

descripcionTablas

Trabajo con los datos de Vicente

Compuestos fenólicos (Procesado?)

Primero, se mide el efecto de agregar un endulzante determinado al zumo y se van haciendo medidas de metabolitos en diferentes tiempos y condiciones. Tenemos entonces dos factores que determinan el valor de los mg/100 ml de zumo, que serían las condiciones de toma de la muestra y el tipo de endulzante que se está añadiendo.

Medimos cuatro compuestos:

- Derivados del Naringenin hexosido
- Eriodictiol hexosa
- Naringenin rutinosido
- Hesperitin rutinosido.

En primer lugar se extrae el área del cromatógrafo y se transforma en mg/100 ml de zumo, con la fórmula siguiente:

$$((AREA - 221,08)/18129) * 610,19 * (0,1)$$

Se repite la medición cada 15 días, en condiciones de luz y oscuridad, a 5° y a 25° celsius de temperatura.

El procesado estadístico realizado en el excel adjuntado consiste en hacer la media y la desviación estándar de cada repetición hecha en las mismas condiciones. Después se calculan los mg totales de compuestos detectados por cada repetición, se hace la media y la SD, además de un porcentaje de pérdida respecto a la primera medida. Se replica en R (faltan los totales y la pérdida)

Procesado y ordenamiento:

```
# Funcion de procesado de tablas

procesado <- function(tabla, endulzante){

  tabla_num <- tabla[-1]

  # Se obtienen las medias de cada medida

  tabla_mean <- do.call(rbind,
                        lapply(seq(1, nrow(tabla_num), 2), function(i){
                          x <- tabla_num[ i:(i + 1), , drop = FALSE]
                          res <- rbind(colSums(x)/2)
                          res
                        })))

  rownames(tabla_mean) <- unique(tabla$Condiciones)

  # La desviación estándar
```

```

tabla_sd <- do.call(rbind,
  lapply(seq(1, nrow(tabla_num), 2), function(i){
    x <- tabla_num[ i:(i + 1), , drop = FALSE]
    res <- rbind(apply(x, 2, sd))
    res
  }))

# Apilamos datos

tabla_1 <- stack(as.data.frame(tabla_mean))
tabla_1$Condiciones <- rep(unique(tabla$Condiciones),4)

tabla_2 <- stack(as.data.frame(tabla_sd))
tabla_2$Condiciones <- rep(unique(tabla$Condiciones),4)

tabla_total <- merge(tabla_1, tabla_2, by = c("Condiciones", "ind"))

tabla_total$Endulzante <- rep(endulzante, nrow(tabla_total))

dplyr::rename(tabla_total, mean = values.x , SD = values.y, Compuesto = ind)
}

# ---- Lectura y procesamiento de las tablas

fenSU <- read.csv("data/Compuestos fenolicos SU VidaUtil Cronico_1.csv", sep = ";", dec = ",")
fenSA <- read.csv("data/Compuestos fenolicos SA VidaUtil Cronico_1.csv", sep = ";", dec = ",")
fenST <- read.csv("data/Compuestos fenolicos ST VidaUtil Cronico_1.csv", sep = ";", dec = ",")

fenSU_total <- procesado(fenSU, "SU")
fenSA_total <- procesado(fenSA, "SA")
fenST_total <- procesado(fenST, "ST")

fenFlavTotal <- rbind(fenSU_total, fenSA_total, fenST_total)

skimr::skim(fenFlavTotal)

```

Table 1: Data summary

Name	fenFlavTotal
Number of rows	192
Number of columns	5
Column type frequency:	
character	2
factor	1
numeric	2
Group variables	None

Variable type: character

skim_variable	n_missing	complete_rate	min	max	empty	n_unique	whitespace
Condiciones	0	1	13	24	0	48	0
Endulzante	0	1	2	2	0	3	0

Variable type: factor

skim_variable	n_missing	complete_rate	ordered	n_unique	top_counts
Compuesto	0	1	FALSE	4	Der: 48, Eri: 48, Nar: 48, Hes: 48

Variable type: numeric

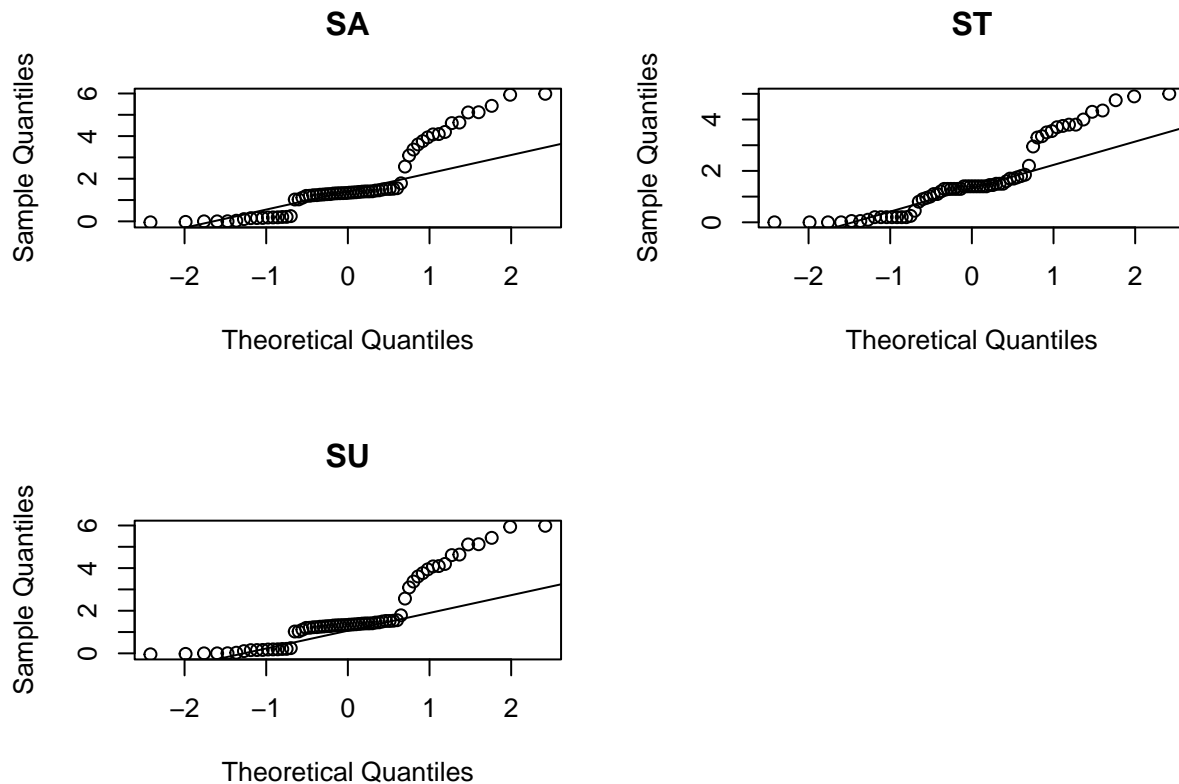
skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100	hist
mean	0	1	1.64	1.47	-0.04	0.56	1.35	1.79	5.98	
SD	0	1	0.03	0.07	0.00	0.00	0.00	0.01	0.57	

Pruebas adicionales

```
shapiro.test(fenFlavTotal$mean)
```

Normalidad

```
##  
##  Shapiro-Wilk normality test  
##  
## data:  fenFlavTotal$mean  
## W = 0.85455, p-value = 1.432e-12  
  
par(mfrow = c(2,2))  
qqnorm(fenFlavTotal[fenFlavTotal$Endulzante == "SA", "mean"], main = "SA")  
qqline(fenFlavTotal[fenFlavTotal$Endulzante == "SA", "mean"])  
qqnorm(fenFlavTotal[fenFlavTotal$Endulzante == "ST", "mean"], main = "ST")  
qqline(fenFlavTotal[fenFlavTotal$Endulzante == "ST", "mean"])  
qqnorm(fenFlavTotal[fenFlavTotal$Endulzante == "SU", "mean"], main = "SU")  
qqline(fenFlavTotal[fenFlavTotal$Endulzante == "SU", "mean"])
```



La media de las medidas no presenta una distribución normal.

Homocedasticidad

```
var.test(x = fenFlavTotal[fenFlavTotal$Endulzante == "SA", "mean"],
        y = fenFlavTotal[fenFlavTotal$Endulzante == "SU", "mean"] )

##
## F test to compare two variances
##
## data:  fenFlavTotal[fenFlavTotal$Endulzante == "SA", "mean"] and fenFlavTotal[fenFlavTotal$Endulzante
## F = 1.4891, num df = 63, denom df = 63, p-value = 0.1167
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.9046806 2.4511295
## sample estimates:
## ratio of variances
##      1.489124

var.test(x = fenFlavTotal[fenFlavTotal$Endulzante == "SA", "mean"],
        y = fenFlavTotal[fenFlavTotal$Endulzante == "ST", "mean"] )

##
## F test to compare two variances
##
## data:  fenFlavTotal[fenFlavTotal$Endulzante == "SA", "mean"] and fenFlavTotal[fenFlavTotal$Endulzante
## F = 1.376, num df = 63, denom df = 63, p-value = 0.208
```

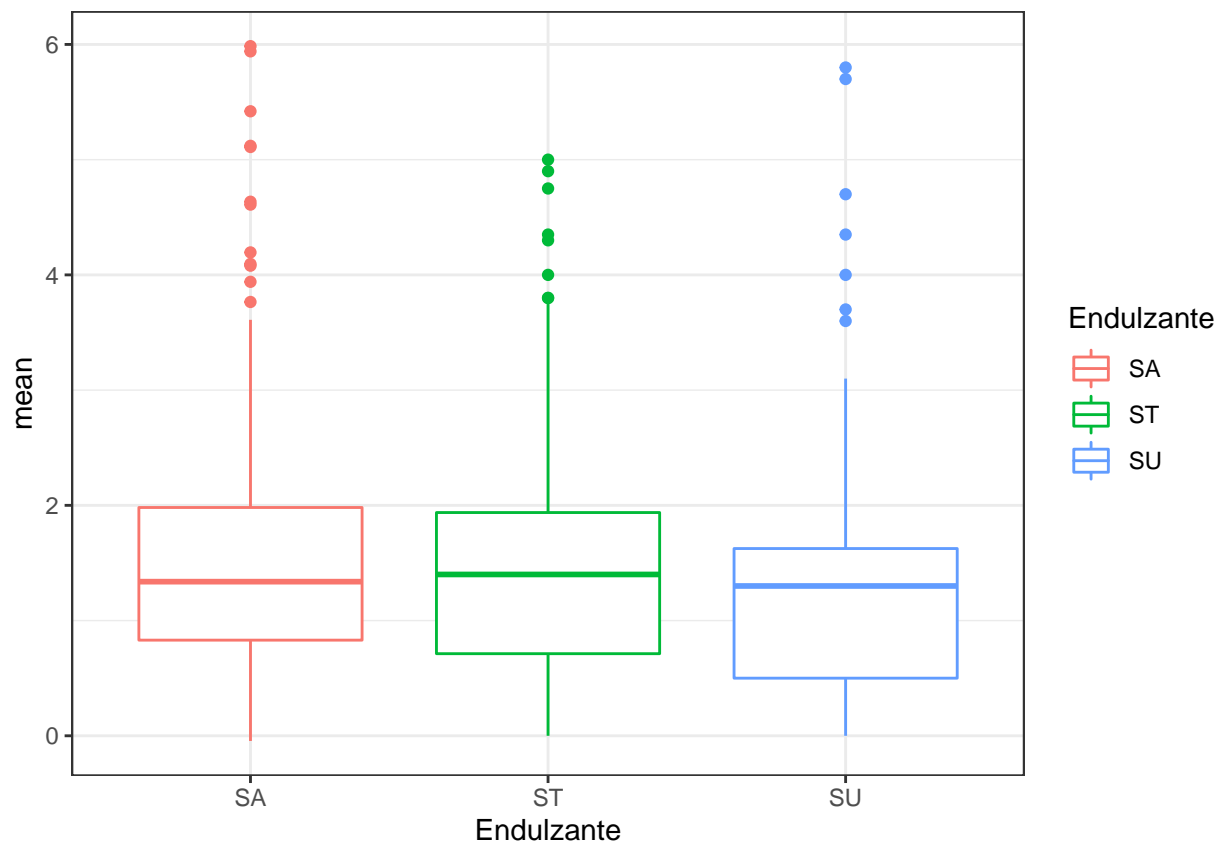
```
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
## 0.8359404 2.2648859
## sample estimates:
## ratio of variances
## 1.375976

var.test(x = fenFlavTotal[fenFlavTotal$Endulzante == "ST", "mean"],
         y = fenFlavTotal[fenFlavTotal$Endulzante == "SU", "mean"] )

##
## F test to compare two variances
##
## data: fenFlavTotal[fenFlavTotal$Endulzante == "ST", "mean"] and fenFlavTotal[fenFlavTotal$Endulzante == "SU", "mean"]
## F = 1.0822, num df = 63, denom df = 63, p-value = 0.7548
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
## 0.6574828 1.7813753
## sample estimates:
## ratio of variances
## 1.082231

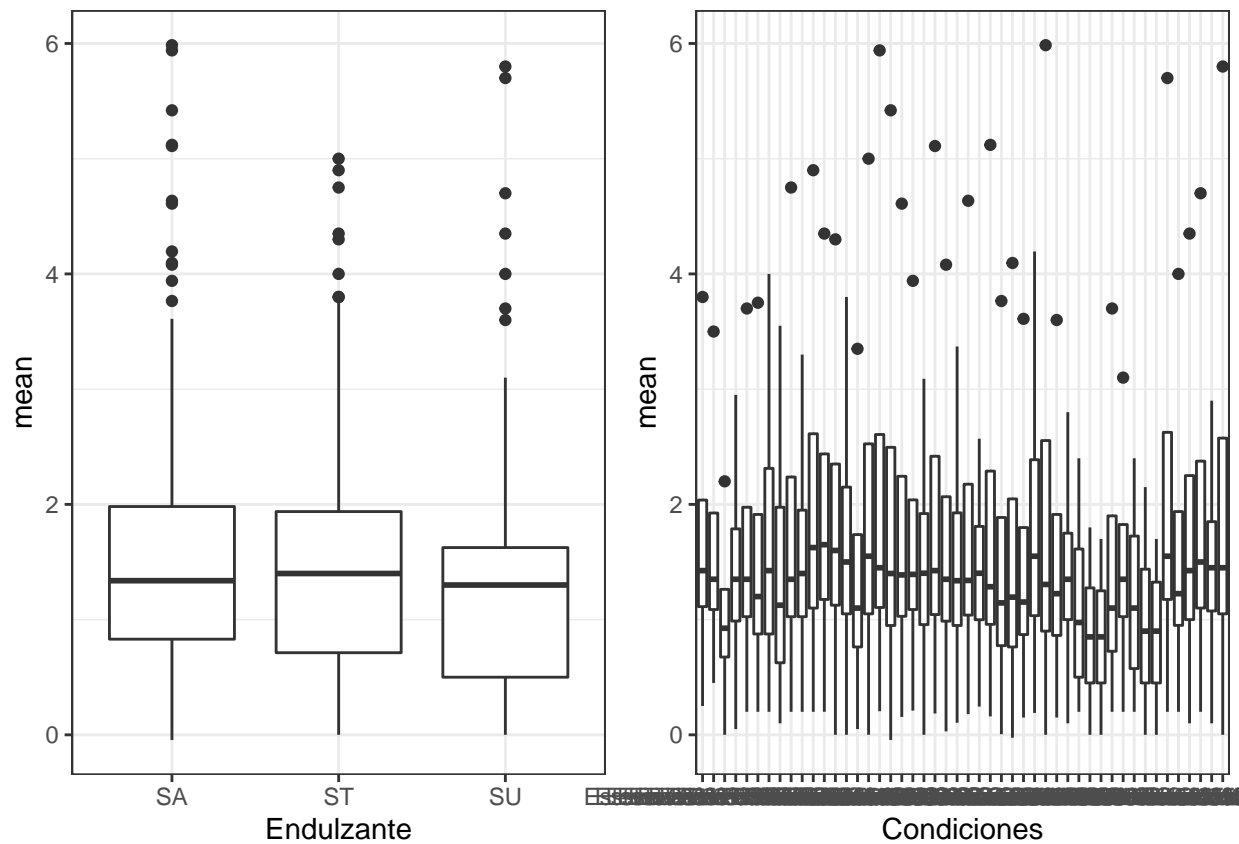
#### Box-plot

ggplot(data = as.data.frame(fenFlavTotal), aes(x = Endulzante, y = mean, color = Endulzante)) +
  geom_boxplot() +
  theme_bw()
```



```
p1 <- ggplot(data = as.data.frame(fenFlavTotal), aes(x = Endulzante, y = mean)) +
  geom_boxplot() + theme_bw()
p2 <- ggplot(data = as.data.frame(fenFlavTotal), aes(x = Condiciones, y = mean)) +
  geom_boxplot() + theme_bw()
p3 <- ggplot(data = as.data.frame(fenFlavTotal), aes(x = Endulzante, y = mean, colour = Condiciones)) +
  geom_boxplot() + theme_bw()

grid.arrange(p1, p2, ncol = 2)
```



p3



Se cumple la condición de homogeneidad de varianzas.

2WAY-ANOVA

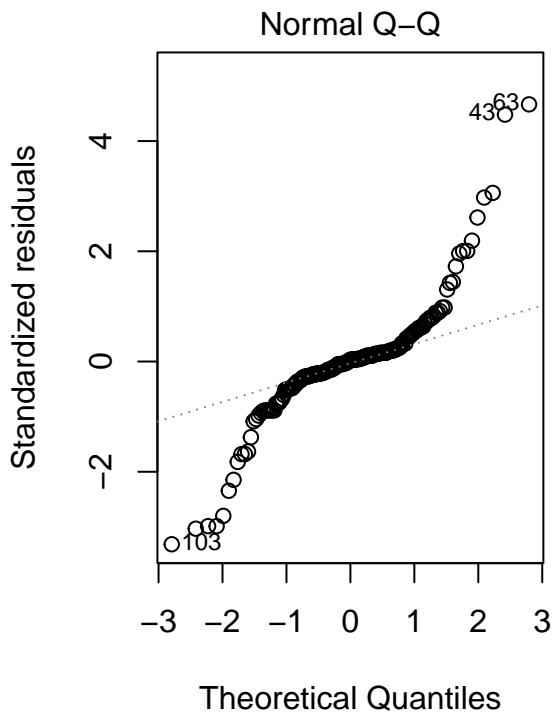
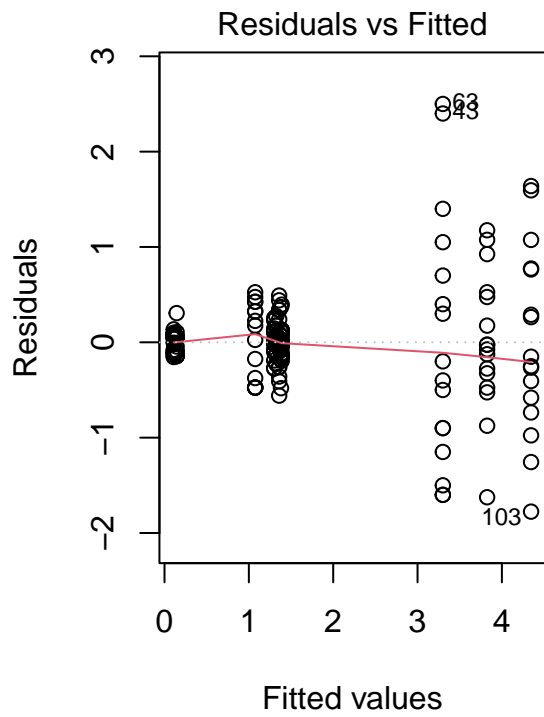
Dado que comparamos dos variables independientes cualitativas y una variable dependiente cuantitativa, la condición de normalidad no se cumple pero existe cierto margen al aproximarse la distribución de la variable dependiente a la normal. Por tanto, realizaremos el estudio con el test paramétrico:

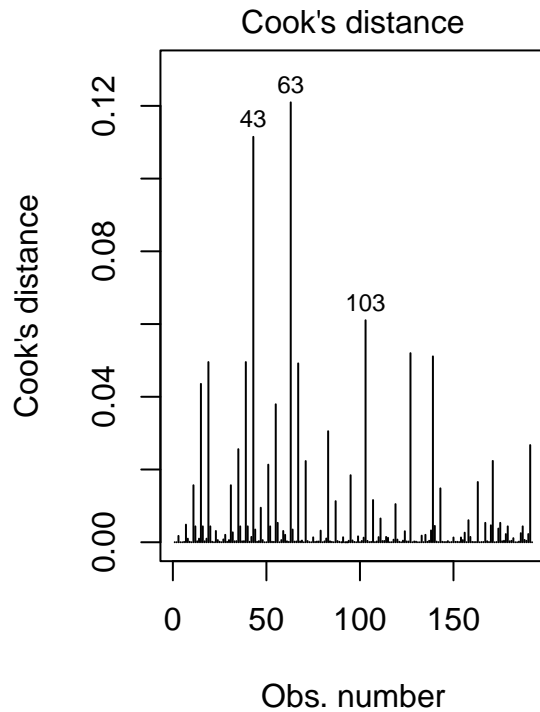
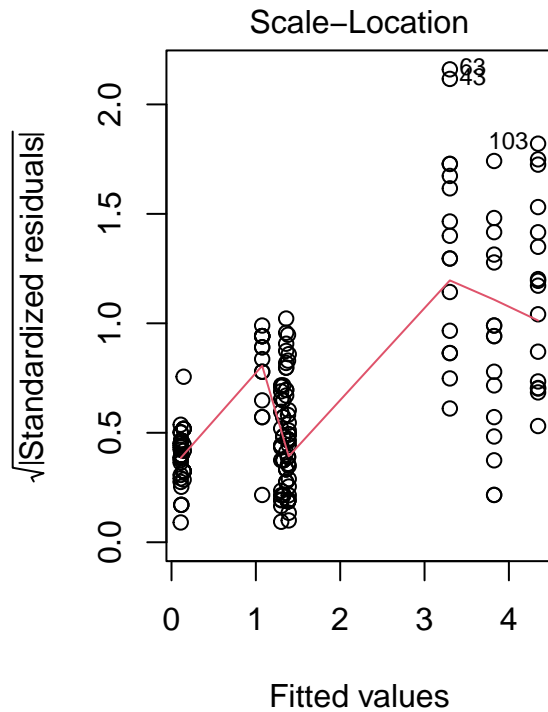
Con interacción entre factores

```
anova <- aov(mean ~ Compuesto * Endulzante, data = fenFlavTotal)
summary(anova)
```

```
##              Df Sum Sq Mean Sq F value    Pr(>F)
## Compuesto      3  350.2   116.72  381.236 < 2e-16 ***
## Endulzante      2    3.7     1.84   6.003 0.00299 **
## Compuesto:Endulzante 6    6.1     1.02   3.344 0.00379 **
## Residuals     180   55.1     0.31
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
par(mfrow=c(1,2))
plot(anova, which=1:4)
```





```
par(mfrow = c(1,1))
lsr::etaSquared(anova)
```

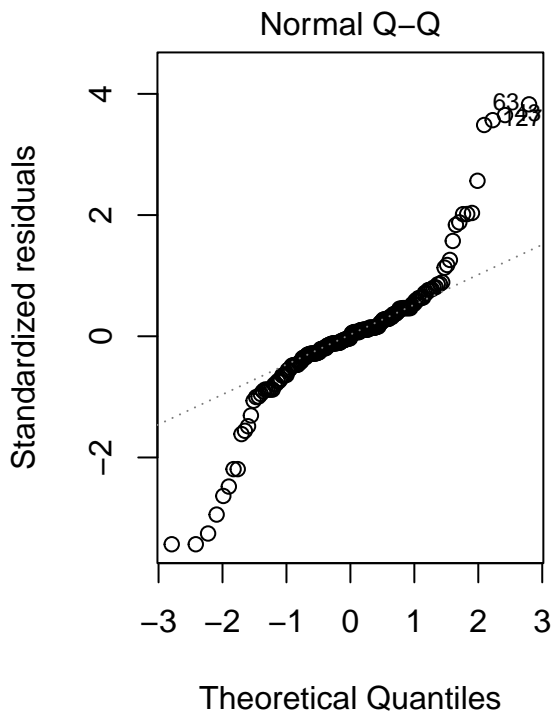
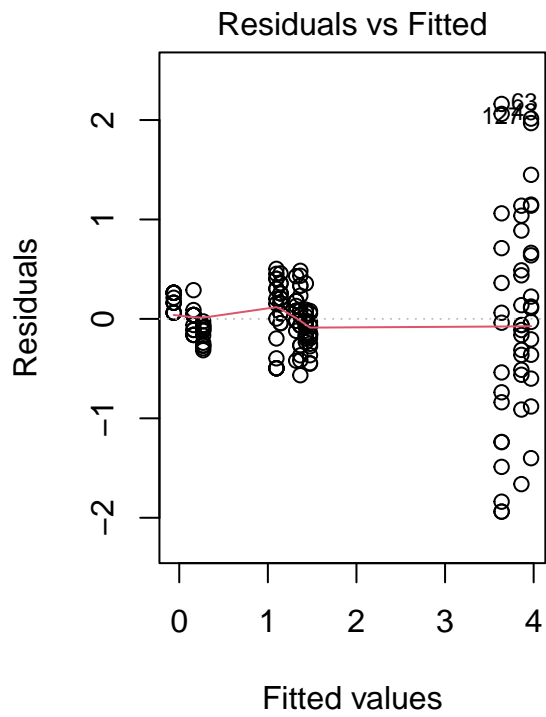
```
##               eta.sq eta.sq.part
## Compuesto      0.843579429  0.86401847
## Endulzante      0.008855346  0.06252886
## Compuesto:Endulzante 0.014800446  0.10029766
```

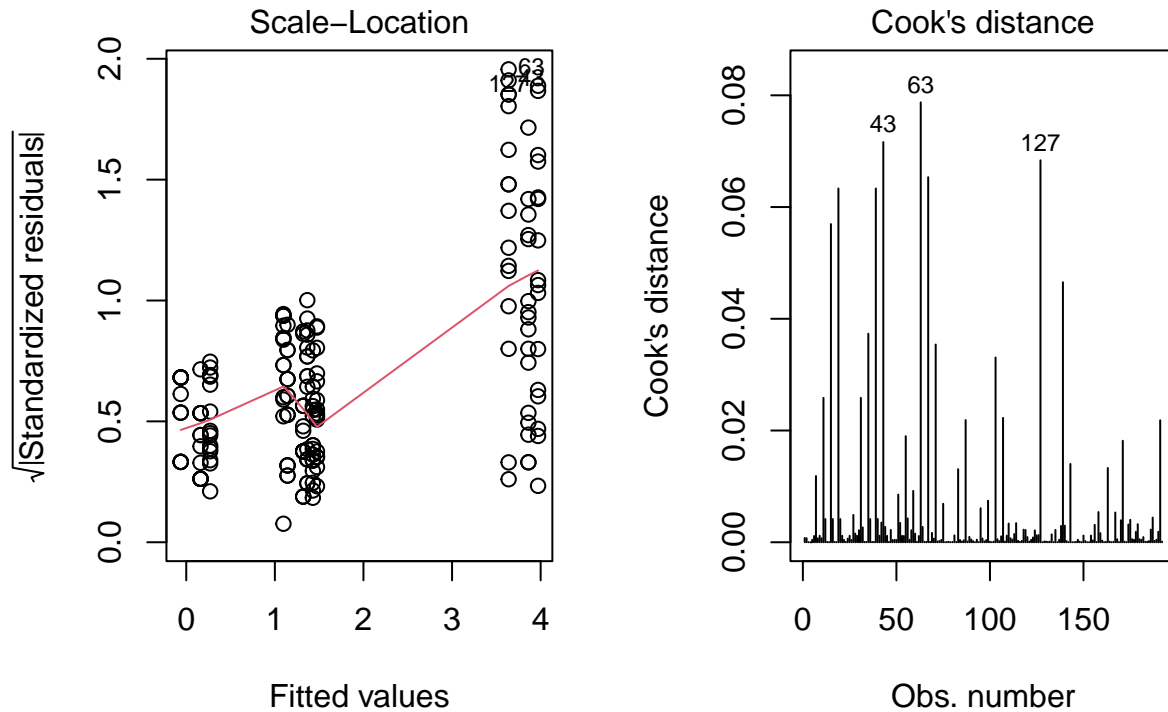
```
# Sin interacción entre factores
```

```
anova2 <- aov(mean ~ Compuesto + Endulzante, data = fenFlavTotal)
summary(anova2)
```

```
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Compuesto   3  350.2   116.72  354.433 < 2e-16 ***
## Endulzante   2    3.7    1.84   5.581 0.00443 **
## Residuals  186   61.3    0.33
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
par(mfrow=c(1,2))
plot(anova2, which=1:4)
```





```
par(mfrow = c(1,1))
lsr::etaSquared(anova2)
```

```
##           eta.sq eta.sq.part
## Compuesto 0.843579429 0.85111636
## Endulzante 0.008855346 0.05661241
```

Se observa que se cumplen las condiciones para la ANOVA, y que existe un efecto significativo sobre la media de la medida por parte del compuesto y del endulzante, así como del efecto de la interacción de ambos factores. Vemos que el efecto es mucho más fuerte en el compuesto que en el endulzante

Dudas

Hesperidina?

Muestras bio

Cronico Orina Atipicos

Cronico plasma Atipicos