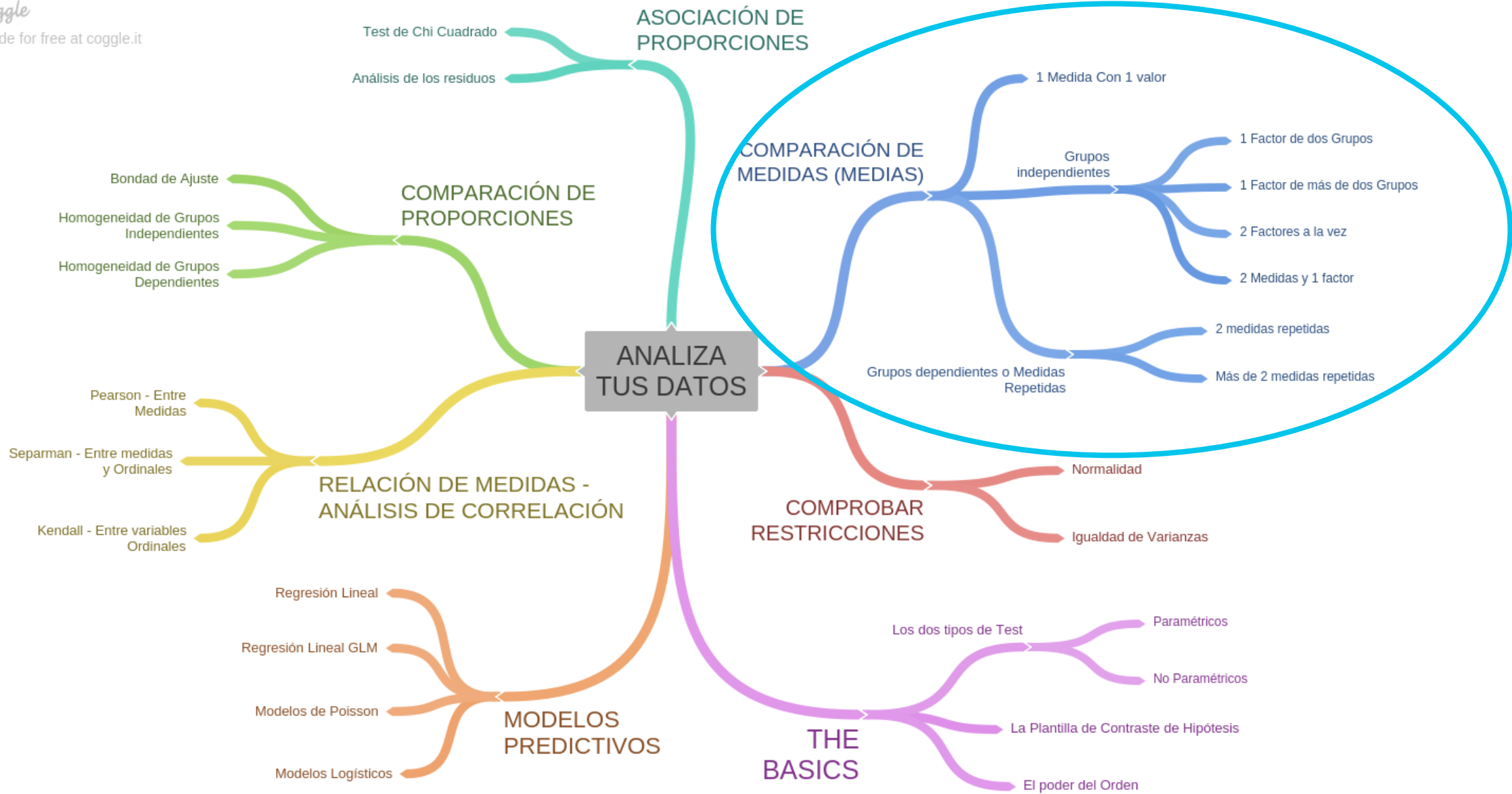


# Comparación de medidas

La comparación de medidas paso a paso

## ¿Que vas a ver en este bloque?

- Qué quiere decir comparar medidas a nivel práctico
- La comparación de medidas en situaciones prácticas
  - Una medida y un valor
  - Una medida por dos grupos de un factor
    - Grupos independientes
    - Medidas repetidas
  - Una medida por más de dos grupos de un factor (one-way ANOVA)
    - Grupos independientes
    - Medidas repetidas
  - Una medida en dos factores (two-way ANOVA)
  - Dos medidas en un factor (MANOVA)

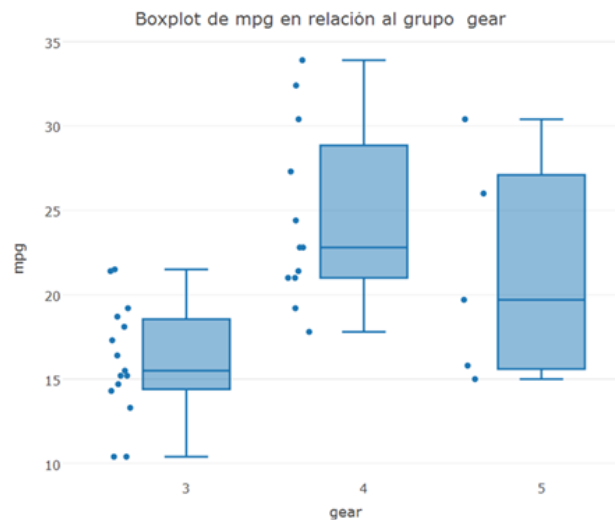
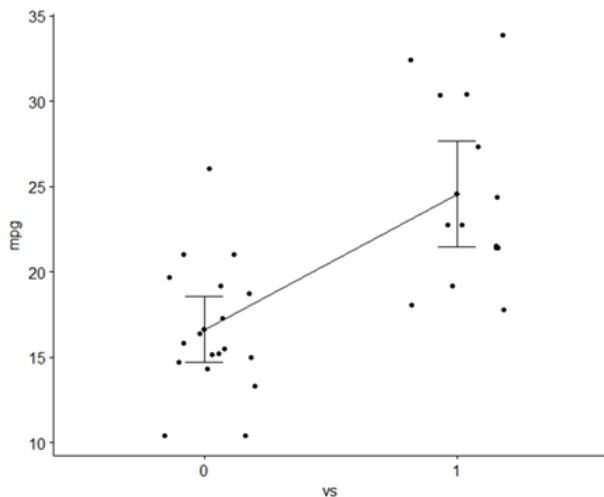


# Qué quiere decir comparar medidas

El significado de la comparación de medidas

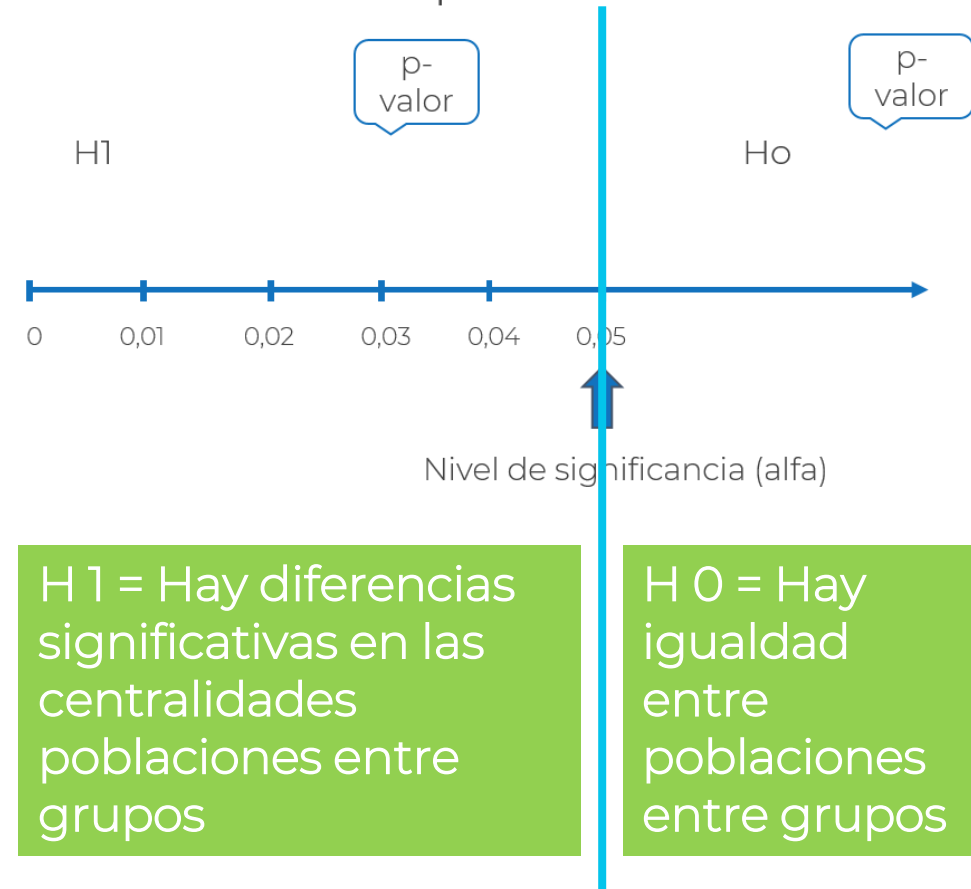
# El significado de la comparación

- Comparar la centralidad entre dos distribuciones
- Variable Dependiente o Respuesta: mpg (medida)
- Variable Independientes o de Estudio: vs (factor)
  - vs = 1 → automático
  - vs = 0 → manual



**OBJETIVO:**  
Comparar una medida con un factor de estudio. Cómo influye el factor en la medida

Regla de decisiones para la prueba de hipótesis



coggle  
One Way ANOVA (P)  
made for free at coggle.it

One Way ANOVA Welch (P)

Kruskall Wallis

Tukey HSD o Pair Wise

Grupos Independientes

1 MEDIDA Y MÁS  
DE 2 GRUPOS DE  
1 FACTOR

ANOVA de medidas repetidas

Friedman Test

Medidas Repetidas

INTERVALOS DE  
CONFIANZA

Two-sided ANOVA -  
caso NO Balanceado  
(P)

Two-sided ANOVA  
- caso Balanceado  
(P)

Scheirer Ray Hare (NP)

1 MEDIDA y 2 o  
más FACTORES

## COMPARACIÓN DE MEDIDAS

1 MEDIDA Y 2  
GRUPOS DE 1  
FACTOR

Grupos Independientes

T-test grupos independientes (P)

T-test con corrección de Welch (P)

Wilcoxon de grupos Independientes (NP)

Medidas repetidas

T-test grupos pareados (P)

Wilcoxon para grupos pareados (NP)

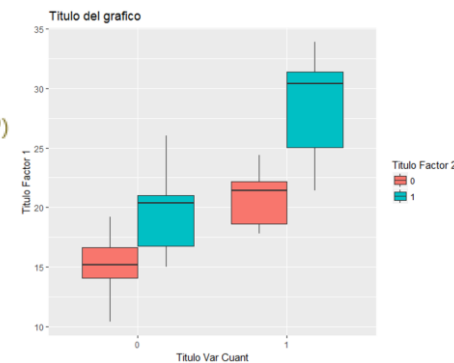
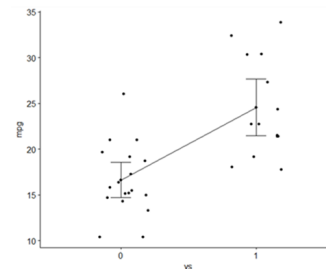
1 MEDIDA Y 1  
VALOR

T-test de una muestra (P)

Wilcoxon de 1 muestra (NP)

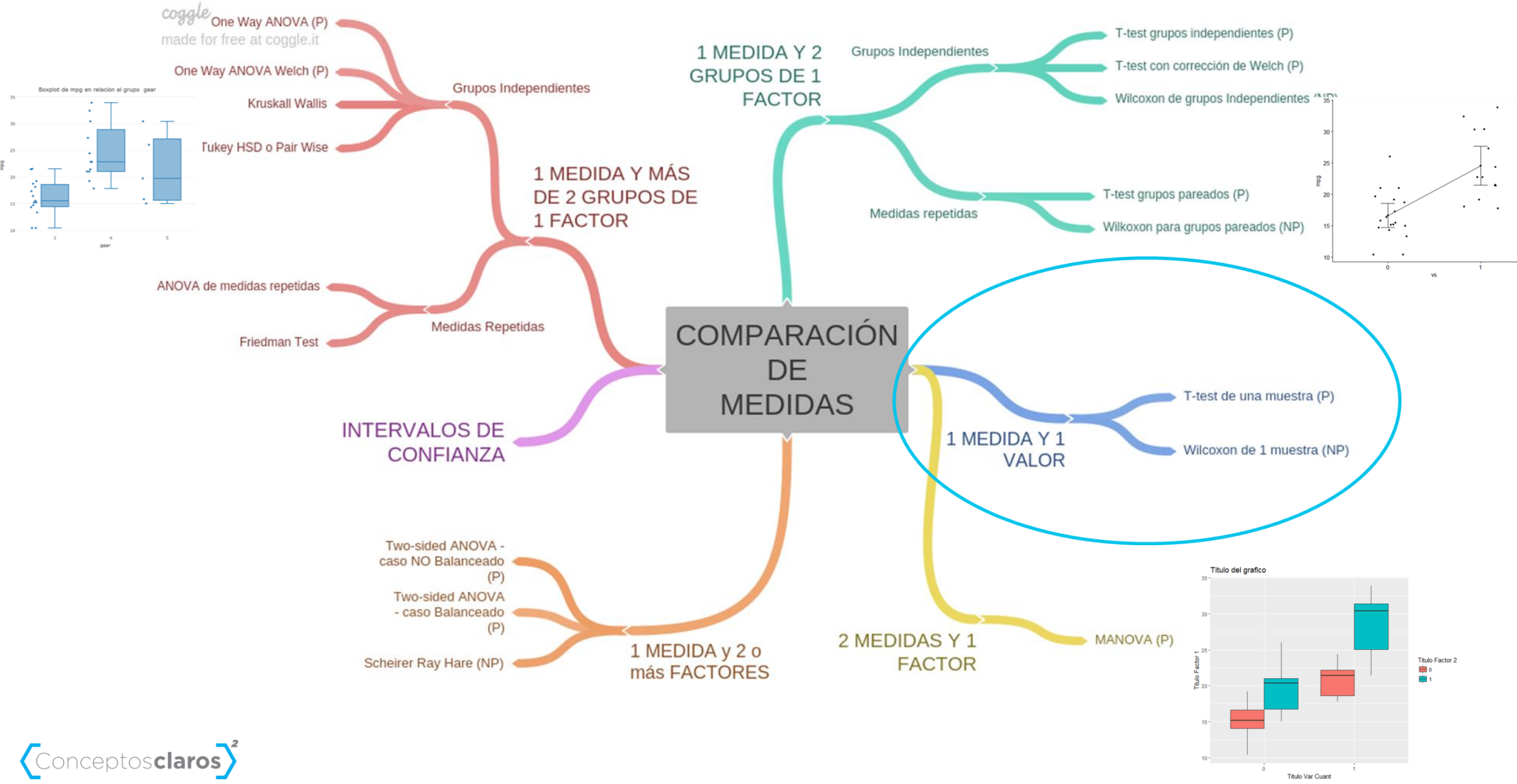
2 MEDIDAS Y 1  
FACTOR

MANOVA (P)



# Comparar 1 medida y 1 valor conocido

El paso a paso para comparar 1 medida con 1 valor conocido

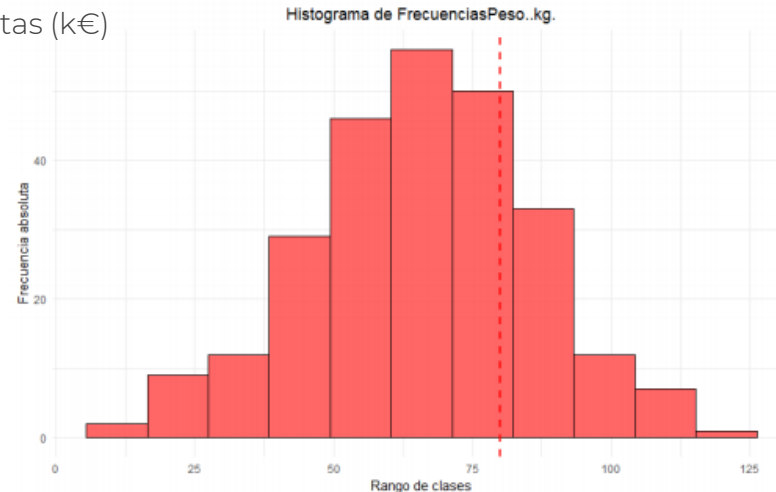




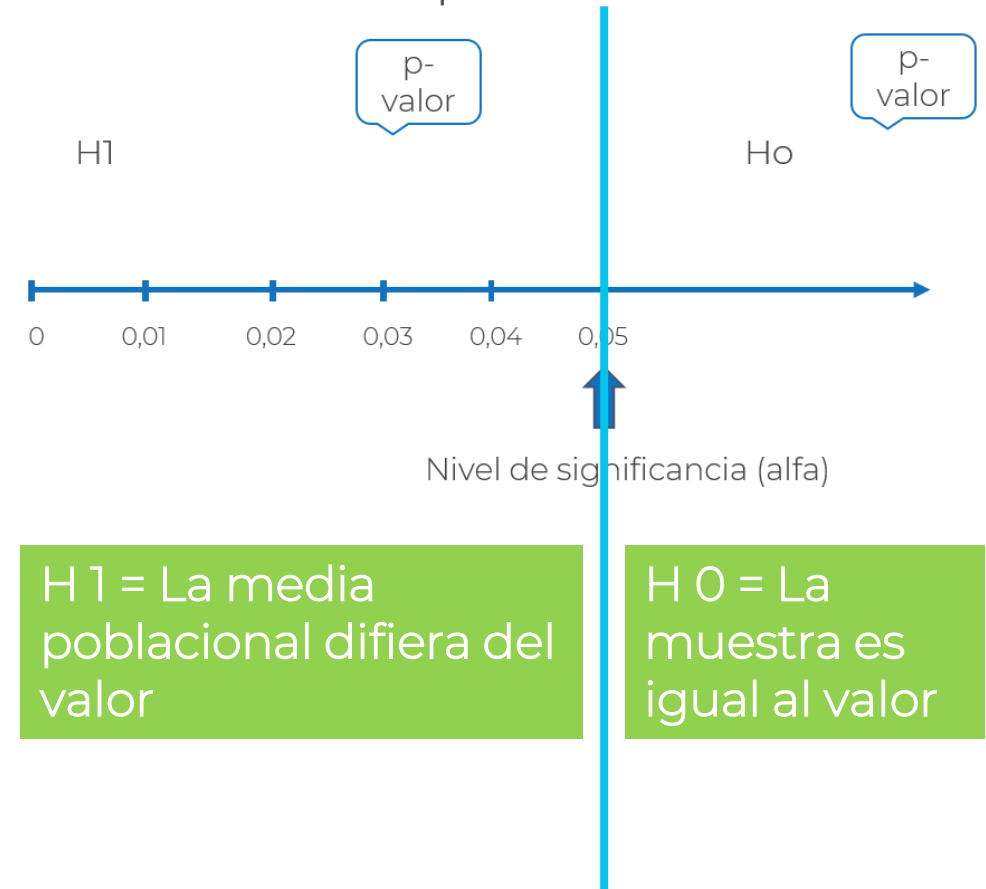
**OBJETIVO:**  
Comparar si la media poblacional es diferente a un valor conocido

## Las claves para comparar la muestra y un valor

- Se utiliza para comparar el aumento entre dos grupos o las diferencias entre dos estados
- Se calcula una variable diferencia:
  - Antes – Después
  - Diferencia entre dos grupos
- Ej1: El método a mejora a los pacientes 50 puntos de índice de discapacidad
  - Variable respuesta: índice discapacidad
  - Valor: 50
- Ej2: Las ventas son más grandes que 40k€ entre los dos grupos de clientes
  - Variable respuesta: ventas (k€)
  - Valor 40 k€
- Descripción:** histograma y valor a comparar



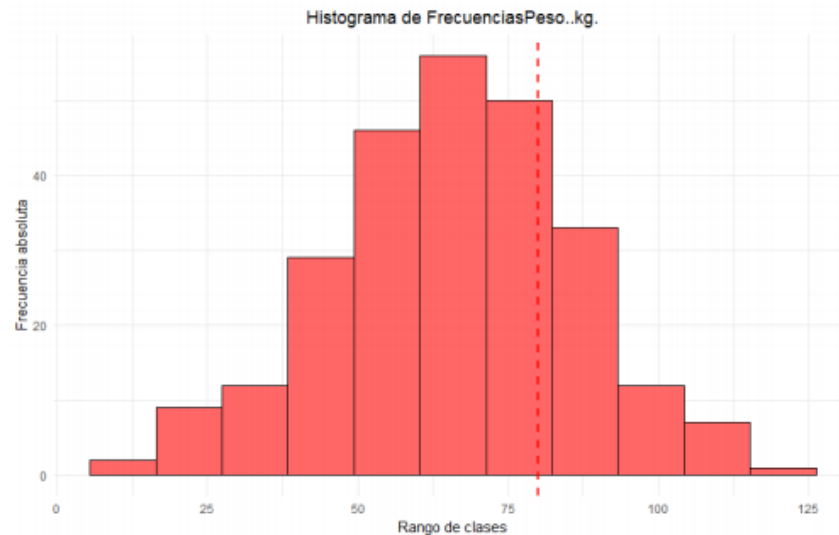
Regla de decisiones para la prueba de hipótesis



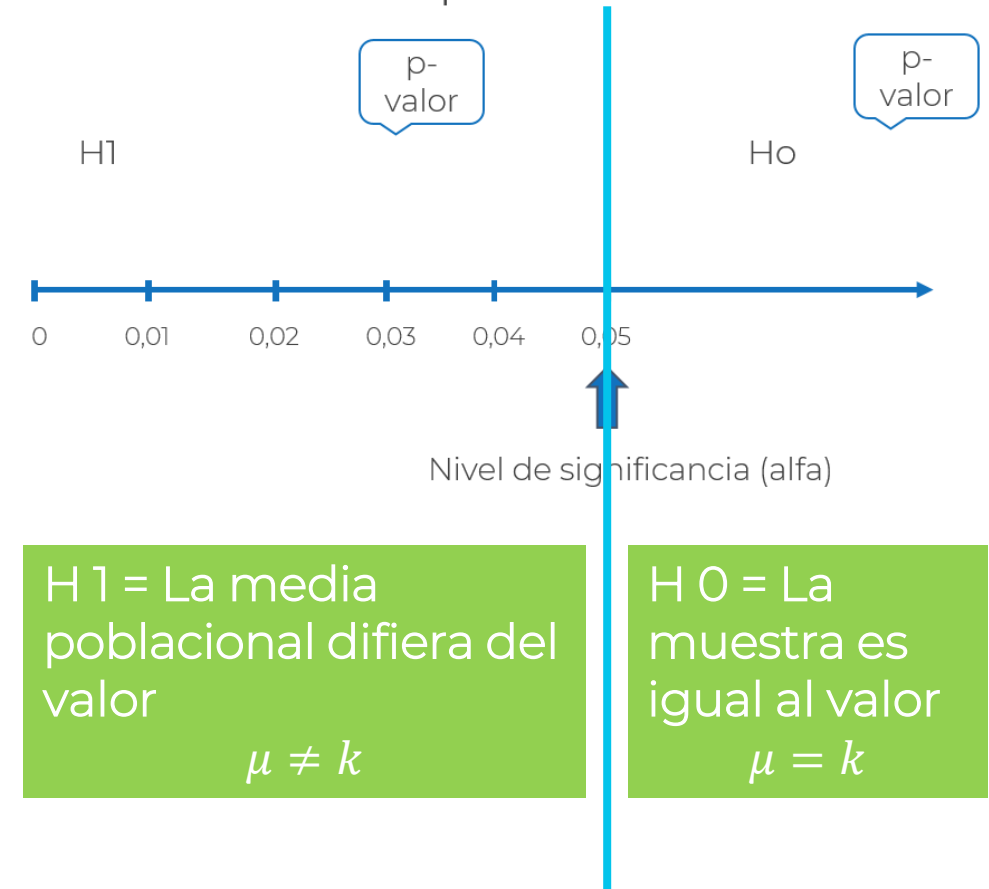
**OBJETIVO:**  
Comparar si la media poblacional es diferente a un valor conocido

# Las claves para comparar la muestra y un valor

- Los dos test posibles son:
  - Si la medida proviene de una distribución normal → T-test de una muestra
  - Si la medida NO proviene de una distr. Normal → Wilcoxon de una muestra

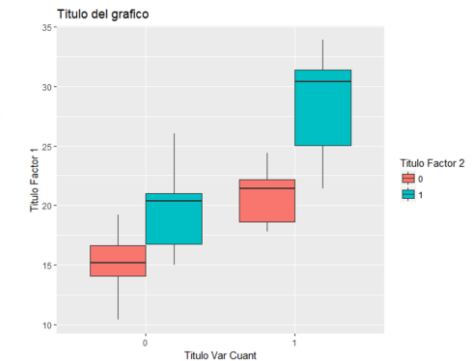
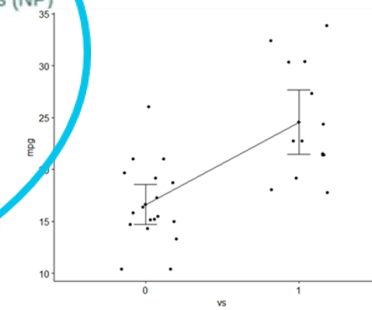
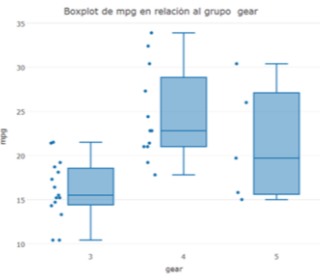
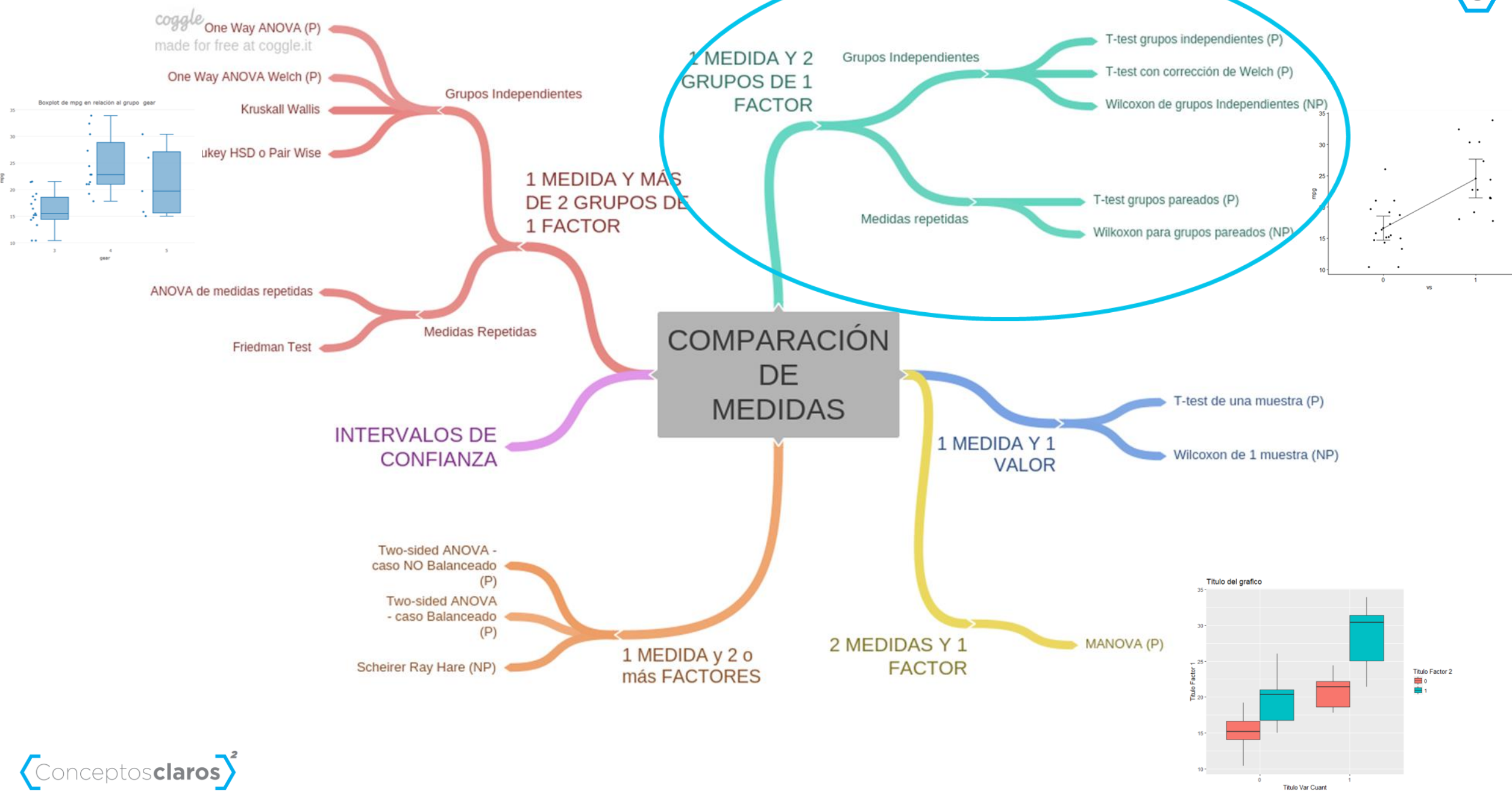


Regla de decisiones para la prueba de hipótesis



# Comparar 1 medida y 1 factor de dos grupos

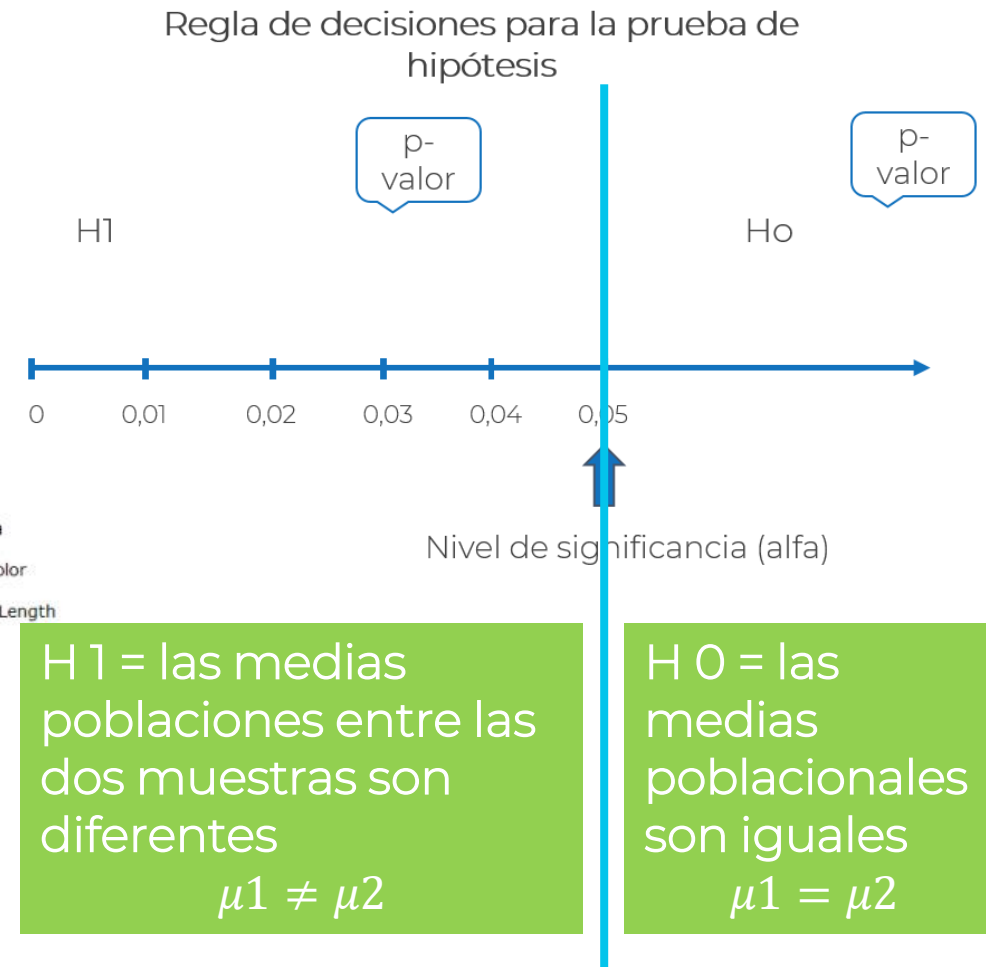
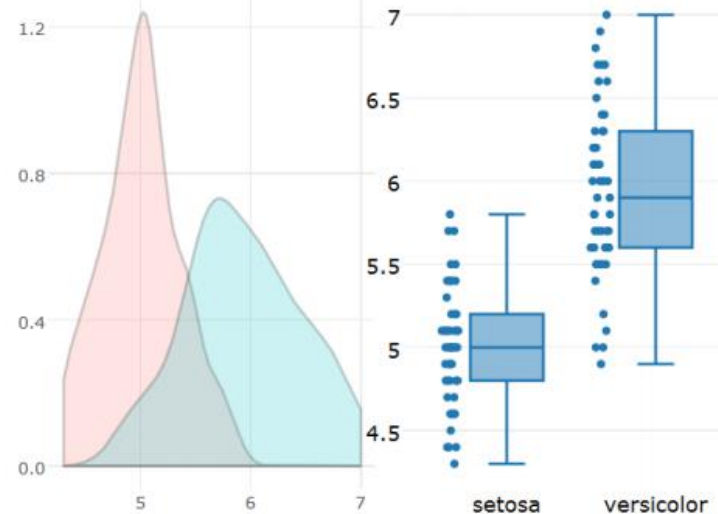
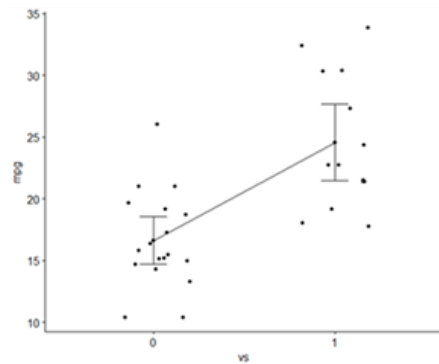
El paso a paso para comparar 1 medida entre dos grupos



# Las claves para comparar una medida y un factor de estudio de dos grupos

**OBJETIVO:**  
Comparar las dos medias poblacionales entre el factor de estudio

- Se utiliza para comparar casos control y de investigación en salud. O cualquier medida que quieras comparar en dos grupos.
- Podemos analizar si un factor influye en la respuesta y cómo
- Variable respuesta: medida → Long. Sépalo
- Variable de estudio: factor → Especie (2 grupos)
- H1: diferencias significativas entre las dos especies
- Descripción:** histograma de densidad, boxplot de 1 factor, diagrama de medias



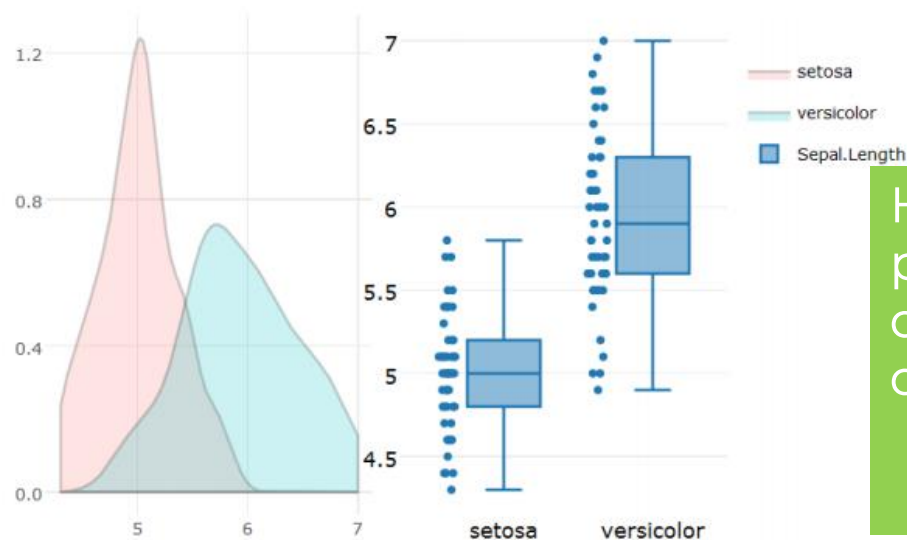
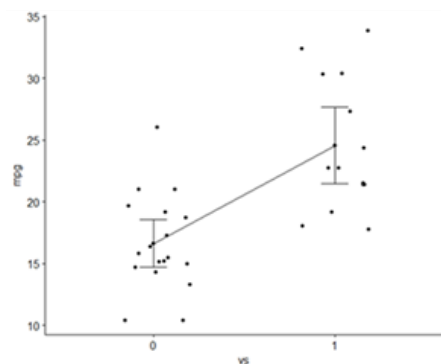
# Las claves para comparar una medida y un factor de estudio de dos grupos

**OBJETIVO:**  
Comparar las dos medias poblacionales entre el factor de estudio

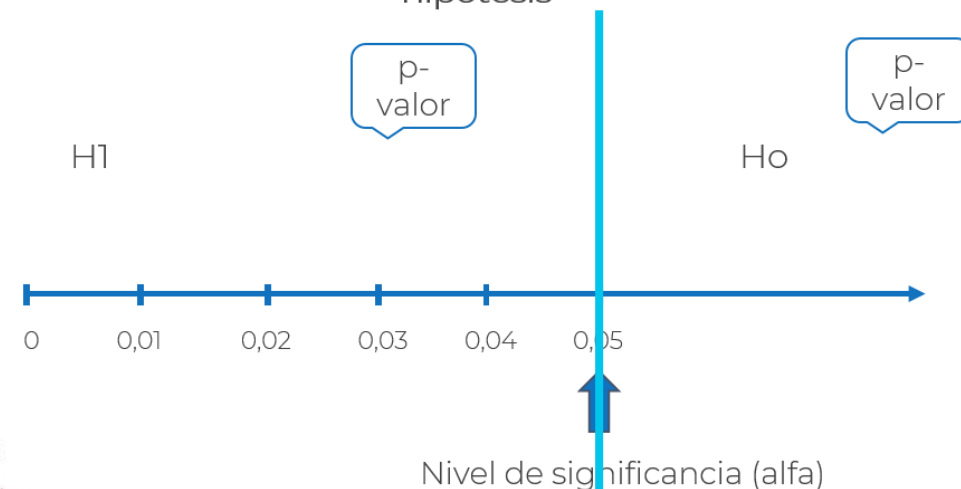
- El Resultado:

welch Two Sample t-test

```
data: var1 and var2
t = -10.521, df = 86.538, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -1.1057074 -0.7542926
sample estimates:
mean of x mean of y
  5.006    5.936
```



Regla de decisiones para la prueba de hipótesis



H 1 = las medias poblacionales entre las dos muestras son diferentes

$$\mu_1 \neq \mu_2$$

$$\mu_1 - \mu_2 \neq 0$$

H 0 = las medias poblacionales son iguales

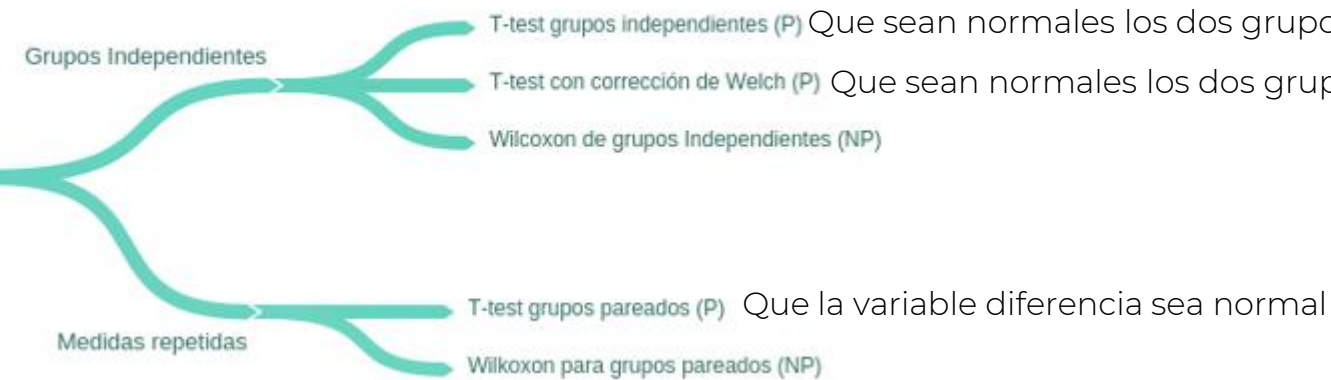
$$\mu_1 = \mu_2$$

$$\mu_1 - \mu_2 = 0$$

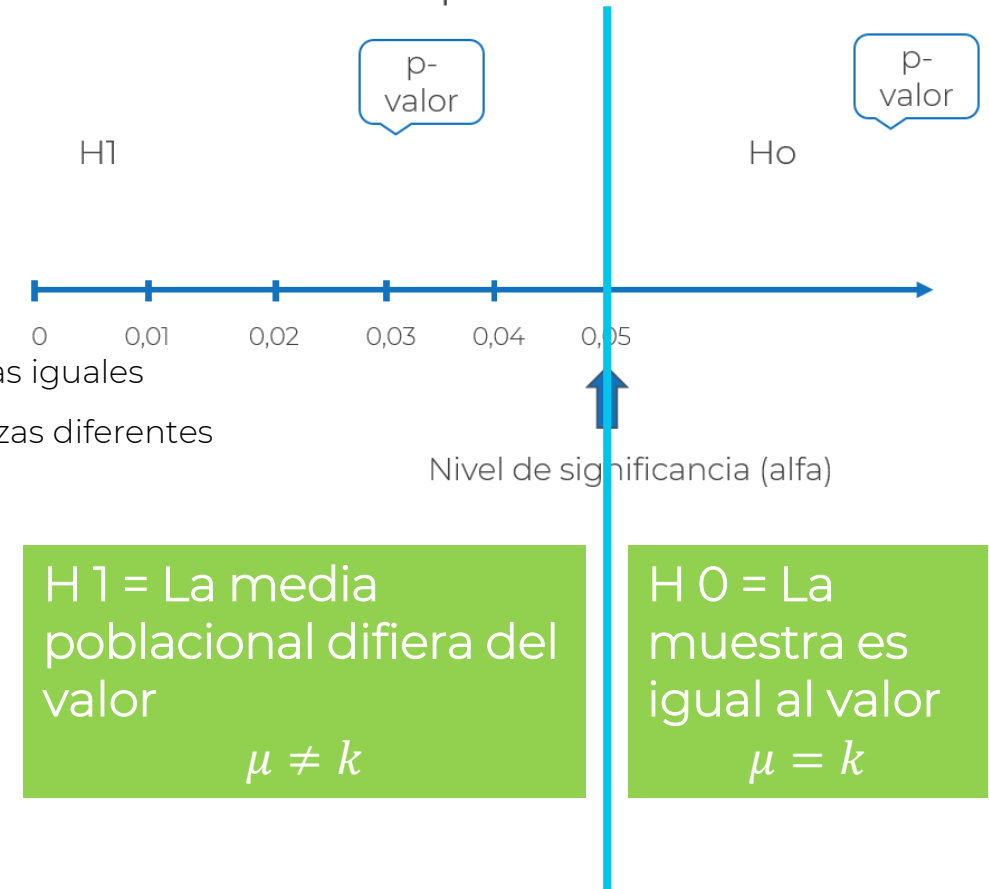
# Las claves para comparar una medida y un factor de estudio de dos grupos

**OBJETIVO:**  
Comparar si la media poblacional es diferente a un valor conocido

- Los dos casos:
  - Medidas repetidas = una misma muestra medida dos veces (antes y después de un tratamiento por ejemplo)
  - Grupos independientes = todos los demás casos, son muestras distintas



Regla de decisiones para la prueba de hipótesis



# Take away

El resumen de la 1/2 lección



## Lo más importante de la lección

- La comparación de medidas es una comparación de medias poblacionales por grupos
- Podemos dividir las comparaciones por casos en la práctica
- Comparar una medida y un valor es utilizado para comparar la diferencia entre dos estados o entre dos grupos controlados
- Comparar medias poblacionales es el pan de cada día en muchos estudios científicos – sobre todo casos control e investigación
  - Grupos dependientes = medida repetidas – misma muestra medida en dos situaciones

# Tú turno

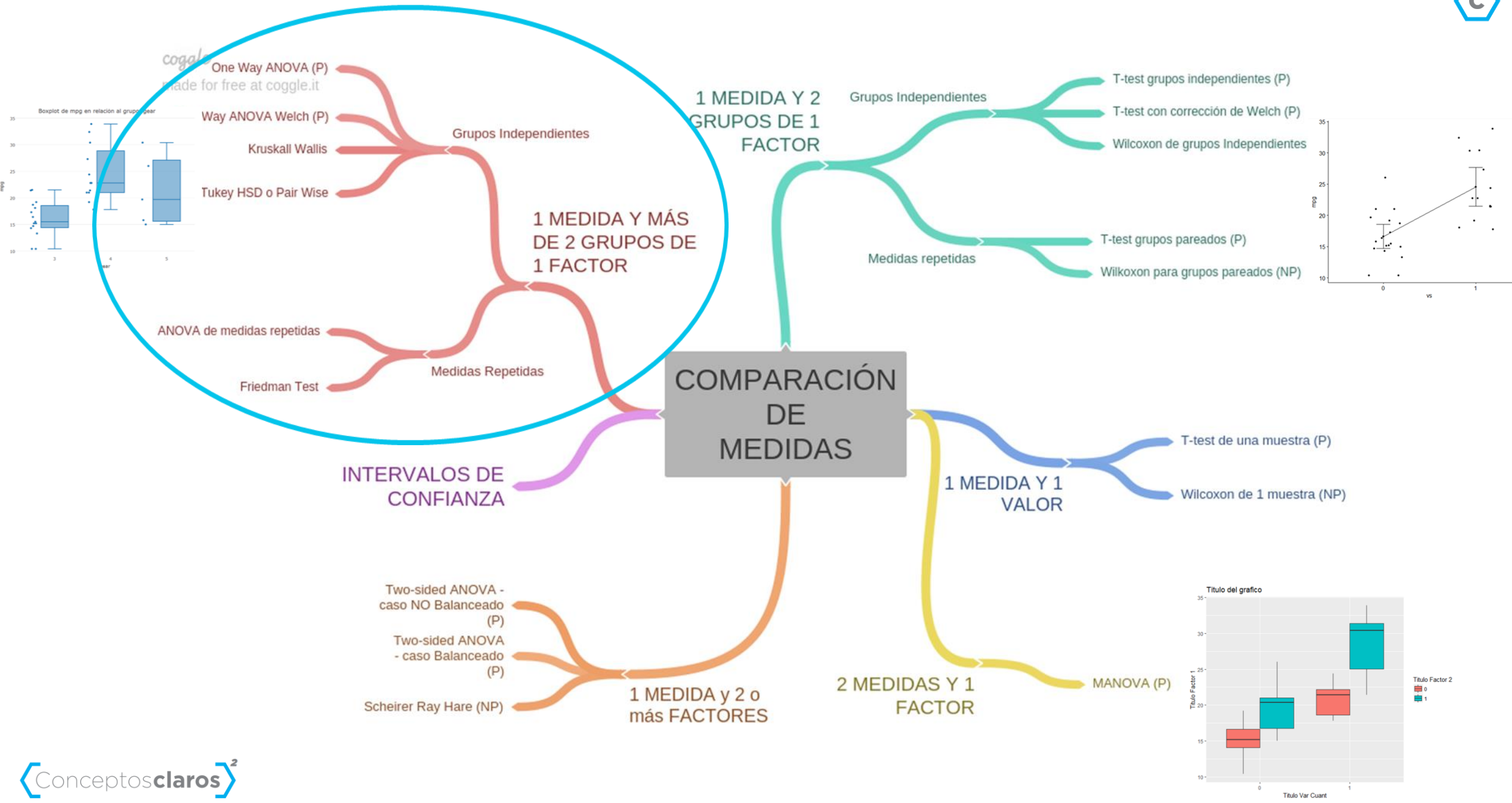
A por tus primeros test estadísticos

## A poner en práctica lo que has visto

- Descarga la hoja de trabajo
- Y utiliza la comparación de dos grupos y la comparación de 1 medida y un valor
  - (para comprobar las restricciones deberás utilizar el test KS y Levene ;) )
- ¡A por ello!

# Comparar 1 medida y 1 factor de más de dos grupos

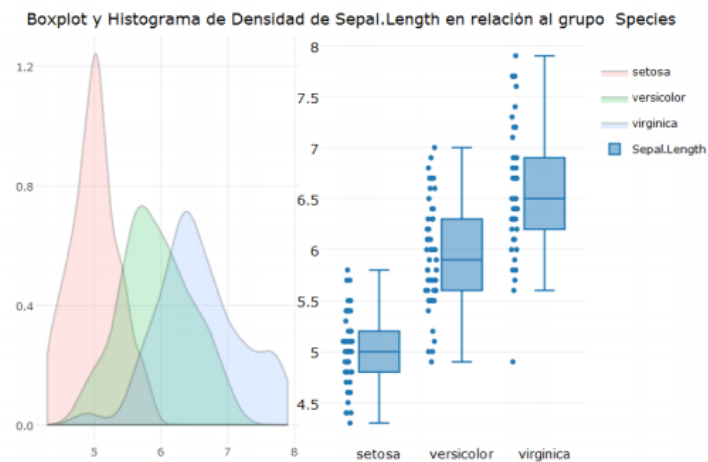
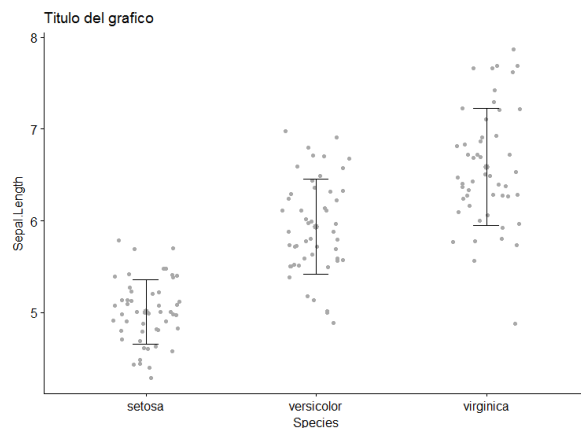
El paso a paso para utilizar el análisis de varianza o tabla  
ANOVA



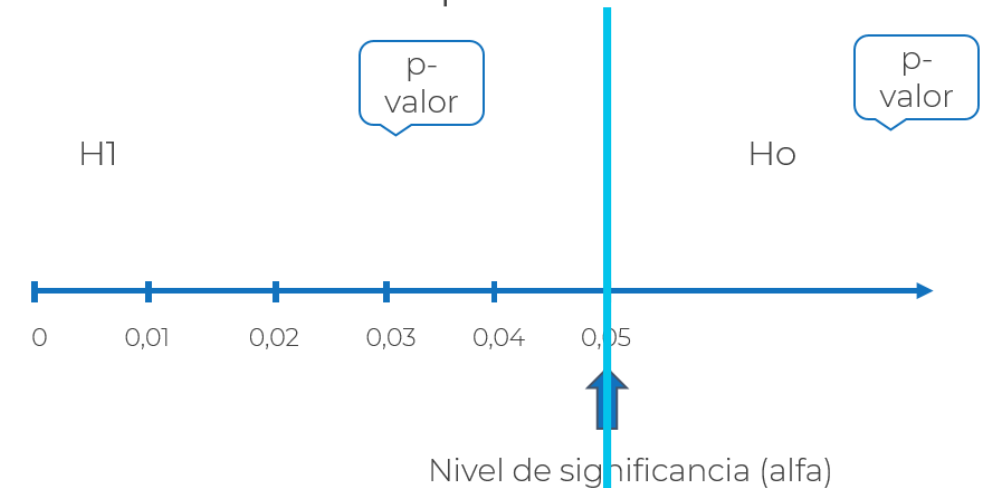
**OBJETIVO:**  
Comparar más de dos medias poblacionales entre el factor de estudio

# Las claves para analizar las varianzas ANOVA

- Se utiliza para comparar las diferencias de una medida en más de dos grupos de estudio
- Podemos analizar si un factor influye en la respuesta y cómo
- Variable respuesta: medida → Long. Sépalo
- Variable de estudio: factor → Especie (3 grupos)
- H1: algún grupo es diferente al resto
- Descripción:** histograma de densidad, boxplot de 1 factor, diagrama de medias / IC



Regla de decisiones para la prueba de hipótesis



H 1 = al menos una media poblacional es diferente al resto de grupos

$$Ej: \mu_1 \neq \mu_2 = \mu_3$$

H 0 = las medias poblacionales son iguales

$$\mu_1 = \mu_2 = \mu_3$$

**OBJETIVO:**  
Comparar más de dos medias poblacionales entre el factor de estudio

# Las claves para analizar las varianzas ANOVA

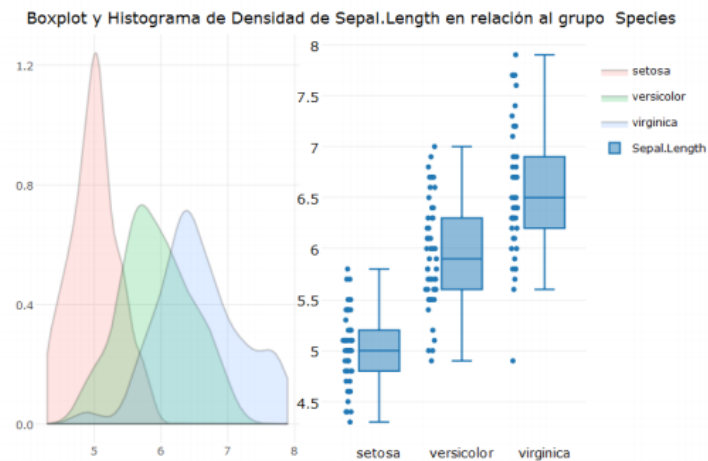
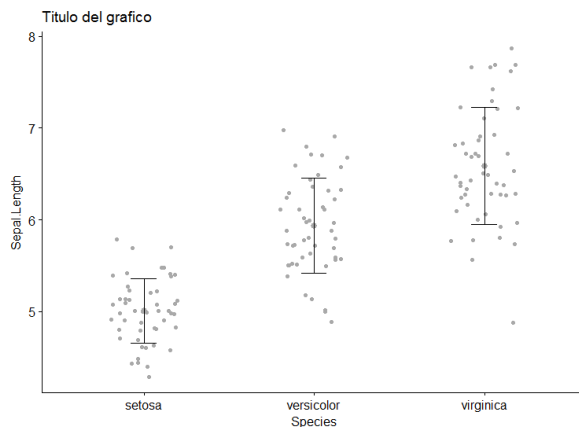
- El resultado:

```

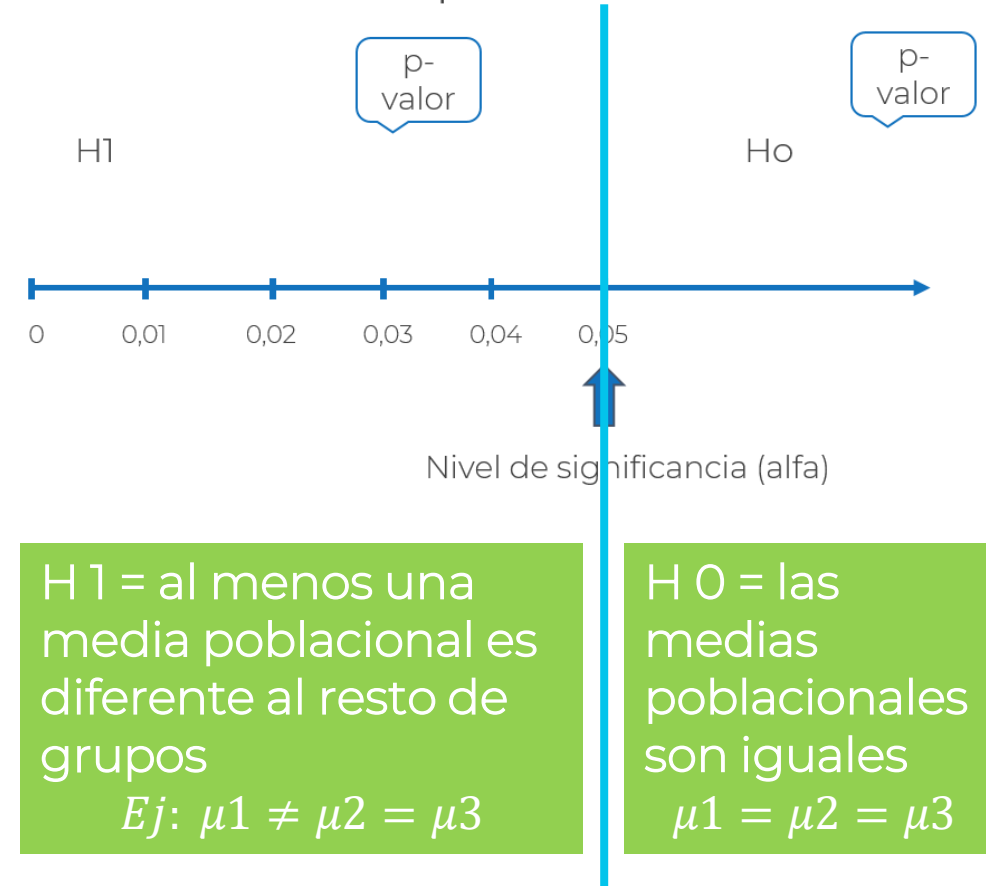
Species      Df Sum Sq Mean Sq F value Pr(>F)
Residuals    98  19.14   0.195

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

```



Regla de decisiones para la prueba de hipótesis



**OBJETIVO:**  
Comparar más de dos medias poblacionales entre el factor de estudio

## Las claves para analizar las varianzas ANOVA

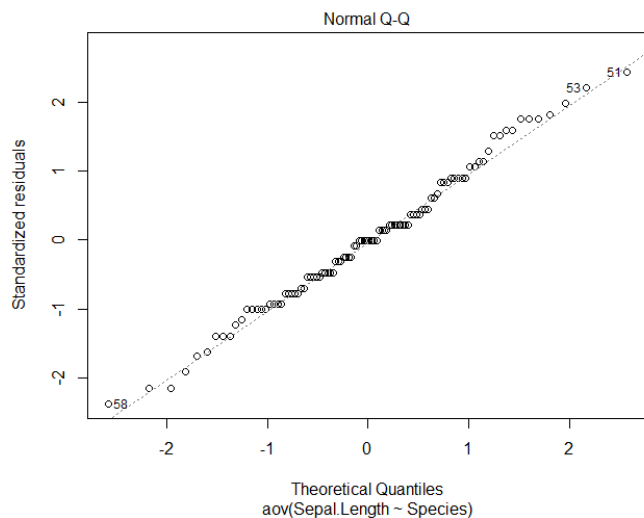
- El resultado de la tabla ANOVA

```

      Df Sum Sq Mean Sq F value Pr(>F)
Species    1   21.62   21.622   110.7 <2e-16 ***
Residuals  98   19.14    0.195
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

- Podemos utilizar la ANOVA si los residuos son normales:



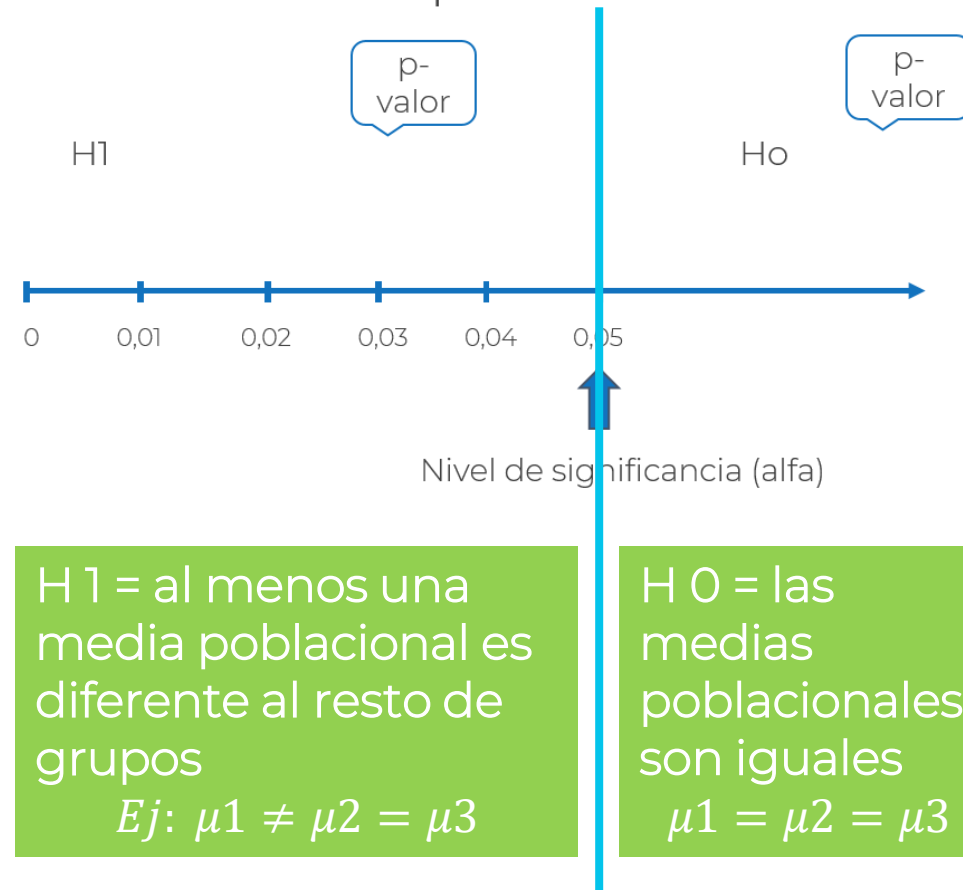
Shapiro-wilk normality test

```

data: aov_residuos
W = 0.99004, p-value = 0.6676

```

Regla de decisiones para la prueba de hipótesis





**OBJETIVO:**  
Comparar más de dos medias poblacionales entre el factor de estudio

# Las claves para analizar las varianzas ANOVA

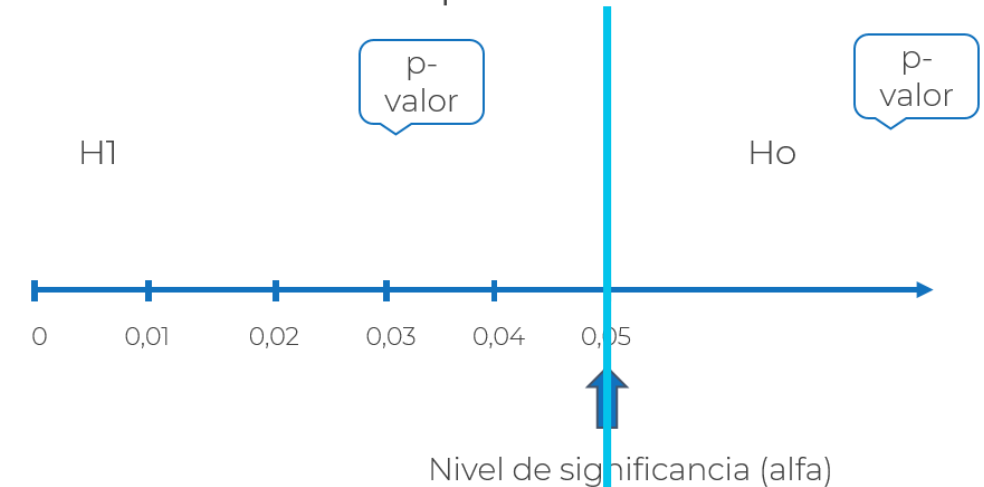
- El siguiente paso es el cálculo del p-valor por parejas para saber qué parejas son diferentes entre sí.
  - Tukey HSD Honest Significant Differences – solamente para ANOVA
  - Pair Wise comparison

Tukey multiple comparisons of means  
95% family-wise confidence level

Fit: aov(formula = Sepal.Length ~ Species, data = varEstudio)

\$Species		diff	lwr	upr	p adj
versicolor-setosa		0.930	0.6862273	1.1737727	0
virginica-setosa		1.582	1.3382273	1.8257727	0
virginica-versicolor		0.652	0.4082273	0.8957727	0

Regla de decisiones para la prueba de hipótesis



H 1 = las dos medias poblacionales son diferentes

$$\mu_1 \neq \mu_2$$

$$\mu_1 - \mu_2 \neq 0$$

H 0 = las medias poblacionales son iguales

$$\mu_1 = \mu_2$$

$$\mu_1 - \mu_2 = 0$$

**OBJETIVO:**  
Comparar más de dos  
medias poblacionales  
entre el factor de estudio

# Las claves para analizar las varianzas ANOVA

- Los dos casos:
  - Medidas repetidas** = una misma muestra medida dos veces (antes y después de un tratamiento por ejemplo)
  - Grupos independientes** = todos los demás casos, son muestras distintas

Los residuos son normales y varianzas iguales

*coggle* One Way ANOVA (P)  
made for free at coggle.it

Los residuos son normales y varianzas diferentes

One Way ANOVA Welch (P)

Kruskall Wallis

Grupos Independientes

1 MEDIDA Y MÁS  
DE 2 GRUPOS DE  
1 FACTOR

Tukey solo para ANOVA, Pair Wise para los tres casos

Tukey HSD o Pair Wise

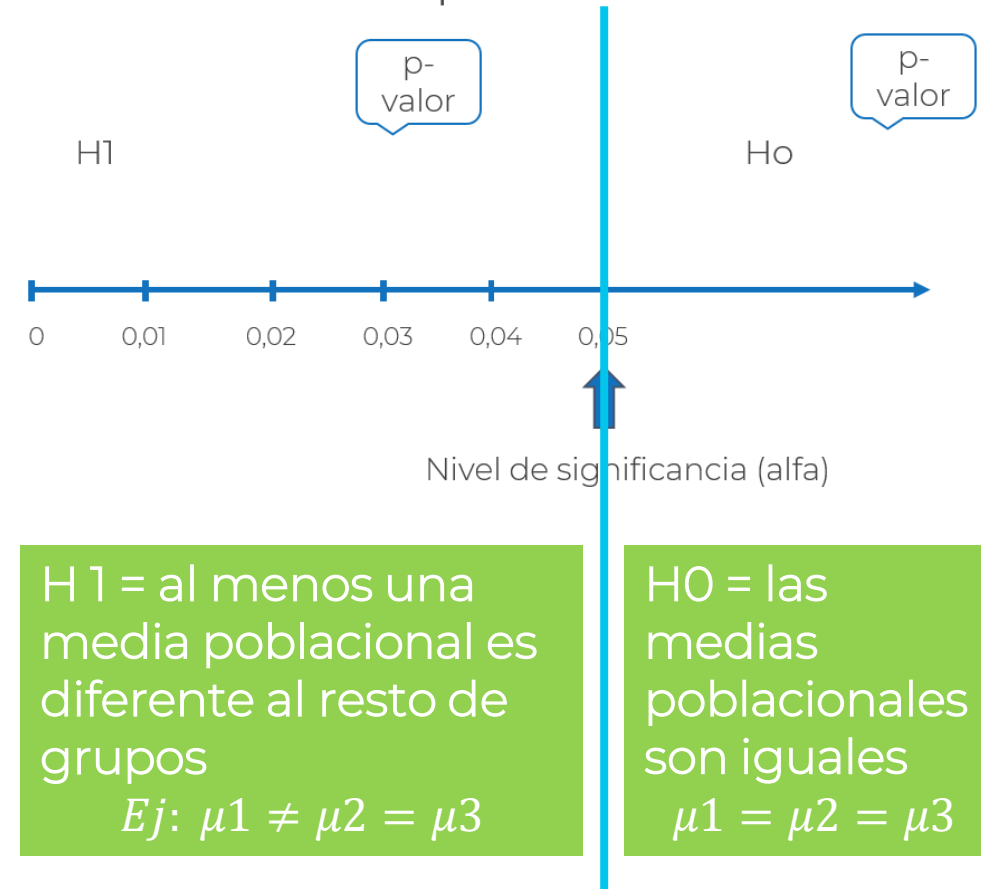
Los residuos son normales

ANOVA de medidas repetidas

Friedman Test

Medidas Repetidas

Regla de decisiones para la prueba de hipótesis



# Take away

El resumen de la lección

## Lo más importante de la lección

- La ANOVA es el ANalysis Of Variance
- El primer paso es describir:
  - Con el boxplot de 1 factor y el diagrama de medias
  - Calcular el IC
- Calcular la ANOVA y mirar:
  - La normalidad de los residuos
  - Igualdad de varianzas
- Decidir el test (paramétrico o NO)
- Calcular el p-valor por parejas

# Tú turno

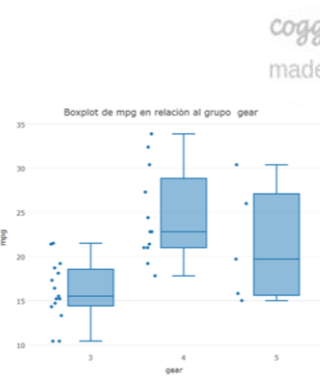
A por tus primeras tablas ANOVA

## A poner en práctica lo que has visto

- Descarga la hoja de trabajo
- Y calcular tu primera tabla ANOVA (ya empiezas a estar muy cerca de los modelos predictivos) 😊
- ¡A por ello!

# Comparar 1 medidas y 2 o más factores

Generalizando la ANOVA para comprobar el efecto de los factores en las medidas



coggle  
One Way ANOVA (P)  
made for free at coggle.it

- Way ANOVA Welch (P)
- Kruskall Wallis
- Tukey HSD o Pair Wise

Grupos Independientes

1 MEDIDA Y MÁS  
DE 2 GRUPOS DE  
1 FACTOR

ANOVA de medidas repetidas

Friedman Test

Medidas Repetidas

INTERVALOS DE  
CONFIANZA

- Two-sided ANOVA -  
caso NO Balanceado  
(P)
- Two-sided ANOVA -  
caso Balanceado  
(P)
- Scheirer Ray Hare (NP)

1 MEDIDA y 2 o  
más FACTORES

## COMPARACIÓN DE MEDIDAS

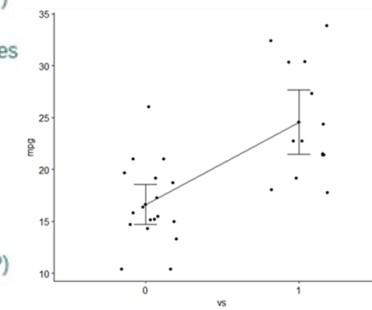
1 MEDIDA Y 2  
GRUPOS DE 1  
FACTOR

Grupos Independientes

- T-test grupos independientes (P)
- T-test con corrección de Welch (P)
- Wilcoxon de grupos Independientes

Medidas repetidas

- T-test grupos pareados (P)
- Wilcoxon para grupos pareados (NP)

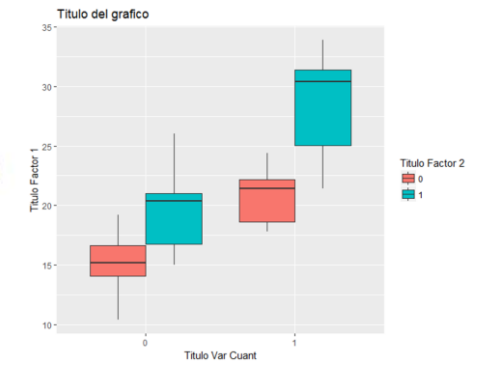


1 MEDIDA Y 1  
VALOR

- T-test de una muestra (P)
- Wilcoxon de 1 muestra (NP)

2 MEDIDAS Y 1  
FACTOR

MANOVA (P)





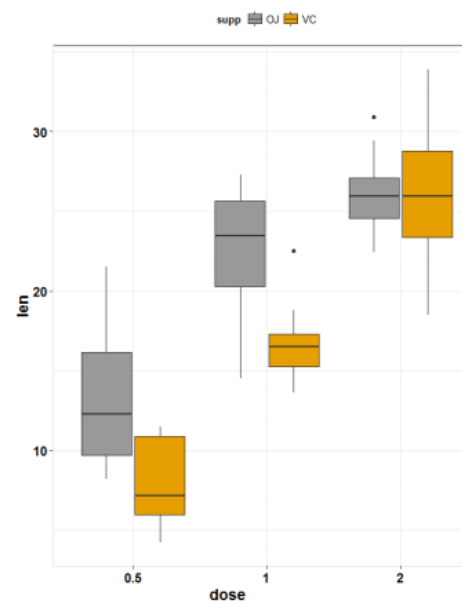
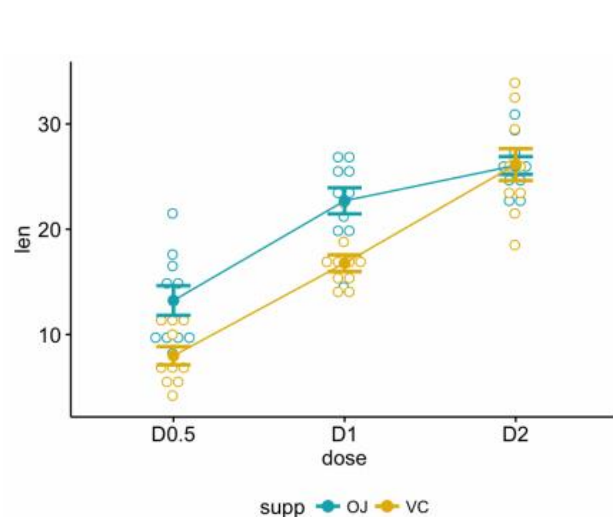
**OBJETIVO:**  
Ver si dos o más factores  
influyen a la medida

# Las claves para analizar la influencia de dos o más factores a una medida

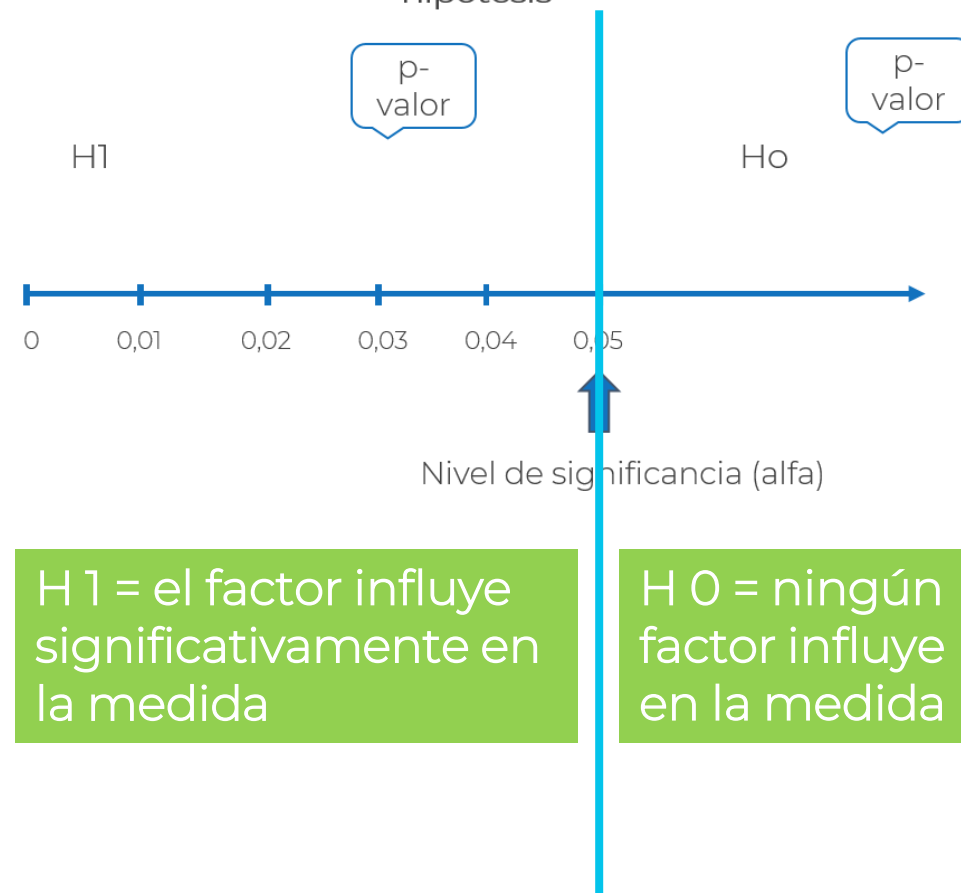
- Se utiliza para cuantificar la influencia de los factores en las medidas
- Variable respuesta: medida → Log. del diente "len"
- Variable de estudio: 2 factores →
  - Tipo de Suplemente "Supp"
  - Dosis suministrada "Dose"
- H1: algún factor influye a la medida

	len	supp	dose
38	9.4	OJ	0.5
36	10.0	OJ	0.5
37	8.2	OJ	0.5
50	27.3	OJ	1.0
59	29.4	OJ	2.0
1	4.2	VC	0.5
13	15.2	VC	1.0
56	30.9	OJ	2.0
27	26.7	VC	2.0
53	22.4	OJ	2.0

- Descripción: boxplot de 2 factores, diagrama de medias de 2 factores



Regla de decisiones para la prueba de hipótesis



$H_1$  = el factor influye significativamente en la medida

$H_0$  = ningún factor influye en la medida

**OBJETIVO:**  
Ver si dos o más factores  
influyen a la medida

# Las claves para analizar la influencia de dos o más factores a una medida

- Caso balanceado –muestras por grupos controladas
  - Tabla de contingencias de los 2 factores
- Resultado:

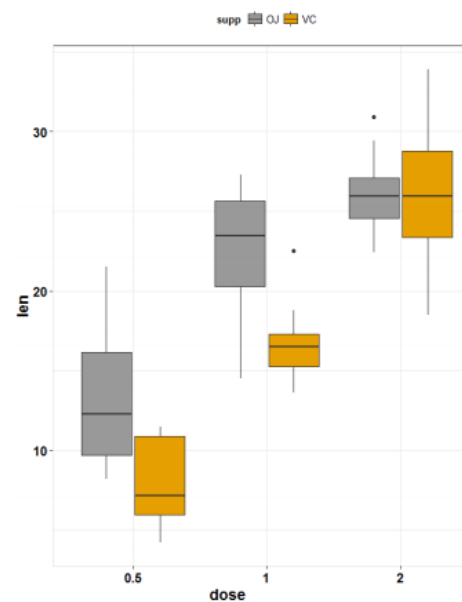
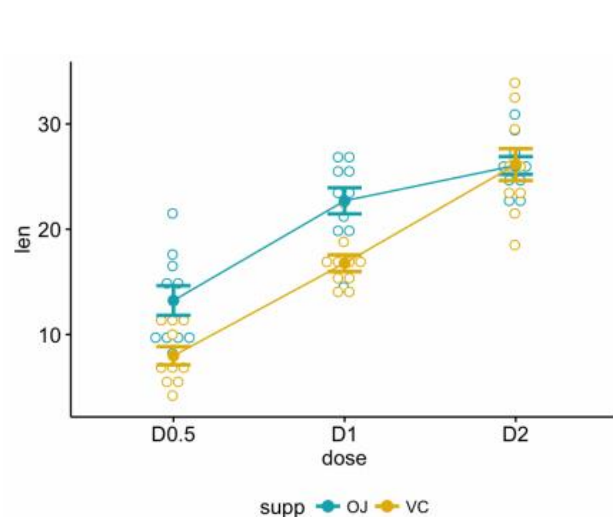
	D0.5	D1	D2
OJ	10	10	10
VC	10	10	10

```

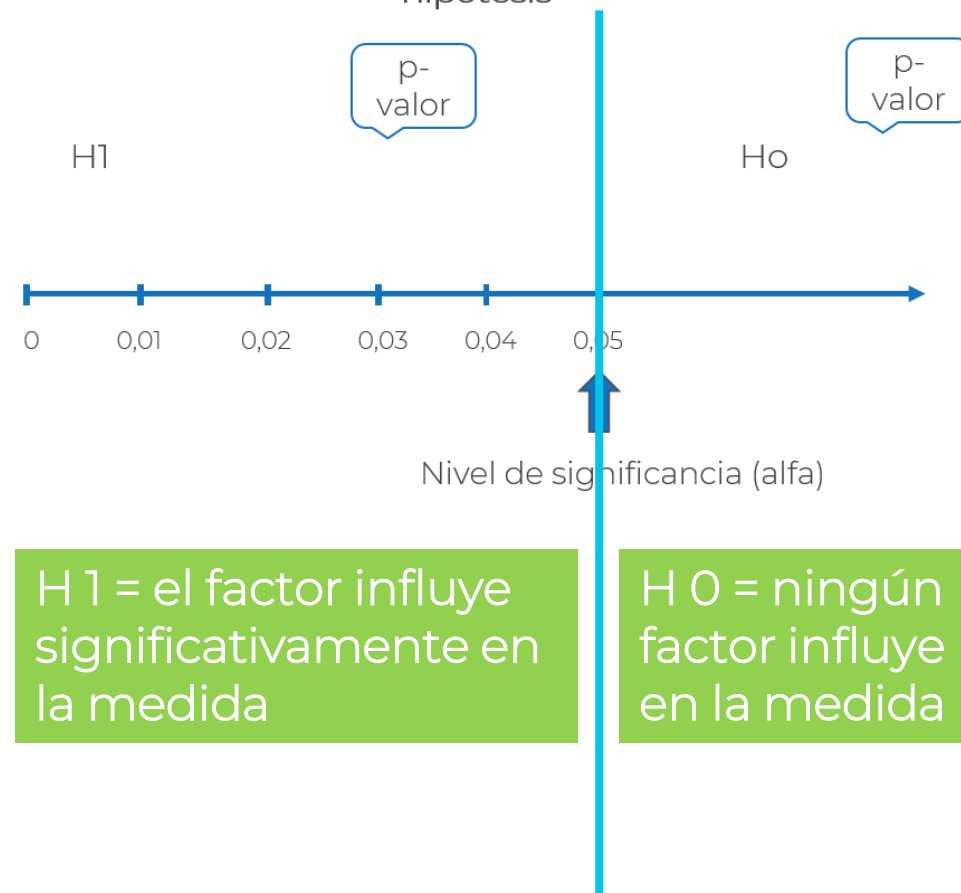
Df Sum Sq Mean Sq F value Pr(>F)
supp    1   205.4    205.4    14.02 0.000429 ***
dose    2  2426.4   1213.2    82.81 < 2e-16 ***
Residuals 56   820.4     14.7

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

```



Regla de decisiones para la prueba de hipótesis



H 1 = el factor influye  
significativamente en  
la medida

H 0 = ningún  
factor influye  
en la medida

OBJETIVO:  
Ver si dos o más factores  
influyen a la medida

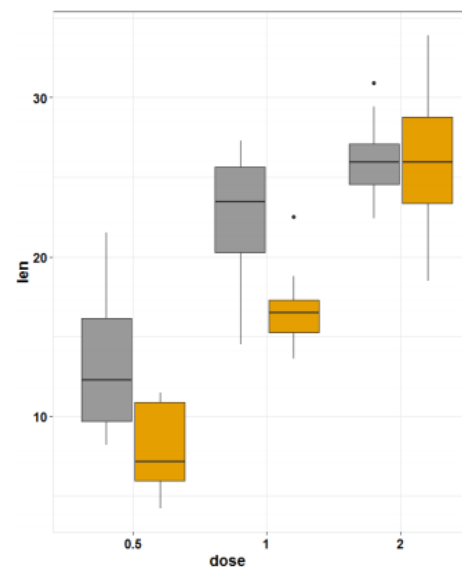
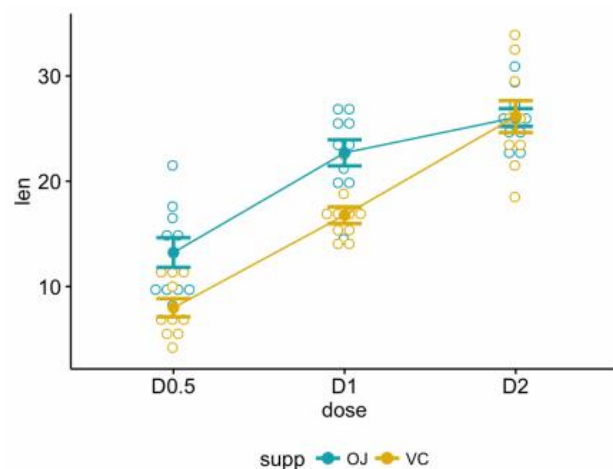
# Las claves para analizar la influencia de dos o más factores a una medida

- Caso balanceado –muestras por grupos controladas
  - Tabla de contingencias de los 2 factores
- Resultado:

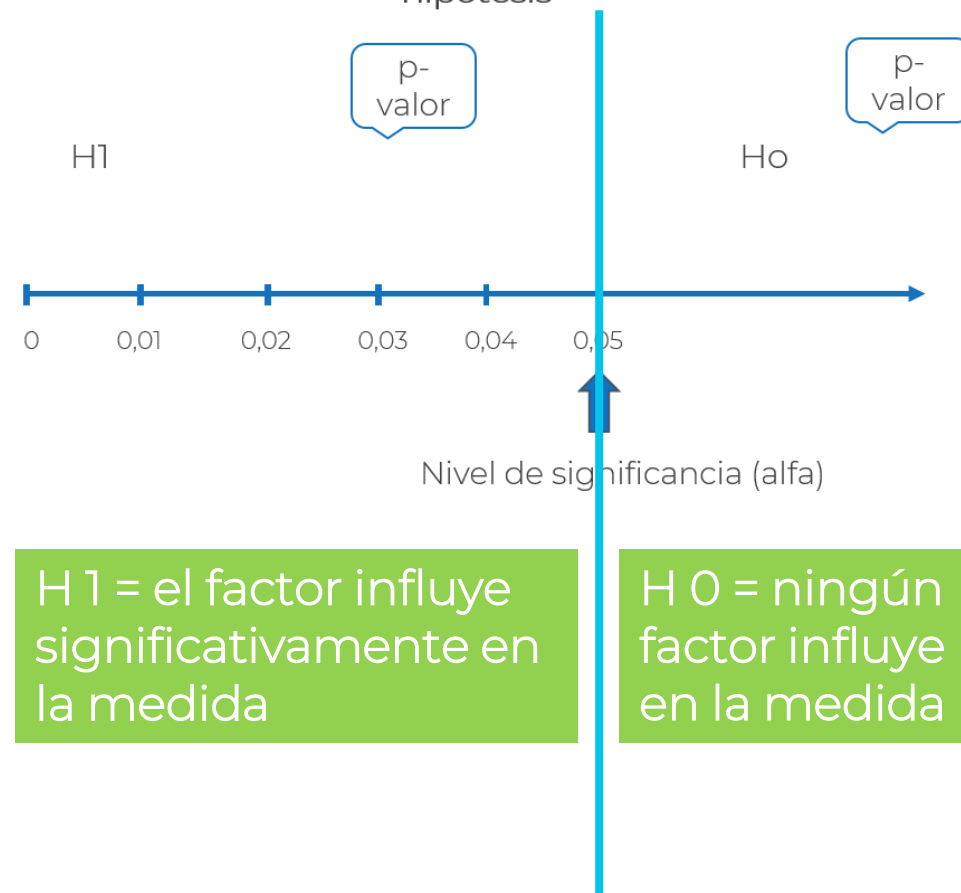
	D0.5	D1	D2
OJ	10	10	10
VC	10	10	10

```

Df Sum Sq Mean Sq F value Pr(>F)
supp      1   205.4      205.4   15.572 0.000231 ***
dose      2  2426.4     1213.2   92.000 < 2e-16 ***
supp:dose  2   108.3       54.2    4.107 0.021860 *
Residuals 54   712.1       13.2
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
  
```



Regla de decisiones para la prueba de hipótesis



H 1 = el factor influye  
significativamente en  
la medida

H 0 = ningún  
factor influye  
en la medida

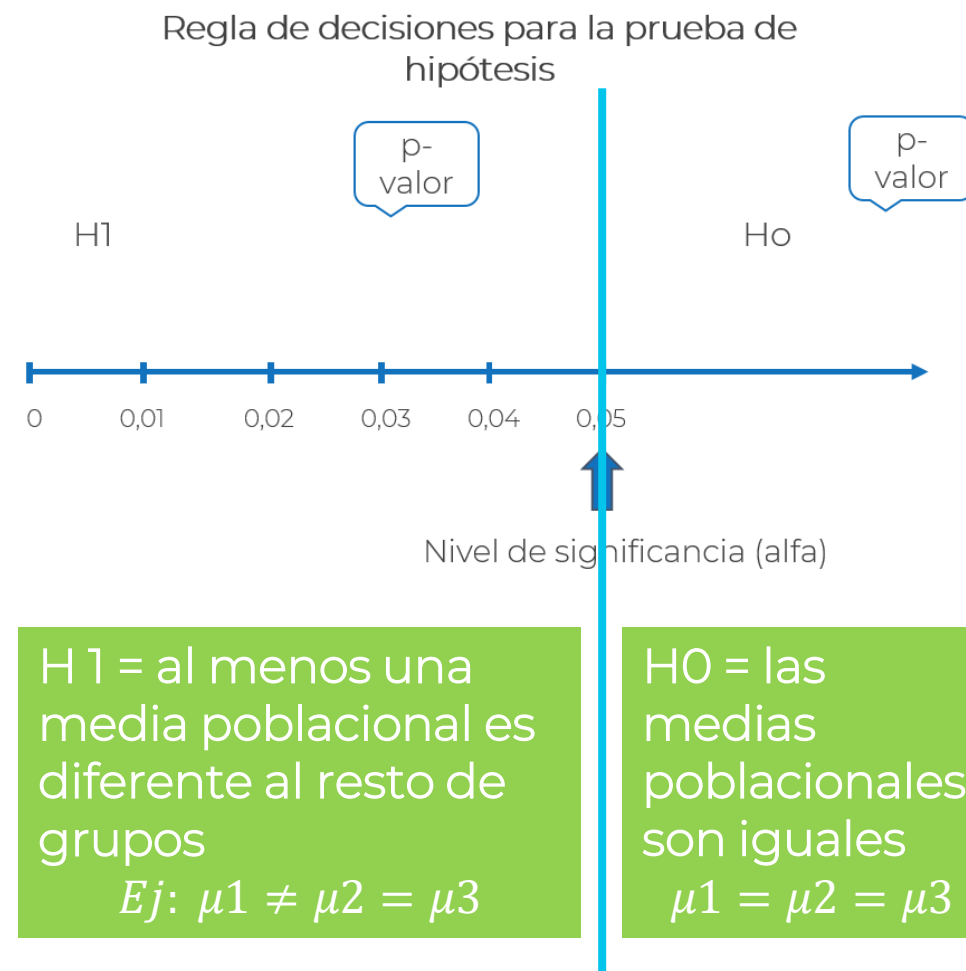
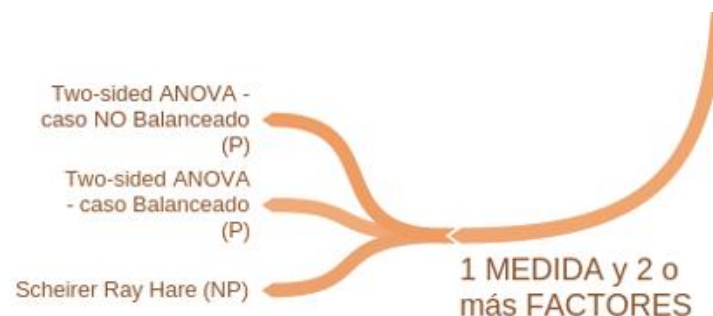
# Las claves para analizar la influencia de dos o más factores a una medida

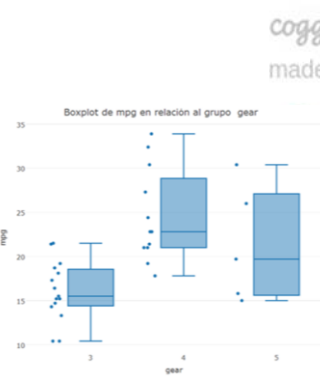
**OBJETIVO:**  
Comparar más de dos medias poblacionales entre el factor de estudio

- Los dos casos:
  - Balanceado** = muestras iguales por factores
  - NO Balanceado** = muestras diferentes por factores

Los residuos son normales

Los residuos son normales





coggle  
One Way ANOVA (P)  
made for free at coggle.it

- Way ANOVA Welch (P)
- Kruskall Wallis
- Tukey HSD o Pair Wise

Grupos Independientes

1 MEDIDA Y MÁS  
DE 2 GRUPOS DE  
1 FACTOR

ANOVA de medidas repetidas

Friedman Test

Medidas Repetidas

INTERVALOS DE  
CONFIANZA

# COMPARACIÓN DE MEDIDAS

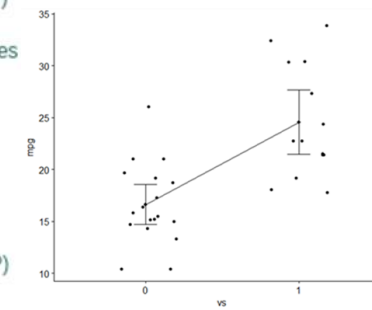
1 MEDIDA Y 2  
GRUPOS DE 1  
FACTOR

Grupos Independientes

- T-test grupos independientes (P)
- T-test con corrección de Welch (P)
- Wilcoxon de grupos Independientes

Medidas repetidas

- T-test grupos pareados (P)
- Wilcoxon para grupos pareados (NP)



1 MEDIDA Y 1  
VALOR

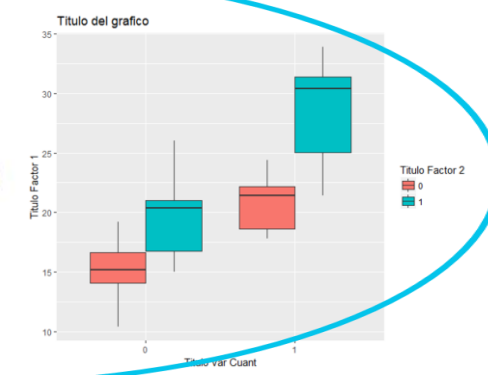
- T-test de una muestra (P)
- Wilcoxon de 1 muestra (NP)

2 MEDIDAS Y 1  
FACTOR

MANOVA (P)

- Two-sided ANOVA - caso NO Balanceado (P)
- Two-sided ANOVA - caso Balanceado (P)
- Scheirer Ray Hare (NP)

1 MEDIDA y 2 o  
más FACTORES



# Take away

El resumen de la lección

## Lo más importante de la lección

- La ANOVA generalizamos a dos factores o más – estamos cerca de los modelos predictivos
- Podemos aplicarlo en experimentos para ver la influencia de factores en las muestras. Incluso podemos añadir la interacción

# Tú turno

Ya puedes aplicar las ANOVA de dos factores



## A poner en práctica lo que has visto

- Descarga la hoja de trabajo
- Y calcular tu primera tabla ANOVA de dos factores (ya empiezas a estar muy cerca de los modelos predictivos) 😊
- ¡A por ello!