Lec 5: Factorial Experiment

Ying Li

November 21, 2011

Example

- Study of the battery life vs the factors temperatures and types of material.
- A: Types of material, 3 levels.
- B: Temperatures, 3 levels.

Example

- Study of the battery life vs the factors temperatures and types of material.
- A: Types of material, 3 levels.
- B: Temperatures, 3 levels.

In general, factorial designs are most efficient for the study of the effects of two or more factors.

Factorial Design

In each complete trial or replication of the experiment all possible combinations of the levels of the factors are investigated.

Factorial Design

In each complete trial or replication of the experiment all possible combinations of the levels of the factors are investigated.

Question

What's the advantage of the factorial design comparing with two-single factor experient?

Advantages of factorial design

- The two-factor factorial gives estimates with higher precision (given the same number of experimental units).

Advantages of factorial design

- The two-factor factorial gives estimates with higher precision (given the same number of experimental units).
- The two-factor factorial experiment makes it possible to detect interactions.

Example

■ 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

■ **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	$y_{a11}, y_{a12}, \ldots, y_{a1n}$	$y_{a21}, y_{a22}, \ldots, y_{a2n}$		$y_{ab1}, y_{ab2}, \ldots, y_{abn}$

■ 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	$a-1$	$MS_A = \frac{SS_A}{a-1}$	$F_0 = \frac{MS_A}{MS_E}$
B treatments	SS_B	$b-1$	$MS_B = \frac{SS_B}{b-1}$	$F_0 = \frac{MS_B}{MS_E}$
Interaction	SS_{AB}	$(a-1)(b-1)$	$MS_{AB} = \frac{SS_{AB}}{(a-1)(b-1)}$	$F_0 = \frac{MS_{AB}}{MS_E}$
Error	SS_E	$ab(n-1)$	$MS_E = \frac{SS_E}{ab(n-1)}$	
Total	SS_T	$abn-1$		

Manual Calculation

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

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

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

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

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

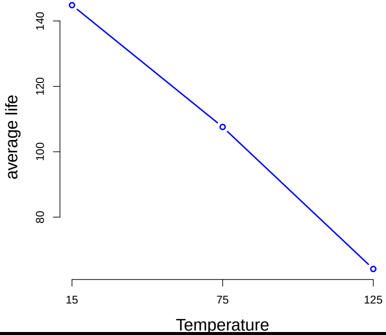
Results for the example

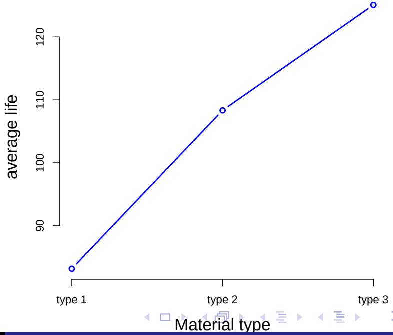
■ **TABLE 5.5**

Analysis of Variance for Battery Life Data

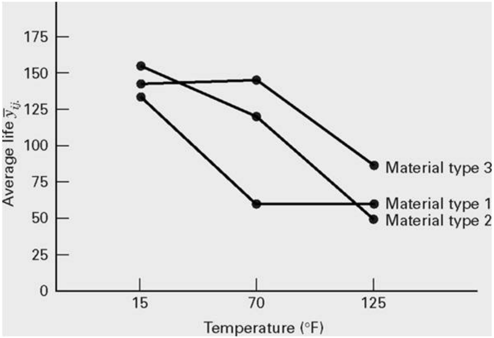
Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0	*P*-Value
Material types	10,683.72	2	5,341.86	7.91	0.0020
Temperature	39,118.72	2	19,559.36	28.97	0.0001
Interaction	9,613.78	4	2,403.44	3.56	0.0186
Error	18,230.75	27	675.21		
Total	77,646.97	35			

	15^o F	70^o F	125^o F	
Type 1	134.75	57.25	57.5	83.16667
Type 2	155.75	119.75	49.5	108.33333
Type 3	144	145.75	85.5	125.08333
	144.83333	107.58333	64.16667	





	15° F	70° F	125° F	
Type 1	134.75	57.25	57.5	83.16667
Type 2	155.75	119.75	49.5	108.33333
Type 3	144	145.75	85.5	125.08333
	144.83333	107.58333	64.16667	



Multiple Comparison

- When interaction is significant, comparisons between the means of one factor (A) may be obscured by the interaction (AB).
- One approach to this situation is to fix factor B at a specific level and apply Tukey's test to the means of factor A at that level.

Multiple Comparison

- When interaction is significant, comparisons between the means of one factor (A) may be obscured by the interaction (AB).
- One approach to this situation is to fix factor B at a specific level and apply Tukey's test to the means of factor A at that level.

Example

$$T_{0.05} = q_{0.05}(a, f)\sqrt{\frac{MS_E}{n}} = q_{0.05}(3, 27)\sqrt{\frac{MS_E}{4}} = 45.71$$

$$\bar{y}_{13.} - \bar{y}_{23.} = 8 \quad \bar{y}_{33.} - \bar{y}_{23.} = 36 \quad \bar{y}_{33.} - \bar{y}_{13.} = 28$$

No interaction in a two-factor model

No interaction in a two-factor model

TABLE 5.5

Analysis of Variance for Battery Life Data

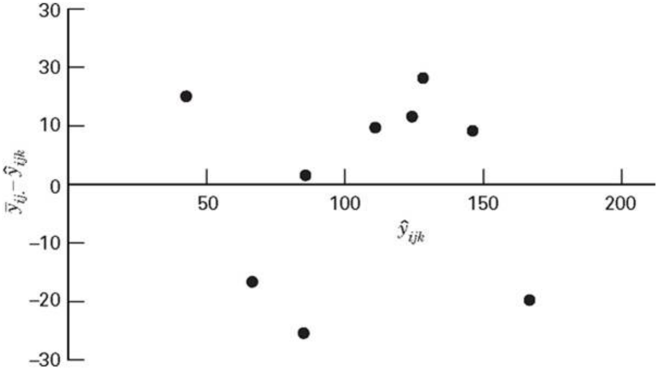
Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0	P-Value
Material types	10,683.72	2	5,341.86	7.91	0.0020
Temperature	39,118.72	2	19,559.36	28.97	0.0001
Interaction	9,613.78	4	2,403.44	3.56	0.0186
Error	18,230.75	27	675.21		
Total	77,646.97	35			

TABLE 5.8

Analysis of Variance for Battery Life Data Assuming No Interaction

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0	P-value
Material types	10,683.72	2	5,341.86	5.95	0.0065
Temperature	39,118.72	2	19,559.36	21.78	0.0001
Error	27,844.52	31	898.21		
Total	77,646.96	35			

Model adequacy checking



One observation per cell case

Model 1

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

$$i = 1, 2, \ldots, a, \quad j = 1, 2, \ldots, b$$

One observation per cell case

Model 1

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

$$i = 1, 2, \ldots, a, \quad j = 1, 2, \ldots, b$$

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	$a-1$	$MS_A = \frac{SS_A}{a-1}$	$F_0 = \frac{MS_A}{MS_E}$
B treatments	SS_B	$b-1$	$MS_B = \frac{SS_B}{b-1}$	$F_0 = \frac{MS_B}{MS_E}$
Interaction	SS_{AB}	$(a-1)(b-1)$	$MS_{AB} = \frac{SS_{AB}}{(a-1)(b-1)}$	$F_0 = \frac{MS_{AB}}{MS_E}$
Error	SS_E	$ab(n-1)$	$MS_E = \frac{SS_E}{ab(n-1)}$	
Total	SS_T	$abn-1$		

One observation per cell case

Model 1

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

$$i = 1, 2, \ldots, a, \quad j = 1, 2, \ldots, b$$

One observation per cell case

Model 1

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

$$i = 1, 2, \ldots, a, \quad j = 1, 2, \ldots, b$$

Model 2

$$y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij}$$

$$i = 1, 2, \ldots, a, \quad j = 1, 2, \ldots, b$$

TABLE 5.9

Analysis of Variance for a Two-Factor Model, One Observation per Cell

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Expected Mean Square
Rows (A)	$\sum_{i=1}^{a} \frac{y_{i.}^2}{b} - \frac{y_{..}^2}{ab}$	$a - 1$	MS_A	$\sigma^2 + \frac{b\sum \tau_i^2}{a-1}$
Columns (B)	$\sum_{j=1}^{b} \frac{y_{.j}^2}{a} - \frac{y_{..}^2}{ab}$	$b - 1$	MS_B	$\sigma^2 + \frac{a\sum \beta_j^2}{b-1}$
Residual or AB	Subtraction	$(a-1)(b-1)$	MS_{Residual}	$\sigma^2 + \frac{\sum\sum (\tau\beta)_{ij}^2}{(a-1)(b-1)}$
Total	$\sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij}^2 - \frac{y_{..}^2}{ab}$	$ab - 1$		

Tukey single-degree-of-freedom test

It is helpful in determining whether interaction is present or not. The procedure assumes $(\tau\beta)_{ij} = \gamma\tau_i\beta_j$. Then the test partitions the residual sum of squares into two part. One is

$$SS_N = \frac{[\sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij}y_{i.}y_{.j} - y_{..}(SS_A + SS_B + \frac{y_{..}^2}{ab})]^2}{abSS_ASS_B},$$

with 1 degree of freedom, and

$$SS_{Error} = SS_{residual} - SS_N,$$

with $(a-1)(b-1)-1$ degrees of freedom. To test the present of interaction, we compute

$$F_0 = \frac{SS_N}{SS_{Error}/[(a-1)(b-1)-1]}$$

If $F_0 > F_{\alpha,1,(a-1)(b-1)-1}$, the hypothesis of no interaction must be rejected.

Question

Is the two-factor factorial model with one observation per cell identical to the randomized complete block model?

Blocking in a Factorial Design

Example

- Study of the battery life vs the factors temperatures and types of material.
- A: Types of material, 3 levels.
- B: Temperatures, 3 levels.
- 9 combinations of A and B.
- 4 observations per combination.
- In total 36 observations.

Example

- Study of the battery life vs the factors temperatures and types of material.
- A: Types of material, 3 levels.
- B: Temperatures, 3 levels.
- 9 combinations of A and B.
- 4 observations per combination.
- In total 36 observations.
- 4 blocks.

■ **TABLE 5.20**

Analysis of Variance for a Two-Factor Factorial in a Randomized Complete Block

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	$\sum_i\sum_j\sum_k y_{ijk}^2 - \frac{y_{...}^2}{abn}$	$abn-1$		

Source	DF	Anova SS	MeanSquare	F Value	Pr > F
temperature	2	39118.722	19559.361	26.26	<.0001
material	2	10683.722	5341.861	7.17	0.0036
Interaction	4	9613.778	2403.444	3.23	0.0297
block	3	354.972	118.324		
Error	24	17875.778	744.824		
Total	35	77646.972			

Three factorial experiment

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

$$i = 1, 2, \cdot, a$$

$$j = 1, 2, \cdot, b$$

$$k = 1, 2, \cdot, c$$

$$l = 1, 2, \cdot, n$$

TABLE 5.12

The Analysis of Variance Table for the Three-Factor Fixed Effects Model

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Expected Mean Square	F_0
A	SS_A	$a-1$	MS_A	$\sigma^2 + \frac{bcn\sum \tau_i^2}{a-1}$	$F_0 = \frac{MS_A}{MS_E}$
B	SS_B	$b-1$	MS_B	$\sigma^2 + \frac{acn\sum \beta_j^2}{b-1}$	$F_0 = \frac{MS_B}{MS_E}$
C	SS_C	$c-1$	MS_C	$\sigma^2 + \frac{abn\sum \gamma_k^2}{c-1}$	$F_0 = \frac{MS_C}{MS_E}$
AB	SS_{AB}	$(a-1)(b-1)$	MS_{AB}	$\sigma^2 + \frac{cn\sum\sum (\tau\beta)_{ij}^2}{(a-1)(b-1)}$	$F_0 = \frac{MS_{AB}}{MS_E}$
AC	SS_{AC}	$(a-1)(c-1)$	MS_{AC}	$\sigma^2 + \frac{bn\sum\sum (\tau\gamma)_{ik}^2}{(a-1)(c-1)}$	$F_0 = \frac{MS_{AC}}{MS_E}$
BC	SS_{BC}	$(b-1)(c-1)$	MS_{BC}	$\sigma^2 + \frac{an\sum\sum (\beta\gamma)_{jk}^2}{(b-1)(c-1)}$	$F_0 = \frac{MS_{BC}}{MS_E}$
ABC	SS_{ABC}	$(a-1)(b-1)(c-1)$	MS_{ABC}	$\sigma^2 + \frac{n\sum\sum\sum (\tau\beta\gamma)_{ijk}^2}{(a-1)(b-1)(c-1)}$	$F_0 = \frac{MS_{ABC}}{MS_E}$
Error	SS_E	$abc(n-1)$	MS_E	σ^2	
Total	SS_T	$abcn-1$			

Example

- Soft drink bottling
- Fill correct volume
- Response variable: difference from correct volume

- Carbonation (10%, 12%, 14%)
- Pressure (25psi, 30psi)
- Line speed (200bpm, 250bpm)

■ **TABLE 5.13**

Fill Height Deviation Data for Example 5.3

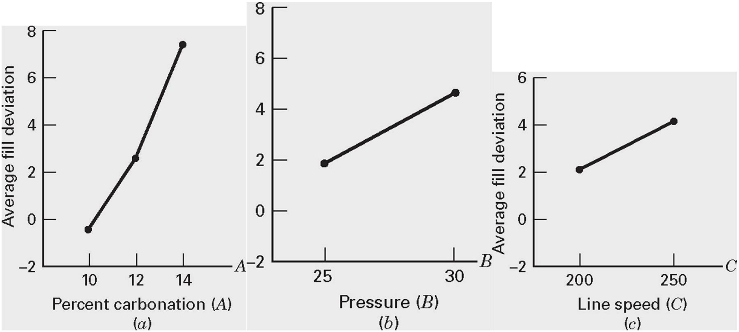
Percent Carbonation (*A*)	**Operating Pressure (*B*): 25 psi, Line Speed (*C*): 200**		**25 psi, Line Speed (*C*): 250**		**30 psi, Line Speed (*C*): 200**		**30 psi, Line Speed (*C*): 250**		$y_{i\ldots}$
	−3		−1		−1		1		
10	−1	(−4)	0	(−1)	0	(−1)	1	(2)	−4
	0		2		2		6		
12	1	(1)	1	(3)	3	(5)	5	(11)	20
	5		7		7		10		
14	4	(9)	6	(13)	9	(16)	11	(21)	59
$B \times C$ Totals $y_{.jk.}$	6		15		20		34		$75 = y_{\ldots}$
$y_{.j..}$	21				54				

$A \times B$ Totals $y_{ij..}$

A	*B* = 25	*B* = 30
10	−5	1
12	4	16
14	22	37

$A \times C$ Totals $y_{i.k.}$

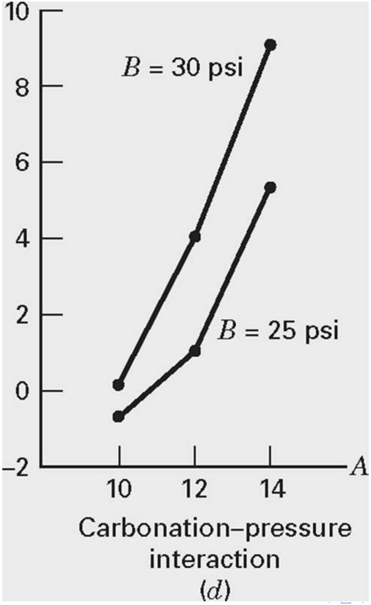
A	*C* = 200	*C* = 250
10	−5	1
12	6	14
14	25	34


8
6
4
2
0
−2
Average fill deviation
10
12
14
A
Percent carbonation (A)
(a)
8
6
4
2
0
−2
25
30
B
Pressure (B)
(b)
6
4
2
0
−2
Average fill deviation
200
250
C
Line speed (C)
(c)

■ TABLE 5.14

Analysis of Variance for Example 5.3

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0	*P*-Value
Percentage of carbonation (*A*)	252.750	2	126.375	178.412	<0.0001
Operating pressure (*B*)	45.375	1	45.375	64.059	<0.0001
Line speed (*C*)	22.042	1	22.042	31.118	0.0001
AB	5.250	2	2.625	3.706	0.0558
AC	0.583	2	0.292	0.412	0.6713
BC	1.042	1	1.042	1.471	0.2485
ABC	1.083	2	0.542	0.765	0.4867
Error	8.500	12	0.708		
Total	336.625	23			



10
8
6
4
2
0
−2
B = 30 psi
B = 25 psi
10
12
14
A
Carbonation–pressure interaction
(d)