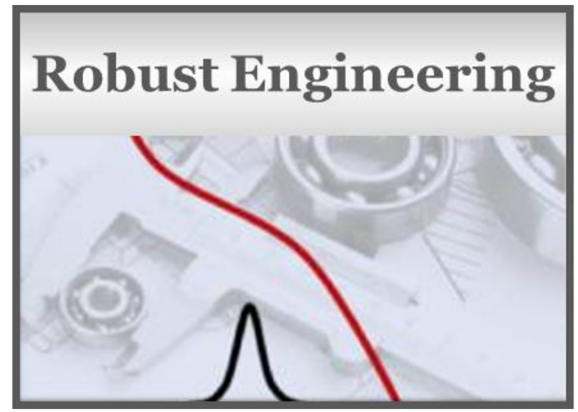
Module: Robust Engineering

Design of Experiments & Response Surface Modelling



Session 5: Planning a Two-Level Experiment- II

In This Session We Will…

- Show that we can save experimental resource by ignoring interactions.
- Explain that main effects are then confounded with interactions (but if interactions are relatively small, this strategy is useful for screening).
- List some designs that are useful for this strategy, including the 12 run Plackett-Burman design.
- Introduce orthogonal arrays and show that they are equivalent to fractional factorials.
- Introduce the concept of resolution number.
- Discuss the role of qualitative factors.
- Give an example of an experiment that is run in 2 blocks.
- Discuss some of the costs and benefits of replication.

Reminder

- If a two-level screening experiment is to be used as the first step in a sequential experimentation strategy, there are three basic ways of designing it:
  - fit a 1st order response surface, ignoring interactions,
  - fit selected interactions,
  - fit a complete '1st order + twist' response surface.
- The last of these options was covered in Session 4

The 'selected interactions' strategy requires detailed knowledge to implement

We will look at the '1st order ' strategy

Fitting a 1st Order Response Surface

- Suppose we have three factors; the 1st order equation is

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3$$

- Here is a design that will fit this equation:

2^{3-1} design

Row	x_1	x_2	x_3
1	−1	−1	+1
2	+1	−1	−1
3	−1	+1	−1
4	+1	+1	+1

The smallest three-factor design

Compared with the 2^3 full factorial, we have saved half our resource!

Confounding

- Although the 2^{3-1} design will allow us to fit the equation,

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3$$

- that is not the end of the story, because our results are vulnerable to the possibility of confounding.
- For example, when we fit the x_1 term we are actually estimating *the sum of the real x_1 and x_2x_3 coefficients.*
- To see why this happens, suppose we try to fit all the 2fi's:

Multiply the x_2 and x_3 columns

Inter-cept	x_1	x_2	x_3	x_1x_2	x_1x_3	x_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

The model matrix contains identical columns

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3$$

Confounding (cont.)

- We cannot fit an equation that has terms in both x_1 and x_2x_3.
- We would usually opt to fit the linear term (an estimate of the gradient of the surface in the x_1 direction).
  - but this doesn't alter the fact that we are actually estimating the sum of the real x_1 and x_2x_3 coefficients, i.e. $\beta_1 + \beta_{23}$.

Remember that we used β's to represent the true (unknown) coefficients

- These two effects are said to be completely confounded (confused).
- The linear terms are often called the main effects,
  - so in the 2^{3-1} design the main effect of x_1 is completely confounded with the interaction of x_2 and x_3.

Some textbooks and software use 'aliased' rather than 'confounded'

Implications of Confounding

- Continuing to focus on x_1, our estimate of the x_1 gradient will be biased unless the real x_2x_3 interaction coefficient β_{23} is zero.
    - the extent of the bias depends on how big β_{23} really is.

Also, a negative coefficient β_{23} could cancel out a positive coefficient β_1 (and vice versa)

On a more optimistic note…

- In many systems the interactions are small (relative to the main effects).
- In other words the system can be approximately represented by a 1st order response surface,
    - this will give us a good idea of which factors have the biggest effect, at minimal cost.

Selecting A Design For The '1st Order' Strategy

- Since the purpose of this '1st order' strategy is usually to minimize the number of runs, we look for the smallest design that will accommodate our list of factors.

No. of factors	Design	No. of runs
3	2^{3-1}	4
4	2^{4-1}	8
5	2^{5-2}	8
6	2^{6-3}	8
7	2^{7-4}	8
8-11	Plackett-Burman	12

Regular fractions

Non-regular fractions

This Plackett-Burman design is given on the next slide

Plackett-Burman Design For 11 Factors

Row	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}
1	+1	+1	–1	+1	+1	+1	–1	–1	–1	+1	–1
2	–1	+1	+1	–1	+1	+1	+1	–1	–1	–1	+1
3	+1	–1	+1	+1	–1	+1	+1	+1	–1	–1	–1
4	–1	+1	–1	+1	+1	–1	+1	+1	+1	–1	–1
5	–1	–1	+1	–1	+1	+1	–1	+1	+1	+1	–1
6	–1	–1	–1	+1	–1	+1	+1	–1	+1	+1	+1
7	+1	–1	–1	–1	+1	–1	+1	+1	–1	+1	+1
8	+1	+1	–1	–1	–1	+1	–1	+1	+1	–1	+1
9	+1	+1	+1	–1	–1	–1	+1	–1	+1	+1	–1
10	–1	+1	+1	+1	–1	–1	–1	+1	–1	+1	+1
11	+1	–1	+1	+1	+1	–1	–1	–1	+1	–1	+1
12	–1	–1	–1	–1	–1	–1	–1	–1	–1	–1	–1

For 8, 9 or 10 factors, use a subset of the columns

Orthogonal Arrays

- Orthogonal arrays (OA's) are mathematically constructed patterns of symbols with special properties.

Genichi Taguchi published tables of OA's and gave them names L4, L8, L12, etc.

- Here is the smallest two-level OA, which Taguchi called L4:

1	1	1
1	2	2
2	1	2
2	2	1

If we think of these as factor levels, any two columns form a full factorial

This definition of 'Orthogonal' is different from 'zero correlations', but very closely related

There are also three-level OA's and mixed-level OA's

Orthogonal Arrays And Fractional Factors

- Compare L4 with the 2^{3-1} design:

1	1	1
1	2	2
2	1	2
2	2	1

−1	−1	+1
+1	−1	−1
−1	+1	−1
+1	+1	+1

- If we identify 1 and 2 with –1 and +1, we don't quite get the same pattern (the final column has signs reversed),
  - but statistically, they are equivalent.

All the fractional factorial designs are statistically equivalent to OA's

The Taguchi notation is popular in some companies – but we only need one or the other!

e.g. the 12 run Plackett-Burman design is equivalent to Taguchi's L12

Higher-Order Effects

- In describing the possible confounding of effects, we have only considered 1st and 2nd order effects (main effects and 2fi's).
- Real systems may have higher-order interaction effects.
  - we will look at a three-factor interaction (3fi) in next session.
- All fractional factorial experiments are subject to some confounding between the effects in our response surface equation and higher-order interactions.
  - in theory the effects in our equation may be biased by higher-order interactions.
  - we hope that this bias is negligible in practical terms.

Resolution Numbers

- In books and software, two-level designs are often classified by their Resolution number (a Roman number III, IV, V, etc.).
- For example, in a Resolution V design the worst confounding is between 2fi's and 3fi's (note that 2 + 3 = 5).
    - this implies that no linear effect is confounded with any 2fi, and there is no confounding between 2fi's.
    - so we can definitely fit a complete '1st order + twist' equation.
- In the tutorial, we will give examples of designs with different resolution numbers.

Quantitative And Qualitative Levels

- So far we have assumed that the factor levels are quantitative, so that it makes sense to predict the response for intermediate levels.
  - e.g. we set temperature at 100 and 200 ℃, and make a prediction for 150 ℃.
- We sometimes want to use qualitative levels (left and right, old and new, supplier A and B, …).
- The useful two-level experimental designs are the same for both types of factor.
- Experiments with multi-level qualitative factors require a different type of design and analysis, and are usually very hard to interpret without making strong assumptions.

e.g. that all interactions are negligible

Blocks

- We may anticipate that discrete ‘jumps’ in the response will occur between different batches of runs, independent of our factor changes.
    - for example, the batches may correspond to different days, different machines, different batches of material.
- The batches are called blocks in general.

- Start with just two blocks.
- This is like having an extra two-level factor, except we usually assume there is ***no interaction*** between this blocking factor and the experimental factors.

2^3 Example

- Suppose the equation we want to fit is:

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_{12} x_1 x_2 + b_{13} x_1 x_3 + b_{23} x_2 x_3$$

- We want to run a 2^3 full factorial but this can't be done in one block
- Here is the best design:

Block	x_1	x_2	x_3
1	–1	–1	–1
1	+1	–1	+1
1	–1	+1	+1
1	+1	+1	–1
2	–1	–1	+1
2	+1	–1	–1
2	–1	+1	–1
2	+1	+1	+1

Both levels occur twice in each block

i.e. more blocks and/or more factor levels; we'll come back to this in later sessions

This pattern could be worked out 'on the back of an envelope' but other cases are more challenging

Replication

- Replication means running through the whole design more than once (not necessarily in the same order).
- This is different from making several observations from the same set-up (usually called repetition).
- It is always better to have more data, but replicating the whole design is a costly strategy.
- We will look at some of the advantages of replication in later sessions.

For replication to work properly, all the potential sources of random variation that operate between different factor settings must also be present when you make two or more observations at the same factor settings

e.g. if part of the random variation comes from setting a fixture, the fixture should be 'un-set' and re-set for every observation

In This Session We Have…

- Shown that we can save experimental resource by ignoring interactions.
- Explained that main effects are then confounded with interactions (but if interactions are relatively small, this strategy is useful for screening).
- Listed some designs that are useful for this strategy, including the 12 run Plackett-Burman design.
- Introduced orthogonal arrays and shown that they are equivalent to fractional factorials.
- Introduced the concept of resolution number.
- Discussed the role of qualitative factors.
- Given an example of an experiment that is run in 2 blocks.
- Discussed some of the costs and benefits of replication.

Session 5: Planning a Two Level Experiment II

Tutorial and Exercise

Tutorial

- **Session TS05: Planning a Two Level Experiment II**
- **Objectives**

  Create a two level fractional factorial experiment.

- **Python Environment**

  A self-guided tutorial has been created as a Colab notebook with pre-designed Python code and notes. For this tutorial, follow the instructions in the notes, upload data files and run the code. No modification of code is required. Interpret the results in accordance with the Technical session.

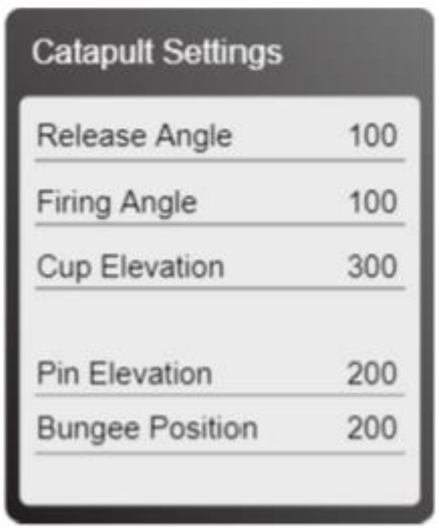
- **Tutorial Task**

  1. Generate a Fractional Factorial design and analyse the design.
  2. Generate a coded 2^{5-1} fractional factorial design.
  3. Generate a coded 2^{5-2} fractional factorial design.
  4. Convert the coded 2^{5-1} fractional factorial design into engineering units.

Exercise

- **Session TS04+05: Planning a Two Level Experiment**
- **Objective**
- To plan a screening experiment for a Virtual Catapult, run the experiment and collect the responses.

https://sigmazone.com/catapult/



Exercise

- **Session TS04+05: Planning a Two Level Experiment**
- **Python Environment**

  The exercise has been created as a Colab notebook with notes. For this exercise, follow the instructions in the notes, and create your own code using the accompanying tutorial as a guide. Interpret the results in accordance with the Technical session. **Continue with Exercise 05, Task 1.**
- **Guidelines**
  - The response is the horizontal distance travelled by the ball.
  - For screening experiment you should include all the 5 factors- show in the "Catapult Settings" dialogue box.
  - In Exercise 06 you will select 3 of them and run a three-level experiment to investigate their effects in more detail.
  - You should first “play” with the simulator to identify some reasonable limits for the DoE space your experiment (i.e. the minimum and maximum setting for each factor).
  - In choosing a design, you need to decide what kind of response surface you want to fit, i.e. 1st order only (ignoring interactions) or 1st order + twist.
  - Then select a design that allows you to fit this equation.