Test Intersección Unión

STELLA MARIS DONATO
RICARDO ANÍBAL LEIVA

RESUMEN

En la actualidad el número de trabajos que emplean el principio Intersección Unión que propone Roy[7] para generar tests de hipótesis se ha incrementado notablemente, pues por sus características esta metodología permite resolver problemas complejos en los que otros procedimientos fracasan.

Este crecimiento vertiginoso ha provocado la aparición de dos efectos no deseados. El primero es debido a que, en ocasiones, el uso de la metodología se produce aprovechando algún aspecto del principio de Intersección-Unión en detrimento de otros, esto va provocando cambios sutiles en la formulación del principio de acuerdo al problema que intenta resolver. El segundo efecto no deseado, que está vinculado con el anterior, es que cuando se contempla el cúmulo de aplicaciones diversas del principio de Intersección Unión se pone en evidencia la carencia de una presentación que destaque las distintas facetas del principio y que permita, de esta manera, abarcar las diferentes aplicaciones.

En este trabajo se elabora una presentación formal del principio Intersección Unión, que apunta a destacar las diferentes facetas que se utilizan en la formulación de tests. Esto permite presentar en forma unificada las numerosas y diversas aplicaciones que se han desarrollado utilizando este principio.

Palabras clave

análisis multivariado / test de hipótesis / principio intersección unión

STELLA MARIS DONATO
Instituto de Ciencias Básicas de la
Universidad Nacional de Cuyo

DEODORO MAGNO BRIGHENTI
Fac. de Ciencias Económicas e
Instituto de Ciencias Básicas
Universidad Nacional de Cuyo y
CONICET

1. INTRODUCCIÓN

Los modelos estadísticos, muchas veces, surgen de investigaciones interdisciplinarias relacionadas con problemáticas de la vida real, y muchas de estas aplicaciones necesitan de un modelo especialmente desarrollado para ellas, ya que las herramientas de inferencia estadística convencional o estándar rara vez se pueden adaptar. Estos estudios, generalmente, están relacionados con diseños complejos, planes de muestreo específicos y leyes de probabilidad subyacentes, que no sólo envuelven parámetros múltiples, sino que también están sujetas a varias restricciones no lineales. Además, muchas veces, la estructura y dimensión de los datos puede complicar considerablemente el análisis estadístico.

En una amplia variedad de problemas multivariados, las herramientas de inferencia estadística clásica, como son los estimadores de máxima verosimilitud y los tests de cociente de verosimilitudes, raras veces se pueden obtener en forma sencilla. Una alternativa muy utilizada, en algunos casos con ventajas computacionales, es el principio de Intersección Unión que presenta Roy en 1953.

Roy[7] propone un método heurístico de construcción de test de hipótesis que se puede describir de la siguiente manera:

Dada una muestra de tamaño n de vectores aleatorios $p-$variados $\mathbf{X}_1,\ldots,\mathbf{X}_n$ se la identifica con el vector aleatorio p $-$variado $\mathbf{X}$, con función de distribución acumulada $F_{\mathbf{X}}$, de cuya función de distribución sólo se sabe que pertenece a una cierta familia $\mathbf{F}$ de distribuciones, o sea, $F_{\mathbf{X}} \in \mathbf{F}$; y se quiere indagar si $F_{\mathbf{X}}$ pertenece a cierto subconjunto no vacío $\mathbf{F}_0$ de $\mathbf{F}$.

Sea $\{\mathbf{F}_j \ : \ j \in J\}$ una colección de subconjuntos de la familia $\mathbf{F}$ de distribuciones, siendo J un conjunto arbitrario de índices.

Para cada $j \in J$ se define la hipótesis nula componente $\mathrm{H}_{0,j}$ como $\mathrm{H}_{0,j} \ : \ F_{\mathbf{X}} \in \mathbf{F}_j$ y la hipótesis alternativa componente $\mathrm{H}_{1,j}$ como $\mathrm{H}_{1,j} \ : \ F_{\mathbf{X}} \in \mathbf{F}'_j = \mathbf{F} - \mathbf{F}_j$.

Se supone, además, que se pueden construir tests adecuados asociados a cada hipótesis componente $\mathrm{H}_{0,j}$. El test Intersección Unión, que se obtiene de acuerdo con el método formulado por Roy, se construye a través de los tests componentes definiendo:

$$\mathrm{H}_0 = \bigcap_{j \in J} \mathrm{H}_{0,j} \qquad \mathrm{H}_1 = \bigcup_{j \in J} \mathrm{H}_{1,j}.$$

De donde se deduce que la hipótesis nula H_0 es verdadera si y sólo si todas las hipótesis nulas componentes $\mathrm{H}_{0,j}$ son verdaderas. Similarmente, la hipótesis alternativa H_1 es verdadera si y sólo si al menos una de las hipótesis alternativas componentes $\mathrm{H}_{1,j}$ es verdadera. El principio Intersección Unión consiste entonces en:

Si $\mathbf{F}_0 \subset \bigcap_{j \in J} \mathbf{F}_j$ entonces se rechaza que $F_{\mathbf{X}} \in \mathbf{F}_0$ si por lo menos una de las $\mathbf{H}_{0,j}$ se rechaza.

Si $\mathbf{F}_0 \supset \bigcap_{j \in J} \mathbf{F}_j$ entonces no se rechaza que $F_{\mathbf{X}} \in \mathbf{F}_0$ si no se rechazan todas las $\mathbf{H}_{0,j}$.

Lo usual en las aplicaciones de este test es que las hipótesis componentes se seleccionen de forma tal que se verifique que $\mathbf{F}_0 = \bigcap_{j \in J} \mathbf{F}_j$.

El método de construcción de los test Intersección Unión, comparado con el clásico test de cociente de verosimilitudes, tiene dos consecuencias prácticas importantes.

La primera es que si la hipótesis nula es rechazada, se puede examinar cuál o cuáles de las hipótesis componentes son las responsables y así tener una idea más clara sobre la causa que llevó a rechazar la hipótesis nula. Desafortunadamente, el test de cociente de verosimilitudes no tiene esta propiedad, si se rechaza la hipótesis nula usando un test de cociente de verosimilitudes no se puede, usando la información brindada por el test, indagar sobre los detalles de las razones de esta decisión.

La segunda es que, en algunos casos, permite obtener de manera más sencilla tests que se presentan en situaciones complejas en contextos multivariados, múltiparámetricos y/o de restricciones de diversos tipos.

En este trabajo se elabora una presentación formal del principio Intersección Unión que propone Roy para el desarrollo de tests de hipótesis que permite abarcar sus diferentes aplicaciones. Se ejemplifica esta presentación formal seleccionando de las referencias algunos casos concretos de aplicación del test Intersección Unión en los que se observan los diferentes aspectos de este principio.

2. CONCEPTOS BÁSICOS

Los conceptos básicos que se definen a continuación son necesarios para la posterior definición y demostración de propiedades de los tests Intersección Unión.

• *Problema estadístico:* Dado el vector aleatorio p −variado $\mathbf{X}$, con función de distribución acumulada $F_{\mathbf{X}}$ (proveniente de identificar a una muestra de tamaño n de vectores aleatorios p− variados $\mathbf{X}_1,\ldots,\mathbf{X}_n$ con el vector $\mathbf{X}$), el problema estadístico consiste en que sólo se sabe que la función de distribución $F_{\mathbf{X}}$ pertenece a una cierta familia $\mathbf{F}$ de distribuciones, o sea, $F_{\mathbf{X}} \in \mathbf{F}$. El problema estadístico es paramétrico cuando la familia de distribuciones depende de un parámetro $q \in \Theta$, o sea, $F_{\mathbf{X}} \in \mathbf{F} = \{F(\cdot,q) : q \in \Theta\}$ La notación que se emplea para destacar la dependencia de la distribución $F_{\mathbf{X}}$ del parámetro q es escribiendo $F_{\mathbf{X}} = F_{\mathbf{X}}(\cdot,q)$

El conjunto de valores que toma el vector aleatorio $\mathbf{X}$, o la muestra $\mathbf{X}_1,\ldots,\mathbf{X}_n$, se denota con el símbolo X, donde se cumple que $\mathrm{X} \subseteq \Re^{\mathcal{T}}$. Usualmente, X se denomina espacio muestral o conjunto potencial de datos de la muestra.

• *Espacio Estadístico:* Sea $\mathtt{P}$ una familia de funciones probabilidad en un espacio medible $(\Omega, \mathtt{A})$, entonces se dice que $(\Omega, \mathtt{A}, \mathtt{P})$ es un espacio estadístico (para consultar estos conceptos ver Burrill[1] y Laha et al.[4]).

• *Hipótesis estadística:* Sean $(\Omega, \mathtt{A}, \mathtt{P})$ un espacio estadístico, $\mathbf{X} : \Omega \to \mathrm{X}$ una muestra y $\mathtt{F}$ la familia de distribuciones a la que pertenece la distribución $F_{\mathbf{X}}$ de la muestra.

Una hipótesis estadística para $\mathtt{F}$ se define como un subconjunto no vacío $\mathtt{F}'$ de $\mathtt{F}$.

Si la familia $\mathtt{F}$ es paramétrica, o sea, $\mathtt{F} = \{F(\cdot, q) : q \in \Theta\}$, y, en consecuencia, $\mathtt{P}$ puede ser identificada con $\mathtt{P} = \{P_q : q \in \Theta\}$, entonces una hipótesis estadística $\mathtt{F}' \subset \mathtt{F}$ se expresa como $\mathtt{F}' = \{\mathrm{F}(\cdot, \theta) : \theta \in \Theta'\}$ con $\varnothing \neq \Theta' \subset \Theta$. En este caso, también se dice simplemente que Θ' es una hipótesis estadística para Θ.

Es habitual designar a una hipótesis estadística con la letra $\mathtt{H}$, en lugar de usar $\mathtt{F}'$.

• *Test estadístico determinístico:* Sean $(\Omega, \mathtt{A}, \mathtt{P})$ un espacio estadístico y $\mathbf{X} : \Omega \to \mathrm{X}$ una muestra. Un test estadístico determinístico es un subconjunto C del conjunto X de valores posibles de la muestra.

• *Función de potencia de un test para familias de distribuciones paramétricas:* Sean $(\Omega, \mathtt{A}, \{P_q : q \in \Theta\})$ un espacio estadístico, $\mathbf{X} : \Omega \to \mathrm{X}$ una muestra, $\mathtt{F} = \{F(\cdot, q) : q \in \Theta\}$ la familia de distribuciones a la que pertenece la distribución F_X de la muestra y C un test estadístico determinístico. La función $\mathrm{b}_C : \Theta \to [0;1]$ que para cada $q \in \Theta$ está dada por $\beta_C(\theta) = P_\theta[\mathbf{X} \in C] = P_\theta[\{w : w \in \Omega \ / \ \mathbf{X}(\mathbf{w}) \in C\}]$ se denomina la función de potencia del test C.

• *Tests de una hipótesis versus otra hipótesis:* Sean $(\Omega, \mathtt{A}, \mathtt{P})$ un espacio estadístico, $\mathbf{X} : \Omega \to \mathrm{X}$ una muestra y $\mathtt{F}$ la familia de distribuciones a la que pertenece la distribución F_X de la muestra. Sean dos hipótesis estadísticas disjuntas $\mathtt{H}_0$ y $\mathtt{H}_1$ (sin pérdida de generalidad se supone que $\mathtt{H}_1 = \mathtt{F} - \mathtt{H}_0$) Se denomina test de $\mathtt{H}_0$ versus $\mathtt{H}_1$ a un test C que tiene asociada la siguiente estrategia: para cada observación $\mathtt{x} \in \mathrm{X}$ se rechaza la hipótesis $\mathtt{H}_0$ si y sólo si $\mathtt{x} \in C$. En este caso se dice que $\mathtt{H}_0$ es la hipótesis nula y que $\mathtt{H}_1$ es la hipótesis alternativa.

En el caso en que $(\Omega, \mathrm{A}, \{P_{\mathrm{q}} : \mathrm{q} \in \Theta\})$ es un espacio estadístico paramétrico, o sea, $\mathrm{F} = \{F(\cdot, \mathrm{q}) : \mathrm{q} \in \Theta\}$, las dos hipótesis estadísticas disjuntas H_0 y H_1 se identifican con dos subconjuntos disjuntos no vacíos Θ_0 y Θ_1 de Θ tales que $\Theta = \Theta_0 \cup \Theta_1$. Se dice entonces que Θ_0 es la hipótesis nula y que Θ_1 es la hipótesis alternativa. La notación que se emplea es: C es un test de $\mathrm{H}_0 : \mathrm{q} \in \Theta_0$ versus $\mathrm{H}_1 : \mathrm{q} \in \Theta_1 = \Theta - \Theta_0$.

Nota: Como siempre se supone que $\mathrm{H}_1 = \mathrm{F} - \mathrm{H}_0$, se puede simplemente decir que C es un test para H_0, en lugar de C es un test de H_0 versus H_1.

• *Nivel de significación:* Sea C un test de $\mathrm{H}_0 : \mathrm{q} \in \Theta_0$. Se denomina nivel de significación de C al número a_C dado por

$$\sup_{j \in J} \alpha_{j,J} \leq \alpha_{UI} \leq \min\left\{1; \sum_{j \in J} \alpha_j\right\},$$

3. FORMALIZACIÓN DEL TEST INTERSECCIÓN UNIÓN

3.1. TESTS NECESARIOS, TESTS SUFICIENTES Y TESTS EQUIVALENTES

Para establecer propiedades entre las funciones de potencia y entre los niveles de significación de diferentes tests, surge la necesidad de definir relaciones entre los tests vinculadas con la función de potencia de los mismos.

• *Definición:* Sean $(\Omega, \mathrm{A}, \{P_{\mathrm{q}} : \mathrm{q} \in \Theta\})$ un espacio estadístico, $\mathbf{X} : \Omega \to \mathrm{X}$ una muestra y $\mathrm{F} = \{F(\cdot, \mathrm{q}) : \mathrm{q} \in \Theta\}$ la familia de distribuciones a la que pertenece la distribución F_X de la muestra. Sean C y C_j tests estadísticos determinísticos con funciones de potencia b_C y b_{C_j}, respectivamente. Se dice que C es suficiente para C_j (o también que C_j es necesario para C) si $\mathrm{b}_C \leq \mathrm{b}_{C_j}$. Si C es a la vez necesario y suficiente para C_j entonces C es equivalente a C_j y en ese caso se verifica que $\mathrm{b}_C = \mathrm{b}_{C_j}$.

Nota: La notación $\mathrm{b}_C \leq \mathrm{b}_{C_j}$ indica que $\forall \mathrm{q} \in \Theta$ se verifica que $\mathrm{b}_C(\mathrm{q}) \leq \mathrm{b}_{C_j}(\mathrm{q})$ De manera análoga se debe interpretar la notación $\mathrm{b}_C = \mathrm{b}_{C_j}$.

• *Propiedad 1:* Sean los tests C con función de potencia b_C y C_j con función de potencia b_{C_j}, tales que C es suficiente para C_j. Para todo Θ_0 y $\Theta_{0,j}$ tales que $\varnothing \neq \Theta_0 \subseteq \Theta_{0,j} \subset \Theta$, si C se considera como test de Θ_0 y si C_j se considera como test de $\Theta_{0,j}$, siendo a y a_j sus respectivos niveles de significación se tiene que $\mathrm{a} \leq \mathrm{a}_j$.

Demostración: Si C es suficiente para C_j, se verifica que $\mathrm{b}_C \leq \mathrm{b}_{C_j}$. Para todo Θ_0 tal que $\varnothing \neq \Theta_0 \subset \Theta$, si C y C_j se consideran ambos como tests de Θ_0, entonces

$$\sup_{\theta\in\Theta_0}\beta_C(\theta)\leq\sup_{\theta\in\Theta_0}\beta_{C_j}(\theta).$$

Por otro lado, siendo C_j un test con función de potencia β_{C_j}. Si Θ_0 y $\Theta_{0,j}$ son dos subconjuntos de Θ tales que $\varnothing\neq\Theta_0\subseteq\Theta_{0,j}\subset\Theta$, entonces

$$\sup_{\theta\in\Theta_0}\beta_{C_j}(\theta)\leq\sup_{\theta\in\Theta_{0,j}}\beta_{C_j}(\theta).$$

En consecuencia,

$$\alpha=\sup_{\theta\in\Theta_0}\beta_C(\theta)\leq\sup_{\theta\in\Theta_0}\beta_{C_j}(\theta)\leq\sup_{\theta\in\Theta_{0,j}}\beta_{C_j}(\theta)=\alpha_j.$$

3.2. EL TEST UNIÓN Y LA HIPÓTESIS INTERSECCIÓN

• *Definición de test unión:* Sea un espacio estadístico $(\Omega,\mathcal{A},\{P_\theta : \theta\in\Theta\})$. Sea $\{C_j : j\in J\}$ una familia arbitraria de tests, se denomina test unión de los tests en $\{C_j : j\in J\}$ al test $C_U=\bigcup_{j\in J}C_j$.

Nota: Los tests C_j se denominan tests componentes del test unión C_U.

• *Definición de hipótesis intersección:* Sea un espacio estadístico $(\Omega,\mathcal{A},\{P_\theta : \theta\in\Theta\})$ Sea $\{\Theta_{0,j} : j\in J\}$ una familia arbitraria de hipótesis para Θ, o sea, $\Theta_{0,j}\subseteq\Theta$ para todo $j\in J$. Se denomina hipótesis intersección de la familia de hipótesis $\{\Theta_{0,j} : j\in J\}$ a $\Theta_0=\bigcap_{j\in J}\Theta_{0,j}$.

Es claro que por ser $\Theta_0=\bigcap_{j\in J}\Theta_{0,j}\subseteq\Theta$ la hipótesis intersección resulta ser una hipótesis para Θ.

La propiedad que sigue, establece cotas para el nivel de significación de los tests unión considerados como test de la hipótesis intersección, cuando la familia de tests C_j es un conjunto numerable.

• *Propiedad 2:* Sea un espacio estadístico $(\Omega,\mathcal{A},\{P_\theta : \theta\in\Theta\})$ Sea $\{C_j : j\in J\}$ una familia arbitraria de tests, con J conjunto numerable, siendo β_{C_j} la función de potencia de C_j, para cada $j\in J$, y donde C_j se considera como un test de $\Theta_{0,j}$, con $\varnothing\neq\Theta_{0,j}\subset\Theta$, con nivel de significación α_j. Sea $C_U=\bigcup_{j\in J}C_j$ con función de potencia β_{C_U} y sea $\Theta_{0,I}=\bigcap_{j\in J}\Theta_{0,j}\neq\varnothing$, entonces si C_U se considera como un test de $\Theta_{0,I}$ tiene nivel de significación α_{UI} tal que

$$\sup_{j\in J}\alpha_{j,I}\leq\alpha_{UI}\leq\min\left\{1;\sum_{j\in J}\alpha_j\right\},$$

donde, para cada $j \in J$, con a $_{j,I}$ se indica el nivel de significación de C_j considerado como un test de $\Theta_{0,J} = \bigcap_{j \in J} \Theta_{0,j}$.

Demostración: Para obtener la desigualdad de la izquierda, se observa primero que

$$C_U = \bigcup_{j \in J} C_j \subseteq \mathrm{X}$$

luego C_U es un test. Dado que cualquiera sea $j \in J,\ C_j \subseteq C_U$, para cualquier $\mathrm{q} \in \Theta$ se verifica

$$\mathrm{b}_{C_U}(\mathrm{q}) = P_{\mathrm{q}}[\mathbf{X} \in C_U] = P_{\mathrm{q}}\left[\mathbf{X} \in \bigcup_{j \in J} C_j\right] \geq P_{\mathrm{q}}[\mathbf{X} \in C_j] = \mathrm{b}_{C_j}(\mathrm{q}) \qquad \forall j \in J.$$

Luego $\mathrm{b}_{C_U}(\mathrm{q}) \geq \mathrm{b}_{C_j}(\mathrm{q})$, de donde se concluye que

$$\mathrm{b}_{C_U}(\mathrm{q}) \geq \sup_{j \in J} \mathrm{b}_{C_j}(\mathrm{q})$$

En consecuencia,

$$\mathrm{a}_{U} = \sup_{\mathrm{q} \in \Theta_{0,J}} \mathrm{b}_{C_U}(\mathrm{q}) \geq \sup_{\mathrm{q} \in \Theta_{0,J}} \sup_{j \in J} \mathrm{b}_{C_j}(\mathrm{q}) = \sup_{j \in J} \sup_{\mathrm{q} \in \Theta_{0,J}} \mathrm{b}_{C_j}(\mathrm{q}) = \sup_{j \in J} \mathrm{a}_{j,J}.$$

Para la desigualdad de la derecha se requiere que J sea un conjunto numerable, pues la desigualdad de Bonferroni lo exige (además en caso contrario $\sum_{j \in J} \mathrm{a}_j = \infty$)

$$\mathrm{a}_{U} = \sup_{\mathrm{q} \in \Theta_{0,J}} \mathrm{b}_{C_U}(\mathrm{q}) = \sup_{\mathrm{q} \in \Theta_{0,J}} P_{\mathrm{q}}\left[\mathbf{X} \in \bigcup_{j \in J} C_j\right] \underset{(*)}{\leq} \sup_{\mathrm{q} \in \Theta_{0,J}} \sum_{j \in J} P_{\mathrm{q}}[\mathbf{X} \in C_j]$$

$$\leq \sum_{j \in J} \sup_{\mathrm{q} \in \Theta_{0,J}} P_{\mathrm{q}}[\mathbf{X} \in C_j] \underset{(*)}{\leq} \sum_{j \in J} \sup_{\mathrm{q} \in \Theta_{0,j}} P_{\mathrm{q}}[\mathbf{X} \in C_j] = \sum_{j \in J} \mathrm{a}_j.$$

Por la desigualdad de Bonferroni (ver Wasserman[12])

Por la propiedad 1

La conclusión se sigue de esta desigualdad y de recordar que $\alpha_{UI} \leq 1$ por ser una probabilidad.

Interesa establecer propiedades similares a la anterior, que relacionen el nivel de significación del test unión con los niveles de significación de los tests C_j considerados, cada uno, como test de la hipótesis $\Theta_{0,j}$ correspondiente. Por este motivo, es necesario definir hipótesis determinante.

3.3. HIPÓTESIS DETERMINANTE O HIPÓTESIS SUPREMAL

Se definen dos tipos de hipótesis determinante, hipótesis determinante de otra hipótesis e hipótesis determinante de una familia de hipótesis.

• *Hipótesis determinante de otra hipótesis:* Sean Θ_0 y Θ_0^* hipótesis estadísticas para Θ. Sea C un test de Θ_0 con nivel de significación α y el mismo C considerado como un test de Θ_0^* tiene nivel de significación α^*. Sea además $\Theta_0^* \subseteq \Theta_0$. Se dice que Θ_0^* es determinante de Θ_0 para el test C o también que Θ_0^* es supremal de Θ_0 para el test C si se verifica que $\alpha^* = \alpha$.

Nota: De acuerdo con la propiedad 1, se cumple que $\alpha^* \leq \alpha$, por lo que la hipótesis determinante constituye un caso particular de lo analizado en la propiedad 1.

• *Hipótesis determinante de una familia de hipótesis:* Sean $\{\Theta_{0,j} \ : \ j \in J\}$ y Θ_0^* hipótesis estadísticas para Θ. Sea $\{C_j \ : \ j \in J\}$ una familia arbitraria de tests donde, para cada $j \in J$, es C_j un test de $\Theta_{0,j}$, con nivel de significación α_j y el mismo C_j considerado como un test de Θ_0^* tiene nivel de significación α_j^*, siendo $\Theta_0^* \subseteq \Theta_{0,j}$. Se dice que Θ_0^* es determinante de $\{\Theta_{0,j} \ : \ j \in J\}$, o también que Θ_0^* es supremal de $\{\Theta_{0,j} \ : \ j \in J\}$, si se verifica que Θ_0^* es determinante de $\Theta_{0,j}$ para el test C_j, para cada $j \in J$. Es decir, para cada $j \in J$ se verifica que $\alpha_j^* = \alpha_j$.

En la siguiente propiedad, se vincula la hipótesis determinante de un conjunto de hipótesis con la hipótesis intersección que se obtiene del mismo conjunto.

• *Propiedad 3:* Sean $\{\Theta_{0,j} \ : \ j \in J\}$ Θ_0^*, $\{C_j \ : \ j \in J\}$ α_j y α_j^* como en la definición anterior. Sea la hipótesis intersección $\Theta_{0,J} = \bigcap_{j \in J} \Theta_{0,j}$ y C_j considerado como un test de $\Theta_{0,J}$ tiene nivel de significación $\alpha_{j,J}$.

a. Si Θ_0^* es determinante de $\{\Theta_{0,j} \ : \ j \in J\}$ entonces Θ_0^* es determinante de $\Theta_{0,J}$ para el test C_j, para cada $j \in J$. Es decir, para cada $j \in J$ se verifica que $\alpha_j^* = \alpha_{j,J}$.

b. Si Θ_0^* es determinante de $\{\Theta_{0,j} \ : \ j \in J\}$ entonces $\Theta_{0,J}$ es determinante de $\{\Theta_{0,j} \ : \ j \in J\}$ Es decir, para cada $j \in J$ se verifica que $\alpha_j = \alpha_{j,J}$.

Demostración:

a. Claramente $\Theta_0^* \subseteq \Theta_{0,J}$, ya que $\Theta_0^* \subseteq \Theta_{0,j}$ y $\Theta_{0,J} = \bigcap_{j \in J} \Theta_{0,j}$, para cada $j \in J$. Por hipótesis Θ_0^* es determinante de $\{\Theta_{0,j} \ : \ j \in J\}$, es decir que, para cada $j \in J$ se verifica que $\alpha_j^* = \alpha_j$.

$$\alpha_j^* = \sup_{\theta \in \Theta_0^*} P_\theta [X \in C_j] \underset{(*)}{\leq} \sup_{\theta \in \Theta_{0,J}} P_\theta [X \in C_j] = \alpha_{j,J}.$$

Por otra parte

$$\alpha_{j,J} = \sup_{\theta \in \Theta_{0,J}} P_\theta [X \in C_j] \underset{(*)}{\leq} \sup_{\theta \in \Theta_{0,j}} P_\theta [X \in C_j] = \alpha_j \underset{(**)}{=} \alpha_j^*.$$

Por la propiedad 1, ya que $\Theta_0^* \subseteq \Theta_{0,I}$.

Por la propiedad 1, ya que $\Theta_{0,I} = \bigcap_{j\in J} \Theta_{0,j} \subseteq \Theta_{0,j}$, para cada $j \in J$.

Por hipótesis, $\alpha_j^* = \alpha_j$.

Lo que implica que:

$$\alpha_j^* = \alpha_{j,I},$$

ya que $\alpha_j^* \leq \alpha_{j,I}$ y $\alpha_j^* \geq \alpha_{j,I}$.

(b) Por hipótesis se cumple que $\alpha_j^* = \alpha_j$, ya que si Θ_0^* es determinante de y de $\{\Theta_{0,j} \ : \ j \in J\}$, la parte (a) se tiene que $\alpha_{j,I} = \alpha_j^*$, lo que implica que $\alpha_j = \alpha_{j,I}$.

En conclusión, esta propiedad afirma que, si una hipótesis Θ_0^* es determinante de un conjunto de hipótesis $\{\Theta_{0,j} \ : \ j \in J\}$ entonces la hipótesis intersección de este conjunto también es determinante del mismo.

A continuación se establece una propiedad que acota el nivel de significación del test unión con los niveles de significación de los tests C_j considerados, cada uno, como test de la hipótesis $\Theta_{0,j}$ correspondiente.

- *Propiedad 4:* Sea $\{\Theta_{0,j} \ : \ j \in J\}$ una familia de hipótesis estadísticas para Θ, donde J es un conjunto numerable. Sea $\{C_j \ : \ j \in J\}$ una familia de tests disjuntos dos a dos, donde, para cada $j \in J$, es C_j un test de $\Theta_{0,j}$, con nivel de significación α_j. Sea $C = C_U = \bigcup_{j\in J} C_j$ el test unión (disjunta) de $\Theta_{0,I}$, con $\varnothing \neq \Theta_{0,I} = \bigcap_{j\in J} \Theta_{0,j}$, con nivel de significación α_U. Si $\Theta_{0,I}$ es determinante para $\{\Theta_{0,j} \ : \ j \in J\}$ entonces se verifica que $\sup_{j\in J} \alpha_j \leq \alpha_U \leq \min\left\{1; \sum_{j\in J} \alpha_j\right\}$.

Demostración: Por la propiedad 2 si se considera a C_j como un test de $\Theta_{0,I}$ con nivel de significación $\alpha_{j,I}$ entonces se verifica que $\sup_{j\in J} \alpha_{j,I} \leq \alpha_U \leq \min\left\{1; \sum_{j\in J} \alpha_j\right\}$.

Si en la propiedad 3 se elige $\Theta_0^* = \Theta_{0,I}$, entonces para cada $j \in J$ se verifica que $\alpha_j = \alpha_{j,I}$, por lo que la desigualdad de la izquierda resulta $\sup_{j\in J} \alpha_j \leq \alpha_U$.

Por otra parte, usando la condición de ser $\{C_j \ : \ j \in J\}$ una familia de tests disjuntos dos a dos resulta

$$\alpha_{UI} = \sup_{\theta\in\Theta_{0,I}} \beta_C(\theta) = \sup_{\theta\in\Theta_{0,I}} P_\theta\left[\mathbf{X} \in \bigcup_{j\in J} C_j\right] = \sup_{\theta\in\Theta_{0,I}} \sum_{j\in J} P_\theta[\mathbf{X} \in C_j]$$

$$\leq \sum_{j\in J} \sup_{\theta\in\Theta_{0,I}} P_\theta[\mathbf{X} \in C_j] = \sum_{j\in J} \alpha_{j,I} = \sum_{j\in J} \alpha_j.$$

De donde se obtiene la desigualdad de la derecha.

Nota: Se observa que en el caso que $\Theta_{0,J} = \bigcap_{j\in J} \Theta_{0,j}$, sea un conjunto unitario, entonces se verifica que

$$\sup_{j\in J} \alpha_j \leq \alpha_{UI} = \min\left\{1; \sum_{j\in J} \alpha_j\right\}.$$

3.4. TESTS INTERSECCIÓN UNIÓN

Sean $(\Omega, \mathbf{A}, \mathbf{P})$ un espacio estadístico, $\mathbf{X} : \Omega \to \mathrm{X}$ una muestra y $\mathbf{F}$ la familia de distribuciones a la que pertenece la distribución $F_{\mathbf{X}}$ de la muestra.

Sea $\{C_j : j \in J\}$ una familia arbitraria de tests, donde C_j se considera como un test de $\mathbf{H}_{0,j} : F_{\mathbf{X}} \in \mathbf{F}_j$, con $\varnothing \neq \mathbf{F}_j \subset \mathbf{F}$, para cada $j \in J$ siendo J algún conjunto de índices. Se obtiene así, una familia arbitraria de hipótesis $\{\mathbf{F}_j : j \in J\}$ para todo $j \in J$.

• *Definición de test Intersección - Unión:* Sea C un test de $\mathbf{H}_0 : F_{\mathbf{X}} \in \mathbf{F}_0$ y sea $\{C_j : j \in J\}$ una familia arbitraria de tests correspondiente a la familia de hipótesis $\{\mathbf{F}_j : j \in J\}$ El test C es un test Intersección-Unión para la hipótesis $\mathbf{H}_0$ si $C = \bigcup_{j\in J} C_j$ y $\mathbf{F}_0 = \bigcap_{j\in J} \mathbf{F}_j$.

En el caso particular que la familia de distribuciones sea paramétrica, se puede definir al test Intersección Unión de manera análoga, de la siguiente forma:

Sea $(\Omega, \mathbf{A}, \{P_\theta : \theta \in \Theta\})$ un espacio estadístico, $\mathbf{X} : \Omega \to \mathrm{X}$ una muestra y $\mathbf{F} = \{F(\cdot, \theta) : \theta \in \Theta\}$ la familia de distribuciones a la que pertenece la distribución $F_{\mathbf{X}}$ de la muestra.

Sea $\{C_j : j \in J\}$ una familia arbitraria de tests, donde C_j se considera como un test de $\mathbf{H}_{0,j} : \theta \in \Theta_{0,j}$, con $\varnothing \neq \Theta_{0,j} \subset \Theta$; para cada $j \in J$ siendo J algún conjunto de índices. Se obtiene así, una familia arbitraria de hipótesis $\{\Theta_{0,j} : j \in J\}$ para Θ.

Sea C un test de $\mathbf{H}_0 : \theta \in \Theta_0$. El test C es un test Intersección-Unión para la hipótesis $\mathbf{H}_0$ si $C = C_U = \bigcup_{j\in J} C_j$ y $\Theta_0 = \bigcap_{j\in J} \Theta_{0,j}$.

La hipótesis nula $\mathbf{H}_0$ y el test C se denominan Hipótesis nula inicial y Test inicial respectivamente. Las hipótesis $\mathbf{H}_{0,j}$ y los test C_j, para cada j, se denominan Hipótesis nula componente y Test componente respectivamente.

Del análisis de cada una de las inclusiones que componen la igualdad $\mathbf{F}_0 = \bigcap_{j\in J} \mathbf{F}_j$ se puede deducir que:

1. Si $\mathbf{F}_0 \subseteq \bigcap_{j\in J} \mathbf{F}_j$, se cumple que si al menos una de las hipótesis componentes $\mathbf{F}_j$ es falsa, entonces $\mathbf{F}_0$ es falsa. O, equivalentemente, si $\mathbf{F}_0$ es verdadera, entonces todas las hipótesis componentes $\mathbf{F}_j$ son verdaderas.

2. Si $\mathbf{F}_0 \supseteq \bigcap_{j \in J} \mathbf{F}_j$, se cumple que si $\mathbf{F}_0$ es falsa, entonces al menos una de las hipótesis componentes $\mathbf{F}_j$ es falsa. O, haciendo uso de la contra-recíproca, si todas las hipótesis componentes $\mathbf{F}_j$ son verdaderas, entonces $\mathbf{F}_0$ es verdadera.

En conclusión:

La hipótesis nula del test inicial es falsa si y sólo si la hipótesis nula de al menos un test componente es falsa.

La hipótesis nula del test inicial es verdadera si y sólo si las hipótesis nulas de todos los test componentes son verdaderas.

Un análisis similar se puede hacer con la igualdad $C = \bigcup_{j \in J} C_j$:

1. Si $C \subseteq \bigcup_{j \in J} C_j$, se cumple que si se rechaza C, entonces al menos uno de los tests componentes C_j se rechaza. Equivalentemente, si ninguno de los tests componentes C_j se rechaza, entonces no se rechaza C.

2. Si $C \supseteq \bigcup_{j \in J} C_j$, se cumple que si se rechaza al menos uno de los tests componentes C_j, entonces se rechaza C. O, Si no se rechaza C, entonces no se rechaza ninguno de los tests componentes C_j.

En consecuencia:

C se rechaza si y sólo si se rechaza al menos uno de los tests componentes C_j.

No se rechaza C si y sólo si ninguno de los tests componentes C_j se rechaza.

Nota: Este método tiene una consecuencia práctica importante que otros métodos no poseen. Si la hipótesis nula se rechaza usando el test Intersección Unión, se puede explorar, utilizando el mismo test, cuáles de los tests componentes son los responsables.

3.4.1. PROCEDIMIENTO DEL TEST INTERSECCIÓN - UNIÓN

Sea $(\Omega, \mathbf{A}, \mathbf{P})$ un espacio estadístico, $\mathbf{X} : \Omega \rightarrow \mathrm{X}$ una muestra y $\mathbf{F}$ la familia de distribuciones a la que pertenece la distribución $F_{\mathbf{X}}$ de la muestra.

Sea la hipótesis inicial: $\mathbf{H}_0 : \mathrm{F}_{\mathbf{X}} \in \mathbf{F}_0$ (en el caso particular de familia de distribuciones paramétricas, la hipótesis es $\mathbf{H}_0 : \mathrm{q} \in \Theta_0$)

1. Especificar una familia de hipótesis componentes $\{\mathbf{F}_j : j \in J\}$ con $\varnothing \neq \mathbf{F}_j \subset \mathbf{F}$, para cada $j \in J$ siendo J algún conjunto de índices tal que $\mathbf{F}_0 = \bigcap_{j \in J} \mathbf{F}_j$. En el caso particular que la familia sea paramétrica, se tiene una familia de hipótesis componentes de la siguiente forma: $\{\Theta_{0,j} : j \in J\}$ con $\varnothing \neq \Theta_{0,j} \subset \Theta$, para cada $j \in J$ siendo J algún conjunto de índices tal que $\Theta_{0,J} = \bigcap_{j \in J} \Theta_{0,j} = \Theta_0$.

2. Especificar la familia de tests componentes $\{C_j : j \in J\}$ donde, para cada $j \in J$ se considera a C_j como un test de $H_{0,j}$, siendo α_j su nivel de significación.

3. Definir el test unión. Como H_0 es verdadera si y sólo si cada hipótesis componente es verdadera, es lógico no rechazar H_0 si y sólo si ninguna de las hipótesis componentes se rechaza; esto es, H_0 se rechaza si y sólo si al menos una hipótesis componente se rechaza. Esto conduce a definir a $C_U = \bigcup_{j \in J} C_j$ como un test de H_0.

4. Obtener una expresión para el test C_U de manera que el nivel de significación α_U sea el deseado.

Teniendo en cuenta las aplicaciones de este procedimiento, los test Intersección Unión se pueden dividir en dos grandes grupos, el primero formado por aquellos tests donde la familia de hipótesis componentes conforma un conjunto numerable, y el segundo integrado por los tests donde la familia de hipótesis componentes es un conjunto infinito no numerable.

4. EJEMPLOS DEL USO DEL TEST INTERSECCIÓN UNIÓN

4.1. EJEMPLO DEL USO DEL TEST INTERSECCIÓN UNIÓN PARA FAMILIAS NUMERABLES DE HIPÓTESIS COMPONENTES

Se considera el problema de contrastar si la media de una población normal es igual a un valor determinado.

Sean $X_1, X_2, \ldots, X_n$ variables aleatorias independientes e idénticamente distribuidas con distribución $N(\mu, \sigma^2)$ con σ^2 conocido $(\sigma^2 = \sigma_0^2)$ y se quiere contrastar $H_0 : \mu = \mu_0$ vs $H_1 : \mu \neq \mu_0$.

Se sabe que el test de cociente de verosimilitudes de nivel α (ver Mood et al.[6]) es

$$C = \left\{ (x_1, \ldots, x_n) : \sqrt{n} \frac{|\bar{x} - \mu_0|}{\sigma_0} \geq z_{1-\alpha/2} \right\},$$

donde $z_{1-\alpha/2}$ es el cuantil $1-\alpha/2$ de la distribución normal estándar, $\bar{X}$ es la media muestral y $\bar{x}$ representa cualquier valor que toma la variable $\bar{X}$.

Este test también se puede obtener a través del test Intersección Unión de la siguiente forma:

Hipótesis inicial:

$$H_0 : \theta = (\mu, \sigma^2) \in \Theta_0 = \{(\mu_0, \sigma_0^2)\}$$

$$\text{vs} \quad H_1 : \theta \in \Theta_1 = \{(\mu, \sigma^2) : \mu \neq \mu_0 \wedge \sigma^2 = \sigma_0^2\},$$

es decir, se tiene que:

$$\Theta_0 = \{(\mu_0, \sigma_0^2)\} \subset \Re^2$$

y $\Theta_1 = \{(\mu, \sigma^2) : \mu \neq \mu_0 \wedge \sigma^2 = \sigma_0^2\} \subset \Re^2.$

Se especifica la familia de hipótesis componentes: (sólo se definen dos hipótesis componentes, o sea $J = \{1, 2\}$)

$$\Theta_{0,1} = \{(\mu, \sigma^2) : \mu \geq \mu_0 \wedge \sigma^2 = \sigma_0^2\} \subset \Re^2$$

vs

$$\Theta_{1,1} = \{(\mu, \sigma^2) : \mu < \mu_0 \wedge \sigma^2 = \sigma_0^2\} \subset \Re^2.$$

Y

$$\Theta_{0,2} = \{(\mu, \sigma^2) : \mu \leq \mu_0 \wedge \sigma^2 = \sigma_0^2\} \subset \Re^2$$

vs

$$\Theta_{1,2} = \{(\mu, \sigma^2) : \mu > \mu_0 \wedge \sigma^2 = \sigma_0^2\} \subset \Re^2.$$

Es evidente que se verifica $\bigcap_{j \in J} \Theta_{0,j} = \Theta_{0,1} \cap \Theta_{0,2} = \Theta_0$. Es decir que se cumple que $\Theta_{0,J} = \Theta_0$.

Se especifica la familia de test componentes:

Sea C_1 el test de $\Theta_{0,1}$ versus $\Theta_{1,1}$, siendo α_1 su nivel de significación. Un test uniformemente más potente para las hipótesis $\Theta_{0,1}$ vs $\Theta_{1,1}$ es el siguiente:

$$C_1 = \left\{ (x_1, \ldots, x_n) : \sqrt{n} \frac{\bar{x} - \mu_0}{\sigma_0} \leq z_{\alpha_1} \right\},$$

donde z_{α_1} es el cuantil α_1 de la distribución normal estándar y $\overline{X}$ es la media muestral.

De igual modo, sea C_2 el test de $\Theta_{0,2}$ versus $\Theta_{1,2}$, siendo α_2 su nivel de significación:

$$C_2 = \left\{ (x_1, \ldots, x_n) : \sqrt{n} \frac{\bar{x} - \mu_0}{\sigma_0} \geq z_{1-\alpha_2} \right\},$$

donde $z_{1-\alpha_2}$ es el cuantil $1-\alpha_2$ de la distribución normal estándar y $\overline{X}$ es la media muestral.

Se define $C_U = \bigcup_{j \in J} C_j = C_1 \cup C_2$ como un test de Θ_0 versus Θ_1:

En este caso, $C_U = C_1 \cup C_2 = \left\{ (x_1, \ldots, x_n) : \sqrt{n} \frac{\bar{x} - \mu_0}{\sigma_0} \leq z_{\alpha_1} \vee \sqrt{n} \frac{\bar{x} - \mu_0}{\sigma_0} \geq z_{1-\alpha_2} \right\}.$

C_1 también es un test uniformemente más potente de nivel α_1 para las hipótesis Θ_0 vs $\Theta_{1,1}$. En consecuencia, como $\Theta_0 \subset \Theta_{0,1}$ y el test C_1 tiene el mismo nivel de significación (a_1) para Θ_0 vs $\Theta_{1,1}$ y para $\Theta_{0,1}$ vs $\Theta_{1,1}$, entonces Θ_0 es determinante para $\Theta_{0,1}$.

Si se hace un razonamiento análogo con el test C_2 y las hipótesis Θ_0 vs $\Theta_{1,2}$ y $\Theta_{0,2}$ vs $\Theta_{1,2}$, se concluye que Θ_0 es determinante para $\Theta_{0,2}$. Por lo tanto, como Θ_0 es determinante para $\Theta_{0,1}$ y $\Theta_{0,2}$, entonces Θ_0 es determinante de $\{\Theta_{0,j} \ : \ j=1,2\}$

Además, se verifica que C_1 y C_2 son tests disjuntos, con nivel de significación a_1 y a_2 respectivamente, por lo que $C_U = C_1 \cup C_2$ es entonces el test unión disjunta de $\Theta_{0,J} = \Theta_0$ vs $\Theta_{1,J} = \Theta_1$ con nivel de significación α_{UI}.

Como Θ_0 es un conjunto unitario y es determinante para $\{\Theta_{0,j} \ : \ j=1,2\}$, entonces de acuerdo con la propiedad 4 se verifica que $\alpha_{UI} = \min\left\{1; \sum_{j\in J} \alpha_j\right\} = \min\{1; \alpha_1 + \alpha_2\}$. Si se considera $\alpha_1 = \alpha_2 = \frac{\alpha}{2}$, entonces $\alpha_{UI} = \min\{1; \alpha_1 + \alpha_2\} = \min\{1; \alpha\} = \alpha$.

Luego

$$C_U = C_1 \cup C_2 = \left\{ (x_1, ..., x_n) \ : \ \sqrt{n} \frac{|\bar{x} - \mu_0|}{\sigma_0} \geq z_{1-\frac{\alpha}{2}} \right\}.$$

En conclusión, si $a_1 = a_2 = \frac{a}{2}$, C_U como test de $\Theta_{0,J} = \Theta_0$ vs $\Theta_{1,J} = \Theta_1$, tiene nivel de significación a y es igual al test que se obtiene a través del cociente de verosimilitudes.

4.2. EJEMPLOS DEL USO DEL TEST INTERSECCIÓN UNIÓN PARA FAMILIAS INFINITAS NO NUMERABLES DE HIPÓTESIS COMPONENTES

En una amplia variedad de aplicaciones de análisis multivariado, el test Intersección Unión se obtiene a partir de una familia infinita de tests componentes, los cuales se obtienen de variables aleatorias que se definen a través de combinaciones lineales de un vector p variado.

Una posible fundamentación de la metodología que se utiliza para obtener estos tests, se basa en el teorema de Cramer-Wold.

4.2.1. EJEMPLIFICACIÓN DE LA HIPÓTESIS INTERSECCIÓN MEDIANTE CRAMER-WOLD. INTERPRETACIÓN DEL TEOREMA DE CRAMER-WOLD

El teorema de Cramer-Wold (ver Mardia[5]) asegura que la distribución de un vector p variado $\mathbf{X}$ está completamente determinada por la familia de distribuciones univariadas de todas las combinaciones lineales $\mathbf{a}'\mathbf{X}$, donde $\mathbf{a} \in \Re_0^p = \Re^p - \{0\}$

.

Sea el vector aleatorio p variado $\mathbf{X}$ con distribución $F_{\mathbf{X}}^{*}:\Re^{p}\to[0;1]$ Es decir, $F_{\mathbf{X}}^{*}\in\mathbf{F}$ con

$$\mathbf{F}=\{F\ :\ F\ :\Re^{p}\to[0;1]\}.$$

Para cada $\mathbf{a}\in\Re_{0}^{p}$ sea $h_{\mathbf{a}}:\Re^{p}\to\Re$ definida como la combinación lineal $h_{\mathbf{a}}(\mathbf{X})=\mathbf{a}'\mathbf{X}$ y sea Y la variable aleatoria definida por $Y=h_{\mathbf{a}}(\mathbf{X})=\mathbf{a}'\mathbf{X}$, con distribución $G_{Y}^{*}=G_{\mathbf{a}'\mathbf{X}}^{*}:\Re\to[0;1]$ que pertenece a la familia de distribuciones univariadas

$$\mathbf{G}=\{G\ :\ G\ :\Re\to[0;1]\}.$$

Se indica con $\mathrm{h}_{\mathbf{a}}$ la función inducida por $h_{\mathbf{a}}$ que a cada distribución $p-$variada F le asigna la función distribución univariada G tal que si el vector aleatorio p variado $\mathbf{Z}$ tiene distribución F entonces la variable $V=h_{\mathbf{a}}(\mathbf{Z})=\mathbf{a}'\mathbf{Z}$ tiene distribución G, o sea, $\mathrm{h}_{\mathbf{a}}\ :\ \mathbf{F}\to\mathbf{G}$.

En particular, $\mathrm{h}_{\mathbf{a}}(F_{\mathbf{X}}^{*})=G_{\mathbf{a}'\mathbf{X}}^{*}$ y es claro que

$$\mathrm{h}_{\mathbf{a}}^{-1}(G_{\mathbf{a}'\mathbf{X}}^{*})=\mathbf{F}_{\mathbf{a}}=\{F\ :\ F\ :\Re^{p}\to[0;1]\ \text{y}\ \mathrm{h}_{\mathbf{a}}(F)=G_{\mathbf{a}'\mathbf{X}}^{*}\}$$

es un subconjunto de $\mathbf{F}$.

Con esta notación, lo que el teorema de Cramer-Wold afirma es simplemente que la intersección de todos los conjuntos $\mathbf{F}_{\mathbf{a}}$ que se obtienen al variar $\mathbf{a}\in\Re_{0}^{p}=\Re^{p}-\{0\}$ es un conjunto esencialmente unitario que se puede representar como

$$\bigcap_{\mathbf{a}\in\Re_{0}^{p}}\mathbf{F}_{\mathbf{a}}=\{F_{\mathbf{X}}^{*}\}$$

En el caso paramétrico, como la forma de la distribución F de $\mathbf{X}$ es conocida, pero no se conoce el valor del parámetro, el teorema de Cramer-Wold afirma que el parámetro de la distribución de $\mathbf{X}$ esta completamente determinado por los parámetros de todas las combinaciones lineales $\mathbf{a}'\mathbf{X}$ donde $\mathbf{a}\in\Re_{0}^{p}$.

Se considera que el vector aleatorio p variado $\mathbf{X}$ tiene distribución $F_{\mathbf{X}}^{*}:\Re^{p}\to[0;1]$ donde $F_{\mathbf{X}}^{*}=F_{\mathbf{X}}^{*}(\cdot,\mathrm{q}^{*})\in\mathbf{F}$ con $\mathbf{F}=\{F(\cdot,\mathrm{q}) : F(\cdot,\mathrm{q})$ función distribución $p-$variada con $\mathrm{q}\in\Theta\}$.

Sea $t_{\mathbf{a}}:\Theta\to\Gamma$ la función que a cada q le asigna $\mathrm{g}_{\mathbf{a}}=g_{\mathbf{a}}(\mathrm{q})$ el parámetro transformado del vector $\mathbf{Z}$ por la transformación lineal, o sea, el parámetro de la variable $V=h_{\mathbf{a}}(\mathbf{Z})=\mathbf{a}'\mathbf{Z}$. En particular, $t_{\mathbf{a}}(\theta^{*})=\gamma_{\mathbf{a}}^{*}$ y $t_{\mathbf{a}}^{-1}(\gamma_{\mathbf{a}}^{*})=\Theta_{0,\mathbf{a}}\subseteq\Theta$. Entonces el teorema de Cramer-Wold establece que

$$\bigcap_{\mathbf{a}\in\Re_{0}^{p}}\Theta_{0,\mathbf{a}}=\{\theta^{*}\}$$

La consecuencia importante de este teorema es que se puede estudiar el comportamiento de un vector p variado a través de la exploración de variables univariadas.

Uso del teorema de Cramer-Wold en los tests Intersección Unión

Usando la notación precedente, dada una hipótesis cualquiera $\mathrm{H} \subset \mathrm{F}$, se tiene que

$$\mathrm{h}_{\mathbf{a}}^{-1}(\mathrm{h}_{\mathbf{a}}(\mathrm{H})) = \mathrm{H}_{\mathbf{a}}$$

y es claro que

$$\mathrm{H} \subset \mathrm{H}_{\mathbf{a}}.$$

El teorema de Cramer-Wold permite concluir que

$$\mathrm{H} = \bigcap_{\mathbf{a} \in \Re_0^p} \mathrm{H}_{\mathbf{a}}.$$

En el caso paramétrico y usando la notación precedente, dada una hipótesis cualquiera $\Theta_0 \subset \Theta$ se tiene que

$$t_{\mathbf{a}}^{-1}(t_{\mathbf{a}}(\Theta_0)) = \Theta_{0,\mathbf{a}} \subseteq \Theta$$

y el teorema de Cramer-Wold permite asegurar que

$$\Theta_0 = \bigcap_{\mathbf{a} \in \Re_0^p} \Theta_{0,\mathbf{a}}.$$

Resulta pertinente aclarar que se trabaja con tres conjuntos diferentes, a saber: Θ_0 la hipótesis nula inicial, $\Theta_{0,\mathbf{a}}$ la hipótesis nula de los tests componentes y $\Theta_{0,\mathbf{a}}^*$ la hipótesis nula de los tests univariados asociados a los tests componentes.

El teorema de Cramer-Wold se puede utilizar para construir tests Intersección Unión, ya que permite asegurar que toda hipótesis sobre la distribución de un vector p variado se puede expresar como la intersección de todos los miembros de una determinada familia de hipótesis, las cuales se obtienen de antitransformar las hipótesis univariadas asociadas a las combinaciones lineales $\mathbf{a}'\mathbf{X}$ donde $\mathbf{a} \in \Re_0^p$.

4.2.1.1. EJEMPLO DEL USO DEL TEOREMA DE CRAMER-WOLD EN LOS TESTS INTERSECCIÓN UNIÓN

Sean $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ vectores aleatorios independientes e idénticamente distribuidos con distribución $N_p(\mathrm{m}, \Sigma)$ con Σ definida positiva $(\Sigma > 0)$.

Sea R una matriz fija de orden $q \times p$ de rango q y se consideran los vectores transformados $\mathbf{H}_1 = R\mathbf{X}_1, \mathbf{H}_2 = R\mathbf{X}_2, \ldots, \mathbf{H}_n = R\mathbf{X}_n$ quienes, por propiedades de la distribución normal, son independientes e idénticamente distribuidos con distribución $N_q(\mathrm{m}_*, \Sigma_*)$ donde $\mathrm{m}_* = R\mathrm{m}$ y $\Sigma_* = R\Sigma R' > 0$.

Para cada $\mathbf{a} \in \Re_0^p$ se definen las variables univariadas $Y_{\mathbf{a}} = \mathbf{a}'\mathbf{H}$, y se tiene entonces que $Y_{1,\mathbf{a}} = \mathbf{a}'\mathbf{H}_1$, $Y_{2,\mathbf{a}} = \mathbf{a}'\mathbf{H}_2,\ldots, Y_{n,\mathbf{a}} = \mathbf{a}'\mathbf{H}_n$ es una muestra aleatoria de una distribución $N(\mu_{\mathbf{a}},\sigma^2_{\mathbf{a}})$, donde $\mu_{\mathbf{a}} = \mathbf{a}'\mu = \mathbf{a}'R\mu$ y $\sigma^2_{\mathbf{a}} = \mathbf{a}'\Sigma_{*}\mathbf{a} = \mathbf{a}'R\Sigma R'\mathbf{a} > 0$.

Se considera el problema de determinar si ciertas combinaciones lineales de las componentes del vector de medias de una población normal multivariada, son iguales a determinados valores fijos, siendo la matriz de covarianzas desconocida.

Sea la hipótesis inicial:

$$\mathrm{H}_0 \; : \; R\mu = \mathbf{r}, \text{siendo } \Sigma > 0 \text{ desconocida}$$

vs

$$\mathrm{H}_1 \; : \; R\mu \neq \mathbf{r}, \text{siendo } \Sigma > 0 \text{ desconocida},$$

se define

$$\Theta_0 = \left\{(\mu,\Sigma) : \; R\mu = \mathbf{r} \wedge 0 < \Sigma \in \Re^{p\times p}\right\} \subset \Re^{p+p^2}$$

y

$$\Theta_1 = \left\{((\mu,\Sigma) : \; R\mu \neq \mathbf{r} \wedge 0 < \Sigma \in \Re^{p\times p}\right\} \subset \Re^{p+p^2},$$

donde, R es una matriz de orden $q \times p$ conocida y $\mathbf{r}$ es un vector q variado conocido.

Para obtener el test, se observa que bajo la hipótesis nula los vectores transformados $\mathbf{H}_1, \mathbf{H}_2,\ldots,\mathbf{H}_n$ son independientes e idénticamente distribuidos con distribución $N_q(\mathbf{r}, R\Sigma R')$ y, por lo tanto, para cada $\mathbf{a} \in \Re_0^p$ son las variables $Y_{1,\mathbf{a}} = \mathbf{a}'\mathbf{H}_1, Y_{2,\mathbf{a}} = \mathbf{a}'\mathbf{H}_2,\ldots,Y_{n,\mathbf{a}} = \mathbf{a}'\mathbf{H}_n$ independientes e idénticamente distribuidas con distribución $N(\mu_{\mathbf{a}},\sigma^2_{\mathbf{a}})$ donde $\mu_{\mathbf{a}} = \mathbf{a}'\mathbf{r}$ y $\sigma^2_{\mathbf{a}} = \mathbf{a}'R\Sigma R'\mathbf{a}$.

Se definen las hipótesis univariadas asociadas

$$\mathrm{H}^*_{0,\mathbf{a}} \; : \; \mu_{\mathbf{a}} = \mathbf{a}'\mathbf{r}, \text{siendo } \sigma^2_{\mathbf{a}} > 0 \text{ desconocida}$$

vs

$$\mathrm{H}^*_{1,\mathbf{a}} \; : \; \mu_{\mathbf{a}} \neq \mathbf{a}'\mathbf{r}, \text{siendo } \sigma^2_{\mathbf{a}} > 0 \text{ desconocida}$$

se tiene

$$\Theta^*_{0,\mathbf{a}} = \left\{\left(\mu_{\mathbf{a}},\sigma^2_{\mathbf{a}}\right) \; : \; \mu_{\mathbf{a}} = \mathbf{a}'\mathbf{r} \wedge \sigma^2_{\mathbf{a}} \in \Re^+\right\} \subset \Re^2$$

y

$$\Theta^*_{1,\mathbf{a}} = \left\{\left(\mu_{\mathbf{a}},\sigma^2_{\mathbf{a}}\right) \; : \; \mu_{\mathbf{a}} \neq \mathbf{a}'\mathbf{r} \wedge \sigma^2_{\mathbf{a}} \in \Re^+\right\} \subset \Re^2.$$

Se definen entonces como hipótesis componentes a

$$\mathrm{H}_{0,\mathbf{a}} \; : \; \mathbf{a}'R\mu = \mathbf{a}'\mathbf{r}, \text{siendo } \Sigma > 0 \text{ desconocida}$$

vs

$$\mathrm{H}_{1,\mathbf{a}} \; : \; \mathbf{a}'R\mu \neq \mathbf{a}'\mathbf{r}, \text{siendo } \Sigma > 0 \text{ desconocida}$$

luego

$$\Theta_{0,\mathbf{a}} = \left\{(\mu,\Sigma) : \; \mathbf{a}'R\mu = \mathbf{a}'\mathbf{r} \wedge 0 < \Sigma \in \Re^{p\times p}\right\} \subset \Re^{p+p^2}$$

y

$$\Theta_{1,\mathbf{a}} = \left\{(\mu,\Sigma) : \; \mathbf{a}'R\mu \neq \mathbf{a}'\mathbf{r} \wedge 0 < \Sigma \in \Re^{p\times p}\right\} \subset \Re^{p+p^2},$$

se cumple que

$$\Theta_0 = \bigcap_{\mathbf{a}\in\Re_0^p} \Theta_{0,\mathbf{a}}.$$

Para cada $\mathbf{a}\in\Re_0^p$ el estadístico $T_{\mathbf{a}} = \frac{\overline{Y}_{\mathbf{a}}-\mu_{\mathbf{a}}}{\sqrt{S_{\mathbf{a}}^2/(n-1)}}$ tiene, bajo las hipótesis univariadas $\mathrm{H}_{0,\mathbf{a}}^*$, una distribución t de student con $n-1$ grados de libertad. Por lo tanto un test para la hipótesis $\mathrm{H}_{0,\mathbf{a}}^*$ es

$$C_{\mathbf{a}}^* = \left\{ \left(y_{1,\mathbf{a}}, y_{2,\mathbf{a}},\ldots, y_{n,\mathbf{a}}\right)^t \ : \ \left| \frac{\overline{y}_{\mathbf{a}} - \mu_{\mathbf{a}}}{\sqrt{S_{\mathbf{a}}^2/(n-1)}} \right| > c_{\mathbf{a}} \right\} \subset \Re^n,$$

donde $c_{\mathbf{a}}$ es algún valor crítico de la t de student con $n-1$ grados de libertad. Recordando que $\mu_{\mathbf{a}} = \mathbf{a}'\mathbf{r}$ y obteniendo

$$\overline{Y}_{\mathbf{a}} = \mathbf{a}'R\overline{\mathbf{X}} \quad \text{y} \quad \mathrm{S}_{\mathbf{a}}^2 = \mathbf{a}'R\mathrm{S}R'\mathbf{a},$$

donde $\mathrm{S} = \frac{1}{n}\sum_{j=1}^{n}\left(\mathbf{X}_j - \overline{\mathbf{X}}\right)\left(\mathbf{X}_j - \overline{\mathbf{X}}\right)$ es el estimador de máxima verosimilitud de Σ, un test equivalente para $\mathrm{H}_{0,\mathbf{a}}$ es

$$C_{\mathbf{a}} = \left\{\left(\mathbf{x}_1', \mathbf{x}_2',\ldots,\mathbf{x}_n'\right)^t \ : \ |t_{\mathbf{a}}| > c_{\mathbf{a}}\right\} = \left\{\left(\mathbf{x}_1', \mathbf{x}_2',\ldots,\mathbf{x}_n'\right)^t \ : \ \left|\frac{\mathbf{a}'\left(R\overline{\mathbf{x}} - \mathbf{r}\right)}{\sqrt{\mathbf{a}'RSR'\mathbf{a}/(n-1)}}\right| > c_{\mathbf{a}}\right\} \subset \Re^{np}.$$

Este test también se puede escribir

$$C_{\mathbf{a}} = \left\{\left(\mathbf{x}_1', \mathbf{x}_2',\ldots,\mathbf{x}_n'\right)^t \ : \ t_{\mathbf{a}}^2 > c_{\mathbf{a}}^2\right\} = \left\{\left(\mathbf{x}_1', \mathbf{x}_2',\ldots,\mathbf{x}_n'\right)^t \ : (n-1)\frac{\mathbf{a}'\left(R\overline{\mathbf{x}} - \mathbf{r}\right)\left(R\overline{\mathbf{x}} - \mathbf{r}\right)^t\mathbf{a}}{\mathbf{a}'RSR'\mathbf{a}} > c_a^2\right\}.$$

Luego, el test Intersección Unión es

$$C_U = \bigcup_{\mathbf{a}\in\Re_0^p} C_{\mathbf{a}}.$$

Es decir, que la hipótesis nula inicial H_0 se rechaza si y sólo si al menos una de las $t_{\mathbf{a}}^2 > c_{\mathbf{a}}^2$; esto es, H_0 no se rechaza si y sólo si $t_{\mathbf{a}}^2 \le c_{\mathbf{a}}^2$, para todo $\mathbf{a}\in\Re_0^p$. Que es lo mismo que decir, H_0 no se rechaza si y sólo si

$$\max_{\mathbf{a}} t_{\mathbf{a}}^2 \le c^2,$$

donde c^2 es algún valor crítico obtenido de la distribución de

$$T^2 = \max_{\mathbf{a}} T_{\mathbf{a}}^2 = (n-1)\left(R\overline{\mathbf{X}} - \mathbf{r}\right)^t (R\mathrm{S}R')^{-1}\left(R\overline{\mathbf{X}} - \mathbf{r}\right)$$

que es Hotelling con parametros q y $n-1$.

Por lo tanto, no se rechaza H_0 si y sólo si

$$t^2 = (n-1)\left(R\overline{\mathbf{x}} - \mathbf{r}\right)^t (RsR')^{-1}\left(R\overline{\mathbf{x}} - \mathbf{r}\right) \le c^2,$$

donde c^2 es un cuantil de la distribución $H(q,n-1)$

Finalmente, resulta que el test Intersección Unión de tamaño a está dado por

$$C_U = \left\{\left(\mathbf{x}_1', \mathbf{x}_2',\ldots,\mathbf{x}_n'\right)^t \ : (n-1)\left(R\overline{\mathbf{x}} - \mathbf{r}\right)^t (RSR')^{-1}\left(R\overline{\mathbf{x}} - \mathbf{r}\right) > H_{1-\alpha}(q,n-1)\right\}.$$

Recordando que si T^2 tiene distribución $H(p,n)$ entonces $F = \frac{n-p+1}{p}T^2$ tiene distribución $F(p, n-p+1)$, podemos definir el test Intersección Unión de tamaño a como

$$C_U = \left\{ (\mathbf{x}_1', \mathbf{x}_2', ..., \mathbf{x}_n')' \ : \ (R\bar{\mathbf{x}} - \mathbf{r})'(RSR')^{-1}(R\bar{\mathbf{x}} - \mathbf{r}) > \frac{q}{n-q} F_{1-\alpha}(q, n-q) \right\}.$$

4.2.2. EXTENSIÓN DEL TEOREMA DE CRAMER-WOLD

Para justificar la obtención del test Intersección Unión del próximo ejemplo es menester demostrar la siguiente extensión del teorema de Cramer-Wold para abarcar el caso de matrices aleatorias de orden $n \times p$, cuyas filas son vectores aleatorios p variados independientes entre sí.

• *Teorema:* La distribución de una matriz U de orden $n \times p$, cuyas filas son vectores aleatorios p variados independientes entre sí, está completamente determinada por la familia de las distribuciones univariadas $\mathbf{a}'U\mathbf{b}$, donde $\mathbf{a} \in \Re_0^n$ recorre todos los vectores n variados fijos no nulos y $\mathbf{b} \in \Re_0^p$ recorre todos los vectores p variados fijos no nulos.

Demostración:

Se considera la matriz U

$$U = \begin{pmatrix} U_{11} & U_{12} & \dots & U_{1p} \\ U_{21} & U_{22} & \dots & U_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ U_{n1} & U_{n2} & \dots & U_{np} \end{pmatrix} = \begin{pmatrix} \mathbf{U}_1' \\ \mathbf{U}_2' \\ \vdots \\ \mathbf{U}_n' \end{pmatrix},$$

donde $\mathbf{U}_i' = (U_{i1}, U_{i2}, ..., U_p)$ representa al $i-ésimo$ vector fila, para $i = 1,2,...,n$.

Para obtener la distribución de la matriz U, se la escribe como el vector p variado que se obtiene colocando ordenadamente las filas transpuestas de la matriz una debajo de la otra. A este vector se lo denota $vec(U)$ esto es,

$$vec(U) = (\mathbf{U}_1^t \quad \mathbf{U}_2^t \quad \cdots \quad \mathbf{U}_n^t)'.$$

Como los vectores p variados $\mathbf{U}_i$ son independientes entre sí, entonces

$$\varphi_{vec(U)}(\mathbf{t}) = \varphi_{\mathbf{U}_1}(\mathbf{t}_1)\varphi_{\mathbf{U}_2}(\mathbf{t}_2)...\varphi_{\mathbf{U}_n}(\mathbf{t}_n) = E\left(e^{i\mathbf{t}_1'\mathbf{U}_1}\right)E\left(e^{i\mathbf{t}_2'\mathbf{U}_2}\right)...E\left(e^{i\mathbf{t}_n'\mathbf{U}_n}\right),$$

donde $\mathbf{t}' = (\mathbf{t}_1', \mathbf{t}_2', ..., \mathbf{t}_n')$.

De donde se observa que si se conocen las distribuciones de todos los vectores $\mathbf{U}_i$ $(i = 1,2,...,n)$ se conoce la distribución de $vec(U)$ y, por lo tanto, la distribución de la matriz U, y viceversa, si se conoce la distribución de la matriz U, se conoce la distribución de todos los $\mathbf{U}_i$.

Se considera el vector n variado $\mathbf{X}_{\mathbf{b}} = U\mathbf{b}$, donde $\mathbf{b}$ es un vector fijo. Claramente las componentes de este vector son variables aleatorias independientes entre sí.

La función característica de $\mathbf{X}_{\mathbf{b}}$ es

$$\begin{aligned}\varphi_{\mathbf{X}_{\mathbf{b}}}(\mathbf{s}) &= \varphi_{\mathbf{U}_1'\mathbf{b}}(s_1)\varphi_{\mathbf{U}_2'\mathbf{b}}(s_2)\ldots\varphi_{\mathbf{U}_n'\mathbf{b}}(s_n) \\ &= E\left(e^{is_1\mathbf{U}_1'\mathbf{b}}\right)E\left(e^{is_2\mathbf{U}_2'\mathbf{b}}\right)\ldots E\left(e^{is_n\mathbf{U}_n'\mathbf{b}}\right) = E\left(e^{i\mathbf{b}_1'\mathbf{U}_1}\right)E\left(e^{i\mathbf{b}_2'\mathbf{U}_2}\right)\cdots E\left(e^{i\mathbf{b}_n'\mathbf{U}_n}\right)\end{aligned}$$

donde $\mathbf{s} = (s_1, s_2, \ldots, s_n)' \in \Re^n$ y $\mathbf{b}_i = s_i\mathbf{b}$. Luego:

$$\varphi_{\mathbf{X}_{\mathbf{b}}}(\mathbf{s}) = \varphi_{\mathbf{X}_{\mathbf{b}}}(\mathbf{b}_1', \mathbf{b}_2', \ldots, \mathbf{b}_n') = \varphi_{\mathbf{U}_1}(\mathbf{b}_1)\varphi_{\mathbf{U}_2}(\mathbf{b}_2)\cdots\varphi_{\mathbf{U}_n}(\mathbf{b}_n).$$

Por lo tanto, el conocimiento de la función característica de todos los vectores $\mathbf{X}_{\mathbf{b}} = U\mathbf{b}$, permite conocer la función característica de todos los $\mathbf{U}_i$ a través de la función característica de la distribución marginal $\varphi_{\mathbf{X}_{\mathbf{b}_i}}(0, \ldots, 0, \mathbf{b}_i, 0, 0, \ldots, 0) = \varphi_{\mathbf{U}_i}(\mathbf{b}_i)$ Y viceversa, si se conoce la distribución de todos los vectores $\mathbf{U}_i$, eligiendo $(\mathbf{b}_1', \mathbf{b}_2', \ldots, \mathbf{b}_n') = ((s_1\mathbf{b})', (s_2\mathbf{b})', \ldots, (s_n\mathbf{b})')$ permite conocer la distribución de $\mathbf{X}_{\mathbf{b}} = U\mathbf{b}$.

De donde, si se conoce la distribución de todos los vectores $\mathbf{X}_{\mathbf{b}} = U\mathbf{b}$, la distribución de los vectores $\mathbf{U}_i$ queda completamente determinada, y por lo tanto, si se conoce la distribución de todos los vectores $\mathbf{X}_{\mathbf{b}} = U\mathbf{b}$, la distribución de la matriz U queda completamente determinada.

De acuerdo con el teorema de Cramer-Wold (ver Mardia[5]), si se conoce la función característica de todas las variables $Y = \mathbf{a}'\mathbf{X}_{\mathbf{b}}$, la distribución del vector $\mathbf{X}_{\mathbf{b}}$ queda completamente determinada.

En conclusión, si se conoce la distribución de todos los vectores $\mathbf{X}_{\mathbf{b}} = U\mathbf{b}$, la distribución de la matriz U queda completamente caracterizada y si se conoce la distribución de todas las variables $Y = \mathbf{a}'\mathbf{X}_{\mathbf{b}}$, la distribución del vector $\mathbf{X}_{\mathbf{b}}$ queda completamente determinada; se sigue que, si se conoce la distribución de todas las variables $\mathbf{a}'U\mathbf{b}$, la distribución de la matriz U queda completamente caracterizada.

4.2.2.1. EJEMPLO DEL USO DE LA EXTENSIÓN DEL TEOREMA DE CRAMER-WOLD EN LOS TESTS INTERSECCIÓN - UNIÓN

Se considera el modelo definido por:

$$Y = X\beta + U,$$

donde Y es una matriz aleatoria observada $(n \times p)$ de p variables respuestas para cada uno de los n individuos, X es una matriz no aleatoria conocida de orden $n \times q$, b es una matriz de parámetros de regresión desconocidos de orden $q \times p$, y U es una matriz de errores aleatorios no observados tal que cada una de sus filas tiene vector de medias $\mathbf{0}$ y matriz de covarianzas común Σ.

Se supone que U tiene distribución normal, esto es, U es una matriz proveniente de una $N_p(0,S)$. En consecuencia, Y es una matriz proveniente de una distribución normal $N_p(X\mathrm{b},\Sigma)$ es decir, que las filas de la matriz Y corresponden a vectores aleatorios $N_p(X\mathrm{b},\Sigma)$ independientes.

Se desea contrastar hipótesis de la forma $C_1\mathrm{b}M_1 = D$, donde $C_1(g\times q)$ $M_1(p\times r)$ y $D(g\times r)$ son matrices dadas y, C_1 y M_1 tienen rango g y r, respectivamente.

Sea la hipótesis inicial:

$$\mathrm{H}_0 \;:\; C_1\beta M_1 = D$$

vs

$$\mathrm{H}_1 \;:\; C_1\beta M_1 \neq D.$$

Para cada $\mathbf{a}\in\Re_0^r$ se definen las variables univariadas $Y_{i,\mathbf{a}} = \mathbf{Y}_i^t M_j\mathbf{a}$, donde $\mathbf{Y}_i^t = (Y_{i1}, Y_{i2}, \ldots, Y_p)$ representa al $i-ésimo$ vector fila de la matriz Y. Es evidente, que las variables $Y_{1,\mathbf{a}}, Y_{2,\mathbf{a}}, \ldots, Y_{n,\mathbf{a}}$ son independientes entre sí. En consecuencia, se puede definir un modelo de regresión lineal univariada de la siguiente forma:

$$\begin{pmatrix} Y_{1,\mathbf{a}} & Y_{2,\mathbf{a}} & \cdots & Y_{n,\mathbf{a}} \end{pmatrix}^t = YM_j\mathbf{a} = X\beta M_j\mathbf{a} + UM_j\mathbf{a}.$$

Se definen las hipótesis univariadas asociadas:

$$\mathrm{H}^*_{0,\mathbf{ab}} \;:\; \mathbf{t}\gamma = \gamma_0$$

vs

$$\mathrm{H}^*_{1,\mathbf{ab}} \;:\; \mathbf{t}\gamma \neq \gamma_0,$$

donde $\mathbf{t} = \mathbf{b}^t C_1$; $\mathrm{g} = \mathrm{b}M_1\mathbf{a}$; $\mathrm{g}_0 = \mathbf{b}^t D\mathbf{a}$ y $\mathbf{b}\in\Re_0^g$.

Se definen las hipótesis componentes de la siguiente manera:

$$\mathrm{H}_{0,\mathbf{ab}} \;:\; \mathbf{b}^t C_1\beta M_1\mathbf{a} = \mathbf{b}^t D\mathbf{a}$$

vs

$$\mathrm{H}_{1,\mathbf{ab}} \;:\; \mathbf{b}^t C_1\beta M_1\mathbf{a} \neq \mathbf{b}^t D\mathbf{a},$$

para todo $\mathbf{a}\in\Re_0^r$ y para todo $\mathbf{b}\in\Re_0^g$.

Se cumple que:

$$\mathrm{H}_0 = \bigcap_{(\mathbf{a}\in\Re_0^r,\, \mathbf{b}\in\Re_0^g)} \mathrm{H}_{0,\mathbf{ab}}.$$

Para contrastar las hipótesis univariadas $\mathrm{H}^*_{0,\mathbf{a}}$, un camino (ver Searle[8]) es considerar el estadístico

$$F^* = \frac{\left\{(\mathbf{t}\hat{\gamma}-\gamma_0)^t\left[\mathbf{t}(X^tX)^{-1}\mathbf{t}^t\right]^{-1}(\mathbf{t}\hat{\gamma}-\gamma_0)\right\}/r}{\left\{(YM_1\mathbf{a})^t\left[I - X(X^tX)^{-1}X^t\right](YM_1\mathbf{a})\right\}/(n-q)},$$

que, bajo hipótesis nula, tiene distribución F de Fisher con r y $n-q$ grados de libertad, donde r representa el número de ecuaciones de $\mathbf{t}\mathrm{g} = \mathrm{g}_0$ independientes, que en este caso, como se tiene sólo una ecuación, $r=1$. Se indica con f^* el valor de este estadístico F^* correspondiente a los valores observados.

Se propone el siguiente test:

$$C^*_{\mathbf{ab}} = \left\{(x_{11},\ldots,x_{nq},y_{1,a},y_{2,a},\ldots,y_{n,a})^t \ : \ f^* > c_{\mathbf{a},\mathbf{b}}\right\} \subset \Re^{nq+n},$$

siendo $c_{\mathbf{a},\mathbf{b}}$ algún valor crítico.

Recordando que $\mathbf{t}=\mathbf{b}'C_1$; $\mathrm{g}=\mathrm{b}M_1\mathbf{a}$; $\mathrm{g}_0=\mathbf{b}'D\mathbf{a}$, un test equivalente para $\mathbf{H}_{0,\mathbf{ab}}$ es:

$$C_{\mathbf{ab}} = \left\{(x_{11},\ldots,x_{nq},y_{11},\ldots,y_{np})^t \ : \ f_{\mathbf{ab}} > c_{\mathbf{a},\mathbf{b}}\right\} \subset \Re^{np+nq}, \text{donde}$$

$$f_{\mathbf{ab}} = \frac{\left\{\mathbf{b}'C_1(X'X)^{-1}X'y_+M_1\mathbf{a}\right\}^2}{\left\{\mathbf{b}'C_1(X'X)^{-1}C_1'\mathbf{b}\right\}\left\{\mathbf{a}'M_1'y'PyM_1\mathbf{a}\right\}}$$ es algún valor observado de

$$F_{\mathbf{ab}} = \frac{\left\{\mathbf{b}'C_1(X'X)^{-1}X'Y_+M_1\mathbf{a}\right\}^2}{\left\{\mathbf{b}'C_1(X'X)^{-1}C_1'\mathbf{b}\right\}\left\{\mathbf{a}'M_1'Y'PYM_1\mathbf{a}\right\}}$$ que tiene, bajo hipótesis nula, una distribución

$(n-q)^{-1}F_{1,n-q}$ para valores fijos de $\mathbf{a}$ y $\mathbf{b}$.

Luego, el test Intersección Unión se define de la siguiente manera:

$$C_U = \bigcup_{\left(\mathbf{a}\in\Re_0^r,\ \mathbf{b}\in\Re_0^g\right)} C_{\mathbf{ab}}.$$

Es decir, que la hipótesis nula inicial $\mathbf{H}_0$ se rechaza si y sólo si al menos una de las $f_{\mathbf{ab}} > c_{\mathbf{a},\mathbf{b}}$. Es lo mismo que decir, $\mathbf{H}_0$ no se rechaza si y sólo si:

$$\max_{\mathbf{a},\mathbf{b}} f_{\mathbf{ab}} \le c_{\mathbf{a},\mathbf{b}},$$

donde $c_{\mathbf{a},\mathbf{b}}$ es algún valor crítico de la distribución de $\max_{\mathbf{a},\mathbf{b}} F_{\mathbf{ab}}$.

Maximizando la expresión de $f_{\mathbf{ab}}$ respecto de $\mathbf{b}$, (ver Mardia[5]) se obtiene:

$$\frac{\mathbf{a}^t H\mathbf{a}}{\mathbf{a}^t E\mathbf{a}},$$

que es el valor observado de $\frac{\mathbf{a}^t\mathbf{H}\mathbf{a}}{\mathbf{a}^t\mathbf{E}\mathbf{a}}$ que tiene una distribución $\frac{g}{n-q}F_{g,n-q}$ para $\mathbf{a}$ fijo, siendo

$$\mathbf{H} = M_1'Y_+'X(X'X)^{-1}C_1'\left[C_1(X'X)^{-1}C_1'\right]^{-1}C_1(X'X)^{-1}X'Y_+M_1,$$

$Y_- = Y - X\mathrm{b}_0$, donde b_0 es cualquier matriz de orden $q\times p$ que satisface $C_1\mathbf{b}_0M_1 = D$ y

$$\mathbf{E} = M_1'Y'PYM_1 = M_1'Y'\left[I - X(X'X)^{-1}X'\right]YM_1.$$

Finalmente, maximizando $\frac{\mathbf{a}^t H\mathbf{a}}{\mathbf{a}^t E\mathbf{a}}$ respecto de $\mathbf{a}$, (ver Mardia[5]) se obtiene:

$$\max_{\mathbf{a},\mathbf{b}} f_{\mathbf{ab}} = \lambda_1(HE^{-1}),$$

donde $l_1(\boldsymbol{B}^{-1})$ denota el mayor autovalor de la matriz B^{-1}.

La hipótesis nula inicial H_0 se rechaza si y sólo si:

$$\lambda_1(HE^{-1}) > c,$$

donde c es un valor crítico que se elige de la distribución de la variable $\Lambda_1 = l_1(\mathbf{B}^{-1})$ de manera que el tamaño del test sea a.

Finalmente, se define el test Intersección Unión de tamaño a de la siguiente manera:

$$C_U = \left\{ (x_{11},\ldots,x_{mq},y_{11},\ldots,y_{mp})^t \;:\; \lambda_1(HE^{-1}) > c \right\}$$

5. DISCUSIÓN

Generalmente, los modelos estadísticos están relacionados con investigaciones multidisciplinarias que conducen a modelos complejos que no siempre permiten computar en formas explícitas y manejables a los estimadores de máxima verosimilitud ni tampoco a los tests de cociente de verosimilitudes. Aún así, se han desarrollado varios algoritmos para solucionar estos inconvenientes de cómputos, pero las propiedades tan convenientes que presentan los estimadores de máxima verosimilitud y los tests de cociente de verosimilitudes en la familia exponencial de densidades y para muestras finitas, no se verifican, en general, en modelos más complejos donde las leyes de probabilidad subyacentes raras veces son miembros de esta familia.

Mientras más complejo es un modelo, más difícil es implementar las herramientas de inferencia estadística clásica. Los estudios estadísticos realizados en el campo genómico es un ejemplo de las grandes dificultades que se presentan al intentar aplicar las herramientas de inferencia estadística convencional a estructuras multidimensionales y en muestras pequeñas.

Frente a esta situación, el principio de Intersección Unión de Roy[7] surge como una alternativa viable, en algunos casos con ventajas computacionales, incrementando las posibilidades de aplicabilidad, mejor adaptación a situaciones no estándares y buenas perspectivas de robustez.

Si bien en sus primeros desarrollos, el procedimiento de Intersección Unión es utilizado sólo en los problemas de comparaciones múltiples e inferencia estadística simultánea, en la actualidad, además de presentarse naturalmente en situaciones de inferencia simultánea, es muy utilizado en el campo de la inferencia estadística con restricciones, e incluso ha motivado una gran cantidad de procedimientos como son, entre otros, los de estimación de matrices de covarianzas bajo un ordenamiento de árbol (Tsai[11]), los de inferencia de correlaciones canónicas restringidas (Das; Sen[2]) y tests no paramétricos para ordenar diversidad de secuencias genómicas (Sen[9]).

La evolución de la inferencia estadística con restricciones durante las últimas cuatro décadas, ha dado mayor furor a los procedimientos de Intersección Unión para estructuras paramétricas, no paramétricas y semiparamétricas. Algunas de las aplicaciones que se pueden nombrar en este sentido son los problemas relacionados con orden estocástico (dominance), los de ensayos clínicos y los de meta análisis, este último es muy usado en el desarrollo de estudios genómicos y bioinformáticos; para ver ejemplos de la utilización del test Intersección Unión en algunas de estas aplicaciones remitirse al trabajo de Sen[10].

Esta evolución, además de incrementar el número de trabajos que emplean el principio propuesto por Roy, ha puesto en evidencia la carencia de una presentación formal que destaque las distintas etapas del principio, ya que en muchas aplicaciones se utiliza algún aspecto del procedimiento en detrimento de otros.

Como se propone en la introducción de este trabajo, es posible unificar las diferentes aplicaciones de los tests Intersección Unión desarrollados por Roy en una presentación formal. En esta presentación formal, se destacan las distintas etapas del procedimiento del test y se clarifican cuáles hipótesis son las que constituyen la familia de hipótesis componentes, ya que en algunos casos de familias infinitas no numerables de hipótesis componentes no se señala claramente la diferencia entre hipótesis univariadas e hipótesis componentes asociadas.

Para realizar esta unificación, es necesario definir los conceptos de hipótesis y de tests como subconjuntos apropiados, definir nociones como las de tests suficientes, hipótesis determinantes, entre otras, y desarrollar propiedades que vinculan las funciones de potencia y los niveles de significación del test unión con sus respectivos tests componentes. Para el caso de familias infinitas no numerables de hipótesis componentes, se propone, además, una posible fundamentación de la metodología que se utiliza basada en el teorema de Cramer-Wold, y para uno de los ejemplos presentados es necesario desarrollar una extensión adecuada del mencionado teorema. Como resultado de estos desarrollos, es posible presentar diversos ejemplos, algunos para familias numerables y otros para familias infinitas no numerables de hipótesis componentes, de una manera formal y unificada.

BIBLIOGRAFÍA

Burrill, C. W. (1972), *Measure, Integration, and Probability*, United States of America, McGraw-Hill, Inc.

Das, S.; Sen, P.K. (1994) "Restricted canonical correlations", *Linear Algebra Application*, 210, 29-47.

Graybill, F.A. (1976), *Theory and Aplication of the Linear Model*, United States of America , Wadsworth & Brooks/Cole Advanced Books & Software.

Laha, R.G.; Rohatgi, V.K. (1979) *Probability Theory*, United States of America, John Wiley & Sons, Inc.

Mardia, K.V.; Kent, J.T.; Bibby, J.M. (1982) *Multivariate Analysis*, United States of America, Academic Press, Inc.

Mood, A.M.; Graybill, F.A.; Boes, D.C. (1974) *Introduction to the Theory of Statistics* (3rd ed.), United States of America, McGraw-Hill, Inc.

Roy, S.N. (1953) "On a heuristic method of test construction and its use in multivariate analysis", *The Annals of Mathematical Statistics*, 24, 220-238.

Searle, S.R. (1997) *Linear Models*, United States of America, John Wiley & Sons, Inc.

Sen, P.K. (2005) "Nonparametric tests for ordered diversity in a genomic sequence", *Journal of Statistical Research*, 39, 1-15.

Sen, P.K. (2007) "Union Intersection principle and constrained statistical inference", *Journal of Statistical Planning and Inference*, 137, 3741-3752.

Tsai, M. (1995) "Estimation of covariances under Lowner order restrictions", *Sankhya A*, 57, 433-439.

Wasserman, L. (2005) *All of Statistics*, United States of America, Springer.