Chapter 6

Tests de Hipótesis

6.1 Introducción

El test de hipótesis es una manera formal de decidir entre dos opciones, o sea, es una manera de distinguir entre distribuciones de probabilidad en base a variables aleatorias generadas por una de ellas. Veamos un ejemplo para tener una idea de lo que significan.

Ejemplo 1. Supongamos que un comerciante debe comprar un cargamento de N manzanas. El comerciante ignora qué parte del cargamento no se encuentra en buen estado. Como inspeccionar todo el cargamento es muy costoso, decide elegir al azar una muestra de n manzanas.

Sea X el número de manzanas en mal estado que observa en la muestra. Luego si D es el número de manzanas en mal estado que hay en el cargamento, se tiene que la distribución de X es hipergeométrica y su función de probabilidad puntual está dada por

$$p(x, D) = \frac{\binom{D}{x}\binom{N-D}{n-x}}{\binom{N}{n}} \quad \text{si } \max(0, D-N+n) \leq x \leq \min(n, D)$$

y D puede tomar valores en el conjunto $\Theta = \{0, 1, 2, \ldots, N\}$.

Supongamos que se hubiese convenido que el cargamento debería tener no más de D_0 manzanas en mal estado. Luego, en base a la variable X, que el comerciante observa, debe decidir si el cargamento satisface los requisitos convenidos. Es decir, debe decidir entre dos alternativas

$$D \in \Theta_1 = \{0, 1, \ldots, D_0\} \qquad \text{o} \qquad D \in \Theta_2 = \{D_0 + 1, \ldots, N\}$$

Esto puede ser expresado como que el comerciante debe decidir entre dos hipótesis:

$$\text{H} : D \in \Theta_1 \qquad \text{contra} \qquad \text{K} : D \in \Theta_2$$

y esta decisión debe hacerla a partir del valor observado X.

Un test será una regla de decisión basada en X. Esto puede ser expresado matemáticamente como una función $\varphi(X)$ que toma dos valores: 1 y 0. 1 significará que rechaza H y por lo tanto acepta K y 0 que acepta H.

Supongamos por ejemplo que $N = 1000$, $n = 100$ y $D_0 = 150$. Un posible test está dado por:

$$\varphi_1(X) = \begin{cases} 1 & \text{si } X > 15 \\ 0 & \text{si } X \leq 15 \;. \end{cases}$$

De acuerdo con este test se rechaza el cargamento, es decir, se decide que $D \in \Theta_2$ si se observa en la muestra más de 15 manzanas en mal estado. Si se quisiera usar un test más seguro para el comprador (en el sentido de que la probabilidad de aceptar un cargamento con más de 150 manzanas en mal estado sea menor) se podría usar, por ejemplo,

$$\varphi_2(X) = \begin{cases} 1 & \text{si } X > 5 \\ 0 & \text{si } X \leq 5 \;. \end{cases}$$

Por ahora, no tenemos ningún criterio para elegir entre dos tests, ni entre los muchos otros que podrían definirse. En los párrafos siguientes atacaremos el problema de definir criterios para comparar diferentes tests, y el de elegir un test óptimo.

Ejemplo 2. Supongamos que para curar una determinada enfermedad se emplea una droga que cura la enfermedad con una probabilidad θ_0 conocida. Se ha obtenido una nueva droga y se quiere determinar si vale la pena cambiar la droga. Para ello se prueba la nueva droga con n pacientes obteniéndose los resultados $X_1, \ldots, X_n$, donde $X_i = 1$ indica que el $i-$ésimo paciente se curó y $X_i = 0$, que no se curó. Sea θ la probabilidad de curar de la nueva droga, la cual no es conocida.

Se está dispuesto a cambiar de droga si la nueva droga es tal que $\theta \geq \theta_0 + 0.05$, es decir si esta última cura al menos un 5% más de pacientes que la vieja. Luego, se tiene que decidir entre dos hipótesis:

$$\mathrm{H} : \theta \leq \theta_0 + 0.05 \qquad \text{y} \qquad \mathrm{K} : \theta > \theta_0 + 0.05$$

Un test será una función $\varphi(X_1, \ldots, X_n)$ que toma valores 1 ó 0. $\varphi(X_1, \ldots, X_n) = 0$ indicará que aceptamos H, es decir, no se continúa usando la droga vieja.

Para ejemplificar, supongamos que $\theta_0 = 0.8$ y $n = 100$. Un posible test sería

$$\varphi(X_1, \ldots, X_n) = \begin{cases} 1 & \text{si } \sum_{i=1}^{100} X_i \geq 85 \\ \\ 0 & \text{si } \sum_{i=1}^{100} X_i < 85 \,. \end{cases}$$

Este test acepta K, es decir, cambia de droga si 85 pacientes o más resultan curados.

Si se quisiera ser más conservador, es decir, estar más seguro que la droga tiene la probabilidad de curar mayor que 0.85 antes de tomar la decisión de cambiarla, se podría usar el test

$$\varphi(X_1, \ldots, X_n) = \begin{cases} 1 & \text{si } \sum_{i=1}^{100} X_i \geq 90 \\ \\ 0 & \text{si } \sum_{i=1}^{100} X_i < 90 \,. \end{cases}$$

6.2 Formulación general del problema del test de hipótesis

Supongamos que se obtiene un vector aleatorio $\mathbf{X} = (X_1, ..., X_n)$ cuya función de distribución pertenece a la familia $F(\mathbf{x}, \boldsymbol{\theta})$ con $\boldsymbol{\theta} \in \Theta \subset I\!R^p$. Sean Θ_1 y Θ_2 tales que $\Theta_1 \cap \Theta_2 = \emptyset$ y $\Theta_1 \cup \Theta_2 = \Theta$. Un test para este problema será una regla basada en $\mathbf{X}$ para decidir entre las dos hipótesis

$$\mathrm{H} : \theta \in \Theta_1 \qquad \text{contra} \qquad \mathrm{K} : \theta \in \Theta_2$$

Definición 1. Se llama *test* a una función $\varphi : I\!R^n \to [0, 1]$.

Se dice que un test φ es *no aleatorizado* si toma solamente los valores 0 ó 1.

Cuando $\varphi(\mathbf{X}) = 1$ se rechazará la hipótesis H y por lo tanto, se aceptará K. En cambio, $\varphi(\mathbf{X}) = 0$ indicará que se acepta H.

Si el test toma valores distintos de 0 y 1 se dice que es un *test aleatorizado*. En este caso, el valor $\varphi(\mathbf{x})$ indica con que probabilidad se rechaza H si se observa $\mathbf{X} = \mathbf{x}$, es decir, $\varphi(\mathbf{x}) = P(\text{rechazar H}|\mathbf{X} = \mathbf{x})$

Por ejemplo, $\varphi(\mathbf{X}) = 1/2$ indicará que si observamos el vector $\mathbf{X}$ debemos rechazar H con probabilidad $1/2$, es decir, podríamos tirar una moneda y si saliera ceca aceptarla, $\varphi(\mathbf{X}) = 1/6$ indicará que si observamos $\mathbf{X}$ debemos rechazar H con probabilidad $1/6$; en este caso podríamos tirar un dado; si saliese 1 rechazaríamos H y en los demás casos la aceptaríamos.

La aleatorización introduce en la decisión un elemento extraño al fenómeno estudiado, como el lanzamiento de una moneda o un dado, con que hemos ejemplificado. Por lo tanto, se evitan en lo posible los tests aleatorizados en los casos prácticos. Sin embargo, desde el punto de vista teórico, conviene como se verá, admitir la posibilidad de tests aleatorizados.

En la mayoría de las situaciones, los tests vienen dados como funciones de un estadístico, llamado *estadístico del test*, que, por ejemplo, como en el caso de la sección anterior, sirven para rechazar H para valores grandes. En general, el estadístico del test sirve para medir la diferencia entre los datos y lo que se espera de ellos bajo H.

Definición 2. La *región crítica* $\mathcal{R}$, de un test φ, es el conjunto de puntos $\mathbf{X}$ que llevan a la decisión de rechazar H y la *región de aceptación* $\mathcal{A}$ es el conjunto de puntos $\mathbf{X}$ que llevan a aceptar H.

Dado un test para un problema de test de hipótesis se podrá incurrir en dos tipos de error.

Definición 3. Se llamará *error de tipo 1* al que se comete al rechazar la hipótesis H, cuando es verdadera. Se llamará *error de tipo 2* al que se comete al aceptar H, cuando esta hipótesis es falsa.

Luego, para un test no aleatorizado, la probabilidad de cometer un error de tipo 1 será $P_{\boldsymbol{\theta}}(\mathcal{R})$, cuando $\boldsymbol{\theta} \in \Theta_1$. Mientras que la probabilidad de error de tipo 2, será $P_{\boldsymbol{\theta}}(\mathcal{A}) = 1 - P_{\boldsymbol{\theta}}(\mathcal{R})$, cuando $\boldsymbol{\theta} \in \Theta_2$.

Ejemplo 1 (donde se visualiza la necesidad de introducir tests aleatorizados). Supongamos que una empresa automotriz sostiene que domina la mitad del mercado, esto es que la mitad de los compradores de automóviles se deciden por alguno de los modelos fabricados por ella. Se desea testear si la afirmación hecha por la empresa es exagerada o no.

Supongamos que se toma una muestra de compradores que, para facilidad en los cálculos, consideraremos de tamaño $n = 6$.

Las hipótesis en cuestión son:

$$\mathrm{H} : \theta = 1/2 \qquad \text{contra} \qquad \mathrm{K} : \theta < 1/2$$

donde θ es la probabilidad de que un comprador tomado al azar compre un automóvil de la empresa.

Consideremos para cada comprador i, la variable X_i tal que $X_i = 1$ si el comprador se decide por un auto fabricado por la empresa; $X_i = 0$ en caso contrario. Luego, cada X_i tendrá distribución $Bi(\theta, 1)$.

Supongamos también que se quiere tener una probabilidad de error de tipo 1 de 0.25, es decir que la probabilidad de rechazar H cuando es verdadera es del 25%.

Parecería intuitivo considerar un test de la forma

$$\varphi_k(\mathbf{X}) = \begin{cases} 1 & \text{si } X < k \\ 0 & \text{si } X \geq k \end{cases}$$

Consideremos los test φ_2 y φ_3. Veamos que ninguno de ellos satisface la exigencia planteada para el error de tipo 1.

Suponiendo que las decisiones de los compradores son independientes entre sí, $T = \sum_{i=1}^{6} X_i$, tiene distribución $Bi(\theta, 6)$.

Calculemos la probabilidad de error de tipo 1 para ambos tests. Para ello usaremos la tabla de la distribución $Bi(6, 1/2)$.

t	0	1	2	3	4	5	6
$P_{\frac{1}{2}}(T = t)$	1/64	6/64	15/64	20/64	15/64	6/64	1/64

Por lo tanto,

$$P_{\frac{1}{2}}(\varphi_2 = 1) = P_{\frac{1}{2}}(T < 2) = 7/64 < 0.25$$

y

$$P_{\frac{1}{2}}(\varphi_3 = 1) = P_{\frac{1}{2}}(T < 3) = 22/64 > 0.25$$

Resulta claro entonces que no podremos elegir un test en la familia de tests no aleatorizados φ_k con un error de tipo 1 igual a 0.25.

Tendría sentido, en esta situación, plantearse un test de la forma

$$\varphi(\mathbf{X}) = \begin{cases} 1 & \text{si } T < 2 \\ \gamma & \text{si } T = 2 \\ 0 & \text{si } T > 2 \end{cases}$$

y tratar de elegir γ de forma tal que tenga el error de tipo I deseado. Para eso se requiere

$$P_{\frac{1}{2}}(\varphi(\mathbf{X}) = 1) = P_{\frac{1}{2}}(T < 2) + \gamma\, P_{\frac{1}{2}}(T = 2) = 0.25\ .$$

Luego, se deberá cumplir

$$\frac{7}{64} + \gamma\,\frac{15}{64} = 0.25,$$

o sea $\gamma = 3/5$.

Una forma de efectivizar el test, en este caso, podría ser la siguiente. Cuando se observa que $T < 2$, se rechaza H; cuando se observa que $T > 2$, se acepta H; cuando se observa $T = 2$ se colocan en una urna tres bolillas rojas y dos bolillas negras y se extrae una al azar; si resulta roja se rechaza H y si no se acepta.

Notemos que si en lugar de pedir que la probabilidad de error de tipo 1 sea 0.25 hubiésemos pedido que fuera 0.10; el test hubiera resultado de la forma

$$\varphi^*(\mathbf{X}) = \begin{cases} 1 & \text{si } T < 1 \\ 0.9 & \text{si } T = 1 \\ 0 & \text{si } T > 1 \end{cases}$$

O sea, cuanto más exigentes somos respecto del error de tipo 1, más estricta es la cota dada para el estadístico del test.

Debemos destacar que en este ejemplo, y en los anteriores, el test se basa en un estadístico cuya distribución es conocida cuando H es cierta. Conocer esa distribución hace posible definir la región de rechazo que tendrá probabilidad α prefijada bajo H. El valor elegido como cota o punto de corte, para tomar la decisión, se llama *valor crítico* y por lo tanto, separa la región de aceptación de la región de rechazo.

Volvamos al problema general de test de hipótesis planteado al comienzo de esta sección. Sea H : $\boldsymbol{\theta} \in \Theta_1$ y K : $\boldsymbol{\theta} \in \Theta_2$; sea $\varphi(\mathbf{X})$ un test para estas dos hipótesis. Entonces

Definición 4. Se llama *función de potencia del test* $\varphi(\mathbf{X})$ a la función

$$\beta_\varphi(\boldsymbol{\theta}) = P_{\boldsymbol{\theta}}(\text{rechazar H}),$$

donde $P_{\boldsymbol{\theta}}$ indica la probabilidad cuando $\boldsymbol{\theta}$ es el valor verdadero.

En el caso que φ es un test no aleatorizado se tiene

$$\beta_\varphi(\boldsymbol{\theta}) = P_{\boldsymbol{\theta}}(\varphi(\mathbf{X}) = 1) = E_{\boldsymbol{\theta}}(\varphi(\mathbf{X}))\ .$$

Si φ es aleatorizado, $\varphi(\mathbf{X})$ puede interpretarse como la probabilidad de rechazar H, condicional a observar $\mathbf{X}$; luego se tiene

$$\varphi(\mathbf{X}) = P(\text{rechazar H}|\mathbf{X})$$

y resulta

$$\beta_\varphi(\boldsymbol{\theta}) = P_{\boldsymbol{\theta}}(\text{ rechazar H}) = E_{\boldsymbol{\theta}}(P(\text{ rechazar H}|\mathbf{X})) = E_{\boldsymbol{\theta}}(\varphi(\mathbf{X}))\ .$$

Por lo tanto, en todos los casos se tiene

$$\beta_\varphi(\boldsymbol{\theta}) = E_{\boldsymbol{\theta}}(\varphi(\mathbf{X}))\ .$$

Expresemos ahora las probabilidades de los errores de un test en términos de $\beta_\varphi(\boldsymbol{\theta})$

- La probabilidad que ocurra un error de tipo 1 será $\beta_\varphi(\boldsymbol{\theta})$ para $\boldsymbol{\theta} \in \Theta_1$.
- La probabilidad que ocurra un error de tipo 2 será $(1 - \beta_\varphi(\boldsymbol{\theta}))$ para $\boldsymbol{\theta} \in \Theta_2$.

Un buen test deberá tener errores de tipo 1 y 2 pequeños, y por lo tanto debe tener una función de potencia $\beta_\varphi(\boldsymbol{\theta})$ que tome valores cercanos a 0 para $\boldsymbol{\theta} \in \Theta_1$ y valores cercanos a 1 para $\boldsymbol{\theta} \in \Theta_2$.

En realidad, no podemos hacer ambos errores pequeños simultáneamente. Más aún, para un tamaño de muestra dado para que decrezca la probabilidad de que ocurra un error de tipo 1, debemos aumentar la probabilidad de que ocurra un error de tipo 2 (o sea disminuir la potencia). Si queremos que ambos sean pequeños debemos aumentar la cantidad de observaciones.

Por ejemplo, en el Ejemplo 1, el test φ^* cumplía $\beta_{\varphi^*}(1/2) = 0.10$. Por otra parte, se verifica que $\beta_{\varphi^*}(\theta) = (1-\theta)^6 + 5.4\,\theta(1-\theta)^5$, con lo cual tenemos la tabla siguiente que da la función de potencia del test φ^* para algunos valores de $\theta \in [0, 1/2]$

θ	0	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5
$\beta_{\varphi^*}(\theta)$	1	0.944	0.85	0.736	0.616	0.498	0.389	0.295	0.215	0.149	0.1

Como vemos, la función de potencia de φ^* es una función decreciente de θ en el intervalo $[0, 1/2]$ que tiende a 1 cuando $\theta \to 0$ y tiende a 0.1 cuando $\theta \to 1/2$. Es decir, la probabilidad de error 2 tiende a 0 cuando $\theta \to 0$ y por lo tanto, se logran detectar bien alternativas lejanas a la hipótesis H.
Para los procedimientos que daremos $1-P(\text{error de tipo 1}) \geq P(\text{error de tipo 2})$. El objetivo será encontrar procedimientos con la menor probabilidad de tipo 2, fijada la probabilidad de tipo 1, es decir, buscaremos procedimientos con potencia grande para $\boldsymbol{\theta} \in \Theta_2$.

6.2.1 Nivel de significación de un test

La teoría clásica de test de hipótesis considera que el error de tipo 1 es mucho más grave que el error de tipo 2. Es decir, la situación de las hipótesis H y K no es simétrica; es mucho más grave rechazar H cuando es cierta que aceptarla cuando es falsa. Esto significa que se debe tener mucha evidencia sobre que H es falsa antes de rechazarla. Se observa en el Ejemplo 2 de la sección 1, que esta simetría corresponde a una situación real, puesto que antes de cambiar de droga, es decir rechazar H, habría que tener un grado de certeza muy alto respecto de que la nueva droga es mejor que la primera. Desde ahora en adelante H se denominará *hipótesis nula* y K *hipótesis alternativa.*

Veamos un ejemplo que servirá para fijar ideas y clarificar la mecánica de elección de H

Ejemplo 1. Supongamos que se quiere decidir si un paciente tiene o no tuberculosis, para proceder, en caso afirmativo, a suministrarle un tratamiento adecuado. Tendremos entonces dos hipótesis:

(A) El señor W está tuberculoso;

(B) El señor W no está tuberculoso.

Es claro que el médico responsable de la decisión considerará mucho más grave rechazar (A) cuando es cierta, que rechazar (B) cuando es cierta (esto es lo mismo que aceptar H cuando es falsa), puesto que en el primer caso se expone al paciente a una agudización grave de su enfermedad, mientras que en el segundo se le aplicará un tratamiento que no necesita y cuyas consecuencias nunca serán comparables con el daño de no tratarlo estando enfermo.

Luego la hipótesis nula será H : "El señor W está tuberculoso"; y la alternativa K : "El señor W no está tuberculoso".

Como dijimos más arriba, supondremos que el error de tipo 1 (rechazar H cuando es cierta), es el más grave. Por lo tanto se va requerir que el error de tipo 1 del test a utilizar no sea mayor que un número $0 < \alpha < 0.5$ prefijado. Este número α es generalmente pequeño (entre 0.01 y 0.10) y se lo determina de acuerdo a la importancia del error de tipo 1. La siguiente definición formaliza este concepto.

Definición 5. El *nivel de significación* de un test φ está definido por

$$\alpha = \sup_{\boldsymbol{\theta} \in \Theta_1} \beta_{\boldsymbol{\theta}}(\varphi)$$

Luego, α es el supremo de la probabilidad de cometer un error de tipo 1.

Por lo tanto, fijado α, se buscará un test que tenga nivel de significación menor o igual que α. Un test con está propiedad asegurará que la probabilidad de rechazar la hipótesis nula H, cuando esta es cierta, no sea mayor que α.

Como existen muchos tests que tienen nivel de significación menor o igual que α para un problema determinado, debemos dar un criterio para elegir uno entre todos ellos. Resulta natural elegir entre todos los tests con la restricción de que su nivel de significación sea menor o igual que α aquel que tenga menor probabilidad de error de tipo 2. Esto motiva la siguiente definición.

Definición 6. Consideremos un problema general de test de hipótesis donde se observa un vector $\mathbf{X}$ con distribución $F(x, \boldsymbol{\theta})$,con $\boldsymbol{\theta} \in \Theta$, y se tiene que decidir entre las hipótesis H: $\boldsymbol{\theta} \in \Theta_1$ y K: $\boldsymbol{\theta} \in \Theta_2$. Diremos que un test φ es el test *más potente* de nivel menor o igual que α para una alternativa fija $\boldsymbol{\theta}_2 \in \Theta_2$ si

(a) $\sup_{\boldsymbol{\theta} \in \Theta_1} \beta_\varphi(\boldsymbol{\theta}) \leq \alpha$, es decir si φ tiene nivel de significación menor o igual que α

(b) Dado otro test φ^* de nivel menor o igual que α entonces se tiene

$$\beta_{\varphi^*}(\boldsymbol{\theta}_2) \leq \beta_\varphi(\boldsymbol{\theta}_2)$$

Es decir, la probabilidad de error cuando $\boldsymbol{\theta}_2$ es el verdadero valor es menor para el test φ que para cualquier otro φ^* de nivel menor o igual que α (o sea, $(1-\beta_\varphi(\boldsymbol{\theta}_2)) \leq (1-\beta_{\varphi^*}(\boldsymbol{\theta}_2))$).

Es claro que si cambiamos la alternativa $\boldsymbol{\theta}_2 \in \Theta_2$ por otro $\boldsymbol{\theta}_2' \in \Theta_2$, el test más potente para esta $\boldsymbol{\theta}_2'$ no tiene porque coincidir con el correspondiente a $\boldsymbol{\theta}_2$. Por ejemplo, si se quiere testear H : $\mu = \mu_0$ contra K : $\mu \neq \mu_0$, para una distribución $N(\mu, \sigma_0^2)$ con σ_0^2 conocida, resultará

$$\Theta_1 = \{\mu_0\} \quad ; \quad \Theta_2 = \{\mu \in I\!R : \mu \neq \mu_0\}.$$

Si se toma una alternativa fija $\mu_1 < \mu_0$, el test más potente de nivel α para esta alternativa no coincide con el test mas potente para una alternativa $\mu_2 > \mu_0$, como veremos más adelante.

Definición 7. Diremos que un φ es un test *uniformemente más potente*, UMP, de nivel menor o igual que α para H : $\boldsymbol{\theta} \in \Theta_1$ contra K : $\boldsymbol{\theta} \in \Theta_2$, si φ es el más potente de nivel menor o igual que α para todo $\boldsymbol{\theta}_2 \in \Theta_2$, es decir, si el mismo test es óptimo *cualquiera* sea la alternativa fija $\boldsymbol{\theta}_2 \in \Theta_2$ considerada.

Lo ideal sería encontrar (cuando existan) tests uniformemente más potentes de nivel menor o igual que α. Estudiaremos casos donde estos tests existen y otros donde no. En estos últimos habrá que elegir otros criterios para seleccionar el test a usar.

Definición 8. El *nivel crítico o p-valor* es el menor valor de significación para el que rechazamos la hipótesis H para una observación dada $\mathbf{x}$.

En el Ejemplo 1 de la sección 2, por ejemplo si observamos $X = 2$ el p-valor del test $\{\varphi_k\}$ que rechaza para valores pequeños de T, será $p = 7/64$.

Prefijado el nivel de significación α, y evaluado el p- valor, p, del test utilizado, rechazaremos H si $p < \alpha$.

A esta altura, la lógica de los tests puede parecer más clara. Es un argumento por contradicción destinado a mostrar que la hipótesis nula lleva a conclusiones absurdas y que por lo tanto, debe ser rechazada.

Supongamos que para un conjunto de datos dado, se evalúa el estadístico del test y se obtiene un p–valor de 0.001. Para interpretarlo, *debemos pensar que la hipótesis nula es cierta* e imaginamos a otros investigadores repitiendo la experiencia en idénticas condiciones. El valor 0.001 dice que sólo un investigador de cada 1000 puede obtener un valor del estadístico tan extremo como el obtenido. Por lo tanto, la diferencia entre los datos y lo que se espera de ellos bajo H no puede atribuirse meramente a variación aleatoria. Este hecho lleva a una contradicción y por lo tanto, a abandonar nuestra hipótesis de que H era cierta.

Es tentador pensar que el p–valor da la probabilidad de que H sea cierta, pero no es así. No importa cuántas veces se repita el experimento, H será siempre cierta o siempre falsa. Es decir, el nivel crítico da la probabilidad de obtener evidencia en contra de la hipótesis nula suponiendo que ésta sea cierta. Por lo tanto, cuanto menor sea el p-valor más evidencia en contra de H tenemos, suponiendo que H es cierta.

6.3 Tests óptimos para el caso de hipótesis simple contra hipótesis simple

El caso más simple de problema de test de hipótesis es la situación donde Θ_1 y Θ_2 contengan cada uno un elemento. En este caso, se dice, H y K son *hipótesis simples*.

Si Θ_1 tuviera más de un elemento, H se llamará *hipótesis compuesta*, y lo mismo vale para K en relación a Θ_2.

En el caso en que H y K sean simples, un problema de test de hipótesis será de la forma

$$\mathrm{H} : \boldsymbol{\theta} = \boldsymbol{\theta}_1 \qquad \text{contra} \qquad \mathrm{K} : \boldsymbol{\theta} = \boldsymbol{\theta}_2$$

Supongamos que $\mathbf{X}$ sea un vector discreto (o continuo) bajo $\boldsymbol{\theta}_1$ y $\boldsymbol{\theta}_2$ y que las funciones de densidad correspondientes sean $p(\mathbf{x}, \boldsymbol{\theta}_1)$ y $p(\mathbf{x}, \boldsymbol{\theta}_2)$. Luego, intuitivamente, parece razonable rechazar H si la "probabilidad" de obtener el valor observado $\mathbf{x}$ bajo $\boldsymbol{\theta}_2$ es grande comparada con la "probabilidad" de obtener $\mathbf{x}$ bajo $\boldsymbol{\theta}_1$, es decir, cuando

$$L_{21} = \frac{p(\mathbf{x}, \boldsymbol{\theta}_2)}{p(\mathbf{x}, \boldsymbol{\theta}_1)} \geq k_\alpha$$

donde k_α es una constante que depende del nivel α. Por lo tanto, se podría pensar en construir un test de la forma

$$\varphi(\mathbf{X}) = \begin{cases} 1 & \text{si } L_{21} > k_\alpha \\ \gamma_\alpha & \text{si } L_{21} = k_\alpha \\ 0 & \text{si } L_{21} < k_\alpha \end{cases}$$

o equivalentemente,

$$\varphi(\mathbf{X}) = \begin{cases} 1 & \text{si } p(\mathbf{x}, \boldsymbol{\theta}_2) > k_\alpha\, p(\mathbf{x}, \boldsymbol{\theta}_1) \\ \gamma_\alpha & \text{si } p(\mathbf{x}, \boldsymbol{\theta}_2) = k_\alpha\, p(\mathbf{x}, \boldsymbol{\theta}_1) \\ 0 & \text{si } p(\mathbf{x}, \boldsymbol{\theta}_2) < k_\alpha\, p(\mathbf{x}, \boldsymbol{\theta}_1) \end{cases} \tag{6.1}$$

donde $0 \leq \gamma_\alpha \leq 1$, correspondiendo el caso $k_\alpha = +\infty$ al test

$$\varphi(\mathbf{X}) = \begin{cases} 1 & \text{si } p(\mathbf{x}, \boldsymbol{\theta}_1) = 0 \\ 0 & \text{si } p(\mathbf{x}, \boldsymbol{\theta}_1) > 0 \end{cases} \tag{6.2}$$

que tiene nivel 0.

Si queremos que el test (6.1) tenga nivel α debemos elegir k_α y β_α tales que se cumpla

$$E_{\boldsymbol{\theta}_1}(\varphi(\mathbf{X})) = \alpha \; . \tag{6.3}$$

Notemos que entonces, en este tipo de test k_α es una función decreciente de α.

Un test de la forma (6.1) se llama *test del cociente de verosimilitud*. El siguiente teorema establece que se pueden elegir k_α y γ_α de manera que se cumpla (6.3) y que usando estos valores en (6.1) se obtiene un test más potente de nivel menor o igual que α. Sin embargo, los tests de la forma (6.1) no garantizan la unicidad y es por ello, que para obtenerla le permitiremos a γ_α depender de $\mathbf{x}$.

Teorema 1 (de Neyman–Pearson)

(i) *Dado $0 \leq \alpha \leq 1$ se pueden elegir k_α y γ_α, $0 \leq \gamma_\alpha \leq 1$, tales que el test de la forma (6.1) satisfaga (6.3).*

(ii) *Sea un test de la forma (6.1) que satisface (6.3) para $\alpha > 0$ y de la forma (6.2) para $\alpha = 0$. Luego ese test es el más potente de nivel menor o igual que α para*

$$H : \boldsymbol{\theta} = \boldsymbol{\theta}_1 \qquad \text{contra} \qquad K : \boldsymbol{\theta} = \boldsymbol{\theta}_2 \; .$$

(iii) Si φ^* es un test uniformemente más potente de nivel $\alpha > 0$ para $H : \boldsymbol{\theta} = \boldsymbol{\theta}_1$ versus $K : \boldsymbol{\theta} = \boldsymbol{\theta}_2$ entonces φ^* es de la forma

$$\varphi(\mathbf{X}) = \begin{cases} 1 & \text{si } p(\mathbf{x}, \boldsymbol{\theta}_2) > k_\alpha \, p(\mathbf{x}, \boldsymbol{\theta}_1) \\ \gamma_\alpha(\mathbf{x}) & \text{si } p(\mathbf{x}, \boldsymbol{\theta}_2) = k_\alpha \, p(\mathbf{x}, \boldsymbol{\theta}_1) \\ 0 & \text{si } p(\mathbf{x}, \boldsymbol{\theta}_2) < k_\alpha \, p(\mathbf{x}, \boldsymbol{\theta}_1) \end{cases} \tag{6.4}$$

excepto quizás en un conjunto $\mathcal{N}$ tal que $P_{\boldsymbol{\theta}_1}(\mathcal{N}) = P_{\boldsymbol{\theta}_2}(\mathcal{N}) = 0$.

Si φ^* es un test uniformemente más potente de nivel 0 para $H : \boldsymbol{\theta} = \boldsymbol{\theta}_1$ versus $K : \boldsymbol{\theta} = \boldsymbol{\theta}_2$ entonces φ^* es de la forma (6.2) excepto quizás en un conjunto $\mathcal{N}$ tal que $P_{\boldsymbol{\theta}_1}(\mathcal{N}) = P_{\boldsymbol{\theta}_2}(\mathcal{N}) = 0$.

DEMOSTRACIÓN: (i) Si $\alpha = 0$ el test (6.2) tiene nivel 0. Sea entonces, $0 < \alpha \leq 1$.

Extendamos la definición de la variable aleatoria L_{21} al caso en que el denominador es 0,

$$L_{21} = \begin{cases} \frac{p(\mathbf{x},\boldsymbol{\theta}_2)}{p(\mathbf{x},\boldsymbol{\theta}_1)} & \text{si } p(\mathbf{x},\boldsymbol{\theta}_1) > 0 \\ 1 & \text{si } p(\mathbf{x},\boldsymbol{\theta}_1) = 0 \end{cases} .$$

Luego,

$$\begin{aligned} E_{\boldsymbol{\theta}_1}(\varphi(\mathbf{X})) &= P_{\boldsymbol{\theta}_1}(L_{21} > k_\alpha) + \gamma\, P_{\boldsymbol{\theta}_1}(L_{21} = k_\alpha) \\ &= 1 - P_{\boldsymbol{\theta}_1}(L_{21} \leq k_\alpha) + \gamma_\alpha\, P_{\boldsymbol{\theta}_1}(L_{21} = k_\alpha) . \end{aligned}$$

Si existe una constante k_0 tal que $P_{\boldsymbol{\theta}_1}(L_{21} \leq k_0) = \alpha$ tomamos $k_\alpha = k_0$ y $\gamma_\alpha = 0$. En caso contrario, siempre existe k_0 tal que

$$P_{\boldsymbol{\theta}_1}(L_{21} < k_0) \leq 1 - \alpha < P_{\boldsymbol{\theta}_1}(L_{21} \leq k_0) \tag{6.5}$$

y se cumple, $P_{\boldsymbol{\theta}_1}(L_{21} = k_0) > 0$. Definamos $k_\alpha = k_0$ y

$$\gamma_\alpha = \frac{P_{\boldsymbol{\theta}_1}(L_{21} \leq k_0) - (1-\alpha)}{P_{\boldsymbol{\theta}_1}(L_{21} = k_0)} .$$

Luego, por (6.5) $0 < \gamma_\alpha \leq 1$ y además $E_{\boldsymbol{\theta}_1}(\varphi(\mathbf{X})) = \alpha$.

Demostraremos (ii) en el caso continuo, el caso discreto es análogo reemplazando las integrales por sumatorias. Supongamos que φ sea de la forma (6.1) y satisfaga (6.3). Luego, por satisfacer (6.3) su nivel es igual a α.

Para mostrar que φ es el test más potente de nivel menor o igual que α, sólo falta mostrar que dado otro test φ^* de nivel menor o igual que α se tiene

$$\beta_\varphi(\boldsymbol{\theta}_2) \geq \beta_{\varphi^*}(\boldsymbol{\theta}_2) \tag{6.6}$$

(a) Supongamos primero $\alpha > 0$ con lo cual $k_\alpha < \infty$ en (6.1). Sea φ^* de nivel menor o igual que α. Consideremos la expresión

$$U(\mathbf{x}) = [\varphi(\mathbf{x}) - \varphi^*(\mathbf{x})]\,[p(\mathbf{x},\boldsymbol{\theta}_2) - k_\alpha\, p(\mathbf{x},\boldsymbol{\theta}_1)] . \tag{6.7}$$

Mostraremos que $U(\mathbf{x}) \geq 0$.

Supongamos primero que

$$p(\mathbf{x},\boldsymbol{\theta}_2) > k_\alpha\, p(\mathbf{x},\boldsymbol{\theta}_1) .$$

Luego, de acuerdo con (6.1), se tendrá $\varphi(\mathbf{x}) = 1$ y por lo tanto $\varphi(\mathbf{x}) \geq \varphi^*(\mathbf{x})$, de donde, $U(\mathbf{x}) \geq 0$.

Si $p(\mathbf{x}, \boldsymbol{\theta}_2) = k_\alpha \, p(\mathbf{x}, \boldsymbol{\theta}_1)$, es claro que $U(\mathbf{x}) = 0$.

Finalmente, si

$$p(\mathbf{x}, \boldsymbol{\theta}_2) < k_\alpha \, p(\mathbf{x}, \boldsymbol{\theta}_1) \ ,$$

entonces usando nuevamente (6.1), $\varphi(\mathbf{x}) = 0$, con lo cual $\varphi(\mathbf{x}) \leq \varphi^*(\mathbf{x})$ y por lo tanto $U(\mathbf{x}) \geq 0$.

Resulta entonces que

$$\int [\varphi(\mathbf{x}) - \varphi^*(\mathbf{x})] \, [p(\mathbf{x}, \boldsymbol{\theta}_2) - k_\alpha \, p(\mathbf{x}, \boldsymbol{\theta}_1)] \, d\mathbf{x} = \int U(\mathbf{x}) d\mathbf{x} \geq 0 \ .$$

Por lo tanto,

$$\int (\varphi(\mathbf{x}) - \varphi^*(\mathbf{x})) p(\mathbf{x}, \boldsymbol{\theta}_2) d\mathbf{x} \geq k_\alpha \int (\varphi(\mathbf{x}) - \varphi^*(\mathbf{x})) p(\mathbf{x}, \boldsymbol{\theta}_1) d\mathbf{x}$$

o equivalentemente,

$$\beta_\varphi(\boldsymbol{\theta}_2) - \beta_{\varphi^*}(\boldsymbol{\theta}_2) \geq k_\alpha (\beta_\varphi(\boldsymbol{\theta}_1) - \beta_{\varphi^*}(\boldsymbol{\theta}_1)) \ .$$

Por (6.3) se tiene $\beta_\varphi(\boldsymbol{\theta}_1) = \alpha$, como φ^* es un test de nivel de significación menor o igual que α, $\beta_{\varphi^*}(\boldsymbol{\theta}_1) \leq \alpha$, y entonces resulta

$$\beta_\varphi(\boldsymbol{\theta}_1) - \beta_{\varphi^*}(\boldsymbol{\theta}_1) \geq 0$$

con lo cual,

$$\beta_\varphi(\boldsymbol{\theta}_2) \geq \beta_{\varphi^*}(\boldsymbol{\theta}_2) \ .$$

Esto demuestra que φ es el test más potente de nivel de significación menor o igual que α si su nivel no es cero.

(b) Si $\alpha = 0$, como el test dado por (6.2) tiene nivel cero queremos ver que dado φ^* con nivel 0 se cumple (6.6). Como φ^* tiene nivel 0,

$$\int \varphi^*(\mathbf{x}) p(\mathbf{x}, \boldsymbol{\theta}_1) d\mathbf{x} = 0 \ .$$

Por lo tanto, $\varphi^*(\mathbf{x}) = 0$ en el conjunto $\{\mathbf{x} \, : \, p(\mathbf{x}, \boldsymbol{\theta}_1) > 0\}$ excepto quizás en un conjunto de medida 0. Por lo tanto, como $\varphi(\mathbf{x}) = 0$ si $p(\mathbf{x}, \boldsymbol{\theta}_1) > 0$ y $\varphi(\mathbf{x}) = 1$ si $p(\mathbf{x}, \boldsymbol{\theta}_1) = 0$ se tiene

$$
\begin{aligned}
\beta_\varphi(\boldsymbol{\theta}_2) - \beta_{\varphi^*}(\boldsymbol{\theta}_2) &= E_{\boldsymbol{\theta}_2}(\varphi(\mathbf{X})) - E_{\boldsymbol{\theta}_2}(\varphi^*(\mathbf{X})) \\
&= \int_{\{\mathbf{x}:\, p(\mathbf{x},\boldsymbol{\theta}_1)=0\}} [\varphi(\mathbf{X}) - \varphi^*(\mathbf{X})]\, p(\mathbf{x},\boldsymbol{\theta}_2) d\mathbf{x} \\
&\quad + \int_{\{\mathbf{x}:\, p(\mathbf{x},\boldsymbol{\theta}_1)>0\}} [\varphi(\mathbf{X}) - \varphi^*(\mathbf{X})]\, p(\mathbf{x},\boldsymbol{\theta}_2) d\mathbf{x} \\
&= \int_{\{\mathbf{x}:\, p(\mathbf{x},\boldsymbol{\theta}_1)=0\}} [1 - \varphi^*(\mathbf{X})]\, p(\mathbf{x},\boldsymbol{\theta}_2) d\mathbf{x} \geq 0 \,.
\end{aligned}
$$

(iii) Haremos primero el caso $\alpha = 0$. Sea φ el test de la forma (6.2) y φ^* un test de nivel 0. Hemos visto que entonces $\varphi^*(\mathbf{x}) = 0$ en el conjunto $\{\mathbf{x} : p(\mathbf{x},\boldsymbol{\theta}_1) > 0\}$ excepto quizás en un conjunto $\mathcal{N}_1$ de medida 0. Luego, $P_{\boldsymbol{\theta}_1}(\mathcal{N}_1) = P_{\boldsymbol{\theta}_2}(\mathcal{N}_1) = 0$ y $\varphi^*(\mathbf{x}) = \varphi(\mathbf{x})$ en $\{\mathbf{x} : p(\mathbf{x},\boldsymbol{\theta}_1) > 0\} - \mathcal{N}_1$.

Falta ver que $\varphi^*(\mathbf{x}) = \varphi(\mathbf{x}) = 1$ en $\{\mathbf{x} : p(\mathbf{x},\boldsymbol{\theta}_1) = 0\}$ excepto quizás un conjunto de medida 0. Como

$$E_{\boldsymbol{\theta}_2}(\varphi(\mathbf{X})) = E_{\boldsymbol{\theta}_2}(\varphi^*(\mathbf{X}))$$

se cumple

$$
\begin{aligned}
0 &= \int_{\{\mathbf{x}:\, p(\mathbf{x},\boldsymbol{\theta}_1)=0\}} [\varphi(\mathbf{X}) - \varphi^*(\mathbf{X})]\, p(\mathbf{x},\boldsymbol{\theta}_2) d\mathbf{x} \\
&\quad + \int_{\{\mathbf{x}:\, p(\mathbf{x},\boldsymbol{\theta}_1)>0\}} [\varphi(\mathbf{X}) - \varphi^*(\mathbf{X})]\, p(\mathbf{x},\boldsymbol{\theta}_2) d\mathbf{x} \\
&= \int_{\{\mathbf{x}:\, p(\mathbf{x},\boldsymbol{\theta}_1)=0\}} [1 - \varphi^*(\mathbf{X})]\, p(\mathbf{x},\boldsymbol{\theta}_2) d\mathbf{x} \,.
\end{aligned}
$$

Pero $\varphi^* \leq 1$ luego el integrando es no negativo y la integral es cero si y solo si $\varphi^* = 1$ en el conjunto $\{\mathbf{x} :\, p(\mathbf{x},\boldsymbol{\theta}_1) = 0\} \cap \{\mathbf{x} :\, p(\mathbf{x},\boldsymbol{\theta}_2) > 0\}$ excepto quizás en un conjunto $\mathcal{N}_2$ de medida 0. Luego si

$$\mathcal{N} = \mathcal{N}_1 \cup \mathcal{N}_2 \cup (\{\mathbf{x} :\, p(\mathbf{x},\boldsymbol{\theta}_1) = 0\} \cap \{\mathbf{x} :\, p(\mathbf{x},\boldsymbol{\theta}_2) = 0\})$$

se tiene $P_{\boldsymbol{\theta}_1}(\mathcal{N}) = P_{\boldsymbol{\theta}_2}(\mathcal{N}) = 0$ y $\varphi^*(\mathbf{x}) = \varphi(\mathbf{x})$ para $\mathbf{x} \notin \mathcal{N}$.

Supongamos ahora $\alpha > 0$. Sea φ^* un test de nivel α uniformemente más potente para H versus K y φ el test dado por (6.1) que también es uniformemente más potente para H versus K por lo visto en (ii). Luego se cumple

$$E_{\boldsymbol{\theta}_1}(\varphi(\mathbf{X})) = E_{\boldsymbol{\theta}_1}(\varphi^*(\mathbf{X})) \qquad \text{y} \qquad E_{\boldsymbol{\theta}_2}(\varphi(\mathbf{X})) = E_{\boldsymbol{\theta}_2}(\varphi^*(\mathbf{X})) \tag{6.8}$$

Por otra parte, la función $U(\mathbf{x})$ definida en (6.7) es no negativa y por (6.8) $\int U(\mathbf{x})d\mathbf{x} = 0$. Luego, $U(\mathbf{x})$ debe ser nula excepto en un conjunto $\mathcal{N}$ de medida 0. Es decir, $(\varphi(\mathbf{x}) - \varphi^*(\mathbf{x}))(p(\mathbf{x}, \boldsymbol{\theta}_2) - k_\alpha\, p(\mathbf{x}, \boldsymbol{\theta}_1)) = 0$ para $\mathbf{x} \notin \mathcal{N}$. Por lo tanto, $\varphi(\mathbf{x}) = \varphi^*(\mathbf{x})$ en el conjunto $\{\mathbf{x} : p(\mathbf{x}, \boldsymbol{\theta}_2) \neq k_\alpha\, p(\mathbf{x}, \boldsymbol{\theta}_1)\} \cap \mathcal{N}^c$ de donde el resultado.

Observación. Si L_{21} es una variable continua no hay que preocuparse por γ_α, ya que $P(L_{21} = k_\alpha) = 0$.

Ejemplo 1. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución perteneciente a $N(\mu, \sigma_0^2)$ donde σ_0^2 es conocido, y supongamos que se quiere decidir entre H : $\mu = \mu_1$ contra K : $\mu = \mu_2$. Supongamos primero que $\mu_2 > \mu_1$. En este caso, el test más potente rechaza H si

$$\frac{p(X_1, \ldots, X_n; \mu_2)}{p(X_1, \ldots, X_n; \mu_1)} \geq k_\alpha$$

donde $p(X_1, \ldots, X_n; \mu)$ indica la densidad conjunta de $\mathbf{X} = (X_1, \ldots, X_n)$ cuando X_i tiene distribución $N(\mu, \sigma_0^2)$. Luego, $\varphi(X_1, \ldots, X_n) = 1$ si

$$L_{21} = \frac{(2\,\pi\sigma_0)^{-n/2} e^{-\Sigma_{i=1}^n (X_i - \mu_2)^2 / 2\sigma_0^2}}{(2\,\pi\sigma_0)^{-n/2} e^{-\Sigma_{i=1}^n (X_i - \mu_1)^2 / 2\sigma_0^2}} \geq k_\alpha$$

o sea $\varphi(X_1, \ldots, X_n) = 1$ si

$$e^{-\Sigma_{i=1}^n (X_i - \mu_2)^2 / 2\sigma_0^2 + \Sigma_{i=1}^n (X_i - \mu_1)^2 / 2\sigma_0^2} \geq k_\alpha$$

o equivalentemente, $\varphi(X_1, \ldots, X_n) = 1$ si

$$-\sum_{i=1}^n (X_i - \mu_2)^2 + \sum_{i=1}^n (X_i - \mu_1)^2 \geq 2\sigma_0^2 \ln k_\alpha\ .$$

Desarrollando el primer miembro de esta desigualdad, se tiene que $\varphi(X_1, \ldots, X_n) = 1$ si

$$2(\mu_2 - \mu_1) \sum_{i=1}^n X_i \geq 2\sigma_0^2 \ln k_\alpha + n\mu_2^2 - n\mu_1^2\ .$$

Como $\mu_2 - \mu_1 > 0$, se tiene que $\varphi(X_1, \ldots, X_n) = 1$ si

$$\sum_{i=1}^n X_i \geq \frac{2\sigma_0^2 \ln k_\alpha + n\mu_2^2 - n\mu_1^2}{2(\mu_2 - \mu_1)}$$

pero el segundo miembro de esta desigualdad es una constante, llamémosla ${k'}_\alpha$.

Luego, el test más potente es de la forma

$$\varphi(X_1, \ldots, X_n) = 1 \qquad \text{si} \qquad \sum_{i=1}^{n} X_i \geq {k'}_\alpha$$

(puesto que las regiones de rechazo planteadas inicialmente y esta última son equivalentes). La constante ${k'}_\alpha$ deberá elegirse de modo que

$$E_{\mu_1}(\varphi(X_1, \ldots, X_n) = \alpha \tag{6.9}$$

Para encontrar el ${k'}_\alpha$ que hace que (6.9) se satisfaga, necesitaríamos una tabla de la distribución $N(n\mu_1, n\sigma_0^2)$, pero para trabajar más cómodamente transformamos el estadístico $\sum_{i=1}^{n} X_i$ en otro cuya distribución sea $N(0,1)$. Para esto escribimos el test de la siguiente forma $\varphi(X_1, \ldots, X_n) = 1$ si

$$\sqrt{n}\frac{(\overline{X}_n - \mu_1)}{\sigma_0} \geq \sqrt{n}\frac{({k'}_\alpha/n - \mu_1)}{\sigma_0}$$

donde $\overline{X}_n = (1/n)\sum_{i=1}^{n} X_i$. Nuevamente $\sqrt{n}({k'}_\alpha/n - \mu_1)/\sigma_0$ es una constante que llamaremos ${k''}_\alpha$. Luego el test puede ser escrito de la forma

$$\varphi(X_1, \ldots, X_n) = 1 \quad \text{si} \ \sqrt{n}\frac{(\overline{X}_n - \mu_1)}{\sigma_0} \geq {k''}_\alpha.$$

Calculemos ${k''}_\alpha$. De acuerdo con el Teorema de Neyman–Pearson, debería tenerse que

$$\begin{aligned}
\alpha &= E_{\mu_1}(\varphi(X_1, \ldots, X_n)) \\
&= P_{\mu_1}(\varphi(X_1, \ldots, X_n) = 1) \\
&= P_{\mu_1}(\sqrt{n}\frac{(\overline{X}_n - \mu_1)}{\sigma_0} \geq {k''}_\alpha) \ .
\end{aligned}$$

Pero cuando μ es μ_1, $\sqrt{n}(\overline{X}_n - \mu_1)/\sigma_0$ es $N(0,1)$. Luego, ${k''}_\alpha$ debe ser igual a z_α.

Finalmente, el test queda como

$$\varphi(X_1, \ldots, X_n) = \begin{cases} 1 & \text{si } \sqrt{n}\frac{(\overline{X} - \mu_1)}{\sigma_0} \geq z_\alpha \\ \\ 0 & \text{si } \sqrt{n}\frac{(\overline{X} - \mu_1)}{\sigma_0} < z_\alpha \end{cases} \tag{6.10}$$

En este caso, no debemos preocuparnos por el caso en que $L_{21} = k_\alpha$ ya que la variable L_{21} es continua.

Si se hubiera tenido que $\mu_2 < \mu_1$, el test más potente de nivel de significación α hubiese resultado

$$\varphi(X_1, \dots, X_n) = \begin{cases} 1 & \text{si } \sqrt{n}\frac{(\overline{X}_n - \mu_1)}{\sigma_0} \leq -z_\alpha \\ \\ 0 & \text{si } \sqrt{n}\frac{(\overline{X}_n - \mu_1)}{\sigma_0} > -z_\alpha \end{cases} \tag{6.11}$$

De (6.10) resulta que el test más potente para H : $\mu = \mu_1$ contra K : $\mu = \mu_2$ no depende de μ_2, es decir es el mismo cualquiera sea $\mu_2 > \mu_1$. Por lo tanto, el test dado por (6.10) es el test *uniformemente más potente* de nivel menor o igual que α para H : $\mu = \mu_1$ contra K : $\mu > \mu_1$.

Análogamente el test dado por (6.11) es el test *uniformemente más potente* de nivel menor o igual que α para H : $\mu = \mu_1$ contra K :$\mu < \mu_1$.

Calculemos ahora la función de potencia del test φ dado por (6.10), el que se puede escribir, haciendo manipuleo algebraico, como

$$\varphi(X_1, \dots, X_n) = \begin{cases} 1 & \text{si } \sqrt{n}\frac{(\overline{X}_n - \mu)}{\sigma_0} \geq z_\alpha + \sqrt{n}\frac{(\mu_1 - \mu)}{\sigma_0} \\ \\ 0 & \text{si } \sqrt{n}\frac{(\overline{X}_n - \mu)}{\sigma_0} < z_\alpha + \sqrt{n}\frac{(\mu_1 - \mu)}{\sigma_0} \end{cases} \tag{6.12}$$

Luego, la función de potencia del test φ definido por (6.10) está dada por

$$\beta_\varphi(\mu) = E_\mu(\varphi(\mathbf{X})) = P_\mu(\sqrt{n}\frac{(\overline{X}_n - \mu)}{\sigma_0} \geq z_\alpha + \sqrt{n}\frac{(\mu_1 - \mu)}{\sigma_0})$$

Pero cuando el valor de la media es μ, $\sqrt{n}(\overline{X}_n - \mu)/\sigma_0$ tiene distribución $N(0,1)$. Luego si Φ es la función de distribución de una variable aleatoria $N(0,1)$, se tendrá

$$\beta_\varphi(\mu) = 1 - \Phi(z_\alpha + \sqrt{n}\frac{(\mu_1 - \mu)}{\sigma_0})\ .$$

Estudiaremos algunas propiedades de $\beta_\varphi(\mu)$.

A. $\beta_\varphi(\mu)$ para n fijo es una función creciente de μ, ya que Φ es una función creciente.

B. $\beta_\varphi(\mu_1) = \alpha$.

C. $\lim_{\mu\to+\infty}\beta_\varphi(\mu) = 1 - \lim_{x\to-\infty}\Phi(x) = 1 - 0 = 1.$

D. $\lim_{\mu\to-\infty}\beta_\varphi(\mu) = 1 - \lim_{x\to+\infty}\Phi(x) = 1 - 1 = 0.$

E. Para μ_2 fijo, $\mu_2 > \mu_1$, se tiene

$$\lim_{n\to\infty}\beta_\varphi(\mu_2) = 1 - \lim_{n\to\infty}\Phi(x) = 1 - 0 = 1.$$

De aquí se deduce que tomando n grande, para un μ_2 fijo, la probabilidad de error de tipo 2 se puede hacer tan pequeño como se quiera.

De A y B resulta que

$$\sup_{\mu\le\mu_1}\beta_\varphi(\mu) = \alpha,$$

y luego φ resulta un test de nivel igual que α para H : $\mu \le \mu_1$ contra K : $\mu > \mu_1$.

Veamos ahora que φ es el test de nivel $\le \alpha$, uniformemente más potente para estas mismas hipótesis. Sea φ^* otro test de nivel menor o igual que α para H : $\mu \le \mu_1$; también tendrá este nivel para H : $\mu = \mu_1$, pero φ es el test uniformemente más potente para H : $\mu = \mu_1$ contra K : $\mu > \mu_1$. Entonces se tiene

$$\beta_\varphi(\mu) \ge \beta_{\varphi^*}(\mu) \quad \forall\, \mu > \mu_1$$

y φ resulta el test más potente de nivel menor o igual que α para H : $\mu \le \mu_1$ contra K : $\mu > \mu_1$.

Luego hemos demostrado el siguiente teorema

Teorema 2.

(i) El test φ dado por (6.10) es el uniformemente más potente de nivel menor o igual que α para
(a) H : $\mu = \mu_1$ contra K : $\mu > \mu_1$
y para
(b) H : $\mu \le \mu_1$ contra K : $\mu > \mu_1$.
Su función de potencia viene dada por

$$\beta_\varphi(\mu) = 1 - \Phi(z_\alpha + \sqrt{n}(\mu_1 - \mu)/\sigma_0).$$

(b) En forma similar el test φ dado por (6.11) es el uniformemente más potente de nivel menor o igual que α para
(a) H : $\mu = \mu_1$ contra K : $\mu < \mu_1$

y para
(b) H : $\mu \geq \mu_1$ contra K : $\mu < \mu_1$.
Su función de potencia viene dada por

$$\beta_\varphi(\mu) = \Phi(-z_\alpha + \sqrt{n}(\mu_1 - \mu)/\sigma_0)$$

Ejemplo 2. Supongamos que se mide el grado de impurezas de un producto químico. El método de medición está afectado por un error que se supone $N(0, \sigma_0^2)$, con σ_0^2 conocida igual a 0.01. Además los errores correspondientes a diferentes mediciones son independientes entre sí. Se sabe que el producto es aceptable si el grado de impurezas es menor o igual que 0.7. Se hacen 64 observaciones, $X_1, \ldots, X_{64}$, y se quiere decidir entre las hipótesis: $\mu < 0.7$ ó $\mu \geq 0.7$. Se quiere encontrar un test de modo que la probabilidad de aceptar el producto, cuando éste no satisfaga las condiciones, sea menor que 0.05. Sabemos que cada X_i puede escribirse

$$X_i = \mu + \varepsilon_i$$

donde μ es el grado de impureza y ε_i el error de medición para la observación $i-$esima. Como los ε_i se supusieron normales e independientes, las X_i serán una muestra aleatoria de la distribución $N(\mu, \sigma_0^2)$.

Lo primero que tenemos que determinar es cuál hipótesis es H y cuál K. Tomamos H : $\mu \geq 0.7$, ya que rechazar esta hipótesis equivale a aceptar el producto, y esto queremos hacerlo solamente si estamos muy seguros de su bondad. Luego, se tiene el problema:

$$\text{H} : \mu \geq 0.7 \qquad \text{contra} \qquad \text{K} : \mu < 0.7$$

y por lo tanto, el test más potente de nivel 0.05 está dado por $\varphi(\mathbf{X}) = 1$ si

$$\sqrt{64}\frac{(\overline{X} - 0.7)}{0.1} \leq -z_{0.05} \; .$$

En las tablas se encuentra que $-z_{0.05} = -1.65$. Así, el test rechaza H, es decir, acepta el producto si

$$\overline{X} \leq \frac{-1.65 \times 0.1}{8} + 0.7 = 0.68 \; .$$

Supongamos ahora que se quiere conocer la probabilidad de cometer error de tipo 2, o sea, de aceptar H cuando es falsa (rechazar el producto cuando cumple la especificación). Tenemos que calcular la función de potencia del test. De acuerdo con lo que hemos visto, será

$$\beta_\varphi(\mu) = \Phi\left(-1.65 - 8\frac{(\mu - 0.7)}{0.1}\right) = \Phi(54.35 - 80\mu) \ .$$

Si queremos, por ejemplo, calcular $\beta_\varphi(0.65)$, esto será uno menos la probabilidad de rechazar el producto cuando $\mu = 0.65$, luego

$$\beta_\varphi(0.65) = \Phi(54.35 - 80 \times 0.65) = \Phi(2.35) = 0.99 \ .$$

Esto quiere decir que la probabilidad de rechazar la droga, cuando $\mu = 0.65$ es 0.01.

6.4 Tests uniformemente más potentes para hipótesis unilaterales

Hemos visto en el parágrafo anterior la forma de encontrar tests más potentes en el caso de hipótesis simples

$$\mathrm{H} : \boldsymbol{\theta} = \boldsymbol{\theta}_0 \qquad \text{contra} \qquad \mathrm{K} : \boldsymbol{\theta} = \boldsymbol{\theta}_1 \ .$$

Esta situación es principalmente de interés teórico puesto que aún las situaciones más simples que se presentan en la práctica, cuando $\theta \in I\!R$, implican problemas de la forma

(a) H : $\theta = \theta_0$ contra K : $\theta > \theta_0$

(b) H : $\theta = \theta_0$ contra K : $\theta < \theta_0$

(c) H : $\theta \leq \theta_0$ contra K : $\theta > \theta_0$

(d) H : $\theta \geq \theta_0$ contra K : $\theta < \theta_0$

(e) H : $\theta = \theta_0$ contra K : $\theta \neq \theta_0$

Los problemas (a) a (d) se denominan *unilaterales* y al (e) *bilateral*. Hemos visto que para el caso $N(\mu, \sigma_0^2)$ con σ_0^2 conocido se puede extender el test de Neyman–Pearson a hipótesis compuestas de la forma

H : $\mu = \mu_0$ contra K : $\mu > \mu_0$
H : $\mu \leq \mu_0$ contra K : $\mu > \mu_0$
H : $\mu = \mu_0$ contra K : $\mu < \mu_0$

H : $\mu \geq \mu_0$ contra K : $\mu < \mu_0$
obteniéndose tests uniformemente más potentes para estos problemas.

La obtención de tests uniformemente más potentes para hipótesis unilaterales a partir de Neyman–Pearson es siempre posible para ciertas familias de distribuciones que tienen una propiedad llamada de *cociente de verosimilitud monótono.*

Definición 1. Una familia de distribuciones discretas o continuas con densidad (o función de probabilidad puntual) $p(\mathbf{x}, \theta)$, $\theta \in \Theta \subset I\!R$ se dice de *cociente de verosimilitud monótono* (CVM) *en el estadístico* $T = r(\mathbf{X})$ donde r toma valores reales, si para todo par $\theta_1 < \theta_2$ se tiene

(i) Las distribuciones correspondientes a $p(\mathbf{x}, \theta_1)$ y $p(\mathbf{x}, \theta_2)$ son distintas

(ii) $p(\mathbf{x}, \theta_2)/p(\mathbf{x}, \theta_1) = g_{\theta_1\theta_2}(r(\mathbf{x}))$, donde $g_{\theta_1\theta_2}(t)$ es una función no decreciente en el conjunto

$$\mathcal{S} = \{t : t = r(\mathbf{x}) \quad \text{con } p(\mathbf{x}, \theta_1) > 0 \text{ó } p(\mathbf{x}, \theta_2) > 0\}$$

Observación. A los efectos de la Definición 1 si $p(\mathbf{x}, \theta_1) = 0$ y $p(\mathbf{x}, \theta_2) > 0$, el cociente $p(\mathbf{x}, \theta_2)/p(\mathbf{x}, \theta_1)$ se considerará igual a ∞.

Es sencillo mostrar que las familias exponenciales a un parámetro con $c(\theta)$ estrictamente monótona son de CVM.

Teorema 1. *Sea la familia exponencial a un parámetro con función de densidad o probabilidad* $p(\mathbf{x}, \theta) = A(\theta)e^{c(\theta)r(\mathbf{x})}h(\mathbf{x})$ *con* $\theta \in \Theta \subset I\!R$. *Luego,*

(i) Si $c(\theta)$ *es estrictamente creciente la familia dada es de CVM en* $r(\mathbf{X})$

(ii) *Si* $c(\theta)$ *es estrictamente decreciente la familia dada es de CVM en* $-r(\mathbf{X})$

Demostración. Sólo demostraremos (i). La parte (ii) se demuestra idénticamente. En este caso se tiene si $\theta_1 < \theta_2$

$$\frac{p(\mathbf{x}, \theta_2)}{p(\mathbf{x}, \theta_1)} = \frac{A(\theta_2)}{A(\theta_1)}\, e^{(c(\theta_2) - c(\theta_1))r(\mathbf{x})} = g_{\theta_1\theta_2}(r(\mathbf{x}))$$

donde

$$g_{\theta_1\theta_2}(t) = \frac{A(\theta_2)}{A(\theta_1)}\, e^{(c(\theta_2) - c(\theta_1))t}$$

es una función creciente.

Por otro lado, por ser c estrictamente monótona, $\theta_1 \neq \theta_2$ implica $c(\theta_1) \neq c(\theta_2)$ y luego $p(\mathbf{x}, \theta_1)$ y $p(\mathbf{x}, \theta_2)$ corresponden a distribuciones diferentes. Luego, la familia dada es de cociente de verosimilitud monótono en $T = r(\mathbf{X})$.

Vamos a mostrar ahora que existen familias de CVM que no son exponenciales. Para ello consideramos el siguiente ejemplo

Ejemplo 1. Consideremos una muestra aleatoria $(X_1, \ldots, X_n)$ de una distribución $U[0, \theta]$ con $\theta \in I\!R^+$.

Luego, la familia de distribuciones conjuntas de $\mathbf{X} = (X_1, \ldots, X_n)$ se puede escribir

$$p(\mathbf{x}, \theta) = \frac{1}{\theta^n} I_{[0,\theta]}(\max_{1 \leq i \leq n} x_i)\, I_{[0,\infty]}(\min_{1 \leq i \leq n} x_i)\ . \tag{6.13}$$

Mostraremos que esta familia es de CVM en $r(\mathbf{X}) = \max_{1 \leq i \leq n} X_i$. Sea $\theta_2 > \theta_1$, luego, el conjunto $\mathcal{S} = \{t : r(\mathbf{x}) \text{ con } p(\mathbf{x}, \theta_1) > 0 \text{ o } p(\mathbf{x}, \theta_2) > 0\}$ resulta igual al intervalo $[0, \theta_2]$. Definiendo

$$g_{\theta_1 \theta_2}(t) = \frac{\theta_1^n}{\theta_2^n} \frac{I_{[0,\theta_2]}(t)}{I_{[0,\theta_1]}(t)}\ ,$$

se tiene que

$$\frac{p(\mathbf{x}, \theta_2)}{p(\mathbf{x}, \theta_1)} = g_{\theta_1 \theta_2}(r(\mathbf{x})).$$

Po lo tanto, bastará mostrar que $g_{\theta_1 \theta_2}(t)$ es monótona en $\mathcal{S}$. Pero

$$g_{\theta_1 \theta_2}(t) = \begin{cases} (\theta_1/\theta_2)^n & \text{si } 0 \leq t \leq \theta_1 \\ \infty & \text{si } \theta_1 \leq t \leq \theta_2\ . \end{cases}$$

Con lo cual, $g_{\theta_1 \theta_2}(t)$ es monótona y la familia dada por (6.13) es de CVM en $r(\mathbf{X})$. Por otro lado, la familia dada por (6.13) no es exponencial de acuerdo a lo visto en el ejercicio 2 del Capítulo 3.

Ejemplo 2. Consideremos una variable aleatoria X con distribución $\mathcal{C}(\theta, 1)$, $\theta \in I\!R$, o sea, su densidad viene dada por

$$p(x, \theta) = \frac{1}{\pi\, [1 + (x - \theta)^2]}\ .$$

Veremos que esta familia no es de cociente de verosimiltud monótono en $r(X) = X$.

Sea $\theta_2 > \theta_1$, luego, se tiene que

$$\frac{p(x,\theta_2)}{p(x,\theta_1)} = \frac{[1 + (x-\theta_1)^2]}{[1 + (x-\theta_2)^2]} = g_{\theta_1\theta_2}(x).$$

Sin embargo, la función $g_{\theta_1\theta_2}(x)$ no es monótona en x ya que $\lim_{x\to-\infty} g_{\theta_1\theta_2}(x) = \lim_{x\to+\infty} g_{\theta_1\theta_2}(x) = 1$.

El siguiente teorema nos permite encontrar tests UMP para familia con la propiedad de CVM.

Teorema 1. *Sea* $\mathbf{X}$ *un vector aleatorio con función de probabilidad o densidad perteneciente a la familia* $p(\mathbf{x},\theta)$ *con* $\theta \in \Theta \subset I\!R$*, que tiene la propiedad de ser de CVM en* $T = r(\mathbf{X})$*. Luego*

(i) *Existen* k_α *y* γ_α *tales que si definimos*

$$\varphi(\mathbf{X}) = \begin{cases} 1 & \text{si } T > k_\alpha \\ \gamma_\alpha & \text{si } T = k_\alpha \\ 0 & \text{si } T < k_\alpha \end{cases} \tag{6.14}$$

se satisface

$$E_{\theta_1}(\varphi(\mathbf{X})) = \alpha\ . \tag{6.15}$$

(ii) *Sea* φ *es un test de la forma (6.14) que satisface (6.15). Luego* φ *es el test uniformemente más potente UMP de nivel menor o igual que* α *para*

$$H : \theta = \theta_1 \text{ contra } K : \theta > \theta_1\ .$$

(iii) $\beta_\varphi(\theta)$ *es monótona no decreciente para todo* θ *y estrictamente creciente para todo* θ *tal que* $0 < \beta_\varphi(\theta) < 1$.

(iv) *Sea* φ *un test de la forma (6.14) que satisface (6.15). Luego,* φ *es el test uniformemente más potente de nivel menor o igual que* α *para*

$$H : \theta \leq \theta_1 \text{ contra } K : \theta > \theta_1\ .$$

DEMOSTRACIÓN: La demostración de (i) es idéntica a la dada en el Teorema de Neyman-Pearson.

Demostraremos (ii) suponiendo que si $\theta_2 > \theta_1$

$$\frac{p(\mathbf{x},\theta_2)}{p(\mathbf{x},\theta_1)} = g_{\theta_1\theta_2}(r(\mathbf{x}))$$

con $g_{\theta_1\theta_2}(t)$ estrictamente creciente. (Esta hipótesis no es necesaria, basta con que sea no decreciente.) En este caso, dado $\theta_2 > \theta_1$ el test dado por (6.14) se puede escribir como

$$\varphi(\mathbf{X}) = \begin{cases} 1 & \text{si} \quad g_{\theta_1\theta_2}(r(\mathbf{X})) > g_{\theta_1\theta_2}(k_\alpha) \\ \gamma_\alpha & \text{si} \quad g_{\theta_1\theta_2}(r(\mathbf{X})) = g_{\theta_1\theta_2}(k_\alpha) \\ 0 & \text{si} \quad g_{\theta_1\theta_2}(r(\mathbf{X})) < g_{\theta_1\theta_2}(k_\alpha) \end{cases}$$

y si llamamos ${k'}_\alpha = g_{\theta_1\theta_2}(k_\alpha)$ resulta

$$\varphi(\mathbf{X}) = \begin{cases} 1 & \text{si} \quad \dfrac{p(\mathbf{X},\theta_2)}{p(\mathbf{X},\theta_1)} > {k'}_\alpha \\ \\ \gamma_\alpha & \text{si} \quad \dfrac{p(\mathbf{X},\theta_2)}{p(\mathbf{X},\theta_1)} = {k'}_\alpha \\ \\ 0 & \text{si} \quad \dfrac{p(\mathbf{X},\theta_2)}{p(\mathbf{X},\theta_1)} < {k'}_\alpha \; . \end{cases}$$

Como $\varphi(\mathbf{X})$ satisface (6.15), usando el Teorema 1 de 6.3 resulta que $\varphi(\mathbf{X})$ es el test más potente de nivel menor o igual que α para $\mathrm{H} : \theta = \theta_1$ contra $\mathrm{K} : \theta = \theta_2$. Como φ no depende de θ_2, este resultado vale para todo $\theta_2 > \theta_1$, luego φ es el test UMP de nivel menor o igual que α para $\mathrm{H} : \theta = \theta_1$ contra $\mathrm{K} : \theta > \theta_2$.

(iii) Sólo demostraremos que $\beta_\varphi(\theta)$ es monótona no decreciente.

Sean θ^* y θ^{**} cualesquiera, tales que $\theta^* < \theta^{**}$. Si llamamos $\alpha^* = E_{\theta^*}(\varphi(\mathbf{X}))$, resulta por (ii) que $\varphi(\mathbf{X})$ es el test más potente a nivel menor o igual que α^* para las hipótesis simples

$$\mathrm{H} : \theta = \theta^* \text{ contra } \mathrm{K} : \theta = \theta^{**} \; .$$

Consideremos ahora el test

$$\varphi^*(\mathbf{X}) = \alpha^* \; .$$

φ^* es un test de nivel α^*, luego φ^* es menos potente que φ en θ^{**}, es decir,

$$E_{\theta^{**}}(\varphi^*(\mathbf{X})) \leq E_{\theta^{**}}(\varphi(\mathbf{X}))$$

pero,

$$E_{\theta^{**}}(\varphi^*(\mathbf{X})) = \alpha^* = E_{\theta^*}(\varphi(\mathbf{X})) = \beta_\varphi(\theta^*)$$

y además

$$E_{\theta^{**}}(\varphi(\mathbf{X})) = \beta_\varphi(\theta^{**})$$

por lo tanto,

$$\beta_\varphi(\theta^*) \leq \beta_\varphi(\theta^{**}) \ ,$$

con lo que queda demostrado que $\beta_\varphi(\theta)$ es monótona no decreciente.

Para demostrar (iv), primero mostraremos que $\varphi(\mathbf{X})$ es un test de nivel menor o igual que α para

$$\mathrm{H} : \theta \leq \theta_1 \text{ contra } \mathrm{K} : \theta > \theta_1$$

o sea que

$$\sup_{\theta \leq \theta_1} \beta_\varphi(\theta) \leq \alpha \ .$$

Como $\beta_\varphi(\theta)$ es monótona creciente se tiene:

$$\sup_{\theta \leq \theta_1} \beta_\varphi(\theta) = \beta_\varphi(\theta_1) = \alpha$$

por (6.15).

Consideremos ahora otro test $\varphi^*(\mathbf{X})$ de nivel menor o igual que α para $\mathrm{H} : \theta \leq \theta_1$ contra $\mathrm{K} : \theta > \theta_1$, luego $\varphi^*(\mathbf{X})$ es de nivel menor o igual que α para $\mathrm{H} : \theta = \theta_1$ contra $\mathrm{K} : \theta > \theta_1$, pero por (ii) $\varphi(\mathbf{X})$ es el test uniformemente más potente para este problema, por lo tanto

$$\beta_\varphi(\theta) \geq \beta_{\varphi^*}(\theta) \quad \forall\, \theta > \theta_1 \ .$$

Análogamente se demuestra el siguiente teorema

Teorema 2. *Sea* $\mathbf{X}$ *un vector aleatorio con función de densidad perteneciente a la familia* $p(\mathbf{x}, \theta)$ *con* $\theta \in \Theta \subset I\!R$. *Supongamos que esta familia es CMV en* $r(\mathbf{X})$. *Luego*

(i) *Existen* k_α *y* γ_α *tales que si definimos*

$$\varphi(\mathbf{X}) = \begin{cases} 1 & \text{si } r(\mathbf{X}) < k_\alpha \\ \gamma_\alpha & \text{si } r(\mathbf{X}) = k_\alpha \\ 0 & \text{si } r(\mathbf{X}) > k_\alpha \end{cases} \tag{6.16}$$

se satisface

$$E_{\theta_1}(\varphi(\mathbf{X})) = \alpha \tag{6.17}$$

(ii) *Sea $\varphi(\mathbf{X})$ es un test de la forma (6.16) que satisface (6.17). Luego φ es el test uniformemente más potente a nivel menor o igual que α para $H : \theta = \theta_1$ contra $K : \theta < \theta_1$.*

(iii) *$\beta_\varphi(\theta)$ es monótona no creciente para todo θ y estrictamente decreciente para todo θ tal que $0 < \beta_\varphi(\theta) < 1$.*

(iv) *Sea φ un test de la forma (6.16) que satisface (6.17). Luego φ es el test uniformemente más potente de nivel menor o igual que α para $H : \theta \geq \theta_1$ contra $K : \theta < \theta_1$.*

Para una versión más completa de este Teorema, ver Teorema 2 de 3.3 en Lehmann [2].

Ejemplo 3. Consideremos una muestra aleatoria $X_1, \ldots, X_n$ de una distribución perteneciente a la familia $N(\mu, \sigma_0^2)$ con σ_0^2 conocido. Luego, es fácil demostrar que la familia de distribuciones de la muestra es exponencial con $r(\mathbf{X}) = \sum_{i=1}^n X_i$ y $c(\mu) = \mu/\sigma_0^2$. Como $c(\mu)$ es creciente de acuerdo al Teorema 1, esta familia es de CMV en $r(\mathbf{X})$. Entonces para testear H : $\mu \leq \mu_1$ contra K : $\mu > \mu_1$, el test UMP de nivel menor o igual que α, es de la forma

$$\varphi(\mathbf{X}) = \begin{cases} 1 & \text{si } \displaystyle\sum_{i=1}^n X_i \geq k_\alpha \\ \\ 0 & \text{si } \displaystyle\sum_{i=1}^n X_i < k_\alpha \end{cases}$$

con $E_{\mu_1}(\varphi(\mathbf{X}) = \alpha$.

En la Sección 6.3 ya habíamos demostrado este resultado y hallado el valor de k_α.

Ejemplo 4. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución perteneciente a la familia $Bi(\theta, 1)$.

En este caso la familia de distribuciones de $X_1, \ldots, X_n$ es exponencial con $T = r(\mathbf{X}) = \sum_{i=1}^n X_i$ y $c(\theta) = \ln(\theta/(1-\theta))$; como $c(\theta)$ es creciente, esta familia es de CMV en $r(\mathbf{X})$.

Luego, el test UMP de nivel menor o igual que α para H : $\theta \leq \theta_1$ contra K : $\theta > \theta_1$ será de la forma

$$\varphi(\mathbf{X}) = \begin{cases} 1 & \text{si } T > k_\alpha \\ \gamma_\alpha & \text{si } T = k_\alpha \\ 0 & \text{si } T < k_\alpha \end{cases}$$

k_α y γ_α deberán ser elegidos de modo que

$$E_{\theta_1}(\varphi(\mathbf{X})) = P_{\theta_1}(T > k_\alpha) + \gamma_\alpha P_{\theta_1}(T = k_\alpha) = \alpha \ . \tag{6.18}$$

Como T tiene distribución $Bi(\theta, n)$ que es discreta, puede suceder que exista o no k tal que

$$P_{\theta_1}(T > k) = \alpha \tag{6.19}$$

Si existe k satisfaciendo (6.19), tomaremos ese valor como k_α y $\gamma_\alpha = 0$.
Si no existe k que verifique (6.19), siempre existirá k tal que

$$P_{\theta_1}(T > k) < \alpha < P_{\theta_1}(T \geq k) \ . \tag{6.20}$$

Este valor k será el k_α que eligiremos y reemplazándolo en (6.18) obtendremos

$$\gamma_\alpha = \frac{\alpha - P_{\theta_1}(T > k_\alpha)}{P_{\theta_1}(T = k_\alpha)} = \frac{\alpha - P_{\theta_1}(T > k_\alpha)}{P_{\theta_1}(T \geq k_\alpha) - P_{\theta_1}(T > k_\alpha)} \ .$$

Por (6.20) resulta que $0 < \gamma_\alpha < 1$.
Para encontrar el k_α que verifica (6.19) o (6.20) deberán usarse tablas binomiales.

Recordemos finalmente que

$$P_{\theta_1}(T \geq k_\alpha) = \sum_{k_\alpha \leq i \leq n} \binom{n}{i} \theta_1^i (1-\theta_1)^{n-i} \ .$$

Supongamos que se tiene una muestra aleatoria X_1, X_2, X_3 de una distribución $Bi(\theta, 1)$ y se quiere testear H : $\theta \leq 1/3$ contra K : $\theta > 1/3$ con nivel de significación menor o igual que 0.1.

Cuando $\theta = 1/3$, la distribución de $T = \sum_{i=1}^3 X_i$ está dada por

t	0	1	2	3
$p_T(t)$	$\frac{8}{27}$	$\frac{12}{27}$	$\frac{6}{27}$	$\frac{1}{27}$

y por lo tanto, tenemos

k	-1	0	1	2	3
$P_{\frac{1}{3}}(T > k)$	1	$\frac{19}{27}$	$\frac{7}{27}$	$\frac{1}{27}$	0

Por lo tanto, no existe k_α que verifique (6.19) y el valor $k_\alpha = 2$ verifica (6.20), pues

$$P_{\frac{1}{3}}(T > 2) = \frac{1}{27} < 0.1 < P_{\frac{1}{3}}(T \geq 2) = P_{\frac{1}{3}}(T > 1) = \frac{7}{27}$$

y γ_α será entonces

$$\gamma_\alpha = \frac{0.1 - \frac{1}{27}}{\frac{6}{27}} = 0.27 \quad .$$

Como ejercicio se sugiere graficar la función de potencia de este test, y siendo el test aleatorizado, sugerir un mecanismo para decidir en caso en que $T = 2$.

Ejemplo 5. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución perteneciente a la familia $U[0,\theta]$.

El test uniformemente más potente para H : $\theta \leq \theta_1$ contra K : $\theta > \theta_1$, será de la forma

$$\varphi(\mathbf{X}) = \begin{cases} 1 & \text{si } \max\limits_{1\leq i\leq n} X_i \geq k_\alpha \\ \\ 0 & \text{si } \max\limits_{1\leq i\leq n} X_i < k_\alpha \end{cases}$$

donde k_α verifica

$$E_{\theta_1}(\varphi(\mathbf{X})) = \alpha \,. \tag{6.21}$$

Teniendo en cuenta que la función de distribución de $T = \max_{1\leq i\leq n} X_i$ es

$$F_T(t) = \begin{cases} 0 & \text{si} \quad t < 0 \\ (t/\theta)^n & \text{con} \quad 0 \leq t \leq \theta \\ 1 & \text{si} \quad t > 1 \end{cases}$$

y que debe cumplirse (6.21), se tiene $0 \leq k_\alpha \leq \theta_1$ y

$$P_{\theta_1}\left(\max_{1\leq i\leq n} X_i \geq k_\alpha\right) = 1 - (k_\alpha/\theta_1)^n = \alpha \,,$$

de donde resulta

$$k_\alpha = \theta_1 \sqrt[n]{1-\alpha} \,.$$

6.5 Tests insesgados

En la mayoría de los casos en que la hipótesis alternativa es una hipótesis compuesta, no existe un test uniformemente más potente.

Ejemplo 1. Supongamos que se tiene una muestra aleatoria $X_1, \ldots, X_n$ de una distribución $N(\mu, \sigma_0^2)$ con σ_0 conocido y se desea testear $\mathrm{H} : \mu = \mu_0$ contra $\mathrm{K} : \mu \neq \mu_0$. Es fácil demostrar que no existe un test uniformemente más potente a nivel menor o igual que α.

Supongamos que tal test existiera y llamémoslo φ; entonces será el test más potente a nivel menor o igual que α para

$$\mathrm{H}_1 : \mu = \mu_0 \text{ contra } \mathrm{K}_1 : \mu = \mu_1 \quad (\mu_1 > \mu_0)$$

y para

$$\mathrm{H}_2 : \mu = \mu_0 \text{ contra } \mathrm{K}_2 : \mu = \mu_2 \quad (\mu_2 < \mu_0)\ .$$

Pero, por el Teorema 3 de la Sección 6.3 el test más potente para H_1 contra K_1 está dado por

$$\varphi_1(\mathbf{X}) = \begin{cases} 1 & \text{si } \sqrt{n}\dfrac{(\overline{X} - \mu_0)}{\sigma_0} \geq z_\alpha \\ \\ 0 & \text{si } \sqrt{n}\dfrac{(\overline{X} - \mu_0)}{\sigma_0} < z_\alpha \end{cases}$$

y el test más potente para H_2 contra K_2 está dado por

$$\varphi_2(\mathbf{X}) = \begin{cases} 1 & \text{si } \sqrt{n}\dfrac{(\overline{X} - \mu_0)}{\sigma_0} \leq -z_\alpha \\ \\ 0 & \text{si } \sqrt{n}\dfrac{(\overline{X} - \mu_0)}{\sigma_0} > -z_\alpha \end{cases}$$

Entonces, por la unicidad dada en el Teorema de Neyman-Pearson, φ debería coincidir con φ_1 y con φ_2 lo cual es imposible.

Recordemos que en el caso de estimadores puntuales tampoco existe en general uno de menor error cuadrático medio. Una manera de poder definir un estimador óptimo que se propuso en el Capítulo 3 fue restringiendo los estimadores a la clase de los insesgados. En el caso de test se procede en forma similar, restringiremos la clase de tests considerados a los que llamaremos insesgados y luego se buscará el test uniformemente más potente en esta clase.

Definición 1. Sea una familia de distribuciones $F(\mathbf{x},\boldsymbol{\theta})$ con $\boldsymbol{\theta} \in \Theta$. Se dirá que un test φ para testear H : $\boldsymbol{\theta} \in \Theta_1$ contra K : $\boldsymbol{\theta} \in \Theta_2$ *es insesgado* si

$$\sup_{\boldsymbol{\theta} \in \Theta_1} \beta_\varphi(\boldsymbol{\theta}) \leq \inf_{\boldsymbol{\theta} \in \Theta_2} \beta_\varphi(\boldsymbol{\theta})$$

El sentido de esta desigualdad es que la probabilidad de rechazar H cuando $\boldsymbol{\theta} \in \Theta_2$, es decir cuando H es falsa, no pude ser menor que cuando $\boldsymbol{\theta} \in \Theta_1$, es decir cuando H es verdadera.

Por lo tanto, un test insesgado de nivel α tiene función de potencia menor o igual que α para $\boldsymbol{\theta} \in \Theta_1$ y mayor o igual que α para $\boldsymbol{\theta} \in \Theta_2$.

Observemos que un test UMP de nivel α es insesgado.

Observación. Si la función de potencia $\beta_\varphi(\boldsymbol{\theta})$ del test φ es una función continua de $\boldsymbol{\theta}$ y φ es un test insesgado de nivel α, entonces $\beta_\varphi(\boldsymbol{\theta})$ debe valer α en la frontera Θ_F entre Θ_1 y Θ_2.
En particular, si $\Theta \subset I\!R$, $\Theta_1 = \{\theta_1\}$ y $\Theta_2 = \Theta - \{\theta_1\}$, o sea, si estamos testeando H : $\theta = \theta_1$ contra K : $\theta \neq \theta_1$, y φ es un test insesgado de nivel α se tiene

$$\begin{aligned} &\beta_\varphi(\theta_1) = \alpha \\ &\beta_\varphi(\theta) \geq \alpha \qquad \forall \theta \neq \theta_1 \ . \end{aligned}$$

Por lo tanto, si la función de potencia $\beta_\varphi(\theta)$ es derivable respecto de θ, φ debe cumplir

$$\beta'_\varphi(\theta_1) = \frac{\partial}{\partial \theta} \beta_\varphi(\theta)|_{\theta=\theta_1} = 0 \ . \tag{6.22}$$

En el caso particular de las familias exponenciales, la función de potencia de cualquier test es derivable y por lo tanto, los tests insesgados cumplen (6.22).

Definición 2. Se dirá que un test φ para testear H : $\boldsymbol{\theta} \in \Theta_1$ contra K : $\boldsymbol{\theta} \in \Theta_2$ *es uniformemente más potente de nivel α entre los insesgados*, IUMP, si

(a) φ tiene nivel α, o sea,

$$\sup_{\boldsymbol{\theta} \in \Theta_1} \beta_\varphi(\boldsymbol{\theta}) = \alpha$$

(b) φ es insesgado, es decir,

$$\beta_\varphi(\boldsymbol{\theta}) \geq \alpha \qquad \forall \boldsymbol{\theta} \in \Theta_2$$

(c) Dado otro test φ^* insesgado y de nivel α se verifica

$$\beta_\varphi(\boldsymbol{\theta}) \geq \beta_{\varphi^*}(\boldsymbol{\theta}) \qquad \forall \boldsymbol{\theta} \in \Theta_2 \ .$$

En la próxima Sección daremos un procedimiento general para encontrar tests para un problema determinado. En muchos casos este procedimiento da como resultado el test insesgado uniformente más potente.

La teoría de los tests insesgados uniformemente más potentes escapa a las posibilidades de este curso y puede verse en Lehmann [3] o en Ferguson [2].

6.6 Test del cociente de máxima verosimilitud

Supongamos que se observa un vector $\mathbf{X}$, cuya distribución tiene función de densidad $p(\mathbf{x}, \boldsymbol{\theta})$, $\boldsymbol{\theta} \in \Theta$ y se quiere testear H : $\boldsymbol{\theta} \in \Theta_1$ contra K : $\boldsymbol{\theta} \in \Theta_2$ $(\Theta_1 \cup \Theta_2 = \Theta)$.

Un procedimiento intuitivamente razonable y que da buenos resultados en una gran variedad de situaciones es el siguiente.

Tomemos estimadores de máxima verosimilitud de $\boldsymbol{\theta}$, suponiendo $\boldsymbol{\theta} \in \Theta_1$, llamémoslo $\widehat{\boldsymbol{\theta}}_1$ y análogamente suponiendo $\boldsymbol{\theta} \in \Theta_2$, $\widehat{\boldsymbol{\theta}}_2$; luego

$$p(\mathbf{X}, \widehat{\boldsymbol{\theta}}_1) = \max_{\boldsymbol{\theta} \in \Theta_1} p(\mathbf{X}, \boldsymbol{\theta})$$

y

$$p(\mathbf{X}, \widehat{\boldsymbol{\theta}}_2) = \max_{\boldsymbol{\theta} \in \Theta_2} p(\mathbf{X}, \boldsymbol{\theta}) \ .$$

Si $\widehat{\boldsymbol{\theta}}_1$ y $\widehat{\boldsymbol{\theta}}_2$ no dependieran de la muestra, podríamos considerar el test más potente para testear H* : $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}_1$ contra K* : $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}_2$, el cual es de la forma

$$\varphi(\mathbf{X}) = \begin{cases} 1 & \text{si } L < k_\alpha \\ \gamma_\alpha & \text{si } L = k_\alpha \\ 0 & \text{si } L > k_\alpha \end{cases}$$

donde

$$L = \frac{1}{L_{21}} = \frac{p(\mathbf{X}, \widehat{\boldsymbol{\theta}}_1)}{p(\mathbf{X}, \widehat{\boldsymbol{\theta}}_2)}$$

y k_α se elige de manera que el test resulte de nivel α.

En algunos casos $\widehat{\boldsymbol{\theta}}_1$ y $\widehat{\boldsymbol{\theta}}_2$ pueden no existir, pero siempre tiene sentido hablar de L definido por

$$L = \frac{\sup_{\boldsymbol{\theta} \in \Theta_1} p(\mathbf{X}, \boldsymbol{\theta})}{\sup_{\boldsymbol{\theta} \in \Theta_2} p(\mathbf{X}, \boldsymbol{\theta})}$$

Intuitivamente, este test puede interpretarse como rechazando $\mathrm{H} : \boldsymbol{\theta} \in \Theta_1$ cuando "el valor más probable de Θ_2" tiene probabilidad considerablemente más grande que "el valor más probable de Θ_1".

En muchos casos, como por ejemplo cuando la dimensión de Θ_1 es menor que la dimensión de $\Theta = \Theta_1 \cup \Theta_2$, y $p(\mathbf{x}, \boldsymbol{\theta})$ es continua, resulta que

$$\sup_{\boldsymbol{\theta} \in \Theta_2} p(\mathbf{X}, \boldsymbol{\theta}) = \sup_{\boldsymbol{\theta} \in \Theta} p(\mathbf{X}, \boldsymbol{\theta}) \tag{6.23}$$

En este caso, el test del cociente de máxima verosimilitud resulta equivalente a

$$\varphi(\mathbf{X}) = \begin{cases} 1 & \text{si } L^* < k_\alpha \\ \gamma_\alpha & \text{si } L^* = k_\alpha \\ 0 & \text{si } L^* > k_\alpha \end{cases}$$

donde

$$L^* = \frac{\sup_{\boldsymbol{\theta} \in \Theta_1} p(\mathbf{X}, \boldsymbol{\theta})}{\sup_{\boldsymbol{\theta} \in \Theta} p(\mathbf{X}, \boldsymbol{\theta})} \; .$$

En general, es más fácil aplicar la forma del test basada en L^* cuando es posible, es decir, cuando (6.23) se cumple.

Ejemplo 1. Se tiene una muestra aleatoria $X_1, \ldots, X_n$ de una distribución $N(\mu, \sigma_0^2)$ con σ_0 conocido y se quiere testear $\mathrm{H} : \mu = \mu_0$ contra $\mathrm{K} : \mu \neq \mu_0$

Como en este caso $\Theta_1 = \{\mu_0\}$ tiene dimensión cero (se reduce a un punto) y $\Theta = \{\mu : -\infty < \mu < +\infty\}$ tiene dimensión uno, podemos usar el test basado en L^*.

Es claro que

$$\sup_{\mu \in \Theta_1} p(\mathbf{X}, \mu) = (2\pi\sigma_0^2)^{-\frac{n}{2}} e^{-\frac{1}{2\,\sigma_0^2} \Sigma_{i=1}^n (X_i - \mu_0)^2}$$

y que

$$\sup_{\mu \in \Theta} p(\mathbf{X}, \mu) = (2\pi\sigma_0^2)^{-\frac{n}{2}} e^{-\frac{1}{2\,\sigma_0^2} \Sigma_{i=1}^n (X_i - \overline{X})^2} \; .$$

Luego,

$$L^* = e^{-\frac{1}{2\sigma_0^2}\left(\Sigma_{i=1}^n (X_i-\mu_0)^2 - \Sigma_{i=1}^n (X_i-\overline{X})^2\right)}$$

y como

$$\sum_{i=1}^n (X_i-\mu_0)^2 - \sum_{i=1}^n (X_i-\overline{X})^2 = n(\overline{X}-\mu_0)^2$$

resulta

$$L^* = e^{-\frac{n}{2\sigma_0^2}(\overline{X}-\mu_0)^2} .$$

Sea $T = \sqrt{n}|\overline{X}-\mu_0|/\sigma_0$. Luego, $L^* = g(T)$ con g decreciente. Luego $\varphi(\mathbf{X})$ es equivalente a

$$\varphi(\mathbf{X}) = \begin{cases} 1 & \text{si} \quad \sqrt{n}\dfrac{|\overline{X}-\mu_0|}{\sigma_0} \geq k'_\alpha \\ \\ 0 & \text{si} \quad \sqrt{n}\frac{|\overline{X}-\mu_0|}{\sigma_0} < k'_\alpha \, . \end{cases}$$

Obsérvese que este test resulta muy razonable intuitivamente, ya que se rechaza la hipótesis de que $\mu = \mu_0$ si $\overline{X}$ difiere sensiblemente de μ_0.

k'_α debe elegirse de modo tal que φ resulte de nivel α, es decir que

$$P_{\mu_0}(\sqrt{n}\frac{|\overline{X}-\mu_0|}{\sigma_0} \geq k_\alpha) = \alpha \, .$$

Pero como, cuando $\mu = \mu_0$ se tiene que$\sqrt{n}(\overline{X}-\mu_0)/\sigma_0$ tiene distribución $N(0,1)$, resulta que $k'_\alpha = z_{\alpha/2}$.

Ejemplo 2. Sea $X_1,\ldots,X_n$ una muestra aleatoria de una distribución $N(\mu,\sigma^2)$ con varianza σ^2 desconocida y se desea testear H : $\mu = \mu_0$ contra K : $\mu \neq \mu_0$. En este caso,

$$\Theta_1 = \{(\mu_0,\sigma^2) : 0 < \sigma^2 < \infty\}$$

resulta de dimensión uno, y

$$\Theta = \{(\mu_1,\sigma^2) : -\infty < \mu < \infty,\ 0 < \sigma^2 < \infty\}$$

es de dimensión dos.

Por lo tanto utilizaremos el test basado en L^*. El estimador de máxima verosimilitud de (μ,σ^2) restringido a Θ_1 es $(\mu_0, \sum_{i=1}^n (X_i-\mu_0)^2/n)$ y el estimador de máxima verosimilitud de (μ,σ^2) sin restricciones es $(\overline{X}, \sum_{i=1}^n (X_i-\overline{X})^2/n)$.

Luego, se tiene

$$\sup_{(\mu,\sigma^2)\in\Theta_1} p(\mathbf{X},\mu,\sigma^2) = \frac{1}{e^{\frac{n}{2}}(2\pi)^{\frac{n}{2}}\left(\frac{\sum_{i=1}^n (X_i-\mu_0)^2}{n}\right)^{\frac{n}{2}}}$$

y

$$\sup_{(\mu,\sigma^2)\in\Theta} p(\mathbf{X},\mu,\sigma^2) = \frac{1}{e^{\frac{n}{2}}(2\pi)^{\frac{n}{2}}\left(\frac{\sum_{i=1}^n (X_i-\overline{X})^2}{n}\right)^{\frac{n}{2}}} .$$

Por lo tanto, L^* está dado por

$$L^* = \left[\frac{\sum_{i=1}^n (X_i-\overline{X})^2}{\sum_{i=1}^n (X_i-\mu_0)^2}\right]^{\frac{n}{2}} .$$

Como

$$\sum_{i=1}^n (X_i-\mu_0)^2 = \sum_{i=1}^n (X_i-\overline{X})^2 + n(\overline{X}-\mu_0)^2$$

se tiene que

$$L^* = \left[1+\frac{n(\overline{X}-\mu_0)^2}{\sum_{i=1}^n (X_i-\overline{X})^2}\right]^{-\frac{n}{2}} .$$

Sea ahora

$$T = \sqrt{n}\frac{(\overline{X}-\mu_0)}{s}$$

donde $s^2 = \sum_{i=1}^n (X_i-\overline{X})^2/(n-1)$. Luego,

$$L^* = \left[\frac{1}{1+\frac{T^2}{n-1}}\right]^{\frac{n}{2}}$$

Como la función $1/(1+t^2/(n-1))$ es monótona decreciente de $|t|$, el test del cociente de máxima verosimilitud resulta equivalente a

$$\varphi(\mathbf{X}) = \begin{cases} 1 & \text{si} \quad |T| \geq k_\alpha \\ 0 & \text{si} \quad |T| < k_\alpha \end{cases}$$

y k_α deberá ser elegido de manera que el test resulte con nivel de significación α, es decir, de manera que

$$P_{\mu_0}(|T| \geq k_\alpha) = \alpha .$$

Como T tiene distribución student con $n-1$ grados de libertad, resulta

$$k_\alpha = t_{n-1,\frac{\alpha}{2}} \ .$$

Obsérvese que este test es completamente análogo al del Ejemplo 1, con la diferencia que se reemplaza σ por s y $z_{\alpha/2}$ por $t_{n-1,\frac{\alpha}{2}}$.

Ejemplo 3. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución $N(\mu, \sigma^2)$ con media y varianza desconocidas. Supongamos que se quiere testear H : $\mu \le \mu_0$ contra K : $\mu > \mu_0$. En este caso,

$$\Theta_1 = \{(\mu, \sigma^2) : \mu \le \mu_0, \ \sigma^2 > 0\}$$

y

$$\Theta_2 = \{(\mu, \sigma^2) : \mu > \mu_0, \ \sigma^2 > 0\} \ .$$

Luego, la dimensión de Θ_1 es igual a la de Θ_2, y el test del cociente de máxima verosimilitud deberá hacerse con L y no con L^*. Como

$$p(\mathbf{X}, \mu, \sigma^2) = (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (X_i-\mu)^2} \tag{6.24}$$

resulta

$$\ln p(\mathbf{X}, \mu, \sigma^2) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^n (X_i - \mu)^2 \ . \tag{6.25}$$

Teniendo en cuenta que

$$\sum_{i=1}^n (X_i - \mu)^2 = \sum_{i=1}^n (X_i - \overline{X})^2 + n(\overline{X} - \mu)^2$$

se obtiene que el estimador de máxima verosimilitud de μ en Θ_1, es igual a

$$\widehat{\mu}_1 = \begin{cases} \overline{X} & \text{si } \overline{X} \le \mu_0 \\ \mu_0 & \text{si } \overline{X} > \mu_0 \end{cases} \tag{6.26}$$

y que el estimador de máxima verosimilitud de μ en Θ_2, es igual a

$$\widehat{\mu}_2 = \begin{cases} \overline{X} & \text{si } \overline{X} > \mu_0 \\ \mu_0 & \text{si } \overline{X} \le \mu_0 \end{cases} \ . \tag{6.27}$$

El estimador de máxima verosimilitud de σ^2, para $\boldsymbol{\theta} \in \Theta_1$ es

$$\widehat{\sigma}_1^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \widehat{\mu}_1)^2$$

y para $\boldsymbol{\theta} \in \Theta_2$ es

$$\widehat{\sigma}_2^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \widehat{\mu}_2)^2 \ .$$

Luego, reemplazando en (6.24) se obtiene

$$\max_{(\mu,\sigma^2)\in\Theta_j} p(\mathbf{X},\mu,\sigma^2) = [2\,e\,\pi\sum_{i=1}^{n}(X_i - \widehat{\mu}_j)^2/n]^{-\frac{n}{2}}$$

para $j = 1, 2$, de donde

$$L = \left[\frac{\sum_{i=1}^n (X_i - \widehat{\mu}_2)^2}{\sum_{i=1}^n (X_i - \widehat{\mu}_1)^2}\right]^{\frac{n}{2}} = \left[\frac{\sum_{i=1}^n (X_i - \overline{X})^2 + n(\overline{X} - \widehat{\mu}_2)^2}{\sum_{i=1}^n (X_i - \overline{X})^2 + (\overline{X} - \widehat{\mu}_1)^2}\right]^{\frac{n}{2}}$$

y usando (6.26) y (6.27) se deduce

$$L = \begin{cases} \left[\dfrac{\sum_{i=1}^n (X_i - \overline{X})^2 + n(\overline{X} - \mu_0)^2}{\sum_{i=1}^n (X_i - \overline{X})^2}\right]^{\frac{n}{2}} & \text{si} \quad \overline{X} \le \mu_0 \\ \left[\dfrac{\sum_{i=1}^n (X_i - \overline{X})^2}{\sum_{i=1}^n (X_i - \overline{X})^2 + n(\overline{X} - \mu_0)^2}\right]^{\frac{n}{2}} & \text{si} \quad \overline{X} > \mu_0 \end{cases} .$$

Si llamamos

$$T = \frac{\sqrt{n}(\overline{X} - \mu_0)}{\sqrt{\frac{\sum_{i=1}^n (X_i - \overline{X})^2}{n-1}}}$$

se tiene

$$L = \begin{cases} (1 + \frac{T^2}{n-1})^{\frac{n}{2}} & \text{si} \quad \overline{X} \le \mu_0 \\ (1 + \frac{T^2}{n-1})^{-\frac{n}{2}} & \text{si} \quad \overline{X} > \mu_0 \end{cases} .$$

Luego, el test del cociente de máxima verosimilitud es de la forma

$$\varphi(\mathbf{X}) = \begin{cases} 1 \quad \text{si} \begin{cases} 1 + \dfrac{T^2}{n-1} \leq k_\alpha \quad \text{y} \quad \overline{X} \leq \mu_0 \quad \text{(A)} \\[2ex] \dfrac{1}{1+\frac{T^2}{n-1}} \leq k_\alpha \quad \text{y} \quad \overline{X} > \mu_0 \quad \text{(B)} \end{cases} \\[6ex] 0 \quad \text{si} \begin{cases} 1 + \dfrac{T^2}{n-1} > k_\alpha \quad \text{y} \quad \overline{X} \leq \mu_0 \quad \text{(C)} \\[2ex] \dfrac{1}{1+\frac{T^2}{n-1}} > k_\alpha \quad \text{y} \quad \overline{X} > \mu_0 \quad \text{(D)} \end{cases} \end{cases} .$$

Tomemos ahora $k_\alpha < 1$ (con $k_\alpha \geq 1$ se llega al mismo resultado), en este caso la primera desigualdad de (A) no puede ocurrir y la primera desigualdad de (C) ocurre siempre, luego $\varphi(\mathbf{X})$ se transforma en

$$\varphi(\mathbf{X}) = \begin{cases} 1 \quad \text{si} \quad \dfrac{1}{1+\frac{T^2}{n-1}} \leq k_\alpha \quad \text{y} \quad \overline{X} > \mu_0 \\[6ex] 0 \quad \text{si} \begin{cases} \overline{X} \leq \mu_0 \\[2ex] \dfrac{1}{1+\frac{T^2}{n-1}} > k_\alpha \text{ y} \quad \overline{X} > \mu_0 \end{cases} \end{cases} .$$

Esto es equivalente a

$$\varphi(\mathbf{X}) = \begin{cases} 1 \quad \text{si } |T| \geq k'_\alpha \quad \text{y} \quad T > 0 \\[4ex] 0 \quad \text{si} \begin{cases} |T| < k'_\alpha \quad \text{y} \quad T > 0 \\[2ex] T < 0 \end{cases} \end{cases} ,$$

de donde, se deduce que

$$\varphi(\mathbf{X}) = \begin{cases} 1 & \text{si } T \geq k'_\alpha \\ 0 & \text{si } T < k'_\alpha . \end{cases}$$

Debemos ver ahora que se puede elegir k'_α de modo que el test resulte de nivel igual α. Esto significa que

$$\sup_{\{\mu \leq \mu_0, \sigma^2 > 0\}} P_{\mu,\sigma^2}(T \geq k'_\alpha) = \alpha .$$

Se puede pensar que el caso más desfavorable, en el cual hay mayor probabilidad de rechazar H, es en el caso límite $\mu = \mu_0$; por lo tanto parece razonable elegir k'_α de manera que

$$P_{\mu_0,\sigma^2}(T \geq k'_\alpha) = \alpha \ .$$

Pero cuando $\mu = \mu_0$, T tiene distribución de Student con $n-1$ grados de libertad, y por lo tanto debemos tomar

$$k'_\alpha = t_{n-1,\alpha} \ .$$

El test φ resulta entonces

$$\varphi(\mathbf{X}) = \begin{cases} 1 & \text{si} \quad T \geq t_{n-1,\alpha} \\ 0 & \text{si} \quad T < t_{n-1,\alpha} \ . \end{cases}$$

Debemos probar ahora que este test tiene realmente nivel α, es decir que,

$$P_{\mu,\sigma^2}(T \geq t_{n-1,\alpha}) \leq \alpha \qquad \forall\, \mu \leq \mu_0 \ .$$

Para ello necesitaremos la siguiente definición.

Definición 1. Llamaremos distribución de Student no central con n grados de libertad y parámetro de no centralidad Δ, $-\infty < \Delta < \infty$, que simbolizaremos por $\mathcal{T}_n(\Delta)$ a la distribución de

$$\frac{U+\Delta}{\sqrt{V/n}}$$

donde U tiene distribución $N(0,1)$ donde V tiene distribución χ^2_n siendo U y V independientes.

Teorema 1. *Sea X_Δ una variable aleatoria con distribución de Student no central $\mathcal{T}_n(\Delta)$, definamos $c_{n,k}(\Delta)$ por*

$$c_{n,k}(\Delta) = P(X \geq k),$$

luego, $c_{n,k}(\Delta)$ es una función monótona creciente de Δ.

Demostración. Como X_Δ tiene distribución $\mathcal{T}_n(\Delta)$; se puede escribir

$$X_\Delta = \frac{U+\Delta}{\sqrt{V/n}}$$

donde U es una variable aleatoria $N(0,1)$ y V tiene distribución χ^2_n, independientes. Luego,

$$c_{n,k}(\Delta) = P(X_\Delta \geq k) = E\left[P(X_\Delta \geq k|V)\right],$$

pero

$$P(X_\Delta \geq k|V = v) = P\left(\frac{U+\Delta}{\sqrt{v/n}} \geq k\,|V = v\right) = 1 - \Phi(k\sqrt{\frac{v}{n}} - \Delta)\,.$$

Luego esta última probabilidad, para k, n y v fijos, es una función creciente de Δ. Por lo tanto, si $\Delta_1 < \Delta_2$ se tiene

$$P(X_{\Delta_1} \geq k|V = v) < P(X_{\Delta_2} \geq k|V = v)$$

con lo cual, tomando esperanza se obtiene

$$E\left(P(X_{\Delta_1} \geq k)|V\right) < E\left(P(X_{\Delta_2} \geq k)|V\right)$$

o sea

$$P(X_{\Delta_1} \geq k) < P(X_{\Delta_2} \geq k),$$

y por lo tanto $c_{n,k}(\Delta)$ es creciente en Δ.

Volvamos ahora al Ejemplo 3. Vamos a mostrar que el test φ dado por

$$\varphi(\mathbf{X}) = \begin{cases} 1 & \text{si} \quad T \geq t_{n-1,\alpha} \\ 0 & \text{si} \quad T < t_{n-1,\alpha} \end{cases}$$

tiene nivel de significación α. Como

$$T = \frac{\sqrt{n}(\overline{X} - \mu_0)}{\sqrt{\frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X})^2}} = \frac{\sqrt{n}\frac{(\overline{X}-\mu_0)}{\sigma}}{\sqrt{\frac{1}{n-1}\frac{\sum_{i=1}^n (X_i-\overline{X})^2}{\sigma^2}}}$$

resulta

$$T = \frac{\sqrt{n}\frac{(\overline{X}-\mu)}{\sigma} + \sqrt{n}\frac{(\mu-\mu_0)}{\sigma}}{\sqrt{\frac{1}{n-1}\frac{\sum_{i=1}^n (X_i-\overline{X})^2}{\sigma^2}}}\,.$$

Llamando $U = \sqrt{n}(\overline{X} - \mu)/\sigma$ y $V = \sum_{i=1}^n (X_i - \overline{X})^2/\sigma^2$ se tiene que U y V son independientes, y cuando los valores de los parámetros son μ y

σ^2, U tiene distribución $N(0,1)$ y V tiene distribución χ^2_{n-1}. Luego T tiene distribución $\mathcal{T}_{n-1}(\Delta)$ donde $\Delta = \sqrt{n}(\mu - \mu_0)/\sigma$. Además,

$$\beta_\varphi(\mu,\sigma^2) = P_{\mu,\sigma^2}(T \geq t_{n-1,\alpha}) = c_{n-1,t_{n-1,\alpha}}(\Delta)\ .$$

Resulta, por el Teorema 1, que $\beta_\varphi(\mu,\sigma^2)$ es una función creciente de μ para cada σ^2 fijo. Como, por otra parte, $\beta_\varphi(\mu_0,\sigma^2) = \alpha$, para todo σ^2, se tiene

$$\beta_\varphi(\mu,\sigma^2) < \alpha \qquad \forall\, \mu < \mu_0$$

y el test φ tiene nivel de significación α. También, a partir de la expresión de $\beta_\varphi(\mu,\sigma^2)$ se obtiene que el test φ es insesgado.

Análogamente, en el caso de testear H : $\mu \geq \mu_0$ contra K : $\mu < \mu_0$, el test del cociente de máxima verosimilitud vendrá dado por

$$\varphi(\mathbf{X}) = \begin{cases} 1 & \text{si} \quad T \leq t_{n-1,\alpha} \\ 0 & \text{si} \quad T > t_{n-1,\alpha}\ . \end{cases}$$

Para calcular la potencia de estos tests se pueden utilizar las tablas construídas por Owen [4].

Ejemplo 4. Supongamos nuevamente que tenemos una muestra aleatoria $X_1,\dots,X_n$ de una distribución $N(\mu,\sigma^2)$ con μ y σ^2 desconocidos. Se desea testear H : $\sigma^2 \leq \sigma_0^2$ contra K : $\sigma^2 > \sigma_0^2$.

Se deduce haciendo un razonamiento análogo al ejemplo anterior que el test del cociente de máxima verosimilitud es de la forma

$$\varphi(\mathbf{X}) = \begin{cases} 1 & \text{si} \ \displaystyle\sum_{i=1}^n (X_i - \overline{X})^2 \geq k_\alpha \\ 0 & \text{si} \ \displaystyle\sum_{i=1}^n (X_i - \overline{X})^2 < k_\alpha\ . \end{cases}$$

La constante k_α se debe elegir de manera que

$$\sup_{\sigma^2 \leq \sigma_0^2} P_{\sigma^2}\left(\sum_{i=1}^n (X_i - \overline{X})^2 \geq k_\alpha\right) = \alpha\ .$$

Determinemos k_α por el valor de σ^2 más desfavorable, o sea, σ_0^2. Luego, debemos elegir k_α tal que

$$P_{\sigma_0^2}\left(\sum_{i=1}^n (X_i - \overline{X})^2 \geq k_\alpha\right) = \alpha$$

o equivalentemente

$$P_{\sigma_0^2}\left(\frac{\sum_{i=1}^n (X_i-\overline{X})^2}{\sigma_0^2} \geq \frac{k_\alpha}{\sigma_0^2}\right) = \alpha \ .$$

Como $\sum_{i=1}^n (X_i-\overline{X})^2/\sigma_0^2$ tiene distribución χ^2_{n-1} cuando $\sigma^2 = \sigma_0^2$, se tiene que

$$k_\alpha = \sigma_0^2 \chi^2_{n-1,\alpha} \quad .$$

Para mostrar que el test tiene realmente nivel de significación α, bastará mostrar que la función de potencia es una función creciente y esto se deduce como sigue. Sea $D_n(k) = P(Y \geq k)$, donde Y es una variable aleatoria con distribución χ^2_n. Luego

$$\begin{aligned} \beta_\varphi(\sigma^2) &= P_{\sigma^2}\left(\sum_{i=1}^n (X_i-\overline{X})^2 \geq \sigma_0^2\chi^2_{n-1,\alpha}\right) \\ &= P_{\sigma^2}\left(\frac{\sum_{i=1}^n (X_i-\overline{X})^2}{\sigma^2} \geq \frac{\sigma_0^2\chi^2_{n-1,\alpha}}{\sigma^2}\right) \\ &= D_{n-1}\left(\frac{\sigma_0^2\chi^2_{n-1,\alpha}}{\sigma^2}\right) \ , \end{aligned}$$

ya que cuando la varianza de cada X_i es σ^2 resulta que $\sum_{i=1}^n (X_i-\overline{X})^2/\sigma^2$ tiene distribución χ^2_{n-1}.

Como $D_n(k)$ es una función decreciente de k, $\beta_\varphi(\sigma^2)$ es una función creciente de σ^2.

Ejemplo 5. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución $N(\mu, \sigma^2)$ con μ y σ^2 desconocidos y supongamos que se quiere testear $\mathrm{H} : \sigma^2 = \sigma_0^2$ contra $\mathrm{K} : \sigma^2 \neq \sigma_0^2$.

En este caso, el test del cociente de máxima verosimilitud es de la forma

$$\varphi(\mathbf{X}) = \begin{cases} 1 & \text{si} \quad \dfrac{\sum_{i=1}^n (X_i-\overline{X})^2}{\sigma_0^2} \geq k'_\alpha \\ \\ 1 & \text{si} \quad \dfrac{\sum_{i=1}^n (X_i-\overline{X})^2}{\sigma_0^2} < k''_\alpha \\ \\ 0 & \text{en cualquier otro caso,} \end{cases}$$

Para que φ tenga nivel de significación α, se debe cumplir que

$$\begin{aligned}\beta_\varphi(\sigma_0^2) &= P_{\sigma_0^2}\left(\frac{\sum_{i=1}^n (X_i-\overline{X})^2}{\sigma_0^2} \geq k'_\alpha\right)\\ &\quad + P_{\sigma_0^2}\left(\frac{\sum_{i=1}^n (X_i-\overline{X})^2}{\sigma_0^2} < k''_\alpha\right) = \alpha \,.\end{aligned}$$

Luego, se debe tener que

$$k'_\alpha = \chi^2_{n-1,\beta} \quad \text{y} \quad k''_\alpha = \chi^2_{n-1,1-\gamma} \tag{6.28}$$

con $\beta+\gamma=\alpha$.

Si queremos que el test resulte insesgado, la derivada de la función de potencia debe ser cero en σ_0. Pero,

$$\begin{aligned}\beta_\varphi(\sigma^2) &= P_{\sigma^2}\left(\frac{\sum_{i=1}^n (X_i-\overline{X})^2}{\sigma^2} \geq \frac{k'_\alpha\sigma_0^2}{\sigma^2}\right)\\ &\quad + P_{\sigma^2}\left(\frac{\sum_{i=1}^n (X_i-\overline{X})^2}{\sigma^2} < \frac{k''_\alpha\sigma_0^2}{\sigma^2}\right) \,,\end{aligned}$$

con lo cual si llamamos Y a una variable con distribución χ^2_{n-1} obtenemos

$$\begin{aligned}\beta_\varphi(\sigma^2) &= P_{\sigma^2}\left(Y \geq \frac{k'_\alpha\sigma_0^2}{\sigma^2}\right) + P_{\sigma^2}\left(Y < \frac{k''_\alpha\sigma_0^2}{\sigma^2}\right)\\ &= 1 - P_{\sigma^2}\left(Y < \frac{k'_\alpha\sigma_0^2}{\sigma^2}\right) + P_{\sigma^2}\left(Y < \frac{k''_\alpha\sigma_0^2}{\sigma^2}\right) \,.\end{aligned}$$

Por lo tanto, si $f_Y(y)$ indica a la densidad de Y, la condición $\beta'_\varphi(\sigma_0^2)=0$ es equivalente a

$$f_Y(k'_\alpha)\, k'_\alpha = f_Y(k''_\alpha)\, k''_\alpha$$

de donde se obtiene que k'_α y k''_α deberán ser elegidos de forma que

$$e^{-k'_\alpha/2}(k'_\alpha)^{\frac{n-1}{2}} = e^{-k''_\alpha/2}(k''_\alpha)^{\frac{n-1}{2}} \tag{6.29}$$

En la práctica se eligen $\gamma=\alpha/2$ y $\beta=\alpha/2$, aunque no satisfaga (6.29). Se puede mostrar que para $n\to\infty$ los β y γ que satisfacen (6.28) y hacen que se satisfaga (6.29) se aproximan a los valores elegidos. En realidad, la aproximación es buena con tal que n no sea muy pequeño. Luego, el test que se usa viene dado por

$$\varphi(\mathbf{X}) = \begin{cases} 1 \quad \text{si} \begin{cases} \sum_{i=1}^{n}(X_i - \overline{X})^2 \geq \sigma_0^2 \chi^2_{n-1,\frac{\alpha}{2}} \\ \\ \sum_{i=1}^{n}(X_i - \overline{X})^2 \leq \sigma_0^2 \chi^2_{n-1,1-\frac{\alpha}{2}} \end{cases} \\ 0 \quad \text{si} \quad \sigma_0^2 \chi^2_{n-1,1-\frac{\alpha}{2}} \leq \sum_{i=1}^{n}(X_i - \overline{X})^2 \leq \sigma_0^2 \chi^2_{n-1,\frac{\alpha}{2}} \end{cases}$$

Se puede mostrar que los tests obtenidos en los Ejemplos 1 a 5 son IUMP. Para estos resultados pueden consultarse el Capítulo 5 de Lehmann [3] o el Capítulo 5 de Ferguson [2].

6.7 Test con nivel de significación asintótico

La mayoría de los test de hipótesis, por ejemplo, los del cociente de verosimilitud, son de la forma

$$\varphi(\mathbf{X}) = \begin{cases} 1 & \text{si } T > k_\alpha \\ \gamma_\alpha & \text{si } T = k_\alpha \\ 0 & \text{si } T < k_\alpha \end{cases}$$

donde T es un estadístico basado en la muestra. Para encontrar k_α se requiere conocer la distribución de T para $\boldsymbol{\theta} \in \Theta_1$. Como en muchos casos esta distribución es muy compleja se puede reemplazar esta distribución por una asintótica. En este caso el test tendra un nivel de significación aproximado al deseado para muestras grandes. Esto motiva la siguiente definición.

Definición 1. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución perteneciente a la familia $F(\mathbf{x}, \boldsymbol{\theta})$, $\boldsymbol{\theta} \in \Theta$ y supongamos que se quiere testear la hipótesis H : $\boldsymbol{\theta} \in \Theta_1$ contra K : $\boldsymbol{\theta} \in \Theta_2$. Se dirá que una sucesión de test $\varphi_n(X_1, \ldots, X_n)$ tiene nivel de significación asintótico α si

$$\lim_{n\to\infty} \sup_{\boldsymbol{\theta} \in \Theta_1} \beta_{\varphi_n}(\boldsymbol{\theta}) = \alpha$$

Es decir, que el nivel del test $\varphi_n(X_1, \ldots, X_n)$ se acerca a α cuando el tamaño de la muestra tiende a infinito.

Ejemplo 1. Supongamos que $X_1, \ldots, X_n$ es una muestra aleatoria de una distribución desconocida con media μ y varianza σ^2.

Supongamos que se quiere testear H : $\mu \leq \mu_0$ contra K : $\mu > \mu_0$. Llamemos

$$\overline{X} = \frac{\sum_{i=1}^n X_i}{n} \qquad \text{y} \qquad s^2 = \frac{\sum_{i=1}^n (X_i - \overline{X})^2}{n-1} \,.$$

Ya hemos demostrado que

$$\sqrt{n}\frac{(\overline{X} - \mu_0)}{s}$$

converge en distribución a la $N(0,1)$ cuando la esperanza de las variables X_i es μ_0. Luego, si definimos

$$\varphi_n(X_1, \ldots, X_n) = \begin{cases} 1 & \text{si} \quad \sqrt{n}\dfrac{(\overline{X} - \mu_0)}{s} \geq z_\alpha \\ & \\ 0 & \text{si} \quad \sqrt{n}\dfrac{(\overline{X} - \mu_0)}{s} < z_\alpha \end{cases}$$

este test tiene nivel de significación asintótico α.

Del mismo modo, si se quiere testear H : $\mu = \mu_0$ contra K : $\mu \neq \mu_0$, un test de nivel de significación asintótico α será

$$\varphi_n(X_1, \ldots, X_n) = \begin{cases} 1 & \text{si} \quad \sqrt{n}\dfrac{|\overline{X} - \mu_0|}{s} \geq z_\alpha \\ & \\ 0 & \text{si} \quad \sqrt{n}\dfrac{|\overline{X} - \mu_0|}{s} < z_\alpha \end{cases}$$

6.7.1 Distribución asintótica del test del cociente de máxima verosimilitud

Sea $\mathbf{X} = (X_1, \ldots, X_n)$ una muestra aleatoria de una distribución de densidad o probabilidad dada por $p(\mathbf{x}, \boldsymbol{\theta})$ con $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_p) \in \Theta$, donde Θ es un conjunto de $\mathbb{R}^p$ que contiene una esfera.

Supongamos que Θ_1 es un conjunto de dimensión menor que p, digamos de dimensión $p - j$, donde $1 \leq j \leq p$. Θ_1 puede venir expresado de varias formas diferentes. Por ejemplo, puede venir dado por j relaciones funcionales entre los parámetros $\theta_1, \ldots, \theta_p$, es decir,

$$\Theta_1 = \{\boldsymbol{\theta} \in \Theta : g_1(\boldsymbol{\theta}) = 0;\; g_2(\boldsymbol{\theta}) = 0, \ldots, g_j(\boldsymbol{\theta}) = 0\}$$

o bien, en forma paramétrica

$$\Theta_1 = \{\boldsymbol{\theta} = (\theta_1, ...\theta_p) : \theta_1 = h_1(\boldsymbol{\lambda}), \ldots, \theta_p = h_p(\boldsymbol{\lambda}),\ \boldsymbol{\lambda} \in \Lambda\},$$

donde $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_{p-j})$ y $\Lambda \subset I\!R^{p-j}$ de dimensión $p - j$.

Supongamos que se está interesado en el siguiente problema de test de hipótesis:

$$\mathrm{H} : \boldsymbol{\theta} \in \Theta_1 \quad \text{contra} \quad \mathrm{K} : \boldsymbol{\theta} \in \Theta_2$$

con $\Theta = \Theta_1 \cup \Theta_2$. Luego, el test del cociente de máxima verosimilitud es de la forma

$$\varphi(\mathbf{X}) = \begin{cases} 1 & \text{si} \quad L^*(\mathbf{X}) \leq k_\alpha \\ 0 & \text{si} \quad L^*(\mathbf{X}) > k_\alpha \end{cases}$$

donde

$$L^*(\mathbf{X}) = \frac{\sup_{\boldsymbol{\theta} \in \Theta_1} p(\mathbf{X}, \boldsymbol{\theta})}{\sup_{\boldsymbol{\theta} \in \Theta} p(\mathbf{X}, \boldsymbol{\theta})}\ .$$

Para determinar k_α es necesario conocer la distribución de $L^*(\mathbf{X})$ bajo H. Muchas veces esta es muy complicada y puede depender del valor particular $\boldsymbol{\theta} \in \Theta_1$ que se considere. Sin embargo, se puede mostrar que, bajo condiciones de regularidad muy generales en $p(\mathbf{x}, \boldsymbol{\theta})$, la distribución asintótica de $Z = -2 \ln L^*$ cuando $\boldsymbol{\theta} \in \Theta_1$ es χ^2_j. Luego un test de nivel de significación asintótico α está dado por

$$\varphi(\mathbf{X}) = \begin{cases} 1 & \text{si} \quad Z \geq \chi^2_{j,\alpha} \\ \\ 0 & \text{si} \quad Z < \chi^2_{j,\alpha} \end{cases}$$

Para ver la teoría asintótica del test del cociente de verosimilitud se puede ver Wald [5] y Chernoff [1]. Nosotros sólo daremos la distribución en el caso particular $\Theta \subset I\!R$ y $\mathrm{H} : \theta = \theta_0$ contra $\mathrm{K} : \theta \neq \theta_0$.

Teorema 1. *Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución discreta o continua con densidad perteneciente a la familia $p(x, \theta)$ con $\theta \in \Theta$ y Θ un abierto en $I\!R$. Indiquemos por $p(\mathbf{x}, \theta)$ la densidad conjunta del vector $\mathbf{X} = (X_1, \ldots, X_n)$.*

Supongamos que se cumplen las siguientes condiciones (en lo que sigue suponemos que $\mathbf{X}$ es continuo, para el caso discreto habrá que reemplazar todos los signos $\int$ por $\sum$):

(A) *El conjunto $\mathcal{S} = \{x : p(x, \theta) > 0\}$ es independiente de θ.*

(B) *Para todo x, $p(x,\theta)$ tiene derivada tercera respecto de θ continua y tal que*

$$\Big|\frac{\partial^3 \ln p(x,\theta)}{\partial\theta^3}\Big| = \Big|\frac{\partial^2 \psi(x,\theta)}{\partial\theta^2}\Big| \leq K$$

para todo $x \in \mathcal{S}$ y para todo $\theta \in \Theta$, donde

$$\psi(x,\theta) = \frac{\partial \ln p(x,\theta)}{\partial\theta} \ .$$

(C) *Si $h(\mathbf{X})$ es un estadístico tal que $E_\theta[|h(\mathbf{X})|] < \infty$ para todo $\theta \in \Theta$ entonces se tiene*

$$\frac{\partial}{\partial\theta}\left[\int_{-\infty}^{\infty}\dots\int_{-\infty}^{\infty} h(\mathbf{x})p(\mathbf{x},\theta)d\mathbf{x}\right] = \int_{-\infty}^{\infty}\dots\int_{-\infty}^{\infty} h(\mathbf{x})\frac{\partial p(\mathbf{x},\theta)}{\partial\theta}d\mathbf{x}$$

donde $d\mathbf{x} = (dx_1, \dots, dx_n)$.

(D)

$$0 < I_1(\theta) = E_\theta\left[\left(\frac{\partial \ln p(X_1,\theta)}{\partial\theta}\right)^2\right] < \infty \ .$$

Sea $\widehat{\theta}_n$ un estimador de máxima verosimilitud de θ consistente, entonces si

$$L^*(\mathbf{X}) = \frac{p(\mathbf{X},\theta_0)}{\sup_{\theta\in\Theta} p(\mathbf{X},\theta)} = \frac{p(\mathbf{X},\theta_0)}{p(\mathbf{X},\widehat{\theta}_n)} \ .$$

se tiene que $Z = -2\ln\left(L^*(\mathbf{X})\right)$ tiene distribución asintótica χ^2_1 con lo cual el test

$$\varphi(\mathbf{X}) = \begin{cases} 1 & si \quad Z \geq \chi^2_{1,\alpha} \\ & \\ 0 & si \quad Z < \chi^2_{1,\alpha} \end{cases}$$

tiene nivel de significación asintótico α.

Demostración. Sea

$$\ell(\theta) = \ln p(\mathbf{X},\theta) = \sum_{i=1}^{n} \ln\left(p(X_i,\theta)\right) \ .$$

Indiquemos además por ℓ', ℓ'' y ℓ''' las derivadas hasta el orden tres respecto de θ de la función ℓ y por

$$\psi'(x,\theta) = \frac{\partial\psi(x,\theta)}{\partial\theta} \quad \text{ y } \quad \psi''(x,\theta) = \frac{\partial^2\psi(x,\theta)}{\partial\theta^2}.$$

Luego, $\widehat{\theta}_n$ verifica

$$\ell'(\widehat{\theta}_n) = \sum_{i=1}^{n} \psi(X_i, \widehat{\theta}_n) = 0 \, .$$

Con lo cual, desarrollando en serie de Taylor alrededor de $\widehat{\theta}_n$ se obtiene:

$$\begin{aligned} \ell(\theta_0) - \ell(\widehat{\theta}_n) &= \ell'(\widehat{\theta}_n)(\theta_0 - \widehat{\theta}_n) + \frac{1}{2}\ell''(\xi_n^1)(\theta_0 - \widehat{\theta}_n)^2 \\ &= \frac{1}{2}(\theta_0 - \widehat{\theta}_n)^2 \left(\sum_{i=1}^{n} \psi'(X_i, \xi_n^1) \right) \\ &= \frac{1}{2} \left(n(\theta_0 - \widehat{\theta}_n)^2 \right) \frac{1}{n} \sum_{i=1}^{n} \psi'(X_i, \theta_0) + R_n \, , \end{aligned}$$

donde ξ_n^1 es un punto intermedio entre $\widehat{\theta}_n$ y θ_0 y

$$R_n = \frac{1}{2} \left(n(\theta_0 - \widehat{\theta}_n)^2 \right) \left(\frac{1}{n} \sum_{i=1}^{n} \psi'(X_i, \xi_n^1) - \frac{1}{n} \sum_{i=1}^{n} \psi'(X_i, \theta_0) \right) \, .$$

Aplicando el Teorema del valor medio se obtiene

$$R_n = \frac{1}{2} \left(n(\theta_0 - \widehat{\theta}_n)^2 \right) \left(\frac{1}{n} \sum_{i=1}^{n} \psi''(X_i, \xi_n^2)(\xi_n^1 - \theta_0) \right) \tag{6.30}$$

donde ξ_n^2 es un punto intermedio entre ξ_n^1 y θ_0. Observemos que por ser $\widehat{\theta}_n$ consistente, se obtiene entonces que $\xi_n^j \to \theta_0$ en probabilidad para $j = 1, 2$.

Reemplazando, obtenemos que

$$Z = 2 \left(\ell(\widehat{\theta}_n) - \ell(\theta_0) \right) = \left(n(\widehat{\theta}_n - \theta_0)^2 \right) A_n - R_n \tag{6.31}$$

donde $A_n = -\frac{1}{n} \sum_{i=1}^{n} \psi'(X_i, \theta_0)$.

Hemos visto en el Teorema 1 de 3.17 que cuando $\theta = \theta_0$

$$\sqrt{n}(\widehat{\theta}_n - \theta_0) \to \mathrm{N}(0, \frac{1}{I_1(\theta_0)}) \quad \text{en distribución,}$$

con lo cual

$$I_1(\theta_0) \, n \, (\widehat{\theta}_n - \theta_0)^2 \to \chi_1^2 \quad \text{en distribución.} \tag{6.32}$$

Por otra parte, la ley de los grandes números implica que

$$\frac{1}{n} \sum_{i=1}^{n} \psi'(X_i, \theta_0) \to E(\psi'(X_1, \theta_0)) \quad \text{en probabilidad.} \tag{6.33}$$

Pero,

$$E_{\theta_0}(\psi'(X_1,\theta_0)) = -I_1(\theta_0)\ ,$$

luego, usando (6.32) y (6.33) se obtiene que

$$\left(n\,(\widehat{\theta}_n-\theta_0)^2\right)\ A_n \to \chi^2_1 \quad \text{en distribución.} \tag{6.34}$$

Por lo tanto, a partir de (6.31) y (6.34) deducimos que bastará probar que

$$R_n \to 0 \qquad \text{en probabilidad.} \tag{6.35}$$

Como $|\psi''(X_i,\theta)| \le K$ para todo θ, se tiene que

$$\left|\frac{1}{2n}\sum_{i=1}^{n}\psi''(X_i,\xi_n^2)(\xi_n^1-\theta_0)\right| \le \frac{K}{2}|(\xi_n^1-\theta_0)|$$

y luego como $\xi_n^1 \to \theta_0$ en probabilidad se deduce que:

$$\frac{1}{n}\sum_{i=1}^{n}\psi''(X_i,\xi_n^2)(\xi_n^1-\theta_0) \to 0 \quad \text{en probabilidad.} \tag{6.36}$$

Pero, (6.32) implica que $n(\widehat{\theta}_n-\theta_0)^2$ está acotado en probabilidad, luego (6.35)se obtiene de (6.30) y (6.36).

Ejemplo 1. Sea $X_1,\ldots,X_n$ una muestra de una distribución perteneciente a la familia $Bi(\theta,1)$, $\ 0<\theta<1$, y supongamos que se quiere testear $\mathrm{H}:\theta=\theta_0$ contra $\mathrm{K}:\theta\neq\theta_0$. Luego el test del cociente de máxima verosimilitud es

$$L^* = \frac{p(\mathbf{x},\theta_0)}{\sup_{\theta\in\Theta}p(x,\theta)} = \frac{\theta_0^T(1-\theta_0)^{n-T}}{\overline{X}^T(1-\overline{X})^{n-T}}\ ,$$

donde $T=\sum_{i=1}^n X_i$. Luego,

$$Z = -2\ln L^* = 2T\ln\frac{\overline{X}}{\theta_0} + 2(n-T)\ln\frac{(1-\overline{X})}{1-\theta_0}$$

tiene una distribución asintótica χ^2_1 bajo H y un test de nivel asintótico estará dado por

$$\varphi(\mathbf{X}) = \begin{cases} 1 & \text{si} \quad Z>\chi^2_{1,\alpha} \\ \\ 0 & \text{si} \quad Z<\chi^2_{1,\alpha}\ . \end{cases}$$

6.8 Relación entre regiones de confianza y test

En esta sección se estudiará la relación que existe entre tests y regiones de confianza.

Supongamos que se tiene un vector aleatorio $\mathbf{X}$ con distribución perteneciente a la familia $F(\mathbf{x}, \boldsymbol{\theta})$ con $\boldsymbol{\theta} \in \Theta$ y supongamos que para cada $\boldsymbol{\theta}_0$ se tiene un test no aleatorizado de nivel α, $\varphi_{\boldsymbol{\theta}_0}$, para H : $\boldsymbol{\theta} = \boldsymbol{\theta}_0$ contra K : $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$.

Se puede construir una región de confianza de nivel $(1-\alpha)$ para $\boldsymbol{\theta}$ definiendo

$$S(\mathbf{X}) = \{\boldsymbol{\theta} : \varphi_{\boldsymbol{\theta}}(\mathbf{X}) = 0\}$$

Es decir, $S(\mathbf{X})$ es el conjunto de todos los $\boldsymbol{\theta} \in \Theta$ tales que la hipótesis de que el valor verdadero es $\boldsymbol{\theta}$, es aceptada cuando se observa $\mathbf{X}$.

Demostraremos que $S(\mathbf{X})$ así definida, es una región de confianza de nivel $1-\alpha$ para $\boldsymbol{\theta}$

$$P_{\boldsymbol{\theta}}(\boldsymbol{\theta} \in S(\mathbf{X})) = P_{\boldsymbol{\theta}}(\varphi_{\boldsymbol{\theta}}(\mathbf{X}) = 0) = 1 - P_{\boldsymbol{\theta}}(\varphi_{\boldsymbol{\theta}}(\mathbf{X}) = 1) = 1 - \alpha \; .$$

Recíprocamente, si se tiene una región de confianza $S(\mathbf{X})$ de nivel $1-\alpha$ para $\boldsymbol{\theta}$, se puede construir un test de nivel α, $\varphi_{\boldsymbol{\theta}_0}$, para H : $\boldsymbol{\theta} = \boldsymbol{\theta}_0$ contra K : $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$.

Definamos

$$\varphi_{\boldsymbol{\theta}_0}(\mathbf{X}) = \begin{cases} 1 & \text{si } \boldsymbol{\theta}_0 \notin S(\mathbf{X}) \\ 0 & \text{si } \boldsymbol{\theta}_0 \in S(\mathbf{X}) \; . \end{cases}$$

Mostraremos que este test tiene realmente nivel de significación α. Efectivamente,

$$P_{\boldsymbol{\theta}_0}(\varphi_{\boldsymbol{\theta}_0}(\mathbf{X}) = 1) = P_{\boldsymbol{\theta}_0}(\boldsymbol{\theta}_0 \notin S(\mathbf{X})) = 1 - P_{\boldsymbol{\theta}_0}(\boldsymbol{\theta}_0 \in S(\mathbf{X})) = 1 - (1-\alpha) = \alpha \; .$$

Ejemplo 1. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución $N(\mu, \sigma^2)$.

En el capítulo anterior hemos demostrado que un intervalo de confianza a nivel $(1-\alpha)$ para μ viene dado por

$$S(\mathbf{X}) = [\overline{X} - t_{n-1,\frac{\alpha}{2}} \frac{s}{\sqrt{n}} \; , \; \overline{X} + t_{n-1,\frac{\alpha}{2}} \frac{s}{\sqrt{n}}]$$

Construyamos el test correspondiente de nivel α para

$$\text{H} : \mu = \mu_0 \quad \text{contra} \quad \text{K}: \mu \neq \mu_0$$

$$\varphi_{\mu_0}(\mathbf{X}) = \begin{cases} 1 & \text{si} \quad \mu_0 \notin S(\mathbf{X}) \\ 0 & \text{si} \quad \mu_0 \in S(\mathbf{X}) \end{cases}$$

pero $\mu_0 \in S(\mathbf{X})$ si y sólo si $|\mu_0 - \overline{X}| \leq t_{n-1,\frac{\alpha}{2}}(s/\sqrt{n})$, luego

$$\varphi_{\mu_0}(\mathbf{X}) = \begin{cases} 1 & \text{si} \quad \sqrt{n}\dfrac{|\overline{X} - \mu_0|}{s} > t_{n-1,\frac{\alpha}{2}} \\ \\ 0 & \text{si} \quad \sqrt{n}\dfrac{|\overline{X} - \mu_0|}{s} \leq t_{n-1,\frac{\alpha}{2}} \ . \end{cases}$$

Por lo tanto, este test coincide con el obtenido en el Ejemplo 2 de 6.6, cuando obtuvimos el test del CMV para este problema. Recíprocamente, a partir de esta familia de tests si se usara el procedimiento indicado anteriormente para obtener intervalos de confianza, se llegará al intervalo inicial.

Ejemplo 2. Sea $X_1, \ldots, X_{n_1}$ una muestra aleatoria de una distribución $N(\mu_1, \sigma^2)$ y sea $Y_1, \ldots, Y_{n_2}$ una muestra aleatoria de una distribución $N(\mu_2, \sigma^2)$ independiente de la primera. Se ha visto en el Capítulo 5 que

$$T = \sqrt{\frac{n_1\, n_2}{n_1 + n_2}}\frac{(\overline{X} - \overline{Y} - (\mu_1 + \mu_2))}{s}$$

donde

$$s^2 = \frac{1}{n_2 + n_1 - 2}\left(\sum_{i=1}^{n_1}(X_i - \overline{X})^2 + \sum_{i=1}^{n_2}(Y_i - \overline{Y})^2\right)$$

tiene distribución de Student con $n_1 + n_2 - 2$ grados de libertad y que un intervalo de confianza para $\mu_1 - \mu_2$ está dado por

$$S(\mathbf{X}) = \left[\overline{X} - \overline{Y} - t_{n_1+n_2-2,\frac{\alpha}{2}}\ s\sqrt{\frac{n_1 + n_2}{n_1\, n_2}}\ , \overline{X} - \overline{Y} + t_{(n_1+n_2-2),\frac{\alpha}{2}}\ s\sqrt{\frac{n_1 + n_2}{n_1\, n_2}}\right]$$

Luego, si se quiere testear H : $\mu_1 - \mu_2 = \lambda_0$ contra K : $\mu_1 - \mu_2 \neq \lambda_0$, con nivel de significación α, se puede obtener un test haciendo

$$\varphi_{\lambda_0}(\mathbf{X}) = \begin{cases} 1 & \text{si } \lambda_0 \notin S(\mathbf{X}) \\ 0 & \text{si } \lambda_0 \in S(\mathbf{X}) \end{cases}$$

pero $\lambda_0 \in S(\mathbf{X})$ si y sólo si

$$\sqrt{\frac{n_1\, n_2}{n_1 + n_2}}\frac{|\overline{X} - \overline{Y} - \lambda_0|}{s} \leq t_{n_1+n_2-2,\frac{\alpha}{2}}.$$

Por lo tanto,

$$\varphi_{\lambda_0}(\mathbf{X}) = \begin{cases} 1 & \text{si} \quad \sqrt{\dfrac{n_1\, n_2}{n_1+n_2}}\dfrac{|\overline{X}-\overline{Y}-\lambda_0|}{s} \geq t_{n_1+n_2-2,\frac{\alpha}{2}} \\ 0 & \text{si} \quad \sqrt{\dfrac{n_1\, n_2}{n_1+n_2}}\dfrac{|\overline{X}-\overline{Y}-\lambda_0|}{s} < t_{(n_1+n_2-2),\frac{\alpha}{2}}\ . \end{cases}$$

Hasta aquí hemos estudiado la relación entre regiones de confianza de nivel $1-\alpha$ para $\boldsymbol{\theta}$ y test de hipótesis para las hipótesis $\mathrm{H} : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ contra $\mathrm{K} : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0$. Esta situación se puede generalizar al caso de

$$\mathrm{H} : \boldsymbol{\theta} = \boldsymbol{\theta}_0 \quad \text{contra} \quad \mathrm{K} : K(\boldsymbol{\theta}_0)$$

donde $K(\boldsymbol{\theta}_0)$ indica una alternativa cualquiera que no contiene a $\boldsymbol{\theta}_0$ si para cada $\boldsymbol{\theta}_0 \in \Theta$ se tiene un test de nivel α, $\varphi_{\boldsymbol{\theta}_0}$, resultará que

$$S(\mathbf{X}) = \{\boldsymbol{\theta} \in \Theta : \varphi_{\boldsymbol{\theta}}(\mathbf{X}) = 0\}$$

será una región con nivel de confianza $1-\alpha$. De la misma forma que antes $S(\mathbf{X})$ será el conjunto de todos los $\boldsymbol{\theta} \in \Theta$ tales que la hipótesis de que $\boldsymbol{\theta}$ es el verdadero valor es aceptada cuando se observa $\mathbf{X}$.

6.9 Cotas de confianza óptimas

Se verá ahora cómo la existencia de tests uniformemente más potentes para hipótesis unilaterales permite la construcción de intervalos de confianza unilaterales óptimas en el sentido definido en la sección 5.9.

Hemos demostrado en 6.4 que para *familias de cociente de verosimilitud monótono* existen tests UMP para las hipótesis:

$$\mathrm{H}_1 : \theta = \theta_0 \quad \text{contra} \quad \mathrm{K}_1 : \theta > \theta_0$$
$$\mathrm{H}_2 : \theta = \theta_0 \quad \text{contra} \quad \mathrm{K}_2 : \theta < \theta_0$$

En estos casos vale el siguiente teorema

Teorema 1. *Sea φ_{θ_0} el test no aleatorizado (si existe) UMP para H_1 contra K_1, de nivel α. Dada $X_1, \ldots, X_n$ y siendo*

$$S(\mathbf{X}) = \{\theta \in \Theta : \varphi_\theta(\mathbf{X}) = 0\}$$

i) *$S(\mathbf{X})$ es una región de confianza de nivel $1-\alpha$ para θ.*

ii) *Si $\varphi^*_{\theta_0}$ es cualquier otro test no aleatorizado de nivel α para esas hipótesis y*
$$S^*(\mathbf{X}) = \{\theta \in \Theta : \varphi^*_\theta(\mathbf{X}) = 0\}$$
entonces $P_\theta\{\theta_0 \in S(\mathbf{X})\} \leq P_\theta\{\theta_0 \in S^*(\mathbf{X})\}$ para todo $\theta > \theta_0$.

Demostración. i) Por la definición de $S(\mathbf{X})$ sabemos que $\theta \in S(\mathbf{X})$ si y sólo si $\varphi_\theta(\mathbf{X}) = 0$, luego

$$P_\theta\{\theta \in S(\mathbf{X})\} = P_\theta\{\varphi_\theta(\mathbf{X}) = 0\} = 1 - \alpha$$

por ser φ_θ de nivel α.
ii) Igual que en i) $S^*(\mathbf{X})$ será una región de confianza de nivel $1-\alpha$. Por ser $\varphi_{\theta_0}(\mathbf{X})$ el test UMP para H_1 contra K_1 resulta que

$$\beta_{\varphi_{\theta_0}}(\theta) \geq \beta_{\varphi^*_{\theta_0}}(\theta) \qquad \forall\, \theta > \theta_0$$

o sea,
$$P_\theta\{\varphi_{\theta_0}(X) = 1\} \geq P_\theta\{\varphi^*_{\theta_0}(X) = 1\} \qquad \forall\, \theta > \theta_0\ .$$
Por lo tanto,

$$P_\theta\{\varphi_{\theta_0}(\mathbf{X}) = 0\} \leq P_\theta\{\varphi^*_{\theta_0}(\mathbf{X}) = 0\} \quad \forall\, \theta > \theta_0\ .$$

pero como $\theta_0 \in S(\mathbf{X})$ si y sólo si $\varphi_{\theta_0}(\mathbf{X}) = 0$ y $\theta_0 \in S^*(\mathbf{X})$ si y sólo si $\varphi^*_{\theta_0}(\mathbf{X}) = 0$, resulta

$$P_\theta\{\theta_0 \in S(\mathbf{X})\} \leq P_\theta\{\theta_0 \in S^*(\mathbf{X})\} \quad \forall\, \theta > \theta_0\ .$$

Un teorema similar puede demostrarse para H_2 contra K_2.

Veamos cómo son las regiones $S(\mathbf{X})$ en el caso del Teorema 1.

Teorema 2. *Sea $\mathbf{X}$ con distribución perteneciente a una familia $F(\mathbf{x}, \theta)$ de cociente de verosimilitud monótono en $T = r(\mathbf{X})$. Supongamos que la función de distribución $F_T(t, \theta)$ de T es continua para todo θ . Sea, para cada $\theta_0 \in \Theta$, $\varphi_{\theta_0}(\mathbf{X})$ el test UMP para $\mathrm{H}_1 : \theta = \theta_0$ contra $\mathrm{K}_1 : \theta > \theta_0$, o sea:*

$$\varphi_{\theta_0}(\mathbf{X}) = \begin{cases} 1 & si \quad T > k_\alpha(\theta_0) \\ 0 & si \quad T \leq k_\alpha(\theta_0) \end{cases}$$

Si además $F_T(t,\theta)$ es continua en θ para cada t fijo, la región de confianza

$$S(\mathbf{X}) = \{\theta \in \Theta : \varphi_\theta(\mathbf{X}) = 0\} = \{\theta \in \Theta : T = r(\mathbf{X}) \leq k_\alpha(\theta)\}$$

es el intervalo $I = [\underline{\theta}(\mathbf{X}), +\infty)$, donde

$$\underline{\theta}(\mathbf{X}) = \inf\{\theta \in \Theta : T \leq k_\alpha(\theta)\} \ .$$

DEMOSTRACIÓN. Ya hemos demostrado que si se tiene una familia de cociente de verosimilitud monótono en $T = r(\mathbf{X})$, el test UMP para $\mathrm{H}_1 : \theta = \theta_0$ contra K_1: $\theta > \theta_0$ es de la forma

$$\varphi_{\theta_0}(\mathbf{X}) = \begin{cases} 1 & \text{si} \quad T > k_\alpha(\theta_0) \\ \gamma_\alpha(\theta_0) & \text{si} \quad T = k_\alpha(\theta_0) \\ 0 & \text{si} \quad T < k_\alpha(\theta_0) \end{cases}$$

con $k_\alpha(\theta_0)$ y $\gamma_\alpha(\theta_0)$ tales que

$$E_{\theta_0}(\varphi_{\theta_0}(\mathbf{X})) = \alpha \ .$$

Como T tiene distribución continua, no es necesario aleatorizar y por lo tanto, el test UMP resulta

$$\varphi_{\theta_0}(\mathbf{X}) = \begin{cases} 1 & \text{si} \quad T > k_\alpha(\theta_0) \\ 0 & \text{si} \quad T \leq k_\alpha(\theta_0) \ . \end{cases}$$

Mostraremos que

(a) $k_\alpha(\theta)$ es una función no decreciente de θ.

(b) $k_\alpha(\theta)$ es una función continua a derecha.

(a) Sabemos que por ser φ_{θ_0} el test UMP de nivel α para H_1 contra K_1, la función de potencia de φ_{θ_0} es mayor o igual que el nivel para todo $\theta > \theta_0$. Luego, dado cualquier $\theta_1 > \theta_0$ se cumple

$$\begin{aligned} \alpha &= E_{\theta_0}(\varphi_{\theta_0}(\mathbf{X})) = P_{\theta_0}(T \geq k_\alpha(\theta_0)) \\ &\leq E_{\theta_1}(\varphi_{\theta_0}(\mathbf{X})) = P_{\theta_1}(T \geq k_\alpha(\theta_0)) \ . \end{aligned}$$

Como además

$$\alpha = E_{\theta_1}(\varphi_{\theta_1}(\mathbf{X})) = P_{\theta_1}(T \geq k_\alpha(\theta_1)) \ ,$$

tendremos

$$\alpha = P_{\theta_1}(T \geq k_\alpha(\theta_1)) \leq P_{\theta_1}(T \geq k_\alpha(\theta_0)) \ ,$$

y por lo tanto, es posible tomar $k_\alpha(\theta_1)$ tal que

$$k_\alpha(\theta_1) \geq k_\alpha(\theta_0) \ .$$

Con lo cual, $k_\alpha(\theta)$ es una función no decreciente de θ.
(b) Sea θ_n una sucesión decreciente que converge a θ, luego como $k_\alpha(.)$ es no decreciente se tiene

$$k_\alpha(\theta_n) \geq k_\alpha(\theta) \tag{6.37}$$

Sea $k = \lim_{n\to\infty} k_\alpha(\theta_n) = \inf_{n\geq 1} k_\alpha(\theta_n)$. Por (6.37) $k \geq k_\alpha(\theta)$, bastará mostrar que $k \leq k_\alpha(\theta)$.

Como $k \leq k_\alpha(\theta_n)$ se cumple

$$P_{\theta_n}(T \leq k) \leq P_{\theta_n}\left(T \leq k_\alpha(\theta_n)\right) = \alpha \ . \tag{6.38}$$

Pero además, como $F_T(k, \theta)$ es continua en θ se tiene

$$P_\theta(T \leq k) = \lim_{n\to\infty} P_{\theta_n}(T \leq k) \ . \tag{6.39}$$

Por lo tanto, (6.38) y (6.39) implican que

$$P_\theta(T \leq k) \leq \alpha = P_\theta\left(T \leq k_\alpha(\theta)\right)$$

luego, es posible tomar $k_\alpha(\theta)$ tal que $k \leq k_\alpha(\theta)$. Con lo cual, $k = k_\alpha(\theta)$ y $k_\alpha(\theta)$ es continua a derecha.

Veamos ahora que $\theta \in S(\mathbf{X})$ si y sólo si $\theta \geq \underline{\theta}(\mathbf{X})$.

Si $\theta \in S(\mathbf{X})$ entonces $T \leq k_\alpha(\theta)$ de donde $\theta \in \{\theta \in \Theta : T \leq k_\alpha(\theta)\}$ y $\theta \geq \underline{\theta}(\mathbf{X})$ que es el ínfimo de este conjunto.

Si $\theta > \underline{\theta}(\mathbf{X})$ entonces existe $\theta' \in \Theta$ tal que $T \leq k_\alpha(\theta')$ con $\underline{\theta}(X) < \theta' \leq \theta$. Pero como $k_\alpha(.)$ es creciente, resulta $T \leq k_\alpha(\theta)$ y por lo tanto, $\theta \in S(\mathbf{X})$.

Si $\theta = \underline{\theta}(\mathbf{X})$, existe una sucesión θ_n decreciente que converge a θ y tal que $\theta_n \in \{\theta \in \Theta : T \leq k_\alpha(\theta)\}$. Por lo tanto, $T \leq k_\alpha(\theta_n)$. Luego, la continuidad a derecha de $k_\alpha(\theta)$ implica que $T \leq k_\alpha(\theta)$ y por lo tanto, $\theta \in S(\mathbf{X})$.

Teorema 3. *Sea* $\mathbf{X} = (X_1, \ldots, X_n)$ *una muestra aleatoria de una distribución perteneciente a una familia* $F(\mathbf{x}, \theta)$ *de cociente de verosimilitud monótono en* $T = r(\mathbf{X})$ *y sea, para cada* $\theta_0 \in \Theta$, $\varphi_{\theta_0}(\mathbf{X})$ *el test UMP para* $\mathrm{H}_1 : \theta = \theta_0$ *contra* $\mathrm{K}_1 : \theta > \theta_0$, *o sea:*

$$\varphi_{\theta_0}(\mathbf{X}) = \begin{cases} 1 & si \quad T > k_\alpha(\theta_0) \\ 0 & si \quad T \leq k_\alpha(\theta_0) \end{cases}$$

suponiendo que la distribución, $F_T(t,\theta)$, de $T(\mathbf{X})$ es continua para todo θ. Supongamos además $F_T(t,\theta)$ es continua en θ para cada t fijo.

En estas condiciones

$$\underline{\theta}(\mathbf{X}) = \inf\{\theta \in \Theta : T \leq k_\alpha(\theta)\}$$

es una cota inferior para θ uniformemente óptima.

Demostración. De acuerdo a la definición de cota inferior con nivel de confianza $1-\alpha$ uniformemente óptima deberá demostrarse que

i) $P_\theta(\theta \geq \underline{\theta}(X)) = 1-\alpha$ para todo θ,

ii) si $\underline{\theta}^*$ es otra cota inferior a nivel α para θ

$$E_\theta(D(\theta,\underline{\theta})) \leq E_\theta(D(\theta,\underline{\theta}^*)) \quad \text{para todo } \theta \tag{6.40}$$

donde D es una medida de la subevaluación de $\underline{\theta}$ respecto de θ, definida por

$$D(\theta,\underline{\theta}) = \begin{cases} \theta - \underline{\theta} & \text{si} \quad \theta > \underline{\theta} \\ 0 & \text{si} \quad \theta \leq \underline{\theta}\ . \end{cases}$$

(i) se deduce del Teorema 1, ya que

$$S(\mathbf{X}) = \{\theta : \theta \geq \underline{\theta}(\mathbf{X})\}$$

es un intervalo de nivel de confianza $1-\alpha$.

(ii) Demostraremos que dada cualquier otra cota $\underline{\theta}^*$ a nivel $1-\alpha$

$$P_\theta\{\theta' \geq \underline{\theta}\} \leq P_\theta\{\theta' \geq \underline{\theta}^*\} \quad \text{para todo } \theta' \leq \theta\ . \tag{6.41}$$

Dado $\theta' \in \Theta$ definamos

$$\varphi^*_{\theta'}(X) = \begin{cases} 1 & \text{si } \theta' \leq \underline{\theta}^* \\ 0 & \text{si } \theta' > \underline{\theta}^*\ . \end{cases}$$

Luego $\varphi^*_{\theta'}(X)$ es un test de nivel α para H : $\theta = \theta'$ contra K : $\theta > \theta'$. Como $\varphi_{\theta'}(X)$ es el UMP para estas hipótesis, por Teorema 1, ii) sabemos que

$$P_\theta\{\theta' \geq \underline{\theta}(X)\} \leq P_\theta\{\theta' \geq \underline{\theta}^*(X)\} \quad \text{para todo } \theta \geq \theta'$$

y como esto se puede hacer para todo $\theta' \in \Theta$ resulta (6.41).

Se podría demostrar que si $\underline{\theta}$ cumple (6.41) entonces $\underline{\theta}$ cumple (6.40). Intuitivamente esto parece razonable, puesto que una cota inferior $\underline{\theta}$ de θ que cumple (6.41) es, en algún sentido, la "mayor" cota inferior y, en este caso, el defecto que presenta $\underline{\theta}$ respecto de θ debería ser lo más pequeño posible. Sin embargo la demostración de esta implicación está fuera de los alcances de este curso. (Para la demostración ver Lehmann [3], ejercicio 21, página 117.)

Ejemplo 1. Sea $X_1, \dots, X_n$ una muestra aleatoria de una distribución $U[0, \theta]$. Sabemos que el test UMP para H : $\theta = \theta_0$ contra K : $\theta > \theta_0$ es de la forma

$$\varphi_{\theta_0}(\mathbf{X}) = \begin{cases} 1 & \text{si} \quad \max\limits_{1 \le i \le n} X_i > \theta_0 \sqrt[n]{1-\alpha} \\ 0 & \text{si} \quad \max\limits_{1 \le i \le n} X_i \le \theta_0 \sqrt[n]{1-\alpha} \end{cases}$$

En este caso, si $T = \max_{1 \le i \le n} X_i$ y $k_\alpha(\theta) = \theta \sqrt[n]{1-\alpha}$

$$S(\mathbf{X}) = \{\theta \in I\!R : \varphi_\theta(X) = 0\} = \{\theta \in I\!R : T \le k_\alpha(\theta)\}$$

resulta igual a

$$\begin{aligned} S(\mathbf{X}) &= \{\theta \in I\!R : \max_{1 \le i \le n} X_i \le \theta \sqrt[n]{1-\alpha}\} = \\ &= \{\theta \in I\!R : \theta \ge \frac{\max\limits_{1 \le i \le n} X_i}{\sqrt[n]{1-\alpha}}\} \end{aligned}$$

y $\underline{\theta}$ será

$$\underline{\theta}(\mathbf{X}) = \frac{\max\limits_{1 \le i \le n} X_i}{\sqrt[n]{1-\alpha}}$$

puesto que este es el menor valor que puede tomar θ que pertenece a $S(\mathbf{X})$.

Resulta entonces que

$$I = [\underline{\theta}(\mathbf{X}), +\infty) = [\frac{\max\limits_{1 \le i \le n} X_i}{\sqrt[n]{1-\alpha}}, +\infty)$$

es un intervalo de confianza unilateral para θ de nivel $1 - \alpha$ y que $\underline{\theta}$ es la mejor cota inferior para θ.

Ejemplo 2. Sea $X_1, \dots, X_n$ una muestra aleatoria de una distribución $N(\mu, \sigma_0^2)$ con σ_0^2 conocido. Sabemos que el test UMP para H : $\mu = \mu_0$ contra

K : $\mu > \mu_0$, es de la forma

$$\varphi_{\mu_0}(\mathbf{X}) = \begin{cases} 1 & \text{si} \quad \sqrt{n}\dfrac{(\overline{X} - \mu_0)}{\sigma_0} > z_\alpha \\ \\ 0 & \text{si} \quad \sqrt{n}\dfrac{(\overline{X} - \mu_0)}{\sigma_0} \leq z_\alpha \end{cases}$$

Procediendo en forma similar a la del Ejemplo 1, resulta

$$S(X) = \{\mu \in I\!R : \mu \geq \overline{X} - z_\alpha \frac{\sigma_0}{\sqrt{n}}\} \ .$$

Luego,

$$\underline{\mu}(\mathbf{X}) = \overline{X} - z_\alpha \frac{\sigma_0}{\sqrt{n}}$$

es la mejor cota inferior para μ y

$$I = [\underline{\mu}(X), +\infty) = [\overline{X} - z_\alpha \frac{\sigma_0}{\sqrt{n}}, +\infty)$$

es un intervalo unilateral de nivel $1 - \alpha$ para μ.

6.10 Relación entre intervalos de confianza con nivel asintótico $1-\alpha$ y test con nivel de significación asintótico α

Supongamos que $X_1, \ldots, X_n$ es una muestra aleatoria de una distribución perteneciente a la familia $F(x, \boldsymbol{\theta})$ y que para cada $\boldsymbol{\theta}_0$ se tenga una sucesión de test $\varphi_{n\boldsymbol{\theta}_0}(X_1, \ldots, X_n)$ con nivel de significación asintótico $1 - \alpha$ para H : $\boldsymbol{\theta} = \boldsymbol{\theta}_0$ contra K : $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$. Luego, puede construirse una sucesión de intervalos de confianza con nivel asintótico $1 - \alpha$ definiendo

$$S_n(X_1, \ldots, X_n) = \{\boldsymbol{\theta} : \varphi_{n\boldsymbol{\theta}}(\mathbf{X}) = 0\} \ .$$

Recíprocamente, dada una sucesión de intervalos de confianza $S_n(X_1, \ldots, X_n)$ de nivel asintótico $1 - \alpha$, si definimos

$$\varphi_{n\boldsymbol{\theta}_0}(\mathbf{X}) = \begin{cases} 1 & \text{si} \quad \boldsymbol{\theta}_0 \notin S(X_1, \ldots, X_n) \\ 0 & \text{si} \quad \boldsymbol{\theta}_0 \in S(X_1, \ldots, X_n) \end{cases}$$

se tiene que $\varphi_{n\boldsymbol{\theta}_0}$ es una sucesión de test con nivel de significación asintótico α para H : $\boldsymbol{\theta} = \boldsymbol{\theta}_0$ contra K : $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$. (Se deja como ejercicio la demostración de estos enunciados.)

Ejemplo 1. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución $Bi(\theta, 1)$. Ya se ha visto que

$$\sqrt{n}\frac{(\overline{X} - \theta_0)}{\sqrt{\theta_0(1-\theta_0)}}$$

converge en distribución a la $N(0,1)$ cuando $\theta = \theta_0$.

Un intervalo de confianza para θ, con nivel asintótico $1-\alpha$ viene dado por

$$S_n(\mathbf{X}) = \{\theta : \sqrt{n}\frac{|\overline{X} - \theta|}{\sqrt{\theta(1-\theta)}} < z_{\frac{\alpha}{2}}\}$$

Luego, un test de significación asintótico α para H : $\theta = \theta_0$ contra K : $\theta \neq \theta_0$, viene dado por

$$\varphi(\mathbf{X}) = \begin{cases} 1 & \text{si} \quad \sqrt{n}\dfrac{|\overline{X} - \theta_0|}{\sqrt{\theta_0(1-\theta_0)}} \geq z_{\frac{\alpha}{2}} \\ \\ 0 & \text{si} \quad \sqrt{n}\dfrac{|\overline{X} - \theta_0|}{\sqrt{\theta_0(1-\theta_0)}} < z_{\frac{\alpha}{2}} \ . \end{cases}$$

REFERENCIAS

1. Chernoff, H. (1954). On the distribution of the likelihood ratio. Ann. Math. Statist. 25: 573-578.

2. Ferguson, T.S. (1967). Mathematical Statistics. A Decision Theoretic Approach. Academic Press.

3. Lehmann, E.L. (1994). Testing Statistical Hypothesis. Chapman and Hall.

4. Owen, D.B. (1965). The power of Student's t test. J. Amer. Statist. Assoc. 60: 320-333.

5. Wald, A. (1943). Tests of statistical hypothesis concerning several parameters when the number of observations is large. Trans. Am. Math. Soc. 54: 426-483.