Case Study : Youden Square or Incomplete latin Square

Prof L. Gentzbittel Skoltech, Project Center for Agro Technologies *

Prof C. Ben, Skoltech, Project Center for Agro Technologies †

Oct. 20th , 2021 - Skoltech

CASE STUDY PRESENTATION

In one of the experiments, the experimenter is interested in making comparisons among 7 treatments and there are 28 experimental units available. These 28 experimental units are arranged in a Youden Square design with 4 rows and 7 columns with one observation per cell.

PREPARATION OF THE WORKING INTERFACE IN R

```
### I. Set working directory  ####
# On RStudio: tab 'Session'-> Set Working Directory -> Choose Directory.
# Choose the directory containing the datafile and the associated R script.

### II. Possibly, installation of new R packages needed for the analysis on RStudio:
# Click on the 'Packages' tab in the bottom-right window of R Studio interface->'Install Packages'
# Comment #1: R package installation requires a connection to internet
# Comment #2: Once packages have been installed,
# no need to re-install them again when you close-open again RStudio.

### III. Initialisation of the working space
# To erase all graphs
graphics.off()
# To erase objects from the working space - Clean up of the memory
rm(list = ls())


# this is a trick to detect which folder contains the R script and the data
main_dir <- dirname(rstudioapi::getSourceEditorContext()$path)
setwd(main_dir)
```

LOADING REQUIRED METHODS FOR ANALYSIS

```
library(agricolae)
library(emmeans)
```

ANALYSIS OF THE CASE STUDY

```
## Load the data and examine
YoudenSq <- read.table('YoudenSquare1.csv', sep = ',', header = TRUE)
str(YoudenSq)
```

*l.gentzbittel@skoltech.ru

†c.ben@skoltech.ru

```
## 'data.frame':    28 obs. of  4 variables:
##  $ row: int  1 1 1 1 1 1 1 2 2 2 ...
##  $ col: int  1 2 3 4 5 6 7 1 2 3 ...
##  $ trt: int  2 3 4 5 6 7 1 7 1 2 ...
##  $ y  : num  4 5.3 1.1 16.9 16.9 10.3 294 17.5 220 12.2 ...
```

```
YoudenSq$row <- factor(YoudenSq$row)
YoudenSq$col <- factor(YoudenSq$col)
YoudenSq$trt <- factor(YoudenSq$trt)
str(YoudenSq)
```

```
## 'data.frame':    28 obs. of  4 variables:
##  $ row: Factor w/ 4 levels "1","2","3","4": 1 1 1 1 1 1 1 2 2 2 ...
##  $ col: Factor w/ 7 levels "1","2","3","4",..: 1 2 3 4 5 6 7 1 2 3 ...
##  $ trt: Factor w/ 7 levels "1","2","3","4",..: 2 3 4 5 6 7 1 7 1 2 ...
##  $ y  : num  4 5.3 1.1 16.9 16.9 10.3 294 17.5 220 12.2 ...
```

```
## What are the characteristics of this design ?
```

The parameters of the design are v (number of treatments) = 7, p (number of rows)= 4, q (number of columns) =7, r (replication of treatments) = 4.

```
YoudenSq
```

```
##    row col trt     y
## 1    1   1   2   4.0
## 2    1   2   3   5.3
## 3    1   3   4   1.1
## 4    1   4   5  16.9
## 5    1   5   6  16.9
## 6    1   6   7  10.3
## 7    1   7   1 294.0
## 8    2   1   7  17.5
## 9    2   2   1 220.0
## 10   2   3   2  12.2
## 11   2   4   3  15.5
## 12   2   5   4  11.0
## 13   2   6   5  26.5
## 14   2   7   6  27.2
## 15   3   1   6  37.0
## 16   3   2   7  26.0
## 17   3   3   1 310.0
## 18   3   4   2  22.7
## 19   3   5   3  24.2
## 20   3   6   4  21.4
## 21   3   7   5  31.3
## 22   4   1   5  46.8
## 23   4   2   6  44.2
## 24   4   3   7  34.3
## 25   4   4   1 282.0
## 26   4   5   2  33.7
## 27   4   6   3  33.7
## 28   4   7   4  30.5
```

```
#################
#################   generate the same Youden square using agricolae
#################

## function to create Youden Square
str(design.youden)
```

```
## function (trt, r, serie = 2, seed = 0, kinds = "Super-Duper", first = TRUE,
```

```
##     randomization = TRUE)
```

```r
# function (trt, r, serie = 2, seed = 0, kinds = "Super-Duper", first = TRUE,
# randomization = TRUE)

trt <- c("T1", "T2", "T3", "T4", 'T5', 'T6', 'T7' )  # treatments
r  <- 4   # size of blocks

## generate the exp. design. Remember the properties of a BIBD
outdesign <- design.youden(trt, r, seed = 3264, serie = 2)


print(outdesign$parameters)
```

```
## $design
## [1] "youden"
##
## $trt
## [1] "T1" "T2" "T3" "T4" "T5" "T6" "T7"
##
## $r
## [1] 4
##
## $serie
## [1] 2
##
## $seed
## [1] 3264
##
## $kinds
## [1] "Super-Duper"
```

```r
book <- outdesign$book
plots <- as.numeric( book[,1] )

print(outdesign$sketch)
```

```
##      [,1] [,2] [,3] [,4]
## [1,] "T2" "T1" "T7" "T5"
## [2,] "T7" "T4" "T6" "T3"
## [3,] "T3" "T6" "T1" "T2"
## [4,] "T4" "T3" "T2" "T6"
## [5,] "T1" "T5" "T4" "T7"
## [6,] "T5" "T7" "T3" "T4"
## [7,] "T6" "T2" "T5" "T1"
```

```r
# write in hard disk
# write.csv(book,"book.csv", row.names=FALSE)
# file.show("book.csv")
#################

#################
#################   Analysis of the YS by hand'
#################

## Because not all treatmnents are tested in all rows, the SS of rows, cols and
## treatment are NOT independant.

## ANOVA for treatment adjusted for row and col effect,
Ana1a <- aov(y ~ row + col + trt, data = YoudenSq)
summary(Ana1a)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)    
## row          3   3302    1101   4.862   0.0194 *  
## col          6  27640    4607  20.347 1.23e-05 ***
## trt          6 194923   32487 143.489 1.85e-10 ***
## Residuals   12   2717     226                     
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
## marginal means ie adjusted means
Fitted.Model <- emmeans(Ana1a, ~ trt)
## LSD for adjusted means
pairs(Fitted.Model, adjust = 'Tukey')
```

```
##  contrast estimate   SE df t.ratio p.value
##  1 - 2     260.443 11.4 12  22.897  <.0001
##  1 - 3     253.307 11.4 12  22.270  <.0001
##  1 - 4     265.221 11.4 12  23.317  <.0001
##  1 - 5     248.721 11.4 12  21.867  <.0001
##  1 - 6     244.229 11.4 12  21.472  <.0001
##  1 - 7     253.479 11.4 12  22.285  <.0001
##  2 - 3      -7.136 11.4 12  -0.627  0.9944
##  2 - 4       4.779 11.4 12   0.420  0.9994
##  2 - 5     -11.721 11.4 12  -1.031  0.9367
##  2 - 6     -16.214 11.4 12  -1.426  0.7794
##  2 - 7      -6.964 11.4 12  -0.612  0.9951
##  3 - 4      11.914 11.4 12   1.047  0.9321
##  3 - 5      -4.586 11.4 12  -0.403  0.9995
##  3 - 6      -9.079 11.4 12  -0.798  0.9807
##  3 - 7       0.171 11.4 12   0.015  1.0000
##  4 - 5     -16.500 11.4 12  -1.451  0.7665
##  4 - 6     -20.993 11.4 12  -1.846  0.5454
##  4 - 7     -11.743 11.4 12  -1.032  0.9362
##  5 - 6      -4.493 11.4 12  -0.395  0.9996
##  5 - 7       4.757 11.4 12   0.418  0.9994
##  6 - 7       9.250 11.4 12   0.813  0.9788
## 
## Results are averaged over the levels of: row, col 
## P value adjustment: tukey method for comparing a family of 7 estimates
```

```r
## be careful : the below is based on **observed** means so may be incorrect
## in some cases (not here because differences are clear)
outHSD <- HSD.test(Ana1a, "trt",console = TRUE)
```

```
## 
## Study: Ana1a ~ "trt"
## 
## HSD Test for y 
## 
## Mean Square Error:  226.4085 
## 
## trt,  means
## 
##         y      std r   Min   Max
## 1 276.500 39.37427 4 220.0 310.0
## 2  18.150 12.88578 4   4.0  33.7
## 3  19.675 12.12776 4   5.3  33.7
## 4  16.000 12.73342 4   1.1  30.5
## 5  30.375 12.47968 4  16.9  46.8
## 6  31.325 11.87529 4  16.9  44.2
## 7  22.025 10.39916 4  10.3  34.3
## 
```

```
## Alpha: 0.05 ; DF Error: 12
## Critical Value of Studentized Range: 4.949594
##
## Minimun Significant Difference: 37.23796
##
## Treatments with the same letter are not significantly different.
##
##         y groups
## 1 276.500      a
## 6  31.325      b
## 5  30.375      b
## 7  22.025      b
## 3  19.675      b
## 2  18.150      b
## 4  16.000      b
```

```
## imagine we just give up with rows
Ana2 <- aov(y ~ col + trt, data = YoudenSq)
summary(Ana2)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## col          6  27640    4607   11.48 7.08e-05 ***
## trt          6 194923   32487   80.96 1.43e-10 ***
## Residuals   15   6019     401
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Comparison of the two MSE show we greatly improve the model by accounting for rows