Design of Experiments

1 Factorial Experiments

Many experiments involve the study of the effects of two or more factors. In general, **factorial designs** are most effcient in this type of experiment. By a factorial design, we mean that in each complete trial or replication of the experiment all possible combinations of the levels of the factors are investigated. For example, if there are a levels of factor A and b levels of factor B, each replicate contains all ab treatment combinations.
The effect of factor is defined to be the change in response produced by a change in the level of the factor. This is called *main effect* because it refers to the primary factors of the experiment. In follwing example, main effect of factor A is

$$A = \frac{40+52}{2} - \frac{20+30}{2} = 21$$

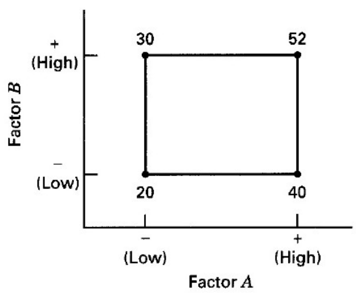

Figure 5-1 A two-factor factorial experiment, with the response (y) shown at the corners.

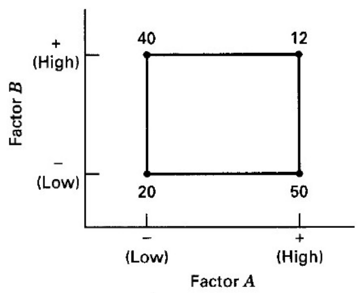

Figure 5-2 A two-factor factorial experiment with interaction.

and main effect of factor B is

$$B = \frac{30+52}{2} - \frac{20+40}{2} = 11$$

In many experiments, we see that the difference of response between the levels of one factor is not the same at all levels of other factor. When this occurs, there is a *interaction* between factors. In the above two factor experiment, the effect of A at low levels of B is

$$50 - 20 = 30$$

and effect of A at high level of B is

$$12 - 40 = -28.$$

The magnitude of the intercation is defined by the average difference between these two effects

$$AB = \frac{-28-30}{2} = -29$$

The Two-Factor Factorial Design

Suppose there are two treatments A and B. There are a levels of factor A and b levels of factor B and these are arranged in a factorial design and each of ab treatment combinations is replicated r times. We can use CRD or RBD for factorial experiment.

If CRD is used then the observations y_{ijk} from k^{th} replicate of i^{th} level of factor A and j^{th} level of factor B can be modeled as

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijk}$$

for $i = 1, \ldots, a$, $j = 1, 2, \ldots, b$ and $k = 1, 2, \ldots, r$. Here μ is the general effect, τ_i is the effect of i^{th} level of factor A ,β_j is the effect of j^{th} level of factor B and $(\tau\beta)_{ij}$ is the interaction effect between τ_i and β_j and $\epsilon_{ijk} \sim N(0, \sigma^2)$. We assume that

$$\sum_i \tau_i = \sum_j \beta_j = \sum_i \sum_j (\tau\beta)_{ij} = 0$$

The null hypothesis to be tested are

(a) equality of different levels of factor A

$$H_0 : \tau_1 = \tau_2 = \ldots = \tau_a = 0$$

against

$$H_1 : \tau_i \neq 0 \qquad \text{for at least one } i$$

(b) equality of different levels of factor B

$$H_0 : \beta_1 = \beta_2 = \ldots = \beta_b = 0$$

against

$$H_1 : \beta_i \neq 0 \qquad \text{for at least one } i$$

and

(c)

$$H_0 : (\tau\beta)_{ij} = 0 \quad \text{for all} i, j$$

against

$$H_1 : (\tau\beta)_{ij} \neq 0 \qquad \text{for at least one pair } i, j$$

Analysis

Define

$$\begin{aligned}
\bar{y}_{i\cdot\cdot} &= \frac{1}{br}\sum_{j=1}^{b}\sum_{k=1}^{r} y_{ijk} \\
\bar{y}_{\cdot j\cdot} &= \frac{1}{ar}\sum_{i=1}^{a}\sum_{k=1}^{r} y_{ijk} \\
\bar{y}_{ij\cdot} &= \frac{1}{r}\sum_{k=1}^{r} y_{ijk} \\
\bar{y}_{\cdots} &= \frac{1}{abr}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r} y_{ijk}
\end{aligned}$$

The total sum of squares can be written as

$$\begin{aligned}
\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r}(y_{ijk} - \bar{y}_{\cdots})^2 &= br\sum_{i=1}^{a}(\bar{y}_{i\cdot\cdot} - \bar{y}_{\cdots})^2 + ar\sum_{j=1}^{b}(\bar{y}_{\cdot j\cdot} - \bar{y}_{\cdots})^2 \\
&\quad + r\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{y}_{ij\cdot} - \bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot j\cdot} + \bar{y}_{\cdots})^2 \\
&\quad + \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r}(y_{ijk} - \bar{y}_{\cdots})^2
\end{aligned}$$

that is

$$TSS = SSA + SSB + SS(AB) + SSE$$

Also. it can be shown that

$$\begin{aligned}
E(MSA) &= E\left(\frac{SSA}{a-1}\right) = \sigma^2 + \frac{br\sum_{i=1}^{a}\tau_i^2}{a-1} \\
E(MSB) &= E\left(\frac{SSA}{b-1}\right) = \sigma^2 + \frac{ar\sum_{j=1}^{b}\beta_j^2}{b-1} \\
E(MS(AB)) &= E\left(\frac{SS(AB)}{(a-1)(b-1)}\right) = \sigma^2 + \frac{r\sum_{i=1}^{a}\sum_{j=1}^{b}(\tau\beta)_{ij}^2}{(a-1)(b-1)} \\
E(MSE) &= E\left(\frac{SSE}{ab(r-1)}\right) = \sigma^2
\end{aligned}$$

The ANOVA for two factor factorial experiment is

Source of variation	Sum of squares	Degrees of freedom	Mean square	F_0
A Treatments	$SSA = \frac{1}{rb}\sum_{i=1}^{a} y_{i..}^2 - \frac{y_{...}^2}{abr}$	$a-1$	$MSA = \frac{SSA}{a-1}$	$\frac{MSA}{MSE}$
B Treatments	$SSB = \frac{1}{ra}\sum_{j=1}^{b} y_{.j.}^2 - \frac{y_{...}^2}{abr}$	$b-1$	$MSB = \frac{SSB}{b-1}$	$\frac{MSB}{MSE}$
Interactions	$SS(AB) = \frac{1}{r}\sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij.}^2 - \frac{y_{...}^2}{abr} - SSA - SSB$	$(a-1)(b-1)$	$MS(AB) = \frac{SS(AB)}{(a-1)(b-1)}$	$\frac{MS(AB)}{MSE}$
Error	$SSE = TSS - SSA - SSB - SS(AB)$	$ab(r-1)$	$MSE = \frac{SSE}{ab(r-1))}$	
Total	$TSS = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r} y_{ijk}^2 - \frac{y_{...}^2}{abr}$	$abr-1$		

If RBD is used with r blocks then model will be

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \delta_k + \epsilon_{ijk}$$

for $i = 1, \ldots, a$, $j = 1, 2, \ldots, b$ and $k = 1, 2, \ldots, r$. Here δ_k is the effect of k^{th} block and $\sum_k \delta_k = 0$. Other parameters are same as CRD model.

The ANOVA table will be

Source of variation	Sum of squares	Degrees of freedom	Mean square	F_0
Blocks	$\text{SS(Blocks)} = \frac{1}{ab}\sum_{k=1}^{r} y_{\cdot\cdot k}^2 - \frac{y_{\cdots}^2}{abr}$	$r-1$	$MS(\text{Blocks}) = \frac{SS(\text{Blocks})}{r-1}$	
A	$SSA = \frac{1}{rb}\sum_{i=1}^{a} y_{i\cdot\cdot}^2 - \frac{y_{\cdots}^2}{abr}$	$a-1$	$MSA = \frac{SSA}{a-1}$	$\frac{MSA}{MSE}$
B	$SSB = \frac{1}{ra}\sum_{j=1}^{b} y_{\cdot j\cdot}^2 - \frac{y_{\cdots}^2}{abr}$	$b-1$	$MSB = \frac{SSB}{b-1}$	$\frac{MSB}{MSE}$
AB	$SS(AB) = \frac{1}{r}\sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij\cdot}^2 - \frac{y_{\cdots}^2}{abr} - SSA - SSB$	$(a-1)(b-1)$	$MS(AB) = \frac{SS(AB)}{(a-1)(b-1)}$	$\frac{MS(AB)}{MSE}$
Error	$SSE = TSS - SSA - SSB - SS(AB)$	$(ab-1)(r-1)$	$MSE = \frac{SSE}{(ab-1)(r-1)}$	
Total	$TSS = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r} y_{ijk}^2 - \frac{y_{\cdots}^2}{abr}$	$abr-1$		

2 2^k Factorial design

Suppose we want to compare k factors each at two levels. The levels may be quantitative such as two values of temparature, preasure or time. A complete replicate of such design requires $2 \times 2 \times \cdots \times 2 = 2^k$ observations and is called 2^k factorial design. We will assume that RBD is used for the design and normality assumption of errors is satisfied.

2.1 The 2^2 design

Suppose there are two factors each run at two levels. This design is called 2^2 design. The levels of the factors may be called "low" and "high". Consider the following example

Factors		Treatment	Replicates			Total
A	B	Combinations	I	II	III	
−	−	A low, B low	28	25	27	80
+	−	A high, B low	36	32	32	100
−	+	A low, B high	18	19	23	60
+	+	A high, B high	31	30	29	90

Now, let us denote the high level of factor in treatment combination by the corresponding lowecase of the letter and low level of factor by the absence of the corresponding letter. Thus, a represents high level of A and low level of B, b represents A at the low level and B at the high level and ab represents both A and B at high level and (1) represents both factors at low level.

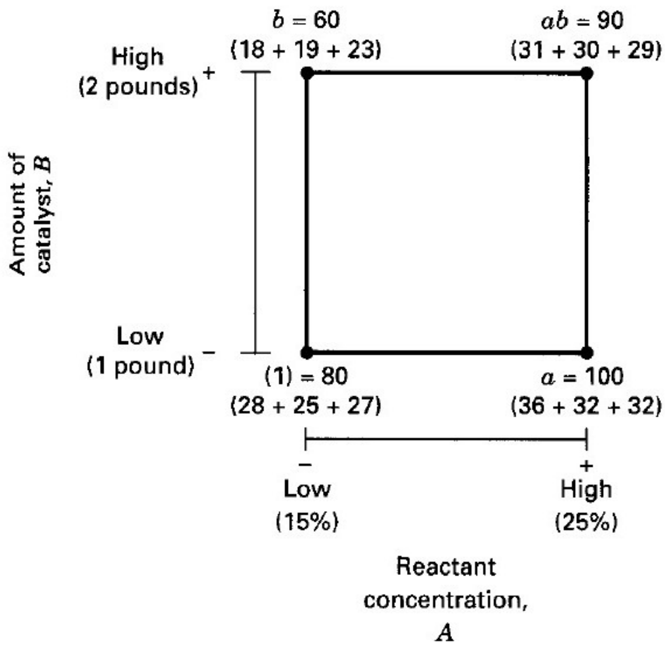

Figure 6-1 Treatment combinations in the 2^2 design.

Let the symbols $[1], [a], [b], [ab]$ represent the total of all r replicates taken at the treatment combinations.

The effect of A at low level of B is $([a]-[1])/r$ and effect of A at high level of B is $([ab]-[b])/r$. The **main effect of** A is obtained by averaging of these two quantities that is the main effect of A is

$$A = \frac{1}{2r}\left([ab]-[b]+[a]-[1]\right) = \frac{1}{2r}\left([a]-[1]\right)\left([b]+[1]\right)$$

Similarly, the effect of B at low level of A is $([b]-[1])/r$ and effect of B at high level of A is $([ab]-[a])/r$. Hence **main effect of** B is

$$B = \frac{1}{2r}\left([ab]-[a]+[b]-[1]\right) = \frac{1}{2r}\left([a]+[1]\right)\left([b]-[1]\right)$$

The **interaction effect** is the average difference between effect of A at high level of B and effect of A at low level of B that is

$$AB = \frac{1}{2r}\left([ab]-[b]-[a]+[1]\right) = \frac{1}{2r}\left([a]-[1]\right)\left([b]-[1]\right)$$

Note that AB and BA are same.

To calculate the sum of squares we note that contrasts are used in estimating A namely

$$Contrast_A = (ab)-(b)+(a)-(1).$$

We can also see that

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

and

$$Contrast_{AB} = (ab)-(a)-(b)+(1)$$

Furthurmore , these three contrasts are orthogonal.

In general, a contrast is a linear combination of the parameters of the form,

$$\Gamma = \sum_{i=1}^{v} c_i \mu_i$$

where $\sum_{i=1}^{v} c_i = 0$. *To test*

$$H_0 : \Gamma = 0$$

we use the statistic

$$t_0 = \frac{\sum_{i=1}^{v} c_i y_i}{\sqrt{nMSE \sum_{i=1}^{v} c_i^2}}$$

We can also use

$$F_0 = \frac{\left(\sum_{i=1}^{v} c_i y_i\right)^2}{nMSE \sum_{i=1}^{v} c_i^2} = \frac{SS_C/1}{MSE}.$$

We reject H_0 *is* $F_0 > F_{\alpha,1,n-v}$. *The sum of squares is*

$$SS_C = \frac{\left(\sum_{i=1}^{v} c_i y_i\right)^2}{n \sum_{i=1}^{v} c_i^2}$$

Using this formula we get the sum of squares of treatment combinations as

$$\begin{aligned} SSA &= \frac{([ab] - [b] + [a] - [1])^2}{4r} \\ SSB &= \frac{([ab] + [b] - [a] - [1])^2}{4r} \\ \text{and} & \\ SS(AB) &= \frac{([ab] - [b] + [a] - [1])^2}{4r} \end{aligned}$$

The total sum of squares is obtained as

$$TSS = \sum_{i=1}^{2} \sum_{j=1}^{2} \sum_{k=1}^{r} y_{ijk}^2 - \frac{y_{...}^2}{4r}$$

where $y_{...} = \sum_{i=1}^{2} \sum_{j=1}^{2} \sum_{k=1}^{r} y_{ijk}$. The total sum of squares has $4r - 1$ degrees of freedom. The error sum of squares is computed as

$$SSE = TSS - SSA - SSB - SS(AB)$$

As we have assumed RBD will be used for the experiment, the ANOVA table is given as

Source of variation	Sum of squares	Degrees of freedom	Mean square	F_0
Blocks	SS(Blocks) $= \frac{1}{4}\sum_{k=1}^{r} y_{..k}^2 - \frac{y_{...}^2}{4r}$	$r-1$	$MS(\text{Blocks}) = \frac{SS(\text{Blocks})}{r-1}$	
A	SSA	1	$MSA = \frac{SSA}{1}$	$\frac{MSA}{MSE}$
B	SSB	1	$MSB = \frac{SSB}{1}$	$\frac{MSB}{MSE}$
AB	$SS(AB)$	1	$MS(AB) = \frac{SS(AB)}{1}$	$\frac{MS(AB)}{MSE}$
Error	$SSE = TSS - SSA - SSB - SS(AB)$	$3(r-1)$	$MSE = \frac{SSE}{3(r-1)}$	
Total	TSS	$4r-1$		

Let M is the mean yield defined by

$$M = \frac{1}{4}\left([ab] + [a] + [b] + [1]\right)$$

The sign table is

Effects	(1)	a	b	ab
M	$+$	$+$	$+$	$+$
A	$-$	$+$	$-$	$+$
B	$-$	$-$	$+$	$+$
AB	$+$	$-$	$-$	$+$

2.2 The 2^3 design

Suppose that three factors , A, B and C each at 2 levels are of interest. The $2^3 = 8$ treatment combinations are (1), a, b, ab, c, ac, bc, abc. These treatment combinations are represented in a cube.

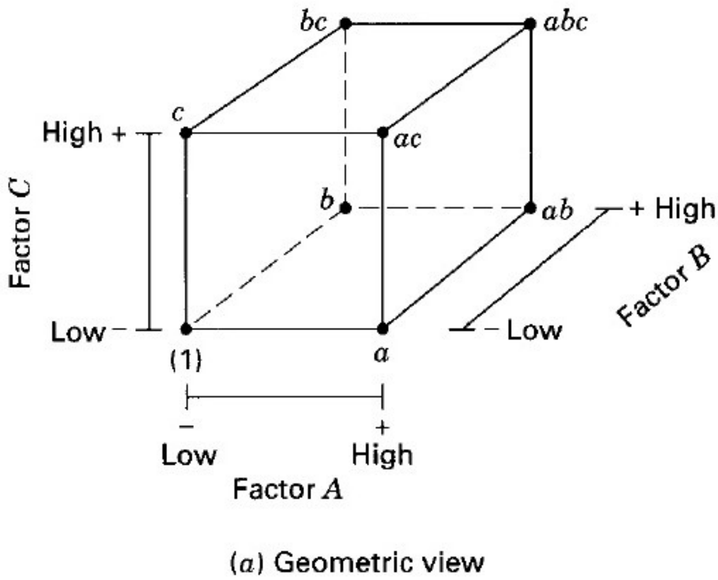

Figure 6-4 The 2^3 factorial design.

There are seven degrees of freedom between eight treatment combinations in 2^3 design. Three degrees of freedom are associated with main effects A, B and C. Four degrees of freedom is associated with interactions, one each with AB, AC, BC and one with ABC.

Let the symbols $[1], [a], [b], [ab], [c], [ac], [bc], [abc]$ represent the total of all r replicates taken at the treatment combinations.

Let us consider estimating main effects. First consider estimation of main effect A. The effect of A at low level of B and C is $([a] - [1])/r$, effect of A at high level of B and low level of C is $([ab] - [b])/r$, effect of A at low level of B and high level of C is $([ac] - [c])/r$ and effect of A at high level of B and high level of C is $([abc] - [ab])/r$. So, main effect of A is

$$A = \frac{1}{4r}\left([abc] - [ab] + [ac] - [c] + [ab] - [b] + [a] - [1]\right)$$

Similarly main effect of B is

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

and main effect of C is

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

The interaction effect AB is estimated by average difference between average of A effects at the two levels of B. At high level of B, the average A effect is $\left([abc]-[bc]+[ab]-[b]\right)/2r$ and at low level of B the average A effect is $\left([ac]-[c]+[a]-[1]\right)/2r$. So, the interaction AB is estimated as

$$AB = \frac{1}{4r}\left([abc]-[bc]+[ab]-[b]-[ac]+[c]-[a]+[1]\right).$$

Using same logic the effects of BC and AC are estimated as

$$BC = \frac{1}{4r}\left([abc]-[ac]+[bc]-[c]-[ab]+[a]-[b]+[1]\right)$$

and

$$AC = \frac{1}{4r}\left([abc]-[bc]+[ac]-[c]-[ab]+[b]-[a]+[1]\right)$$

The ABC interaction is defined as the average difference between AB interaction at two different levels of C.

$$\begin{aligned} ABC &= \frac{1}{4r}\left(([abc]-[bc])-([ac]-[c])-([ab]-[b])+([a]-[1])\right) \\ &= \frac{1}{4r}\left([abc]-[bc]-[ac]+[c]-[ab]+[b]+[a]-[1]\right) \end{aligned}$$

The sign table

Effects	(1)	a	b	ab	c	ac	bc	abc
M	+	+	+	+	+	+	+	+
A	−	+	−	+	−	+	−	+
B	−	−	+	+	−	−	+	+
C	−	−	−	−	+	+	+	+
AB	+	−	−	+	+	−	−	+
BC	+	+	−	−	−	−	+	+
AC	+	−	+	−	−	+	−	+
ABC	−	+	+	−	+	−	−	+

The three -factor analysis of variance model is

$$y_{ijkl} = \mu + \tau_i + \beta_j + \gamma_k + (\tau\beta)_{ij} + (\tau\gamma)_{ik} + (\beta\gamma)_{jk} + (\tau\beta\gamma)_{ijk} + \epsilon_{ijkl}$$

for $i, j, k = 1, 2$ and $l = 1, 2, \ldots, r$. Here μ is the general effect, τ_i is the effect of i^{th} level of factor A ,β_j is the effect of j^{th} level of factor B, γ_k is the effect of k^{th} level of factor C, $(\tau\beta)_{ij}$ is the interaction effect between τ_i and β_j, $(\tau\gamma)_{ik}$ is the interaction effect between τ_i and γ_k, $(\beta\gamma)_{jk}$ is the interaction effect between β_j and γ_k, $(\tau\beta\gamma)_{ijk}$ is the interaction effect between τ_i, β_j, γ_k and $\epsilon_{ijkl} \sim N(0, \sigma^2)$.
We assume that

$$\sum_i \tau_i = \sum_j \beta_j = \sum_k \gamma_k = 0$$

$$\sum_i \sum_j (\tau\beta)_{ij} = \sum_i \sum_k (\tau\gamma)_{ik} = \sum_j \sum_k (\beta\gamma)_{jk} = 0$$

$$\sum_i \sum_j \sum_k (\tau\beta\gamma)_{ijk} = 0$$

As contrasts are used estimate the effects A, B, C, AB, AC, BC, ABC, the sum of squares for the effects corresponding to single degree of freedom contrast is given by

$$SS = \frac{(\text{Contrast})^2}{8r}$$

$$Contrast_A = (abc) - (ab) + (ac) - (c) + (ab) - (b) + (a) - (1).$$

$$Contrast_B = (abc) - (ac) + (ab) - (a) + (bc) - (c) + (b) - (1).$$

$$Contrast_{AB} = (abc) - (bc) + (ac) - (c) - (ab) + (b) - (a) + (1).$$

$$Contrast_C = (abc) - (ab) + (ac) - (a) + (bc) - (b) + (c) - (1).$$

$$Contrast_{BC} = (abc) - (ac) + (bc) - (c) - (ab) + (a) - (b) + (1).$$

$$Contrast_{AC} = (abc) - (bc) + (ac) - (c) - (ab) + (b) - (a) + (1).$$

$$Contrast_{ABC} = (abc) - (ab) - (ac) + (c) - (ab) + (b) + (a) - (1).$$

As we have assumed RBD will be used for the experiment, the ANOVA table is given as

Source of variation	Sum of squares	Degrees of freedom	Mean square	F_0
Blocks	SS(Blocks) $= \frac{1}{8}\sum_{l=1}^{r} B_l^2 - \frac{y_{....}^2}{8r}$	$r-1$	$MS(\text{Blocks}) = \frac{SS(\text{Blocks})}{r-1}$	
A	SSA	1	$MSA = \frac{SSA}{1}$	$\frac{MSA}{MSE}$
B	SSB	1	$MSB = \frac{SSB}{1}$	$\frac{MSB}{MSE}$
AB	$SS(AB)$	1	$MS(AB) = \frac{SS(AB)}{1}$	$\frac{MS(AB)}{MSE}$
C	SSC	1	$MSC = \frac{SSC}{1}$	$\frac{MSC}{MSE}$
BC	$SS(BC)$	1	$MS(BC) = \frac{SS(BC)}{1}$	$\frac{MS(BC)}{MSE}$
AC	$SS(AC)$	1	$MS(AC) = \frac{SS(AC)}{1}$	$\frac{MS(AC)}{MSE}$
ABC	$SS(ABC)$	1	$MS(ABC) = \frac{SS(ABC)}{1}$	$\frac{MS(ABC)}{MSE}$
Error	$SSE = TSS - SSA - SSB - SS(AB)$	$7(r-1)$	$MSE = \frac{SSE}{7(r-1)}$	
Total	TSS	$8r-1$		

where B_l is the sum of lth block.

3 Confounding in the 2^k factorial design

There are many problems for which it is impossible to perform a complete replicate of factorial design in a single block. As the number treatment combinations get larger the blocks may not remain homogeneous. **Confounding** is a design technique for arranging a complete factorial experiment in blocks, where block size is smaller the number of treatment combinations in one replicate. The technique causes information about certain treatment effects (usually higher order interactions) to be *indistinguishable from* or *confounded with*, blocks.

Here we consider construction and analysis 2^k design in 2^p incomplete blocks, where $p < k$.

3.1 Confounding 2^k factorial design in two blocks

Suppose in 2^2 design we wish to confound the effect of the interaction AB. In a 2^2 design there are 4 treatment combinations (1), a, b, ab. Let a single replicate is divided into blocks. Block 1 contains the treatment combinations (1) and ab and block 2 contains a and b. The *order* in which the treatment combinations are run within a block is randomly determined.
Suppose we estimate the main effects of A and B just as no blocking had occurred. So, esimates of A and B are

$$A = \frac{1}{2}\left([ab] - [b] + [a] - [1]\right)$$

and

$$B = \frac{1}{2}\left([ab] - [a] + [b] - [1]\right)$$

Note that estimates of A and B are unaffected by block affects as in each estimate there is one plus and one minus treatment combination from each block. Now consider the AB interaction

$$AB = \frac{1}{2}\left([ab] - [b] - [a] + [1]\right)$$

Because the two treatment combination with plus sign are in block 1 and two treatment combinations with minus sign are in block 2, the block effect = block 1 - block 2 and AB interactions are identical. That is , AB is **confounded** with blocks.

We see that all treatment combinations with plus sign is assigned to block 1 and treatment combinations with minus sign is assigned to block 2 to confound the effect AB with blocks. This approach can be used to confound any effects (A, B, AB) with blocks. If treatment combinations ab, a is assigned to block 1 and b and (1) is assigned to block 2, then effect A will be confounded. The usual practice is to confound the highest order interaction with blocks.

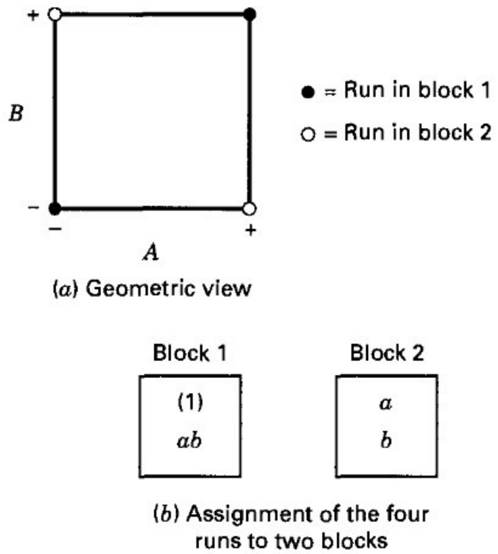

Figure 7-1 A 2^2 design in two blocks.

Suppose in a 2^3 design we wish to confound the interaction effect ABC with blocks. Then treatment combinations with plus sign on ABC in sign table a, b, c, abc are assigned to block 1and treatment combinations with minus sign on ABC, $(1), ab, bc, ac$ are assigned to block 2. The treatment combinations within the block are run in random order.

Another method of constructing blocks

Let in general the treatments in 2^k are denoted as $A^{\alpha_1}B^{\alpha_2}C^{\alpha_3}\cdots Z^{\alpha_k}$ for $\alpha_i = 1$ or 0.

This method uses the linear combination

$$L = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_k x_k$$

where x_i is the level of i^{th} factor appearing in a treatment combination and α_i is the exponent appearing on the i^{th} factor in the effect to be confounded. We have $x_i = 0$(low level) or $x_i = 1$(high level). The linear combination L is called a **defining contrast.** Treatment combinations that produces same value of L (mod 2) will be placed in same block. Because the only possible values of $L(\text{mod } 2)$ are 0 and 1, this will assign 2^k treatment combinations into two blocks.

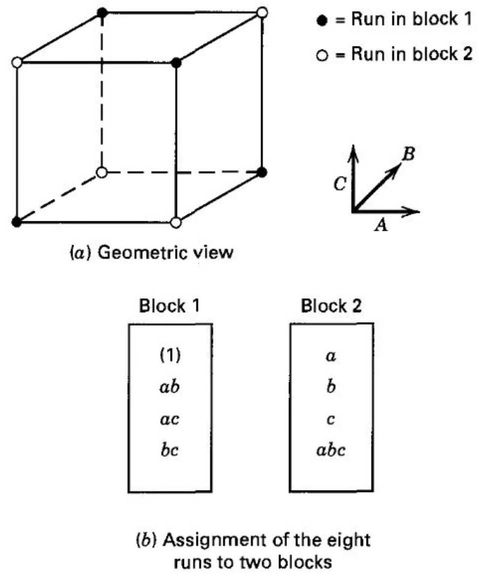

Figure 7-2 The 2^3 design in two blocks with ABC

Let us consider a 2^3 experiment where the effect ABC to be confounded. Here $\alpha_1 = \alpha_2 = \alpha_3 = 1$. So the defining contrast corresponds to ABC is

$$L = x_1 + x_2 + x_3$$

The treatment combination (1) can be written as 000 in $(0, 1)$ notation, therefore

$$L = 1.0 + 1.0 + 1.0 = 0 = 0(\text{mod}2)$$

Similarly treatment combination a is 100 and henc

$$L = 1.1 + 1.0 + 1.0 = 1 = 1(\text{mod}2)$$

Hence the treatment combintions (1) and a go to different blocks.

The remaining treatment combinations are

$$\begin{aligned}
b:L &= 1.0+1.1+1.0=1=1(\text{mod}2)\\
c:L &= 1.0+1.0+1.1=1=1(\text{mod}2)\\
ab:L &= 1.1+1.1+1.0=2=0(\text{mod}2)\\
ac:L &= 1.1+1.0+1.1=2=0(\text{mod}2)\\
bc:L &= 1.0+1.1+1.1=2=0(\text{mod}2)\\
abc:L &= 1.1+1.1+1.1=3=1(\text{mod}2)
\end{aligned}$$

Thus $(1), ab, ac, bc$ are run in block 1 and a, b, c, abc are run in block 2.

The block containing the treatment combination (1) is called the **principal block**. The treatment combinations in this block form a group with respect to multiplication modulas 2. This implies any element [except (1)] in the principal block can be generated by multiplying the two other elements in the principal block modulas 2. For example, consider a 2^3 design with ABC confounded. Note that

$$\begin{aligned}
ab.ac &= a^2bc=bc\\
ab.bc &= ab^2c=ac\\
bc.ac &= abc^2=ab
\end{aligned}$$

Treatment combinations in other can be obtained by multiplying the elements of principal block modulas 2 with one element of other block. In 2^3 design if ABC is confounded then principal block contains $(1), ab, bc, ac$. We know a is in the other block. Thus the elements in the other block are

$$\begin{aligned}(1).a &= a\\ ab.a = a^2b &= b\\ bc.a = abc &= abc\\ ac.a = a^2c &= c\end{aligned}$$

This agrees with previous results.

Estimation of Error

To estimate the error it is necessary to replicate the experiment. For example, a 2^3 factorial experiment with ABC must be run in two blocks and the experimenter decides to replicate the design r times.

There are $8r$ observations and $8r - 1$ total degrees of freedom. Because there are $2r$ blocks $2r - 1$ degrees of freedom must be associated with these blocks. The Analysis of variance table for four replicates of a 2^3 design with ABC confounded is

Sources of variation	Degrees of freedom
Blocks	2r - 1
A	1
B	1
C	1
AB	1
AC	1
BC	1
Error	6(r-1)
Total	8r - 1

3.2 Confounding 2^k Factorial Design in Four Blocks

It is possible to 2^k factorial design confounded in four blocks of 2^{k-2} observations each. This design is useful when the number of factors is moderately large, say $k \geq 4$ and block sizes are relatively small.
As an example, consider 2^5 design. If each block hold 8 treatment combinations, then four blocks must be used. Suppose we want to confound the effects of ADE and BCE. The effects have two defining contrasts

$$\begin{aligned} L_1 &= x_1 + x_4 + x_5 \\ L_2 &= x_2 + x_3 + x_5 \end{aligned}$$

associated with them. Now every treatment combination will yield a particualr pair of values of $L_1(\text{mod}2)$ and $L_2(\text{mod}2)$. The set of all posiisble values of the pair (L_1, L_2) is $\{(0,0),(0,1),(1,0),(1,1)\}$. Treatment combinations yielding same values of (L_1, L_2) are assigned to same block. In our example we find

$$\begin{aligned} L_1 = 0, L_2 = 0 \quad &\text{for} \quad (1), ad, bc, abe, ace, bde, cde, abcd \\ L_1 = 1, L_2 = 0 \quad &\text{for} \quad a, d, abc, be, ce, abde, acde, bcd \\ L_1 = 0, L_2 = 1 \quad &\text{for} \quad b, abd, c, ae, abce, de, bcde, acd \\ L_1 = 1, L_2 = 1 \quad &\text{for} \quad e, ade, bce, ab, ac, bd, cd, abcde \end{aligned}$$

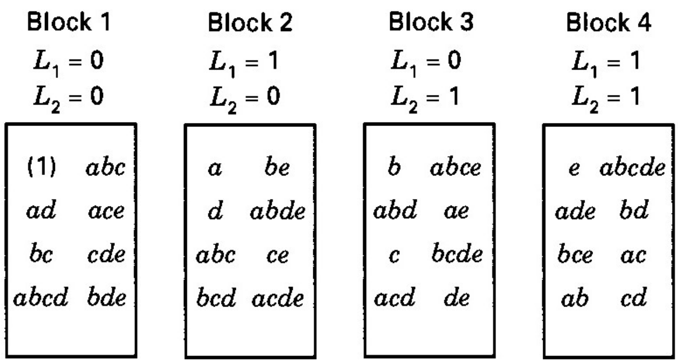

Figure 7-5 The 2^5 design in four blocks with *ADE*, *BCE*, and *ABCD* confounded.

As there are four blocks, there must be three degrees of freedom between them. But ADE and BCE have only one degrees of freedom each, so an additional effect with one degrees of freedom must be confounded. This effect is called **generalized interaction** of ADE and BCE which is defined to be product of ADE and BCE modulas 2. Here the generalized interaction is $(ADE)(BCE) = ABCDE^2 = ABCD$ is also confounded with blocks.

3.3 Confounding the 2^k Factorial Design in 2^p Blocks

The method described above can be extended to the construction of a 2^k factorial design confounded in 2^p blocks, $(p < k)$ where each block conatains 2^{k-p} runs. We select p independent effects to be confounded where by "independent" we mean that no effect is generalized interaction of others.
The blocks may be generated by use of the p defining contrasts $L_1, L_2, \cdots L_p$ associated with effects. The treatment combinations with same value of $(L_1, L_2, \ldots, L_p)$is assigned to same blocks. In addition to these p effects exactly $2^p - p - 1$ will be confounded with blocks, these being generalized interaction of the p independent effects initially chosen to be confounded.
For example, we want to construct 2^6 design confounded in $2^3 = 8$ blocks each with $2^3 = 8$ runs. The three independent effects $ABEF$, $ABCD$ and ACE are chosen to be confounded. Then their $2^p - p - 1 = 4$ generalized interactions will also be confounded. They are

$$\begin{aligned}
(ABEF)(ABCD) &= A^2B^2CDEF = CDEF \\
(ABEF)(ACE) &= A^2BCE^2F = BCF \\
(ABCD)(ACE) &= A^2BC^2DE = BDE \\
(ABEF)(ABCD)(ACE) &= A^3B^2C^2DE^2F = ADF
\end{aligned}$$

Exercise: Construct the 8 blocks.

3.4 Partial Confounding

We have seen that in a example in section 3.1 that 2^3 conducted in two blocks with ABC confounded with blocks. Suppose to estimate the error the experi-

ment is replicated replicated 4 times. If ABC is confounded in each replicate, then design is said to be **completely confounded**. In that no information about the intercation ABC can not be retrived.
Suppose instead of confounding ABC in each of replicates we confound different interaction in each replicate. That is ABC is confounded in replicate I, AB is confounded in replicate II, BC is confounded in replicate III and AC is confounded in replicate IV.

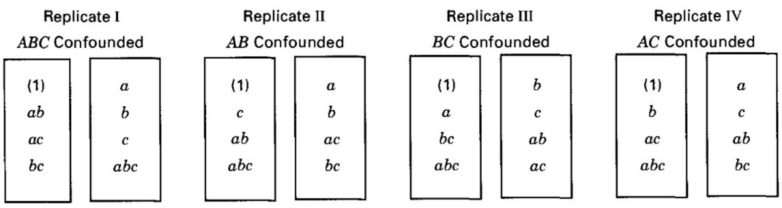

Figure 7-6 Partial confounding in the 2^3 design.

As a result information about ABC can be obtained from the data in replicates II, III and IV., information on AB can be obtained from replicates I, III, IV, information on AC can be obtained from replicates I, II, III and information on BC can be obtained from replicates I, II, IV. We say that 3/4th information can be obtained on the interactions as they unconfounded in 3 replicates. Yates called the ratio 3/4 the **relative information for confounded effects**. This design is said to be **partially confounded**.

The ANOVA table for a 2^3 design with partial confounding is

Sources of variation	Degrees of freedom
Blocks	7
A	1
B	1
C	1
AB	1
AC	1
BC	1
ABC	1
Error	17
Total	31

Comparison of unconfounded, completely confounded and partially confounded 2^k design

In an unconfounded design, the replicate itself a block and inthis case we shall denote the error variance by σ^2. In a design with complete confounding in two blocks, a block is half replicate, two blocks makes up a replicate. In this we denote the error variance as $\sigma^2_{1/2}$. It is expected that $\sigma^2_{1/2} < \sigma^2$.
The variance of the estimator of an effect, main or interaction, in a 2^k experiment in r replicates without confounding is $\sigma^2/r2^{k-2}$, whereas the variance of the estimator of each unconfounded effect in 2^k experiment with r replicates is $\sigma^2_{1/2}/r2^{k-2}$. The information about each effects in an unconfounded design is $r2^{k-2}/\sigma^2$ and the information about each unconfounded effect in a completely confounded design is $r2^{k-2}/\sigma^2_{1/2}$. Since, $\sigma^2_{1/2} < \sigma^2$, the completely confounded design contains more information about unconfounded effects than unconfounded design. But completely confounded design contains no information about effect that has been confounded.
In a partially confounded design with 4 replicates, the information about each of unconfounded effects is $8/\sigma^2_{1/2}$ and as only 3 replicates contain information about confounded replication, the amount of information for them is $6/\sigma^2_{1/2}$.

Effect	Amount of information		
	Unconfounded design	*ABC* completely confounded	*AB*, *AC*, *BC* and *ABC* partially confounded
A	$8/\sigma^2$	$8/\sigma^2_{1/2}$	$8/\sigma^2_{1/2}$
B	$8/\sigma^2$	$8/\sigma^2_{1/2}$	$8/\sigma^2_{1/2}$
C	$8/\sigma^2$	$8/\sigma^2_{1/2}$	$8/\sigma^2_{1/2}$
AB	$8/\sigma^2$	$8/\sigma^2_{1/2}$	$6/\sigma^2_{1/2}$
AC	$8/\sigma^2$	$8/\sigma^2_{1/2}$	$6/\sigma^2_{1/2}$
BC	$8/\sigma^2$	$8/\sigma^2_{1/2}$	$6/\sigma^2_{1/2}$
ABC	$8/\sigma^2$	0	$6/\sigma^2_{1/2}$