Factorial Designs

Two (or more) factors, say $A$ and $B$, with $a$ and $b$ levels, respectively.

A factorial design uses all $ab$ combinations of levels of $A$ and $B$, for a total of $ab$ treatments.

When both (all) factors have 2 levels, we have a $2 \times 2$ ($2^k$) design.
E.g. a $2 \times 2$ experiment:

<table>
<thead>
<tr>
<th>Factor A</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor B</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>30</td>
</tr>
</tbody>
</table>

Main effect of $A$ is

Average response for high level of $A$  
= $\frac{40 + 52}{2} - \frac{20 + 30}{2} = 21$

Similarly, main effect of $B$ is 11.
Interaction

The *simple* effect of $A$ at $B^-$ is $40 - 20 = 20$, and the simple effect of $A$ at $B^+$ is $52 - 30 = 22$.

The difference between these simple effects is the *interaction* $AB$ (actually, one half the difference).

Graphical presentation: the interaction plot.
In R; lines are parallel if interaction is 0:

```r
twobytwo <- data.frame(A = c("-", "+", "-", "+") ,
                        B = c("+", "+", "-", "-") ,
                        y = c(30, 52, 20, 40))
interaction.plot(twobytwo$A, twobytwo$B, twobytwo$y)
# or, saving some typing:
with(twobytwo, interaction.plot(A, B, y))
```

![Graph showing interaction plot](image-url)
The other way; lines are still parallel if interaction is 0:

\[
\text{with(twobytwo, interaction.plot(B, A, y))}
\]
Interaction with Two Quantitative Factors

Write the general level of factor $A$ as $x_1$, the general level of factor $B$ as $x_2$, and the response as $y$.

Code $x_1$ so that $x_1 = -1$ at $A^-$ and $x_1 = +1$ at $A^+$, and $x_2$ similarly.

Estimate the regression model representation

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{1,2} x_1 x_2 + \epsilon.$$
Least squares estimates of the $\beta$’s can be found from the main effects and interaction.

For the simple example:

```
twobytwo$x1 <- ifelse (twobytwo$A == "+", 1, -1)
twobytwo$x2 <- ifelse (twobytwo$B == "+", 1, -1)
summary(lm(y ~ x1 * x2, twobytwo))
```

from which we get

$$\hat{y} = 35.5 + 10.5x_1 + 5.5x_2 + 0.5x_1x_2.$$ 

A graph of $\hat{y}$ versus $x_1$ and $x_2$ is called a response surface plot.
Factorial Designs
Definitions and Principles
Making the Response Surface Plot in R

Use `expand.grid()` to set up a grid of values of $x_1$ and $x_2$, use the `predict()` method to evaluate $\hat{y}$ on the grid, and use `persp()` to make a surface plot of it:

```r
ngrid <- 20
x1 <- with(twobytwo, seq(min(x1), max(x1), length = ngrid))
x2 <- with(twobytwo, seq(min(x2), max(x2), length = ngrid))
grid <- expand.grid(x1 = x1, x2 = x2)
yhat <- predict(lm(y ~ x1 * x2, twobytwo), grid)
yhat <- matrix(yhat, length(x1), length(x2))
persp(x1, x2, yhat, theta = 25, expand = 0.75, ticktype = "detailed")
```
## A Two-Factor Example

Battery life data; $A = \text{Material}, \ a = 3, \ B = \text{Temperature}, \ b = 3,$ replications $n = 4$ (battery-life.txt):

<table>
<thead>
<tr>
<th>Material</th>
<th>Temperature</th>
<th>Life</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15</td>
<td>130</td>
</tr>
<tr>
<td>1</td>
<td>15</td>
<td>155</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>125</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>15</td>
<td>150</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>125</td>
<td>45</td>
</tr>
<tr>
<td>3</td>
<td>15</td>
<td>138</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>125</td>
<td>60</td>
</tr>
</tbody>
</table>
boxplot(Life ~ factor(Material) * factor(Temperature),
batteryLife)
Statistical model

\[ y_{i,j,k} = \mu + \tau_i + \beta_j + (\tau \beta)_{i,j} + \epsilon_{i,j,k} \]

is the response for Material \( i \), Temperature \( j \), replicate \( k \).

\( \tau \)'s, \( \beta \)'s, and \( (\tau \beta) \)'s satisfy the usual constraints (natural or computer).

Hypotheses

- No differences among Materials; \( H_0 : \tau_i = 0 \), all \( i \);
- No effect of Temperature; \( H_0 : \beta_j = 0 \), all \( j \);
- No interaction; \( H_0 : (\tau \beta)_{i,j} = 0 \), all \( i \) and \( j \).
R command

batteryLife <- read.table("data/battery-life.txt", header = TRUE)
summary(aov(Life ~ factor(Material) * factor(Temperature),
        batteryLife))

Output

<table>
<thead>
<tr>
<th>Df</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F value</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>factor(Material)</td>
<td>2</td>
<td>10684</td>
<td>5342</td>
<td>7.9114</td>
</tr>
<tr>
<td>factor(Temperature)</td>
<td>2</td>
<td>39119</td>
<td>19559</td>
<td>28.9677</td>
</tr>
<tr>
<td>factor(Material):factor(Temperature)</td>
<td>4</td>
<td>9614</td>
<td>2403</td>
<td>3.5595</td>
</tr>
</tbody>
</table>

Residuals 27 18231 675

---

Signif. codes: 0 *** 0.001 ** 0.01 * 0.05 . 0.1 1
Interaction plots

Can be made in two ways:

- Lifetime versus Temperature, with one curve for each type of Material;
- Lifetime versus Material, with one curve for each level of Temperature.

Same information either way, but usually easier to interpret with a quantitative factor on the X-axis.

Here Temperature is quantitative, but Material is qualitative.
with(batteryLife, interaction.plot(Temperature, Material, Life,
  type = "b"))
with(batteryLife, interaction.plot(Material, Temperature, Life, type = "b"))
### Pairwise comparisons

```r
TukeyHSD(aov(Life ~ factor(Material) * factor(Temperature),
             batteryLife))
```

### Output

Tukey multiple comparisons of means
95% family-wise confidence level

Fit: `aov(formula = Life ~ factor(Material) * factor(Temperature),
         data = batteryLife)`

<table>
<thead>
<tr>
<th><code>factor(Material)</code></th>
<th>diff</th>
<th>lwr</th>
<th>upr</th>
<th>p adj</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-1</td>
<td>25.16667</td>
<td>-1.135677</td>
<td>51.46901</td>
<td>0.0627571</td>
</tr>
<tr>
<td>3-1</td>
<td>41.91667</td>
<td>15.614323</td>
<td>68.21901</td>
<td>0.0014162</td>
</tr>
<tr>
<td>3-2</td>
<td>16.75000</td>
<td>-9.552344</td>
<td>43.05234</td>
<td>0.2717815</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><code>factor(Temperature)</code></th>
<th>diff</th>
<th>lwr</th>
<th>upr</th>
<th>p adj</th>
</tr>
</thead>
<tbody>
<tr>
<td>70-15</td>
<td>-37.25000</td>
<td>-63.55234</td>
<td>-10.94766</td>
<td>0.0043788</td>
</tr>
<tr>
<td>125-15</td>
<td>-80.66667</td>
<td>-106.96901</td>
<td>-54.36432</td>
<td>0.0000001</td>
</tr>
<tr>
<td>125-70</td>
<td>-43.41667</td>
<td>-69.71901</td>
<td>-17.11432</td>
<td>0.0009787</td>
</tr>
</tbody>
</table>
$\text{'factor(Material):factor(Temperature)'}$

<table>
<thead>
<tr>
<th></th>
<th>diff</th>
<th>lwr</th>
<th>upr</th>
<th>p</th>
<th>adj</th>
</tr>
</thead>
<tbody>
<tr>
<td>2:15-1:15</td>
<td>21.00</td>
<td>-40.823184</td>
<td>82.823184</td>
<td>0.9616404</td>
<td></td>
</tr>
<tr>
<td>3:15-1:15</td>
<td>9.25</td>
<td>-52.573184</td>
<td>71.073184</td>
<td>0.9998527</td>
<td></td>
</tr>
<tr>
<td>1:70-1:15</td>
<td>-77.50</td>
<td>-139.323184</td>
<td>-15.676816</td>
<td>0.0065212</td>
<td></td>
</tr>
<tr>
<td>2:70-1:15</td>
<td>-15.00</td>
<td>-76.823184</td>
<td>46.823184</td>
<td>0.9953182</td>
<td></td>
</tr>
<tr>
<td>3:70-1:15</td>
<td>11.00</td>
<td>-50.823184</td>
<td>72.823184</td>
<td>0.9994703</td>
<td></td>
</tr>
<tr>
<td>1:125-1:15</td>
<td>-77.25</td>
<td>-139.073184</td>
<td>-15.426816</td>
<td>0.0067471</td>
<td></td>
</tr>
<tr>
<td>2:125-1:15</td>
<td>-85.25</td>
<td>-147.073184</td>
<td>-23.426816</td>
<td>0.0022351</td>
<td></td>
</tr>
<tr>
<td>3:125-1:15</td>
<td>-49.25</td>
<td>-111.073184</td>
<td>12.573184</td>
<td>0.2016535</td>
<td></td>
</tr>
<tr>
<td>3:15-2:15</td>
<td>-11.75</td>
<td>-73.573184</td>
<td>50.073184</td>
<td>0.9991463</td>
<td></td>
</tr>
<tr>
<td>1:70-2:15</td>
<td>-98.50</td>
<td>-160.323184</td>
<td>-36.676816</td>
<td>0.0003449</td>
<td></td>
</tr>
<tr>
<td>2:70-2:15</td>
<td>-36.00</td>
<td>-97.823184</td>
<td>25.823184</td>
<td>0.5819453</td>
<td></td>
</tr>
<tr>
<td>3:70-2:15</td>
<td>-10.00</td>
<td>-71.823184</td>
<td>51.823184</td>
<td>0.9997369</td>
<td></td>
</tr>
<tr>
<td>1:125-2:15</td>
<td>-98.25</td>
<td>-160.073184</td>
<td>-36.426816</td>
<td>0.0003574</td>
<td></td>
</tr>
<tr>
<td>2:125-2:15</td>
<td>-106.25</td>
<td>-168.073184</td>
<td>-44.426816</td>
<td>0.0001152</td>
<td></td>
</tr>
<tr>
<td>3:125-2:15</td>
<td>-70.25</td>
<td>-132.073184</td>
<td>-8.426816</td>
<td>0.0172076</td>
<td></td>
</tr>
<tr>
<td>1:70-3:15</td>
<td>-86.75</td>
<td>-148.573184</td>
<td>-24.926816</td>
<td>0.0018119</td>
<td></td>
</tr>
<tr>
<td>2:70-3:15</td>
<td>-24.25</td>
<td>-86.073184</td>
<td>37.573184</td>
<td>0.9165175</td>
<td></td>
</tr>
</tbody>
</table>

...
Residual Plots

\[
\text{plot(aov(Life ~ factor(Material) * factor(Temperature), batteryLife))};
\]
Theoretical Quantiles
Standardized residuals
aov(Life ~ factor(Material) * factor(Temperature))
Factors Design

Two-Factor Design