

---

# Statistical Analysis for Speedup

Roberto Carrasco

---

---

# Objetivos

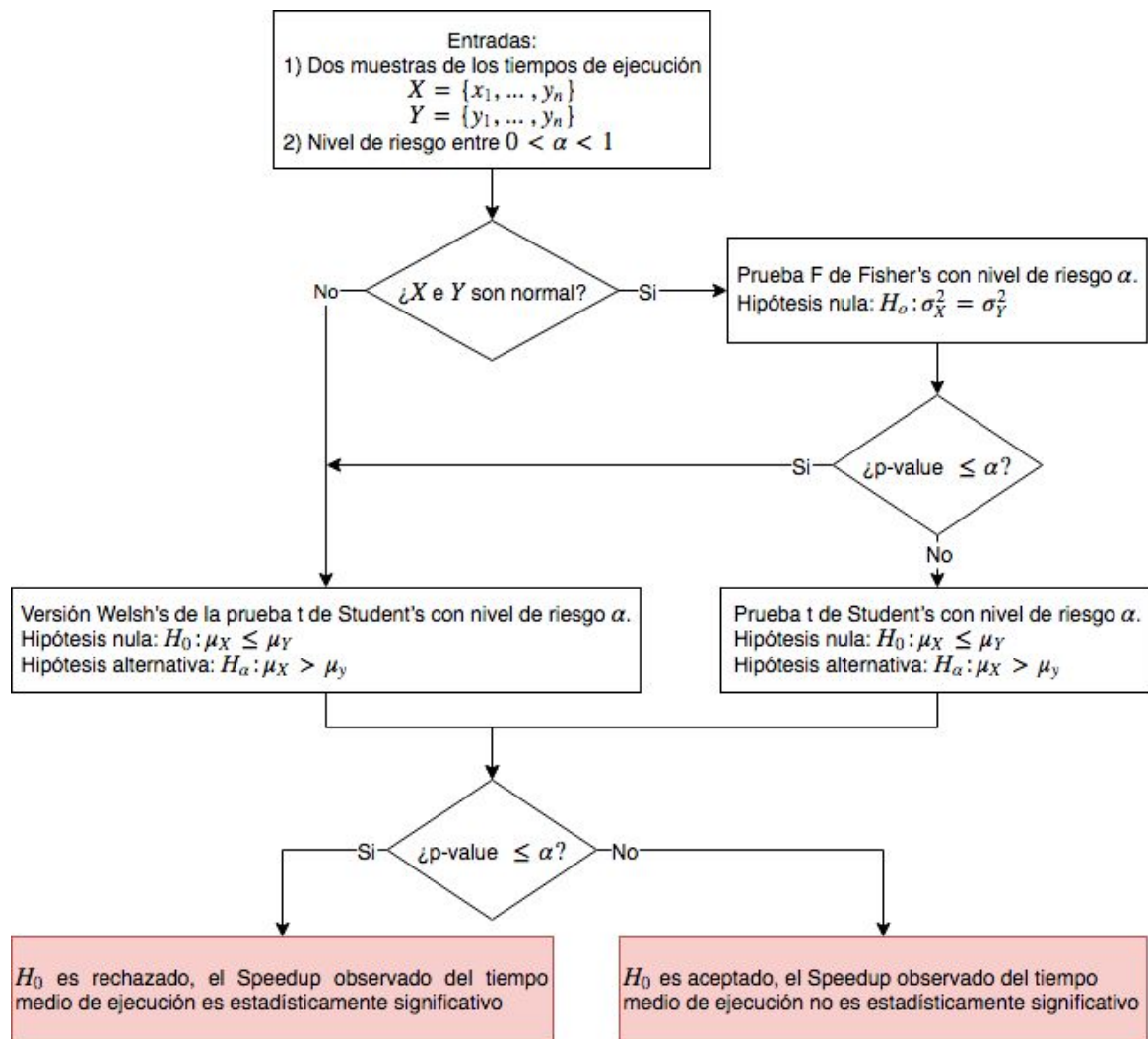
- Análisis de media por medio de una prueba t de Student's
  - Análisis de mediana por medio de una prueba Wilcoxon-Mann-Whitney
  - Speedup
-

---

# Objetivos

- **Análisis de media por medio de una prueba t de Student's**
  - Análisis de mediana por medio de una prueba Wilcoxon-Mann-Whitney
  - Speedup
-

# Análisis de medias



## Entradas:

X e Y son muestras con  $n$  observaciones.  
Se asume un nivel de riesgo  $\alpha$  ( $\alpha: 0.05$ ).

Entradas:

1) Dos muestras de los tiempos de ejecución

$$X = \{x_1, \dots, x_n\}$$

$$Y = \{y_1, \dots, y_n\}$$

2) Nivel de riesgo entre  $0 < \alpha < 1$

Entradas:

- 1) Dos muestras de los tiempos de ejecución  
 $X = \{x_1, \dots, x_n\}$   
 $Y = \{y_1, \dots, y_n\}$
- 2) Nivel de riesgo entre  $0 < \alpha < 1$

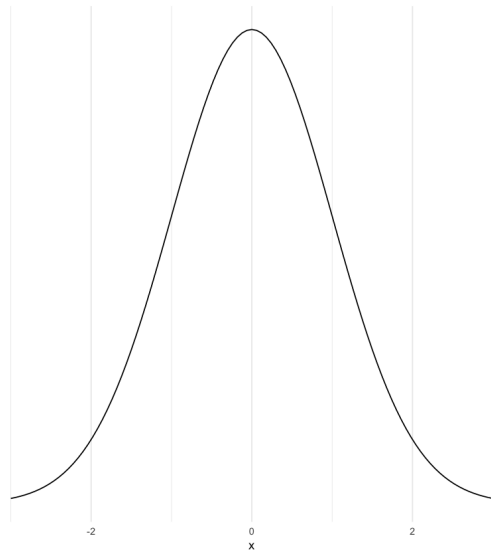
$\alpha: 0.05$

## Muestra X

0.054538 0.054415 0.054354 0.054456 0.054529 0.054426 0.054467 0.054456 0.0542 0.05462 0.054364 0.054569 0.054559 0.054303  
 0.054579 0.054282 0.054446 0.054313 0.054538 0.054558 0.054538 0.054467 0.054548 0.054385 0.05463 0.054426 0.054528  
 0.054364 0.054528 0.054548 0.054282 0.054426 0.054548 0.054467 0.054405 0.054364 0.054508 0.054405 0.054497 0.054405  
 0.054333 0.054548 0.05462 0.054518 0.054405 0.054641 0.054262 0.054436 0.054641 0.054496 0.054508 0.054415 0.054385  
 0.054497 0.054651 0.054497 0.054344 0.054753 0.054497 0.054528 0.054395 0.054764 0.054497 0.054272 0.054477 0.054395  
 0.054774 0.054518 0.054436 0.054559 0.054733 0.054467 0.054456 0.054508 0.054405 0.054497 0.0546 0.054395 0.054487  
 0.054385 0.054487 0.054498 0.054426 0.054518 0.054486 0.054374 0.054579 0.054415 0.054467 0.054426 0.054537 0.054374  
 0.05463 0.054405 0.054456 0.054764 0.054374 0.054497 0.054416 0.054579

## Muestra Y

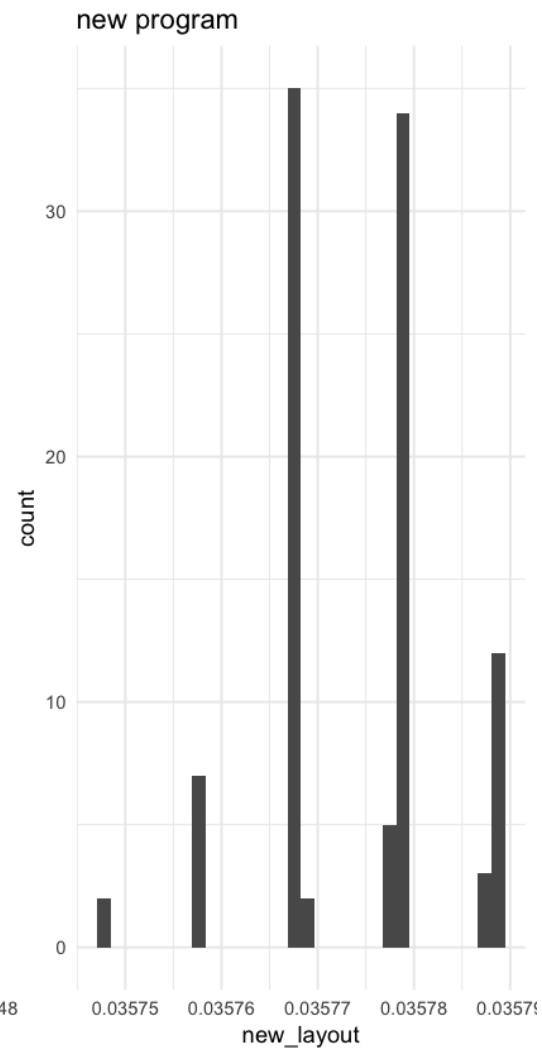
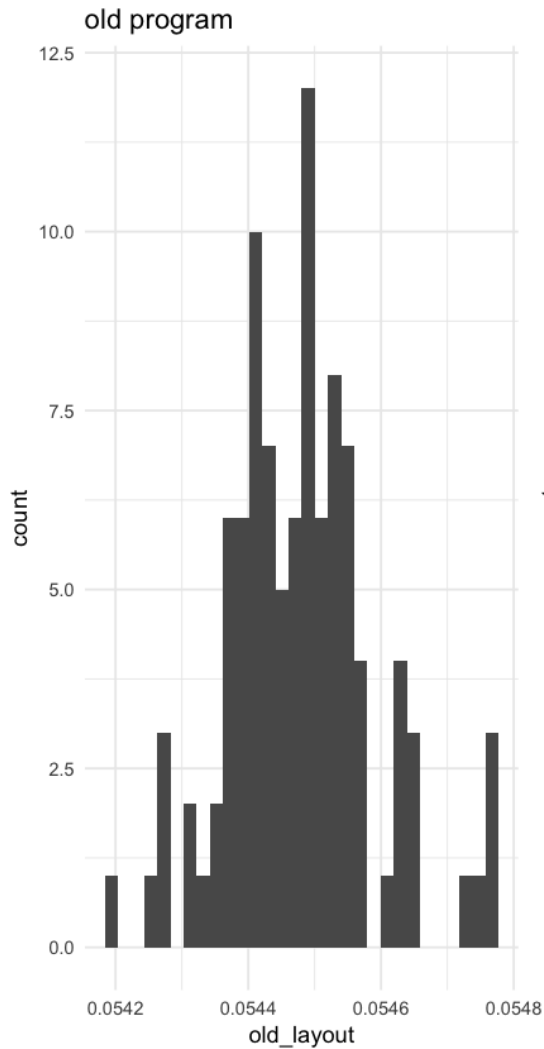
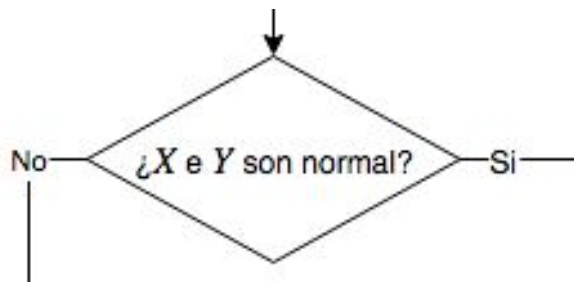
0.035768 0.035779 0.035758 0.035768 0.035748 0.035779 0.035779 0.035779 0.035779 0.035779 0.035779 0.035758 0.035779  
 0.035758 0.035768 0.035768 0.035779 0.035779 0.035768 0.035768 0.035768 0.035758 0.035768 0.035768 0.035779 0.035779  
 0.035778 0.035768 0.035768 0.035779 0.035768 0.035779 0.035789 0.035769 0.035789 0.035768 0.035768 0.035789 0.035789  
 0.035779 0.035779 0.035768 0.035779 0.035768 0.035789 0.035779 0.035789 0.035789 0.035769 0.035789 0.035779 0.035779  
 0.035779 0.035779 0.035758 0.035779 0.035748 0.035779 0.035779 0.035768 0.035779 0.035779 0.035768 0.035788 0.035789  
 0.035779 0.035778 0.035779 0.035768 0.035768 0.035767 0.035768 0.035768 0.035767 0.035779 0.035789 0.035788 0.035768  
 0.035789 0.035778 0.035768 0.035768 0.035768 0.035778 0.035768 0.035779 0.035788 0.035768 0.035768 0.035768 0.035779  
 0.035779 0.035777 0.035758 0.035768 0.035779 0.035758 0.035768 0.035789 0.035768



- Normalidad de las muestras
- Tamaño de las muestras

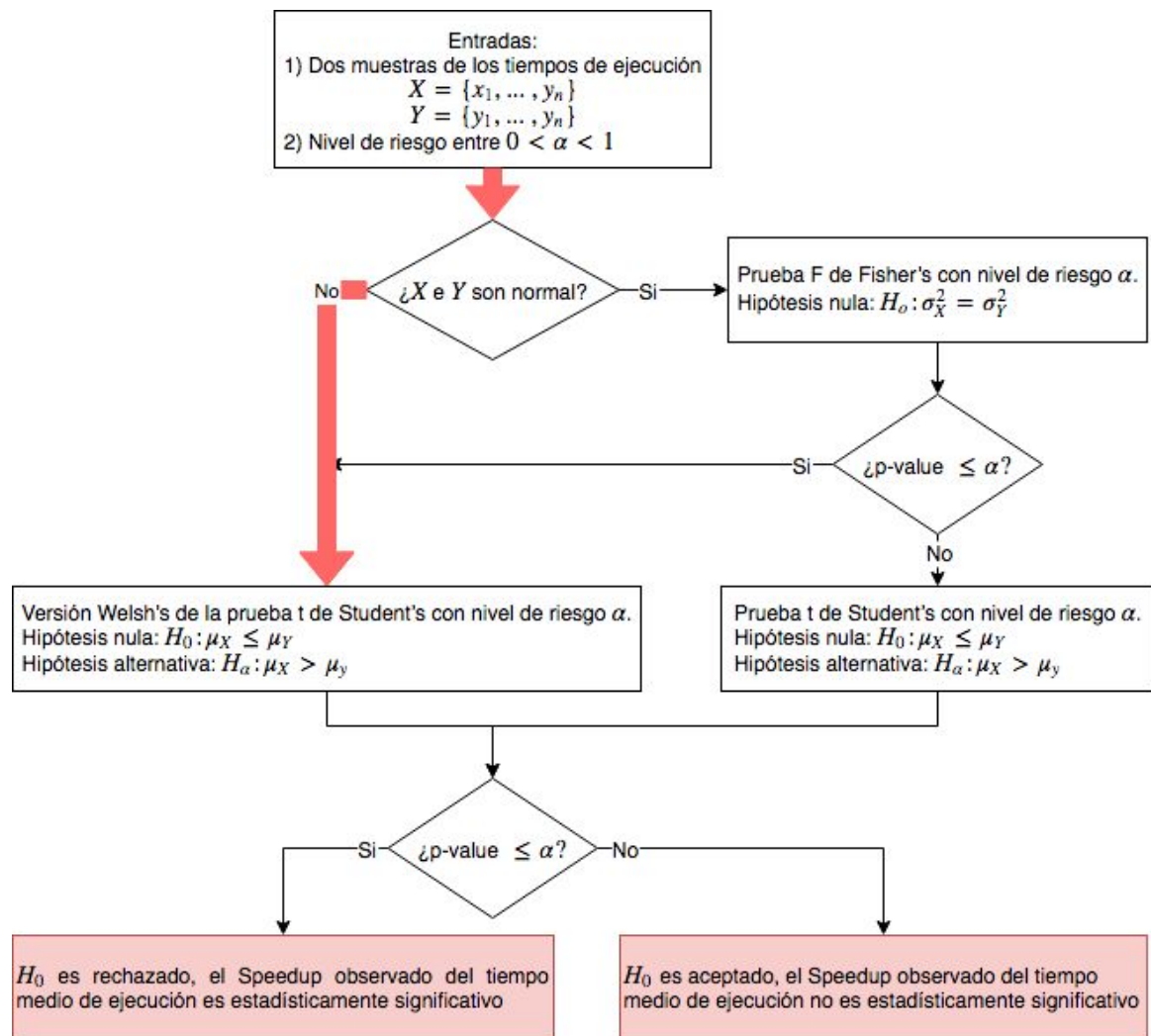
(n>30)	Normal	No Normal
n ≤ 30	Si	-
n > 30	Si	No

# Normalidad de las muestras





# Análisis de medias



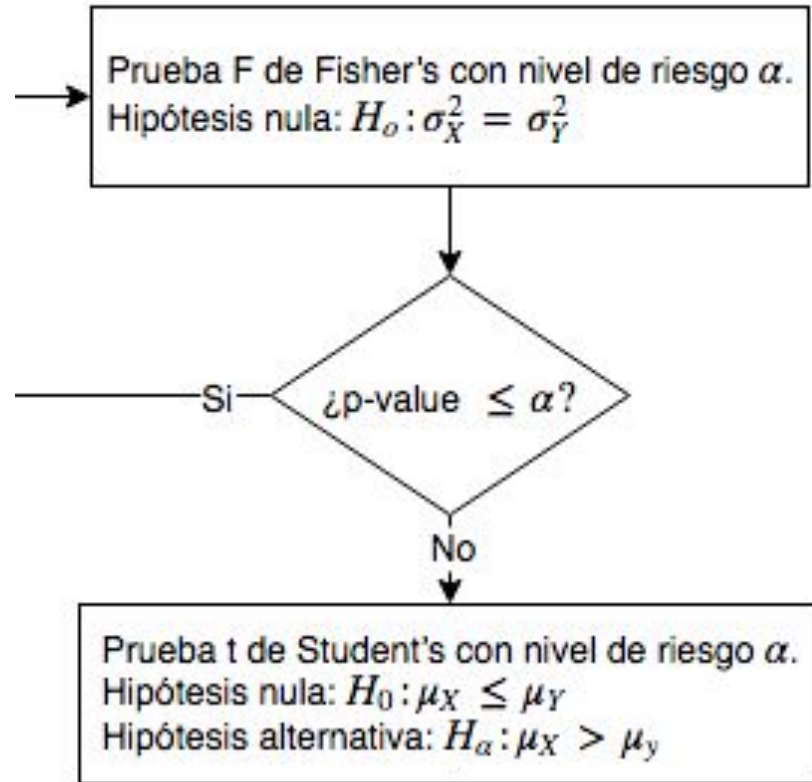
Versión Welsh's de la prueba t de Student's con nivel de riesgo  $\alpha$ .  
Hipótesis nula:  $H_0: \mu_X \leq \mu_Y$   
Hipótesis alternativa:  $H_a: \mu_X > \mu_Y$

Welch Two Sample t-test

data: old\_layout and new\_layout  
t = 373.71, df = 99, p-value < 2.2e-16 (0.00000000000000022204)  
alternative hypothesis: true difference in means is not equal to 0  
95 percent confidence interval:  
1.83023e-05 1.84977e-05  
sample estimates:  
mean of x mean of y  
5.44e-05 3.60e-05

## T de welch:

- Lo usamos para probar la hipótesis de que dos poblaciones tienen medias iguales.
- Es más confiable cuando las dos muestras tienen varianzas desiguales y tamaños de muestra desiguales.



## F de Fisher's:

- Probar que las desviaciones estándar de dos poblaciones normalmente distribuidas tienen diferencia significativas.

## T de Student's:

- Probar que las medias de dos poblaciones normalmente distribuidas tienen diferencia significativas.

# Hipótesis:

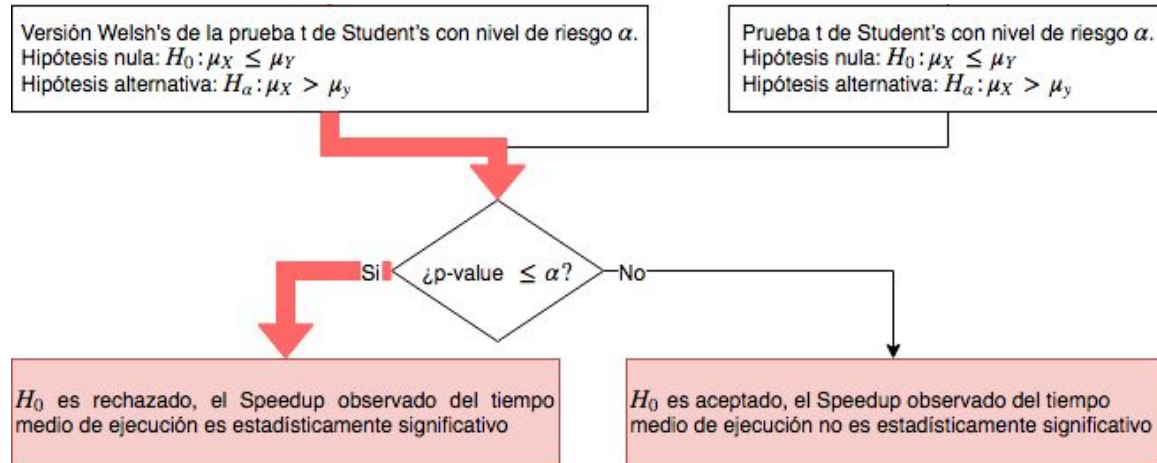
Hipótesis nula

Hipótesis alternativa

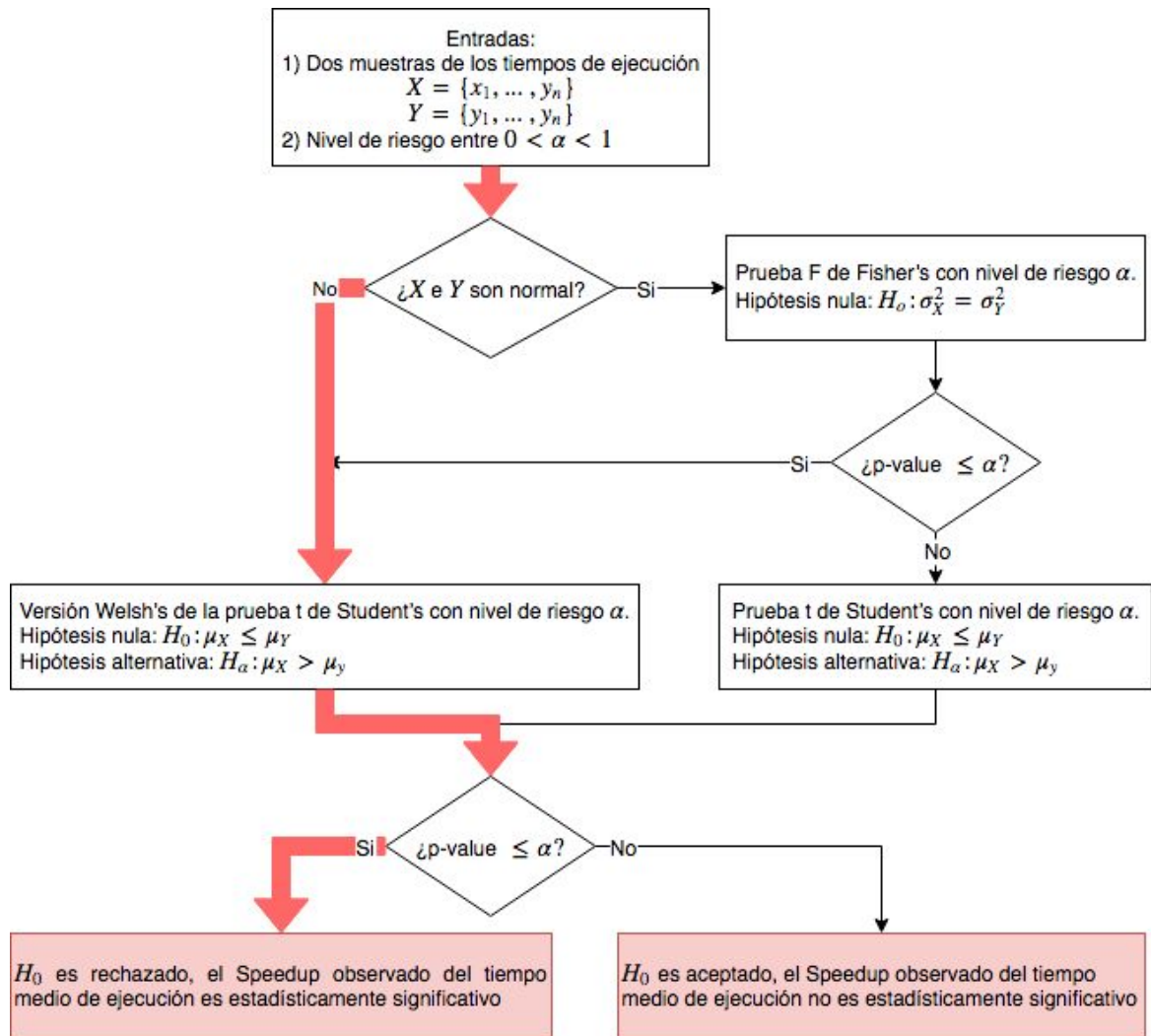
“si el valor p es inferior al nivel de significación, entonces la hipótesis nula es rechazada”

p-value:  $2.2e-16$

$\alpha$ : 0.05



# Resumen: Análisis de medias

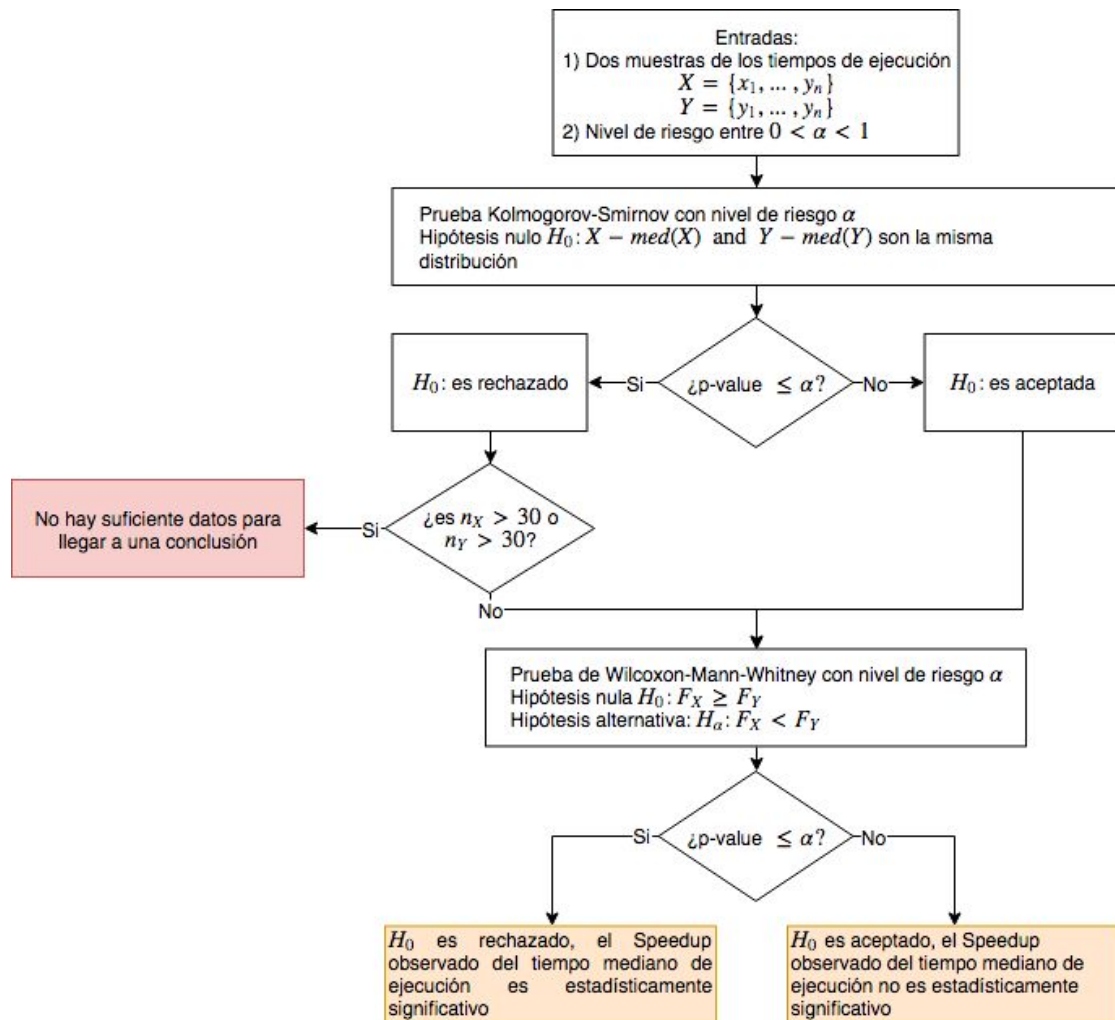


---

# Objetivos

- Análisis de media por medio de una prueba t de Student's
  - **Análisis de mediana por medio de una prueba Wilcoxon-Mann-Whitney**
  - Speedup
-

# Análisis de medianas



Entradas:

- 1) Dos muestras de los tiempos de ejecución  
 $X = \{x_1, \dots, x_n\}$   
 $Y = \{y_1, \dots, y_n\}$
- 2) Nivel de riesgo entre  $0 < \alpha < 1$

$\alpha: 0.05$

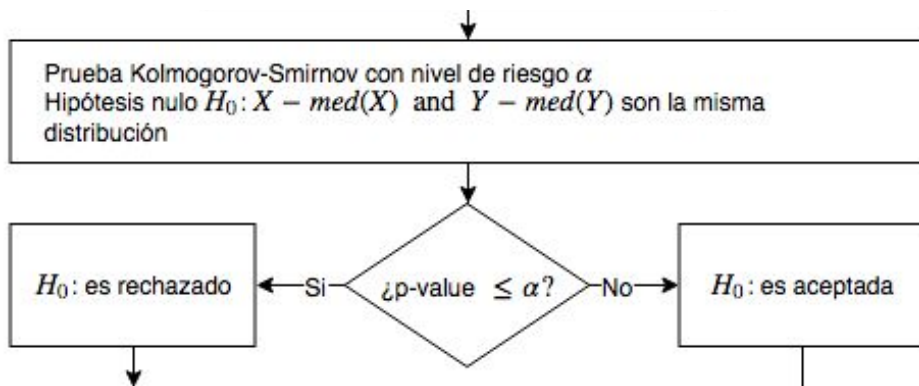
## Muestra X

0.054538 0.054415 0.054354 0.054456 0.054529 0.054426 0.054467 0.054456 0.0542 0.05462 0.054364 0.054569 0.054559 0.054303  
 0.054579 0.054282 0.054446 0.054313 0.054538 0.054558 0.054538 0.054467 0.054548 0.054385 0.05463 0.054426 0.054528  
 0.054364 0.054528 0.054548 0.054282 0.054426 0.054548 0.054467 0.054405 0.054364 0.054508 0.054405 0.054497 0.054405  
 0.054333 0.054548 0.05462 0.054518 0.054405 0.054641 0.054262 0.054436 0.054641 0.054496 0.054508 0.054415 0.054385  
 0.054497 0.054651 0.054497 0.054344 0.054753 0.054497 0.054528 0.054395 0.054764 0.054497 0.054272 0.054477 0.054395  
 0.054774 0.054518 0.054436 0.054559 0.054733 0.054467 0.054456 0.054508 0.054405 0.054497 0.0546 0.054395 0.054487  
 0.054385 0.054487 0.054498 0.054426 0.054518 0.054486 0.054374 0.054579 0.054415 0.054467 0.054426 0.054537 0.054374  
 0.05463 0.054405 0.054456 0.054764 0.054374 0.054497 0.054416 0.054579

## Muestra Y

0.035768 0.035779 0.035758 0.035768 0.035748 0.035779 0.035779 0.035779 0.035779 0.035779 0.035779 0.035758 0.035779  
 0.035758 0.035768 0.035768 0.035779 0.035779 0.035768 0.035768 0.035768 0.035758 0.035768 0.035768 0.035779 0.035779  
 0.035778 0.035768 0.035768 0.035779 0.035768 0.035779 0.035789 0.035769 0.035789 0.035768 0.035768 0.035789 0.035789  
 0.035779 0.035779 0.035768 0.035779 0.035768 0.035789 0.035779 0.035789 0.035789 0.035769 0.035789 0.035779 0.035779  
 0.035779 0.035779 0.035758 0.035779 0.035748 0.035779 0.035779 0.035768 0.035779 0.035779 0.035768 0.035788 0.035789  
 0.035779 0.035778 0.035779 0.035768 0.035768 0.035767 0.035768 0.035768 0.035767 0.035779 0.035789 0.035788 0.035768  
 0.035789 0.035778 0.035768 0.035768 0.035768 0.035778 0.035768 0.035779 0.035788 0.035768 0.035768 0.035768 0.035779  
 0.035779 0.035777 0.035758 0.035768 0.035779 0.035758 0.035768 0.035789 0.035768





Two-sample Kolmogorov-Smirnov test

data: ds\$old\_layout and ds\$new\_layout

D = 1, p-value < 2.2e-16

alternative hypothesis: two-sided

## Kolmogorov-Smirnov test

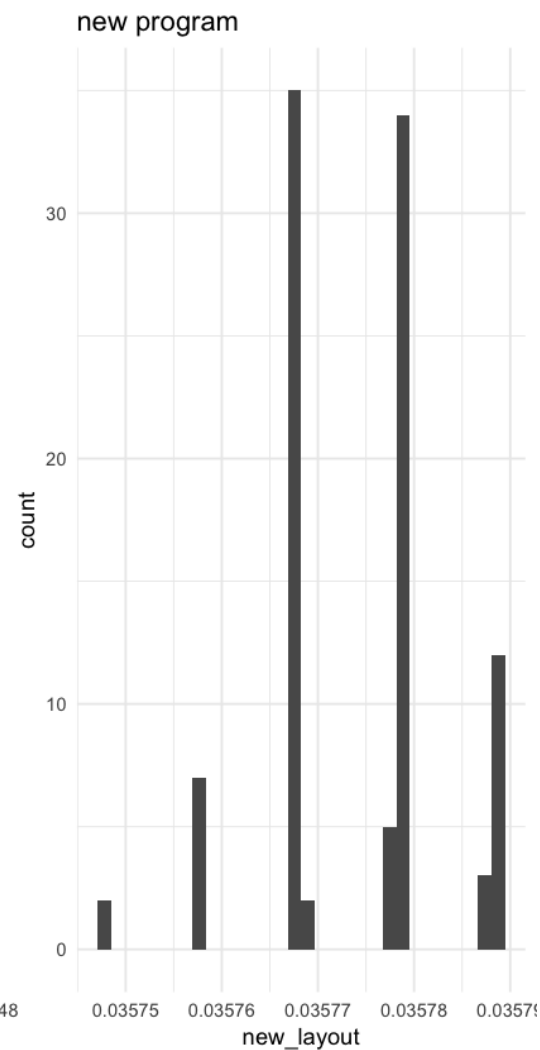
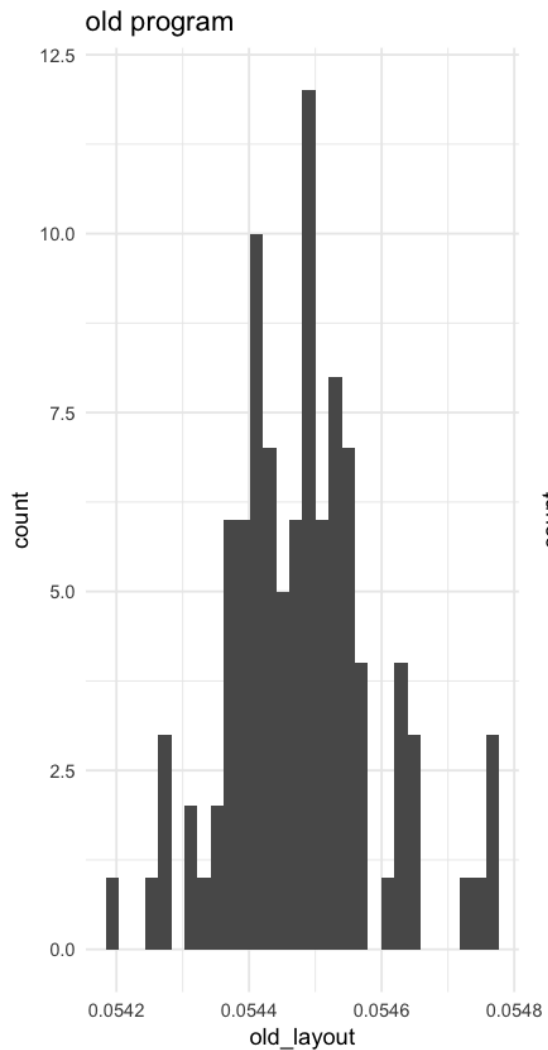
- Prueba no paramétrica que determina la bondad de ajuste.
- $H : X-med(X) \wedge Y-med(Y)$

# Distribución de las muestras

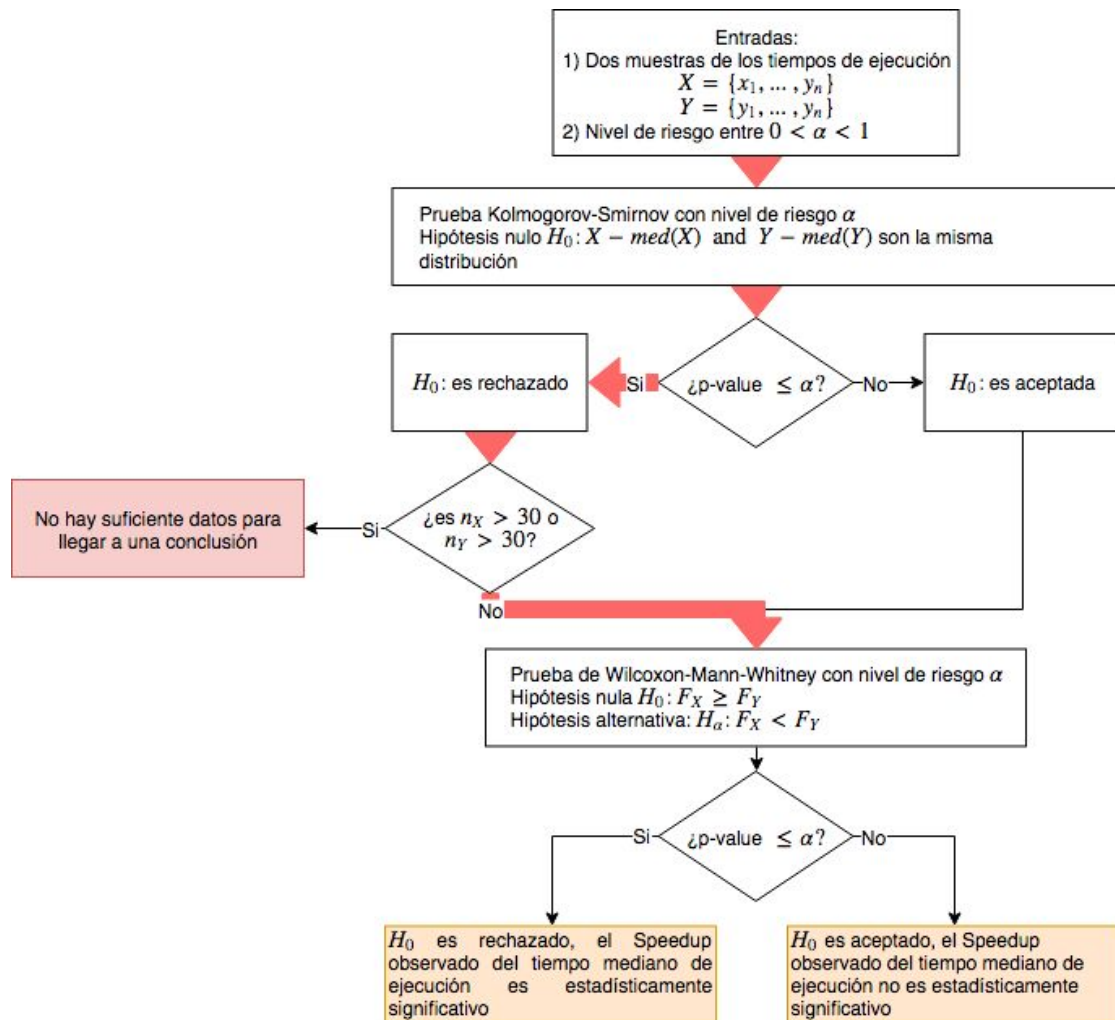
No hay suficiente datos para llegar a una conclusión

Si

¿es  $n_X > 30$  o  $n_Y > 30$ ?



# Análisis de medianas



Prueba de Wilcoxon-Mann-Whitney con nivel de riesgo  $\alpha$

Hipótesis nula  $H_0: F_X \geq F_Y$

Hipótesis alternativa:  $H_a: F_X < F_Y$

Wilcoxon rank sum test with continuity correction

data: ds\$old\_layout and ds\$new\_layout

W = 10000, p-value < 2.2e-16

alternative hypothesis: true location shift is not equal to 0

95 percent confidence interval:

1.8e-05 1.9e-05

sample estimates:

difference in location

1.8e-05

## Wilcoxon-Mann-Whitney:

- Lo usamos para probar la hipótesis de que dos poblaciones tienen mediana iguales.
- No requiere que las muestras se distribuyen normalmente

# Hipótesis:

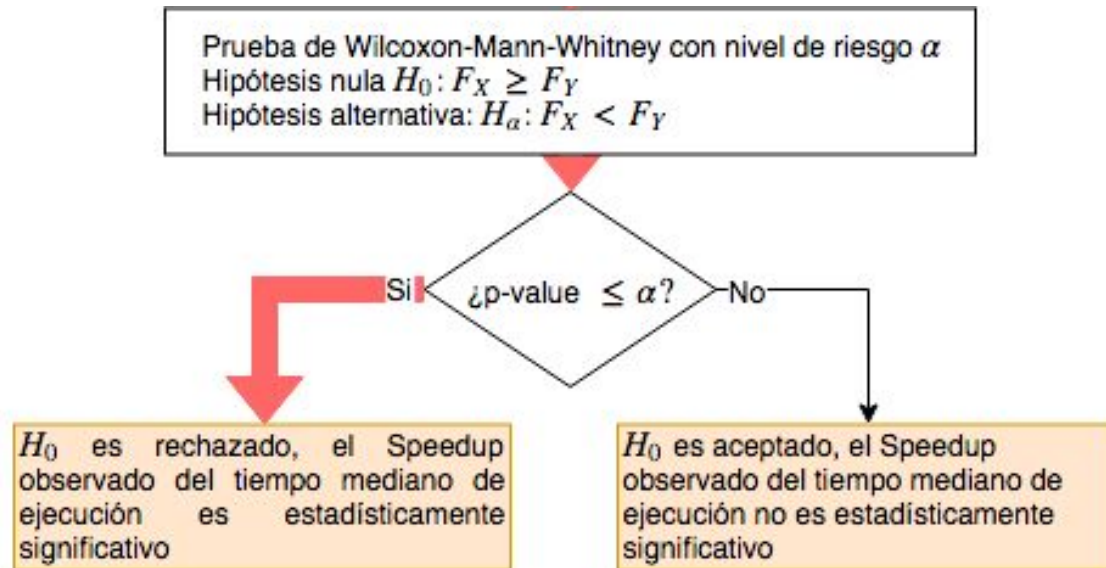
Hipótesis nula

Hipótesis alternativa

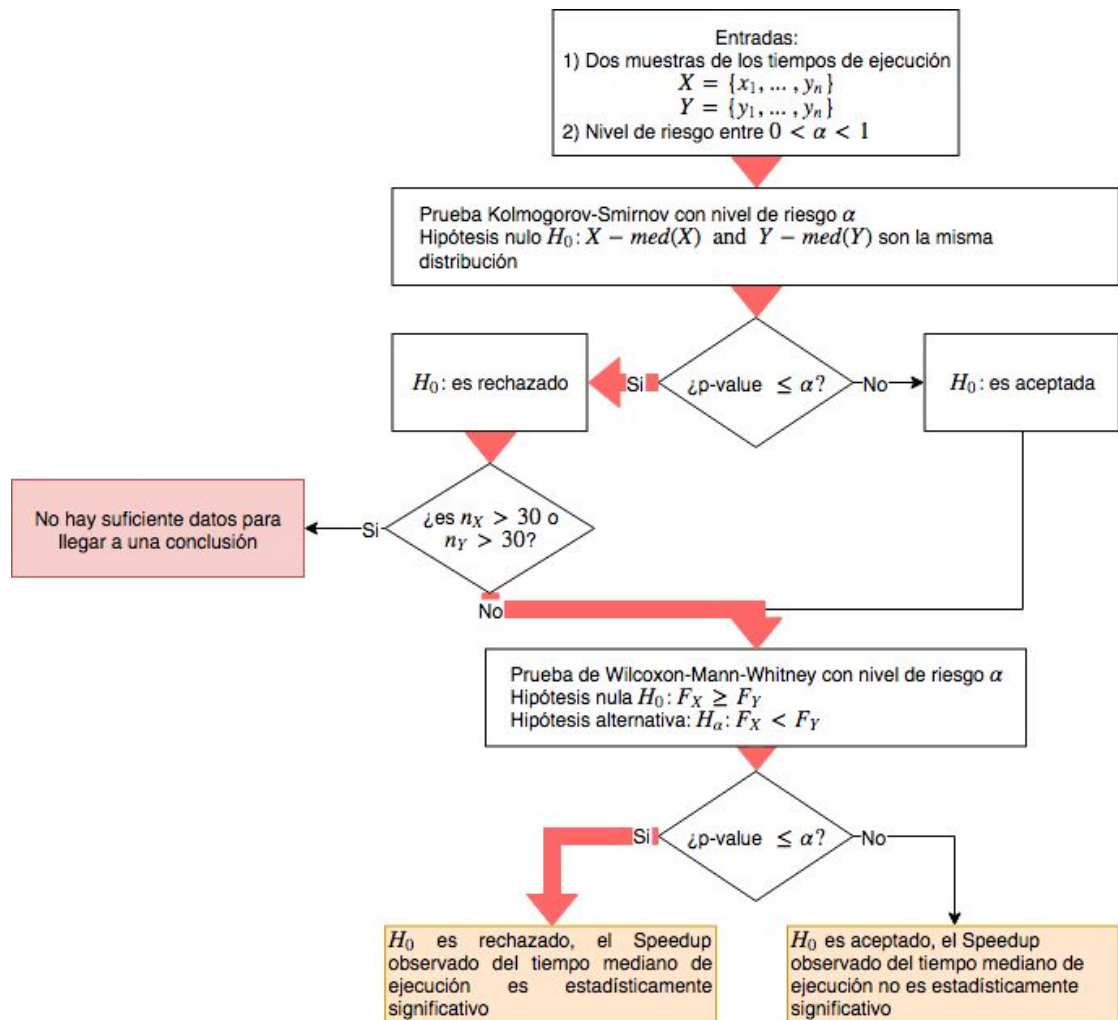
“si el valor p es inferior al nivel de significación, entonces la hipótesis nula es rechazada”

p-value: 2.2e-16

$\alpha$ : 0.05



# Resumen: Análisis de medianas



---

# Objetivos

- Análisis de media por medio de una prueba t de Student's
  - Análisis de mediana por medio de una prueba Wilcoxon-Mann-Whitney
  - **Speedup**
-

# Speedup

“Fracción de aceleración de un programa  $a$  sobre un programa  $b$ ”

$$S = \frac{a}{b}$$

Siguiendo con el ejemplo se obtiene:

- Speedup obtenido usando media: 1.5229
- Speedup obtenido usando mediana: 1.5228



—

**Fin**