

3

ESTADÍSTICA

CARACTERIZACIÓN MULTIVARIANTE DE LOS DELITOS EN PANAMÁ A TRAVÉS DEL MÉTODO HJ-BIPLLOT

Carmen C. Rodríguez¹, Mitzi I. Cubilla¹ y Estelina Ortega-Gómez¹

¹ Universidad de Panamá, Facultad de Ciencias Naturales Exactas y Tecnología, Departamento de Estadística, E-mail: ccrm07@gmail.com

Resumen

La delincuencia, como fenómeno social es multifactorial. Para este estudio hemos seleccionado cinco variables de tipo delictivo, que corresponden a los delitos denunciados con mayor frecuencia en las diferentes regiones del país durante el año 2012. El interés por el comportamiento delincuencia en Panamá, nos ha llevado a realizar este estudio. Los objetivos fueron: Ver la utilidad del método Biplot en el análisis de datos sobre delitos. Determinar si existen diferencias entre las provincias con respecto a los tipos de delitos que se cometen; Explicar la formación de conglomerados, según las variables que los caracterizan. Se aplicó en método HJ-Biplot, el cual es una herramienta multivariante potente en el análisis de datos. Los resultados muestran cuatro grupos o clúster.

Abstract

Crime as a social phenomenon is multifactorial. For this study we selected five variables of type of crime, which correspond to the most frequently reported crimes in different regions of the country during 2012. The interest for delinquent behavior in Panama, has led us to conduct this study. The objectives were: View the usefulness of Biplot method in analyzing crime data. Determine if there are differences between provinces with respect to the types of crimes committed; explain the formation of clusters, according to the variables that characterize them. Was applied HJ-Biplot method, which is a powerful tool for analyzing multivariate data. The results show four or cluster groups.

Keywords: Biplot, Crime, Panama, Multivariate Analysis.

Citación: C.C. Rodríguez, M.I. Cubillas y E. Ortega-Gómez. 2014. Caracterización Multivariante de los Delitos en Panamá a Través del Método HJ-Biplot. Revista Colón Ciencias, Tecnología y Negocios 1 (2): 18-29

Recibido: 5 de diciembre de 2014 **Aceptado:** 19 de diciembre de 2014 **Publicado:** 31 de diciembre de 2014

Correspondencia al autor: ccrm07@gmail.com (Carmen Cecilia Rodríguez Martínez)

INFORMACIÓN DEL ARTÍCULO

Palabras clave:

Biplot,

Delincuencia,

Panamá,

Análisis Multivariante.

INTRODUCCIÓN

Los delitos son una de las mayores preocupaciones para los ciudadanos y asunto de interés de los gobernantes en Panamá. Desde la invasión a Panamá en el año 1989, el incremento de los hechos delictivos ha ido en aumento. Se señala que en los últimos años las pandillas han ido apoderándose del país, distribuyéndose a nivel nacional. La asociación de los grupos delincuenciales con el crimen organizado ha generado la ocurrencia de delitos violentos sobre todo en la capital del país y la ciudad de Colón (Camacho y Prettel, 2013).

En estas dos décadas los medios de comunicación señalan un aumento del pandillerismo y del tráfico de drogas, entre otros actos delincuenciales, que se han convertido en un problema nacional y que puede llevar a un problema de salud pública (Cueva, 2012).

De acuerdo a un estudio realizado por el Observatorio de Seguridad Ciudadana de Panamá 2013, el 64% de la población panameña considera que la delincuencia ha aumentado. EL 62% de los ciudadanos perciben que el grado de violencia de la delincuencia en las provincias ha aumentado. Con respecto al temor al delito se encontró que el mismo es diferente en las ciudades de Panamá y Colón. En la ciudad de Panamá, 49% de los encuestados, manifestó sentir temor cuando salen de su casa, mientras que en el interior del

país, el 39% de los ciudadanos siente mayor temor al volver en la noche a sus casas. De acuerdo al estudio, la percepción de los panameños a futuro sobre los niveles de delincuencia, es que aumentará (49%). Los encuestados también señalan que la pérdida de valores y el desinterés de los padres hacia sus hijos son, entre otras, las causas de este aumento en la delincuencia.

El interés por el comportamiento delincriminal en Panamá, nos ha llevado a realizar este estudio aplicando la técnica HJ-Biplot.

El término Biplot se le debe a Gabriel (1971), los métodos Biplot son una representación gráfica a baja dimensión de una matriz de datos multivariante, X (I individuos, J variables).

El Biplot es una técnicas factoriales de reducción de dimensionalidad (desde el punto de vista algebraico), por ende se basan en la descomposición en valores y vectores singulares (DVS) de una matriz. Básicamente la diferencia se da en que, en los métodos Biplot se trata de reproducir el dato y a la vez de integra una representación simultanea de individuos y variables.

Con el método HJ-Biplot, (Galindo Villardón, 1986), se demuestra que con la elección adecuada de los marcadores, es posible representar las filas y las columnas simultáneas sobre un mismo sistema de coordenadas, obteniendo una alta calidad de representación tanto para filas como para columnas.

El HJ Biplot se define para una matriz de datos $X_{(n \times p)}$ de rango r ($r \leq \min(n, p)$), como una representación gráfica multivariante mediante marcadores (vectores) g_1, g_2, \dots, g_n para las filas y h_1, h_2, \dots, h_p para las columnas de X , elegidos de forma que ambos marcadores puedan superponerse en el mismo sistema de referencia con máxima calidad de representación.

Los objetivos fueron primeramente, ver la utilidad del método Biplot en el análisis del datos sobre delitos; igualmente el determinar si existen diferencias entre las provincias con respecto a los tipos de delitos que se cometen; y por último explicar la formación de conglomerados, según las variables que los caracterizan.

MATERIALES Y MÉTODOS

Los datos se refieren a la información registrada por el Sistema Integrado de Estadística Criminales (SIEC) y el Instituto Nacional de Estadística y Censo de Panamá, para el año 2012. La matriz de datos está formada por nueve filas (provincias/individuos) y cinco columnas (vectores/delitos).

Para el análisis se aplicó el método HJ-Biplot el cual permite realizar un análisis exploratorio de datos. El objetivo principal de este Biplot, es el de conseguir una interpretación simultánea de las relaciones entre individuos y variables, filas y columnas respectivamente, de la matriz de datos. Desarrollado por Galindo Villardón en 1985 (Galindo, 1986; Galindo y Cuadras, 1986). Corresponde a una extensión de los biplots clásicos introducidos por Gabriel en 1971 (Gabriel, 1971, 1972; Gabriel y Odoroff, 1990).

Al igual que los otros métodos Biplot, el HJ-Biplot, se basa en la descomposición en valores singulares (DVS). Es decir que cualquier matriz real $X_{n \times p}$ de rango r se puede factorizar como el producto de tres matrices de tal forma que:

$$X = U\Lambda V^T \quad \text{con} \quad U^T U = V^T V = I$$

donde:

U es la matriz de vectores propios de XX^T .

V es la matriz de vectores propios de $X^T X$.

Λ es la matriz diagonal de $\lambda_1, \lambda_2, \dots, \lambda_r$, correspondiente a los r valores propios de XX^T o $X^T X$.

Los elementos de X vienen dados por :

$$x_{ij} = \sum_{k=1}^r \sqrt{\lambda_k} u_{ik} v_{jk} \quad \text{para } i= 1, 2, \dots, n, \quad ; \quad j= 1, 2, \dots, p$$

De esta manera y partiendo de la DVS, la elección de marcadores en la dimensión q para las filas y columnas de la matriz X son:

$$\mathbf{J} = \mathbf{U}\mathbf{\Lambda} \text{ y } \mathbf{H} = \mathbf{V}\mathbf{\Lambda}$$

La calidad de representación para las filas y para las columnas de la matriz de datos X es la misma. Los marcadores fila representados por $\mathbf{J}_{n \times r} = \mathbf{U}_{n \times r} \mathbf{\Lambda}_{r \times r}$ y los marcadores columna como $\mathbf{H}_{p \times r} = \mathbf{V}_{p \times r} \mathbf{\Lambda}_{r \times r}$, lo que le da el nombre de **HJ-Biplot**.

Para este tipo de representación se interpretan las relaciones fila-columna a través de los ejes factoriales, es decir que ya no se aproximan los elementos de la matriz original.

Las posiciones de las filas, de las columnas y las relaciones fila-columna a través de las contribuciones relativas del factor al elemento y del elemento al factor, ya que tanto la filas como las columnas tienen la misma calidad de representación, (Galindo Villardón y Cuadras, 1986). Las propiedades del método HJ-Biplot son las siguientes:

- Esta representación proporciona la mejor representación simultánea.
- Los productos escalares de las columnas de la matriz X , coinciden con los productos escalares de los marcadores H , es decir:

$$\mathbf{X}^T \mathbf{X} = (\mathbf{U}\mathbf{\Lambda}\mathbf{V}^T)^T (\mathbf{U}\mathbf{\Lambda}\mathbf{V}^T) = (\mathbf{V}\mathbf{\Lambda})(\mathbf{V}\mathbf{\Lambda})^T = \mathbf{H}\mathbf{H}^T$$

- La longitud al cuadrado de los vectores h_j es proporcional a la varianza de la variable x_j .
- El coseno del ángulo entre dos vectores h_i , h_j representa la correlación entre las variables x_i y x_j
- Los productos escalares de las filas de la matriz X , coinciden con los productos escalares de los marcadores j , es decir:

$$\mathbf{X}\mathbf{X}^T = (\mathbf{U}\mathbf{\Lambda}\mathbf{V}^T)(\mathbf{U}\mathbf{\Lambda}\mathbf{V}^T)^T = (\mathbf{U}\mathbf{\Lambda})(\mathbf{U}\mathbf{\Lambda})^T = \mathbf{J}\mathbf{J}^T$$

- La distancia Euclídea entre dos filas de la matriz X, coincide con la distancia Euclídea entre los marcadores j del HJ-Biplot.
- Los marcadores para las filas coinciden con las coordenadas de los individuos en el espacio de las componentes principales de las variables.
- Los marcadores para las columnas coinciden con las coordenadas de las variables en el espacio de las componentes de las filas. Si una variable toma un valor preponderante para un individuo, el punto que representa a esa variable estará próximo al punto que representa al individuo.
- La proximidad entre los individuos se interpreta en términos de similitud. Cuando más distantes aparecen los puntos que representan a los marcadores columna del centro de gravedad, mayor variabilidad habrá en el estudio.
- Cuando menor sea el ángulo que forman los vectores que unen los puntos que representan a dos variables con el centro de gravedad, más correlacionada estarán las variables. La calidad de representación para las filas y columnas es la misma y viene expresada por:

$$\left[\frac{\sum_{i=1}^d \lambda_i^2}{\sum_{i=1}^r \lambda_i^2} \right] * 100$$

- El ángulo que forma una variable con un eje principal representa la importancia que tiene esa variable en el poder discriminante del eje.

Como el HJ-Biplot realiza una representación simultánea tanto de filas como de columnas, también recibe el nombre de CRMP-Biplot-Row Column Metric Preserving-, (Cárdenas *et al.*, 2007). Se utilizó, para el análisis de los datos, el programa MultBiplot (Vicente Villardón, 2010).

RESULTADOS

Como lo indica Galindo (1986), varias medidas son esenciales para una correcta implementación del algoritmo HJ-Biplot. Los valores propios y la varianza explicada (Cuadro 1), en especial, y la contribución relativa del factor al elemento (Cuadro 2). Con esta información es posible detectar las variables responsables de la posición de los ejes, de tal forma de obtener la configuración de ellos. El primer eje presenta un valor de absorción un poco más alto. La absorción de inercia en los dos primeros ejes es de un 89.66. El eje 1 es dominante ya que tiene un 66.28% de absorción de inercia.

Cuadro 1. Valores Propios y Varianza Explicada

Ejes	Valor Propio	Varianza explicada	Varianza Acumulada
1	26.51	66.28	66.28
2	9.35	23.38	89.66
3	2.85	7.13	96.79
4	1.26	3.14	99.93
5	0.03	0.07	100.00

Se han retenido dos ejes pues se consigue una inercia acumulada de 89.66%, suficiente para caracterizar la actividad delictiva de las provincias del país en relación a las variables (tasas de delitos) que fueron consideradas (Cuadro 1). Atendiendo a las contribuciones del factor al elemento para las columnas (Cuadro 2), se observa que todas las variables han de interpretarse en el primer plano factorial, plano 1-2. En cuanto a las filas, solamente una no ha quedado bien recogida en los tres primeros ejes, es esta la Provincia de Chiriquí.

Cuadro 2. Calidad de Representación para las Columnas

Variable	Eje 1	Eje2	Eje 3
Tasa de Homicidios	818	48	83
Tasa de Hurtos	470	454	4
Tasa de Robos	987	2	4
Tasa de lesiones personales	752	53	194
Tasa delitos en contra del orden familiar	288	612	71

En este estudio se ha analizado la distribución de las provincias del país en función de los las tasas de delitos registradas mediante el método HJ-Biplot que permite realizar una representación gráfica de datos multivariantes en el que filas y columnas pueden ser superpuestas en un mismo sistema de referencia con máxima calidad de representación. Se pone de manifiesto la relación entre variables, observándose una correlación directa entre las tasas de homicidio, tasas de robo y tasa de lesiones personales, es decir que la tasa de homicidio está asociada con las tasas de robo y lesiones personales. El estudio permite observar que las características analizadas son, en parte, dependientes de cada provincia. El análisis de los datos ha puesto de manifiesto que la provincia de Panamá presenta altas tasas de delitos, como homicidios y robos (Figura 1).

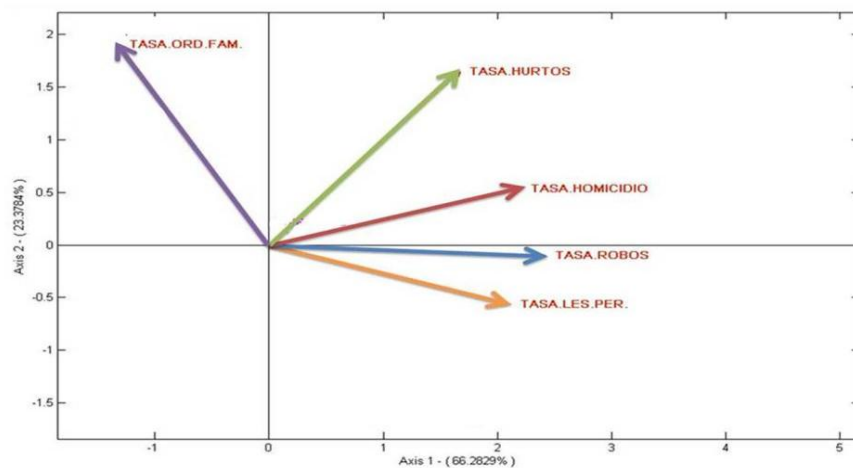


Figura 1. Representación Biplot de los tipos de delitos

Las mayores tasas de delitos contra el Orden Jurídico, Familiar y del Estado Civil se dan para la provincia de Los Santos y le sigue la provincia de Herrera. En el caso de Lesiones Personales la provincia de Colón es la que presenta mayores tasas, se aprecian diferencias entre las provincias de Panamá con el resto de las provincias. Las provincias de Coclé, Veraguas y Darién presentan similitud en sus comportamientos e igualmente muy bajas tasas con respecto a los delitos estudiados. La provincia de Bocas del Toro presenta bajas tasa de delito contra el Orden Jurídico, Familiar y del Estado Civil. La provincia de Chiriquí no presentó buena calidad de representación en ninguno de los planos estudiados.

Cuadro 3. Conglomerados y calidades de representación en el plano 1-2.

Cluster	N	CLR-1	CLR-2	CLR1-2
1	2	37.66	61.26	98.91
2	5	69.75	25.79	95.55
3	1	43.78	35.15	78.94
4	1	94.91	3.14	98.05

A través de las coordenadas Biplot se han calculado los clusters (método K-means, distancia euclídea). Se observa en el gráfico factorial (Figura 2) que los centros forman conglomerados. Las calidades de representación para cada conglomerado en el primer plano factorial se exponen en el Cuadro 3.

Se definen cuatro grupos o clúster: (1) Provincias de Darién, Coclé y Veraguas con muy bajas tasas de los delitos estudiados; (2) Provincias de Bocas del Toro, Herrera y Los Santos con bajas, medias y altas tasas de delitos contra el Orden Jurídico, Familiar y del Estado Civil,

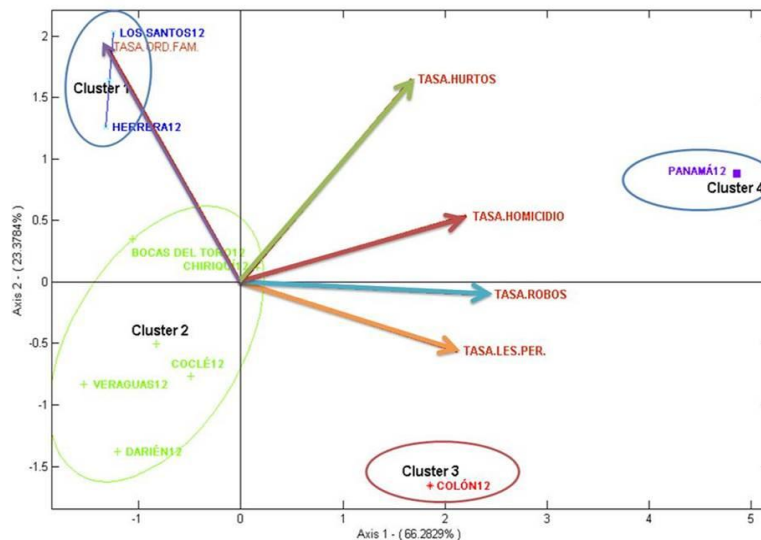


Figura 2 Cluster y calidades de representación en el plano 1-2

respectivamente; (3) Panamá con altas tasas de homicidio y robos; y (4) La provincia de Colón con altas tasas de lesiones personales. Puede decirse no se muestra relación entre las tasas de homicidios y las tasas de delito contra el Orden Jurídico, Familiar y del Estado Civil; al igual que entre las tasas de hurtos y lesiones personales.

CONCLUSION

El gráfico factorial que se da como resultado del análisis HJ-Biplot, confirma la información generada de los comportamientos delictivos en Panamá. Con este trabajo se ha mostrado la manera en que el HJ-Biplot (Galindo, 1985, 1986), una técnica que no ha sido muy utilizada en nuestro ámbito, puede ser una herramienta potente para el análisis de datos delictivos.

REFERENCIAS BIBLIOGRAFICAS

ABRIL DONOSO, M., y A. TUPIZA ALDAZ. 2009. **Delitos contra vehículos: el caso de Quito y Guayaquil**. Boletín Ciudad Segura, 32. Quito, Ecuador.

CAMACHO, A. y A. AMARIS. 2013. **Situación de la seguridad pública en Panamá**. Anuario 2013 de la seguridad regional en América Latina y el Caribe.

CÁRDENAS O., P. GALINDO-VILLARDÓN y J.L. VICENTE-VILLARDÓN. 2001. Estimación máximo verosímil en aproximación biplot. **Revista Venezolana de Análisis de Coyuntura**. 1, pp 333-347.

CÁRDENAS O., P. GALINDO-VILLARDÓN y J.L.VICENTE-VILLARDÓN. 2007. Los Métodos Biplot: Evolución y Aplicaciones. **Revista Venezolana de Análisis de Coyuntura**. 13(1), pp 279-303.

GABRIEL, K.R. 1971. The biplot graphic of matrices with application to principal component analysis, *Biometrika*, 58.

GABRIEL, K.R. y C.L. ODOROFF. 1990. Biplots in biomedical research, *Statistics in Medicine*, 9(5).

GALINDO, M.P. 1986. Una alternativa de representación simultánea: HJ-Bplot. *Qüestió*, 10(1), 13-23.

GALINDO, M.P. y C.M CUADRAS. 1986. **Una extensión del método Biplot y su relación con otras técnicas**. *Publicaciones de Bioestadística y Biomatemática*. Universidad de Barcelona, 17.

Informe de Criminalidad 2012. **Sistema Integrado de Estadísticas Criminales**. Ministerio de Seguridad Pública. Panamá.

IV Informe de Seguridad Ciudadana. Cámara de Comercio, Industrias y Agricultura de

Panamá. Observatorio de Seguridad Ciudadana. 2013.

MEJÍA ACOSTA, M. A., y J. RAMIREZ FIGUEROA. 2009. **Análisis estadístico multivariado de la criminalidad en la ciudad de Guayaquil**. Disponible en el sitio Web: <http://www.dspace.espol.edu.ec/bitstream/123456789/2209/1/4366.pdf> . (Ultimo acceso: 20/agosto/2014).

TREJO, A. C. A., y MSC, J. T. C. C. **La delincuencia en la ciudad de guayaquil, un análisis espacial de su distribución por delito**. Disponible en el sitio Web: <http://xn--caribea-9za.eumed.net/delincuencia-guayaquil/>. (Ultimo acceso: 20/agosto/2014).

VARELA, F., y H. SCHWADERER. 2010. **Determinantes del temor al delito en Chile**. Santiago, Chile: fundación paz ciudadana.

VICENTE VILLARDÓN, J.L. 2010. **MULTBILOT: A package for Multivariate Analysis using Biplots**. Departamento de Estadística. Universidad de Salamanca. Disponible en el sitio Web: <http://biplot.usal.es/ClassicalBiplot/index.html>.