

Análisis numérico multivariante en parámetros métricos mandibulares equinos (raza *Cavall Pirinenc Català*) - Multivariant numerical analysis with equine mandibular metric parameters (*Cavall Pirinenc Català* breed)

Pere-Miquel Parés i Casanova

Dept. de Anatomía Animal. Ciència i Salut Animal. Universitat de Lleida. España

peremiquelp@prodan.udl.cat

Resumen

Se analizan 9 variables métricas de 16 mandíbulas procedente de animales adultos de la raza equina *Cavall Pirinenc Català*. A partir de estas variables se realiza una estadística descriptiva y de normalidad, obteniendo posteriormente los coeficientes de correlación para todas las variables. Se realizan las pruebas de adecuación muestral de Kaiser-Meyer-Olkin y de esfericidad de Bartlett y, finalmente, el análisis de componentes principales (PCA). La selección del PCA como método de extracción demuestra ser el adecuado. El uso de 8 variables métricas conlleva una mínima pérdida de información, pudiendo excluir la variable "longitud mandibular máxima". Las variables relacionadas con la rama y sección angular aparecen sensiblemente diferenciadas del resto. La longitud del diastema y la sección angular se muestran como las más discriminantes.

Palabras clave: raza, mandíbula, osteometría, *Cavall Pirinenc Català*

Abstract

Nine metric parameters were taken on 16 mandibles from adult horses belonging to the *Cavall Pirinenc Català* breed. Based on these parameters, we develop a descriptive and normality statistics and we obtain correlation coefficients for all the parameters. Thereafter we apply the Kaiser-Meyer-Olkin and Bartlett's tests and we carry out a Principal Component Analysis (PCA). The PCA selection as the extraction method turns out to be the right one. The use of 8 metric parameters involves a minimal loss of information, and it is possible to exclude the parameter "maximum mandibular length". The parameters related with the branch and angular section appear sensibly differentiated from

the rest. The diastem length and the angular section are the more discriminant parameters

Keywords: breed, mandible, osteometry, *Cavall Pirinenc Català*

INTRODUCCIÓN

El análisis de componentes principales (PCA) construye una transformación lineal que escoge un nuevo sistema de coordenadas para el conjunto original de datos en el cual la varianza de mayor tamaño del conjunto de datos es capturada en el primer eje (llamado el Primer Componente Principal), la segunda varianza más grande es el segundo eje, y así sucesivamente. Para construir esta transformación lineal debe construirse primero la matriz de covarianza o matriz de coeficientes de correlación. Una de las ventajas del PCA para reducir la dimensionalidad de un grupo de datos, es que retiene aquellas características del conjunto de datos que contribuyen más a su varianza, manteniendo un orden de bajo nivel de los componentes principales e ignorando los de alto nivel. El objetivo es que esos componentes de bajo orden a veces contienen el "más importante" aspecto de esa información.

Ahora bien el objetivo de síntesis no se conseguiría si se seleccionasen todos los componentes principales. Lo que se persigue es reducir el número de variables con la menor pérdida de información, a través de sintetizarlos en los componentes principales o factores.

Los ensayos factoriales, por otro lado, son de utilidad por contribuir a depurar la matriz de información. Su importancia consiste en aportar las variables que conforman la matriz para la extracción de factores, así como evaluar la validez del método. El ensayo utilizado parte del cálculo de la matriz de correlación, puesto que se considera de importancia que todos los indicadores tengan al menos un coeficiente de correlación significativo en la matriz. Por tal motivo se considera que si las correlaciones entre los indicadores son pequeñas, resulta bastante improbable que originen factores comunes, y por el contrario, si el valor del coeficiente de correlación entre dos indicadores es muy alto, significa que uno de ellos aporta información redundante.

El objetivo de este punto es pues obtener las variables que contribuyen más a la descripción métrica en mandíbulas equinas.

MATERIAL Y MÉTODOS

Se han analizado 9 variables métricas de 16 mandíbulas (maxilares inferiores), el material óseo, procedente de animales adultos (datados a partir de la serie dental completa) de la raza equina *Cavall Pirinenc Català*¹. En las **figuras 1 y 2** se exponen los puntos topográficos de referencia utilizados; todos ellos siguiendo el protocolo de Eisenmann (1980)². La mayoría de las medidas se han tomado con un pie de rey ($\pm 0,1$ mm). Para las longitudes de las partes más largas se ha utilizado un calibre de gruesos ($\pm 0,05$ cm). Todas las medidas se ofrecen en mm. No se consideró necesaria la autorización ética puesto que se trabajó con material procedente de cadáveres de animales muertos por causas naturales.

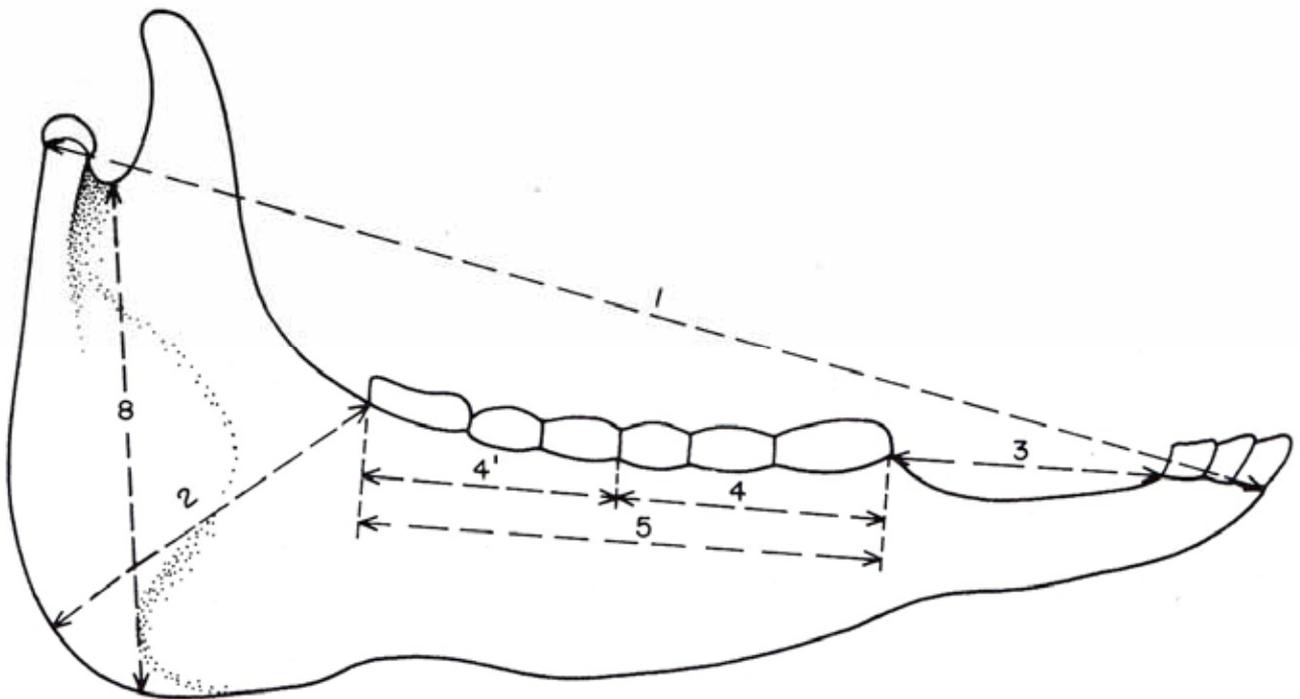
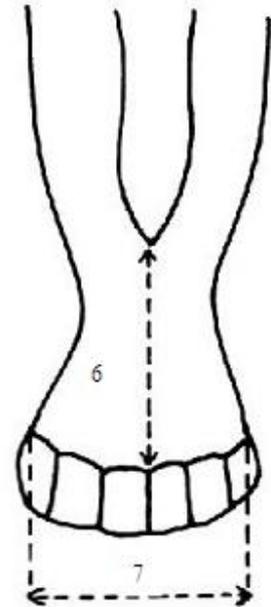


Figura 1. Puntos topográficos utilizados (visión lateral)

- 1 longitud mandibular máxima
- 2 sección del ángulo mandibular
- 3 diastema
- 4 longitud de la serie premolar
- 4' longitud de la serie molar
- 5 longitud de la serie yugal
- 8 altura de la rama mandibular

Figura 2. Puntos topográficos utilizados (visión dorsal del extremo rostral)

6 longitud de la sínfisis mandibular
7 anchura del cuerpo del incisivo



A partir de estas variables se ha realizado una estadística descriptiva (N, mínimo y máximo, error estándar, media, error estándar, varianza, desviación estándar, mediana, sesgo y curtosis) y se aplica la prueba de normalidad de Shapiro-Wilk, obteniendo posteriormente los coeficientes de correlación para todas las variables. Un dendrograma de correlaciones, que se realiza aplicando el método de Ward -que minimiza la suma de los cuadrados de los residuos de cada dos hipotéticos clusters que pueden ser formados en cada paso- nos permite a la vez evaluar la distancia entre grupos. A fin de comprobar la pertinencia del análisis dimensional de los datos se realizan las pruebas de adecuación muestral de Kaiser-Meyer-Olkin (KMO) y de esfericidad de Bartlett y, finalmente, el análisis de componentes principales (PCA).

El cálculo de los estadísticos y de los tests de adecuación muestral se obtienen mediante el paquete FACTOR (Lorenzo-Seva & Ferrando, 2007); con el paquete PAST (Hammer *et al.*, 2001) se efectúan la prueba de normalidad, la matriz de correlación entre variables, el PCA, y el establecimiento del dendrograma de correlaciones.

RESULTADOS Y DISCUSIÓN

a) Estadística descriptiva y prueba de normalidad de Shapiro-Wilk

Los resultados de la estadística descriptiva se muestran en la **tabla 1**. Todas las variables presentan una distribución normal.

Tabla 1. Estadística descriptiva de las variables métricas obtenidas¹

	1	2	3	4	4'	5	6	7	8
N	16	16	16	16	15	15	16	16	16
Mínimo	446	134,5	93,75	89,4	89,25	182,0	84,0	68,8	233,0
Máximo	509,0	161,0	120,8	108,0	108,0	212,0	98,8	81,3	274,0
Media	485,8	145,7	106,4	97,8	94,9	192,0	91,4	75,3	251,3
E.S.	4,650	2,071	2,000	1,239	1,243	2,141	1,045	0,873	3,162
Varianza	346,03	68,66	64,02	24,56	23,18	68,78	17,47	12,20	160,05
d.e.	18,60	8,28	8,00	4,95	4,81	8,29	4,17	3,49	12,65
Mediana	488,5	143,5	105,1	97,1	94,2	191,0	91,7	75,0	252,2
Sesgo	-0,748	0,457	0,100	0,409	1,140	0,799	-0,022	0,045	0,341
Curtosis	-0,582	-1,182	-1,074	-0,687	1,012	-0,129	-1,103	-0,943	-0,972
W	0,914	0,935	0,956	0,971	0,889	0,926	0,980	0,975	0,947

E.S.: Error Estándar

d.e.: desviación estándar

W: coeficientes de Shapiro-Wilk

¹ De una mandíbula no se dispone el valor correspondiente a la serie premolar, por lo que para las variables 4' y 5 el universo muestral se reduce.

b) Cálculo de la matriz de correlación

Los coeficientes de correlación analizados para las 9 variables se muestran en la **tabla 2**. Los resultados muestran que el grado de correlación entre las variables morfométricas analizadas es muy disimilar ($0,217 \pm 0,371$). La mayoría de los coeficientes de correlación resultan ser positivos, siendo significativos los relacionados con el ángulo mandibular (medida 2) y el diastema (medida 3), e igualmente la longitud de la serie yugal (medida 5) con la anchura del cuerpo del incisivo (medida 7). Los valores negativos son todos débiles, destacando el de la serie molar con el diastema (lo que es lógico: mandíbulas con mayor espacio diastemal ofrecen menos espacio para la serie yugal). En la **figura 3** aparece el dendrograma de correlaciones. La variable 1 es la que presenta un mayor grado de independencia. Las variables 2, 5 y 8 forman a su vez un cluster separado.

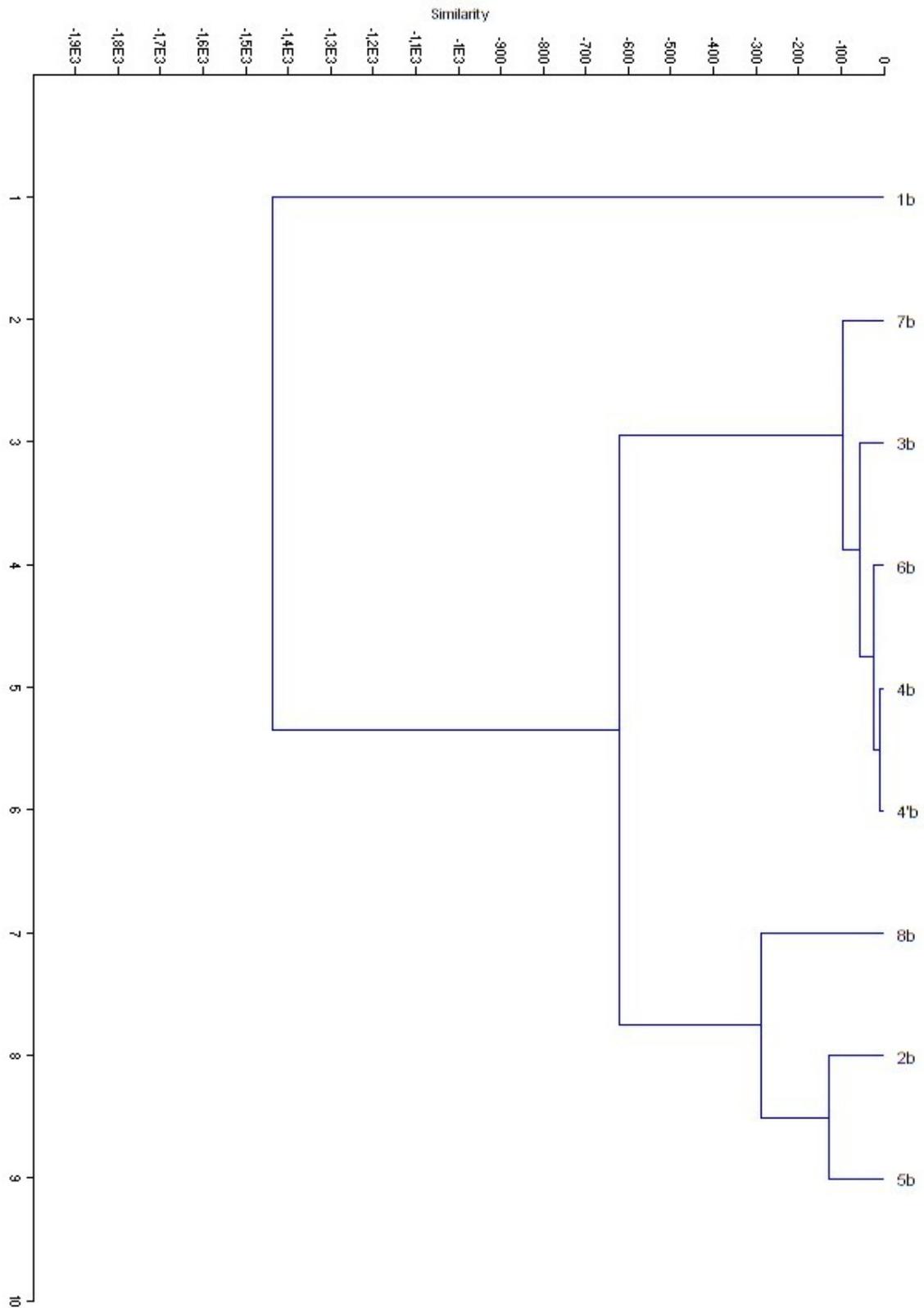


Figura 3. Dendrograma de correlaciones para 2 variables ambientales

Tabla 2. Coeficientes de correlación

	1	2	3	4	4'	5	6	7	8
1	1,000								
2	0,237	1,000							
3	0,374	0,682**	1,000						
4	0,375	-0,394	-0,456	1,000					
4'	0,359	-0,045	-0,369	0,680**	1,000				
5	0,431	-0,336	-0,466	0,925**	0,867*	1,000			
6	0,107	0,408	0,461	-0,089	0,140	-0,043	1,000		
7	0,430	-0,226	-0,163	0,580*	0,450	0,618*	-0,040	1,000	
8	0,439	0,517*	0,411	0,010	0,369	0,152	0,104	0,341	1,000

* P < 0,05

** P < 0,01

c) Pruebas de adecuación muestral

Para nuestro estudio se ha tomado como criterio para la eliminación de las variables redundantes la variable que no alcance ningún valor de correlación igual o mayor a 0,441 (P=0,10), indicando que tienen redundancia en la información. En este caso se suprime la variable 1, longitud mandibular máxima.

Se aplica seguidamente el test KMO. Si al calcular este índice su valor se aproxima a la unidad, indica que la suma de los coeficientes de correlación parcial entre todos los pares de indicadores de la matriz es pequeña, lo cual constituye una señal de que la elección del análisis factorial como técnica para el estudio y la síntesis de los datos ha sido adecuada. En el estudio, este test toma un valor muy bueno, de 0,999.

Posteriormente, se aplica el test de esfericidad de Bartlett, que comprueba si la matriz de correlaciones es una matriz identidad. Puesto que $X^2_{0,01[15]} = 30,58$, concluimos que las varianzas de la muestra son homogéneas.

En la **tabla 3** aparecen los ensayos del test KMO y de Bartlett. No ha sido necesaria rotación de factores para conseguir la validez.

Tabla 3. Ensayos de valoración del análisis factorial

Test KMO	0,999
Test de esfericidad de Bartlett	-258,7
Nivel de significación	0,5

d) Extracción de los componentes ambientales principales

Aplicamos el PCA para examinar la relación existente entre las 8 variables cuantitativas finalmente elegidas. En nuestro estudio, y para dos variables ambientales, encontramos que el primer factor principal explica el 63,86 % de la variación total de la muestra y el segundo, el 15,24 % (**tabla 4**). Las 8 variables sintetizadas en los dos factores principales explican en conjunto el 79,00 % de la varianza total, umbral que debe considerarse idóneo. El primer PC, que explica la mayor parte de la variación, permite constatar como son las variables 2 (sección del ángulo mandibular), 3 (longitud del diastema) y 8 (altura de la rama mandibular) las más discriminantes, mostrando en este caso que es la conformación general de la mandíbula la que aparece como más discriminante. En la **figura 4** aparecen las representaciones PCA para 2 variables ambientales.

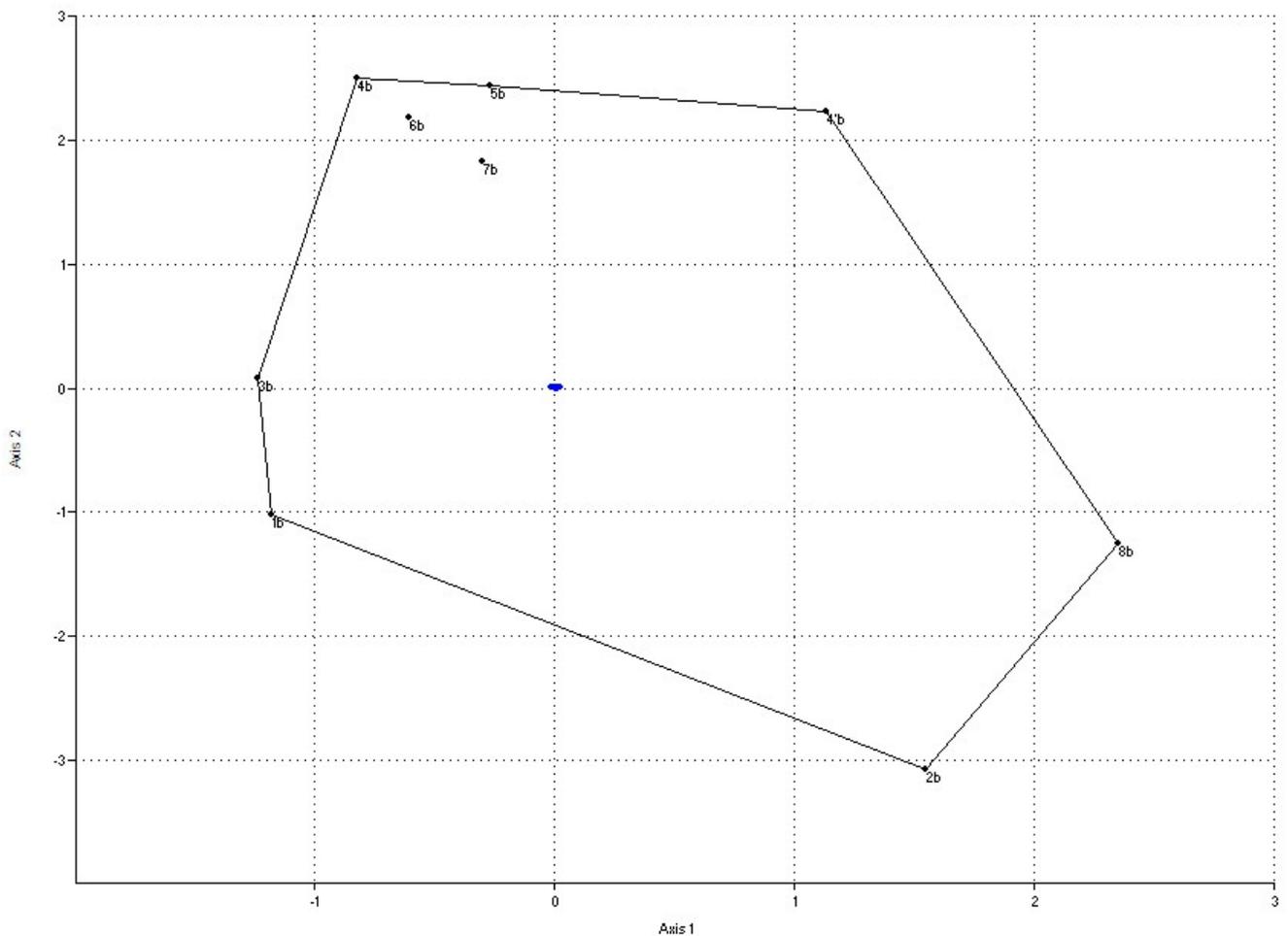


Figura 4. Representació PCA

Tabla 4. Análisis de componentes principales utilizando 2 variables ambientales

Variable	Can1	Can2
2	-7,082	1,507
3	-6,277	2,908
4	1,578	-3,256
4'	0,096	-4,130
5	2,316	-7,250
6	0,050	-2,405
7	0,050	-2,405
8	-9,111	-5,694

Can 1: 63,86 %
Can 2: 15,24 %

e) Determinación de la comunalidad para cada variable

Por último, otro instrumento que contribuye a a nuestro estudio ha sido la determinación de la comunalidad para cada variable. Ésta indica hasta qué punto las variables que se determinan ayudan a explicar el indicador en cuestión, la conformación mandibular en nuestro caso. Aquellos que poseen baja comunalidad en el proceso primario de su selección se eliminan, pues serán poco explicados por los factores.

Los valores obtenidos para las comunalidades para cada uno de las variables estudiadas que conforman la matriz de información inicial oscilan del 115,443 al 2,075, resultando especialmente altos para las variables 2, 3, 5 y 8 b (**tabla 5**).

Tabla 5. Valores obtenidos para las comunalidades

Variable	Comunalidad
2	52,427
3	47,861
4	13,094
4'	17,062
5	57,930
6	2,075
7	5,786
8	115,443

CONCLUSIÓN

La selección del PCA como método de extracción ha sido el adecuado para sintetizar los indicadores considerados en pocas variables ambientales, perdiendo la cantidad mínima de información. El uso de 8 variables métricas conlleva una mínima pérdida de información en el estudio mandibular; la longitud mandibular máxima la podemos excluir. Las variables relacionadas con la rama y sección angular aparecen sensiblemente diferenciadas del resto y la longitud del diastema y la sección del ángulo las más discriminantes. La serie yugal es poco discriminante.

De todos modos, debería recordarse que un análisis multivariado requiere un alto número de muestras para ser confiable. Como guía, y de acuerdo a la relación de Henry, para estudio de este tipo utilizando 8 variables lineales debería contarse con un universo muestral de al menos 68 ejemplares.

NOTAS

¹ Raza nominalmente reciente, cuyo estándar racial se aprobó en Diciembre del 2007. Corresponde al hasta ahora equivocadamente conocido como "Hispano-Bretón" del Pirineo Catalán; y más concretamente, el material estudiado procede exclusivamente de animales del topotipo ceretano, caracterizado por su mayor tamaño y armonicidad.

² No seguiremos siempre la nomenclatura de la N.A.V. (1968), a fin de evitar utilizar términos ciertamente de rigor anatómico, pero inusuales en la expresión zoológico-anatómica a que todos estamos habituados.

BIBLIOGRAFÍA CITADA

- EISENMANN, V., 1980. Les Chevaux (*Equus sensu lato*) fossiles et actuels: Crânes et Dents Jugales Supérieures. Centre Nat. de la Recherche Scientifique. París
- HAMMER, Ø.; HARPER, D.A.T.; RYAN, P. D., 2001. PAST. Paleontological Statistics Software Package for Education and Data Analysis
- LORENZO-SEVA, U.; FERRANDO, P.J., 2007. Unrestricted Factor Analysis. Release v. 7.00
- NOMINA ANATOMICA VETERINARIA, 1968. International Committee on Veterinary Anatomical Nomenclature. Adolf Holzhausen's Successors. Viena

REDVET: 2010 Vol. 11, Nº 1

Recibido 04.08.09 - Ref.prov. AGO0908B - Revisado 06.12.09 - Aceptado 17.12.09
Ref. def. 011001_REDVET - Publicado: 01.01.10

Este artículo está disponible en <http://www.veterinaria.org/revistas/redvet/n010110.html> concretamente en <http://www.veterinaria.org/revistas/redvet/n010110/011001.pdf>

REDVET® Revista Electrónica de Veterinaria está editada por Veterinaria Organización®. Se autoriza la difusión y reenvío siempre que enlace con Veterinaria.org® <http://www.veterinaria.org> y con REDVET® - <http://www.veterinaria.org/revistas/redvet> - <http://revista.veterinaria.org>