

# Prueba de hipótesis estadísticas

## 1.1 Introducción

La prueba de una hipótesis estadística tiene una fuerte relación con el concepto de estimación.

Una hipótesis estadística es una afirmación con respecto a alguna característica desconocida de una población de interés. La esencia de probar una hipótesis estadística es el de decidir si la afirmación se encuentra apoyada por la evidencia experimental que se obtiene a través de una muestra aleatoria.

## 1.2 Conceptos básicos para la prueba de hipótesis estadísticas

La idea principal es determinar si el valor de un parámetro es igual a un valor específico o si es diferente de ese valor. A la afirmación igual a un valor específico se le llama *hipótesis nula* y se escribe como:

$$H_0 : \mu = 10.$$

Debemos tomar en cuenta las consecuencias que puede originarse como resultado del verdadero estado de la naturaleza al tomar alguna de las posibles decisiones. Es decir, existe dos posibles decisiones con respecto a  $H_0$ : Rechazar o equivocarse al rechazar  $H_0$ . Sin embargo, cada una de estas decisiones tiene las siguientes dos consecuencias con respecto al estado de la naturaleza:

$$\begin{array}{cc} \text{Rechazar } H_0 \left\{ \begin{array}{l} \text{cuando } H_0 \text{ es cierta} \\ \text{cuando } H_0 \text{ es falsa} \end{array} \right. & \text{Equivocarse al rechazar } H_0 \left\{ \begin{array}{l} \text{cuando } H_0 \text{ es cierta} \\ \text{cuando } H_0 \text{ es falsa} \end{array} \right. \end{array}$$

Cuando se toma una decisión con respecto a una hipótesis nula, dos de las posibles consecuencias relativas al verdadero estado de la naturaleza conducen a errores inferenciales. El rechazo de la hipótesis  $H_0$  cuando en realidad es cierta, constituye lo que se denomina *error de tipo I*. Equivocarse al rechazar  $H_0$  cuando en realidad  $H_0$  es falsa, constituye lo que se denomina *error de tipo II*. Debemos tomar en cuenta que sólo es posible el error de tipo I cuando la decisión es el rechazar la hipótesis nula, mientras que el error de tipo II sólo es posible cuando la decisión es el no rechazar  $H_0$ . En otras palabras, si la hipótesis nula es falsa, sólo puede cometerse un error de tipo I; si la hipótesis nula realmente es falsa, sólo puede cometerse un error de tipo II.

Es necesario tener alguna cantidad que mida la posibilidad de cometer alguno de estos errores. Esta medida es una probabilidad.

**Definición 1.1** La probabilidad de rechazar  $H_0$ , dado que  $H_0$  es cierta, se define como la probabilidad (o tamaño) de error de tipo I y se denota por  $\alpha$ ,  $0 \leq \alpha \leq 1$ .

**Definición 1.2** La probabilidad de no rechazar  $H_0$ , dado que  $H_0$  es falsa, se define como la probabilidad (o tamaño) de error de tipo II y se denota por  $\beta$ ,  $0 \leq \beta \leq 1$ .

Por lo tanto, las probabilidades de error de tipo I y tipo II están dadas por las proposiciones

$$P(\text{rechazar } H_0 | H_0 \text{ es cierta}) = \alpha.$$

$$P(\text{no poder rechazar } H_0 | H_0 \text{ es falsa}) = \beta.$$

Cuando una afirmación se incorpora en la proposición de la hipótesis nula, se necesita una regla que indique que decisión tomar con respecto a  $H_0$  una vez que se encuentra disponible la evidencia muestral. Esta regla recibe el nombre de prueba de una hipótesis estadística.

**Definición 1.3** Una *prueba* de una hipótesis estadística con respecto a alguna característica desconocida de la población de interés es cualquier regla para decidir si se rechaza la hipótesis nula con base en una muestra aleatoria de la población.

Por ejemplo, se decide rechazar  $H_0 : \mu = 10$ . Para un tamaño  $n$  dado de la muestra, supóngase que se decide rechazar  $H_0$  si se observa un valor de la media muestral  $\bar{X}$  que sea más grande que 12. Entonces,  $\bar{X}$  es la estadística de prueba, el valor  $\bar{X} = 12$  es el *valor crítico*, y el conjunto de valores mayores que 12 constituyen la región crítica de la prueba.

El área de la región crítica es igual al tamaño del error de tipo I. Es decir,

$$P(\bar{X} > 12 | \mu = 10) = \alpha.$$

La probabilidad de  $\alpha$  es sólo una referencia con respecto a la región  $\bar{X} > 12$ . Pero la decisión de rechazar  $H_0$  se tomará con base en una sola muestra de tamaño  $n$ , a partir de la cual se calculará el estimador de  $\bar{x}$ . De esta forma, si  $\bar{x} > 12$ , esto no significa que la probabilidad de que  $H_0$  sea correcta es  $\alpha$ ; más bien, esto implica una interpretación de frecuencia para  $\alpha$  cuando se toman muchas muestras. En otras palabras, si el valor de  $\mu$  es realmente 10, y si se tomasen en forma repetida muestras de tamaño  $n$  de la población, debe esperarse que en un  $100\alpha\%$  de las veces, se encuentre un valor de la estadística de prueba  $\bar{X}$  mayor que 12 y de esta forma debe rechazarse la hipótesis nula. Sólo en este sentido puede decirse que la confiabilidad al rechazar  $H_0$ , cuando el estimador  $\bar{X} > 12$  es igual al complemento del error  $\alpha$  de tipo I, o  $1 - \alpha$ .

También es necesario establecer una *hipótesis alternativa* que refleje el valor posible o intervalo de valores del parámetro de interés si la hipótesis nula ( $H_1$ ) es falsa. Esto es,  $H_1$  representa alguna forma de negación de  $H_0$ .

Cabe mencionar que el rechazo de la hipótesis nula implicaría que ha sido capaz encontrar evidencia suficiente para garantizar la hipótesis alternativa. Por otro lado, si no se presenta evidencia sustancial, se aceptará la hipótesis nula. Esta decisión no implica necesariamente que "el acusado sea inocente", más bien hace énfasis en la falta de evidencia sustancial necesaria para condenar al acusado. Por lo tanto, en cierto sentido un veredicto de culpable (rechazo de  $H_0$ ) debe ser considerado como una decisión más fuerte que un veredicto de inocente (no rechazar  $H_0$ ). A lo que, el error de tipo I se considera un error mucho más grave que el error de tipo II.

Como resultado se tiene que muchas veces se selecciona con anticipación el tamaño del error de tipo I que puede tolerarse y se intenta construir un procedimiento de prueba que minimice el tamaño del error de tipo II. Es decir, no es posible fijar tanto a  $\alpha$  como a  $\beta$  y diseñar alguna regla de decisión para probar  $H_0$  contra  $H_1$ , dada una muestra aleatoria de tamaño  $n$ . Es por esta razón que se dice "equivocación al

rechazar  $H_0$ " mas que aceptar  $H_0$  cuando la evidencia muestral no apoya el rechazo de la hipótesis nula.

También cabe mencionar que el error de tipo II aumenta conforme el error de tipo I disminuya. La probabilidad  $\alpha$  del error de tipo I también se conoce como el *nivel de significancia estadístico*. En este contexto la palabra "significancia" sólo implica que la evidencia muestral es tal que garantiza el rechazo de  $H_0$  a un nivel dado  $\alpha$ .