

Examen - Planificación de la Investigación

Juan Cantero Jimenez

2/24/2022

Lectura de datos

```
load("cultiv.Rdata")
```

```
str(cultiv)
```

```
## 'data.frame': 15 obs. of 3 variables:
## $ matriz : Factor w/ 3 levels "FN","MG","SSTF": 2 2 2 2 1 1 1 1 1 ...
## $ oveja : int 1 2 3 4 5 1 2 3 4 5 ...
## $ recuento: num 3922 4084 3873 4225 3880 ...
```

```
cultiv$oveja <- as.factor(cultiv$oveja)
str(cultiv)
```

```
## 'data.frame': 15 obs. of 3 variables:
## $ matriz : Factor w/ 3 levels "FN","MG","SSTF": 2 2 2 2 1 1 1 1 1 ...
## $ oveja : Factor w/ 5 levels "1","2","3","4",...: 1 2 3 4 5 1 2 3 4 5 ...
## $ recuento: num 3922 4084 3873 4225 3880 ...
```

Ejercicio 1

La variable de interés sería el conteo de células/ml al cabo de 24 horas en función de las distintas matrices, además estas distintas matrices son factores fijos de los que queremos estudiar como afectan al conteo de células/ml al cabo de 24 horas. Por último como solo estamos interesados en conocer como afectan las distintas matrices al crecimiento celular pero sospechamos que la oveja de donde proceda la muestra puede tener un efecto, esta última se considerará como un bloque. Esto permitirá poder estudiar la base idónea para maximizar el crecimiento de CS sin que se vea afectado por la variabilidad aportada por las distintas muestras.

Ejercicio 2

En este caso tenemos un diseño experimental de un factor fijo y un bloque, es decir Randomized complete block, RCB.

El modelo estadístico sería el siguiente:

$$\text{recuento}_i = \mu + \beta_{M,MG} * I(M = MG) + \beta_{M,FN} * I(M = FN) + \beta_{M,SSTF} * I(M = SSTF) + \beta_{O,1} * I(O = 1) + \beta_{O,2} * I(O = 2) + \beta_{O,3} * I(O = 3) + \beta_{O,4} * I(O = 4) + \beta_{O,5} * I(O = 5)$$

Notese que la función indicadora $I(M = MG)$ valdrá 1 cuando la observación pertenezca al nivel MG y valdrá 0 si es distinta de este, lo mismo se aplica a las otras funciones indicadoras con los distintos niveles.

Las restricciones para este modelo son:

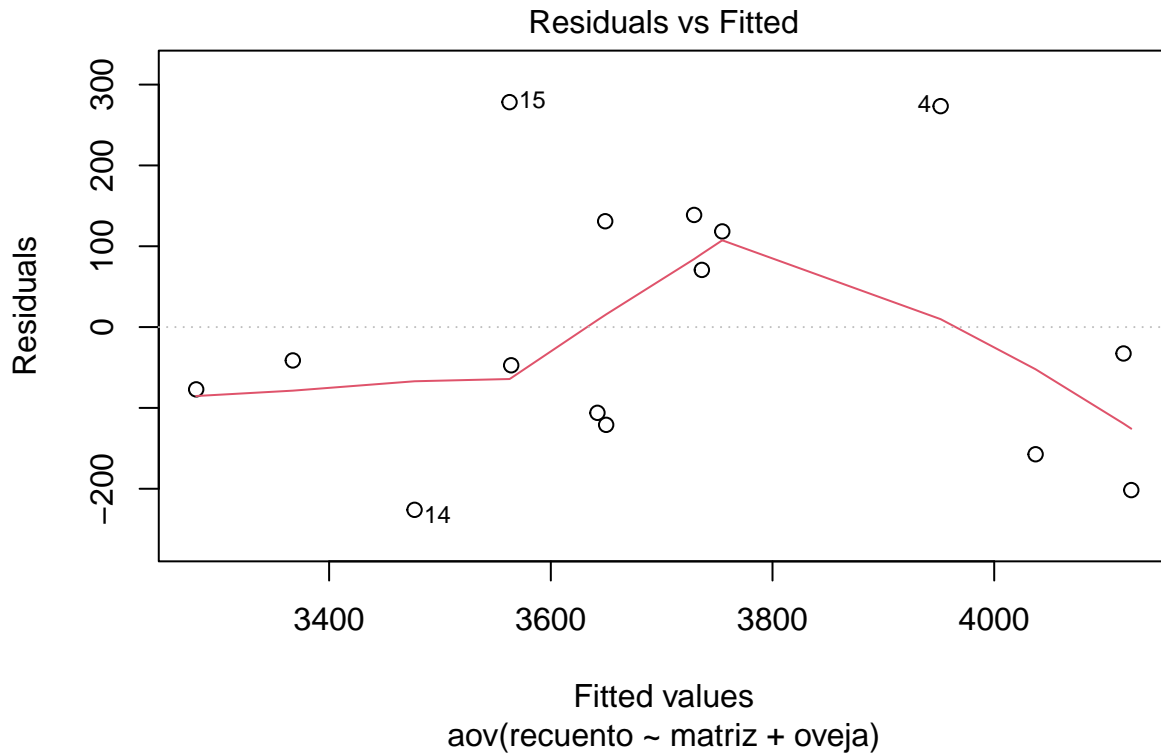
$$\sum \beta_m = \sum \beta_o = 0$$

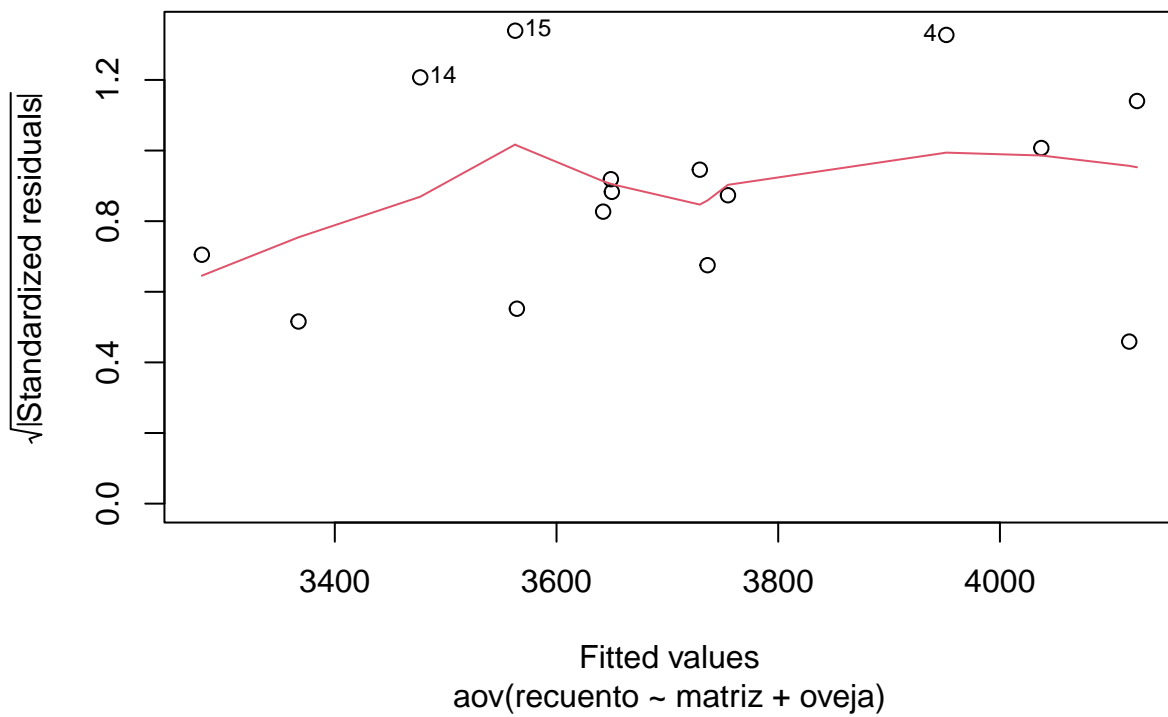
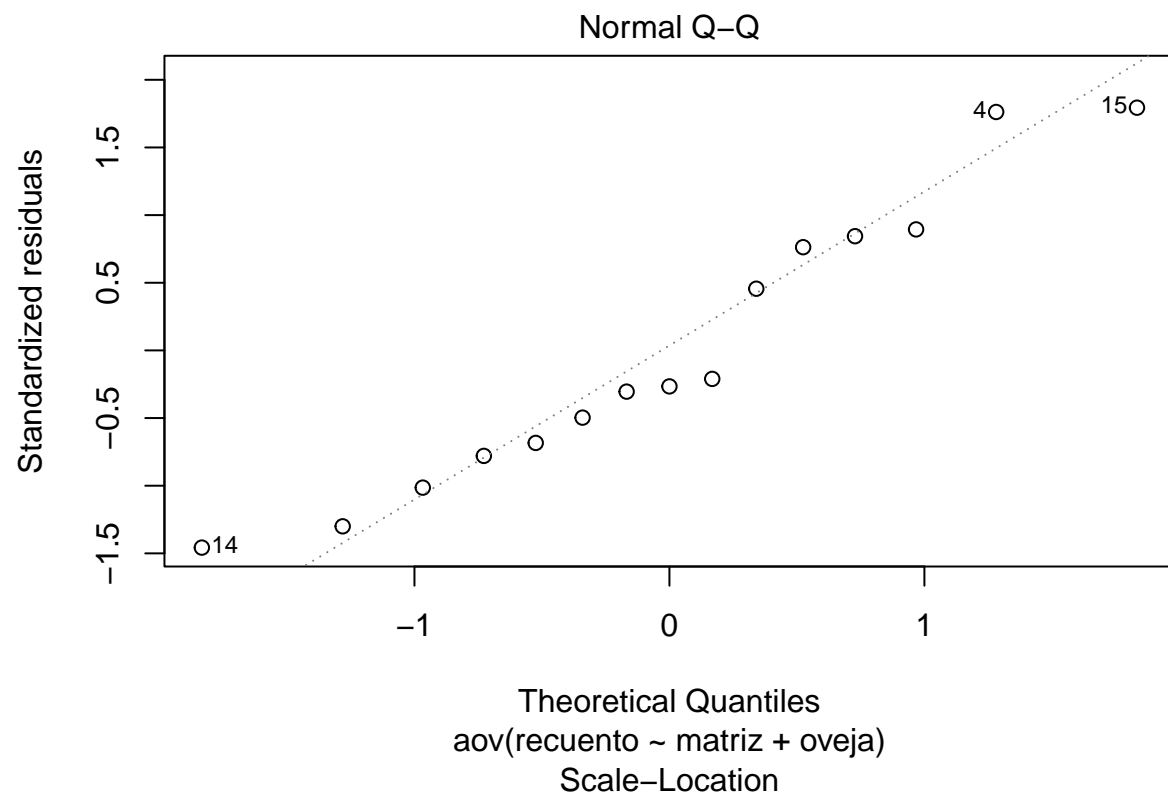
Donde μ representa la media global así como M y O, matriz y oveja respectivamente.

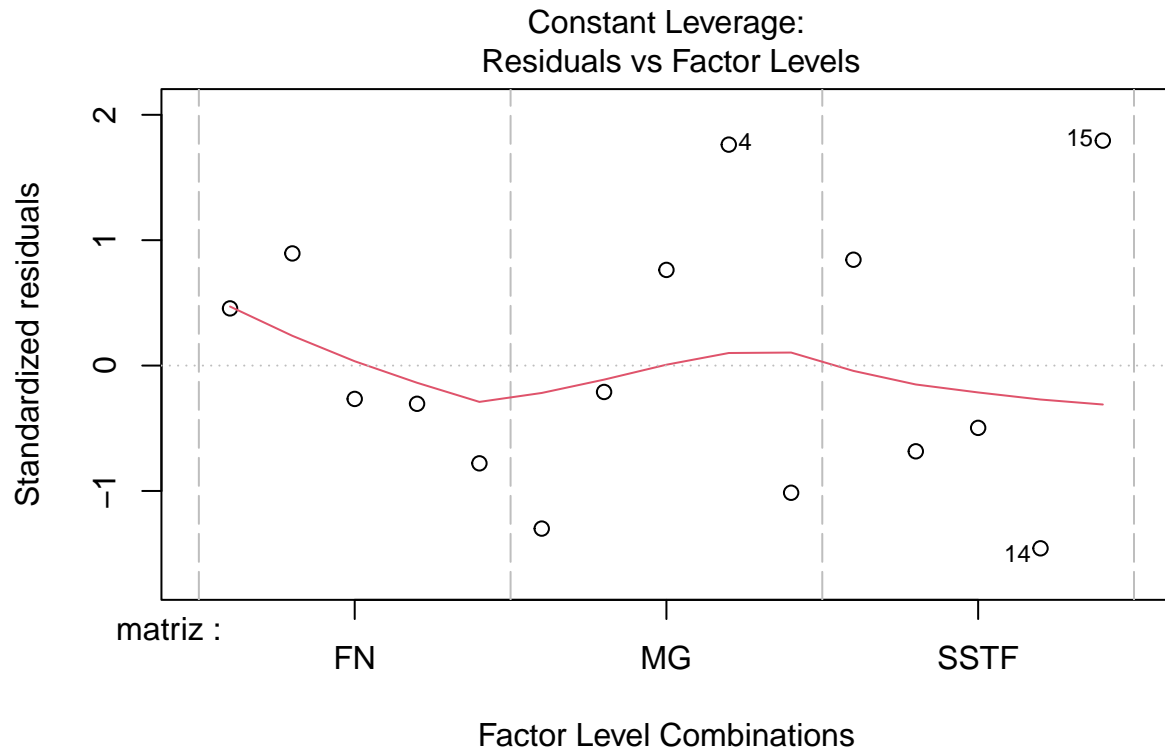
```
aov1 <- aov(recuento ~ matriz + oveja, data=cultiv)
summary(aov1)
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## matriz        2 638213   319106   7.074  0.017 *
## oveja         4 278315    69579   1.542  0.279
## Residuals     8 360888    45111
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
plot(aov1)
```







Para que el modelo sea aplicable los residuos deben de comportarse de forma normal así como la varianza entre los grupos debe de ser homogénea. Así si observamos el qqplot mostrado anteriormente podremos ver como los percentiles muestrales coinciden en mayor o menor medida con los percentiles muestrales por lo que podemos concluir en base a estos gráficos que el modelo sigue una distribución normal. Si atendemos ahora a la igualdad de varianza, en el plot de Residual vs Fitted así como Scale-Location se puede ver como la nube de puntos se distribuye de forma homogénea por lo que se puede concluir que la varianza entre los grupos es homogénea.

Ejercicio 3

```
model.tables(aov1, type="effects")
```

```
## Tables of effects
##
## matriz
## matriz
##      FN      MG      SSTF
## -100.07 287.33 -187.27
##
## oveja
## oveja
##      1      2      3      4      5
## 126.87 119.87 -242.13 -45.13  40.53
```

```
mean(cultiv$recuento)
```

```
## [1] 3709.467
```

Así los coeficientes para el modelo

$$\text{recuento}_i = \mu + \beta_{M,MG} * I(M = MG) + \beta_{M,FN} * I(M = FN) + \beta_{M,SSTF} * I(M = SSTF) + \beta_{O,1} * I(O = 1) + \beta_{O,2} * I(O = 2) + \beta_{O,3} * I(O = 3) + \beta_{O,4} * I(O = 4) + \beta_{O,5} * I(O = 5)$$

son $\mu = media\ global = 3709.467$, $\beta_{M, FN} = -100.07$, $\beta_{M, MG} = 287.33$, $\beta_{M, STF} = -187.27$, $\beta_{O, 1} = 126.87$,
 $\beta_{O, 2} = -119.87$, $\beta_{O, 3} = -242.13$, $\beta_{O, 4} = -45.13$, $\beta_{O, 5} = 40.53$