Factorial Design: Significance of Main and Interaction Effects

Note that for calculating the confidence intervals, we assumed that the true variance σ^2 is the same for all observations and that the observations are independent.

- How can we check if this assumption is valid? **Bartlett Test**

$H_0 : \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_m^2 = \sigma_y^2$

H_1 : at least one $\sigma_i^2 \neq \sigma_j^2, i \neq j$

$$\boxed{\chi^2_{v=m-1} = \frac{M}{C}}$$

where

$$M = (N-m)\ln(s_p^2) - \sum_{i=1}^{m}(n_i - 1)\ln(s_i^2)$$

$$C = 1 + \frac{1}{\underset{6}{3(m-1)}}\left[\left(\sum_{i=1}^{m}\frac{1}{n_i - 1}\right) - \frac{1}{N-m}\right]$$

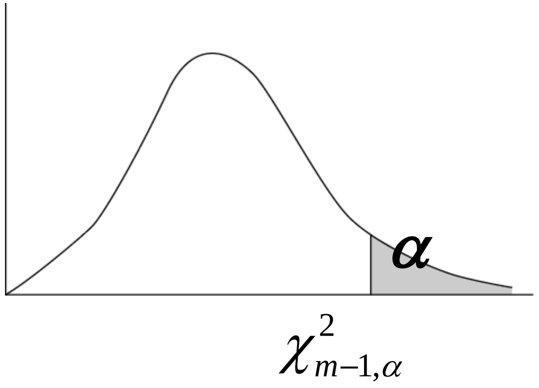
$m = 2^k$ = 8 *(Total experiments)*

Sample size = n $N = \overset{2}{n_1} + \overset{2}{n_2} + \ldots + \overset{2}{n_m}$ = 16

The value of M will be large if the sample variances s_i^2 differ greatly in magnitude, and will be zero if all the sample variances are exactly equal.

We will reject H_0 if χ^2_{cal} is too large, i.e.,

$$\chi^2_{cal} > \underline{\chi^2_{m-1,\alpha}}$$



Example: Bartlett Test

$$\chi^2_{v=m-1} = \frac{M}{C} \quad \text{where, } M = (N-m)\ln(s_p^2) - \sum_{i=1}^{m}(n_i-1)\ln(s_i^2), \quad \text{and} \quad C = 1 + \frac{1}{3(m-1)}\left[\left(\sum_{i=1}^{m}\frac{1}{n_i-1}\right) - \frac{1}{N-m}\right]$$

Here, N = 16 m = $2^3 = 8$

here $n_i = 2$ for all i

$S_1^2 = 24.5$, $S_2^2=21.78$, $S_3^2=134.48$, $S_4^2=242.0$,

$S_5^2=3.92$, $S_6^2=8.82$, $S_7^2=33.62$, $S_8^2=72.00$

The value of M will be large if the sample variances s_i^2 differ greatly in magnitude, and will be zero if all the sample variances are exactly equal.

We will reject H_0 if χ^2_{cal} is too large, i.e., $\chi^2_{\text{cal}} > \chi^2_{m-1,\alpha}$

$$S_p^{\,2} = \frac{[(y_{a1}-\bar{y}_1)^2+(y_{b1}-\bar{y}_1)^2+\ldots\ldots+(y_{a8}-\bar{y}_8)^2+(y_{b8}-\bar{y}_8)^2]}{(2-1)+\ldots\ldots+(2-1)} = 67.64$$

$$M = (16-8)\ln 67.64 - [(2-1)\ln 24.5 + (2-1)\ln 21.78 + (2-1)\ln 134.48 + (2-1)\ln 242 + (2-1)\ln 3.92 + (2-1)\ln 8.82 + (2-1)33.62 + (2-1)\ln 72.0] = 5.713$$

$$C = 1 + \frac{1}{3(8-1)}\left[\left(\sum_{i=1}^{8}\frac{1}{2-1}\right) - \frac{1}{16-8}\right] = 1 + \frac{1}{21}[8-5] = 1.357$$

$$\chi^2_{\text{cal}} = \frac{5.713}{1.357} = 4.21 \qquad \chi^2_{7,\alpha=0.05} = 14.1$$

Ho is valid

Test #	X1	X2	X3	Y_{ai} (kpsi)	Y_{bi} (kpsi)	Average (kpsi)
1	-1	-1	-1	84	91	87.5
2	1	-1	-1	90.6	84	87.3
3	-1	1	-1	69.6	86	77.8
4	1	1	-1	76	98	87
5	-1	-1	1	77.7	80.5	79.1
6	1	-1	1	99.7	95.5	97.6
7	-1	1	1	82.7	74.5	78.6
8	1	1	1	93.7	81.7	87.7

Factorial Design: Examples

The yield form, a certain chemical depends on

either the **chemical formulation of the input materials** or **the mixer speed, or both**.

A 2-level factorial design was run with three replicates and the yield data are shown below.

x_1	x_2	y_a	y_b	y_c	$\bar{y}$
-1	-1	10	20	30	20
1	-1	40	30	50	40
-1	1	60	30	60	50
1	1	40	45	50	45

main and interaction effects

confidence intervals

significance using ANOVA

Factorial Design: Examples

x_1	x_2	y_a	y_b	y_c	$\bar{y}$
-1	-1	10	20	30	20
1	-1	40	30	50	40
-1	1	60	30	60	50
1	1	40	45	50	45

$E_1 = \frac{1}{2}\left(-20+40+(-50)+45\right)$

$E_1 = 7.5$

$E_2 = \frac{1}{2}\left(-20-40+50+45\right) = 17.5$

$E_{12} = \frac{1}{2}\left(20-40-50+45\right) = -12.5$

What are sample variances?

How to find confidence intervals for E_1, E_2, E_{12}?

$S_1^2 = \sum \frac{(y_i-\bar{y})^2}{n-1} = \frac{(10-20)^2+(20-20)^2+(30-20)^2}{2} = 100$

$S_2^2 = \frac{(40-40)^2+(30-40)^2+(50-40)^2}{2} = 100$

$S_3^2 = 300$

$S_4^2 = 25$

$V(E_1) = ?$

Pooled variance

$S_p^2 = \frac{\sum DOF \times S_i^2}{TDOF} = \frac{2(100)+2(100)+2(300)+2(25)}{2+2+2+2}$

$= 131.25$

x_1	x_2	y_a	y_b	y_c	$\bar{y}$
-1	-1	10	20	30	20
1	-1	40	30	50	40
-1	1	60	30	60	50
1	1	40	45	50	45

y_1, y_2, y_3, y_4

Confidence interval for E_1 $V(cy) = c^2 V(y)$

$$V(E_1) = V\left(\frac{\bar{y}_2 - \bar{y}_1 + \bar{y}_4 - \bar{y}_3}{2}\right)$$

$$= \frac{1}{4} V\left(\frac{y_{2a}+y_{2b}+y_{2c}}{3} - \frac{y_{1a}+y_{1b}+y_{1c}}{3} + \frac{y_{4a}+y_{4b}+y_{4c}}{3} - \frac{y_{3a}+y_{3b}+y_{3c}}{3}\right)$$

$$= \frac{1}{36}\, 12\, V(y_{1a}) = \frac{12}{36}\sigma^2 = \frac{1}{3}\sigma^2$$

$V(E_1) = \sigma^2/3$

$V(E_2) = \sigma^2/3 = V(E_{12})$

Confidence interval $E_i \pm t_{\nu,\alpha}\sqrt{\frac{S_p^2}{3}}$

at 95% confidence $t_{\nu,\alpha} = t_{8,0.025} = 2.306$ from table

$E_i \pm 2.306\sqrt{131.25/3} = E_i \pm 15.25$

Factorial Design: Examples

x_1	x_2	y_a	y_b	y_c	$\bar{y}$
-1	-1	10	20	30	20
1	-1	40	30	50	40
-1	1	60	30	60	50
1	1	40	45	50	45

How to do 2-factor ANOVA?

effects model?

$\bar{y} = \mu + \text{Effect } X_1 + \text{Effect } X_2 + \text{Interaction} + \varepsilon$

DOF

12 SS_{Total} ✓ $10^2 + 20^2 + 30^2 + \dots$

1 SS_{mean} ✓ $= 12 \times (38.75)^2$

1 $SS_{x1} = 2\times3\left[(35 - 38.75)^2 + (42.5 - 38.75)^2\right]$

1 $SS_{x2} = 2\times3\left((30 - 38.75)^2 + (47.5 - 38.75)^2\right]$

1 SS_{x1x2} (subtraction) ✓

8 $SS_{error} = (10-20)^2 + (20-20)^2 + (30-20)^2 + (60-50)^2 + (30-50)^2 + \dots$

12 terms

$x_1 \downarrow$ / $x_2 \rightarrow$	x_2 low	x_2 high	
x_1 low	10, 20, 30 ($\boxed{20}$)	60, 30, 60 ($\boxed{50}$)	35
x_1 high	40, 30, 50 ($\boxed{40}$)	40, 45, 50 ($\boxed{45}$)	42.5

What is one observation is missing?

x_1	x_2	y_a	y_b	y_c	$\bar{y}$
-1	-1	10	20	30	20
1	-1	40	X	50	45
-1	1	60	30	60	50
1	1	40	45	50	45

$$E_1 = 1/2\,(-20+45-50+45) = 10$$

$$E_2 = 1/2\,(-20-45+50+45) = 15$$

$$E_{12} = 1/2\,(20-45-50+45) = -15$$

Sample Variances?

$$s_1^2 = \left(\frac{(10-20)^2+(20-20)^2+(30-20)^2}{2}\right)$$

$$s_2^2 = \frac{(40-45)^2+(50-45)^2}{1}$$

$$s_3^2 = \quad , \; s_4^2$$

$$S_p^2 = \frac{2s_1^2 + 1s_2^2 + 2s_3^2 + 2s_4^2}{2+1+2+2} = \frac{900}{7}$$

Confidence interval on E_1

$$V(E_1) = V\left(1/2\,(\bar{y}_2-\bar{y}_1+\bar{y}_4-\bar{y}_3)\right)$$

$$= 1/4\left(\frac{y_{2a}+y_{2b}}{2} - \frac{y_{1a}+y_{1b}+y_{1c}}{3} + \ldots\right)$$

$$= 1/4\left[1/4\;2\sigma^2 + 1/9\;9\sigma^2\right] = 3/8\,\sigma^2$$

Factorial Design: Missing Data

What is one observation is missing?

x_1	x_2	y_a	y_b	y_c	$\bar{y}$
-1	-1	10	20	30	
1	-1	40	X	50	
-1	1	60	30	60	
1	1	40	45	50	

$V(E_1) = \frac{3}{8}\sigma^2$

$E_1 \pm t_{\nu,\alpha}\sqrt{S_p^2 \, 3/8}$

7

$S_p^2 = 900/7$

$E_1 \pm \underline{16.42}$ for 95% confidence

$\alpha = 0.025$

Factorial Design: Examples

Consider an investigation into the **effect of the concentration of the reactant and the amount of the catalyst on the conversion (yield)** in a chemical process.

The objective of the experiment was to determine if adjustments to either of these two factors would increase the yield.

Let the reactant **concentration be factor A** and let the **two levels of interest be 15 and 25 percent**. The **catalyst is factor B**, with the high level denoting the use of **2 pounds** of the catalyst and the low level denoting the use of only **1 pound.**

The experiment is **replicated three times**, so there are 12 runs. The order in which the runs are made is random, so this is a completely randomized experiment. The data obtained are as follows:

Factor *A*	Factor *B*	Treatment Combination	Replicate I	Replicate II	Replicate III	Total
−	−	*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

Factorial Design: Examples

Factor		Treatment Combination	Replicate			Total
A	*B*		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

Factorial Design: Examples

Factor		Treatment	Replicate			
A	B	Combination	I	II	III	Total
−	−	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

Fractional Factorial Designs (FFD)

- In Full Factorial designs, a large amount of resources are expended in estimating interaction terms. That is, the ratio of the number of main effects to the total number of effects reduces rapidly as the number of variables ‘k’ increases.

$y_i = \mu + \underline{\tau_1 + \tau_2 + \tau_3 + \tau_4 + \tau_5 + \tau_6} + (\tau_1\tau_2)$

- For example, in a full 2^6 experiment, only 9.5% (6/64) of the effects calculated are main/average effects.
- The remaining 90.5% (58/64) effects relate to interaction effects.

- The negligible magnitude of the many higher-order interactions, together with the fact that only a few variables will have significant influence on the response, a tremendous amount of redundancy exists in two-level factorial designs.
- To reduce the problem of estimating large numbers of possibly unimportant interaction effects, **fractional factorial designs** are created by replacing some of the higher-order interaction terms by additional experimental factors.

2^6

1	Mean
6	Main effects
15 ${}^6C_2$	Two-factor interaction
20 ${}^6C_3$	Three-factor interaction
15 ${}^6C_4$	Fourth order
6 ${}^6C_5$	Fifth order
1 ${}^6C_6$	Sixth order
64	Total Tests

Creating a Fractional Factorial Design (FFD)

Example: Suppose that you want to study four factors, X1, X2, X3 and X4. The total number of tests in a full factorial design using 2 levels of each variable will be $2^4 = 16$

- BUT you are resource constrained and suspect that many higher order interaction terms may be insignificant
- You want to use only 8 tests instead of 16 -> **Create a fractional factorial design**
- To do this, first write down a full factorial design with 3 variables $2^3 = 8$

Base design

note exactly 4 (+) 4 (-)

no 2 columns same

Test	X_1	X_2	X_3	X_1X_2	X_1X_3	X_2X_3	$X_1X_2X_3$
1	-1	-1	-1	+1	+1	+1	-1
2	+1	-1	-1	-1	-1	+1	+1
3	-1	+1	-1	-1	+1	-1	+1
4	+1	+1	-1	+1	-1	-1	-1
5	-1	-1	+1	+1	-1	-1	+1
6	+1	-1	+1	-1	+1	-1	-1
7	-1	+1	+1	-1	-1	+1	-1
8	+1	+1	+1	+1	+1	+1	+1

Creating a Fractional Factorial Design (FFD)

- Next, since the highest-order interaction is least likely to be important, replace the X1X2X3 column by the letter X4. This is abbreviated by writing X4 = X1X2X3

X_1	X_2	X_3	X_1X_2	X_1X_3	X_2X_3	$X_1X_2X_3$
-1	-1	-1	+1	+1	+1	-1
+1	-1	-1	-1	-1	+1	+1
-1	+1	-1	-1	+1	-1	+1
+1	+1	-1	+1	-1	-1	-1
-1	-1	+1	+1	-1	-1	+1
+1	-1	+1	-1	+1	-1	-1
-1	+1	+1	-1	-1	+1	-1
+1	+1	+1	+1	+1	+1	+1

X_1	X_2	X_3	$X4 = X_1X_2X_3$
-1	-1	-1	-1
+1	-1	-1	+1
-1	+1	-1	+1
+1	+1	-1	-1
-1	-1	+1	+1
+1	-1	+1	-1
-1	+1	+1	-1
+1	+1	+1	+1

Creating a Fractional Factorial Design (FFD)

- Note that the fractional factorial design uses 8 rows from total 16 rows of full factorial design

$2^4 = 16$

X_1	X_2	X_3	X4
-1	-1	-1	-1
+1	-1	-1	+1
-1	+1	-1	+1
+1	+1	-1	-1
-1	-1	+1	+1
+1	-1	+1	-1
-1	+1	+1	-1
+1	+1	+1	+1

4 factor 8

Run	X_1	X_2	X_3	X_4	
1	-1	-1	-1	-1	1
2	+1	-1	-1	-1	
3	-1	+1	-1	-1	
4	+1	+1	-1	-1	4
5	-1	-1	+1	-1	
6	+1	-1	+1	-1	6
7	-1	+1	+1	-1	7
8	+1	+1	+1	-1	
9	-1	-1	-1	+1	
10	+1	-1	-1	+1	2
11	-1	+1	-1	+1	3
12	+1	+1	-1	+1	
13	-1	-1	+1	+1	5
14	+1	-1	+1	+1	
15	-1	+1	+1	+1	
16	+1	+1	+1	+1	8

16

Creating a Fractional Factorial Design (FFD)

- Because the 8 test runs comprise only a fraction of the 16 runs required in a full 2^4 design, we say that the 8-run experiment is a fractional factorial experiments.
- Furthermore, since this design uses only half of the 16 runs, we say that it is a half fraction of the full factorial design based on four factors.
- Once the tests are conducted in accordance with the test recipes defined by the design matrix, the calculation matrix is determined to provide for the estimation of the interaction effects.
- Expanding the design matrix, we obtain the calculation matrix by forming all possible products of columns 1 through 4.

Creating a Fractional Factorial Design (FFD)

- We obtain the calculation matrix by forming all possible products of columns 1 through 4.

main 6 4 1

X_1	X_2	X_3	X4
-1	-1	-1	-1
+1	-1	-1	+1
-1	+1	-1	+1
+1	+1	-1	-1
-1	-1	+1	+1
+1	-1	+1	-1
-1	+1	+1	-1
+1	+1	+1	+1

FFD

Test	*I*	1	2	3	4	12	13	14	23	24	34	123	124	134	234	1234
1	+	−	−	−	−	+	+	+	+	+	+	−	−	−	−	+
2	+	+	−	−	+	−	−	+	+	−	−	+	−	−	+	+
3	+	−	+	−	+	−	−	−	−	+	−	+	−	+		+
4	+	+	+	−	−	+	−	−	−	−	+	−	−	+	+	+
5	+	−	−	+	+	+	−	−	−	−	+	+	+	−	−	+
6	+	+	−	+	−	−	+	−	−	+	−	−	+	−	+	+
7	+	−	+	+	−	−	−	+	+	−	−	−	+	+	−	+
8	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

Consequences of fractioning a full factorial design

- Examination of the calculation matrix reveals that many of the columns are identical.
- In particular, of the 16 columns, only eight are unique; each unique column appears twice.

The following pairs of variable effects are represented in the calculation matrix by the same column of plus and minus signs:

1 and 234

2 and 134

3 and 124

4 and 123

12 and 34

13 and 24

23 and 14

avg. (I) and 1234

Test	*I*	**1**	**2**	**3**	**4**	12	13	14	23	24	34	123	124	134	234	1234
1	+	−	−	−	−	+	+	+	+	+	+	−	−	−	−	+
2	+	+	−	−	+	−	−	+	+	−	−	+	−	−	+	+
3	+	−	+	−	+	−	−	−	−	+	−	+	−	+		+
4	+	+	+	−	−	+	−	−	−	−	+	−	−	+	+	+
5	+	−	−	+	+	+	−	−	−	−	+	+	+	−	−	+
6	+	+	−	+	−	−	+	−	−	+	−	−	+	−	+	+
7	+	−	+	+	−	−	−	+	+	−	−	−	+	+	−	+
8	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

Consequences of fractioning a full factorial design

- Examination of the calculation matrix reveals that many of the columns are identical. **What does that mean?**
- When you multiply, for example, the 12 column by the data -> sum -> divided by 4; do you get an estimate of the two-factor interaction 12? Or the two-factor interaction 34? Or both?
- **The interactions 12 and 34 are said to be confounded or confused.**
- The interaction 12 and 23 are said to be aliases of the unique column of plus and minus signs defined by (+−−++−−+).
- Use of this column for effect estimation produces a number (estimate) that is actually the sum of the two-factor interaction effects 12 and 34.
- Similarly, 1 and 234 are confounded effects, 2 and 134 are confounded effects, and so on…

Test	*I*	**1**	**2**	**3**	**4**	12	13	14	23	24	34	123	124	134	234	1234
1	+	−	−	−	−	+	+	+	+	+	+	−	−	−	−	+
2	+	+	−	−	+	−	−	+	+	−	−	+	−	−	+	+
3	+	−	+	−	+	−	−	−	−	+	−	+	−	+	−	+
4	+	+	+	−	−	+	−	−	−	−	+	−	−	+	+	+
5	+	−	−	+	+	+	−	−	−	−	+	+	+	−	−	+
6	+	+	−	+	−	−	+	−	−	+	−	−	+	−	+	+
7	+	−	+	+	−	−	−	+	+	−	−	−	+	+	−	+
8	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

Consequences of fractioning a full factorial design

- It seems that the innocent act of using the 123 column to introduce a fourth variable into a 2^3 full factorial scheme has created a lot of confounding among the variable effects.
- The eight unique columns in the calculation matrix are used to obtain the linear combinations l_0, l_1, …, l_{123} of confounded effects when their signs are applied to the data, and the result is summed and then divided by 4 (divide by 8 for l_0).

E0 l_0 estimates mean + (1/2)(1234) ✓

E1 l_1 estimates 1 + 234

E2 l_2 estimates 2 + 134

l_3 estimates 3 + 124

E12 l_{12} estimates 12 + 34

E13 l_{13} estimates 13 + 24

l_{23} estimates 23 + 14

l_{123} estimates 4 + 123

Test	*I*	**1**	**2**	**3**	**4**	12	13	14	23	24	34	123	124	134	234	1234
1	+	−	−	−	−	+	+	+	+	+	+	−	−	−	−	+
2	+	+	−	−	+	−	−	+	+	−	−	+	−	−	+	+
3	+	−	+	−	+	−	−	−	−	+	−	+	−	+		+
4	+	+	+	−	−	+	−	−	−	−	+	−	−	+	+	+
5	+	−	−	+	+	+	−	−	−	−	+	+	+	−	−	+
6	+	+	−	+	−	−	+	−	−	+	−	−	+	−	+	+
7	+	−	+	+	−	−	−	+	+	−	−	−	+	+	−	+
8	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

Consequences of fractioning a full factorial design

- Some of this confounding can be eliminated by invoking the assumption that third- and higher-order effects are negligible, leading to clear estimates of all main effects.
- But the six two-factor interactions are still hopelessly confounded.

l_0 estimates mean + (1/2)(1234) ≈mean

l_1 estimates 1 + 234 ≈ 1

l_2 estimates 2 + 134 ≈ 2

l_3 estimates 3 + 124 ≈ 3

l_{12} estimates 12 + 34

l_{13} estimates 13 + 24

l_{23} estimates 23 + 14

l_{123} estimates 4 + 123 ≈ 4

Test	*I*	**1**	**2**	**3**	**4**	12	13	14	23	24	34	123	124	134	234	1234
1	+	−	−	−	−	+	+	+	+	+	+	−	−	−	−	+
2	+	+	−	−	+	−	−	+	+	−	−	+	−	−	+	+
3	+	−	+	−	+	−	−	−	−	+	−	+	−	+		+
4	+	+	+	−	−	+	−	−	−	−	+	−	−	+	+	+
5	+	−	−	+	+	+	−	−	−	−	+	+	+	−	−	+
6	+	+	−	+	−	−	+	−	−	+	−	−	+	−	+	+
7	+	−	+	+	−	−	−	+	+	−	−	−	+	+	−	+
8	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

Fractioning a full factorial design (2^{k-p})

- The four-variable, two-level, eight-test experiment discussed thus far is referred to as a two-level fractional factorial design since it considers only a fraction of the tests defined by the full factorial.
- **In this case we have created a one-half fraction design.** It is commonly referred to as a 2^{4-1} fractional factorial design. **It is a member of the general class of 2^{k-p} Fractional Factorial Designs.**

For 2^{k-p} designs

- **k variables** are examined in 2^{k-p} tests
- Requiring that the **'p'** of the variables be introduced into the full factorial in k–p variables
- By assigning them to interaction effects in the first k–p variables.

Example:

In **2^{4-1} FFD**, **k = 4 variables** are studied in **2^{4-1} = 8 tests**

p = 1 of the variables is introduced into a 2^3 full factorial

By assigning it to the interaction 123 (i.e., let 4 = 123)

2^4

4

4-2

2^2

2^4

2^{4-1}

2^{k-p}

$2^4 = 16$

$2^{4-2} = 4$

Fractional Factorial Design (2^{k-p})

- Many other useful fractional factorials can be developed, some dealing with rather large numbers of variables in relatively few tests.
- The 2^{4-1} fractional factorial design just examined is one of the more simple fractional factorial designs. **It can get much worse. Therefore, we need a system to set up such designs easily and to determine quickly the precise nature/pattern of the confounding of the variable effects.**

Example: Suppose that an investigator wishes to study the potential effects that five variables may have on the output of a certain process using some type of two-level factorial experiment.

- If all possible combinations of five variables at two levels each are to be considered, then 2^5 = 32 tests must be conducted.

$2^{5-2} = 2^3 = 8$

- His boss informs him that due to time and budget limitations he will only be able to run 8 tests, not 32.
- How might the investigator reconsider his original test plan and gain some useful information about the five variables?
- **If only eight tests are to be considered using a two-level scheme, only three variables can be examined in a full two-level factorial test plan.**

Fractional Factorial Design (2^{k-p})

If only eight tests are to be considered using a two-level scheme, only three variables can be examined in a full two-level factorial test plan.

Test	Average I	1	2	3	12	13	23	123	y
1	+	−	−	−	+	+	+	−	y_1
2	+	+	−	−	−	−	+	+	y_2
3	+	−	+	−	−	+	−	+	y_3
4	+	+	+	−	+	−	−	−	y_4
5	+	−	−	+	+	−	−	+	y_5
6	+	+	−	+	−	+	−	−	y_6
7	+	−	+	+	−	−	+	−	y_7
8	+	+	+	+	+	+	+	+	y_8

Fractional Factorial Design (2^{k-p})

BUT we want to study FIVE factors with only eight tests -> Fractional Factorial (8 out of 32)

- As discussed earlier, in designing fractional factorial experiments, we introduce additional variables into the base design by borrowing columns initially assigned to interaction effects in the base design variables.
- The base design is the full factorial design associated with the number of tests we wish to run.

For the case under consideration:

- Five variables will be studied using only eight tests. Therefore, a **2^3 design is the base design**.
- **Two variables must be further introduced in the 2^3 base design**. Columns 12, 13, 23 and 123 are available to introduce these two additional variables.
- The new test plan will be called a **2^{5-2} fractional factorial design**.
- Two levels of each variable. Five variables under study.
- 2^{5-2} = 8 tests to be run.

Fractional Factorial Design (2^{k-p})

- For the five-variable, eight-test fractional factorial under study, let us introduce variables 4 and 5 into the 2^3 base design by assigning them to the 12 and 13 columns, respectively.

w x

Test	*I*	1	2	3	4 12	5 13	23	123	y
1	+	−	−	−	+	+	+	−	y_1
2	+	+	−	−	−	−	+	+	y_2
3	+	−	+	−	−	+	−	+	y_3
4	+	+	+	−	+	−	−	−	y_4
5	+	−	−	+	+	−	−	+	y_5
6	+	+	−	+	−	+	−	−	y_6
7	+	−	+	+	−	−	+	−	y_7
8	+	+	+	+	+	+	+	+	y_8

F F D

Fractional Factorial Design (2^{k-p})

- **The question that remains is to determine exactly which effects are confounded with each other.**
- From now on, when we refer to a column heading (e.g., 1 or 23 or 123) we should imagine a column of + and − signs directly under it. Our 2^{5-2} fractional factorial design was generated by setting the 4-column equal to the 12-column and the 5-column equal to the 13-column.

Test	*I*	1	2	3	4 12	5 13	23	123	y
1	+	−	−	−	+	+	+	−	y_1
2	+	+	−	−	−	−	+	+	y_2
3	+	−	+	−	−	+	−	+	y_3
4	+	+	+	−	+	−	−	−	y_4
5	+	−	−	+	+	−	−	+	y_5
6	+	+	−	+	−	+	−	−	y_6
7	+	−	+	+	−	−	+	−	y_7
8	+	+	+	+	+	+	+	+	y_8

Confounding Effects in Fractional Factorial Design (2^{k-p})

1) In the interest of convenience we will denote these as

$4 = 12$ and $5 = 13$

where the = sign really implies an identity between columns of + and − signs, for example.

2) Now, if any column of + and − signs is multiplied by itself, a column of all + signs is produced. We will denote such a column by the heading $\mathbf{I}$.

3) IF we multiply both sides of the two "equations" above by 4 and 5, respectively, then

$$4\times4 = 12\times4 \qquad 5\times5 = 13\times5$$

which reduces to, **I = 124** **I = 135**

These two identities are referred to as our design generators.

While both the left- and right-hand sides of the equation above represent columns of all + signs, the right-hand side retains the individual column headings that produced the column of all + signs by their product.

Test	I	1	2	3	4 12	5 13	y
1	+	−	−	−	+	+	y_1
2	+	+	−	−	−	−	y_2
3	+	−	+	−	−	+	y_3
4	+	+	+	−	+	−	y_4
5	+	−	−	+	+	−	y_5
6	+	+	−	+	−	+	y_6
7	+	−	+	+	−	−	y_7
8	+	+	+	+	+	+	y_8

Confounding Effects in Fractional Factorial Design (2^{k-p})

$5C_2 = 10$
$5C_3 = 10$
$5C_4 = 5$
1

main   5 main

Design Generators: I = 124   I = 135

- Multiply the design generator equations together,

I x I = 124 x 135 = (11) x 2345 = I x 2345

So, I = 2345

Now, $\boxed{\text{I = 124 = 135 = 2345}}$ * (DR)

The identity I = 124 = 135 = 2345 is referred to as the defining relation of this 2^{5-2} fractional factorial design, and through it we can reveal the complete aliasing/confounding structure of this fractional factorial design.

- Multiply the defining relation by 1

I x 1 = 124 x 1 = 135 x 1 = 2345 x 1

So, 1 = 24 = 35 = 12345

Which means, E1 will estimate (1+24+35+12345)

Test	I	1	2	3	4 12	5 13	y
1	+	−	−	−	+	+	y_1
2	+	+	−	−	−	−	y_2
3	+	−	+	−	−	+	y_3
4	+	+	+	−	+	−	y_4
5	+	−	−	+	+	−	y_5
6	+	+	−	+	−	+	y_6
7	+	−	+	+	−	−	y_7
8	+	+	+	+	+	+	y_8

5 var   full factor $2^5 = 32$

FFD 8 test

Confounding Effects in Fractional Factorial Design (2^{k-p})

Design Generators: **I = 124** **I = 135**

Defining Relation: **I = 124 = 135 = 2345**

- Multiply the defining relation by 2

I x 2 = 124 x 2 = 135 x 2 = 2345 x 2

So, 2 = 14 = 1235 = 345

Which means, E2 will estimate (2+14+1235+345)

- Similarly, we can find,

1 = 24 = 35 = 12345

2 = 14 = 345 = 1235

3 = 1234 = 15 = 245

4 = 12 = 235 = 1345

5 = 13 = 234 = 1245

23 = 134 = 125 = 45

123 = 34 = 25 = 145.

I = 124 = 135 = 2345

Note: Only 8 unique columns out of possible 32

Test	I	1	2	3	4 12	5 13	y
1	+	−	−	−	+	+	y_1
2	+	+	−	−	−	−	y_2
3	+	−	+	−	−	+	y_3
4	+	+	+	−	+	−	y_4
5	+	−	−	+	+	−	y_5
6	+	+	−	+	−	+	y_6
7	+	−	+	+	−	−	y_7
8	+	+	+	+	+	+	y_8

Confounding Effects in Fractional Factorial Design (2^{k-p})

The eight columns (including **I**) in the 2^{5-2} fractional factorial design produce the following linear combinations of effects which can be estimated:

E0 estimates mean + (1/2)(124 + 135 + 2345)

~ mean

E1 estimates 1 + 24 + 35 + 12345

E2 estimates 2 + 14 + 1235 + 345

E3 estimates 3 + 1234 + 15 + 245

E12 estimates 12 + 4 + 235 + 1345

E13 estimates 13 + 234 + 5 + 1245

E23 estimates 23 + 134 + 125 + 45

E123 estimates 123 + 34 + 25 + 145.

Test	I	1	2	3	4 12	5 13	y
1	+	−	−	−	+	+	y_1
2	+	+	−	−	−	−	y_2
3	+	−	+	−	−	+	y_3
4	+	+	+	−	+	−	y_4
5	+	−	−	+	+	−	y_5
6	+	+	−	+	−	+	y_6
7	+	−	+	+	−	−	y_7
8	+	+	+	+	+	+	y_8

Confounding Effects in Fractional Factorial Design (2^{k-p})

If **we assume** that the majority of the variability in the data can be explained by the presence of main effects and two-factor interaction effects, the linear combinations of effects are

E0 estimates mean + (1/2)(~~124~~ + ~~135~~ + ~~2345~~)

E1 estimates 1 + 24 + 35 + ~~12345~~

E2 estimates 2 + 14 + ~~1235~~ + ~~345~~

E3 estimates 3 + ~~1234~~ + 15 + ~~245~~

E12 estimates 12 + 4 + ~~235~~ + ~~1345~~

E13 estimates 13 + ~~234~~ + 5 + ~~1245~~

E23 estimates 23 + ~~134~~ + ~~125~~ + 45

E123 estimates ~~123~~ + 34 + 25 + ~~145~~.

Test	I	1	2	3	4 12	5 13	y
1	+	−	−	−	+	+	y_1
2	+	+	−	−	−	−	y_2
3	+	−	+	−	−	+	y_3
4	+	+	+	−	+	−	y_4
5	+	−	−	+	+	−	y_5
6	+	+	−	+	−	+	y_6
7	+	−	+	+	−	−	y_7
8	+	+	+	+	+	+	y_8

Confounding Effects in Fractional Factorial Design (2^{k-p})

- We have previously seen that the introduction of additional variables into two-level full factorials gives rise to confounding or aliasing of variable effects.
- **It would be desirable to make this introduction in such a way as to confound low order effects (main effects and two-factor interactions) not with each other but with higher order interactions that are considered unimportant.**

Example:

Consider the study of five variables in just sixteen tests (the full factorial would required $2^5 = 32$ tests). One additional variable – the fifth variable – must be introduced into a $2^4 = 16$ run base design. Any of the interactions in the first four variables could be used for this purpose.

- 12, 13, 14, 23, 24, 34
- 123, 124, 134, 234
- 1234.

Confounding Effects in Fractional Factorial Design (2^{k-p})

Example: Consider the study of five variables in just sixteen tests (the full factorial would required $2^5 = 32$ tests). One additional variable – the fifth variable – must be introduced into a $2^4 = 16$ run base design. Any of the interactions in the first four variables could be used for this purpose.

$5 \times 5 = 12 \times 5$

$I = 125$

(A) If any one of the two-factor interactions are used, say, 5 = 12, then the design generator becomes,

And the Defining Relationship will be $I = 125$ —— *

Therefore, at least some of the average/main effects will be confounded with two-factor interactions, viz.,

1 = 25, 2 = 15, 5 = 12 $E_1 \sim 1 + 25$

(B) BUT if we use any of the three-factor interactions to introduce the fifth variable, the situation is greatly improved, at least for the estimation of average/main effects. For example, if we let 5 = 123, then

Design Generator: $I = 1235$ Defining Relationship: $I = 1235$

So, some main effects are confounded with, at worst, three-factor interactions, while two-factor interactions are confounded with each other, e.g., 1 = 235, 2 = 135, 3 = 125, 5 = 123, and 12 = 35, 13 = 25, 23 = 15

Confounding Effects in Fractional Factorial Design (2^{k-p})

Another Option

(C) NOW, If the four-factor interaction is used to introduce the fifth variable, i.e., 5 = 1234, an **even more desirable result is obtained** (the best under these circumstances).

The generator and defining relationship for this situation is I = 12345

Therefore,

1 = 2345,	2 = 1345,	3 = 1245,	4= 1235,	5 = 1234,
12 = 345,	13 = 245,	14 = 235,	15 = 234,	23 = 145,
24 = 135,	25 = 134,	34 = 125,	35 = 124,	45 = 123.

In this last case, all main effects are confounded with four-factor interactions. All two-factor interactions are confounded with three-factor interactions.

The varying confounding structures produced by using different orders of variable interactions to introduce the fifth variable in the example above are described by the concept of the resolution of fractional factorial designs.

Design Resolution of Fractional Factorial Design (2^{k-p})

"The resolution of a two-level fractional factorial design is defined to be equal to the number of letters (numbers) in the shortest length word (term) in the defining relationship, *excluding I*."

Examples:

- If the defining relationship of a certain design is **I = 124 = 135 = 2345** then the design is of **resolution three**, denoted as a Resolution III, since the words "123" and "135" have three letters each.
- If the defining relation of a certain design is **I = 1235 = 2345 = 1456** then the design is of **Resolution IV** ("1235", "2346" and "1456" each have four letters).
- Similarly, the design with defining relationship (**I = 12345**), is a **Resolution V design**.

Important Note:

- If a design is of Resolution III, this means that at least some main effects are confounded with two-factor interactions.
- If a design is of Resolution IV, this means that at least some main effects are confounded with three-factor interactions while at least some two-factor interactions are confounded with other two-factor interactions.
- If a design is of Resolution V, this means that at least some main effects are confounded with four-factor interactions and some two-factor interactions are confounded with three-factor interactions.

Design Resolution of Fractional Factorial Design (2^{k-p})

- It may be noted, that the number of words in the defining relationship for a 2^{k-p} fractional factorial design is equal to 2^p. Thus, for a 2^{6-3} fractional factorial ($k=6$ and $p=3$), there are $2^3 = 8$ words in the defining relationship.

Example: A 2^{6-2} fractional factorial design is set up by introducing variable 5 and 6 via $5 = 123$, $6 = 1234$.

- What is the resolution of this design? The design generators are: $I = 1235$, $I = 12346$
- The defining relationship is: $I = 456 = 1235 = 12346$
- Design resolution is: III

What would the resolution be if the generators were $5 = 123$, $6 = 124$?

design gen $\quad I = 1235 \quad I = 1246$

Def. rel $\quad I = 1235 = 1246 = 3456$

Design res $\quad$ IV

Design Resolution of Fractional Factorial Design (2^{k-p})

- **Higher resolution designs seem more desirable** since they provide the opportunity for low order effect estimates to be determined in an unconfounded state, assuming higher order interaction effects can be neglected.
- **The more variables considered in a fixed number of tests, the lower the resolution of the design becomes.**
- There is a limit to the number of variables that can be considered in a fixed number of tests while maintaining a pre-specified resolution requirement.
- No more than **(n-1)** variables can be examined in **n** tests (n is a power of 2, e.g., 4, 8, 16, 32, …) to maintain a design resolution of at least III. Such designs are commonly referred to as saturated design.
  - Examples are 2^{3-1}, 2^{7-4}, 2^{15-11}, 2^{31-26}.
- **For saturated designs all interactions in the base design variables are used to introduce additional variables.**

4th var 12, 13, 23, 123

4 variables ? 2^{4-1}

5 variables ? 2^{5-2} 4th, 5th 12, 13, 23, 123

$2^3 = 8$ test

7 variables ? 2^{7-4} 4th, 5th, 6th, 7th 12, 13, 23, 123

8 variables → 2^{8-5}