Case Study : Balanced Incomplete Block Design - example 1

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

The objective is to analyse a small Balanced Incomplete Block Design, and to see how to generate such a design to implement it in the field.

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
BIBD1 <- read.table('BIBD1.txt', sep = ' ', header = TRUE)
str(BIBD1)
```


*l.gentzbittel@skoltech.ru

†c.ben@skoltech.ru

```
## 'data.frame':    12 obs. of  3 variables:
##  $ blk  : int  1 1 1 2 2 2 3 3 3 4 ...
##  $ trt  : chr  "A" "C" "D" "A" ...
##  $ yield: int  69 62 40 77 85 60 72 63 55 70 ...
```

```
BIBD1$blk <- factor(BIBD1$blk)
str(BIBD1)
```

```
## 'data.frame':    12 obs. of  3 variables:
##  $ blk  : Factor w/ 4 levels "1","2","3","4": 1 1 1 2 2 2 3 3 3 4 ...
##  $ trt  : chr  "A" "C" "D" "A" ...
##  $ yield: int  69 62 40 77 85 60 72 63 55 70 ...
```

```
## Draw at the white board. What are the characteristics of this design ?


#################
#################   generate the same BIBD using agricolae
#################

## function to create BIBD
str(design.bib)
```

```
## function (trt, k, r = NULL, serie = 2, seed = 0, kinds = "Super-Duper",
##     maxRep = 20, randomization = TRUE)
```

```
# function (trt, k, r = NULL, serie = 2, seed = 0, kinds = "Super-Duper",
#           maxRep = 20, randomization = TRUE)

trt <- c("A", "B", "C", "D" )  # treatments
k <- 3   # size of blocks

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

```
##
## Parameters BIB
## ==============
## Lambda     : 2
## treatmeans : 4
## Block size : 3
## Blocks     : 4
## Replication: 3
##
## Efficiency factor 0.8888889
##
## <<< Book >>>
```

```
print(outdesign$parameters)
```

```
## $design
## [1] "bib"
##
## $trt
## [1] "A" "B" "C" "D"
##
## $k
## [1] 3
##
## $serie
## [1] 2
##
```

```
## $seed
## [1] 3264
##
## $kinds
## [1] "Super-Duper"
```

```r
book <- outdesign$book
plots <- as.numeric( book[,1] )
matrix(plots, byrow = TRUE, ncol = k)
```

```
##      [,1] [,2] [,3]
## [1,]  101  102  103
## [2,]  201  202  203
## [3,]  301  302  303
## [4,]  401  402  403
```

```r
print(outdesign$sketch)
```

```
##      [,1] [,2] [,3]
## [1,] "C"  "A"  "D"
## [2,] "D"  "C"  "B"
## [3,] "D"  "A"  "B"
## [4,] "C"  "B"  "A"
```

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


#################
#################   Analysis of the BIBD using agricolae
#################

Analysis <- BIB.test(block = BIBD1$blk,
         trt = BIBD1$trt,
         y = BIBD1$yield,
         test = c("tukey"),
         alpha = 0.05, group = TRUE,  # compute groups of means
         console = TRUE)
```

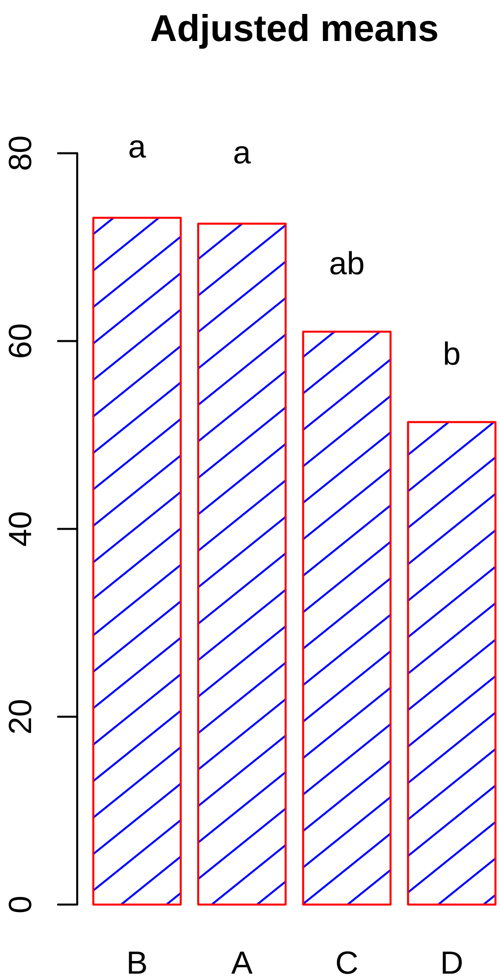
```
##
## ANALYSIS BIB:  BIBD1$yield
## Class level information
##
## Block:  1 2 3 4
## Trt  :  A C D B
##
## Number of observations:  12
##
## Analysis of Variance Table
##
## Response: BIBD1$yield
##             Df Sum Sq Mean Sq F value  Pr(>F)
## block.unadj  3 445.67  148.56  3.4995 0.10554
## trt.adj      3 861.08  287.03  6.7615 0.03282 *
## Residuals    5 212.25   42.45
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## coefficient of variation: 10.1 %
```

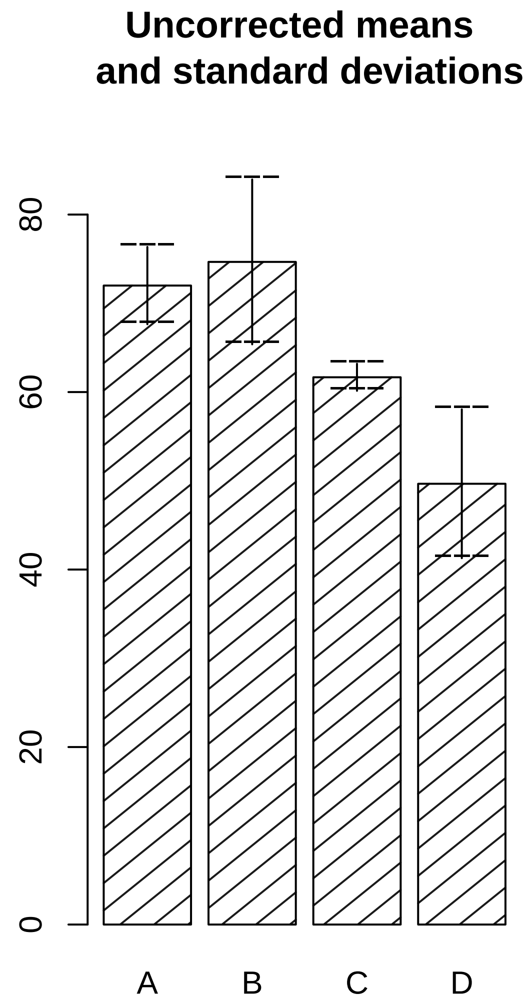
```
## BIBD1$yield Means: 64.5
##
## BIBD1$trt,  statistics
##
##   BIBD1$yield mean.adj       SE r      std Min Max
## A    72.00000   72.500 3.934026 3 4.358899  69  77
## B    74.66667   73.125 3.934026 3 9.291573  67  85
## C    61.66667   61.000 3.934026 3 1.527525  60  63
## D    49.66667   51.375 3.934026 3 8.386497  40  55
##
## Tukey
## Alpha      : 0.05
## Std.err    : 3.989831
## HSD        : 20.82023
## Parameters BIB
## Lambda     : 2
## treatmeans : 4
## Block size : 3
## Blocks     : 4
## Replication: 3
##
## Efficiency factor 0.8888889
##
## <<< Book >>>
##
## Comparison between treatments means
##       Difference pvalue sig.
## A - B     -0.625 0.9994
## A - C     11.500 0.2884
## A - D     21.125 0.0474    *
## B - C     12.125 0.2559
## B - D     21.750 0.0426    *
## C - D      9.625 0.4088
##
## Treatments with the same letter are not significantly different.
##
##   BIBD1$yield groups
## B      73.125      a
## A      72.500      a
## C      61.000     ab
## D      51.375      b
```

```
## use of agricolae::bar.group() function
x11()
par(mfrow=c(1,2), cex = 1)  ## two graphics within the same figure

bar.group(Analysis$groups,
          col = "blue", border = "red", density = 6,
         ylim = c(0, max(BIBD1$yield)),
         main = 'Adjusted means'
          )

bar.err(Analysis$means, variation="SD",
        col = 'grey10', density=8,
        ylim=c(0,90),
        main="Uncorrected means \n and standard deviations" )
```






```
#################
#################   Analysis of the BIBD by hand'
#################

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

## ANOVA for treatment adjusted for block effect, block unadjusted
Ana1a <- aov(yield ~ blk + trt, data = BIBD1)
summary(Ana1a)
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## blk          3  445.7  148.56   3.500 0.1055
## trt          3  861.1  287.03   6.762 0.0328 *
## Residuals    5  212.3   42.45
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## marginal means ie adjusted means
emmeans(Ana1a, ~ trt)
```

```
##  trt emmean   SE df lower.CL upper.CL
##  A     72.5 3.93  5     62.4     82.6
##  B     73.1 3.93  5     63.0     83.2
##  C     61.0 3.93  5     50.9     71.1
##  D     51.4 3.93  5     41.3     61.5
##
## Results are averaged over the levels of: blk
```

```
## Confidence level used: 0.95
```

```r
## ANOVA for blocks adjusted for treatment effect, treatment unadjusted
## ** NOT really useful **
Ana1b <- aov(yield ~ trt + blk, data = BIBD1)
summary(Ana1b)
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## trt          3 1163.0   387.7   9.132  0.018 *
## blk          3  143.7    47.9   1.129  0.421
## Residuals    5  212.3    42.5
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
## marginal means ie adjusted means
emmeans(Ana1b, ~ trt)
```

```
##  trt emmean   SE df lower.CL upper.CL
##  A     72.5 3.93  5     62.4     82.6
##  B     73.1 3.93  5     63.0     83.2
##  C     61.0 3.93  5     50.9     71.1
##  D     51.4 3.93  5     41.3     61.5
##
## Results are averaged over the levels of: blk
## Confidence level used: 0.95
```