

# Clase5.R

Usuario

2019-08-09

```
#Alumno:Emmanuel T. Ferrer
#Matricula:2031741
#Fecha: 09/08/2019

#Clase No.5

#Analisis de varianza
#comparar efecto 3 concentracion de fertilizantes
#comparar media de los tratamientos
#normalidad de los datos (shapiro)
#que las varianzas sean iguales de los diferentes tratamiento
#ho= que no existe diferencia significativa entre los tratamientos
#h1: al menos una media existe diferencia entre los tratamientos
#Xij= media observacion de los tres tratamientos mas alfa, -error
#i es tratamiento y j numero de observacion
#m es media de la observacion
#error no explicada
#gl generar niveles (3 tratamientos)

# Ingresar datos para analisis de varianza -----
arena <- c(6,10,8,6,14,17,9,11,7,11)
arena

## [1] 6 10 8 6 14 17 9 11 7 11
arcilla <- c(17,15,3,11,14,12,12,8,10,13)
arcilla

## [1] 17 15 3 11 14 12 12 8 10 13
limo <- c(13,16,9,12,15,16,17,13,18,14)
limo

## [1] 13 16 9 12 15 16 17 13 18 14
y.ton <-c(arena, arcilla, limo)
y.ton

## [1] 6 10 8 6 14 17 9 11 7 11 17 15 3 11 14 12 12 8 10 13 13 16 9
## [24] 12 15 16 17 13 18 14
suelo <-gl(3,10,30,labels = c("arena", "arcilla", "limo"))
suelo

## [1] arena arena arena arena arena arena arena arena
## [9] arena arena arcilla arcilla arcilla arcilla arcilla arcilla
## [17] arcilla arcilla arcilla arcilla limo limo limo limo
## [25] limo limo limo limo limo limo
## Levels: arena arcilla limo
```

```
prod <-data.frame(suelo, y.ton)
prod
```

```
##      suelo y.ton
## 1  arena     6
## 2  arena    10
## 3  arena     8
## 4  arena     6
## 5  arena    14
## 6  arena    17
## 7  arena     9
## 8  arena    11
## 9  arena     7
## 10 arena    11
## 11 arcilla  17
## 12 arcilla  15
## 13 arcilla   3
## 14 arcilla  11
## 15 arcilla  14
## 16 arcilla  12
## 17 arcilla  12
## 18 arcilla   8
## 19 arcilla  10
## 20 arcilla  13
## 21  limo    13
## 22  limo    16
## 23  limo     9
## 24  limo    12
## 25  limo    15
## 26  limo    16
## 27  limo    17
## 28  limo    13
## 29  limo    18
## 30  limo    14
```

```
head(prod)
```

```
##      suelo y.ton
## 1 arena     6
## 2 arena    10
## 3 arena     8
## 4 arena     6
## 5 arena    14
## 6 arena    17
```

```
#media de la productividad para cada uno de los tratamientos
tapply(prod$y.ton, prod$suelo, mean)
```

```
##      arena arcilla  limo
##      9.9   11.5   14.3
```

```
#varianza de cada uno de los grupos
tapply(prod$y.ton, prod$suelo, var)
```

```
##      arena  arcilla    limo
## 12.544444 15.388889  7.122222
```

```

#normalidad de los datos
shapiro.test(prod$y.ton)

##
##  Shapiro-Wilk normality test
##
## data:  prod$y.ton
## W = 0.97214, p-value = 0.5993
#son normales porque es 0.5993

#varianza prueba comparar
bartlett.test(prod$y.ton, prod$suelo)

##
##  Bartlett test of homogeneity of variances
##
## data:  prod$y.ton and prod$suelo
## Bartlett's K-squared = 1.2764, df = 2, p-value = 0.5283
#No existe diferencias entre las varianzas
#hay homogeneidad

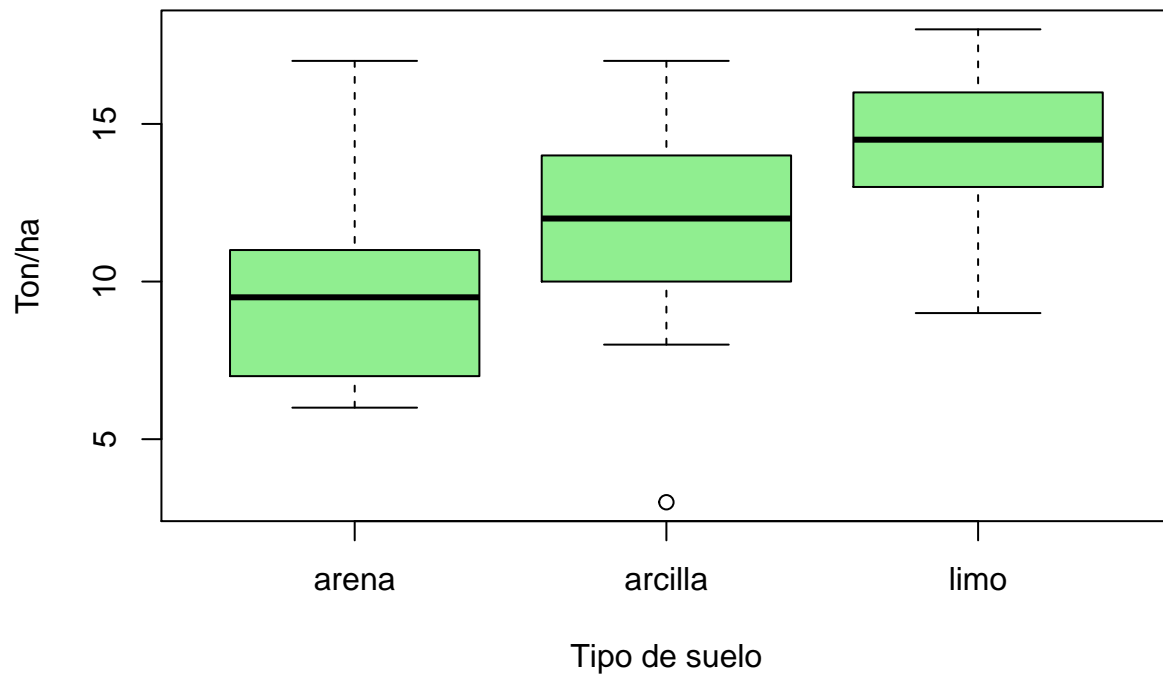
#determinar homogeneidad de las varianzas
fligner.test(prod$y.ton, prod$suelo)

##
##  Fligner-Killeen test of homogeneity of variances
##
## data:  prod$y.ton and prod$suelo
## Fligner-Killeen:med chi-squared = 0.36507, df = 2, p-value =
## 0.8332
#no existe diferencia
#hay homogeneidad

#no hay tanta variacion entre bartlett vs fligner
#puede ser cualquiera aplica

#inspeccion de los datos
boxplot(prod$y.ton ~ prod$suelo, xlab="Tipo de suelo",
        ylab="Ton/ha", col="light green")

```



*#no hay diferencias significativas, no  
#es tan visibles las diferencias*

*#fuente de variacion es el suelo  
#error*

*#grado de libertad 2*

*#error= numero de observacion menos tres trataimientos es igua a 27*

*#total es 29*

```
aov.suelo <- aov(prod$y.ton ~ prod$suelo)
aov.suelo
```

## Call:

```
## aov(formula = prod$y.ton ~ prod$suelo)
```

##

## Terms:

```
##          prod$suelo Residuals
```

```
## Sum of Squares      99.2    315.5
```

```
## Deg. of Freedom        2        27
```

##

```
## Residual standard error: 3.41836
```

```
## Estimated effects may be unbalanced
```

```
summary(aov.suelo)
```

```
##          Df Sum Sq Mean Sq F value Pr(>F)
## prod$suelo  2   99.2   49.60   4.245  0.025 *
```

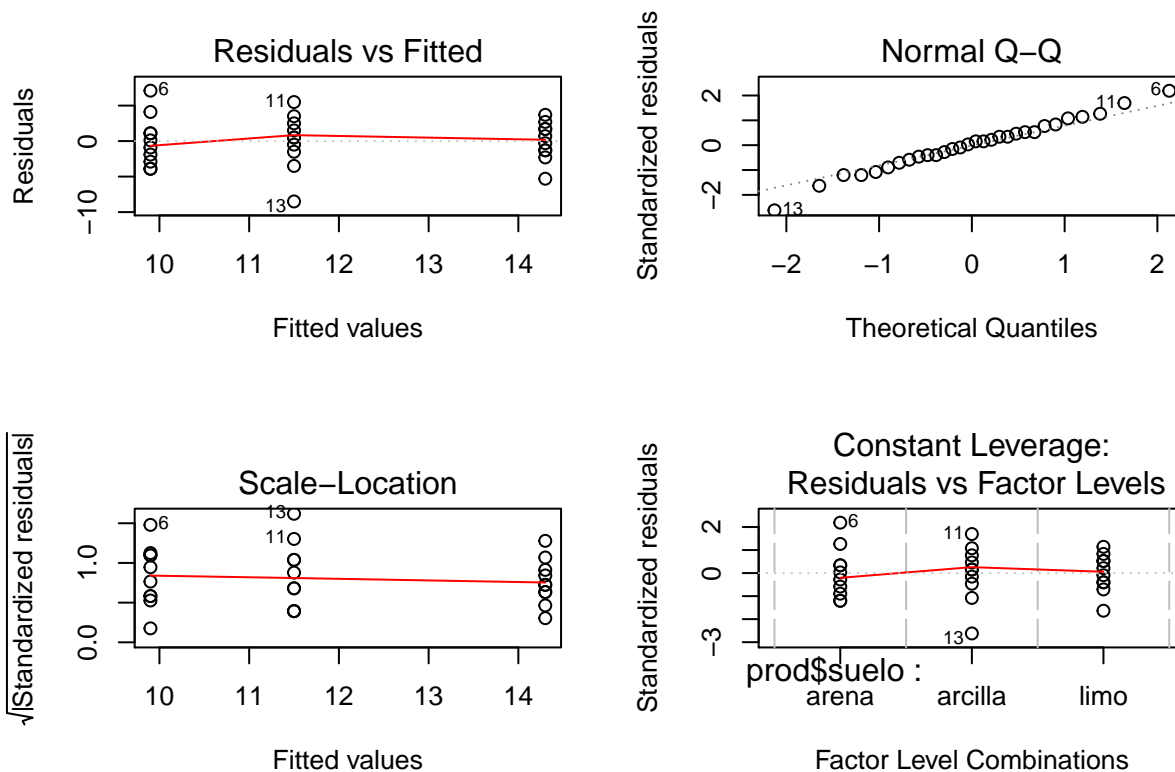
```
## Residuals    27   315.5    11.69
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

#nos interesa descomponer las varianzas
#cual de los tres bloques es mas variables
#arena es mas variable
#hay que ver si la variacion es entre o fuera
#si hay diferencia entre los tratamientos porque
#dio 0.025 pr(>F)
#son ligeramente significativo
#que procede
#visualizacion

par(mfrow=c(2,2))
par

## function (... , no.readonly = FALSE)
## {
##   .Pars.readonly <- c("cin", "cra", "csi", "cxy", "din", "page")
##   single <- FALSE
##   args <- list(...)
##   if (!length(args))
##     args <- as.list(if (no.readonly)
##       .Pars[-match(.Pars.readonly, .Pars)]
##     else .Pars)
##   else {
##     if (all(unlist(lapply(args, is.character))))
##       args <- as.list(unlist(args))
##     if (length(args) == 1) {
##       if (is.list(args[[1L]]) | is.null(args[[1L]]))
##         args <- args[[1L]]
##       else if (is.null(names(args)))
##         single <- TRUE
##     }
##   }
##   value <- .External2(C_par, args)
##   if (single)
##     value <- value[[1L]]
##   if (!is.null(names(args)))
##     invisible(value)
##   else value
## }
## <bytecode: 0x000000001544b8d0>
## <environment: namespace:graphics>

plot(aov(prod$y.ton ~ prod$suelo))
```



```
par(mfrow=c(1,1))
#no hay ningun indicio que existe diferencia
#entre los tratamientos es muy variables
#que son normales

#como mostramos que son diferente a quien
#despues de encontrar diferencia significativo

#la prueba de tukey (para tres tratamiento o mas)
#para ver que tratamiento es diferente a los otros

TukeyHSD(aov.suelo, conf.level = 0.95)
```

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = prod$y.ton ~ prod$suelo)
##
## $`prod$suelo`
##      diff      lwr      upr    p adj
## arcilla-arena 1.6 -2.1903777 5.390378 0.5546301
## limo-arena    4.4  0.6096223 8.190378 0.0204414
## limo-arcilla  2.8 -0.9903777 6.590378 0.1785489

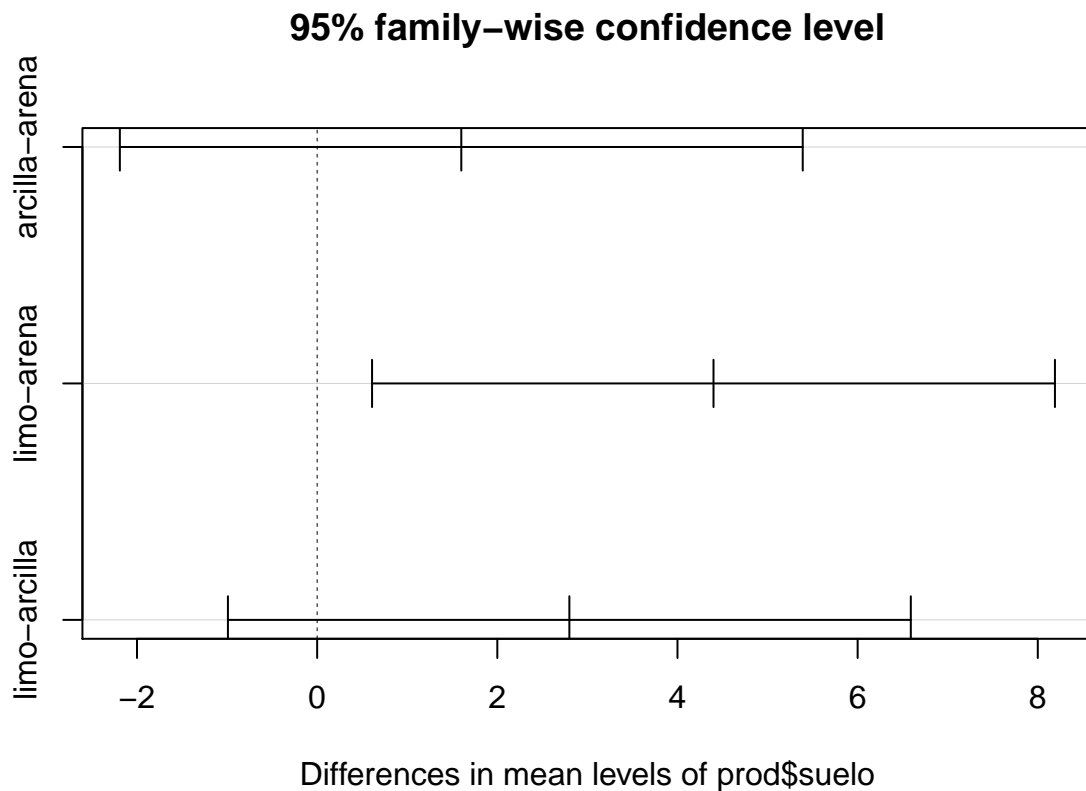
#arcilla-arena diferencia es 1.6 padj 0.5546301
#no hay diferencia significativa
```

```
#alimo-arena diff 4.4 padj 00.0204414
#si hay diferencia significativo

#limo-arcilla dfi 2.8 p adj. 0.1785489
#no hay diferencia significativa

#cual es el mejor tipo de suelo para producir
#limo-arcilla

#graficar
plot(TukeyHSD(aov.suelo))
```



```
summary(aov.suelo)
```

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## prod$suelo  2   99.2   49.60   4.245  0.025 *
## Residuals 27  315.5   11.69
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary.lm(aov.suelo)
```

```
##
## Call:
## aov(formula = prod$y.ton ~ prod$suelo)
##
## Residuals:
```

```
##      Min      1Q Median      3Q      Max
##    -8.5    -1.8    0.3    1.7    7.1
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      9.900      1.081   9.158 9.04e-10 ***
## prod$sueloarcilla  1.600      1.529   1.047  0.30456
## prod$suelolimo    4.400      1.529   2.878  0.00773 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.418 on 27 degrees of freedom
## Multiple R-squared:  0.2392, Adjusted R-squared:  0.1829
## F-statistic: 4.245 on 2 and 27 DF,  p-value: 0.02495
```

*#el analisis de varianza es significativa*

*#p-value: 0.02495*

*#Descripcion de hipotesis Nula y Alternativas*

*#Ho.la media de produccion en toneladas*

*#es la misma en los tratamientos.*

*#H1.la media de produccion en toneladas*

*#al menos uno es diferente a los demas tratamientos.*