Design of Experiments:  
Factorial Designs

Prof. Daniel A. Menasce  
Dept. of Computer Science  
George Mason University

Basic Concepts

• Factorial design: more than one factor is studied simultaneously.

\[2^k\]  
\[\text{number of factors}\]  
\[\text{number of levels of each factor}\]

2\(^3\) design: three factors, each with two levels. Total of 8 (2\(^3\)) combinations
Two-factor Design with Equal Number of Replicates (n’)

Notation

\( r \) : number of levels of factor A  
\( c \) : number of levels of factor B  
\( n’ \) : number of replications for each cell  
\( n \) : total number of observations (\( n = rcn’ \))  
\( X_{ijk} \) : k-th observation for level i of factor A and level j of factor B
Means

\[
\bar{X} = \sum_{i=1}^{r} \sum_{j=1}^{c} \sum_{k=1}^{n'} X_{ijk} \quad \text{(overall or grand mean)}
\]

\[
\bar{X}_{i..} = \frac{\sum_{j=1}^{c} \sum_{k=1}^{n'} X_{ijk}}{cn} \quad \text{(mean of i-th level of factor A)}
\]

\[
\bar{X}_{.j} = \frac{\sum_{i=1}^{r} \sum_{k=1}^{n'} X_{ijk}}{rn} \quad \text{(mean of j-th level of factor B)}
\]

\[
\bar{X}_{ij.} = \frac{\sum_{k=1}^{n'} X_{ijk}}{n} \quad \text{(mean of cell i,j)}
\]

Area for Grand Mean

\[
\bar{X} = \frac{\sum_{i=1}^{r} \sum_{j=1}^{c} \sum_{k=1}^{n'} X_{ijk}}{rcn}
\]
Area for Mean of a Level of Factor A

\[ \bar{X}_{i..} = \frac{\sum_{j=1}^{c} \sum_{k=1}^{n} X_{ijk}}{cn} \]

<table>
<thead>
<tr>
<th>Factor B</th>
<th>1</th>
<th>2</th>
<th>...</th>
<th>c</th>
</tr>
</thead>
<tbody>
<tr>
<td>X111</td>
<td>X121</td>
<td>...</td>
<td>X1c1</td>
<td></td>
</tr>
<tr>
<td>X112</td>
<td>X122</td>
<td>...</td>
<td>X1c2</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X11n'</td>
<td>X12n'</td>
<td>...</td>
<td>X1cn'</td>
<td></td>
</tr>
<tr>
<td>X211</td>
<td>X221</td>
<td>...</td>
<td>X2c1</td>
<td></td>
</tr>
<tr>
<td>X212</td>
<td>X222</td>
<td>...</td>
<td>X2c2</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X21n'</td>
<td>X22n'</td>
<td>...</td>
<td>X2cn'</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xr11</td>
<td>Xr21</td>
<td>...</td>
<td>Xrc1</td>
<td></td>
</tr>
<tr>
<td>Xr12</td>
<td>Xr22</td>
<td>...</td>
<td>Xrc2</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xr1n'</td>
<td>Xr2n'</td>
<td>...</td>
<td>Xrcn'</td>
<td></td>
</tr>
</tbody>
</table>

Area for Mean of a Level of Factor B

\[ \bar{X}_{..j} = \frac{\sum_{i=1}^{r} \sum_{k=1}^{n} X_{ijk}}{rn} \]

<table>
<thead>
<tr>
<th>Factor B</th>
<th>1</th>
<th>2</th>
<th>...</th>
<th>c</th>
</tr>
</thead>
<tbody>
<tr>
<td>X111</td>
<td>X121</td>
<td>...</td>
<td>X1c1</td>
<td></td>
</tr>
<tr>
<td>X112</td>
<td>X122</td>
<td>...</td>
<td>X1c2</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X11n'</td>
<td>X12n'</td>
<td>...</td>
<td>X1cn'</td>
<td></td>
</tr>
<tr>
<td>X211</td>
<td>X221</td>
<td>...</td>
<td>X2c1</td>
<td></td>
</tr>
<tr>
<td>X212</td>
<td>X222</td>
<td>...</td>
<td>X2c2</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X21n'</td>
<td>X22n'</td>
<td>...</td>
<td>X2cn'</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xr11</td>
<td>Xr21</td>
<td>...</td>
<td>Xrc1</td>
<td></td>
</tr>
<tr>
<td>Xr12</td>
<td>Xr22</td>
<td>...</td>
<td>Xrc2</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xr1n'</td>
<td>Xr2n'</td>
<td>...</td>
<td>Xrcn'</td>
<td></td>
</tr>
</tbody>
</table>

© 2001 D. A. Menascé. All Rights Reserved.
Area for Mean of a Cell

\[ \bar{X}_{ij} = \frac{1}{n} \sum_{k=1}^{n} X_{ijk} \]

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>2</th>
<th>...</th>
<th>c</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X111</td>
<td>X121</td>
<td>...</td>
<td>X1c1</td>
<td></td>
</tr>
<tr>
<td>X112</td>
<td>X122</td>
<td>...</td>
<td>X1c2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X1n'</td>
<td>X12n'</td>
<td>...</td>
<td>X1cn'</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X211</td>
<td>X221</td>
<td>...</td>
<td>X2c1</td>
<td></td>
</tr>
<tr>
<td>X212</td>
<td>X222</td>
<td>...</td>
<td>X2c2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X2n'</td>
<td>X22n'</td>
<td>...</td>
<td>X2cn'</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>...</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Xr11</td>
<td>Xr21</td>
<td>...</td>
<td>Xrc1</td>
</tr>
<tr>
<td></td>
<td>Xr12</td>
<td>Xr22</td>
<td>...</td>
<td>Xrc2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Xrn'</td>
<td>Xr2n'</td>
<td>...</td>
<td>Xrcn'</td>
</tr>
</tbody>
</table>

Partitioning the Variation

Total Variation (SST)
\[ df = n - 1 \]

Factor A Variation (SSA)
\[ df = r - 1 \]

Factor B Variation (SSB)
\[ df = c - 1 \]

Interaction (SSAB)
\[ df = (r - 1)(c - 1) \]

Random Variation (SSE)
\[ df = rc(n' - 1) \]
Partitioning the Variation

**Total Variation (SST)**
\[ \text{df} = n - 1 \]

**Factor A Variation (SSA)**
\[ \text{df} = r - 1 \]

**Factor B Variation (SSB)**
\[ \text{df} = c - 1 \]

**Interaction (SSAB)**
\[ \text{df} = (r - 1)(c - 1) \]

**Random Variation (SSE)**
\[ \text{df} = rc(n' - 1) \]

**Total Variation (SST)**
\[
SST = \sum_{i=1}^{r} \sum_{j=1}^{c} \sum_{k=1}^{n'} \left( X_{ijk} - \bar{X} \right)^2
\]
Partitioning the Variation

Total Variation (SST)
\[ df = n - 1 \]

Factor A Variation (SSA)
\[ df = r - 1 \]

Factor B Variation (SSB)
\[ df = c - 1 \]

Interaction (SSAB)
\[ df = (r - 1)(c - 1) \]

Random Variation (SSE)
\[ df = rc(n' - 1) \]

Factor A Variation (SSA)
\[
SSA = cn' \sum_{i=1}^{r} \left( \bar{X}_{i..} - \bar{X} \right)^2
\]

Partitioning the Variation

Total Variation (SST)  
\[ \text{df} = n - 1 \]

Factor A Variation (SSA)  
\[ \text{df} = r - 1 \]

Factor B Variation (SSB)  
\[ \text{df} = c - 1 \]

Interaction (SSAB)  
\[ \text{df} = (r - 1)(c - 1) \]

Random Variation (SSE)  
\[ \text{df} = rc(n' - 1) \]

Factor B Variation (SSB)

\[ SSB = r n' \sum_{j=1}^{c} \left( \bar{X}_{j} - \bar{X} \right)^2 \]
Partitioning the Variation

Total Variation (SST)
\[ df = n - 1 \]

Factor A
Variation (SSA)
\[ df = r - 1 \]

Factor B
Variation (SSB)
\[ df = c - 1 \]

Interaction (SSAB)
\[ df = (r - 1)(c-1) \]

Random
Variation (SSE)
\[ df = rc(n' - 1) \]

Variation due to Interaction (SSAB)

\[ SSAB = n' \sum_{i=1}^{r} \sum_{j=1}^{c} (\bar{X}_{ij}. - \bar{X}_{i..} - \bar{X}_{.j} + \bar{X})^2 \]
Partitioning the Variation

Total Variation (SST)\[ \text{df} = n - 1 \]

Factor A Variation (SSA)\[ \text{df} = r - 1 \]

Factor B Variation (SSB)\[ \text{df} = c - 1 \]

Interaction (SSAB)\[ \text{df} = (r - 1)(c - 1) \]

Random Variation (SSE)\[ \text{df} = rc(n' - 1) \]

Random Error (SSE)

\[ SSE = \sum_{i=1}^{r} \sum_{j=1}^{c} \sum_{k=1}^{n'} (X_{ijk} - \bar{X}_{ij})^2 \]
Mean Squares

\[ MSA = \frac{SSA}{r-1} \]
\[ MSB = \frac{SSB}{c-1} \]
\[ MSAB = \frac{SSAB}{(r-1)(c-1)} \]
\[ MSE = \frac{SSE}{rc(n-1)} \]

Two-Factor ANOVA Model

No Difference Due to Factor A

\[ H_0 : \mu_{1..} = \mu_{2..} = \ldots = \mu_{r..} \]
\[ H_1 : \text{Not all} \mu_{i..} (i = 1, \ldots, r) \text{ are equal.} \]

F-Test statistic for Factor A: \[ F = \frac{MSA}{MSE} \]

The F-test statistic follows an F distribution with (r-1) degrees of freedom in the numerator and rc(n'-1) in the denominator.

Reject \( H_0 \) if \( F > Fu \)
Two-Factor ANOVA Model

No Difference Due to Factor B

\[ H_0 : \mu_{1,1}, \mu_{1,2}, \ldots, \mu_{1,c} = \ldots = \mu_{c,c}. \]

\[ H_1 : \text{Not all} \mu_{j,i}, (j = 1, \ldots, c) \text{ are equal.} \]

F-Test statistic for Factor B: \[ F = \frac{MSB}{MSE} \]

The F-test statistic follows an F distribution with \((c-1)\) degrees of freedom in the numerator and \(rc(n'-1)\) in the denominator.

Reject \( H_0 \) if \( F > Fu \)

---

Two-Factor ANOVA Model

No Interaction of Factors A and B

\[ H_0 : \text{the interaction of A and B is 0.} \]

\[ H_1 : \text{the interaction of A and B \( \neq \) 0.} \]

F-Test statistic for the interaction: \[ F = \frac{MSAB}{MSE} \]

The F-test statistic follows an F distribution with \((r-1)(c-1)\) degrees of freedom in the numerator and \(rc(n'-1)\) in the denominator.

Reject \( H_0 \) if \( F > Fu \)
Example of Two Factor Design Analysis

Response time (in msec) of a Web Site.

<table>
<thead>
<tr>
<th></th>
<th>1 CPU</th>
<th>2 CPUs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Server</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Server</td>
<td>101.0</td>
<td>98.0</td>
</tr>
<tr>
<td>1 Server</td>
<td>103.0</td>
<td>97.5</td>
</tr>
<tr>
<td>1 Server</td>
<td>102.4</td>
<td>99.3</td>
</tr>
<tr>
<td>1 Server</td>
<td>104.0</td>
<td>100.0</td>
</tr>
<tr>
<td>2 Servers</td>
<td>43.0</td>
<td>41.0</td>
</tr>
<tr>
<td>2 Servers</td>
<td>46.0</td>
<td>44.0</td>
</tr>
<tr>
<td>2 Servers</td>
<td>45.0</td>
<td>42.0</td>
</tr>
<tr>
<td>2 Servers</td>
<td>49.0</td>
<td>46.0</td>
</tr>
</tbody>
</table>

Anova: Two-Factor With Replication

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P-value</th>
<th>F crit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Servers</td>
<td>12611.29</td>
<td>1</td>
<td>12611.29</td>
<td>3586.149</td>
<td>3.11E-16</td>
<td>4.747221</td>
</tr>
<tr>
<td>Number of CPUs</td>
<td>40.96</td>
<td>1</td>
<td>40.96</td>
<td>11.64739</td>
<td>0.005146</td>
<td>4.747221</td>
</tr>
<tr>
<td>Interaction</td>
<td>1.96</td>
<td>1</td>
<td>1.96</td>
<td>0.557346</td>
<td>0.469706</td>
<td>4.747221</td>
</tr>
<tr>
<td>Within</td>
<td>42.2</td>
<td>12</td>
<td>3.51667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>12696.41</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Reject Hypothesis that there is no difference due to number of servers.
Reject Hypothesis that there is no difference due to number of CPUs.
Accept hypothesis that there is no interaction between number of servers and number of CPUs.
Lines are almost parallel: no interaction between factors!

Lines are not parallel: factors interact!