

Chapter 8

Factorial Experiments

Factorial experiments involve simultaneously more than one factor and each factor is at two or more levels. Several factors affect simultaneously the characteristic under study in factorial experiments and the experimenter is interested in the main effects and the interaction effects among different factors.

First, we consider an example to understand the utility of factorial experiments.

Example: Suppose the yield from different plots in an agricultural experiment depends upon

1. (i) variety of crop and
- (ii) type of fertilizer.

Both the factors are in the control of the experimenter.

2. (iii) Soil fertility. This factor is not in the control of the experimenter.

In order to compare different crop varieties

- assign it to different plots keeping other factors like irrigation, fertilizer, etc. fixed and the same for all the plots.
- The conclusions for this will be valid only for the crops grown under similar conditions with respect to the factors like fertilizer, irrigation etc.

In order to compare different fertilizers (or different dosage of fertilizers)

- sow single crop on all the plots and vary the quantity of fertilizer from plot to plot.
- The conclusions will become invalid if different varieties of the crop are sown.
- It is quite possible that one variety may respond differently than another to a particular type of fertilizer.

Suppose we wish to compare

- two crop varieties – a and b , keeping the fertilizer fixed and
- three varieties of fertilizers – A , B and C .

This can be accomplished with two randomized block designs (RBD) by assigning the treatments at random to three plots in any block and two crop varieties at random.

The possible arrangement of the treatments may appear can be as follows.

<i>bB</i>	<i>bA</i>	<i>bC</i>
<i>bC</i>	<i>bB</i>	<i>bA</i>
<i>bA</i>	<i>bC</i>	<i>bB</i>

and

<i>aA</i>	<i>aB</i>	<i>aC</i>
<i>aC</i>	<i>aA</i>	<i>aB</i>
<i>aB</i>	<i>aC</i>	<i>aA</i>

With these two *RBDs*,

- the difference among two fertilizers can be estimated
- but the difference among the crop varieties cannot be estimated. The difference among the crop varieties is entangled with the difference in blocks.

On the other hand, if we use three sets of three blocks each and each block having two plots, then

- randomize the varieties inside each block and
- assign treatments at random to three sets.

The possible arrangement of treatment combinations in blocks can be as follows:

<i>bB</i>	<i>aB</i>
<i>aB</i>	<i>bB</i>
<i>bB</i>	<i>aB</i>

,

<i>aC</i>	<i>bC</i>
<i>bC</i>	<i>aC</i>
<i>aC</i>	<i>bC</i>

and

<i>aA</i>	<i>bA</i>
<i>bA</i>	<i>aA</i>
<i>bA</i>	<i>aA</i>

Here the difference between crop varieties is estimable but the difference between fertilizer treatment is not estimable.

Factorial experiments overcome this difficulty and combine each crop with each fertilizer treatment. There are six treatment combinations as

aA, aB, aC, bA, bB, bC.

Keeping the total number of observations to be 18 (as earlier), we can use *RBD* with three blocks with six plots each, e.g.

<i>bA</i>	<i>aC</i>	<i>aB</i>	<i>bB</i>	<i>aA</i>	<i>bC</i>
<i>aA</i>	<i>aC</i>	<i>bC</i>	<i>aB</i>	<i>bB</i>	<i>bA</i>
<i>bB</i>	<i>aB</i>	<i>bA</i>	<i>aC</i>	<i>aA</i>	<i>bC</i>

Now we can estimate

- the difference between crop varieties and
- the difference between fertilizer treatments. Factorial experiments involve simultaneously more than one factor each at two or more levels.

If the number of levels for each factor is the same, we call it is a **symmetrical factorial experiment**.

If the number of levels of each factor is not the same, then we call it as a **symmetrical or mixed factorial experiment**.

We consider only symmetrical factorial experiments.

Through the factorial experiments, we can study

- the individual effect of each factor and
- interaction effect.

Now we consider a 2^2 factorial experiment with an example and try to develop and understand the theory and notations through this example.

General notation for representing the factors is to use capital letters, e.g., A , B , C etc. and levels of a factor are represented in small letters. For example, if there are two levels of A , they are denoted as a_0 and a_1 . Similarly, the two levels of B are represented as b_0 and b_1 . Another alternative representation to indicate the two levels of A is 0 (for a_0) and 1 (for a_1). The factors of B are then 0 (for b_0) and 1 (for b_1).

Note: An important point to remember is that the factorial experiments are conducted in the design of an experiment. For example, the factorial experiment is conducted as an RBD.

Factorial experiments with factors at two levels (2^2 factorial experiment):

Suppose in an experiment, the values of current and voltage in an experiment affect the rotation per minutes (*rpm*) of fan speed. Suppose there are two levels of current.

- 5 Ampere, call it as level 1 (C_1) and denote it as a_0
- 10 Ampere, call it as level 2 (C_1) and denote it as a_1 .

Similarly, the two levels of voltage are

- 200 volts, call it as level 1 (V_0) and denote it as b_0
- 220 volts, call it as level 2 (V_1) and denote it as b_1 .

The two factors are denoted as A , say for current and B , say for voltage.

In order to make an experiment, there are 4 different combinations of values of current and voltage.

1. Current = 5 Ampere and Voltage = 200 Volts, denoted as $C_0V_0 \equiv a_0b_0$
2. Current = 5 Ampere and Voltage = 220 Volts, denoted as $C_0V_1 \equiv a_0b_1$.
3. Current = 10 Ampere and Voltage = 200 Volts, denoted as $C_1V_0 \equiv a_1b_0$
4. Current = 10 Ampere and Voltage = 220 Volts, denoted as $C_1V_1 \equiv a_1b_1$

The responses from those treatment combinations are represented by $a_0b_0 \equiv (1)$, $(a_0b_1) \equiv (b)$, $(a_1b_0) \equiv (a)$ and $(a_1b_1) \equiv (ab)$, respectively.

Now consider the following:

$$\text{I. } \frac{(C_0V_0) + (C_0V_1)}{2} : \text{Average effect of voltage for the current level } C_0$$

$$: \frac{(a_0b_0) + (a_0b_1)}{2} \equiv \frac{(1) + (b)}{2}$$

$$\text{II. } \frac{(C_1V_0) + (C_1V_1)}{2} : \text{Average effect of voltage for the current level } C_1$$

$$: \frac{(a_1b_0) + (a_1b_1)}{2} \equiv \frac{(a) + (ab)}{2}$$

Compare these two group means (or totals) as follows:

Average effect of V_1 level – Average effect at V_0 level

$$= \frac{(b) + (ab)}{2} - \frac{(1) + (a)}{2}$$

$$= \text{Main effect of voltage}$$

$$= \text{Main effect of } B.$$

Comparison like

$$(C_0V_1) - (C_0V_0) \equiv (a) - (1): \text{ indicate the effect of voltage at current level } C_0$$

and

$$(C_1V_1) - (C_1V_0) \equiv (ab) - (b): \text{ indicate the effect of voltage at current level } C_1.$$

The average interaction effect of voltage and current can be obtained as

$$\begin{aligned} & \left(\text{Average effect of voltage} \right)_{\text{at current level } I_o} - \left(\text{Average effect of voltage} \right)_{\text{at current level } I_1} \\ &= \text{Average effect of voltage at different levels of current.} \\ &= \frac{(C_1V_1) - (C_1V_o)}{2} - \frac{(C_oV_1) - (C_oV_o)}{2} \\ &= \frac{(ab) - (b)}{2} - \frac{(a) - (1)}{2} \\ &= \text{Average interaction effect.} \end{aligned}$$

Similarly

$$\begin{aligned} \frac{(C_oV_o) + (C_1V_o)}{2} &= \frac{(1) + (b)}{2} : \text{Average effect of current at voltage level } V_o. \\ \frac{(C_oV_1) + (C_1V_1)}{2} &= \frac{(a) + (ab)}{2} : \text{Average effect of current at voltage level } V_1 \end{aligned}$$

Comparison of these two as

$$\begin{aligned} & \left(\text{Average effect of current} \right)_{\text{at voltage level } V_o} - \left(\text{Average effect of current} \right)_{\text{at voltage level } V_1} \\ &= \frac{(C_oV_1) + (C_1V_1)}{2} - \frac{(C_oV_o) + (C_1V_o)}{2} \\ &= \frac{(a) + (ab)}{2} - \frac{(1) + (b)}{2} \\ &= \text{Main effect of current} \\ &= \text{Main effect of A.} \end{aligned}$$

Comparison like

$$\begin{aligned} (C_1V_o) - (C_oV_o) &= (b) - (1) : \text{Effect of current at voltage level } V_o \\ (C_1V_1) - (C_oV_1) &= (ab) - (a) : \text{Effect of current at voltage level } V_1 \end{aligned}$$

The average interact effect of current and voltage can be obtained as

$$\begin{aligned} & \left(\text{Average effect of current} \right)_{\text{at voltage level } V_o} - \left(\text{Average effect of current} \right)_{\text{at voltage level } V_1} \\ &= \text{Average effect of current at different levels of voltage} \\ &= \frac{(C_1V_1) - (C_oV_1)}{2} - \frac{(C_1V_o) - (C_oV_o)}{2} \\ &= \frac{(ab) - (a)}{2} - \frac{(b) - (1)}{2} \\ &= \text{Average interaction effect} \\ &= \text{Same as average effects of voltage at different levels of current.} \end{aligned}$$

(It is expected *that* the interaction effect of current and voltage is same as the interaction effect of voltage and current).

The quantity

$$\frac{(C_0V_0) + (C_1V_0) + (C_0V_1) + (C_1V_1)}{4} = \frac{(1) + (a) + (b) + (ab)}{4}$$

gives the **general mean effect** of all the treatment combination.

Treating (ab) as $(a)(b)$ symbolically (mathematically and conceptually, it is incorrect), we can now express all the main effects, interaction effect and general mean effect as follows:

$$\text{Main effect of } A = \frac{(a) + (ab)}{2} - \frac{(1) + (b)}{2} = \frac{1}{2}[(ab) - (b) + (a) - (1)] = \frac{(a-1)(b+1)}{2}$$

$$\text{Main effect of } B = \frac{(b) + (ab)}{2} - \frac{(1) + (a)}{2} = \frac{1}{2}[(ab) - (a) + (b) - (1)] = \frac{(a+1)(b-1)}{2}$$

$$\text{Interaction effect of } A \text{ and } B = \frac{(ab) - (b)}{2} - \frac{(a) - (1)}{2} = \frac{1}{2}[(ab) - (a) + (1) - (b)] = \frac{(a-1)(b-1)}{2}$$

$$\text{General mean effect } (M) = \frac{(1) + (a) + (b) + (ab)}{4} = \frac{1}{4}[(1) + (a) + (b) + (ab)] = \frac{(a+1)(b+1)}{4}$$

Notice the roles of + and – signs as well as the divisor.

- There are two effects related to A and B .
- To obtain the effect of a factor, write the corresponding factor with – sign and others with + sign. For example, in the main effect of A , a occurs with – sign as in $(a - 1)$ and b occurs with + sign as in $(b + 1)$.
- In AB , both the effects are present so a and b both occur with + signs as in $(a + 1)(b + 1)$.
- Also note that the main and interaction effects are obtained by considering the typical differences of averages, so they have divisor 2 whereas the general mean effect is based on all the treatment combinations and so it has divisor 4.
- There is a well defined statistical theory behind this logic but this logic helps in writing the final treatment combination easily. This is demonstrated later with appropriate reasoning.

Other popular notations of treatment combinations are as follows:

$$a_0b_0 \equiv 0 \quad 0 \equiv I$$

$$a_0b_1 \equiv 0 \quad 1 \equiv a$$

$$a_1b_0 \equiv 1 \quad 0 \equiv b$$

$$a_1b_1 \equiv 1 \quad 1 \equiv ab.$$

Sometimes 0 is referred to as ‘low level’ and 1 is referred to ‘high level’.

Here I denote that both factors are at lower levels (a_0b_0 or 00). This is called as the **control treatment**.

These effects can be represented in the following table

Factorial effects	Treatment combinations				Divisor
	(1)	(a)	(b)	(ab)	
<i>M</i>	+	+	+	+	4
<i>A</i>	-	+	-	+	2
<i>B</i>	-	-	+	+	2
<i>AB</i>	+	-	-	+	2

The model corresponding to 2^2 factorial experiment is

$$y_{ijk} = \mu + A_i + B_j + (AB)_{ij} + \varepsilon_{ijk}, \quad i = 1, 2, j = 1, 2, k = 1, 2, \dots, n$$

where n observations are obtained for each treatment combinations.

When the experiments are conducted factor by factor, then much more resources are required in comparison to the factorial experiment. For example, if we conduct *RBD* for three-levels of voltage V_0, V_1 and V_2 and two levels of current I_0 and I_1 , then to have 10 degrees of freedom for the error variance, we need

- 6 replications on voltage
- 11 replications on current.

So the total number of fans needed is 40.

For the factorial experiment with 6 combinations of 2 factors, the total number of fans needed are 18 for the same precision.

We have considered the situation up to now by assuming only one observation for each treatment combination, i.e., no replication. If r replicated observations for each of the treatment combinations are obtained, then the expressions for the main and interaction effects can be expressed as

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

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

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

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

Now we detail the statistical theory and concepts related to these expressions.

Let $Y_* = ((1), a, b, ab)'$ be the vector of total response values. Then

$$A = \frac{1}{2r} \ell'_A Y_* = \frac{1}{2r} (-1 \ 1 \ -1 \ 1) Y_*$$

$$B = \frac{1}{2r} \ell'_B Y_* = \frac{1}{2r} (-1 \ -1 \ 1 \ 1) Y_*$$

$$AB = \frac{1}{2r} \ell'_{AB} Y_* = \frac{1}{2r} (1 \ -1 \ -1 \ 1) Y_*$$

Note that A , B and AB are the linear contrasts. Recall that a linear parametric function is estimable only when it is in the form of linear contrast. Moreover, A , B and AB are the linear orthogonal contrasts in the total response values $(1), a, b, ab$ except for the factor $1/2r$.

The sum of squares of a linear parametric function $\ell' y$ is given by $\frac{(\ell' y)^2}{\ell' \ell}$. If there are r replicates, then the sum of squares is $\frac{(\ell' y)^2}{r \ell' \ell}$. It may also be recalled under the normality of y 's, this sum of squares has a Chi-square distribution with one degree of freedom (χ_1^2). Thus the various associated sum of squares due to A, B and AB are given by the following:

$$SSA = \frac{(\ell'_A Y_*)^2}{r \ell'_A \ell_A} = \frac{1}{4r} (ab + a - b - (1))^2$$

$$SSB = \frac{(\ell'_B Y_*)^2}{r \ell'_B \ell_B} = \frac{1}{4r} (ab + b - a - (1))^2$$

$$SSAB = \frac{(\ell'_{AB} Y_*)^2}{r \ell'_{AB} \ell_{AB}} = \frac{1}{4r} (ab + (1) - a - b)^2.$$

Each of SSA , SSB and $SSAB$ has χ_1^2 under normality of Y_* .

The sum of squares due to total is computed as usual

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

where

$$G = \sum_{i=1}^2 \sum_{j=1}^2 \sum_{k=1}^r y_{ijk}$$

is the grand total of all the observations.

The TSS has χ^2 distribution with $(2^2 r - 1)$ degrees of freedom.

The sum of squares due to error is also computed as usual as

$$SSE = TSS - SSA - SSB - SSAB$$

which has χ^2 distribution with

$$(4r - 1) - 1 - 1 - 1 = 4(r - 1)$$

degrees of freedom.

The mean squares are

$$MSA = \frac{SSA}{1},$$

$$MSB = \frac{SSB}{1},$$

$$MSAB = \frac{SSAB}{1},$$

$$MSE = \frac{SSE}{4(r - 1)}.$$

The F -statistic corresponding to A, B and AB are

$$F_A = \frac{MSA}{MSE} \sim F(1, 4(r - 1)) \text{ under } H_0,$$

$$F_B = \frac{MSB}{MSE} \sim F(1, 4(r - 1)) \text{ under } H_0,$$

$$F_{AB} = \frac{MSAB}{MSE} \sim F(1, 4(r - 1)) \text{ under } H_0.$$

The ANOVA table is case of 2^2 factorial experiment is given as follows:

Source	Sum of squares	Degrees of freedom	Mean squares	F
A	SSA	1	MSA	$F_A = \frac{MSA}{MSE}$ $F_B = \frac{MSB}{MSE}$ $F_{AB} = \frac{MSAB}{MSE}$
B	SSB	1	MSB	
AB	$SSAB$	1	$MSAB$	
$Error$	SSE	$4(r - 1)$	MSE	
Total	TSS	$4r - 1$		

The decision rule is to reject the concerned null hypothesis when the value of the concerned F statistic

$$F_{\text{effect}} > F_{1-\alpha}(1, 4(r - 1)).$$

2³ Factorial experiment:

Suppose that in a complete factorial experiment, there are three factors - A, B and C , each at two levels, viz., $a_0, a_1; b_0, b_1$ and c_0, c_1 respectively. There are a total of eight number of combinations:

$$a_0b_0c_0, a_0b_0c_1, a_0b_1c_0, a_0b_1c_1, \\ a_1b_0c_0, a_1b_0c_1, a_1b_1c_0, a_1b_1c_1.$$

Each treatment combination has r replicates, so the total number of observations are $N = 2^3 r = 8r$ that are to be analyzed for their influence on the response.

Assume the total response values are

$$Y_* = [(1), a, b, ab, c, ac, bc, abc]'$$

The response values can be arranged in a three-dimensional contingency table. The effects are determined by the linear contrasts

$$\ell'_{effect} Y_* = \ell'_{effect} ((1), a, b, ab, c, ac, bc, abc)$$

using the following table:

Factorial effect	Treatment combinations							
	(1)	a	b	ab	c	ac	bc	abc
I	+	+	+	+	+	+	+	+
A	-	+	-	+	-	+	-	+
B	-	-	+	+	-	-	+	+
AB	+	-	-	+	+	-	-	+
C	-	-	-	-	+	+	+	+
AC	+	-	+	-	-	+	-	+
BC	+	+	-	-	-	-	+	+
ABC	-	+	+	-	+	-	-	+

Note that once a few rows have been determined in this table, rest can be obtained by simple multiplication of the symbols. For example, consider the column corresponding to a , we note that A has + sign, B has - sign, so AB has - sign (sign of $A \times$ sign of B). Once AB has - sign, C has - sign then ABC has (sign of $AB \times$ sign of C) which is + sign and so on.

The first row is a basic element. With this $a = 1'Y_*$ can be computed where 1 is a column vector of all elements unity. If other rows are multiplied with the first row, they stay unchanged (therefore we call it as identity and denoted as I). Every other row has the same number of + and - signs. If + is replaced by 1 and - is replaced by -1, we obtain the vectors of orthogonal contrasts with the norm $8 (= 2^3)$.

If each row is multiplied by itself, we obtain I (first row). The product of any two rows leads to a different row in the table. For example

$$A.B = AB$$

$$AB.B = AB^2 = A$$

$$AC.BC = A.C^2.BB = AB.$$

The structure in the table helps in estimating the average effect.

For example, the average effect of A is

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

which has the following explanation.

- (i) Average effect of A at low level of B and $C \equiv (a_1b_0c_0) - (a_0b_0c_0) \equiv \frac{[(a) - (1)]}{r}$.
- (ii) Average effect of A at high level of B and low level of $C \equiv (a_1b_1c_0) - (a_0b_1c_0) \equiv \frac{[(ab) - (b)]}{r}$.
- (iii) Average effect of A at low level of B and high level of $C \equiv (a_1b_0c_1) - (a_0b_0c_1) \equiv \frac{[(ac) - (c)]}{r}$.
- (iv) Average effect of A at high level of B and $C \equiv (a_1b_1c_1) - (a_0b_1c_1) \equiv \frac{[(abc) - (bc)]}{r}$.

Hence for all combinations of B and C , the average effect of A is the average of all the average effects in (i)-(iv).

Similarly, other main and interaction effects are as follows:

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

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

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

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

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

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

Various sum of squares in the 2^3 factorial experiment are obtained as

$$SS(\text{Effect}) = \frac{(\text{linear contrast})^2}{8r} = \frac{(\ell'_{\text{effect}} Y_*)^2}{r \ell'_{\text{effect}} \ell_{\text{effect}}}$$

which follow a Chi-square distribution with one degree of freedom under normality of Y_* . The corresponding mean sum of squares is obtained as

$$MS(\text{Effect}) = \frac{SS(\text{Effect})}{\text{Degrees of freedom}}.$$

The corresponding F -statistics are obtained by

$$F_{\text{Effect}} = \frac{MS(\text{Effect})}{MS(\text{Error})}$$

which follows an F -distribution with degrees of freedoms 1 and error degrees of freedom under the respective null hypothesis. The decision rule is to reject the corresponding null hypothesis at the α level of significance whenever

$$F_{\text{effect}} > F_{1-\alpha}(1, df_{\text{error}}).$$

These outcomes are presented in the following ANOVA table

Sources	Sum of squares	Degrees of freedom	Mean sum of squares	F
A	SSA	1	$MSA = SSA / 1$	F_A
B	SSB	1	$MSB = SSB / 1$	F_B
AB	$SSAB$	1	$MSAB = SSAB / 1$	F_{AB}
C	SSC	1	$MSC = SSC / 1$	F_C
AC	$SSAC$	1	$MSAC = SSAC / 1$	F_{AC}
BC	$SSBC$	1	$MSBC = SSBC / 1$	F_{BC}
ABC	$SSABC$	1	$MSABC = SSABC / 1$	F_{ABC}
Error	$SS(\text{Error})$	$8(r-1)$	$MS(\text{Error}) = SS(\text{Error}) / \{8(r-1)\}$	
Total	TSS	$8r-1$		

2ⁿ Factorial experiment:

Based on the theory developed for 2² and 2³ factorial experiments, we now extend them for the 2ⁿ factorial experiment.

- Capital letters A, B, C, \dots denote the **factors**. They are the main effect contrast for the factors A, B, C, \dots
- AB, AC, BC, \dots denote the first order or **2-factor interactions**
- ABC, ABD, BCD, \dots denote the **second-order** or **3-factor interactions** and so on.
- Each of the main effect and interaction effect carries one degree of freedom.
- Total number of main effects = $\binom{n}{1} = n$.
- Total number of first-order interactions = $\binom{n}{2}$.
- Total number of second-order interactions = $\binom{n}{3}$

and so on.

Standard order for treatment combinations:

The list of treatments can be expressed in a standard order.

- For one factor A , the standard order is $(1), a$.
- For two factors A and B , the standard order is obtained by adding b and ab in the standard order of one factor A . This is derived by multiplying (1) and a by b , i.e.
 $b \times \{(1), a\} = (1), a, b, ab$.
- For three factors, add c, ac, bc and abc which are derived by multiplying the standard order of A and B by c , i.e.
 $c \times \{(1), a, b, ab\} = (1), a, b, ab, c, ac, bc, abc$.

Thus the standard order of any factor is obtained step by step by multiplying it with an additional letter to preceding standard order.

For example, the standard order of A, B, C and D is 2⁴ factorial experiment is

$$\begin{aligned} & (1), a, b, ab, c, ac, bc, abc, d \times \{(1), a, b, ab, c, ac, bc, abc\} \\ & = (1), a, b, ab, c, ac, bc, abc, d, ad, bd, abd, cd, acd, bcd, abcd. \end{aligned}$$

How to find the contrasts for main effects and interaction effect:

Recall that earlier, we had illustrated the concept in writing the contrasts for main and interaction effects.

For example, in a 2^2 factorial experiment, we had expressed

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

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

Note that each effect has two component - divisor and contrast. When the order of factorial increases, it is cumbersome to derive such expressions. Some methods have been suggested to write the expressions for factorial effects. First, we detail how to write divisor and then illustrate the methods for obtaining the contrasts.

How to write divisor:

In a 2^n factorial experiment,

- the general mean effect has divisor 2^n and
- any effect (main or interaction) has divisor 2^{n-1} .

For example, in a 2^6 factorial experiment, the general mean effect has divisor 2^6 and any main effect or interaction effect of any order has divisor $2^{6-1} = 2^5$.

If r replicates of each effect are available, then

- the general mean effect has divisor $r2^n$ and
- any main effect or interaction effect of any order has a divisor $r2^{n-1}$.

How to write contrasts:

Method 1:

Contrast belonging to the main effects and the interaction effects are written as follows:

$$A = (a-1)(b+1)(c+1)\dots(z+1)$$

$$B = (a+1)(b-1)(c+1)\dots(z+1)$$

$$C = (a+1)(b+1)(c-1)\dots(z+1)$$

⋮

$$AB = (a-1)(b-1)(c+1)\dots(z+1)$$

$$BC = (a+1)(b-1)(c-1)\dots(z+1)$$

⋮

$$ABC = (a-1)(b-1)(c-1)\dots(z+1).$$

⋮

$$ABC\dots Z = (a-1)(b-1)(c-1)\dots(z-1)$$

Look at the pattern of assigning + and – signs on the right-hand side. The letters common on left and right hand sides of the equality (=) sign (irrespective of small or capital letters) contain – sign and rest contain + sign.

The expression on right-hand side when simplified algebraically give the contrasts in terms of treatment combination. For example, in a 2^3 factorial

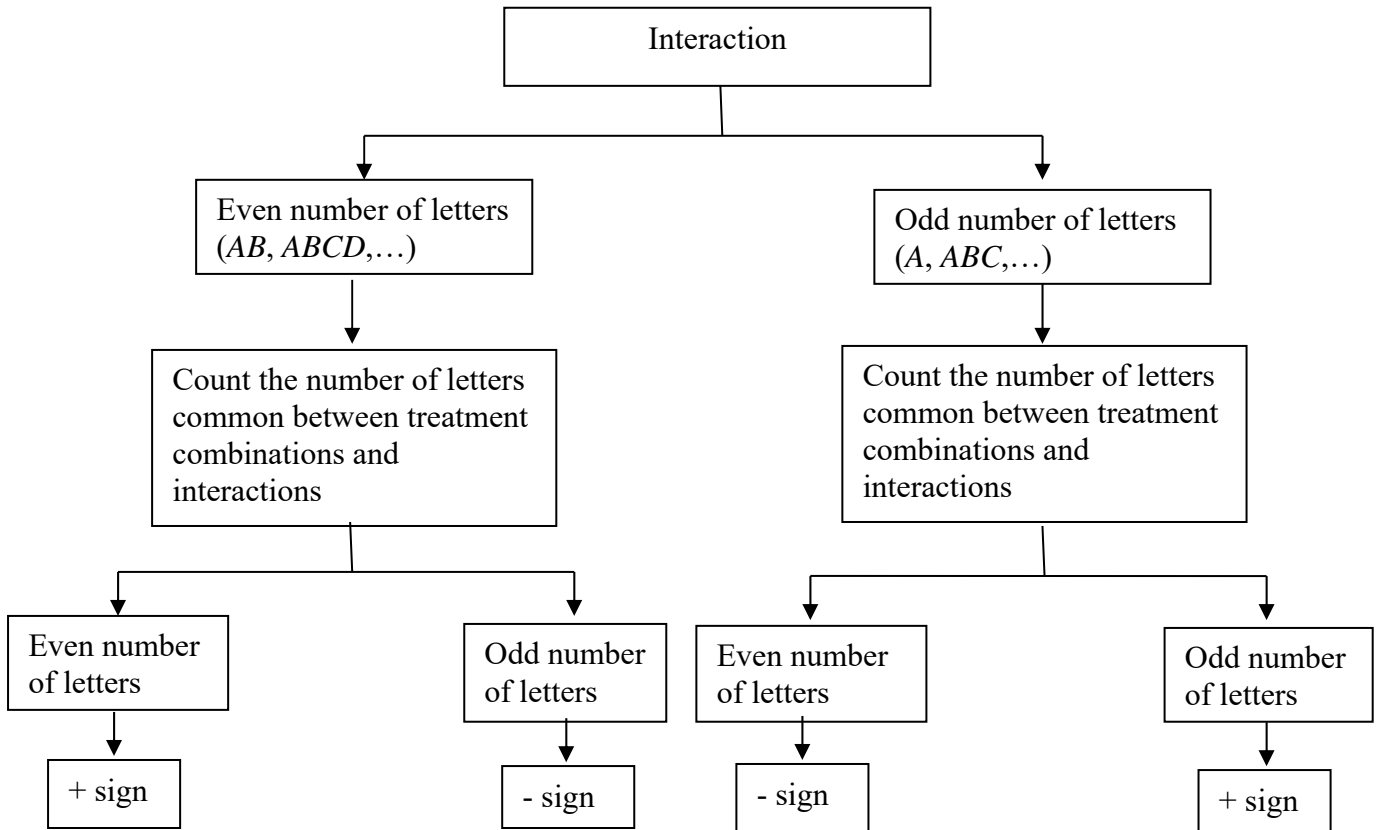
$$\begin{aligned}
 A &= \frac{1}{2^{3-1}}(a-1)(b+1)(c+1) \\
 &= \frac{1}{4}[-(1)+(a)-(b)+(ab)-(c)+(ac)-(bc)+(abc)] \\
 M &= \frac{1}{2^3}(a+1)(b+1)(c+1) \\
 &= \frac{1}{8}[(1)+(a)+(b)+(ab)+(c)+(ac)+(bc)+(abc)]
 \end{aligned}$$

Method 2

- Form a table such that
 - rows correspond to the main or interaction effect and
 - columns correspond to treatment combinations (or another way round)
- + and – signs in the table indicate the sign of the treatment combinations of main and interaction effects.
- Signs are determined by the “rule of odds and evens” given as follows:
 - if the interaction has an even number of letters ($AB, ABCD, \dots$), a treatment combination having an even number of letters common with the interaction enters with a + sign and one with an odd number of letters common enters with a – sign.
 - if the interaction has an odd number of letters (A, ABC, \dots), the rule is reversed.
- Once few rows are filled up, others can be obtained through multiplication rule. For example, the sign of $ABCD$ is obtained as

$$(\text{sign of } A \times \text{sign of } BCD) \quad \text{or} \quad (\text{sign of } AB \times \text{sign of } CD).$$
- Treatment combination (1) is taken to have an even number (zero) of letters common with every interaction.

This rule of assignment of + or – is illustrated in the following flow diagram:



For example, in a 2^3 factorial experiment, write

- rows for main and interaction effects and
- columns for treatment combinations in standard order.
- Take treatment combination (1) to have an even number (zero) of letter common with every interaction.

This gives the following table

Factorial effect	Treatment combinations							
	(1)	<i>a</i>	<i>b</i>	<i>ab</i>	<i>c</i>	<i>ac</i>	<i>bc</i>	<i>abc</i>
<i>I</i>	+	+	+	+	+	+	+	+
<i>A</i>	-	+	-	+	-	+	-	+
<i>B</i>	-	-	+	+	-	-	+	+
<i>AB</i>	+	-	-	+	+	-	-	+
<i>C</i>	-	-	-	-	+	+	+	+
<i>AC</i>	+	-	+	-	-	+	-	+
<i>BC</i>	+	+	-	-	-	-	+	+
<i>ABC</i>	-	+	+	-	+	-	-	+

Sums of squares:

Suppose 2^n factorial experiment is carried out in a randomized block design with r replicates.

Denote the total yield (output) from r plots (experimental units) receiving a particular treatment combination by the same symbol within a square bracket. For example, $[ab]$ denotes the total yield from the plots receiving the treatment combination (ab) .

In a 2^2 factorial experiment, the factorial effect totals are

$$[A] = [ab] - [b] + [a] - [1]$$

$[ab]$ = treatment total, i.e. the sum of r observations in which both the factors A and B are at the second level.

$[a]$ = treatment total, i.e., the sum of r observations in which factor A is at the second level and factor B is at the first level

$[b]$ = treatment total, i.e., the sum of r observations in which factor A is at the first level and factor B is at the second level.

$[1]$ = treatment total, i.e. the sum of r observations in which both the factors A and B are at the first level.

$$[A] = \sum_{i=1}^r [y_{i(ab)} - y_{i(b)} + y_{i(a)} - y_{i(1)}]$$

$$= \ell'_A y_A \text{ (say).}$$

where ℓ_A is a vector of +1 and =1 and y_A is a vector denoting the responses from ab, b, a and 1 .

Similarly, other effects can also be found.

Similarly, other effects can also be found.

The sum of squares due to a particular effect is obtained as

$$\frac{[\text{Total yield}]^2}{\text{Total number of observations}}.$$

In a 2^2 factorial experiment in an RBD, the sum of squares due to A is

$$SSA = \frac{(\ell'_A y_A)^2}{r2^2}.$$

In a 2^n factorial experiment in an RBD, the divisor will be $r.2^n$. If Latin square design is used based on $2^n \times 2^n$ Latin square, then r is replaced by 2^n .

Yates method of computation of the sum of squares:

Yates method gives a systematic approach to find the sum of squares. We are not presenting here the complete method. Only the part which is used for computing only the sum of squares is presented and the method to verify them is not presented.

It has the following steps

1. First, write the treatment combinations in the standard order in the column at the beginning of the table, called a **treatment column**.
2. Find the total yield for each treatment. Write this as the second column of the table, called a **yield column**.
3. Obtain columns (1),(2),...,(*n*) successively
 - (i) obtain column (1) from yield column
 - (a) upper half is obtained by adding yields in pairs.
 - (b) the second half is obtained by taking differences in pairs, the difference obtained by subtracting the first term of pairs from the second term.
 - (ii) The columns (2),(3),...,(*n*) are obtained from preceding ones in the same manner as used for getting (1) from the yield columns.
4. This process of finding columns is repeated *n* times in 2^n factorial experiment.
5. Sum of squares due to interaction = $\frac{[\text{column}(n)]^2}{\text{Total number of observations}}$

Example: Yates procedure for a 2^2 factorial experiment

Treatment Combination	Yield (total from all <i>r</i> replicates)	(1)	(2)
(1)	(1)	(1) + (<i>a</i>)	(1) + (<i>a</i>) + (<i>b</i>) + (<i>ab</i>) = [<i>M</i>]
<i>a</i>	(<i>a</i>)	(<i>b</i>) + (<i>ab</i>)	-(1) + (<i>a</i>) - (<i>b</i>) + (<i>ab</i>) = [<i>A</i>]
<i>b</i>	(<i>b</i>)	(<i>a</i>) - (1)	-(1) - (<i>a</i>) + (<i>b</i>) + (<i>ab</i>) = [<i>B</i>]
<i>ab</i>	(<i>ab</i>)	(<i>ab</i>) - (<i>b</i>)	(1) - (<i>a</i>) - (<i>b</i>) + (<i>ab</i>) = [<i>AB</i>]

Note: The columns are repeatedly obtained 2 times due to 2^2 factorial experiment.

$$\text{Now } SSA = \frac{[A]^2}{4r}; \quad SSB = \frac{[B]^2}{4r}; \quad SSAB = \frac{[AB]^2}{4r}$$

Example: Yates procedure for 2^3 factorial experiment.

Treatment	Yield (total from all r replicates)				
(1)	(2)	(3)	(4)	(5)	(6)
1	(1)	$u_1 = (1) + (a)$	$v_1 = u_1 + u_2$	$w_1 = v_1 + v_2$	[M]
a	(a)	$u_2 = (b) + (ab)$	$v_2 = u_3 + u_4$	$w_2 = v_3 + v_4$	[A]
b	(b)	$u_3 = (c) + (ac)$	$v_3 = u_5 + u_6$	$w_3 = v_5 + v_6$	[B]
ab	(ab)	$u_4 = (bc) + (abc)$	$v_4 = u_7 + u_8$	$w_4 = v_7 + v_8$	[AB]
c	(ac)	$u_5 = (a) - (1)$	$v_5 = u_2 - u_7$	$w_5 = v_2 - v_1$	[C]
ac	(ac)	$u_6 = (ab) - (b)$	$v_6 = u_4 - u_3$	$w_6 = v_4 - v_3$	[AC]
bc	(bc)	$u_7 = (ac) - (c)$	$v_7 = v_6 - u_5$	$w_7 = v_6 - v_5$	[BC]
abc	(abc)	$u_8 = (abc) - (bc)$	$v_8 = u_8 - u_7$	$w_8 = v_8 - v_7$	[ABC]

The sum of squares are obtained as follows when the design is RBD.

$$SS(Effect) = \frac{[Effect]^2}{r \cdot 2^3}.$$

For analysis of 2^n factorial experiment, the analysis of variance involves the partitioning of treatment sum of squares so as to obtain sum of squares due to main and interaction effects of factors. These sum of squares are mutually orthogonal, so

Treatment $SS =$ Total of SS due to main and interaction effects.

For example:

In 2^2 factorial experiment in an RBD with r replications, the division of degrees of freedom and the treatment sum of squares are as follows:

Source	Degree of freedom	Sum of squares
Replications	$r - 1$	
Treatments	$4 - 1 = 3$	
- A	1	$[A]^2 / 4r$
- B	1	$[B]^2 / 4r$
- AB	1	$[AB]^2 / 4r$
Error	$3(r - 1)$	
Total	$4r - 1$	

The decision rule is to reject the concerned null hypothesis when the related F -statistic $F_{effect} > F_{1-\alpha}(1, 3(r-1))$.