Estadística Inferencial

Capítulo X - Ejercicio 51

Aaric Llerena Medina

El agente de compras de la compañía "PC" se vio confrontado con dos marcas de computadoras para su adquisición. Se le permitió probar ambas marcas asignando una misma tarea a 50 máquinas de cada marca, resultando las medias respectivas 55 y 50 minutos. Suponga las dos poblaciones tienen varianza homogénea igual a 100. Para el nivel de significación $\alpha = 0.05$:

a) ¿Excede el tiempo promedio de la marca 1 al de la marca 2 en al menos 9 minutos?

b) Halle la potencia de la prueba cuando la diferencia real entre promedios de tiempo de marca 1 menos marca 2 sea 3 minutos.

c) ¿Qué tan grande debe ser la muestra si la potencia de la prueba es 0.95, cuando la diferencia real entre promedios de tiempo marca 1 menos marca 2 es 3 minutos?

Solución:

Los datos del problema:

- Media muestral de la marca 1: $\bar{X}_1 = 55$.
- Media muestral de la marca 2: $\bar{X}_2 = 50$.
- Varianza común: $\sigma^2 = 100$.
- Tamaño de la muestra para cada marca: $n_1 = n_2 = 50$.
- Nivel de significación: $\alpha = 0.05$.

Definiendo las hipótesis:

$$H_0 : \mu_1 - \mu_2 \geq 9 \quad \text{contra} \quad H_1 : \mu_1 - \mu_2 < 9$$

donde μ_1 y μ_2 son los tiempos promedio de las marcas 1 y 2, respectivamente.

a) El estadístico de prueba para la diferencia de medias con varianzas iguales se calcula como:

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{2\sigma^2}{n}}}$$

donde $(\mu_1 - \mu_2)_0 = 9$ es la diferencia bajo la hipótesis nula.

Sustituyendo los valores:

$$Z = \frac{(55-50)-9}{\sqrt{\frac{2\times 100}{50}}} = \frac{5-9}{\sqrt{4}} = \frac{-4}{2} = -2$$

Para un nivel de significación $\alpha = 0.05$, el valor crítico $Z_{0.05}$ es aproximadamente -1.645 (obtenido de la tabla de distribución normal estándar).

Ahora, se establece la regla de decisión:

- Si $Z_{\text{calc}} < Z_{\alpha}$, se rechaza H_0.
- Si $Z_{\text{calc}} \geq Z_{\alpha}$, no se rechaza H_0.

Dado que $Z = -2 < -1,645$, se rechaza H_0. Por lo tanto, hay pruebas suficientes para concluir que el tiempo medio de la marca 1 no supera al de la marca 2 en al menos 9 minutos.

b) Dado que la diferencia verdadera es $\mu_1 - \mu_2 = 3$. Además, la desviación estándar de la diferencia de medias es:

$$\sigma = \sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}} = \sqrt{\frac{100}{50} + \frac{100}{50}} = \sqrt{2+2} = 2$$

Dado que la diferencia y la varianza, la distribución que sigue $\bar{X}_1 - \bar{X}_2$ es:

$$\bar{X}_1 - \bar{X}_2 \sim N\left(3, 4\right)$$

El estadístico de prueba Z para H_0 es:

$$Z = \frac{\left(\bar{X}_1 - \bar{X}_2\right) - 9}{\sigma} = \frac{\left(\bar{X}_1 - \bar{X}_2\right) - 9}{2}$$

Estandarizando:

$$Z = \frac{\left(\bar{X}_1 - \bar{X}_2\right) - 9}{2} \sim N\left(\frac{3-9}{2}, \frac{4}{4}\right) = N(-3, 1)$$

La potencia es la probabilidad de rechazar H_0 cuando H_1 es verdadera y la región crítica para un $\alpha = 0.05$ en una cola izquierda es $Z < -1.645$. y con $Z \sim N\left(-3, 1\right)$, estandarizando el valor crítico para la distribución de H_1:

$$Z_{\text{crítico}} = \frac{-1.645 - (-3)}{1} = 1.355$$

La potencia es:

$$\text{Potencia} = P(Z < 1.36) = \Phi(1.36) \approx 0.9131$$

Por lo tanto, la potencia de la prueba es aproximadamente 0.9131 (91.31 %) cuando la diferencia real entre los promedios es de 3 minutos.

c) Para calcular el tamaño de la muestra con una potencia de la prueba de 0.95, para ello, se emplea la fórmula:

$$n = \frac{(Z_{1-\alpha} + Z_{\beta})^2 \cdot \sigma^2}{(\mu_1 - \mu_2)^2} = \frac{(Z_{1-0.05} + Z_{0.95})^2 \cdot 100}{3^2} = \frac{(1.645 + 1.645)^2 \cdot 100}{9} = 120.2678 \approx 121$$

Por lo tanto, el tamaño de muestra necesario es aproximadamente 121.

Otra forma de abordar la solución, es usando la fórmula de la potencia para una prueba Z de una cola:

$$\text{Potencia} = \Phi\left(-Z_{\alpha} + \frac{|\Delta - \delta|}{\sigma}\right)$$

Donde:

- $\Delta = 9$ (diferencia para H_0).
- $\delta = 3$ (diferencia para H_1).
- $\sigma = \sqrt{\frac{\sigma^2}{n} + \frac{\sigma^2}{n}} = \sqrt{\frac{200}{n}}$
- Φ es la función acumulada de la normal estándar.

Igualando a la potencia de 0.95 y se quiere despejar n, entonces:

$$\Phi\left(-Z_{0.05} + \frac{9-3}{\sqrt{\frac{200}{n}}}\right) = 0.95$$

Por lo que calculando el valor Φ^{-1} a ambos lados y buscando $\Phi^{-1}(0.95) = 1.645$, asimismo, $-Z_{0.05} = 1.645$, por lo que se obtiene:

$$-1.645 + \frac{6}{\sqrt{\frac{200}{n}}} = 1.645$$

$$\frac{6}{\sqrt{\frac{200}{n}}} = 3.290$$

$$\sqrt{\frac{200}{n}} = \frac{6}{3.290}$$

$$n = \left(\frac{3.290}{6}\right)^2 \times 200$$

$$n \approx 60.1339$$

Por lo tanto, se requiere una muestra de 61 computadoras por marca para alcanzar una potencia del 95 % cuando la diferencia real entre los promedios es de 3 minutos. Es decir, la muestra total es 121 computadoras.