

CONTRIBUCIONES AL MANOVA-BIPLLOT: REGIONES DE CONFIANZA ALTERNATIVAS

I. R. Amaro^{1*}, J. L. Vicente-Villardón^{**}, M^a. P. Galindo-Villardón^{**}.

*Universidad Central de Venezuela, Facultad de Ingeniería, Núcleo de Cagua. Cagua 2122. Venezuela.

**Universidad de Salamanca. Departamento de Estadística. Salamanca, España.

RESUMEN

En este trabajo se proponen regiones de confianza elípticas para el Manova-Biplot, en lugar de las regiones circulares tradicionalmente usadas. En la construcción de estas regiones se considera la variabilidad muestral de los ejes de la representación y se toma como punto de partida el desarrollo matemático usado Por KRZANOWSKI (1989) en el contexto del Análisis de Variables Canónicas. En una aplicación de los resultados teóricos a un conjunto de datos se observa que el área de las elipses es considerablemente más grande que el área de los círculos. Intuitivamente es de esperarse entonces que las regiones elípticas tengan más exactitud probabilística que las circulares.

ABSTRACT

We propose elliptical confidence regions for the Manova-Biplot, instead of the traditionally used confidence circles. These regions are constructed taken into account the sample variability of the representation axes and are based on the mathematical development used by Krzanowski (1989). An application of the theoretical results to a data set, shows that the areas of the ellipses are considerably greater than the areas of the circles. These suggest that elliptical confidence regions are probabilistically more accurate than the traditional circles.

KEY WORDS: Biplot, MANOVA, Manova-Biplot, Confidence regions.

MSC 62H25

1. INTRODUCCIÓN

El Manova-Biplot es una técnica estadística multivariante utilizada en situaciones experimentales donde se dispone de varias variables respuesta y se quiere buscar las diferencias entre varios grupos. Fue denominada Manova-Biplot por Gabriel (1972, 1995), aunque también se le conoce como Biplots Canónicos por Vicente-Villardón (1992) y Gower Y Hand (1996). El Manova-Biplot para diseños de dos vías fue introducido por Amaro et al. (2004).

En el presente trabajo se proponen regiones de confianza para el Manova-Biplot construidas considerando la variabilidad de los ejes y tomando como punto de partida los desarrollos matemáticos usados por Krzanowski (1989) en el contexto del Análisis de Variables Canónicas.

A diferencia del Análisis de Variables Canónicas, en el Manova-Biplot, además de interpretar las diferencias- semejanzas entre los grupos; también podemos interpretar las relaciones entre las variables; y las relaciones entre grupos y variables. Adicionalmente, el Manova-Biplot nos proporciona medidas de la calidad de representación tanto de medias de grupo como de las variables, que nos permiten una mejor interpretación de los resultados.

En los casos en los que se tienen arreglos de tratamientos con dos factores o más, es necesario el uso del Manova, y por lo tanto el Manova-Biplot además permite estudiar la interacción, el efecto del factor fila y el efecto del factor columna.

2. MÉTODOS

¹ :amaror@hotmail.com

El Análisis Multivariante de la Varianza o MANOVA, por sus siglas en ingles, es una técnica muy conocida que comienza con una matriz de datos \mathbf{X} que se supone particionada en k grupos, el i -ésimo de los cuales esta formado por n_i individuos, $i=1,\dots,k$, $n = \sum_{i=1}^k n_i$; y sea \mathbf{x}_{ij} el vector de p observaciones de las variables x_1, x_2, \dots, x_p sobre el j -ésimo individuo ($j=1,\dots,n_i$) del i -ésimo grupo que se supone es una muestra aleatoria de una población $N(\boldsymbol{\mu}_i, \boldsymbol{\Sigma})$.

Los k grupos se consideran muestras que dependen de parámetros desconocidos y siguen el Modelo Lineal General Multivariante $\mathbf{X} = \mathbf{A}\mathbf{B} + \mathbf{U}$, para una clasificación de una vía, y se asume que tienen varianza común, estimada por \mathbf{E} , que es la matriz de sumas de cuadrados y productos “dentro” de grupos y \mathbf{H} es la matriz de sumas de cuadrados y productos “entre” grupos.

Se quiere probar la Hipótesis Lineal General: $H_0 : \mathbf{C}\mathbf{B}\mathbf{M} = \mathbf{0}$.

El Manova-Biplot se construye a partir de la descomposición en valores singulares generalizada:

$$\mathbf{R}^{-1/2} \hat{\mathbf{D}} \mathbf{E}^{-1/2} = \mathbf{U} \mathbf{D}_\lambda \mathbf{V}^T \quad (1)$$

donde:

$$\mathbf{R} = \mathbf{C}(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{C}^T$$

$$\hat{\mathbf{D}} = \mathbf{C} \hat{\mathbf{B}} \mathbf{M} = \mathbf{C}(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{X} \mathbf{M}$$

Tomando como marcadores para las filas (grupos):

$$\mathbf{P} = \mathbf{R}^{-1/2} \mathbf{U} \mathbf{D}_\lambda$$

y como marcadores para las columnas (variables) a:

$$\mathbf{Q} = \mathbf{E}^{-1/2} \mathbf{V}$$

En dos dimensiones, las dos primeras columnas de la matriz \mathbf{P} contienen las coordenadas de las medias de grupo. Las regiones de confianza más usadas han sido círculos alrededor de estas medias.

3. REGIONES DE CONFIANZA ALTERNATIVAS

El objetivo es construir regiones de confianza para las medias de grupo en el Manova-Biplot, tomando en consideración la variabilidad muestral de los ejes.

La descomposición Manova-Biplot de la ecuación (1) establece que la matriz $\mathbf{P} = \mathbf{R}^{-1/2} \mathbf{U} \mathbf{D}_\lambda$ contiene los marcadores de los grupos en la representación Biplot. En la proyección en dos dimensiones tenemos que las coordenadas muestrales del centroide del grupo i son:

$$\mathbf{p}_i^T = \left(P_{i,1}, P_{i,2} \right) \quad i=1,\dots,k$$

Donde $P_{i,1}$ son las coordenadas sobre el primer eje y $P_{i,2}$ las coordenadas sobre el segundo eje, del centroide del grupo i .

Entonces se quiere construir una región de confianza para \mathbf{p}_i^T .

Cuadras (1996), pág. 352, establece que si $\mathbf{a} = (a_1, \dots, a_n)$ es un punto del espacio Euclídeo \mathfrak{R}^n , \mathbf{C} es una matriz definida positiva y b es una constante, el conjunto de puntos $\mathbf{x} = (x_1, \dots, x_n)^T$ tales que:

$$(\mathbf{x} - \mathbf{a})^T \mathbf{C}^{-1} (\mathbf{x} - \mathbf{a}) \leq b$$

constituyen un elipsoide de \mathfrak{R}^n de centro en \mathbf{a} .

Basado en esto, la teoría estándar de la distribución normal multivariante, que puede verse por ejemplo en Mardia et al. (1979), pág. 38, Morrison (1978), pág. 87, Seber (1984), pág. 19; establece el siguiente resultado general: si ζ_i es asintóticamente normal con media cero y matriz de varianzas y covarianzas Σ entonces:

$$(\zeta_i - \mathbf{a})^T \Sigma^{-1} (\zeta_i - \mathbf{a}) \cong \chi_{\alpha}^2 \quad (2)$$

Donde χ_{α}^2 es el punto crítico de la distribución Chi-cuadrado. De tal forma que una región de confianza del $100(1 - \alpha)\%$ para la i -ésima media de grupo, es un elipsoide centrado en \mathbf{a} y definido por:

$$(\zeta_i - \mathbf{a})^T \Sigma^{-1} (\zeta_i - \mathbf{a}) \leq \chi_{\alpha}^2$$

Si se supone que los ejes son fijos, entonces la matriz Σ es la identidad, Cuadras (1996), pág. 356, y se tiene:

$$n_i (\zeta_i - \mathbf{a})^T (\zeta_i - \mathbf{a}) \leq \chi_{\alpha}^2$$

que es la tradicional región de confianza, una hiperesfera centrada en \mathbf{a} .

Cuando se proyectan sobre los dos primeros ejes resultan círculos de radio $\chi_{2,\alpha}^2 / n_i$. Estas son las regiones más usadas en estos análisis.

Algunas aplicaciones relativamente recientes del Manova-Biplot usando regiones circulares pueden verse en: Vicente, M. A. et al. (2001), Iñigo, A. C. et al. (2004) y VARAS, M. J. et al. (2005).

Los ejes Biplot de la representación gráfica se construyen a partir de los vectores propios asociados a los valores propios λ_i de la descomposición (1).

Los $P_{i,j}$ son las proyecciones de las medias del i -ésimo grupo sobre el eje j , por lo tanto pueden escribirse en función de los vectores propios asociados, luego la varianza $Var(P_{i,j})$, dependerá de la variabilidad de cada eje.

Cuando los ejes **no** se consideran fijos la distribución muestral exacta de los valores propios λ_i y de los vectores propios asociados \mathbf{v}_i es muy compleja para ser usada en la práctica, aunque Anderson (1984), pág. 545, estudia sus distribuciones asintóticas y encuentra que, si $(n - k) \mathbf{E}$ sigue una distribución de Wishart con $f_1 = (n - k)$ grados de libertad y parámetro Σ , mientras que $(k - 1) \mathbf{H}$ sigue una distribución de Wishart independiente no central con $f_2 = (k - 1)$ grados de libertad y parámetro

$\Phi = \sum n_i (\boldsymbol{\mu}_i - \boldsymbol{\mu})(\boldsymbol{\mu}_i - \boldsymbol{\mu})^T$, y $f_2/f_1 \rightarrow \gamma > 0$, entonces los momentos asintóticos aproximados son:

$$E(\lambda_i) = \theta_i, \quad E(\mathbf{v}_i) = \mathbf{a}_i,$$

$$\text{var}(\lambda_i) = \frac{2\theta_i^2(1+\gamma)}{\gamma f_1}, \quad \text{cov}(\mathbf{v}_i, \mathbf{v}_j) = -\frac{\theta_i \theta_j (1+\gamma)}{f_1 \gamma (\theta_j - \theta_i)} \mathbf{a}_j \mathbf{a}_i^T \quad (i \neq j)$$

A partir de aquí se demuestra que: $E(P_{i,j}) = \xi_{i,j}$, la posición de la i -ésima media poblacional con referencia al j -ésimo eje, y, utilizando expresiones para el cálculo de esperanzas condicionales que pueden verse, por ejemplo, en Rao (1973), pág. 96, 97, Krzanowski (1989), encuentra que, para $i=1, \dots, k$, $l=1, \dots, s$:

$$\text{Var}(P_{i,l}) = \frac{1}{n_i} + \frac{\left\{ \frac{1}{2} \xi_{i,l}^2 + \sum_{j \neq l=1}^s \frac{\theta_l (\theta_j + \gamma \theta_l)}{\gamma (\theta_j + \theta_l)^2} \xi_{i,j}^2 \right\}}{(n-k)} \quad (3)$$

y

$$\text{Cov}(P_{i,j}, P_{i,l}) = -\frac{\theta_j \theta_l (1+\gamma)}{(n-k) \gamma (\theta_j - \theta_l)^2} \xi_{i,j} \xi_{i,l}, \quad (j \neq l) \quad (4)$$

Donde los parámetros $\theta_i, \xi_{i,j}$ son los parámetros poblacionales que son estimados, usando distribuciones aproximadas, por $1 + \lambda_i / (k-1)$, y $P_{i,j}$, respectivamente, y $\gamma = (k-1)/(n-k)$, $s = \min(k-1, p)$.

Suponiendo normalidad asintótica de los \mathbf{v}_i y apelando al teorema central del límite tomando en cuenta que p no debe ser muy pequeño, se tiene entonces que los $P_{i,j}$ tienen aproximadamente una distribución normal con media dada por: $E(P_{i,j}) = \xi_{i,j}$, varianza dada por (3) y covarianza por (4).

De tal forma que para construir las regiones de confianza de las medias de grupo en las dos primeras variables como ejes, el vector de medias del i -ésimo grupo $(P_{i,1}, P_{i,2})$ tiene una distribución normal con media $(\xi_{i,1}, \xi_{i,2})^T$ y matriz de dispersión:

$$\begin{bmatrix} v_{i1} & c_i \\ c_i & v_{i2} \end{bmatrix}$$

Donde v_{i1}, v_{i2} están dados por las ecuaciones (3) con $l=1, 2$ respectivamente; mientras que c_i es dado por (4) con $j, l=1, 2$, respectivamente.

Entonces, de acuerdo a la ecuación (2), la región de confianza del $100(1-\alpha)\%$ para la media del grupo i -ésimo, proyectado sobre los dos primeros ejes, es una elipse centrada en $(P_{i,1}, P_{i,2})$ y definida por:

$$(\xi_{i,1} - P_{i,1}, \xi_{i,2} - P_{i,2}) \begin{bmatrix} v_{i1} & c_i \\ c_i & v_{i2} \end{bmatrix}^{-1} \begin{pmatrix} \xi_{i,1} - P_{i,1} \\ \xi_{i,2} - P_{i,2} \end{pmatrix} \cong \chi_{2,\alpha}^2$$

Donde $\chi_{2,\alpha}^2$ es el punto crítico de la distribución chi-cuadrado con dos grados de libertad.

Haciendo:

$$\begin{bmatrix} v_{i1} & c_i \\ c_i & v_{i2} \end{bmatrix}^{-1} = \frac{1}{v_{i1}v_{i2} - c_i^2} \begin{bmatrix} v_{i2} & -c_i \\ -c_i & v_{i1} \end{bmatrix}$$

Llamando x , y a las coordenadas de los puntos sobre la elipse. Después de un desarrollo matricial, se llega a que la ecuación de esa elipse puede ser escrita en la forma:

$$\begin{aligned} v_{i2}(x - P_{i,1})^2 - 2c_i(x - P_{i,1})(y - P_{i,2}) + v_{i1}(y - P_{i,2})^2 \\ = (v_{i1}v_{i2} - c_i^2) \chi_{2,\alpha}^2 \end{aligned} \quad (5)$$

Se puede demostrar que la excentricidad de esta elipse es $(v_{i1}v_{i2} - c_i^2)^{1/2} / v_{i2}$ y que su área interior es $\pi (v_{i1}v_{i2} - c_i^2)^{1/2} \chi_{2,\alpha}^2$.

Estas regiones de confianza, que son elipses con los ejes rotados, se han construido usando las varianzas y covarianzas expresadas por las ecuaciones (3) y (4), que usan en su construcción la variabilidad de los ejes, que es lo que se buscaba.

4. APLICACIÓN

Con el fin de ilustrar el Manova-Biplot con regiones de confianza elípticas se aplicará esta técnica al famoso conjunto de datos de Fisher (ver MARDIA et al. (1979)), que contienen 50 observaciones de 4 características de la flor de *Lis* (variables): V1:Longitud del sépalo (SepalLen), V2:Ancho del sépalo (SepalWid), V3:Longitud del pétalo (PetalLen) y V4:Ancho del pétalo (PetalWid), para tres especies (grupos) del genero *lis*: G1: *setosa*, G2: *versicolor* y G3: *virginia*. Total de observaciones: 150.

Antes de estudiar la aplicación del método es necesario considerar ciertos aspectos como son: el examen previo de la matriz de datos multivariantes, donde se incluye la detección de outliers y el tratamiento de datos ausentes y la verificación de los supuestos del análisis multivariante de la varianza. Los resultados de estos análisis no se muestran aquí ya que este conjunto de datos ha sido ampliamente usado en la literatura para ejemplificar varias técnicas estadísticas, como por ejemplo el Análisis Discriminante y MANOVA, ver, por ejemplo, MARDIA et al. (1979), MORRISON (1978).

Es necesario señalar que la normalidad multivariante, además de ser un requisito para MANOVA, también es necesaria para la construcción de las regiones de confianza elípticas que se proponen en este trabajo, y este requisito es cumplido por este conjunto de datos.

5. MANOVA-BILOT

Todos los análisis fueron realizados utilizando programas desarrollados por los autores bajo MATLAB™ ver. 7.

En primer lugar se presentan, en la tabla 1, los valores propios, inercias e inercias acumuladas para el Biplot.

Tabla 1: Valores propios, inercias e inercias acumuladas.

Eje	Valores propios	Porcentaje de inercia	Inercia acumulada
1	5.67	99.12	99.12
2	0.53	0.87	100
3	0.00	0.00	100

Se puede observar que el porcentaje de variabilidad explicado por el primer plano factorial es 100 %, que, es igual a la bondad de ajuste en el plano.

Las calidades de representación para las medias se muestran en la siguiente tabla:

Tabla 2: Calidades de representación de las medias en el Biplot.

Grupos	Eje 1	Eje2	Acumulada
G1	99.92	0.08	100
G2	86.28	13.72	100
G3	99.22	0.78	100

Podemos observar, en la tabla 2, que, para cada grupo, los dos primeros ejes recogen el 100 % de la variabilidad de las medias, por lo que las diferencias entre los grupos pueden interpretarse sobre el primer plano.

Tabla 3: Calidades de representación de las variables en el Biplot.

Variables	Eje 1	Eje 2	Eje 3	Acumulada
V1	4.96	9.66	83.29	97.91
V2	1.42	13.49	74.59	89.50
V3	49.85	40.89	2.81	93.55
V4	40.09	54.35	0.18	94.62

. Las calidades de representación para las variables se muestran en la siguiente tabla:

Las variables V3 y V4 tienen una alta calidad de representación en los dos primeros ejes: 90.74 y 94.44, respectivamente. Por lo que estas dos variables pueden ser interpretadas en estos ejes. No así, las

variables V1 y V2 cuyas calidades de representación son bajas en el primer plano, 14.62 y 14.91, respectivamente, por lo que no se interpretan

En la tabla 4 se muestran las varianzas calculadas usando la ecuación 3.

Tabla 4: Varianzas según ecuación 3.

Grupo	Eje 1	Eje 2
G1	0.0215	0.1643
G2	0.0213	0.2830
G3	0.0214	0.1034

Y las covarianzas calculadas con la ecuación 4 se muestran en la tabla 5.

Tabla 5: Covarianzas según ecuación 4.

Grupo	Covarianza
G1	-0.0042
G2	-0.0034
G3	0.0075

Con los valores mostrados en las tablas 4 y 5, se construyen las ecuaciones de las elipses de confianza del 95 % para los centroides de grupo según el procedimiento descrito antes, ecuación (5). En la tabla 6 se muestran las ecuaciones de las elipses de confianza para cada grupo.

Tabla 6: Ecuaciones de las elipses de confianza.

Grupo	Ecuación de la elipse
G1	$0.164(x + 2.3)^2 + 0.008(x + 2.3)(y + 0.06) + 0.021(y + 0.06)^2 = 0.021$
G2	$0.028(x - 0.5)^2 + 0.006(x - 0.5)(y - 0.22) + 0.021(y - 0.22)^2 = 0.003$
G3	$0.103(x - 1.7)^2 - 0.015(x - 1.7)(y + 0.15) + 0.021(y + 0.15)^2 = 0.012$

Se puede observar en la tabla 7, que las elipses tienen diferentes ángulos de inclinación variando desde el grupo G1 con un ángulo muy pequeño de 1.7°, hasta el grupo G3 con 23.1°. Así mismo, las áreas de las elipses son diferentes.

En la figura 1 se muestra la representación Biplot en el primer plano factorial.

Usando las propiedades del Manova-Biplot, en la figura 1 podemos observar que las elipses de confianza no se interceptan, lo que significa, que hay diferencias entre los tres grupos, es decir que, las tres especies son diferentes en las cuatro variables estudiadas simultáneamente.

Las características más importantes de estas elipses están mostradas en la tabla 7.

Tabla 7: Características de las elipses de confianza.

Grupo	Centro	Ángulo De Inclinación	Longitud del eje mayor	Longitud del eje menor	Área
G1	(-2.3, -0.06)	1.7°	0.99	0.35	1.114
G2	(0.5, 0.22)	23.1°	0.42	0.34	0.418
G3	(1.7, -0.15)	5.21°	0.67	0.35	0.873

Siguiendo con las propiedades, se puede comparar un grupo con el promedio de otros dos. Así, por ejemplo, el promedio de las medias de los grupos G1 y G2, es el punto medio del segmento que une los centroides de G1 y G2. Y así se pueden comparar los otros casos.

MANOVA BIPLLOT DE UNA VIA

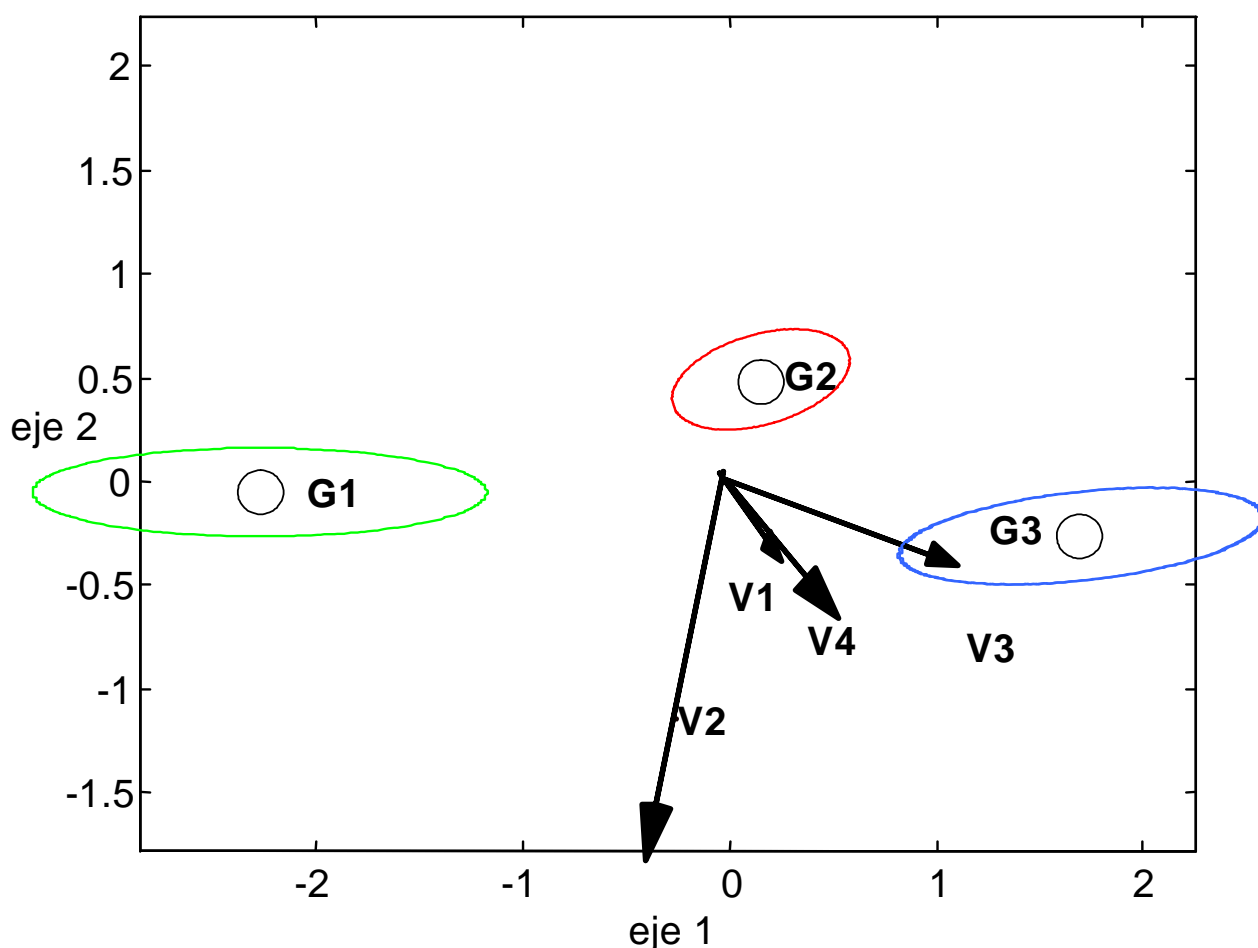


Figura 1.-Representación Biplot con Elipses y Círculos.

De la representación se deduce que los tres grupos son diferentes en la variable V3 (longitud de los pétalos), porque al proyectar sus elipses de confianza sobre el vector que representa a esa variable, las proyecciones son disjuntas. Con este mismo argumento, también se tiene que los tres grupos son diferentes en la variable V4 (ancho de los pétalos).

De la tabla 1 se tiene que el primer eje explica el 99.12 % de la variación entre grupos, y en la tabla 9 se observa que la variable V3 es la que tiene la mas alta correlación con el eje 1 (97.01), es por esto que en la figura 1 aparece el grupo G1 opuesto a los otros dos grupos indicando que observaciones (flores) con

valores altos para la variable V3 (pétalos más largos) aparecen situadas a la derecha del gráfico. Es decir, la especie *lis setosa* tiene los pétalos más cortos, seguida de la especie *Lis versicolor* y finalmente, con los pétalos más largos, la especie *Lis virginia*.

Se ve entonces en el lado izquierdo del gráfico la especie *lis setosa* con longitud y ancho del pétalo pequeños, mientras que en el otro extremo, a la derecha, la especie *lis virginia* con longitud y ancho de los pétalos más grandes.

También en la figura 1, observamos que la variable que más influye en la separación de los grupos G2 y G3 es, V3, porque el vector que representa a esta variable es casi paralelo a la recta que pasa por los centroides de G2 y G3.

La desviación estándar de la variable V4 es mayor que la de la variable V3, esto es: el vector que corresponde a V4 tiene mayor longitud que el que corresponde a V3. A pesar de que en la figura 1 aparece más largo el vector que representa a V2, esa variable no es interpretable en este plano por tener muy baja calidad de representación.

Otro aspecto importante es el que tiene que ver con el área de las elipses. Por la forma como se han construido, el área de las elipses está relacionada con la variabilidad muestral de cada grupo y con el tamaño muestral. Así observamos que la elipse correspondiente al grupo G1 tiene mayor área y este es el grupo con mayor variabilidad.

Tabla 8: Leyenda para la figura 1.

GRUPOS	VARIABLES
G1: <i>Lis setosa</i> G2: <i>Lis versicolor</i> G3: <i>Lis virginia</i>	V1: Longitud del sépalo V2: Ancho del sépalo V3: Longitud del pétalo V4: Ancho del pétalo

En la figura 1 están representadas las regiones elípticas y las circulares, cuyos radios vienen dados por la expresión:

$$\chi_{2,\alpha}^2/n_i$$

Se observa, en esta figura, que todos los círculos tienen el mismo radio porque los datos son balanceados. Hay una gran diferencia entre el área de los círculos y el de las elipses: el área de las elipses es considerablemente más grande que el área de los círculos. Intuitivamente es de esperarse entonces que bajo diferentes muestras de estas poblaciones, en un mayor número de veces las regiones elípticas contendrán a la verdadera media poblacional en lugar de las regiones circulares.

Tabla 9: Correlaciones de las variables con los ejes.

Variables	Correlaciones		Correlaciones al cuadrado (x 100)	
	Eje 1	Eje 2	Eje 1	Eje 2
V1	0.7919	0.2176	62.71	4.73
V2	-0.5308	0.7580	28.17	57.45
V3	0.9850	0.0460	97.01	0.21
V4	0.9728	0.2229	94.64	4.97

6. DISCUSIÓN

En el ejemplo presentado se ha mostrado una metodología para construir regiones de confianza elípticas para los centroides en un Manova-Biplot considerando la variabilidad muestral de los ejes de la representación.

La incorporación de estas regiones debe conducir a análisis más precisos que los obtenidos con las regiones circulares clásicas.

Las regiones de confianza propuestas están basadas en una serie de aproximaciones y los resultados arrojan considerables diferencias entre las áreas de las elipses y de los círculos, por lo que se plantea la necesidad de futuras investigaciones en este sentido. Se recomienda realizar un estudio comparativo, posiblemente utilizando técnicas de simulación, Bootstrap, etc.

La metodología propuesta deja abierta la posibilidad de futuras investigaciones en otros aspectos, como son: generalización a diseños de dos factores o más, Manova-Biplot Robustos, Análisis de la Covarianza, Análisis de Perfiles, Modelos Mixtos, etc.

7. CONCLUSIONES

1.-Es posible construir regiones de confianza elípticas para los centroides de los grupos en un Manova-Biplot, y a partir de aquí, interpretar gráficamente diferencias entre grupos, las variables responsables de las diferencias y otros contrastes de interés.

2.-Las regiones de confianza elípticas se han construido tomando en consideración la variabilidad muestral de los ejes, lo que incorpora información inferencial que permitirá hacer análisis más precisos.

3.-Se aplicó el Manova-Biplot, con regiones de confianza elípticas, para analizar un conjunto de datos de la bibliografía y, comparándolo con las regiones de confianza circulares clásicas, intuitivamente pareciera que las nuevas regiones tienen más exactitud probabilística que las circulares.

4.-Se pone de manifiesto, en la aplicación, las ventajas del Manova-Biplot respecto al Análisis de Variables Canónicas, sobre todo en cuanto a aspectos como son: la representación en el mismo gráfico de los centroides de los grupos, de las variables y de las regiones de confianza; y el uso de medidas de las calidades de representación.

Received March 2007

Revised July 2008

REFERENCIAS

- [1] ANDERSON, T. W. (1984): **An Introduction to Multivariate Statistical Analysis**. 2nd ed. Wiley, New York,.
- [2] AMARO, I. R., VICENTE-VILLARDÓN, J. L. y GALINDO-VILLARDÓN, M. P. (2004): Manova Biplot para arreglos de tratamientos con dos factores basado en modelos lineales generales multivariantes. **Interciencia**, 29, 26-32.
- [3] CUADRAS, C. M. (1996): **Métodos de análisis multivariante**. Eunibar, Barcelona.
- [4] GABRIEL, K.R. (1972): Analysis of meteorological data by means of canonical decomposition and Biplots. **Journal of Applied Meteorology**, 11: 1071-1077.
- [5] GABRIEL, K. R. (1995): MANOVA Biplots for two-way contingency tables. En: **Recent Advances in Descriptive Multivariate Analysis**. (W. KRZANOWSKI, ed.). Clarendon Press, Oxford, 227-268.
- [6] GOWER, J. C. y HAND, D. J. (1996): **Biplots**. Chapman and Hall, London,
- [7] IÑIGO, A. C., VICENTE-TAVERA, S y RIVES, V. (2004): MANOVA-Biplot statistical analysis of the effect of artificial ageing (Freezing/Thawing) on the colour of treated granite stones. **Color research and application**. . 29, 2, 115-120.
- [8] KRZANOWSKI, W. J. (1989): On confidence regions in canonical variate analysis. **Biometrika**, 76, 107-116.

- [9] MARDIA, K. V.; KENT, J. T. y BIBBY, J. M. (1979): **Multivariate Analysis**. Academic Press, London,
- [10] The Math Works (2007): **MATLAB™ version 7**. The Math Works, inc., Natick, MA, USA.
- [11] MORRISON, D. F. (1978): **Multivariate Statistical Methods**. McGraw-Hill, London-
- [12] RAO, C. R. (1973): **Linear statistical inference and its applications**. Wiley, New York.
- [13] SEBER, G. A. F. (1984): **Multivariate Observations**. Wiley, New York.
- [14] VARAS, M. J., VICENTE-TAVERA, S., MOLINA, E. y VICENTE-VILLARDÓN, J. L. (2005): Role of canonical biplot method in the study of building stones: an example from Spanish monumental heritage. **ENVIRONMETRICS**, 16, 1-15.
- [15] VICENTE-VILLARDÓN, J. L. (1992). Una alternativa a las técnicas factoriales basada en una generalización de los métodos Biplot. **Tesis Doctoral. Universidad de Salamanca**.
- [16] VICENTE, M. A., y VICENTE-TAVERA, S. (2001): Clay poultices in salt extraction from ornamental stones: a statistical approach. **Clays and Clay Minerals**, . 49,, 227-235.