

# Herramientas de Estadística (No)Paramétrica

## (Parte 3)

Juan C. Correa

Material de uso exclusivo para  
INGENIO PANTALEON, S.A.  
Diagonal 6, 10-31, Zona 10

Ciudad de Guatemala



- 1 Comparación entre dos grupos
- 2 Consideraciones Conceptuales
- 3 Ejemplos



# Comparación entre dos grupos

Una manera de establecer la relación entre variables es a través de la comparación de grupos.

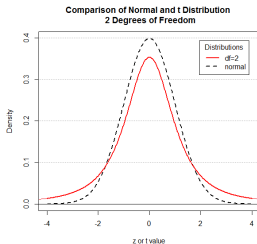
Hay varias técnicas que nos permiten comparar grupos, entre ellas están:

- t de Student: Es especialmente útil cuando se quiere saber la asociación entre una variable nominal/ordinal dicotómica y una variable intervalo.
- U de Mann-Whitney: Es la alternativa no paramétrica a la t de Student, para conocer la relación entre una variable nominal/ordinal y una variable ordinal.
- Prueba de rangos con signo de Wilcoxon: Es una variante no-paramétrica que se aplica para conocer la asociación entre una variable nominal/ordinal dicotómica y una variable ordinal.



# Consideraciones Conceptuales

La asociación entre variables, al comparar dos o más grupos, se mide a través de un estadístico que nos va a indicar cuánto se asocian esas dos variables.



Por lo general, siempre se va a asumir que el valor de ese estadístico será igual a cero lo cual indica que las dos variables son independientes entre sí, o no guardan absolutamente ninguna relación o asociación.



Los estadísticos más empleados para evaluar la asociación entre dos variables tienen propiedades de distribuciones bien conocidas:

- Distribución t de Student
- Distribución F de Snedecor-Fisher
- Distribución de  $\chi^2$

Para juzgar si hay o no una asociación estadística entre variables, hay que evaluar el resultado numérico del estadístico y compararlo con sus correspondientes **grados de libertad**.



# Consideraciones Conceptuales

Los **grados de libertad** representan el número de formas posibles en las que puede cambiar una variable aleatoria según las observaciones empíricas que se han hecho de ella. Usualmente se calcula como  $n - r$ , donde  $n$  es el número de observaciones que pueden tomar un valor y  $r$  es el número de sujetos cuyo valor dependerá del que tomen los miembros de la muestra que son libres. También pueden ser representados por  $k - r$ , donde  $k$  número de grupos cuando se realizan operaciones con grupos y no con sujetos individuales.

**Ejemplo:** Si al observar una muestra de sujetos a los que se le pregunta el “Nivel Educativo” y el primer sujeto afirma tener educación universitaria completa, entonces se podría inferir que los grados de libertad serían 5, asumiendo que estos son todos los valores posibles: educación primaria incompleta, educación primaria completa, educación secundaria incompleta, educación secundaria completa, educación universitaria incompleta, y educación universitara completa.



# Consideraciones Conceptuales

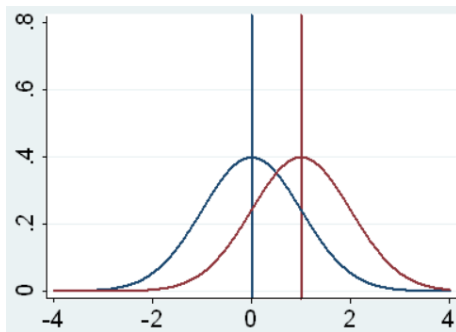
En general, se asumirá que dos variables están relacionadas si:

- El estadístico es muy distinto de cero
- Si el valor del estadístico, en función de sus grados de libertad y distribución subyacente, refleja un valor de probabilidad lo suficientemente pequeño que sobrepasa nuestra tolerancia de admitir que el azar ha intervenido.
- Use recursos de visualización de datos para complementar la interpretación del cálculo.
- Si se hace uso de estadística Bayesiana, la interpretación se hace en función del valor que obtenga el factor de Bayes.

([https://es.qaz.wiki/wiki/Bayes\\_factor](https://es.qaz.wiki/wiki/Bayes_factor))



# Consideraciones Conceptuales



Cuánto más distanciadas están las distribuciones de los grupos, tanto más asociación o relación existe entre la variable que define dichos grupos (representada por las líneas rojas y azul) y la variable de intervalo que se representa en el eje X.





# Comparación entre dos grupos

Veamos algunos ejemplos de comparación entre dos grupos en el siguiente enlace.

<https://github.com/jcorrean/Pantaleon/blob/main/Comparaci%C3%B3nDeGrupos.Rmd>

