IE 361 Module 50
Design and Analysis of Experiments Part 10
(Fractional Factorial Studies With 2-Level Factors)

Reading: Section 6.1 Statistical Methods for Quality Assurance

ISU and Analytics Iowa LLC
In this module we discuss what can be done in a many-factor context where the number of possible combinations of levels of $p$ factors is so large that doing a complete factorial experiment is not practically possible. For $p$ factors, even $2^p$ gets big fast, for example.

$$\text{for } p = 10, \quad 2^p = 1024$$

So in practice, for $p$ of any size, one must make do with information from only some fraction of a complete factorial in $p$ factors.

We consider rational approaches to design and analysis of fractional factorial studies, and here limit our discussion to fractions of $2^p$ studies. We will first consider half fractions of these, and then other fractions that are powers of $1/2$. 
A 15-factor chemical experiment had factors and levels as

<table>
<thead>
<tr>
<th>Factor</th>
<th>Levels</th>
<th>Factor</th>
<th>Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-Coating Roll Temp</td>
<td>115° vs 125°</td>
<td>J-Feed Air to Dryer Preheat</td>
<td>Yes vs No</td>
</tr>
<tr>
<td>B-Solvent</td>
<td>Recycled vs Refined</td>
<td>K-Dibutylfutile in Formula</td>
<td>12% vs 15%</td>
</tr>
<tr>
<td>C-Polymer X-12 Preheat</td>
<td>No vs Yes</td>
<td>L-Surfactant in Formula</td>
<td>.5% vs 1%</td>
</tr>
<tr>
<td>D-Web Type</td>
<td>LX-14 vs LB-17</td>
<td>M-Dispersant in Formula</td>
<td>.1% vs .2%</td>
</tr>
<tr>
<td>E-Coating Roll Tension</td>
<td>30 vs 40</td>
<td>N-Wetting Agent in Formula</td>
<td>1.5% vs 2.5%</td>
</tr>
<tr>
<td>F-Number of Chill Roll</td>
<td>1 vs 2</td>
<td>O-Time Lapse</td>
<td>10min vs 30min</td>
</tr>
<tr>
<td>G-Drying Roll Temp</td>
<td>75° vs 80°</td>
<td>P-Mixer Agitation Speed</td>
<td>100rpm vs 250rpm</td>
</tr>
<tr>
<td>H-Humidity of Air Feed</td>
<td>75% vs 90%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

and response variable

\[ y = \text{a measure of product cold crack resistance} \]

A full factorial would require at least \(2^{15} = 32,768\) runs of the process! An obvious "solution" is to collect data for only some (a fraction) of all possible combinations of levels of the factors.
Qualitative points that ought to be “obvious” \textit{a priori} if a fraction of a factorial is used as an experimental design are that

- there must be some information loss (relative to the full factorial),
- some ambiguity must inevitably follow because of the loss, and
- careful planning and wise analysis are needed to hold these to a minimum.
"Obvious" Limitations and Objective

Example 50-2 (Hypothetical/unrealistic but instructive half fraction of a 2X2 Factorial)

The full $2^2$ factorial structure is shown in below. Here, if combination (1) is used, combination ab must be employed or else one learns nothing about the action of one of the factors. (If a is used, then b must be employed or one learns nothing about the action of one of the factors.) On the other hand, if there is a big difference in response between the two combinations included in the half fraction, one doesn’t know whether to attribute this to one or the other or to the interaction effect of the factors A and B.

**Figure**: A $2^2$ Full Factorial Layout
Potential and Problems

Example 50-3 (Hypothetical)

Suppose that $2^3$ factorial effects are

\[
\begin{align*}
\mu_{...} &= 10, \quad \alpha_2 = 3, \quad \beta_2 = 1, \quad \gamma_2 = 2, \quad \alpha\beta_{22} = 2, \\
\alpha\gamma_{22} &= 0, \quad \beta\gamma_{22} = 0, \quad \alpha\beta\gamma_{222} = 0
\end{align*}
\]

The corresponding means are then as pictured below.

**Figure:** A $2^3$ Factorial With a "Good" Half Fraction Indicated With Circled Corners
Suppose further that one gets data adequate to essentially reveal the mean responses for combinations a, b, c and abc (the 4 corners circled on panel 6) but has no data on the other combinations. How then might one try to assess the $2^3$ factorial effects if presented with information from the circled corners on panel 6?

Remember that with the $2^3$ as pictured on panel 6,

$$\alpha_2 = \text{"right face average"} - \text{"grand average"} = 3$$

A "half-fraction version" of this might be

$$\alpha^*_2 = \text{"available right face average"} - \text{"available grand average"} = 13 - 10 = 3$$

Here $\alpha^*_2 = \alpha_2$!!! Is this something for nothing? Can we learn about the A main effect using data from only "4 corners"?
But note that a similar calculation for the C main effect gives
\[ \gamma_2^* = " \text{available back face average} - \text{available grand average} \]
\[ = 14 - 10 = 4 \]
and \[ 4 = \gamma_2^* \neq \gamma_2 = 2 \quad ??? \]

The general story here is that for this \( 2^{3-1} \) fractional factorial
\[ \alpha_2^* = \alpha_2 + \beta \gamma_{22} \quad \text{and} \quad \gamma_2^* = \gamma_2 + \alpha \beta_{22} \]

We were able to "recover" \( \alpha_2 \) using only \( \alpha_2^* \) (based on only half of the corners of the cube) because \( \beta \gamma_{22} = 0 \). We were unable to "recover" \( \gamma_2 \) using only \( \gamma_2^* \) (based on only half of the corners of the cube) because \( \alpha \beta_{22} = 2 \neq 0 \). We can really only know \( \alpha_2 + \beta \gamma_{22} \) (and not \( \alpha_2 \) alone) and \( \gamma_2 + \alpha \beta_{22} \) (and not \( \gamma_2 \) alone) based on the half fraction. This is an example of confounding/aliasing/ambiguity that of necessity comes with use of only a fractional factorial data collection plan.
As a bit of an aside at this point, notice that the choice of 4 corners on panel 6 has admirable symmetry. If one collapses the cube in any direction one is left with a complete factorial arrangement in the two other factors. That means that if, in fact, one of the 3 factors is "inert" doing nothing to affect response, one has a full factorial in the 2 factors that matter.