

# Pruebas estadísticas en R: X2, t y ANOVA

RLadies Cuernavaca, junio de 2022

Aurora Labastida

## Pruebas de hipótesis

¿La altura promedio es distinta según el sexo?



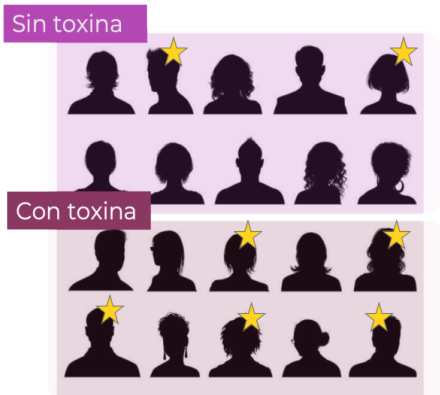
Promedios iguales

Hipótesis nula  $H_0$

Promedios diferentes

Hipótesis alternativa  $H_1$

## Pruebas de hipótesis



¿La toxina está asociada con la enfermedad?

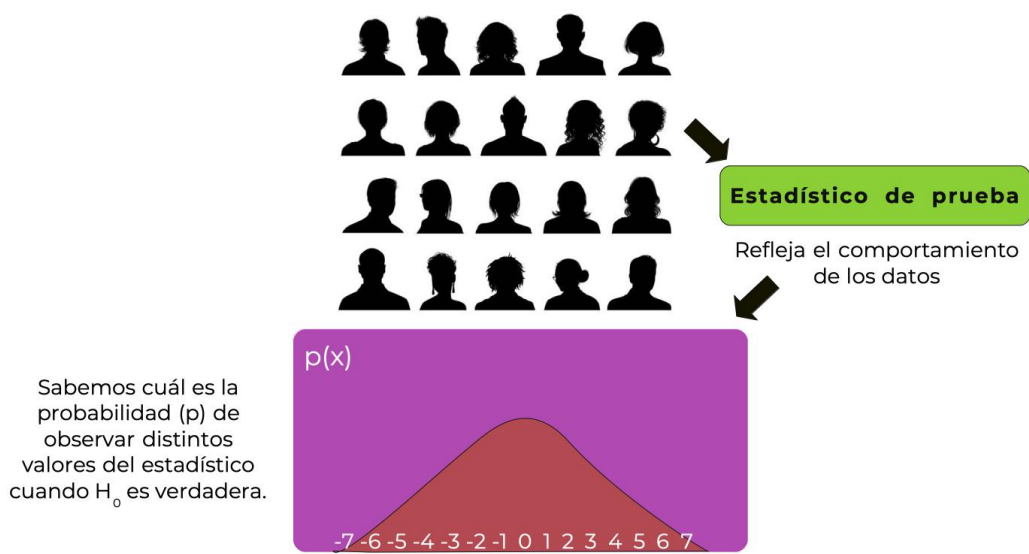
Hipótesis nula  $H_0$

Variables no asociadas

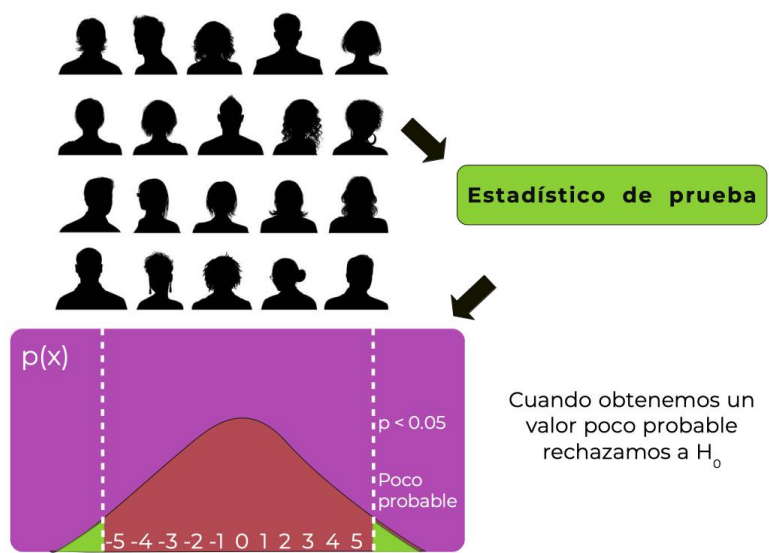
Hipótesis alternativa  $H_1$

Variables asociadas

Pruebas de hipótesis



Pruebas de hipótesis



Pruebas de hipótesis

		No Paramétrica			Paramétrica
		Nominal: 2 categorías	Nominal: 3 o más categorías	Numérica no normal (u ordinal)	Numérica
Muestras Independientes	1 Grupo	$\chi^2$ Bondad de ajuste	$\chi^2$ Bondad de ajuste	$\chi^2$ Bondad de ajuste	T de Student para una muestra
	2 Grupos	$\chi^2$ de homogeneidad + corrección de Yates o test exacto de Fisher	$\chi^2$ de homogeneidad	U Mann-Whitney	T de Student para muestras independientes
	3 o más Grupos	$\chi^2$ de homogeneidad	$\chi^2$ de homogeneidad	H Kruskal-Wallis	ANOVA
Muestras relacionadas	2 Mediciones	Mc Nemar	Q de Cochran	Wilcoxon	T de Student para muestras no independientes
(Estudio longitudinal)	3 o más Mediciones	Q de Cochran	Q de Cochran	Friedman	ANOVA para muestras no independientes

# Test X²

		Enfermedad		
		NO	SI	
Toxina	NO	100	100	200
	SI	50	150	200
		150	250	400

H<sub>0</sub>: Las variables son independientes entre sí

H<sub>1</sub>: Las variables están asociadas entre sí

# Test X²

Valores observados (o)

		Enfermedad		
		NO	SI	
Toxina	NO	100	100	200
	SI	50	150	200
		150	250	400

Valores esperados (e)

75	125	200
75	125	200
150	250	400

$$e = \frac{\text{total fila} * \text{total columna}}{\text{total global}}$$
$$e = \frac{200 * 250}{400}$$

# Test X²

Valores observados (o)

		Enfermedad		
		NO	SI	
Toxina	NO	100	100	200
	SI	50	150	200
		150	250	400

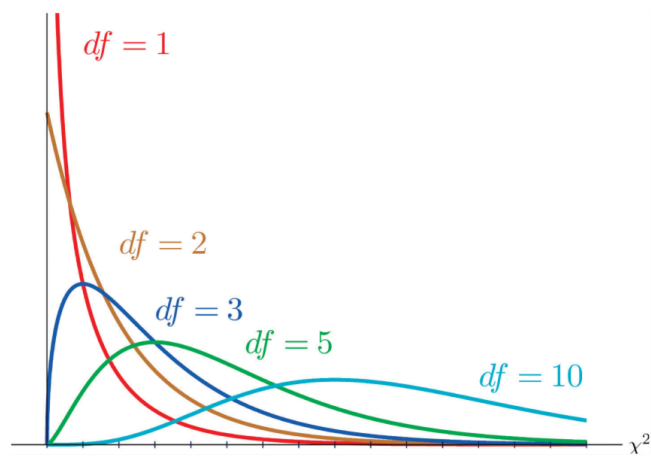
Valores esperados (e)

75	125	200
75	125	200
150	25	400

$$\chi^2 = \sum \frac{(o-e)^2}{e} = \frac{(100-75)^2}{100} + \frac{(100-125)^2}{125} + \frac{(50 - 75)^2}{75} + \frac{(150 - 125)^2}{125}$$

Test X²

Grados de libertad (k) = (Número de filas - 1) x (Número de columnas - 1)



Test X²

		Valores observados (o)		Valores esperados (e)		Residuales de Pearson	
		Enfermedad					
		NO	SI				
Toxina	NO	100	100	75	125	2.9	-2.2
	SI	50	150	75	125	-2.9	2.2
		150	250	150	250	150	250

Test X²

		Valores observados (o)		Valores esperados (e)		Residuales de Pearson	
		Enfermedad					
		NO	SI				
Toxina	NO	100	100	75	125	2.9	-2.2
	SI	50	150	75	125	-2.9	2.2
		150	250	150	250	150	250

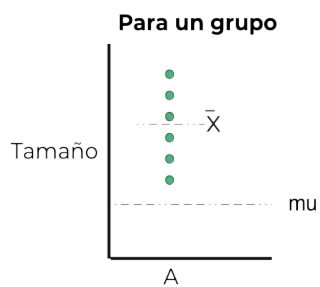
$$\chi^2 = \sum \frac{(o-e)^2}{e}$$

$$r = \frac{o - e}{\sqrt{e}}$$

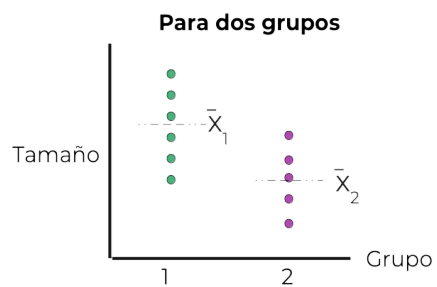
contribución del residual= 
$$\frac{r^2}{\chi^2}$$



## t de Student

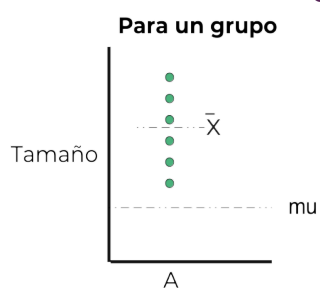


$H_0$ : El promedio es distinto a  $\mu$



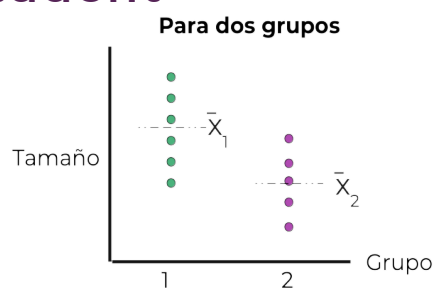
$H_0$ : Los promedios de los grupos A y B son iguales

## t de Student



$H_0$ : El promedio es distinto a  $\mu$

$$t = \frac{(\bar{X} - \mu)^2}{s / \sqrt{n}}$$



$H_0$ : Los promedios de los grupos A y B son iguales

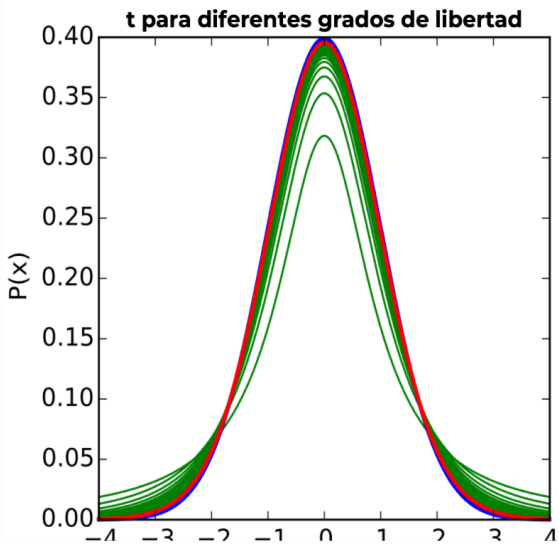
$$t = \frac{(\bar{X}_1 - \bar{X}_2)^2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

# t de Student

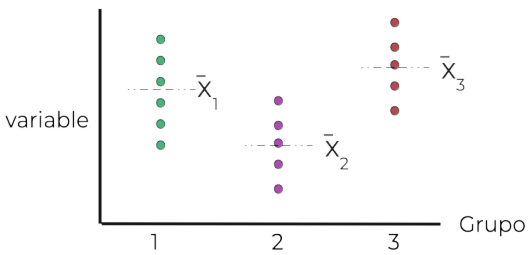
Grados de libertad (k)

Para un grupo  
 $k = n - 1$

Para dos grupos  
 $k = n_1 + n_2 - 2$



# ANOVA



$H_0$ : Los promedios de los grupos son iguales entre si

$H_1$ : Por lo menos un promedio es distinto a los demás

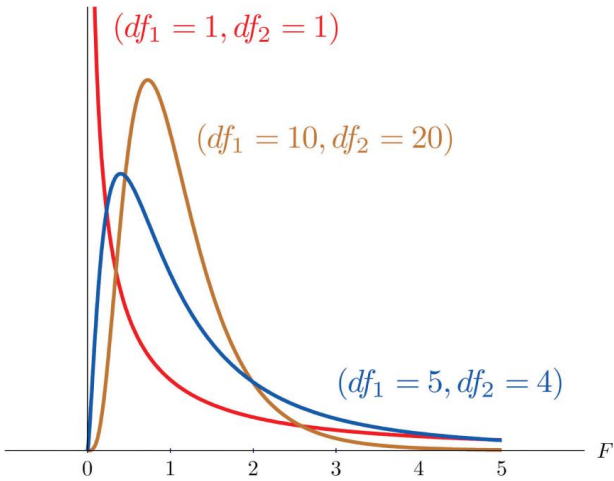
$$F = \frac{\text{Variación explicada por los grupos}}{\text{Variación no explicada por los grupos}}$$

$$F = \frac{\text{Variación entre el promedio de los grupos}}{\text{Variación al interior de los grupos}}$$

# ANOVA

$df_1 = \text{Grupos} - 1$

$df_2 = n - 1$



**¡Muchas gracias!**