

**UTILIZACIÓN DE MÉTRICAS RIEMANNIANAS
EN ANÁLISIS DE DATOS MULTIDIMENSIONALES
Y SU APLICACIÓN A LA BIOLOGÍA**

JOSE M^a OLLER SALA

BARCELONA, 25 de NOVIEMBRE de 1982.

Sea $E = \{P_1, P_2, \dots, P_k\}$. Una preordenación sobre E , asociada a una distancia d , es una relación binaria sobre $E \times E$ definida por:

$$(P_i, P_j) \leq (P_k, P_h) \Leftrightarrow d(P_i, P_j) \leq d(P_k, P_h) \quad (19)$$

Es inmediato demostrar que dicha relación es un preorden, al verificar las propiedades reflexiva y transitiva.

Exigiremos que la preordenación asociada a D sea la misma que la preordenación asociada a \hat{D} .

Existen varios procedimientos para obtener \hat{D} , Shepard (1962), Lingoes (1971), Mardia (1978), uno de los más generales es el denominado método de la transformación monótona, definido a través de:

$$\hat{d}_{ij} = f(d_{ij}) \quad (20)$$

donde f es una función monótona lineal o no lineal. El resultado de aplicar varias veces (20) es una configuración de puntos euclídeos que realiza la preordenación asociada a D . Más detalles pueden encontrarse en Cuadras (1981).

6.2.2. Caso de no disponer de una expresión analítica de la distancia.

Este es un caso más desfavorable que el anterior y puede intentar resolverse de varias formas posibles.

6.2.2.a. Una primera forma consiste en utilizar técnicas de cálculo numérico para hallar la distancia entre dos puntos de la variedad. El problema a resolver es un problema de contorno a saber: dados dos puntos de la variedad hallar una geodésica que los una e integrar el elemento de línea a lo largo de la curva hallada.

Uno de los procedimientos posibles consiste en trabajar con otro tensor métrico definido por la siguiente expresión matricial:

$$H = \lambda G + (1-\lambda)I \quad \lambda \in [0,1] \quad (21)$$

de forma que si $\lambda=0$ la métrica sea la euclídea y si $\lambda=1$ la métrica resultante sea la definida en el espacio paramétrico originalmente.

El principal problema para hallar una geodésica que parta del punto A ($s=0$) y llegue al punto B, consiste en determinar los valores, en el punto A, de las derivadas de las coordenadas de la curva, como funciones de s , $\dot{x}^\gamma(0)$, $\gamma=1, \dots, n$, ya que una vez dados estos, junto con las coordenadas del punto A, la geodésica queda bien determinada.

Para $\lambda=0$, la geodésica buscada será una recta, y dados los puntos A: (a^1, \dots, a^n) y B: (b^1, \dots, b^n) es inmediato hallar las derivadas en el origen:

$$\dot{x}^\gamma(0) = \frac{b^\gamma - a^\gamma}{\sqrt{\sum_{i=1}^n (b^i - a^i)^2}} \quad \gamma=1, \dots, n \quad (21)$$

El siguiente paso consiste en incrementar ligeramente λ , en $\Delta\lambda$, y utilizar como derivadas en el origen las obtenidas en (21). El resultado es una geodésica que, generalmente, no pasará por el punto B, aunque se acercará a él, tanto más cuanto menor sea $\Delta\lambda$ y más próximo sea A de B. Deberá, a continuación, iniciarse un proceso iterativo consistente en modificar los valores de las derivadas en el origen hasta obtener una geodésica que partiendo de A alcance el punto B. Una vez obtenida, se incrementa de nuevo λ en $\Delta\lambda$ y se repite el proceso iterativo anteriormente citado, usando como aproximación inicial de los valores de las derivadas en el origen, los obtenidos al final del proceso, iterativo del ajuste anterior.

El proceso de incrementar λ se repite las veces necesarias hasta que $\lambda=1$. En esta última etapa se obtiene la geodésica buscada y asociada a ésta la distancia, medida sobre la geodésica, entre A y B.

Con éste método, u otros parecidos, pueden evaluarse, numéricamente, las distancias entre los puntos P_1, \dots, P_k , obtener la matriz de interdistancias D , y proceder como en el apartado 6.2.1.

Veamos a continuación otra forma de proceder:

6.2.2.b. Aunque no sea posible obtener una transformación de coordenadas tal que deje constante al tensor métrico en toda la variedad (espacio no euclídeo), si es posible encontrar un sistema de coordenadas tal que en un punto cualquiera P se anulen todas las derivadas parcia-

les del tensor métrico, y por tanto todos los símbolos de Christoffel en este punto. Ello significa que en un entorno de P podemos considerar al tensor métrico aproximadamente constante, y calcular las distancias entre puntos como si de un espacio euclídeo se tratara. Procedamos a continuación a hallar las ecuaciones que definirán dicho sistema de coordenadas. Sea la transformación:

$$(x^1, \dots, x^n) \rightarrow (y^1, \dots, y^n) \quad (22)$$

tal que bajo el sistema de referencia (y^1, \dots, y^n) las derivadas parciales del tensor métrico, y por tanto los símbolos de Christoffel, se anulen en P .

Entonces, si designamos por $(\Gamma_{\beta\gamma}^{\alpha})_P$ el valor de los símbolos de Christoffel de segunda especie, respecto (x^1, \dots, x^n) , en el punto P , teniendo en cuenta el convenio de sumación de los índices repetidos y (41) del capítulo segundo, se verifica:

$$\frac{\partial^2 x^{\alpha}}{\partial y^{\lambda} \partial y^{\mu}} + (\Gamma_{\beta\gamma}^{\alpha})_P \frac{\partial x^{\beta}}{\partial y^{\lambda}} \frac{\partial x^{\gamma}}{\partial y^{\mu}} = 0 \quad (23)$$

$$\alpha, \lambda, \mu = 1, \dots, n$$

dicho sistema de ecuaciones diferenciales en derivadas parciales admite infinitas soluciones, una de las cuales, Sokolnikoff (1971), viene definida por el sistema algebraico:

$$x^{\alpha} = x_P^{\alpha} + y^{\alpha} - \frac{1}{2} (\Gamma_{\beta\gamma}^{\alpha})_P y^{\beta} y^{\gamma} \quad \alpha=1, \dots, n \quad (24)$$

Nótese que las coordenadas del punto P, en el nuevo sistema de referencia, son $(0, \dots, 0)$. Además,

$$\frac{\partial x^\alpha}{\partial y^\mu} = \delta_\mu^\alpha - (\Gamma_{\beta\mu}^\alpha)_P y^\beta \quad (25)$$

por tanto el tensor métrico, bajo el nuevo sistema de coordenadas, viene dado por:

$$\bar{g}_{\mu\nu} = (\delta_\mu^\alpha - (\Gamma_{i\mu}^\alpha)_P y^i) (\delta_\nu^\beta - (\Gamma_{j\nu}^\beta)_P y^j) g_{\alpha\beta} \quad (26)$$

y en el punto P:

$$\bar{g}_{\mu\nu}(0, \dots, 0) = g_{\mu\nu}(x_P^1, \dots, x_P^n) \quad (27)$$

Consideremos ahora que tenemos k puntos P_1, \dots, P_k , de coordenadas, respecto el sistema (x^1, \dots, x^n) , $P_i: (x_i^1, \dots, x_i^n)$ $i=1, \dots, k$.

Un procedimiento razonable, para evaluar las distancias ente los puntos, consiste en considerar al punto P: $(\bar{x}^1, \dots, \bar{x}^n)$ con:

$$\bar{x}^i = \frac{1}{k} \sum_{j=1}^k x_j^i \quad i=1, \dots, n \quad (28)$$

y aplicar la transformación definida por (24), con P definido en (28). Así, en dicho punto P("centroide") se anularan, en el nuevo sistema de coordenadas, todos los símbolos de Christoffel y, equivalentemente, todas las derivadas parciales de las componentes del tensor métrico. En-

tonces, si los puntos P_1, \dots, P_k son próximos a D , podemos calcular las distancias entre ellos considerando las componentes del tensor métrico como constantes, iguales al valor que toman en el punto P , (27). Equivale a considerar que los puntos P_i , referidos al sistema de coordenadas (y^1, \dots, y^n) son puntos de un espacio euclídeo, cuyo producto escalar, en este sistema de referencia, viene dado por la matriz $G = (\bar{g}_{\mu\nu})$. En este punto es posible efectuar una transformación lineal de coordenadas para reducir al tensor métrico en el punto P , a la identidad. Como G es simétrica podemos escribir:

$$G = T D T^t \quad (29)$$

por tanto una matriz de cambio de base que reduzca al tensor métrico a la identidad vendrá dada por:

$$P = T D^{-1/2} \quad (30)$$

Efectuada dicha transformación, los k puntos están referidos a un sistema de coordenadas ortonormal. En estas condiciones el problema de la reducción de la dimensión se resuelve, como en 6.1, mediante un Análisis de Componentes Principales.

El proceso lo podemos resumir de la siguiente forma:

- Cálculo del punto P .
- Evaluación de los símbolos de Christoffel, en P , referidos al sistema de coordenadas, (x^1, \dots, x^n) .

- Resolución del sistema (24), para cada punto, por técnicas numéricas, por ejemplo, el método de Newton-Raphson.

- Con los P_i referidos al nuevo sistema de coordenadas (y^1, \dots, y^n) , pueden seguirse dos alternativas:

a) Cálculo de la matriz de cambio de base P , definida en (30) y efectuar la transformación de coordenadas correspondiente. Tendremos ahora los P_i referidos a un sistema ortonormal (z^1, \dots, z^n) . Posteriormente se aplica el Análisis de Componentes Principales en orden a efectuar una representación en dimensión reducida.

b) Formar la matriz $k \times n$ $Y = (y_i^j)$ y a partir de ésta, proceder como en (2) y (3), obteniendo la matriz A . En este punto al proceso continúa efectuando una diagonalización de A respecto G^{-1} (a diferencia de (4)) y hallando los vectores propios de norma unitaria respecto G^{-1} . Posteriormente se procede como en el apartado 6.1. Este enfoque es equivalente al anterior.

Finalmente, vamos a ver una última forma de tratar el problema relacionando la distancia obtenida a través de la matriz de información de Fisher, con el test de la razón de verosimilitud.

6.2.2.c. Dadas dos poblaciones estadísticas A y B asociadas a una misma clase de funciones de densidad paramétricas, $f(x^1, \dots, x^m, \theta^1, \dots, \theta^n)$, ca

da una de ellas vendrá caracterizada por un conjunto de parámetros $(\theta_A^1, \dots, \theta_A^n)$ y $(\theta_B^1, \dots, \theta_B^n)$. Podemos, a partir de dos muestras aleatorias simples, de tamaños N_A y N_B , de las poblaciones A y B respectivamente, plantear el siguiente contraste:

$$H_0: \quad \theta_A^i = \theta_B^i \quad i=1, \dots, n \quad (31)$$

$$H_1: \quad \exists j \in \{1, \dots, n\} / \theta_A^j \neq \theta_B^j$$

que puede ser resuelto por el método de la razón de verosimilitud.

Vamos a suponer en primer lugar que el espacio paramétrico:

$$E = \{(\theta^1, \dots, \theta^n) / f(x^1, \dots, x^m, \theta^1, \dots, \theta^n) \text{ es función de densidad}\} \quad (32)$$

con el tensor métrico definido a través de la matriz de información de Fisher, es euclideo. En estas condiciones, sin pérdida de generalidad podemos suponer:

$$g_{\mu\nu} = \delta_{\mu\nu} \quad \mu, \nu = 1, \dots, n \quad (33)$$

Hallemos, a continuación, el desarrollo de Taylor del logaritmo de la función de verosimilitud asociada a la muestra A, en el punto definido por la estimación máximo-verosímil de los parámetros, a partir de dicha muestra. Despreciando los términos de orden tres en adelante:

$$\ln L_A \approx \ln \hat{L}_A + \frac{1}{2} \sum_{\mu=1}^n \sum_{\nu=1}^n \frac{\partial^2 \ln L_A}{\partial \theta^\mu \partial \theta^\nu} (\theta^\mu - \hat{\theta}_A^\mu) (\theta^\nu - \hat{\theta}_A^\nu) \quad (34)$$

pero para tamaños muestrales grandes, $N_A \rightarrow \infty$, resulta que:

$$\frac{\partial^2 \ln L_A}{\partial \theta^\mu \partial \theta^\nu} \approx N_A E \left(\frac{\partial^2 \ln f}{\partial \theta^\mu \partial \theta^\nu} \right) \quad (35)$$

por tanto, podemos escribir:

$$\begin{aligned} \ln L_A &\approx \ln \hat{L}_A - \frac{1}{2} N_A \sum_{\mu=1}^n \sum_{\nu=1}^n \delta_{\mu\nu} (\theta_A^\mu - \hat{\theta}_A^\mu) (\theta_A^\nu - \hat{\theta}_A^\nu) = \\ &= \ln \hat{L}_A - \frac{1}{2} N_A \sum_{\mu=1}^n (\theta_A^\mu - \hat{\theta}_A^\mu)^2 \end{aligned} \quad (36)$$

Análogamente, para la población B, resulta:

$$\ln L_B \approx \ln \hat{L}_B - \frac{1}{2} N_B \sum_{\mu=1}^n (\theta_B^\mu - \hat{\theta}_B^\mu)^2 \quad (37)$$

El logaritmo de la función de verosimilitud conjunta será pues aproximadamente:

$$\ln(L_A L_B) = \ln L_A + \ln L_B = \ln \hat{L}_A + \ln \hat{L}_B - \frac{1}{2} \sum_{\mu=1}^n (N_A (\theta_A^\mu - \hat{\theta}_A^\mu)^2 + N_B (\theta_B^\mu - \hat{\theta}_B^\mu)^2) \quad (38)$$

hallems ahora un máximo de ésta función bajo la hipótesis nula, $\theta_A^\nu = \theta_B^\nu = \theta^\nu$, $\nu=1, \dots, n$:

$$\frac{\partial \ln L_A L_B}{\partial \theta^\nu} = - (N_A (\theta^\nu - \hat{\theta}_A^\nu) + N_B (\theta^\nu - \hat{\theta}_B^\nu)) \quad \nu=1, \dots, n \quad (39)$$

igualando a cero y despejando $\hat{\theta}^{*v}$ resulta:

$$\hat{\theta}^{*v} = \frac{1}{N_A + N_B} (N_A \hat{\theta}_A^v + N_B \hat{\theta}_B^v) \quad v=1, \dots, n \quad (40)$$

solución que corresponde efectivamente a un máximo, ya que el hessiano:

$$\frac{\partial^2 \ln L_A L_B}{\partial \theta^\lambda \partial \theta^\nu} = -(N_A + N_B) \delta_{\lambda\nu} \quad \lambda, \nu = 1, \dots, n \quad (41)$$

es diagonal y todos sus elementos son negativos, por tanto es definido negativo.

Luego resulta:

$$\begin{aligned} \frac{1}{2} \sum_{\mu=1}^n (N_A (\hat{\theta}_A^{*\mu} - \hat{\theta}_A^\mu)^2 + N_B (\hat{\theta}_B^{*\mu} - \hat{\theta}_B^\mu)^2) &= \frac{1}{2} \sum_{\mu=1}^n (N_A \left(\frac{-N_B \hat{\theta}_A^\mu + N_A \hat{\theta}_B^\mu}{N_A + N_B} \right)^2 + \\ &+ N_B \left(\frac{N_A \hat{\theta}_A^\mu - N_B \hat{\theta}_B^\mu}{N_A + N_B} \right)^2) = \frac{N_A N_B}{2(N_A + N_B)} \sum_{\mu=1}^n (\hat{\theta}_A^\mu - \hat{\theta}_B^\mu)^2 = \\ &= \frac{N_A N_B}{2(N_A + N_B)} \hat{D}^2 \end{aligned} \quad (42)$$

En estas condiciones podemos escribir, si denominamos $L(\hat{\omega})$ al supremo de la función de verosimilitud bajo la hipótesis nula,

$$\ln L(\hat{\omega}) \approx \ln \hat{L}_A + \ln \hat{L}_B - \frac{1}{2} \frac{N_A N_B}{N_A + N_B} \hat{D}^2 \quad (43)$$

y si designamos por $L(\hat{\Omega})$ al supremo de la función de verosimilitud

fuera de la hipótesis nula, sin restricciones, resulta:

$$\ln L(\hat{\Omega}) \approx \ln \hat{L}_A + \ln \hat{L}_B \quad (44)$$

por tanto si Λ es la razón de verosimilitud:

$$-2 \ln \Lambda = 2 \left[\ln(\hat{\Omega}) - \ln(\hat{\omega}) \right] \approx \frac{N_A N_B}{N_A + N_B} \hat{D}^2 \quad (45)$$

Nótese que $-2 \ln \Lambda$ sigue asintóticamente una distribución χ^2 con n grados de libertad, resultado que concuerda con (116) del capítulo 3.

A partir de (45) podemos escribir:

$$\hat{D} \approx \sqrt{-\frac{2(N_A + N_B)}{N_A N_B} \ln \Lambda} \quad (47)$$

y si $N_A = N_B$:

$$\hat{D} \approx 2 \sqrt{-\frac{1}{N} \ln \Lambda} \quad (48)$$

resultado cuya validez depende de dos hipótesis, que el espacio paramétrico E , (32) sea euclídeo y que los tamaños muestrales N_A y N_B sean grandes.

En caso que el espacio paramétrico no sea euclídeo las anteriores expresiones pueden considerarse como aproximadas siempre que los

puntos A y B no sean muy distantes. En estas condiciones podemos utilizar la métrica del espacio tangente, en A ó B, euclídeo, como aproximación a la métrica riemanniana.

Las expresiones (47) y (48) permiten pues obtener, a partir de los valores muestrales, una medida aproximada de la distancia, definida a través de la matriz de información de Fisher, entre dos o más puntos. Una vez obtenida la matriz de interdistancias entre ellos, el problema es equivalente al planteado en el apartado 6.2.1.

7. UNA APLICACIÓN A LA METODOLOGÍA ETOLÓGICA.

Resumen:

En el presente capítulo se sugiere una metodología basada en las distancias estadísticas y la representación de datos, para analizar las secuencias de comportamiento animal y se aplica al estudio de la conducta agonística del Lúgano.

Sumario:

7.1. PLANTEO GENERAL.

7.2. ESTUDIO DE LA CONDUCTA AGONISTICA DEL LUGANO.

7.3. RESULTADOS.

7.4. DISCUSION.

7.1. PLANTEO GENERAL

Es frecuente en las investigaciones etológicas clasificar el comportamiento animal en determinadas pautas bien definidas y posteriormente observar y registrar las secuencias de dichas pautas en las relaciones o enfrentamientos entre animales. La frecuencia con que se dan dichas secuencias permite sacar conclusiones acerca de las causas más importantes que influyen en el comportamiento animal. Hay varios métodos matemáticos para el análisis de dichos resultados, basados en la estadística, teoría de la información, teoría de la decisión, etc., Colgan & Smith (1978), Hazlett (1977), Sokal & Rohlf (1969). En el presente capítulo se sugiere una nueva metodología matemática, cuya aplicación sistemática puede tener interés para la interpretación de los resultados experimentales.

La información experimental puede disponerse en forma de una tabla de frecuencias multidimensional. Así, por ejemplo, si en un encuentro entre dos animales, uno de ellos puede efectuar una pauta $A_i \in \{A_1, \dots, A_p\}$, el otro responderle con otra, $B_j \in \{B_1, \dots, B_q\}$ y el primero replicar a su vez a éste último con la pauta $C_k \in \{C_1, \dots, C_r\}$, para cada encuentro habremos observado un suceso de la forma $A_i \times B_j \times C_k$ donde los índices varían según: $i = 1, \dots, p$, $j = 1, \dots, q$ y $k = 1, \dots, r$. Después de haber observado N encuentros, si llamamos n_{ijk} a la frecuencia absoluta del suceso $A_i \times B_j \times C_k$, podremos disponer los datos en una tabla tridimensional:

$$\begin{array}{ccccccc}
 & & & n_{11r} & n_{12r} & \dots & n_{1qr} \\
 & & & \cdot & \cdot & & \cdot \\
 & & n_{112} & n_{122} & \dots & n_{1q2} & \\
 & n_{111} & n_{121} & \dots & n_{1q1} & & \\
 & \cdot & \cdot & & \cdot & & \cdot \\
 & & n_{p1r} & n_{p2r} & \dots & n_{pqr} & \\
 & \cdot & \cdot & & \cdot & & \cdot \\
 n_{p11} & n_{p21} & \dots & n_{pq1} & & &
 \end{array} \quad (1)$$

verificándose:

$$\sum_{i=1}^p \sum_{j=1}^q \sum_{k=1}^r n_{ijk} = N \quad (2)$$

Podemos estimar las probabilidades de los sucesos $A_i \times B_j \times C_k$ a través de:

$$P(A_i \times B_j \times C_k) = \frac{n_{ijk}}{N} \quad (3)$$

Tiene también interés estudiar las distribuciones de probabilidad condicionadas a uno o varios sucesos. En el ejemplo antes citado podríamos considerar las siguientes colecciones de distribuciones de probabilidad:

a) Distribuciones de probabilidad condicionadas a un suceso A_i (p en total). Las estimaciones de las mismas vienen dadas por:

$$P(B_j \times C_k / A_i) = \frac{n_{ijk}}{n_{i..}} \quad \begin{array}{l} j=1, \dots, q \\ k=1, \dots, r \end{array} \quad (4)$$

siendo $n_{i..} = \sum_{j=1}^q \sum_{k=1}^r n_{ijk}$ (supuesto que $n_{i..} \neq 0$).

b) Distribuciones de probabilidad condicionadas a un suceso B_j (q en total). Las estimaciones de las mismas vienen dadas por:

$$P(A_i \times C_k / B_j) = \frac{n_{ijk}}{n_{.j.}} \quad \begin{array}{l} i=1, \dots, p \\ k=1, \dots, r \end{array} \quad (5)$$

siendo $n_{.j.} = \sum_{i=1}^p \sum_{k=1}^r n_{ijk}$ (supuesto que $n_{.j.} \neq 0$).

c) Distribuciones de probabilidad condicionadas a un sucesos C_k (r en total). Las estimaciones de las mismas vienen dadas por:

$$P(A_i \times B_j / C_k) = \frac{n_{ijk}}{n_{..k}} \quad \begin{array}{l} i=1, \dots, p \\ j=1, \dots, q \end{array} \quad (6)$$

siendo $n_{..k} = \sum_{i=1}^p \sum_{j=1}^q n_{ijk}$ (supuesto que $n_{..k} \neq 0$).

d) Distribuciones de probabilidad condicionadas a un producto de sucesos $A_i \times B_j$ ($p \cdot q$ en total). Las estimaciones de las mismas vienen dadas por:

$$P(C_k / A_i \times B_j) = \frac{n_{ijk}}{n_{ij.}} \quad k=1, \dots, r \quad (7)$$

siendo $n_{ij.} = \sum_{k=1}^r n_{ijk}$ (supuesto que $n_{ij.} \neq 0$).

e) Distribuciones de probabilidad condicionadas a un producto de sucesos $A_i \times C_k$ (p.r en total). Las estimaciones de las mismas vienen dadas por:

$$P(B_j/A_i \times C_k) = \frac{n_{ijk}}{n_{i \cdot k}} \quad j=1, \dots, q \quad (8)$$

siendo $n_{i \cdot k} = \sum_{j=1}^q n_{ijk}$ (supuesto que $n_{i \cdot k} \neq 0$).

f) Distribuciones de probabilidad condicionadas, a un producto de sucesos $B_j \times C_k$ (q.r en total). Las estimaciones de las mismas vienen dadas por:

$$P(A_i/B_j \times C_k) = \frac{n_{ijk}}{n_{\cdot jk}} \quad i=1, \dots, p \quad (9)$$

siendo $n_{\cdot jk} = \sum_{i=1}^p n_{ijk}$ (supuesto $n_{\cdot jk} \neq 0$).

En general, si hubiésemos observado secuencias de m pautas, el número de colecciones de distribuciones de probabilidad condicionada, C , vendría dado por:

$$C = \binom{m}{1} + \dots + \binom{m}{m-1} = 2^m - 2 = 2(2^{m-1} - 1) \quad (10)$$

Se sugiere, como metodología standard, el comparar las distribuciones de probabilidad, de cada una de las colecciones asociadas a la tabla de frecuencias m -dimensional, por separado, a través

del uso de una función distancia.

Dentro de cada colección de distribuciones de probabilidad, podemos usar la distancia definida en el capítulo IV para la distribución multinomial, con $N=1$. Si p_1, \dots, p_h y q_1, \dots, q_h son dos distribuciones de una colección, la distancia entre ambas vendrá dada por:

$$d = \sqrt{2} \arccos \left(\sum_{i=1}^h \sqrt{p_i q_i} \right) \quad (11)$$

Una vez distanciadas, puede efectuarse una representación gráfica de las distribuciones de probabilidad (como puntos en un plano), de cada colección por separado, usando un análisis de coordenadas principales o bien técnicas de Multidimensional Scaling (MDS), Cuadras (1981).

Una vez definida esta distancia, es posible también obtener una clasificación jerárquica de las mismas, utilizando los métodos habituales de la taxonomía numérica. Como sea que la distancia definida a través de (11) no cumplirá, generalmente, la propiedad ultramétrica, es preciso aplicar algún algoritmo que la deforme hasta obtener una distancia ultramétrica, lo más parecida posible a la original, hecho que se traducirá por tener un coeficiente de correlación cofenética elevado. Dentro de una clase muy general de algoritmos que efectúan dicho proceso, es el conocido método UPGMA el que proporciona mayor correlación cofenética, Arcas (1982) y por ésto es recomendable su empleo.

Tanto las representaciones gráficas de las distribuciones de probabilidad como puntos en un plano, como los dendogramas obtenidos al clasificarlas jerárquicamente, pueden ayudar a resumir la información experimental de la tabla de frecuencias multidimensional, y facilitar la interpretación de los resultados.

7.2. ESTUDIO DE LA CONDUCTA AGONISTICA DEL LUGANO.

El Lúgano (*Carduelis spinus L.*) es un pájaro fringílido, del tamaño de un jilguero, de plumas verdosas, manchado de negro y ceniza, amarillo en el cuello, pecho y extremidades de las plumas remeras y timoneras, siendo la hembra de un color más ceniciento. Su hábitat natural son los bosques de coníferas, prefiriendo el bosque bajo al bosque alto, pero se adapta fácilmente a la cautividad.

El concepto de "conducta agonística" se utiliza para designar al conjunto de comportamientos que aparecen en la lucha y la conducta territorial. Los cuatro comportamientos más característicos de la conducta agonística son: el despliegue agresivo, el ataque, la huida o la sumisión, Manning (1972). Todos estos comportamientos tienen en común que se desencadenan al ser sobrepasada una cierta distancia individual. La magnitud de dicha distancia depende del contexto en que se halla el animal, pequeña, por ejemplo, si se aproxima un subordinado, grande si se acerca un macho dominante a otro macho en periodo reproductor. Este territorio individual proporciona al animal el espacio mínimo para utilizar eficazmente

los recursos del medio, Marler (1957), Senar (1982).

La conducta agonística será utilizada tanto para mandar como para recoger información, la cual se utilizará en última instancia para aumentar la eficiencia de la utilización de los recursos, reduciendo, por ejemplo, el número de luchas inútiles o disminuyendo la peligrosidad de éstas, Senar (1982).

Los datos utilizados en el presente trabajo han sido obtenidos a partir de una comunidad cautiva de lúganos, cuatro machos y tres hembras, instalados en una pajarera de $1.25 \times 0.60 \times 1.25$ metros. Las ventajas e inconvenientes de la investigación en cautividad han sido ampliamente discutidas Hindel (1954), Tordoff (1954), Dilger (1960), encontrándose que la única variación observada, respecto al comportamiento en el campo, es un aumento en la frecuencia de encuentros agresivos, debido probablemente a la reducción del espacio vital. Por otra parte, los estudios en cautividad ofrecen posibilidades de observación que difícilmente pueden encontrarse en el campo. El trabajo experimental fue llevado a cabo por Senar y colaboradores (1982), consistiendo en la observación de casi 3000 enfrentamientos entre lúganos. En cada enfrentamiento se anotaba la pauta que realizaba el individuo que iniciaba el despliegue agresivo, que llamaremos activo (ACT1), lo que contestaba el individuo receptor, al que llamaremos pasivo (PAS) y la réplica del primero (ACT2).

El individuo activo puede iniciar el enfrentamiento (ACT1)

con una serie de pautas, descritas en Senar (1982) y que resumimos a continuación.

A0. El pájaro alarga la cabeza y dirige el pico cerrado hacia el oponente. Cuerpo en posición normal, aunque con tendencia a la horizontalidad. Alas plegadas, plumaje normal. Patas en la misma posición que se encontraban antes de la amenaza.

A1. El pájaro dirige la cabeza, con el pico semiabierto, hacia el oponente. La cabeza y el cuello se estiran, acercándose al otro pájaro. Plumaje normal. Alas plegadas, aunque no cruzadas sobre el obispillo. Cola siguiendo la línea del cuerpo. Patas en posición normal.

A2. El pájaro dirige la cabeza, con el pico semiabierto, hacia el oponente. Cabeza y cuello estirados. Plumaje normal, aunque mas bien liso. Cola en la horizontal del cuerpo; cuerpo en posición horizontal, paralelo al suelo. Alas semiabiertas, no desplegadas; codos algo hacia delante. Se despliega el antebrazo, apareciendo el ala algo caída, aunque puede levantarla hasta que su borde exterior quede prácticamente horizontal.

A3. Igual que A2, pero en éste caso las alas vibran en un movimiento de arriba a abajo, aunque no de gran amplitud.

A4. El pájaro dirige la cabeza, con el pico semiabierto, hacia el

oponente. Cabeza y cuello muy estirados, plumaje liso. Cuerpo horizontal, al igual que la cola. Alas totalmente desplegadas y elevadas. En ocasiones, avanza hacia el oponente moviendo algo las alas.

A7. Es el "ataque suplantador". En ocasiones el pájaro salta o se posa después de un vuelo dirigido, en un punto que milésimas de segundo antes estaba ocupado, y que es abandonado por su anterior ocupante. A veces el antiguo ocupante huye ya al ver al "suplantador" preparándose para el vuelo.

A9. El pájaro al tener cerca a su oponente, estira la cabeza y la pinza alguna parte del cuerpo.

Estas son las siete pautas consideradas en el activo, al iniciar el enfrentamiento. Las respuestas tipificadas del pasivo son: el ataque (A), el despliegue agresivo (D), la huida (H), la sumisión (S), o bien un comportamiento no relacionado con la conducta agonística (comportamiento indiferente), el pájaro se limita a "estar" (E).

Finalmente la réplica del activo (ACT2) es también evaluada como ataque (A), despliegue agresivo (D), huida (H), sumisión (S) y estar (E).

7.3. RESULTADOS

Los datos experimentales se adaptan bien al ejemplo expuesto en el apartado 7.1, con $p=7$, $q=5$ y $r=5$. La tabla de frecuencias tridimensional puede escribirse como:

TABLA I (Senar (1982), inédita)

Pauta ACT1: A0

		ACT2				
		A	D	E	H	S
PAS	A	0	0	0	0	0
	D	0	0	3	0	0
	E	1	2	13	0	0
	H	0	1	90	0	0
	S	0	1	29	0	0

Pauta ACT1: A1

		ACT2				
		A	D	E	H	S
PAS	A	36	0	1	11	0
	D	11	4	7	26	6
	E	30	16	65	2	0
	H	10	0	779	0	0
	S	0	3	158	0	0

Pauta ACT1: A2

		ACT2				
		A	D	E	H	S
PAS	A	10	1	0	6	0
	D	1	1	3	9	0
	E	6	7	14	0	0
	H	5	0	229	0	0
	S	0	0	18	0	0

TABLA I (Continuación)

Pauta ACT1: A3

		ACT2				
		A	D	E	H	S
	A	1	0	0	0	0
	D	2	0	0	0	0
PAS	E	0	0	2	0	0
	H	1	0	25	0	0
	S	0	0	1	0	0

Pauta ACT1: A4

		ACT2				
		A	D	E	H	S
	A	9	0	0	1	0
	D	4	1	0	4	0
PAS	E	5	0	0	0	0
	H	5	0	96	0	0
	S	0	0	6	0	0

Pauta ACT1: A7

		ACT2				
		A	D	E	H	S
	A	39	0	0	2	1
	D	1	0	1	15	0
PAS	E	1	0	1	1	0
	H	9	4	882	0	0
	S	0	0	0	0	0

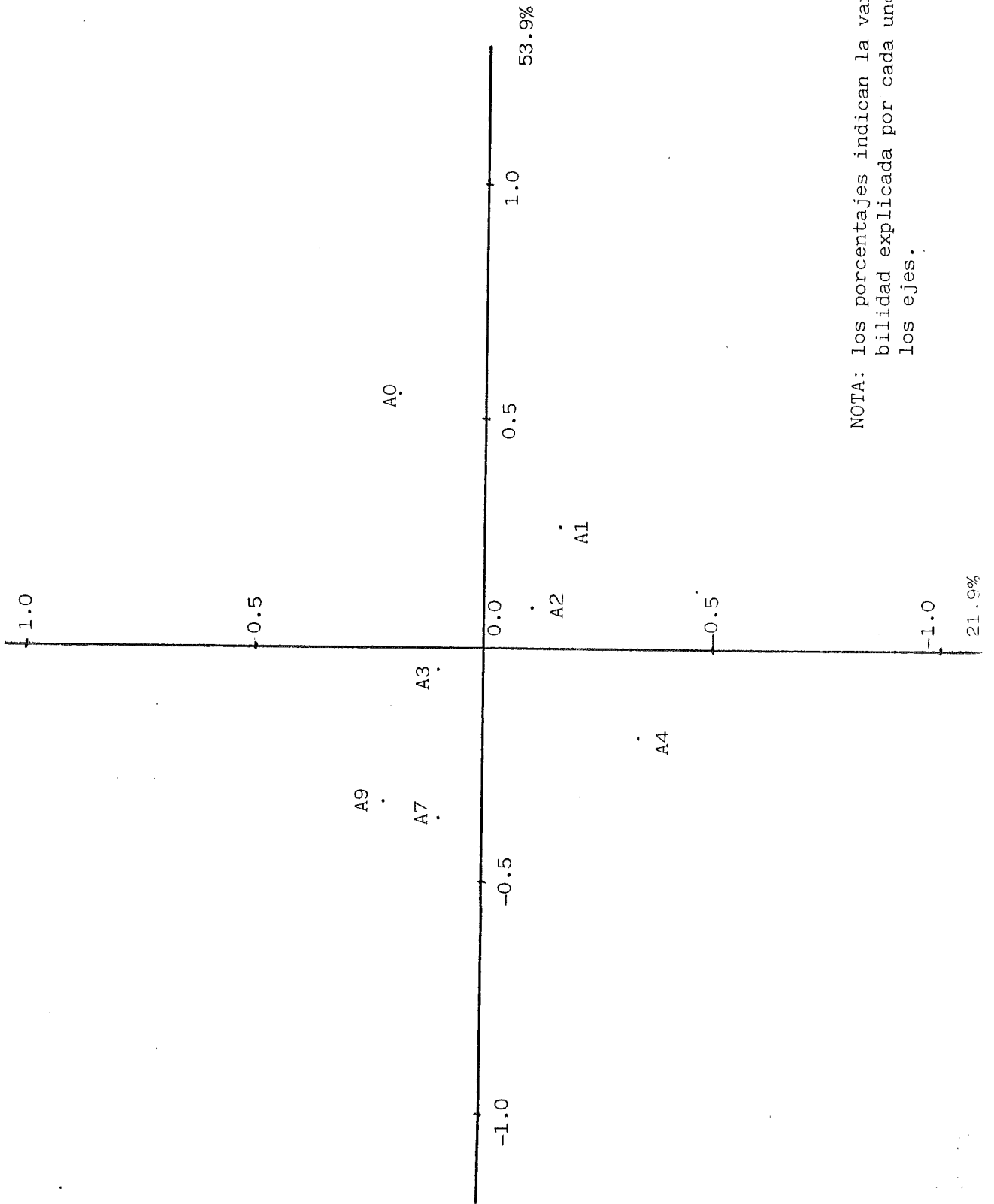
TABLA I (Continuación)

Pauta ACT1: A9

		ACT2				
		A	D	E	H	S
	A	3	0	0	1	0
	D	1	0	1	2	0
PAS	E	0	0	2	0	0
	H	1	1	81	3	0
	S	0	0	0	0	0

A continuación siguen las representaciones gráficas y los dendogramas de cada una de las seis colecciones de funciones de densidad condicionadas (a,b,c,...,f) expuestas en el apartado 7.1. Se han suprimido, en el análisis, aquellos casos en que la estimación de la función de densidad condicionada se obtenía a partir de una muestra demasiado pequeña (cinco a siete casos).

Representación gráfica de las funciones de densidad condicionadas a cada una de las pautas que inicialmente efectúa el activo (ACT1)



NOTA: Los porcentajes indican la variabilidad explicada por cada uno de los ejes.

Dendograma de las funciones de densidad condicionadas a cada una de las pautas que inicialmente efectúa el activo (ACT1).

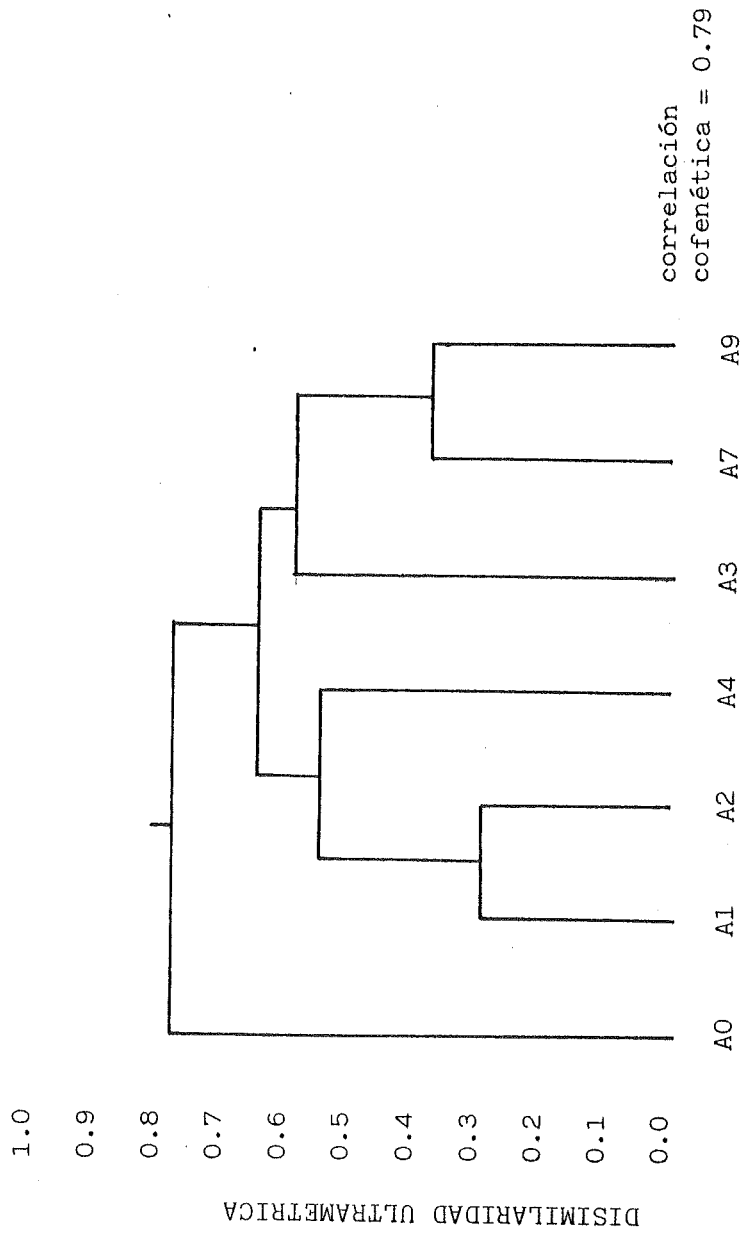
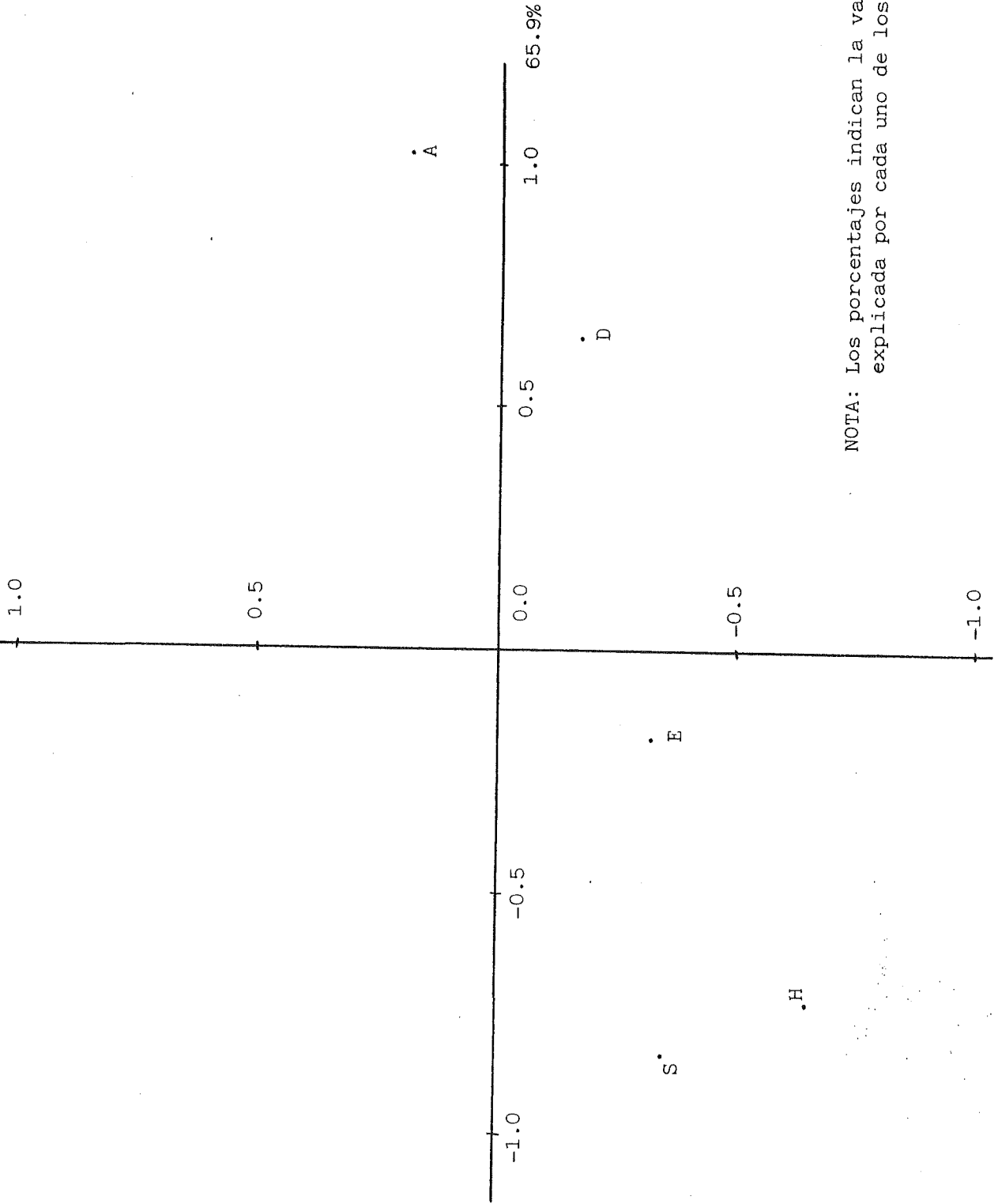


Figura 1.b

Representación gráfica de las funciones de densidad condicionadas a cada una de las pautas que efectúa el pasivo (PAS).



NOTA: Los porcentajes indican la variabilidad explicada por cada uno de los ejes.

Figura 2.a 15.7%