ANALISIS ESTADISTICO MULTIVARIANTE Y REPRESENTACION CANONICA DE FUNCIONES ESTIMABLES

TESIS PARA OPTAR AL GRADO
DE DOCTOR EN CIENCIAS, SECCION
DE MATEMATICAS, PRESENTADA POR

CARLOS M. CUADRAS AVELLANA

UNIVERSIDAD DE BARCELONA

ANALISIS ESTADISTICO MULTIVARIANTE
Y REPRESENTACION CANONICA
DE FUNCIONES ESTIMABLES

Memoria presentada por
CARLOS M. CUADRAS AVELLANA
para optar al Grado de Doctor
en Ciencias, Sección de Matemáticas.

V° B° El Director de la Tesis

Dr. E. GARDEÑES MARTIN
Profesor Agregado de Análisis Numérico
Facultad de Ciencias. Universidad de Barcelona

Barcelona, Mayo de 1973.

PROLOGO

Poco después de terminar la Licenciatura (1968), me integré al recién creado Laboratorio de Cálculo de la Facultad de Ciencias. En este centro, mi labor de analista de aplicaciones científicas, se fue concretando en la estadística aplicada, sea para cubrir las necesidades de los usuarios, sea por la relación que he tenido con la enseñanza de la biometría y estadística.

He tenido ocasión de conocer, a través de mi trabajo, numerosas aplicaciones de la estadística, y sobretodo, del Análisis Multivariante. Pero me he debido enfrentar algunas veces, con problemas que no podían ser resueltos satisfactoriamente por los métodos conocidos.

De esta forma, he llegado a tomar interés por un problema que me ha sido planteado a menudo: cuando las medidas de unos caracteres sobre una muestra de individuos, dependen de los efectos distintos de unas determinadas situaciones o factores externos (temperatura, profundidad, tratamiento, etc.), ¿cuáles son las causas de estas diferencias, y cómo se distribuyen las diferencias entre los efectos de cada factor?

El caso más sencillo en que los niveles de un factor son simplemente grupos o poblaciones, la diferenciación fue resuelta por RAO (1952), introduciendo la técnica de la representación canónica de poblaciones. No se me ha presentado,pues,dificultad alguna en programarla y aplicarla (véase CUADRAS,CAMPA,MONTORIOL,1972a,1972b; PETITPIERRE,1972), consultando la obra de COOLEY y LOHNES(1962) y SEAL(1964). Pero cuando las medidas de cada individuo de la muestra, dependían de los efectos de más de un factor que se pretendían representar, el análisis canónico de poblaciones resultaba insuficiente.

Generalizar el análisis canónico de RAO a diseños más complicados que el de simples poblaciones, es el objetivo central de esta memoria.Consideraré haberlo cumplido si he llegado a las conclusiones con todo rigor, si tienen aplicación práctica y si pueden ser fácilmente interpretadas por aquel que las utilice.

AGRADECIMIENTOS

Deseo hacer constar mi agradecimiento al Departamento de Ecuaciones Funcionales y al Laboratorio de Cálculo, por su incondicional colaboración y estímulo; al Departamento de Cálculo de Probabilidades y Estadística Matemática, y a la Cátedra "Artigas Sanz" de Estadística Fundamental y Aplicada, a los que debo parte de mi formación.

Esta memoria se ha ido elaborando como forma de solución a algunas consultas de los Departamentos de Genética, Psicología, Microbiología, Ecología y Pedagogía, los cuales, al plantearme su problemática en estadística, han ejercido una función indirecta de guía, orientación y crítica, que me ha sido muy útil.

1 INTRODUCCION

1.1. POSICION DEL PROBLEMA

Cuando, realizado un análisis de la varianza, el efecto de un factor resulta significativo, para poder tener información sobre las causas de esta significación, se hace necesario conocer la forma en que se diferencian los niveles del factor. Algo parecido podría decirse respecto a las interacciones, con efecto significativo, entre dos o más factores del diseño.

El problema, desde un punto de vista general, podría formularse así: dado un diseño multivariante de experimentos, y obtenido un sistema de funciones paramétricas estimables, si son significativamente distintas, ¿cómo se diferencian entre sí?.

Un primer camino para diferenciar los efectos de un factor (o unas interacciones, o un sistema de funciones paramétricas estimables en general), podría consistir en realizar análisis parciales de la varianza, tomando algunos niveles y prescindiendo de los demás. Esta solución, que exigiría un análisis para cada una de las posibles combinaciones, es muy engorroso y prácticamente inviable.

La utilización de contrastes ortogonales, T-contrastes ó S-contrastes en el caso más general (SCHEFFE,1959) permite comparar combinaciones lineales de los niveles. Son muy útiles, pero tienen el inconveniente de que exigen la elección de los coeficientes, que puede ser complicado si los tamaños de las muestras de cada una de las celdas del diseño son distintos. Además, su aplicación se limita en la práctica, a comparar los efectos de un factor principal.

Las dificultades y limitaciones de las soluciones anteriores, y, en cambio, la sencillez de interpretación que proporciona el análisis canónico de RAO (1952) para representar y diferenciar poblaciones, sugieren la conveniencia de generalizarlo a diseños más complicados, para poder hacer lo mismo con un sistema de funciones paramétricas estimables.

1.2. OBJETIVOS

La generalización del análisis canónico de poblaciones a un sistema de funciones paramétricas estimables, es el principal objetivo de esta memoria. En realidad, no abordaremos el problema sin antes revisar algunos conceptos de análisis multivariante.

En primer lugar, presentaremos la memoria según una notación actualizada, siguiendo a DEMPSTER (1969), que utilice los recursos del álgebra lineal moderna. Nos permitirá, entre otras cosas, exponer de forma algebraica la teoría y estimación de funciones paramétricas, independientemente de cualquier parametrización.

Desde luego, debemos analizar con todo detalle, las relaciones entre la versión algebraica y la versión paramétrica de una función estimable, estableciendo la forma de pasar de una a la otra. Además, deseamos obtener en función de un muestreo, la expresión que optimice la estimación de una función estimable, y relacionarla con la estimación que proporciona el teorema de Gauss-Markov.

La representación canónica de un sistema de funciones estimables deberá tener las mismas propiedades métricas que el análisis discriminante de Rao (1952).

Otro objetivo importante será obtener una región confidencial <u>exacta</u> (fijado un coeficiente de confianza) para cada función estimable.

Finalmente, nos proponemos estudiar la representación canónica en el caso de que existan variables concomitantes que influyan en las funciones estimables, y la conexión entre el análisis canónico generalizado y el análisis de coordenadas principales de GOWER (1966).

2 REPRESENTACION BORELIANA MULTIVARIANTE DE POBLACIONES

Se plantean las hipótesis fundamentales que regirán en la memoria, introduciendo los conceptos por álgebra lineal, siguiendo y ampliando la línea que para el análisis multivariante fue iniciada por DEMPSTER (1969).

La representación multivariante de una población,se extiende a un conjunto finito de poblaciones. Para esta parte, hemos tomado como referencia algunas obras sobre álgebras y espacios de probabilidades(CRAMER, 1968; LOEVE,1963; NEVEU,1970; KAPPOS,1969).

2.1 EL ESPACIO DE LAS VARIABLES

Sea Ω un conjunto fundamental de sucesos elementales, que también llamaremos individuos. De momento, identificaremos Ω con una población (aunque más adelante será reunión de poblaciones disjuntas).

Supongamos que sobre Ω tenemos planteada una cierta experiencia aleatoria, que identificaremos con una σ-álgebra $\mathcal{A}$, siendo $\mathcal{A} \subset \mathcal{P}(\Omega)$. Finalmente, suponemos definida una probabilidad P sobre los sucesos de $\mathcal{A}$.

A menudo desconoceremos las características de $\mathcal{A}$, pero podremos resolver en parte este problema, si conocemos un conjunto de variables aleatorias sobre el espacio de probabilidades $(\Omega, \mathcal{A}, P)$.

Sean, pues, $(Y_1, Y_2, \ldots, Y_p)$ p variables aleatorias observables sobre Ω,

$$Y_i : \Omega \longrightarrow \mathbb{R} \qquad [Y_i \leq r] \in \mathcal{A} \qquad \forall r \in \mathbb{R} \qquad i=1,\ldots,p.$$

Indiquemos por E el espacio vectorial sobre R, generado por todas las combinaciones lineales de $(Y_1,\ldots,Y_p)$

$$E = \langle Y_1,\ldots,Y_p \rangle = \{ Y | Y = \sum_1^p a_i Y_i \}$$

Definimos en E un producto escalar

$$\forall Y, Y' \in E \qquad (Y, Y') = Y \cdot Y' = \operatorname{cov}(Y, Y')$$

La matriz asociada a este producto escalar en la base $(Y_1,\ldots,Y_p)$ es la matriz Σ de varianzas-covarianzas entre estas variables aleatorias.

Como primeras hipótesis fundamentales, supondremos:

1) Las esperanzas, varianzas y covarianzas de las v.a. $(Y_1,\ldots,Y_p)$ existen y son finitas, es decir, Σ existe.

2) No existen relaciones de dependencia lineal entre las v.a. $(Y_1,\ldots,Y_p)$, es decir,

$$\operatorname{rang}(\Sigma) = p, \quad \text{o también, } \dim E = p.$$

3) La matriz simétrica Σ es definida positiva.

2.2 EL ESPACIO DUAL DE LOS INDIVIDUOS

Sea E^* el espacio vectorial dual de E. Dado un individuo $\omega \in \Omega$, y una v.a. $Y \in E$, al par ordenado (ω, Y) le podemos asociar el nº real $Y(\omega)$. Existe, pues, una aplicación,

$$\delta : \Omega \longrightarrow E^*$$

$$\omega \longrightarrow \omega^* = \delta(\omega) \quad \text{tal que} \quad \omega^*(Y) = Y(\omega)$$

que permite identificar Ω con un subconjunto de E^*. A los elementos de $\delta(\Omega) \subset E^*$ les llamaremos también individuos, indicando por ω^* al elemento $\omega \in \Omega$ considerado dentro de E^*.

A partir de ahora trabajaremos con una <u>nueva hipótesis fundamental</u>:

(2.2.1) $\quad \delta(\Omega) = E^* \quad$ (δ es aplicación exhaustiva)

Esta aplicación no será, en general, biyectiva, es decir, pueden haber dos o más individuos que alcancen los mismos valores para las v.a. $(Y_1, \dots, Y_p)$.

Sobre E^* nos será posible definir un espacio de probabilidades, por un lado, y una métrica, al considerarlo como espacio vectorial. Estas dos estructuras sobre E^*, que trasladaremos a Ω utilizando (2.2.1) son:

1) Sea $F(y_1, \dots, y_p)$ la función de distribución conjunta de las v.a. $\{Y_i\}$. Podemos construir sobre E^*, que se identifica con un espacio euclídeo R^p, una estructura de espacio de probabilidades,

$$(E^*, \mathcal{B}^p, P_F)$$

siendo $\mathcal{B}^p$ la σ-álgebra de los borelianos de R^p, y P_F la probabilidad que la función F define sobre $\mathcal{B}^p$,

(2.2.2) $\quad P_F(B) = \int_B dF(y_1, \dots, y_p) \qquad \forall B \in \mathcal{B}^p.$

2) Consideremos el isomorfismo:

$$\psi : E \longrightarrow E^*$$

$$Y \longrightarrow \omega^*_Y$$

tal que $\psi(Y)(Y_1) = \omega^*_Y(Y_1) = Y \cdot Y_1 \qquad \forall Y_1 \in E$

A través de ψ se introduce de una forma natural, un producto escalar en E^*,

$$\omega_1^* \cdot \omega_2^* = \psi^{-1}(\omega_1^*) \cdot \psi^{-1}(\omega_2^*) \qquad \forall \omega_1^*, \omega_2^* \in E$$

La matriz asociada a este producto escalar, con referencia a la base dual $(Y_1^*, \ldots, Y_p^*)$, es Σ^{-1}.

2.3 REPRESENTACION MULTIVARIANTE DE UNA POBLACION

Para poder identificar Ω con E^*, definimos la relación siguiente:

Diremos que ω_1 y ω_2 son indistinguibles (métricamente) lo que indicaremos por $\omega_1 \text{ IN } \omega_2$, si $\delta(\omega_1) = \delta(\omega_2)$.

La relación IN significa que las v.a. $\{Y_i\}$ alcanzan los mismos valores sobre ω_1 y ω_2. Naturalmente, IN es una relación de equivalencia, y el conjunto cociente Ω/IN está en correspondencia biyectiva con E^*,

(2.3.1) $\qquad \Omega/\text{IN} \longleftrightarrow E^*$

Indiquemos por $\{\bar{\omega}_i\}_{i \in I}$ la partición de Ω de las clases de equivalencia

$$\Omega = \bigcup_{i \in I} \bar{\omega}_i$$

Sean A y B dos sucesos de $\mathcal{A}$. Diremos que A y B son indistingibles, A IN B, si $\delta(A) = \delta(B)$.

Al ser

$$A = \bigcup_{i \in I} A \cap \bar{\omega}_i, \qquad B = \bigcup_{i \in I} B \cap \bar{\omega}_i$$

la relación IN en $\mathcal{A}$, es de equivalencia. A IN B implica que, para cada $i \in I$, o bien $A \cap \bar{\omega}_i = B \cap \bar{\omega}_i = \emptyset$, o bien existe un $\omega_i' \in A \cap B$ tal que $\omega_i' \text{ IN } \omega_i$. Entonces, los elementos del cociente $\mathcal{A}/\text{IN}$ son de la forma

$$\tilde{A} \in \mathcal{A}/\text{IN}, \qquad \tilde{A} = \left\{ \bar{\omega}_i,\ i \in I_A \right\}$$

siendo $I_A \subset I$ un subconjunto de índices que no depende del representante A, pues $A \; IN \; B \Rightarrow I_A = I_B$.

<u>Teorema 2.3.1</u>

$\mathcal{A}/IN$ es una σ-álgebra de sucesos.

<u>Demostración</u>

El conjunto $\bar{\omega}_i$ es la antiimagen de $(Y_1(\omega_i)\dots Y_p(\omega_i))$, que es un boreliano de $\mathcal{B}^p$. Por ser $\{Y_i\}$ v.a., $\bar{\omega}_i$ es un suceso de $\mathcal{A}$.

El suceso seguro de $\mathcal{A}/IN$ es Ω/IN . Además, si $\bar{A}_n \in \mathcal{A}/IN$, $n \in N$, es una colección finita o numerable de elementos de $\mathcal{A}/IN$,

$$\bigcup_{n\in N} \bar{A}_n = \bigcup_{n\in N}\{\bar{\omega}_{n_i},\ n_i \in I_{A_n}\} = \bigcup_{i\in I_B} \bar{\omega}_i = \overline{\bigcup_{n\in N} A_n}$$

siendo $B = \bigcup_{n\in N} A_n$. Análogamente

$$\bigcap_{n\in N} \bar{A}_n = \overline{\bigcap_{n\in N} A_n}$$

Finalmente,

$$(\bar{A})^c = (\{\bar{\omega}_i,\ i\in I_A\})^c = \{\bar{\omega}_i,\ i\in I-I_A\} = \overline{A^c}$$

<u>Definición 2.3.1</u>

Existe una correspondencia biyectiva

$$\mathcal{A}/IN \longrightarrow \mathcal{B}^p$$

$$\bar{A} \longrightarrow \delta(A)$$

que nos permite definir una probabilidad $\bar{P}$ sobre el espacio medible $(\Omega/IN, \mathcal{A}/IN)$, que, de acuerdo con (2.2.2), será

(2.3.2) $\quad \bar{P}(\bar{A}) = P_{\mathbb{F}}(\delta(A))$

Teorema 2.3.2

1) La σ-álgebra $\mathcal{A}/\mathbb{N}$ es isomorfa con la σ-álgebra $\mathcal{A}'$, subálgebra de $\mathcal{A}$, generada por las clases de equivalencia de Ω,

$$\mathcal{A}/\mathbb{N} \simeq \mathcal{A}' = g(\{\bar{\omega}_i,\ i\in I\})$$

2) Los espacios de probabilidades

$$(\Omega/\mathbb{N}, \mathcal{A}/\mathbb{N}, \bar{P}) \text{ y } (E^*, \mathcal{B}^p, P_F)$$

son isométricos.

Demostración:

Consideremos la colección de sucesos

$$\mathcal{A}' = \{A' \mid A' = \bigcup_{i\in I_A} \bar{\omega}_i,\ A\in\mathcal{A}\}$$

que verifica $\bar{\omega}_i \in \mathcal{A}'$, $i\in I$. Es fácil demostrar que $\mathcal{A}'$ es una σ-álgebra, y que es el álgebra generada por los sucesos $\{\bar{\omega}_i,\ i\in I\}$.

Sea ahora la aplicación

$$\mathcal{A}/\mathbb{N} \longrightarrow \mathcal{A}'$$

$$\bar{A} = \{\bar{\omega}_i,\ i\in I_A\} \qquad A' = \bigcup_{i\in I_A} \bar{\omega}_i$$

Es biyectiva, porque, dentro de $\mathcal{A}$, A y A' son equivalentes A $\mathbb{N}$ A', y por lo tanto $I_A = I_{A'}$. Es fácil demostrar, además, que esta aplicación conserva las operaciones de reunión, intersección y complementación de sucesos.

La isometría entre $(\Omega/\mathbb{N}, \mathcal{A}/\mathbb{N}, \bar{P})$ y $(E^*, \mathcal{B}^p, P_F)$ es consecuencia inmediata de la definición 2.3.1 .

c. q. d.

Vemos así, que la representación de Ω a partir de las v.a. $\{Y_i\}$, nos obliga a identificar elementos distintos de Ω, y a operar con una σ-álgebra $\mathcal{A}/\mathbb{N}$, menos fina que $\mathcal{A}$.

2.4 DESCOMPOSICION EN POBLACIONES

De ahora en adelante supondremos que Ω se descompone en k sucesos mutuamente excluyentes:

$$\Omega = H_1 + \ldots + H_k, \quad H_i \cap H_j = \emptyset \quad H_i \in \mathcal{A}, \quad i=1,\ldots,k,$$

de probabilidades $P(H_1) = p_1, \ldots, P(H_k) = p_k$, a los que llamaremos poblaciones (o también celdas, grupos, casillas, etc.).

Admitiremos que las hipótesis fundamentales 1), 2) y 3) siguen cumpliéndose, restringiendo las v.a. $\{Y_i\}$ a cada una de las celdas o poblaciones, y que la matriz Σ de varianzas-covarianzas es común a todas ellas. Supondremos, además, que la retricción de la aplicación δ a cada H_i verifica:

(2.4.1) $\qquad \delta(H_i) = E^* \qquad i=1,\ldots,k.$

Dado H_i, y construyendo el álgebra condicionada $\mathcal{A}_{H_i} = H_i \cap \mathcal{A}$, la relación de equivalencia IN y todo lo dicho en §2.2 y §2.3, se puede extender al espacio $(H_i, \mathcal{A}_{H_i}, P_{H_i})$.

En particular, el espacio $(H_i/IN, \mathcal{A}_{H_i}/IN, \tilde{P}_{H_i})$ será isométrico a $(E^*, \mathcal{B}^p, P_{F_i})$, siendo F_i la función de distribución conjunta de las v.a. $\{Y_i\}$ restringidas a la celda H_i.

Nuestro propósito es extender estos conceptos a Ω.

2.5 REPRESENTACION MULTIVARIANTE DE k POBLACIONES

La proyección $H_i \longrightarrow H_i/IN \quad (i=1,\ldots,k)$, nos dará:

$$\Omega_{IN} = H_1/IN + \ldots + H_k/IN$$

Asímismo, $\mathcal{A}_{H_i} \longrightarrow \mathcal{A}_{H_i}/IN$, extendida a $\mathcal{A}$, nos da

$$\mathcal{A} \longrightarrow \mathcal{A}_{IN} = \mathcal{A}_{H_1}/IN + \dots + \mathcal{A}_{H_k}/IN$$

Es importante observar que Ω_{IN} no es Ω/IN, y $\mathcal{A}_{IN}$ no es $\mathcal{A}/IN$. Los elementos de $\mathcal{A}_{IN}$ son de la forma:

$$\tilde{A} \in \mathcal{A}_{IN}, \quad \tilde{A} = \tilde{A}_1 + \dots + \tilde{A}_k \qquad \tilde{A}_i \in \mathcal{A}_{H_i}/IN$$

Es fácil verificar que $\mathcal{A}_{IN} \subset \mathcal{P}(\Omega_{IN})$ es una σ-álgebra, y que

$$\tilde{P}(\tilde{A}) = p_1 \tilde{P}_{H_1}(\tilde{A}_1) + \dots + p_k \tilde{P}_{H_k}(\tilde{A}_k)$$

es una probabilidad $\tilde{P}$ sobre $\mathcal{A}_{IN}$. Hemos construido de esta manera, un espacio de probabilidades

$$(\Omega_{IN}, \mathcal{A}_{IN}, \tilde{P}).$$

Consideremos ahora los espacios de probabilidades

$$(\Omega, \mathcal{A}_H, p) \text{ y } (E^*, \mathcal{B}^p, P_{F_i}) \qquad (i=1,\dots,k)$$

siendo $\mathcal{A}_H = g(H_1,\dots,H_k)$

el álgebra finita generada por las k celdas o poblaciones, y F_i $(i=1,\dots,k)$, la función de distribución conjunta de las v.a. restringidas a la celda H_i.

Tomemos como suceso seguro $E^* \times \Omega$, y sea el subconjunto de $\mathcal{P}(E^* \times \Omega)$,

$$\mathcal{B}^p \circ \mathcal{A}_H = \left\{ (B_1,H_1) \cup \dots \cup (B_k,H_k), \quad B_i \in \mathcal{B}^p, \quad (i=1,\dots,k) \right\}$$

Tiene estructura de σ-álgebra porque:

$$E^* \times \Omega = (E^*,H_1) + \dots + (E^*,H_k)$$

$$\bigcup_{n \in N} \bigcup_{i=1}^{k} (B_{n_i}, H_i) = \bigcup_{i=1}^{k} \left(\bigcup_{n \in N} B_{n_i}, H_i \right)$$

$$\bigcap_{n \in N} \bigcup_{i=1}^{k} (B_{n_i}, H_i) = \bigcup_{i=1}^{k} \left(\bigcap_{n \in N} B_{n_i}, H_i \right)$$

$$\left(\bigcup_{i=1}^{k} (B_i, H_i) \right)^c = \bigcup_{i=1}^{k} (B_i^c, H_i)$$

Definamos sobre $\mathcal{B}^P \circ \mathcal{A}_H$ la probabilidad P_f

$$P_f\left(\bigcup_{i=1}^{k} (B_i, H_i)\right) = \sum_{i=1}^{k} P_{F_i}(B_i)\, p_i$$

Es una probabilidad, pues:

$$P_f(E^* \times \Omega) = \sum_{i=1}^{k} 1\, p_i = 1$$

y P_f es σ-aditiva, ya que dado la colección de sucesos disjuntos:

$$P_f \bigcup_{n=1}^{\infty} \bigcup_{i=1}^{k} (B_{n_i}, H_i) = P_f\left(\bigcup_{i=1}^{k}\left(\bigcup_{n=1}^{\infty} B_{n_i}, H_i\right)\right) =$$

$$= \sum_{i=1}^{k}\left(\sum_{n=1}^{\infty} P_{F_i}(B_{n_i})\right) p_i =$$

por tratarse de series absolutamente convergentes

$$= \sum_{n=1}^{\infty} \sum_{i=1}^{k} P_{F_i}(B_{n_i})\, p_i = \sum_{n=1}^{\infty} P_f\left(\bigcup_{i=1}^{k} (B_{n_i}, H_i)\right)$$

Hemos construido así el espacio de probabilidades:

$$(E^* \times \Omega,\ \mathcal{B}^P \circ \mathcal{A}_H,\ P_f)$$

Teorema 2.5.1

Las σ-álgebras $\mathcal{A}_{\mathbb{N}}$ y $(\mathcal{B}^P \circ \mathcal{A}_H)$ son isomorfas, y los espacios de probabilidades

$$(\Omega_{\mathbb{N}},\ \mathcal{A}_{\mathbb{N}},\ \tilde{P}) \quad y \quad (E^* \times \Omega,\ \mathcal{B}^P \circ \mathcal{A}_H,\ P_f)$$

son isométricos.

Demostración:

Dado $\bar{A} \in \mathcal{A}_{\mathbb{N}}$, este suceso es de la forma:

$$\bar{A} = \bar{A}_1 + \ldots + \bar{A}_k \qquad \bar{A}_i \in \mathcal{A}_{H_i}/\mathbb{N}$$

Le podemos hacer corresponder el suceso de $\mathcal{B}^P \circ \mathcal{A}_H$:

$$(B_1, H_1) + \ldots + (B_k, H_k)$$

siendo B_i la imagen de $\tilde{A}_i$ por el isomorfismo entre $\mathcal{A}_{H_i}$/IN y $\mathcal{B}^p$ (teorema 2.3.2). Queda así construida una aplicación

$$\mathcal{A}_{IN} \longrightarrow \mathcal{B}^p \circ \mathcal{A}_H$$

que evidentemente es biyectiva y conserva la probabilidad,

$$\tilde{P}(\tilde{A}) = p_1 \tilde{P}_{H_1}(\tilde{A}_1) + \ldots + p_k \tilde{P}_{H_k}(\tilde{A}_k) =$$

$$= p_1 P_{F_1}(B_1) + \ldots + p_k P_{F_k}(B_k)$$

c.q.d.

La isometría descrita en el teorema anterior viene definida por la aplicación δ, que depende de las v.a. $\{Y_i\}$. Diremos entonces que el espacio $(E^* \times \Omega, \mathcal{B}^p \circ \mathcal{A}_H, P_f)$ es una δ-realización del espacio $(\Omega, \mathcal{A}, P)$.

De ahora en adelante, cuando hablemos de un indivíduo $\omega \in H_i$, nos refiriremos explícitamente a la clase de equivalencia $\tilde{\omega}$. Al correspondiente (ω^*, H_i), lo indicaremos por el indivíduo del dual ω^*, si no hay duda de la celda de donde procede.

3 DEFINICION, PROPIEDADES Y ESTIMACION ESTADISTICA DE FUNCIONES ESTIMABLES

Al interpretar el individuo medio de una población como elemento del espacio dual de las variables aleatorias observables (DEMPSTER,1969), llegamos al concepto de función estimable, que puede considerarse una versión algebraica, independiente de cualquier parametrización, de la definición introducida por BOSE (1944).

Entre las posibles expresiones sobre los individuos medios de una función estimable, se razona la forma de obtener la que llamamos Δ-óptima, que es la mejor estimación a partir de un muestreo.

3.1 INDIVIDUOS MEDIOS

La <u>esperanza matemática</u> de una v.a., es un operador lineal que define una aplicación:

$$\mathcal{E}: E \longrightarrow R$$
$$Y \longrightarrow \mathcal{E}(Y)$$

que, restringida a cada celda H_i, determina k elementos del espacio dual E^*:

(3.1.1) $\quad m_1^*,\ldots,m_k^* \qquad m_i^*(Y) = \mathcal{E}_{H_i}(Y) \qquad i=1,\ldots,k$

a los que llamaremos <u>indivíduos medios</u>. También llamaremos así a sus antiimágenes por δ (que es aplicación exhaustiva). Existen, pues, dentro de Ω, k indivíduos (clases de equivalencia):

(3.1.2) $\quad m_1,\ldots,m_k$

en donde $m_i \in H_i$ es el indivíduo medio de H_i (es decir, $\bar{m}_i \in H_i/\mathrm{IN}$ es una clase de equivalencia).

Es importante observar que entre los indivíduos medios (3.1.1) pueden existir relaciones lineales (las habrá necesariamente si $k > p = \dim E^*$), incluso ser iguales, pero, por el contrario, los indivíduos medios (3.1.2), son totalmente distintos.

Estudiaremos las relaciones métricas entre las poblaciones H_i a través de sus respectivos indivíduos medios. Pero antes vamos a generalizar el concepto de indivíduo medio.

3.2 FUNCIONES ESTIMABLES

Sea $EM^* \subset E^*$ el subespacio generado por los k indivíduos medios

$$EM^* = \langle m_1^*, \ldots, m_k^* \rangle$$

Definición 3.2.1

A los elementos de EM^*,

$$\psi^* \in EM^* \qquad \psi^* = d_1 m_1^* + \ldots + d_k m_k^*$$

les llamaremos funciones estimables (lineales).

Consideremos ahora el espacio vectorial formal generado por los k indivíduos medios (3.1.2)

$$M = \langle m_1, \ldots, m_k \rangle$$

Es fácil extender la aplicación δ (§ 2.2) al espacio M,

$$\delta : M \longrightarrow EM^* \subset E^*$$

$$\psi = \sum_{i=1}^{k} d_i m_i \longrightarrow \psi^* = \sum_{i=1}^{k} d_i m_i^*$$

Al elemento ψ de M le seguiremos llamando función estimable.

En la práctica, los indivídios medios son desconocidos, y también las funciones estimables. Un indivíduo medio es en particular función estimable. Como veremos más adelante, podrán ser estimados mediante un muestreo. De aquí la denominación.

En cuanto al espacio M deberemos considerar dos casos:

3.3.1. Caso en que dim M = k.- Si los indivíduos medios son desconocidos, y no sabemos que verifiquen ninguna relación lineal, asignaremos a M una dimensión k.

3.2.2. Caso en que dim M = r < k.- Si los indivíduos medios son desconocidos, pero, sea por razones experimentales, sea por verificación a posteriori de ciertas hipótesis lineales, se conocen s relaciones lineales, linealmente independientes, entre los individuos medios,

(3.2.1) $\qquad c_{i1}m_1 + \ldots + c_{ik}m_k = 0 \qquad i=1,\ldots,s$

Esto significa que existen r=k-s individuos, funciones estimables, $\tilde{\Psi}_1,\ldots,\tilde{\Psi}_r$, linealmente independientes en M, de forma que los k individuos medios pueden expresarse como combinación lineal de ellos,

(3.2.2) $\qquad m_i = a_{i1}\tilde{\Psi}_1 + \ldots + a_{ir}\tilde{\Psi}_r \qquad i=1,\ldots,k$

Asignaremos entonces, al espacio M, la dimensión r

$$\dim M = r$$

Es fácil hallar una matriz A

$$A = \begin{pmatrix} a_{11} \ldots & a_{1r} \\ \ldots\ldots & \ldots \\ a_{k1} \ldots & a_{kr} \end{pmatrix} \qquad \text{rang } A = r$$

a partir de $C = (c_{ij})$ de (3.2.1). Basta completar las s filas de C en otros r vectores fila ortogonales a los de C, es decir, la matriz A debe verificar:

(3.2.3) $\qquad C.A = 0$

Dada entonces una función estimable $\Psi = \sum_{i=1}^{k} d_i m_i$, tendrá una expresión única en la base $(\tilde{\Psi}_1,\ldots,\tilde{\Psi}_r)$,

(3.2.4) $\qquad \Psi = b_1\tilde{\Psi}_1 + \ldots + b_r\tilde{\Psi}_r$

Cualquiera que sea la expresión $D = (d_1\ldots d_k)'$ de Ψ en función de los individuos medios, deberá verificar:

(3.2.5) $\qquad B = A'.D \qquad$ siendo $B = (b_1\ldots b_k)'$

<u>Definición 3.2.1</u>

Diremos que las funciones estimables $\psi^*_1,\dots,\psi^*_m$ son <u>M-linealmente independientes</u>, si es linealmente independiente, dentro de M, el sistema

$$\psi_1 = \delta^{-1}(\psi^*_1),\dots,\psi_m = \delta^{-1}(\psi^*_m)$$

El siguiente teorema, cuya demostración es inmediata, nos proporciona las primeras propiedades de las funciones estimables.

<u>Teorema 3.2.1.</u>

Se verifica:

1) Toda combinación lineal de funciones estimables, es también función estimable.

2) El nº máximo de funciones estimables, linealmente independientes en el espacio EM^*, es mín (p,r), siendo $r = \dim M$.

3.3 LA APLICACION ε^*

Dado un indivíduo $i \in \Omega$, tendremos

(3.3.1) $$i^* = m^*_j + e^*_{ji} \quad \text{si} \quad i \in H_j,$$

siendo $e^*_{ji} = i^* - m^*_j$. Un indivíduo, dentro del dual E^*, se descompone en la suma de un indivíduo medio más un <u>indivíduo error</u> e^*_{ji}. Se interpreta (3.3.1) en el sentido de que i^* está "alrededor" del indivíduo medio m^*_j, en relación con la métrica en E^*.

<u>Definamos</u>, pues, la aplicación:

$$\varepsilon^* : \Omega \longrightarrow EM^*$$

(3.3.2) $$\varepsilon^*(i) = m^*_j \quad \text{si} \quad i \in H_j$$

que a cada indivíduo le hace corresponder el indivíduo medio de la celda a la cual pertenece.

3.4. MUESTREO

Para muestras de tamaños $n_1, n_2, \dots, n_k$ de cada una de las poblaciones que integran Ω, consideremos el espacio muestral:

$$(3.4.1) \qquad H_1^n \times \dots \times H_k^n$$

siendo $H_i^n = H_i \times \overset{n}{\dots} \times H_i$ el espacio muestral sobre la celda H_i. Una muestra aleatoria simple de $n = \sum n_i$ indivíduos es el resultado de una <u>extracción</u> (ó de la prueba de una experiencia)

$$\mathfrak{m} = (i_{11} \dots i_{1n_1}; \dots ; i_{k1} \dots i_{kn_k}) \in H_1^{n_1} \times \dots \times H_k^{n_k}$$

Indicaremos por $\mathfrak{m}_\Delta$ el conjunto de las posibles extracciones de tamaño $\Delta = (n_1, \dots, n_k)$, e introduciremos n elementos,

$$(\rho_{11} \dots \rho_{1n_1}; \dots ; \rho_{k1} \dots \rho_{kn_k})$$

de modo que, dada una extracción $\mathfrak{m} \in \mathfrak{m}_\Delta$, a partir de ρ_{jh} se obtiene el indivíduo muestral $i_{jh} = \rho_{jh}(\mathfrak{m})$.

A los elementos $\{\rho_{jh}\}$ les llamaremos <u>indivíduos posibles</u>.

Sea ahora N el espacio vectorial formal generado por los indivíduos posibles, de dimensión n, y con un producto escalar de modo que, <u>por definición</u>, la base $\{\rho_{jh}\}$ sea <u>ortonormal</u>.

Desde luego, no hay dificultad en extender la aplicación δ de Ω a $H_1^{n_1} \times \dots \times H_k^{n_k}$, de modo que a la muestra $\mathfrak{m}$ le hace corresponder n indivíduos de E*,

$$\mathfrak{m} \xrightarrow{\delta} \mathfrak{m}^* = (i_{11}^* \dots i_{1n_1}^*; \dots ; i_{k1}^* \dots i_{kn_k}^*)$$

Análogamente, la aplicación $\mathcal{E}^*$ también se puede extender de Ω al espacio N,

$$\mathcal{E}^*: N \longrightarrow EM^*$$

$$\mathcal{E}^*\left(\sum_{i=1}^{k} \sum_{h=1}^{n_i} c_{ih}\rho_{ih}\right) = \sum_{i=1}^{k} \left(\sum_{h=1}^{n_i} c_{ih}\right) m_i^*$$

3.5. ESTIMACION ESTADISTICA DE FUNCIONES ESTIMABLES

Propongámonos ahora estimar una función estimable ψ^* a partir de un muestreo.

Definición 3.5.1

Diremos que $\rho \in N$, $\rho = \sum_{i=1}^{k} \sum_{h=1}^{n_i} c_{ih}\rho_{ih}$ es un estimador lineal insesgado de la función estimable $\psi^* \in EM^*$, si

$$(3.5.1) \qquad \mathcal{E}^*(\rho) = \psi^*$$

Lamaremos dispersión del estimador insesgado ρ, al módulo (al cuadrado) $\|\rho\|^2$, de acuerdo con la métrica en N.

Teorema 3.5.1

Supongamos que dim M = k. Sea $\psi^* = \sum_{i=1}^{k} d_i m_i^*$ una función estimable. Para un muestreo de tamaño $n = \sum n_i$, indiquemos,

$$\bar{\rho}_i = \frac{1}{n_i} \sum_{h=1}^{n_i} \rho_{ih} \qquad i=1,\ldots,k$$

Entonces $\hat{\psi} = \sum_{i=1}^{k} d_i \bar{\rho}_i$ es un estimador lineal insesgado de ψ^*, con dispersión mínima.

Demostración:

Para que $\rho = \sum_{i=1}^{k} \sum_{h=1}^{n_i} c_{ih}\rho_{ih}$ sea un estimador lineal insesgado de ψ^*, deberá verificarse (3.5.1), es decir,

$$(3.5.2) \qquad \mathcal{E}^*(\rho) = \sum_{i=1}^{k} \left(\sum_{h=1}^{n_i} c_{ih}\right) m_i^* = \sum_{i=1}^{k} d_i m_i^*$$

$$d_i = \sum_{h=1}^{n_i} c_{ih} \qquad (i=1,\ldots,k)$$

La dispersión de ρ es función de los coeficientes c_{ih}

$$\|\rho\|^2 = \sum_{i=1}^{k} \sum_{h=1}^{n_i} c_{ih}^2 = F(c_{11}\ldots c_{1n_1};\ldots;c_{k1}\ldots c_{kn_k})$$

De la función F se pueden eliminar k variables poniendo:

$$c_{in_i} = d_i - \sum_{h=1}^{n_i-1} c_{ih} \qquad (i=1,\dots,k)$$

Entonces, para localizar el mínimo de F,

$$\frac{\partial F}{\partial c_{ih}} = 0 \Rightarrow 2a_{ih} - 2(d_i - \sum_{h=1}^{n_i-1} c_{ih}) = 0 \qquad \begin{array}{l} i=1,\dots,k \\ h=1,\dots,n_i-1 \end{array}$$

obtenemos (n-k) ecuaciones, cuya solución es:

$$c_{11} = \dots = c_{1n_1} = \frac{d_1}{n_1}; \dots; c_{k1} = \dots = c_{kn_k} = \frac{d_k}{n_k}$$

c. q. d.

La mínima dispersión que se puede alcanzar, suponiendo dim M= k, para estimar ψ^*, es:

$$(3.5.3) \qquad \|\rho\|^2_{mín} = \frac{d_1^2}{n_2} + \dots + \frac{d_k^2}{n_k}$$

Corolario 3.5.1

Si dim M=k, $\bar{\rho}_i = \frac{1}{n_i} \sum_{h=1}^{n_i} \rho_{ih}$ es el estimador insesgado, de dispersión mínima, del individuo medio m_i^*, (i=1,...,k).

El teorema 3.5.1 nos da una estimación centrada, con dispersión mínima, entre todas las estimaciones lineales, de la función estimable ψ^*. Hemos supuesto dim M = k, lo que implica que la expresión de ψ^* en función de los individuos medios es única.

Si dim M = r < k, entonces ψ^* admite distintas expresiones en función de los individuos medios. El problema de la estimación de ψ^* se complica, entonces, un poco más porque, entre todas estas expresiones habrá que hallar aquella que minimice la dispersión (3.5.3).

Teorema 3.5.2

Supongamos dim M = r < k, y sea $\Psi^* = \sum_{i=1}^{k} d_i m_i^*$ una función estimable. De un muestreo de tamaño $\Delta = (n_1, \ldots, n_k)$, construyamos la matriz diagonal:

$$\Delta = \begin{pmatrix} n_1 & & \\ & \ddots & \\ & & n_k \end{pmatrix}$$

Entonces, si A es una matriz tal que C.A=O (§ 3.2), un estimador lineal insesgado de Ψ^*, con dispersión mínima es:

$$(3.5.4) \qquad \hat{\Psi} = \hat{d}_1 \bar{\rho}_1 + \ldots + \hat{d}_k \bar{\rho}_k$$

en donde $\hat{D} = (\hat{d}_1 \ldots \hat{d}_k)'$ verifica:

$$(3.5.5) \qquad \hat{D} = \Delta A (A' \Delta A)^{-1} A.D \qquad (D = (d_1 \ldots d_k)')$$

La expresión (3.5.4) es única.

Demostración:

Fijando los coeficientes $(d_1, \ldots d_k)$, la dispersión mínima será:

$$F(d_1, \ldots, d_k) = \frac{d_1^2}{n_1} + \ldots + \frac{d_k^2}{n_k}$$

Debemos hallar unos nuevos coeficientes $(\hat{d}_1 \ldots \hat{d}_k)$ que minimicen esta función, con la condición $B = A'.D$ (§3.2.5).

Introduciendo r multiplicadores de Lagrange $\lambda_1, \ldots, \lambda_r$, consideremos la función:

$$F = \sum_{i=1}^{k} \frac{d_i^2}{n_i} + 2 \sum_{i=1}^{r} \lambda_i (b_i - a_{1i} d_1 - \ldots - a_{ki} d_k)$$

$$\frac{\partial F}{\partial d_i} = 0 \Rightarrow \frac{d_i}{n_i} - \lambda_1 a_{1i} - \ldots - \lambda_r a_{ri} = 0 \qquad i = 1, \ldots, k$$

Es notación matricial: $D = \Delta . A\lambda$, siendo $\Delta = \begin{pmatrix} n_1 & & \\ & \ddots & \\ & & n_k \end{pmatrix}$

y $\lambda = (\lambda_1 \ldots \lambda_r)'$. Combinando esta ecuación matricial con $B = A'.D$, tendremos:

$$B = A'.D = A' \Delta A \lambda \Rightarrow \lambda = (A' \Delta A)^{-1}.B, \text{ luego}$$

$$\hat{D} = \Delta A(A' \Delta A)^{-1}B \;, \qquad B = A'D$$

Demostremos ahora la unicidad de $\hat{D}$. Si en lugar de A, utilizamos otra matriz $A_1 = A.T$ tal que $C.A_1 = 0$, entonces, indicando $B_1 = A_1'.D = T'A'D$,

$$\hat{D}_1 = \Delta A_1 (A_1' \Delta A_1)^{-1} B_1 = \Delta A.T(T'A' \Delta A.T)^{-1}T'A'D =$$

$$= \Delta A.T.T^{-1}(A'\Delta A)^{-1}(T^{-1})'T'A'D = \hat{D}$$

c. q. d.

3.6. EXPRESION Δ-OPTIMA DE UNA FUNCION ESTIMABLE

La estimación insesgada y de dispersión mínima de la función estimable ψ^* la indicaremos por

$$\hat{\psi} = \sum_{i=1}^{k} \hat{d}_i \bar{\rho}_i$$

Depende de la muestra de tamaño $\Delta = (n_1 \ldots n_k)$.

LLamaremos <u>expresión Δ-óptima</u> de ψ^* a

$$(3.6.1) \qquad \psi^* = \sum_{i=1}^{k} \hat{d}_i m_i^*$$

Aunque ψ^* pueda expresarse de más de una manera, como combinación lineal de los individuos medios (si dim $M=r < k$), la expresión Δ-óptima <u>es única</u>.

Vamos a obtener ahora las componentes en E^* de la imagen de $\hat{\psi}$ por la aplicación que define un muestreo $m \in \mathfrak{m}_\Delta$

$$\hat{\psi} = \sum_{i=1}^{k} \hat{d}_i \bar{\rho}_i \xrightarrow{m} \sum_{i=1}^{k} \hat{d}_i \left(\frac{1}{n_i} \sum_{h=1}^{n_i} i_{ih}\right) = \hat{\psi}^* \in E^*$$

Este muestreo, aplicado a una v.a. $Y \in E$, nos dará:

$$(y_{11},\ldots,y_{1n_1};\ldots;y_{k1},\ldots,y_{kn_k})$$

siendo $y_{ih} = t^*_{ih}(Y)$.

En particular, indicaremos $y_{jih} = t^*_{ih}(Y_j) \quad j=1,\ldots,p$

La expresión de $\hat{\psi}^* = m(\hat{\psi})$ en la base dual (Y^*_i) será:

$$\hat{\psi}^* = \hat{\psi}^*(Y_1)Y_1 + \ldots + \hat{\psi}^*(Y_p)Y_p$$

es decir, indicando

$$\bar{y}_{ji.} = \frac{1}{n_i} \sum_{h=1}^{n_i} y_{jih}$$

tendremos:

(3.6.2) $$\hat{\psi}^* = (\sum_i d_i \bar{y}_{1i.})Y^*_1 + \ldots + (\sum_i d_i \bar{y}_{pi.})Y^*_p$$

4 RELACIONES CON LA TEORIA DE FUNCIONES PARAMETRICAS ESTIMABLES

Introduciendo un diseño factorial, y por tanto, una parametrización, se expone la noción de función paramétrica, y las condiciones de estimabilidad en relación con la teoría de funciones estimables. Para este fin, hemos utilizado la obra de SCHEFFE (1959).

La matriz del diseño se define solamente dependiente de las poblaciones, y según esta matriz reducida, se determinan las ecuaciones normales para hallar las estimaciones LS de los parámetros. Consultando a RAO (1965), hemos establecido la conexión entre la estimación paramétrica de una función estimable, según el teorema de Gauss-Markov, y la expresión Δ-óptima obtenida en el capítulo anterior.

Otras referencias consultadas: COCHRAN y COX(1966), MOOD(1955), MORRISON(1967), RAO(1945a,1945b), SEAL(1964).

4.1. DISEÑO FACTORIAL

En §3.2, caso en que dim $M = r$, se han expresado los k individuos medios en función de r funciones estimables desconocidas

(4.1.1) $$m_i = a_{i1}\tilde{\psi}_1 + \dots + a_{ir}\tilde{\psi}_r \qquad i=1,\dots,k$$

La correspondiente expresión, imagen por δ, en el espacio dual E^*, será

(4.1.2) $$m_i^* = a_{i1}\tilde{\psi}_1^* + \dots + a_{ir}\tilde{\psi}_r^*$$

Generalizando (4.1.1), supongamos que tenemos m individuos desconocidos de P,

$$\beta_1, \dots, \beta_m$$

a los que llamaremos <u>indivíduos paramétricos</u>.

Sea $E\beta = \langle\beta_1, \dots, \beta_m\rangle$ el espacio vectorial formal generado por los indivíduos paramétricos, con dim $E\beta = m$.

La aplicación δ nos proporcionará otro subespacio de E^*,

$$E\beta^* = \langle\beta_1^*, \dots, \beta_m^*\rangle \text{ , siendo } \delta(\beta_i) = \beta_i^* \in E^*.$$

<u>Definición 4.1.1</u>

Un <u>diseño factorial</u> de matriz de diseño X, es una aplicación lineal

$$M \xrightarrow{X} E\beta \qquad \text{rang } X = r = \dim M$$

tal que

(4.1.3) $$m_i = x_{i1}\beta_1 + \dots + x_{im}\beta_m \qquad i=1,\dots,k$$

La matriz $X = (x_{ij})$, de orden (k,m) y rango r, es <u>conocida</u>, y se construye de acuerdo con las condiciones experimentales del diseño.

Por δ, (4.1.3) se transforma en:

(4.1.4) $$m_i^* = x_{i1}\beta_1^* + \dots + x_{im}\beta_m^* \qquad i=1,\dots,k$$

La aplicación que define X presupone $M \subset E^\beta$.

Puesto que $M = \mathrm{im}X$, si $m=r$, podremos identificar los indivíduos paramétricos $(\beta_1,\ldots,\beta_r)$ con r funciones estimables $(\psi_1,\ldots,\psi_r)$. Si $m>r$, un indivíduo paramétrico no podrá ser considerado función estimable.

Sea ahora Y una variable aleatoria de E. Indiquemos por:

$$\beta_{Yj} = \beta_j^*(Y) \qquad j=1,\ldots,m$$

los parámetros correspondientes a la v.a. Y.

Indiquemos, asimismo, el vector paramétrico

$$\beta_Y = (\beta_{Y1},\ldots,\beta_{Ym})'$$

En particular, para una v.a. de la base $\{Y_i\}$ pondremos $\beta_{Y_i} = (\beta_{i1},\ldots,\beta_{im})'$

Entonces, si $X_1',\ldots,X_k'$ son los vectores fila de la matriz X, la expresión de los indivíduos medios en la base dual $\{Y_i^*\}$, será:

$$(4.1.5) \qquad m_i^* = (X_i'\beta_{Y_1})Y_1^* + \ldots + (X_i'\beta_{Y_p})Y_p^* \qquad i=1,\ldots,k$$

Además,

$$(4.1.6) \qquad m_i^*(Y) = x_{i1}\beta_{Y1} + \ldots + x_{im}\beta_{Ym}, \quad \forall\, Y \in E$$

4.2. FUNCIONES PARAMETRICAS ESTIMABLES

<u>Definición 4.2.1</u>

A los elementos de $E^{\beta *}$ les llamaremos <u>funciones paramétricas</u>

$$\psi^* \in E^{\beta *}, \qquad \psi^* = p_1\beta_1^* + \ldots + p_m\beta_m^*$$

Refiriendo ψ^* a la base dual, si $P = (p_1,\ldots,p_m)'$,

$$(4.2.1) \qquad \Psi^* = (P'\beta_{Y_1})Y_1^* + \ldots + (P'\beta_{Y_p})Y_p^*$$

y una función paramétrica es una generalización de (4.1.5).

Ψ^* tendrá la correspondiente expresión en $E\beta$,

$$\Psi = p_1\beta_1 + \ldots + p_m\beta_m$$

<u>Definición 4.2.2</u>

Diremos que la función paramétrica $\psi^* \in E\beta^*$ es <u>estimable</u>, si Ψ^* es función estimable, es decir, si $\psi^* \in EM^*$.

Toda función paramétrica estimable Ψ^* (f.p.e.), se expresará como combinación lineal de los indivíduos medios:

$$\Psi^* = \sum_{i=1}^{m} p_i\beta_i^* = \sum_{i=1}^{k} d_i m_i^*$$

Recíprocamente, a través de la aplicación X, toda función estimable $\Psi^* = \sum_{i=1}^{k} d_i m_i^*$, es también una función paramétrica estimable. De esta manera, hemos identificado ambos conceptos, uno introducido linealmente a partir de los indivíduos medios, y el otro a partir de los parámetros.

Podremos aplicar a toda f.p.e. las propiedades de las funciones estimables de § 3.2, y la estimación, cuando se conoce su expresión en función de los indivíduos medios, expuesta en § 3.5.

4.3. CONDICIONES DE ESTIMABILIDAD

Una función paramétrica ψ^* viene definida por el vector P. Para que sea estimable, debe existir alguna relación entre P y la matriz del diseño factorial X, que nos pueda permitir expresarla en función de los individuos medios.

Teorema 4.3.1

Una condición necesaria y suficiente para que la función paramétrica

$$\psi^* = \sum_{i=1}^{p} P'\beta_{Y_i} Y_i^*$$

sea estimable, es que exista un estimador lineal insesgado:

$$\hat{\psi} = \sum_{j,h} c_{jh}\rho_{jh} \quad \text{de } \psi^*.$$

Demostración:

La necesidad es consecuencia del teorema 3.5.2, que además nos proporciona el estimador de dispersión mínima.

Supongamos ahora que existe $\hat{\psi} = \sum_{j,h} c_{jh}\rho_{jh}$ tal que

$$\mathcal{E}^*(\hat{\psi}) = \psi^*$$

Puesto que

$$\mathcal{E}^*(\hat{\psi}) = \sum_{j}\left(\sum_{h} c_{jh}\right)m_j^*$$

deducimos que $\psi^* \in EM_s^*$ y la condición es también suficiente.

c. q. d.

Teorema 4.3.2

Sea $\psi^* = \sum_{i=1}^{p} p'\beta_{Y_i} Y_i^*$ una función paramétrica. La condición necesaria y suficiente para que ψ^* sea estimable es que P' sea combinación lineal de los vectores fila de X,

(4.3.1) $P' = d_1 X_1' + \ldots + d_k X_k'$

y entonces la expresión de ψ^* en función de los individuos medios será:

(4.3.2) $\psi^* = d_1 m_1^* + \ldots + d_k m_k^*$

Demostración:

Si ψ^* es estimable, y $\psi^* = \sum_i d_i m_i^*$, entonces, aplicando (4.1.5),

$$\psi^* = \sum_{i=1}^{p} p'\beta_{Y_i} Y_i^* = \sum_{j=1}^{k} d_j m_j^* = \sum_j d_j \left(\sum_i X_j' \beta_{Y_j} Y_j^* \right) =$$

$$= \sum_{i=1}^{p} \left(\sum_j d_j X_j' \right) \beta_{Y_i} Y_i^*$$

Comparando términos deducimos $P' = \sum_j d_j X_j'$

Recíprocamente, si $P' = \sum_j d_j X_j'$, entonces, invirtiendo el razonamiento, obtendríamos $\psi^* = \sum_j d_j m_j^*$

c. q. d.

Corolario 4.3.1

Si la función paramétrica $\psi^* = \sum_{i=1}^{p} p\beta_{Y_i} Y_i^*$ es estimable, las componentes $D = (d_1, \ldots, d_k)'$ de (4.3.2) son solución de la ecuación:

(4.3.3) $P = X'D$

Corolario 4.3.2

Sea $r = \text{rang}(X)$. Entonces:

1) Si $r = m$, toda función paramétrica es estimable.

2) Si $r < m$, existen funciones paramétricas no estimables

3) Si $r = k$, una función paramétrica se expresa de forma única en función de los indivíduos medios. Si $r < k$, podrá expresarse entonces de infinitas maneras.

4.4. ESTIMACIÓN POR MÍNIMOS CUADRADOS

Propongámonos ahora, estimar los indivíduos paramétricos $\beta_1^*,\dots,\beta_m^*$, que son en particular funciones paramétricas (no necesariamente estimables), a partir de una muestra de indivíduos de tamaño $\Delta = (n_1,\dots,n_k)$,

$$\Pi^* = (i_{11},\dots,i_{1n_1};\dots;i_{k1},\dots,i_{kn_k})$$

Utilizando el criterio de los mínimos cuadrados, definamos la aplicación:

$$(4.4.1) \qquad \mathcal{L}(\beta^*): E \longrightarrow R \qquad \mathcal{L}(\beta^*) = \sum_{i,h} (i_{ih}^* - m_i^*)^2$$

$$(4.4.2) \qquad \mathcal{L}(\beta^*)(Y) = \sum_i \sum_h (y_{ih} - x_{i1}\beta_{Y1} - \dots - x_{im}\beta_{Ym})^2$$

Deseamos obtener la estimación $\hat{\beta}^*$ de los indivíduos paramétricos

$$\beta^* = (\beta_1^*,\dots,\beta_m^*)$$

a partir de la muestra Π^*, de modo que

$$\mathcal{L}(\hat{\beta}^*)(Y) = \min \mathcal{L}(\beta^*)(Y) \qquad \forall\ Y \in E$$

El método de los mínimos cuadrados nos llevará a las ecuaciones normales (4.4.4), que aquí presentaremos en la forma más simplificada (4.4.5), que está más en conexión con lo expuesto en el capítulo 3, y es más fácil de resolver. En efecto, en (4.4.5), en lugar de considerar a todos los $\Delta = (n_1,\dots,n_k)$ valores muestrales, se utilizan solamente los individuos medios muestrales de cada celda.

El muestreo sobre una v.a. Y, lo indicaremos

$$mY = (y_{11},\dots,y_{1n_1};\dots;y_{k1},\dots,y_{kn_k})$$

Indicaremos, además, por $Y^{(n)}$, el <u>vector</u> de n v.a. independientes que Y define sobre

$$H_1^{n_1} \times \dots \times H_k^{n_k}$$

Ampliaremos la matriz X para obtener la <u>matriz del diseño ampliada</u>:

$$X_a = \begin{pmatrix} x_{11}\dots x_{1m} \\ \dots\dots\dots \\ x_{11}\dots x_{1m} \\ \dots\dots\dots \\ x_{k1}\dots x_{km} \\ \dots\dots\dots \\ x_{k1}\dots x_{km} \end{pmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} n_1 \\ \vdots \\ \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} n_k \end{matrix}$$

Aplicando el diseño factorial (4.1.3) a cada una de las v.a. del vector $Y^{(n)}$, obtendremos el siguiente modelo en función de los parámetros

$$(4.4.3) \qquad \mathcal{E}(Y^{(n)}) = X_a\beta_Y \qquad \forall\ Y \in E$$

<u>Teorema 4.4.1</u>

El vector de parámetros $\hat{\beta}^* = (\hat{\beta}^*_1,\dots,\hat{\beta}^*_m)$ que minimiza

$$\mathcal{L}(\beta^*)(Y) = \sum_i \sum_h (y_{ih} - x_{i1}\beta_{Y1} - \dots - x_{im}\beta_{Ym})^2$$

es solución de la ecuación

$$X' \Delta X \beta^* = X' \Delta \overline{\Pi}^*$$

siendo

$$\Delta = \begin{pmatrix} n_1 & & \\ & \ddots & \\ & & n_k \end{pmatrix} \quad \bar{\Pi}^* = \left(\frac{1}{n_1}\sum_h i^*_{1h},\ldots,\frac{1}{n_k}\sum_h i^*_{kh}\right),$$

Demostración:

Resolviendo las m ecuaciones:

$$\frac{\partial\mathscr{L}(\beta^*)}{\partial\beta_{Y1}} = \ldots = \frac{\partial\mathscr{L}(\beta^*)}{\partial\beta_{Ym}} = 0$$

a fin de hallar el mínimo de (4.4.2), se obtienen las ecuaciones normales,

(4.4.4) $\quad X'_a X_a \beta_Y = X'_a mY$

Formalmente, podemos poner:

$$X'_a X_a \beta^*(Y) = X'_a X_a \beta_Y$$

$$X'_a \Pi^*(Y) = X'_a mY$$

luego, (4.4.4) nos dará

$$X'_a X_a \beta^* = X'_a \Pi^*$$

Es fácil comprobar que

$$X'_a X_a = X' \Delta X$$
$$X'_a \Pi^* = X' \Delta \bar{\Pi}^*$$

de donde

(4.4.5) $\quad X' \Delta X \beta^* = X' \Delta \bar{\Pi}^*$ $\qquad$ c. q. d.

A la estimación $\hat{\beta}^*$ de β^* que verifica (4.4.5) , la llamaremos estimación LS. Los indivíduos $(\hat{\beta}^*_1,\ldots,\hat{\beta}^*_m)$ son combinación lineal de los n indivíduos Π^* obtenidos después de un muestreo. Indicaremos por $\hat{\beta}=(\hat{\beta}_1,\ldots,\hat{\beta}_m)$ la estimación LS, en la que se sustituyen los indivíduos muestrales por los indivíduos posibles. Los elementos $(\hat{\beta}_1,\ldots,\hat{\beta}_m)$ son aleatorios, combinación lineal, según (4.4.5), de los indivíduos medios posibles $\bar{\rho}_1,\ldots,\bar{\rho}_k$. Desde luego, si $m > r$, los indivíduos paramétricos no son funciones estimables, y las estimaciones LS no son insesgadas.

4.5. TEOREMA DE GAUSS-MARKOV Y OBTENCION DE LA EXPRESION Δ-OPTIMA

El teorema 3.5.2 nos proporcionaba un estimador insesgado y de dispersión mínima de una función estimable, a través de la expresión Δ-óptima, en función de los indivíduos medios. El teorema de Gauss-Markov, contenido en la primera parte del teorema 4.5.1, da otra forma equivalente de obtener este estimador, pero en función de las estimaciones LS de los parámetros.

Lema 4.5.1

Si $P = X'D$ para algún D, entonces existe un vector D_1 tal que $P = X' \Delta X D_1$, y recíprocamente.

Además, el vector $D = \Delta X D_1$, es único.

En efecto: $P = X'D$ significa que P pertenece al subespacio S generado por los vectores columna de X'. Sea ahora V un vector columna ortogonal a S

$$V \perp S \Rightarrow V'X' = 0 \Rightarrow V'X' \Delta X = 0$$

luego V es también ortogonal al subespacio T generados por los vectores columna de $X' \Delta X$. Si ahora tenemos un vector V ortogonal a T,

$$V \perp T \Rightarrow V'X' \Delta X = 0 \Rightarrow V'(X' \Delta^{-1/2})(\Delta^{-1/2} X)V = 0$$

$$\Rightarrow \Delta^{-1/2} X . V = 0 \Rightarrow X . V = 0$$

luego, los ortocomplementarios de S y T coinciden, y por tanto S = T. Esto significa que existe un vector D_1 tal que

$$P = X' \Delta X D_1$$