

Test de Hipótesis para
dos muestras



Test de Hipótesis para dos Muestras

1. Test para la diferencia de dos medias poblaciones, **muestras grandes e independientes**.
2. Test para la diferencia de dos medias poblaciones, **muestras pequeñas e independientes**, varianzas desconocidas.
3. Test para la diferencia de dos medias poblaciones, **datos apareados**.
4. Test para la diferencia entre dos **proporciones**, muestras independientes.

1. Se tiene la siguiente prueba de hipótesis: $H_0: \mu_1 - \mu_2 = 0$ vs. $H_A: \mu_1 - \mu_2 \neq 0$. Los resultados siguientes son para dos muestras independientes tomadas de dos poblaciones.

Muestra 1	Muestra 2
$n_1 = 110$	$n_2 = 120$
$\bar{x}_1 = 104$	$\bar{x}_2 = 106$
$\sigma_1 = 8.4$	$\sigma_2 = 7.6$

$$\alpha = 0.05$$

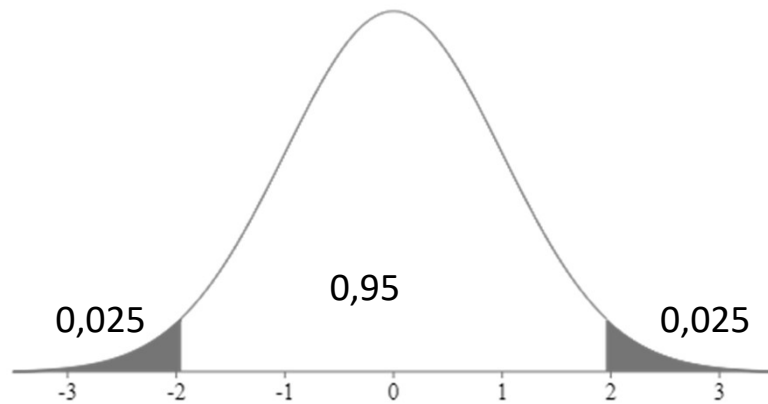
1. Se tiene la siguiente prueba de hipótesis: $H_0: \mu_1 - \mu_2 = 0$ vs. $H_A: \mu_1 - \mu_2 \neq 0$. Los resultados siguientes son para dos muestras independientes tomadas de dos poblaciones.

Muestra 1	Muestra 2
$n_1 = 110$	$n_2 = 120$
$\bar{x}_1 = 104$	$\bar{x}_2 = 106$
$\sigma_1 = 8.4$	$\sigma_2 = 7.6$

$$\alpha = 0.05$$

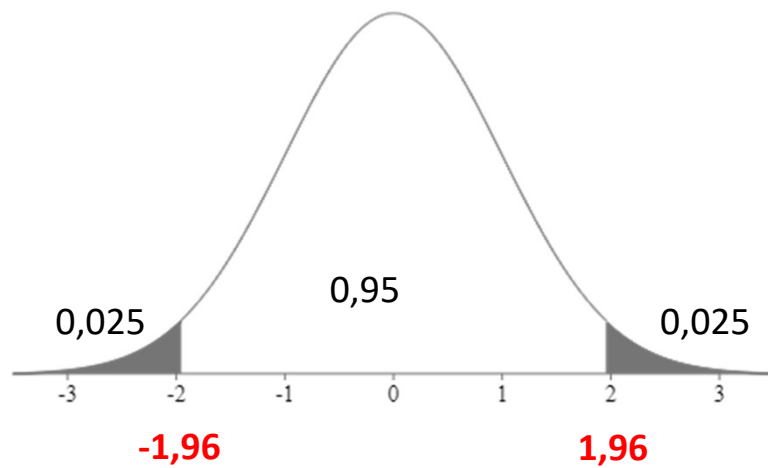
Como las muestras son grandes, $\bar{x} - \bar{y}$ se distribuye de manera aproximadamente, con

$$E(\bar{x} - \bar{y}) = \mu_x - \mu_y \qquad \sigma_{\bar{x} - \bar{y}} = \sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}$$



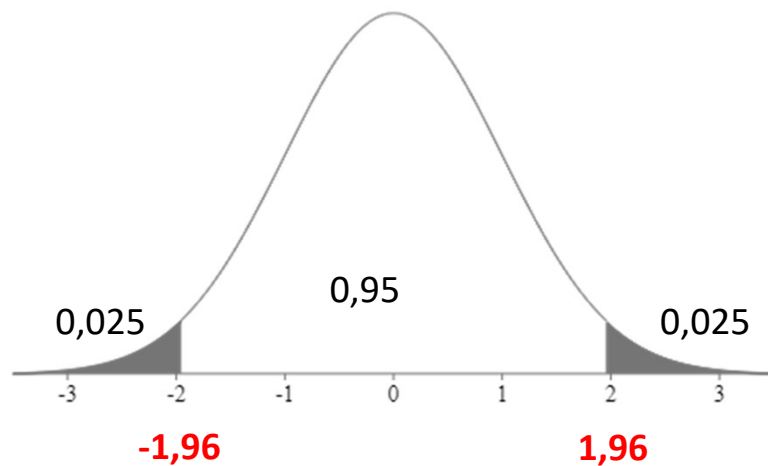
Si $-z^c \leq z^{obs} \leq z^c$
no tengo evidencia para rechazar H_o

Si $z^{obs} < -z^c$ o $z^{obs} > z^c$
Tengo evidencia suficiente para rechazar
 H_o con una confianza del 95%.



Si $-z^c \leq z^{obs} \leq z^c$
no tengo evidencia para rechazar H_o

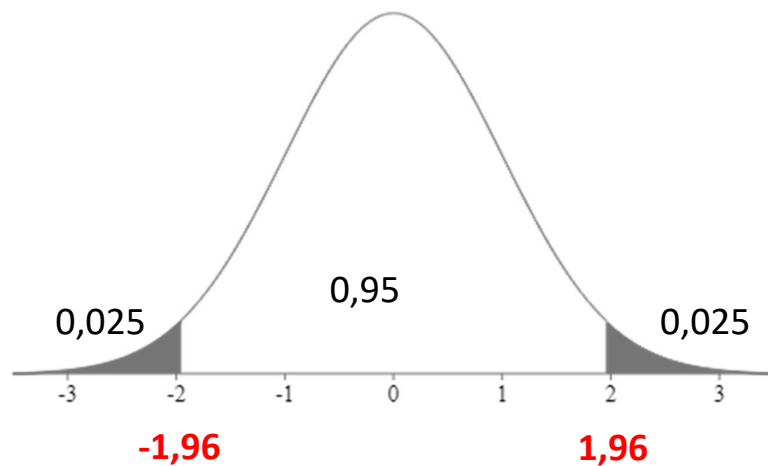
Si $z^{obs} < -z^c$ o $z^{obs} > z^c$
Tengo evidencia suficiente para rechazar
 H_o con una confianza del 95%.



Si $-z^c \leq z^{obs} \leq z^c$
no tengo evidencia para rechazar H_o

Si $z^{obs} < -z^c$ o $z^{obs} > z^c$
Tengo evidencia suficiente para rechazar H_o con una confianza del 95%.

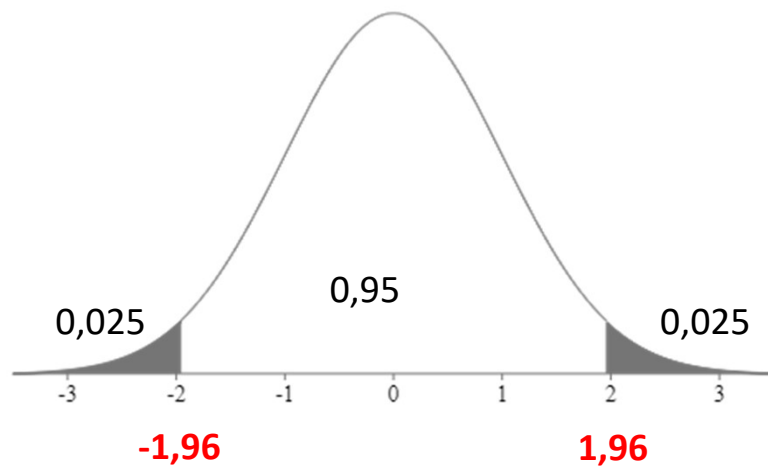
$$z^{obs} = \frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}}$$



Si $-z^c \leq z^{obs} \leq z^c$
no tengo evidencia para rechazar H_o

Si $z^{obs} < -z^c$ o $z^{obs} > z^c$
Tengo evidencia suficiente para rechazar H_o con una confianza del 95%.

$$z^{obs} = \frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}} = \frac{(104 - 106) - 0}{\sqrt{\frac{8,4^2}{110} + \frac{7,6^2}{120}}}$$



Si $-z^c \leq z^{obs} \leq z^c$
no tengo evidencia para rechazar H_o

Si $z^{obs} < -z^c$ o $z^{obs} > z^c$
Tengo evidencia suficiente para rechazar H_o con una confianza del 95%.

$$z^{obs} = \frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}} = \frac{(104 - 106) - 0}{\sqrt{\frac{8,4^2}{110} + \frac{7,6^2}{120}}} = -1,89$$

Valor p =

$$z^{obs} = \frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}} = \frac{(104 - 106) - 0}{\sqrt{\frac{8,4^2}{110} + \frac{7,6^2}{120}}} = -1,89$$

$$\text{Valor } p = P(Z < -1,89) + P(Z > 1,89)$$

$$z^{obs} = \frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}} = \frac{(104 - 106) - 0}{\sqrt{\frac{8,4^2}{110} + \frac{7,6^2}{120}}} = -1,89$$

$$\text{Valor } p = P(Z < -1,89) + P(Z > 1,89) = 2 * P(Z > 1,89)$$

$$z^{obs} = \frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}} = \frac{(104 - 106) - 0}{\sqrt{\frac{8,4^2}{110} + \frac{7,6^2}{120}}} = -1,89$$

$$\text{Valor } p = P(Z < -1,89) + P(Z > 1,89) = 2 * P(Z > 1,89) =$$

$$2 * (1 - P(Z < 1,89))$$

$$z^{obs} = \frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}} = \frac{(104 - 106) - 0}{\sqrt{\frac{8,4^2}{110} + \frac{7,6^2}{120}}} = -1,89$$

$$\begin{aligned} \text{Valor } p &= P(Z < -1,89) + P(Z > 1,89) = 2 * P(Z > 1,89) = \\ &2 * (1 - P(Z < 1,89)) = 2 * (1 - 0,9706) \end{aligned}$$

$$z^{obs} = \frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}} = \frac{(104 - 106) - 0}{\sqrt{\frac{8,4^2}{110} + \frac{7,6^2}{120}}} = -1,89$$

$$\text{Valor } p = P(Z < -1,89) + P(Z > 1,89) = 2 * P(Z > 1,89) =$$

$$2 * (1 - P(Z < 1,89)) = 2 * (1 - 0,9706) = \mathbf{0,0588}$$

$$z^{obs} = \frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}} = \frac{(104 - 106) - 0}{\sqrt{\frac{8,4^2}{110} + \frac{7,6^2}{120}}} = \mathbf{-1,89}$$

$$\begin{aligned} \text{Valor } p &= P(Z < -1,89) + P(Z > 1,89) = 2 * P(Z > 1,89) = \\ &2 * (1 - P(Z < 1,89)) = 2 * (1 - 0,9706) = \mathbf{0,0588} \end{aligned}$$

Valor $p > \alpha \rightarrow$ no tengo evidencia para rechazar H_0

$$z^{obs} = \frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}} = \frac{(104 - 106) - 0}{\sqrt{\frac{8,4^2}{110} + \frac{7,6^2}{120}}} = \mathbf{-1,89}$$

2. Se tiene la siguiente prueba de hipótesis: $H_0: \mu_1 - \mu_2 \leq 0$ vs. $H_A: \mu_1 - \mu_2 > 0$. Los resultados siguientes son para dos muestras independientes tomadas de dos poblaciones normales.

Muestra 1

Muestra 2

$$n_1 = 40$$

$$n_2 = 50$$

$$\bar{x}_1 = 25,2$$

$$\bar{x}_2 = 22,8$$

$$s_1 = 5,2$$

$$s_2 = 6,0$$

$$\alpha = 0.05$$

2. Se tiene la siguiente prueba de hipótesis: $H_0: \mu_1 - \mu_2 \leq 0$ vs. $H_A: \mu_1 - \mu_2 > 0$. Los resultados siguientes son para dos muestras independientes tomadas de dos poblaciones normales.

Muestra 1	Muestra 2
$n_1 = 40$	$n_2 = 50$
$\bar{x}_1 = 25,2$	$\bar{x}_2 = 22,8$
$s_1 = 5,2$	$s_2 = 6,0$

$$\alpha = 0.05$$

Como las muestras son chicas y la varianza desconocida, hay que usar la t de Student. Se necesita que las poblaciones sean Normales.

2. Se tiene la siguiente prueba de hipótesis: $H_0: \mu_1 - \mu_2 \leq 0$ vs. $H_A: \mu_1 - \mu_2 > 0$. Los resultados siguientes son para dos muestras independientes tomadas de dos poblaciones normales.

Muestra 1	Muestra 2
$n_1 = 40$	$n_2 = 50$
$\bar{x}_1 = 25,2$	$\bar{x}_2 = 22,8$
$s_1 = 5,2$	$s_2 = 6,0$

$$\alpha = 0.05$$

Como las muestras son chicas y la varianza desconocida, hay que usar la t de Student. Se necesita que las poblaciones sean Normales.

También hay que asumir que ambas poblaciones tienen la misma varianza y se construye un estimador de la varianza poblacional:

$$s_p^2 = \frac{(n_x - 1)s_x^2 + (n_y - 1)s_y^2}{(n_x + n_y - 2)}$$

2. Se tiene la siguiente prueba de hipótesis: $H_0: \mu_1 - \mu_2 \leq 0$ vs. $H_A: \mu_1 - \mu_2 > 0$. Los resultados siguientes son para dos muestras independientes tomadas de dos poblaciones normales.

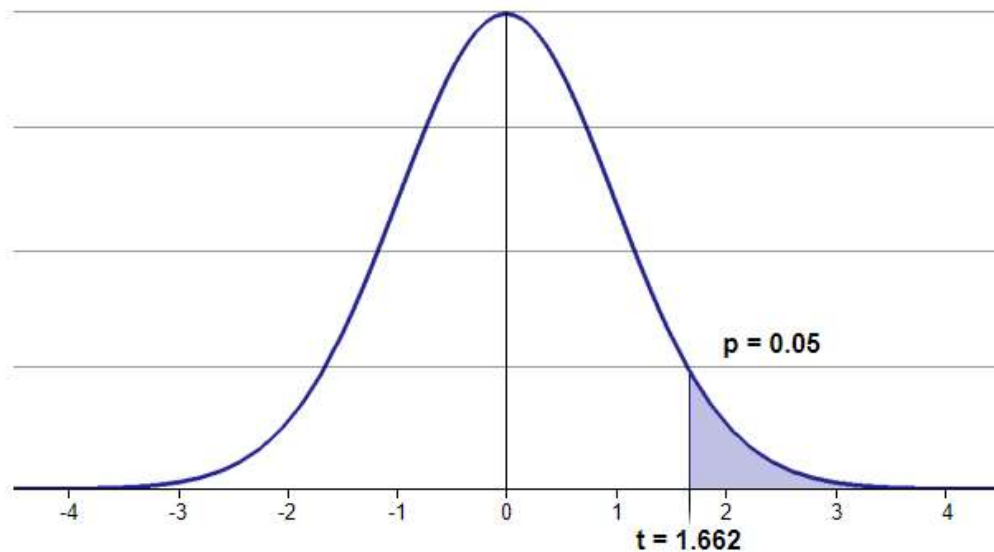
Muestra 1	Muestra 2
$n_1 = 40$	$n_2 = 50$
$\bar{x}_1 = 25,2$	$\bar{x}_2 = 22,8$
$s_1 = 5,2$	$s_2 = 6,0$

$$\alpha = 0.05$$

Como las muestras son chicas y la varianza desconocida, hay que usar la t de Student. Se necesita que las poblaciones sean Normales.

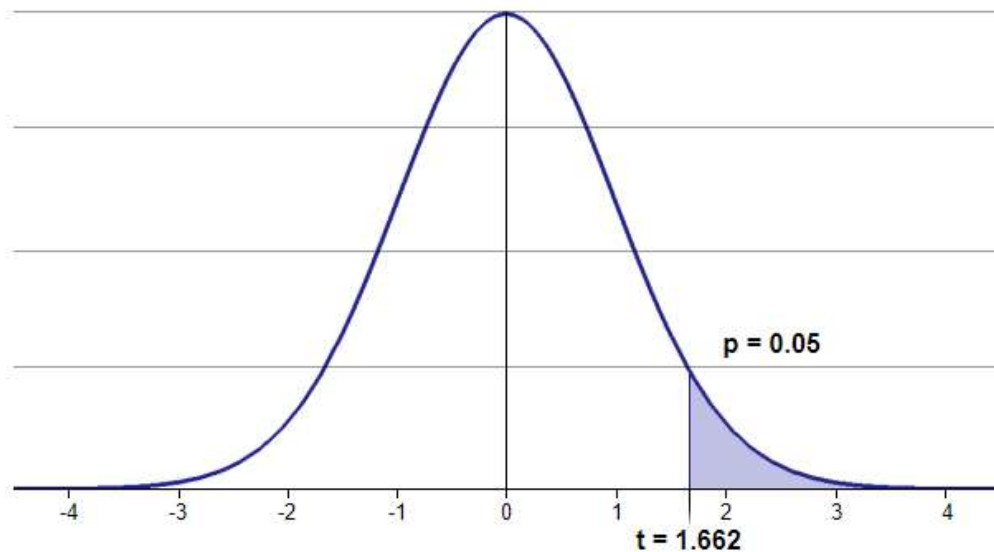
También hay que asumir que ambas poblaciones tienen la misma varianza y se construye un estimador de la varianza poblacional:

$$s_p^2 = \frac{(n_x - 1)s_x^2 + (n_y - 1)s_y^2}{(n_x + n_y - 2)} = \frac{(40 - 1)5,2^2 + (50 - 1)6^2}{(40 + 50 - 2)} = 32,029$$



Si $t^{obs} \leq t_{88,0,05}$
no tengo evidencia para rechazar H_0

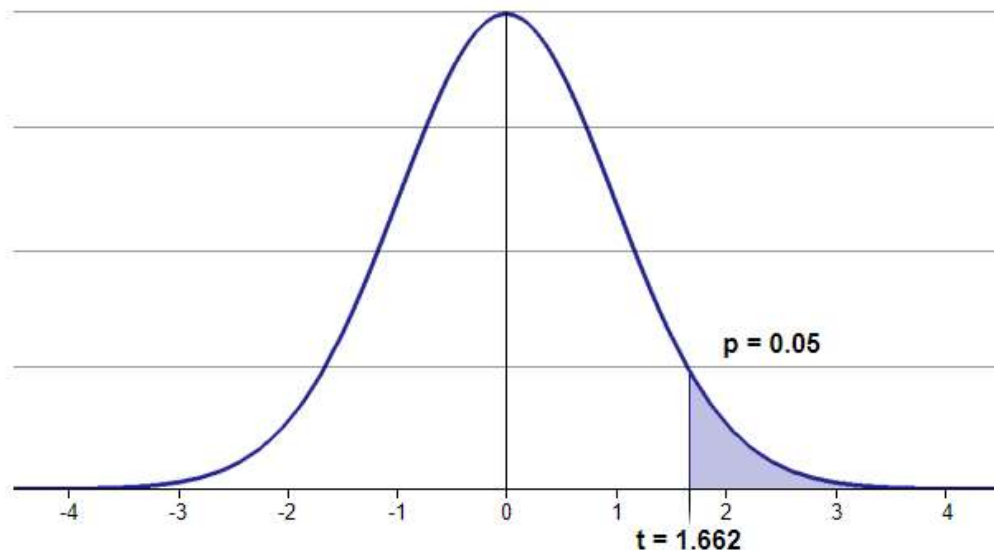
Si $t^{obs} > t_{88,0,05}$
Tengo evidencia suficiente para rechazar H_0 con una confianza del 95%.



Si $t^{obs} \leq t_{88,0,05}$
no tengo evidencia para rechazar H_o

Si $t^{obs} > t_{88,0,05}$
Tengo evidencia suficiente para rechazar H_o con una confianza del 95%.

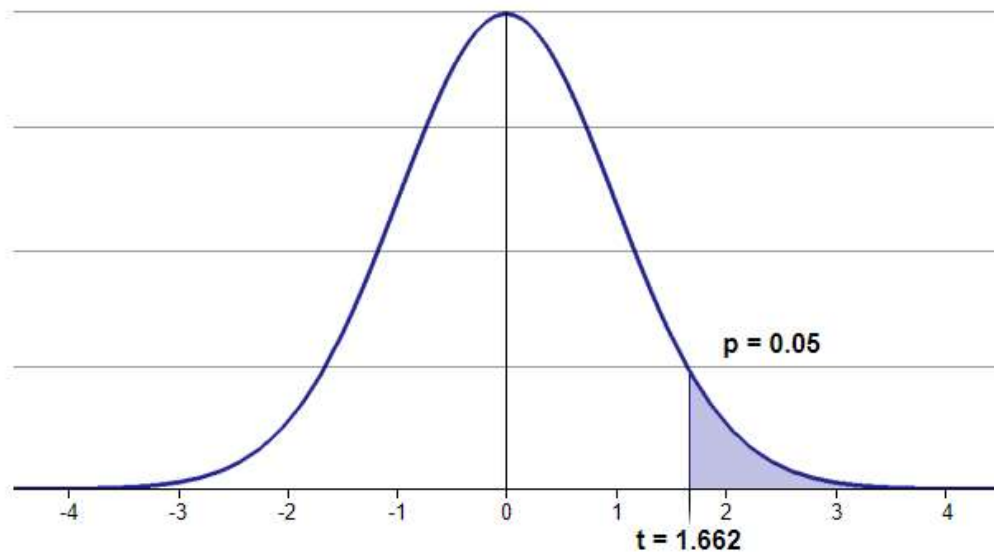
$$t^{obs} = \frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{\sqrt{\frac{s_p^2}{n_x} + \frac{s_p^2}{n_y}}}$$



Si $t^{obs} \leq t_{88,0,05}$
no tengo evidencia para rechazar H_o

Si $t^{obs} > t_{88,0,05}$
Tengo evidencia suficiente para rechazar H_o con una confianza del 95%.

$$t^{obs} = \frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{\sqrt{\frac{s_p^2}{n_x} + \frac{s_p^2}{n_y}}} = \frac{(25,2 - 22,8) - 0}{\sqrt{\frac{32,029}{40} + \frac{32,029}{50}}}$$



Si $t^{obs} \leq t_{88,0,05}$
no tengo evidencia para rechazar H_o

Si $t^{obs} > t_{88,0,05}$
Tengo evidencia suficiente para rechazar H_o con una confianza del 95%.

$$t^{obs} = \frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{\sqrt{\frac{s_p^2}{n_x} + \frac{s_p^2}{n_y}}} = \frac{(25,2 - 22,8) - 0}{\sqrt{\frac{32,029}{40} + \frac{32,029}{50}}} = 2,00$$

Valor $p =$

$$t^{obs} = \frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{\sqrt{\frac{s_p^2}{n_x} + \frac{s_p^2}{n_y}}} = \frac{(25,2 - 22,8) - 0}{\sqrt{\frac{32,029}{40} + \frac{32,029}{50}}} = 2,00$$

Valor $p = P(t_{88} > 2,00)$

$$t^{obs} = \frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{\sqrt{\frac{s_p^2}{n_x} + \frac{s_p^2}{n_y}}} = \frac{(25,2 - 22,8) - 0}{\sqrt{\frac{32,029}{40} + \frac{32,029}{50}}} = 2,00$$

$$\text{Valor } p = P(t_{88} > 2,00)$$

$$\text{Por tabla: } P(t_{90} < 1,9867) = 0,975 \quad y \quad P(t_{90} < 2,3685) = 0,990$$

$$t^{obs} = \frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{\sqrt{\frac{s_p^2}{n_x} + \frac{s_p^2}{n_y}}} = \frac{(25,2 - 22,8) - 0}{\sqrt{\frac{32,029}{40} + \frac{32,029}{50}}} = 2,00$$

$$\text{Valor } p = P(t_{88} > 2,00)$$

$$\text{Por tabla: } P(t_{90} < 1,9867) = 0,975 \quad y \quad P(t_{90} < 2,3685) = 0,990$$

$$P(t_{90} > 1,9867) = 0,025 \quad y \quad P(t_{90} > 2,3685) = 0,001$$

$$t^{obs} = \frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{\sqrt{\frac{s_p^2}{n_x} + \frac{s_p^2}{n_y}}} = \frac{(25,2 - 22,8) - 0}{\sqrt{\frac{32,029}{40} + \frac{32,029}{50}}} = 2,00$$

$$\text{Valor } p = P(t_{88} > 2,00)$$


$$\text{Por tabla: } P(t_{90} < 1,9867) = 0,975 \quad y \quad P(t_{90} < 2,3685) = 0,990$$

$$P(t_{90} > 1,9867) = 0,025 \quad y \quad P(t_{90} > 2,3685) = 0,001$$

$$0,001 < \text{Valor } p < 0,025 < \alpha$$

Tengo evidencia suficiente para rechazar H_0

$$t^{obs} = \frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{\sqrt{\frac{s_p^2}{n_x} + \frac{s_p^2}{n_y}}} = \frac{(25,2 - 22,8) - 0}{\sqrt{\frac{32,029}{40} + \frac{32,029}{50}}} = 2,00$$



7. Le han pedido que averigüe si dos procesos de producción diferentes producen una media diferente de unidades por hora. El proceso 1 tiene una media μ_1 y el proceso 2 una media μ_2 . La hipótesis nula y la hipótesis alternativa son: $H_0: \mu_1 - \mu_2 \geq 0$ vs. $H_1: \mu_1 - \mu_2 < 0$. Utilizando una muestra aleatoria de 25 observaciones apareadas, la desviación típica de la diferencia entre las medias muestrales es 25. ¿Puede rechazar la hipótesis nula utilizando una probabilidad de cometer el error de tipo I con $\alpha=0.05$, en cada uno de los siguientes casos?

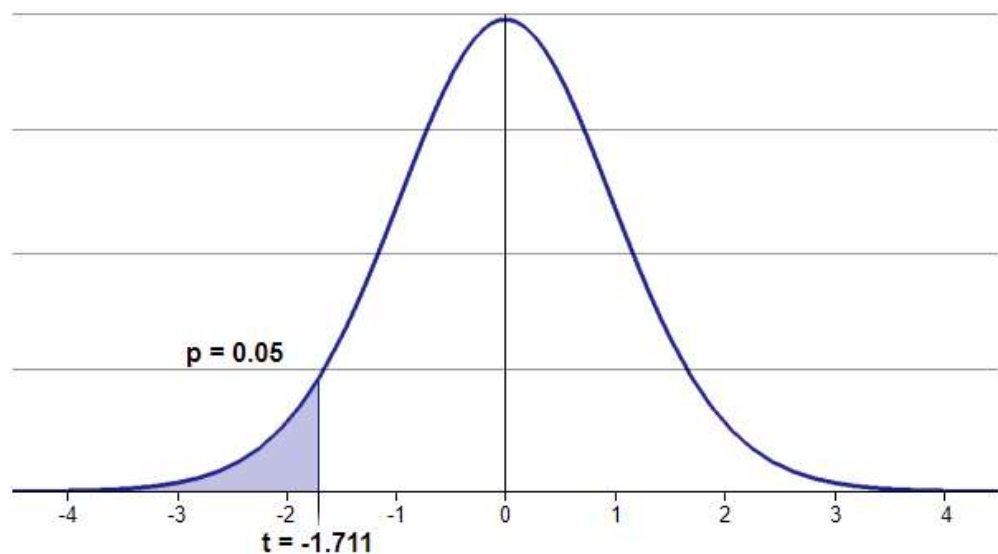
- a) Las medias muestrales son 56 y 50
- b) Las medias muestrales son 59 y 60
- c) Las medias muestrales son 48 y 57

7. Le han pedido que averigüe si dos procesos de producción diferentes producen una media diferente de unidades por hora. El proceso 1 tiene una media μ_1 y el proceso 2 una media μ_2 . La hipótesis nula y la hipótesis alternativa son: $H_0: \mu_1 - \mu_2 \geq 0$ vs. $H_1: \mu_1 - \mu_2 < 0$. Utilizando una muestra aleatoria de 25 observaciones apareadas, la desviación típica de la diferencia entre las medias muestrales es 25. ¿Puede rechazar la hipótesis nula utilizando una probabilidad de cometer el error de tipo I con $\alpha=0.05$, en cada uno de los siguientes casos?

- a) Las medias muestrales son 56 y 50
- b) Las medias muestrales son 59 y 60
- c) Las medias muestrales son 48 y 57

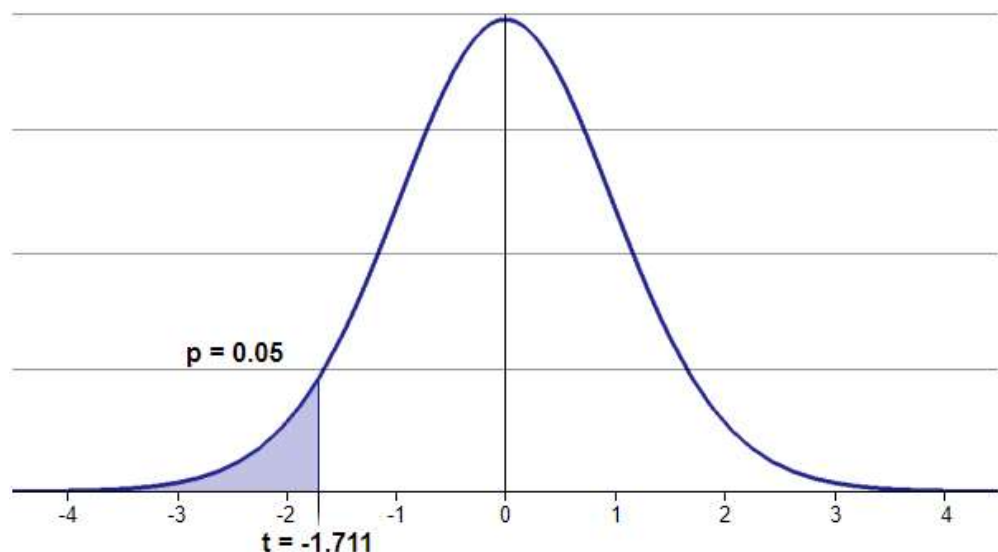
Datos Apareados. Se utiliza la distribución t de Student. El estadístico tiene la siguiente forma:

$$t^{obs} = \frac{\bar{d} - \mu_{\bar{d}}}{S_{\bar{d}} / \sqrt{n}}$$



Si $t^{obs} \geq -t_{24,0,05}$
no tengo evidencia para rechazar H_0

Si $t^{obs} < -t_{24,0,05}$
Tengo evidencia suficiente para rechazar H_0 con una confianza del 95%.

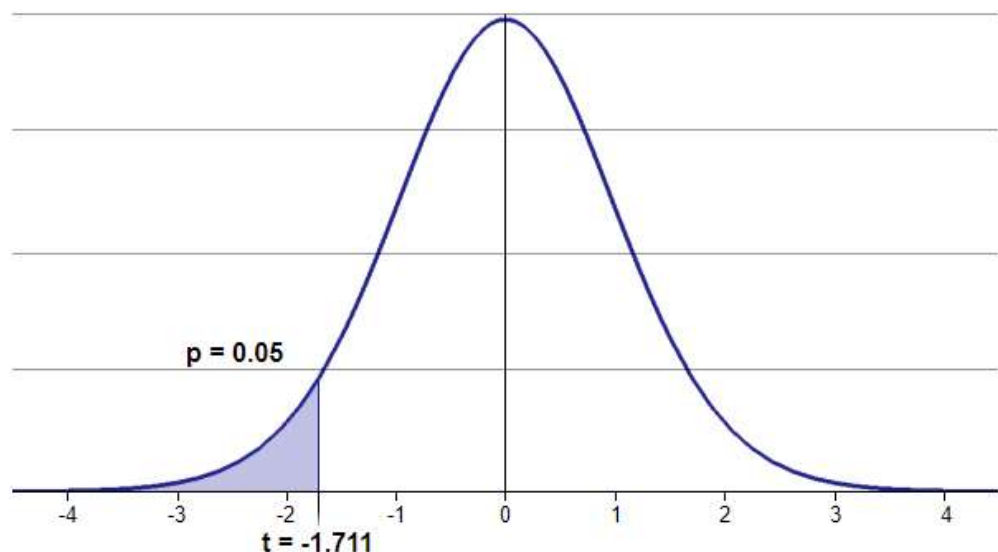


Si $t^{obs} \geq -t_{24,0,05}$
no tengo evidencia para rechazar H_o

Si $t^{obs} < -t_{24,0,05}$
Tengo evidencia suficiente para rechazar H_o con una confianza del 95%.

a)

$$t^{obs} = \frac{\bar{d} - \mu_{\bar{d}}}{S_{\bar{d}} / \sqrt{n}} = \frac{(56 - 50) - 0}{25 / \sqrt{5}} = 1,2$$

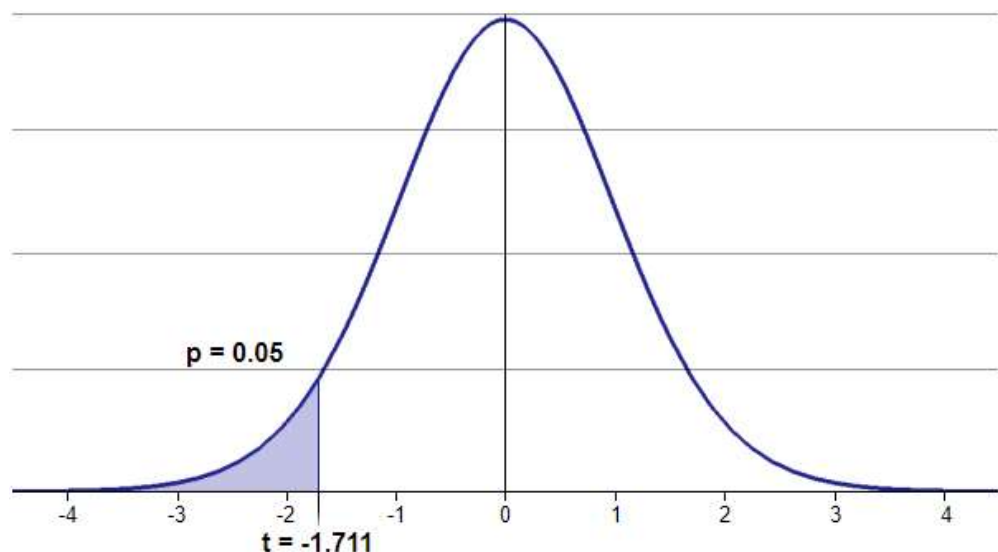


Si $t^{obs} \geq -t_{24,0,05}$
no tengo evidencia para rechazar H_o

Si $t^{obs} < -t_{24,0,05}$
Tengo evidencia suficiente para rechazar H_o con una confianza del 95%.

b)

$$t^{obs} = \frac{\bar{d} - \mu_{\bar{d}}}{S_{\bar{d}} / \sqrt{n}} = \frac{(59 - 60) - 0}{25 / \sqrt{5}} = -0,2$$



Si $t^{obs} \geq -t_{24,0,05}$
no tengo evidencia para rechazar H_o

Si $t^{obs} < -t_{24,0,05}$
Tengo evidencia suficiente para rechazar H_o con una confianza del 95%.

c)

$$t^{obs} = \frac{\bar{d} - \mu_{\bar{d}}}{S_{\bar{d}} / \sqrt{n}} = \frac{(48 - 57) - 0}{25 / \sqrt{5}} = -1,8$$




11. Las muestras aleatorias de 900 personas de EE.UU. y de Gran Bretaña indican que el 60% de los estadounidenses ve con optimismo el futuro de la economía, mientras que la cifra es del 66% en el caso de los británicos. ¿Es esta información una prueba contundente de que los británicos ven con más optimismo el futuro de la economía? Considere un error de tipo I del 1%.



11. Las muestras aleatorias de 900 personas de EE.UU. y de Gran Bretaña indican que el 60% de los estadounidenses ve con optimismo el futuro de la economía, mientras que la cifra es del 66% en el caso de los británicos. ¿Es esta información una prueba contundente de que los británicos ven con más optimismo el futuro de la economía? Considere un error de tipo I del 1%.

Test de diferencia de proporciones.

- 
11. Las muestras aleatorias de 900 personas de EE.UU. y de Gran Bretaña indican que el 60% de los estadounidenses ve con optimismo el futuro de la economía, mientras que la cifra es del 66% en el caso de los británicos. ¿Es esta información una prueba contundente de que los británicos ven con más optimismo el futuro de la economía? Considere un error de tipo I del 1%.

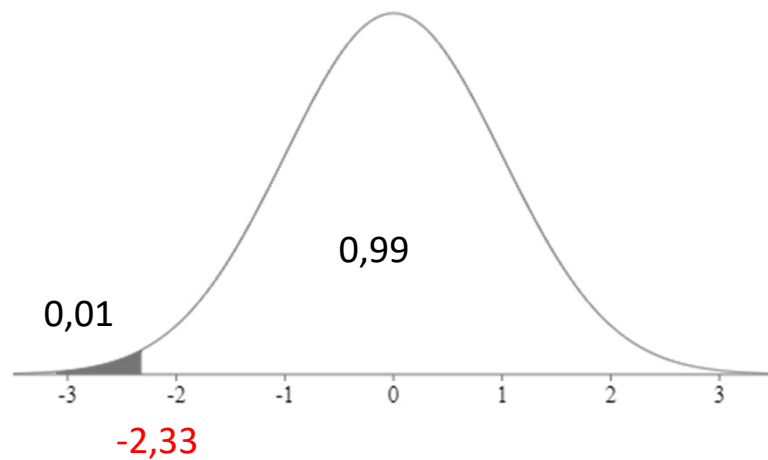
Test de diferencia de proporciones. Quiero ver si hay evidencia suficiente de que los británicos están más optimistas, entonces planteo las hipótesis de esta forma:

$$H_0: p_1 - p_2 \geq 0$$

$$H_1: p_1 - p_2 < 0$$

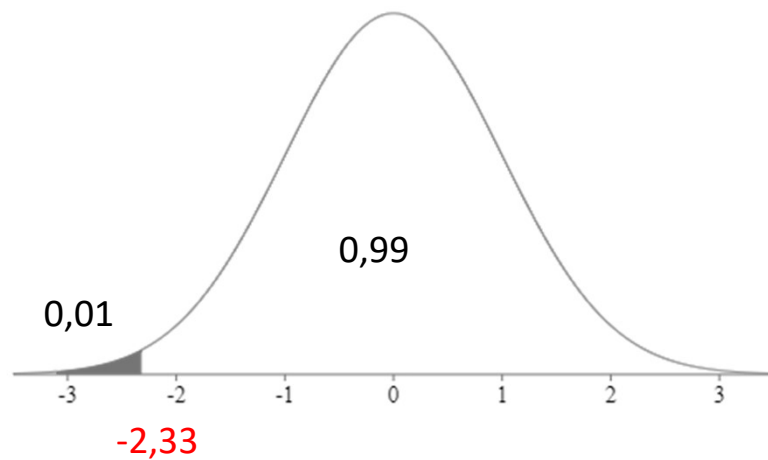
1: EEUU

2: Gran Bretaña



Si $-z^c \leq z^{obs}$
no tengo evidencia para rechazar H_o

Si $z^{obs} < -z^c$
Tengo evidencia suficiente para rechazar H_o con una confianza del 99%.



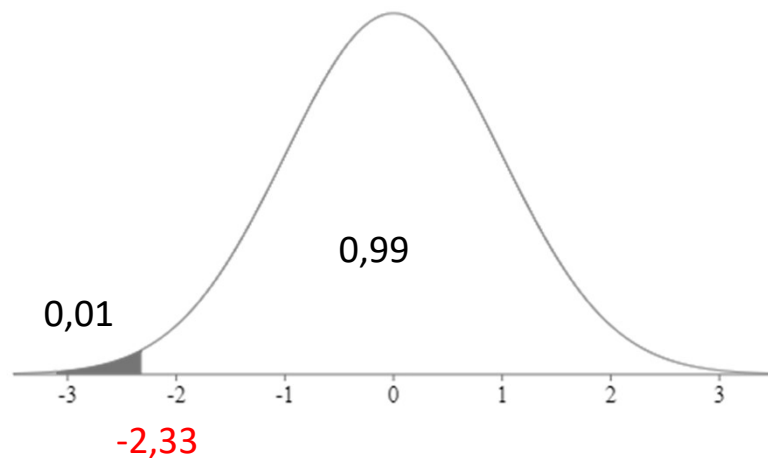
Si $-z^c \leq z^{obs}$

no tengo evidencia para rechazar H_o

Si $z^{obs} < -z^c$

Tengo evidencia suficiente para rechazar H_o con una confianza del 99%.

$$z^{obs} = \frac{(\hat{p}_x - \hat{p}_y)}{\sqrt{\frac{p_0(1-p_0)}{n_x} + \frac{p_0(1-p_0)}{n_y}}}$$



Si $-z^c \leq z^{obs}$

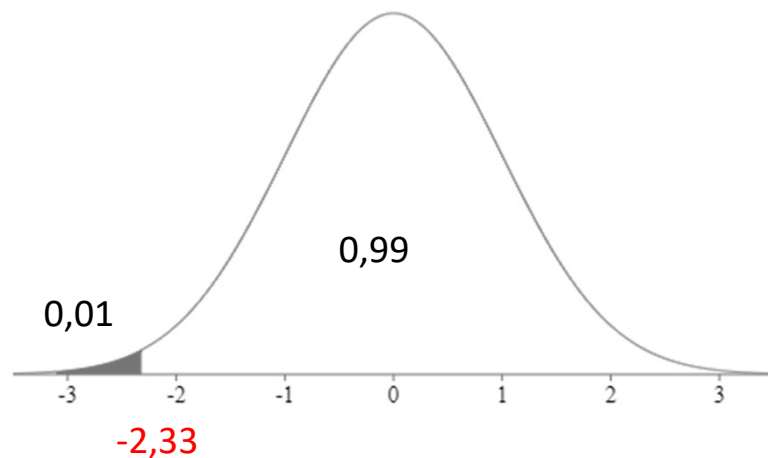
no tengo evidencia para rechazar H_o

Si $z^{obs} < -z^c$

Tengo evidencia suficiente para rechazar H_o con una confianza del 99%.

$$z^{obs} = \frac{(\hat{p}_x - \hat{p}_y)}{\sqrt{\frac{p_0(1-p_0)}{n_x} + \frac{p_0(1-p_0)}{n_y}}}$$

$$p_0 = \frac{n_x \hat{p}_x + n_y \hat{p}_y}{n_x + n_y}$$



Si $-z^c \leq z^{obs}$

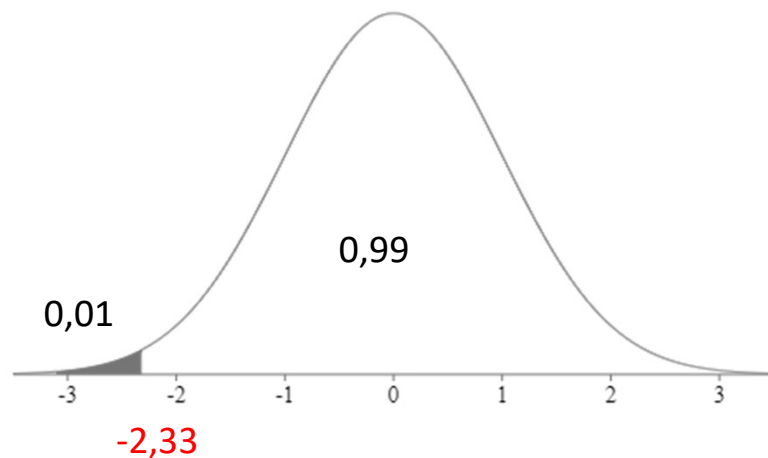
no tengo evidencia para rechazar H_0

Si $z^{obs} < -z^c$

Tengo evidencia suficiente para rechazar H_0 con una confianza del 99%.

$$z^{obs} = \frac{(\hat{p}_x - \hat{p}_y)}{\sqrt{\frac{p_0(1-p_0)}{n_x} + \frac{p_0(1-p_0)}{n_y}}}$$

$$p_0 = \frac{n_x \hat{p}_x + n_y \hat{p}_y}{n_x + n_y} = \frac{900 * 0,60 + 900 * 0,66}{900 + 900} = 0,63$$



Si $-z^c \leq z^{obs}$

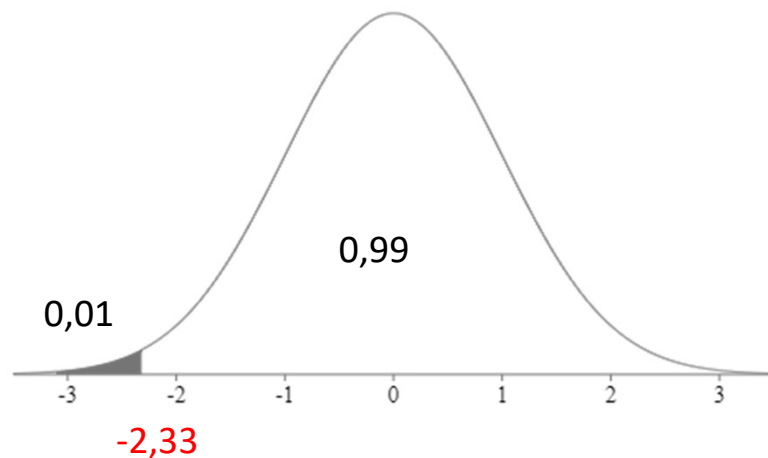
no tengo evidencia para rechazar H_0

Si $z^{obs} < -z^c$

Tengo evidencia suficiente para rechazar H_0 con una confianza del 99%.

$$z^{obs} = \frac{(\hat{p}_x - \hat{p}_y)}{\sqrt{\frac{p_0(1-p_0)}{n_x} + \frac{p_0(1-p_0)}{n_y}}} = \frac{(0,60 - 0,66)}{\sqrt{\frac{0,63(1-0,63)}{900} + \frac{0,66(1-0,66)}{900}}}$$

$$p_0 = \frac{n_x \hat{p}_x + n_y \hat{p}_y}{n_x + n_y} = \frac{900 * 0,60 + 900 * 0,66}{900 + 900} = 0,63$$



Si $-z^c \leq z^{obs}$

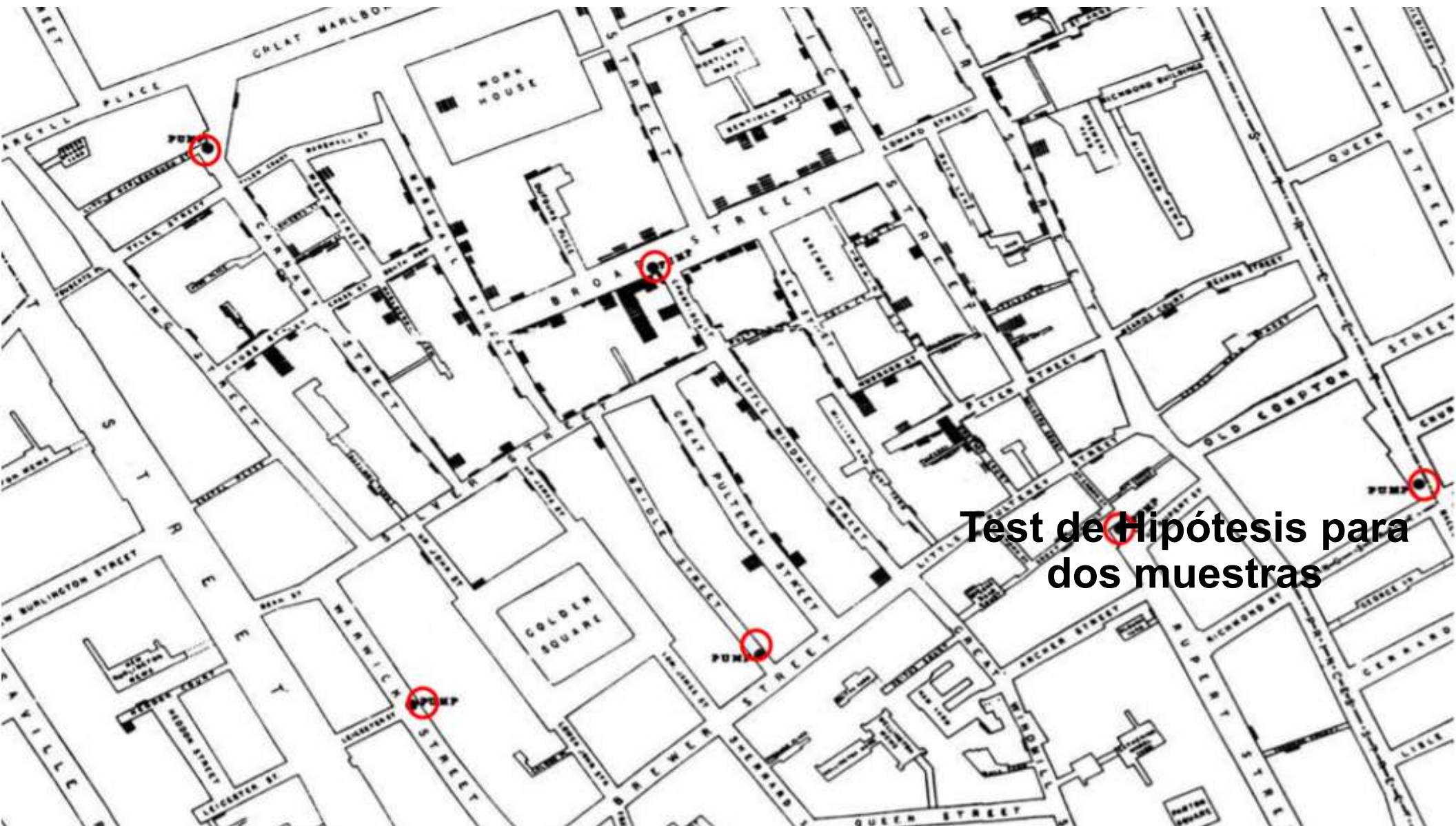
no tengo evidencia para rechazar H_o

Si $z^{obs} < -z^c$

Tengo evidencia suficiente para rechazar H_o con una confianza del 99%.

$$z^{obs} = \frac{(\hat{p}_x - \hat{p}_y)}{\sqrt{\frac{p_0(1-p_0)}{n_x} + \frac{p_0(1-p_0)}{n_y}}} = \frac{(0,60 - 0,66)}{\sqrt{\frac{0,63(1-0,63)}{900} + \frac{0,66(1-0,66)}{900}}} = -2,64$$

$$p_0 = \frac{n_x \hat{p}_x + n_y \hat{p}_y}{n_x + n_y} = \frac{900 * 0,60 + 900 * 0,66}{900 + 900} = 0,63$$



Test de Hipótesis para
dos muestras