



Ejemplo

En 15 estanques del Estado de México se quiere medir la eficacia de los humedales para disminuir el pH, se tomaron mediciones de pH antes de su colocación, después de un año de su colocación se tomaron mediciones de pH:

Puntos	pH abril del 2020	pH abril 2021
1	8.32	7.21
2	6.43	6.53
3	8.67	6.77
4	7.28	6.05
5	8.59	7.36
6	7.15	6.45
7	7.85	5.59
8	6.48	6.33
9	8.95	6.28
10	8.45	7.23
11	8.97	8.01
12	6.77	6.01
13	8.57	8.06
14	7.99	6.55
15	8.72	7.31

Tamaño de la muestra:

$$n_1 = 15$$

$$n_2 = 15$$

Tamaño de la muestra (diferencia):

$$n_d = 15$$

a. Establecer hipótesis

$H_0: \bar{d} = 0$ La diferencia promedio es igual a cero (NO hay diferencia entre las medidas de un año a otro)

$H_1: \bar{d} \neq 0$ La diferencia promedio es diferente de cero (SI hay diferencia entre las medidas de un año a otro)

b. Dibujar la región de rechazo o no rechazo.

Es una prueba de hipótesis de **dos colas** por lo que el alfanúmero es $\frac{\alpha}{2}$

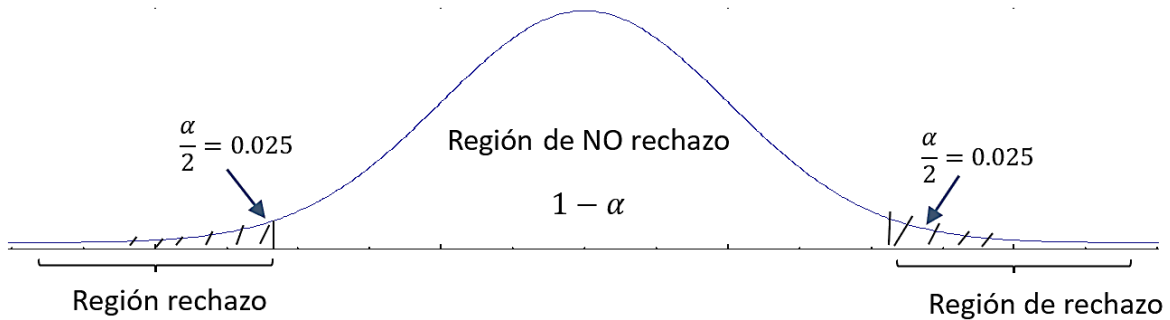


Fig. 1 Región de rechazo y no rechazo para una prueba bilateral con $\frac{\alpha}{2} = 0.025$

c. Calcular la diferencia de $x_1 - x_2 = d$ y calcular el promedio de la diferencia

Puntos	pH 04.2020 x_1	pH 04.2021 x_2	$x_1 - x_2 = d$
1	8.32	7.21	$(8.32-7.21) = 1.11$
2	6.43	6.53	$(6.43-6.53) = -0.1$
3	8.67	6.77	$(8.67-6.77) = 1.9$
4	7.28	6.05	$(7.28-6.05) = 1.23$
5	8.59	7.36	$(8.59-7.36) = 1.23$
6	7.15	6.45	$(7.15-6.45) = 0.7$
7	7.85	5.59	$(7.85-5.59) = 2.26$
8	6.48	6.33	$(6.48-6.33) = 0.15$
9	8.95	6.28	$(8.95-6.28) = 2.67$
10	8.45	7.23	$(8.45-7.23) = 1.22$
11	8.97	8.01	$(8.97-8.01) = 0.96$
12	6.77	6.01	$(6.77-6.01) = 0.76$
13	8.57	8.06	$(8.57-8.06) = 0.51$
14	7.99	6.55	$(7.99-6.55) = 1.44$
15	8.72	7.31	$(8.72-7.31) = 1.41$
			$\bar{d} = 1.16333333$

d. Calcular desviación estándar

$$Sd = \sqrt{\frac{\sum(d-\bar{d})^2}{n_d-1}} = \sqrt{\frac{7.63613333}{15-1}} = 0.738537$$

e. Obtener estadístico de prueba

$$t_{cal} = \frac{\frac{\bar{d}}{Sd}}{\frac{\sqrt{n_d}}{\sqrt{15}}} = \frac{1.1633333}{0.738537} = 6.100670126 \approx 6.1006$$

Calcular valor teórico en tablas

$$t_{teo} = t_{(gl, \frac{\alpha}{2})} = t_{(n_d-1, 0.025)} = 2.624$$

Nota: para buscar el valor teórico, se requieren tablas de distribución, éstas son cálculos que representan valores estandarizados. Por lo general los grados de libertad se encuentran en la parte izquierda de la tabla y se representan con una ν , mientras que la significancia se encuentra en la parte superior representados por α o bien $\frac{\alpha}{2}$, según sea el caso. Una vez se conozcan las tablas que se van a utilizar, se busca la intersección entre los grados de libertad (el valor que se calculó con la fórmula) con el valor de α .

ν	0.1	0.05	0.025	0.01	0.005	α
	0.20	0.10	0.05	0.02	0.01	$\frac{\alpha}{2}$
1	3.078	6.314	12.706	31.821	63.657	
2	1886	2.920	4.303	6.965	9.9925	
...						
14	1.345	1.761	2.145	2.624	2.977	
...						
∞	1.282	1.645	1.960	2.326	2.576	

f. Toma de decisión

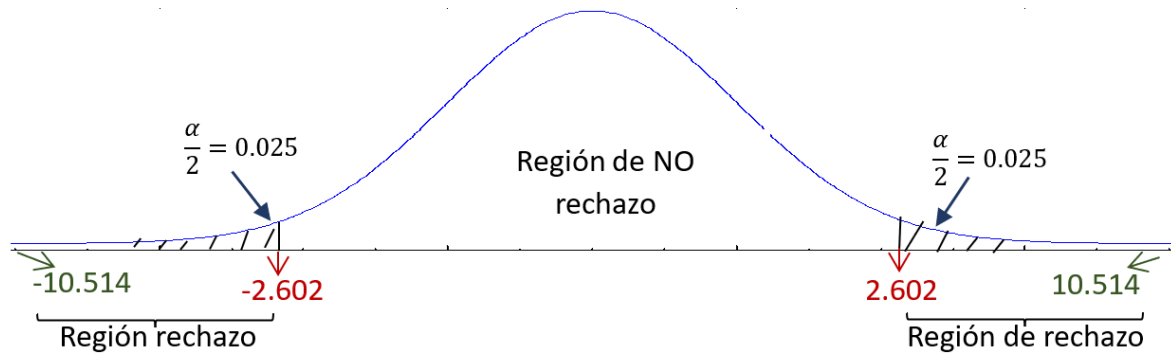


Fig. 2 La t calculada cae en las regiones de rechazo.

Entonces, como $t_{cal} = 6.100670 > t_{teo} = 2.624$ se rechaza la hipótesis nula.

g. Concluir texto del problema

Con un 95% de confianza se puede afirmar que existe diferencia significativa entre el promedio de las diferencias de pH del año 2020 y del 2021.