UFMG
UNIVERSIDADE FEDERAL
DE MINAS GERAIS

Design and Analysis of Experiments

12 - Factorial Designs

Version 2.11

Felipe Campelo
http://www.cpdee.ufmg.br/~fcampelo

Graduate Program in Electrical Engineering

Belo Horizonte
May 2015

“*We did not evolve to understand or comprehend reality. We evolved to survive it. For understanding we need science.*”

Mark A. Crislip
1952-
American infectologist

Image: `http://boards.medscape.com/.29f3af03/`

Factorial Designs

Basic definitions

Many experiments involve more than a single *factor* of interest - that is, multiple independent variables that can influence a response variable.

In general, an effective way to explore the main effects and interactions of multiple factors is the use of a *factorial design* in which all level combinations are evaluated at each experimental replicate;

In this context, the **main effect** of a factor quantifies the mean change in the response variable due to changing between the levels of that factor;

An **interaction effect** represents the mean change in the response variable due to the simultaneous change of levels of two or more factors.

Factorial Designs

Example: Electrical current in motors

Two engineers wish to investigate factors that may affect the electrical current demanded by the single-phase motors used for ventilation in an industrial chicken coop.

Previous observations suggest that the current varies considerably from motor to motor, and process knowledge suggests two likely candidates for explaining this variability: the *Manufacturer* (A, B or C) and the *State* (original or rewinded) of each motor.

To investigate this question, the engineers decide to sample 40 motors from each manufacturer, with 20 in the original state and 20 being rewinded motors.

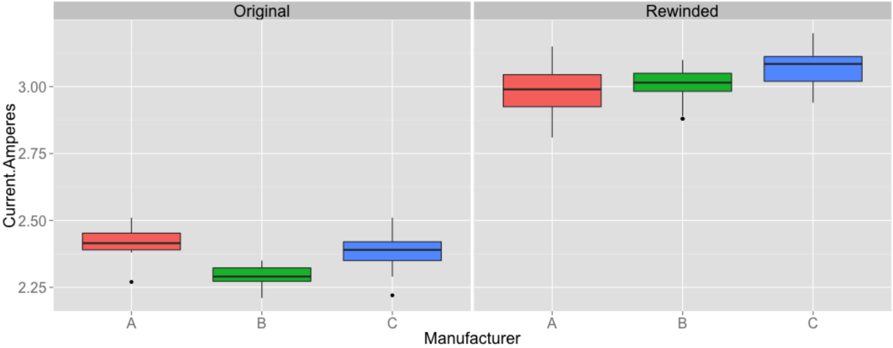
Adapted from M.H.Costa and T.L. Vieira's course project for the Design and Analysis of Experiments Course, PPGEE-UFMG, November 2013. The data used in this example is not necessarily the original one.

Image: `http://refrigelms.com.br/ventilador-para-aviario-qla85-grade-p-1734.html`

Exploratory data analysis

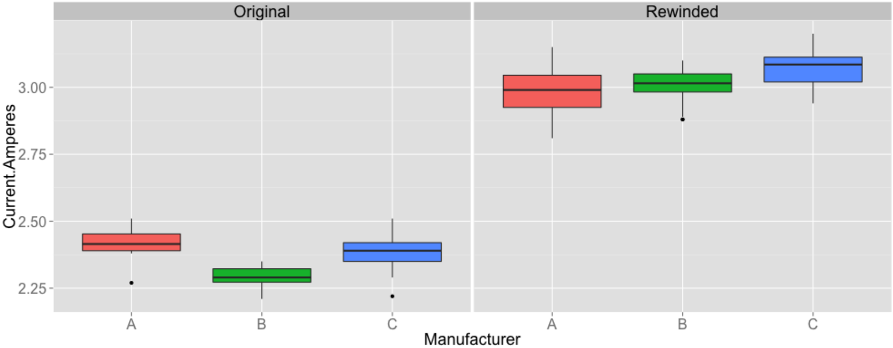
Example: Electrical current in motors

```
> data <- read.table("../data files/motors.txt", header = TRUE)
> library(ggplot2)
> p <- ggplot(data, aes(x = Manufacturer, y = Current.Amperes,
                        fill = Manufacturer))
> p + geom_boxplot() + facet_grid(.~State) + ...
```



The exploratory plot suggests a relatively large effect for the *State* factor, but is inconclusive with regards to the *Manufacturer* effect. Any interaction effect is also likely to by small.

Lets assume for this example that the engineers want $\alpha = 0.05$, $\beta = 0.2$ and $\delta^* = 0.1A$.



Factorial Designs

Statistical model for two factors

In the general case for a completely randomized factorial design we have:

- a levels for factor **A**;
- b levels for the factor **B**;
- n replicates within each combination of levels;
- Completely randomized collection of observations;

The effects model for a set of observations collected following this design can be expressed as:

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijk} \begin{cases} i = 1, \ldots, a \\ j = 1, \ldots, b \\ k = 1, \ldots, n \end{cases}$$

Factorial Designs

Statistical model for two factors

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijk} \begin{cases} i = 1, \ldots, a \\ j = 1, \ldots, b \\ k = 1, \ldots, n \end{cases}$$

As before, effects are treated as deviations from the grand mean.
By construction:

$$\sum_{i=1}^{a} \tau_i = 0 \qquad \sum_{j=1}^{b} \beta_j = 0$$

$$\sum_{i=1}^{a} (\tau\beta)_{ij} = \sum_{j=1}^{b} (\tau\beta)_{ij} = 0$$

The factorial design emerges whenever we wish to model both the main and the interaction effects of multiple factors. This means that, for the two-factor case, the hypotheses that can be tested are:

$$\text{Factor } A\text{, main effect:} \begin{cases} H_0 : \tau_i = 0,\ \forall i \\ H_1 : \exists \tau_i \neq 0 \end{cases}$$

$$\text{Factor } B\text{, main effect:} \begin{cases} H_0 : \beta_j = 0,\ \forall j \\ H_1 : \exists \beta_j \neq 0 \end{cases}$$

$$\text{Interaction effect, } AB\text{:} \begin{cases} H_0 : (\tau\beta)_{ij} = 0,\ \forall i, j \\ H_1 : \exists (\tau\beta)_{ij} \neq 0 \end{cases}$$

The test statistics for these hypotheses will, as usual, be derived from the partition of the total variability into specific components:

$$
\begin{aligned}
SST &= \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(y_{ijk}-\bar{y}_{\cdots}\right)^2 \\
&= \underbrace{bn\sum_{i=1}^{a}(\bar{y}_{i\cdot\cdot}-\bar{y}_{\cdots})^2}_{SS_A} + \underbrace{an\sum_{j=1}^{b}\left(\bar{y}_{\cdot j\cdot}-\bar{y}_{\cdots}\right)^2}_{SS_B} \\
&\quad + \underbrace{n\sum_{i=1}^{a}\sum_{j=1}^{b}\left(\bar{y}_{ij\cdot}-\bar{y}_{i\cdot\cdot}-\bar{y}_{\cdot j\cdot}+\bar{y}_{\cdots}\right)^2}_{SS_{AB}} + \underbrace{\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(y_{ijk}-\bar{y}_{ij\cdot}\right)^2}_{SS_E}
\end{aligned}
$$

To distinguish SS_{AB} from SS_E we need $n \geq 2$.

The mean squares are also calculated as usual:

$$MS_A = \frac{SS_A}{a-1} \qquad E\left[MS_A\right] = \sigma^2 + \frac{bn\sum_{i=1}^{a}\tau_i^2}{a-1}$$

$$MS_B = \frac{SS_B}{b-1} \qquad E\left[MS_B\right] = \sigma^2 + \frac{an\sum_{j=1}^{b}\beta_j^2}{b-1}$$

$$MS_{AB} = \frac{SS_{AB}}{(a-1)(b-1)} \qquad E\left[MS_{AB}\right] = \sigma^2 + \frac{n\sum_{i=1}^{a}\sum_{j=1}^{b}(\tau\beta)_{ij}^2}{(a-1)(b-1)}$$

$$MS_E = \frac{SS_E}{ab(n-1)} \qquad E\left[MS_E\right] = \sigma^2$$

Factorial Designs

Statistical model for two factors

If the usual assumptions (ϵ_{ijk} i.i.d. $\mathcal{N}(0, \sigma^2)$) hold, the fractions:

$$F_0^{(A)} = \frac{MS_A}{MS_E}$$

$$F_0^{(B)} = \frac{MS_B}{MS_E}$$

$$F_0^{(AB)} = \frac{MS_{AB}}{MS_E}$$

are distributed under their respective null hypotheses as F variables (each with their respective degrees of freedom), and the hypotheses can be tested in the usual manner (i.e., comparing the obtained value of F_0 against the critical value of $F_{\alpha;df_1;df_2}$).

Example: Electrical current in motors

Statistical model for two factors

```
> model <- aov(Current.Amperes~State*Manufacturer,
+              data = data)
> summary(model)
                    Df Sum Sq Mean Sq F value   Pr(>F)
State                1 12.956  12.956 2798.41  < 2e-16 ***
Manufacturer         2  0.118   0.059   12.71 1.04e-05 ***
State:Manufacturer   2  0.114   0.057   12.27 1.49e-05 ***
Residuals          114  0.528   0.005
---

> summary.lm(model)$r.squared
[1] 0.9615174
```

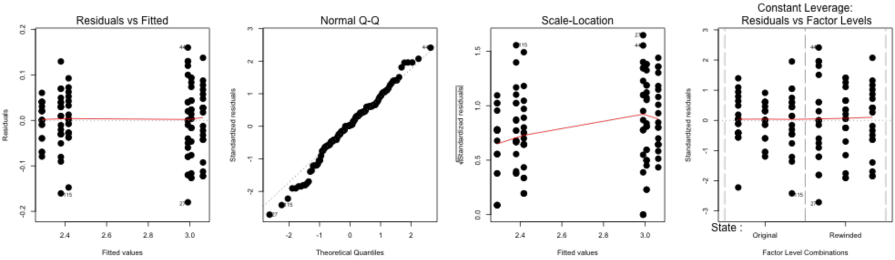
Example: Electrical current in motors

Statistical model for two factors

As usual, the assumptions can be verified by means of residual analysis, like in the one-way ANOVA (except for a little adjustment needed for the Fligner-Killeen test)

```
> shapiro.test(model$residuals)
W = 0.9857, p-value = 0.2392

> fligner.test(Current.Amperes ~ interaction(State, Manufacturer),
+                 data = data)
med chi-squared = 10.1721, df = 5, p-value = 0.0705
```



If the ANOVA indicates the existence of significant effects, we can perform pairwise comparisons between levels to investigate specific differences;

When the interaction effect is not significant, the comparisons between factor levels can be done in a straightforward manner, using the estimated level means. For instance, the test statistic for comparing the means of levels 2 and 3 of factor A could be calculated as:

$$t_0 = \frac{\bar{y}_{2\cdot\cdot} - \bar{y}_{3\cdot\cdot}}{\sqrt{2\frac{MS_E}{n'}}}$$

where n' is the number of specific replicates for the comparison under consideration.

More generally,

$$t_0 = \frac{\Delta \bar{y}}{\sqrt{2\frac{MS_E}{n'}}}$$

For comparisons of factor levels (main effects), the value of n' is the total number of observations under that level;

For comparisons of level combinations (interaction effects), it is the number of observations within each combination group;

```
> replications(Current ~ State*Manufacturer,
+                 data = data)
              State       Manufacturer State:Manufacturer
                 60                 40                 20
```

Also, the α value for the comparisons has to be adjusted to prevent inflation of the type-I error rate.

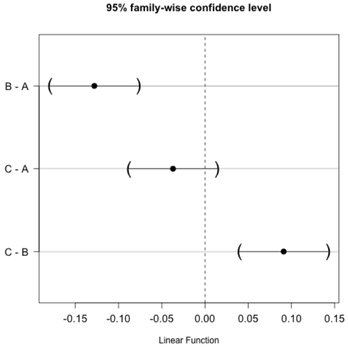
Factorial designs

Multiple Comparisons

The usual routines for performing multiple comparisons in *R* are applicable. For instance, performing *all vs all* comparisons using Tukey's method yields, for the *Manufacturer* factor:

```
> mcp.manuf <- glht(model, linfct = mcp(Manufacturer = "Tukey"))
Warning message:
In mcp2matrix(model, linfct = linfct) :
  covariate interactions found
  -- default contrast might be inappropriate

> plot(confint(mcp.manuf),
+      cex.axis   = 1.2,
+      cex        = 2)
```



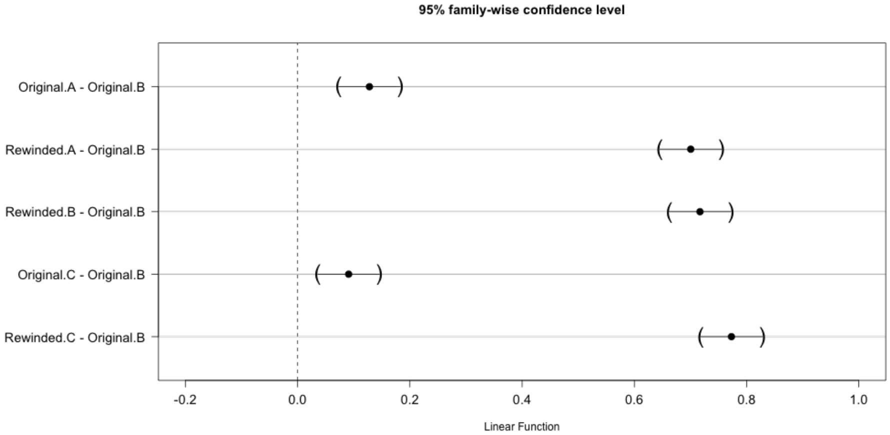
The comparison of means for the interaction groups requires a little more work, but nothing too complex. Here, we assume that we want to compare all groups versus the one with the smallest sample average.

```
> # Create meta-factor for interaction groups
> interfac <- with(data,
+                  interaction(State, Manufacturer))
>
> # Use group with the smallest sample mean as the reference
> with(data, which.min(tapply(Current.Amperes, interfac, mean)))
Original.B
         3
> interfac <- relevel(interfac, ref = "Original.B")
>
> # ReFit model
> model2 <- aov(Current.Amperes ~ interfac,
+               data = data)
>
> # Multiple comparisons
> mcp.inter <- glht(model2,
+                 linfct = mcp(interfac = "Dunnett"))
```

Factorial designs

Multiple Comparisons

```
> par(mar = c(5,12,4,2))
> plot(confint(mcp.inter), xlim = c(-0.2, 1), cex.axis = 1.2, cex = 2)
```



For this example, the engineers would have now enough data to draw recommendations. For example, the data clearly shows that rewinded motors result in much larger currents drawn, which results in extra operational and structural (wiring, protection equipment, etc.) costs.

The engineers will factor economic and performance factors to reach conclusions about whether they should start gradually replacing all ventilation motors for new ones from manufacturer B, and whether it is better to fix or scrap the motors in need of rewinding.

General Factorial Designs

Experiments with more than 2 factors

In the general case, the factorial design assumes:

- a levels of the factor **A**;
- b levels of the factor **B**;
- c levels of the factor **C**;
- $\ldots$

If we consider an experiment with $n \geq 2$ replicates, the total number of observations required is given by $abc \ldots n$;

It is actually possible to design an experiment with a single replicate (particularly for larger designs). This will be discussed later.

General Factorial Designs

Experiments with more than 2 factors

The modeling and analysis of these experiments are easily obtained from the generalization of the design with 2 factors. For example, for 3 factors we have:

$$
\begin{aligned}
y_{ijkl} &= \mu && \longleftarrow \text{Grand mean} \\
&\quad + \tau_i + \beta_j + \gamma_k && \longleftarrow \text{Main effects} \\
&\quad + (\tau\beta)_{ij} + (\tau\gamma)_{ik} + (\beta\gamma)_{jk} && \longleftarrow \text{2nd order interactions} \\
&\quad + (\tau\beta\gamma)_{ijk} && \longleftarrow \text{3rd order interaction} \\
&\quad + \epsilon_{ijkl} && \longleftarrow \text{Residual} \\
&i = 1, \ldots, a; \\
&j = 1, \ldots, b; \\
&k = 1, \ldots, c; \\
&l = 1, \ldots, n;
\end{aligned}
$$

General Factorial Designs

Sum of squares: total and main effects

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{n} y_{ijkl}^2 - \frac{y_{....}^2}{abcn}$$

$$SS_A = \frac{1}{bcn}\sum_{i=1}^{a} y_{i...}^2 - \frac{y_{....}^2}{abcn}$$

$$SS_B = \frac{1}{acn}\sum_{j=1}^{b} y_{.j..}^2 - \frac{y_{....}^2}{abcn}$$

$$SS_C = \frac{1}{abn}\sum_{k=1}^{c} y_{..k.}^2 - \frac{y_{....}^2}{abcn}$$

General Factorial Designs

Sum of squares: 2nd order interactions

$$SS_{AB} = \frac{1}{cn} \sum_{i=1}^{a} \sum_{j=1}^{b} y_{ij\cdot\cdot}^2 - \frac{y_{\cdot\cdot\cdot\cdot}^2}{abcn} - SS_A - SS_B$$

$$SS_{AC} = \frac{1}{bn} \sum_{i=1}^{a} \sum_{k=1}^{c} y_{i\cdot k\cdot}^2 - \frac{y_{\cdot\cdot\cdot\cdot}^2}{abcn} - SS_A - SS_C$$

$$SS_{BC} = \frac{1}{an} \sum_{j=1}^{b} \sum_{k=1}^{c} y_{\cdot jk\cdot}^2 - \frac{y_{\cdot\cdot\cdot\cdot}^2}{abcn} - SS_B - SS_C$$

General Factorial Designs

Sum of squares: 3rd order interaction and residual

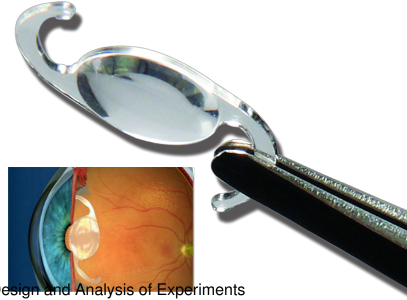
$$
\begin{aligned}
SS_{ABC} =& \frac{1}{n} \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{c} y_{ijk\cdot}^2 - \frac{y_{\cdots\cdot}^2}{abcn} \\
&- SS_A - SS_B - SS_C \\
&- SS_{AB} - SS_{AC} - SS_{BC}
\end{aligned}
$$

$$
\begin{aligned}
SS_E =& SS_T \\
&- SS_A - SS_B - SS_C \\
&- SS_{AB} - SS_{AC} - SS_{BC} \\
&- SS_{ABC}
\end{aligned}
$$

General Factorial Designs

Example: intraocular lenses

The standard surgical intervention for the treatment of cataracts consists in the removal of the crystalline lens and implantation of an artificial intraocular lens (IOL). IOLs are generally manufactured using a high precision CNC lathe, in which a circular piece of biocompatible material is carved to the desired lens shape with a diamond cutting tool.



Adapted from L.M. Carvalho e D.F. Filgueiras' course project for the Design and Analysis of Experiments Course, PPGEE-UFMG, June 2013. The data used in this example is not necessarily the original one.

Eye image: http://www.peruenvideos.com/implante-lentes-intraoculares-curacion-cataratas/

Lens image: http://www.allaboutvision.com/conditions/iols.htm

Before being marketed each lens is tested for the compliance of their optical properties, and the ones that fail to meet the required specifications are discarded.

Based on their knowledge of this process, two engineers designed an experiment for the preliminary investigation of the influential factors on the percentage of lenses that meet the specifications.

Three factors were selected for this preliminary study, each one with two levels. The resources allocated to the study were enough for the execution of exactly eight batches of lenses - in other words, a *single replicate* for each combination of levels.

Factors and levels:

- Lathe time (in minutes): $[2.35; 3.15]$;
- Polishing time (in days): $[5; 7]$;
- Age of the cutting tool (in cycles): $[\approx 400; \approx 1200]$;

For each combination of levels a batch of 30 lenses was produced, and the proportion of lenses in conformity with specification was recorded as the response variable.

The experiment was conducted in a completely randomized way, with partial blinding (lathe operators and technical inspectors did not know which level combination they were dealing with).

The significance level was set as $\alpha = 0.05$, and the researchers were interested in detecting any effects equal or larger than 0.1 with a power of 0.8.

General Factorial Designs

Example: intraocular lenses

Since there is only one replicate, there are not enough degrees of freedom to calculate MS_E. Consequently, the test of hypotheses becomes unfeasible.

```
> data <- read.table("../data files/lio.txt", header = TRUE)

> model<-aov(Conf.rate ~ .^3, data = data)
> summary(model)
                                        Df Sum Sq Mean Sq
CNCTime.min                              1 0.5151  0.5151
PolTime.days                             1 0.0861  0.0861
ToolAge.cycles                           1 0.0105  0.0105
CNCTime.min:PolTime.days                 1 0.0036  0.0036
CNCTime.min:ToolAge.cycles               1 0.0001  0.0001
PolTime.days:ToolAge.cycles              1 0.0001  0.0001
CNCTime.min:PolTime.days:ToolAge.cycles  1 0.0036  0.0036
```

To perform the test we need some degrees of freedom for the error term. In cases with single replicates, the most usual way of doing this is by discarding low-influence terms from the model. But which ones should be discarded?

A good way to proceed in these cases is to start by removing the highest-order interactions from the model, so that these terms are absorbed for the calculation of MS_E.

This heuristic is based on the *sparsity principle*, which states that most systems are dominated by main effects and low-order interactions;

A qualitative way of verifying the possibility of excluding some effects is the examination of a plot known as *Daniel's effects plot*, which consists on plotting effect estimators obtained from a *saturated* model on a normal QQ plot.

Strong effects will appear as outliers, while weak or insignificant effects will apper around the expected Normal line. By examining this plot we can obtain a simplified model, containing only the relevant effects.

Daniel plots work only in designs with only 2 levels per factor (2^k *designs*).

For a more general effects plot, check Whitcomb and Oehlert (2007), *Graphical Selection of Effects in General Factorials*: `http://goo.gl/6dw7dn`

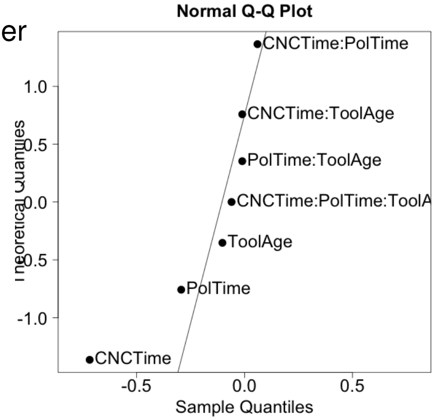
Example: intraocular lenses

Model simplification

```
> effect.est <- as.numeric(model$effects[-1])
> qq.text    <- rownames(summary.aov(model)[[1]])
> qq.obj     <- qqnorm(effect.est, datax = TRUE, ...)
> qqline(effect.est, datax = TRUE)
> text(qq.obj$x, qq.obj$y, labels = qq.text,...)
```

The effects plot suggests that the higher order effects have little influence over the response variable;

Factor **CNCTime** seems to be the most important, with **PolTime** also a possibly interesting effect.



Example: intraocular lenses

Model simplification

Discarding the interaction effects, we can suggest a simplified model:

```
> model2 <- aov(Conf.rate ~ ., data = data)
> summary(model2)
            Df Sum Sq Mean Sq F value   Pr(>F)
CNCTime      1 0.5151  0.5151 276.570 7.66e-05 ***
PolTime      1 0.0861  0.0861  46.235  0.00244 **
ToolAge      1 0.0105  0.0105   5.644  0.07634 .
Residuals    4 0.0074  0.0019

> summary.lm(model2)$r.squared
[1] 0.9879681

> shapiro.test(model2$residuals)
W = 0.9271, p-value = 0.4902
```

Example: intraocular lenses

Exploring Specific Differences

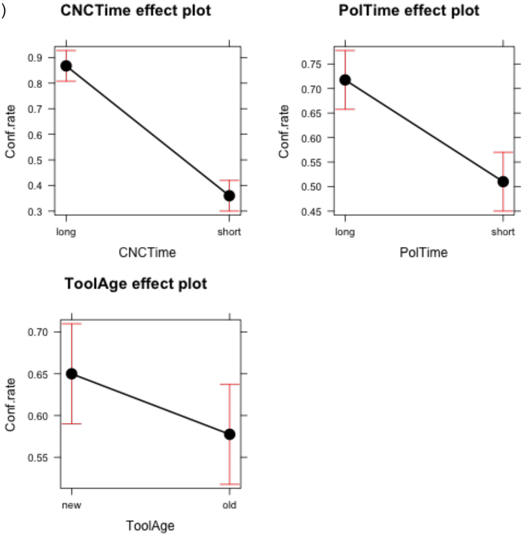
```
> summary(glht(model2, linfct = mcp(CNCTime = "Tukey")))
Linear Hypotheses:
                  Estimate Std. Error t value Pr(>|t|)
short - long == 0 -0.50750    0.03052  -16.63 7.66e-05 ***

> summary(glht(model2, linfct = mcp(PolTime = "Tukey")))
Linear Hypotheses:
                  Estimate Std. Error t value Pr(>|t|)
short - long == 0 -0.20750    0.03052    -6.8  0.00244 **
```

Example: intraocular lenses

Exploring Specific Differences

```
> library(effects)
> lio.effs <- allEffects(model2)
> plot(lio.effs)
```



The effect of greatest impact on the quality of the process if the lathe time. The proportion of lenses in conformity with the specifications goes from 0.36 to 0.87, which strongly suggests the use of larger lathe times as a good strategy.

The polishing time also presented a significant impact, with a jump from 0.51 (5 days) to 0.71 (7 days) on the proportion of compliant lenses.

No significant difference was detected between “old” and “new” cutting tools. This may have been due to an absence of effect, or due to the low sample size employed in this test.

It is probably interesting to explore CNC lathe times further, and to include manufacturing cost considerations into this discussion.

The inclusion of blocking variables in factorial designs essentially as simple as the single-factor case.

The RCBD will contain one full experimental replicate per block. The modeling and analysis aspects can be easily derived from the last two chapters.

Bibliography

Required reading

1. D.C. Montgomery, G.C. Runger, *Applied Statistics and Probability for Engineers*, 3rd ed., 2003 - Chapter 14;
2. P. Hoff, *Applied Statistics and Experimental Design*, Chapter 6 (Factorial Designs), `http://goo.gl/NiyVCX`

Recommended reading

1. R. Feynman, *Surely you're joking, Mr. Feynman*, W.W. Norton&Company, 1997.
2. H. Wickham, *ggplot2: Elegant Graphics for Data Analysis*, Springer 2009.

About this material

Conditions of use and referencing

Please reference this work as:

Felipe Campelo (2015), *Lecture Notes on Design and Analysis of Experiments*. Online: `https://github.com/fcampelo/Design-and-Analysis-of-Experiments` Version 2.11, Chapter 12; Creative Commons BY-NC-SA 4.0.

```
@Misc{Campelo2015-01,
  title={Lecture Notes on Design and Analysis of Experiments},
  author={Felipe Campelo},
  howPublished={\url{https://github.com/fcampelo/Design-and-Analysis-of-Experiments}},
  year={2015},
  note={Version 2.11, Chapter 12; Creative Commons BY-NC-SA 4.0.},
}
```