Working Paper - Design and Analysis of Multivariate Experiments

Diego Vallarino

15/12/2021

Contents

---

1 Introduction

Certain experimental conditions give rise to designs that contain crossed and nested factors. In this exercise we can identify a design with three factors, a fixed *Spectrometer* factor with *3* levels crossed with a random *Plots* factor with *3* levels, and a third *Samples* factor with *2* levels as well. random that is crossed with *Spectrometer* but nested with *Plots*.

Therefore, the analysis of the experiment is as follows:

1.1 Descriptive data analysis

```
##     parcela     muestra    espectrometro     medida
##  Min.   :1   Min.   :1.0   Min.   :1     Min.   :3.700
##  1st Qu.:1   1st Qu.:1.0   1st Qu.:1     1st Qu.:3.778
##  Median :2   Median :1.5   Median :2     Median :3.860
##  Mean   :2   Mean   :1.5   Mean   :2     Mean   :3.841
##  3rd Qu.:3   3rd Qu.:2.0   3rd Qu.:3     3rd Qu.:3.881
##  Max.   :3   Max.   :2.0   Max.   :3     Max.   :3.988

## 'data.frame':    36 obs. of  4 variables:
##  $ parcela      : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ muestra      : int  1 1 2 2 1 1 2 2 1 1 ...
##  $ espectrometro: int  1 1 1 1 2 2 2 2 3 3 ...
##  $ medida       : num  3.83 3.87 3.82 3.85 3.93 ...
```

```
##                vars  n mean   sd median trimmed  mad min  max range skew
## parcela*          1 36 2.00 0.83   2.00    2.00 1.48 1.0 3.00  2.00 0.00
## muestra*          2 36 1.50 0.51   1.50    1.50 0.74 1.0 2.00  1.00 0.00
## espectrometro*    3 36 2.00 0.83   2.00    2.00 1.48 1.0 3.00  2.00 0.00
## medida            4 36 3.84 0.08   3.86    3.84 0.12 3.7 3.99  0.29 0.15
##                kurtosis   se
## parcela*          -1.58 0.14
## muestra*          -2.05 0.08
## espectrometro*    -1.58 0.14
## medida            -0.98 0.01
```

1.2 Linear Model for the Experiment

The model is as follows:

modelo<-lm(medida ~ espectrometro + r(parcela) + espectrometro:r(parcela) + r(muestra)%in%r(parcela) + espectrometro:(r(muestra)%in%r(parcela)), datos)

The calculations are presented below:

```
##
##  Shapiro-Wilk normality test
##
## data:  modelo$residuals
## W = 0.88116, p-value = 0.001084
```

```
## Analysis of variance (unrestricted model)
## Response: medida
##                               Mean Sq Sum Sq Df F value Pr(>F)
## espectrometro                  0.0626  0.125  2    9.06 0.0327
## parcela                        0.0175  0.035  2    2.14 0.2397
## espectrometro:parcela          0.0069  0.028  4    2.71 0.1326
## parcela:muestra                0.0038  0.011  3    1.50 0.3076
## espectrometro:parcela:muestra  0.0025  0.015  6    7.38 0.0004
## Residuals                      0.0003  0.006 18       -      -
##
##                                     Err.term(s) Err.df   VC(SS)
## 1 espectrometro                             (3)      4    fixed
## 2 parcela                     (3) + (4) - (5)   3.74 0.000779
## 3 espectrometro:parcela                     (5)      6 0.001089
## 4 parcela:muestra                           (5)      6 0.000211
## 5 espectrometro:parcela:muestra             (6)     18 0.001101
## 6 Residuals                                   -      - 0.000345
## (VC = variance component)
##
##                                          Expected mean squares
## espectrometro                (6) + 2 (5) + 4 (3) + 12 Q[1]
## parcela                      (6) + 2 (5) + 6 (4) + 4 (3) + 12 (2)
## espectrometro:parcela        (6) + 2 (5) + 4 (3)
## parcela:muestra              (6) + 2 (5) + 6 (4)
## espectrometro:parcela:muestra (6) + 2 (5)
## Residuals                    (6)
```

1.3 Model Hypothesis

For the Spectrometer Fixed factor, A:

$H_0 : \alpha_1 = \alpha_2 = \alpha_3 = 0$
$H_1 : \alpha_i \neq 0$

For Random Factor Plots, B

$H_0 : \sigma_B^2 \neq 0$
$H_1 : \sigma_B^2 > 0$

For the interaction Spectrometer and Plot, AB

$H_0 : \sigma_{AB}^2 \neq 0$
$H_1 : \sigma_{AB}^2 > 0$

For the random factor Plots nested to Samples, C(B):

$H_0 : \sigma_C^2 \neq 0$
$H_1 : \sigma_C^2 > 0$

For the Spectrometer $\times$ Samples interaction nested to Plots, AC(B)

$H_0 : \sigma_{AC}^2 \neq 0$
$H_1 : \sigma_{AC}^2 > 0$

1.4 Significant factors

According to the evidence presented with the ANOVA analysis, only the fixed factor **Spectometer** is significant, together with its *interaction* with the Plot.

1.5 Variance components

a) Error variance: 0.0003
b) Variance of the parcel factor: 0.000779
c) Variance of the spectrometer:plot interaction: 0.001089
d) Variance of the plot:sample interaction: 0.000211
e) Variance of the spectrometer:plot:sample interaction: 0.001101

The estimate of the variance of the response variable is = 0.0003 + 0.000779 + 0.001089 + 0.000211 + 0.001101 = **0.00348** and therefore, the percentage of the variance components are:

a) Error variance: 9%
b) Variance of the parcel factor: 22%
c) Variance of the spectrometer:plot interaction:31%
d) Variance of plot:sample interaction: 6%
e) Variance of the spectrometer:plot:sample interaction: 32%

We conclude that the *spectrometer:plot:sample interaction factor* is the main cause of variability in the response (32%), followed by the *spectrometer:plot factor* (31%).

1.6 Perform multiple comparisons

```r
HSD.test(modelo, "fert", console = TRUE)
```

```
## Name:  fert
##  espectrometro parcela muestra
```

```r
HSD.test(modelo,"espectrometro", group=TRUE,console=TRUE)
```

```
##
## Study: modelo ~ "espectrometro"
##
## HSD Test for medida
##
## Mean Square Error:  0.0003449722
##
## espectrometro,  means
##
##     medida        std  r   Min   Max
## 1 3.802667 0.06176986 12 3.720 3.887
## 2 3.923917 0.04659976 12 3.872 3.988
## 3 3.795333 0.05199184 12 3.700 3.872
##
## Alpha: 0.05 ; DF Error: 18
## Critical Value of Studentized Range: 3.609304
##
## Minimun Significant Difference: 0.01935196
##
## Treatments with the same letter are not significantly different.
##
##     medida groups
## 2 3.923917      a
## 1 3.802667      b
## 3 3.795333      b
```

We can identify that **spectrometers 1 and 3** have similar performances, while spectrometer 2 has a different performance than the other two. Therefore, we can identify 2 homogeneous groups of spectrometers.

1.7 Confidence Intervals with 95%

```r
HSD.test(modelo,"espectrometro", group=FALSE,console=TRUE)
```

```
##
## Study: modelo ~ "espectrometro"
##
## HSD Test for medida
##
## Mean Square Error:  0.0003449722
##
## espectrometro,  means
##
```

```
##     medida        std  r   Min   Max
## 1 3.802667 0.06176986 12 3.720 3.887
## 2 3.923917 0.04659976 12 3.872 3.988
## 3 3.795333 0.05199184 12 3.700 3.872
##
## Alpha: 0.05 ; DF Error: 18
## Critical Value of Studentized Range: 3.609304
##
## Comparison between treatments means
##
##         difference pvalue signif.         LCL         UCL
## 1 - 2 -0.121250000 0.0000     *** -0.14060196 -0.10189804
## 1 - 3  0.007333333 0.6063         -0.01201862  0.02668529
## 2 - 3  0.128583333 0.0000     ***  0.10923138  0.14793529
```