



# Prueba $T$ y $Z$

Inferencia estadística

Instructor: Juan Luis Palacios Soto

palacios.s.j.l@gmail.com



scidata  
Matemáticas para  
la ciencia de datos

- 1 Concepto de inferencia estadística
- 2 La Prueba  $T$
- 3 La Prueba  $Z$
- 4 La prueba  $T$  apareada

scidata  
matemáticas para  
la ciencia de datos

## Definición (Inferencia Estadística)

*Es una rama de la Estadística encargada de hacer generalizaciones, predicciones, inducciones a partir del análisis muestral.*

$$\bar{x} \approx \mu$$

$$S^2 \approx \sigma^2$$

$$S \approx \sigma$$

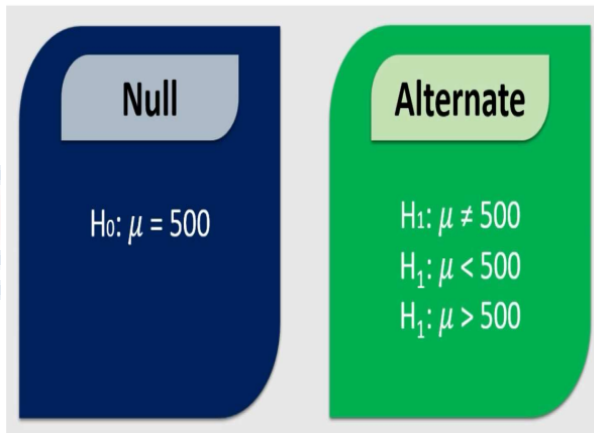


## Uso de la prueba $T$

Se trata de una prueba estadística para demostrar una hipótesis sobre la media poblacional. A saber:

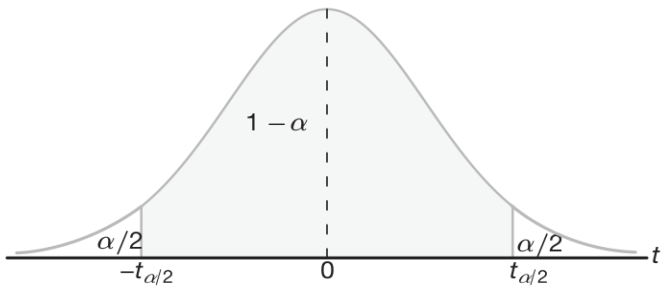
- Se utiliza la prueba  $T$  de una muestra para averiguar si la media poblacional toma o no cierto valor. Es decir, se toma como hipótesis nula el enunciado " $H_0$ : la media poblacional real vale A" ( $H_0 : \mu = \mu_0$ ).
- Se utiliza cuando la muestra proviene de una **distribución normal**, hay **menos de 30** elementos y se **desconoce la desviación estándar poblacional**  $\sigma$ .
- También se puede aplicar cuando el tamaño de la muestra es mayor o igual a 30, pero se sigue desconociendo la desviación estándar poblacional. El supuesto de normalidad se cumple por el Teorema del Límite Central
- En general la prueba  $T$  tiene colas más pesadas.

Ver Geogebra.



Los pasos para la aplicación de la prueba  $T$  son:

- 1 Seleccionar el nivel de significación. Este es denotado por  $\alpha$ . Generalmente  $\alpha = 0.01$ ,  $\alpha = 0.05$  o  $\alpha = 0.10$ . Esto representa la probabilidad de rechazar la hipótesis nula cuando es verdadera (Error Tipo I). Por ejemplo, un nivel de significación de 0.05 indica un riesgo del 5 % de concluir que existe una diferencia entre los resultados del estudio y la hipótesis nula cuando en realidad no hay ninguna diferencia. Al número  $1 - \alpha$  se le llama **nivel de confianza**.
- 2 Encontrar el **valor crítico**. Este es denotado por  $t_{n-1,\alpha}$ . Antigüamente se usaba una tabla de  $T$ . Se seleccionaba la columna basado en  $\alpha$  y la fila basado en los grados de libertad, que en este caso es  $n - 1$  (siendo  $n$  el tamaño de la muestra).
- 3 Calcular el valor  $t$ . Se refiere a calcular el número  $t = \frac{\bar{x} - \mu}{S/\sqrt{n}}$ , donde  $\bar{x}$  es la media muestral,  $\mu$  la media poblacional desconocida,  $S$  la desviación estándar muestral y  $n$  el tamaño de la muestra.
- 4 Comparar y decidir. Si  $|t| \geq t_{n-1,\alpha}$ , rechazamos  $H_0$ . En caso contrario, “aceptamos”  $H_0$  (porque en realidad diríamos que no hay evidencia suficiente en la muestra para rechazar la hipótesis nula).





# Prueba $T$ en Excel

Excel (Error de activación de productos)

Inicio Fórmulas Datos Revisar Vista Desarrollador ¿Qué desea hacer?

Conexiones Propiedades Editar vínculos Conexiones

Ordenar Filtro Avanzadas Ordenar y filtrar

Reemplazo de texto Rellenar rápido Consolidar Relaciones Administración de datos Herramientas de datos

Validación de datos Administración de datos

Previsión

Esquema

Análisis de datos

Country Name

	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
de	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	5.106	3.061	0.491	1.726				
	1.399	2.433	8.801	9.566	4.855	9.837	3.184	3.192	-5.243	1.603	6.93	4.485	5.053	4.284								
	1.931	4.001	2.986	4.056	3.204	2.795	3.844	3.658	1.937	2.067	2.463	3.918	2.585	2.533								
	1.1	1.707	1.038	3.571	2.322	2.552	3.677	0.447	-2.021	2.864	1.694	0.739	0.459	1.579								
	3.77	5.938	5.156	6.435	7.124	6.874	7.344	6.072	-3.586	1.374	1.915	0.031	0.494	1.838								
2																						

Funciones para análisis

- Análisis de Fourier
- Histograma
- Media móvil
- Generación de números aleatorios
- Jerarquía y percentil
- Regresión
- Muestra
- Prueba t para medias de dos muestras emparejadas
- Prueba t para dos muestras suponiendo varianzas iguales
- Prueba t para dos muestras suponiendo varianzas desiguales

Prueba t para dos muestras suponiendo varianzas desiguales

Entrada

Rango para la variable 1:

Rango para la variable 2:

Diferencia hipotética entre las medias:

Alfa: 0.05

Opciones de salida

☐ Rango de salida:

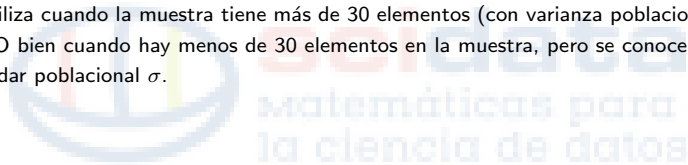
☒ En una hoja nueva:

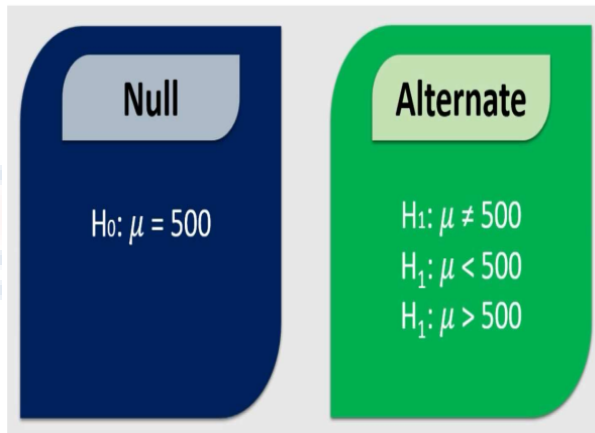
☐ En un libro nuevo

## Uso de la prueba $Z$

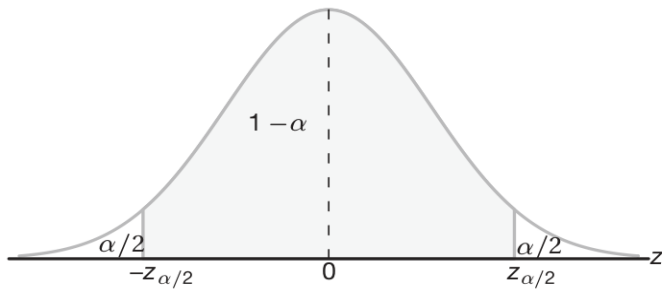
Se trata de una prueba estadística para demostrar una hipótesis sobre la media poblacional. Un supuesto es que la muestra tomada sigue una distribución normal.

- 1 Se utiliza la prueba  $Z$  de una muestra para averiguar si la media poblacional toma o no cierto valor. Es decir, se toma como hipótesis nula el enunciado ' $H_0$  : la media poblacional real vale  $A$ '.
- 2 Se utiliza cuando la muestra tiene más de 30 elementos (con varianza poblacional conocida o no). O bien cuando hay menos de 30 elementos en la muestra, pero se conoce la desviación estándar poblacional  $\sigma$ .

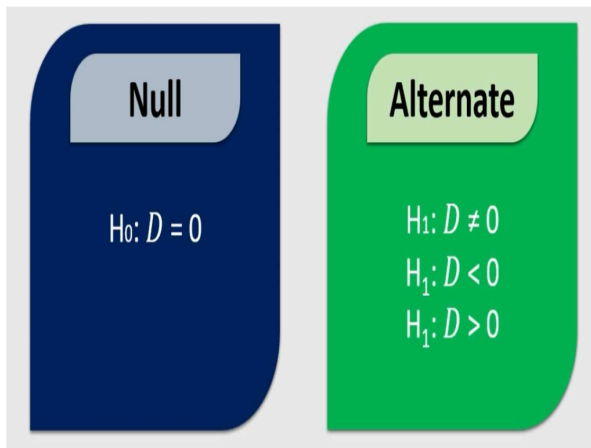




- **Seleccionar el nivel de significación.** Este es denotado por  $\alpha$ . Generalmente  $\alpha = 0.01$ ,  $\alpha = 0.05$  o  $\alpha = 0.10$ . Esto representa la probabilidad de rechazar la hipótesis nula cuando es verdadera. Por ejemplo, un nivel de significación de 0.05 indica un riesgo del 5 % de concluir que existe una diferencia entre los resultados del estudio y la hipótesis nula cuando en realidad no hay ninguna diferencia. Al número  $1 - \alpha$  se le llama nivel de confianza.
- **Encontrar el valor crítico.** Este es denotado por  $z_{\alpha}$ . Antiguamente se usaba una tabla de  $Z$ . Se seleccionaba la columna basado en  $\alpha$ .
- **Calcular un parámetro.** Se refiere a calcular el número  $Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$  con los datos de la muestra.
- **Comparar y decidir.** Si  $|Z| \geq z_{\alpha}$ , rechazamos  $H_0$ . En caso contrario, diremos que no hay suficiente evidencia para rechazar la hipótesis nula a partir de la información muestral.



[illegible]



Los pasos para la aplicación de la prueba  $T$  pareada son:

- **Seleccionar el nivel de significación.** Este es denotado por  $\alpha$ . Generalmente  $\alpha = 0.01$ ,  $\alpha = 0.05$  o  $\alpha = 0.10$ . Esto representa la probabilidad de rechazar la hipótesis nula cuando es verdadera.
- **Encontrar el valor crítico.** Este es denotado por  $t_{n-1,\alpha}$ .
- **Calcular un parámetro.** Se refiere a calcular el número  $T = \frac{\bar{d} - D}{S_{diff}/\sqrt{n}}$  con los datos de la muestra.
- **Comparar y decidir.** Si  $|T| \geq t_{n-1,\alpha}$ , rechazamos  $H_0$ . En caso contrario, diremos que no hay suficiente evidencia para rechazar la hipótesis nula a partir de la información muestral.



Diseño de página   Fórmulas   Datos   Revisar   Vista   Desarrollador   ¿Qué desea hacer?

ar consultas   Conexiones   Actualizar todo   Propiedades   Editar vínculos   Conexiones   Ordenar y filtrar   Ordenar   Filtrar   Avanzadas   Texto en columnas   Rellenar rápido   Quitar duplicados   Validación de datos   Herramientas de datos   Consolidar   Relaciones   Administrar modelo de datos   Previsión   Análisis de datos   Agrupar   Desagrupar   Subtotal   Esquema   Análisis

fx   Country Name

	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
Country Code	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018					
AE	40.71	43.57	46.38	53.05	51.97	50.84	64.41	69.65	73.81	61.64	61.59	64	64.73	68.93	74.35	75.74	76.99	67.98					
AL	22.14	20.78	21.2	19.92	21.01	21.71	21.85	22.71	22.79	20.84	20.47	21.73	21.25	21.47	21.51	21.53	20.58	21.39					
AR	67.96	65.25	63.68	66.01	70.25	72.76	74.17	80.21	66.57	74.15	80.8	80.34	78.53										
BR	44.54	41.95	44.94	52.51	57.63	64.54	71.21	72.3	50.61	53.03	58.69	63.99	65.32										
IS	38.43	34.1	35.25	37.79	42.64	49.24	47.92	47.94	42.33	43.72	50.9	46.12	45.53										




Análisis de datos   ?   X

Funciones para análisis

- Suavización exponencial
- Prueba F para varianzas de dos muestras
- Análisis de Fourier
- Histograma
- Media móvil
- Generación de números aleatorios
- Jerarquía y percentil
- Regresión
- Muestra
- Prueba t para medias de dos muestras emparejadas

Aceptar   Cancelar   Ayuda

Prueba t para medias de dos muestras emparejadas   ?   X

Entrada  
 Rango para la variable 1:    
 Rango para la variable 2:    
 Diferencia hipotética entre las medias:   
☐ Botulos  
 Alfa:  0.05  
 Opciones de salida  
☐ Rango de salida:    
☒ En una hoja nueva:   
☐ En un libro nuevo

Aceptar   Cancelar   Ayuda

	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218					
V	24.39	23.65	22.85	23.42	24.7	26.99	27.7	27.63	23.02	26.93	28.27	27.28	26.23										
N	9.569	9.667	9.944	10.95	12.5	14.46	15.6	16.97	11.97	13.58	15.47	16.09	18.23	20.01	18.03	15.26	16.84	18.19					
A	43.57	41.42	40.68	44.15	41.27	41.13	39.49	38.53	27.82	26.81	34.09	31.68	28.94	29.16	28.55	28.46	28.86	30.13					