



## Estadística con Excel

Héctor Manuel Garduño Castañeda

Noviembre, 2021



# Contenido

ANOVA de una vía



## Uso del ANOVA de una vía

Se trata del Análisis de la varianza (de ahí su nombre). En particular, para ANOVA de una vía tenemos:

- ▶ Se utiliza para comparar la varianza entre diferentes muestras elegidas de una misma población.
- ▶ La idea es comparar dos o más muestras basados en la diferencia de las varianzas y se trata de una prueba de hipótesis sobre las medias poblacionales.
- ▶ Puede ser usado cuando tenemos al menos dos variables, donde una es categórica y la otra es continua. A la categórica la llamamos *variable independiente* y a la continua la llamamos *variable de respuesta*.



## Condiciones para ser aplicable

Supongamos que tenemos la variable independiente  $X$  y la variable de respuesta  $Y$ . Para poder aplicar correctamente el ANOVA de una vía se debe cumplir:

- ▶ Los grupos son independientes.
- ▶  $Y$  debe ser aproximadamente normal en cada grupo (siendo menos estricta esta condición cuanto mayor sea el tamaño de cada grupo).
- ▶ Todos los grupos tienen la misma varianza (esta condición es más importante cuanto menor es el tamaño de los grupos).
- ▶ No tener datos atípicos.



# Hipótesis

Supongamos que la muestra se divide, según la variable independiente, en los grupos  $G_1, G_2, \dots, G_k$ . Sea  $\mu_m$  la media poblacional del grupo  $m$ . Entonces la prueba de hipótesis se establece como:

$$\begin{cases} H_0 : & \mu_1 = \mu_2 = \dots = \mu_k \\ H_a : & \text{hay al menos dos medias diferentes} \end{cases}$$

Por lo tanto, en caso de rechazar  $H_0$ , se tiene que proceder a lo que se conoce como *pruebas post hoc*, siendo la prueba por parejas la más conocida. Ver <https://github.com/scidatmath2020/Inferencia-Estadistica/blob/master/C08.%20ANOVA.ipynb>



## Proceso de investigación

Los pasos para la aplicación del ANOVA de una vía son:

1. **Seleccionar el nivel de significación.** Este es denotado por  $\alpha$ . Generalmente  $\alpha = 0.01$ ,  $\alpha = 0.05$  o  $\alpha = 0.10$ .
2. **Encontrar el valor crítico.** Este es denotado por  $f$ . Antiguamente se usaba una tabla de  $F$ . Se seleccionaba la columna basado en  $\alpha$  y la fila basado en los **grados de libertad**.
3. **Calcular un parámetro.** Se refiere a calcular el número  $F$  con los datos de la muestra como la media de la suma de los cuadrados.
4. **Comparar y decidir.** Si  $F < f$ , aceptamos  $H_0$ . En caso contrario, rechazamos la hipótesis nula.



## Proceso en Excel

