Treball aliments (Quarto)

Joan Vicenç

# Introducción

El objetivo de este proyecto es determinar las diferencias entre las propiedades de ciertos alimentos –por ejemplo, el % de humedad.

Genericament vamos a aplicar los siguientes tests:

* [Shapiro–Wilk](https://es.wikipedia.org/wiki/Prueba_de_Shapiro%E2%80%93Wilk#:~:text=En%20estad%C3%ADstica%2C%20la%20prueba%20de%20Shapiro%E2%80%93Wilk%20se%20usa,pruebas%20m%C3%A1s%20potentes%20para%20el%20contraste%20de%20normalidad.): “En estadística, la prueba de Shapiro–Wilk se usa para contrastar la normalidad de un conjunto de datos. Se plantea como hipótesis nula que una muestra proviene de una población normalmente distribuida. Fue publicado en 1965 por Samuel Shapiro y Martin Wilk. Se considera uno de las pruebas más potentes para el contraste de normalidad.”
* Si suponemos normalidad:
  + [Anova](https://es.wikipedia.org/wiki/An%C3%A1lisis_de_la_varianza): Sirve para estudiar la varianza
  + [T-Student](https://es.wikipedia.org/wiki/Prueba_t_de_Student) + [Bonferroni](https://es.wikipedia.org/wiki/Correcci%C3%B3n_de_Bonferroni):“es cualquier prueba en la que el estadístico utilizado tiene una distribución t de Student si la hipótesis nula es cierta. Se aplica cuando la población estudiada sigue una distribución normal, pero el tamaño muestral es demasiado pequeño como para que el estadístico en el que está basada la inferencia esté normalmente distribuido, utilizándose una estimación de la desviación típica en lugar del valor real” + “La prueba de hipótesis estadística se basa en rechazar la hipótesis nula si la probabilidad de los datos observados en las hipótesis nulas es baja. Si se prueban múltiples hipótesis, aumenta la probabilidad de un evento raro y, por lo tanto, aumenta la probabilidad de rechazar incorrectamente una hipótesis nula (es decir, cometer un error de tipo I.)”
* Si NO suponemos normalidad:
  + [Kruskal-Wallis](https://es.wikipedia.org/wiki/Prueba_de_Kruskal-Wallis): Com la ANOVA pero no assume normalidad
  + [Wilcoxon](https://es.wikipedia.org/wiki/Prueba_de_los_rangos_con_signo_de_Wilcoxon): “La prueba de los rangos con signo de Wilcoxon es una prueba no paramétrica para comparar el rango medio de dos muestras relacionadas y determinar si existen diferencias entre ellas. Se utiliza como alternativa a la prueba t de Student cuando no se puede suponer la normalidad de dichas muestras.”

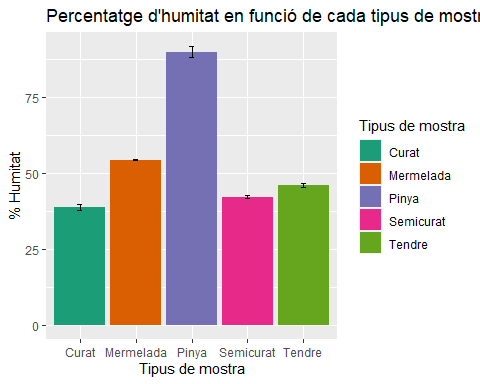
# Anàlisis

## Anàlisis de la base de datos

La mida de la muestra es de 18 observaciones agrupadas en 5 grupos con las siguientes propiedades:

# A tibble: 5 x 3  
 `Tipus de mostra` Mitjana Error  
 <chr> <dbl> <dbl>  
1 Tendre 46.0 0.634  
2 Semicurat 42.2 0.402  
3 Curat 38.7 0.981  
4 Pinya 89.9 1.82   
5 Mermelada 54.3 0.174

## Gráfico



## Test

# Comporvam normalitat  
shapiro.test(data$Contingut)

Shapiro-Wilk normality test  
  
data: data$Contingut  
W = 0.71897, p-value = 0.0001338

Al tener un p.valor infimo, no podemos assumir la normalidad de los datos.

# Si no es normal: Kruskal test  
kruskal.test(data$Contingut ~ data$Mostra)

Kruskal-Wallis rank sum test  
  
data: data$Contingut by data$Mostra  
Kruskal-Wallis chi-squared = 16.07, df = 4, p-value = 0.002926

Al tener un p.valor infimo, assumimos que al menos un parámetro es diferente al resto.

# Si no es normal: Wilcoxon test  
pairwise.wilcox.test(data$Contingut, data$Mostra, paired = FALSE, p.adjust.method = "holm")

Pairwise comparisons using Wilcoxon rank sum exact test   
  
data: data$Contingut and data$Mostra   
  
 Curat Mermelada Pinya Semicurat  
Mermelada 0.46 - - -   
Pinya 0.46 0.46 - -   
Semicurat 0.46 0.46 0.46 -   
Tendre 0.29 0.46 0.46 0.29   
  
P value adjustment method: holm

Al tener p.valor “alto” podemos afirmar que hay diferencias signficativas entre las muestras.