

T.151

53023

T. 157



ESPECIALIDAD EN ESTADISTICA
ORIENTACION DISEÑO EXPERIMENTAL

MONOGRAFIA

ANALISIS DE CORRELACION CANONICA
Elementos teóricos básicos - Un ejemplo de aplicación a datos reales

Director: Dr. Oscar H. BUSTOS

Silvia G. VALDANO

1998

82027

MEN:
Credit:

INDICE	Página
RESUMEN	1
OBJETIVO de la MONOGRAFIA	1
1. INTRODUCCIÓN	1
2. EL MODELO	
2.1. La Matemática del Análisis de Correlación Canónica	2
2.1.1. Notas Ampliatorias	4
2.1.2. Variables y Correlaciones Canónicas Muestrales	6
2.2. Inferencia para muestras de gran tamaño	
2.2.1. Test de cociente de verosimilitud	7
2.2.2. Test de Bartlett	8
2.2.3. Test de Lawley	10
2.2.4. Otros tests	10
2.2.5. Problemas para la aplicación de los tests	10
2.3. La Interpretación	
2.3.1. Identificación de las variables canónicas poblacionales	10
2.3.2. La Correlación Canónica como generalización de otros coeficientes de correlación	12
2.3.3. Información de las primeras p Variables Canónicas sobre la variabilidad	14
2.3.4. Interpretación Geométrica del Análisis Canónico Poblacional	14
2.4. Medidas Descriptivas Adicionales	
2.4.1. Matrices de error de aproximación	15
2.4.2. Proporción de la varianza muestral explicada	17
2.5. Más de dos conjuntos de variables	18
3. APLICACIÓN A DATOS REALES	
3.1. Evaluación del componente ambiental de la variación fenotípica en poblaciones prehistóricas	19
3.1.1. Profundización del Análisis	26
4. BIBLIOGRAFÍA CITADA	29
5. AGRADECIMIENTOS	30
APÉNDICE	
1. CONCEPTOS DE ALGEBRA MATRICIAL	
1.1. Matrices Definidas Positivas y Formas Cuadráticas Asociadas	31
1.2. Descomposición Espectral	31
1.3. Matriz Raíz Cuadrada	31
1.4. Desigualdad de Cauchy-Schwarz (C-S)	31
1.4.1. Extensión de la Desigualdad de C-S a matrices definidas positivas	32
1.5. Principios de Maximización	32
1.5.1. Nota Ampliatoria	34
2. UN POCO DE HISTORIA...	34
3. ELEMENTOS DE GENETICA CUANTITATIVA	
3.1. Conceptos Básicos de Genética	35
3.1.1. Glosario	35
3.2. Caracteres Métricos y su Variación	36
3.3. El Material Osteológico y la Metodología Propuesta	41
3.4. Resultados del Análisis Estadístico Univariado para la estimación de la repetibilidad de caracteres métricos	43
3.5. Variables Utilizadas en el Análisis	44

ANÁLISIS DE CORRELACIÓN CANÓNICA

Elementos teóricos básicos - Un ejemplo de aplicación a datos reales.

RESUMEN

Con este trabajo monográfico se ofrece una presentación de una técnica de análisis multivariado: el Análisis de Correlación Canónica (ACC).

Luego de una somera introducción histórica, se expone la descripción algebraica de la naturaleza de la correlación canónica y se estudia el empleo y la significación de este tipo de análisis en ciencias biológicas.

Se analizan medidas descriptivas adicionales como las matrices de error de aproximación y la proporción de la varianza muestral explicada.

Para propósitos de inferencia, se consideran varios tests de significación (Cociente de verosimilitud, Bartlett, Lawley, entre otros). Se presentan las bases para la extensión del ACC a tres o más conjuntos de variables.

Como aplicación se analiza un caso real (población prehistórica de San Pedro de Atacama, Norte de Chile) para lograr una estimación de la componente genética asociada a variables métricas, de gran importancia en estudios de biología evolutiva en poblaciones antiguas. Se concluye que las variables canónicas obtenidas son buenas predictoras de las variables originales y las que presentan mayor correlación entre ambos lados del cráneo acumulan mayor variación genética y ambiental general.

OBJETIVOS de la MONOGRAFIA

1. Elaborar, a partir de la revisión bibliográfica, una presentación didáctica en español de los principales elementos teóricos del Análisis de Correlación Canónica.
2. A través del análisis de un caso concreto de aplicación, mostrar la importancia del empleo de algunas herramientas propias de esta técnica.

1. INTRODUCCIÓN

El Análisis de Correlación Canónica, técnica desarrollada por Hotelling (1935, 1936) vio la luz en un trabajo publicado en el Journal of Educational Psychology sobre la relación entre la velocidad y la capacidad de lectura, y la velocidad y la capacidad de cálculo aritmético en 140 estudiantes del séptimo grado. Divididas naturalmente las variables en dos grupos, Hotelling investigó la asociación existente entre las habilidades lectoras y de cálculo, llegando a la conclusión de que efectivamente los niños con grandes diferencias entre las variables de lectura presentaban también grandes diferencias entre las variables de cálculo aritmético.

Podemos decir que, en general, *el Análisis de Correlación Canónica mide la potencia de la asociación entre dos conjuntos de variables*, determinando primero pares de combinaciones lineales de dichos conjuntos de variables que presenten la mayor

correlación posible, luego eligiendo otros pares de combinaciones lineales no correlacionados con los anteriores y que maximicen dicha correlación, y así sucesivamente. Estos pares de combinaciones lineales se denominan *variables canónicas* y sus correlaciones reciben el nombre de *correlaciones canónicas*.

Como todos los aspectos de maximización representan un intento de concentrar una relación multidimensional en pocos pares de variables canónicas, esta técnica, al igual que las de Componentes Principales y de Análisis Discriminante, puede conducir a una reducción en las dimensiones del problema.

La naturaleza de la correlación canónica puede ser descripta algebraicamente, como se presenta a continuación.

2. EL MODELO

2.1. La Matemática del Análisis de Correlación Canónica

Sean dos conjuntos de variables aleatorias $\{X_1, X_2, \dots, X_p\}$ e $\{Y_1, Y_2, \dots, Y_q\}$, con $p \leq q$. Pueden representarse como vectores aleatorios $\mathbf{x}' = [X_1, X_2, \dots, X_p]$ e $\mathbf{y}' = [Y_1, Y_2, \dots, Y_q]$, para los cuales se cumple que:

$$E(\mathbf{x}) = \boldsymbol{\mu}_1, \quad \text{Var}(\mathbf{x}) = \boldsymbol{\Sigma}_1, \quad E(\mathbf{y}) = \boldsymbol{\mu}_2, \quad \text{Var}(\mathbf{y}) = \boldsymbol{\Sigma}_2, \quad \text{Cov}(\mathbf{x}, \mathbf{y}) = \boldsymbol{\Sigma}_{12} = \boldsymbol{\Sigma}'_{21}$$

Si consideramos conjuntamente ambos vectores, es decir $\mathbf{z}' = [\mathbf{x}' \ \mathbf{y}']$, es válido que:

$$E(\mathbf{z})' = [\boldsymbol{\mu}'_1 \ \boldsymbol{\mu}'_2] \quad \text{y} \quad \text{Var}(\mathbf{z}) = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix} = \boldsymbol{\Sigma}$$

destacando que esta partición de $\boldsymbol{\Sigma}$ no es necesariamente cuadrada, ya que $p \leq q$.

Por medio de $\boldsymbol{\Sigma}_{12}$ podemos medir la asociación entre los dos conjuntos de variables. Si p y q toman valores relativamente grandes, la interpretación colectiva de los elementos de $\boldsymbol{\Sigma}_{12}$ suele ser dificultosa, por ello es preferible considerar combinaciones lineales de las variables, y con ellas pocas covarianzas (o correlaciones); es decir, en vez de pensar en términos de \mathbf{x} y de \mathbf{y} , pensamos en términos de variables unidimensionales.

Sean pues, $U = \mathbf{a}'\mathbf{x}$ y $V = \mathbf{b}'\mathbf{y}$, para ciertos vectores de coeficientes \mathbf{a} y \mathbf{b} , donde:

$$\begin{aligned} E(U) &= \mathbf{a}' \boldsymbol{\mu}_1 & E(V) &= \mathbf{b}' \boldsymbol{\mu}_2 \\ \text{Var}(U) &= \sigma_U^2 = \mathbf{a}' \text{Var}(\mathbf{x}) \mathbf{a} = \mathbf{a}' \boldsymbol{\Sigma}_{11} \mathbf{a} & \text{Var}(V) &= \sigma_V^2 = \mathbf{b}' \text{Var}(\mathbf{y}) \mathbf{b} = \mathbf{b}' \boldsymbol{\Sigma}_{22} \mathbf{b} \end{aligned}$$

$$\text{Cov}(U, V) = \mathbf{a}' \text{Cov}(\mathbf{x}, \mathbf{y}) \mathbf{b} = \mathbf{a}' \boldsymbol{\Sigma}_{12} \mathbf{b}$$

Debemos hallar los vectores de coeficientes \mathbf{a} y \mathbf{b} que permitan obtener la máxima correlación posible entre las nuevas variables U y V , y definiremos *al primer par de variables canónicas* como al par de combinaciones lineales (U_1, V_1) , con varianza unidad, que maximice su correlación. El *segundo par de variables canónicas* será el par de combinaciones lineales (U_2, V_2) , con varianza unidad, que maximice su correlación entre

todas las elecciones posibles de pares que no estén correlacionados con el primer par de variables canónicas. Así, en el k -ésimo paso, el k -ésimo par de variables canónicas será el par de combinaciones lineales (U_k, V_k) , con varianzas unidad, que maximice su correlación entre todas las elecciones posibles de pares que no estén correlacionados con los primeros $(k-1)$ pares de variables canónicas. La correlación entre las k -ésimas variables canónicas recibe el nombre de k -ésima correlación canónica.

El coeficiente de correlación está dado por la siguiente expresión:

$$\text{Corr}(U, V) = \frac{\text{Cov}(U, V)}{\sqrt{\text{Var}(U) \cdot \text{Var}(V)}} = \frac{\mathbf{a}' \Sigma_{12} \mathbf{b}}{\sqrt{\mathbf{a}' \Sigma_{11} \mathbf{a}} \sqrt{\mathbf{b}' \Sigma_{22} \mathbf{b}}} = \rho$$

la cual puede reescribirse considerando las siguientes identidades (ver Apéndice, 1.3.), bajo el supuesto de que Σ_{11} y Σ_{22} son no singulares, de lo contrario puede reducirse el número de variables de manera apropiada y las combinaciones lineales U y V serán expresadas en términos del conjunto reducido de variables que tendrá una matriz de varianzas-covarianzas no singular:

$$\Sigma_{11} = \Sigma_{11}^{1/2} \Sigma_{11}^{1/2}, \quad \Sigma_{11}^{-1} = \Sigma_{11}^{-1/2} \Sigma_{11}^{-1/2}, \quad \Sigma_{22} = \Sigma_{22}^{1/2} \Sigma_{22}^{1/2}, \quad \Sigma_{22}^{-1} = \Sigma_{22}^{-1/2} \Sigma_{22}^{-1/2}$$

y los vectores $\mathbf{c} = \Sigma_{11}^{1/2} \mathbf{a}$, $\mathbf{d} = \Sigma_{22}^{1/2} \mathbf{b}$, de modo que

$$\boxed{\mathbf{a} = \Sigma_{11}^{-1/2} \mathbf{c}} \quad \text{y} \quad \boxed{\mathbf{b} = \Sigma_{22}^{-1/2} \mathbf{d}}$$

Entonces:

$$\frac{\mathbf{a}' \Sigma_{12} \mathbf{b}}{\sqrt{\mathbf{a}' \Sigma_{11} \mathbf{a}} \sqrt{\mathbf{b}' \Sigma_{22} \mathbf{b}}} = \frac{\mathbf{c}' \Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1/2} \mathbf{d}}{\sqrt{\mathbf{c}' \mathbf{c}} \sqrt{\mathbf{d}' \mathbf{d}}} \quad (1)$$

si consideramos el numerador $\underbrace{\mathbf{c}' \Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1/2} \mathbf{d}}_{\mathbf{h}' \mathbf{d}}$, utilizando la desigualdad de Cauchy-Schwarz (Apéndice, 1.4.), es:

$$\mathbf{h}' \mathbf{d} \leq (\mathbf{h}' \mathbf{h})^{1/2} (\mathbf{d}' \mathbf{d})^{1/2} = (\mathbf{c}' \Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1/2} \Sigma_{22}^{-1/2} \Sigma_{21} \Sigma_{11}^{-1/2} \mathbf{c})^{1/2} (\mathbf{d}' \mathbf{d})^{1/2}$$

Retomando la expresión (1):

$$\frac{\mathbf{c}' \Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1/2} \mathbf{d}}{\sqrt{\mathbf{c}' \mathbf{c}} \sqrt{\mathbf{d}' \mathbf{d}}} \leq \frac{(\mathbf{c}' \mathbf{S} \mathbf{c})^{1/2} (\mathbf{d}' \mathbf{d})^{1/2}}{(\mathbf{c}' \mathbf{c})^{1/2} (\mathbf{d}' \mathbf{d})^{1/2}} = \frac{(\mathbf{c}' \mathbf{S} \mathbf{c})^{1/2}}{(\mathbf{c}' \mathbf{c})^{1/2}} \leq \sqrt{\lambda_1}$$

o, equivalentemente, $\mathbf{c}' \mathbf{S} \mathbf{c} \leq \lambda_1 \mathbf{c}' \mathbf{c}$, con $\mathbf{S} = \Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1/2}$ matriz simétrica; la igualdad resulta de aplicar el segundo principio de maximización (Apéndice, 1.5.), cuando λ_1 y $\mathbf{c} = \mathbf{e}_1$ son los primeros autovalor (el mayor) y autovector respectivamente asociados a \mathbf{S} . También obtendríamos la igualdad si tomáramos $\mathbf{d} = \mathbf{f}_1$, el primer vector propio de $\mathbf{F} = \Sigma_{22}^{-1/2} \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1/2}$ (con \mathbf{f}_1 forma normalizada de $\Sigma_{22}^{-1/2} \Sigma_{21} \Sigma_{11}^{-1/2} \mathbf{e}_1$) ya que se verifica la siguiente correspondencia al multiplicar $\mathbf{S} \mathbf{e}_1 = \lambda_1 \mathbf{e}_1$ por $\Sigma_{22}^{-1/2} \Sigma_{21} \Sigma_{11}^{-1/2}$.

En efecto:

$$\begin{aligned} \mathbf{S}\mathbf{e}_1 = \lambda_1\mathbf{e}_1 &\Rightarrow (\Sigma_{22}^{-1/2} \Sigma_{21} \Sigma_{11}^{-1/2})\mathbf{S}\mathbf{e}_1 = \lambda_1 (\Sigma_{22}^{-1/2} \Sigma_{21} \Sigma_{11}^{-1/2}) \mathbf{e}_1 \\ &\Rightarrow (\Sigma_{22}^{-1/2} \Sigma_{21} \Sigma_{11}^{-1/2}) \Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1/2} \mathbf{e}_1 = \lambda_1 \mathbf{f}_1 \end{aligned}$$

Como $\Sigma_{22}^{-1} = \Sigma_{22}^{-1/2} \Sigma_{22}^{-1/2}$:

$$\begin{aligned} &\Rightarrow \Sigma_{22}^{-1/2} \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1/2} (\Sigma_{22}^{-1/2} \Sigma_{21} \Sigma_{11}^{-1/2} \mathbf{e}_1) = \lambda_1 \mathbf{f}_1 \\ &\Rightarrow \mathbf{F} \mathbf{f}_1 = \lambda_1 \mathbf{f}_1 \end{aligned}$$

De este modo, si $(\lambda_1, \mathbf{e}_1)$ es el par formado por los primeros autovalor y autovector de \mathbf{S} , resulta $(\lambda_1, \mathbf{f}_1)$ el par formado por los primeros autovalor y autovector de \mathbf{F} , donde el signo de \mathbf{f}_1 se elige para tener una correlación positiva.

En síntesis, queda demostrado que $U_1 = \mathbf{e}'_1 \Sigma_{11}^{-1/2} \mathbf{x}$ y $V_1 = \mathbf{f}'_1 \Sigma_{22}^{-1/2} \mathbf{y}$ forman el primer par de variables canónicas, siendo su correlación $\rho_1 = \sqrt{\lambda_1}$. Además:

$$\text{Var}(U_1) = \mathbf{e}'_1 \Sigma_{11}^{-1/2} \Sigma_{11} \Sigma_{11}^{-1/2} \mathbf{e}_1 = \mathbf{e}'_1 \mathbf{e}_1 = 1$$

$$\text{Var}(V_1) = \mathbf{f}'_1 \Sigma_{22}^{-1/2} \Sigma_{22} \Sigma_{22}^{-1/2} \mathbf{f}_1 = \mathbf{f}'_1 \mathbf{f}_1 = 1$$

U_1 y cualquier combinación lineal arbitraria $\mathbf{a}'\mathbf{x} = \mathbf{c}' \Sigma_{11}^{-1/2} \mathbf{x}$ no estarán correlacionados si:

$$\text{Cov}(U_1, \mathbf{c}' \Sigma_{11}^{-1/2} \mathbf{x}) = \mathbf{e}'_1 \Sigma_{11}^{-1/2} \Sigma_{11} \Sigma_{11}^{-1/2} \mathbf{c} = \mathbf{e}'_1 \mathbf{c} = 0 \quad \text{o} \quad \mathbf{c} \perp \mathbf{e}_1$$

El paso k-ésimo requiere que $\mathbf{c} \perp \mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_{k-1}$ y, como indica el inciso c) del segundo principio de maximización del Apéndice, resulta $\mathbf{c}' \Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1/2} \mathbf{c} \leq \lambda_k \mathbf{c}'\mathbf{c}$ para $\mathbf{c} \perp \mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_{k-1}$, y la correlación de las combinaciones lineales k-ésimas será:

$$\text{Corr}(\mathbf{a}'\mathbf{x}, \mathbf{b}'\mathbf{y}) = \frac{\mathbf{a}'\Sigma_{12}\mathbf{b}}{\sqrt{\mathbf{a}'\Sigma_{11}\mathbf{a}}\sqrt{\mathbf{b}'\Sigma_{22}\mathbf{b}}} = \frac{\mathbf{c}'\Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1/2}\mathbf{d}}{\sqrt{\mathbf{c}'\mathbf{c}}\sqrt{\mathbf{d}'\mathbf{d}}} \leq \sqrt{\lambda_k}$$

(la igualdad se verifica cuando $\mathbf{c} = \mathbf{e}_k$ o $\mathbf{a} = \Sigma_{11}^{-1/2} \mathbf{e}_k$ y $\mathbf{b} = \Sigma_{22}^{-1/2} \mathbf{f}_k$).

Entonces $U_k = \mathbf{e}'_k \Sigma_{11}^{-1/2} \mathbf{x}$ y $V_k = \mathbf{f}'_k \Sigma_{22}^{-1/2} \mathbf{y}$ son el k-ésimo par canónico, cuya correlación está dada por $\sqrt{\lambda_k} = \rho_k$ y no están correlacionados con los demás U_j o V_j ya que, para $k \neq j \leq p$:

$$\text{Cov}(V_k, V_j) = \mathbf{f}'_k \Sigma_{22}^{-1/2} \Sigma_{22} \Sigma_{22}^{-1/2} \mathbf{f}_j = \mathbf{f}'_k \mathbf{f}_j = 0$$

y

$$\text{Cov}(U_k, V_j) = \mathbf{e}'_k \Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1/2} \mathbf{f}_j = \mathbf{e}'_k \mathbf{f}_j = 0 \quad (\text{por ser } \mathbf{f}'_j \text{ múltiplo de } \mathbf{e}'_j \Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1/2})$$

2.1.1. Notas Ampliatorias

a) Si \mathbf{x} e \mathbf{y} son vectores de variables aleatorias estandarizadas, las matrices Σ involucradas se convierten en matrices de correlación ρ : $\rho_{11}, \rho_{12}, \rho_{21}$ y ρ_{22} . Los

vectores \mathbf{e}_k y \mathbf{f}_k están asociados a $\rho_{11}^{-1/2} \rho_{12} \rho_{22}^{-1} \rho_{21} \rho_{11}^{-1/2}$ y $\rho_{22}^{-1/2} \rho_{21} \rho_{11}^{-1} \rho_{12} \rho_{22}^{-1/2}$ respectivamente, para $k = 1, \dots, p$.

b) Las correlaciones canónicas NO cambian con la estandarización. Dado que:

$$\begin{aligned} \mathbf{a}'_k (\mathbf{x} - \boldsymbol{\mu}_1) &= a_{k1} (X_1 - \mu_{11}) + a_{k2} (X_2 - \mu_{12}) + \dots + a_{kp} (X_p - \mu_{1p}) \\ &= a_{k1} \sqrt{\sigma_{11}} \frac{(X_1 - \mu_{11})}{\sqrt{\sigma_{11}}} + a_{k2} \sqrt{\sigma_{22}} \frac{(X_2 - \mu_{12})}{\sqrt{\sigma_{22}}} + \dots + a_{kp} \sqrt{\sigma_{pp}} \frac{(X_p - \mu_{1p})}{\sqrt{\sigma_{pp}}} \end{aligned}$$

y similarmente para \mathbf{y} :

$$\begin{aligned} \mathbf{b}'_k (\mathbf{y} - \boldsymbol{\mu}_2) &= b_{k1} (Y_1 - \mu_{21}) + b_{k2} (Y_2 - \mu_{22}) + \dots + b_{kp} (Y_p - \mu_{2p}) \\ &= b_{k1} \sqrt{\sigma_{11}} \frac{(Y_1 - \mu_{21})}{\sqrt{\sigma_{11}}} + b_{k2} \sqrt{\sigma_{22}} \frac{(Y_2 - \mu_{22})}{\sqrt{\sigma_{22}}} + \dots + b_{kp} \sqrt{\sigma_{pp}} \frac{(Y_p - \mu_{2p})}{\sqrt{\sigma_{pp}}} \end{aligned}$$

si llamamos $\mathbf{V}_{11}^{1/2}$ y $\mathbf{V}_{22}^{1/2}$ a las matrices diagonales cuyos elementos no nulos son, respectivamente, $\sqrt{\sigma_{ii}} = \sqrt{\text{Var}(X_i)}$ y $\sqrt{\sigma_{ii}} = \sqrt{\text{Var}(Y_i)}$; tendremos que, si para las k -ésimas variables canónicas U_k y V_k sus vectores de coeficientes eran \mathbf{a}'_k y \mathbf{b}'_k los que las relacionaban con $\{X_1, X_2, \dots, X_p\}$ y $\{Y_1, Y_2, \dots, Y_p\}$, a partir de las variables estandarizadas serán $\mathbf{a}'_k \mathbf{V}_{11}^{1/2}$ y $\mathbf{b}'_k \mathbf{V}_{22}^{1/2}$ los vectores de coeficientes correspondientes, pero no cambiarán las correlaciones canónicas.

c) Sólo si hay autovalores con multiplicidad mayor a 1, la elección de los vectores de coeficientes \mathbf{a}'_k y \mathbf{b}'_k no será única.

d) George Seber (1984) propone maximizar la *correlación cuadrática*.

El máximo $\lambda_1 = \rho_1$ se alcanza cuando $\mathbf{a} = \mathbf{e}_1$, autovector de $\Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}$ y $\mathbf{b} = \mathbf{f}_1$, autovector de $\Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12}$.

Demostración:

$$\begin{aligned} \max_{\mathbf{a}, \mathbf{b}} \frac{(\mathbf{a}' \Sigma_{12} \mathbf{b})^2}{\mathbf{a}' \Sigma_{11} \mathbf{a} \cdot \mathbf{b}' \Sigma_{22} \mathbf{b}} &= \max_{\mathbf{b}} \left(\frac{1}{\mathbf{b}' \Sigma_{22} \mathbf{b}} \cdot \max_{\mathbf{a}} \frac{(\mathbf{a}' \Sigma_{12} \mathbf{b})^2}{\mathbf{a}' \Sigma_{11} \mathbf{a}} \right) \quad \text{Sea } \mathbf{y} = \Sigma_{12} \mathbf{b}, \text{ entonces:} \\ &= \max_{\mathbf{b}} \left(\frac{1}{\mathbf{b}' \Sigma_{22} \mathbf{b}} \cdot \max_{\mathbf{a}} \frac{(\mathbf{a}' \mathbf{y})^2}{\mathbf{a}' \Sigma_{11} \mathbf{a}} \right) \end{aligned}$$

por el Primer Principio de Maximización (Apéndice, 1.5.) para $\mathbf{a} = \mathbf{c} \Sigma_{11}^{-1} \mathbf{y}$, resulta:

$$\max_{\mathbf{b}} \left(\frac{1}{\mathbf{b}' \Sigma_{22} \mathbf{b}} \cdot \max_{\mathbf{a}} \frac{(\mathbf{a}' \mathbf{y})^2}{\mathbf{a}' \Sigma_{11} \mathbf{a}} \right) = \max_{\mathbf{b}} \left(\frac{\mathbf{y}' \Sigma_{11}^{-1} \mathbf{y}}{\mathbf{b}' \Sigma_{22} \mathbf{b}} \right) = \max_{\mathbf{b}} \left(\frac{\mathbf{b}' \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} \mathbf{b}}{\mathbf{b}' \Sigma_{22} \mathbf{b}} \right)$$

por la Nota Ampliatoria de Principios de Maximización (Apéndice, 1.5.1), se obtiene:

$$\max_{\mathbf{b}} \left(\frac{\mathbf{b}' \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} \mathbf{b}}{\mathbf{b}' \Sigma_{22} \mathbf{b}} \right) = \lambda_1 \quad \text{para } \mathbf{b} = \mathbf{e}_1 \text{ de } \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12}, \text{ siendo } \sqrt{\lambda_1} \text{ la primera correlación canónica entre } U_1 \text{ y } V_1.$$

¿Por qué preferir un algoritmo sobre el otro? Pues, porque estas matrices generalmente no son simétricas.

¿Son realmente equivalentes los algoritmos? Sí, ya que:

$$\frac{\mathbf{b}'\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\mathbf{b}}{\mathbf{b}'\Sigma_{22}\mathbf{b}} = \frac{\mathbf{b}'\mathbf{I}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\mathbf{Ib}}{\mathbf{b}'\Sigma_{22}\mathbf{b}} = \frac{\mathbf{b}'\Sigma_{22}^{1/2}\Sigma_{22}^{-1/2}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1/2}\Sigma_{22}^{1/2}\mathbf{b}}{\mathbf{b}'\Sigma_{22}^{1/2}\Sigma_{22}^{1/2}\mathbf{b}}$$

Sea $\mathbf{z} = \Sigma_{22}^{1/2}\mathbf{b}$. Luego:

$$\frac{\mathbf{b}'\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\mathbf{b}}{\mathbf{b}'\Sigma_{22}\mathbf{b}} = \frac{\mathbf{z}'\Sigma_{22}^{-1/2}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1/2}\mathbf{z}}{\mathbf{z}'\mathbf{z}} = \frac{\mathbf{z}'\mathbf{Cz}}{\mathbf{z}'\mathbf{z}}$$

alcanza su máximo en λ_1 con \mathbf{z}_1 su vector propio asociado a la matriz \mathbf{C} que coincide con la matriz \mathbf{F} presentada en la página 3, y $\text{Corr}(U_1, V_1) = \sqrt{\lambda_1}$.

2.1.2. Variables y Correlaciones Canónicas Muestrales

Una muestra aleatoria de n observaciones para cada una de las $(p+q)$ variables: $\{X_1, X_2, \dots, X_p, Y_1, Y_2, \dots, Y_q\}$, puede considerarse dada por la matriz \mathbf{Z} :

$$\mathbf{Z} = \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \dots & \vdots \\ x_{p1} & x_{p2} & \dots & x_{pn} \\ y_{11} & y_{12} & \dots & y_{1n} \\ y_{21} & y_{22} & \dots & y_{2n} \\ \vdots & \vdots & \dots & \vdots \\ y_{q1} & y_{q2} & \dots & y_{qn} \end{bmatrix} = [\mathbf{z}_1 \quad \mathbf{z}_2 \quad \dots \quad \mathbf{z}_n]$$

con $\mathbf{z}'_i = [x_{1i}, x_{2i}, \dots, x_{pi}, y_{1i}, y_{2i}, \dots, y_{qi}]$. El vector de medias muestrales se denota por $\bar{\mathbf{z}}' = [\bar{\mathbf{x}}', \bar{\mathbf{y}}']$, con $\bar{\mathbf{x}}'$ e $\bar{\mathbf{y}}'$ vectores de medias muestrales para \mathbf{X} e \mathbf{Y} respectivamente.

La matriz de varianzas-covarianzas muestral está dada por:

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{bmatrix}$$

donde \mathbf{S}_{11} y \mathbf{S}_{22} representan las matrices de varianzas-covarianzas para \mathbf{X} e \mathbf{Y} respectivamente, mientras que $\mathbf{S}_{12} = \mathbf{S}'_{21} = \text{Cov}(\mathbf{X}, \mathbf{Y})$.

Las combinaciones lineales $\hat{U} = \hat{\mathbf{a}}'\mathbf{x}$ y $\hat{V} = \hat{\mathbf{b}}'\mathbf{y}$ tienen como coeficiente de correlación muestral a:

$$r_{\hat{U}, \hat{V}} = \frac{\hat{\mathbf{a}}'\mathbf{S}_{12}\hat{\mathbf{b}}}{\sqrt{\hat{\mathbf{a}}'\mathbf{S}_{11}\hat{\mathbf{a}}}\sqrt{\hat{\mathbf{b}}'\mathbf{S}_{22}\hat{\mathbf{b}}}}$$

Como para las variables canónicas poblacionales, llamaremos *primer par de variables canónicas muestrales* al par de combinaciones lineales (\hat{U}_1, \hat{V}_1) con varianza unidad, que maximice su correlación. El *k-ésimo par de variables canónicas muestrales* será el par de combinaciones lineales (\hat{U}_k, \hat{V}_k) , con varianza unidad, que maximice su correlación entre todas las elecciones posibles de pares que no estén correlacionados con los primeros (k-1) pares de variables canónicas muestrales. La correlación muestral entre las k-ésimas variables canónicas muestrales recibe el nombre de *k-ésima correlación canónica muestral*.

Tanto las variables canónicas muestrales como los coeficientes de correlación muestral pueden obtenerse a partir de S_{11} , S_{22} y S_{12} de manera consistente a la ya presentada para el caso poblacional: Sean $\hat{\lambda}_i$ y \hat{e}_i (para $i=1, \dots, p$) los correspondientes autovalores y autovectores de $S_{11}^{-1/2} S_{12} S_{22}^{-1} S_{21} S_{11}^{-1/2}$; mientras que $\hat{\lambda}_k$ y \hat{f}_k (para $k=1, \dots, q$) lo son para $S_{22}^{-1/2} S_{21} S_{11}^{-1} S_{12} S_{22}^{-1/2}$. Los p primeros vectores propios pueden obtenerse de $\hat{f}_k = \frac{1}{\hat{\lambda}_k} S_{22}^{-1/2} S_{21} S_{11}^{-1/2} \hat{e}_k$ (en este caso, $k=1, \dots, p$).

De este modo, $\hat{U}_k = \hat{e}_k' S_{11}^{-1/2} \mathbf{x}$ y $\hat{V}_k = \hat{f}_k' S_{22}^{-1/2} \mathbf{y}$, donde \mathbf{x} e \mathbf{y} en este caso representan los valores para una unidad experimental en particular. La mayor correlación muestral será entonces $r_{\hat{U}_1, \hat{V}_1} = \sqrt{\hat{\lambda}_1}$ y la k-ésima será: $r_{\hat{U}_k, \hat{V}_k} = \sqrt{\hat{\lambda}_k}$, la mayor posible entre todos los pares de combinaciones lineales no correlacionadas con los (k-1) pares anteriores de variables canónicas muestrales.

Como se expresó en 2.1.1.b), los coeficientes para las variables estandarizadas (en este caso, muestrales) estarán dados por:

$\hat{\mathbf{u}} = \hat{\mathbf{A}}' \mathbf{D}_{11}^{1/2} \mathbf{x}$ y $\hat{\mathbf{v}} = \hat{\mathbf{B}}' \mathbf{D}_{22}^{1/2} \mathbf{y}$ con $\hat{\mathbf{A}}' = [\hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2, \dots, \hat{\mathbf{a}}_p]'$, $\hat{\mathbf{B}}' = [\hat{\mathbf{b}}_1, \hat{\mathbf{b}}_2, \dots, \hat{\mathbf{b}}_q]'$, $\mathbf{D}_{11}^{1/2}$ y $\mathbf{D}_{22}^{1/2}$ matrices diagonales (de dimensión $p \times p$ y $q \times q$ respectivamente) cuyos i-ésimos elementos son las raíces de las varianzas muestrales para \mathbf{x} e \mathbf{y} .

2.2. Inferencia para muestras de gran tamaño

2.2.1. Test de Cociente de Verosimilitud

Cuando $\Sigma_{12} = \mathbf{0}$, entonces $\text{Cov}(\mathbf{a}'\mathbf{x}, \mathbf{b}'\mathbf{y}) = \mathbf{a}'\Sigma_{12}\mathbf{b} = 0$ para cualesquiera vectores \mathbf{a} y \mathbf{b} . Por ello, todas las correlaciones canónicas serán nulas y no tendrá objeto intentar un análisis de correlación canónica. Surge entonces la pregunta: ¿Cómo testar $\Sigma_{12} = \mathbf{0}$ para muestras grandes? A continuación la respuesta brindada por Kshirsagar (en Seber, 1984).

Sea $\mathbf{z}'_j = [\mathbf{x}'_j \ \mathbf{y}'_j]$, $j = 1, \dots, n$ una muestra aleatoria de una población $\sim N_{p+q}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ con

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}$$

(p x p) (p x q)
(q x p) (q x q)

El test de cociente de verosimilitud para testar $H_0: \boldsymbol{\Sigma}_{12} = \mathbf{0}$ versus $H_1: \boldsymbol{\Sigma}_{12} \neq \mathbf{0}$ rechaza la hipótesis nula para valores grandes de

$$-2\text{Log}(\Lambda) = n \text{Log} \left(\frac{|\mathbf{S}_{11}| |\mathbf{S}_{22}|}{|\mathbf{S}|} \right) = -n \text{Log} \prod_{i=1}^p (1 - \hat{\lambda}_i)$$

el cual se distribuye aproximadamente como χ^2_{pq} cuando el tamaño muestral es suficientemente grande.

\mathbf{S} es un estimador insesgado de $\boldsymbol{\Sigma}$, con

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{bmatrix}$$

(p x p) (p x q)
(q x p) (q x q)

El cociente de verosimilitud compara el determinante de la varianza muestral generalizada bajo H_0 :

$$\begin{vmatrix} \mathbf{S}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_{22} \end{vmatrix} = |\mathbf{S}_{11}| |\mathbf{S}_{22}|$$

con el correspondiente a la varianza generalizada 'no restringida' $|\mathbf{S}|$

2.2.2. Test de Bartlett

En 1939, Bartlett sugirió reemplazar el factor "n" en el estadístico anterior por " $n - 1 - \frac{1}{2}(p + q + 1)$ " para mejorar la aproximación de la distribución muestral de $-2\text{Log}(\Lambda)$ a la distribución χ^2 (Johnson y Wichern, 1992). Así, para valores grandes de n y de $n - 1 - \frac{1}{2}(p + q + 1)$, podremos rechazar $H_0: \boldsymbol{\Sigma}_{12} = \mathbf{0}$ ($\rho_1 = \rho_2 = \dots = \rho_p = 0$) con nivel de significación α si:

$$\phi_0^2 = -\left(n - 1 - \frac{1}{2}(p + q + 1)\right) \text{Log} \prod_{i=1}^p (1 - \hat{\rho}_i^2) > \chi_{pq}^2(\alpha)$$

con $\chi_{pq}^2(\alpha)$ percentil superior (100α) de una distribución χ^2 con pq grados de libertad.

Si ϕ_0^2 es suficientemente grande se concluye que existe al menos una correlación canónica significativa. Si ϕ_0^2 no es suficientemente grande, entonces no hay evidencias de algún tipo de relación entre las variables $\{X_1, X_2, \dots, X_p\}$ e $\{Y_1, Y_2, \dots, Y_q\}$.

Asumiendo que ϕ_0^2 es significativo, el próximo paso consiste en remover el efecto de la primera correlación canónica del estadístico y considerar:

$$\phi_1^2 = -\left(n - 1 - \frac{1}{2}(p + q + 1)\right) \sum_{i=2}^p \text{Log}(1 - \hat{\lambda}_i)$$

(equivalente a la fórmula anterior por aplicar logaritmo natural al producto y reemplazando el cuadrado de cada correlación canónica por el autovalor correspondiente). Este nuevo estadístico se distribuye aproximadamente como χ^2 con $(p-1)(q-1)$ grados de libertad.

Si ϕ_1^2 es significativo, diremos que existen al menos dos correlaciones canónicas significativas. Si ϕ_1^2 no lo es, entonces la primera correlación canónica será considerada como representativa de toda la relación que pueda existir entre las variables X e Y.

Si ambos, ϕ_0^2 y ϕ_1^2 son significativos, entonces deberá removerse el efecto de las dos primeras correlaciones canónicas para ver si las correlaciones remanentes son significativas. Este proceso continúa hasta que se encuentra que dichas correlaciones no son significativas y pueden despreciarse.

El estadístico del test después de que las primeras j correlaciones canónicas han sido removidas es:

$$\phi_j^2 = -\left(n - 1 - \frac{1}{2}(p + q + 1)\right) \sum_{i=j+1}^p \text{Log}(1 - \hat{\lambda}_i)$$

con $(p-j)(q-j)$ grados de libertad.

Por ser un test relativamente simple es muy utilizado, pero debemos efectuar algunas observaciones al mismo:

- a) este encadenamiento (o secuencia) de tests asume que la Hipótesis Nula es: $H_0^{(j)}: \rho_1 \neq 0, \rho_2 \neq 0, \dots, \rho_j \neq 0, \rho_{j+1} = 0, \dots, \rho_p = 0$ versus la Hipótesis Alternativa $H_1^{(j)}: \rho_i \neq 0$, para algún $i \geq j+1$, que es testada por el criterio de máxima verosimilitud. Es decir: $\phi_j^2 > \chi_{(p-j)(q-j)}^2(\alpha)$, donde el nivel global de significación deja de ser α (de hecho, es muy difícil determinarlo con precisión).
- b) Este procedimiento tiende a concluir que la hipótesis nula es correcta simplemente porque no es rechazada, por lo que se pide precaución al trabajar con este test, tomando los resultados obtenidos sólo como indicadores “gruesos” para seleccionar el número de variables canónicas de importancia, y comparar en lo posible con resultados ya avalados.

2.2.3. Test de Lawley

Lawley propone una nueva modificación al estadístico original para mejorar la aproximación χ^2 adoptando el factor:

$$-\left(n - 1 - k - \frac{1}{2}(p + q + 1) + \sum_{i=1}^k \hat{\rho}_i^2\right)$$

Para determinar k debemos testar la secuencia de hipótesis como propone Bartlett hasta obtener un resultado no significativo para H_{0r} , entonces se elige $k = r$ (según Glynn y Muirhead, en Seber, 1984).

Sin embargo, este último autor comenta la indicación de Harris acerca de que el estadístico que utiliza la modificación de Lawley no es aproximadamente χ^2 , y sugiere por tanto abandonarlo.

2.2.4. Otros Tests

En su compilación, Seber (1984) brinda referencias sobre otros tests desarrollados por varios autores, entre ellos: a) Fujikoshi y Veitch, quienes proponen otros procedimientos usando extensiones del estadístico C_p de Mallows; b) Fujikoshi (presenta una expansión de la distribución bajo la hipótesis alternativa del estadístico de Lawley para muestras grandes) y c) Akaike, quien expone un procedimiento basado en el criterio de información.

2.2.5. Problemas para la aplicación de los tests

Al igual que otros tests basados en matrices de dispersión muestrales, el estadístico $-2\text{Log}(\Lambda)$ no es robusto y se ve afectado por la no normalidad, especialmente si hay 'outliers' o se trata de distribuciones de colas pesadas. Se han propuesto alternativas basadas en las distribuciones asintóticas de $\sqrt{n}(\hat{\lambda}_i - \lambda_i)$ (normales con media cero), pero el problema consiste en que la matriz de varianzas-covarianzas asintótica ahora contiene términos adicionales, por lo que $-2\text{Log}(\Lambda)$ se convierte en una suma ponderada de variables χ^2 , tal como demuestran Muirhead y Waternaux (Seber, 1984).

2.3. La Interpretación

2.3.1. Identificación de las variables canónicas poblacionales

Una vez que se ha decidido cuáles variables canónicas (si las hay) son significativas para un conjunto de datos bajo estudio, el siguiente aspecto a considerar es *cómo interpretar* dichas variables canónicas.

Las variables canónicas son, en general, artificiales, sin sentido físico. Si se utilizan las variables originales $\mathbf{x} = \{X_1, X_2, \dots, X_p\}$ e $\mathbf{y} = \{Y_1, Y_2, \dots, Y_q\}$, los coeficientes canónicos \mathbf{a} y \mathbf{b} tienen unidades proporcionales a las de \mathbf{x} e \mathbf{y} . Si se utilizan variables estandarizadas, \mathbf{a} y \mathbf{b} no tienen unidades de medida y deben ser interpretados en términos de las variables estandarizadas.

A primera vista, puede parecer fácil describir los elementos de los vectores \mathbf{u} y \mathbf{v} ($\mathbf{u} = \{U_1, U_2, \dots, U_p\}$ y $\mathbf{v} = \{V_1, V_2, \dots, V_p\}$) en términos de \mathbf{x} y de \mathbf{y} si sus coeficientes son grandes (ora positivos, ora negativos). Desafortunadamente, las correlaciones entre las variables originales pueden alterar la interpretación: por ejemplo, podría ocurrir que a_{i1} sea positivo y sin embargo, la correlación entre U_i y X_1 sea negativa. Esta aparente contradicción puede darse porque X_1 esté altamente correlacionada con otras X_i y parte del efecto de X_1 está contado en los restantes coeficientes.

De hecho, si alguna X_i es (o casi es) combinación lineal de otras X , habrá infinitas variedades de combinaciones lineales de \mathbf{x} , algunas con a_{ij} muy diferentes, que virtualmente darán los mismos valores para U_i . Lo mismo puede ocurrir con \mathbf{y} . Entonces, se debe ser muy prudente con el análisis ya que si \mathbf{x} o \mathbf{y} están altamente correlacionadas, será muy difícil poder “desenredar” la información de cuánto contribuyen a las variables canónicas.

Algunos autores han sugerido que es mejor estudiar las correlaciones de \mathbf{u} y \mathbf{v} con \mathbf{x} e \mathbf{y} en vez de decidir sobre \mathbf{a} y \mathbf{b} . De este modo, por ejemplo, si U_i está altamente correlacionada con X_1 (en sentido positivo), puede decirse que U_i refleja en gran medida el comportamiento de X_1 ; si V_i está altamente correlacionada con Y_1 (en sentido negativo), diremos que V_i refleja al opuesto de Y_1 . Esta aproximación por lo menos tiene el mérito de mostrar todas las variables con las cuales las variables canónicas parecen estar relacionadas.

Sin embargo, todavía hay que tener cuidado al interpretar estas correlaciones, pues ellas sólo darán *información univariada* en el sentido de que *no indican cómo* las variables originales *contribuyen conjuntamente* al análisis canónico. Por ello, muchos investigadores prefieren declarar o informar la contribución de las variables originales directamente desde los coeficientes estandarizados. Veamos por qué:

Como $\mathbf{u} = \mathbf{A}'\mathbf{x}$, con $\mathbf{A}' = [a_1, a_2, \dots, a_p]'$ resulta:

$$\text{Cov}(\mathbf{u}, \mathbf{x}) = \text{Cov}(\mathbf{A}'\mathbf{x}, \mathbf{x}) = \mathbf{A}' \text{Cov}(\mathbf{x}, \mathbf{x}) = \mathbf{A}'\Sigma_{11}$$

Además:

$$\text{Corr}(U_i, X_k) = \frac{\text{Cov}(U_i, X_k)}{\sqrt{\text{Var}(U_i)}\sqrt{\sigma_{kk}}} = \frac{\text{Cov}(U_i, X_k)}{\sqrt{\sigma_{kk}}} = \text{Cov}(U_i, (\sigma_{kk})^{-1/2} X_k)$$

Por lo tanto:

$$\rho_{\mathbf{u}, \mathbf{x}} = \text{Corr}(\mathbf{u}, \mathbf{x}) = \text{Cov}(\mathbf{u}, \mathbf{V}_{11}^{-1/2} \mathbf{x}) = \text{Cov}(\mathbf{A}'\mathbf{x}, \mathbf{V}_{11}^{-1/2} \mathbf{x}) = \mathbf{A}'\Sigma_{11} \mathbf{V}_{11}^{-1/2}$$

Por un razonamiento similar, para $\mathbf{v} = \mathbf{B}'\mathbf{y}$, con $\mathbf{B}' = [b_1, b_2, \dots, b_q]'$:

$$\text{tr}(\mathbf{R}_{22|r}) = \text{tr}(\hat{\mathbf{b}}'_{z_1} \hat{\mathbf{b}}_{z_1} + \hat{\mathbf{b}}'_{z_2} \hat{\mathbf{b}}_{z_2} + \dots + \hat{\mathbf{b}}'_{z_r} \hat{\mathbf{b}}_{z_r}) = \sum_{i=1}^r \sum_{k=1}^p r_{\hat{v}_i, Y_{zk}}^2$$

La proporción de la varianza muestral total estandarizada explicada por las r primeras variables canónicas es:

$$\text{tr}(\mathbf{R}_{11|r}) / \text{tr}(\mathbf{R}_{11}) = \frac{\sum_{i=1}^r \sum_{k=1}^p r_{\hat{u}_i, X_{zk}}^2}{p} = R_{x_z | \hat{u}_1, \hat{u}_2, \dots, \hat{u}_r}^2$$

$$\text{tr}(\mathbf{R}_{22|r}) / \text{tr}(\mathbf{R}_{22}) = \frac{\sum_{i=1}^r \sum_{k=1}^p r_{\hat{v}_i, Y_{zk}}^2}{q} = R_{y_z | \hat{v}_1, \hat{v}_2, \dots, \hat{v}_r}^2$$

Estas proporciones son medidas descriptivas que dan idea de *cuán buena es la representación* de las variables canónicas de sus respectivos conjuntos de variables originales. También permiten describir las matrices de error de aproximación, ya que:

$$\text{tr}(\tilde{\mathbf{S}}_{11})/p = 1 - R_{x_z | \hat{u}_1, \hat{u}_2, \dots, \hat{u}_r}^2 \quad \text{y} \quad \text{tr}(\tilde{\mathbf{S}}_{22})/q = 1 - R_{y_z | \hat{v}_1, \hat{v}_2, \dots, \hat{v}_r}^2$$

2.5. Más de dos conjuntos de variables

La extensión de dos a tres o más conjuntos de variables comienza con considerar los siguientes conjuntos de variables aleatorias: $\{X_1, X_2, \dots, X_p\}$, $\{Y_1, Y_2, \dots, Y_q\}$ y $\{Z_1, Z_2, \dots, Z_r\}$ con $p \leq q \leq r$. Pueden representarse como vectores aleatorios: $\mathbf{x}' = [X_1, X_2, \dots, X_p]$, $\mathbf{y}' = [Y_1, Y_2, \dots, Y_q]$ y $\mathbf{z}' = [Z_1, Z_2, \dots, Z_r]$, para los cuales se cumple que:

$$E(\mathbf{x}) = \boldsymbol{\mu}_1, \quad \text{Var}(\mathbf{x}) = \boldsymbol{\Sigma}_1, \quad E(\mathbf{y}) = \boldsymbol{\mu}_2, \quad \text{Var}(\mathbf{y}) = \boldsymbol{\Sigma}_2, \quad E(\mathbf{z}) = \boldsymbol{\mu}_3, \quad \text{Var}(\mathbf{z}) = \boldsymbol{\Sigma}_3$$

$$\text{Cov}(\mathbf{x}, \mathbf{y}) = \boldsymbol{\Sigma}_{12} = \boldsymbol{\Sigma}'_{21}, \quad \text{Cov}(\mathbf{x}, \mathbf{z}) = \boldsymbol{\Sigma}_{13} = \boldsymbol{\Sigma}'_{31}, \quad \text{Cov}(\mathbf{y}, \mathbf{z}) = \boldsymbol{\Sigma}_{23} = \boldsymbol{\Sigma}'_{32}$$

Si consideramos conjuntamente los tres vectores, es decir $\mathbf{w}' = [\mathbf{x}' \ \mathbf{y}' \ \mathbf{z}']$, es válido que:

$$E(\mathbf{w})' = [\boldsymbol{\mu}'_1 \quad \boldsymbol{\mu}'_2 \quad \boldsymbol{\mu}'_3] \quad \text{y} \quad \text{Var}(\mathbf{w}) = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} & \boldsymbol{\Sigma}_{13} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} & \boldsymbol{\Sigma}_{23} \\ \boldsymbol{\Sigma}_{31} & \boldsymbol{\Sigma}_{32} & \boldsymbol{\Sigma}_{33} \end{bmatrix} = \boldsymbol{\Sigma}$$

destacando que esta partición de $\boldsymbol{\Sigma}$ no es necesariamente regular, ya que para $p \leq q \leq r$, puede ser que no todas las matrices $\boldsymbol{\Sigma}_{ij}$ ($i \neq j$) sean cuadradas.

Las combinaciones lineales que se trabajarán serán de la forma: $\mathbf{u} = \mathbf{A}\mathbf{x}$, $\mathbf{v} = \mathbf{B}\mathbf{y}$ y $\mathbf{t} = \mathbf{C}\mathbf{z}$. Las ternas del tipo (U_i, V_i, T_i) representarán la i -ésima terna de variables canónicas para cada grupo de variables originales y maximizarán la correlación entre todos las ternas posibles no correlacionadas con las primeras $(i-1)$ ternas de variables canónicas (Timm, 1975). Gnanadesikan desarrolla el tema ampliamente, ofreciendo ejemplos y detalles que exceden el alcance de libros clásicos de análisis multivariado (Seber, 1984).

3. APLICACIÓN A DATOS REALES

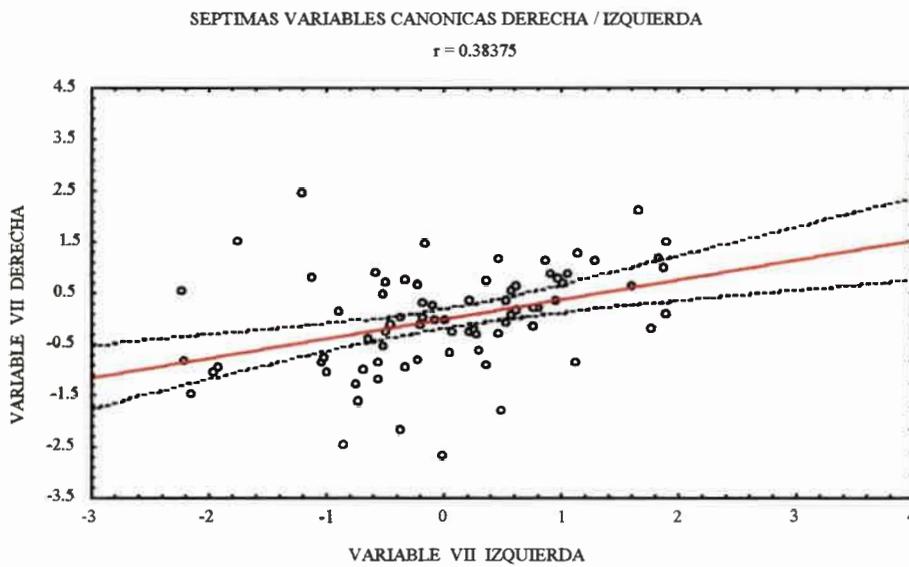
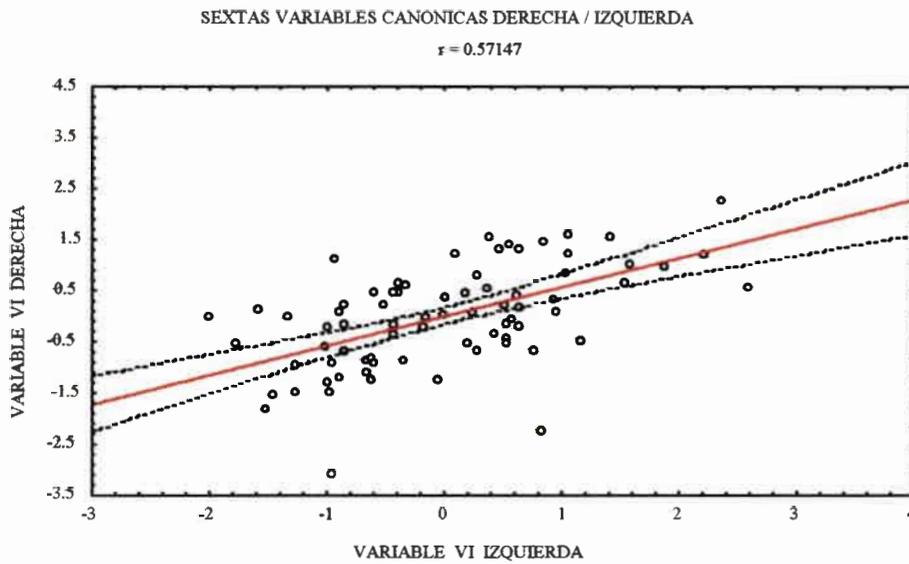
Si bien existe en la bibliografía consultada abundante material de ejercitación de nivel teórico (Seber, 1984; Johnson y Wichern, 1992), y a nivel de aplicación puede trabajarse con datos presentados en ejemplos y/o ejercicios propuestos (Johnson y Wichern, 1992; Timm, 1975; Manly, 1986; SAS, Guía del Usuario Vol. 1, 1990; etc.), creo oportuno presentar un caso de aplicación con datos reales a partir de lo cual es posible ofrecer, además de las conclusiones que permite el análisis estadístico, la interpretación del fenómeno biológico concreto.

3.1. Evaluación de la componente ambiental de la variación fenotípica en poblaciones prehistóricas

Durante 1995 se llevó a cabo un estudio sobre la población prehistórica de San Pedro de Atacama (norte de Chile) con el objetivo de lograr una estimación de la componente genética asociada a variables métricas, de gran importancia en estudios de biología evolutiva en poblaciones antiguas. Se encuentra en prensa en la revista HOMO (Stuttgart, Alemania) el artículo titulado "Evaluation of the Environmental Component of the Phenotypic Variance in Prehistoric Populations", con la autoría de los Dres. H.H. Varela y J.A. Cocilovo, bajo cuya dirección integro un grupo de investigación. Agradezco a ambos autores la gentileza de facilitarme los datos y resultados del proceso estadístico realizado para presentar en esta monografía.

En el Apéndice (3.1 y 3.2), se ofrece, bajo el título "Elementos de Genética Cuantitativa", un glosario de conceptos básicos de Genética, y se exponen brevemente los fundamentos del análisis de la varianza fenotípica para caracteres métricos respectivamente. Además se presenta cómo el aislar la componente ambiental en la misma permite estimar el límite superior de la heredabilidad o repetibilidad, es decir, permite obtener una estimación de la componente genética asociada a variables morfométricas.

Cuando se utilizan caracteres cuantitativos para efectuar estudios sobre la estructura de poblaciones antiguas, el investigador se enfrenta a la imposibilidad de conocer cuánto de lo que se mide es genético y cuánto es ambiental. Por ejemplo, *"...el cálculo de valores D^2 de Mahalanobis siempre va acompañado de sospechas acerca de la verdadera naturaleza de la diferencia que refleja. Aunque es posible hacer inferencias evolutivas a partir de correlaciones y de varianzas fenotípicas, debido a la alta correlación comprobada entre matrices de varianzas y covarianzas genéticas y fenotípicas (Cheverud, 1988), es recomendable que dichas inferencias se realicen después de la estimación de la proporción de varianza ambiental representada en la varianza fenotípica de cada población. La estimación de parámetros asociados con modelos genético-poblacionales que registran las propiedades de las poblaciones y del proceso microevolutivo en una determinada región depende del conocimiento de la variación genética existente en el sistema..."* (Varela y Cocilovo, ep).



3.1.1. Profundización del Análisis

Si bien suele recurrirse a los cuadrados de las correlaciones canónicas, es decir, a los autovalores de la matriz $\Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12}$ para indicar la proporción de varianza de las variables originales explicada por las variables canónicas; según Stewart y Love (en Timm, 1975) ellos tienden a indicar un mayor grado de solapamiento entre ambos grupos de variables, y prefieren el uso de los llamados *Indíces de Redundancia* los cuales presentan una mayor riqueza en su interpretación, pues permiten averiguar cuál grupo de variables es más redundante que el otro.

Sus expresiones generales son:

$$V_{x|V_i}^2 = U_x^2 \hat{\lambda}_i = U_x^2 \hat{\rho}_i^2 \quad y \quad U_{y|U_i}^2 = V_y^2 \hat{\lambda}_i = V_y^2 \hat{\rho}_i^2 \quad i=1, \dots, p$$

con U_x^2 y V_y^2 promedios de los cuadrados de las correlaciones expuestas en las tablas 3A y 3B para cada variable canónica.

Ellos informan del porcentaje de variación en x que puede explicarse por medio de la i -ésima combinación lineal de las variables de y , y viceversa. Si hay varias correlaciones canónicas significativas (como en este caso) se suman los índices de redundancia y se obtiene un índice general:

$$V_{x|V_1, \dots, V_k}^2 = \sum_{i=1}^k V_{x|V_i}^2 \quad \text{y} \quad U_{y|U_1, \dots, U_k}^2 = \sum_{i=1}^k U_{y|U_i}^2$$

Los valores correspondientes para esta ocasión figuran en la próxima tabla:

Tabla 4: Análisis de Redundancia Canónico

	$\hat{\lambda}_i = \hat{\rho}_i^2$	U_x^2	$V_{x V_i}^2$	V_x^2	$U_{y U_i}^2$
I	0.8626	0.38035	0.32809	0.33677	0.29050
II	0.8254	0.19998	0.16506	0.18997	0.15680
III	0.7575	0.12316	0.09329	0.15189	0.11505
IV	0.4793	0.09724	0.04661	0.11305	0.05419
V	0.4007	0.06010	0.02408	0.06131	0.02457
VI	0.3266	0.07738	0.02527	0.08220	0.02685
VII	0.1473	0.06177	0.00910	0.06479	0.00954
Índice General			0.6915		0.6775

Observando los resultados de esta tabla podemos concluir que:

- Las tres primeras variables canónicas (U_1 , U_2 y U_3) resumen el 70% de la variación en el lado izquierdo del cráneo, así como casi el 68% correspondiente a la variación en el lado derecho es informada por V_1 , V_2 y V_3 , según indican los valores de las tercera y quinta columnas.
- Los valores $V_{x|V_i}^2$ y $U_{y|U_i}^2$ son ligeramente menores para las tres primeras variables canónicas (informan respectivamente casi el 59% y el 56% de la variación de las variables del lado opuesto) y, globalmente, dan cuenta aproximadamente del 70% (valores en negrita para ambos índices), comprobando la efectividad del análisis realizado.

Al utilizar el programa SAS, solicitando la opción REDUNDANCY en el PROC CANCORR se obtiene (en *itálica* se destacan los valores comunes a la tabla 4):

Tabla 5. Resultados del Análisis de Redundancia Canónico obtenidos por SAS

Canonical Redundancy Analysis
Standardized Variance of the 'Variables de la Izquierda'
Explained by

	Their Own Canonical Variables			The Opposite Canonical Variables	
	Proportion	Cumulative Proportion	Canonical R-Squared	Proportion	Cumulative Proportion
1	0.3804	0.3804	0.8626	0.3281	0.3281
2	0.2000	0.5804	0.8254	0.1651	0.4932
3	0.1232	0.7035	0.7575	0.0933	0.5864
4	0.0972	0.8007	0.4793	0.0466	0.6330
5	0.0601	0.8609	0.4007	0.0241	0.6571
6	0.0774	0.9382	0.3266	0.0253	0.6824
7	0.0618	1.0000	0.1473	0.0091	0.6915

Canonical Redundancy Analysis
Standardized Variance of the 'Variables de la Derecha'
Explained by

	Their Own Canonical Variables			The Opposite Canonical Variables	
	Proportion	Cumulative Proportion	Canonical R-Squared	Proportion	Cumulative Proportion
1	0.3368	0.3368	0.8626	0.2905	0.2905
2	0.1900	0.5267	0.8254	0.1568	0.4473
3	0.1519	0.6786	0.7575	0.1151	0.5623
4	0.1131	0.7917	0.4793	0.0542	0.6165
5	0.0613	0.8530	0.4007	0.0246	0.6411
6	0.0822	0.9352	0.3266	0.0268	0.6679
7	0.0648	1.0000	0.1473	0.0095	0.6775

Canonical Redundancy Analysis
Squared Multiple Correlations Between the 'Variables de la Izquierda'
and the First 'M' Canonical Variables of the 'Variables de la Derecha'

	I	II	III	IV	V	VI	VII
<i>altura del pómulo</i>	0.1589	0.5685	0.7695	0.7806	0.7819	0.7821	0.7861
<i>ancho de la órbita</i>	0.6250	0.6917	0.7001	0.7062	0.7292	0.7605	0.7631
<i>altura de la órbita</i>	0.2286	0.6187	0.7445	0.7701	0.7828	0.7852	0.7858
<i>altura órbita-alveolar</i>	0.1688	0.2610	0.4278	0.4396	0.4545	0.4752	0.5263
<i>curva horizontal</i>	0.4022	0.4543	0.4688	0.4718	0.4971	0.6112	0.6160
<i>cuerda porion-basion</i>	0.4856	0.5356	0.6647	0.6675	0.7457	0.7465	0.7469
<i>cuerda porion-bregma</i>	0.2276	0.3222	0.3298	0.5956	0.6086	0.6161	0.6163

Canonical Redundancy Analysis
Squared Multiple Correlations Between the 'Variables de la Derecha'
and the First 'M' Canonical Variables of the 'Variables de la Izquierda'

	I	II	III	IV	V	VI	VII
<i>altura del pómulo</i>	0.1649	0.5030	0.7620	0.7825	0.7831	0.7831	0.7850
<i>ancho de la órbita</i>	0.4735	0.5998	0.6143	0.6199	0.7000	0.7214	0.7217
<i>altura de la órbita</i>	0.2283	0.6733	0.7506	0.7595	0.7800	0.7868	0.7873
<i>altura órbita-alveolar</i>	0.1379	0.2031	0.3993	0.4193	0.4249	0.4384	0.4981
<i>curva horizontal</i>	0.2980	0.3271	0.4123	0.4145	0.4341	0.5790	0.5804
<i>cuerda porion-basion</i>	0.5083	0.5765	0.7450	0.7474	0.7870	0.7872	0.7874
<i>cuerda porion-bregma</i>	0.2226	0.2483	0.2528	0.5725	0.5787	0.5797	0.5825

De los dos últimos cuadros (cuadrados de coeficientes de correlación múltiple “cruzados”) se deduce que todas las variables canónicas son buenas predictores de las variables originales (los *dos* valores más pequeños son mayores a 0.10 y se observan en la primera variable canónica para la altura del pómulo y la altura órbito-alveolar).

En síntesis, tanto desde el punto de vista del análisis univariado cuanto del análisis multivariado, pueden obtenerse estimaciones de la componente ambiental especial y cotas superiores para la heredabilidad. El análisis minucioso de los resultados obtenidos para cada una de las siete variables realizado por los autores encuentra en los hábitos alimentarios y en la práctica cultural de la deformación artificial del cráneo (según su incidencia particular en cada individuo) las fuentes atribuibles a la variación ambiental especial: en el caso de la altura órbito alveolar debido a la pérdida de piezas dentarias en edades tempranas, y en el caso de la deformación por los efectos localizados de la técnica o del aparato deformatario aplicado durante el desarrollo del individuo, tal como evidencian la curva horizontal y la cuerda porion-bregma.

En el caso del Análisis de Correlación Canónica, como todas las correlaciones obtenidas son estadísticamente significativas, revisten importancia para representar la asociación existente entre las variables de lateralidad. El elevado índice de redundancia obtenido muestra que, aunque en menor medida, la información provista por un grupo de las variables (por ejemplo, las del lado izquierdo) resume la información que brindan las variables tomadas en el lado opuesto, y es en esa “pequeña falla” de predicción que encontramos el efecto debido a la variación ambiental especial.

4. BIBLIOGRAFÍA CITADA

- Cavalli-Sforza, L.L. y W.F. Bodmer (1981) *Genética de las Poblaciones Humanas*. Ed. Omega, Barcelona, España.
- Comas, J. (1966) *Manual de Antropología Física*. Universidad Nacional Autónoma de México, Instituto de Investigaciones Históricas, Sección de Antropología, México.
- Cooley, W.W and P.R. Lohnes (1962) *Multivariate Procedures for the Behavioral Sciences*. Ed. John Wiley & Sons. New York, USA.
- Droessler, J. (1981) *Craneometry and biological distance. Biocultural continuity and change at the late-Woodland-Mississippian interface*. Center for American Archeology, At Northwestern University, Evanston, Illinois, USA.
- Falconer, D.S. (1990) *Introducción a la Genética Cuantitativa*, 3° edición, CESCO, México.
- Hotelling, H. (1935) The Most Predictable Criterion. *Journal of Educational Psychology*, 26:139-142.
- Hotelling, H. (1936) Relations between two sets of variates. *Biometrika*, 28:321-327.
- Howells, W.W. (1973) Cranial variation in man: a study by multivariate analysis of patterns of difference among recent human population *Pap. Peabody Mus. Am. Archaeol. Ethnol.* 67.
- Johnson, R. y D. Wichern (1992) *Applied Multivariate Statistical Analysis*. 3° Edición. Ed. Prentice Hall. New Jersey, USA.

- Kendall, M.G. (1972) *A Course in Multivariate Analysis*. Ed. Charles Griffin and Co. London.
- Kendall, M.G. (1975) *Multivariate Analysis*. Ed. Charles Griffin and Co.
- Manly, B.F.J. (1986) *Multivariate Statistical Methods. A Primer*. Ed. Champan and Hall. London, New York, USA.
- Rothhammer, F. y R. Cruz-Coke (1977) *Curso Básico de Genética Humana*. Ed. Universitaria, Santiago, Chile.
- SAS/STAT (1990) *User Guide, Version 6, Fourth Edition*. Vol. 1. Ed. SAS Institute. NC, USA.
- Seber, G.A. (1984) *Multivariate Observations*. Ed. John Wiley & Sons. New York, USA.
- Timm, N.H. (1975) *Multivariate Analysis with Applications in Education and Psychology*. Ed. Wadsworth Publishing Company. California, USA.
- Varela, H.H. and J.A. Cocilovo (ep) Evaluation of the Environmental Component of the Phenotypic Variance in Prehistoric Populations. *HOMO*. Gustav Fisher, Stuttgart, Alemania.

5. AGRADECIMIENTOS

Durante mis años de estudio universitario y posterior desempeño profesional, siempre me he sentido apoyada por personas que ofrecieron desinteresadamente su amistad y colaboración. Por ello, no puedo menos que expresar mi sincero agradecimiento a varias de ellas: al Dr. Oscar Bustos quien me honró aceptando ser mi Director en este trabajo monográfico, brindándome desde la época en que fui alumna de "Probabilidades y Estadística" su gran experiencia y constante dedicación; al Dr. José A. Cocilovo, quien es y ha sido mi Director en varios proyectos de investigación, y a lo largo de estos años me ayudó a conocer e interpretar la complejidad de la investigación bioantropológica.

A mis compañeras/o de estudio y trabajo: Susana Ferrero, Gabriela Palacio, Mery Picco y Hugo Varela. A ellas por compartir todas las vivencias del cursado de la Especialidad y a él por 'soportar' nuestras discusiones de estudio en el mismo ámbito de trabajo, además de la colaboración recibida en todo momento y la generosa entrega del material, junto a Tito Cocilovo, analizado en esta monografía.

A mi esposo y a mi hija, quienes sufrieron las consecuencias de mi dedicación al estudio y al trabajo (ya desde épocas del Profesorado), pero siempre me brindaron su amor y apoyo incondicional.

Finalmente, debo mi reconocimiento a todos y cada uno de los profesores/as que me guiaron en el camino del conocimiento, pues de otro modo, este trabajo no habría salido a la luz. Gracias por todo.


13-11-98

APENDICE

1. CONCEPTOS DE ALGEBRA MATRICIAL

1.1. Matrices Definidas Positivas y Formas Cuadráticas Asociadas

Cuando para una matriz simétrica \mathbf{B} de dimensión $(p \times p)$ se verifica que $\mathbf{x}'\mathbf{B}\mathbf{x} > 0$ para todo $\mathbf{x}' = [x_1, x_2, \dots, x_p] \neq \mathbf{0}$, se la denomina *matriz definida positiva*.

Como $\mathbf{x}'\mathbf{B}\mathbf{x}$ sólo contiene términos cuadráticos (x_i^2) y términos con productos ($x_i x_j$), recibe el nombre de *forma cuadrática*.

Nota: La matriz de varianzas-covarianzas Σ de un vector aleatorio \mathbf{x} es, pues, una matriz definida positiva.

1.2. Descomposición Espectral

La descomposición espectral de una matriz simétrica \mathbf{B} de dimensión $(p \times p)$ está dada por:

$$\mathbf{B} = \lambda_1 \mathbf{e}_1 \mathbf{e}_1' + \lambda_2 \mathbf{e}_2 \mathbf{e}_2' + \dots + \lambda_p \mathbf{e}_p \mathbf{e}_p' = \sum_{i=1}^p \lambda_i \mathbf{e}_i \mathbf{e}_i'$$

con λ_i autovalores (o valores propios) de \mathbf{B} y \mathbf{e}_i autovectores (o vectores propios) normalizados asociados a \mathbf{B} , para $i = 1, \dots, p$.

1.3. Matriz Raíz Cuadrada

La descomposición espectral permite que se exprese la inversa de una matriz cuadrada en términos de sus valores y vectores propios pues si llamamos Λ a la matriz diagonal cuyos elementos no nulos son los autovalores λ_i para $i = 1, \dots, p$, y llamamos \mathbf{P} a $[\mathbf{e}_1 \ \mathbf{e}_2 \ \dots \ \mathbf{e}_p]$, tendremos:

$$\mathbf{B} = \sum_{i=1}^p \lambda_i \mathbf{e}_i \mathbf{e}_i' = \mathbf{P}\Lambda\mathbf{P}' \quad \text{con } \mathbf{P}\mathbf{P}' = \mathbf{I} \quad \text{y entonces:}$$

$$\mathbf{B}^{-1} = \mathbf{P}\Lambda^{-1}\mathbf{P}' = \sum_{i=1}^p \frac{1}{\lambda_i} \mathbf{e}_i \mathbf{e}_i' \quad \text{dado que } (\mathbf{P}\Lambda^{-1}\mathbf{P}')\mathbf{P}\Lambda\mathbf{P}' = \mathbf{P}\Lambda\mathbf{P}'(\mathbf{P}\Lambda^{-1}\mathbf{P}') = \mathbf{P}\mathbf{P}' = \mathbf{I}$$

Con $\Lambda^{1/2}$ denotamos la matriz diagonal con $\sqrt{\lambda_i}$ como i -ésimo elemento diagonal. La matriz $\sum_{i=1}^p \sqrt{\lambda_i} \mathbf{e}_i \mathbf{e}_i' = \mathbf{P}\Lambda^{1/2}\mathbf{P}'$ es conocida entonces como *matriz raíz cuadrada de \mathbf{B}* y se denota $\mathbf{B}^{1/2}$, y goza de las siguientes propiedades:

- Es simétrica, es decir: $(\mathbf{B}^{1/2})' = \mathbf{B}^{1/2}$
- $\mathbf{B}^{1/2} \mathbf{B}^{1/2} = \mathbf{B}$
- $(\mathbf{B}^{1/2})^{-1} = \mathbf{P}\Lambda^{-1/2}\mathbf{P}'$, donde $\Lambda^{-1/2}$ es matriz diagonal con $1/\sqrt{\lambda_i}$ como i -ésimo elemento diagonal
- $\mathbf{B}^{1/2} \mathbf{B}^{-1/2} = \mathbf{B}^{-1/2} \mathbf{B}^{1/2} = \mathbf{I}$ y $\mathbf{B}^{-1/2} \mathbf{B}^{-1/2} = \mathbf{B}^{-1}$, donde $\mathbf{B}^{-1/2} = (\mathbf{B}^{1/2})^{-1}$

1.4. Desigualdad de Cauchy-Schwarz

Sean \mathbf{x} y \mathbf{y} dos vectores de dimensión $(p \times 1)$, entonces: $(\mathbf{x}'\mathbf{y})^2 \leq (\mathbf{x}'\mathbf{x})(\mathbf{y}'\mathbf{y})$, valiendo la igualdad si y solo si $\mathbf{x} = c\mathbf{y}$ (o $\mathbf{y} = c\mathbf{x}$) para alguna constante c .

Demostración:

La desigualdad es obvia para el caso en que alguno de los vectores (o ambos) es nulo. Excluyendo esta posibilidad, consideremos el vector $\mathbf{x} - c\mathbf{y}$, con c un escalar arbitrario.

Si $\mathbf{x} - c\mathbf{y} \neq \mathbf{0}$, $(\mathbf{x} - c\mathbf{y})'(\mathbf{x} - c\mathbf{y}) > 0$, es decir:

$0 < (\mathbf{x} - c\mathbf{y})'(\mathbf{x} - c\mathbf{y}) = (\mathbf{x}' - c\mathbf{y}')(\mathbf{x} - c\mathbf{y}) = \mathbf{x}'\mathbf{x} - c\mathbf{x}'\mathbf{y} - c\mathbf{y}'\mathbf{x} + c^2\mathbf{y}'\mathbf{y} = \mathbf{x}'\mathbf{x} - 2c\mathbf{x}'\mathbf{y} + c^2\mathbf{y}'\mathbf{y}$ que resulta ser una inecuación cuadrática en c .

Así, completando cuadrados:

$$\begin{aligned} \mathbf{x}'\mathbf{x} - 2c\mathbf{x}'\mathbf{y} + c^2\mathbf{y}'\mathbf{y} &= \mathbf{x}'\mathbf{x} - (\mathbf{x}'\mathbf{y})^2/(\mathbf{y}'\mathbf{y}) + (\mathbf{x}'\mathbf{y})^2/(\mathbf{y}'\mathbf{y}) - 2c\mathbf{x}'\mathbf{y} + c^2\mathbf{y}'\mathbf{y} \\ &= \mathbf{x}'\mathbf{x} - (\mathbf{x}'\mathbf{y})^2/(\mathbf{y}'\mathbf{y}) + \mathbf{y}'\mathbf{y} [(\mathbf{x}'\mathbf{y})^2/(\mathbf{y}'\mathbf{y})^2 - 2(c\mathbf{x}'\mathbf{y})/(\mathbf{y}'\mathbf{y}) + c^2] \\ &= \mathbf{x}'\mathbf{x} - (\mathbf{x}'\mathbf{y})^2/(\mathbf{y}'\mathbf{y}) + \mathbf{y}'\mathbf{y} [c - (\mathbf{x}'\mathbf{y})/(\mathbf{y}'\mathbf{y})]^2 \end{aligned}$$

Como para $c = (\mathbf{x}'\mathbf{y})/(\mathbf{y}'\mathbf{y})$ el último término es nulo, resulta:

$$0 < \mathbf{x}'\mathbf{x} - (\mathbf{x}'\mathbf{y})^2/(\mathbf{y}'\mathbf{y}) \Rightarrow \mathbf{x}'\mathbf{x} > (\mathbf{x}'\mathbf{y})^2/(\mathbf{y}'\mathbf{y}) \Rightarrow \boxed{(\mathbf{x}'\mathbf{y})^2 < (\mathbf{x}'\mathbf{x})(\mathbf{y}'\mathbf{y})} \text{ si } \mathbf{x} \neq c\mathbf{y} \text{ para algún } c.$$

Si $\mathbf{x} = c\mathbf{y}$ para algún c , $0 = (\mathbf{x} - c\mathbf{y})'(\mathbf{x} - c\mathbf{y})$ y con igual argumentación se obtiene

$$\boxed{(\mathbf{x}'\mathbf{y})^2 = (\mathbf{x}'\mathbf{x})(\mathbf{y}'\mathbf{y})}$$

1.4.1. Extensión de la Desigualdad de Cauchy-Schwarz a matrices definidas positivas

Sean \mathbf{x} e \mathbf{y} dos vectores de dimensión $(p \times 1)$ y sea \mathbf{B} una matriz definida positiva de dimensión $(p \times p)$, entonces: $(\mathbf{x}'\mathbf{y})^2 \leq (\mathbf{x}'\mathbf{B}\mathbf{x})(\mathbf{y}'\mathbf{B}^{-1}\mathbf{y})$, valiendo la igualdad si y solo si $\mathbf{x} = c\mathbf{B}^{-1}\mathbf{y}$ (o $\mathbf{y} = c\mathbf{B}\mathbf{x}$) para alguna constante c .

Demostración:

La desigualdad es obvia cuando $\mathbf{x} = \mathbf{0}$ o $\mathbf{y} = \mathbf{0}$.

A partir de $\mathbf{B}^{1/2}$ y $\mathbf{B}^{-1/2}$ se obtiene que:

$$\mathbf{x}'\mathbf{y} = \mathbf{x}'\mathbf{I}\mathbf{y} = \mathbf{x}'\mathbf{B}^{1/2}\mathbf{B}^{-1/2}\mathbf{y} = (\mathbf{B}^{1/2}\mathbf{x})'(\mathbf{B}^{-1/2}\mathbf{y}) = \mathbf{z}'\mathbf{w}$$

Por lo tanto:

$$(\mathbf{x}'\mathbf{y})^2 = (\mathbf{z}'\mathbf{w})^2 \leq (\mathbf{z}'\mathbf{z})(\mathbf{w}'\mathbf{w}) = (\mathbf{x}'\mathbf{B}^{1/2}\mathbf{B}^{1/2}\mathbf{x})(\mathbf{y}'\mathbf{B}^{-1/2}\mathbf{B}^{-1/2}\mathbf{y}) = \mathbf{x}'\mathbf{B}\mathbf{x} \mathbf{y}'\mathbf{B}^{-1}\mathbf{y}$$

↑
Des.C-S

Corolario: $\frac{(\mathbf{x}'\mathbf{y})^2}{\mathbf{x}'\mathbf{B}\mathbf{x}} \leq \mathbf{y}'\mathbf{B}^{-1}\mathbf{y}$

1.5. Principios de Maximización

P1. Para una matriz definida positiva \mathbf{B} y un vector \mathbf{y} dados, la cota superior $\mathbf{y}'\mathbf{B}^{-1}\mathbf{y}$ que indica el corolario anterior se alcanza cuando $\boxed{\mathbf{x} = c\mathbf{B}^{-1}\mathbf{y}}$ ($c \neq 0$):

$$\frac{(\mathbf{x}'\mathbf{y})^2}{\mathbf{x}'\mathbf{B}\mathbf{x}} = \frac{(c\mathbf{y}'\mathbf{B}^{-1}\mathbf{y})^2}{c^2\mathbf{y}'\mathbf{B}^{-1}\mathbf{B}\mathbf{B}^{-1}\mathbf{y}} = \frac{c^2(\mathbf{y}'\mathbf{B}^{-1}\mathbf{y})^2}{c^2\mathbf{y}'\mathbf{B}^{-1}\mathbf{y}} = \mathbf{y}'\mathbf{B}^{-1}\mathbf{y}$$

P2. ¿Cuándo alcanzan su máximo expresiones de la forma $\frac{\mathbf{x}'\mathbf{B}\mathbf{x}}{\mathbf{x}'\mathbf{x}}$? Veamos que se cumplen:

$$\text{a) } \max_{\mathbf{x} \neq \mathbf{0}} \frac{\mathbf{x}'\mathbf{B}\mathbf{x}}{\mathbf{x}'\mathbf{x}} = \lambda_1, \quad \text{b) } \min_{\mathbf{x} \neq \mathbf{0}} \frac{\mathbf{x}'\mathbf{B}\mathbf{x}}{\mathbf{x}'\mathbf{x}} = \lambda_p \quad \text{y} \quad \text{c) } \max_{\mathbf{x} \perp \mathbf{e}_1, \dots, \mathbf{e}_k} \frac{\mathbf{x}'\mathbf{B}\mathbf{x}}{\mathbf{x}'\mathbf{x}} = \lambda_{k+1} \quad k=1, \dots, p-1$$

Demostración:

a) Por la descomposición espectral (conocida también como Teorema de Diagonalización) $\mathbf{B} = \mathbf{P} \mathbf{D} \mathbf{P}'$, con \mathbf{P} matriz de autovectores y \mathbf{D} matriz diagonal de autovalores λ_i . Sea además $\mathbf{y} = \mathbf{P}'\mathbf{x}$.

$\mathbf{x} \neq \mathbf{0}$ implica que $\mathbf{y} \neq \mathbf{0}$, entonces:

$$\frac{\mathbf{x}'\mathbf{B}\mathbf{x}}{\mathbf{x}'\mathbf{x}} = \frac{\mathbf{x}'\mathbf{P}\mathbf{D}\mathbf{P}'\mathbf{x}}{\mathbf{x}'\mathbf{P}\mathbf{P}'\mathbf{x}} = \frac{(\mathbf{P}'\mathbf{x})'\mathbf{D}(\mathbf{P}'\mathbf{x})}{(\mathbf{P}'\mathbf{x})'(\mathbf{P}'\mathbf{x})} = \frac{\mathbf{y}'\mathbf{D}\mathbf{y}}{\mathbf{y}'\mathbf{y}} = \frac{\sum_{i=1}^p \lambda_i y_i^2}{\sum_{i=1}^p y_i^2} \leq \frac{\sum_{i=1}^p \lambda_1 y_i^2}{\sum_{i=1}^p y_i^2} = \lambda_1$$

Por lo tanto $\frac{\mathbf{x}'\mathbf{B}\mathbf{x}}{\mathbf{x}'\mathbf{x}} \leq \lambda_1$, siendo entonces el primer valor propio cota superior (se lo llama máximo incondicional)

¿Cuándo se da la igualdad? Pues, cuando $\mathbf{x} = \mathbf{e}_1$ de \mathbf{B} (primer vector propio) ya que

$$\frac{\mathbf{e}_1'\mathbf{B}\mathbf{e}_1}{\mathbf{e}_1'\mathbf{e}_1} = \mathbf{e}_1'\mathbf{P}\mathbf{D}\mathbf{P}'\mathbf{e}_1 = \begin{bmatrix} \sum \mathbf{e}_{ii}^2 & \dots & 0 \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_p \end{bmatrix} \cdot \begin{bmatrix} \sum \mathbf{e}_{ii}^2 \\ \vdots \\ 0 \end{bmatrix} = \lambda_1 \quad (\text{pues } \mathbf{e}_1'\mathbf{e}_1 = 1)$$

b) se obtiene el resultado con un argumento similar al de a).

c) Sea ahora $\mathbf{x} = \mathbf{P}\mathbf{y} = y_1 \mathbf{e}_1 + y_2 \mathbf{e}_2 + \dots + y_p \mathbf{e}_p$, de modo que $\mathbf{x} \perp \mathbf{e}_1, \dots, \mathbf{e}_k$ implica que $0 = \mathbf{e}_i'\mathbf{x} = y_1 \mathbf{e}_i'\mathbf{e}_1 + y_2 \mathbf{e}_i'\mathbf{e}_2 + \dots + y_p \mathbf{e}_i'\mathbf{e}_p = y_i, \quad i \leq k$

Es decir, que para \mathbf{x} perpendicular a los primeros k vectores propios \mathbf{e}_i , $\frac{\mathbf{x}'\mathbf{B}\mathbf{x}}{\mathbf{x}'\mathbf{x}}$ da lugar a la siguiente expresión:

$$\frac{\mathbf{x}'\mathbf{B}\mathbf{x}}{\mathbf{x}'\mathbf{x}} = \frac{\sum_{i=k+1}^p \lambda_i y_i^2}{\sum_{i=k+1}^p y_i^2}$$

y tomando $y_{k+1} = 1, y_{k+2} = \dots = y_p = 0$ se llega al resultado.

Para cualquier $\mathbf{x}_0 \neq \mathbf{0}$, puede construirse $\mathbf{x}' = \frac{\mathbf{x}'_0}{\sqrt{\mathbf{x}'_0 \mathbf{x}_0}}$ (de tamaño unitario) de modo

que $\frac{\mathbf{x}'_0 \mathbf{B} \mathbf{x}_0}{\mathbf{x}'_0 \mathbf{x}_0}$ tiene el mismo valor que $\frac{\mathbf{x}' \mathbf{B} \mathbf{x}}{\mathbf{x}' \mathbf{x}}$. Podemos entonces decir que los valores

propios λ_1 y λ_p son, respectivamente, el máximo y el menor valor de la forma cuadrática $\mathbf{x}'\mathbf{B}\mathbf{x}$ para todos los puntos \mathbf{x} cuya distancia al origen es la unidad. Representan los valores extremos de $\mathbf{x}'\mathbf{B}\mathbf{x}$ para los puntos de la esfera unidad. Los autovalores intermedios (λ_2 a λ_{p-1}) también se interpretan como valores extremos cuando \mathbf{x} está restringido a ser ortogonal a los anteriores (máximos condicionales).

1.5.1. Nota Ampliatoria:

También se verifica para el cociente de formas cuadráticas de la forma $\frac{\mathbf{x}'\mathbf{B}\mathbf{x}}{\mathbf{x}'\mathbf{A}\mathbf{x}}$, con \mathbf{A} matriz de dimensión $p \times p$, que:

$$\text{a) } \max_{\mathbf{x} \neq 0} \frac{\mathbf{x}'\mathbf{B}\mathbf{x}}{\mathbf{x}'\mathbf{A}\mathbf{x}} = \gamma_1 \quad \text{y} \quad \text{b) } \min_{\mathbf{x} \neq 0} \frac{\mathbf{x}'\mathbf{B}\mathbf{x}}{\mathbf{x}'\mathbf{A}\mathbf{x}} = \gamma_p, \text{ con } \gamma_1, \dots, \gamma_p \text{ autovalores de } \mathbf{A}^{-1}\mathbf{B},$$

alcanzándose el máximo cuando $\mathbf{x} = \mathbf{v}_1$ y el mínimo cuando $\mathbf{x} = \mathbf{v}_p$ respectivamente, siendo \mathbf{v}_i el i -ésimo autovector normalizado de $\mathbf{A}^{-1}\mathbf{B}$.

2. UN POCO DE HISTORIA ...

En la Introducción se presentó escuetamente el origen del Análisis de Correlación Canónica. La investigación llevada a cabo por Hotelling en 1935 con un grupo de 140 alumnos del séptimo grado es desarrollada con más detalle a continuación:

Se tomó a los niños una batería de cuatro tests para investigar la performance de los mismos en cuatro habilidades escolares consideradas básicas para un aprendizaje efectivo: la velocidad y capacidad de lectura, y la velocidad y capacidad de cálculo aritmético.

Sean las variables X_1 : velocidad de lectura y X_2 : capacidad de lectura, como así también Y_1 : velocidad de cálculo aritmético e Y_2 : capacidad de cálculo aritmético; luego: $\mathbf{x} = [X_1, X_2]$ e $\mathbf{y} = [Y_1, Y_2]$.

La correlación entre los dos grupos de variables está dada por la siguiente matriz:

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{R}_{21} & \mathbf{R}_{22} \end{bmatrix} = \begin{bmatrix} 1.00 & 0.6328 & 0.2412 & 0.0586 \\ 0.6328 & 1.00 & -0.0553 & 0.0655 \\ 0.2412 & -0.0553 & 1.00 & 0.4248 \\ 0.0586 & 0.0655 & 0.4248 & 1.00 \end{bmatrix}$$

La matriz \mathbf{R} es una matriz cuadrada, simétrica, de orden 4×4 . Si adoptamos para trabajar la matriz $\mathbf{F} = \mathbf{R}_{22}^{-1} \mathbf{R}_{21} \mathbf{R}_{11}^{-1} \mathbf{R}_{12}$, al resolver la ecuación canónica

$$(\mathbf{R}_{22}^{-1} \mathbf{R}_{21} \mathbf{R}_{11}^{-1} \mathbf{R}_{12} - \lambda_i \mathbf{I}) \mathbf{f}_i = 0$$

para $i=1,2$, obtendremos los pares $(\lambda_i, \mathbf{f}_i)$ autovalores y autovectores de \mathbf{F} (2.1., pág.3). El determinante correspondiente es:

$$\begin{vmatrix} 0.156759 - \lambda_i & 2.74291\text{E-}3 \\ -0.0623125 & 3.61571\text{E-}3 - \lambda_i \end{vmatrix} = 0$$

y los valores finales para los vectores propios son:

$$\hat{\mathbf{e}}'_1 = [-2.7772 \quad 2.2655] \quad \text{y} \quad \hat{\mathbf{f}}_1 = [-2.4404 \quad 1] \quad \text{con} \quad \hat{\lambda}_1 \approx 0.157867 \quad \text{y} \quad \hat{\lambda}_2 \approx 0.002508$$

es decir, con $\hat{\rho}_1 \approx 0.3973$ y $\hat{\rho}_2 \approx 0.0501$. La segunda correlación canónica es casi nula, por lo que la primera correlación canónica se encarga de representar aproximadamente toda la relación existente entre las habilidades estudiadas.

Como \hat{U} mide la diferencia entre las habilidades de lectura y \hat{V} mide la diferencia entre las habilidades del cálculo aritmético, puede apreciarse que aquellos niños con grandes diferencias entre X_1 y X_2 , también tienden a presentar grandes diferencias entre Y_1 e Y_2 . Lamentablemente, como los datos están disponibles a partir de la matriz de correlación, no es posible realizar ninguno de los gráficos sugeridos en 2.3.1, ni intentar 'validar' con una mitad de los datos (eran ¡140!) los resultados que podrían haberse obtenido con la otra mitad.

Podemos utilizar el test de Bartlett para testar $H_0: \Sigma_{12} = \mathbf{0}$ vs. $H_1: \Sigma_{12} \neq \mathbf{0}$

$$\text{Así: } \phi_0^2 = -\left(n - 1 - \frac{1}{2}(p + q + 1)\right) \text{Log} \prod_{i=1}^p (1 - \hat{\rho}_i^2) = -(136.5) \times (-0.174328) \approx 23.7958,$$

altamente significativo al compararlo con $\chi_{4,0.01}^2 = 13.277$. Si removemos el efecto de la primera correlación canónica, resulta:

$\phi_1^2 = -(136.5) \times (-0.00251082) \approx 0.3427$, obviamente no significativo para $\chi_{1,0.05}^2 = 3.841$, por lo que es a la primera correlación canónica la que debemos considerar representativa de toda la relación que pueda existir entre las habilidades lectoras y de cálculo.

3. ELEMENTOS DE GENÉTICA CUANTITATIVA

3.1. Conceptos Básicos de Genética

Para facilitar la lectura del tema posterior en este apéndice y de 3.1 en el texto principal, se ofrecen, a modo de glosario, algunos conceptos básicos. En vez de utilizar el clásico orden alfabético, los términos aparecen en la secuencia según la cual son necesarios para entender los conceptos aquí presentados.

3.1.1. Glosario

Gen: Unidad de la herencia que está en el cromosoma.

Cromosoma: Uno de los cuerpos del núcleo celular a lo largo del cual están los genes, se visualizan como filamentos. Todo cromosoma que no sea un cromosoma sexual se denomina *autosoma* (El ser humano tiene 22 pares de autosomas y 1 par de cromosomas sexuales -X: femenino, Y: masculino-)

Alelos: Dos o más formas alternativas de un gen. Los alelos ocupan la misma posición (locus, en plural: loci) en cromosomas homólogos y se separan uno de otro cuando ocurre la meiosis. Ejemplo: alelo A y alelo a.

cefalogramas laterales de 72 pares de mellizos, calculó el cociente entre la varianza dentro de mellizos dicigóticos y la varianza dentro de mellizos monocigóticos como una medida de la heredabilidad relativa para un conjunto de caracteres craneofaciales. En este trabajo se comprueba que las alturas tienen una componente genética mayor que las longitudes. Da Rocha et al. (1972) hicieron lo mismo en una muestra de mellizos mono y dicigóticos de Porto Alegre, Brasil. Sus resultados muestran valores de h^2 mayores de 0.6 en el 75% de las 16 variables antropométricas empleadas...”.

Además, destacan que “...El cálculo de la heredabilidad empleando un diseño experimental con mellizos requiere la comparación de los monocigóticos con los dicigóticos. Debido a la dificultad de aislar la varianza correspondiente al ambiente común entre mellizos monocigóticos, esta estimación se realiza bajo el supuesto de que la componente de la varianza ambiental es igual en ambos tipos de mellizos (Falconer, 1990 y Wright, 1969)...”

“...Morton y Green (1972) también estimaron la heredabilidad de 12 caracteres y seis índices antropométricos empleando grupos de hermanos mokilese y pingelapese, obteniendo un h^2 promedio de 0.565. Droessler (1981) realiza el comentario de varios trabajos en los que se estima la heredabilidad de caracteres antropométricos, como el de Mc Henry y Giles (1971) en hermandades de pueblos Waffa de Nueva Guinea, cuyos resultados coinciden con los obtenidos por Osborne y De George (1959) en mellizos di y monocigóticos de una población moderna de América del Norte. Droessler (1981) destaca la importancia de este tipo de investigaciones para los estudios de variación craneométrica entre poblaciones por el aporte de evidencias que demuestran una alta componente genética para muchos caracteres antropométricos en diferentes poblaciones. Si bien esto es real, es preciso advertir también que h^2 es una propiedad del carácter y de la población por lo tanto dicho valor varía entre caracteres y entre poblaciones...” (Varela y Cocilovo, ep).

“...Cuando se trabaja con poblaciones humanas antiguas no es posible calcular h^2 debido a que no se conoce ningún tipo de relación familiar entre los individuos de una muestra dada. Pero sí es factible aislar parte de la varianza ambiental de la varianza fenotípica y estimar el límite superior de la heredabilidad o repetibilidad. Esto se puede lograr mediante mediciones múltiples, es decir, haciendo más de una medición del carácter en cada individuo. Los estudios de repetibilidad han sido muy empleados por su importancia práctica en los programas de mejoramiento vegetal y animal (Falconer 1990, Lush 1965)...”(Varela y Cocilovo, ep).

Por lo tanto, considerando que se realizan múltiples observaciones del carácter en estudio para cada individuo, podemos admitir otra partición de la varianza fenotípica total, cuyas componentes son la varianza entre individuos (between, de allí el subíndice), V_B , y la varianza entre mediciones -de la misma variable- dentro de individuos, V_W (within). Entonces:

$$V_F = V_B + V_W, \quad \text{con} \quad V_B = V_G + V_{Eg} \quad \text{y} \quad V_W = V_{Ee}$$

donde V_{Eg} es la varianza ambiental *general* que afecta a todos los individuos y V_{Ee} es la varianza ambiental *especial* dentro de individuos debida a efectos localizados que actúan durante el desarrollo ontogenético, de modo que la varianza dentro de individuos es completamente ambiental en su origen. Según este modelo de partición de la varianza, se puede separar y medir la V_{Ee} , ya que el cociente:

$$(V_G + V_{Eg}) / V_F = V_B / (V_B + V_W)$$

no es más que la correlación intraclase r . Como es la correlación entre medidas repetidas del mismo individuo, recibe el nombre de *repetibilidad del carácter* y expresa la **proporción de la varianza** de mediciones simples debida a diferencias permanentes o no localizadas entre los individuos, ya sea de naturaleza ambiental o genética (Falconer, 1990).

Luego, por complemento, $1 - r = V_{Ee} / V_F$ será la **proporción de la varianza** debida al ambiente especial. Esto significa que r es límite superior para h^2 sea ésta calculada en sentido amplio o estricto, permitiendo entonces estimarla de manera aproximada cuando no se disponga de datos específicos (condiciones de laboratorio, etc.). Los supuestos implícitos en la idea de repetibilidad son dos: i) las varianzas de las diferentes mediciones son las mismas y sus componentes se presentan en las mismas proporciones; y ii) las distintas mediciones reflejan genéticamente lo que es el carácter.

Se han efectuado diversos experimentos -Morley, 1951, Barker y Robertson, 1966, Reeve y Robertson, 1954, entre otros- que mostraron, dado que la componente ambiental general es bastante pequeña, que la repetibilidad es un buen estimador tanto del grado de determinación genética como de la heredabilidad (Falconer, 1990).

Los restantes elementos teóricos fueron presentados por Varela y Cocilovo (ep), en el acápite “Material y Métodos”, que se transcribe a continuación.

3.3. El Material osteológico y la Metodología propuesta (Transcripción de Varela y Cocilovo, ep)

“...La muestra empleada en el presente trabajo pertenece a los sitios de Quitar 5, Quitar 6 y Quitar Conde Duque Bajo de la población prehistórica de San Pedro de Atacama (Norte de Chile), depositada en el Instituto de Investigaciones arqueológicas R.P. Gustavo Le Paige, S.J. de la misma localidad. Los tres sitios son contiguos espacialmente y corresponden a la fase II (300 dC - 1000 dC) y a la fase III (1000 dC - 1300 dC) del desarrollo biocultural establecido para esta población...”

“...La muestra está constituida por 86 cráneos deformados artificialmente (tabular erecto y oblicuo; circular erecto y oblicuo) y no deformados, de ambos sexos (51 masculinos y 35 femeninos) y de diferentes edades (adultos, maduros y seniles)...” (Figura 1). Se midieron las mismas 7 variables del lado izquierdo y derecho de cada cráneo (Apéndice, 3.5), es decir 14 mediciones por cada uno de ellos. Las observaciones fueron realizadas por nosotros

(H.H.V. y J.A.C.) en proporciones iguales de ejemplares, en el mismo momento y con un estricto control del error interobservador...”

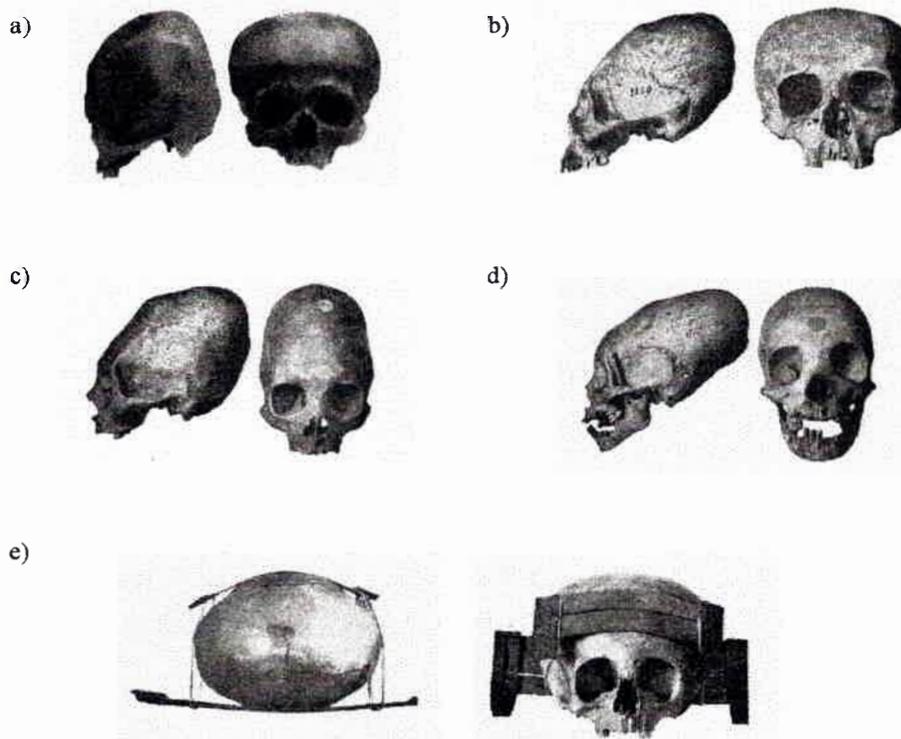


Figura 1. Tipos y variedades de deformación artificial del cráneo. a) Tabular Erecta, b) Tabular Oblicua, c) Circular Erecta, d) Circular Oblicua, e) Aparato deformatorio utilizado para lograr la deformación tabular oblicua.

“...El archivo de datos métricos originales fue sometido a un tratamiento numérico (Seber 1984) por medio del cual se extrajo la variación producida por la edad, el sexo y la deformación artificial, quedando de esta forma una variación residual a partir de la cual se realizaron los cálculos correspondientes a la estimación de la repetibilidad. Mediante un análisis de la varianza de dos vías (Rao, 1952), se comprobó que no existen diferencias estadísticamente significativas entre las fases cronológicas II y III, y entre los materiales exhumados de los tres cementerios antes citados. Además, en cada carácter métrico se puso a prueba el supuesto de normalidad (Shapiro y Wilk 1965, David et al. 1954) y el de homogeneidad de las varianzas (Bliss 1967) entre repeticiones...”

“...La estimación de la repetibilidad de cada variable se efectuó de acuerdo a Falconer (1990) como la correlación intraclase entre ambas repeticiones,

$$r = \frac{V_G + V_{Eg}}{V_F}$$

A continuación se presenta el modelo del análisis de la varianza correspondiente:

Fuente de variación	S.C.	g.l.	C.M.	C.M. Esperado
Entre individuos	entre	$N - 1$	$S.C./ (N - 1)$	$s^2_W + ns^2_B$
Dentro de individuos	dentro	$N(n - 1)$	$S.C./ [N(n - 1)]$	s^2_W

con S.C.=suma de cuadrados, C.M.=Cuadrado Medio, g.l.=grados de libertad, N=número total de individuos, n = número de repeticiones por individuo, $s^2_W \sim V_W = V_{Ee}$, $s^2_B \sim V_B = V_G + V_{Eg}$.

“...El error estándar (ES) de r se calculó siguiendo a Becker (1975) según lo propuesto por Fisher (1954) como,

$$ES = \sqrt{\frac{2[1 + (n - 1)r]^2 (1 - r)^2}{n(n - 1)(N - 1)}}$$

bajo el supuesto de que el número total de observaciones es suficientemente grande, de modo que r se distribuye normalmente...”

3.4. Resultados del Análisis Estadístico Univariado para el Estimación de la Repetibilidad de Caracteres Métricos (Transcripción de Varela y Cocilovo, ep)

“...Los resultados correspondientes a los cálculos de repetibilidad de los siete caracteres métricos del cráneo figuran en la Tabla 1, conjuntamente con el error estándar y la proporción de la varianza ambiental especial. En ella se observa que los valores de r son altos y presentan bajos errores de estimación. Todas las variables tienen repetibilidades mayores del 75% con la única excepción de la altura órbito-alveolar que es del 65%. La media de r, teniendo en cuenta las siete variables fue 0.791 y la mediana de 0.822...”

“...En ningún caso el error estándar de la estimación llega al 7%, y en cuatro variables (altura del pómulo, ancho y alto de la órbita, y cuerda porion-basion) no supera el 3.5%. El error estándar promedio calculado para todas las variables es de 0.043. Las proporciones de la varianza ambiental especial tienen una relación inversa con respecto a la repetibilidad, siendo la altura órbito-alveolar y la altura del pómulo, los caracteres que presentan la mayor y la menor proporción de la componente ambiental especial, 35% y 11,6% respectivamente...”

Tabla 1: Correlación entre las mediciones de la izquierda y la derecha de siete variables métricas del cráneo en una población prehistórica del Norte de Chile.

Variabes	r*	E.S.	1 - r
altura del pómulo	0.884	0.024	0.116
ancho de la órbita	0.822	0.035	0.178
altura de la órbita	0.853	0.030	0.147
altura órbito-alveolar	0.650	0.063	0.350
curva horizontal	0.747	0.064	0.253
cuerda porion-basion	0.832	0.034	0.168
cuerda porion-bregma	0.748	0.048	0.252

r = repetibilidad o coeficiente de correlación, E.S. = error estándar de r,

1-r = varianza ambiental especial como proporción de la varianza total.

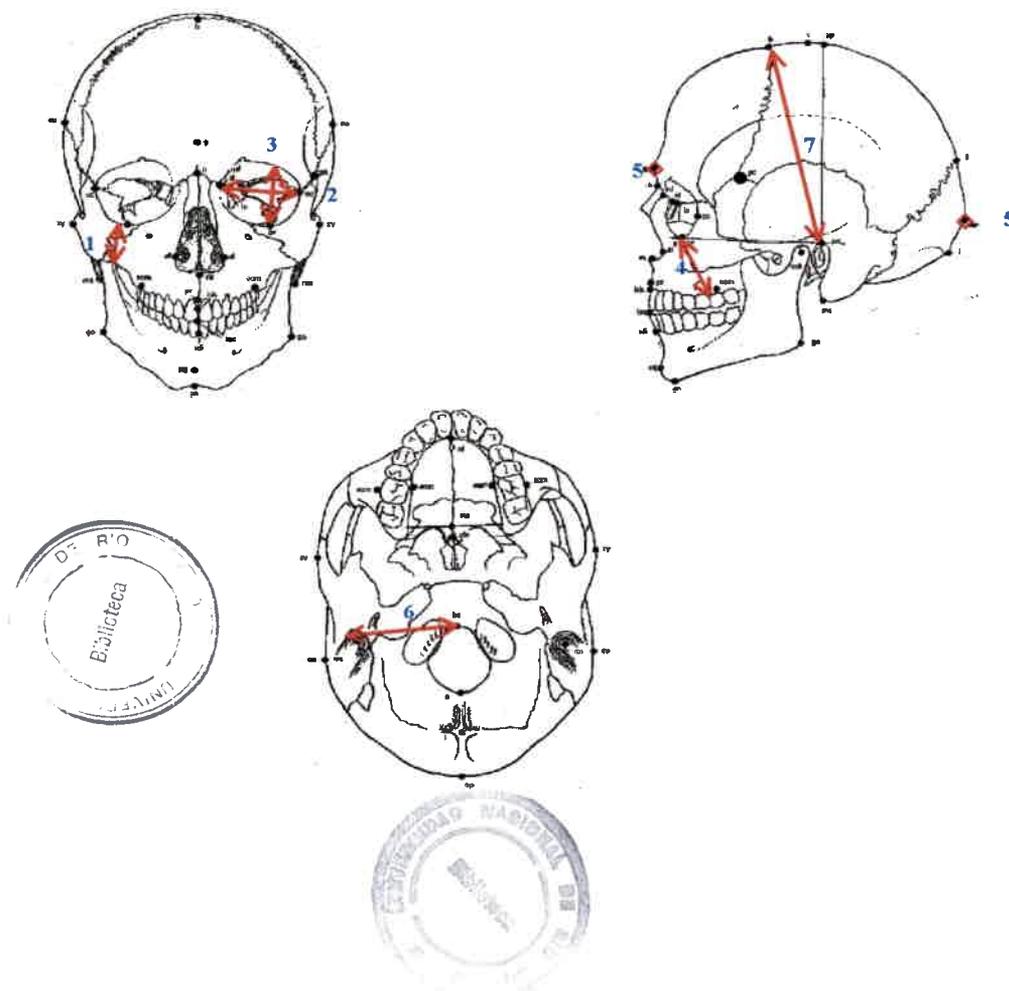
* Todos los valores r son distintos de cero con $p < 0.0000$

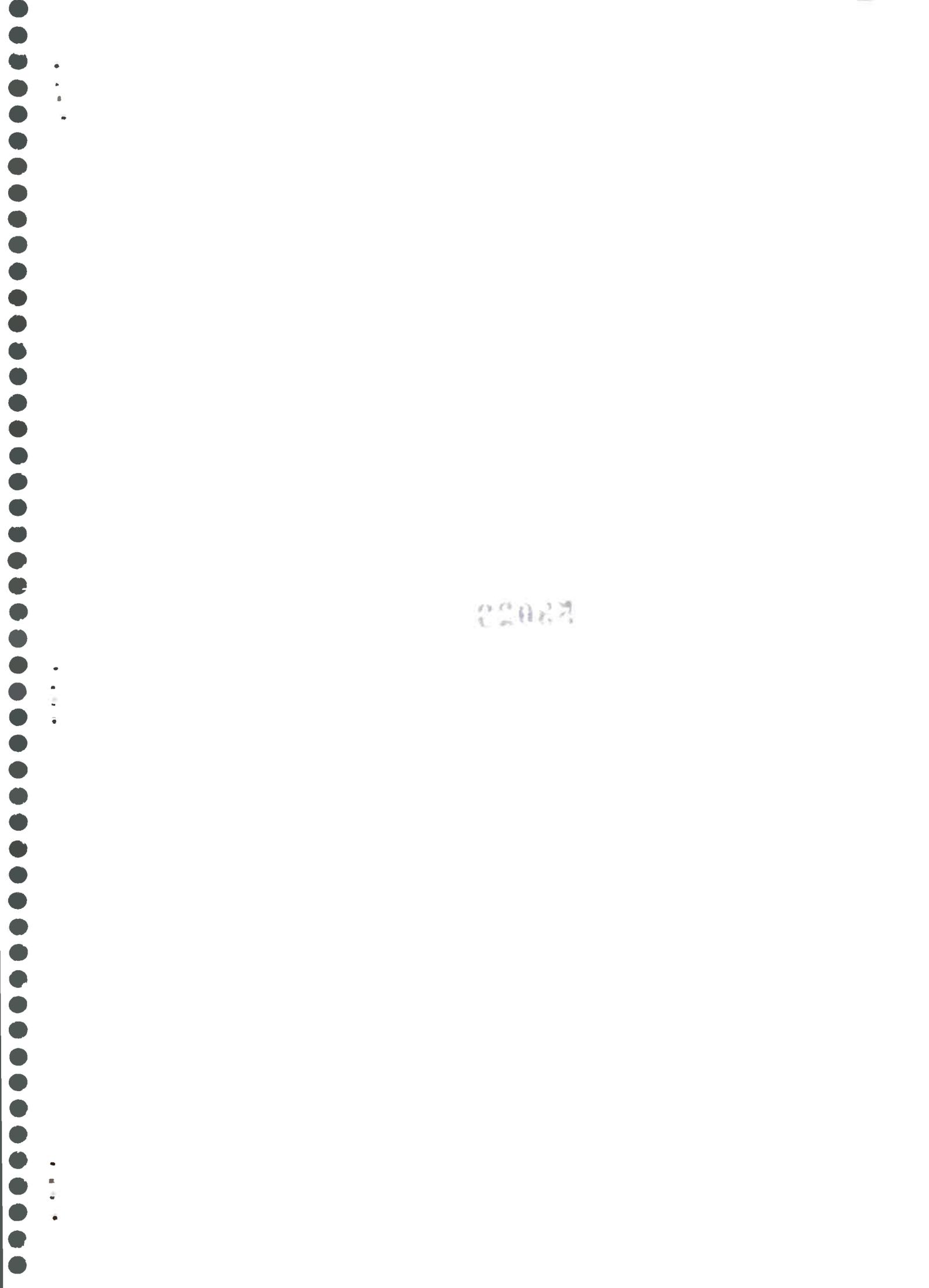
3.5. Variables Utilizadas en el Análisis

Definición de las variables medidas en ambos lados (izquierdo y derecho) del cráneo:

Nº	Nombre	Definición	Referencia
1	altura del pómulos	cuerda mínima desde el borde inferior de la órbita al borde inferior del maxilar	Howells (1973) Droessler (1981)
2	ancho de la órbita	desde el dacrion hasta el borde externo de la órbita, en el punto donde termina su eje transversal	Convención Internacional de Mónaco de 1906 (Comas, 1966)
3	altura de la órbita	separación máxima entre el borde superior e inferior de la órbita, perpendicular a la anchura	Convención Internacional de Mónaco de 1906 (Comas, 1966)
4	altura órbita-alveolar	distancia mínima entre el borde inferior de la órbita y el borde alveolar	Convención Internacional de Mónaco de 1906 (Comas, 1966)
5	arco horizontal	desde la glabella hasta el opistocráneo	definida por JAC y HHV
6	cuerda porion-basion	desde el porion hasta el basion	definida por JAC y HHV
7	cuerda porion-bregma	desde el porion hasta el bregma	definida por JAC y HHV

JAC: José A. Cocilovo, HHV: H.Hugo Varela





53023