

## Modelo estadístico utilizando análisis discriminante para clasificar problemas refractivos oculares en la ciudad de barranquilla

### Statistical model using discriminant analysis to classify ocular refractive problems in the city of Barranquilla

Jose Barraza <sup>1</sup>

<sup>1</sup>Magíster en Estadística Aplicada, Universidad del Norte, Barranquilla - Colombia.  
Especialista en Estadística Universidad Nacional de Bogotá  
e-mail: josebarraza@mail.uniatlantico.edu.co - jobaest@hotmail.com

Hernando Hernández <sup>2</sup>

<sup>2</sup> Optómetra de la Universidad Metropolitana, Barranquilla - Colombia  
e-mail: optoacademico@unimetro.edu.co – lealernandez@gmail.com

Recibido: 04/07/2014 - Aceptado: 13/11/2014

---

#### Resumen

Este trabajo tiene como propósito construir un modelo estadístico multivariado para clasificar pacientes con problemas refractivos oculares, en términos de un conjunto de variables que especifican un tipo de defecto refractivo ocular. El estudio se desarrolló en Barranquilla, Colombia, con las historias clínicas de 499 pacientes, de diferentes profesiones, con edades de 8 a 82 años. Los pacientes fueron seleccionados al azar de una base de datos facilitada para este fin por un consultorio optométrico de la ciudad. La base de datos consta de 15 variables cuantitativas y 5 variables cualitativas. Las técnicas de análisis multivariados aplicadas al estudio fueron: el Análisis Factorial por Componentes Principales, con el objetivo de crear nuevas variables que resuman óptimamente toda la información contenida en las variables estudiadas, y luego el Análisis Discriminante, con el propósito de clasificar a los pacientes en algunos de los grupos definidos por el tipo de defecto refractivo ocular manifestado en su diagnóstico. El análisis multivariado es una herramienta que permite descubrir la interdependencia que pueda existir entre variables que miden un defecto refractivo visual (dioptrías en el lente esférico, dioptrías en el lente cilíndrico, eje del cilindro, agudeza visual visión próxima), la dependencia del tipo de defecto refractivo ocular diagnosticado por el médico, y este conjunto de variables o subconjunto generado de ellas, a fin de efectuar la clasificación de grupos de pacientes según el tipo de defecto refractivo visual que manifieste en su diagnóstico, y la predicción o asignación de un paciente en uno de entre varios grupos definidos.

*Palabras claves:* Problemas refractivos oculares, historias clínicas, diagnóstico, análisis factorial, análisis discriminante.

#### Abstract

The present work intends to build a multivariate statistical model to classify patients with ocular refractive problems (hyperopia, myopia, astigmatism, hyperopic astigmatism and myopic astigmatism) in terms of a set of variables that specify a type of eye refractive defect. This study was developed in the North of Barranquilla, analyzing the medical histories provided by a clinic in Optometry. In total 499 patients were selected at random to analyze their diagnostic. Their ages were in the range of 8 to 82, with different professions. The database has 15 quantitative variables and 5 qualitative variables. The Multivariate analyzing techniques than were applied to the study, were the principal components factorial analysis with the aim of creating new variables than summarize all information which it may have in the original variables and the discriminant analysis with the purpose of classifying patients in some of the groups defined by the type of defects refractive manifested in his diagnostic.

*Keywords:* Ocular refractive problems, medical histories, diagnostic, principal components factorial analysis, discriminant analysis

---

## 1. Introducción

En el Caribe colombiano, los oftalmólogos y optómetras recomiendan utilizar anteojos con filtro de protección contra los rayos ultravioletas del sol. La alta incidencia de personas que están padeciendo de defectos refractivos oculares llama mucho la atención. Algunas enfermedades visuales son de carácter multifactorial (M, Tamayo, 2001); por ejemplo, el glaucoma y la hipermetropía elevada. La ambliopía o deterioro de la agudeza visual se produce con mayor frecuencia cuando hay estrabismo. Este trastorno puede ocurrir en ambos ojos, si los dos ojos tienen visión borrosa. Ello puede ocurrir cuando hay un alto grado de miopía, hipermetropía o astigmatismo. Enfermedades como la catarata pueden generar una ambliopía; y cualquier factor que impida que una imagen clara pueda ser enfocada sobre la retina en la parte posterior del ojo, puede llevar al desarrollo de ambliopía en una persona. Con respecto a la frecuencia de aparición de miopía, es de destacar la influencia racial: en los países asiáticos hay un alto porcentaje de esta ametropía. La miopía es muy común en las personas con una excesiva longitud axial. En un estudio comparativo con muestras de 1000 pacientes de dos ciudades colombianas, Bogotá y Barranquilla, realizado por la Academia de Medicina, se encontró que el 56 % del grupo estudiado eran miopes en Bogotá, mientras que en Barranquilla era del 49 %. En este estudio realizado recientemente en el Norte de Barranquilla, en personas de diferentes edades, oficios y ocupaciones, se encontró que el 35,07 % padecían de hipermetropía y astigmatismo, el 5,85 % solo padecía de hipermetropía, el 20,24 % de astigmatismo miópico, el 12,02 % de miopía y el 6,81 % de astigmatismo puro en uno o en ambos ojos. De los 60 pacientes que en sus histo-

rias se les diagnosticó miopía en uno o en ambos ojos, el 80 % tienen en alguna forma un grado de escolaridad (docentes o estudiantes, empleados de oficina, empleados de la salud). En el caso de la hipermetropía sola o combinada con astigmatismo, se encontró que el 60,92 % de los pacientes la padecen. Este defecto se presenta en la práctica con una frecuencia muy elevada (55 %), según estudios realizados por los Dres. M.L. Tamayo Fernández y G. Tamayo Fernández. En las historias clínicas analizadas se encontraron pacientes con antecedentes familiares y antecedentes personales de todo tipo (médico, traumático, quirúrgico y alérgico). Se encontraron, también, pacientes con baja visión, con dificultad para leer, ver televisión o desplazarse independientemente. Entre las causas de la baja visión están las enfermedades transmitidas por los padres, como la diabetes o la hipertensión arterial. Entre las enfermedades más comunes que causan baja visión se tienen: glaucoma, retinopatía diabética o hipertensiva, y catarata. La presbicia, por ejemplo, es una condición de la visión que consiste en una disminución gradual de la capacidad de enfocar objetos cercanos (30 a 40 cm), debida a una pérdida de la elasticidad del cristalino, que comienza a manifestarse después de los 40 años.

## 2. Técnicas estadísticas multivariadas aplicadas

Al aplicar pruebas de normalidad multivariada y de igualdad de matrices de varianzas-covarianzas, se encontró que en ambos casos se rechazan estos supuestos en cada uno de los grupos. Véase los anexos (Tabla 1). Cuando se aplica análisis factorial por componentes principales (AFCP) sobre todas las variables en el estudio, se observó un índice de esfericidad KMO ma-

yor que 0.5, indicio de existencia de correlación significativa entre variables. Identificadas las variables latentes (con AFPC) que especifican el tipo de defecto refractivo ocular, se utilizó análisis discriminante (AD) lineal y cuadrático por el método de inclusión por pasos, para construir los modelos mediante los cuales se puedan clasificar los pacientes en una de las siguientes cinco categorías de la variable diagnóstico: astigmatismo (grupo 1), astigmatismo hipermetrope (grupo 2), astigmatismo miópico (grupo 3), hipermetropía (grupo 4) y miopía (grupo 5). En la muestra, después de excluidos los datos atípicos, quedaron 376 pacientes, de los cuales 26 pertenecen al grupo 1, 129 al grupo 2, 65 al grupo 3, 113 al grupo 4 y 43 al grupo 5. Para especificar un defecto refractivo ocular, se utilizan diferentes tipos de variables, en su mayoría cuantitativas, como lo son: dioptrías en el lente esférico ojo derecho e izquierdo (Lesld y Lesli); adición en las lentes ojo derecho e izquierdo (ADDd y ADDi); lente cilíndrico ojo derecho e izquierdo (Lcild, Lcili); eje del cilindro ojo derecho e izquierdo (Ejed, Ejei); agudeza visual para visión de lejos ojo derecho e izquierdo (Avld, Avli), y agudeza visual para visión próxima ojo derecho e izquierdo (Avpd, Avpi).

Las últimas variables que se registran en los diagnósticos son las distancias pupilares para visión de cerca y de lejos (Dpc, Dpl). La variable edad es utilizada como ilustrativa.

En la búsqueda de determinar unos pocos factores que retengan la mayor variabilidad contenida en los datos, utilizando el SPSS y aplicando el AFPC sobre la matriz de correlaciones, se tiene que con seis componentes se explica un 93,846 % de la variabilidad total. Al agrupar las variables en torno a cada componente aplicando el método de rotación Varimax, se observa, en la tabla 2, que:

- En la primera componente están bien re-

presentadas las variables ADDd, ADDi, Avpd y Avpi. Cuando se precisa hacer una adición dióptrica en una lente, es porque hay presencia de presbicia; esta componente será denominada de "corrección présbita" (CP). Los anteojos de lectura son el modo más común de corregir los problemas de presbicia. Para evaluar la agudeza visual de cerca se utiliza la Cartilla de Jaeger. Si el paciente comienza a alejarse, buscando mejor iluminación, es señal de presbicia.

- En la segunda componente están bien representadas las variables Lcild, Lcili, Ejed y Ejei. Este es el tipo de lentes utilizados para corregir astigmatismo, y presentan una curva más amplia en una dirección que en la otra. La segunda componente será denominada "Corrección Astigmática (CA)".
- En la tercera componente están bien representadas las variables Avld y Avli. Como la agudeza visual para ver de lejos es una medida de la capacidad del sistema visual para detectar, reconocer o detallar situaciones espaciales, la primera componente será denominada de "agudeza visual de lejos" (AVL).
- En la cuarta componente están bien representadas las variables Lesld y Lesli. La formulación de lentes esféricas se da para corregir problemas de hipermetropías o de miopías; la hipermetropía se corrige con lente convergente, y la miopía con lente divergente; por ello, la segunda componente principal es denominada de "ubicación del objeto respecto a la retina" (UORR).
- En la quinta componente están bien representadas las variables Dpc y Dpl. Esta com-

ponente será denominada distancia inter-pupilar (DIP).

En la gráfica 1(Ver Anexos) se observa con claridad, qué grupos de variables están asociados a los componentes rotados por ejemplo:

- La primera componente está correlacionada positivamente con las variables Lesd, Lesli, ADDd, ADDi, y la variable ilustrativa Edad; y negativamente, con las variables Avpd, Avpi, Avli y Avld.
- Las variables Lcild, Lcili, Ejed y Ejei están correlacionadas con la segunda componente, las dos primeras negativamente y las otras positivamente.
- Las variables Dpc y Dpl se encuentran muy cerca al origen, indicando que tienen una baja contribución en la formación de las dos primeras componentes.

### 2.1. *Aplicación del análisis discriminante en los problemas refractivos oculares*

El AD es utilizado para clasificar a los pacientes en uno de los cinco grupos:

1. Astigmáticos (A).
2. Astigmáticos hipermétropes (AH).
3. Astigmáticos miópicos (AM).
4. Hipermétropes (H).
5. Miopes (M).

Esta técnica proporciona los métodos que permiten establecer reglas que se puedan emplear para clasificar a otros pacientes en uno de los grupos anteriormente mencionados. En el AD con una variable dependiente con cinco categorías y 14 variables explicativas, el número máximo de ejes discriminante es cuatro. La clasificación de los pacientes entre los grupos se

hace considerando las variables que más y mejor caracterizan el tipo de defecto refractivo ocular, y que en consecuencia diferencian a los grupos. Estas variables denominadas discriminantes, se presentan como combinaciones lineales (o cuadráticas) de las variables originales, y se expresan por una o más funciones discriminantes. Al aplicar AD por pasos, sobre la variable diagnóstico con las variables latentes generadas por el AFPC, los resultados son los que se muestran en la tabla 2.

Los resultados muestran que las variables de la Tabla 3, son las que más discriminan en la clasificación de los pacientes en las diferentes categorías de la variable diagnóstico. El estadístico lambda de Wilks, que contrasta la hipótesis nula igualdad de medias entre los grupos, es rechazado muy significativamente, ya que el p-valor es igual a 0,000. Por lo tanto, se concluye que la información aportada por las variables latentes es estadísticamente significativa en la discriminación de las categorías. El estadístico lambda de Wilks muestra que su menor valor se tiene cuando se incluyen las seis variables.

### 2.2. *Construcción del modelo con las variables discriminantes*

Identificadas las variables discriminantes, el paso siguiente es la construcción del modelo. En el caso lineal, la comparación de las funciones discriminantes genera desigualdades de la forma:

$$(\vec{x}_i - \vec{x}_j)' S^{-1} \vec{x} - \frac{1}{2} (\vec{x}_i - \vec{x}_j)' S^{-1} (\vec{x}_i + \vec{x}_j) > \frac{1}{2} \ln \frac{\Pi_j}{\Pi_i}$$

Donde  $\Pi_i$  es la probabilidad a priori para la categoría  $i = 1, \dots, 5$  y  $S$  es la matriz de varianzas-covarianzas combinadas; en este caso, esa matriz tiene como elementos los mostrados en la tabla 4.

En el caso cuadrático, estas funciones generan desigualdades de la forma:

$$\vec{x}'(S_j^{-1} - S_i^{-1})\vec{x} - 2\vec{x}'(S_j^{-1}\vec{x}_j - S_i^{-1}\vec{x}_i) + (\vec{x}_j'S_j^{-1}\vec{x}_j - \vec{x}_i'S_i^{-1}\vec{x}_i) > \ln \frac{|S_i|}{|S_j|} + \ln \frac{\Pi_i}{\Pi_j}$$

En ambos modelos, se clasifica al individuo  $\vec{x}$ , en la población  $i$ , si se satisface la desigualdad para cada  $j, j = 1, \dots, 5, j \neq i, i = 1, \dots, 5$ . Las  $S_j, j = 1, \dots, 5$ , son las matrices de varianzas-covarianzas muestrales de cada categoría.

Construidas estas funciones, el interés se centra en determinar su importancia en la correcta clasificación de los pacientes. Para ello, bastaría comparar el porcentaje de personas correctamente clasificadas en cada grupo, en relación con el porcentaje de pacientes que se esperaba fueran correctamente clasificados aleatoriamente, teniendo en cuenta los tamaños de cada grupo.

Las probabilidades previas de clasificación de pacientes en astigmáticos, astigmáticos hipermétropes, astigmáticos miópicos, hipermétropes y miopes fueron: 0.071, 0.333, 0.181, 0.286 y 0.129, respectivamente. Estos resultados se deducen de la tabla 5, Grupo 1 (Astigmáticos);  $\Pi_1 = \frac{27}{(27+127+69+109+49)} = 0,071$ , etc. No tiene sentido utilizar las funciones discriminantes para clasificar pacientes si el porcentaje de pacientes bien clasificados son valores cercanos a los anteriores. Las diferencias en los resultados de clasificación son notorias. Mientras que en el modelo lineal el porcentaje clasificado correctamente respecto a los casos agrupados originales fue de 79,8 %, para el modelo cuadrático fue del 87,9 %, como se observa en la tabla 5 (ver anexos).

La discriminación cuadrática muchas veces resulta bastante inestable, salvo que se tenga muestras grandes, además de que con frecuencia se presentan mejores resultados con el modelo lineal; en este caso, sucedió todo lo contrario.

Se observa, asimismo, la estabilidad de los dos modelos. Comparando los resultados de clasificación al aplicar estos modelos en comparación con una clasificación al azar, mediante el índice de significancia práctica (ISP), se observa en el caso lineal un 73,84 % de pacientes bien clasificados, en comparación con el porcentaje de pacientes que se esperaba fueran bien clasificados, al azar, en caso cuadrático: este porcentaje es más del 82 %. El ISP es un criterio que se utiliza para determinar qué tan bueno es un modelo en la correcta clasificación de los pacientes en uno de los grupos previamente definidos según la característica de su tipo de defecto refractivo, y en relación con la cantidad de pacientes que se esperaba fueran correctamente clasificados en cada grupo al azar.

### 3. Conclusiones

El AFCP, aplicado a este estudio en el Norte de Barranquilla, facilitó descubrir la alta intercorrelación entre las variables que tipifican este tipo de deficiencias visuales. Con seis factores seleccionados, se logra explicar más del 94 % de la variabilidad total contenida en las variables originales.

Al clasificar a los pacientes en una población adecuada, comparando los modelos generados utilizando AD lineal con el modelo utilizando AD cuadrático, se observó un mejor ajuste en el modelo cuadrático.

El AD es muy sensible a las desviaciones de la normalidad de los datos (Daniel Peña, 2002) y, a pesar de que la discriminación lineal es más robusta, en este estudio, con el AD cuadrático se obtuvo un mejor ajuste. Con el AD lineal, se obtiene una reducción de más del 74 % en el error de clasificación, en relación con el porcentaje obtenido por una clasificación al azar. Con el AD cuadrático, se obtiene una reducción, de más del 82 %, en el error de clasificación cuando

este análisis se realiza sobre las variables latentes generadas por el AF. Y con un valor agregado en relación con el modelo lineal —como es el hecho de la correcta clasificación de más elementos en cada grupo en relación con el modelo lineal—, más aún, la tabla 7 muestra en las celdas sombreadas que solo 10 de los 306 pacientes bien clasificados por el modelo lineal fueron mal clasificados por el modelo cuadrático. Es decir, de los 306 elementos bien clasificados por el mejor modelo lineal, fueron bien clasificados por el mejor modelo cuadrático 296 (más del 96,7%).

Para concluir, el mejor modelo lineal de clasificación viene dado por:

$$(\vec{x}_i - \vec{x}_j)' S^{-1} \vec{x} - \frac{1}{2} (\vec{x}_i - \vec{x}_j)' S^{-1} (\vec{x}_i + \vec{x}_j) > \frac{1}{2} \ln \frac{\Pi_j}{\Pi_i}$$

En el caso cuadrático, viene dado por:

$$\vec{x}' (S_j^{-1} - S_i^{-1}) \vec{x} - 2\vec{x}' (S_j^{-1} \vec{x}_j - S_i^{-1} \vec{x}_i) + (\vec{x}_j' S_j^{-1} \vec{x}_j - \vec{x}_i' S_i^{-1} \vec{x}_i) > \ln \frac{|S_i|}{|S_j|} + \ln \frac{\Pi_i}{\Pi_j}$$

donde se clasifica al individuo  $\vec{x}$ , en la población  $i$ , si se satisface la desigualdad anterior para cada  $j, j = 1, \dots, 5, j \neq i$ .

**Prueba de Kolmogorov-Smirnov para una muestra**

	Lesld	Lesli	ADDd	ADDi	Lcild	Lcili	Ejed	Ejei	Avld	Avli	Avpd	Avpi	Dpc	Dpl
N	376	376	376	376	376	376	376	376	376	376	376	376	376	376
Parámetros normales <sup>a, b</sup>														
Media	,4965	,4980	1,6064	1,6064	-,3132	-,3247	30,07	31,31	,6152	,6146	,6045	,6061	61,44	63,58
Desviación típica	1,58106	1,57178	1,14382	1,14382	,41972	,43927	45,006	47,467	,36127	,36210	,39747	,39664	3,810	3,780
Diferencias más extremas														
Absoluta	,105	,091	,215	,215	,293	,291	,408	,405	,237	,237	,282	,281	,069	,080
Positiva	,055	,048	,215	,215	,225	,225	,408	,405	,151	,145	,183	,180	,064	,072
Negativa	-,105	-,091	-,190	-,190	-,293	-,291	-,252	-,255	-,237	-,237	-,282	-,281	-,069	-,080
Z de Kolmogorov-Smirnov	2,045	1,767	4,171	4,171	5,691	5,650	7,902	7,850	4,594	4,590	5,461	5,452	1,347	1,551
Sig. asintót. (bilateral)	,000	,004	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,053	,016

a. La distribución de contraste es la Normal.  
 b. Se han calculado a partir de los datos.

Figura 1. Prueba de normalidad

**Matriz de componentes rotados<sup>a</sup>**

	Componente				
	1	2	3	4	5
Lesld	,327	,031	-,071	,925	,045
Lesli	,339	,044	-,088	,918	,034
ADDd	,831	,094	-,046	,398	,060
ADDi	,831	,094	-,046	,398	,060
Lcild	,131	-,871	,067	-,112	-,063
Lcili	,081	-,849	,011	,027	-,078
Ejed	,050	,774	-,075	,081	-,031
Ejei	,241	,743	,017	-,069	,000
Avld	-,202	-,073	,948	-,064	-,027
Avli	-,204	-,082	,944	-,057	-,030
Avpd	-,778	,061	,465	-,122	-,019
Avpi	-,775	,065	,471	-,126	-,020
Dpc	,046	,030	-,028	,043	,994
Dpl	,048	,056	-,027	,031	,994

Método de extracción: Análisis de componentes principales.

Método de rotación: Normalización Varimax con Kaiser.

a. La rotación ha convergido en 6 iteraciones.

□

Figura 2. Extracción de componentes principales

**Variables introducidas/excluidas<sup>a, b, c, d</sup>**

Paso	Introducidas	Lambdas de Wilks											
		Estadístico	gl1	gl2	gl3	F exacta				F aproximada			
						Estadístico	gl1	gl2	Sig.	Estadístico	gl1	gl2	Sig.
1	Corrección astigmática	.492	1	4	371,000	95,822	4	371,000	.000				
2	Corrección esférica	.267	2	4	371,000	86,498	8	740,000	.000				
3	Corrección presbita	.189	3	4	371,000					71,317	12	976,574	.000

En cada paso se introduce la variable que minimiza la lambda de Wilks global.

- a. El número máximo de pasos es 10.
- b. La F parcial mínima para entrar es 3.84.
- c. La F parcial máxima para salir es 2.71
- d. El nivel de F, la tolerancia o el VN son insuficientes para continuar los cálculos.

Figura 3. Variables en el modelo

**Matrices de covarianzas**

Diag	Agudeza visual para ver de lejos	Ubicación del Objeto Respecto a la Retina	Corrección presbita	Agudeza visual para visión próxima	Lente cilíndrico	Eje del cilindro
1	Agudeza visual para ver de lejos	Ubicación del Objeto Respecto a la Retina	Corrección presbita	Agudeza visual para visión próxima	Lente cilíndrico	Eje del cilindro
	.483	-.093	.350	-.242	.259	-.133
		-.093	.102	-.272	.154	.056
			.350	-.272	1.023	-.184
				-.242	.154	-.184
					.259	-.026
						.938
						.628
						.180
						.628
						1.792
2	Agudeza visual para ver de lejos	Ubicación del Objeto Respecto a la Retina	Corrección presbita	Agudeza visual para visión próxima	Lente cilíndrico	Eje del cilindro
	.823	-.104	.039	.237	.160	-.135
		-.104	.532	-.142	.206	-.188
			.039	-.142	.800	.195
				.237	.206	.195
					.160	-.108
						1.357
						.383
						.871
						.113
						.383
						.871
3	Agudeza visual para ver de lejos	Ubicación del Objeto Respecto a la Retina	Corrección presbita	Agudeza visual para visión próxima	Lente cilíndrico	Eje del cilindro
	1.342	.286	.006	-.376	-.065	-.182
		.286	.706	-.315	.035	-.068
			.006	-.315	1.070	.045
				-.376	.035	.045
					.065	.051
						-.061
						.132
						1.210
						.530
						.069
						.530
						1.605
4	Agudeza visual para ver de lejos	Ubicación del Objeto Respecto a la Retina	Corrección presbita	Agudeza visual para visión próxima	Lente cilíndrico	Eje del cilindro
	.857	-.129	-.166	.179	-.032	.088
		-.129	.450	-.128	.286	.037
			-.166	-.128	.755	.152
				.179	.286	.152
					-.032	-.007
						.054
						.006
						.000
						.016
						.040
						.000
						.000
						.016
						.016
						.053
5	Agudeza visual para ver de lejos	Ubicación del Objeto Respecto a la Retina	Corrección presbita	Agudeza visual para visión próxima	Lente cilíndrico	Eje del cilindro
	1.386	.129	.145	-.352	-.093	.030
		.129	.611	-.456	.035	.031
			.145	-.456	1.168	-.208
				-.352	.035	-.208
					.093	.042
						.042
						.016
						.018
						.018
						.053
						.053
						.053

Figura 4. Matriz de varianzas-covarianzas combinadas

Modelo lineal: Resultados de la clasificación <sup>a,b,d</sup>									
Diag				Grupo de pertenencia pronosticado					Total
				1	2	3	4	5	
Casos seleccionados	Original	Recuento	1	7	10	5	5	0	27
			2	7	96	4	19	1	127
			3	3	5	47	2	12	69
			4	0	0	0	109	0	109
			5	0	0	0	4	45	49
	%	1	25,9	37,0	18,5	18,5	0,0	100,0	
		2	5,5	75,6	3,1	15,0	,8	100,0	
		3	4,3	7,2	68,1	2,9	17,4	100,0	
		4	0,0	0,0	0,0	100,0	0,0	100,0	
		5	0,0	0,0	0,0	8,2	91,8	100,0	
	Validación cruzada <sup>c</sup>	Recuento	1	7	10	5	5	0	27
			2	8	94	4	20	1	127
			3	4	5	44	2	14	69
			4	0	0	0	109	0	109
			5	0	0	0	4	45	49
%	1	25,9	37,0	18,5	18,5	0,0	100,0		
	2	6,3	74,0	3,1	15,7	,8	100,0		
	3	5,8	7,2	63,8	2,9	20,3	100,0		
	4	0,0	0,0	0,0	100,0	0,0	100,0		
	5	0,0	0,0	0,0	8,2	91,8	100,0		
Casos no seleccionados	Original	Recuento	1	1	2	3	1	0	7
			2	0	37	1	10	0	48
			3	0	4	25	0	3	32
			4	0	0	0	20	0	20
			5	0	0	0	3	8	11
	%	1	14,3	28,6	42,9	14,3	0,0	100,0	
		2	0,0	77,1	2,1	20,8	0,0	100,0	
		3	0,0	12,5	78,1	0,0	9,4	100,0	
		4	0,0	0,0	0,0	100,0	0,0	100,0	
		5	0,0	0,0	0,0	27,3	72,7	100,0	

- a. Clasificados correctamente el 79,8% de los casos agrupados originales seleccionados.
- b. Clasificados correctamente el 77,1% de casos agrupados originales no seleccionados.
- c. La validación cruzada sólo se aplica a los casos del análisis. En la validación cruzada, cada caso se clasifica mediante las funciones
- d. Clasificados correctamente el 78,5% de los casos agrupados validados mediante validación cruzada seleccionados.

Modelo cuadrático: Resultados de la clasificación <sup>a,b</sup>									
Diag				Grupo de pertenencia pronosticado					Total
				1	2	3	4	5	
Casos seleccionados	Original	Recuento	1	16	8	2	1	0	27
			2	8	111	7	1	0	127
			3	5	5	55	1	3	69
			4	0	0	0	109	0	109
			5	0	1	1	3	44	49
	%	1	59,3	29,6	7,4	3,7	0,0	100,0	
		2	6,3	87,4	5,5	,8	0,0	100,0	
		3	7,2	7,2	79,7	1,4	4,3	100,0	
		4	0,0	0,0	0,0	100,0	0,0	100,0	
		5	0,0	2,0	2,0	6,1	89,8	100,0	
Casos no seleccionados	Original	Recuento	1	1	3	3	0	0	7
			2	0	46	0	2	0	48
			3	2	3	27	0	0	32
			4	0	0	0	19	1	20
			5	0	0	0	1	10	11
	%	1	14,3	42,9	42,9	0,0	0,0	100,0	
		2	0,0	95,8	0,0	4,2	0,0	100,0	
		3	6,3	9,4	84,4	0,0	0,0	100,0	
		4	0,0	0,0	0,0	95,0	5,0	100,0	
		5	0,0	0,0	0,0	9,1	90,9	100,0	

- a. Clasificados correctamente el 87,9% de los casos agrupados originales seleccionados.
- b. Clasificados correctamente el 87,3% de casos agrupados originales no seleccionados.

Figura 5. Matrices de clasificación



## Matrices de covarianzas

Diagn		Corrección presbita	Corrección astigmática	Agudeza visual de lejos	Corrección esférica	Distancia interpupilar
1	Corrección presbita	1,210	-,421	,503	-,352	-,262
	Corrección astigmática	-,421	,991	-,390	,048	-,002
	Agudeza visual de lejos	,503	-,390	,566	-,113	-,058
	Corrección esférica	-,352	,048	-,113	,132	,122
	Distancia interpupilar	-,262	-,002	-,058	,122	1,490
2	Corrección presbita	,916	,035	-,096	-,366	-,031
	Corrección astigmática	,035	,747	-,125	,183	-,039
	Agudeza visual de lejos	-,096	-,125	1,024	-,101	-,011
	Corrección esférica	-,366	,183	-,101	,680	-,014
	Distancia interpupilar	-,031	-,039	-,011	-,014	,986
3	Corrección presbita	,893	,181	,255	-,207	-,038
	Corrección astigmática	,181	,913	-,153	,003	,014
	Agudeza visual de lejos	,255	-,153	1,035	,202	,123
	Corrección esférica	-,207	,003	,202	,579	,081
	Distancia interpupilar	-,038	,014	,123	,081	,975
4	Corrección presbita	,642	-,043	-,221	-,218	,107
	Corrección astigmática	-,043	,038	,074	-,013	-,028
	Agudeza visual de lejos	-,221	,074	,952	-,041	-,043
	Corrección esférica	-,218	-,013	-,041	,588	-,078
	Distancia interpupilar	,107	-,028	-,043	-,078	,949
5	Corrección presbita	,869	,009	,237	-,123	,096
	Corrección astigmática	,009	,009	,002	-,014	-,032
	Agudeza visual de lejos	,237	,002	1,120	,216	-,052
	Corrección esférica	-,123	-,014	,216	,322	,088
	Distancia interpupilar	,096	-,032	-,052	,088	,916

Figura 6. Matrices de varianzas-covarianzas por grupo

Frecuencia			Grupo pronosticado por el modelo cuadrático * Grupo pronosticado por el modelo lineal * Diagnóstico					
Diagnóstico			Grupo pronosticado por el modelo lineal					
			1	2	3	4	5	Total
1	Grupo pronosticado por el modelo cuadrático	1	6	6	5	2		19
		2	0	4	0	2		6
		3	0	0	1	0		1
		4	0	0	0	1		1
		Total	6	10	6	5		27
2	Grupo pronosticado por el modelo cuadrático	1	9	3	0	0	0	12
		2	0	94	0	15	1	110
		3	0	0	2	1	0	3
		4	0	0	0	2	0	2
		Total	9	97	2	18	1	127
3	Grupo pronosticado por el modelo cuadrático	1	1	1	5	0	1	8
		2	0	3	1	1	1	6
		3	0	0	43	0	7	50
		4	0	0	0	1	0	1
		5	0	0	0	0	4	4
Total	1	4	49	2	13	69		
4	Grupo pronosticado por el modelo cuadrático	4				109		109
		Total				109		109
5	Grupo pronosticado por el modelo	3				0	1	1
		4				4	0	4
		5				0	44	44
		Total				4	45	49
Total	Grupo pronosticado por el modelo cuadrático	1	16	10	10	2	1	39
		2	0	101	1	18	2	122
		3	0	0	46	1	8	55
		4	0	0	0	117	0	117
		5	0	0	0	0	48	48
Total	16	111	57	138	59	381		

Figura 7. Pacientes pronosticados por cada modelo

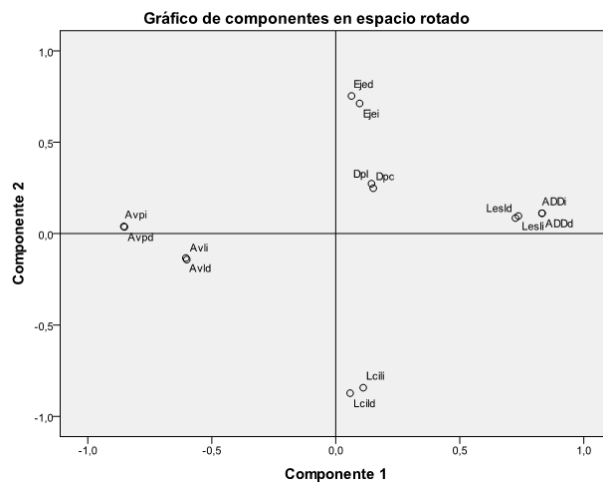


Figura 8. Componentes en espacio rotulado

## Referencias

- [1] <http://www.institutodelavision.com/patologias.php>
- [2] <http://www.quirof.cl/lentes-y-lentes-de-contacto-drjrosasg.htm>
- [3] <http://www.quirof.cl/astigmatismo-drjrosasg.htm>
- [4] HAIR, J. F; ANDERSON, R. E; TATHAM, R. L; BLACK, W. C (1999). *Análisis Multivariable*. 5ª Ed. Prentice-Hall, Madrid.
- [5] PEÑA, D. (2002). *Análisis de datos multivariantes*. Ed. McGraw-Hill, Interamericana de España.
- [6] DÍAZ, R.M.; GONZÁLEZ, A. A.; HENAO, P.A.; DÍAZ, M.M.E. (2012). *Una Introducción al Análisis Estadístico Multivariado Aplicado*. Ediciones Uninorte, Barranquilla.

Para citar este artículo: J. Barraza H. Hernández. 2014, "Modelo estadístico utilizando análisis discriminante para clasificar problemas refractivos oculares en la ciudad de barranquilla". Disponible en Revistas y Publicaciones de la Universidad del Atlántico en <http://investigaciones.uniatlantico.edu.co/revistas/index.php/MATUA>.