Overview

$2^k r$ Designs

$2^2 r$ Designs
Effects
Analysis of Variance
Confidence Intervals
Predictions
Verification

Multiplicative Models
Example

General $2^k r$ Designs
\(2^k\) Factorial Designs With Replications

- \(2^k\) factorial designs do not allow for estimation of experimental error
  - No experiment is ever repeated
- Error is usually present
  - And usually important
- Handle issue by replicating experiments
- But which to replicate, and how often?
$2^k r$ Factorial Designs

- Replicate each experiment $r$ times
- Allows quantifying experimental error
- Again, easiest to first look at case of only 2 factors
2\(^r\) Factorial Designs

- 2 factors, 2 levels each, with \(r\) replications at each of the four combinations
- \(y = q_0 + q_A x_A + q_B x_B + q_{AB} x_A x_B + e\)
- Now we need to compute effects, estimate errors, and allocate variation
- Can also produce confidence intervals for effects and predicted responses
Computing Effects for $2^r$ Factorial Experiments

- We can use sign table, as before
- But instead of single observations, regress off mean of the $r$ observations
- Compute errors for each replication using similar tabular method
  - Sum of errors must be zero
  - $e_{ij} = y_{ij} - \hat{y}_i$
- Similar methods used for allocation of variance and calculating confidence intervals

The tabular method for errors is as follows: after computing the effects, multiply the effects by the sign table to get the estimated response. Enter that into the table and then subtract from each measured response to get errors.
Example of $2^r$ Factorial Design With Replications

- Same parallel system as before, but with 4 replications at each point ($r = 4$)
- No DLM, 8 nodes: 820, 822, 813, 809
- DLM, 8 nodes: 776, 798, 750, 755
- No DLM, 64 nodes: 217, 228, 215, 221
- DLM, 64 nodes: 197, 180, 220, 185
### $2^2r$ Factorial Example Analysis Matrix

<table>
<thead>
<tr>
<th>I</th>
<th>A</th>
<th>B</th>
<th>AB</th>
<th>y</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>(820,822,813,809)</td>
<td>816.00</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>(217,228,215,221)</td>
<td>220.25</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>(776,798,750,755)</td>
<td>769.75</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>(197,180,220,185)</td>
<td>195.50</td>
</tr>
<tr>
<td>2001.5</td>
<td>-1170.0</td>
<td>-71.00</td>
<td>21.5</td>
<td>Total</td>
<td></td>
</tr>
<tr>
<td>500.4</td>
<td>-292.5</td>
<td>-17.75</td>
<td>5.4</td>
<td>Total/4</td>
<td></td>
</tr>
</tbody>
</table>

$q_0 = 500.40 \quad q_A = -292.5$

$q_B = -17.75 \quad q_{AB} = 5.4$
Estimation of Errors for $2^2r$ Factorial Example

Figure differences between predicted and observed values for each replication:

$$e_{ij} = y_{ij} - \hat{y}_i = y_{ij} - q_0 - q_A x_{Ai} - q_B x_{Bi} - q_{AB} x_{Ai} x_{Bi}$$

Now calculate SSE:

$$SSE = \sum_{i=1}^{2^2} \sum_{j=1}^{r} e_{ij}^2 = 2606$$
Allocating Variation

- We can determine percentage of variation due to each factor’s impact
  - Just like $2^k$ designs without replication
- But we can also isolate variation due to experimental errors
- Methods are similar to other regression techniques for allocating variation
We’ve already figured SSE
We also need SST, SSA, SSB, and SSAB

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

Also, SST = SSA + SSB + SSAB + SSE
Use same formulae as before for SSA, SSB, and SSAB
Sums of Squares for Example

- SST = SSY – SS0 = 1,377,009.75
- SSA = 1,368,900
- SSB = 5041
- SSAB = 462.25
- Percentage of variation for A is 99.4%
- Percentage of variation for B is 0.4%
- Percentage of variation for A/B interaction is 0.03%
- And 0.2% (approx.) is due to experimental errors
Confidence Intervals for Effects

- Computed effects are random variables
- Thus would like to specify how confident we are that they are correct
- Usual confidence-interval methods
- First, must figure Mean Square of Errors
  \[ s^2_e = \frac{SSE}{2^2(r - 1)} \]
- \( r - 1 \) is because errors add up to zero
  ⇒ Only \( r - 1 \) can be chosen independently
Calculated Variances of Effects

- Variance (due to errors) of all effects is the same:

\[ s^2_{q_0} = s^2_{q_A} = s^2_{q_B} = s^2_{q_{AB}} = \frac{s^2_e}{2^r} \]

- So standard deviation is also the same
- In calculations, use \( t \)- or \( z \)-value for \( 2^r(r-1) \) degrees of freedom
Calculating Confidence Intervals for Example

- At 90% level, using $t$-value for 12 degrees of freedom, 1.782
- Standard deviation of effects is 3.68
- Confidence intervals are $q_i \mp (1.782)(3.68)$
- $q_0$ is (493.8, 506.9)
- $q_A$ is (-299.1, -285.9)
- $q_B$ is (-24.3, -11.2)
- $q_{AB}$ is (-1.2, 11.9)
 ► We already have predicted all the means we can predict from this kind of model
 ► We measured four, we can “predict” four
 ► However, we can predict how close we would get to true sample mean if we ran \( m \) more experiments
For $m$ future experiments, predicted mean is

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

Where

$$s_{\hat{y}m} = s_e \left( \frac{1}{n_{\text{eff}}} + \frac{1}{m} \right)^{1/2}$$

$$n_{\text{eff}} = \frac{\text{Total number of runs}}{1 + \text{sum of DFs of parameters used in } \hat{y}}$$
Example of Predicted Means

- What would we predict as confidence interval of response for no dynamic load management at 8 nodes for 7 more tests?

\[ s^\hat{y} = 3.68 \left( \frac{1}{16/5} + \frac{1}{7} \right)^{1/2} = 2.49 \]

- 90% confidence interval is (811.6, 820.4)
- We’re 90% confident that mean would be in this range
Visual Tests for Verifying Assumptions

- What assumptions have we been making?
  - Model errors are statistically independent
  - Model errors are additive
  - Errors are normally distributed
  - Errors have constant standard deviation
  - Effects of errors are additive

- All boils down to independent, normally distributed observations with constant variance
Testing for Independent Errors

- Compute residuals and make scatter plot
- Trends indicate dependence of errors on factor levels
  - But if residuals order of magnitude below predicted response, trends can be ignored
- Usually good idea to plot residuals vs. experiment number
Example Plot of Residuals vs. Predicted Response

Predicted Response

Error Residual

0 200 400 600 800 1000

0 10 20 30

-20

2015-06-15
CS147
2^k Designs
Verification
Example Plot of Residuals vs. Predicted Response
Example Plot of Residuals vs. Experiment Number

Experiment Number

Error Residual

0 5 10 15
-20 -10 0 10 20 30
As usual, do quantile-quantile chart against normal distribution
If close to linear, normality assumption is good
Assumption of Constant Variance

- Checking homoscedasticity
- Go back to scatter plot of residuals vs. prediction and check for even spread
The Scatter Plot, Again
Multiplicative Models for $2^r$ Experiments

- Assumptions of additive models
- Example of a multiplicative situation
- Handling a multiplicative model
- When to choose multiplicative model
- Multiplicative example
Previous analysis used additive model:

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

Assumes all effects are additive:

- Factors
- Interactions
- Errors

This assumption must be validated!
Testing processors with different workloads
Most common multiplicative case
Consider 2 processors, 2 workloads
  Use $2^2$ $r$ design
Response is time to execute $w_j$ instructions on processor that requires $v_i$ seconds/instruction
Without interactions, time is $y_{ij} = v_i w_j$
Handling a Multiplicative Model

- Take logarithm of both sides:
  \[ y_{ij} = v_i w_j \]
  so \[ \log y_{ij} = \log v_i + \log w_j \]
- Now easy to solve using previous methods
- Resulting model is:
  \[ y = 10^{q_0} 10^{q_A x_A} 10^{q_B x_B} 10^{q_{AB} x_{AB}} 10^{e} \]
Meaning of a Multiplicative Model

- Model is $10^{q_0} 10^{q_A x_A} 10^{q_B x_B} 10^{q_{AB} x_{AB}} 10^e$
- Here, $\mu_A = 10^{q_A}$ is inverse of ratio of MIPS ratings of processors; $\mu_B = 10^{q_B}$ is ratio of workload sizes
- Antilog of $q_0$ is geometric mean of responses:
  $$\hat{y} = 10^{q_0} = \sqrt[n]{y_1 y_2 \cdots y_n}$$
  where $n = 2^2 r$
When to Choose a Multiplicative Model?

- Physical considerations (see previous slides)
- Range of $y$ is large
  - Making arithmetic mean unreasonable
  - Calling for log transformation
- Plot of residuals shows large values and increasing spread
- Quantile-quantile plot doesn’t look like normal distribution
Consider additive model of processors $A_1$ & $A_2$ running benchmarks $B_1$ and $B_2$:

<table>
<thead>
<tr>
<th>$y_1$</th>
<th>$y_2$</th>
<th>$y_3$ Mean</th>
<th>I</th>
<th>A</th>
<th>B</th>
<th>AB</th>
</tr>
</thead>
<tbody>
<tr>
<td>85.1</td>
<td>79.5</td>
<td>147.9</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>0.891</td>
<td>1.047</td>
<td>1.072</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>0.955</td>
<td>0.933</td>
<td>1.122</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0.015</td>
<td>0.013</td>
<td>0.012</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Total: 106.19 -104.15 -104.15 102.17

Note large range of $y$ values
Quantile-Quantile Plot of Additive Model
Multiplicative Models

- Taking logs of everything, the model is:

<table>
<thead>
<tr>
<th></th>
<th>$y_1$</th>
<th>$y_2$</th>
<th>$y_3$</th>
<th>Mean</th>
<th>I</th>
<th>A</th>
<th>B</th>
<th>AB</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.93</td>
<td>1.9</td>
<td>2.17</td>
<td>2.000</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>-0.05</td>
<td>0.02</td>
<td>0.0302</td>
<td>0.000</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.02</td>
<td>-0.03</td>
<td>0.05</td>
<td>0.000</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-1.83</td>
<td>-1.9</td>
<td>-1.928</td>
<td>-1.886</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>0.11</td>
<td>-3.89</td>
<td>-3.89</td>
<td>0.11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total/4</td>
<td>0.03</td>
<td>-0.97</td>
<td>-0.97</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Error Residuals of Multiplicative Model

-2 -1 0 1 2
Predicted Response
-0.1 0.0 0.1 0.2
Error Residual

2015-06-15
CS147
Quantile-Quantile Plot for Multiplicative Model
### Summary of the Two Models

<table>
<thead>
<tr>
<th>Factor</th>
<th>Additive Model</th>
<th>Multiplicative Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pct of Variation</td>
<td>Confidence Interval</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>30.15 - 36.23 - 15.85</td>
</tr>
<tr>
<td>B</td>
<td>-26.04</td>
<td>30.15 - 36.23 - 15.85</td>
</tr>
<tr>
<td>AB</td>
<td>25.54</td>
<td>29.01 - 15.35 - 35.74</td>
</tr>
<tr>
<td>e</td>
<td>10.69</td>
<td>15.35 - 35.74</td>
</tr>
</tbody>
</table>

2015-06-15
General $2^k r$ Factorial Design

- Simple extension of $2^2 r$
- See Box 18.1 in book for summary
- Always do visual tests
- Remember to consider multiplicative model as alternative
Consider a $2^3$ design:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<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>14</td>
<td>16</td>
<td>12</td>
<td>14</td>
<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>22</td>
<td>18</td>
<td>20</td>
<td>20</td>
<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>11</td>
<td>15</td>
<td>19</td>
<td>15</td>
<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>34</td>
<td>30</td>
<td>35</td>
<td>33</td>
<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>46</td>
<td>42</td>
<td>44</td>
<td>44</td>
<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>58</td>
<td>62</td>
<td>60</td>
<td>60</td>
<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>50</td>
<td>55</td>
<td>54</td>
<td>53</td>
<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>86</td>
<td>80</td>
<td>74</td>
<td>80</td>
<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>Total</td>
<td>319</td>
<td>67</td>
<td>43</td>
<td>155</td>
<td>23</td>
<td>19</td>
<td>15</td>
<td>-1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total/8</td>
<td>39.88</td>
<td>8.38</td>
<td>5.38</td>
<td>19.38</td>
<td>2.88</td>
<td>2.38</td>
<td>1.88</td>
<td>-0.13</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
ANOVA for $2^3$ Design

- Percent variation explained:
  
<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>AB</th>
<th>AC</th>
<th>BC</th>
<th>ABC</th>
<th>Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>14.1</td>
<td>5.8</td>
<td>75.3</td>
<td>1.7</td>
<td>1.1</td>
<td>0.7</td>
<td>0.7</td>
<td>1.37</td>
</tr>
</tbody>
</table>

- 90% confidence intervals

<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>38.7</td>
<td>7.2</td>
<td>4.2</td>
<td>18.2</td>
<td>1.7</td>
<td>1.2</td>
<td>0.7</td>
<td>-1.3</td>
</tr>
<tr>
<td>41.0</td>
<td>9.5</td>
<td>6.5</td>
<td>20.5</td>
<td>4.0</td>
<td>3.5</td>
<td>3.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>
Error Residuals for $2^3 \times 3$ Design

Predicted Response

Error Residual

0 20 40 60 80

2015-06-15
Quantile-Quantile Plot for $2^33$ Design