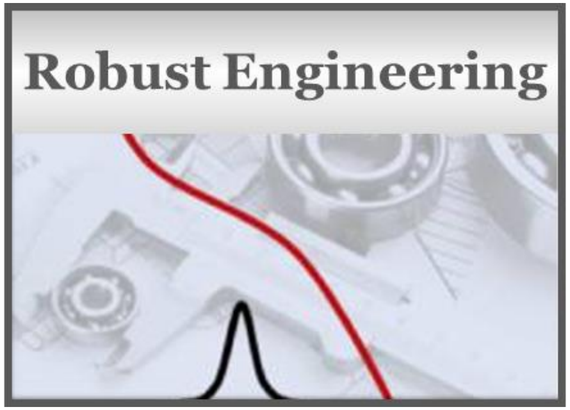
UNIVERSITY of BRADFORD

Statistiques Appliquées Formation Industrielle

Module: Robust Engineering

Design of Experiments & Response Surface Modelling



Session 7: Analysing a Two-Level Screening Experiment- II

In This Session We Will…

- Use the 2^3 design to explain the construction and interpretation of the Half Normal plot.
- Explain that a Half Normal plot can include checking effects such as the three-factor interaction.
- Compare two methods of calibrating the plot:
  - by eye,
  - using Lenth’s PSE.
- Apply these methods to a larger (2^{5-1}) design.
- Compare Half Normal and Full Normal plots.

The Half Normal Plot

This plot was invented by Cuthbert Daniel

- This can always be used as an alternative to t tests.
- It is the only method available when ResDF = 0.
- It provides a simple way of investigating higher-order interactions that are not in our ‘1st order + interactions’ regression equation.

Constructing A Half Normal Plot

- The plot is constructed from the effects that we calculated in the previous session (TS6).
  - one effect for each column of the model matrix M, apart from the intercept
- …. together with some extra effects known as checking effects, which are usually higher-order effects than the ones in our regression equation.
  - one checking effect for each residual degree of freedom.
- Altogether, there will be $(n-1)$ points on the plot.
- The method:
  - arrange the $(n-1)$ effects in increasing order of absolute value (size),
  - plot the ordered absolute effects against a set of Half Normal scores (percentiles of the Half Normal distribution).

Example

- These are the effects we calculated earlier for the 2^3 experiment:

	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3
Ave (+)	41.01	41.25	42.39	41.41	41.20	41.37
Ave (–)	41.48	41.24	40.09	41.08	41.29	41.11
Effect	-0.47	0.01	2.30	0.33	-0.09	0.26

- For the 2^3 experiment we have 1 residual degrees of freedom, so 1 checking effect.
    - it is $x_1x_2x_3$, the three-factor interaction (3fi) between x_1, x_2 and x_3.

A 3fi measures the extent to which (e.g.) the x_1x_2 interaction varies with the level of x_3

The 3fi is calculated on the next slide

Calculating The Three-factor Interaction

- The method is the same as before:
  - construct the $x_1x_2x_3$ column by multiplying the x_1, x_2 and x_3 columns together row by row,
  - pick out the rows with a '$+1$' and average the y values,
  - pick out the rows with a '-1' and average the y values,
  - subtract the '-1' average from the '$+1$' average.

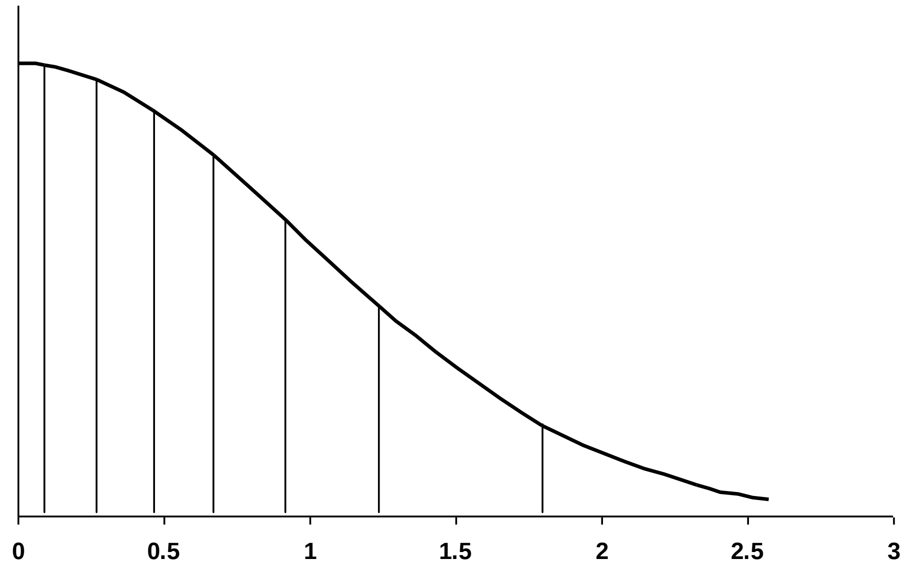
	x_1	x_2	x_3	$x_1x_2x_3$	y
	-1	-1	-1	–1	40.49
	+1	–1	–1	+1	39.94
	–1	+1	–1	+1	40.07
	+1	+1	–1	–1	39.86
	–1	–1	+1	+1	42.78
	+1	–1	+1	–1	41.74
	–1	+1	+1	–1	42.56
	+1	+1	+1	+1	42.49
Ave (+)				41.32	
Ave (-)				41.16	
Effect				0.16	

Constructing The Plot

- Here are the 7 effects arranged in order of absolute value:

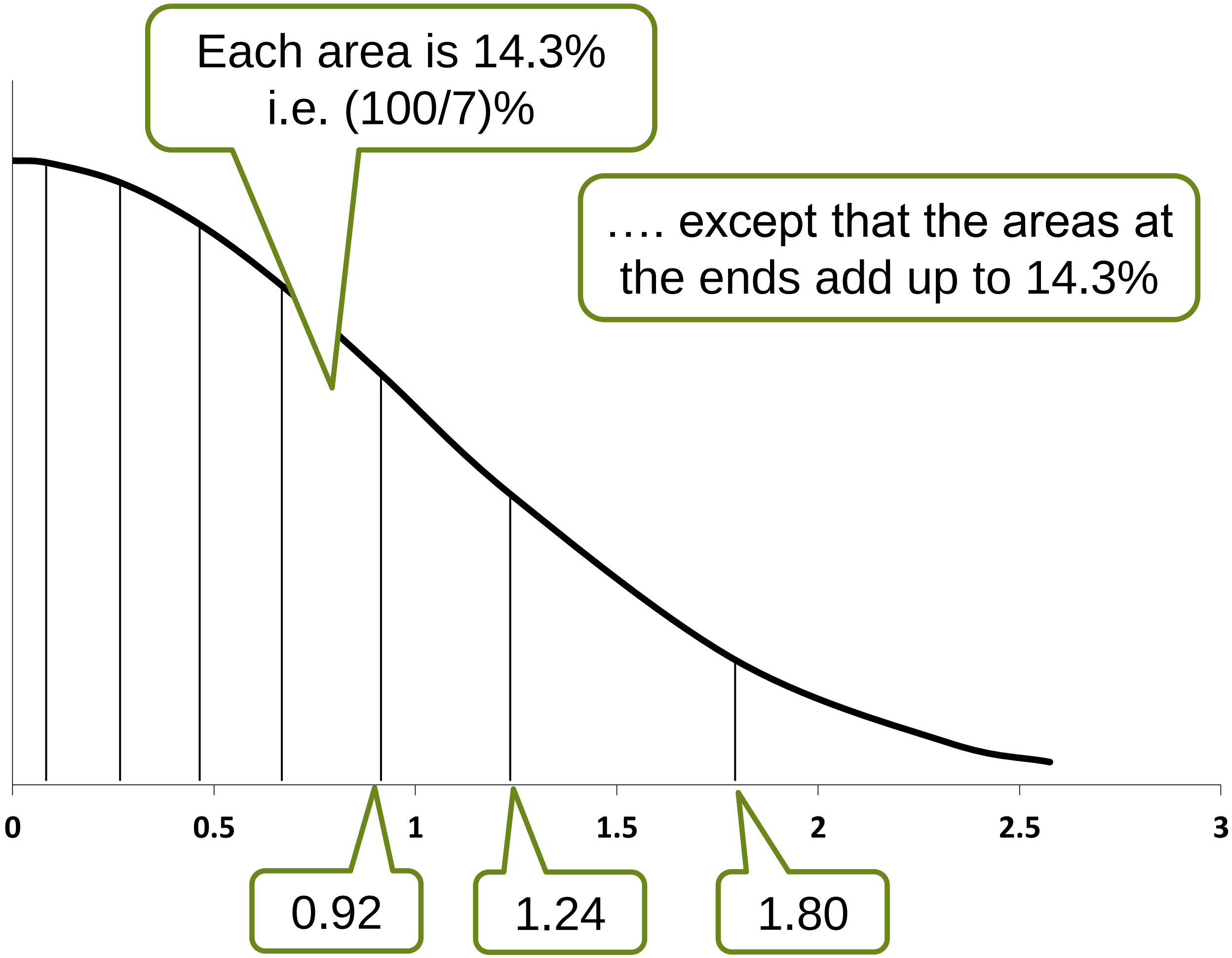
x_2	x_1x_3	$x_1x_2x_3$	x_2x_3	x_1x_2	x_1	x_3
0.01	0.09	0.16	0.26	0.33	0.47	2.30

- We find the 7 scores (percentiles) by dividing the Half Normal distribution into equal areas.



The Half Normal Scores (cont.)

We use the Half Normal distribution because we are plotting absolute values

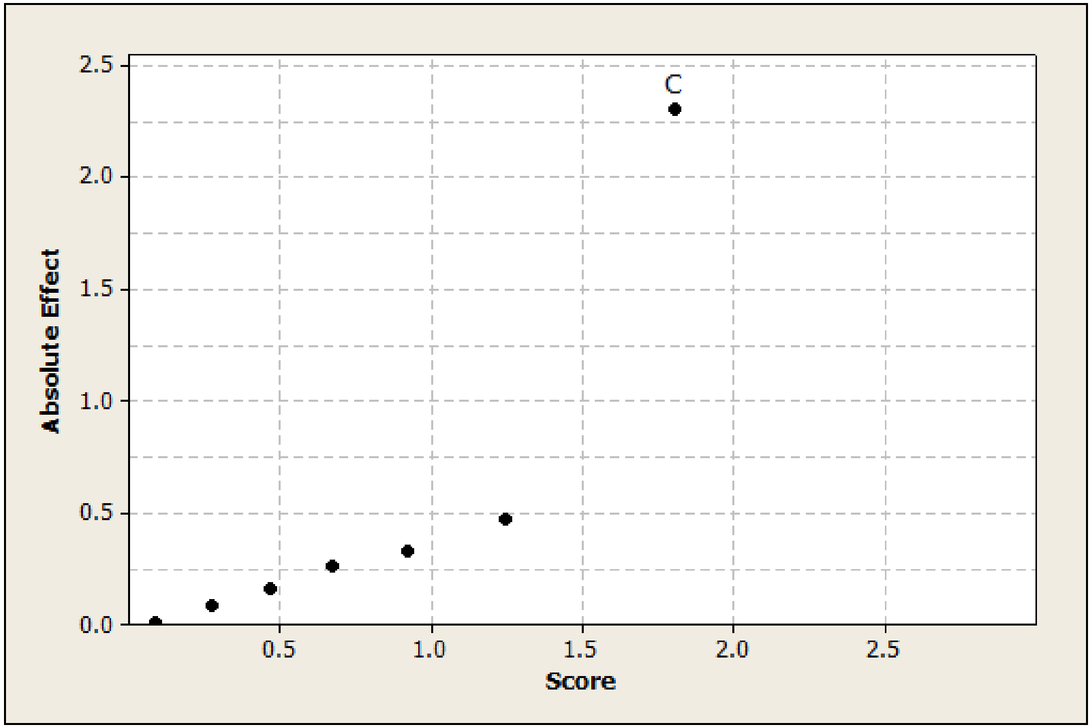


Half Normal Plot For Our Example

	x_2	x_1x_3	$x_1x_2x_3$	x_2x_3	x_1x_2	x_1	x_3
Abs. effect	0.01	0.09	0.16	0.26	0.33	0.47	2.30
Score	0.09	0.27	0.46	0.67	0.92	1.24	1.80

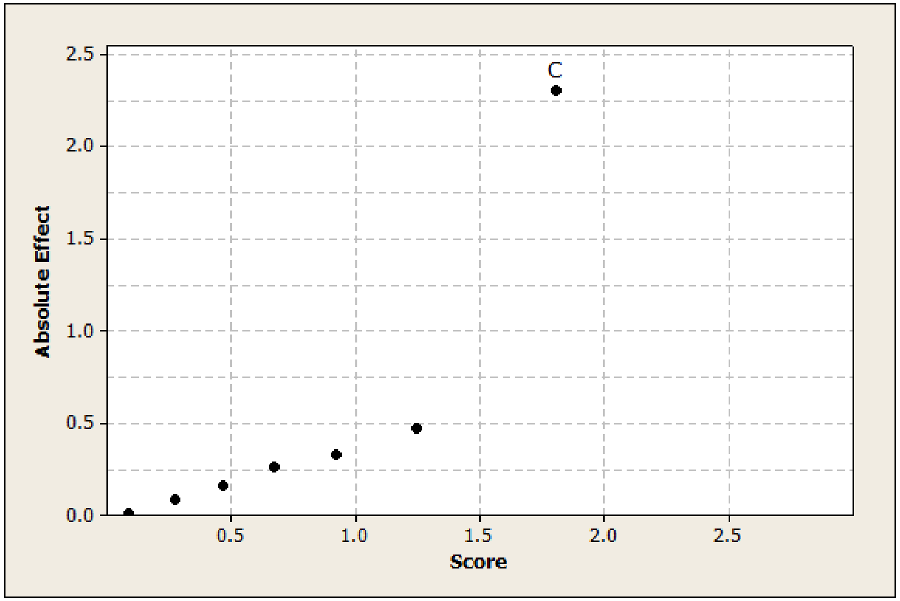
Letters are used here to denote the factors

$C \equiv x_3$



Interpretation

- We are hoping to see a group of points which 'fit' a straight line pointing to the origin.
- .... and another group which lie to the right and clearly above the line.
- The second group are taken to be real effects.

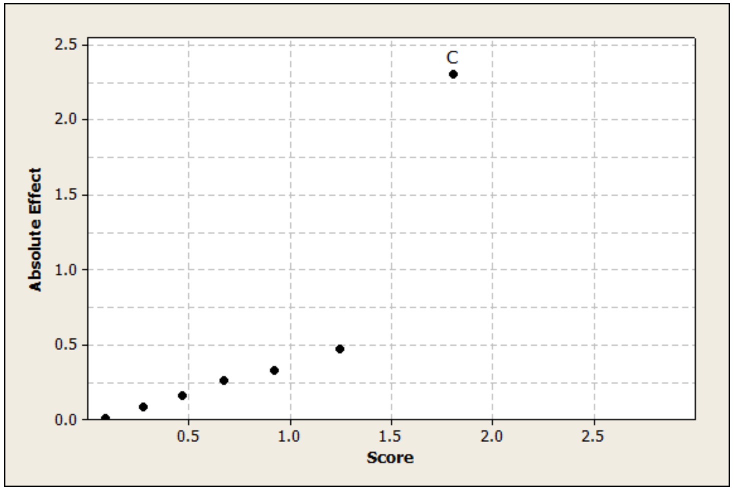


In this case the x_3 linear effect is the only detectable effect

The logic of the plot

- If there were no real effects, the points would on average form a straight line with gradient equal to the SE of the effects (Box, Hunter & Hunter, Chapter 2).
- For example, if there were no real effects the 5th largest effect would on average equal;

$$(5^{th}\ score) \times (SE\ of\ effects)$$

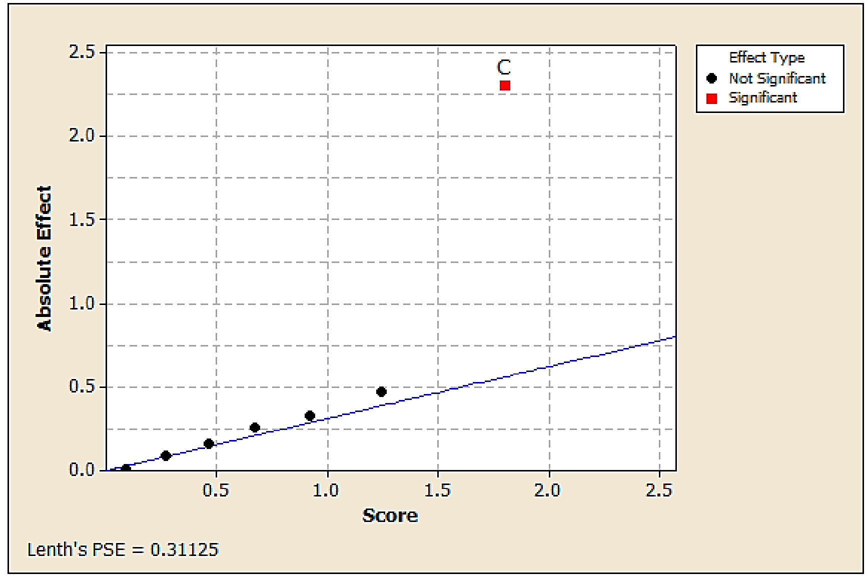


- If we have a combination of random effects and large real effects, the real effects would all plot to the right and above the line
- In practice the real and random effects will be mixed together but we hope that most of those 'off the line' are real.

The Logic of The Plot

There are two schools of thought:

- The plot is self-calibrating: a line drawn by eye through the smaller effects tells us where to expect the larger effects if everything is random.
  - in doubtful cases look for the 'lowest plausible line'.
- Calibrate the plot using Lenth's pseudo-standard error (PSE).
  - draw a line through the origin with gradient equal to the PSE.



Rather than use an 'eyeball' test, it is possible to carry out 'pseudo t tests' based on the PSE – in the plot above we have used this method to select C.

Lenth's PSE

	x_2	x_1x_3	$x_1x_2x_3$	x_2x_3	x_1x_2	x_1	x_3
Abs. effect	0.01	0.09	0.16	0.26	0.33	0.47	2.30

- Find the median of the absolute effects (0.26);
- Discard all effects $> \ 3.75 \ \times \ median,$

  $3.75 \ \times 0.26 = 0.966,$ so discard 2.30;

- PSE $= 1.5 \ \times \ median$ of the remaining effects,

$$1.5 \times \frac{1}{2} \left(0.16 + 0.26\right) = 0.31125$$

Comments On The Half Normal Plot

- Like all statistical methods, using the plot to identify real effects can result in;
  - false positives (Type I error): points are off the line purely by chance.
  - false negatives (Type II error): small real effects that cannot be distinguished from random effects.
- The power (sensitivity) of the plot can be increased by:
  - increasing the size of the experiment.
  - reducing random variation (better experimental technique).
- The role of the checking effects: if we have fitted the right form of surface equation these effects should be purely random.
  - if so, they help us to calibrate the plot.
  - but if they are ‘off the line’, we have not got the right equation!

A Larger Experiment On The Moulded Bumper Brackets

- We introduced this 2^{5-1} design in TS4, and we will now analyse the results of an experiment.

Factor	Name	Units	–1	+1
x_1	Melt temp	℃	230	270
x_2	Screw speed	rpm	50	300
x_3	Hold pressure	MPa	50	250
x_4	Hold time	sec	15	35
x_5	Injection rate	mm/sec	50	300

x_1	x_2	x_3	x_4	x_5	Strength (MPa)
–1	–1	–1	–1	+1	41.23
+1	–1	–1	–1	–1	37.72
–1	+1	–1	–1	–1	39.15
+1	+1	–1	–1	+1	42.61
–1	–1	+1	–1	–1	41.27
+1	–1	+1	–1	+1	43.50
–1	+1	+1	–1	+1	43.05
+1	+1	+1	–1	–1	41.51
–1	–1	–1	+1	–1	38.62
+1	–1	–1	+1	+1	42.71
–1	+1	–1	+1	+1	41.61
+1	+1	–1	+1	–1	37.94
–1	–1	+1	+1	+1	43.02
+1	–1	+1	+1	–1	40.88
–1	+1	+1	+1	–1	42.11
+1	+1	+1	+1	+1	44.00

Sample mean from 3 parts

Fitting The Response Surface

- The 2nd order equation has 16 terms: the intercept, 5 linear terms, 10 2fi's.
- We can estimate the coded regression coefficients by halving the effects:

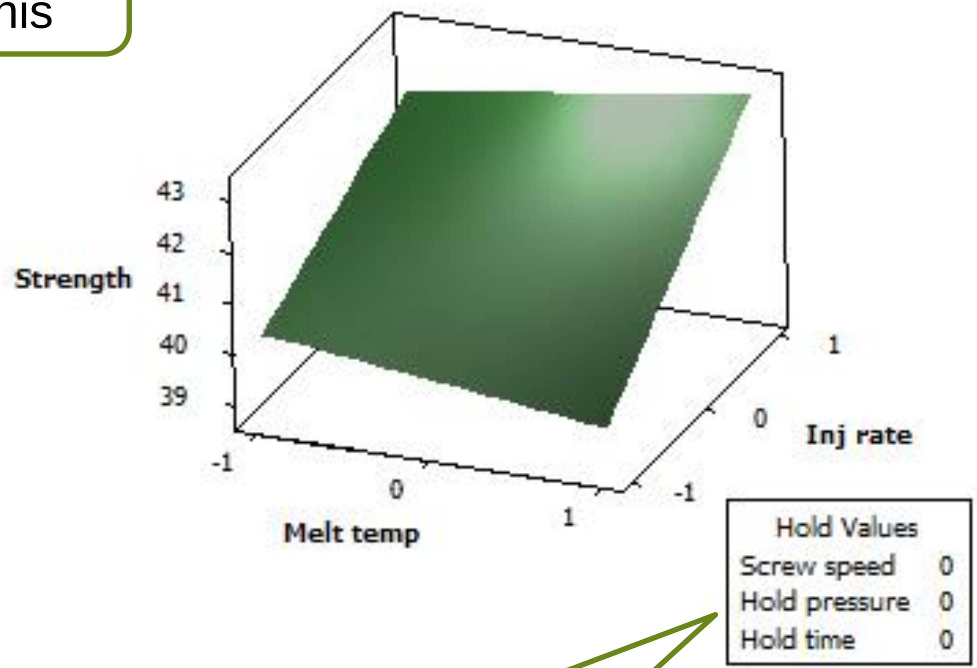
	x_1	x_2	x_3	x_4	x_5	x_1x_2	x_1x_3	x_1x_4	x_1x_5
Ave (+)	41.36	41.50	42.42	41.36	42.72	41.28	41.31	41.28	41.75
Ave (-)	41.26	41.12	40.20	41.26	39.90	41.34	41.30	41.34	40.87
Effect	0.10	0.38	2.22	0.11	2.82	-0.07	0.01	-0.06	0.88
Coefficient	0.05	0.19	1.11	0.06	1.41	-0.04	0.01	-0.03	0.44

	x_2x_3	x_2x_4	x_2x_5	x_3x_4	x_3x_5	x_4x_5
Ave (+)	41.37	41.17	41.22	41.34	40.88	41.37
Ave (-)	41.25	41.44	41.40	41.28	41.74	41.24
Effect	0.12	-0.27	-0.18	0.06	-0.87	0.13
Coefficient	0.06	-0.14	-0.09	0.03	-0.44	0.07

The intercept is the overall mean strength (41.31)

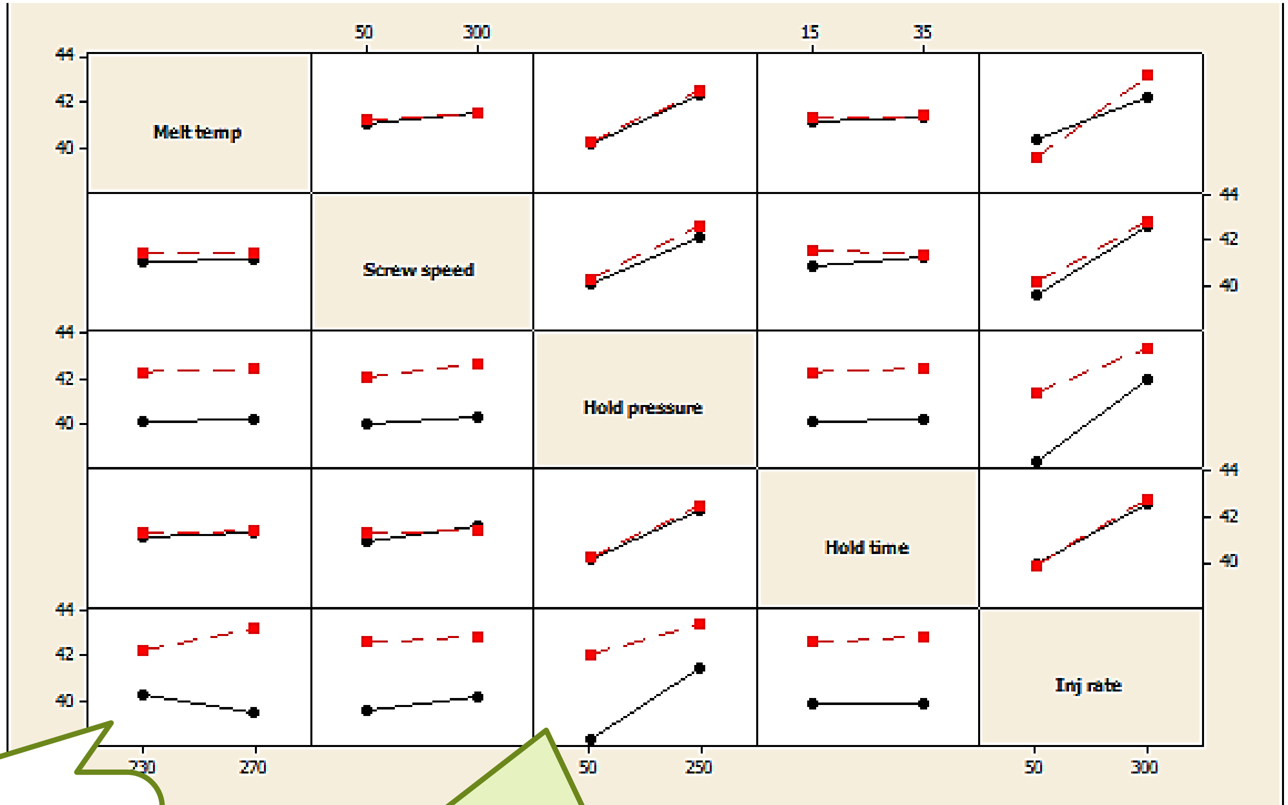
Plotting The Response Surface

There are 9 other plots like this



3 factors held constant at coded 0

Interaction Plots



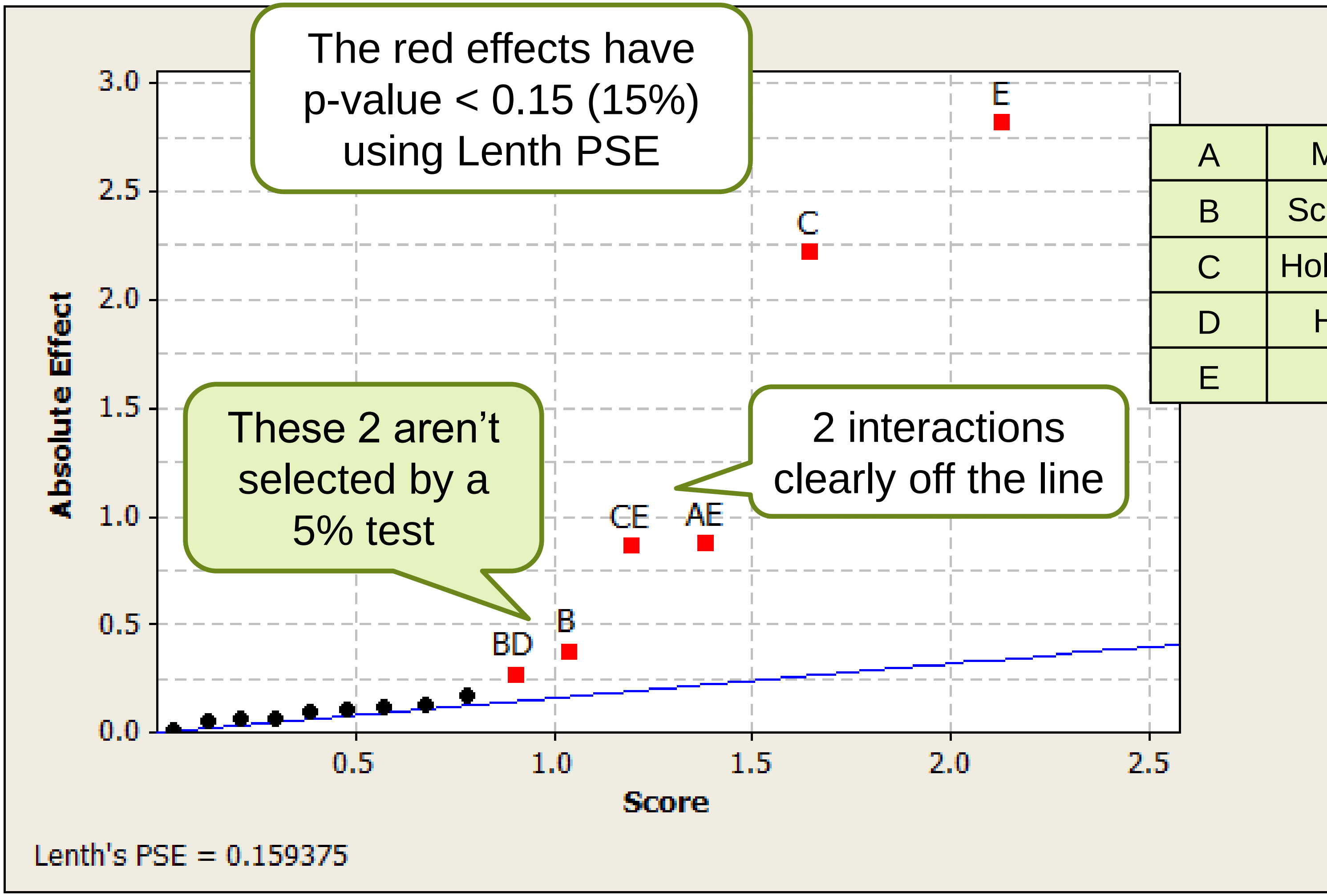
Interaction between Melt temp and Injection rate

Interaction between Hold pressure and Injection rate

ANOVA Table, R^2, etc

- We have fitted 16 terms to 16 runs, so in this experiment **ResDF is 0.**
- We cannot construct an ANOVA table or calculate R^2.
- The residuals are all 0.
- The leverages are all exactly 1, so the leave-one-out error $\frac{\text{residual}}{(1-\text{leverage})}$ is undefined.
- We cannot use *t* tests to identify the important effects.

Selecting The Important Factors

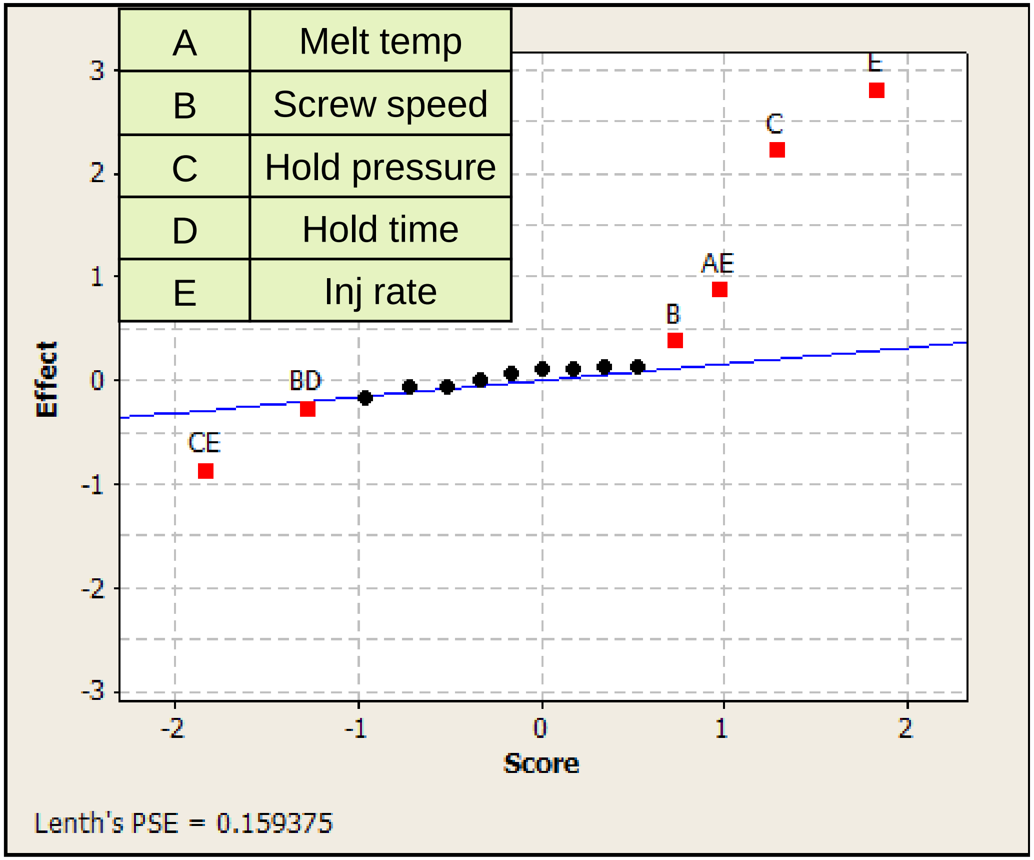


A	Melt temp
B	Screw speed
C	Hold pressure
D	Hold time
E	Inj rate

Full Normal Plot

- We can use an ordinary Normal plot instead of the Half Normal.
- Most analysts prefer the Half Normal because all the potentially significant effects are plotted at the same end.

Full Normal plots can be useful for spotting outliers in screening experiments



In This Session We Have…

- Used the 2^3 design to explain the construction and interpretation of the Half Normal plot.
- Explained that a Half Normal plot can include checking effects such as the three-factor interaction.
- Compared two methods of calibrating the plot:
  - by eye,
  - using Lenth's PSE.
- Applied these methods to a 2^{5-1} design.
- Compared Half Normal and Full Normal plots.

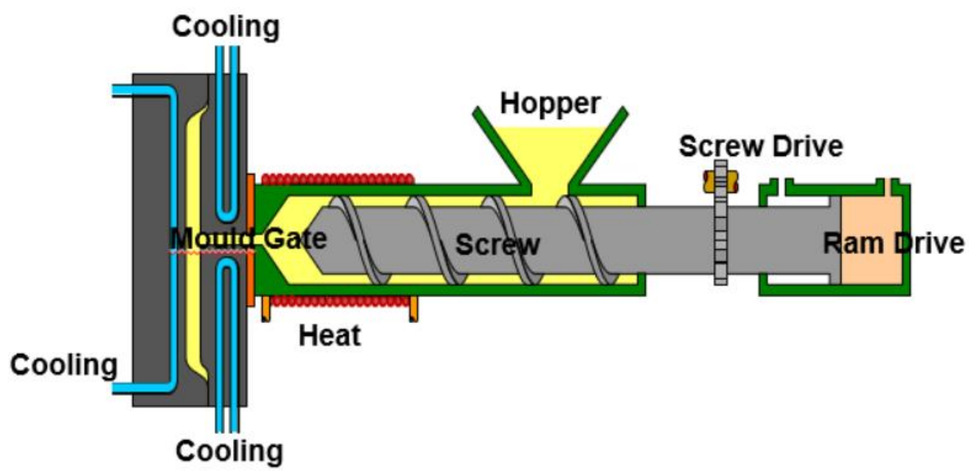
Session 7: Analysing a Two Level Experiment II

Tutorial and Exercise

Tutorial

- **Session TS06+07: Analysing a Two Level Experiment**
- **Objectives**

  Analyse a two level screening experiment.
- **Engineering Scenario**
- A study of potential alternative materials for bumper bracket mouldings.
- The brackets are injection moulded (IM).
- In our example, the factors are settings on the IM machine.



Tutorial

- **Session TS06+07: Analysing a Two Level 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. **Continue from Tutorial 07, Task 4.0.**

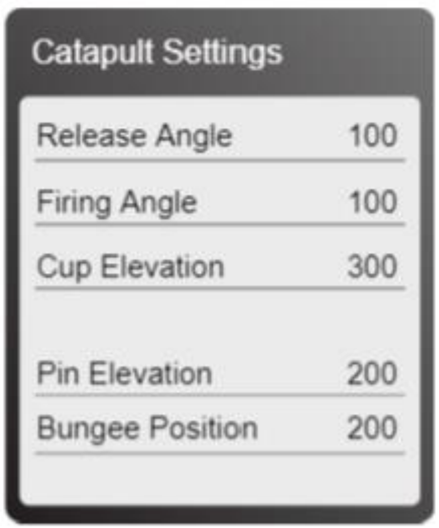
- **Tutorial Task**

  1. Generate a 2^3 coded design.
  2. Export the design to add the response of the system and then import the edited file.
  3. Generate regression output from the 2^3 design.
  4. Convert the coded design to uncoded units
  5. Examine main effect plots and analyse interaction plots
  6. Examine Plot of the response surface.
  7. Generate and analyse a half normal plot

Exercise

- **Session TS06+07: Analysing a Two Level Experiment**
- **Objective**
- To analyse and discuss the results of a screening experiment for the Virtual Catapult.

https://sigmazone.com/catapult/



Exercise

- **Session TS06+07: Analysing 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 from Exercise 07, Task 6.**
- **Guidelines**

  Use the relevant cells from the Python Colab notebook in Tutorials 06 and 07 to analyse your experiment using regression to fit the appropriate (first order or first order with two-factors interactions) model, and draw conclusions about which factors are significant and you would consider for inclusion in a follow-up 3-level experiment.