TRABAJOS DE ESTADÍSTICA Y
DE INVESTIGACIÓN OPERATIVA
Vol. 35. Núm. 2, 1984, pp. 215 a 221

SOBRE EL TEST DE MÁXIMA POTENCIA PARA HIPÓTESIS NULA SIMPLE CONTRA HIPÓTESIS ALTERNATIVA SIMPLE

Pedro José Paúl Escolano
Departamento de Matemáticas
de la E.T.S. Ingenieros Industriales de Sevilla

RESUMEN

El propósito de este trabajo es dar una construcción explícita del test de máxima potencia para un constraste de hipótesis paramétrico en el que tanto la hipótesis nual como la hipótesis alternativa son simples; utilizando para ello técnicas de Análisis Funcional y de Programación Matemática.

PLANTEAMIENTO DEL PROBLEMA Y NOMENCLATURA (1)

Sea una variable aleatoria unidimensional con función de distribución $F(x; \theta)$. Supongamos conocida la forma funcional de $F(x; \theta)$ salvo el parámetro θ. Con una muestra aleatoria simple de tamaño n: $X = (x_1, x_2, \ldots, x_n)$, cuya función de verosimilitud es:

$$dF(X; \theta) = \prod_{i=1}^{n} dF(x_i; \theta)$$

(*) Recibido. Mayo. 1982

vamos a contrastar la hipótesis nula:

$$H_0: \ \theta = \theta_0$$

contra la hipótesis alternativa:

$$H_1: \ \theta = \theta_1$$

Si representamos el funcional *tamaño* de una función

$$\rho \in L^1 \quad (R^n, dF(X; \theta_0))$$

por:

$$T(\rho) = E[\rho/\theta = \theta_0] = \int_{R^n} \rho \, dF(X; \theta_0)$$

y el funcional *potencia* de una función ψ del conjunto $L^1(R^n, dF(X; \theta_1))$ por:

$$P(\psi) = E[\psi/\theta = \theta_1] = \int_{R^n} \psi \, dF(X; \theta_1)$$

entonces la clase de los *test de tamaño* α para el contraste anterior con $0 < \alpha < 1$ es:

$$\phi_\alpha = \{ \rho\colon \ R^n \to R/\rho \text{ es medible; } 0 \leqslant \rho(X) \leqslant 1; \quad T(\rho) \leqslant \alpha \}$$

Entonces, el problema de hallar el test más potente de tamaño α para contrastar H_0 contra H_1, puede formularse:

Maximizar $P(\rho)$
Sujeto a $\rho \in \phi_\alpha$

o, más desarrollado:

Maximizar $P(\rho)$
Sujeto a $T(\rho) \leqslant \alpha$
$0 \leqslant \rho(X) \leqslant 1$ para todo $X \in R^n$

PROPOSICIONES BÁSICAS

Proposición 1. ϕ_α es un conjunto convexo y cerrado en $L^1(R^n, dF(X; \theta_0))$.

DEMOSTRACIÓN. Si llamamos ν a la medida de Borel-Stieltjes inducida en R^n por la función de distribución de probabilidad $F(X; \theta_0)$, $\nu(R^n) = 1$; entonces:

a) Si $\rho \in \phi_\alpha \Rightarrow 0 \leqslant \rho(X) \leqslant 1 \quad \forall X \in R^n \Rightarrow \rho$ está acotada: $\rho \in L^\infty(R^n, dF(X, \theta_0))$. Pero $\nu(R^n) = 1$ y en ese caso: $L^\infty(R^n, dF(X; \theta_0)) \subseteq \subseteq L^1(R^n, dF(X; \theta_0)) \Rightarrow \rho \in L^1(R^n, dF(X; \theta_0))$ de donde $\phi_\alpha \subseteq L^1(R^n, dF(X; \theta_0))$.

b) Si $\rho_1, \rho_2 \in \phi_\alpha$ y $0 \leqslant \lambda \leqslant 1$ al ser ρ_1 y ρ_2 medibles, es $\lambda\rho_1 + (1 - \lambda)\rho_2$ medible.

Por otro lado:

$$T(\lambda\rho_1 + (1 - \lambda)\rho_2) = \lambda T(\rho_1) + (1 - \lambda)T(\rho_2) \leqslant \lambda\alpha + (1 - \lambda)\alpha = \alpha$$

y por último:

$$0 \leqslant \lambda\rho_1(X) + (1 - \lambda)\rho_2(X) = (\lambda\rho_1 + (1 - \lambda)\rho_2)(X) \leqslant \lambda + (1 - \lambda) = 1$$

para todo $X \in R^n$.

Por tanto ϕ_α es convexo.

c) En $L^1(R^n, dF(X; \theta_0))$ tenemos definida la norma:

$$||\psi||_1 = \int_{R^n} |\psi| \, dF(X; \theta_0)$$

Sea $(\rho_n)_{n=1}^\infty \subseteq \phi_\alpha$ una sucesión convergente a ρ, veamos que $\rho \in \phi_\alpha$.

Como (ρ_n) converge a ρ en $\|\cdot\|_1$, existe una subsucesión (ρ_{n_k}) que converge a ρ casi uniformemente, por tanto converge en casi todo, de donde ρ es medible.

Luego, salvo $Z \subseteq R^n$ con $\nu(Z) = 0$ es

$$\lim_{n \to \infty} \rho_n(X) = \rho(X) \Rightarrow \rho(X) \in [0, 1] \quad \forall X \in R^n \backslash Z.$$

Eligiendo un representante de ρ que verifique $0 \leqslant \rho(X) \leqslant 1$ para $X \in Z$ tendremos:

$$0 \leqslant \rho(X) \leqslant 1 \quad \forall X \in R^n$$

Además $0 \leqslant T(\rho_n) = \|\rho_n\|_1 \leqslant \alpha$ para todo $n \in N$ por ser ρ_n positivas, como se tiene

$$\lim_{n \to \infty} \|\rho_n - \rho\|_1 = 0 \quad \text{es} \quad 0 \leqslant T(\rho) \leqslant \alpha$$

por ser ρ también positiva. □

Proposición 2. El funcional P: $\phi_\alpha \to R$ definido:

$$P(\rho) = \int_{R^n} \rho \, dF(X; \theta_1)$$

verifica:

$P(\phi_\alpha)$ es un cerrado contenido en $[0, 1]$.

DEMOSTRACIÓN.

$$\rho \in \phi_\alpha \Rightarrow 0 \leqslant \rho(X) \leqslant 1 \quad \text{para todo} \quad X \in R^n \Rightarrow 0 \leqslant \int_{R^n} \rho \, dF(X: \theta_1) \leqslant$$
$$\leqslant \int_{R^n} dF(X; \theta_1) = 1 \Rightarrow 0 \leqslant P(\rho) \leqslant 1 \Rightarrow P(\phi_\alpha) \subseteq [0, 1]$$

Consideremos el espacio de medidas $(R^n, \mathcal{B}, \mu)$ siendo $\mu = dF(X; \theta_0) + dF(X; \theta_1)$, entonces $dF(X; \theta_0)$ y $dF(X; \theta_1)$ son absolutamente continuas respecto de μ; por tanto, existen dos funciones $g, h \in L^1(R^n, \mu)$ tales que:

$$dF(X; \theta_0) = g \, d\mu \text{ y}$$
$$dF(X; \theta_1) = h \, d\mu$$

(como se sigue del teorema de Radon-Nikodym: (2)) con lo que

$$P(\rho) = \int_{R^n} \rho \, dF(X; \theta_1) = \int_{R^n} \rho h \, d\mu \text{ y}$$

$$T(\rho) = \int_{R^n} \rho \, dF(X; \theta_0) = \int_{R^n} \rho g \, d\mu$$

verificándose además que P es un funcional lineal y continuo de $L^\infty(R^n, \mu)$ en $R(|P(\rho)| \leqslant \|\rho\| \cdot \|h\|_1)$.

Dotemos a $L^\infty(R^n, \mu)$ de la topología débil $\sigma[L^\infty(R^n, \mu), L^1(R^n, \mu)]$ que denotamos $\sigma[L^\infty, L^1]$ entonces P es «$\sigma[L^\infty, L^1]$-continuo». Sea además $\phi_\alpha^0 = \{f\colon |\langle \rho, f \rangle| \leqslant 1 \text{ para todo } \rho \in \phi_\alpha\}$ el polar de ϕ_α en la dualidad $[L^\infty, L^1]$.

Si $\|f\|_1 \leqslant 1 \Rightarrow \int_{R^n} |f| \, d\mu \leqslant 1 \Rightarrow \left| \int_{R^n} \rho f \, d\mu \right| \leqslant \int_{R^n} |f| \, d\mu \leqslant 1 \quad \forall \rho \in \phi_\alpha$

$\Rightarrow f \in \phi_\alpha^0$. O sea, la bola $B(0, 1) \subseteq \phi_\alpha^0$, podemos aplicar el teorema de Alaoglu-Bourbaki (3) y concluir que ϕ_α^{00} es $\sigma[L^\infty, L^1]$-compacto.

Probemos que ϕ_α es $\sigma[L^\infty, L^1]$-cerrado: como $\bar{\phi}_\alpha^\sigma = \bar{\phi}_\alpha$ (3), si $(\rho_i)_{i\in I}$ es una red convergente a ρ en la topología $\sigma[L^\infty, L^1]$ como $g \in L^1(R^n, \mu)$ es:

$$\int_{R^n} |\rho_i - \rho| \, dF(X; \theta_0) = \int_{R^n} |\rho_i - \rho| g \, d\mu \to 0$$

y, por la Proposición 1, es $\rho \in \phi_\alpha \Rightarrow \phi_\alpha$ es $\sigma[L^\infty, L^1]$-cerrado, ϕ_α^{00} $\sigma[L^\infty, L^1]$-compacto y $\phi_\alpha \subseteq \phi_\alpha^{00} \Rightarrow \phi_\alpha$ es $\sigma[L^\infty, L^1]$-compacto $\Rightarrow P(\phi_\alpha)$ es cerrado en $[0, 1]$. □

CONSTRUCCIÓN DEL TEST MÁS POTENTE

Vamos a aplicar técnicas de programación al problema:

$$\begin{array}{l} \text{Maximizar } P(\rho) \\ \text{Sujeto a } \rho \in \phi_\alpha \end{array}$$

Lema 1. Existe un test ρ^* que alcanza el máximo del problema anterior.

DEMOSTRACIÓN. Al ser $P(\phi_\alpha)$ cerrado en $[0, 1]$ tiene máximo $\Rightarrow$ existe un test ρ^* que alcanza dicho máximo. □

Lema 2. El test ρ^* obtenido en el lema anterior verifica $T(\rho^*) = \alpha$.

DEMOSTRACIÓN. Los funcionales T y P son lineales, por tanto son Gateaux-diferenciables y su G-diferencial coincide con ellos. R tiene un cono positivo $[0, +\infty)$, así que podemos aplicar el teorema de Kuhn y Tucker. (4), y entonces para la solución, ρ^*, del problema

$$\begin{array}{ll} \text{Maximizar } P(\rho) & \\ \text{Sujeto a} & T(\rho) \leqslant \alpha \\ & 0 \leqslant \rho(X) \leqslant 1 \quad \text{para todo} \quad X \in R^n \end{array}$$

existen tres multiplicadores de Lagrange: $\lambda_1, \lambda_2, \lambda_3 \in R$ tales que:

$$\lambda_1(T(\rho^*) - \alpha) + \lambda_2(\rho^* - 1) + \lambda_3(-\rho^*) = 0$$
$$\lambda_i \geqslant 0 \quad i = 1, 2, 3, \quad \text{y no todos nulos.}$$

Si $\lambda_1 = 0$ resulta $\lambda_2(\rho^* - 1) + \lambda_3(-\rho^*) = 0$, pero $\rho^* - 1$ y $-\rho^*$ son números negativos no nulos a la vez por tanto, al ser λ_2 y λ_3 positivos, alguno estrictamente, llegaríamos a unan contradicción. Así pues, $\lambda_1 > 0$.

Si $T(\rho^*) < \alpha$, $\lambda_1(T(\rho^*) - \alpha) + \lambda_2(\rho^* - 1) + \lambda_3(-\rho^*) < 0$ luego debe ser $T(\rho^*) = \alpha$ y, entonces, nos queda:

$$\lambda_2(\rho^* - 1) = \lambda_3(-\rho^*) = 0.$$

$\rho^*(X)$ no puede ser siempre igual a 1 pues se tendría $T(\rho^*) = 1 > \alpha$ así pues existe $Y \in R^n$ tal que $\rho^*(Y) - 1 < 0 \Rightarrow \lambda_2 = 0$. Asimismo si $\rho^*(X) = 0$ para todo $X \in R^n$ es $T(\rho^*) = 0 < \alpha$ contra $T(\rho^*) = \alpha > 0$, entonces existe $Z \in R^n$ tal que $\rho^*(Z) > 0$ de donde $\lambda_3 = 0$.

Pero el teorema de Kuhn y Tucker también asegura que ρ^* es también solución del problema:

$$\begin{array}{ll} \textbf{Maximizar} & H(\rho) = P(\rho) - \lambda T(\rho) \\ \textbf{Sujeto a} & T(\rho) = \alpha \\ & 0 \leqslant \rho(X) \leqslant 1 \quad \text{para todo} \quad X \in R^n \end{array}$$

Notamos:

$$f_0(X)\,dX = dF(X;\theta_0)$$
$$f_1(X)\,dX = dF(X;\theta_1)$$

(f_i serán funciones de densidad o de probabilidad según la variable aleatoria sea continua o discreta).

Entonces:

$$H(\rho) = \int_{R^n} \rho(f_1 - \lambda f_0)\,dX$$

Si $f_1(X) > \lambda f_0(X)$, ρ^* debe tomar el mayor valor posible: $\rho^*(X) = 1$.

Si $f_1(X) < \lambda f_0(X)$, ρ^* debe tomar el menor valor posible: $\rho^*(X) = 0$.

Si $f_1(X) = \lambda f_0(X)$, ρ^* puede tomar cualquier valor en el intervalo $[0, 1]$. Veamos que existe un valor $\gamma \in [0, 1]$ tal que si $\rho^*(X) = \gamma$ para todo $X / f_1(X) = \lambda f_0(X) \neq 0$ es $T(\rho^*) = \alpha$ (hemos supuesto $f_0(X) \neq 0$ porque si no fuera así en este último caso no influye el valor $\rho(X)$ ni en el funcional P ni en el T)

$$\alpha = T(\rho^*) = \int_{R^n} \rho^* f_0(X)\,dX =$$

$$= \text{Prob}\{X: f_1(X) > \lambda f_0(X)/\theta = \theta_0\} + \text{Prob}\{X: f_1(X) = \lambda f_0(X)/\theta = \theta_0\}$$
$$= 1 - \text{Prob}\{X:\ (f_1(X)/f_0(X)) \leqslant \lambda;\quad \theta = \theta_0\} +$$
$$+ \gamma\,\text{Prob}\{X:\ (f_1(X)/f_0(X)) = \lambda;\quad \theta = \theta_0\}$$

Si existe $\lambda = \lambda_0$ tal que:

$$\text{Prob}\{X:\ (f_1(X)/f_0(X)) \leqslant \lambda_0;\quad \theta = \theta_0\} = 1 - \alpha$$

se toma $\gamma = 0$. (Esto ocurre con funciones de distribución continuas).

Si no existe el valor anterior sí existe otro, que seguimos notando λ_0 que verifica:

$$\text{Prob}\{X:\ (f_1(X)/f_0(X)) < \lambda_0;\quad \theta = \theta_0\} \leqslant 1 - \alpha <$$
$$< \text{Prob}\{X:\ (f_1(X)/f_0(X)) \leqslant \lambda_0;\quad \theta = \theta_0\}$$

y tomamos:

$$\gamma = \frac{\text{Prob}\{X:\ (f_1(X)/f_0(X) \leqslant \lambda_0;\quad \theta = \theta_0\} - (1 - \alpha)}{\text{Prob}\{X:\ (f_1(X)/f_0(X)) = \lambda_0;\quad \theta = \theta_0\}}$$

Así que el test de máxima potencia buscado es:

$$\rho^*(X) = \begin{cases} 1 & \text{si}\quad f_1(X) > \lambda_0 f_0(X) \\ \gamma & \text{si}\quad f_1(X) = \lambda_0 f_0(X) \\ 0 & \text{si}\quad f_1(X) < \lambda_0 f_0(X) \end{cases}$$

elegidos γ y λ_0 como se ha expuesto. $\square$

BIBLIOGRAFÍA

(1) Graybill, F. A. y Mood, A. M. (1963): *Introduction to the theory of Statistics*. Ed. MacGraw-Hill Book Co. Inc.

(2) Halmos, P. (1970): *Measure Theory*. Ed. Springer-Verlag.

(3) Kothe, G. (1969): *Topological vector spaces, I*. Ed. Springer-Verlag.

(4) Luenberger, D. (1968): *Optimization by vector spaces methods*. Ed. John Wiley & Sons.