Diferencia de Medias para dos Poblaciones con Sigma Conocida

Una tienda tiene mejores ventas en su tienda 1 que en su tienda 2. Se cree que esta diferencia se debe a que la edad media de los clientes de la zona 1 es distinta a la edad media de los clientes de la zona 2.

$$\mu_1 = Media\ de\ edades\ de\ la\ tienda\ 1$$

$$\mu_2 = Media\ de\ edades\ de\ la\ tienda\ 2$$

Como no se cuenta con las edades de todos los clientes, el gerente decide hacer una encuesta donde pregunta a sus clientes su edad.

$$\bar{x}_1 = Media\ muestral\ de\ edades\ de\ la\ tienda\ 1$$

$$\bar{x}_2 = Media\ muestral\ de\ edades\ de\ la\ tienda\ 2$$

$$n_1 = Tamaño\ de\ la\ muestra\ de\ la\ tienda\ 1$$

$$n_2 = Tamaño\ de\ la\ muestra\ de\ la\ tienda\ 2$$

A ti te interesa saber si hay alguna diferencia entre la media de edades de las tiendas. Es decir, el parámetro que te interesa es:

$$\mu_1 - \mu_2$$

Para estimar $\mu_1 - \mu_2$ las medias muestrales. Tu estimación puntual será:

$$\bar{x}_1 - \bar{x}_2$$

1. Intervalo de Confianza

- Recuerda que:

$$Intervalo = Estimación\ Puntual \pm Margen\ de\ Error$$

$$Margen\ de\ Error = Estadistico\ Critico * Error\ Estandar$$

- En este caso:

$$Estimación\ Puntual = \bar{x}_1 - \bar{x}_2$$

$$Estadistico\ Critico = Z_{\alpha/2}$$

$$Error\ Estandar = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$Intervalo = (\bar{x}_1 - \bar{x}_2) \pm Z_{\alpha/2} * \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Utiliza los siguientes datos:

	Zona 1	**Zona 2**
Tamaño de muestra	$n_1 = 36$	$n_2 = 49$
Media muestral	$\bar{x}_1 = 40$	$\bar{x}_2 = 35$
Desv. Estándar Conocida	$\sigma_1 = 9$	$\sigma_2 = 10$

Cuál es la estimación por intervalo utilizando el 95% de confianza

$$Intervalo = (\bar{x}_1 - \bar{x}_2) \pm Z_{\alpha/2} * \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$Intervalo = (40 - 35) \pm 1.96 * \sqrt{\frac{9^2}{36} + \frac{10^2}{49}}$$

$$Intervalo = 5 \pm 4.06 = \begin{cases} 0.94\ años \\ 9.06\ años \end{cases}$$

2. Pruebas de Hipótesis

3 tipos de pruebas:

1	2	3
$H_o: \mu_1 - \mu_2 \geq D_o$	$H_o: \mu_1 - \mu_2 \leq D_o$	$H_o: \mu_1 - \mu_2 = D_o$
$H_A: \mu_1 - \mu_2 < D_o$	$H_A: \mu_1 - \mu_2 > D_o$	$H_A: \mu_1 - \mu_2 \neq D_o$

$$D_o = Diferencia\ Hipotetica$$

Estadístico de Prueba:

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - D_o}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Siguiendo con el ejemplo anterior y con un nivel de significancia del 5%, **¿rechazarías la hipótesis nula de que la diferencia entre las edades de las dos tiendas es de cero?**

Las Hipótesis Nula y Alternativa son:

$$H_o: \mu_1 - \mu_2 = 0$$

$$H_A: \mu_1 - \mu_2 \neq 0$$

Como ya obtuvimos el intervalo de confianza del 95% y observamos que el cero no está dentro del intervalo, sabemos que se va a rechazar la nula. Sin embargo, vamos a solucionar este problema mediante el método de pruebas de hipótesis.

El estadístico de prueba es:

$$Z = \frac{(40 - 35) - 0}{\sqrt{\frac{9^2}{36} + \frac{10^2}{49}}} = 2.41$$

- **Método PValue**

$$Pvalue = 2 * F(-2.41) = 1.58\%$$

  - Donde F(x) es la función acumulada normal

Por lo tanto se rechaza la hipótesis nula y concluimos que hay una diferencia de edades entre los clientes de las dos tiendas.

- **Mediante el Método del Estadístico Critico**

$$F^{-1}(0.025) = -1.96 \ \ y \ \ F^{-1}(0.975) = +1.96$$

$$2.41 > 1.96 \rightarrow Rechazar\ la\ hipotesis\ nula$$