TEMA 6

CONTRASTE DE HIPÓTESIS

6.1 INTRODUCCIÓN

Un contraste de hipótesis es una técnica estadística para analizar la validez empírica de una afirmación hecha con respecto a alguna característica desconocida de una población a la cual denominaremos *hipótesis nula* y denotaremos como H_0. Para ello se juzga si la afirmación realizada es compatible con una muestra **X** extraída de dicha población.

El objetivo será confirmar o refutar la hipótesis planteada en función del grado de compatibilidad de la muestra observada con H_0. Para ello se construye una medida de de discrepancia $D(\mathbf{X},H_0)$ entre los datos y la hipótesis a contrastar, de forma que si dicha medida de discrepancia es suficientemente grande se entiende que la hipótesis H_0 es muy improbable que explique los datos observados y, por lo tanto, se rechaza y en caso contrario se acepta como válida.

A partir de este planteamiento aparecen varios problemas.

1) Cómo elegir D

2) Cómo determinar cuando D es suficientemente grande

3) Si H_0 se rechaza qué hipótesis alternativas podrían explicar los datos observados teniendo en cuenta los resultados del contraste

En este tema abordaremos dichas cuestiones cuando conocemos a priori la distribución de la variable aleatoria poblacional X y nuestras hipótesis se refieren a algún parámetro θ desconocido de dicha distribución.

6.2 CONCEPTOS BÁSICOS

A continuación vamos a realizar una clasificación de las hipótesis. En primer lugar, según su papel en el contraste:

Llamaremos **hipótesis nula** y la denotamos por H_0 a la hipótesis que planteamos como cierta y queremos confirmar o rechazar. Suele representar a la hipótesis estable o conservadora, es decir, la situación sigue como hasta ahora. Llamaremos **hipótesis alternativa** y la denotamos por H_1 a la hipótesis que nos creeremos en caso de rechazar la hipótesis nula. Suele representar a la hipótesis innovadora, es decir, la situación cambia o se modifica.

En segundo lugar una hipótesis se clasifica según el tipo de información contrastada. Si la hipótesis hace referencia a un parámetro poblacional estaremos ante un contraste de **significación.** Si hace referencia a un modelo, estaremos ante un **contraste de bondad de ajuste o especificación.**

En tercer lugar, una hipótesis se clasifica según la información aportada:

Llamaremos **hipótesis simple** a la que determina perfectamente la distribución de probabilidad de la variable aleatoria poblacional, es decir, proporciona una única distribución bajo ella (una única función de cuantía o función de densidad). Llamaremos **hipótesis compuesta** a la que no determina la distribución de probabilidad de la variable aleatoria poblacional, es decir, proporciona una familia finita o no de distribuciones (una familia de funciones de cuantía o de funciones de densidad).

Un **contraste de hipótesis** es una regla de decisión mediante la cual optamos por creer en una o en otra hipótesis. Por lo tanto, un contraste proporciona una partición del espacio muestral en dos subconjuntos disjuntos S y S^C, de tal forma que si la muestra observada pertenece a S rechazaremos la hipótesis nula (creeremos en la hipótesis alternativa) y si pertenece a S^C (es decir, no pertenece a S) seguiremos pensando que es cierta la hipótesis nula (rechazaremos la alternativa). Al subconjunto S le denominamos región crítica o región de rechazo y a su complementario, S^C, región de aceptación. La construcción de S la realizaremos a partir de estadísticos muestrales, T, a los cuales denominaremos *estadísticos del contraste*, que canalizan la información proporcionada por los datos en la dirección adecuada para llevarlo a cabo. De esta forma la regla de decisión del contraste tomará la siguiente forma:

Si $T(X_1, \ldots, X_n) \in S$ se rechaza H_0 y se acepta H_1

Si $T(X_1, \ldots, X_n) \notin S$ se acepta H_0

En toda toma de decisiones podemos cometer errores de asignación. Así, en un contraste podemos equivocarnos al rechazar una hipótesis cierta o al admitir una hipótesis falsa, es decir, hay dos tipos de errores. El error que se comete al rechazar la hipótesis nula cuando realmente es cierta se denomina **error de tipo I**. El error que se produce al aceptar la hipótesis nula cuando realmente es falsa recibe el nombre de **error de tipo II**. Estos comentarios se resumen en la siguiente tabla:

		¿Estado de la naturaleza?	
		H_0	H_1
¿Decisión adoptada?	H_0	Decisión correcta	**Error tipo II**
	H_1	**Error tipo I**	Decisión correcta

Puesto que en un test de hipótesis se pueden cometer errores tendremos que cuantificarlos de alguna forma para su estudio posterior. El modo de cuantificarlos será calculando la probabilidad de que se produzcan, estas probabilidades deben ser condicionales puesto que no conocemos la veracidad de las hipótesis estadísticas. Así pues, las probabilidades de cometer un error vienen dadas por:

$$P\{\text{Error tipo I}\} = P\{T(X_1, X_2, ..., X_n) \in S \mid H_0 \text{ cierta}\}$$
$$P\{\text{Error tipo II}\} = P\{T(X_1, X_2, ..., X_n) \notin S \mid H_1 \text{ cierta}\}$$

Estas probabilidades pueden ser funciones si una o ambas hipótesis son compuestas porque en ese caso hay una familia de distribuciones de probabilidad para la variable aleatoria poblacional y para cada una de ellas tendríamos un valor concreto de la probabilidad de cometer el error de tipo I o el error de tipo II. Por este motivo, con objeto de agrupar las probabilidades de cometer algún error, se define la **función de potencia** como una función entre el espacio paramétrico y el intervalo [0,1], asignando a cada posible valor del parámetro la probabilidad de rechazar la hipótesis nula condicionado a dicho valor. Esta función de potencia para un contraste cuya función es φ se denota por β.

$$\beta_T : \Omega \to [0,1]$$
$$\theta \to \beta_T(\theta) = P\{\text{Rechazar } H_0 \mid \theta\} = P(T(X_1,\ldots,X_n) \in S|\theta)$$

De esta forma expresamos, en una misma función, ambas probabilidades de cometer alguno de los errores del contraste de hipótesis. Así, por ejemplo, si el parámetro pertenece a la hipótesis nula estamos calculando la probabilidad del error de tipo I y si el parámetro pertenece a la hipótesis alternativa estamos calculando la probabilidad complementaria al error de tipo II.

Llamaremos **nivel de significación** o **tamaño del contraste**, denotado como α, a la máxima probabilidad de cometer el error de tipo I:

$$\beta_T(\theta)\leq\alpha \quad \forall\theta\in H_0$$

El investigador fijará a priori el nivel de significación α y dentro de todos los contrastes de nivel α buscará, si existe, aquel que maximiza la potencia del contraste en todos los valores paramétricos que pertenecen a la hipótesis nula. Así se dice que un contraste con estadístico del test T* es un contraste uniformemente de máxima potencia (UMP) para un nivel de significación α si:

$$\beta_{T^*}(\theta)\geq\beta_T(\theta) \quad \forall\theta\in H_1, \forall T \text{ tal que } \beta_T(\theta)\leq\alpha \quad \forall\theta\in H_0$$

6.3 CONTRASTE DE HIPÓTESIS SIMPLES

En esta situación cada hipótesis sólo contiene una función de cuantía o de densidad, por lo tanto, la probabilidad del error de tipo I y la probabilidad del error de tipo II son dos únicos puntos.

6.3.1. Lema de Neyman-Pearson

Sea X la variable aleatoria poblacional cuya distribución es conocida salvo un único parámetro θ. Planteamos un contraste de hipótesis simples: la hipótesis nula H_0 se expresa como $\theta=\theta_0$ y la hipótesis alternativa H_1 como $\theta=\theta_1$. Se toma una muestra aleatoria simple $(X_1,X_2,...,X_n)$, cuya función de verosimilitud es $L(\theta)=L(X_1,X_2,...,X_n;\theta)$, entonces el contraste que tiene como estadístico del contraste $T(X_1, \ldots, X_n) = \dfrac{L(\theta_0)}{L(\theta_1)}$ y cuya función de decisión viene dada por:

Rechazar H_0 si $\dfrac{L(\theta_0)}{L(\theta_1)}<k$

Aceptar H_0 si $\dfrac{L(\theta_0)}{L(\theta_1)}\geq k$

donde k>0 tal que $P\left(\dfrac{L(\theta_0)}{L(\theta_1)}<k \mid \theta=\theta_0\right)=\alpha$, es el contraste uniformemente de máxima potencia dentro de los de su nivel de significación α.

6.4 CONTRASTE DE HIPÓTESIS COMPUESTAS

En esta situación al menos una hipótesis es compuesta y, por lo tanto, al menos estamos contrastando una familia de funciones de cuantía o de densidad. Del mismo modo, la función de potencia contiene un conjunto de valores, es realmente una función continua. Para estos casos no existe en general un test UMP y se utilizan diversas formas para construir contrastes. En Econometría se explicará algunas de ellas. En un curso introductorio como este explicamos el llamado contraste de razón de verosimilitudes que es uno de los más utilizados.

6.4.1. Test de la razón de verosimilitudes

Sea X la variable aleatoria poblacional cuya distribución es conocida salvo un único parámetro θ. Planteamos un contraste de hipótesis compuestas: la hipótesis nula H_0 se expresa como $\theta \in \Omega_0$ y la hipótesis alternativa H_1 como $\theta \in \Omega_1$. Denotaremos por Ω al espacio paramétrico y viene dado por la unión de los espacios de la hipótesis nula y alternativa: $\Omega = \Omega_0 \cup \Omega_1$. Se toma una muestra aleatoria simple $(X_1, X_2, ..., X_n)$, cuya función de verosimilitud es $L(\theta) = L(X_1, X_2, ..., X_n; \theta)$, el estadístico del contraste viene dado por $T(X_1, \ldots, X_n) = \dfrac{\sup_{\theta \in \Omega_0} L(\theta)}{\sup_{\theta \in \Omega} L(\theta)}$ y la regla de decisión viene dada por:

Rechazar H_0 si $\dfrac{\sup_{\theta \in \Omega_0} L(\theta)}{\sup_{\theta \in \Omega} L(\theta)} < k$

Aceptar H_0 si $\dfrac{\sup_{\theta \in \Omega_0} L(\theta)}{\sup_{\theta \in \Omega} L(\theta)} \geq k$

donde k>0 si $\sup_{\theta \in \Omega_0} P\left(\dfrac{\sup_{\theta \in \Omega_0} L(\theta)}{\sup_{\theta \in \Omega} L(\theta)} < k \mid \theta \right) = \alpha$.

es el contraste de la razón de verosimilitud. La constante k vendrá calculada a partir del nivel de significación α establecido por el investigador a priori.

En los casos más habituales, la razón de verosimilitudes queda expresada mediante un estadístico cuya distribución es conocida, por lo tanto, es posible fijar el nivel de significación α y delimitar la región crítica.

6.5 CONTRASTES NOTABLES

En este epígrafe vamos a presentar los contrastes más habituales respecto de la media (proporción) y de la varianza.

6.5.1. Contrastes respecto a la media

Sea X la variable aleatoria poblacional cuya media μ es desconocida y cuya desviación típica es σ. Extraemos una m.a.s. de dicha población que denotamos por $(X_1,X_2,\ldots,X_n)$. El estimador de la media es la media muestral $\overline{X}$ y el de la varianza, en caso de necesitar estimarla, es la cuasivarianza muestral s_1^2.

En el caso particular de que la variable poblacional sigue una distribución Bernoulli entonces la media poblacional es la proporción p y su estimador es la proporción muestral $\hat{p}$.

El estadístico media muestral sigue una distribución normal en cualquiera de estas dos situaciones: la distribución poblacional sigue una normal o el tamaño muestral es elevado y podemos aplicar el teorema central del límite.

El estadístico media muestral sigue una distribución t de Student si la distribución poblacional es normal y hemos tenido que estimar la desviación típica porque era desconocida.

Bajo estas premisas, la tabla siguiente presenta las diferentes regiones críticas, dependiendo de las condiciones previas que tengamos de la variable aleatoria poblacional y de las hipótesis planteadas. Las zonas de rechazo se han obtenido aplicando el test de la razón de verosimilitudes.

Condiciones	H_0	H_1	Región Crítica	Estadístico
Normalidad o n grande Varianza conocida	$\mu=\mu_0$ $\mu\geq\mu_0$ $\mu=\mu_0$ $\mu\leq\mu_0$ $\mu=\mu_0$	$\mu<\mu_0$ $\mu<\mu_0$ $\mu>\mu_0$ $\mu>\mu_0$ $\mu\neq\mu_0$	$Z<-z_\alpha$ $Z<-z_\alpha$ $Z>z_\alpha$ $Z>z_\alpha$ $\lvert Z>z_{\alpha/2}\rvert$	$Z=\frac{\overline{X}-\mu_0}{\sigma}\sqrt{n}$
n grande Varianza desconocida	$\mu=\mu_0$ $\mu\geq\mu_0$ $\mu=\mu_0$ $\mu\leq\mu_0$ $\mu=\mu_0$	$\mu<\mu_0$ $\mu<\mu_0$ $\mu>\mu_0$ $\mu>\mu_0$ $\mu\neq\mu_0$	$Z<-z_\alpha$ $Z<-z_\alpha$ $Z>z_\alpha$ $Z>z_\alpha$ $\lvert Z>z_{\alpha/2}\rvert$	$Z=\frac{\overline{X}-\mu_0}{s_1}\sqrt{n}$
Normalidad Varianza desconocida	$\mu=\mu_0$ $\mu\geq\mu_0$ $\mu=\mu_0$ $\mu\leq\mu_0$ $\mu=\mu_0$	$\mu<\mu_0$ $\mu<\mu_0$ $\mu>\mu_0$ $\mu>\mu_0$ $\mu\neq\mu_0$	$T<-t_{n-1,\alpha}$ $T<-t_{n-1,\alpha}$ $T>t_{n-1,\alpha}$ $T>t_{n-1,\alpha}$ $\lvert T\rvert>\lvert t_{n-1,\alpha/2}\rvert$	$T=\frac{\overline{X}-\mu_0}{s_1}\sqrt{n}$
Distribución Bernoulli n grande	$p=p_0$ $p\geq p_0$ $p=p_0$ $p\leq p_0$ $p=p_0$	$p<p_0$ $p<p_0$ $p>p_0$ $p>p_0$ $p\neq p_0$	$Z<-z_\alpha$ $Z<-z_\alpha$ $Z>z_\alpha$ $Z>z_\alpha$ $\lvert Z>z_{\alpha/2}\rvert$	$Z=\frac{\hat{p}-p_0}{\sqrt{p_0(1-p_0)}}\sqrt{n}$ o $Z=\frac{\hat{p}-p_0}{\sqrt{\hat{p}(1-\hat{p})}}\sqrt{n}$

6.5.2. Contrastes sobre la varianza de una población normal

Sea X la variable aleatoria poblacional con distribución normal cuya media μ y cuya desviación típica σ son desconocidas. Extraemos una m.a.s. de dicha población que denotamos por $(X_1,X_2,...X_n)$. El estimador de la media es la media muestral $\overline{X}$ y el de la varianza es la cuasivarianza muestral s_1^2.

Utilizando el Teorema de Fisher obtenemos que la cuasivarianza muestral, excepto constantes, se distribuye según una ji-cuadrado cuando la distribución poblacional es normal.

Bajo esta premisa, la siguiente tabla presenta la región crítica, dependiendo de las hipótesis planteadas. Las zonas de rechazo se han obtenido aplicando el test de la razón de verosimilitudes.

Condiciones	H_0	H_1	Región Crítica	Estadístico
Normalidad	$\sigma=\sigma_0$	$\sigma<\sigma_0$	$\chi^2 < \chi^2_{n-1;1-\alpha}$	$\chi^2 = \frac{(n-1)s_1^2}{\sigma_0^2} = \frac{ns^2}{\sigma_0^2}$
	$\sigma\geq\sigma_0$	$\sigma<\sigma_0$	$\chi^2 < \chi^2_{n-1;1-\alpha}$	
	$\sigma=\sigma_0$	$\sigma>\sigma_0$	$\chi^2 > \chi^2_{n-1;\alpha}$	
	$\sigma\leq\sigma_0$	$\sigma>\sigma_0$	$\chi^2 > \chi^2_{n-1;\alpha}$	
	$\sigma=\sigma_0$	$\sigma\neq\sigma_0$	$\chi^2 < \chi^2_{n-1;1-\alpha/2} \cup \chi^2 > \chi^2_{n-1;\alpha/2}$	

6.6 RELACIÓN ENTRE LOS CONTRASTES DE HIPÓTESIS Y LOS INTERVALOS DE CONFIANZA

Ambas son técnicas inferenciales cuyos objetivos y estrategias de resolución son diferentes pero que pueden aportar unas mismas conclusiones, siempre bajo ciertas garantías o afirmaciones probabilísticas.

En primer lugar, el objetivo de un intervalo de confianza es estimar o aproximar el valor de un parámetro desconocido mientras que un contraste pretende confirmar o rechazar un supuesto o una hipótesis previa del investigador. Además, el intervalo de confianza utiliza sólo la información muestral y un contraste se basa en la información contenida en la hipótesis planteada. Así, por ejemplo, un contraste tendrá en cuenta la información dada por la hipótesis de trabajo mientras que un intervalo de confianza sólo utilizará la información previa y los datos de la muestra.

Sin embargo, desde el punto de vista del investigador, ambas técnicas proporcionan información para formular las conclusiones. En concreto, el intervalo de confianza contiene todos los valores que el investigador puede creer ciertos a priori y que no serán rechazados por un contraste de hipótesis bilateral que toma como hipótesis nula un valor concreto del parámetro. Por lo tanto, si un investigador cree en un valor del parámetro y no está contenido en el intervalo de confianza entonces tendremos evidencias para rechazarlo.

Por ejemplo, supongamos que tenemos una variable aleatoria poblacional cuya distribución es normal y con desviación típica σ conocida. Supongamos que el investigador cree que el valor de la media poblacional es μ_0, nos preguntamos: ¿qué valores de μ_0 no serán rechazados por un contraste de hipótesis? Planteado el contraste de hipótesis:

$$\begin{gathered} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{gathered}$$

La región de aceptación viene expresada por:

$$-z_{\alpha/2} \leq Z = \frac{\overline{X} - \mu_0}{\sigma}\sqrt{n} \leq z_{\alpha/2}$$

Si despejamos μ_0 para responder a la cuestión, descubrimos que los valores no rechazados son:

$$\mu_0 \in \left(\overline{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \overline{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \right) = IC_{1-\alpha}(\mu)$$

El resultado son los valores que están incluidos en el intervalo de confianza correspondiente. Así pues, desde un punto de vista instrumental los resultados conseguidos son válidos, aunque conceptualmente son dos técnicas diferentes porque el Intervalo de Confianza no utiliza la información contenida en la hipótesis planteada. Por lo tanto, cada técnica debe utilizarse en el contexto adecuado.

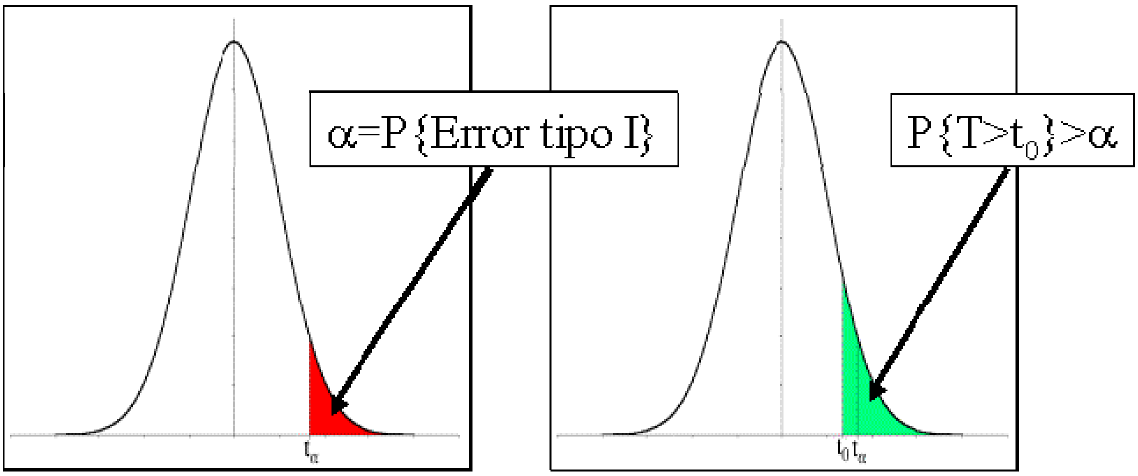
6.7 NIVEL DE SIGNIFICACIÓN VERSUS P-VALOR

En el planteamiento del contraste de hipótesis se debe fijar el nivel de significación α, cuestión que resuelve el investigador según su grado de creencia en la hipótesis nula y en las posibles consecuencias al rechazar dicha hipótesis. Sin embargo, parece más lógico dejar que los propios datos muestren su compatibilidad o no con la hipótesis nula planteada. En este contexto, surge el concepto de p-valor que se define como el valor del nivel de significación tal que para valores inferiores a él, la decisión del contraste sería aceptar H_0 y para valores superiores, la decisión del contraste sería rechazarla. Así si la región crítica del contraste es de la forma $T>t_\alpha$, el pvalor vendría dado por la expresión

$$\text{p} - \text{valor} = P\{T > t_{obs} | \text{cierta } H_0\}$$

siendo t_{obs} el valor del estadístico para la muestra observada.

De esta forma, un p-valor muy pequeño indicará que casi es imposible observar un valor del estadístico peor que el dado siendo cierta la hipótesis nula y, por lo tanto, tenemos evidencia para rechazar la hipótesis nula. Por el contrario, un p-valor alto indicará que es un valor probable bajo la hipótesis nula y, por lo tanto, no podemos dudar de ella. En el gráfico siguiente se muestra el planteamiento habitual frente al razonamiento utilizando el p-valor:



Las ventajas de este nuevo razonamiento radica en:

- Sólo utilizamos la información muestral y no tenemos que fijar a priori el nivel de significación.
- Permite comparar y cuantificar diferentes contrastes
- Conocemos la "significatividad" de nuestra conclusión

Sin embargo hay que tener claro que:

- Un pvalor es una medida calculada ex–post pero basada en una probabilidad calculada ex–ante
- No es una probabilidad a posteriori

- Está calculado suponiendo cierta la hipótesis nula y, por tanto, no da información clara sobre la hipótesis alternativa.
- Por último, no hay reglas pero existe un acuerdo sobre la actuación con respecto al p-valor:

Si $p<0,10$ entonces tenemos evidencia límite o muy débil en contra de H_0

Si $p<0,05$ entonces tenemos evidencia razonable o débil en contra de H_0

Si $p<0,025$ entonces tenemos evidencia fuerte en contra de H_0

Si $p<0,01$ entonces tenemos evidencia muy fuerte en contra de H_0