MAT 458-Design of Experiments
(V) Factorial Designs

Fuxia Cheng

Outline

1. Introduction

2. The Two-Factor Factorial Design

- An Example
- The Two-Factor Factorial Design with Fixed Effects
- The Two-Factor Factorial Design with Random Effects
- The Two-Factor Factorial Design with Mixed Effects

3. The ANOVA for factorials Extensions to more than two factors

1. Introduction

Many experiments involve the study of the effects of two or more factors. In general, **factorial designs** are most efficient for this type of experiment.

- General Principles

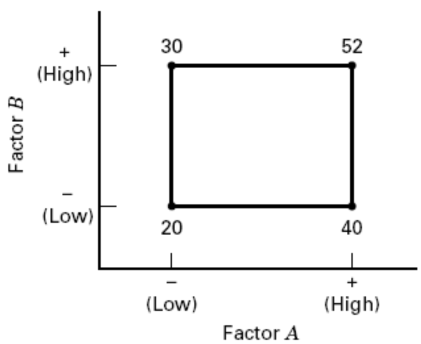
By a **factorial design**, we mean that in each complete trial or replicate of the experiment all possible combinations of the levels of the factors are investigated.

For example, if there are a levels of factor A and b levels of factor B, each replicate contains all ab treatment combinations. When factors are arranged in a factorial design, they are often said to be **crossed**.

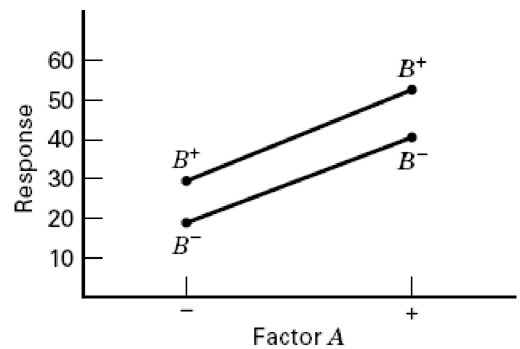
The effect of a factor is defined to be the change in response produced by a change in the level of the factor. This is frequently called a **main effect** because it refers to the primary factors of interest in the experiment.

For example, consider the simple experiment in Figure 5.1.

Some Basic Definitions



■ **FIGURE 5.1** A two-factor factorial experiment, with the response (y) shown at the corners



■ **FIGURE 5.3** A factorial experiment without interaction

Definition of a factor effect:
The change in the mean response when the factor is changed from low to high

$$A = \bar{y}_{A^+} - \bar{y}_{A^-} = \frac{40+52}{2} - \frac{20+30}{2} = 21$$

$$B = \bar{y}_{B^+} - \bar{y}_{B^-} = \frac{30+52}{2} - \frac{20+40}{2} = 11$$

$$AB = \frac{52+20}{2} - \frac{30+40}{2} = +1$$

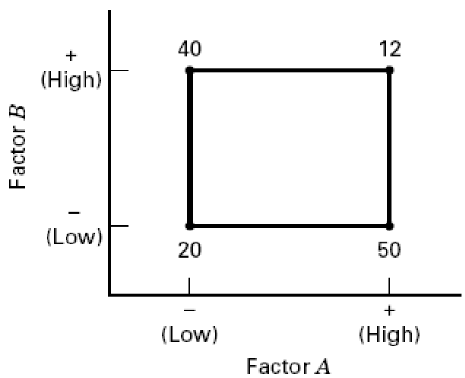
This is a two-factor factorial experiment with both design factors at two levels. We have called these levels “low” and “high” and denoted them “-” and “+”, respectively.

The main effect of factor A (or B) in this two-level design can be thought of as the difference between the average response at the low level of A (or B) and the average response at the high level of A (or B).

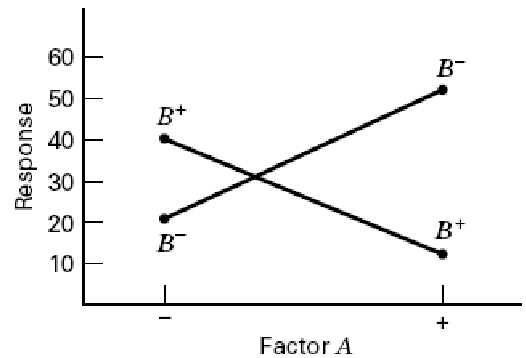
In some experiments, we may find that the difference in response between the levels of one factor is not the same at all levels of the other factors. When this occurs, there is an **interaction** between the factors.

For example, consider the two-factor factorial experiment shown in Figure 5.2.

The Case of Interaction:



■ FIGURE 5.2 A two-factor factorial experiment with interaction



■ FIGURE 5.4 A factorial experiment with interaction

$$A = \bar{y}_{A^+} - \bar{y}_{A^-} = \frac{50+12}{2} - \frac{20+40}{2} = 1$$

$$B = \bar{y}_{B^+} - \bar{y}_{B^-} = \frac{40+12}{2} - \frac{20+50}{2} = -9$$

$$AB = \frac{12+20}{2} - \frac{40+50}{2} = -29$$

At the low level of factor B, the A effect is

$$50 - 20 = 30$$

and at the high level of factor B , the A effect is $12 - 40 = -28$.

Because the effect of A depends on the level chosen for factor B, we see that there is interaction between A and B.

The magnitude of the interaction effect is the average difference in these two A effects, i.e.

$$(-28 - 30)/2 = -29.$$

Clearly, the interaction is large in this experiment. These ideas may be illustrated graphically.

Figure 5.3 plots the response data in Figure 5.1 against factor A for both levels of factor

B. Note that the B^+ and B^- lines are approximately parallel, indicating a lack of interaction between factors A and B.

Similarly, Figure 5.4 plots the response data in Figure 5.2. Here we see that the B^+ and B^- lines are not parallel. This indicates an interaction between factors A and B.

Remarks: Two-factor interaction graphs such as above are frequently very useful in interpreting significant interactions and in reporting results to nonstatistically trained personnel. **However,** they should not be utilized as the sole technique of data analysis because their interpretation is subjective and their appearance is often misleading.

Generally, when an interaction is large, the corresponding main effects have little practical meaning. These points are clearly indicated by the interaction plot in Figure 5.4.

In the presence of **significant interaction**, the experimenter must usually examine the levels of one factor, say A, with levels of the other factors fixed to draw conclusions about the main effect of A.

- Some Advantages of Factorials

Factorial designs have several advantages.

(i) They are more efficient than one-factor-at-a-time experiments.

(ii) Furthermore, a factorial design is necessary when interactions may be present to avoid misleading conclusions.

(iii) Finally, factorial designs allow the effects of a factor to be estimated at several levels of the other factors, yielding conclusions that are valid over a range of experimental conditions.

2. The Two-Factor Factorial Design

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. And, in general, there are n replicates.

Remark: We must have at least two replicates ($n \geq 2$) to determine a sum of squares due to error if the interaction is included in the model.

Question: What is N ?

$$N = abn$$

- An Example

Let's look at the following battery life experiment, which is an example of a factorial design involving two factors.

Example 5.1: An engineer is designing a battery for use in a device that will be subjected to some extreme variations in temperature. The only design parameter that he can select at this point is the plate material for the battery, and he has three possible choices. When the device is manufactured and is shipped to the field, the engineer has no control over the temperature extremes that the device will encounter, and he knows from experience that temperature will probably affect the effective battery life. However, temperature can be controlled in the product development laboratory for the purposes of a test.

The engineer decides to test all three plate materials at three temperature levels ($15, 70$, and 125^{o}F). Because there are two factors at three levels, this design is sometimes called a 3^2 **factorial design.**

Four batteries are tested at each combination of plate material and temperature, and all 36 tests are run in random order.

The experiment and the resulting observed battery life data are given in Table 5.1.

Example 5.1 The Battery Life Experiment

■ TABLE 5.1

Life (in hours) Data for the Battery Design Example

Material Type	Temperature (°F)					
	15		70		125	
1	130	155	34	40	20	70
	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

In this problem, the engineer wants to answer the following questions:

1. What effects do material type and temperature have on the life of the battery ?

2. Is there a choice of material that would give uniformly long life regardless of temperature (a robust product) ?

In order to answer the above questions, let's consider the general two-factor factorial design.

The above design is a specific example of the general case of a two-factor factorial.

For the general case, let Y_{ijk} be the observed response when factor A is at the i^{th} level ($i = 1, 2, \cdots, a$) and factor B is at the j^{th} level ($j = 1, 2, \cdots, b$) for the k^{th} replicate ($k = 1, 2, \cdots, n$).

See the following Table 5.2.

The order in which the abn observations are taken is selected at random so that this design is a **completely randomized design**.

The General Two-Factor Factorial Experiment

■ TABLE 5.2

General Arrangement for a Two-Factor Factorial Design

		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	$y_{211}, y_{212}, \ldots, y_{21n}$	$y_{221}, y_{222}, \ldots, y_{22n}$		$y_{2b1}, y_{2b2}, \ldots, y_{2bn}$
	$\vdots$				
	a	$y_{a11}, y_{a12}, \ldots, y_{a1n}$	$y_{a21}, y_{a22}, \ldots, y_{a2n}$		$y_{ab1}, y_{ab2}, \ldots, y_{abn}$

a levels of factor *A*; *b* levels of factor *B*; *n* replicates

This is a **completely randomized design**

The observations in a factorial experiment can be described by a model. There are several ways to write the model for a factorial experiment, for example, effects model, means model or regression model.

We will mainly consider the effects model.

The effects model for a factorial experiment is

$$Y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk}\,, \quad \begin{cases} i = 1, 2, \cdots, a \\ j = 1, 2, \cdots, b \\ k = 1, 2, \cdots, n \end{cases}$$

In the two-factor factorial, both row and column factors (or treatments), A and B, are of equal interest.

Specifically, we are interested in testing hypotheses about the equality of row treatment effects, say

$$
\begin{aligned}
H_0: &\ \ \tau_1 = \tau_2 = \cdots = \tau_a = 0 \\
H_1: &\ \ \text{at least one}\, \tau_i \neq 0
\end{aligned}
$$

and the equality of column treatment effects, say

$$
\begin{aligned}
H_0: &\ \ \beta_1 = \beta_2 = \cdots = \beta_b = 0 \\
H_1: &\ \ \text{at least one}\, \beta_j \neq 0
\end{aligned}
$$

We are also interested in determining whether row and column treatments interact.

Thus, we also wish to test

$$
\begin{aligned}
H_0: &\ \ (\tau\beta)_{ij} = 0 \text{ for all } i, j \\
H_1: &\ \ \text{at least one}\, (\tau\beta)_{ij} \neq 0
\end{aligned}
$$

We now discuss how these hypotheses are tested using a two-factor ANOVA.

- The Two-Factor Factorial Design with Fixed Effects

Let μ_{ij} denote the mean of the ij^{th} cell. Then we have $\mu = \bar{\mu}_{..}$ and

(i) Main effect of A$_i$ ($i = 1, 2, \cdots a$)

$$\tau_i = \bar{\mu}_{i.} - \bar{\mu}_{..}$$

(ii) Main effect of B$_j$ ($j = 1, 2, \cdots b$)

$$\beta_j = \bar{\mu}_{.j} - \bar{\mu}_{..}$$

(iii) Specific effect of A$_i$ for B$_j$ is

$$\tau_{i(j)} = \mu_{ij} - \bar{\mu}_{.j}$$

(iv) Specific effect of B$_j$ for A$_i$ is

$$\beta_{j(i)} = \mu_{ij} - \bar{\mu}_{i.}$$

If the effect of one factor depends on the level of the others, i.e, specific effects are different from the main effect, factor interact.

The **interaction** $(\tau\beta)_{ij}$ (between A_i and B_j) is the difference in specific effect and the main effect.

Question: What is $(\tau\beta)_{ij}$?

$$(\tau\beta)_{ij} = \mu_{ij} - \bar{\mu}_{i.} - \bar{\mu}_{.j} + \bar{\mu}_{..}$$

Rewrite $Y_{ijk} - \bar{Y}_{...}$

$$Y_{ijk} - \bar{Y}_{...} = (\bar{Y}_{i..} - \bar{Y}_{...}) + (\bar{Y}_{.j.} - \bar{Y}_{...}) + (\bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...}) + (Y_{ijk} - \bar{Y}_{ij.})$$

Let's extend the ANOVA to the two-factor factorials with fixed effects.

Extension of the ANOVA to Factorials (Fixed Effects Case)

$$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk}-\bar{y}_{...})^2 = bn\sum_{i=1}^{a}(\bar{y}_{i..}-\bar{y}_{...})^2 + an\sum_{j=1}^{b}(\bar{y}_{.j.}-\bar{y}_{...})^2$$

$$+n\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{y}_{ij.}-\bar{y}_{i..}-\bar{y}_{.j.}+\bar{y}_{...})^2+\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk}-\bar{y}_{ij.})^2$$

$$SS_T = SS_A + SS_B + SS_{AB} + SS_E$$

df breakdown:

$$abn-1 = a-1+b-1+(a-1)(b-1)+ab(n-1)$$

The interaction degrees of freedom are simply the number of degrees of freedom for cells (which is $ab-1$) minus the number of degrees of freedom for the two main effects A and B; i.e.,

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

Under H_0 and the usual assumption that ε_{ijk} is $NID(0, \sigma^2)$, each sum of squares on the right-hand side of the above decomposition is, upon division by σ^2, an independently distributed chi-square random variable.

Mean squares are as follows.

$$\begin{aligned}
MS_A &= \frac{SS_A}{a-1}, \quad MS_B = \frac{SS_B}{b-1} \\
MS_{AB} &= \frac{SS_{AB}}{(a-1)(b-1)} \\
&\text{and} \\
MS_E &= \frac{SS_E}{ab(n-1)}
\end{aligned}$$

The expected value of the mean squares can be shown to be

$$E(MS_A) = \sigma^2 + \frac{bn\sum_{i=1}^{a}\tau_i^2}{a-1}$$

$$E(MS_B) = \sigma^2 + \frac{an\sum_{j=1}^{b}\beta_j^2}{b-1}$$

$$E(MS_{AB}) = \sigma^2 + \frac{n\sum_{i=1}^{a}\sum_{j=1}^{b}(\tau\beta)_{ij}^2}{(a-1)(b-1)}$$

$$E(MS_E) = \sigma^2$$

Under the null hypothesis, each of the ratios of mean squares MS_A/MS_E, MS_B/MS_E, and MS_{AB}/MS_E is distributed as F with $a-1, b-1$, and $(a-1)(b-1)$ numerator degrees of freedom, respectively, and $ab(n-1)$ denominator degrees of freedom, and the rejection region would be the upper tail of the F distribution.

See the ANOVA table, and manual computing of the sums of squares.

ANOVA Table – Fixed Effects Case

■ **TABLE 5.3**

The Analysis of Variance Table for the Two-Factor Factorial, Fixed Effects Model

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
A treatments	SS_A	$a-1$	$MS_A = \dfrac{SS_A}{a-1}$	$F_0 = \dfrac{MS_A}{MS_E}$
B treatments	SS_B	$b-1$	$MS_B = \dfrac{SS_B}{b-1}$	$F_0 = \dfrac{MS_B}{MS_E}$
Interaction	SS_{AB}	$(a-1)(b-1)$	$MS_{AB} = \dfrac{SS_{AB}}{(a-1)(b-1)}$	$F_0 = \dfrac{MS_{AB}}{MS_E}$
Error	SS_E	$ab(n-1)$	$MS_E = \dfrac{SS_E}{ab(n-1)}$	
Total	SS_T	$abn-1$		

The sums of squares for the main effects are

$$SS_A = \frac{1}{bn}\sum_{i=1}^{a} y_{i..}^2 - \frac{y_{...}^2}{abn} \tag{5.7}$$

and

$$SS_B = \frac{1}{an}\sum_{j=1}^{b} y_{.j.}^2 - \frac{y_{...}^2}{abn} \tag{5.8}$$

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

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

This sum of squares also contains SS_A and SS_B. Therefore, the second step is to compute SS_{AB} as

$$SS_{AB} = SS_{\text{Subtotals}} - SS_A - SS_B \tag{5.9}$$

We may compute SS_E by subtraction as

$$SS_E = SS_T - SS_{AB} - SS_A - SS_B \tag{5.10}$$

or

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

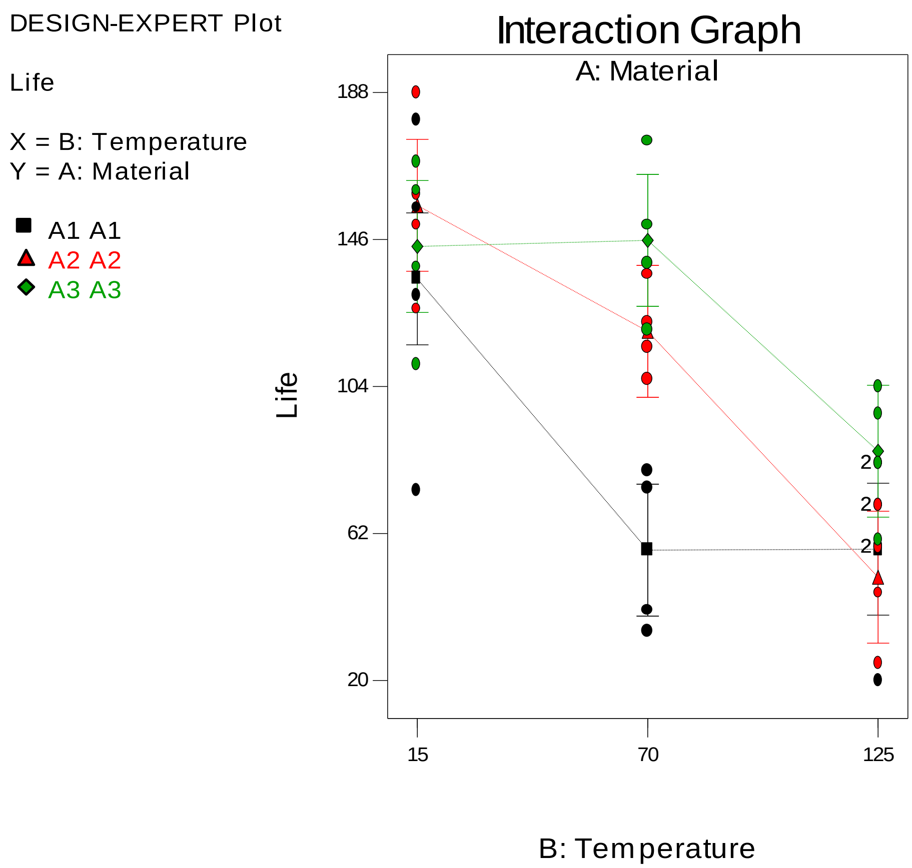
The ANOVA table for battery life experiment

Source of variation	df	SS	Mean Square	f	P_{value}
Material	2	10683.72	5341.86	7.91	0.002
Temperature	2	39118.72	19559.36	28.97	$<$0.001
Interaction	4	9613.78	2403.45	3.56	0.0186
Error	27	18230.75	675.21		
Total	35	77646.97			

Note that interaction is significant at level $\alpha = 0.05$.

Next, using the $\bar{y}_{ij.}$ (i.e. cell mean), we may draw interaction graph to see the interaction.

Interaction Plot



Multiple Comparison:

When the ANOVA indicates that row or column means differ, it is usually of interest to make comparisons between the individual row or column means to discover the specific differences.

Here we illustrate the use of Tukey's CI on the battery life data in Example 5.1.

Note that in this experiment, interaction is significant at level $\alpha = 0.05$.

When interaction is significant, comparisons between the means of one factor (e.g., A) may be obscured by the AB interaction.

One approach to this situation is to fix factor B at a specific level and apply Tukey's CI for the the difference in any two means of factor A at that level.

To illustrate, suppose that in Example 5.1 we are interested in detecting differences among the means of the three material types. Because interaction is significant, we make this comparison at just one level of temperature, say level 2 (70^oF).

We use $MS_E = 675.21$ to estimate σ^2.

The three material type averages at 70^oF are

$$\begin{aligned} \bar{y}_{12.} &= 57.25 \quad \text{(material type 1)} \\ \bar{y}_{22.} &= 119.75 \quad \text{(material type 2)} \\ \bar{y}_{32.} &= 145.75 \quad \text{(material type 3)} \end{aligned}$$

and the $\pm$ part of the Tukey's 95% CI is

$$q_{0.05}(3, 27)\sqrt{\frac{MS_E}{n}} = 3.50\sqrt{\frac{675.21}{4}} = 45.47,$$

where we obtained $q_{0.05}(3,27) \approx 3.50$ by interpolation in Appendix Table V.

Notice that

$$\bar{y}_{32.} - \bar{y}_{12.} = 145.75 - 57.25 = 88.50 > 45.47$$

$$\bar{y}_{32.} - \bar{y}_{22.} = 145.75 - 119.75 = 26.00 < 45.47$$

$$\bar{y}_{22.} - \bar{y}_{12.} = 119.75 - 57.25 = 62.50 > 45.47$$

We have that at the temperature level 70^{o}F, the mean battery life is the same for material types 2 and 3, and that the mean battery life for material type 1 is significantly lower in comparison to both types 2 and 3.

If interaction is significant, the experimenter could compare all ab cell means to determine which ones differ significantly. In this analysis, differences between cell means include interaction effects as well as both main effects.

In Example 5.1, this would give 36 comparisons between all possible pairs of the nine cell means.

Model Adequacy Checking:

As before, the primary diagnostic tool is residual analysis. The residuals for the two-factor factorial model with interaction are

$$e_{ijk} = Y_{ijk} - \widehat{Y}_{ijk} = Y_{ijk} - \bar{Y}_{ij.},$$

where the fitted value $\hat{Y}_{ijk} = \bar{Y}_{ij.}$ (i.e., the average of the observations in the ij^{th} cell).

The normal probability plot of these residuals (the battery life data in Example 5.1) does not reveal anything particularly troublesome.

We may plot the residuals versus material types and temperature, respectively. Both plots indicate mild inequality of variance, with the treatment combination of 15^{o}F and material type one possibly having larger variance than the others.