El principio de invariancia

Por VICENTE QUESADA PALOMA y MANUEL DEL RIO BUENO

Instituto Nacional de Estadística.
Departamento de Estadística e I. O. Facultad de Ciencias
de la Universidad Complutense.
Departamento de Matemáticas.
Facultad de Ciencias de la Universidad Autónoma.

ELEMENTOS DE UN PROBLEMA DE DECISION

Los elementos de un problema de decisión son: un espacio paramétrico Θ, un conjunto A de acciones o decisiones y una función de pérdida $L(\theta, a) \geqq 0$, definida sobre $\Theta \times A$. En los problemas de decisión estadísticos es posible observar una variable aleatoria **X** que toma valores en el espacio muestral (**X**, **A**) y cuya distribución depende del valor $\theta \in \Theta$; debido a ello nos referiremos a la familia $\mathbf{P} = \{P_\theta;\ \theta \in \Theta\}$ de distribuciones de X.

Una vez observado $\mathbf{X} = x$ consideraremos dos formas de actuar. Una de ellas tiene en cuenta el conjunto D de las funciones definidas en **X** y con valores en A, llamadas reglas de decisión no aleatorizadas; una vez elegida $d \in D$, se tomará la acción $d(x) \in A$. La segunda, que comprende como caso particular la ya expuesta, tiene en cuenta el conjunto Π de las funciones definidas en **X** y cuyos valores son distribuciones de probabilidad sobre A; con este fin supondremos que hemos seleccionado en A una σ-álgebra **F** que contiene sus puntos individuales. En este caso, una vez elegida $\pi \in \Pi$, tomaremos una acción o un conjunto de acciones de A mediante la distribución $\pi(x)$, que denotaremos en lo sucesivo por π_x.

A los elementos de Π les llamaremos reglas de decisión aleatorizadas. (Hacemos notar que en la bibliografía habitual estas reglas reciben el nombre de reglas de comportamiento, reservándose el adjetivo aleatorizadas para aquellas obtenidas mediante otro tipo de aleatorización. En nuestro estudio, al considerar un único método de aleatorización, no habrá lugar a confusión.)

Para comparar las diferentes reglas de decisión consideraremos las siguientes funciones de riesgo:

$$R(\theta, d) = E_\theta \{ L[\theta, d(x)] \} = \int L[\theta, d(x)]\, d P_\theta(x) \quad ; \quad d \in D, \theta \in \Theta$$
$$R(\theta, \pi) = E_\theta \{ L(\theta, \pi_x) \} = \int L(\theta, \pi_x)\, d P_\theta(x) \quad ; \quad \pi \in \Pi \quad ; \quad \theta \in \Theta$$

estando definida $L(\theta, \pi_x)$ del siguiente modo:

$$L(\theta, \pi_x) = \int L(\theta, a)\, d \pi_x(a)$$

Supondremos siempre que las clases D y Π están formaaas por aquellas reglas para las cuales las anteriores funciones existen y son finitas.

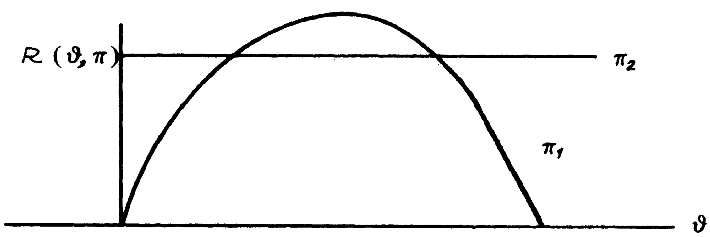
Una regla de decisión óptima sería aquella que proporcionara el mínimo riesgo, independientemente del valor del parámetro; sin embargo, esta situación se presenta en pocos casos. Para orientar el problema de modo matemático se pueden seguir los siguientes caminos: 1) Restringir el criterio de óptimo, introduciendo una cierta ordenación entre las reglas de decisión (la introducción de los principios Bayes o minimax conduce a esta alternativa); 2) tratar de reducir el problema sin pérdida de información importante, por ejemplo, mediante el criterio de dominancia o la consideración de estadísticos suficientes; 3) restringir la clase de las reglas a una clase menor, considerando principios como el de invariancia.

Nos proponemos considerar el principio de invariancia y estudiar, en una primera aproximación, sus relaciones con los caminos indicados en 1) y 2).

Recordemos lo que se entiende por principio minimax y por criterio de dominancia.

Si no conocemos la distribución *a priori* sobre Θ podemos ordenar las reglas de decisión basándonos en el criterio pesimista de asociar a cada regla la mayor pérdida que sea posible alcanzar, $\sup_{\Theta} R(\theta, d)$ o $\sup_{\Theta} R(\theta, \pi)$; de modo general, una regla óptima será aquella $\pi^* \in \Pi$ al que $\sup_{\theta} R(\theta, \pi^*) = \inf_{\pi} \sup_{\theta} R(\theta, \pi)$; a tal regla se le llama regla de decisión minimax.

Las reglas minimax proporcionan la mayor protección posible frente a la situación más adversa. El principio minimax no siempre es plausible, como muestra la figura siguiente.



Una regla $\pi_1 \in \Pi$ se dice que es tan buena como $\pi_2 \in \Pi$ si

$$R(\theta, \pi_1) \leq R(\theta, \pi_2) \qquad \forall\, \theta \in \Theta$$

Se dice que $\pi_1 \in \Pi$ es mejor que $\pi_2 \in \Pi$ o que π_1 domina a π_2 si

$$R(\theta, \pi_1) \leq R(\theta, \pi_2) \qquad \forall\, \theta \in \Theta$$

y

$$R(\theta_0\ \pi_1) < R(\theta_0\ \pi_2)$$

para algún $\theta_0 \in \Theta$.

Una regla de decisión se dice que es admisible si no está dominada.

PRINCIPIO DE INVARIANCIA

Si un problema de decisión es simétrico respecto a algunos valores del parámetro, parece razonable exigir que se actúe de modo simétrico respecto de esos valores. La expresión matemática de la simetría es el concepto de invariancia bajo transformaciones.

El principio de invariancia lleva asociados grupos de transformaciones en los tres espacios sobre los que se construye la teoría de la decisión: el espacio paramétrico Θ, el espacio de acciones A, y el espacio muestral X, que suponemos es un subconjunto del espacio euclídeo n-dimensional R_n.

Sea G un grupo de transformaciones medibles de X en sí mismo, siendo la operación del grupo la composición: Si g_1 y $g_2 \in G$, $g_2 g_1$ es la transformación que a $x \longrightarrow g_2 g_1(x)$. El que G sea grupo implica que todas las transformaciones habrán de ser inyectivas y sobre. Por otra parte, el que sean transformaciones medibles asegura que si X es variable aleatoria, g (X) también lo sea.

Definición 1

La familia de distribuciones $\{P_\theta \ ; \ \theta \in \Theta\}$ se dice invariante por el grupo G, si para $\forall\, g \in G$ y $\forall\, \theta \in \Theta$ existe un único $\theta' \in \Theta$ tal que la distribución de $g(x)$ viene dada por $P_{\theta'}$, siempre que la distribución de X venga dada por P_θ:

Plantearemos $\theta' = \bar{g}(\theta)$ y obtendremos la importante fórmula

$$P_\theta \{g(X) \in B\} = P_{\bar{g}(\theta)} \{X \in B\} \qquad \forall\, B \in \mathrm{A}$$

Identificabilidad.—Un parámetro θ se dice identificable para una familia de distribuciones $\{P_\theta \, ; \, \theta \in \Theta\}$ si $\theta \neq \theta' \Longrightarrow P_\theta \neq P_{\theta'}$. Si en la definición anterior el parámetro es identificable, la unicidad de θ' se satisface automáticamente.

Recíprocamente, si $\{P_\theta \, ; \, \theta \in \Theta\}$ es invariante bajo G, la unicidad de θ' implica que θ es identificable.

Lema 1

Si una familia $\{P_\theta ; \theta \in \Theta\}$ es invariante por G, entonces $\bar{G} = \{\bar{g} ; g \in G\}$ es un grupo de transformaciones de Θ en sí mismo.

Demostración.—Si la distribución de X es P_θ, la de $g_1(\theta)$ es $P_{\bar{g}(\theta)}$; de aquí la de $g_2[g_1(x)] = g_2 g_1(x)$ será $P_{\bar{g}_2[\bar{g}_1(\theta)]}$, pero también será la de $P_{\overline{g_2 g_1}(\theta)}$ por ser G grupo. Por la unicidad resulta que $\bar{g}_2[\bar{g}_1(\theta)] = \overline{g_2 g_1}(\theta)$, lo cual implica que $\bar{G}$ es cerrado por la composición.

Si e es la identidad en G, $\bar{e}$ es la identidad en $\bar{G}$, y tomando $g_2 = g_1^{-1}$ tenemos $\bar{g}_1^{-1}\bar{g}_1 = \bar{e}$ y, por tanto, $\overline{g_1^{-1}} = \bar{g}_1^{-1}$, y el enunciado queda probado.

Observemos que las transformaciones $\bar{g}$ son biyectivas, pues si $\theta' \in \Theta$ y X tiene la distribución P_θ, la distribución de $g^{-1}(x)$ será P_θ, siendo $\theta = \bar{g}^{-1}(\theta^1)$; por tanto, $(\bar{g})^{-1}(\theta') = \theta$ de donde $\theta' = \bar{g}(\theta)$. Por otra parte, si $\bar{g}(\theta_1) = \bar{g}(\theta_2)$, tendremos que

$$P_{\bar{g}(\theta_1)}(X \in B) = P_{\bar{g}(\theta_2)}(X \in B) \qquad \forall\, B \in A$$

lo que implica que

$$P_{\theta_1}[g(x) \in B] = P_{\theta_2}[g(x) \in B] \Longrightarrow P_{\theta_1}(B) = P_{\theta_2}(B)$$

y por la identificabilidad, $\theta_1 = \theta_2$.

Definición 2

Diremos que un problema de decisión (Θ, A, L) con variable aleatoria observable X, de distribución perteneciente a la familia $\{P_\theta \; ; \; \theta \in \Theta\}$ es invariante por el grupo G, si la familia $\{P_\theta \; ; \; \theta \in \Theta\}$ es invariante por G y si la función de pérdida L es invariante por G en el sentido de que para todo $g \in G$ y $a \in A$ existe un único $a' \in A$ tal que $L(\theta, a) = L[\bar{g}(\theta), a']$, $\forall\, \theta \in \Theta$.

A este a' determinado por g y a lo indicaremos por $\tilde{g}(a)$.

Lema 2

Si un problema de decisión es invariante para G, entonces $\tilde{G} = \{\tilde{g} \; ; \; g \in G\}$ es un grupo de transformaciones de A en sí mismo.

Demostración.—Hemos de demostrar que $\widetilde{g_2 g_1} = \tilde{g}_2 \tilde{g}_1$ dentro del conjunto G. Utilizando el resultado del lema anterior tendremos:

$$L(\theta, a) = L[\bar{g}_1(\theta), \tilde{g}_1(a)] = L[\bar{g}_2 \bar{g}_1(\theta), \tilde{g}_2 \tilde{g}_1(a)] =$$

$$= L[\overline{g_2 g_1}(\theta), \widetilde{g_2 g_1}(a)] \Longrightarrow \widetilde{g_2 g_1}(a) = \tilde{g}_2 \tilde{g}_1(a)$$

con lo cual $\widetilde{g_2 g_1} = \tilde{g}_2 \tilde{g}_1$ por la unicidad del transformado.

Tenemos, pues, que $\tilde{G}$ es cerrado bajo la composición. Por otra parte, si e es la identidad en G y $\tilde{e}$ es la identidad en $\tilde{G}$, tomando $g_2 = g_1^{-1}$ tendremos

$$(\widetilde{g_1^{-1}})\,\tilde{g}_1 = \tilde{e}$$

y, por tanto,

$$(\widetilde{g_1^{-1}}) = (\tilde{g}_1)^{-1}$$

Si un problema de decisión expresado en términos de $X, \theta \in \Theta$ y $a \in A$, es invariante bajo G, el problema transformado expresado en términos de $x' = g(x)$, $\theta' = \bar{g}(\theta) \in \Theta$ y $a' = \tilde{g}(a)$ es formalmente el mismo problema inicial, pues $L(\theta, a) = L(\theta', a')$; todo ocurre como si $g \in G$ efectuara un cambio de coordenadas en los elementos. Si al observar $X = x$ se toma la acción $a = d(X)$ y al observar $X' = x'$ se toma la acción $a' = d(x')$, cuyo nombre en el problema inicial es $\tilde{g}^{-1}[d(x')]$, al ser la acción a tomar independiente de las coordenadas elegidas, se debe tener $d(x) = g^{-1}\{d[g(x)]\}$, o bien, $\tilde{g}[d(x)] = d[g(x)]$; este razonamiento nos justifica la siguiente definición:

Definición 3

Dado un problema de decisión (Θ, A, L) con variable aleatoria observable X de distribución perteneciente a la familia $\{P_\theta, \theta \in \Theta\}$, invariante bajo G, decimos que la regla no aleatorizada $d \in D$ es invariante bajo G si se tiene que:

$$d[g(x)] = \tilde{g}[d(x)] \quad ; \quad \forall\, x \in \mathbf{X} \quad ; \quad \forall\, g \in G.$$

Análogamente, se dice que la regla aleatorizada $\pi \in \Pi$ es invariante bajo G si se tiene que

$$\pi_{g(x)} = \pi_x \tilde{g} \quad ; \quad \forall\, x \in \mathbf{X} \quad ; \quad \forall\, g \in G$$

siendo $\pi_x \tilde{g}$ la distribución de $\tilde{g}Z$ cuando la variable Z tiene la distribución π_x, es decir,

$$\pi_x \tilde{g}(F) = Pr(\tilde{g}\, Z \in F) = \pi_x[\tilde{g}^{-1}(F)] \qquad \forall\, F \in \mathbf{F}$$

Por tanto, la condición de invariancia de $\pi \in \Pi$ se puede escribir:

$$\pi_x[\tilde{g}^{-1}(F)] = \pi_{g(x)}(F) \qquad \forall\, x \in \mathbf{X},\ \forall\, g \in G,\ \forall\, F \in \mathbf{F}$$

o bien de un modo más simple:

$$\pi_x(F) = \pi_{g(x)}[\tilde{g}(F)] \qquad \forall\, x \in \mathbf{X},\ \forall\, g \in G\ \ \forall\, F \in \mathbf{F}$$

Definición 4

Dos puntos $\theta_1, \theta_2 \in \Theta$ se dice que son equivalentes respecto a G si existe $\bar{g} \in \bar{G}$ tal que $\theta_1 = \bar{g}(\theta_2)$.

Puede comprobarse que esta relación es de equivalencia; a las clases de equivalencia en que queda dividido las llamaremos «órbitas».

Nuestro próximo objetivo es probar que la función de riesgo de una regla invariante es constante sobre cada órbita, resultado que permite una cierta reducción a la búsqueda de reglas invariantes. Con este fin introducimos la siguiente notación: si $\pi \in \Pi$, designaremos por π^g una regla tal que $\forall\, x \in \mathbf{X}$, π^g_x es la distribución de $\tilde{g}^{-1} Z$ cuando la distribución de Z es $\pi_g(x)$; manteniendo el convenio de la definición podemos escribir $\pi^g_x = \pi_{(g)x}\, \tilde{g}^{-1}$; es decir:

$$\pi^g_x(F) = P(g^{-1} Z \in F) = \pi_{g(x)}[\tilde{g}(F)] \qquad \forall\, x \in \mathbf{X}, \quad \forall\, g \in G, \quad \forall\, F \in \mathbf{F} \qquad [4.1]$$

Se tiene que $\pi \in \Pi$ es invariante si, y sólo si, $\pi = \pi^g$, $\forall\, g \in G$.

LEMA 3

Si un problema de decisión es invariante bajo G, entonces para cada regla $\pi \in \Pi$ y cada $g \in G$, se tiene:

$$R(\theta, \pi^g) = R[\bar{g}(\theta), \pi], \quad \forall\, \theta \in \Theta$$

Demostración.—Por definición,

$$R(\theta, \pi) = \int L(\theta, \pi_x)\, d P_\theta(x) = \int \left[\int L(\theta, a)\, d\pi_x(a)\right] d P_\theta(x) \quad , \quad \pi \in \Pi$$

Al ser el problema invariante,

$$L(\theta, a) = L[\bar{g}(\theta), \tilde{g}(a)] \qquad \forall\, \theta \in \Theta$$

de donde se deduce, debido a

$$L(\theta, \pi_x) = \int L(\theta, a)\, d\pi_x(a) = \int L[\bar{g}(\theta), \tilde{g}(a)]\, d\pi_x(a) = \int L[\bar{g}(\theta), a']\, d(\pi_x \tilde{g})(a') =$$

$$= L[\bar{g}(\theta), \pi_x \tilde{g}] = L(\theta, \pi_{g(x)}\, \tilde{g}^{-1}) = L[\bar{g}(\theta), \pi_{g(x)}]$$

y por [4.1],

$$R(\theta, \pi^g) = \int L(\theta, \pi^g_x)\, d P_\theta(x) = \int L[\bar{g}(\theta), \pi_{g(x)}]\, d P_\theta(x) = \int L[\bar{g}(\theta), \pi_x]\, d P_{\bar{g}\theta}(x) =$$

$$R[\bar{g}(\theta), \pi] \quad , \quad \forall\, \theta \in \Theta \quad , \quad \text{c. q. d.}$$

TEOREMA 1. Si un problema de decisión es invariante bajo G, entonces:

a) Para toda regla de decisión invariante $d \in D$ se verifica:

$$R(\theta, d) = R[\bar{g}(\theta), d] \quad ; \quad \forall\, \theta \in \Theta \quad ; \quad \forall\, \bar{g} \in \bar{G}$$

b) Para toda regla de decisión invariante $\pi \in \Pi$ se verifica:

$$R(\theta, \pi) = R[\bar{g}(\theta), \pi] \quad ; \quad \forall\, \theta \in \Theta \quad ; \quad \forall\, \bar{g} \in \bar{G}$$

Demostración.—a) Utilizando la invariancia de la función de pérdida de $d \in D$ y de **P** en este orden, se tiene:

$$R(\theta, d) = \int L[\theta, d(x)]\, dP_\theta(x) = \int L[\bar{g}(\theta), \tilde{g}\, d(x)]\, dP_\theta(x) =$$
$$= \int L[\bar{g}(\theta), d(g(x))]\, dP_\theta(x) = \int L[\bar{g}(\theta), d(x)]\, dP_{\bar{g}(\theta)}(x) = R[\bar{g}(\theta), d]$$

b) Utilizando el lema 3 y la caracterización [4.1] resulta:

$$R(\theta, \pi) = \mathrm{R}(\theta, \pi^g) = R[\bar{g}(\theta), \pi]$$

El siguiente teorema proporciona resultados relativos a reglas minimax invariantes, mostrando cómo la invariancia puede proporcionar una simplificación en la búsqueda de reglas minimax.

TEOREMA 2. Supongamos que un problema es invariante bajo el grupo finito $G = \{g_1, \ldots, g_n\}$. Si existe una regla minimax, existe una regla aleatorizada que es minimax e invariante. Si una regla es minimax en la clase de las reglas aleatorizadas invariantes, es minimax.

Demostración.—Basta probar que para cada regla $\pi \in \Pi$ existe una regla invariante $\pi^I \in \Pi$ tal que $\sup_\theta R(\theta, \pi^I) \leq \sup_\theta R(\theta, \pi)$.

Definamos

$$\pi^I_x(F) = \frac{1}{n} \sum_{i=1}^{n} \pi^{g_i}_x = \frac{1}{n} \sum_{i=1}^{n} \pi_{g_i(x)}\, [\tilde{g}_i(F)] \qquad \forall\, x \in \mathbf{X} \quad , \quad \forall\, F \in \mathbf{F}$$

π^I es invariante, pues

$$\pi^I_{g(x)}\, [\tilde{g}(F)] = \frac{1}{n} \sum_{i=1}^{n} \pi_{g_i[g(x)]}\, \tilde{g}_i[\,,(F)] = \frac{1}{n} \sum \pi_{gi(x)}\, [\tilde{g}_i(F)] =$$
$$= \pi^I_x(F) \qquad \forall\, F \in \mathbf{F}$$

Tenemos, debido a las definiciones de π^I y de $R(\theta, \pi^I)$, que

$$R(\theta, \pi^I) = \frac{1}{n} \sum_{i=1}^{n} R(\theta, \pi^{g_i})$$

Por tanto, aplicando el lema 3 y teniendo en cuenta que $\bar{g}_i\ 1 \leq i \leq n$ es biyectiva, tenemos:

$$\sup_{\Theta} R(\theta, \pi^I) = \sup_{\Theta} \frac{1}{n} \sum R(\theta, \pi^{g_i}) = \sup_{\Theta} \frac{1}{n} \sum R[\bar{g}_i(\theta), \pi] \leq$$

$$\leq \frac{1}{n} \sum_{i=1}^{n} \sup_{\Theta} R[\bar{g}(\theta), \pi] = \frac{1}{n} \sum_{i=1}^{n} \sup R(\theta, \pi) = \sup_{\Theta} R(\theta, \pi)$$

y el teorema queda probado.

La admisibilidad de reglas basada en el principio de invariancia viene dado por el siguiente teorema.

TEOREMA 3. Si un problema de decisión es invariante bajo un grupo finito $G = \{g_1, ..., g_n\}$, toda regla invariante admisible en la clase de las reglas invariantes es admisible.

Demostración.—Sea $\pi_0^I \in \Pi$ invariante y admisible en la clase de reglas invariantes. Si π_0^I fuera inadmisible existiría $\pi \in \Pi$ tal que

$$R(\theta, \pi) \leq R(\theta, \pi_0^I)\ \forall\ \theta \in \Theta \qquad \text{y} \qquad R(\theta_0, \pi) < R(\theta_0; \pi_0^I) \qquad \theta_0 \in \Theta$$

Definamos π^I como en el teorema 2; utilizando el lema 3 y el teorema 2, se tiene:

$$R(\theta, \pi^I) = \frac{1}{n} \sum_{i=1}^{n} R(\theta, \pi^{g_i}) = \frac{1}{n} \sum_{i=1}^{n} R[\bar{g}_i(\theta), \pi] \leq \frac{1}{n} \sum_{i=1}^{n} R[\bar{g}_i(\theta), \pi_0^I] =$$

$$= \frac{1}{n} \sum_{i=1}^{n} R(\theta, \pi_0^I) = R(\theta, \pi_0^I) \qquad \forall\ \theta \in \Theta$$

y para θ_0, como uno de los g_i es la identidad

$$R(\theta^0, \pi_0^I) = \frac{1}{n} \sum R[\bar{g}_i(\theta_0), \pi] < \frac{1}{n} \sum_{i=1}^{n} R[\bar{g}_i(\theta_0), \pi_0^I] = R(\theta_0, \pi_0^I) \qquad \text{y} \qquad \pi_0^I$$

no sería admisible en la clase de las reglas invariantes.

Ejemplos.—Consideremos el problema de estimar el parámetro en la densidad exponencial negativa $\theta e^{-\theta x}$, $x > 0$, $\theta > 0$.

Con una muestra aleatoria simple, y con función de pérdida $L(\theta, a) = \dfrac{(a - \theta)^2}{\theta^2}$. La acción a es hacer la inferencia de que el parámetro es a, y la función de decisión aplica el punto muestral x en su correspondiente estimador $a = d(x)$. Sabemos que la exponencial negativa es el modelo del intervalo entre sucesos de un Proceso de Poisson y el parámetro θ es el número medio de sucesos por unidad de tiempo; el recíproco $\dfrac{1}{\theta}$ es el tiempo medio entre sucesos. Si X se mide en minutos, θ se mide en recíproco de minutos.

Supongamos ahora que X está en minutos, pero las medidas se hacen en segundos: núm. de segs. = 60 núm. de min., si $X' = 60\,X$, y entonces la densidad de X' es

$$\frac{1}{60}\,\theta\, e^{-\theta\frac{x}{60}} = \theta'\, e^{-\theta' x'} \qquad \theta' = \frac{\theta}{60}$$

La acción correcta a' si θ' es el verdadero estado, es 1/60 de la acción correcta a para el estado θ. La función de pérdida puede expresarse en términos de a' y θ':

$$\frac{(a-\theta)^2}{\theta^2} = \frac{\left(\dfrac{a}{60} - \dfrac{\theta}{60}\right)^2}{\left(\dfrac{\theta}{60}\right)^2} = \frac{(a'-\theta')^2}{\theta^2}$$

Entonces el problema transformado es: dadas las medidas X_1', ..., X_n' construir un estimador a' de θ' para la densidad $\theta' e^{-\theta' x'}$ con función de pérdida $\dfrac{(a'-\theta')^2}{(\theta')^2}$. Pero éste, como vemos, es el mismo problema anterior con nuevos niveles. El problema vemos es invariante bajo la transformación $X' = 60\,X$ (es decir, cambio de escala $X' = KX$).

Por esto resultará que una función de decisión que es buena para el problema original será buena para el nuevo (aplicando la transformación a los datos X'); esto es, $d\,(X')$ será usada en el nuevo problema:

$$d\,(X') = d\,(X'_1, \ldots, X'_n) = d\,(60\,X_1, \ldots, 60\,X_n)$$

y a causa de que $d\,(X')$ es un estimador de θ' en segundos recíprocos, el correspondiente estimador de θ en recíproco de minutos será de $60\,d\,(60\,X)$. Una función de decisión que tenga esta propiedad se dice que es invariante bajo la transformación.

$$d\,(X) = \frac{1}{\bar{x}} \text{ tiene esta propiedad.}$$

Otro ejemplo lo tenemos con la estimación de la variancia de una distribución: al ser la varianza invariante por una traslación, está justificado exigir que el procedimiento de decisión sea independiente del origen de medidas.

APLICACIONES

Estimación de parámetros de posición y de escala

El problema de Estimación puede considerarse como un problema de decisión estadística; en éste, la decisión tomada por el estadístico es la estimación del valor de algún parámetro θ que toma valores en un subconjunto Θ de R. En este caso el espacio D de decisiones coincide con Θ. Supondremos que $D = \Theta = R$, aunque pueda ser nula la probabilidad de que θ pertenezca a ciertos subconjuntos de R. En este caso la pérdida $L(\theta, d)$ reflejará la discrepancia entre θ y su estimación d. Frecuentemente, la función de pérdida L se supone de la forma $L(\theta, d) = \gamma(\theta)\,\Lambda(\theta - d)$, siendo Λ una función no negativa de la diferencia $(\theta - d)$ y tal que $\Lambda(0) = 0$, y γ es una función no negativa que indica la gravedad del error $(\theta - d)$ según los valores de θ; muchas veces γ se puede tomar como constante sobre Θ.

Nuestro propósito es el de exponer la aplicación del Principio de Invariancia al estudio de la teoría de la Estimación, cuando se trata de problemas en los que intervienen parámetros de posición y de escala.

Consideremos una variable aleatoria real X, cuya distribución P_θ depende de un parámetro real $\theta \in \Theta \subset R$ eje real. Sea $F_X(x/\theta)$ la función de distribución de X cuando θ es el verdadero valor del parámetro.

Definición

Un parámetro real $\theta \in \Theta$ se dice parámetro de posición para la distribución de una variable aleatoria X si $F_x(X/\theta)$ es una función sólo de $(x - \theta)$, es decir, $F_X(x/\theta) = F(x - \theta)$, donde F es una función de distribución.

El siguiente lema da definiciones alternativas de parámetro de posición.

Lema *a)* θ es un parámetro de posición para X si, y sólo si, la distribución de $(X - \theta)$, cuando θ es el verdadero valor del parámetro, es independiente de θ.

Lema *b)* Si la distribución de X es absolutamente continua con función de densidad $f_X(x/\theta)$, θ es parámetro de posición si $f_X(x/\theta) = f(x - \theta)$ para alguna densidad $f(x)$.

Ejemplo: Si X tiene la distribución $N(\theta, \sigma)$ con σ conocido, θ es parámetro de posición.

Definición

Un parámetro real $\theta \in \Theta$ se dice que es un parámetro de escala, si $F_X(x/\theta)$ es función solo de $\dfrac{x}{\theta}$, es decir, $F_X(x/\theta) = F\left(\dfrac{x}{\theta}\right)$, siendo F una función de distribución.

Los siguientes lemas dan definiciones alternativas de parámetro de escala.

Lema *a)* θ es parámetro de escala para X si, y sólo si, la distribución de $\dfrac{X}{\theta}$, cuando θ es el verdadero valor del parámetro, es independiente de θ.

Lema *b)* Si la distribución de X es absolutamente continua con densidad $f_X(x/\theta)$, entonces θ es un parámetro de escala para X y sólo si $f_X(x/\theta) = \dfrac{1}{\theta} f\left(\dfrac{x}{\theta}\right)$ siendo f una función de densidad.

Observaciones

1. Si θ es un parámetro de escala para X, $\log\theta$ es un parámetro de posición para $\log X$ si $X > 0$, y para $\log(-X)$ si $X < 0$.
2. Si θ es un parámetro de posición para X, $\sigma = e^{\theta}$ es de escala para $Y = e^{X}$.

Esto permite pasar de los problemas con parámetros de posición a los problemas con parámetros de escala.

DEFINICIÓN

Un parámetro bidimensional (μ, σ) con $\sigma > 0$ se dice de posición-escala para X si $F_X(x/\mu, \sigma)$ es función sólo de $\dfrac{x-\mu}{\theta}$, es decir, $F_X(x/\mu, \sigma) = F\left(\dfrac{x-\mu}{\sigma}\right)$; siendo F una función de distribución.

Estimación minimax de parámetros de posición

Aplicando por una parte la definición de regla invariante no aleatorizada, y por otra el que la función de riesgo correspondiente a una regla invariante es en este caso independiente de θ, resulta:

TEOREMA.—En el problema de estimar un parámetro de posición cuando Θ = eje real y $L(\theta, a) = L(a - \theta)$; en cuyo caso el problema es invariante, por el grupo de traslaciones, si $E_0 L(x - b)$ (pérdida esperada correspondiente a $\theta = 0$, $a = b$) existe y es finita para algún b, y existe un b_0 tal que

$$E_0 L(x - b_0) = \inf_b E_0 L(x - b) \qquad [1]$$

Tomando el ínfimo sobre el conjunto de b para los cuales existe $E_0 L(x - b)$, entonces $d(x) = x - b_0$ es la mejor regla invariante y tiene riesgo constante R_0 igual al primer miembro de [1].

El hecho de que la regla hallada sea una regla igualizadora sugiere la posibilidad de que sea minimax. Sin embargo, esto no es siempre cierto (los resultados obtenidos sobre reglas invariantes no son aquí aplicables al no ser G finito). Vamos, pues, a dar una condición para que la regla obtenida sea minimax.

TEOREMA.—Bajo las hipótesis del teorema anterior, si L está acotada inferiormente y para $\forall\, \varepsilon > 0$ existe N tal que

$$\int_{-N}^{N} L(x - b)\, dF(x) \geq R_0 - \epsilon \qquad \forall\, b$$

entonces la mejor regla invariante es minimax.

Estimación minimax de parámetros de escala

Supongamos que nos planteamos el mismo problema anterior, pero para parámetros de escala.

Sea θ un parámetro de escala de la distribución de X y sea la función de pérdida función de $\frac{a}{\theta}$; es decir: $L(\theta, a) = L\left(\frac{a}{\theta}\right)$. Este problema es invariante bajo el grupo de cambios de escala $g_c(x) = Cx$; $c > 0$.

Utilizando la citada equivalencia entre parámetros de escala y de posición y aplicando los resultados del párrafo anterior llegamos a que el mejor estimador invariante de θ es $\frac{x}{b_0}$, donde b_0 es el valor de b que minimiza $E\left[L\left(\frac{x}{b}\right)/\theta = 1\right]$.

En el caso de la distribución normal $N(\theta, \sigma)$, en que es un parámetro de posición, y (θ, σ) es un parámetro de posición-escala, son aplicables los dos apartados anteriores.

Estimador de Pitman

Un caso de estimador invariante de un parámetro de posición es el célebre estimador de Pitman, estimador invariante de riesgo mínimo respecto al grupo de traslaciones cuando la función de pérdida es cuadrática. Girshick y Savage vieron que el estimador era no sólo de riesgo mínimo, sino también minimax.

El citado estimador fue propuesto por Pitman en 1939, seis años antes de que Hunt y Stein introdujeran el principio general de invariancia.

Forma del estimador:

Sea $(x_1, \ldots, x_n)$ una variable n dimensional para la cual θ es un parámetro de posición.

Sea la función de densidad

$$f_{X_1, \ldots, X_n}^{(x_1, \ldots, x_n)} = f(x_1 - \theta, \ldots, x_n - \theta) \qquad f_{X_1, \ldots, X_n}$$

y la función de pérdida cuadrática.

Aplicando a este caso la construcción dada para el mejor estimador invariante, queda

$$d_0(x) = \frac{\int \sigma f(x_1 - \theta, \ldots, x_n - \theta)\, d\theta}{\int f(x_1 - \theta, \ldots, x_n - \theta)\, d\theta}$$

«Test» de hipótesis

Los problemas de decisión en los que el espacio de acciones consta de dos elementos, $A = \{a_0, a_1\}$, son los llamados problemas de contraste o «test» de hipótesis; esencialmente consisten en decidir si es cierta o no alguna hipótesis que ha sido formulada, estando basada nuestra decisión en una variable aleatoria X que toma valores en un espacio muestral $(\mathbf{X}, \mathbf{A},)$ y cuya distribución pertenece a la familia $\mathbf{P} = \{ P_\theta, \theta \in \Theta \}$.

En la versión clásica del problema, a la cual nos referiremos con el fin de estudiar y utilizar sus resultados, el espacio paramétrico Θ está dividido en dos subconjuntos disjuntos, Θ_0, Θ_1, y asociaremos la acción a_0 a aceptar la hipótesis H_0 inicialmente establecida, llamada «hipótesis nula», de que el parámetro está en Θ_0 y la acción a_1 irá asociada a rechazar la hipótesis inicial y aceptar la «hipótesis alternativa» de que el parámetro está en Θ_1.

Una regla de decisión para estos problemas, regla que en adelante llamaremos «test», no aleatorizada $d(x)$, asigna a cada valor $X = x$ una de las posibles acciones y, por tanto, divide a $\mathbf{X}$ en dos regiones complementarias, S_0, S_1, de modo que $d(x) = a_0$ si $x \in S_0$ y $d(x) = a_1$ si $x \in S_1$. La región S_1 se llama «región crítica».

Sea $L(\theta, a)$ la función de pérdida; si hacemos $L_i(\theta) = L(\theta, a_i)$, $i = 0,1$, la función de riesgo para el «test» $d(x)$ será $R(\theta, d) = L_0(\theta) Pr_\theta \{ X \in S_0 \} + L_1(\theta) Pr_\theta \{ X \in S_1 \} = $ $= L_0(\theta) + Pr_\theta \{ X \in S_1 \} [L_1(\theta) - L_0(\theta)]$.

En los problemas de contraste los «test» aleatorizados juegan un papel importante. Debido a la estructura del espacio A, un «test» aleatorizado (en lo sucesivo usaremos sólo la expresión «test») estará determinado conociendo la probabilidad de una de las dos acciones. Es habitual indicar la probabilidad de la acción a_1; de este modo, un «test» es una función medible Φ definida en $\mathbf{X}$ y con valores en $[0,1]$; dado un valor $X = x$, $\Phi(x)$ representa la probabilidad de tomar la acción a_1. La función de riesgo para el «test» Φ será

$$R(\theta, \Phi) = L_0(\theta) \int [1 - \Phi(x)] \, d P_\theta(x) + L_1(\theta) \int \Phi(x) \, d P_\theta(x) =$$
$$= L_0(\theta) + \int \Phi(x) \, d P_\theta(x) [L_1(\theta) - L_0(\theta)]$$

A la función $E_\theta \Phi(x) = \int \Phi(x) \, d P_\theta(x)$ se le llama «función de potencia» o «potencia» del «test» Φ y la denotaremos por $P_\Phi(\theta)$. Por tanto,

$$R(\theta, \Phi) = L_0(\theta) + P_\Phi(\theta) [L_1(\theta) - L_0(\theta)] \qquad [2]$$

Es posible en estos problemas cometer dos errores: rechazar la hipótesis cuando es cierta, «error de tipo I», y aceptar la hipótesis cuando es falsa, «error de tipo II». El problema se centra en obtener «test» que minimicen estos errores; esto es, en general, difícil, pues las probabilidades de los errores no se pueden controlar simultáneamente.

Supongamos que si $\theta \in \Theta_0$, al ser a_0 la decisión correcta, se tiene $L_0(\theta) = 0$, y que, análogamente, $L_1(\theta) = 0$ si $\theta \in \Theta_1$. Entonces, de [2] obtenemos

$$R(\theta, \Phi) = \begin{cases} L_1(\theta) P_\Phi(\theta) & \text{si} \quad \theta \in \Theta_0 \\ L_0(\theta) [1 - P_\Phi(\theta)] & \text{si} \quad \theta \in \Theta_1 \end{cases} \qquad [3]$$

Vemos cómo, en este caso, la función de riesgo de un «test» depende de su potencia y si, además, $L_1(\theta)$ y $L_0(\theta)$ son constantes en Θ_0 y Θ_1, respectivamente, la función de riesgo depende únicamente da la potencia. A ello se debe el que la mayor parte de la teoría clásica de contraste esté basada en la función de potencia y no considere la función de pérdida.

A la vista de [3], un «test» óptimo sería aquel que minimizara $P_\Phi(\theta)$ en Θ_0 y maximizara $P_\Phi(\theta)$ en Θ_1; esto no es posible en general, es preciso imponer ciertas restricciones. Una restricción habitual consiste en fijar una cota α a la probabilidad de cometer un error de tipo I, y entonces un «test» óptimo sería uno de esta clase que maximizara $P_\Phi(\theta)$ en Θ_1.

Un «test» Φ para contrastar H_0: $\theta \in \Theta_0$ frente a H_1: $\theta \in \Theta_1$, se dice que tiene nivel α si $P_\Phi(\theta) \leq \alpha \ \forall\ \theta \in \Theta_0$. Un «test» se dice que tiene tamaño α si $\sup\limits_{\Theta_0} P_\Phi(\theta) = \alpha$.

Un «test» Φ_0 se dice que es uniformemente más potente (U. M. P.) de tamaño α para contrastar H_0 frente a H_1 si Φ_0 es de tamaño α y para cualquier otro «test» Φ de nivel α se verifica $P_\Phi(\theta) \leq P\Phi_0(\theta)$, $\forall\ \theta \in \Theta_1$.

Hipótesis y alternativa simple

Supongamos que $\Theta_0 = \{\theta_0\}$ y $\Theta_1 = \{\theta_1\}$, es decir, que la hipótesis y la alternativa son simples. La siguiente definición permite concretar nuestra búsqueda en este caso.

Un «test» Φ_0 se dice que es el más potente de tamaño α para contrastar H_0: $\theta = \theta_0$ frente a H_1: $\theta = \theta_1$ si $P\Phi_0(\theta_0) = \alpha$, y para cualquier otro «test» de nivel α se verifica $P_\Phi(\theta_1) \leq P\Phi_0(\theta_1)$.

Obsérvese que si Φ_0 es admisible, entonces Φ_0 es el más potente de su tamaño; el recíproco, sin embargo, no siempre es cierto (Ferguson).

TEOREMA: (*Lema fundamental de Neyman-Pearson*).—Sean P_0 y P_1 dos distribuciones de probabilidad sobre (X, A) con densidades respectivas $p_0(x)$ y $p_1(x)$, respecto a una medida σ-finita μ. Entonces:

a) *Existencia.*—Para cada $0 \leq \alpha \leq 1$ o y para contrastar H_0: p_0 frente a H_1: p_1 existen un «test» Φ y constantes C, γ tales que

$$\Phi(x) = \begin{cases} 1 & \text{si} \quad p_1(x) > Cp_0(x) \\ \gamma & \text{si} \quad p_1(x) = Cp_0(x) \\ 0 & \text{si} \quad p_1(x) < Cp_0(x) \end{cases} \qquad [4]$$

b) *Suficiencia.*—Si un «test» es de la forma [4] para algún $C \geq 0$ es el más potente de su tamaño para contrastar H_0 frente a H_1. Correspondiente a $C = \infty$, el «test»

$$\Phi(x) = \begin{cases} 1 & \text{si} \quad p_0(x) = 0 \\ 0 & \text{si} \quad p_0(x) > 0 \end{cases} \qquad [5]$$

es el mejor de tamaño α para contrastar H_0 frente a H_1.

c) *Unicidad.*—Si Φ_0 es el «test» más potente de tamaño α para contrastar H_0 frente a H_1, entonces Φ_0 tiene la forma [4] o [5] excepto, a lo sumo, en un conjunto μ-nulo.

Su demostración, que omitimos por ser bien conocida y encontrarse en todos los textos clásicos, puede verse, por ejemplo, en Ferguson, Fraser y Lehmann.

«Test» de hipótesis invariantes

Supongamos que la función de pérdida para el problema de contraste H_0: $\theta \in \Theta_0$ frente a H_1: $\theta \in \Theta_1$ es de la forma:

$$L_0(\theta) = \begin{cases} L_{00} & \text{si} \quad \theta \in \Theta_0 \\ L_{01} & \text{si} \quad \theta \in \Theta_1 \end{cases}$$

$$L_1(\theta) = \begin{cases} L_{10} & \text{si} \quad \theta \in \Theta_0 \\ L_{11} & \text{si} \quad \theta \in \Theta_1 \end{cases}$$

donde

$$L_{00} < L_{10} \quad \text{y} \quad L_{11} < L_{01}$$

Sea G un grupo de transformaciones sobre $\mathbf{X}$ que deja invariante este problema, y sean $\bar{G}$ y $\tilde{G}$ los grupos de transformaciones inducidos sobre Θ y A. Debido a la invariancia de la función de pérdida, se tiene $L(\theta, a) = L[\bar{g}(\theta), \tilde{g}(a)]$, $\forall\, \theta \in \Theta$, $\forall\, a \in A$, $\forall\, g \in G$. Expresión que para este problema particular puede darse de modo más sencillo, pues si los L_{ij} son diferentes, la pérdida será invariente si, y sólo si,

$$\tilde{g}(a) = a, \quad \forall\, a \in A, \quad \forall\, \tilde{g} \in \tilde{G}$$

$$\bar{g}(\theta_0) = \theta_0, \quad \bar{g}(\theta_1) = \theta_1, \quad \forall\, \bar{g} \in \bar{G} \qquad [6]$$

Debido a ello, en los problemas de contraste tomaremos $\tilde{G} = \{ \tilde{e} \}$ y reemplazaremos la condición de invariancia de la función de pérdida por la condición [6]. Así, pues, diremos que un grupo G de transformaciones sobre $\mathbf{X}$ deja invariante un problema de contraste si G deja invariantes las familias de distribuciones $\{ P_\theta, \theta \in \Theta_1 \}$ y $\{ P_\theta\, \theta \in \Theta \}$, es decir, si $\bar{G}$ preserva los subconjuntos Θ_0 y Θ_1. Teniendo en cuenta el concepto de regla invariante, podemos dar en este caso particular la siguiente definición:

Un «test» Φ es invariante bajo si se verifica $\Phi[g(x)] = \Phi(x)$, $\forall\, x \in \mathbf{X}$, $\forall\, \bar{g} \in G$.

Invariantes maximales

Para cada problema de contraste invariante el principio de invariancia permite restringir nuestra atención a los «test» invariantes; surge así la necesidad de caracterizar estos «tests». La principal simplificación que proporciona la aplicación del principio de invariancia a los problemas de contraste radica en el hecho de que los «tests» invariantes pueden ser caracterizados de una forma muy simple, mediante unos estadísticos particulares que pasamos a definir a continuación.

Sea G un grupo de transformaciones sobre $\mathbf{X}$. Se dice que un estadístico $T(x)$ definido sobre $\mathbf{X}$ es un invariante maximal (I. M.) respecto a G si $T[g(x)] = T(x)$, $\forall x \in \mathbf{X}$, $\forall g \in G$ y si $T(x_1) = T(x_2)$ implica la existencia de $g \in G$ tal que $x_1 = g(x_2)$.

Un invariante maximal es constante sobre cada órbita y toma valores diferentes en cada una de ellas, es decir, distingue las órbitas.

Ejemplo: Sea $\mathbf{X} = R^n$, definamos para cada $c \in R$, $g_c(x) = g_c(x_1, \ldots, x_n) = (x_1 + c, \ldots, x_n + c)$, y consideremos $G = \{g_c, c \in R\}$.

La función $T(x) = (x_1 - x_n, \ldots, x_{n-1} - x_n)$ es invariante bajo G, pues $T[g_c(x)] = T(x_1 + c, \ldots, x_n + c, \ldots, x_n + c) = (x_1 - x_n, \ldots, x_{n-1} - x_n) = T(x)$, $\forall g_c \in G$; además, $T(x) = T(x')$ implica $x_i - x_n = x'_i - x'_n$, $1 \leq i \leq n-1$, y si $c = x_n - x'_n$ resulta $x_i = x'_i + c$, $1 \leq i \leq n$, es decir, $g_c(x') = x$. Por tanto, $T(x)$ es I. M. respecto a G.

Ejemplo: Sea $\mathbf{X} = R^n$ y G el grupo de las permutaciones de n elementos. La función $T(x) = (x^{(1)}, \ldots, x^{(n)})$, donde $x^{(1)}$ es el menor x_i, $x^{(2)}$ el siguiente menor, ..., $x^{(n)}$ el mayor x_i, es decir, el estadístico de orden (Ríos, p. 261; Wilks, p. 234), es I. M. En efecto, cada permutación de las coordenadas x_i no cambia el orden existente entre ellas, y dos puntos que tienen el mismo conjunto ordenado de coordenadas pueden obtenerse uno del otro mediante una permutación de sus coordenadas.

El siguiente teorema muestra como los «tests» invariantes quedan caracterizados mediante un I. M.

TEOREMA.—Sea G un grupo de transformaciones sobre un espacio $\mathbf{X}$ y $T(x)$ un I. M. respecto a G; entonces, una condición necesaria y suficiente para que Φ sea invariante es que Φ sea una función de $T(x)$.

Demostración.—Sea Φ invariante y $T(x_1) = T(x_2)$, existe $g \in G$ tal que $x_1 = g(x_2)$; por tanto, $\Phi(x_1) = \Phi[g(x_2)] = \Phi(x_2)$, lo que prueba que Φ es función de $T(x)$.

Si $\Phi(x) = h[T(x)]$, $\forall x \in \mathbf{X}$, se tiene $\Phi[g(x)] = h\{T[g(x)]\} = h[T(x)] = \Phi(x)$, $\forall x \in \mathbf{X}$, $\forall g \in G$ y Φ es invariante.

Parece que podemos deducir que la clase de los «tests» invariantes está formada por los «tests» que dependen del I. M.; esta deducción es correcta en los casos usuales, pero es preciso hacer la siguiente observación: Para determinar un estadístico es preciso especificar la clase β de los conjuntos medibles de su espacio imagen. Si, en nuestro caso, β es la clase de los conjuntos B tales que $T^{-1}(B) \in \mathbf{A}$, todo es correcto, pues si $\Phi(x) = h[T(x)]$, al ser Φ A-medible, para todo conjunto C de Borel se tiene $\Phi^{-1}(C) = T^{-1}[h^{-1}(C)] \in \mathbf{A}$; por tanto, $h^{-1}(C) \in \beta$; h es β-medible y Φ es un «test» basado en el estadístico $T(x)$. En muchas aplicaciones, $T(x)$ es una función que toma valores en R^m y conviene tomar como σ-álgebra β la clase de los conjuntos de Borel, ya que en este caso, si $\mathbf{X} = R^n$ $\mathbf{A}$ es la clase de los conjuntos de Borel, y $\Phi(x) = h[T(x)]$ es una función medible que depende sólo de $T(x)$, se verifica que h *es* β-medible.

A la vista de lo expuesto, la aplicación del principio de invariancia para el contraste de hipótesis puede esquematizarse de la siguiente forma: en primer lugar, considerar un grupo G de transformaciones que deje invariante el problema; a continuación, obtener un I. M. respecto a G, calcular las distribuciones inducidas para dicho I. M. y considerar el problema para ellas; finalmente, buscar el mejor «test» para este nuevo problema; este «test», expresado en términos del problema inicial mediante el I. M., será el mejor «test» invariante.

El siguiente resultado muestra cómo las distribuciones inducidas del I. M. pueden ser expresadas mediante un I. M. respecto al grupo $\bar{G}$.

TEOREMA.—Si $T(x)$ es invariante respecto al grupo G y si $U(\theta)$ es I. M. respecto al grupo $\bar{G}$, entonces la distribución de $T(x)$ depende de θ a través de $U(\theta)$.

Demostración.—Si $U(\theta_1) = U(\theta_2)$, existe $\bar{g} \in \bar{G}$ tal que $\theta_1 = \bar{g}(\theta_2)$, y entonces

$$Pr_{\theta_2}\{T(x) \in B\} = Pr_{\theta_2}\{T[g(x)] \in B\} = Pr_{\bar{g}(\theta_2)}\{T(X) \in B\} = Pr_{\theta_1}\{T(x) \in B\}$$

La invariancia reduce los datos a un estadístico I. M. cuya distribución depende únicamente de una función del parámetro; por tanto reduce, en cierto sentido, el espacio paramétrico, identificando los puntos que son equivalentes respecto a $\bar{G}$.

CONCLUSIONES

En los apartados anteriores se han desarrollado la aplicación del Principio de la Invariancia a problemas de estimación y «test» de hipótesis. Hemos de señalar, no obstante, que no sólo es en estos casos donde el citado Principio tiene aplicación, es en problemas de Análisis de la varianza y Estadística no paramétrica donde alcanza su máxima eficacia.

Ha sido nuestra intención presentar y divulgar sólo algunos aspectos de la importancia del mismo, dejando para futuras exposiciones los temas anteriormente citados.

BIBLIOGRAFIA

BLACWELL, D., y GIRSHICK, M. A.: *Theory of Games and Statistical Decisions*. John Wiley, 1954.

DE GROOT, M. H.: *Optimal Statistical Decisions*. McGraw-Hill, 1970.

FERGUSON, T. S.: *Mathematical Statistics: A Decision Theoretic Approach*. Academic Press, 1967.

FRASER, D. A. S.: *Nonparametric Methods in Statistics*. John Wiley, 1957.

HALMOS, P. R.: *Measure Theory*. Van Nostrand, 1950.

HOEL, P. G.: *Mathematical Statistic* (3.ª edic.). John Wiley, 1966.

KARLIN, S., y RUBIR, H.: «The theory of decision procedures for distributions with monotone likelihood ratio». *Ann. Math. Statist.*, 27, 1956.

LEHMANN, E. L.: *Testing Statistical Hypotheses*. John Wiley, 1959.

LOÉVE, M.: *Probability Theory* (3.ª edic.). Van Nostrand, 1963.

RÍOS, S.: *Métodos Estadísticos* (5.ª edic.). Ed. del Castillo, 1967.

RÍOS, S.: «Procesos dinámicos de decisión en concurrencia», *Rev. de la R. Acad. de Ciencias*, 1967. Análisis de decisiones. Madrid, 1973.

WILKS, S.: *Mathematical Statistics*. John Wiley, 1963.

ZACKS, S.: *The Theory of Statistical Inference*. John Wiley, 1971.