Estadística Inferencial

Capítulo X - Ejercicio 47

Aaric Llerena Medina

Un inversionista está por decidir entre dos localidades para abrir un centro comercial. Para esto debe probar la hipótesis de que hay diferencia en la media de los ingresos mensuales de los hogares de las dos provincias. Se escogió una muestra aleatoria de cada lugar y se obtiene la tabla de resultados en dólares:

	Localidad	
	A	B
Tamaño muestral	300	400
Media muestral	400	420
Varianza muestral	8,100	14,400

a) ¿Qué estadística es la apropiada para esta prueba de hipótesis?

b) Para un nivel de significación de 0.05, ¿puede el inversionista concluir que le es indiferente construir en cualquiera de las dos localidades?, si no es así, ¿en cuál de las localidades debería abrir el centro comercial?

Solución:

Los datos muestrales para las localidades son:

- **A:** Tamaño muestral $n_A = 300$, media muestral $\bar{x}_A = 400$, varianza muestral $s_A^2 = 8,100$
- **B:** Tamaño muestral $n_B = 400$, media muestral $\bar{x}_B = 420$, varianza muestral $s_B^2 = 14,400$

Se plantean las hipótesis:

$$H_0 : \mu_A = \mu_B \quad \text{vs} \quad H_1 : \mu_A \neq \mu_B$$

a) Dado que las varianzas poblacionales son desconocidas y posiblemente diferentes, utilizamos la prueba t de Welch para la diferencia de medias. El estadístico de prueba es:

$$t = \frac{(\bar{X}_A - \bar{X}_B) - (\mu_A - \mu_A)}{\sqrt{\dfrac{s_A^2}{n_A} + \dfrac{s_B^2}{n_B}}} = \frac{400 - 420}{\sqrt{\dfrac{8,100}{300} + \dfrac{1,400}{400}}} = \frac{-20}{\sqrt{27 + 36}} = \frac{-20}{\sqrt{63}} \approx \frac{-20}{7.937} \approx -2.52$$

Asimismo, los grados de libertad para la prueba t de Welch se calculan con la fórmula:

$$\text{gl} = \frac{\left(\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}\right)^2}{\frac{(s_A^2/n_A)^2}{n_A - 1} + \frac{(s_B^2/n_B)^2}{n_B - 1}} = \frac{\left(\frac{8,100}{300} + \frac{14,400}{400}\right)^2}{\frac{(8,100/300)^2}{300 - 1} + \frac{(14,400/400)^2}{400 - 1}} = \frac{3,969}{5.6862} \approx 698$$

b) Para un nivel de significación del 5 % y una prueba bilateral, los valores críticos de t con 698 grados de libertad son aproximadamente ± 1.96. La región de rechazo es:

$$t < -1.96 \quad \text{o} \quad t > 1.96$$

Por lo tanto, el valor calculado del estadístico de prueba es $t \approx -2.52$ es menor que -1.96, por lo que el valor calculado cae en la región de rechazo, es decir, se rechaza la hipótesis nula H_0.

Además, se puede calcular valor-P bilateral para confirmar la conclusión anterior. Para ello, dado que distribución t con 698 grados de libertad es prácticamente indistinguible de la distribución normal estándar. El valor absoluto del estadístico de prueba es: $|t| = 2.52$ Buscando la probabilidad de que una variable aleatoria Z (distribución normal estándar) sea mayor que 2.52:

$$\begin{aligned} P(Z > 2.52) &= 1 - P(Z < 2.52) \\ &= 1 - 0.9941 \\ &= 0.0059 \end{aligned}$$

Como la prueba es bilateral, se multiplica por 2:

$$\text{Valor-}p = 2 \times 0.0059 = 0.0118$$

Como el valor-p es menor que el nivel de significación 0.05, por lo que rechazamos la hipótesis nula. Esto confirma nuestra conclusión anterior de que hay una diferencia significativa en las medias de los ingresos mensuales entre las dos localidades.