$2^{k-p}$ Fractional Factorial Design

Wu & Hamada: Chapter 4
Montgomery: Chapter 8
Fundamental Principles for Factorial Effects

Suppose there are \( k \) factors \((A,B,...,J,K)\) in an experiment. All possible factorial effects include

- effects of order 1: \( A, B, ..., K \) (main effects)

- effects of order 2: \( AB, AC, ..., JK \) (2-factor interactions)


- Hierarchical Ordering principle
  - Lower order effects are more likely to be important than high order effects.
  - Effects of the same order are equally likely to be important

- Effect Sparsity Principle (Pareto principle)
  - The number of relatively important effects in a factorial experiment is small

- Effect Heredity Principle
  - In order for an interaction to be significant, at least one of its parent factors should be significant.
Fractional Factorials

• May not have sources for full factorial design

• Number of runs required for full factorial grows quickly
  – Consider $2^k$ design
  – If $k = 7$ → 128 runs required
  – Can estimate 127 effects
  – Only 7 df for main effects, 21 for 2-factor interactions
  – 99 df are for interactions of order $\geq 3$

• Often low order effects are important

• Full factorial design may not be necessary based on
  – Hierarchical ordering principle
  – Effect Sparsity Principle

• A fraction of the full factorial design (i.e. a subset of all possible treatment combinations) is sufficient.

  Fractional Factorial Design
Example 1

• Suppose you were designing a new car

• Wanted to consider the following factors each with 2 levels
  – Engine Size
  – Number of cylinders
  – Drag
  – Weight
  – Automatic vs Manual
  – Shape
  – Tires
  – Suspension
  – Gas Tank Size

• Only have resources for conduct $2^6 = 64$ runs
  – If you drop three factors for a $2^6$ full factorial design, those factor and their interactions with other factors cannot be investigated.

  – Want investigate all nine factors in the experiment

  – A fraction of $2^9$ full factorial design will be used.

  – Confounding (aliasing) will happen because using a subset

    How to choose (or construct) the fraction?
Example 2

Filtration rate experiment:

Recall that there are four factors in the experiment (A, B, C and D), each of 2 levels. Suppose the available resource is enough for conducting eight 8 runs. $2^4$ full factorial design consists of all the 16 treatment combinations of the four factors. We need to choose half of them.

$2^4$ Full Factorial Design

<table>
<thead>
<tr>
<th>factor</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

The chosen half is called $2^{4-1}$ fractional factorial design. Which half we should select (construct)?
2⁴⁻¹ Fractional Factorial Design

- the number of factors: 4
- the fraction index: 1
- the number of runs (treatment combinations): \( \frac{2^4}{2^1} = 8 \)
- Construct 2^4⁻¹ designs:
  - select 3 factors (e.g. A, B, C) to form a 2^3 full factorial (basic design)
  - confound (alias) with a high order interaction of A, B and C. For example,

\[
D = ABC
\]

<table>
<thead>
<tr>
<th>I</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>AB</th>
<th>AC</th>
<th>BC</th>
<th>ABC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Verify:
1. columns A, B, C, and D form a half of 2^4 design (note that -1 and -, 1 and + are interchangeable.).
2. the product of column A, B, C and D equals 1, i.e.,

\[
I = ABCD
\]

which is called the defining relation, or \( ABCD \) is called a defining word (contrast, effect).
Aliasing in $2^{4-1}$ Design

For four factors $A$, $B$, $C$ and $D$, there are $2^4 - 1$ effects: $A$, $B$, $C$, $D$, $AB$, $AC$, $AD$, $BC$, $BD$, $CD$, $ABC$, $ABD$, $ACD$, $BCD$, $ABCD$

<table>
<thead>
<tr>
<th>Response</th>
<th>I</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>AB</th>
<th>CD</th>
<th>ABC</th>
<th>...</th>
<th>ABCD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y_1$</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>..</td>
<td>1</td>
<td>-1</td>
<td>...</td>
<td>1</td>
</tr>
<tr>
<td>$y_2$</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>..</td>
<td>-1</td>
<td>1</td>
<td>...</td>
<td>1</td>
</tr>
<tr>
<td>$y_3$</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>..</td>
<td>-1</td>
<td>1</td>
<td>...</td>
<td>1</td>
</tr>
<tr>
<td>$y_4$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>..</td>
<td>1</td>
<td>...</td>
<td>1</td>
</tr>
<tr>
<td>$y_5$</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>..</td>
<td>1</td>
<td>1</td>
<td>...</td>
<td>1</td>
</tr>
<tr>
<td>$y_6$</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>..</td>
<td>-1</td>
<td>1</td>
<td>...</td>
<td>1</td>
</tr>
<tr>
<td>$y_7$</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>..</td>
<td>-1</td>
<td>1</td>
<td>...</td>
<td>1</td>
</tr>
<tr>
<td>$y_8$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>..</td>
<td>1</td>
<td>1</td>
<td>...</td>
<td>1</td>
</tr>
</tbody>
</table>

Contrasts for main effects by $-$ to $-1$ and $+$ to $1$; contrasts for other effects by multiplication.

$$A = \bar{y}_{A+} - \bar{y}_{A-} = \frac{1}{4}(-y_1 + y_2 - y_3 + y_4 - y_5 + y_6 - y_7 + y_8)$$

$$BCD = \frac{1}{4}(-y_1 + y_2 - y_3 + y_4 - y_5 + y_6 - y_7 + y_8)$$

$A$, $BCD$ are aliases or aliased. The contrast is for $A+BCD$, $A$ and $BCD$ are not distinguishable.

$$AB = \bar{y}_{AB+} - \bar{y}_{AB-} = \frac{1}{4}(y_1 - y_2 - y_3 + y_4 + y_5 - y_6 - y_7 + y_8)$$

$$CD = \bar{y}_{CD+} - \bar{y}_{CD-} = \frac{1}{4}(y_1 - y_2 - y_3 + y_4 + y_5 - y_6 - y_7 + y_8)$$

$AB$, $CD$ are aliases or aliased. The contrast is for $AB+CD$, $AB$ and $CD$ are not distinguishable.

There are other 5 pairs, all of which are caused by the defining relation

$$I = ABCD$$

that says, $I$ and 4-factor interaction $ABCD$ are aliased because of the design.

1-6
**Alias Structure in** $2^{4-1}$ **with** $I = ABCD$

- **Alias Structure:**
  
  $I = ABCD$
  
  $A = A \ast I = A \ast ABCD = BCD$
  
  $B = \ldots \ldots = AC\, D$
  
  $C = \ldots \ldots = AB\, D$
  
  $D = \ldots \ldots = ABC$
  
  $AB = AB \ast I = AB \ast ABCD = CD$
  
  $AC = \ldots \ldots = BD$
  
  $AD = \ldots \ldots = BC$

  all 16 factorial effects for $A, B, C$ and $D$ are partitioned into 8 groups each with 2 aliased effects.

- **Another defining relation:** $d_2$: $I = ABD$, it implies the following alias structure:

  $I = ABD$
  
  $A = BD$
  
  $B = AD$
  
  $C = ABCD$
  
  $D = AB$
  
  $ABC = CD$
  
  $ACD = BC$
  
  $BCD = AC$

- **Let** $d_1$ **be** $I = ABCD$. **Comparing** $d_1$ **and** $d_2$, **which one is better?**

  1. Length of a defining word is defined to be the number of factors it consists of
  
  2. Resolution is defined to be the minimum length of defining words, usually denoted by Roman numbers, III, IV, V, etc...
Resolution and Maximum Resolution Criterion

• $d_1$: $I = ABCD$ is a resolution IV design denoted by $2^4_{IV}$.

• $d_2$: $I = ABC$ is a resolution III design denoted by $2^4_{III}$.

• A design is of resolution $R$ if no $i$-factor effect is aliased with another effect containing less than $R - i$ factors.

  $d_1$: main effects are not aliased with other main effects and 2-factor interactions
  $d_2$: main effects are not aliased with main effects

  $d_1$ is better. In fact, $d_1$ is optimal.

• Maximum Resolution Criterion
  fractional factorial design with maximum resolution is optimal
Analysis for $2^{4-1}$ Design: Filtration Experiment

Recall that the filtration rate experiment was originally a $2^4$ full factorial experiment. We pretend that only half of the combinations were run. It became a $2^{4-1}$ fractional factorial design. Suppose the defining relation is $I = ABCD$.

<table>
<thead>
<tr>
<th>basic design</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D = ABC</th>
<th>filtration rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>45</td>
</tr>
<tr>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>100</td>
</tr>
<tr>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>45</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>65</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>75</td>
</tr>
<tr>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>60</td>
</tr>
<tr>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>80</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>96</td>
</tr>
</tbody>
</table>

Let $\mathcal{L}_{\text{effect}}$ denote the estimate for the effect (based on the corresponding contrast). Due to aliasing,

\[
\begin{align*}
\mathcal{L}_I & \rightarrow I + ABCD \\
\mathcal{L}_A & \rightarrow A + BCD \\
\mathcal{L}_B & \rightarrow B + ACD \\
\mathcal{L}_C & \rightarrow C + ABD \\
\mathcal{L}_D & \rightarrow D + ABC \\
\mathcal{L}_{AB} & \rightarrow AB + CD \\
\mathcal{L}_{AC} & \rightarrow AC + BD \\
\mathcal{L}_{AD} & \rightarrow AD + BC
\end{align*}
\]
SAS file for $2^{4-1}$ Filtration Experiment

goption colors=(none);

data filter;
do C = -1 to 1 by 2;
do B = -1 to 1 by 2; do A = -1 to 1 by 2; D=A*B*C;
input y @@; output; end; end; end;
datalines;
45 100 45 65 75 60 80 96
;
data inter;    /* Define Interaction Terms */
set filter;
AB=A*B; AC=A*C; AD=A*D;

proc glm data=inter;    /* GLM Proc to Obtain Effects */
class A B C D AB AC AD;
model y=A B C D AB AC AD;
estimate 'A' A -1 1; estimate 'B' B -1 1; estimate 'C' C -1 1;
estimate 'D' D -1 1; estimate 'AB' AB -1 1; estimate 'AC' AC -1 1;
estimate 'AD' AD -1 1; run;

proc reg outest=effects data=inter;    /* REG Proc to Obtain Effects */
model y=A B C D AB AC AD;

data effect2; set effects;
drop y intercept _RMSE_
proc transpose data=effect2 out=effect3;
data effect4; set effect3; effect=col1*2;
proc sort data=effect4; by effect;
proc print data=effect4;
proc rank data=effect4 normal=blom;
var effect; ranks neff;
symbol1 v=circle;
proc gplot;
plot effect*neff=_NAME_;
run;

1-10
### SAS Output

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>7</td>
<td>3071.500000</td>
<td>438.785714</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>Error</td>
<td>0</td>
<td>0.000000</td>
<td>.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CoTotal</td>
<td>7</td>
<td>3071.500000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Type I SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>1</td>
<td>4.500000</td>
<td>4.500000</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>392.000000</td>
<td>392.000000</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>D</td>
<td>1</td>
<td>544.500000</td>
<td>544.500000</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>AB</td>
<td>1</td>
<td>2.000000</td>
<td>2.000000</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>AC</td>
<td>1</td>
<td>684.500000</td>
<td>684.500000</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>AD</td>
<td>1</td>
<td>722.000000</td>
<td>722.000000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Obs</th>
<th><em>NAME</em></th>
<th>COL1</th>
<th>effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AC</td>
<td>-9.25</td>
<td>-18.5</td>
</tr>
<tr>
<td>2</td>
<td>AB</td>
<td>-0.50</td>
<td>-1.0</td>
</tr>
<tr>
<td>3</td>
<td>B</td>
<td>0.75</td>
<td>1.5</td>
</tr>
<tr>
<td>4</td>
<td>C</td>
<td>7.00</td>
<td>14.0</td>
</tr>
<tr>
<td>5</td>
<td>D</td>
<td>8.25</td>
<td>16.5</td>
</tr>
<tr>
<td>6</td>
<td>A</td>
<td>9.50</td>
<td>19.0</td>
</tr>
<tr>
<td>7</td>
<td>AD</td>
<td>9.50</td>
<td>19.0</td>
</tr>
</tbody>
</table>
Let $x_A, x_C, x_D$ be the variables for factor $A, C$ and $D$. A possible regression model is

$$y = 70.75 + 9.50x_A + 7.00x_C + 8.25x_D - 9.25x_Ax_C + 9.50x_Ax_D$$
Aliased effects and Techniques for Resolving the Ambiguities

The estimates are for the sum of aliased factorial effects.

\[ \mathcal{L}_I = 70.75 \rightarrow I + ABCD \]
\[ \mathcal{L}_A = 19.0 \rightarrow A + BCD \]
\[ \mathcal{L}_B = 1.5 \rightarrow B + ACD \]
\[ \mathcal{L}_C = 14.0 \rightarrow C + ABD \]
\[ \mathcal{L}_D = 16.5 \rightarrow D + ABC \]
\[ \mathcal{L}_{AB} = -1.0 \rightarrow AB + CD \]
\[ \mathcal{L}_{AC} = -18.5 \rightarrow AC + BD \]
\[ \mathcal{L}_{AD} = 19.0 \rightarrow AD + BC \]

Techniques for resolving the ambiguities in aliased effects

- Use the fundamental principles (Slide 1)
- Follow-up design
  - add orthogonal runs
  - optimal design approach
  - fold-over design
- Sequential design
  In the $2^{4-1}$ filtration experiment, we have used a half of all 16 runs defined by $I = ABCD$. The remaining 8 runs are indeed defined by the following relationship
  \[ D = -ABC, \text{ or } I = -ABCD \]
Sequential Experiment

<table>
<thead>
<tr>
<th>basic design</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D = −ABC</th>
<th>filtration rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>−</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>43</td>
<td></td>
</tr>
<tr>
<td>+</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td>−</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td>+</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>104</td>
<td></td>
</tr>
<tr>
<td>−</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>68</td>
<td></td>
</tr>
<tr>
<td>+</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>86</td>
<td></td>
</tr>
<tr>
<td>−</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>+</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td>65</td>
<td></td>
</tr>
</tbody>
</table>

\( I = −ABCD \) implies that: \( A = −BCD, B = −ACD, \ldots, AB = −CD \) …

Similarly, we can derive the following estimates (\( \tilde{\mathcal{L}}_{\text{effect}} \)) and aliasing structure

\[
\begin{align*}
\tilde{\mathcal{L}}_I &= 69.375 \quad \rightarrow \quad I - ABCD \\
\tilde{\mathcal{L}}_A &= 24.25 \quad \rightarrow \quad A - BCD \\
\tilde{\mathcal{L}}_B &= 4.75 \quad \rightarrow \quad B - ACD \\
\tilde{\mathcal{L}}_C &= 5.75 \quad \rightarrow \quad C - ABD \\
\tilde{\mathcal{L}}_D &= 12.75 \quad \rightarrow \quad D - ABC \\
\tilde{\mathcal{L}}_{AB} &= 1.25 \quad \rightarrow \quad AB - CD \\
\tilde{\mathcal{L}}_{AC} &= −17.75 \quad \rightarrow \quad AC - BD \\
\tilde{\mathcal{L}}_{AD} &= 14.25 \quad \rightarrow \quad AD - BC
\end{align*}
\]


**Combine Sequential Experiments**

Combining two experiments $\Rightarrow 2^4$ full factorial experiment
Combining the estimates from these two experiments $\Rightarrow$ estimates based on the full experiment

$$\mathcal{L}_A = 19.0 \rightarrow A + BCD$$

$$\mathcal{L}_A = 24.25 \rightarrow A - BCD$$

$$A = \frac{1}{2}(\mathcal{L}_A + \mathcal{L}_A) = 21.63$$

$$ABC = \frac{1}{2}(\mathcal{L}_A - \mathcal{L}_A) = -2.63$$

Other effects are summarized in the following table

<table>
<thead>
<tr>
<th>$i$</th>
<th>$\frac{1}{2}(\mathcal{L}_i + \mathcal{L}_i)$</th>
<th>$\frac{1}{2}(\mathcal{L}_i - \mathcal{L}_i)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A$</td>
<td>21.63 $\rightarrow$ $A$</td>
<td>-2.63 $\rightarrow$ $BCD$</td>
</tr>
<tr>
<td>$B$</td>
<td>3.13 $\rightarrow$ $B$</td>
<td>-1.63 $\rightarrow$ $ACD$</td>
</tr>
<tr>
<td>$C$</td>
<td>9.88 $\rightarrow$ $C$</td>
<td>4.13 $\rightarrow$ $ABD$</td>
</tr>
<tr>
<td>$D$</td>
<td>14.63 $\rightarrow$ $D$</td>
<td>1.88 $\rightarrow$ $ABC$</td>
</tr>
<tr>
<td>$AB$</td>
<td>1.13 $\rightarrow$ $AB$</td>
<td>-1.13 $\rightarrow$ $CD$</td>
</tr>
<tr>
<td>$AC$</td>
<td>-18.13 $\rightarrow$ $AC$</td>
<td>-0.38 $\rightarrow$ $BD$</td>
</tr>
<tr>
<td>$AD$</td>
<td>16.63 $\rightarrow$ $AD$</td>
<td>2.38 $\rightarrow$ $BC$</td>
</tr>
</tbody>
</table>

We know the combined experiment is not a completely randomized experiment. Is there any underlying factor we need consider? what is it?
General $2^{k-1}$ Design

- $k$ factors: $A, B, \ldots, K$

- can only afford half of all the combinations ($2^{k-1}$)

- Basic design: a $2^{k-1}$ full factorial for $k - 1$ factors: $A, B, \ldots, J$.

- The setting of $k$th factor is determined by alasing $K$ with the $ABC\ldots J$, i.e., $K = ABC \cdots JK$

- Defining relation: $I = ABCD\ldots JK$. Resolution=V

- $2^k$ factorial effects are partitioned into $2^{k-1}$ groups each with two aliased effects.

- only one effect from each group (the representative) should be included in ANOVA or regression model.

- Use fundamental principles, domain knowledge, follow-up experiment to de-alias.
One Quarter Fraction: $2^{k-2}$ Design

Parts manufactured in an injection molding process are showing excessive shrinkage. A quality improvement team has decided to use a designed experiment to study the injection molding process so that shrinkage can be reduced. The team decides to investigate six factors

- A: mold temperature
- B: screw speed
- C: holding time
- D: cycle time
- E: gate size
- F: holding pressure

each at two levels, with the objective of learning about main effects and interactions.

They decide to use 16-run fractional factorial desing.

- a full factorial has $2^6=64$ runs.
- 16-run is one quarter of the full factorial
- How to construct the fraction?
Injection Molding Experiment: \(2^{6-2}\) Design

<table>
<thead>
<tr>
<th>basic design</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E = ABC</th>
<th>F = BCD</th>
<th>shrinkage</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>6</td>
</tr>
<tr>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>10</td>
</tr>
<tr>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>32</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>60</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>4</td>
</tr>
<tr>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>15</td>
</tr>
<tr>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>26</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>60</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>8</td>
</tr>
<tr>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>12</td>
</tr>
<tr>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>34</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>60</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>16</td>
</tr>
<tr>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>5</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>37</td>
</tr>
</tbody>
</table>

Two defining relations are used to generate the column for \(E\) and \(F\).

\[ I = ABCE, \text{ and } I = BCDF \]

They induce another defining relation:

\[ I = ABCE \ast BCDF = AB^2C^2DEF = ADEF \]

The complete defining relation:

\[ I = ABCE = BCDF = ADEF \]

Defining contrasts subgroup: \(\{I, ABCE, BCDF, ADEF\}\)
Alias Structure for $2^{6-2}$ with

$I = ABCE = BCDF = ADEF$

$I = ABCE = BCDF = ADEF$ implies $A = BCE = ABCDF = ADEF$

Similarly, we can derive the other groups of aliased effects.

<table>
<thead>
<tr>
<th>A = BCE = DEF = ABCDF</th>
<th>AB = CE = ACDF = BDEF</th>
</tr>
</thead>
<tbody>
<tr>
<td>B = ACE = CDF = ABDEF</td>
<td>AC = BE = ABDF = CDEF</td>
</tr>
<tr>
<td>C = ABE = BDF = ACD</td>
<td>AD = EF = BCDE = ABCF</td>
</tr>
<tr>
<td>D = BCF = AEF = ABCDE</td>
<td>AE = BC = DF = ABCDEF</td>
</tr>
<tr>
<td>E = ABC = ADF = BCDF</td>
<td>AF = DE = BCEF = ABCD</td>
</tr>
<tr>
<td>F = BCD = ADE = ABCEF</td>
<td>BD = CF = ACDE = ABEF</td>
</tr>
<tr>
<td>$ABD = CDE = ACF = BEF$</td>
<td>$BF = CD = ACEF = ABDE$</td>
</tr>
</tbody>
</table>

Wordlength pattern $W = (W_0, W_1, \ldots, W_6)$, where $W_i$ is the number of defining words of length $i$ (i.e., involving $i$ factors)

$$W = (1, 0, 0, 0, 3, 0, 0)$$

Resolution is the smallest $i$ such that $i > 0$ and $W_i > 0$. Hence it is a $2^{6-2}_{IV}$ design
$2^{4-2}$ Design: an Alternative

- Basic Design: $A$, $B$, $C$, $D$

- $E = ABCD$, $F = ABC$, i.e., $I = ABCDE$, and $I = ABCF$

- which induces: $I = DEF$

- complete defining relation: $I = ABCDE = ABCF = DEF$

- wordlength pattern: $W =$

- Alias structure (ignore effects of order 3 or higher)

\[
\begin{array}{c|c}
A = .. & AB = CF = .. \\
B = .. & AC = BF = .. \\
C = .. & AD = .. \\
D = EF = .. & AE = .. \\
E = DF = .. & AF = BC = .. \\
F = DE = .. & BD = .. \\
& BE = .. \\
& CD = .. \\
& CE = .. \\
\end{array}
\]

- an effect is said to be clearly estimable if it is not aliased with main effect or two-factor interactions.

- Which design is better?
Injection Molding Experiment Analysis

goption colors=(none);
data molding;
do D = -1 to 1 by 2;
do C = -1 to 1 by 2;
do B = -1 to 1 by 2; do A = -1 to 1 by 2; E=A*B*C; F=B*C*D;
input y @@; output; end; end; end; end;
datalines;
6 10 32 60 4 15 26 60 8 12 34 60 16 5 37 52 ;
data inter; /* Define Interaction Terms */
set molding;
AB=A*B; AC=A*C; AD=A*D; AE=A*E; AF=A*F; BD=B*D; BF=B*F; ABD=A*B*D;
ACD=A*C*D;

proc glm data=inter; /* GLM Proc to Obtain Effects */
class A B C D E F AB AC AD AE AF BD BF ABD ACD;
model y=A B C D E F AB AC AD AE AF BD BF ABD ACD;
estimate 'A' A -1 1; estimate 'B' B -1 1; estimate 'C' C -1 1;
estimate 'D' D -1 1; estimate 'E' E -1 1; estimate 'F' F -1 1;
estimate 'AB' AB -1 1; estimate 'AC' AC -1 1; estimate 'AD' AD -1 1;
estimate 'AE' AE -1 1; estimate 'AF' AF -1 1; estimate 'BD' BD -1 1;
estimate 'BF' BF -1 1; estimate 'ABD' ABD -1 1; estimate 'ACD' ACD -1 1;
run;

proc reg outest=effects data=inter; /* REG Proc to Obtain Effects */
model y=A B C D E F AB AC AD AE AF BD BF ABD ACD;
data effect2; set effects; drop y intercept _RMSE_;
proc transpose data=effect2 out=effect3;
data effect4; set effect3; effect=col1*2;
proc sort data=effect4; by effect;
proc print data=effect4;
proc rank data=effect4 normal=blom; var effect; ranks neff;
symbol1 v=circle;
proc gplot; plot effect*neff=_NAME_; run;

1-21
Estimates of factorial effects

<table>
<thead>
<tr>
<th>Obs</th>
<th><em>NAME</em></th>
<th>COL1</th>
<th>effect</th>
<th>aliases</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AD</td>
<td>-2.6875</td>
<td>-5.375</td>
<td>AD+EF</td>
</tr>
<tr>
<td>2</td>
<td>ACD</td>
<td>-2.4375</td>
<td>-4.875</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>AE</td>
<td>-0.9375</td>
<td>-1.875</td>
<td>AE+BC+DF</td>
</tr>
<tr>
<td>4</td>
<td>AC</td>
<td>-0.8125</td>
<td>-1.625</td>
<td>AC+BE</td>
</tr>
<tr>
<td>5</td>
<td>C</td>
<td>-0.4375</td>
<td>-0.875</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>BD</td>
<td>-0.0625</td>
<td>-0.125</td>
<td>BD+CF</td>
</tr>
<tr>
<td>7</td>
<td>BF</td>
<td>-0.0625</td>
<td>-0.125</td>
<td>BF+CD</td>
</tr>
<tr>
<td>8</td>
<td>ABD</td>
<td>0.0625</td>
<td>0.125</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>E</td>
<td>0.1875</td>
<td>0.375</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>F</td>
<td>0.1875</td>
<td>0.375</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>AF</td>
<td>0.3125</td>
<td>0.625</td>
<td>AF+DE</td>
</tr>
<tr>
<td>12</td>
<td>D</td>
<td>0.6875</td>
<td>1.375</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>AB</td>
<td>5.9375</td>
<td>11.875</td>
<td>AB+CE</td>
</tr>
<tr>
<td>14</td>
<td>A</td>
<td>6.9375</td>
<td>13.875</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>B</td>
<td>17.8125</td>
<td>35.625</td>
<td></td>
</tr>
</tbody>
</table>

Effects B, A, AB, AD, ACD, are large.
**QQ plot to Identify Important Effects**

Effects $B$, $A$, $AB$ appear to be important; effects $AD$ and $ACD$ are suspicious.
De-aliasing and Model Identification

Model 1:
proc reg data=inter;
model y=A B AB AD ACD;
run;

Root MSE 1.95256  R-Square 0.9943
Dependent Mean 27.31250  Adj R-Sq 0.9914
Coeff Var 7.14897

| Parameter Estimates | Parameter | Standard Error | t Value | Pr > |t| |
|---------------------|-----------|----------------|---------|------|---------|
| Intercept           | 27.31250  | 0.48814        | 55.95   | <.0001|
| A                   | 6.93750   | 0.48814        | 14.21   | <.0001|
| B                   | 17.81250  | 0.48814        | 36.49   | <.0001|
| AB                  | 5.93750   | 0.48814        | 12.16   | <.0001|
| AD                  | -2.68750  | 0.48814        | -5.51   | 0.0003|
| ACD                 | -2.43750  | 0.48814        | -4.99   | 0.0005|

Model 2:
proc reg data=inter;
model y=A B AB;

Root MSE 4.55293  R-Square 0.9626
Dependent Mean 27.31250  Adj R-Sq 0.9533
Coeff Var 16.66976

| Parameter Estimates | Parameter | Standard Error | t Value | Pr > |t| |
|---------------------|-----------|----------------|---------|------|---------|
| Intercept           | 27.31250  | 1.13823        | 24.00   | <.0001|
| A                   | 6.93750   | 1.13823        | 6.09    | <.0001|
| B                   | 17.81250  | 1.13823        | 15.65   | <.0001|
| AB                  | 5.93750   | 1.13823        | 5.22    | 0.0002|
Three-Factor Interaction: SAS code

$ACD$ or $ABF$?

data inter; /* Define Interaction Terms */
set molding;
AB=A*B; AC=A*C; AD=A*D; AE=A*E; AF=A*F; BD=B*D; BF=B*F; ABD=A*B*D;
ACD=A*C*D;
if B=-1 and F=-1 then SBF='B-F-';
if B=-1 and F=1 then SBF='B-F+';
if B=1 and F=-1 then SBF='B+F-';
if B=1 and F=1 then SBF='B+F+';

proc sort data=inter; by A SBF;
proc means noprint;
var y; by A SBF;
output out=ymeanabf mean=mn;

symbol1 v=circle i=join; symbol2 v=square i=join;
symbol3 v=diamond i=join; symbol4 v=dot i=join;
proc gplot data=ymeanabf;
plot mn*A=SBF

1-25
3-Factor Interaction Plot
General $2^{k-p}$ Fractional Factorial Effects

- $k$ factors, $2^k$ level combinations, but want to run a $2^{-p}$ fraction only.
- Select the first $k - p$ factors to form a full factorial design (basic design).
- Alias the remainig $p$ factors with some high order interactions of the basic design.
- There are $p$ defining relation, which induces other $2^p - p - 1$ defining relations. The complete defining relation is $I = \ldots = ... = \ldots$.
- Defining contrasts subgroup: $G = \{ \text{defining words} \}$
- Wordlength pattern: $W = (W_i)$ $W_i$=the number of defining words of length $i$.
- Alias structure: $2^k$ factorial effects are partitioned into $2^{k-p}$ groups of effects, each of which contains $2^p$ effects. Effects in the same group are aliased (aliases).
- Use maximum resolution and minimum aberration to choose the optimal design.

- In analysis, only select one effect from each group to be included in the full model.
- Choose important effect to form models, pool unimportant effects into error component
- De-aliasing and model selection.
Minimin Aberration Criterion

Recall $2^{k-p}$ with maximum resolution should be preferred. But, we can have two designs both have the maximum resolution. How should we further distinguish them?

For example, consider $2^{7-2}$ fractional factorial design

$d_1$: basic design: $A, B, C, D, E; F = ABC, G = ABDE$
complete defining relation: $I = ABCF = ABDEG = CDEFG$
wordlength pattern: $W = (1, 0, 0, 0, 1, 2, 0, 0)$
Resolution: IV

$d_2$: basic design: $A, B, C, D, E; F = ABC, G = ADE$
complete defining relation: $I = ABCF = ADEG = BCDEGF$
wordlength pattern: $W = (1, 0, 0, 0, 2, 0, 1, 0)$
Resolution: IV

$d_1$ and $d_2$, which is better?

**Minimum Aberration Criterion**
Definition: For any $2^{k-p}$ designs $d_1$ and $d_2$, let $r$ be the smallest positive integer such that $W_r(d_1) \neq W_r(d_2)$.
If $W_r(d_1) < W_r(d_2)$, then $d_1$ is said to have less aberration than $d_2$. If there is no design with less aberration than $d_1$, then $d_1$ has minimum aberration.

Small Minimum Aberration Designs are used a lot in practice. They are tabulated in most design books. For the most comprehensive table, consult Wu&Hamada.