Example: Change of Variables

The stationary point we found was $x_{1,s} = 0.389$ and $x_{2,s} = 0.306$

In terms of original time and temperature units, $\xi_1 = 86.95 \simeq 87$ minutes of reaction time and $\xi_2 = 176.53 \simeq 176.5°F$.

What would you do if for some reason (e.g. cost) we cannot operate at this point?

- We can back away slightly from this optimum point and see if any other point in vicinity can work
- Where to go?

$$\hat{y} = 80.21 - 0.9641w_1^2 - 1.4147w_2^2$$

at x_s

$w_1 = w_2 = 0$

you would move first in dir where $|\lambda_i|$ is smallest

Ridge Systems

- Consider 2nd order model with canonical form $\hat{y} = \hat{y}_s + \lambda_1 w_1^2 + \lambda_2 w_2^2 + \cdots + \lambda_k w_k^2$
- Suppose the stationary point (x_s) is in the region of experimentation, and some λ_i are small, i.e., $\lambda_i \approx 0$
- Then, the response variable y is very insensitive to variables with small λ

Is small

$$\hat{y} = \hat{y}_s + \lambda_1 w_1^2 + \lambda_2 w_2^2$$

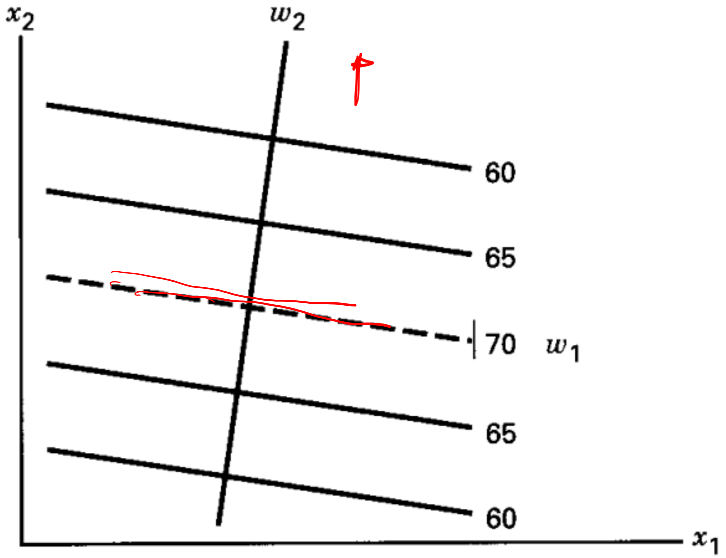


Figure 11-12 A contour plot of a stationary ridge system.

- Because of small λ_1, the optimum can be taken anywhere along the line of y = 70
- This type of response surface is called 'stationary ridge system'

Ridge System

- If the stationary point (x_s) is far outside the region of experimentation, and some λ_i are small, i.e., $\lambda_i \approx 0$
- Then the response surface could be a 'rising ridge' or 'falling ridge'
- In such type of systems, we can NOT draw conclusions about the true surface or the stationary point
- BECAUSE the stationary point is far outside the region where we fitted the model
- In this example, further exploration is needed in the w1 direction

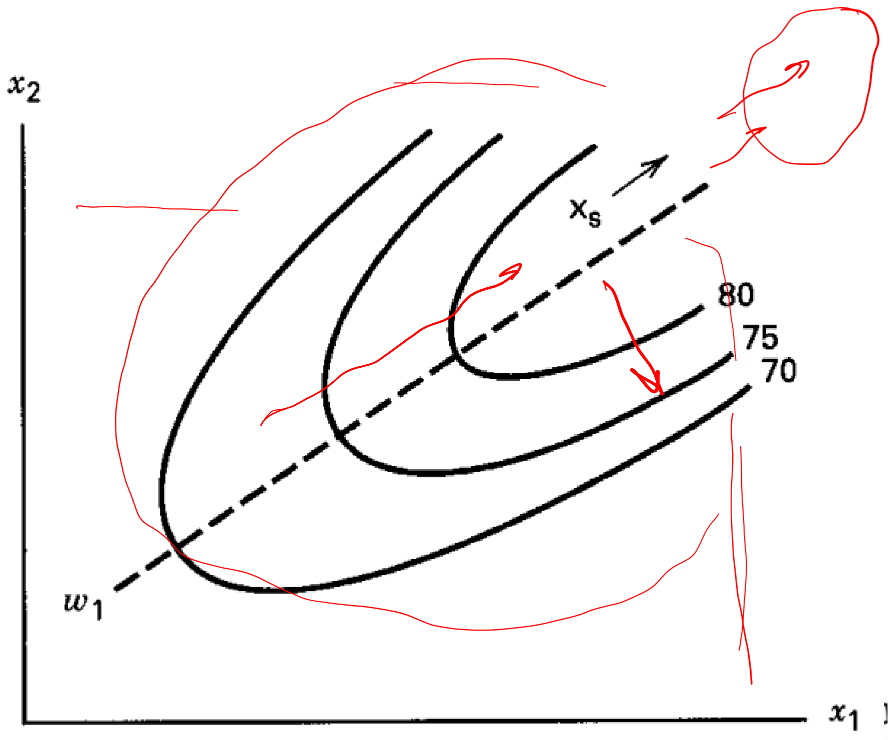


Figure 11-13 A contour plot of a rising ridge system.

Multiple Responses

- Consider the previous example

- Similar to yield, we can also obtain models for viscosity and molecular weight

We only found a response surface for ONE of the responses - Yield

Final Equation in Terms of Coded Factors:

yield =
+79.94
+0.99 * A
+0.52 * B
−1.38 * A²
−1.00 * B²
+0.25 * A * B

$$\hat{y}_2 = 70.00 - 0.16x_1 - 0.95x_2 - 0.69x_1^2 - 6.69x_2^2 - 1.25x_1x_2$$
$$\hat{y}_3 = 3386.2 + 205.1x_1 + 17.4x_2$$

In terms of the natural levels of time (ξ_1) and temperature (ξ_2), these models are

$$\hat{y}_2 = -9030.74 + 13.393\xi_1 + 97.708\xi_2 - 2.75 \times 10^{-2}\xi_1^2 - 0.26757\xi_2^2 - 5 \times 10^{-2}\xi_1\xi_2$$

and

$$\hat{y}_3 = -6308.8 + 41.025\xi_1 + 35.473\xi_2$$

Time, Temp

Natural Variables		Coded Variables		Responses		
ξ_1	ξ_2	x_1	x_2	y_1 (yield)	y_2 (viscosity)	y_3 (molecular weight)
80	170	−1	−1	76.5	62	2940
80	180	−1	1	77.0	60	3470
90	170	1	−1	78.0	66	3680
90	180	1	1	79.5	59	3890
85	175	0	0	79.9	72	3480
85	175	0	0	80.3	69	3200
85	175	0	0	80.0	68	3410
85	175	0	0	79.7	70	3290
85	175	0	0	79.8	71	3500
92.07	175	1.414	0	78.4	68	3360
77.93	175	−1.414	0	75.6	71	3020
85	182.07	0	1.414	78.5	58	3630
85	167.93	0	−1.414	77.0	57	3150

How would you optimize multiple responses?

A relatively straightforward approach to optimizing several responses that works well when there are only a few process variables is to **overlay the contour plots** for each response. Figure 11-16 (page 451) shows an overlay plot for the three responses in Example 11-2, with contours for which y_1 (yield) $\geqslant 78.5$, $62 \leqslant y_2$ (viscosity) $\leqslant 68$, and y_3 (molecular weight Mn) $\leqslant 3400$. If these boundaries represent important conditions that must be met by the process, then as the unshaded portion of Figure 11-16 shows, there are a number of combinations of time and temperature that will result in a satisfactory process. The experimenter can visually examine the contour plot to determine appropriate operating conditions. For example, it is likely that the experimenter would be most interested in the larger of the two feasible operating regions shown in Figure 11-16.

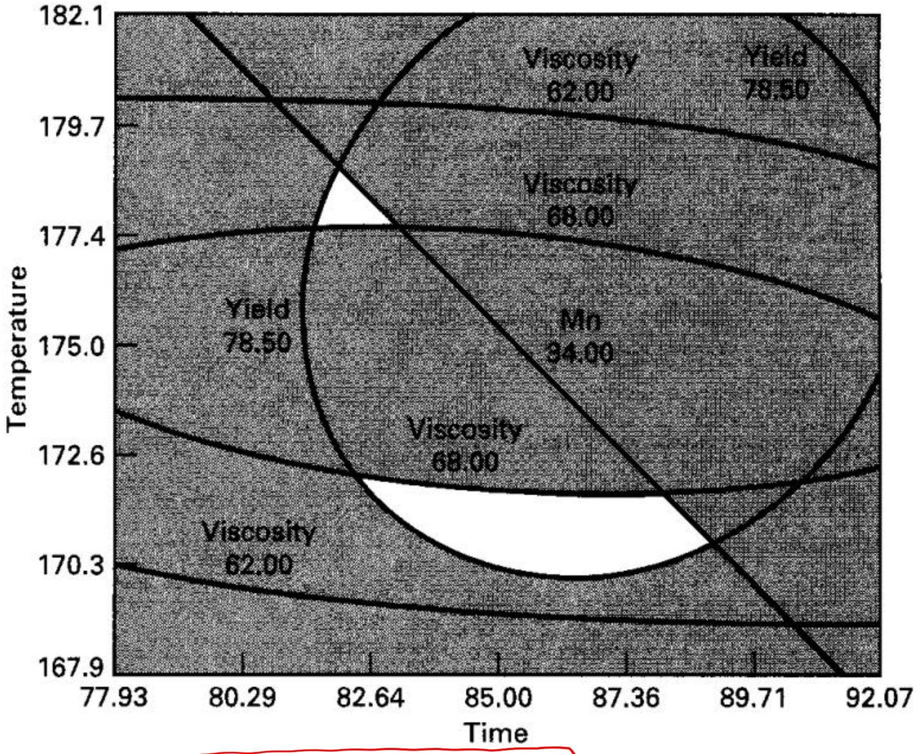


Figure 11-16 Region of the optimum found by overlaying yield, viscosity, and molecular weight response surfaces, Example 11-2.

But what will you do if there are even more responses OR if there are more than two independent variables?

Graphical method won't work!

Constrained Optimization

- Popular approach is to formulate the problem as constrained optimization problem
- For example,

$$\begin{gathered}\text{Max } y_1 \checkmark \\ \text{subject to} \\ 62 \leqslant y_2 \leqslant 68 \checkmark \\ y_3 \leqslant 3400 \checkmark\end{gathered}$$

- Then one can use numerical techniques to solve such a problem ('non-linear programming methods')

The two solutions found are

$$\text{time} = 83.5 \qquad \text{temp} = 177.1 \qquad \hat{y}_1 = 79.5$$

$$\text{time} = 86.6 \qquad \text{temp} = 172.25 \qquad \hat{y}_1 = 79.5$$

Desirability Functions

Another useful approach to optimization of multiple responses is to use the simultaneous optimization technique popularized by Derringer and Suich (1980). Their procedure makes use of **desirability functions**. The general approach is to first convert each response y_i into an individual desirability function d_i that varies over the range

$$0 \leqslant d_i \leqslant 1$$

where if the response y_i is at its goal or target, then $d_i = 1$, and if the response is outside an acceptable region, $d_i = 0$. Then the design variables are chosen to maximize the overall desirability

$$D = (d_1 \cdot d_2 \cdot \cdots \cdot d_m)^{1/m}$$

where there are m responses.

yield | visc | mol. wt

y_1 | y_2 | y_3

$y_1 = 80$

$d_1 = 1$

$y_1 \not> 70\%$

$y_1 \leq 70$

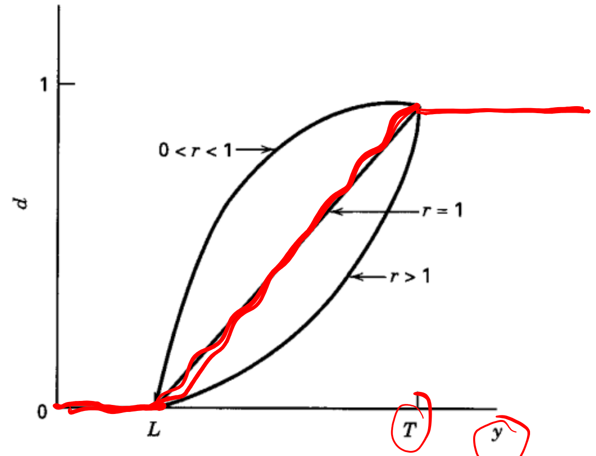
$d_1 = 0$

d_1

Desirability Functions

If the objective or target T for the response y is a maximum value,

$$d = \begin{cases} 0 & y < L \\ \left(\dfrac{y-L}{T-L}\right)^r & L \leqslant y \leqslant T \\ 1, & y > T \end{cases}$$

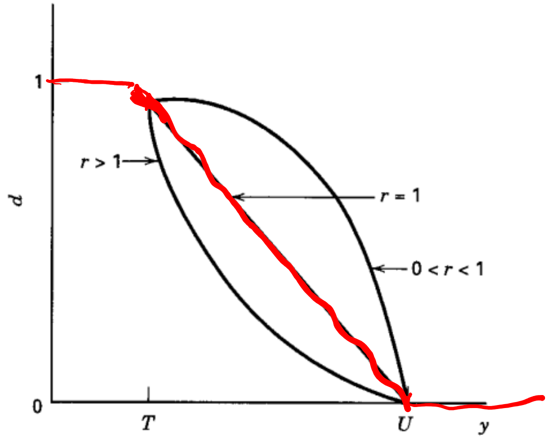


(a) Objective (target) is to maximize y

$y_1 \to d_1$

If the target for the response is a minimum value,

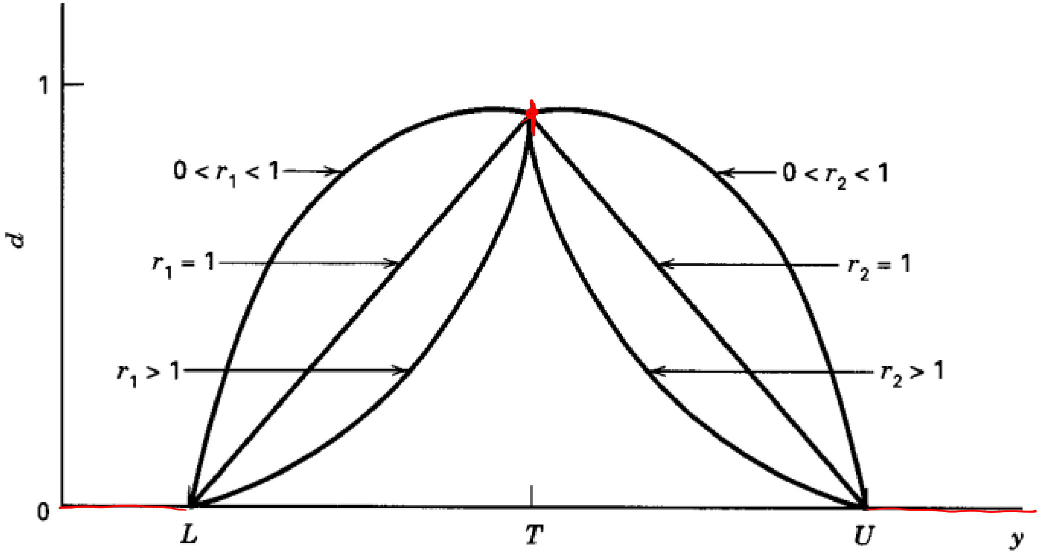
$$d = \begin{cases} 1 & y < T \\ \left(\dfrac{U-y}{U-T}\right)^r & T \leqslant y \leqslant U \\ 0 & y > U \end{cases}$$



(b) Objective (target) is to minimize y

The two-sided desirability function shown in Figure 11-17(c) assumes that the target is located between the lower (L) and upper (U) limits, and is defined as

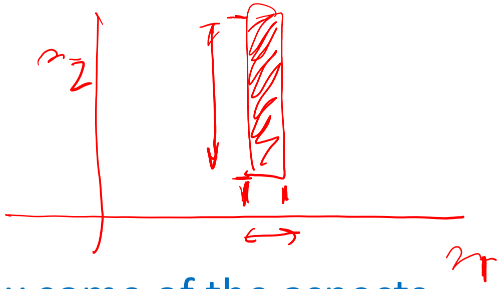
$$d = \begin{cases} 0 & y < L \\ \left(\dfrac{y-L}{T-L}\right)^{r_1} & L \leqslant y \leqslant T \\ \left(\dfrac{U-y}{U-T}\right)^{r_2} & T \leqslant y \leqslant U \\ 0 & y > U \end{cases} \qquad (11\text{-}13)$$



(c) Objective is for y to be as close as posible to the target

Design of Experiments for Fitting Response Surfaces

- Fitting and analysing a response surface can be made very effective by *proper choice of experimental design* to collect the data
- What is a 'good experimental design'?



1. Provides a reasonable distribution of data points (and hence information) throughout the region of interest
2. Allows model adequacy, including lack of fit, to be investigated
3. Allows experiments to be performed in blocks
4. Allows designs of higher order to be built up sequentially
5. Provides an internal estimate of error
6. Provides precise estimates of the model coefficients
7. Provides a good profile of the prediction variance throughout the experimental region
8. Provides reasonable robustness against outliers or missing values
9. Does not require a large number of runs
10. Does not require too many levels of the independent variables
11. Ensures simplicity of calculation of the model parameters

Note how some of the aspects are conflicting -> We need to apply our judgement

DOE for fitting First-Order Model

- First order model with ‘k’ variables: $y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \epsilon$
- Using the experimental data, one would find the regression coefficients β_i
- There is a unique class of experimental designs that can minimize the variance of β_i ->

Orthogonal First-Order Designs

- A first-order design is orthogonal if the off-diagonal elements of (X’X) matrix are ALL ZERO
- That is the <cross> products of the columns of the X matrix add to ZERO

$$\sum x_{1i} x_{2i} = 0$$

2^2

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$$

$$y_1 = \beta_0 - \beta_1 - \beta_2$$
$$y_2 = \beta_0 + \beta_1 - \beta_2$$
$$y_3 = \beta_0 - \beta_1 + \beta_2$$
$$y_4 = \beta_0 + \beta_1 + \beta_2$$

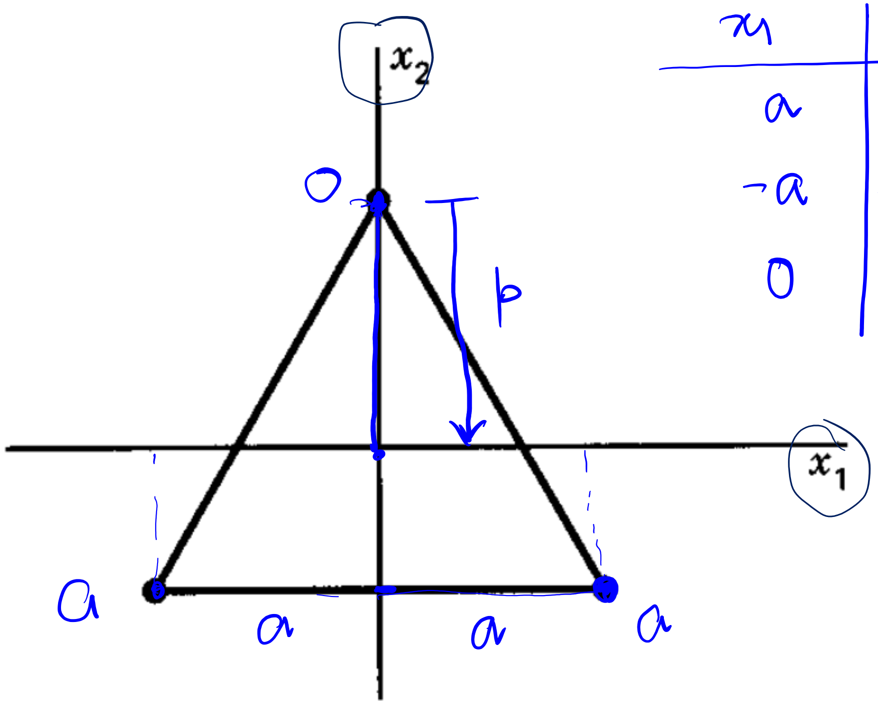
x_1	x_2	y
−1	−1	y_1
+1	−1	y_2
−1	+1	y_3
+1	+1	y_4

Orthogonal First-order Designs

- Are 2^k factorial designs orthogonal?
- Yes – they are a part of orthogonal first-order designs
  - BUT, 2^k design does not afford an estimate of experimental error unless we do replication
  - A common method is to augment the 2^k design with several observations at the centre point
  - The addition of centre point doesn't influence the $\beta_1, \beta_2, \ldots$, BUT β_0 becomes a grand average of ALL the observations
  - Also, note how addition of centre point does not affect the orthogonality of the design

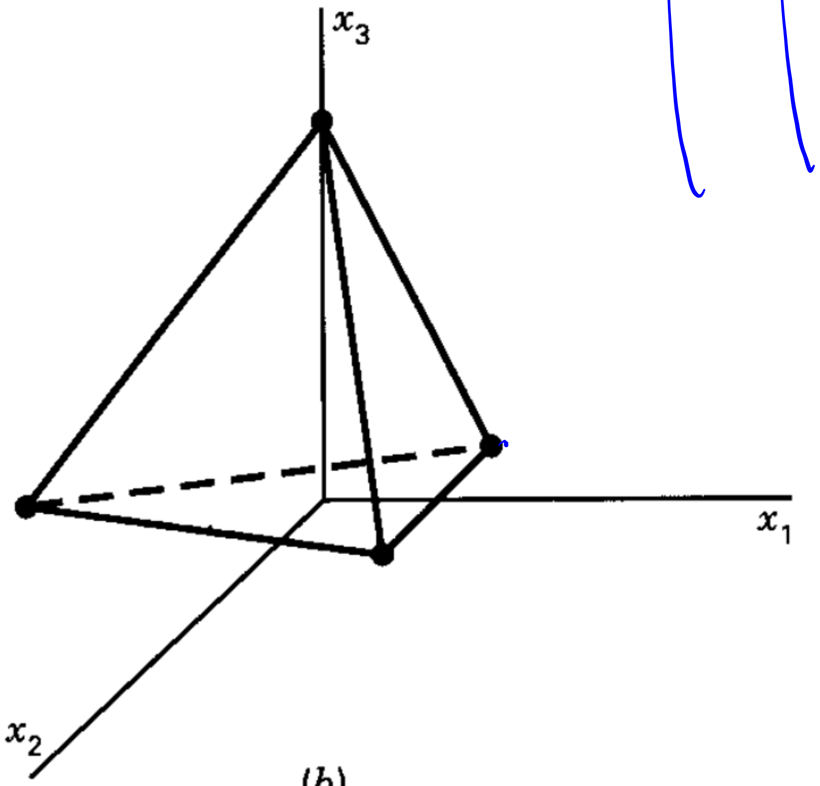
Orthogonal First-order Designs

Another orthogonal first-order design is the **simplex**. The simplex is a regularly sided figure with $k + 1$ vertices in k dimensions. Thus, for $k = 2$ the simplex design is an equilateral triangle and for $k = 3$ it is a regular tetrahedron. Simplex designs in two and three dimensions are shown in Figure 11-19.



(a)

x_1	x_2
a	$-\sqrt{3}/2\, a$
$-a$	$-\sqrt{3}/2\, a$
0	0



(b)

x_1	x_2	x_3

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.
- 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.

1	Mean
6	Main effects
15	Two-factor interaction
20	Three-factor interaction
15	Fourth order
6	Fifth order
1	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

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

note exactly 4 (+)
4 (-)

no 2 columns same

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.

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

main 6 4 1

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).

E_0 l_0 estimates mean + (1/2)(1234) ✓

E_1 l_1 estimates 1 + 234 0

E_2 l_2 estimates 2 + 134 0

l_3 estimates 3 + 124 0

E_{12} l_{12} estimates 12 + 34

E_{13} 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)

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.
- 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.

					4	5			
Test	***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

FFD

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

$$4 = 12 \quad \text{and} \quad 5 = 13$$

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 **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, $I = 124$ $\qquad$ $I = 135$

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

main 5 main

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

Design Generators: I = 124 I = 135

- Multiply the design generator equations together,

$$I \times I = 124 \times 135 = (11) \times 2345 = I \times 2345$$

So, I = 2345

Now, 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

$$I \times 1 = 124 \times 1 = 135 \times 1 = 2345 \times 1$$

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 factr $2^5 = 32$

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~~)

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

~ mean

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.

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

$5 \times 5 = 12 \times 5$

$I = 125$

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

(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

$E_1 \sim 1 + 25$

Design Generator: Defining Relationship:

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

$I = 1235$ $I = 1235$

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.