



PI DATA STRATEGY & CONSULTING

“I don’t know what a p-value is... and now I’m too afraid to ask”

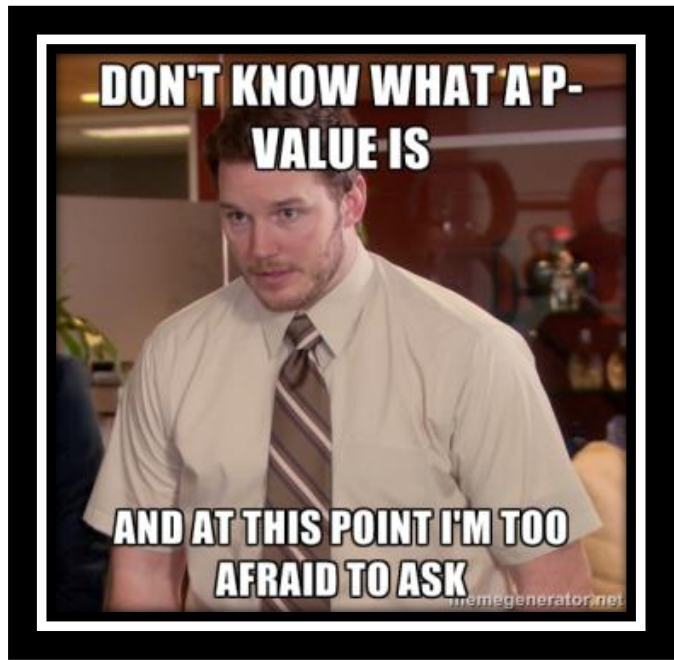


Nindirí Armenta

Data Scientist, Pi Data Strategy & Consulting



**“I don’t know what a p-value is... and now
I’m too afraid to ask”**



Estadística para Data Science

2º encuentro

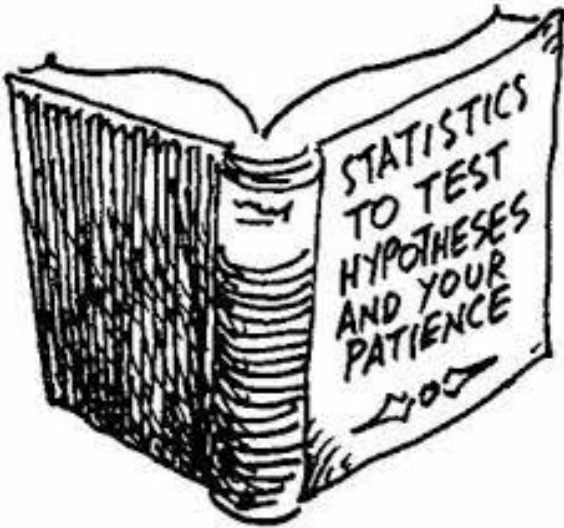


Agenda

- Introducción
 - ¿Qué es una hipótesis?
 - ¿En qué consiste un test estadístico?
- Métodos para probar una hipótesis
- Tipos de test estadísticos
- Test estadísticos de parámetros cuantitativos
- Pasos para realizar un test estadístico
- Selección de un test estadístico
- Errores Tipo I y Tipo II
- Ejemplos con Python en Jupyter Notebook



Introducción: ¿Qué es una hipótesis?



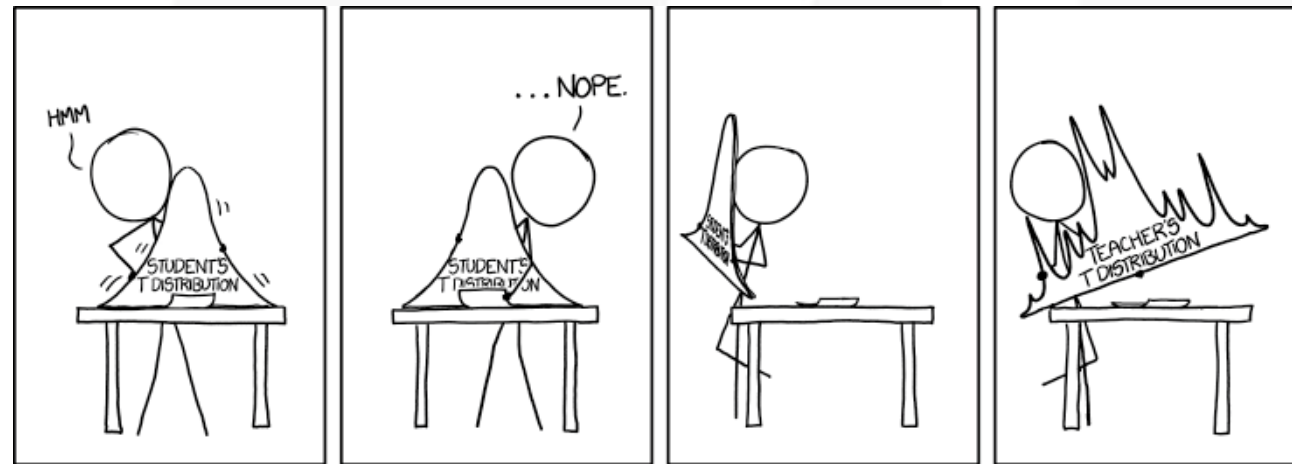
Definición:

Suposición hecha a partir de unos datos que sirve de base para iniciar una investigación o una argumentación.



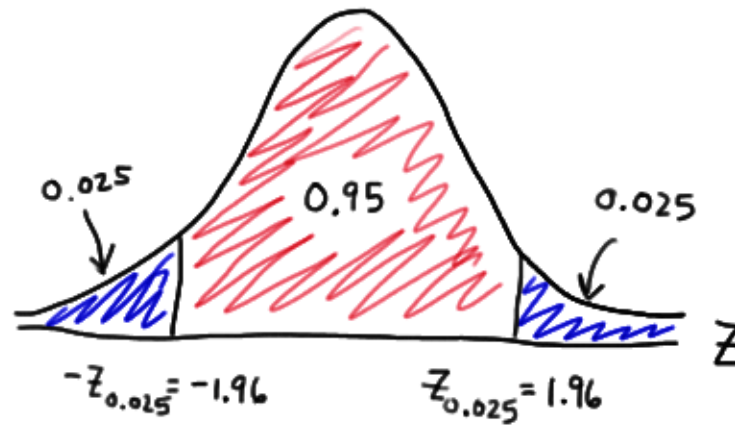
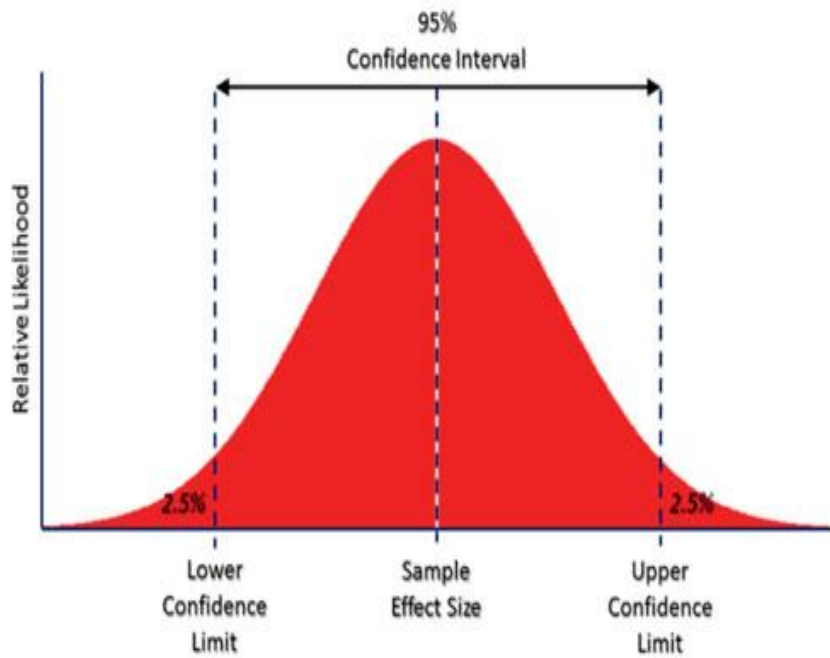
Introducción: ¿En qué consiste un test estadístico?

Es un procedimiento estadístico que permite aceptar o rechazar una afirmación hecha con respecto a un fenómeno o suceso.



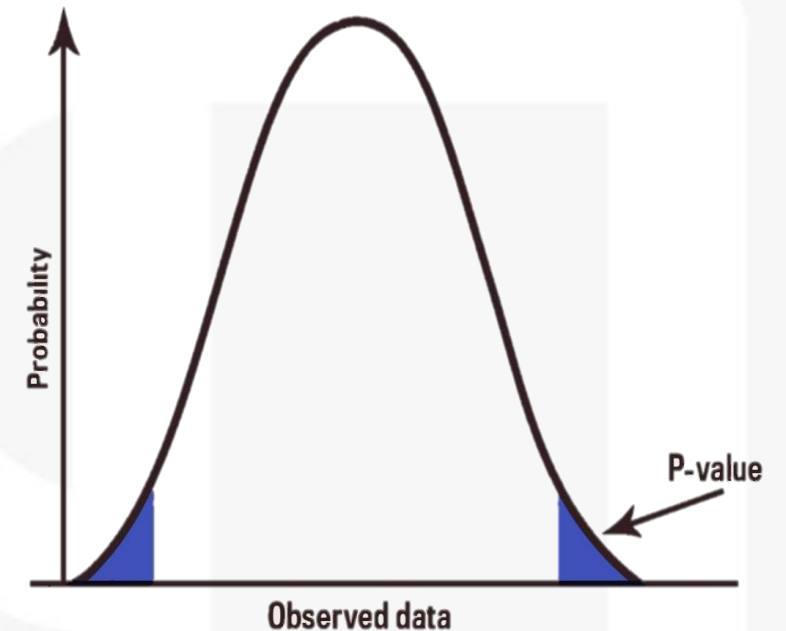
Métodos para probar una hipótesis

Intervalo de confianza



Estadístico de Prueba

P-valor



Tipos de test estadísticos



Paramétricos

- Media
- Varianza
- Proporción
- Comparación de medias
- Comparación de varianzas
- Comparación de proporciones

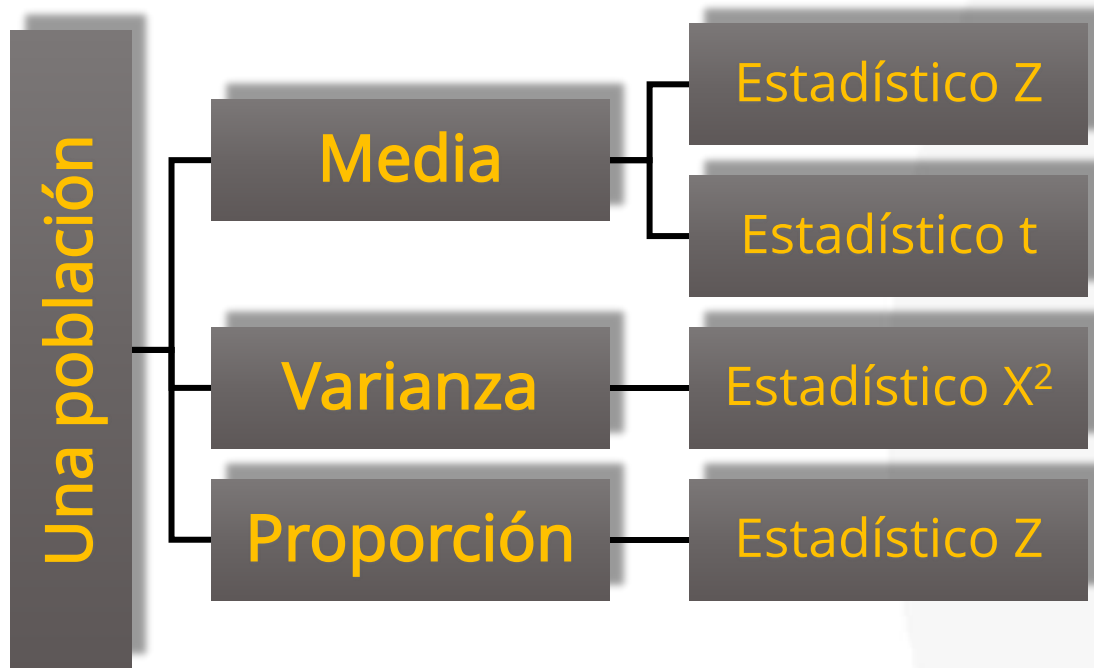
No Paramétricos

- Comparación de medias
- Comparación de varianzas
- Bondad de ajuste

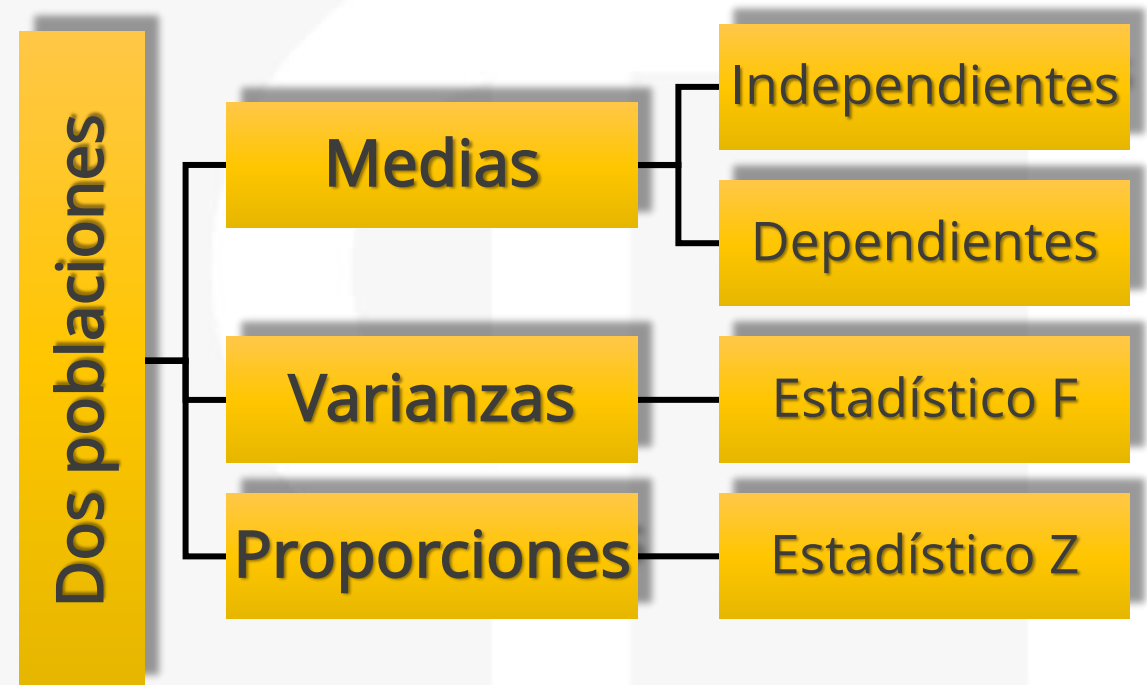


Test estadísticos de parámetros cuantitativos

Comparación con un valor determinado



Comparación entre valores poblacionales



Pasos para realizar un test estadístico

Planteamiento de hipótesis:

Ho: Hipótesis nula - H1: Hipótesis alternativa

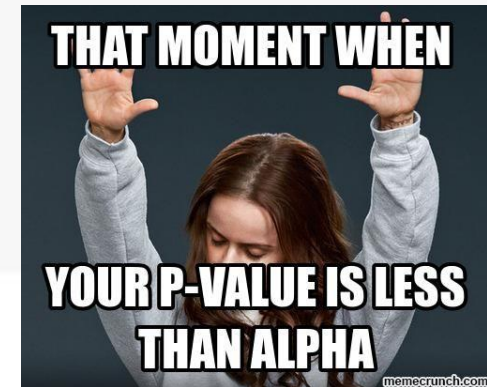
Elegir un nivel de significancia (α : alpha)

Selección del test estadístico

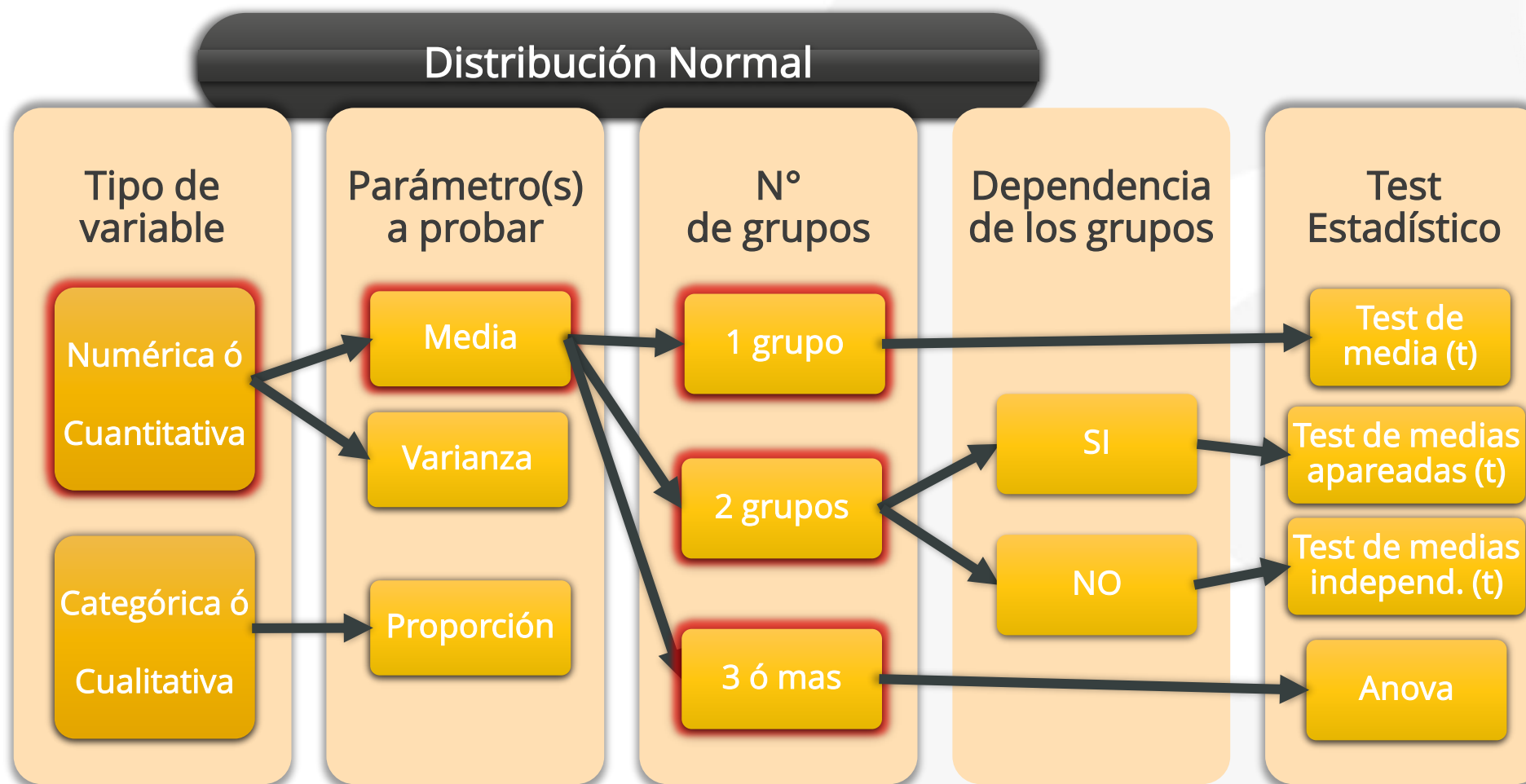
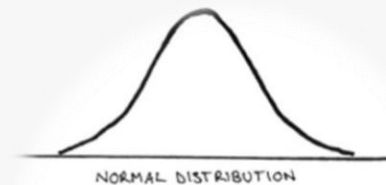
(selección del estadístico de prueba)

Comparar el p-valor con α

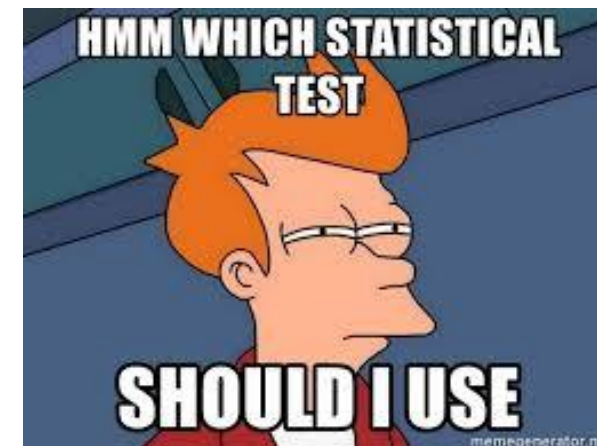
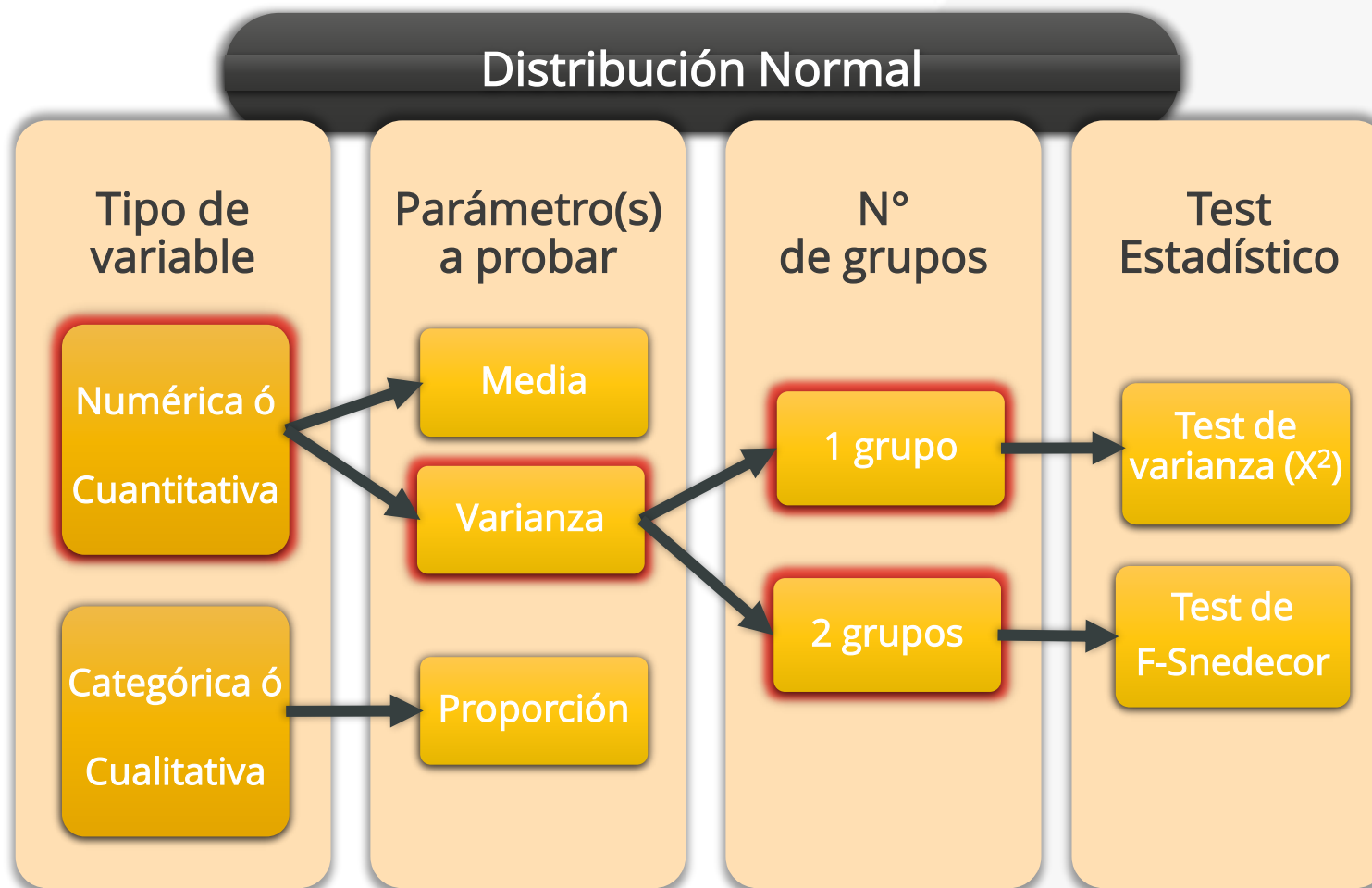
Realizar la conclusión correspondiente



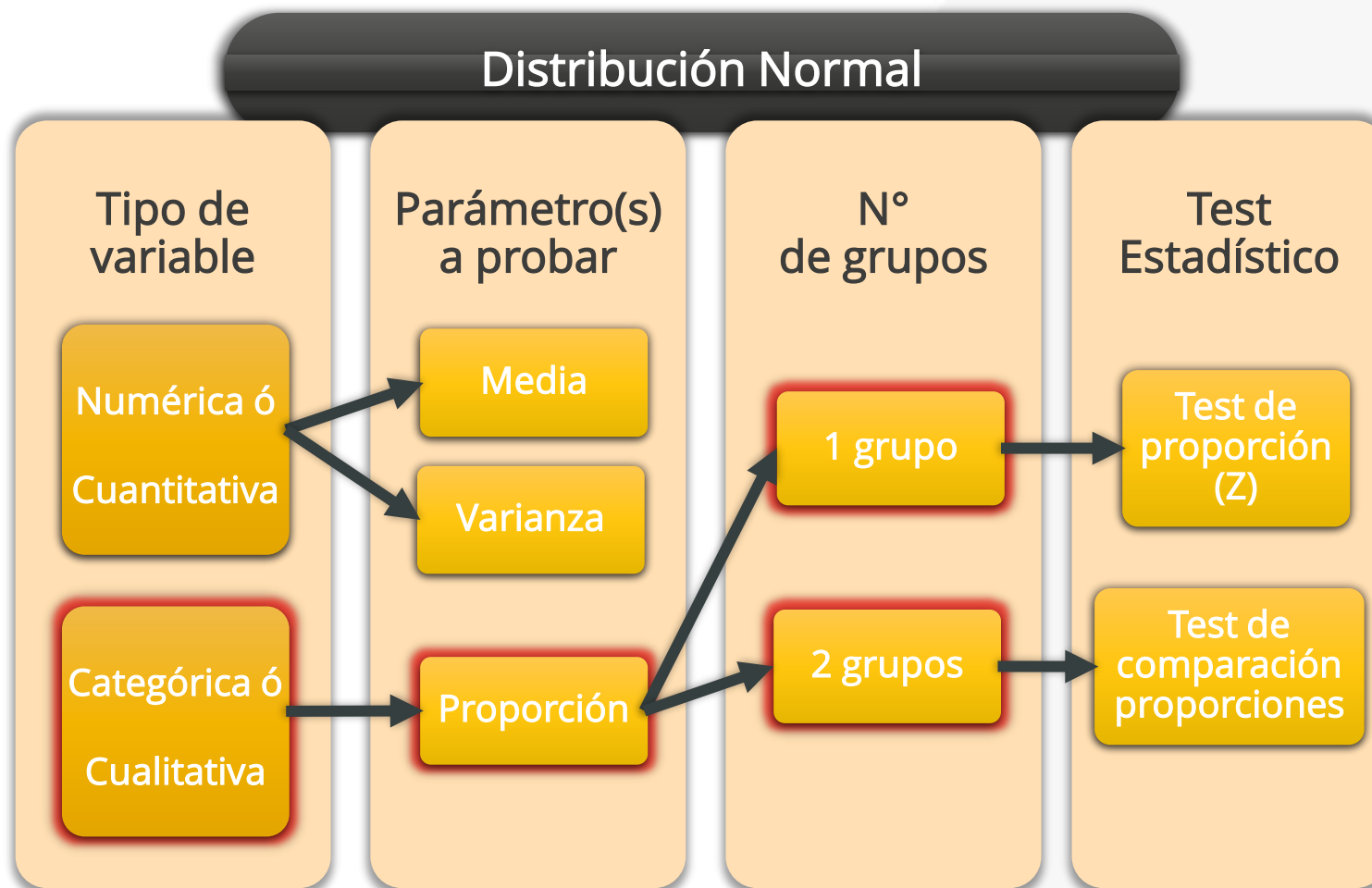
Selección del test estadístico



Selección del test estadístico



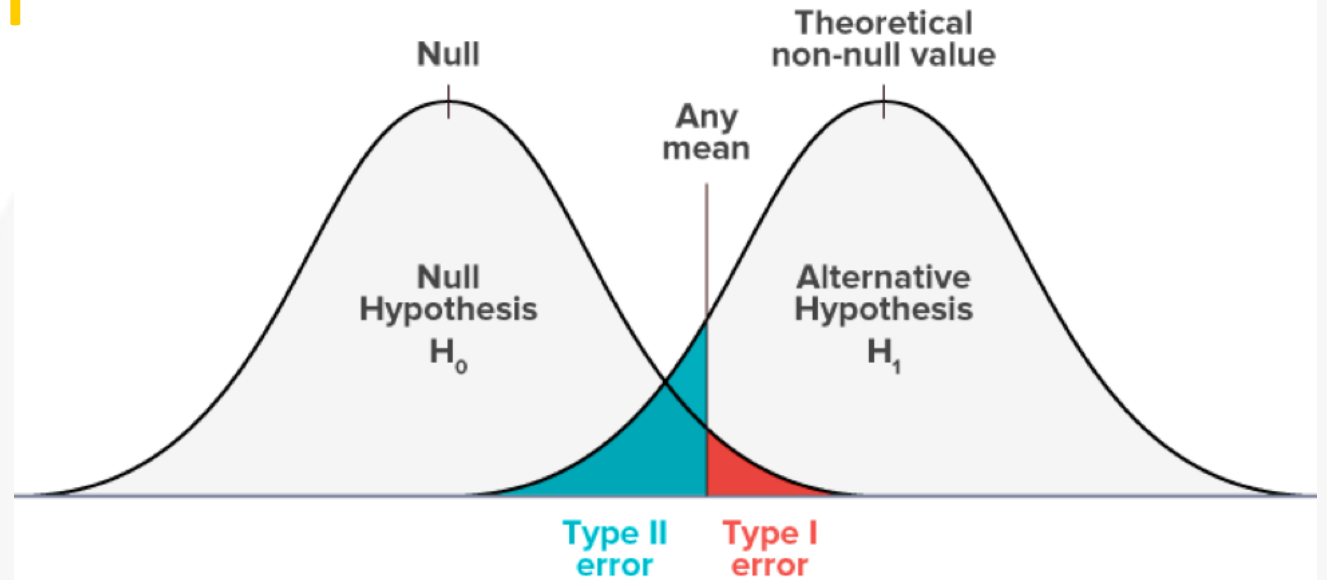
Selección del test estadístico



Errores Tipo I y Tipo II

| | | The Person is | |
|----------------|----------|---------------|--------------|
| | | Innocent | Guilty |
| The Judge Says | Innocent | No Error | Type 2 error |
| | Guilty | Type 1 error | No Error |

<https://www.analyticsvidhya.com/blog/2020/06/statistics-analytics-hypothesis-testing-z-test-t-test/>



| | H0 Verdadera | H1 Falsa |
|----------------|---------------------------------|--------------------------------|
| No rechazar H0 | Acierto $1 - \alpha$ | Error tipo 2 β |
| Rechazar H0 | Error tipo 1 α | Acierto $1 - \beta$ |



■ Algunas consideraciones

- Un valor bajo del p-value no implica que una variable no sea más o menos importante.
- Para una comunicación sólida, es necesario informar no solo el valor p, sino también el tamaño de la muestra con el que se trabajó.
- Los tests estadísticos son una herramienta útil que proporciona un estándar objetivo que todos pueden evaluar.



Ejemplos con Python en Jupyter Notebook

