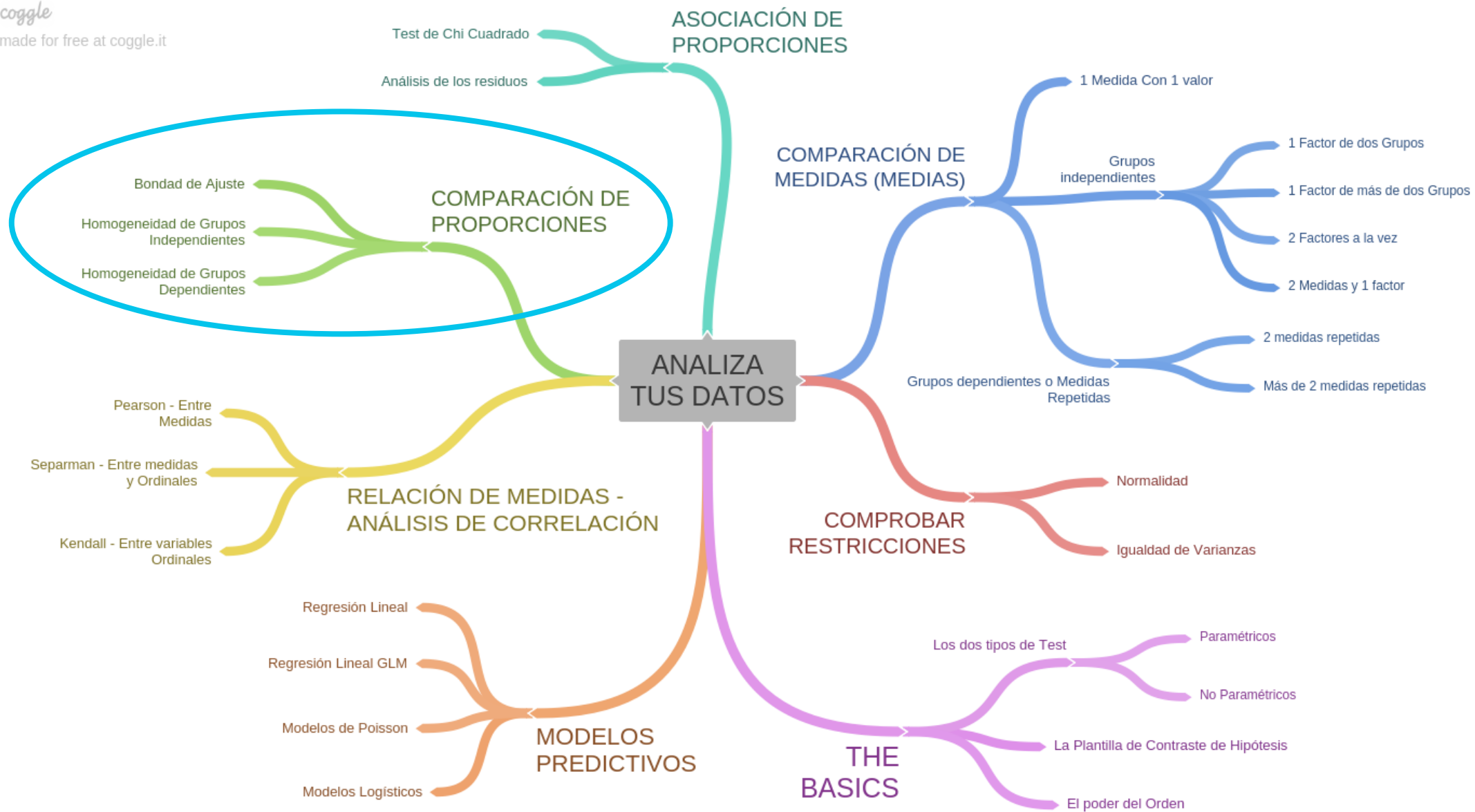


# Comparación de proporciones

La comparación de proporciones paso a paso

## ¿Que vas a ver en este bloque?

- Qué quiere decir comparar proporciones
- Los valores observados y los valores esperados
- Los casos de la comparación de proporciones en la práctica
  - Bondad de Ajuste
  - Homogeneidad
    - Grupos independientes
    - Medidas repetidas



# Qué quiere decir comparar proporciones

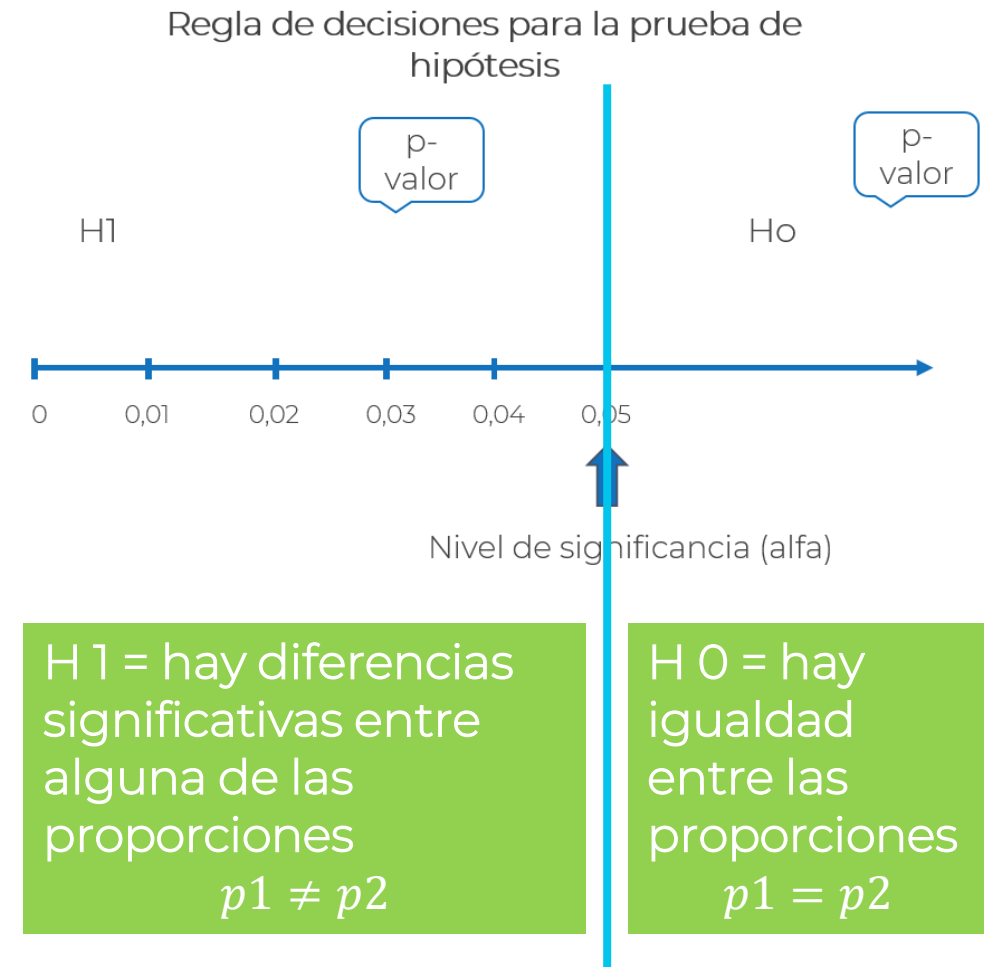
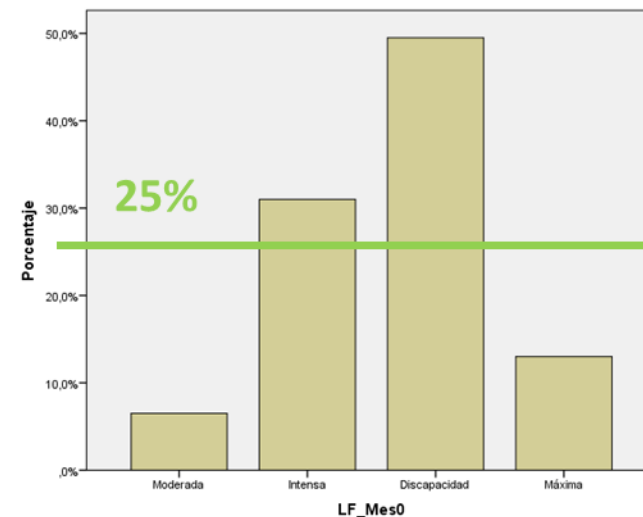
El significado y el uso de la comparación de proporciones

OBJETIVO:  
Comparar las proporciones estadísticamente entre grupos

# El significado de la comparación de proporciones

- Comparar las proporciones estadísticamente entre grupos
- Variable Dependiente o Respuesta: LF\_mes0 (factor)
- Variable Independientes:
  - En este caso no aplica pero podríamos poder otro factor (ahora lo veremos)

LF_Mes0				
		Frecuencia	Porcentaje	Porcentaje válido
Válido	Moderada	13	6,5	6,5
	Intensa	62	31,0	31,0
	Discapacidad	99	49,5	49,5
	Máxima	26	13,0	13,0
	Total	200	100,0	100,0



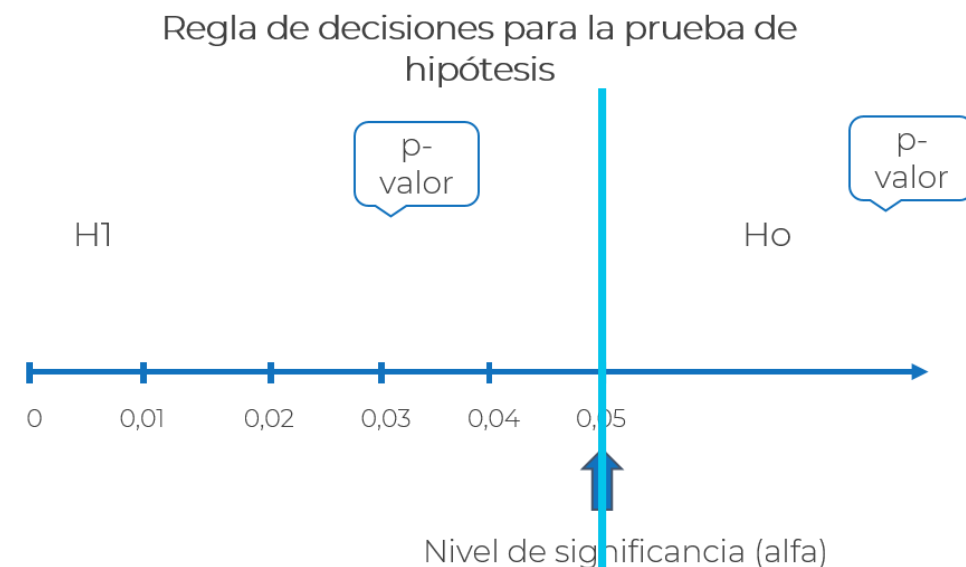
# El significado de la comparación de proporciones

- Comparar las proporciones estadísticamente entre grupos
- **Variable Dependiente o Respuesta:** Estrés Laboral (factor)
- **Variable Independientes o de estudio:** área académica (factor)

Tabla cruzada Estrés laboral / Área académica

			Área académica			Total
			Ciencias de la Salud	Ciencias Sociales	Ingenierías	
Estrés laboral	Si	Recuento	39	11	12	62
		% dentro de Área académica	88,6%	40,7%	41,4%	62,0%
	No	Recuento	5	16	17	38
		% dentro de Área académica	11,4%	59,3%	58,6%	38,0%
Total		Recuento	44	27	29	100
		% dentro de Área académica	100,0%	100,0%	100,0%	100,0%

**OBJETIVO:**  
Comparar las proporciones estadísticamente entre grupos



H 1 = hay diferencias significativas entre alguna de las proporciones

$$p_{SI} \neq p_{NO}$$

Para alguno de los grupos de área académica

H 0 = hay igualdad entre las proporciones

$$p_{SI} = p_{NO}$$

Para todos los grupos de área académica

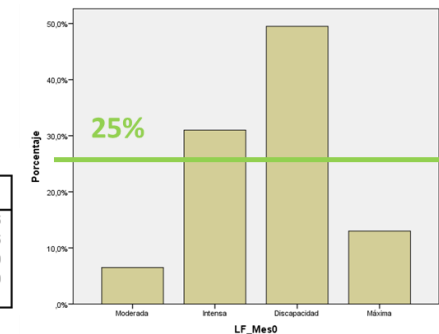


**Tabla cruzada Estrés laboral / Área académica**

			Área académica			Total
			Ciencias de la Salud	Ciencias Sociales	Ingenierías	
Estrés laboral	Sí	Recuento	39	11	12	62
		% dentro de Área académica	88,6%	40,7%	41,4%	62,0%
	No	Recuento	5	16	17	38
		% dentro de Área académica	11,4%	59,3%	58,6%	38,0%
Total		Recuento	44	27	29	100
		% dentro de Área académica	100,0%	100,0%	100,0%	100,0%

**LF\_Mes0**

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Moderada	13	6,5	6,5	6,5
	Intensa	62	31,0	31,0	37,5
	Discapacidad	99	49,5	49,5	87,0
	Máxima	26	13,0	13,0	100,0
Total		200	100,0	100,0	



# Valores esperados y valores observados

La clave en la comparación de proporciones



## Los detalles de la comparación de proporciones

- El test Chi Cuadrado se utiliza para comparar proporciones y para la prueba de bondad de ajuste
- Para calcular el p-valor se utiliza una función de densidad de probabilidad denominada Chi Cuadrado
- Para calcularla se utilizan los valores esperados y observados

La frecuencia es el valor que hemos observado

El valor esperado es la proporción que quieres comparar en la prueba de bondad de ajuste. (En este caso hemos supuesto que los grupos tengan la misma proporción, 200 pacientes en 4 grupos = 50)

Nota: El valor de Chi Cuadrado se utiliza para calcular el P-valor

Cuando calculamos el t-student utilizamos el estadístico t para obtener el p-valor

Grupos	Frecuencia	Esperado	Diferencia	Diferencia ^2
Moderada	13	50	-37	1369
Intensa	62	50	12	144
Discapacidad	99	50	49	2401
Máxima	26	50	-24	576
Total	200		Chi Cuadrado :	89.8

El valor de Chi Cuadrado es la suma de las diferencias al cuadrado entre el valor observado y el valor esperado ponderado por el valor esperado

$$X^2 = \sum \frac{(o - e)^2}{e}$$

## Los detalles de la comparación de proporciones

Otro Ejemplo  
"tonto" para ilustrar  
el estadístico Chi  
Cuadrado

Lanzamos la moneda  
20 veces y  
comparamos la  
proporción de caras y  
cruz (grupos)  
observados y  
esperados

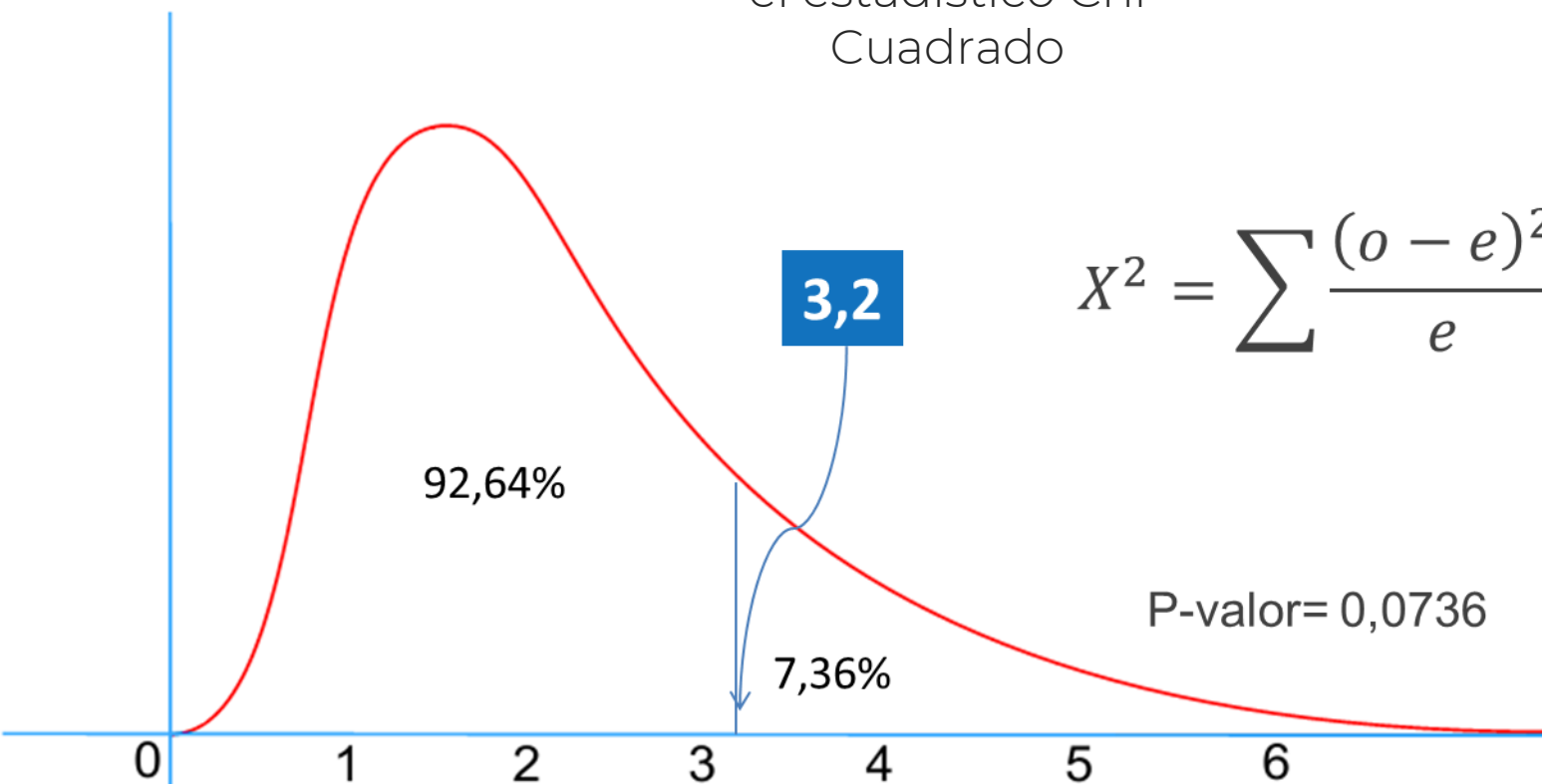


20 lanzamientos

$$X^2 = \sum \frac{(o - e)^2}{e}$$

	Observado	Esperado	Diferencia	Diferencia <sup>2</sup>	Dif <sup>2</sup> /Esperado
Cara	6	10	-4	16	1,6
Cruz	14	10	+4	16	1,6

Suma = 3,2



## Los detalles de la comparación de proporciones

- El valor esperado para una celda de una tabla de contingencias se calcula sumando las filas y columna de la casilla correspondiente y se divide por el total de individuos de la muestra

**Tabla cruzada Estrés laboral\*Área académica**

Recuento		Área académica			Total
		Ciencias de la Salud	Ciencias Sociales	Ingenierías	
Estrés laboral	Si	39	11	12	62
	No	5	16	17	38
Total		44	27	29	100

Suma por la fila  $\rightarrow 39 + 11 + 12 = 62$

Suma por la columna  $\rightarrow 39 + 5 = 44$

$$\frac{\text{Suma por la fila} * \text{Suma por la Columna}}{\text{TOTAL}} = \frac{62 * 44}{100} = 27.28$$

- Una vez se calcular el valor Esperado puedes calcular la Chi Cuadrado con la misma fórmula:

$$\chi^2 = \sum \frac{(o - e)^2}{e}$$

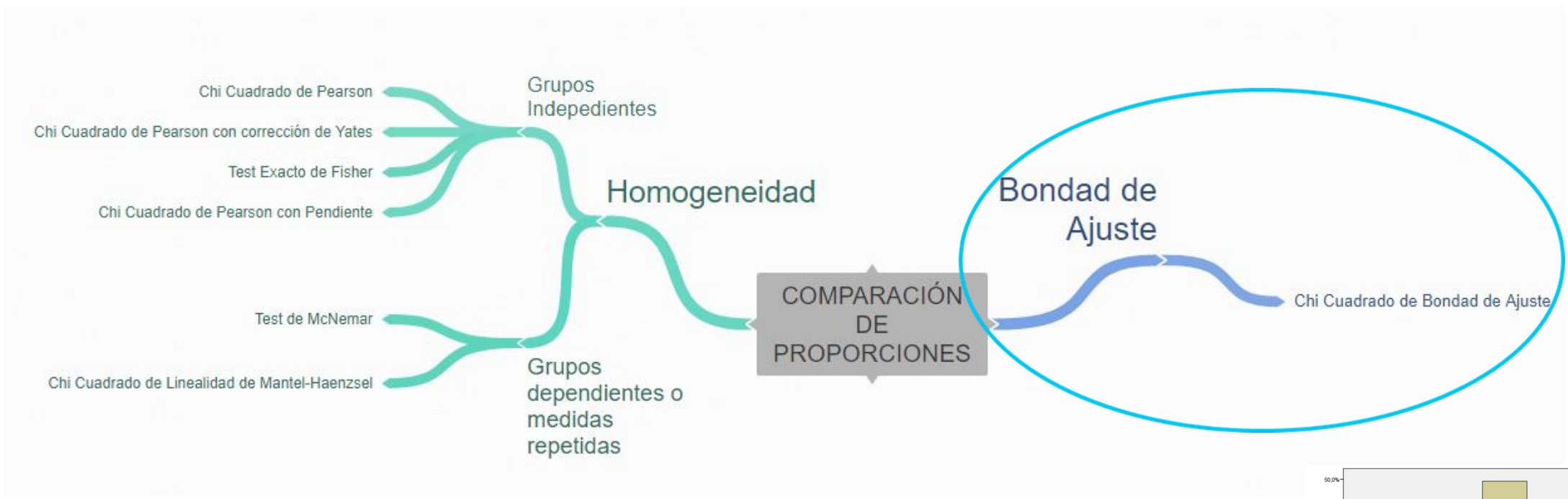
El requisito del test Chi Cuadrado de Pearson es que el valor esperado de todas las casillas sea mayor a 5

**Tabla cruzada Estrés laboral\*Área académica**

			Área académica			Total
			Ciencias de la Salud	Ciencias Sociales	Ingenierías	
Estrés laboral	Si	Recuento	39	11	12	62
		Recuento esperado	27,3	16,7	18,0	62,0
	No	Recuento	5	16	17	38
		Recuento esperado	16,7	10,3	11,0	38,0
Total		Recuento	44	27	29	100
		Recuento esperado	44,0	27,0	29,0	100,0

# Bondad de ajuste de proporciones

Cómo podemos comparar proporciones entre grupos

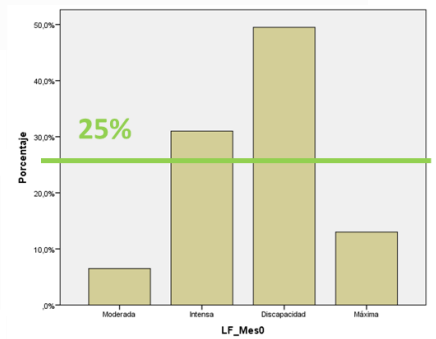


**Tabla cruzada Estrés laboral / Área académica**

			Área académica			Total
			Ciencias de la Salud	Ciencias Sociales	Ingenierías	
Estrés laboral	Sí	Recuento	39	11	12	62
		% dentro de Área académica	88,6%	40,7%	41,4%	62,0%
	No	Recuento	5	16	17	38
		% dentro de Área académica	11,4%	59,3%	58,6%	38,0%
Total		Recuento	44	27	29	100
		% dentro de Área académica	100,0%	100,0%	100,0%	100,0%

**LF\_Mes0**

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido Moderada	13	6,5	6,5	6,5
Intensa	62	31,0	31,0	37,5
Discapacidad	99	49,5	49,5	87,0
Máxima	26	13,0	13,0	100,0
Total	200	100,0	100,0	

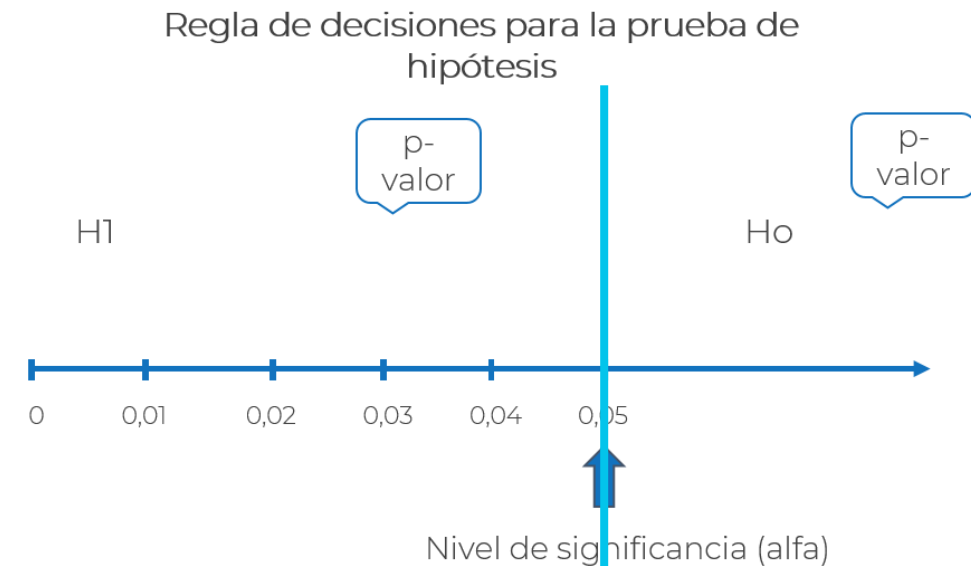
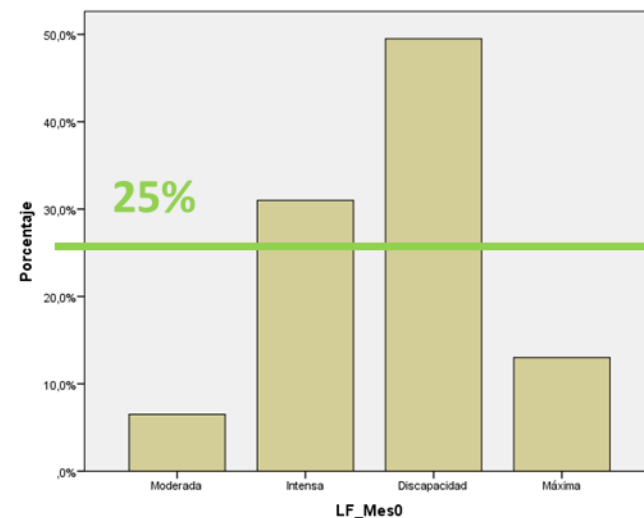


**OBJETIVO:**  
Comparar las proporciones estadísticamente entre grupos

## El paso a paso de la bondad de ajuste

- Comparar las proporciones estadísticamente entre grupos
- Variable Dependiente o Respuesta: LF\_mes0 (factor)
- Variable Independientes:
  - No tenemos, pero si valor esperado = 25% para todos los casos

LF_Mes0				
		Frecuencia	Porcentaje	Porcentaje válido
Válido	Moderada	13	6,5	6,5
	Intensa	62	31,0	31,0
	Discapacidad	99	49,5	49,5
	Máxima	26	13,0	13,0
	Total	200	100,0	100,0

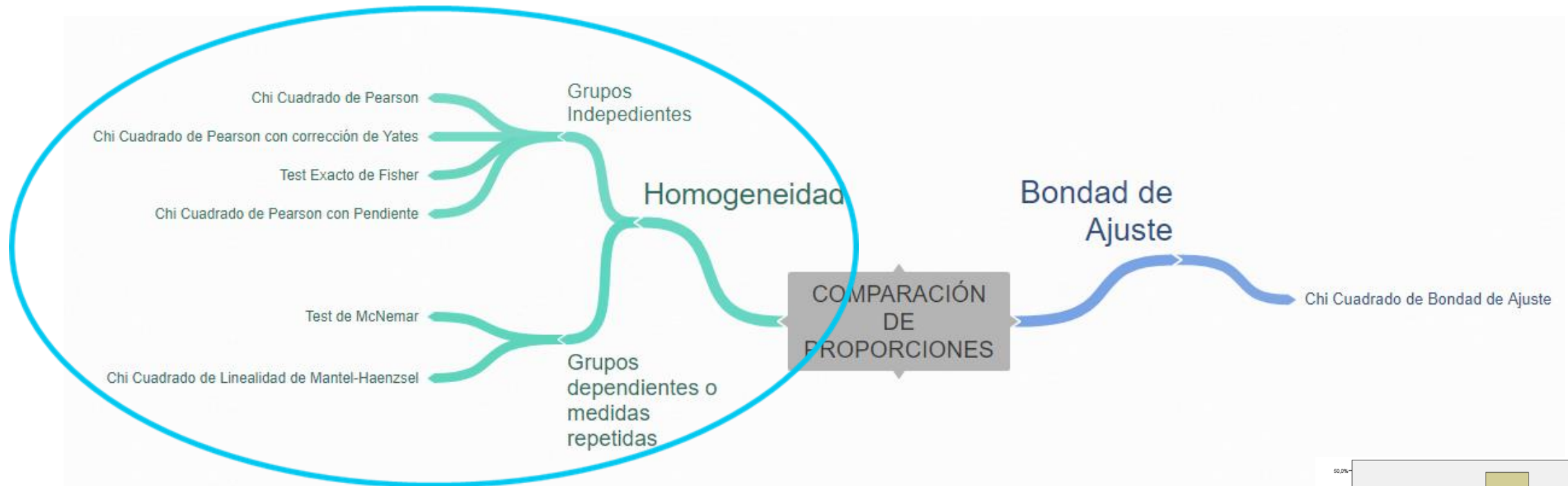


H 1 = hay diferencias significativas entre alguna de las proporciones

H 0 = hay igualdad entre las proporciones  
 $p1 = p2 = p3 = p4 = 25\%$

# Homogeneidad entre grupos de estudio

El paso a paso para comparar proporciones entre dos grupos de un factor de estudio

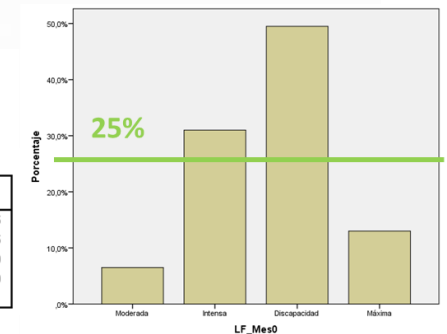


**Tabla cruzada Estrés laboral / Área académica**

			Área académica			Total
			Ciencias de la Salud	Ciencias Sociales	Ingenierías	
Estrés laboral	Si	Recuento	39	11	12	62
		% dentro de Área académica	88,6%	40,7%	41,4%	62,0%
	No	Recuento	5	16	17	38
		% dentro de Área académica	11,4%	59,3%	58,6%	38,0%
Total		Recuento	44	27	29	100
		% dentro de Área académica	100,0%	100,0%	100,0%	100,0%

**LF\_Mes0**

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido Moderada	13	6,5	6,5	6,5
Intensa	62	31,0	31,0	37,5
Discapacidad	99	49,5	49,5	87,0
Máxima	26	13,0	13,0	100,0
Total	200	100,0	100,0	





**OBJETIVO:**  
Comparar las proporciones estadísticamente entre grupos

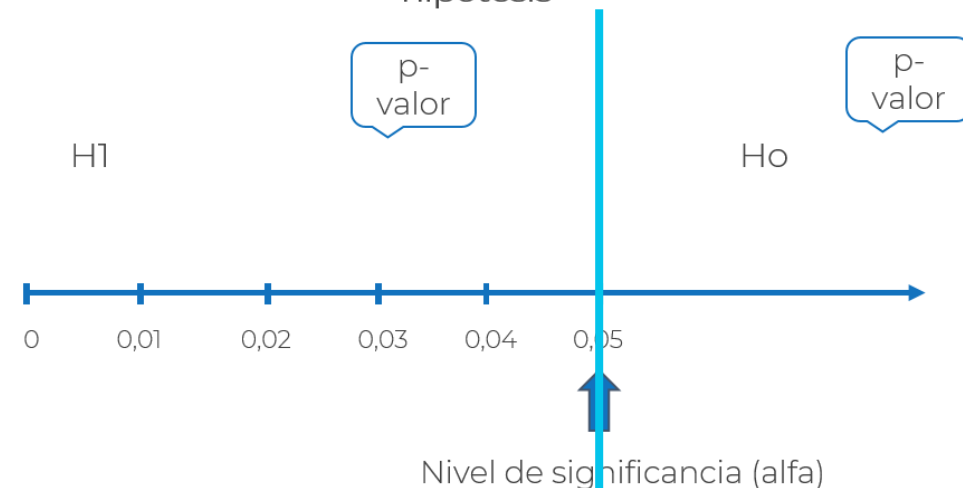
## El paso a paso de la homogeneidad

- Comparar las proporciones estadísticamente entre grupos
- **Variable Dependiente o Respuesta:** Estrés Laboral (factor) - Filas
- **Variable Independientes o de estudio:** área académica (factor) - Columnas

Tabla cruzada Estrés laboral / Área académica

			Área académica			Total
			Ciencias de la Salud	Ciencias Sociales	Ingenierías	
Estrés laboral	Si	Recuento	39	11	12	62
		% dentro de Área académica	88,6%	40,7%	41,4%	62,0%
	No	Recuento	5	16	17	38
		% dentro de Área académica	11,4%	59,3%	58,6%	38,0%
Total		Recuento	44	27	29	100
		% dentro de Área académica	100,0%	100,0%	100,0%	100,0%

Regla de decisiones para la prueba de hipótesis



H 1 = hay diferencias significativas entre alguna de las proporciones

$$pSI \neq pNO$$

Para alguno de los grupos de área académica

H 0 = hay igualdad entre las proporciones

$$pSI = pNO$$

Para todos los grupos de área académica

# OBJETIVO: Comparar las proporciones estadísticamente entre grupos

## El paso a paso de la homogeneidad

- Comparar las proporciones estadísticamente entre grupos
- Variable Dependiente o Respuesta: Estrés Laboral (factor) - Filas
- Variable Independientes o de estudio: área académica (factor) - Columnas

Tabla cruzada Estrés laboral'Área académica

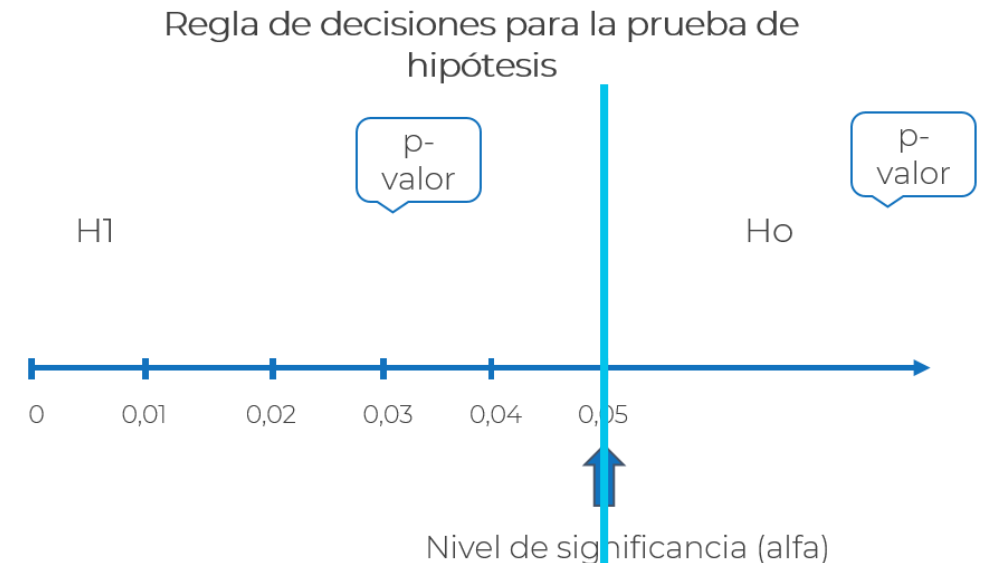
			Área académica			Total
			Ciencias de la Salud	Ciencias Sociales	Ingenierías	
Estrés laboral	Si	Recuento	39	11	12	62
		Recuento esperado	27,3	16,7	18,0	62,0
	No	Recuento	5	16	17	38
		Recuento esperado	16,7	10,3	11,0	38,0
Total		Recuento	44	27	29	100
		Recuento esperado	44,0	27,0	29,0	100,0

Pruebas de chi-cuadrado

	Valor	gl	Significación asintótica (bilateral)	Significación exacta (bilateral)	Significación exacta (unilateral)
Chi-cuadrado de Pearson	,000 <sup>a</sup>	1	,988		
Corrección de continuidad <sup>b</sup>	,000	1	1,000		
Razón de verosimilitud	,000	1	,988		
Prueba exacta de Fisher				1,000	,550
Asociación lineal por lineal	,000	1	,988		
N de casos válidos	200				

a. 0 casillas (0,0%) han esperado un recuento menor que 5. El recuento mínimo esperado es 44,06.

b. Sólo se ha calculado para una tabla 2x2



H 1 = hay diferencias significativas entre alguna de las proporciones

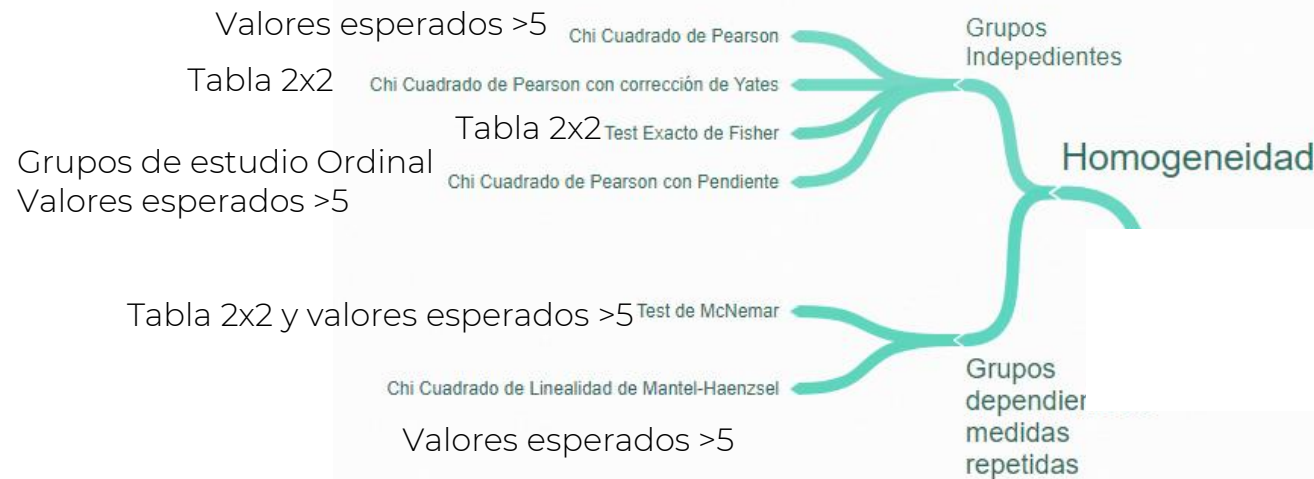
$pSI \neq pNO$   
Para alguno de los grupos de área académica

H 0 = hay igualdad entre las proporciones

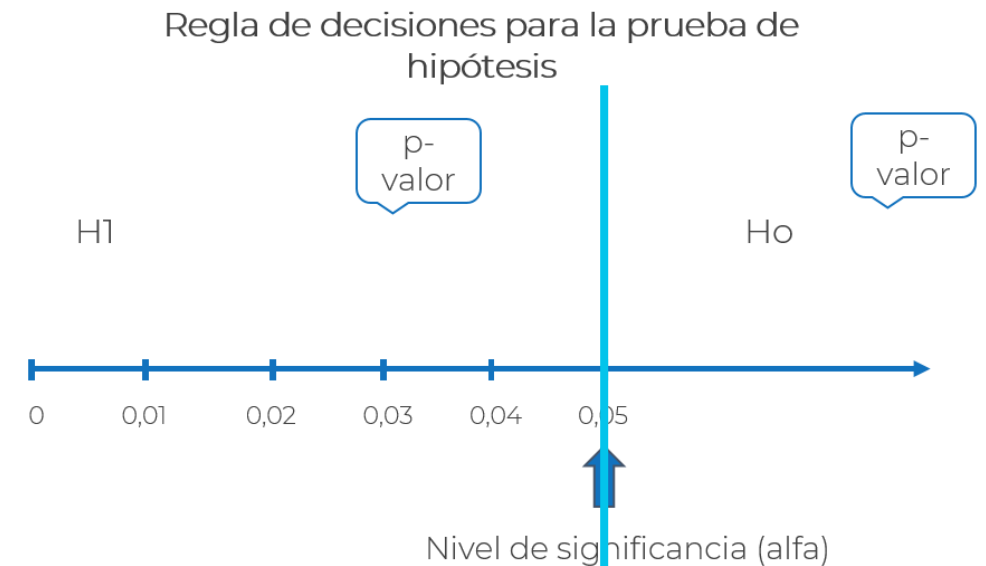
$pSI = pNO$   
Para todos los grupos de área académica

# El paso a paso de la homogeneidad

- 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



**OBJETIVO:**  
Comparar las proporciones estadísticamente entre grupos de estudio



H 1 = hay diferencias significativas entre alguna de las proporciones

$$pSI \neq pNO$$

Para alguno de los grupos de área académica

H 0 = hay igualdad entre las proporciones

$$pSI = pNO$$

Para todos los grupos de área académica

# Take away

El resumen de la 1/2 lección

## Lo más importante de la lección

- La comparación de proporciones es una práctica habitual en estudios científicos. Sobre todo de casos controles
- Podemos dividir los problemas en:
  - Bondad de ajuste – comprobar una distribución observada con una teórica
  - Homogeneidad – comprobar si grupos tienen la misma distribución de proporciones por factor de estudio

# Tú turno

A por tus primeros test estadísticos

## A poner en práctica lo que has visto

- Descarga la hoja de trabajo
- Compara las proporciones con la ayuda de esta lección.
- No te olvides de describir los datos con la ayuda del bloque 4
- ¡Go go go!