniently coded by minus and plus signs. It does not matter here which is associated with the plus as long as the labeling is consistent. A display of levels to be run in a design such as is given in Table 10.1 is called a *design matrix*.

Exercise 10.1. How many variables and how many runs are there in a $2 \times 4 \times 3 \times 2$ factorial design? *Answer*: 4 variables, 48 runs.

Exercise 10.2. If four variables are to be studied using a three-level factorial design (all variables are studied at three levels), how will we designate the design and how many runs will it require? *Answer*: 3^4 design, 81.

There are other notations in common use for the design matrix of a two-level factorial. One identifies the "upper" level of each factor by the use of the corresponding lower case letter. Another notation uses a 0 and 1 in place of our − and + signs. The three alternative notations for a 2^3 factorial design are as follows:

Run	T	C	K		T	C	K
1	−	−	−	1	0	0	0
2	+	−	−	t	1	0	0
3	−	+	−	c	0	1	0
4	+	+	−	tc	1	1	0
5	−	−	+	k	0	0	1
6	+	−	+	tk	1	0	1
7	−	+	+	ck	0	1	1
8	+	+	+	tck	1	1	1

TABLE 10.1. Data from a 2^3 factorial design, pilot plant example

test condition number	temperature (°C) **T**	concentration (%) **C**	catalyst (*A* or *B*) **K**	yield (grams) **y**
		a. Original units of variables		
1	160	20	*A*	60
2	180	20	*A*	72
3	160	40	*A*	54
4	180	40	*A*	68
5	160	20	*B*	52
6	180	20	*B*	83
7	160	40	*B*	45
8	180	40	*B*	80
		b. Coded units of variables		
1	−	−	−	60
2	+	−	−	72
3	−	+	−	54
4	+	+	−	68
5	−	−	+	52
6	+	−	+	83
7	−	+	+	45
8	+	+	+	80

temperature (°C)		*concentration (%)*		*catalyst*	
−	+	−	+	−	+
160	180	20	40	*A*	*B*

We prefer the ± notation because it relates to a geometric view of the design and, as we will see later, is also readily applicable to regression analysis and to the construction of fractional factorial designs.

10.3. CALCULATION OF MAIN EFFECTS

Averaging Individual Measures of Effects

What can we find from this factorial design? For example, what does it tell us about the *effect* of temperature on yield? (By the "effect" of a factor we mean the change in the response as we move from the − to the + version of that factor, here from the low to the high level of temperature.) Consider the first two tests in Table 10.1. Aside from experimental error the corresponding yields (60 and 72) differ only because of temperature. The concentration (20%) and the catalyst (*A*) are the same for both of these conditions. Altogether there are four measures of the temperature effect at each of the four combinations of conditions of the other variables as listed below.

individual measure of the effect of changing temperature from 160 to 180°C	condition at which comparison is made: concentration **C**	catalyst **K**
$y_2 - y_1 = 72 - 60 = 12$	20	*A*
$y_4 - y_3 = 68 - 54 = 14$	40	*A*
$y_6 - y_5 = 83 - 52 = 31$	20	*B*
$y_8 - y_7 = 80 - 45 = \underline{35}$	40	*B*
main effect of temperature $T = 23$		

The average of these four measures (+23 for this example) is called the *main effect* of temperature and is denoted by *T*. It measures the *average* effect of temperature over all conditions of the other variables.

Because of the general symmetry of the design (see Figure 10.1), there is a similar set of four measures for the effect of concentration, for each of which the levels of the remaining variables are constant.

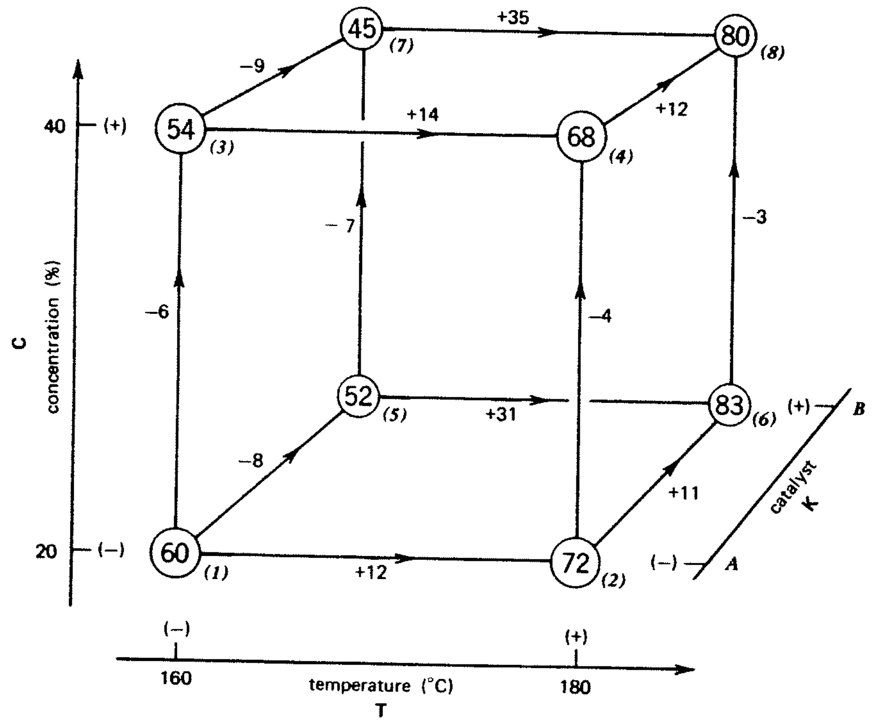


FIGURE 10.1. 2^3 factorial design, pilot plant example.

individual measure of the effect of changing concentration from 20 to 40%	condition at which comparison is made: temperature **T**	catalyst **K**
$y_3 - y_1 = 54 - 60 = -6$	160	*A*
$y_4 - y_2 = 68 - 72 = -4$	180	*A*
$y_7 - y_5 = 45 - 52 = -7$	160	*B*
$y_8 - y_6 = 80 - 83 = \underline{-3}$	180	*B*
main effect of concentration $C = -5$		

Finally (see Figure 10.1), there are four measures of the effect of catalyst.

individual measures of the effect of changing from catalyst *A* to catalyst *B*	condition at which comparison is made: temperature **T**	concentration **C**
$y_5 - y_1 = 52 - 60 = -8$	160	20
$y_6 - y_2 = 83 - 72 = 11$	180	20
$y_7 - y_3 = 45 - 54 = -9$	160	40
$y_8 - y_4 = 80 - 68 = \underline{12}$	180	40
main effect of catalyst $K = 1.5$		

Difference between Two Averages

The main effect for each of the variables is seen to be the difference between two averages:

$$\text{main effect} = \bar{y}_+ - \bar{y}_- \qquad (10.1)$$

where $\bar{y}_+$ is the average response for the plus level of the variable and $\bar{y}_-$ is the average response for the minus level. Thus (see Figure 10.2*a*)

temperature effect $$T = \frac{72 + 68 + 83 + 80}{4} - \frac{60 + 54 + 52 + 45}{4}$$
$$= 75.75 - 52.75 = 23$$

concentration effect $$C = \frac{54 + 68 + 45 + 80}{4} - \frac{60 + 72 + 52 + 83}{4}$$
$$= 61.75 - 66.75 = -5$$

catalyst effect $$K = \frac{52 + 83 + 45 + 80}{4} - \frac{60 + 72 + 54 + 68}{4}$$
$$= 65.0 - 63.5 = 1.5$$

Notice that (1) *all* the observations are being used to supply information on *each* of the main effects, and (2) each effect is determined with the precision of a fourfold replicated difference.

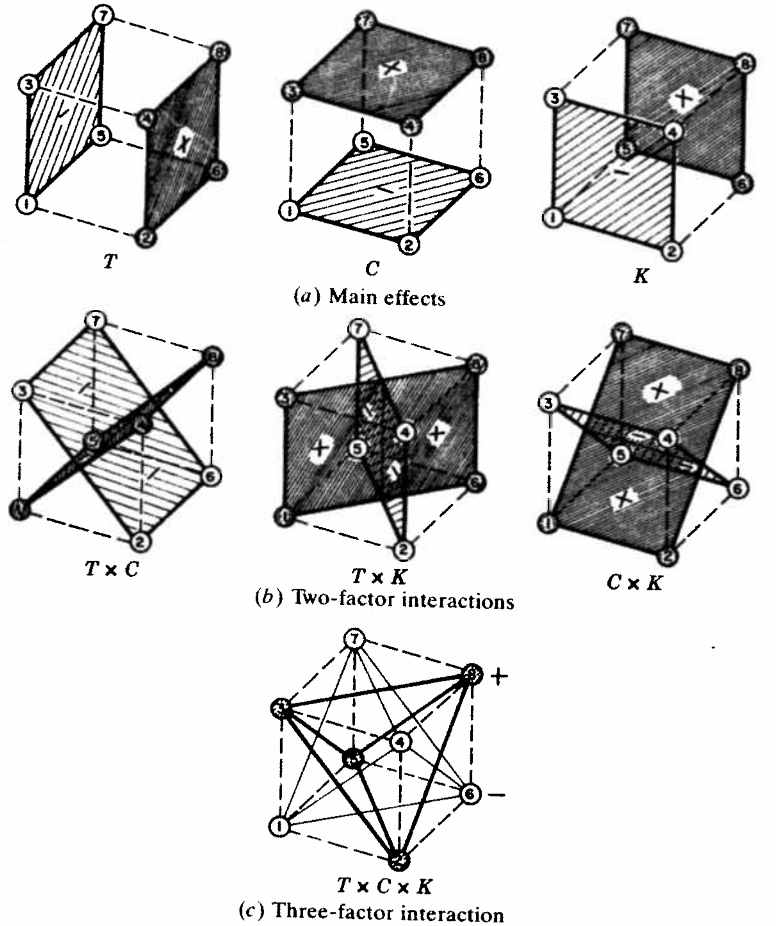


FIGURE 10.2. Geometric representation of contrasts corresponding to main effects and interactions.

Advantages over the "One-Factor-at-a-Time" Method

Suppose that in the above investigation, instead of a factorial arrangement, the "one-factor-at-a-time" method had been used. This method, in which experimental factors are varied one at a time, with the remaining factors held constant, was formerly regarded as the only correct way to conduct research (see Section 15.1 for further discussion of this procedure). The method provides an estimate of the effect of a single variable at selected *fixed* conditions of the other variables. However, for such an estimate to have general relevance it is necessary to assume that the effect would be the same at other settings of the other variables—that, over the ranges of current

interest, the variables act on the response additively. However, (1) if the variables *do* act additively, the factorial does the job with more precision; and (2) if the variables do *not* act additively, the factorial, unlike the one-factor-at-a-time design, can detect and estimate interactions that *measure* the nonadditivity.

Gain in Precision If Variables Act Additively

To secure the same precision for the estimate of the temperature effect the one-factor-at-a-time experiment would need to employ eight runs, four at each level of temperature, with all the observations made at some arbitrarily *fixed* levels of concentration and catalyst. In a similar manner two further sets of eight runs would be required to study concentration and catalyst. Thus to obtain estimates of the main effects of three variables with the same precision as is provided by the 2^3 factorial design, the one-factor-at-a-time method would require 24 runs—a threefold increase. In general, for k factors a k-fold increase would be required. Some economy can be introduced by using a single experimental condition from which to make all changes; even with this arrangement, however, the one-factor-at-a-time design requires $(k + 1)/2$ times as many runs as the factorial.

10.4. INTERACTION EFFECTS

Two-Factor Interactions

In the example, the average effect of temperature is 23. It is obvious from the data, however, that the temperature effect is much greater with catalyst B than with catalyst A. The variables temperature and catalyst do not behave additively and are therefore said to "interact." A measure of this interaction is supplied by the difference between the average temperature effect with catalyst A and the average temperature effect with catalyst B. By convention, *half* the difference is called the *temperature by catalyst interaction* or, in symbols, the $T \times K$ interaction.

catalyst	average temperature effect
$(+)B$	33
$(-)A$	13
difference	20

$$T \times K \text{ interaction} = \frac{20}{2} = 10 \qquad (10.2)$$

The $T \times K$ interaction may equally well be thought of as one-half the difference in the average *catalyst* effects at the two levels of *temperature*.

temperature (°C)	average catalyst effect
(+)180	11.5
(−)160	−8.5
difference	20.0

$$T \times K \text{ interaction} = \frac{20}{2} = 10 \tag{10.3}$$

That the results must be equivalent is seen as soon as we consider the way the interaction employs the observations. Consider the calculations leading to Equation 10.2.

catalyst	average temperature effect
(+)B	$33 = \frac{31 + 35}{2} = \frac{1}{2}(y_6 - y_5 + y_8 - y_7)$
(−)A	$13 = \frac{12 + 14}{2} = \frac{1}{2}(y_2 - y_1 + y_4 - y_3)$

$$T \times K \text{ interaction} = \tfrac{1}{2}\text{ difference} = \frac{33 - 13}{2} = 10$$

$$= \tfrac{1}{4}(y_1 - y_2 + y_3 - y_4 - y_5 + y_6 - y_7 + y_8)$$

$$= \frac{y_1 + y_3 + y_6 + y_8}{4} - \frac{y_2 + y_4 + y_5 + y_7}{4} \tag{10.4}$$

The reader should confirm that exactly the same result is obtained by carrying through the same argument for Equation 10.3.

Like the main effects, the interaction effect is seen to be a difference between two averages, half of the eight results being included in one average and half in the other. Just as main effects may be viewed as a *contrast* between observations on parallel faces of the cube, as shown in Figure 10.2*a*, the $T \times K$ interaction is a contrast between results on two *diagonal planes*, as shown in Figure 10.2*b*. The $T \times C$ and the $C \times K$ interactions are obtained in a similar way.

concentration	average temperature effect	catalyst	average concentration
(+) 40%	24.5	(+)B	−5.0
(−) 20%	21.5	(−)A	−5.0
difference	3.0	difference	0.0

$$T \times C \text{ interaction} = \frac{3}{2} = 1.5 \qquad C \times K \text{ interaction} = \frac{0}{2} = 0 \tag{10.5}$$

Also, as illustrated in Figure 10.2*b*,

$$T \times C \text{ interaction} = \frac{y_1 + y_4 + y_5 + y_8}{4} - \frac{y_2 + y_3 + y_6 + y_7}{4}$$
$$C \times K \text{ interaction} = \frac{y_1 + y_2 + y_7 + y_8}{4} - \frac{y_3 + y_4 + y_5 + y_6}{4} \tag{10.6}$$

Three-Factor Interaction

Consider the temperature by concentration ($T \times C$) interaction. Two measures of the $T \times C$ interaction are available from the experiment, one for each catalyst.

$T \times C$ interaction with catalyst B (+):

$$\frac{(y_8 - y_7) - (y_6 - y_5)}{2} = \frac{(80 - 45) - (83 - 52)}{2} = \frac{35 - 31}{2} = 2 \tag{10.7}$$

$T \times C$ interaction with catalyst A (−):

$$\frac{(y_4 - y_3) - (y_2 - y_1)}{2} = \frac{(68 - 54) - (72 - 60)}{2} = \frac{14 - 12}{2} = 1 \tag{10.8}$$

The difference measures the consistency of the *temperature by concentration interaction* for the two catalysts. Half this difference is defined as the *three-factor interaction* of temperature, concentration, and catalyst, denoted as the $T \times C \times K$ interaction. Thus

$$T \times C \times K \text{ interaction} = \frac{2 - 1}{2} = 0.5 \tag{10.9}$$

As before, this interaction is symmetric in all the variables. For example, it could equally well have been defined as half the difference between the temperature by catalyst interactions at each of the two concentrations. The estimate of this three-factor interaction can again be reduced to the difference between two averages.* If the experimental points contributing to the two averages are isolated, they define the vertices of the two tetrahedra in Figure 10.2*c*, which together comprise the cube. It is helpful to remember that for any main effect or interaction the $\bar{y}_+$ average *always* contains the observation from the run in which all variables are at their plus levels, for example, for a 2^3 factorial, the $(+\ +\ +)$ run.

Exercise 10.3. Calculate the main effects and interactions for the following data:

test condition number	brand of popcorn **1**	ratio of popcorn to oil **2**	batch size (cup) **3**	yield of popcorn **y**
1	− (ordinary)	− (low)	− ($\frac{1}{3}$)	$6\frac{1}{4}$
2	+ (gourmet)	− (low)	− ($\frac{1}{3}$)	8
3	− (ordinary)	+ (high)	− ($\frac{1}{3}$)	6
4	+ (gourmet)	+ (high)	− ($\frac{1}{3}$)	$9\frac{1}{2}$
5	− (ordinary)	− (low)	+ ($\frac{2}{3}$)	8
6	+ (gourmet)	− (low)	+ ($\frac{2}{3}$)	15
7	− (ordinary)	+ (high)	+ ($\frac{2}{3}$)	9
8	+ (gourmet)	+ (high)	+ ($\frac{2}{3}$)	17

Answer: $1 = 5.1$, $2 = 1.1$, $3 = 4.8$, $1 \times 2 = 0.7$, $1 \times 3 = 2.4$, $2 \times 3 = 0.4$, $1 \times 2 \times 3 = -0.2$.

Notation

In what follows italic type is used to denote the effects of factors (variables). Bold face type is used to denote the factors themselves and the columns of signs that determine their levels. Thus T is the main effect of variable **T**.

* $T \times C \times K$ interaction $= \frac{1}{4}(-y_1 + y_2 + y_3 - y_4 + y_5 - y_6 - y_7 + y_8)$

$$= \frac{y_2 + y_3 + y_5 + y_8}{4} - \frac{y_1 + y_4 + y_6 + y_7}{4}$$

10.5. INTERPRETATION OF RESULTS

The results of the factorial analysis are collected in Table 10.2. A point we have not yet mentioned about the data is that each of the eight yield values in Table 10.1 was actually an average of two replicate runs. We show in Section 10.6 how, by using this fact, we can obtain the standard errors shown in Table 10.2. For the moment we use these standard errors to complete an analysis of the data. Comparison of the estimates with their standard errors (see also Figure 10.3) suggests that the circled items T, C, and the interaction $T \times K$ require interpretation, while the apparent effects remaining could be generated by the noise.

TABLE 10.2. Calculated effects and standard errors for the 2^3 factorial design, pilot plant example

effect	estimate ± standard error
average	64.25 ± 0.7
main effects	
temperature T	(23.0 ± 1.4)
concentration C	(−5.0 ± 1.4)
catalyst K	1.5 ± 1.4
two-factor interactions	
$T \times C$	1.5 ± 1.4
$T \times K$	(10.0 ± 1.4)
$C \times K$	0.0 ± 1.4
three-factor interaction	
$T \times C \times K$	0.5 ± 1.4

The main effect of a variable should be individually interpreted only if there is no evidence that the variable interacts with other variables. When

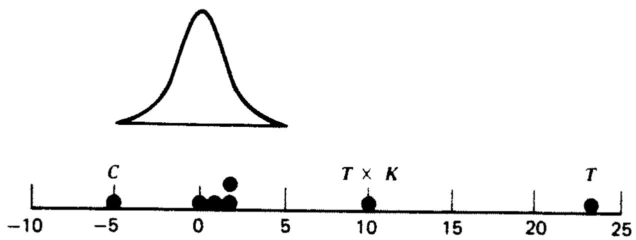


FIGURE 10.3. Main effects and interactions in relation to a reference t distribution with eight degrees of freedom and scale factor 1.4, pilot plant example.

there is evidence of one or more such interaction effects, the interacting variables should be considered jointly.

In Table 10.2 there is a large temperature effect, 23.0 ± 1.4. But since temperature interacts with catalyst type (the $T \times K$ interaction is 10.0 ± 1.4), we make no statement about the effect of temperature alone. The main effect of concentration is −5.0 ± 1.4, and in this case there is no evidence of any interactions involving concentration. Thus we draw the following tentative conclusions:

1. The effect of concentration (**C**) is to reduce the yield by about five units, and this is approximately so irrespective of the tested levels of the other variables.
2. The effects of temperature (**T**) and catalyst (**K**) cannot be interpreted separately because of the large $T \times K$ interaction, and can best be considered using the two-way table shown in Figure 10.4. The interaction evidently arises from a difference in sensitivity to temperature change for the two catalysts. With catalyst A the temperature effect is 13 units, but with catalyst B it is 33 units.

The result of most practical interest was the very different behaviors of the two "catalyst types" in response to temperature. The effect was unexpected, for although obtained from two different suppliers, the catalysts

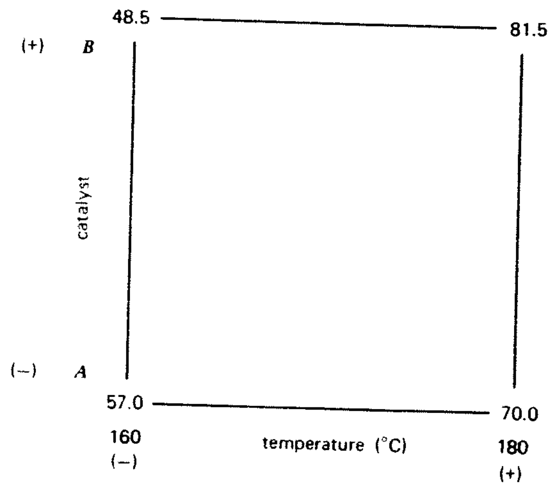


FIGURE 10.4. The temperature–catalyst interaction, pilot plant example.

were supposedly identical. Also, the yield from catalyst B at 180°C was the highest that, up to that time, had been seen. The finding led to a very careful study of the catalyst in further iterations of the investigation.

10.6. CALCULATION OF STANDARD ERRORS FOR EFFECTS USING REPLICATED RUNS

When genuine run replicates are made under a given set of experimental conditions, the variation between their associated observations may be used to estimate the standard deviation of a single observation and hence the standard deviation of the effects. By *genuine* run replicates we mean that variation between runs made at the same experimental conditions is a reflection of the *total* variability afflicting runs made at different experimental conditions. This point requires careful consideration.

Randomization of run order usually ensures that replicates are genuine. Consider the pilot plant example. A pilot plant run consisted of (1) cleaning the reactor, (2) inserting the appropriate catalyst charge, (3) running the apparatus at a given temperature and a given feed concentration for 3 hours to allow the process to settle down at the chosen experimental conditions, (4) sampling the output every 15 minutes during the final hour of running, and (5) combining chemical analyses made on these samples. A genuine run replicate must involve the taking of all these steps again. In particular, several chemical analyses from a single run would provide only an estimate of *analytical* variance, usually only a small part of the run-to-run variance. Similarly, several samples from the same run could provide only an estimate of sampling plus analytical variance. Generally this problem of wrongly assessing experimental error variance has been particularly troublesome. It has led, for instance, to gross underestimates of the errors associated with such quantities as the astronomical unit (see Youden, 1972, and the references therein).

In general, if g sets of experimental conditions are genuinely replicated and the n_i replicate runs made at the ith set yield an estimate s_i^2 of σ^2 having $\nu_i = n_i - 1$ degrees of freedom, the pooled estimate of run variance is

$$s^2 = \frac{\nu_1 s_1^2 + \nu_2 s_2^2 + \cdots + \nu_g s_g^2}{\nu_1 + \nu_2 + \cdots + \nu_g} \qquad (10.10)$$

with $\nu = \nu_1 + \nu_2 + \cdots + \nu_g$ degrees of freedom.

With only $n_i = 2$ replicates at each of the g sets of conditions, the formula for the ith variance reduces to $s_i^2 = d_i^2/2$ with $\nu_i = 1$, where d_i is the difference between the duplicate observations for the ith set of conditions. Thus

TABLE 10.3. Estimation of the variance, pilot plant example

average response value (previously used in the analysis)	T	C	K	results from individual runs*		difference of duplicate	estimated variance at each set of conditions s_i^2 = (difference)^2/2
60	−	−	−	$59^{(6)}$	$61^{(13)}$	−2	2
72	+	−	−	$74^{(2)}$	$70^{(4)}$	4	8
54	−	+	−	$50^{(1)}$	$58^{(16)}$	−8	32
68	+	+	−	$69^{(5)}$	$67^{(10)}$	2	2
52	−	−	+	$50^{(8)}$	$54^{(12)}$	−4	8
83	+	−	+	$81^{(9)}$	$85^{(14)}$	−4	8
45	−	+	+	$46^{(3)}$	$44^{(11)}$	2	2
80	+	+	+	$79^{(7)}$	$81^{(15)}$	−2	2
						total	64

s^2 = pooled estimate of σ^2 = average of estimated variances
= $\frac{64}{8}$ = 8 with ν = 8 degrees of freedom

* Superscripts give the order in which the runs were made.

Equation 10.10 yields $s^2 = \sum d_i^2/2g$. Using the replicated pilot plant data displayed in Table 10.3, where $g = 8$, we obtain $s^2 = 128/(2 \times 8) = 8$ and $s = 2.8$ with $\nu = 8$ degrees of freedom.

Since each main effect and interaction is a statistic of the form $\bar{y}_+ - \bar{y}_-$, where each average contains eight observations, the variance of each effect (assuming independent errors) is given by

$$V(\text{effect}) = V(\bar{y}_+ - \bar{y}_-) = (\tfrac{1}{8} + \tfrac{1}{8})\sigma^2 = \tfrac{1}{4}\sigma^2$$

In general, if a total of N runs is made in conducting a two-level factorial or replicated factorial design, then

$$V(\text{effect}) = \frac{4}{N}\sigma^2$$

For the pilot plant data, on substituting for σ^2 the estimate $s^2 = 8$, the estimated variance of an effect is $(\frac{1}{8} + \frac{1}{8})8 = 2$ (or, equivalently, $\frac{4}{16} \times 8 = 2$). Thus the estimated standard error of an effect is $\sqrt{2.0} = 1.4$, the value used in Table 10.2.

From equation 3.16, the variance of a mean is σ^2/N. Thus the estimated standard error of the mean is $s/\sqrt{N} = 2.8/\sqrt{16} = 0.7$, the value used in Table 10.2.

Exercise 10.4. For the following data calculate the main effects and interactions and their standard errors:

test condition number	depth of planting (in.) 1	watering (times daily) 2	type of lima bean 3	yield replication 1	yield replication 2	yield replication 3
1	− (½)	− (once)	− (baby)	6	7	6
2	+ (3/2)	− (once)	− (baby)	4	5	5
3	− (½)	+ (twice)	− (baby)	10	9	8
4	+ (3/2)	+ (twice)	− (baby)	7	7	6
5	− (½)	− (once)	+ (large)	4	5	4
6	+ (3/2)	− (once)	+ (large)	3	3	1
7	− (½)	+ (twice)	+ (large)	8	7	7
8	+ (3/2)	+ (twice)	+ (large)	5	5	4

Answer: *1* = −2.2, *2* = 2.5, *3* = −2.0, *1* × *2* = −0.3, *1* × *3* = −0.2, *2* × *3* = 0.2, *1* × *2* × *3* = 0.0, standard error of effect = 0.3.

Exercise 10.5. Repeat Exercise 10.4, assuming that data for replication 3 are unavailable.
Answer: *1* = −2.1, *2* = 2.6, *3* = −1.9, *1* × *2* = −0.4, *1* × *3* = 0.1, *2* × *3* = −0.1, *1* × *2* × *3* = −0.1, standard error of effect = 0.28.

Exercise 10.6. Suppose that in addition to the eight runs in Exercise 10.3 the following popcorn yields were obtained from genuine replicate runs made at the center conditions **1** = a mixture of 50% ordinary and 50% gourmet popcorn, **2** = medium, **3** = ½ cup: 9, 8, 9½, and 10 cups. Use these extra runs to calculate standard errors for main effects and interactions. Plot the effects in relation to an appropriate reference distribution, and draw conclusions.
Answer: Main effects and interactions remain unchanged. They are given in Exercise 10.3. Estimate of σ^2 from center points $= s^2 = 0.73$ with three degrees of freedom. Standard error of an effect = 0.60. Reference t distribution has three degrees of freedom, scale factor = 0.60.

Exercise 10.7. Show that the sum of squares in the analysis of variance table associated with any effect from any two-level factorial design containing (including possible replication) a total of N runs = N × (estimated effect)^2/4.

Exercise 10.8. Given the data in Table 10.3, an alternative way to estimate the variance is to set up a one-way classification analysis of variance table. For the 16 observations and 8 treatments of this example, complete the analysis of variance table and show that $s^2 = 8$ with $\nu = 8$ degrees of freedom.

Economy in Experimentation

In most situations there are more factors to be investigated than can conveniently be accommodated with the time and budget available. Rather than duplicate a 2^3 factorial as was done in the pilot plant study, it is usually better to include a fourth variable and run an unreplicated 2^4 design. Or, as we shall see in Chapter 12, it may be even better to run a half-replicated 2^5 design and use the 16 runs to study five factors. The reader at this point may be worried about obtaining an estimate of error if there is no replication. We show in Section 10.8 how it is usually possible to overcome this difficulty.

10.7. QUICKER METHODS FOR CALCULATING EFFECTS

It would be extremely tedious if effects had to be calculated from first principles whenever a factorial design was analyzed. Fortunately this is unnecessary. We now describe two quicker methods: one employs a table of contrast coefficients; the other, Yates's algorithm.

Table of Contrast Coefficients

The calculations performed to obtain the various effects can be characterized by the table of signs in Table 10.4.

TABLE 10.4. Signs for calculating effects from the 2^3 factorial design, pilot plant example

	mean	T	C	K	TC	TK	CK	TCK	yield averages
	+	−	−	−	+	+	+	−	60
	+	+	−	−	−	−	+	+	72
	+	−	+	−	−	+	−	+	54
	+	+	+	−	+	−	−	−	68
	+	−	−	+	+	−	−	+	52
	+	+	−	+	−	+	−	−	83
	+	−	+	+	−	−	+	−	45
	+	+	+	+	+	+	+	+	80
divisor	8	4	4	4	4	4	4	4	

Thus the estimate of the mean is calculated from the first column,

$$\frac{+60 + 72 + 54 + 68 + 52 + 83 + 45 + 80}{8} = \frac{514}{8} = 64.25 \quad (10.11)$$

The T main effect is calculated from the second column,

$$\frac{-60 + 72 - 54 + 68 - 52 + 83 - 45 + 80}{4} = \frac{92}{4} = 23.0 \quad (10.12)$$

and the $T \times K$ interaction from the sixth column,

$$\frac{+60 - 72 + 54 - 68 - 52 + 83 - 45 + 80}{4} = \frac{40}{4} = 10.0 \quad (10.13)$$

The remaining effects are obtained similarly to yield the estimates already displayed in Table 10.2.

The signs for the interactions reveal a remarkable fact: they can be obtained by directly multiplying the signs of their respective variables! Thus the array of signs for the $T \times K$ interaction is obtained by multiplying together the signs for T and K. The method is quite general, and effects for any 2^k factorial design may be calculated using a table of signs like Table 10.4. We note further that each effect is a contrast, and that they are all mutually orthogonal.

However, although such a table of signs has many uses, the most rapid way of calculating effects is by means of an algorithm due to Yates.

Yates's Algorithm

Yates's algorithm is applied to the observations after they have been rearranged* in what is called *standard order*. A 2^k factorial design is in standard order when, as in Table 10.1, the first column of the design matrix consists of successive minus and plus signs, the second column of successive pairs of minus and plus signs, the third column of four minus signs followed by four plus signs, and so forth. In general, the kth column consists of 2^{k-1} minus signs followed by 2^{k-1} plus signs.

The Yates calculations for the pilot plant data are shown in Table 10.5. In this table the design matrix gives the experimental conditions in standard order. Column **y** contains the corresponding average yields for each run. (Of course, if the runs have not been replicated, each average would be the single observation recorded for that run.) These averages are now considered in successive pairs. The first four entries in column (1) are obtained by adding the pairs together. Thus 60 + 72 = 132, 54 + 68 = 122, and so on. The second four entries in column (1) are obtained by subtracting the *top number from the bottom number* of each pair. Thus 72 − 60 = 12, 68 − 54 = 14, and so on.

In just the same way that column (1) is obtained from column **y**, column (2) is obtained from column (1). Finally, column (3) is obtained from column

* The order of actual running should, of course, be random.

TABLE 10.5. Yates's algorithm, pilot plant example

test condition number	design matrix variables T	C	K	algorithm: run average y	(1)	(2)	(3)	divisor	estimate	identification
1	−	−	−	60	132	254	514	8	64.25	average
2	+	−	−	72	122	260	92	4	23.0	*T*
3	−	+	−	54	135	26	−20	4	−5.0	*C*
4	+	+	−	68	125	66	6	4	1.5	*TC*
5	−	−	+	52	12	−10	6	4	1.5	*K*
6	+	−	+	83	14	−10	40	4	10.0	*TK*
7	−	+	+	45	31	2	0	4	0.0	*CK*
8	+	+	+	80	35	4	2	4	0.5	*TCK*

(2) in the same manner. The entries in column (3) are precisely the values that are obtained by combining the averages with the appropriate column of signs in Table 10.4. To obtain the effects one has only to divide as before by the appropriate divisor, which is 8 for the first entry and 4 for the others. The first estimate is the grand average of all the observations. The remaining effects are identified by locating the plus signs in the design matrix. Thus in the second row a plus sign occurs only in the **T** column, so that the effect in that row is the T effect. In the seventh row plus signs occur in both the **C** and **K** columns, so that the effect in that row is the $C \times K$ interaction.

In general, for a 2^k factorial design, whether working with individual observations or averages of observations, k columns (1), (2), ..., (k) will be generated by adding and subtracting appropriate pairs of numbers. The first divisor will be 2^k, and the remaining divisors will be 2^{k-1}. Yates's algorithm is explained in more detail in Appendix 10A, which includes a description of ways to check the calculations as they are done.

Exercise 10.9. For the data in Exercise 10.3 and 10.4, compute the main effects and interactions, using the table of contrast coefficients and Yates's algorithm.

Answer: See answers for Exercises 10.3 and 10.4.

10.8. A 2^4 FACTORIAL DESIGN: PROCESS DEVELOPMENT STUDY

Results from a 2^4 design employed in a process development study are shown in Table 10.6. A visual display of the data is given in Figure 10.5, which repays careful study, particularly after the analysis has been carried out.

TABLE 10.6. Data obtained in a process development study arranged in standard (Yates) order

observation number	1	2	3	4	conversion (%)	order of runs
1	−	−	−	−	71	(8)
2	+	−	−	−	61	(2)
3	−	+	−	−	90	(10)
4	+	+	−	−	82	(4)
5	−	−	+	−	68	(15)
6	+	−	+	−	61	(9)
7	−	+	+	−	87	(1)
8	+	+	+	−	80	(13)
9	−	−	−	+	61	(16)
10	+	−	−	+	50	(5)
11	−	+	−	+	89	(11)
12	+	+	−	+	83	(14)
13	−	−	+	+	59	(3)
14	+	−	+	+	51	(12)
15	−	+	+	+	85	(6)
16	+	+	+	+	78	(7)

variable	−	+
1 catalyst charge (lb)	10	15
2 temperature (°C)	220	240
3 pressure (psi)	50	80
4 concentration (%)	10	12

Table 10.7 is obtained by multiplying the elements of the columns of the design matrix, first two at a time, next three at a time, then four at a time.

Calculation of Effects

The effects calculated from the table of signs are shown in Table 10.8. Note that, for conciseness, we now denote an interaction as *12* rather than *1* × *2*.

Exercise 10.10. Recalculate the effects using Yates's algorithm. Check your calculations by the methods described in Appendix 10A.

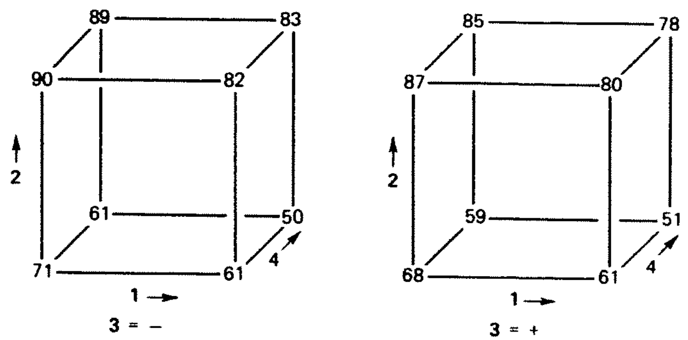


FIGURE 10.5. A 2^4 design with data displayed geometrically, process development example.

TABLE 10.7. **Signs for calculating effects for a 2^4 factorial, process development example**

	I	1	2	3	4	12	13	14	23	24	34	123	124	134	234	1234	conversion (%)
	+	−	−	−	−	+	+	+	+	+	+	−	−	−	−	+	71
	+	+	−	−	−	−	−	−	+	+	+	+	+	+	−	−	61
	+	−	+	−	−	−	+	+	−	−	+	+	+	−	+	−	90
	+	+	+	−	−	+	−	−	−	−	+	−	−	+	+	+	82
	+	−	−	+	−	+	−	+	−	+	−	+	−	+	+	−	68
	+	+	−	+	−	−	+	−	−	+	−	−	+	−	+	+	61
	+	−	+	+	−	−	−	+	+	−	−	−	+	+	−	+	87
	+	+	+	+	−	+	+	−	+	−	−	+	−	−	−	−	80
	+	−	−	−	+	+	+	−	+	−	−	−	+	+	+	−	61
	+	+	−	−	+	−	−	+	+	−	−	+	−	−	+	+	50
	+	−	+	−	+	−	+	−	−	+	−	+	−	+	−	+	89
	+	+	+	−	+	+	−	+	−	+	−	−	+	−	−	−	83
	+	−	−	+	+	+	−	−	−	−	+	+	+	−	−	+	59
	+	+	−	+	+	−	+	+	−	−	+	−	−	+	−	−	51
	+	−	+	+	+	−	−	−	+	+	+	−	−	−	+	−	85
	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	78
divisor	16	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	

TABLE 10.8. Estimated effects from a 2^4 factorial design, process development example

effects	estimated effects ± standard error
average	72.25 ± 0.27
1	(−8.00) ± 0.55
2	(24.00) ± 0.55
3	−2.25 ± 0.55
4	(−5.50) ± 0.55
12	1.00 ± 0.55
13	0.75 ± 0.55
14	0.00 ± 0.55
23	−1.25 ± 0.55
24	(4.50) ± 0.55
34	−0.25 ± 0.55
123	−0.75 ± 0.55
124	0.50 ± 0.55
134	−0.25 ± 0.55
234	−0.75 ± 0.55
1234	−0.25 ± 0.55

Calculation of Standard Errors for Effects Using Higher-Order Interactions

No direct estimate of σ^2 is available from these 16 runs since there were no replicates. However we can obtain such an estimate if certain assumptions are made. In particular, if all three- and four-factor interactions are supposed negligible (an assumption made plausible by the earlier discussion of smoothness and similarity of response functions) these higher-order interactions would measure differences arising principally from experimental error.

They could thus provide an appropriate reference set for the remaining effects. We find:

	effect	effect^2
123	−0.75	0.5625
124	0.50	0.2500
134	−0.25	0.0625
234	−0.75	0.5625
1234	−0.25	0.0625
	sum	1.5000

Accordingly an estimated value for the variance of an effect, having five degrees of freedom, is $1.50/5 = 0.30$. The estimated standard error of an effect is therefore $\sqrt{0.30} = 0.55$, which is the value used in Table 10.8.

Exercise 10.11. The following are data in standard order from a 2^4 factorial design in which the variables were (**1**) brand of tape deck ($- = A$, $+ = B$), (**2**) bass level (low, high), (**3**) treble level (low, high), and (**4**) synthesizer (with, without), and the response was judged quality of sound: 58, 44, 55, 45, 55, 42, 56, 46, 51, 45, 58, 44, 60, 46, 54, 45. Calculate the effects by means of Yates's algorithm, and, using the three and four-factor interactions, calculate the standard error of an effect. *Answer*: 1.64.

A diagram showing the effects, together with a reference t distribution centered at zero and having five degrees of freedom and a scale factor of 0.55, is shown in Figure 10.6. It appears that main effects *1*, *2*, and *4* and interaction *24* (circled in Table 10.8) are distinguishable from the noise. Possibly main effect *3* is also in this category.

Interpretation of the Process Development Data

Proceeding as before, we find:

1. An increase in catalyst charge (variable **1**) from 10 to 15 pounds reduces conversion by about 8%, and the effect is consistent over the levels of the other factors tested.
2. With much less certainty it seems that an increase in pressure (variable **3**) from 50 to 80 psi may possibly reduce conversion by about 2%.
3. Since there is an appreciable interaction between temperature (variable **2**) and concentration (variable **4**), the effects of these variables must be considered jointly. The nature of the interaction is indicated by the two-way table in Figure 10.7, obtained by averaging over levels of the other variables. Evidently high temperature produces high conversion. The interaction occurs because at the low temperature an increase in concentration reduces conversion, whereas an increase in concentration at a higher temperature produces no appreciable effect.

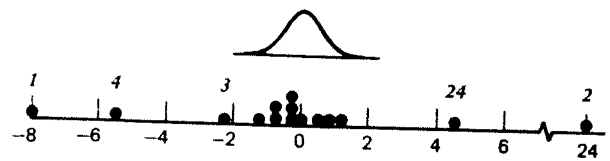


FIGURE 10.6. Main effects and interactions in relation to a reference t distribution with five degrees of freedom and scale factor 0.55, process development example.

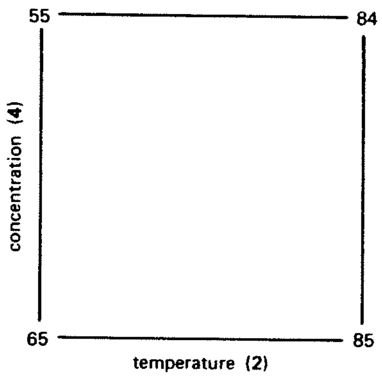


FIGURE 10.7. Two-way table for process development data.

10.9. ANALYSIS OF FACTORIALS USING NORMAL PROBABILITY PAPER

Two problems arise in the assessment of effects from unreplicated factorials: (1) occasionally real and meaningful high-order interactions occur, and (2) it is necessary to allow for selection. The assessment of effects using a reference distribution (Figure 10.6) scaled by an error estimate obtained from higher order interactions does not confront the first of these problems. However, a method (Daniel, 1959) by which *effects* are plotted on normal probability paper often provides an effective way of overcoming both difficulties.

The reader is urged to study Daniel's book (1976) *Applications of Statistics in Industrial Experimentation,* for a penetrating discussion and criticism of factorial designs in general. In particular, this book takes an appropriately skeptical attitude toward mechanical analysis of data and provides many interesting methods for diagnostic checking, including the plotting of *residuals* on normal probability paper.

Normal Probability Paper

A normal distribution is shown in Figure 10.8*a*. The percentage probability of the occurrence of some value less than X is given by the shaded area P. If we plot P against X, we obtain the sigmoid cumulative normal curve shown in Figure 10.8*b*. Normal probability paper is obtained by adjusting the vertical scale in the manner shown in Figure 10.8*c*, so that P versus X plots as a straight line. Suppose that the dots in Figure 10.8*a* represent a random sample of 10 observations from the normal distribution. Since the sample size is $n = 10$, the observation at the extreme left can be taken to represent the first 10% of the cumulative distribution. We plot that first

Accordingly an estimated value for the variance of an effect, having five degrees of freedom, is $1.50/5 = 0.30$. The estimated standard error of an effect is therefore $\sqrt{0.30} = 0.55$, which is the value used in Table 10.8.

Exercise 10.11. The following are data in standard order from a 2^4 factorial design in which the variables were (**1**) brand of tape deck ($- = A$, $+ = B$), (**2**) bass level (low, high), (**3**) treble level (low, high), and (**4**) synthesizer (with, without), and the response was judged quality of sound: 58, 44, 55, 45, 55, 42, 56, 46, 51, 45, 58, 44, 60, 46, 54, 45. Calculate the effects by means of Yates's algorithm, and, using the three and four-factor interactions, calculate the standard error of an effect. *Answer*: 1.64.

A diagram showing the effects, together with a reference t distribution centered at zero and having five degrees of freedom and a scale factor of 0.55, is shown in Figure 10.6. It appears that main effects *1*, *2*, and *4* and interaction *24* (circled in Table 10.8) are distinguishable from the noise. Possibly main effect *3* is also in this category.

Interpretation of the Process Development Data

Proceeding as before, we find:

1. An increase in catalyst charge (variable **1**) from 10 to 15 pounds reduces conversion by about 8%, and the effect is consistent over the levels of the other factors tested.
2. With much less certainty it seems that an increase in pressure (variable **3**) from 50 to 80 psi may possibly reduce conversion by about 2%.
3. Since there is an appreciable interaction between temperature (variable **2**) and concentration (variable **4**), the effects of these variables must be considered jointly. The nature of the interaction is indicated by the two-way table in Figure 10.7, obtained by averaging over levels of the other variables. Evidently high temperature produces high conversion. The interaction occurs because at the low temperature an increase in concentration reduces conversion, whereas an increase in concentration at a higher temperature produces no appreciable effect.

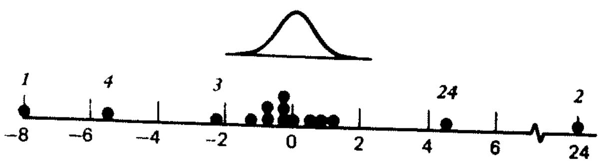


FIGURE 10.6. Main effects and interactions in relation to a reference t distribution with five degrees of freedom and scale factor 0.55, process development example.

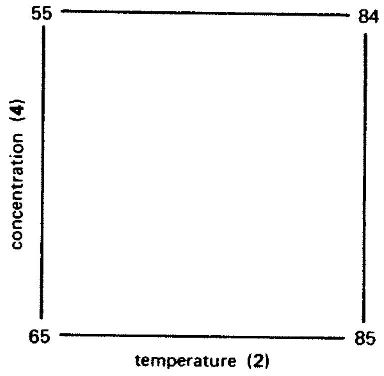


FIGURE 10.7. Two-way table for process development data.

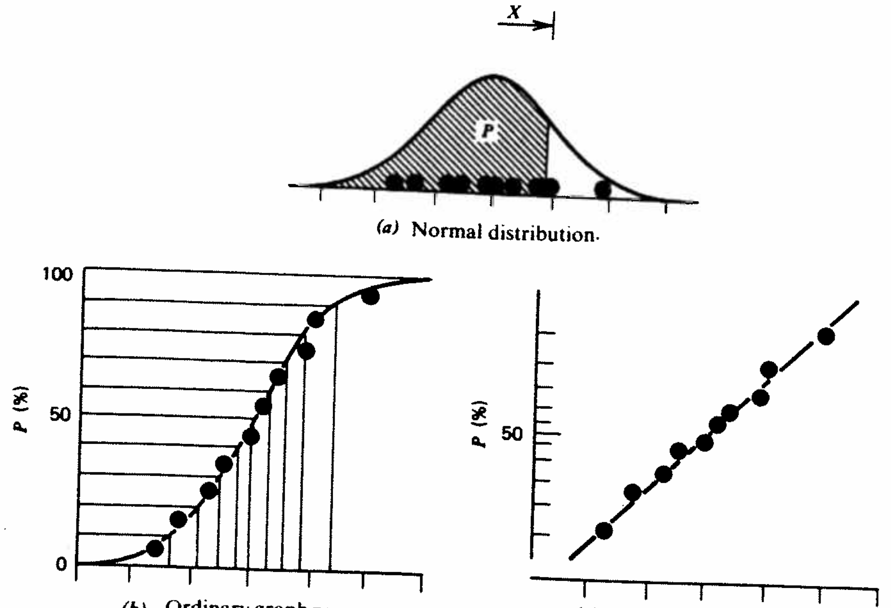
10.9. ANALYSIS OF FACTORIALS USING NORMAL PROBABILITY PAPER

Two problems arise in the assessment of effects from unreplicated factorials: (1) occasionally real and meaningful high-order interactions occur, and (2) it is necessary to allow for selection. The assessment of effects using a reference distribution (Figure 10.6) scaled by an error estimate obtained from higher order interactions does not confront the first of these problems. However, a method (Daniel, 1959) by which *effects* are plotted on normal probability paper often provides an effective way of overcoming both difficulties.

The reader is urged to study Daniel's book (1976) *Applications of Statistics in Industrial Experimentation,* for a penetrating discussion and criticism of factorial designs in general. In particular, this book takes an appropriately skeptical attitude toward mechanical analysis of data and provides many interesting methods for diagnostic checking, including the plotting of *residuals* on normal probability paper.

Normal Probability Paper

A normal distribution is shown in Figure 10.8*a*. The percentage probability of the occurrence of some value less than X is given by the shaded area P. If we plot P against X, we obtain the sigmoid cumulative normal curve shown in Figure 10.8*b*. Normal probability paper is obtained by adjusting the vertical scale in the manner shown in Figure 10.8*c*, so that P versus X plots as a straight line. Suppose that the dots in Figure 10.8*a* represent a random sample of 10 observations from the normal distribution. Since the sample size is $n = 10$, the observation at the extreme left can be taken to represent the first 10% of the cumulative distribution. We plot that first



(a) Normal distribution.

(b) Ordinary graph paper.

(c) Normal probability paper.

FIGURE 10.8. Normal probability plots.

observation, therefore, in Figure 10.8*b* midway between zero and 10%, that is, at the value 5%. Similarly, the second observation can be taken to represent the second 10% of the cumulative distribution, between 10 and 20%, and is plotted at the intermediate value 15%. The sample values approximately trace out the sigmoid curve, as expected. Furthermore, when these same points are plotted on normal probability paper in Figure 10.8*c*, they plot roughly as a straight line. Scales for making your own normal probability plots are given in Table E at the end of the volume. Intercepts such that $P_i = 100(i - \frac{1}{2})/m$ for $i = 1, 2, \ldots, m$ are given for the frequently needed values N = 16, 32, 64, 15, 31, 63.

Process Development Data

For illustration refer to the results from the 2^4 process development experiment considered before. Suppose that these data had occurred simply as the result of random (roughly normal) variation about a fixed mean, and the changes in levels of the variables had had no real effect at all on the percent conversion. Then the $m = 15$ effects (main effects and interactions), rep-

TABLE 10.9. The 15 ordered effects and the probability points *P*, process development example

order number *i*	effect	identity of effects	$P = 100(i - \frac{1}{2})/15$
1	−8.0	*1*	3.3
2	−5.5	*4*	10.0
3	−2.25	*3*	16.7
4	−1.25	*23*	23.3
5	−0.75	*123*	30.0
6	−0.75	*234*	36.7
7	−0.25	*34*	43.3
8	−0.25	*134*	50.0
9	−0.25	*1234*	56.7
10	0.00	*14*	63.3
11	0.50	*124*	70.0
12	0.75	*13*	76.7
13	1.00	*12*	83.3
14	4.50	*24*	90.0
15	24.00	*2*	96.7

resenting 15 contrasts between pairs of averages containing eight observations each, would have been roughly normal and would have been distributed about zero. They would therefore plot on normal probability paper as a straight line. To see whether they do, we order the 15 effects as in Table 10.9 and plot with the appropriate scale from Table E for $m = 15$. It happens that for these data the estimated main effect of factor 1 represents the first $\frac{1}{15} = 6.7\%$ of the cumulative distribution and should therefore be plotted against the value $\frac{1}{2}/15 = 3.3\%$. The next largest estimate is the main effect of factor 4, which is plotted at $1\frac{1}{2}/15 = 10\%$. Beginning with effects with magnitudes close to zero, 11 of the estimates fit reasonably well on a straight line. Those corresponding to *1*, *4*, *24*, and *2* do not. As before, we conclude that these effects are not easily explained as chance occurrences.

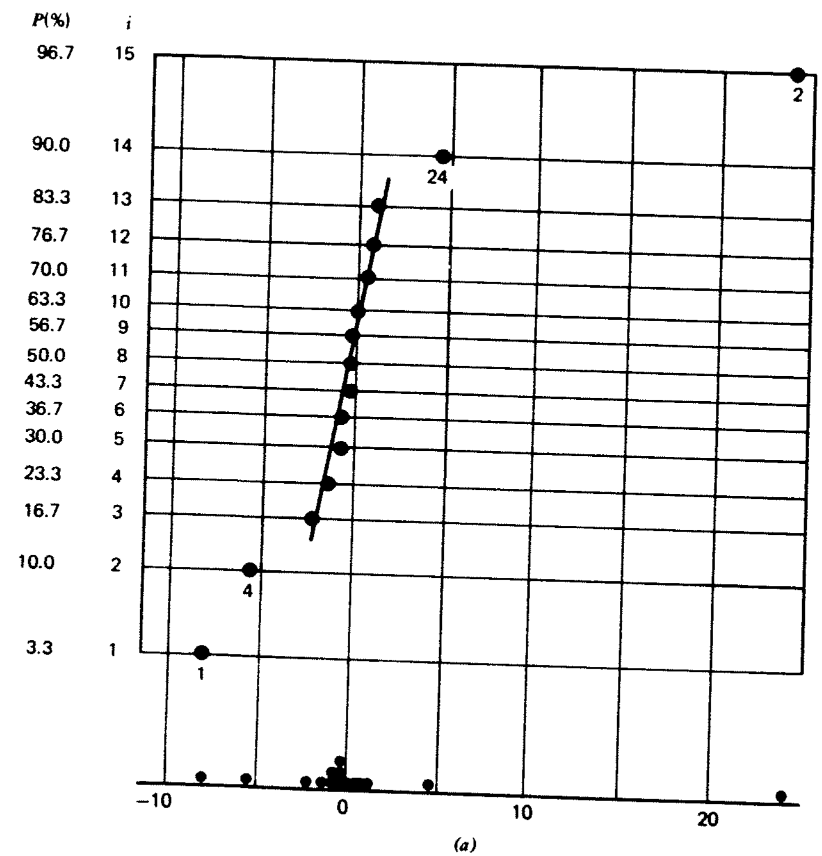


FIGURE 10.9. (*a*) Normal plot of effects, process development example.

Exercise 10.12. Add nine units to observations 2, 3, 6, 7, 9, 12, 13, 16 in Table 10.6. Recalculate the effects, and construct a normal plot.
Answer: Notice that *124* is the only effect whose numerical value is changed. This three-factor interaction is now picked out as deviating from the line.

Diagnostic Checks

Normal plotting of *residuals* provides a diagnostic check for any tentatively entertained model. For example, the plot in Figure 10.9*a* suggests that all effects, with the exception of the *average* 72.25, *1* = −8.0, *2* = 24.0, *4* = −5.5, and *24* = 4.5, can be explained by noise. If this is true, the estimated percent

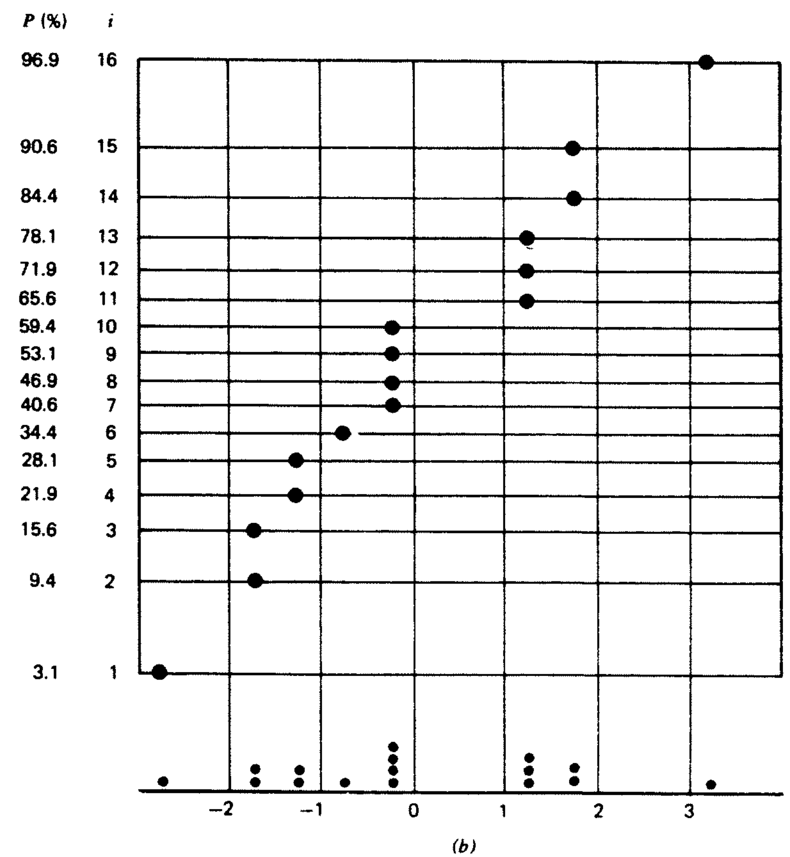


FIGURE 10.9. (*b*) Normal plot of residuals, on normal probability paper, process development example.

conversion for the process development data are given at the vertices of the design by*

$$\hat{y} = 72.25 + \left(\frac{-8.0}{2}\right)x_1 + \left(\frac{24.0}{2}\right)x_2 + \left(\frac{-5.5}{2}\right)x_4 + \left(\frac{4.5}{2}\right)x_2x_4 \quad (10.14)$$

where x_1, x_2, x_4 take the value -1 or $+1$ according to the columns of signs in Tables 10.6 and 10.7. Notice that the coefficients that appear in the equations are *half* the calculated effects. This is so because a change from $x = -1$ to $x = +1$ is a change of *two* units along the x axis.

The values of y, $\hat{y}$, and $y - \hat{y}$ are then as follows:

y	71	61	90	82	68	61	87	80
$\hat{y}$	69.25	61.25	88.75	80.75	69.25	61.25	88.75	80.75
$y - \hat{y}$	1.75	−0.25	1.25	1.25	−1.25	−0.25	−1.75	−0.75
y	61	50	89	83	59	51	85	78
$\hat{y}$	59.25	51.25	87.75	79.75	59.25	51.25	87.75	79.75
$y - \hat{y}$	1.75	−1.25	1.25	3.25	−0.25	−0.25	−2.75	−1.75

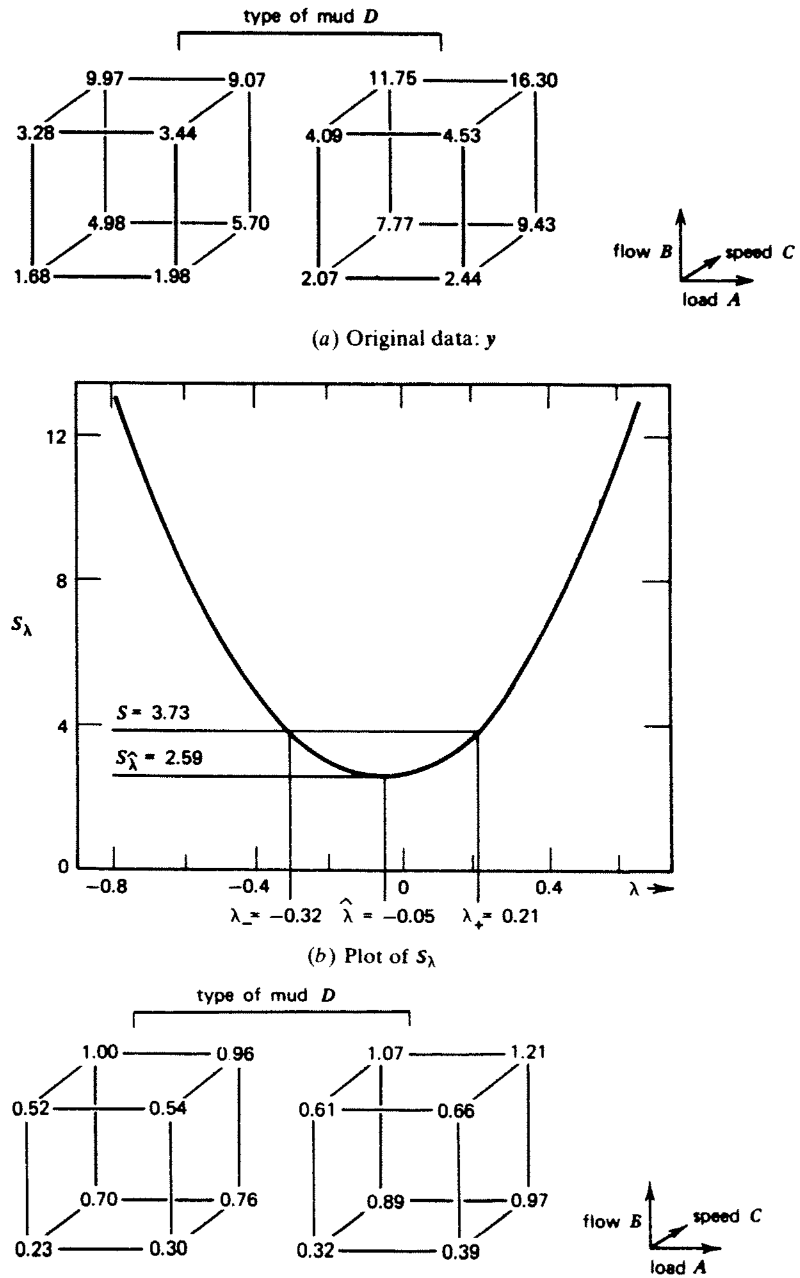
(Alternatively, these residuals may be obtained by using the reverse Yates algorithm described in Appendix 10A.)

The model may now be checked by plotting these residuals on normal probability paper (see Figure 10.9*b*) for $m = 16$. Unlike the original plot of the effects, all the points from this residual plot now lie close to the line, confirming the conjecture that effects other than *1*, *2*, *4*, and *24* are readily explained by random noise. This residual check is valuable provided that the number of effects eliminated (four in this case) is fairly small compared to m.

10.10. TRANSFORMATION OF DATA FROM FACTORIAL DESIGNS

When y_{max}/y_{min} is large, the possibility of simplified and more efficient representation as a result of appropriate transformation (see Section 7.8) should always be kept in mind. A striking example of simplification and

* The occurrence of a cross-product term in this expression suggests that quadratic terms (involving x_2^2 and x_4^2) might be needed in a model which adequately represented the conversion at *interpolated* levels of the variables. Such terms cannot be estimated using two-level designs. The fitting and use of approximating functions called *response surfaces* are outlined in Chapter 15 where methods are discussed for augmenting two-level designs to allow estimation of quadratic effects.



(*a*) Original data: y

(*b*) Plot of s_λ

(*c*) Logged data: $\log y$

FIGURE 10.10. Daniel's 2^4 data with transformation plot.

increased sensitivity achieved by transformation of data from a 3^3 factorial design is given, for example, by Box and Cox (1964).

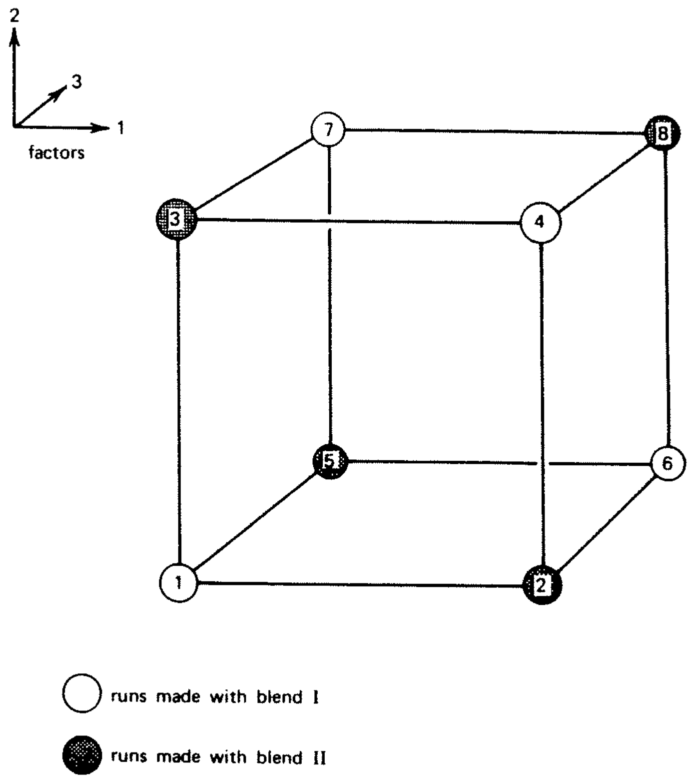
The data plotted in Figure 10.10*a* are from a 2^4 factorial given by Daniel (1976, p. 72). The factors were A, the load on a small stone drill; B, the flow rate through it; C, its rotational speed; and D, the "type of mud used in drilling." The response y was the rate of advance of the drill. The plot of the data y in Figure 10.10*a* shows that the elementary comparisons made along the edges of the cubes are remarkably consistent in sign.

There is, however, a clear tendency for these comparisons to increase in magnitude as the level of the response increases. This will produce interactions (nonadditivity) which might be removed by transformation. In a very interesting and thorough treatment of these data, Daniel shows how such a transformation may be chosen using normal plotting techniques. To illustrate the likelihood approach discussed in Section 7.8 we tentatively assume that, for some suitable power transformation $Y = y^{\lambda}$ of the original data, only main effects are needed to account for the response. As before, supposing that the tentative assumption is true, we carry through an analysis of $y^{(\lambda)} = (y^{\lambda} - 1)/\lambda \dot{y}^{\lambda - 1}$ for various values of λ. The sum of squares of residuals S_{λ} after the four main effects have been eliminated has 11 degrees of freedom. It is plotted against λ in Figure 10.10*b*. We see that, among the class of models considered, $\hat{\lambda}$ is close to zero, suggesting a transformation to logarithms or to some small negative power. As would be expected, this is in agreement with Daniel's analysis, and the log plot in Figure 10.10*c* suggests that a remarkably succinct summary of the major influences of the variables is indeed possible in this metric. Confidence limits (95%) obtained as before are $\lambda_{-} = -0.32$, $\lambda_{+} = 0.21$.

In a more extensive likelihood analysis other models and other transformations could be entertained if it was thought worthwhile to do so. As always, residuals from the various transformations should be studied, and the possibility of bad values kept in mind. We will not take this matter further here, but instead refer the reader to the comprehensive analysis in Daniel's book.

10.11. BLOCKING

Suppose that a trial is to be conducted using a 2^3 factorial design, and, to make the eight runs under conditions as homogeneous as possible, it is desirable that batches of raw material sufficient for the complete experiment be blended together. Suppose, however, that the available blender is only large enough to accommodate material for four runs. This means that two different blends will have to be used. Figure 10.11 shows how the 2^3 factorial



run	1	2	3	12	13	23	123	block		1	2	3	run
1	−	−	−	+	+	+	−	I	block I	−	−	−	1
2	+	−	−	−	−	+	+	II		+	+	−	4
3	−	+	−	−	+	−	+	II		+	−	+	6
4	+	+	−	+	−	−	−	I		−	+	+	7
5	−	−	+	+	−	−	+	II	block II	+	−	−	2
6	+	−	+	−	+	−	−	I		−	+	−	3
7	−	+	+	−	−	+	−	I		−	−	+	5
8	+	+	+	+	+	+	+	II		+	+	+	8

FIGURE 10.11. Arranging a 2^3 factorial design in two blocks of size four.

design can be arranged in two blocks of four runs to neutralize the effect of possible blend differences. Runs indicated by open dots and numbered 1, 4, 6, 7 use blend I, and runs indicated by black dots and numbered 2, 3, 5, 8 use blend II.

The main effects of the factors are contrasts between averages on opposite faces of the cube. But two black dots and two white dots occur on each face, so that any additive effect associated with blends is eliminated from each of the main effects. Now look again at Figure 10.2, and consider the diagonal contrasts, which correspond to the two-factor interactions. Again two black dots and two white dots appear on each side of the contrast. Consequently any additive effect associated with blends is eliminated from each of the two-factor interactions.

The design is blocked in this way by placing all runs in which **123** is minus in one block and all the other runs, in which **123** is plus, in the other block. If all the runs in block II were higher by amount *h* than they would have been if they had been performed in block I, then, whatever the value of *h*, it will sum out in the calculation of effects *1, 2, 3, 12, 13,* and *23.*

Notice that we have had to give up something to get something. We have deliberately *confounded* (i.e., confused) the three-factor interaction and the blend difference. Therefore with this design we cannot estimate the three-factor interaction separately from the blend effect. However, we would usually expect this interaction to be unimportant. In exchange we have been able to run the design in two blocks, which can ensure that main effects and two-factor interactions are much more precisely measured than would otherwise be the case. This important idea of using confounding in the design of experiments is again due to Fisher.

In the 2^3 factorial example, suppose that we give the block variable the identifying number **4**. Then we could think of our experiment as containing four variables, the last of which is assumed to possess the rather special property that it does not interact with the others. If this new variable is introduced by making its levels coincide exactly with the plus and minus signs attributed to the **123** interaction, the blocking may be said to be "generated" by the relationship **4** = **123**. This idea may be used to derive more sophisticated blocking arrangements.

A 2^3 *in Four Blocks of Size Two*

Suppose that in the 2^3 experiment the blends had been large enough for only *two* runs each. How could the design be arranged in four blocks of two runs so as to do as little damage as possible to the estimates of the important effects?

How Not To Do It. Let us introduce two block factors, which we will call **4** and **5**. At first glance it seems reasonable to associate block factor **4** with the **123** interaction

and block factor **5** with some expendable two-factor interaction, say the **23** interaction. The consequences are illustrated in Table 10.10. We assign runs to the different blocks, depending on the signs of the block variables in the columns for **4** and **5**. Thus runs for which these signs are (− −) go in one block, and those that have (− +) signs in another. The (+ −) runs go in a third block, and the (+ +) runs in a fourth.

TABLE 10.10. A 2^3 design in blocks of size two, an undesirable arrangement

run	experimental variable 1	2	3	block variable 4 = 123	5 = 23
1	−	−	−	−	+
2	+	−	−	+	+
3	−	+	−	+	−
4	+	+	−	−	−
5	−	−	+	+	−
6	+	−	+	−	−
7	−	+	+	−	+
8	+	+	+	+	+

experiment arranged in four blocks

block	1	2	3	run
I	+	+	−	4
	+	−	+	6
II	−	−	−	1
	−	+	+	7
III	−	+	−	3
	−	−	+	5
IV	+	−	−	2
	+	+	+	8

variable 5	variable 4: −	+
+	1, 7	2, 8
−	4, 6	3, 5

There is a serious weakness in the design in Table 10.10. We have confounded block variables **4** and **5** with interactions **123** and **23**. But there are three degrees of freedom between four blocks. With what contrast is the third degree of freedom associated? Inspection of Table 10.10 shows that it is associated with the *45* interaction, but this is the "diagonal" contrast between runs 2, 4, 6, 8 and runs 1, 3, 5, 7 measuring the main effect of variable **1** as well as the contrast of blocks (I and IV) with blocks (II and III). Thus the arrangement we have chosen results in confounding of the main effect for factor **1** with block differences!

Clearly caution is necessary. Fortunately a simple calculus is available to show immediately the consequences of any proposed blocking arrangement. If the elements of any column in Table 10.10 are multiplied by themselves, a column of plus signs is obtained. We denote a column of all plus signs by **I**. Thus we write

$$\mathbf{I} = \mathbf{11} = \mathbf{22} = \mathbf{33} = \mathbf{44} = \mathbf{55} \qquad (10.15)$$

where, for example, by **33** we mean the product of the elements in column **3** with themselves. The effect of multiplying the elements in any column by the elements in column **I** is to leave those elements unchanged. Now in the blocking arrangement just considered

$$\mathbf{4} = \mathbf{123}, \qquad \mathbf{5} = \mathbf{23} \qquad (10.16)$$

The **45** column is thus

$$\mathbf{45} = \mathbf{123} \cdot \mathbf{23} = \mathbf{12233} = \mathbf{1II} = \mathbf{1} \qquad (10.17)$$

which shows at once that column **45** is identical to column **1**. That means that interaction *45* and main effect *1* are confounded.

How To Do It. A better arrangement is obtained by confounding the two block variables with any two of the two-factor interactions. The third degree of freedom between blocks is then confounded with the third two-factor interaction. Thus, for

$$\mathbf{4} = \mathbf{12}, \qquad \mathbf{5} = \mathbf{13} \qquad (10.18)$$

interaction *45* is confounded with interaction *23* since

$$\mathbf{45} = \mathbf{1123} = \mathbf{23} \qquad (10.19)$$

The organization of the experiment in four blocks using this arrangement is indicated in Table 10.11, the blocks being typified as before by the pairs of observations for which **4** and **5** take the signatures (− −), (− +), (+ −), and (+ +).

The two runs comprising each block in the above example are complementary in the sense that the plus and minus levels of one run are exactly reversed in the second. Each block is said to consist of a *fold-over* pair. For example, in block I, the plus and minus signs for the pair of runs are (+ − −) and (− + +). Any 2^k factorial may be broken into 2^{k-1} blocks of size two by this method. *Such blocking arrangements leave the main effects of the* k *variables unconfounded with block variables.* All the two-factor interactions, however, are confounded with blocks.

Partial Confounding

When, to achieve sufficient accuracy, replication is necessary, an opportunity is presented to confound different effects in different replicates. Suppose, for example, that four replicates of the 2^3 factorial were to be run in 16 blocks of size two. Then we might run the pattern shown in Table 10.12.

This arrangement would estimate main effects with greatest precision, providing less precision in the estimates of two-factor interactions and still less for three-factor interaction. The reader may find it entertaining to invent other schemes that place different degrees of emphasis on the various effects.

TABLE 10.11. A better arrangement for a 2^3 factorial design in blocks of size two

	experimental variable			block variable	
run	1	2	3	4 = 12	5 = 13
1	−	−	−	+	+
2	+	−	−	−	−
3	−	+	−	−	+
4	+	+	−	+	−
5	−	−	+	+	−
6	+	−	+	−	+
7	−	+	+	−	−
8	+	+	+	+	+

experiment arranged in four blocks

block	1	2	3	run
I	+	−	−	2
	−	+	+	7
II	−	+	−	3
	+	−	+	6
III	+	+	−	4
	−	−	+	5
IV	−	−	−	1
	+	+	+	8

Recovery of Interblock Information

In designs of this kind, estimates confounded with blocks, rather than being regarded as lost, can, following Yates, be thought of merely as having a different (usually larger) variance. Now, from a design such as that in Table 10.12, separate variances for unconfounded and confounded effects can be estimated from the replication of the calculated effects. Suppose that these estimates are s_u^2 and s_c^2 and that for a particular effect we have two estimates: E_u, based on an average of n_u unconfounded estimates, and E_c, based on an average of n_c confounded estimates. Then a combined estimate, using all the information,

TABLE 10.12. Partial confounding: 2^3 design in four replicates

Effects confounded with blocks are indicated by c.
Effects not confounded with blocks are indicated by a u.

replicate	1	2	3	12	13	23	123
first	u	u	u	c	c	c	u
second	c	u	u	u	u	c	c
third	u	c	u	u	c	u	c
fourth	u	u	c	c	u	u	c
number of unconfounded replicates	3	3	3	2	2	2	1
number of confounded replicates	1	1	1	2	2	2	3

is provided by the weighted average of the estimates $E = (w_u E_u + w_c E_c)/(w_u + w_c)$, where the weights are given by $w_u = n_u/s_u^2$, $w_c = n_c/s_c^2$. For a more general discussion of the recovery of interblock information see Cochran and Cox (1957).

10.12. SUMMARY

Two-level factorial designs estimate not only main effects, but also interactions, with maximum precision. The significance of effects may be judged from: an estimate of variance obtained by genuine replication when available; from higher order interactions (assumed due to noise); or by plotting effects on normal probability paper. Rapid calculation of effects is possible using Yates's algorithm. The designs may be run in blocks by associating effects of supposedly lesser importance with block differences.

APPENDIX 10A. YATES'S ALGORITHM

This appendix provides more details on Yates's algorithm, introduced in Section 10.7. Table 10A.1 shows diagramatically the addition and subtraction steps required to generate columns (1), (2), and (3). It also shows how the calculations can be checked after each new column of numbers has been computed. This check is provided by the sum of squares of the entries in each column. The total sum of squares of the eight average results is 34,342. The sum of squares of the entries in column (1) is 2(34,342); in column

TABLE 10A.1. Calculations for carrying out and checking Yates's algorithm: pilot plant example (see Table 10.5)

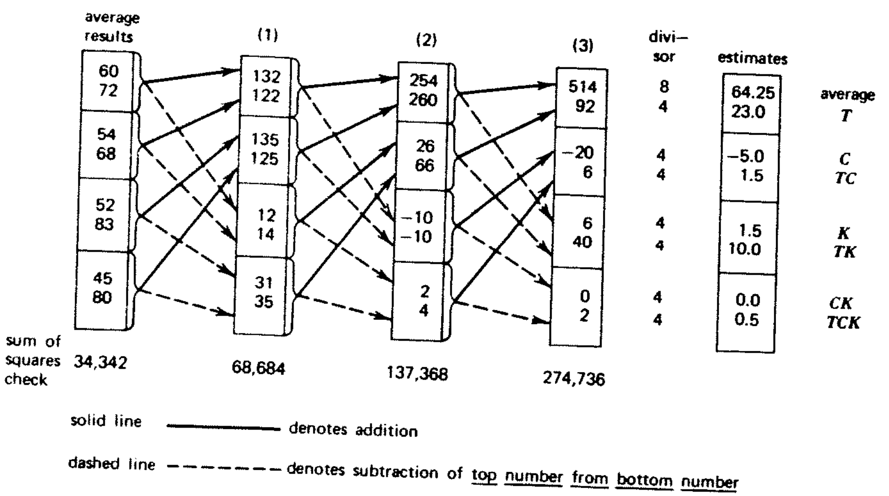


TABLE 10A.2. Reverse Yates's algorithm for obtaining estimated values $\hat{y}$ and residuals $y - \hat{y}$

effect	modified column (3) from forward Yates	(1)	(2)	(3)	divisor	$\hat{y}$	y	$y - \hat{y}$	identification T	C	K
TCK	0	0	40	626	8	78.25	80	1.75	+	+	+
CK	0	40	586	362	8	45.25	45	−0.25	−	+	+
TK	40	−20	−40	666	8	83.25	83	−0.25	+	−	+
K	0	606	402	402	8	50.25	52	1.75	−	−	+
TC	0	0	40	546	8	68.25	68	−0.25	+	+	−
C	−20	−40	626	442	8	55.25	54	−1.25	−	+	−
T	92	−20	−40	586	8	73.25	72	−1.25	+	−	−
average	514	422	442	482	8	60.25	60	−0.25	−	−	−
sum of squares	274,660	549,320	1,098,640	2,197,280		34,332.5	34,342	9.5			

(2), $2^2(34, 342)$; in column (3), $2^3(34,342)$. In general, for a 2^k factorial design, in column (m) the sum of squares should be $2^m \sum y^2$.

Note that the sum of the entries in column (2) is 592. The sum of every second entry in column (1) is $122 + 125 + 14 + 35 = 296$, which, if multiplied by 2, gives 592, that is, $2 \times 296 = 592$. If the calculations are correct, it will be true in general that twice the sum of every second entry in column (i) is equal to the sum of the entries in column ($i + 1$). This provides an additional way to check calculations as they are done.

Yates's algorithm can be used for a replicated 2^k factorial design by starting with *totals* rather than averages, that is, for this example with $59 + 61 = 120$ for the first test condition ($- - -$), $74 + 70 = 144$ for the second test condition ($+ - -$), and so forth. (Recall from Table 10.3 that each test condition was run twice.) When this is done, it will always be true that the main effects and interactions can be obtained by dividing the entries in column (k) by $N/2$, except the first entry, which is divided by N, where N is the total number of runs performed (e.g., in this example $N = 16$ and $k = 3$).

Reverse Yates's Algorithm to Give Estimated Values $\hat{y}$ *and residuals* $y - \hat{y}$

Once again following Cuthbert Daniel, we may obtain estimated values $\hat{y}$ and hence residuals for the 2^k factorials by reversing Yates's algorithm. For example, suppose that in the pilot plant example we wished to obtain estimated values $\hat{y}$ on the supposition that the effects K, TC, CK, and TCK arose only from noise. Then in column (3) of Table 10A.1 these entries would be replaced by zeros. Now, if to this column of entries in *reversed standard order* [see column (1) of Table 10A.2] Yates's algorithm is applied, the $\hat{y}$'s, multiplied by 8, are obtained in reverse standard order at the end of the calculation.

In general, for a 2^k design (a) the kth column of the forward calculation is written in reverse order, appropriately modified by substitution of zeros or other values of interest; (b) the entries are operated on k times with the standard Yates algorithm; and (c) the quantities produced are $2^k \times \hat{y}$ arranged in reverse standard order. Notice once again that the sum of squares checks for the columns. Thus $2^k \sum \hat{y}^2 = 8 \sum \hat{y}^2 = 8 \times 34{,}332.5 = 274{,}660$, which is the sum of squares of the initial entries. Also, since $\sum y^2 = \sum \hat{y}^2 + \sum (y - \hat{y})^2$, $34{,}342 = 34{,}332.5 + 9.5$. An example in which the reverse Yates was employed to uncover a maverick observation in a 2^4 factorial is given in Hunter (1966).

Exercise 10A.1. Recompute the estimated values and residuals for the process development example in Table 10.6, using the reverse Yates algorithm.

APPENDIX 10B. MORE ON BLOCKING FACTORIAL DESIGNS

We have used the 2^3 factorial design to illustrate the principle of blocking designs by means of confounding. These principles, however, are equally applicable to larger examples. Table 10B.1 provides a list of particularly useful arrangements. To understand how this table is used consider a complete 2^6 factorial design containing 64 runs.

Suppose that the experimenter wishes to arrange this program in eight blocks of eight runs each. The eight blocks, for example, might be associated with eight periods of time or eight blends of raw material. We choose to confound three three-factor or higher order interactions among the experimental variables with the block variables $\mathbf{B}_1$, $\mathbf{B}_2$, and $\mathbf{B}_3$. In addition, we require that the interactions between the block variables be themselves confounded with three-factor or higher order interactions among the experimental variables. The table suggests that we use the arrangement

$$\mathbf{B}_1 = \mathbf{135}, \quad \mathbf{B}_2 = \mathbf{1256}, \quad \mathbf{B}_3 = \mathbf{1234}$$

Using the rule $\mathbf{11} = \mathbf{22} = \mathbf{33} = \mathbf{44} = \mathbf{55} = \mathbf{66} = \mathbf{I}$, we obtain the following relationships after multiplication of the block variables by one another:

$$\begin{aligned} \mathbf{B}_1 &= \mathbf{135} \\ \mathbf{B}_2 &= \mathbf{1256} \\ \mathbf{B}_3 &= \mathbf{1234} \\ \mathbf{B}_1\mathbf{B}_2 &= \mathbf{236} \\ \mathbf{B}_1\mathbf{B}_3 &= \mathbf{245} \\ \mathbf{B}_2\mathbf{B}_3 &= \mathbf{3456} \\ \mathbf{B}_1\mathbf{B}_2\mathbf{B}_3 &= \mathbf{146} \end{aligned}$$

which show that only high-order (three-factor and higher order) interactions are confounded with the seven degrees of freedom associated with the eight blocks.

To allocate the experiments to the eight blocks we follow the same procedure as before: we write down the plus and minus levels corresponding to $\mathbf{B}_1$, $\mathbf{B}_2$, and $\mathbf{B}_3$ and then combine the individual runs into blocks that have the same signs for the block variables. As illustrated in Table 10B.2, the signs of ($\mathbf{B}_1$, $\mathbf{B}_2$, $\mathbf{B}_3$) are associated with the eight blocks according to the following scheme:

block	1	2	3	4	5	6	7	8
($\mathbf{B}_1$, $\mathbf{B}_2$, $\mathbf{B}_3$)	($- - -$)	($+ - -$)	($- + -$)	($+ + -$)	($- - +$)	($+ - +$)	($- + +$)	($+ + +$)

REFERENCES AND FURTHER READING

Yates, F. (1937). *The Design and Analysis of Factorial Experiments*, Bulletin 35, Imperial Bureau of Soil Science, Harpenden, Herts, England, Hafner (Macmillan).

Yates, F. (1970). *Selected Papers*, Hafner (Macmillan).

Daniel, C. (1976). *Applications of Statistics to Industrial Experimentation*, Wiley. (This excellent text carries considerably further than is done here the analysis of 2^k factorial designs and, in particular, the detection and interpretation of discrepant results.)

Davies, O. L. (Ed.) (1971). *The Design and Analysis of Industrial Experiments*, Hafner (Macmillan).

Cochran, W. G. and G. M. Cox (1957). *Experimental Designs*, 2nd ed., Wiley.

Youden, W. J. (1972). Enduring values, *Technometrics*, **14**, 1.

Box, G. E. P. and D. R. Cox (1964). An analysis of transformations, *J. Roy. Stat. Soc.* Series B, **26**, 211.

Daniel, C. (1959). Use of half-normal plot in interpreting factorial two-level experiments, *Technometrics*, **1**, 149.

Hunter, J. S. (1966). Inverse Yates algorithm, *Technometrics*, **8**, 177.