

## Estadística (Q) - Clase 18

Resolución ejercicio clase 18 (Práctica 7) Anova

---

1. Primero cargamos los datos en R

```
tiempo<-scan()  
51 64 70 63 78 55 67 75 82 61 53 60 62 83 77 90 85 60 70 58 40 61 66 37 54 77 75 57 85  
42 42 66 58 44 71 39 39 54 69 47 66 44 56 55 67 47 58 39 42 45 72 72 69 75 57 54 34 62  
80 85 78 64 80 80 57 75 76 78 83 74 78 84
```

Luego generamos un vector de la misma longitud que el vector tiempo que en cada coordenada nos indicará a qué grupo pertenece dicho dato

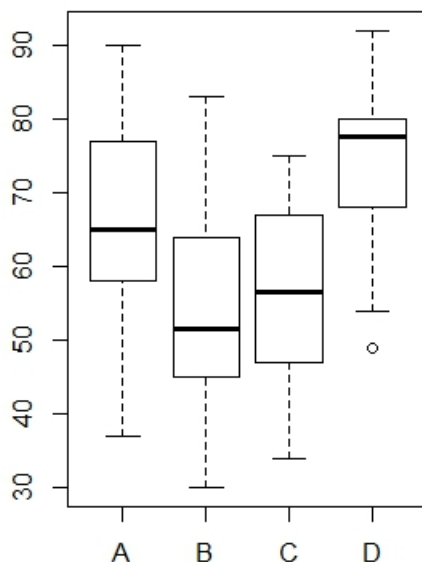
```
pais<-as.factor(c(rep(1,30),rep(2,30),rep(3,30),rep(4,30)))
```

Es muy importante que guardemos dicho vector como un factor. Esto permitirá realizar el test Anova de manera correcta.

Vamos a graficar juntos los boxplot de las 4 variedades de yeso.

```
boxplot(tiempo~pais,names=c("A","B","C","D"),  
main="boxplot de los tiempos de fraguado")
```

**boxplot de los tiempos de fraguado**



Aplicamos el test Anova a los datos del tiempo y lo guardamos con el nombre salida

```
salida<- aov(tiempo~pais)
```

Vemos un resumen del resultado del test

```
summary(salida)
```

```
> summary(salida)
```

```
          Df Sum Sq Mean Sq F value    Pr(>F)
pais         3   8773   2924.5    20.79 7.54e-11 ***
Residuals  116  16317    140.7
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Veamos cual es la información que guarda el test:

```
> attributes(salida)
```

```
$names
```

```
[1] "coefficients" "residuals"    "effects"      "rank"
[5] "fitted.values" "assign"        "qr"           "df.residual"
[9] "contrasts"    "xlevels"      "call"         "terms"
[13] "model"
```

```
$class
```

```
[1] "aov" "lm"
```

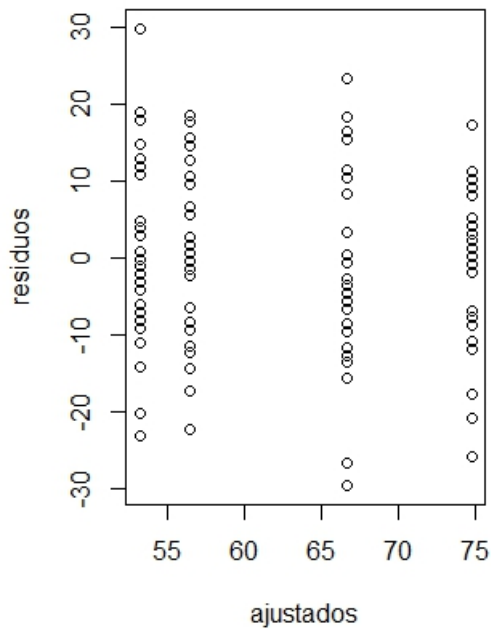
Notar toda la información que se consigue al realizar un test anova todo queda guardado con el nombre que lo guardé. Y escribiendo nombre+\$+variable vemos la variable deseada que guarda el programa Por ejemplo:

```
residuos<-salida$residuals
```

```
ajustados<-salida$fitted.values
```

Ver cuál es el vector de residuos y el de valores ajustados

```
plot(ajustados,residuos)
```



nos da una idea sobre la homogeneidad de varianzas De estar cargado el paquete "lawstat", puedo realizar el test de homogeneidad de varianzas (El test de Levene)

```
> levene.test(tiempo,pais)
```

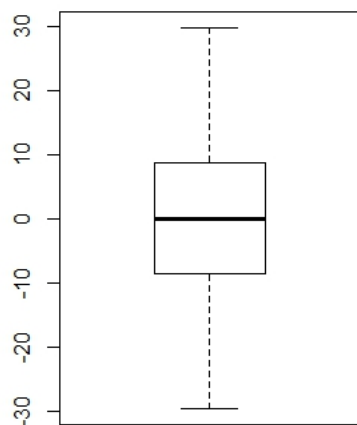
```
modified robust Brown-Forsythe Levene-type test based on the
absolute deviations from the median
```

```
data: tiempo
Test Statistic = 1.3867, p-value = 0.2504
```

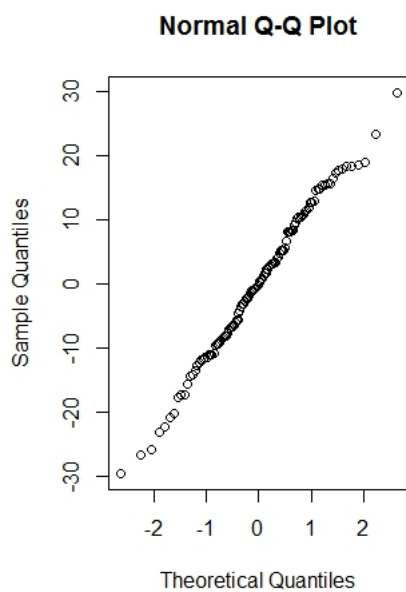
Si el p-valor es alto no rechazo homoscedasticidad.

Una vez que no rechacé homoscedasticidad, resulta razonable para chequear normalidad de los datos, directamente mirar si los residuos son normales.

```
boxplot(residuos)
```



```
qqnorm(residuos)
```



```
> shapiro.test(residuos)
```

Shapiro-Wilk normality test

data: residuos

W = 0.9921, p-value = 0.7312