Factorial Designs

Basic Definitions and Principles

By a factorial design, we mean that in each complete trial or replicate of the experiment all possible combinations of the levels of the factors are investigated.

The effect of a factor is defined to be the change in response produced by a change in the level of the factor. This is frequently called a main effect because it refers to the primary factors of interest in the experiment.

In some experiments, we may find that the difference in response between the levels of one factor is not the same at all levels of the other factors. When this occurs, there is an interaction between the factors.

Advantages of Factorial Designs

- They are more efficient than one-factor-at a time experiment
- Factorial design is necessary when interactions may be present to avoid misleading conclusions
- Factorial design allows the effects of a factor to be estimated at several levels of the other factors yielding conclusions that are a valid over the range of experimental conditions.

Two-Factor Factorial Designs

The simplest types of factorial designs involve only two factors or sets of treatments. There are a levels of factor A and b levels of factor B, and these are arranged in a factorial design; that is, each replicate of the experiment contains all ab treatment combinations. In general, there are n replicates.

The observations in a factorial experiment can be described in a model. There are several ways to write the model for a factorial experiment. The **effects model** is

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijk}, \quad i = 1, ..., a, j = 1, ..., b, k = 1, ..., n$$

where μ is the overall mean effect, τ_i is the effect of the $i-th$ level of the row factor A, β_j is the effect of the $j-th$ level of column factor B, $(\tau\beta)_{ij}$ is the effect of the interactions between τ_i and β_j and ϵ_{ijk} is a random error component.

Both factors are assumed to be **fixed** and the treatment effects are defined as deviations from the overall mean so $\sum_{i=1}^{a} \tau_i = 0$ and $\sum_{j=1}^{b} \beta_j = 0$.

Similarly the interaction effects are fixed and are defined such that $\sum_{i=1}^{a} (\tau\beta)_{ij} = \sum_{i=1}^{b} (\tau\beta)_{ij} = 0$

Because there are n replicates of the experiment there are abn total observations.

The General arrangement for a two factor factorial design is:

		Factor B			
		1	2	. . .	b
Factor A	1	$y_{111}, y_{112}, \ldots, y_{11n}$	$y_{121}, y_{122}, \ldots, y_{12n}$		$y_{1b1}, y_{1b2}, \ldots, y_{1bn}$
	2	$y_{211}, y_{212}, \ldots, y_{21n}$	$y_{221}, y_{222}, \ldots, y_{22n}$		$y_{2b1}, y_{2b2}, \ldots, y_{2bn}$
	⋮				
	a	$y_{a11}, y_{a12}, \ldots, y_{a1n}$	$y_{a21}, y_{a22}, \ldots, y_{a2n}$		$y_{ab1}, y_{ab2}, \ldots, y_{abn}$

Statistical Analysis of the Fixed Effects Model

Let $y_{..}$ denote the total of all observations under the $i-th$ level of factor A. $y_{.j.}$ denote the total of all observations under the $j-th$ level of factor B, $y_{ij.}$ denote the total of all observations in the $ij-th$ cell and $y_{...}$ denote the grand total of all observations.

Define $\bar{y_{i..}}, \bar{y_{.j.}}, \bar{y_{ijk}}$ and $\bar{y_{...}}$ as the corresponding row, column,cell and grand averages. Expressed mathematically as:

$$y_{i.} = \sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk}, \bar{y_{i..}} = \frac{y_{i..}}{bn}$$

$$y_{.j.} = \sum_{i=1}^{a}\sum_{k=1}^{n} y_{ijk}, \bar{y_{.j.}} = \frac{y_{.j.}}{an}$$

$$y_{ijk.} = \sum_{k=1}^{n} y_{ijk}, \bar{y_{ij.}} = \frac{y_{ij.}}{n}$$

$$y_{...} = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk}, \bar{y_{...}} = \frac{y_{...}}{abn}$$

The total sum of squares is computed as:

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk}^2 - \frac{y_{...}^2}{abn}$$

The sum of squares for the main effects are

$$SS_A = \frac{1}{bn}\sum_{i=1}^{a} y_{i..}^2 - \frac{y_{...}^2}{abn}$$

and

$$SS_B = \frac{1}{an}\sum_{j=1}^{b} y_{.j.}^2 - \frac{y_{...}^2}{abn}$$

It is convenient to compute SS_{AB} in two stages. First we compute the sum of squares between ab cell totals which is called the sum of squares due to subtotals

$$SS_{Subtotals} = \frac{1}{n}\sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij.}^2 - \frac{y_{...}^2}{abn}$$

This sum of squares also contains SS_A and SS_B Therefore

$$SS_{AB} = SS_{Subtotals} - SS_A - SS_B$$

Then SS_E is constructed as:

$$SS_E = SS_T - SS_{Subtotals}$$

The ANOVA Table is given as:

Source	SS	df	MS	F
A treatments	SS_A	a-1	MS_A	$\frac{MS_A}{MS_E}$
B treatments	SS_B	b-1	MS_B	$\frac{MS_B}{MS_E}$
Interactions	SS_{AB}	(a-1)(b-1)	MS_{AB}	$\frac{MS_{AB}}{MS_E}$
Error	SS_E	ab(n-1)	MS_E	
Total	SS_T	abn-1		

Example

Analyze the following data involving two factors an engineer is designing a battery for use in a device that will be subjected to some extreme variations in temperature.

	Temperature					
Material Type	15		70		125	
1	130	155	34	40	20	20
	74	180	80	75	82	58
2	150	188	136	122	25	70
	159	126	106	115	58	45
3	138	110	174	120	96	104
	168	160	150	139	82	60

Solution

	Temperature									
Material Type	15			70			125			$y_{i..}$
1	130	155	539	34	40	229	20	20	230	998
	74	180		80	75		82	58		
2	150	188	623	136	122	479	25	70	198	1300
	159	126		106	115		58	45		
3	138	110	576	174	120	583	96	104	342	1501
	168	160		150	139		82	60		
$y_{.j.}$	1738			1291			770			$y_{...} = 3799$

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk}^2 - \frac{y_{...}^2}{abn} = (130)^2 + ... + (60)^2 - \frac{(3799)^2}{36} = 77646.97$$

$$SS_A = \frac{1}{bn}\sum_{i=1}^{a} y_{i..}^2 - \frac{y_{...}^2}{abn} = \frac{1}{3 \times 4}[(998)^2 + (1300)^2 + (1501)^2] - \frac{3799^2}{36} = 10683.72$$

$$SS_B = \frac{1}{an}\sum_{j=1}^{b} y_{.j.}^2 - \frac{y_{...}^2}{abn} = \frac{1}{12}[(1738)^2 + (1291)^2 + (770)^2] - \frac{3799^2}{36} = 39118.72$$

$$SS_{AB} = \frac{1}{n}\sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij.}^2 - \frac{y_{...}^2}{abn} - SS_A - SS_B = \frac{1}{4}[(539)^2 + ... + (342)^2] - \frac{3799^2}{36} - 10683.72 - 39118.72 = 9613.78$$

$$SS_E = SS_T - SS_A - SS_B - SS_{AB} = 77646.97 - 10683.72 - 39118.72 - 9613.78 = 18230.75$$

Source	df	SS	MSS	F
A	2	10683.72	5341.86	7.91
B	2	39118.72	19559.36	28.97
AB	4	9613.78	2403.44	3.56
Error	27	18230.75	675.21	
Total	35	77646.97		

The ANOVA is shown in Table 5.5. Because $F_{0.05,4,27} = 2.73$ we conclude that there is a significant interaction between material types and temperature. Furthermore, $F_{0.05,2,27} = 3.35$ so the main effects of material type and temperature are also significant.

Blocking in a Factorial Experiment

Consider a factorial experiment with two factors (A and B) and n replicates with blocking, the effects model for this new design is given as:

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

where δ_k is the effect of the $k-th$ block.

The ANOVA table for a two-factor factorial design is given as:

Source of Variation	Sum of Squares	Degrees of Freedom	Expected Mean Square	F_0
Blocks	$\frac{1}{ab}\sum_k y_{..k}^2 - \frac{y_{...}^2}{abn}$	$n-1$	$\sigma^2 + ab\sigma_\delta^2$	
A	$\frac{1}{bn}\sum_i y_{i..}^2 - \frac{y_{...}^2}{abn}$	$a-1$	$\sigma^2 + \frac{bn\sum \tau_i^2}{a-1}$	$\frac{MS_A}{MS_E}$
B	$\frac{1}{an}\sum_j y_{.j.}^2 - \frac{y_{...}^2}{abn}$	$b-1$	$\sigma^2 + \frac{an\sum \beta_j^2}{b-1}$	$\frac{MS_B}{MS_E}$
AB	$\frac{1}{n}\sum_i\sum_j y_{ij.}^2 - \frac{y_{...}^2}{abn} - SS_A - SS_B$	$(a-1)(b-1)$	$\sigma^2 + \frac{n\sum\sum(\tau\beta)_{ij}^2}{(a-1)(b-1)}$	$\frac{MS_{AB}}{MS_E}$
Error	Subtraction	$(ab-1)(n-1)$	σ^2	
Total	$\sum_i\sum_j\sum_k y_{ijk}^2 - \frac{y_{...}^2}{abn}$	$abn-1$		

Example

Analyze the following two factor factorial design with blocking.

Operators (blocks)	1		2		3		4	
Filter Type	1	2	1	2	1	2	1	2
Ground clutter								
Low	90	86	96	84	100	92	92	81
Medium	102	87	106	90	105	97	96	80
High	114	93	112	91	108	95	98	83

Solution

The $SS_T, SS_A, SS_B, SS_{AB}$ are calculated as before.

The

$$SS_{Blocks} = \frac{1}{ab}\sum_{k=1}^{n} y^2_{..k} - \frac{y^2_{...}}{abn} = \frac{1}{6}[(572)^2 + (597)^2 + (530)^2] - \frac{(2278)^2}{24} = 402.17$$

The ANOVA table is given as:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0	*P*-Value
Ground clutter (*G*)	335.58	2	167.79	15.13	0.0003
Filter type (*F*)	1066.67	1	1066.67	96.19	<0.0001
GF	77.08	2	38.54	3.48	0.0573
Blocks	402.17	3	134.06		
Error	166.33	15	11.09		
Total	2047.83	23			

$2^k Factorial Design$

Factorial designs are widely used in experiments involving several factors where it is necessary to study the joint effect of the factors on a response.

The most important of these special cases is that of k factors each at only two levels.

These levels may be quantitative, such as two values of temperature, pressure, or time; or they may be qualitative, such as two machines, two operators, the "high" and "low" levels of a factor, or perhaps the presence and absence of a factor.

A complete replicate of such a design requires $2 \times 2 \times ... \times 2 = 2^k$ observations and is called a 2^k factorial design.

2^2 Factorial Design

The 2^2 factorial design is a design with two factors A and B each at two levels.

The levels of the factors may be arbitrarily called low and high.

Example

In a chemical reaction, the reactant concentration is factor A run at two levels, 15% and 25%, and the catalyst is factor B, with two levels, one bag used or two bags used. The experiment is replicated three times, and the data are

The effect of treatment A at the low level of B is $\frac{[a-(1)]}{n}$ and the effect of A at high level of B is $\frac{[ab-b]}{n}$. Therefore the main effect of treatment A is the average of the this two quantities

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