El contraste de hipótesis compuesta frente a alternativa simple como problema de programación matemática infinita

Por RAMIRO MELENDRERAS GIMENO y EDUARDO RAMOS MÉNDEZ

Recibido: 6 octubre 1981

PRESENTADO POR EL ACADÉMICO NUMERARIO D. SIXTO RÍOS

Título abreviado: *Contraste de hipótesis y programación infinita*

Resumen

Se plantea el problema de contraste de hipótesis compuesta frente a alternativa simple como un problema de programación matemática infinita. Suponiendo que las restricciones de nivel pertenecen al espacio de Banach de las funciones continuas en el espacio paramétrico se encuentra como variable dual óptima una distribución a priori sobre docho espacio paramétrico. Cambiando la hipótesis de continuidad por la de integrabilidad de las funciones respecto a una medida σ-finita positiva se obtiene una función de densidad cuyo papel es decisivo en la determinación del test óptimo.

English title: *Testing a composite hypothesis against a simple alternative via infinite mathematical programming*

Summary

The duality theory of infinite mathematical programming is used to solve the problem of testing a composite hypothesis against a simple alternative. The optimal test and the optimal multiplier are obtained in the cases: *a*) the level constraints belongs to $C(\Theta)$ where Θ is a compact Hausdorff parameter space; *b*) the level constraints belongs to $L_1(\Theta, \nu)$, where ν is a σ-finite positive measure.

1. INTRODUCCION

El importante lema de Neyman y Pearson ha significado la base teórica para el estudio del contraste de hipótesis en la estadística matemática. A partir de él se han tratado de abordar problemas más generales, así como obtener condiciones necesarias y suficientes de existencia de tests más potentes. Una de las contribuciones de mayor interés en este sentido es la debida a G. B. Dantzig y A. Wald (1951), donde por primera vez se ponen de manifiesto las conexiones entre la estadística matemática y la teoría de la optimización.

El punto de partida de nuestro artículo es un resultado que aparece en el excelente libro de E. L. Lehmann (1959) sobre la distribución menos favorable en la teoría de tests. El propio Lehmann hace un comentario sobre la importancia que tiene la determinación de los multiplicadores, que en este problema «toman la forma de una distribución arbitraria» (*op. cit.*, pág. 91).

Sobre el mismo tema pueden consultarse el trabajo-resumen de O. Krafft (1970) que analiza una versión diferente del problema, así como el artículo de K. Isii (1964), que lo estudia dentro de un contexto más general y parte de cuyos resultados utilizaremos para obtener soluciones más concretas, en particular la existencia de una función de densidad a priori.

2. PROGRAMACION MATEMATICA INFINITA Y TESTS DE HIPOTESIS

Sea X un espacio muestral y $B(X)$ una σ-álgebra de subconjuntos de X. Sea Θ un espacio paramétrico y denotemos $P = \{P_\theta; \theta \in \Theta\}$ *al conjunto de distribuciones de probabilidad definidas en* $(X, B(X))$. Suponemos, además, que P está dominada por una medida σ-finita μ. Representaremos por $f(x;\theta)$ a la función de densidad de un elemento arbitrario $P_\theta \in P$ respecto a la medida μ. Sea $g(x)$ una función de densidad con respecto a μ que no pertenezca a la familia anterior.

El problema de contraste de hipótesis que nos proponemos estudiar es el siguiente:

2.1. Problema de contraste de hipótesis 1 (PCH1)

$$H_0: f(x;\theta),\ \theta \in \Theta$$
$$H_1: g(x)$$

donde la hipótesis nula es compuesta y la alternativa es simple.

La solución de este problema bajo el punto de vista de la teoría de Neyman y Pearson consiste en encontrar un test perteneciente a $\Phi = \{\phi: X \to [0,1]\}$ que maximice la potencia sujeto a que su nivel sea $\alpha \in (0,1)$. Formalmente esto puede expresarse como un problema de programación matemática infinita.

2.2. Problema de contraste de hipótesis 2 (PCH2)

Hallar ϕ^*, si existe, tal que

$$\int_X g(x)\,\phi^*(x)\,d\mu(x) = \sup_{\phi \in \Phi(\alpha)} \int_X g(x)\,\phi(x)\,d\mu(x)$$

donde

$$\Phi(\alpha) = \left\{\phi \in \Phi: \int_X f(x;\theta)\,\phi(x)\,d\mu(x) \leqslant \alpha,\ \forall\,\theta \in \Theta\right\}$$

Lehmann (1959), introduciendo una distribución a priori sobre el espacio paramétrico Θ, convierte el problema 2.1 en otro de contraste de hipótesis simple contra alternativa simple, siendo la solución de éste igual a la del primero bajo ciertas condiciones. O. Krafft (1970) formula un problema lineal dual del problema 2.2, cuyas variables son las distribuciones a priori sobre el espacio paramétrico; utilizando cadenas de desigualdades y una técnica debida a Fan (*vid.* Krafft, 1970), demuestra que se verifica un teorema de dualidad entre ambos; al test óptimo ϕ^*, solución de 2.2, viene asociada una distribución a priori λ^*, solución del problema dual, obteniendo una generalización del teorema de Lehmann.

En el apartado siguiente daremos un resultado paralelo al de Krafft, usando la teoría de dualidad de programación matemática infinita general debida a K. Isii (1964). Por este motivo resumimos a continuación algunos de los resultados de este autor que serán utilizados en los apartados siguientes.

Sea V un espacio vectorial, no necesariamente topológico, X un subconjunto convexo no vacío de V y Z un espacio vectorial topológico sobre los reales. Sea $K \subseteq Z$ un cono convexo.

2.3 Definición. En Z definimos un pseudoorden dado por

$$z_1, z_2 \in Z, \quad z_1 \geqq_K z_2 \quad \text{si y sólo si} \quad z_1 - z_2 \in K$$

Sea G una aplicación cóncava en el sentido del orden inducido por K con dominio en X y rango un subconjunto de Z, y sea F un funcional real cóncavo con dominio en X.

2.4. Problema de programación infinita (PM1)

Hallar $v^* \in V$, si existe, tal que

$$F(v^*) = \sup_{v \in S} F(v)$$

donde

$$S = \{v \in V\colon v \in X, G(v) \geqq_K 0\}$$

El cero representa el origen del espacio Z.

En orden a establecer condiciones necesarias y suficientes de óptimo para este problema se precisan las siguientes hipótesis.

2.5. Condición de regularidad A

El cono K tiene interior no vacío y el origen de Z es un punto interior de $G(X) - K$.

2.6. Condición de regularidad B

Existe un entorno E del origen de Z y una constante real M tal que para cualquier $z \in E$,

$$\sup\{F(v)\colon v \in X, G(v) \geqq_K z\} > M$$

No es difícil demostrar que la condición de regularidad A es equivalente a la siguiente condición de regularidad de Slater generalizada.

2.7. Condición de regularidad C (Slater generalizada)

Existe $v \in X$ tal que $G(v) >_K 0$ (i.e.: $G(v) \in \operatorname{int} K$).

2.8 Teorema. Bajo una cualquiera de las hipótesis A, B o C se verifica que

$$\begin{aligned} &\sup\{F(v)\colon v \in X, G(v) \geqq_K 0\} = \\ &= \inf\left\{\sup_{v \in X}\left[F(v) + z^*(G(v))\right]\colon z^* \in Z^*, z^* \geqq 0^*\right\} \end{aligned} \qquad [1]$$

en donde Z^* es el espacio dual de Z y $z^* \geqq 0^*$ significa que $z^*(z) \geqq 0$ para todo $z \in K$. Además, cuando el primer miembro de la igualdad [1] es finito, existe siempre un $z_0^* \geqq 0^*$ que alcanza el ínfimo del segundo miembro.

La condición necesaria y suficiente para que un punto $c \in X$, $G(v) \geqq_X 0$ alcance el supremo del primer miembro de [1] es que para algún $z^* \geqq 0^*$ se alcance el supremo entre corchetes del segundo miembro de [1] y que $z^*(G(v)) = 0$.

Demostración. Véase K. Isii (1964).

Cuando no se verifican ninguna de las condiciones de regularidad se puede emplear el teorema siguiente.

2.9. **Teorema.** Sea V un espacio vectorial topológico sobre los reales, X un subconjunto convexo de V y Z un espacio vectorial topológico localmente convexo sobre los reales, en el cual está definido un pseudoorden por medio de un cono cerrado convexo K. Sea G una aplicación continua cóncava con dominio X y rango un subconjunto de Z y F un funcional real cóncavo semicontinuo superiormente. Supongamos, asimismo, que existe un entorno E del origen de Z tal que $G^{-1}(\overline{E + K}) \cap X$ es un conjunto compacto en V. Entonces, si el origen pertenece al conjunto $G(X) - K$, se verifica [1]. En este caso, sin embargo, no siempre existe un z^* que alcance el ínfimo del segundo miembro, pero existe $v_0 \in X$, $G(v_0) \geqq_K 0$, que alcanza el supremo del primer miembro.

Demostración. Véase K. Isii (1964).

3. ESPACIO PARAMETRICO COMPACTO Y HAUSDORFF

Consideremos $\Phi = \{\phi: X \to [0, 1]\}$, es decir, el conjunto de los tests aleatorizados; trivialmente es convexo. Este conjunto va a desempeñar el papel de X en los teoremas anteriores.

Sea Θ un espacio compacto y Haussdorff. Denotamos por $C(\Theta)$ el conjunto de las funciones continuas definidas en Θ sobre los reales, dotado de la norma

$$r \in C(\Theta) \quad ; \quad \|r\| = \sup_{\theta} |r(\theta)|$$

De esta forma $C(\Theta)$ es un espacio de Banach. En este apartado el Z del teorema 2 será $C(\Theta)$.

El cono K viene definido por

$$K = \{r \in C(\Theta): r(\theta) \geqslant 0, \forall \theta \in \Theta\}$$

Es sencillo comprobar que K es un cono convexo y con interior no vacío en la topología inducida por la norma anterior.

Identificaremos con F y G del apartado anterior las expresiones siguientes:

$$F(\phi) = \int_X g(x)\, \phi(x)\, d\mu(x)$$

$$G(\phi) = \alpha - \int_X f(x; \theta)\, \phi(x)\, d\mu(x)$$

Claramente F es un funcional real lineal (y, por tanto, cóncavo). Supondremos que $G(\Phi) \subseteq C(\Theta)$; con ello es evidente que G es un funcional cóncavo en el sentido del pseudoorden inducido por K. Con estas notaciones el PMI 2.3 coincide con el PCH2 2.2.

3.1. Lema. Se verifican las condiciones de regularidad A, B y C.

Demostración. Trivial.

Estamos ahora en condiciones de aplicar el teorema 2 que nos permite escribir la siguiente igualdad:

$$\sup\left\{\int_X g(x)\,\phi(x)\,d\mu(x);\ \phi\in\Phi,\ \alpha-\int_X f(x;\theta)\,\phi(x)\,d\mu(x)\geq 0,\ \forall\theta\in\Theta\right\}=$$

$$=\inf\left\{\sup_{\phi\in\Phi}\left[\int_X g(x)\,\phi(x)\,d\mu(x)+r^*\left(\alpha-\int_X f(x;\theta)\,\phi(x)\,d\mu(x)\right)\right]:r^*\geqq 0^*;\ r^*\in C^*(\Theta)\right\}$$

El teorema de representación de Riesz expresa los funcionales $r^*\in C^*(\Theta)$ de la forma siguiente:

$$r^*(r)=\int_\Theta r(\theta)\,d\lambda(\theta)$$

donde $\lambda\in\mathrm{rca}(\Theta)$ espacio de las funciones de conjunto regulares numerablemente aditivas definidas en un σ-álgebra $B(\Theta)$. El cono polar K^* viene dado por

$$K^*=\left\{\lambda\in\mathrm{rca}(\Theta):\int_\Theta r(\theta)\,d\lambda(\theta)\geqslant 0,\ \forall r\in K\right\}$$

Además, las medidas $\lambda\in K^*$ son no negativas (Dunford y Schwartz, 1958, págs. 262-265).

El segundo miembro de [2], utilizando los resultados anteriores, puede ser escrito de la siguiente forma:

$$\inf_{\lambda\in K^*}\left\{\sup_{\Phi\in\alpha}\left[\int_X g(x)\,\phi(x)\,d\mu(x)+\int_\Theta\left(\alpha-\int_X f(x;\theta)\,\phi(x)\,d\mu(x)\right)d\lambda(\theta)\right]\right\}=$$

$$=\inf_{\lambda\in K^*}\left\{\sup_{\phi\in\Phi}\left[\alpha\lambda(\Theta)+\int_X\left(g(x)-\int_\Theta f(x;\theta)\,d\lambda(\theta)\right)\phi(x)\,d\mu(x)\right]\right\}\qquad[3]$$

El superior de la expresión anterior se alcanza para

$$\phi^*(x)=\begin{cases}1 & \text{si}\quad g(x)>\displaystyle\int_\Theta f(x;\theta)\,d\lambda(\theta)\\[2ex] 0 & \text{si}\quad g(x)<\displaystyle\int_\Theta f(x;\theta)\,d\lambda(\theta)\end{cases}\qquad[4]$$

y el interior para una medida $\lambda^* \in K^*$, con lo que la expresión [3] se convierte en

$$\int_{X+} g(x)\, d\mu(x) + \int_{\Theta} \left(\alpha - \int_{X+} f(x;\theta)\, d\mu(x)\right) d\lambda^*(\theta)$$

donde

$$X^+ = \left\{x \in X : g(x) > \int_{\Theta} f(x;\theta)\, d\lambda^*(\theta)\right\}$$

El teorema de K. Isii 2 nos asegura también que en esta situación

$$\int_{\Theta} \left(\alpha - \int_{X+} f(x;\theta)\, d\mu(x)\right) d\lambda^*(\theta) = 0$$

o lo que es igual

$$\alpha = \int_{X+} f(x;\theta)\, d\mu(x) \quad , \quad \forall \theta \in \Theta, \text{ c.s. } -\lambda^*$$

Obsérvese que el multiplicador de Lagrange λ^* asociado al PMI coincide con la idea expuesta por Lehmann.

Los resultados anteriores pueden resumirse en el teorema siguiente:

3.2. Teorema. Dado el problema de contraste de hipótesis en la forma 2.2 y bajo las condiciones especificadas en este apartado existe una distribución a priori sobre el espacio paramétrico que alcanza el ínfimo en [3]. La condición necesaria y suficiente para que un test de la forma [4] sea óptimo para el problema 2.2 es que exista $\lambda^* \in K^*$ tal que alcance el supremo en [3] y verifique [5].

4. EXISTENCIA DE FUNCION DE DENSIDAD A PRIORI

En orden a obtener la función de densidad a priori, si existe, se van a definir nuevos espacios de funciones que hagan las veces de V y Z del apartado 2.

Sea $(X, B(X))$ un espacio medible y supongamos que $B(X)$ tiene un conjunto numerable de generadores. Consideremos el espacio V de todas las medidas sobre $(X, B(X))$ engendradas por funciones medibles uniformemente acotadas y sea X el subconjunto de V de todas las medidas no negativas generadas por Φ respecto a la medida μ. Obviamente este conjunto es convexo no vacío; también es compacto con respecto a la convergencia $\phi_n \to \phi$ definida por la propiedad

$$\int_X m(x)\, \phi_n(x)\, d\mu(x) \to \int_X m(x)\, \phi(x)\, d\mu(x)$$

para todas las funciones $m(x)$ que son integrables μ (Lehmann, 1959, pág. 354). X puede identificarse con Φ de forma natural.

Sobre el espacio paramétrico Θ no supondremos ahora ninguna estructura topológica. Sea dada una medida σ-finita positiva ν definida en el espacio medible $(\Theta, B(\Theta))$. Ello permite introducir el espacio $L_1(\Theta, \nu)$ de todas las funciones medibles e integrables ν, dotado de la norma.

$$r \in L_1(\Theta, \nu) \quad ; \quad \|r\| = \int_\Theta |r(\theta)| \, d\nu(\theta)$$

que es un espacio de Banach. Este será ahora el espacio Z.

El cono

$$K = \{r \in L_1(\Theta, \nu) \colon r(\theta) \geqslant 0, \text{ c.s. } -\nu\}$$

define el pseudoorden usual en $L_1(\Theta, \nu)$.

4.1. Lema. K es cerrado y convexo.

Demostración. La convexidad es evidente. Para demostrar que es cerrado consideremos $\{r_n(\theta)\}_{n \in N}$, $r_n \in K$, tal que

$$\lim_{n \to \infty} r_n(\theta) = r(\theta)$$

Demostraremos que $r \in K$. Supongamos que existiese

$$\Theta_0 = \{\theta \colon r(\theta) < 0\} \quad \text{con} \quad \nu(\Theta_0) > 0$$

Puesto que $r_n(\theta) \geqslant 0$ c.s. $-\nu$ para todo n, entonces

$$r_n(\theta) - r(\theta) > 0$$

para todo n y para todo $\theta \in \Theta_0$, de donde se deduce que existe $\delta > 0$ tal que

$$\int_{\Theta_0} |r_n(\theta) - r(\theta)| \, d\nu(\theta) > \delta > 0$$

Esto contradice que r_n converja a r.

Sean $g(x)$ y $f(x; \theta)$ integrables con respecto a todo elemento de X para todo $\theta \in \Theta$ y llamemos de nuevo

$$F(\phi) = \int_X g(x) \, \phi(x) \, d\mu(x)$$

$$G(\phi) = \alpha - \int_X f(x; \theta) \, \phi(x) \, d\mu(x)$$

Tenemos que suponer que $G(\Phi) \subseteq L_1(\Theta, \nu)$. Es de notar que ésta es una de las situaciones en que no se verifican en general las condiciiones de regularidad de K. Isii anteriores. Sin embargo, si se cumplen las hipótesis del teorema 2.9; en efecto: F y G

son funcionales continuos en la topología, además son lineales; debido a que X es compacto, entonces $\overline{G^{-1}(E + K)} \cap X$ es compacto, siendo E un entorno del origen; también el origen pertenece a $G(X) - K$. Así pues, podemos escribir que

$$\sup \left\{ \int_X g(x)\, \phi(x)\, d\mu(x);\ \phi \in \Phi, \right.$$

$$\left. \alpha - \int_X f(x;\theta)\, \phi(x)\, d\mu(x) \geqslant 0,\ \forall \theta \in \Theta,\ \text{c.s.}\ -v \right\} =$$

$$= \inf \left\{ \sup_{\phi \in \Theta} \left[\int_X g(x)\, \phi(x)\, d\mu(x) + r^* \left(\alpha - \int_X f(x;\theta)\, \phi(x)\, d\mu(x) \right) \right] : \right.$$

$$\left. r^* \geqq 0^*,\ r^* \in L_1^*(\Theta, v) \right\}$$

El teorema de representación de Riesz identifica los $r^* \in L_1^*(\Theta, v)$ con los $\ell \in L_\infty(\Theta, v)$ mediante la identidad

$$r^*(r) = \int_\Theta r(\theta)\, \ell(\theta)\, dv(\theta)$$

para todo $r \in L_1(\Theta, v)$ (Dunford y Schwartz, 1958, pág. 289).

El cono polar es

$$K^* = \left\{ \ell \in L_\infty(\Theta, v) :\ \int_\Theta r(\theta)\, \ell(\theta)\, dv(\theta) \geqslant 0,\ \forall r \in K \right\}$$

4.2. Lema. Si $\ell \in K^*$, entonces es no negativa c.s. $-v$.

Demostración. Supongamos que existiese un conjunto $A \in B(\Theta)$ con $v(A) > 0$ y $\ell(\theta) < 0$ para $\theta \in A$. Sea $\chi_A(\theta)$ el indicador del conjunto A y $\chi_\phi(\theta)$ el indicador del vacío; entonces:

i) $\chi_A(\theta)$, $\chi_A(\theta) \in L_1(\Theta, v)$
ii) $\chi_A(\theta) \geqslant \chi_\phi(\theta)$ para todo $\theta \in \Theta$

Además, como $\chi_A - \chi_\phi \in K$, al ser $\ell \in K^*$ se verifica

$$\int_\Theta (\chi_A(\theta) - \chi_\phi(\theta))\, \ell(\theta)\, dv(\theta) \geqslant 0$$

esto implica que

$$\int_A \ell(\theta)\, dv(\theta) \geqslant 0$$

pero por la hipótesis supuesta es

$$\int_A \ell(\theta)\, d\nu(\theta) < 0$$

llegando a una contradicción.

Si definimos

$$\lambda(A) = \int_A \ell(\theta)\, d\nu(\theta)$$

ℓ es la densidad a priori de la medida λ.

En el miembro derecho de [6] podemos escribir:

$$\inf_{\ell\in K^*} \left\{ \sup_{\phi\in\Phi} \left[\int_X g(x)\,(x)\, d\mu(x) + \right.\right.$$

$$\left.\left. + \int_\Theta \left(\alpha - \int_X f(x;\theta)\, \phi(x)\, d\mu(x) \right) \ell(\theta)\, d\nu(\theta) \right] \right\} =$$

$$= \inf_{\ell\in K^*} \left\{ \sup_{\phi\in\Phi} \left[\alpha\lambda(\Theta) + \int_X \left(g(x) - \int_\Theta f(x;\theta)\, \ell(\theta)\, d\nu(\theta) \right) \phi(x)\, d\mu(x) \right] \right\}$$

El superior se alcanza para

$$\phi^*(x) = \begin{cases} 1 & \text{si} \quad g(x) > \int_\Theta f(x;\theta)\, \ell(\theta)\, d\nu(\theta) \\ 0 & \text{si} \quad g(x) < \int_\Theta f(x;\theta)\, \ell(\theta)\, d\nu(\theta) \end{cases} \qquad [8]$$

Podemos escribir [7] de la forma

$$\inf_{\ell\in K^*} \left\{ \int_{X_\ell} g(x)\, d\mu(x) + \int_\Theta \left(\alpha - \int_{X_\ell} f(x;\theta)\, d\mu(x) \right) \ell(\theta)\, d\nu(\theta) \right\}$$

donde

$$X_\ell^+ = \left\{ x \in X \colon g(x) > \int_\Theta f(x;\theta)\, \ell(\theta)\, d\nu(\theta) \right\}$$

Todo lo anterior permite enunciar el siguiente resultado:

4.3. Teorema. Supuestas las hipótesis de este apartado si existe una función de densidad a priori ℓ^* de una medida λ^* que alcance el ínfimo de [7] la solución del problema 2.2 tiene la forma [8].

BIBLIOGRAFIA

DANTZIG, G. B., y WALD, A. (1951): «On the fundamental lemma of Neymann and Pearson», *Ann. Math. Stat.*, *22*, p. 87-93.

DUNFORD, F., y SCHWARTZ, J. T. (1958): *Linear Operators*, Part I, Interscience, Nueva York.

ISII, K. (1964): «Inequalities of the types of Chebyshev and Crámer-Rao and Mathematical Programming», *Ann. Inst. Stat. Math.*, *16*, págs. 277-293.

KRAFFT, O. (1970): «Programming methods in statistics and probability theory», en J. B. Rosen, O. L. Mangasarian y K. Ritter (eds.), *Nonlinear Programming*, Academic Press, Nueva York.

LEHMANN, E. L. (1959): *Testing Statistical Hypotheses*, John Wiley, Nueva York.

Departamento de Estadística
Facultad de Ciencias
Universidad de Granada