2^kr Factorial Designs

Raj Jain
Washington University in Saint Louis
Saint Louis, MO 63130
Jain@cse.wustl.edu

These slides are available on-line at:
http://www.cse.wustl.edu/~jain/cse567-15/
Overview

- Computation of Effects
- Estimation of Experimental Errors
- Allocation of Variation
- Confidence Intervals for Effects
- Confidence Intervals for Predicted Responses
- Visual Tests for Verifying the assumptions
- Multiplicative Models
2^kr Factorial Designs

- r replications of 2^k Experiments
  ⇒ 2^kr observations.
  ⇒ Allows estimation of experimental errors.

- Model:
  \[ y = q_0 + q_A x_A + q_B x_B + q_{AB} x_A x_B + e \]

- e = Experimental error
Computation of Effects

Simply use means of r measurements

<table>
<thead>
<tr>
<th>I</th>
<th>A</th>
<th>B</th>
<th>A</th>
<th>B</th>
<th>y</th>
<th>Mean</th>
<th>(\bar{y})</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td></td>
<td>(15, 18, 12)</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td></td>
<td>(45, 48, 51)</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td></td>
<td>(25, 28, 19)</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td>(75, 75, 81)</td>
<td>77</td>
<td></td>
</tr>
<tr>
<td>164</td>
<td>86</td>
<td>38</td>
<td>20</td>
<td></td>
<td>total</td>
<td></td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>21.5</td>
<td>9.5</td>
<td>5</td>
<td></td>
<td>total/4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Effects: \(q_0 = 41, \ q_A = 21.5, \ q_B = 9.5, \ q_{AB} = 5\).
Estimation of Experimental Errors

- Estimated Response:

\[ \hat{y}_i = q_0 + q_A x_{Ai} + q_B x_{Bi} + q_{AB} x_{Ai} x_{Bi} \]

Experimental Error = Measured - Estimated

\[ e_{ij} = y_{ij} - \hat{y}_i \]
\[ = y_{ij} - q_0 - q_A x_{Ai} - q_B x_{Bi} - q_{AB} x_{Ai} x_{Bi} \]
\[ \sum_{i,j} e_{ij} = 0 \]

- Sum of Squared Errors: \[ SSE = \sum_{i=1}^{2^2} \sum_{j=1}^{r} e_{ij}^2 \]
Experimental Errors: Example

- Estimated Response:
  \[ \hat{y}_1 = q_0 - q_A - q_B + q_{AB} = 41 - 21.5 - 9.5 + 5 = 15 \]

- Experimental errors:
  \[ e_{11} = y_{11} - \hat{y}_1 = 15 - 15 = 0 \]

<table>
<thead>
<tr>
<th>i</th>
<th>I</th>
<th>A</th>
<th>B</th>
<th>A</th>
<th>B</th>
<th>Estimated Response</th>
<th>Measured Responses</th>
<th>Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>41</td>
<td>21.5</td>
<td>9.5</td>
<td>5</td>
<td></td>
<td>(\hat{y}_i)</td>
<td>(y_{i1})</td>
<td>(y_{i2})</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td></td>
<td>15</td>
<td>15</td>
<td>18</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td></td>
<td>48</td>
<td>45</td>
<td>48</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td></td>
<td>24</td>
<td>25</td>
<td>28</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td>77</td>
<td>75</td>
<td>75</td>
</tr>
</tbody>
</table>
Allocation of Variation

- Total variation or total sum of squares:

\[
SST = \sum_{i,j} (y_{ij} - \bar{y}.)^2
\]

\[
y_{ij} = q_0 + q_A x_{Ai} + q_B x_{Bi} + q_{AB} x_{Ai} x_{Bi} + e_{ij}
\]

\[
\sum_{i,j} (y_{ij} - \bar{y}.)^2 = 2^2 r q_A^2 + 2^2 r q_B^2 + 2^2 r q_{AB}^2 + \sum_{i,j} e_{ij}^2
\]

\[
SST = SSA + SSB + SSAB + SSE
\]
Example 18.3: Memory-Cache Study

\[
\begin{align*}
SSY &= 15^2 + 18^2 + 12^2 + 45^2 + \cdots + 75^2 + 75^2 + 81^2 \\ &= 27204 \\
SS0 &= 2^2 r q_0^2 = 12 \times 41^2 = 20172 \\
SSA &= 2^2 r q_A^2 = 12 \times (21.5)^2 = 5547 \\
SSB &= 2^2 r q_B^2 = 12 \times (9.5)^2 = 1083 \\
SSAB &= 2^2 r q_{AB}^2 = 12 \times 5^2 = 300 \\
SSE &= 27204 - 2^2 \times 3(41^2 + 21.5^2 + 9.5^2 + 5^2) \\ &= 102 \\
SST &= SSY - SS0 \\ &= 27204 - 20172 = 7032
\end{align*}
\]
Example 18.3 (Cont)

\[ SSA + SSB + SSAB + SSE \]
\[ = 5547 + 1083 + 300 + 102 \]
\[ = 7032 = SST \]

Factor A explains \( \frac{5547}{7032} \) or 78.88%
Factor B explains 15.40%
Interaction AB explains 4.27%
1.45% is unexplained and is attributed to errors.
Confidence Intervals For Effects

- Effects are random variables.
- Errors \( \sim N(0, \sigma_e) \Rightarrow y \sim N(\bar{y}; \sigma_e) \)
  \[
  q_0 = \frac{1}{2^2 r} \sum_{i,j} y_{i,j}
  \]
- \( q_0 \) = Linear combination of normal variates
  \( \Rightarrow q_0 \) is normal with variance \( \sigma_e^2 / (2^2 r) \)

Variance of errors:

\[
se^2 = \frac{1}{2^2(r - 1)} \sum_{i,j} e_{ij}^2 = \frac{SSE}{2^2(r - 1)} \triangle \text{MSE}
\]

- Denominator = \( 2^2(r-1) \) = # of independent terms in SSE
  \( \Rightarrow \) SSE has \( 2^2(r-1) \) degrees of freedom.
- Estimated variance of \( q_0 \): \( s_{q_0}^2 = s_e^2 / (2^2 r) \)
Conf. Intervals For Effects (Cont)

- Similarly,
  \[ s_{q_A} = s_{q_B} = s_{q_{AB}} = \frac{s_e}{\sqrt{2^2 r}} \]

- Confidence intervals (CI) for the effects:
  \[ q_i \pm t[1-\alpha/2; 2^2(r-1)] s_{q_i} \]

- CI does not include a zero \( \Rightarrow \) significant
Example 18.4

- For Memory-cache study: Standard deviation of errors:
  \[ s_e = \sqrt{\frac{\text{SSE}}{2^2(r-1)}} = \sqrt{\frac{102}{8}} = \sqrt{12.75} = 3.57 \]

- Standard deviation of effects:
  \[ s_{q_i} = s_e / \sqrt{(2^2r)} = 3.57 / \sqrt{12} = 1.03 \]

- For 90% Confidence: \( t_{[0.95,8]} = 1.86 \)

- Confidence intervals: \( q_i \mp (1.86)(1.03) = q_i \mp 1.92 \)
  - \( q_0 = (39.08, 42.91) \)
  - \( q_A = (19.58, 23.41) \)
  - \( q_B = (7.58, 11.41) \)
  - \( q_{AB} = (3.08, 6.91) \)

- No zero crossing \( \Rightarrow \) All effects are significant.
Confidence Intervals for Contrasts

- Contrast: Linear combination with $\sum$ coefficients = 0
  $$\sum h_i q_i \text{ with } \sum h_i = 0$$
  For example, $q_A - q_B$ or $q_A + q_B - 2q_{AB}$

- Mean of $\sum h_i q_i = \sum h_i E[q_i]$}

- Variance of $\sum h_i q_i$
  $$s^2_{\sum h_i q_i} = \frac{s^2_e \sum h_i^2}{2^{2r}}$$

- For 100(1-$\alpha$)% confidence interval, use $t_{[1-\alpha/2; 2^{2(r-1)}]}$. 
Example 18.5

Memory-cache study

\[ u = q_A + q_B - 2q_{AB} \]

Coefficients = 0, 1, 1, and -2 \( \Rightarrow \) Contrast

Mean \( \bar{u} = 21.5 + 9.5 - 2 \times 5 = 21 \)

Variance \( s_u^2 = \frac{s_e^2 \times 6}{2^2 \times 3} = 6.375 \)

Standard deviation \( s_u = \sqrt{6.375} = 2.52 \)

\( t_{[0.95;8]} = 1.86 \)

90% Confidence interval for \( u: \)

\[ \bar{u} \mp ts_u = 21 \mp 1.86 \times 2.52 = (16.31, 25.69) \]
Conf. Interval For Predictions

- Mean response \( \hat{y} \):
  \[
  \hat{y} = q_0 + q_A x_A + q_B x_B + q_{AB} x_A x_B
  \]

- The standard deviation of the mean of m responses:
  \[
  s_{\hat{y}_m} = s_e \left( \frac{1}{n_{\text{eff}}} + \frac{1}{m} \right)^{1/2}
  \]

\( n_{\text{eff}} \) = Effective deg of freedom

\( n_{\text{eff}} = \frac{\text{Total number of runs}}{1 + \text{Sum of DFs of params used in } \hat{y}} 
\]

\( n_{\text{eff}} = \frac{2^2 r}{5} \)
Conf. Interval for Predictions (Cont)

$100(1-\alpha)\%$ confidence interval:

$$\hat{y} \pm t_{[1-\alpha/2;2^2(r-1)]} s_{\hat{y}_m}$$

- A single run ($m=1$): $s_{\hat{y}_1} = s_e \left( \frac{5}{2^2 r} + 1 \right)^{1/2}$

- Population mean ($m=\infty$): $s_{\hat{y}} = s_e \left( \frac{5}{2^2 r} \right)^{1/2}$
Example 18.6: Memory-cache Study

- For $x_A = -1$ and $x_B = -1$:

- A single confirmation experiment:

  \[
  \hat{y}_1 = q_0 - q_A - q_B + q_{AB} \\
  = 41 - 21.5 - 9.5 + 5 = 15
  \]

- Standard deviation of the prediction:

  \[
  s_{\hat{y}_1} = s_e \left( \frac{5}{2^2r} + 1 \right)^{1/2} = 3.57 \sqrt{\frac{5}{12}} + 1 = 4.25
  \]

- Using $t_{[0.95;8]} = 1.86$, the 90% confidence interval is:

  \[
  15 \pm 1.86 \times 4.25 = (7.09, 22.91)
  \]
Mean response for 5 experiments in future:

\[ s_{\hat{y}_1} = s_e \left( \frac{5}{2^2 r} + \frac{1}{m} \right)^{1/2} \]

\[ = 3.57 \sqrt{\frac{5}{12} + \frac{1}{5}} = 2.80 \]

The 90% confidence interval is:

\[ 15 \mp 1.86 \times 2.80 = (9.79, 20.21) \]
Example 18.6 (Cont)

- Mean response for a large number of experiments in future:

$$s_{\hat{y_1}} = s_e \left( \frac{5}{2^2 r} \right)^{1/2} = 3.57 \sqrt{\frac{5}{12}} = 2.30$$

- The 90% confidence interval is:

$$15 \mp 1.86 \times 2.30 = (10.72, 19.28)$$

- Current mean response: Not for future. Use contrasts formula.

$$s_{\hat{y_1}} = \sqrt{\frac{s^2_e \sum h_i^2}{2^2 r}} = \sqrt{\frac{12.75 \times 4}{12}} = 2.06$$

- 90% confidence interval:

$$15 \mp 1.86 \times 2.06 = (11.17, 18.83)$$
Updated Exercise 18.1: The following table lists measured CPU times for two processors on two workloads. Each experiment was repeated three times. Determine the effects.

Table 18.12 $2^2$ 3 Experimental Design Exercise

<table>
<thead>
<tr>
<th>Workload</th>
<th>Processor</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>I</td>
<td>(41.16, 39.02, 42.56)</td>
</tr>
<tr>
<td>J</td>
<td>(53.50, 55.50, 50.50)</td>
</tr>
</tbody>
</table>
Assumptions

1. Errors are statistically independent.
2. Errors are additive.
3. Errors are normally distributed.
4. Errors have a constant standard deviation \( \sigma_e \).
5. Effects of factors are additive  
   \( \Rightarrow \) observations are independent and normally distributed with constant variance.
Visual Tests

1. **Independent Errors:**
   - Scatter plot of residuals versus the predicted response $\hat{y}_i$
   - Magnitude of residuals $< \text{Magnitude of responses}/10$
     $\Rightarrow$ Ignore trends
   - Plot the residuals as a function of the experiment number
   - Trend up or down $\Rightarrow$ other factors or side effects

2. **Normally distributed errors:**
   Normal quantile-quantile plot of errors

3. **Constant Standard Deviation of Errors:**
   Scatter plot of $y$ for various levels of the factor
   Spread at one level significantly different than that at other
   $\Rightarrow$ Need transformation
Example 18.7: Memory-cache

[Graphs showing predicted response vs. residuals and normal quantile vs. residuals]
Multiplicative Models

- Additive model:
  \[ y_{ij} = q_0 + q_A x_A + q_B x_B + q_{AB} x_A x_B + e_{ij} \]

- Not valid if effects do not add.
  E.g., execution time of workloads.
  \[ i \text{th processor speed} = v_i \text{ instructions/second.} \]
  \[ j \text{th workload Size} = w_j \text{ instructions} \]

- The two effects multiply. Logarithm \( \Rightarrow \) additive model:
  Execution Time \( y_{ij} = w_j / v_i \)
  \[ \log(y_{ij}) = \log(w_j) - \log(v_i) \]

- Correct Model:
  \[ y'_{ij} = q_0 + q_A x_A + q_B x_B + q_{AB} x_A x_B + e_{ij} \]
  Where, \( y'_{ij} = \log(y_{ij}) \)
Multiplicative Model (Cont)

- Taking an antilog of effects:
  \[ u_A = 10^{q_A}, \quad u_B = 10^{q_B}, \quad \text{and} \quad u_{AB} = 10^{q_{AB}} \]

- \( u_A \) = ratio of MIPS rating of the two processors
- \( u_B \) = ratio of the size of the two workloads.
- Antilog of additive mean \( q_0 \) \( \Rightarrow \) geometric mean

\[
y = 10^{q_0} = (y_1 y_2 \cdots y_n)^{1/n} \quad n = 2^r
\]
Example 18.8: Execution Times

Analysis Using an Additive Model

<table>
<thead>
<tr>
<th>I</th>
<th>A</th>
<th>B</th>
<th>AB</th>
<th>y</th>
<th>Mean $\bar{y}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>(85.10, 79.50, 147.90)</td>
<td>104.170</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>(0.891, 1.047, 1.072)</td>
<td>1.003</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>(0.955, 0.933, 1.122)</td>
<td>1.003</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>(0.0148, 0.0126, 0.0118)</td>
<td>0.013</td>
</tr>
<tr>
<td>106.19</td>
<td>-104.15</td>
<td>-104.15</td>
<td>102.17</td>
<td>total</td>
<td></td>
</tr>
<tr>
<td>26.55</td>
<td>-26.04</td>
<td>-26.04</td>
<td>25.54</td>
<td>total/4</td>
<td></td>
</tr>
</tbody>
</table>

Additive model is not valid because:

- Physical consideration $\Rightarrow$ effects of workload and processors do not add. They multiply.
- Large range for $y$. $y_{\text{max}}/y_{\text{min}} = 147.90/0.0118$ or 12,534 $\Rightarrow$ log transformation
- Taking an arithmetic mean of 114.17 and 0.013 is inappropriate.
Example 18.8 (Cont)

- The residuals are not small as compared to the response.
  
  - The spread of residuals is large at larger value of the response.
    
    \[ \Rightarrow \text{log transformation} \]
Example 18.8 (Cont)

- Residual distribution has a longer tail than normal
Analysis Using Multiplicative Model

<table>
<thead>
<tr>
<th>I</th>
<th>A</th>
<th>B</th>
<th>AB</th>
<th>y</th>
<th>Mean</th>
<th>$\bar{y}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>(1.93, 1.90, 2.17)</td>
<td>2.00</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>(-0.05, 0.02, 0.03)</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>(-0.02, -0.03, 0.05)</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>(-1.83, -1.90, -1.93)</td>
<td>-1.89</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.11</td>
<td>-3.89</td>
<td>-3.89</td>
<td>0.11</td>
<td>total</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.03</td>
<td>-0.97</td>
<td>-0.97</td>
<td>0.03</td>
<td>total/4</td>
<td></td>
</tr>
</tbody>
</table>
Variation Explained by the Two Models

<table>
<thead>
<tr>
<th>Factor</th>
<th>Additive Model</th>
<th>Multiplicative Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Effect</td>
<td>% Var.</td>
</tr>
<tr>
<td>I</td>
<td>26.55</td>
<td>(16.35, 36.74)</td>
</tr>
<tr>
<td>A</td>
<td>-26.04</td>
<td>30.1% (36.23, -15.84)</td>
</tr>
<tr>
<td>B</td>
<td>-26.04</td>
<td>30.1% (36.23, -15.84)</td>
</tr>
<tr>
<td>AB</td>
<td>25.54</td>
<td>29.0% (15.35, 35.74)</td>
</tr>
<tr>
<td>e</td>
<td>10.8%</td>
<td></td>
</tr>
</tbody>
</table>

† ⇒ Not Significant

- With multiplicative model:
  - Interaction is almost zero.
  - Unexplained variation is only 0.2%
Visual Tests

- **Conclusion**: Multiplicative model is better than the additive model.
Interpretation of Results

\[
\log(y) = q_0 + q_A x_A + q_B x_B + q_{AB} x_A x_B + e
\]

\[\Rightarrow y = 10^{q_0} 10^{q_A x_A} 10^{q_B x_B} 10^{q_{AB} x_A x_B} 10^e \]

\[= 10^{0.03} 10^{-0.97 x_A} 10^{-0.97 x_B} 10^{0.03 x_A x_B} 10^e \]

\[= 1.07 \times 0.107^{x_A} \times 0.107^{x_B} \times 1.07^{x_A x_B} 10^e \]

- The time for an average processor on an average benchmark is 1.07.
- The time on processor A_1 is 9.35 times (0.107^{-1}) that on an average processor. The time on A_2 is 1/9.35 (0.107^1) of that on an average processor.
- MIPS rate for A_2 is 87.4 times that of A_1.
- Benchmark B_1 executes 87.4 times more instructions than B_2.
- The interaction is negligible.

\[\Rightarrow \text{Results apply to all benchmarks and processors.}\]
Transformation Considerations

- $y_{\text{max}}/y_{\text{min}}$ small $\Rightarrow$ Multiplicative model results similar to additive model.
- Many other transformations possible.
- Box-Cox family of transformations:
  \[
  w = \begin{cases} 
  \frac{y^a - 1}{ag^{a-1}}, & a \neq 0 \\
  (\ln y)g, & a = 0 
  \end{cases}
  \]
- Where $g$ is the geometric mean of the responses:
  \[g = (y_1y_2\cdots y_n)^{1/n}\]
- $w$ has the same units as $y$.
- $a$ can have any real value, positive, negative, or zero.
- Plot SSE as a function of $a \Rightarrow$ optimal $a$
- Knowledge about the system behavior should always take precedence over statistical considerations.
General \(2^k r\) Factorial Design

- **Model:**
  \[ y_{ij} = q_0 + q_A x_{A_i} + q_B x_{B_i} + q_{AB} x_{A_i} x_{B_i} + \cdots + e_{ij} \]

- **Parameter estimation:**
  \[ q_j = \frac{1}{2^k} \sum_{i=1}^{2^k} S_{ij} \bar{y}_i \]
  \( S_{ij} = (i,j)\)th entry in the sign table.

- **Sum of squares:**
  \[ SS_Y = \sum_{i=1}^{2^k} \sum_{j=1}^{r} y_{ij}^2 \]
  \[ SS_0 = 2^k r q_0^2 \]
  \[ SST = SS_Y - SS_0 \]
  \[ SS_j = 2^k r q_j^2 \quad j = 1, 2, \ldots, 2^k - 1 \]
  \[ SSE = SST - \sum_{j=1}^{2^k-1} SS_j \]
General $2^k r$ Factorial Design (Cont)

- Percentage of $y$'s variation explained by $j$th effect =
  \[ \frac{SS_j}{SST} \times 100\% \]

- Standard deviation of errors:
  \[ s_e = \sqrt{\frac{SSE}{2^k(r-1)}} \]

- Standard deviation of effects:
  \[ s_{q_0} = s_{q_A} = s_{q_B} = s_{q_{AB}} = \frac{s_e}{\sqrt{2^k r}} \]

- Variance of contrast $\sum h_i q_i$, where $\sum h_i = 0$ is:
  \[ s^2_{\sum h_i q_i} = \left( \frac{s_e^2 \sum h_i^2}{2^k r} \right) \]
General $2^k r$ Factorial Design (Cont)

- Standard deviation of the mean of $m$ future responses:
  \[ s_{\hat{y}_p} = s_e \left( \frac{1 + 2^k}{2^k r} + \frac{1}{m} \right)^{1/2} \]

- Confidence intervals are calculated using $t_{[1-\alpha/2;2^k(r-1)]}$.

- Modeling assumptions:
  - Errors are IID normal variates with zero mean.
  - Errors have the same variance for all values of the predictors.
  - Effects and errors are additive.
Visual Tests for $2^k r$ Designs

- The scatter plot of errors versus predicted responses should not have any trend.
- The normal quantile-quantile plot of errors should be linear.
- Spread of y values in all experiments should be comparable.
### Example 18.9: A $2^3$ Design

<table>
<thead>
<tr>
<th>I</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>A B</th>
<th>A C</th>
<th>B C</th>
<th>A B C</th>
<th>y</th>
<th>Mean</th>
<th>y</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>
<td>14</td>
<td>16</td>
<td>12</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>
<td>22</td>
<td>18</td>
<td>20</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>
<td>11</td>
<td>15</td>
<td>19</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>
<td>34</td>
<td>30</td>
<td>35</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>
<td>46</td>
<td>42</td>
<td>44</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>
<td>58</td>
<td>62</td>
<td>60</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>
<td>50</td>
<td>55</td>
<td>54</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>
<td>86</td>
<td>80</td>
<td>74</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>319</th>
<th>67</th>
<th>43</th>
<th>155</th>
<th>23</th>
<th>19</th>
<th>15</th>
<th>-1</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>39.87</td>
<td>8.375</td>
<td>5.375</td>
<td>19.37</td>
<td>2.875</td>
<td>2.375</td>
<td>1.875</td>
<td>-0.125</td>
<td>total/8</td>
</tr>
</tbody>
</table>
Example 18.9 (Cont)

- Sum of Squares:

<table>
<thead>
<tr>
<th>Component</th>
<th>Sum of Squares</th>
<th>Percent Variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y$</td>
<td>4.9E4</td>
<td></td>
</tr>
<tr>
<td>$\bar{y}$</td>
<td>3.8E4</td>
<td></td>
</tr>
<tr>
<td>$y-\bar{y}$</td>
<td>1.1E4</td>
<td>100.00%</td>
</tr>
<tr>
<td>A</td>
<td>1683.0</td>
<td>14.06%</td>
</tr>
<tr>
<td>B</td>
<td>693.3</td>
<td>5.79%</td>
</tr>
<tr>
<td>C</td>
<td>9009.0</td>
<td>75.27%</td>
</tr>
<tr>
<td>AB</td>
<td>198.3</td>
<td>1.66%</td>
</tr>
<tr>
<td>AC</td>
<td>135.4</td>
<td>1.13%</td>
</tr>
<tr>
<td>BC</td>
<td>84.4</td>
<td>0.70%</td>
</tr>
<tr>
<td>ABC</td>
<td>0.4</td>
<td>0.00%</td>
</tr>
<tr>
<td>Errors</td>
<td>164.0</td>
<td>1.37%</td>
</tr>
</tbody>
</table>
Example 18.9 (Cont)

- The errors have $2^3(3-1)$ or 16 degrees of freedom. Standard deviation of errors:

$$s_e = \sqrt{\frac{\text{SSE}}{2^k(r - 1)}} = \sqrt{\frac{164}{16}} = 3.20$$

- Standard deviation of effects:

$$s_{qi} = \frac{s_e}{\sqrt{2^33}} = \frac{3.20}{\sqrt{24}} = 0.654$$
Example 18.9 (Cont)

- % Variation:

<table>
<thead>
<tr>
<th>Component</th>
<th>Sum of Squares</th>
<th>Percent Variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y$</td>
<td>4.9E4</td>
<td></td>
</tr>
<tr>
<td>$\bar{y}$</td>
<td>3.8E4</td>
<td></td>
</tr>
<tr>
<td>$y-\bar{y}$</td>
<td>1.1E4</td>
<td>100.00%</td>
</tr>
<tr>
<td>A</td>
<td>1683.0</td>
<td>14.06%</td>
</tr>
<tr>
<td>B</td>
<td>693.3</td>
<td>5.79%</td>
</tr>
<tr>
<td>C</td>
<td>9009.0</td>
<td>75.27%</td>
</tr>
<tr>
<td>AB</td>
<td>198.3</td>
<td>1.66%</td>
</tr>
<tr>
<td>AC</td>
<td>135.4</td>
<td>1.13%</td>
</tr>
<tr>
<td>BC</td>
<td>84.4</td>
<td>0.70%</td>
</tr>
<tr>
<td>ABC</td>
<td>0.4</td>
<td>0.00%</td>
</tr>
<tr>
<td>Errors</td>
<td>164.0</td>
<td>1.37%</td>
</tr>
</tbody>
</table>
Example 18.9 (Cont)

- $t_{[0.90, 16]} = 1.337$
- 80% confidence intervals for parameters: $q_i \mp (1.337)(0.654) = q_i \mp 0.874$
  
  - $q_0 = (39.00, 40.74)$
  - $q_A = (7.50, 9.25)$
  - $q_B = (4.50, 6.25)$
  - $q_C = (18.50, 20.24)$
  - $q_{AB} = (2.00, 3.75)$
  - $q_{AC} = (1.50, 3.25)$
  - $q_{BC} = (1.00, 2.75)$
  - $q_{ABC} = (-1.00, 0.75)$

- All effects except $q_{ABC}$ are significant.
Example 18.9 (Cont)

- For a single confirmation experiment \((m = 1)\)
  With \(A = B = C = -1\):

\[
\hat{y} = 14
\]

\[
s_{\hat{y}} = s_e \left( \frac{9}{2^k r} + \frac{1}{m} \right)^{1/2}
\]

\[
= 3.2 \left( \frac{9}{24} + 1 \right)^{1/2}
\]

\[
= 3.75
\]

- 80% confidence interval:

\[
14 \pm 1.337 \times 3.75 = 14 \pm 5.02 = (8.98, 19.02)
\]
Importance vs. Significance

- Important = Parameters or models that explain a high percent of variation
- Significant = Parameters or models whose confidence interval does not include zero

**Important**
- Run more experiments

**Significant**
- Ignore the parameter and see if the model is still significant

**Good**
### Case Study 18.1: Garbage collection

#### Factors and Levels

<table>
<thead>
<tr>
<th>Variable</th>
<th>Factor</th>
<th>-1</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Workload</td>
<td>Single Task</td>
<td>Several parallel tasks</td>
</tr>
<tr>
<td>B</td>
<td>Compiler</td>
<td>Simple</td>
<td>Deallocation</td>
</tr>
<tr>
<td>C</td>
<td>Limbo List</td>
<td>Enabled</td>
<td>Disabled</td>
</tr>
<tr>
<td>D</td>
<td>Chunk Size</td>
<td>4K bytes</td>
<td>16K bytes</td>
</tr>
</tbody>
</table>
## Case Study 18.1 (Cont)

<table>
<thead>
<tr>
<th>I</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>y</th>
<th>Mean $\bar{y}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>(97, 97, 97)</td>
<td>97.00</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>(31, 31, 32)</td>
<td>31.33</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>(97, 97, 97)</td>
<td>97.00</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>(31, 32, 31)</td>
<td>31.33</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>(97, 97, 97)</td>
<td>97.00</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>(32, 32, 31)</td>
<td>31.67</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>(97, 97, 97)</td>
<td>97.00</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>(32, 32, 32)</td>
<td>32.00</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>(407, 407, 407)</td>
<td>407.00</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>(135, 136, 135)</td>
<td>135.33</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>(409, 409, 409)</td>
<td>409.00</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>(135, 135, 136)</td>
<td>135.33</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>(407, 407, 407)</td>
<td>407.00</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>(139, 140, 139)</td>
<td>139.33</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>(409, 409, 409)</td>
<td>409.00</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>(139, 139, 140)</td>
<td>139.33</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Total</th>
<th>2695.67</th>
<th>-1344.33</th>
<th>4.33</th>
<th>9.00</th>
<th>1667.00</th>
<th>Total/16</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>168.48</td>
<td>-84.02</td>
<td>0.27</td>
<td>0.56</td>
<td>104.19</td>
<td></td>
</tr>
</tbody>
</table>
## Case Study 18.1 (Cont)

<table>
<thead>
<tr>
<th>Factor</th>
<th>Effect</th>
<th>% Variation</th>
<th>Conf. Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>168.48</td>
<td></td>
<td>(168.386, 168.573)</td>
</tr>
<tr>
<td>A</td>
<td>-84.02</td>
<td>34.4%</td>
<td>(-84.114, -83.927)</td>
</tr>
<tr>
<td>B</td>
<td>0.27</td>
<td>0.0%</td>
<td>(0.177, 0.364)</td>
</tr>
<tr>
<td>C</td>
<td>0.56</td>
<td>0.0%</td>
<td>(0.469, 0.656)</td>
</tr>
<tr>
<td>D</td>
<td>104.19</td>
<td>52.8%</td>
<td>(104.094, 104.281)</td>
</tr>
<tr>
<td>AB</td>
<td>-0.23</td>
<td>0.0%</td>
<td>(-0.323, -0.136)</td>
</tr>
<tr>
<td>AC</td>
<td>0.56</td>
<td>0.0%</td>
<td>(0.469, 0.656)</td>
</tr>
<tr>
<td>AD</td>
<td>-51.31</td>
<td>12.8%</td>
<td>(-51.406, -51.219)</td>
</tr>
<tr>
<td>BC</td>
<td>0.02</td>
<td>0.0%</td>
<td>(-0.073, 0.114)†</td>
</tr>
<tr>
<td>BD</td>
<td>0.23</td>
<td>0.0%</td>
<td>(0.136, 0.323)</td>
</tr>
<tr>
<td>CD</td>
<td>0.44</td>
<td>0.0%</td>
<td>(0.344, 0.531)</td>
</tr>
<tr>
<td>ABC</td>
<td>0.02</td>
<td>0.0%</td>
<td>(-0.073, 0.114)†</td>
</tr>
<tr>
<td>ABD</td>
<td>-0.27</td>
<td>0.0%</td>
<td>(-0.364, -0.177)</td>
</tr>
<tr>
<td>ACD</td>
<td>0.44</td>
<td>0.0%</td>
<td>(0.344, 0.531)</td>
</tr>
<tr>
<td>BCD</td>
<td>-0.02</td>
<td>0.0%</td>
<td>(-0.114, 0.073)†</td>
</tr>
<tr>
<td>ABCD</td>
<td>-0.02</td>
<td>0.0%</td>
<td>(-0.114, 0.073)†</td>
</tr>
</tbody>
</table>

† ⇒ Not Significant
Case Study 18.1: Conclusions

- Most of the variation is explained by factors A (Workload), D (Chunk size), and the interaction A D between the two.

- The variation due to experimental error is small
  \[ \Rightarrow \] Several effects that explain less than 0.05% of variation (listed as 0.0%) are statistically significant.

- Only effects A, D, and AD are both practically significant and statistically significant.
Summary

- Replications allow estimation of measurement errors
  ⇒ Confidence Intervals of parameters
  ⇒ Confidence Intervals of predicted responses
- Allocation of variation is proportional to square of effects
- Multiplicative models are appropriate if the factors multiply
- Visual tests for independence normal errors
Updated Exercise 18.1: For the data of Homework 18A, determine percentage of variation explained, find confidence intervals of the effects, and conduct visual tests.