

# “ANÁLISIS ESTADÍSTICO MULTIVARIADO DE LA CRIMINALIDAD EN LA CIUDAD DE GUAYAQUIL”

Mariana Antonieta Mejía Acosta<sup>1</sup>, Jhon Ramírez Figueroa<sup>2</sup>

<sup>1</sup>Ingeniera en Estadística Informática 2002

<sup>2</sup>Director de Tesis. Matemático especializado en la Escuela Politécnica Nacional año 1996, profesor de la ESPOL desde 1996

## RESUMEN

El presente trabajo trata de un Análisis Multivariante de la criminalidad en la ciudad de Guayaquil, lo cual nos ha permitido verificar el alto índice de violencia en cuanto a crímenes existen en la ciudad y las razones por las que más influyen para que dichos actos se cometan.

En el Capítulo 1 se presenta la evolución histórica de la criminalística, definición y descripción de la criminalística; También descripción de la ciudad de Guayaquil por parroquias. En el Capítulo 2 se presenta el Marco Teórico que se siguió para efectuar nuestro análisis, descripción de variables.

En el Capítulo 3 se presenta un Análisis Multivariado realizado a cada una de las variables que pertenecen a las causas a los lugares y la hora en que se dan este tipo de delito. Finalmente las Conclusiones y Recomendaciones basándonos en el análisis realizado y sugerencias de acuerdo a la información analizada.

## INTRODUCCION

El trabajo a presentarse trata de un análisis estadístico multivariado de la criminalística en la ciudad de Guayaquil, en donde la unidad OID en unión con la PJ hacen referencia a éste tipo de anomalías; esto es lo que respecta a la ciudad del Guayas, pero el presente trabajo solo hace referencia a la ciudad de Guayaquil.

Con los adelantos modernos, la Policía se ha transformado cambiando el sistema empírico de la adivinación de las personas peligrosas y de los delincuentes, por el conocimiento científico y el análisis de todos los fenómenos del delito, mediante el sometimiento de la investigación a una técnica por un sistema racional y científico debidamente sistematizado y que hoy se practica en todos los pueblos civilizados.

La Criminalística, con sus disciplinas científicas, auxilia técnicamente a la Función Judicial, con dictámenes periciales, reconstrucción de los hechos, terceros peritos en discordia y en otras aclaraciones sobre cuestiones criminalísticas, que los jueces, Magistrados y Ministros Jueces, deseen aclarar o complementar en pruebas supervenientes; constituyéndose la criminalística como soporte esencial en los dictámenes Judiciales.

## CONTENIDO

### CAPITULO 1

#### LA EVOLUCIÓN HISTÓRICA DE LA CRIMINALÍSTICA

El objeto de la criminalística es analizar elementos para la investigación y buscar la relación con posibles delitos, esto se basa en la tecnificación de los lamentos, por el cual, se pasa por un primer principio que es el hombre, de la persona humana que utilizó sus manos como herramienta.

#### DEFINICIÓN

Se suele denominarse “criminalística” a la disciplina que tiene por finalidad el descubrimiento de los delitos. En la práctica policial y judicial no es suficiente saber que un hecho punible ha tenido ejecución sino que se necesita averiguar., como donde, cuando y quien lo realizó a efecto de poder imponer la adecuada sanción penal. En la técnica moderna, la criminalística se vale de ciencias y artes diversa, tales como : la física, la química, la medicina legal, la antropometría, la fotografía, la dactiloscopia, la balística, etc, que permiten en cada caso determinar el valor probatorio de los rastros e indicios que han sido advertidos.

## Descripción de la Ciudad de Guayaquil

El cantón de Guayaquil es la cabecera municipal de la Provincia del Guayas, es la ciudad más grande del país, cuenta con aproximadamente 2'100.000 habitantes. Predomina el clima cálido tropical o Tropical húmedo con una temperatura de 25° C; y en épocas de calor va desde 36° C a 38° C. Promedio anual. Además es uno de los puertos más importantes del país y del Océano Pacífico por esa razón llamada la “Perla del Pacífico”.

Es una ciudad de contrastes, pulmón económico, financiero, comercial, pasando principales productos de exportación e importación, generando riquezas, divisas para el desarrollo del Ecuador y así también con un alto índice delincencial.

Figura 1

### Mapa geográfico de la ciudad de Guayaquil



### Descripción de la Ciudad de Guayaquil por parroquias

La ciudad de Guayaquil, está conformada por diez y seis (16) parroquias urbanas y por cinco (5) parroquias rurales.

Las parroquias urbanas son las siguientes:

- |                 |              |                    |              |
|-----------------|--------------|--------------------|--------------|
| * Ayacucho      | * Bolívar    | * Carbo            | * F. Cordero |
| * García Moreno | * Letamendi  | * Nueve de octubre | * Olmedo     |
| * Roca          | * Rocafuerte | * Sucre            | * Tarqui     |
| * Urdaneta      | * Ximena     | * Chongón          | * Pascuales  |

Las parroquias rurales son las siguientes:

- |                |            |           |        |
|----------------|------------|-----------|--------|
| * Gómez Rendón | * El Morro | * Posorja | * Puná |
| * Tenguel      |            |           |        |

### CAUSAS

La Policía Judicial es la aplicación práctica de la Criminología. Estudia siempre en concreto los siguientes puntos:

- Las causas de la criminalidad respecto de cada delincuente, y en caso concreto si son de orden interno o antropológico o de orden externo o sociológico;
- Las formas de la criminalidad con ocasión de determinado delito, psicológicas o de orden interno, externo o sociales;

- c) La represión considerada en sus relaciones con la acción criminal y la acción penal, a través de la responsabilidad en sus causas particulares, fase que comprende la investigación del delito y el tratamiento de los delincuentes;
- d) La prevención estudiada por el aspecto del peligro social, de las medidas de defensa preventiva y de seguridad social, y en cuanto a los detalles pertinentes a la investigación de los lugares y de los peligrosos, la aplicación de las medidas, la vigilancia especial de la Policía y la tutela penal.

A los métodos científicos que han venido a llamarse Policía Judicial se debe la creación de los Laboratorios de Policía o de Investigación Criminal que hoy existe en varias capitales del mundo, mediante los cuales se organiza la investigación de una manera metódica, transformando el sistema empírico de la adivinación o intuición de la Policía, en el sistema científico de la comprobación de laboratorio.

### **Situación actual**

Los tiempos violentos que en la actualidad atravesamos, reflejan claramente el alto grado de inseguridad que se vive y que afecta no solo al Ecuador sino a toda América Latina. Al interior del país podemos presenciar que la delincuencia ha tomado cuerpo y se ha concentrado en las grandes ciudades, esto se debe a la situación de crisis que afectan a la sociedad ecuatoriana en todos sus sectores.

Basados en la información, datos estadísticos, etc, podemos decir que en los últimos tiempos se nota un incremento del porcentaje de detenidos y de igual forma un aumento del porcentaje de denuncias, los unos y los otros en los diferentes ambientes delincuenciales.

## **CAPITULO 2**

### **MARCO TEÓRICO DE COMPONENTES PRINCIPALES**

El análisis estadístico multivariante recoge varias técnicas de análisis de datos, tipología, análisis discriminante, análisis factorial de correspondencias, clasificación. etc., todas ellas tratan de describir, clasificar y clarificar los datos.

Debido a que la mayoría de las variables son categóricas se va a realizar un Análisis de Componentes Principales No Lineal.

En este análisis primero se reponderan o recodifican las variables aleatorias mediante un algoritmo no lineal y luego estas variables reponderadas son analizadas mediante un ACP clásico. En este capítulo describiremos tanto el ACP clásico como el método de reponderación.

### **Objetivos del Análisis de Componentes Principales**

1. reducir la dimensionalidad del conjunto de datos
2. identificar nuevas variables subyacentes

Las nuevas variables (componentes principales) presentan un orden decreciente de importancia.

1. no están correlacionadas
2. la primera componente principal explica tanto de la variabilidad de los datos como sea posible.
3. cada componente subsiguiente tome en cuenta de la variabilidad restante como sea posible.

Un análisis de componentes principales frecuentemente revela la relación de que no fueron previamente supuesto y con relación a eso permiten interpretaciones que no deberían resultar ordinarias. Este análisis es un medio para llegar al resultado final de los mismos en el presente trabajo.

### **Métodos de ponderación no lineales**

En esta parte se describe un tratamiento no estándar del análisis multivariante, el denominado análisis no lineal multivariante, el cual no solo trabaja con variables métricas sino también con variables categóricas o nominales. Este análisis se divide en:

- Codificación de los datos categóricos.
- Concepto de homogeneidad.
- Un algoritmo que maximiza la homogeneidad de un grupo de variables.

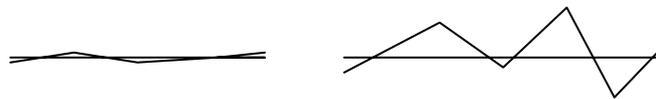
### Codificación de datos categóricos

Sea un conjunto de  $n$  objetos o individuos. Una variable  $h_j$  hace corresponder al conjunto de los individuos un conjunto finito de  $k_j$  categorías, este conjunto de categorías se denomina el rango de  $h_j$ . Vamos a asumir que existe un número finito de  $m$  variables  $h_j$  ( $j=1, \dots, m$ ). El producto cartesiano de todas estas categorías se denomina rango multivariante, sus elementos son todas las posibles combinaciones de las  $m$  categorías, y se denominan perfiles. La matriz de datos  $H$  es una matriz  $n \times m$  con elementos  $h_{ij}$  que nos indican la categoría de la variable  $h_j$  para el individuo  $i$ . Estos elementos no necesariamente son números.

### Homogeneidad de variables

La idea de homogeneidad es cercana a la idea de que diferentes variables pueden medir la “misma cosa”, si esto fuese cierto, la matriz de datos, teniendo en cuenta que las variables son estandarizadas, podrían dar valores idénticos en cada fila o, si dibujamos las observaciones como perfiles, cada uno de ellos sería una recta (Figura 2.1). Si se pone la idea de tener la misma medición de variables no fuese muy exacta porque las variables miden la “misma cosa” pero con un error aleatorio, las filas de la matriz de datos pueden tener elementos que varían un poco. Un gráfico de los perfiles sería una línea quebrada.

**Figura 2**  
**Gráfico de perfiles de la homogeneidad de variables**



Si reemplazamos tales perfiles por líneas rectas (la línea media) tendremos una pérdida de información. Las variables son homogéneas si la pérdida de información es relativamente pequeña.

### Algoritmo de puntajes normalizados

En este algoritmo los puntajes de los individuos se sujetan a la restricción  $\mathbf{x}^t \mathbf{x} = 1$ . El algoritmo requiere un vector de pesos inicial y arbitrario  $\mathbf{a} \neq 0$ :

$$\text{Actualización de puntajes: } \mathbf{x}^0 \leftarrow \frac{1}{m} H \mathbf{a}^0 \quad \text{Paso 1}$$

$$\text{Normalización: } \mathbf{x}^+ \leftarrow \frac{\mathbf{x}^0}{\|\mathbf{x}^0\|} \quad \text{Paso 2}$$

$$\text{Actualización de pesos: } \mathbf{a}^+ \leftarrow H^t \mathbf{x}^+ \quad \text{Paso 3}$$

Test de convergencia: Regresar a (paso 1), hacer que  $\mathbf{a}^0 \leftarrow \mathbf{a}^+$  mientras los valores de  $\mathbf{x}^+$  y  $\mathbf{a}^+$  no estén suficientemente estabilizados (de acuerdo a algún criterio de exactitud previamente establecido).

### Descripción del algoritmo:

Corresponde al mínimo condicional no restringido de la función de pérdida (2.1) para un  $\mathbf{a}^0$  fijo.

Notemos que  $H \mathbf{a}^0 / m$  es un vector que contiene los promedios de las filas reescaladas por  $a_j^0$ . Los puntajes actualizados  $\mathbf{x}_0$  por consiguiente también minimizan la pérdida relativa  $W/T$  para  $H$  reescalada con pesos fijos  $\mathbf{a}_0$ .

Es la proyección de  $\mathbf{x}_0$  sobre la hiper-esfera de todos los “ $\mathbf{x}$ ” normalizados, lo que transfiere la restricción de minimización a una región factible (la región que contiene todas las soluciones que satisfacen la restricción).

Corresponde al mínimo condicional no restringido de la función de pérdida (2.1) para un “x+” fijo. Puesto que “x+” y las columnas de H son centradas y normalizadas, “a+” es un vector de correlaciones.

El algoritmo converge monótonamente, puesto que los pasos 1 y 2 conjuntamente y el paso 3, siempre dan un pequeño valor de la función de pérdida, la cual está acotada inferiormente por 0.

### CAPITULO 3 ANÁLISIS ESTADÍSTICO UNIVARIADO INTRODUCCION

En el siguiente análisis Univariado se presenta una Estadística Descriptiva cuyo objetivo es ilustrar las estadísticas básicas como: Porcentaje, Mínimo, Máximo, Media, Mediana, Varianza, Desviación Estándar, Sesgo, Kurtosis. También se presenta ilustraciones gráficas como: Histograma de Frecuencias, Diagrama de Cajas.

**Variable 1: Fecha de muerte.-** es la fecha que la víctima muere.

**Variable 2: Género.-** aquí se describe si la víctima es de género masculino o femenino.

**Variable 3: Edad.-** la edad de la víctima.

**Variable 4: Clase de víctima .-** aquí se describe si la víctima es una persona civil, guardia, policía o es una persona que pertenece a una de las ramas Armadas del Ecuador, que se lo denota como un milita

**Variable 5: Estado civil.-** aquí se describe si la víctima estuvo casado o permanecía soltero.

**Variable 6: Lugar.-** es la dirección donde ocurrió el delito, que está descrito por parroquias.

**Variable 7: Hora.-** es la hora que ocurrió el delito.

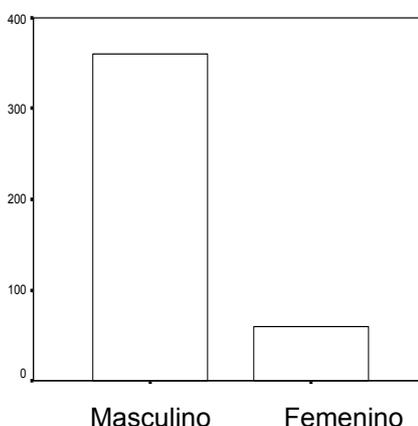
**Variable 8 :Arma.-** ésta variable describe que tipo de arma utilizó el criminal para cometer el delito, que puede ser un arma blanca, un arma de fuego, un objeto contundente, a golpes, puede ser por medio de una mala práctica médica, asfixiado, por medio de algún estupefaciente, con una soga.

**Variable 9: Causa.-** es la causa por la que el criminal ha cometido tal acto en contra de la víctima, este tipo de variable puede ser por venganza, por robo, por bromista, por no pagar, caso fortuito, acribillado , por celos, por confusión, por discusión, por defender a pariente, por alguna balacera.

**TABLA I**  
**Estadística descriptiva para el género de la víctima**

Frecuencia	
• Masculino	85,9%
• Femenino	14,1%

**Figura 3**  
**Histograma de Frecuencias del género de la víctima**



Aproximadamente 350 personas, fueron de género masculino y menos de 100 personas, fueron de género femenino, es la representación que se observa en la Figura 3.1. A continuación se verá el Análisis Univariado de la variable:

Edad de la víctima:

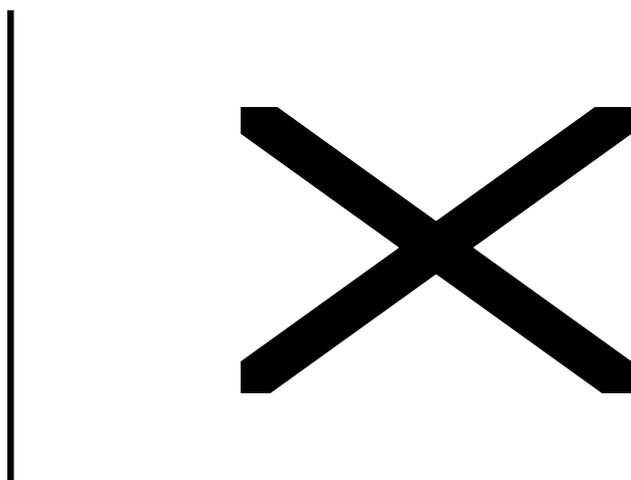
**TABLA II**  
**Estadística descriptiva para la edad de la víctima**

Media	30,16
Mediana	27,00
Desviación típica	12,04
Varianza	144,87
Mínimo	6
Máximo	90

Vemos que la media de edad para las personas víctimas de los crímenes es de 30 años con un mínimo de edad de 06 años, un máximo de edad de 90 años.

Vemos que la mediana de edad para las personas víctimas de los crímenes es de 27 años.

**Figura 4**  
**Histograma de frecuencias de la edad de la víctima**



Claramente observamos las frecuencias en cuanto a la edad de las personas que son víctimas de los maleantes al cometer los crímenes.

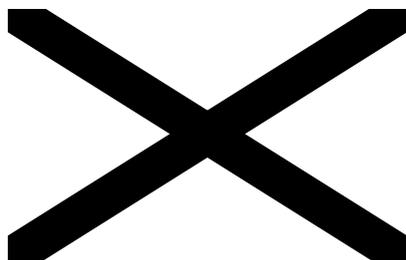
Variable: lugar del crimen

**TABLA III**  
**Estadística descriptiva para el lugar del crimen**

	<b>Frecuencia</b>	<b>Porcentaje</b>
9 de octubre	6	1,4%
ayacucho	8	1,9%
bolivar	1	0,2%
carbo	3	0,7%
chongon	7	1,7%
el morro	7	1,7%
febres cordero	66	15,8%
garcia moreno	11	2,6%
gomez rendon	6	1,4%
letamendi	14	3,3%
olmedo	4	1,0%
pascuales	27	6,4%
posorja	9	2,1%
puna	6	1,4%
roca	5	1,2%
rocafuerte	5	1,2%
sucre	7	1,7%
tarqui	110	<b>26,3%</b>
tenguel	5	1,2%
urdaneta	10	2,4%
ximena	102	<b>24,3%</b>

Se observan las dos principales parroquias donde se dan muchos crímenes que son la parroquia tarqui con un 26,3% y la ximena con un 24,3%. También se presentan crímenes pero en menor cantidad en la parroquia bolívar con un 0,2%.

**Figura 5**  
**Histograma de frecuencias del lugar donde existe índice criminal**



1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21

**Tabla IV**  
**Codificación de las parroquias por número**

1	9 de octubre	11	olmedo
2	Ayacucho	12	Pascuales
3	Bolívar	13	Posorja
4	Carbo	14	Puná
5	Chongón	15	Roca
6	El morro	16	Rocafuerte
7	Febres cordero	17	Sucre
8	García moreno	18	Tarqui
9	Gómez rendón	19	Tenguel
10	Letamendi	20	Urdaneta
		21	Ximena

Se confirma claramente la afluencia que tiene la parroquia tarqui para que los maleantes cometan crímenes por esas zonas.

Variable: arma de ataque

**TABLA V**  
**Estadística descriptiva para la variable arma de ataque**

	Frecuencia	Porcentaje
A golpes	8	1,9
Arma blanca	141	33,7
Asfixiado	1	,2
De fuego	242	57,8
Drogado	2	,5
Objeto contundente	17	4,1
Mala práctica médica	2	,5
Soga	6	1,4

Por medio de estadística descriptiva observamos que la principal arma de ataque preferida por los maleantes para cometer sus fechorías es el arma de fuego con un 57,8%; se tiene también que el arma blanca es otra de las herramientas preferidas por los maleantes para cometer crímenes y se lo puede verificar por medio de la Figura 5.

## **ANÁLISIS MULTIVARIADO DE LA CRIMINALÍSTICA EN LA CIUDAD DE GUAYAQUIL**

Se realizará un análisis Multivariado de las variables que se han encontrado en los datos proporcionados. Los datos proporcionados para nuestro estudio fueron tomados desde enero del 2002 hasta febrero del 2003, y en todas las partes que respecta solo la ciudad de Guayaquil, en donde se ha sectorizado por parroquias.

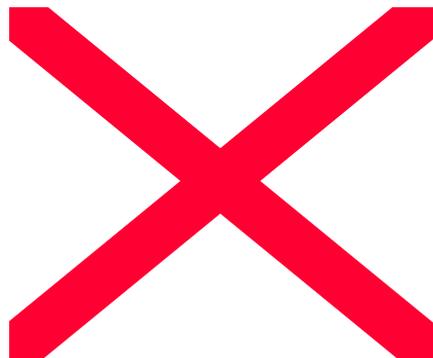
El Análisis de Componentes Principales No Lineal, arroja los resultados numéricos no en forma algebraica como lo hacen los métodos comunes, tampoco especifica el criterio de información explicado a través de la varianza como el Análisis de Componentes Principales clásico.

**TABLA VI**  
**Varianza explicada y perdidas de los resultados**  
**según CATPA**

Dimensión	Varianza Explicada (%)		Pérdida	
1	19.263	Total	Coordenadas de centroide	Restricción del centroide a las coordenadas del vector
2	9.689			
Total	28,952164	61,047836	59,061798	1,986038

Según la tabla XI muestra que con dos dimensiones se obtiene una varianza explicada del 28.95%, y la perdida total, producto del promedio de las puntuaciones de los objetos (coordenadas del centroide) más la restricción de las coordenadas del centroide a las coordenadas del vector (variables categóricas simples).

**Figura 6**  
**Representación de las cargas de las Variables**  
**en dos dimensiones según CATPA**



Observamos en la Figura 3.12 que las líneas relativamente largas aportan mayor información explicando más la variabilidad de todas las variables, mientras que las líneas cortas no aportan mayor información, ya que no presentan una variación significativa.

Observamos tres grupos de variables en la Figura 3.12 la variable ARMATAQ (Arma de ataque) se encuentra en dirección opuesta a las variables HORA (hora del crimen), y con la variable CAUSA (causa o motivo del crimen) que forman aproximadamente un ángulo de 180°, lo que indican que se hallan correlacionadas negativamente.

La variable LUGAR se encuentra en dirección opuesta a la variable MES formando casi un ángulo de 180°, lo que indica que están correlacionadas negativamente.

Se observa que el grupo de variables TIPOVICT (tipo de víctima), GENERO, EDAD, ESTADOCIV(estado civil de la persona) tienen un aporte poco significativo.

En efecto esto demuestra que el arma de ataque utilizada dependerá de la hora y de la causa, ya que si se ataca en el día, por ejemplo, entre la mañana hasta la tarde, el arma que utiliza el individuo será diferente a la que utilice en la noche o madrugada; y si se trata de una venganza o por celos, es decir, la causa, el arma de ataque también será diferente.

Se puede afirmar que la causa del crimen es por discusión, entonces su arma de ataque será una de fuego, si el arma de ataque fue por golpes, seguramente como se ve en la explicación de las componentes se habría dado por un arranque de celos.

## CONCLUSIONES

Con la información analizada se verifica que en nuestro medio existe un porcentaje alto de personas víctimas de crímenes que son de género masculino (85.9%), y que su edad tiene una fluctuación desde 22 años hasta los 28 años. En menor cantidad para las personas de género femenino (14.1%) y con fluctuaciones de edad de 22 a 28 años.

Así mismo se aprecia que las personas que pertenecen a una rama de las fuerzas armadas son víctimas de crímenes, pero en mayor porcentaje (88.1%) son personas civiles, es decir, que no pertenezcan a alguna rama de las fuerzas armadas.

El Análisis de Componentes Principales No Lineal, nos proporciona un reducido número de componentes, facilitando la interpretación y análisis de las variables en estudio.

Según el método de CATPA se observa que la varianza explicada es del 28.95%, representado para las dos componentes principales, tomando como base dos dimensiones que es lo recomendado.

La primera componente principal se denomina se denomina “ TIPO DE ATAQUE ”, ya que las variables de mayor significancia o mayor peso están dadas por la CAUSA, ARMA DE ATAQUE y a la HORA, de ahí que se la identifica a la componente.

Para la denominación de la segunda componente se muestra las variables de mayor peso, a saber, MES y LUGAR, donde la denominación para esta componente es “ ESPACIO TIEMPO ”.

Así mismo se observan los tres grupos de variables claramente diferenciados:

Para el primer grupo:

**ARMA DE ATAQUE** correlacionado negativamente con **HORA** y **CAUSA**.

Para el segundo grupo:

**MES** correlacionada negativamente con **LUGAR**

Para el tercer grupo :

**TIPO DE VICTIMA, EDAD, GENERO, ESTADO CIVIL**, las cuales son variables cercanas a cero y no son significativas para nuestro estudio.

Especificando que la variable **ARMA DE ATAQUE** se encuentra correlacionada negativamente con la variable **HORA** y **CAUSA**, formando aproximadamente un ángulo de 180°. Lo que significa que el **ARMA DE ATAQUE** utilizada por los maleantes depende de la **HORA** para cometer el crimen, un arma de fuego será mas utilizada en horas de la madrugada y no en horas de la mañana, aunque si sucede en porcentajes bajos pero los maleantes prefieren horas de la madrugada y utilizando éste tipo de arma. Del mismo modo el **ARMA DE ATAQUE** depende mucho de la **CAUSA** por la que se cometa el crimen, si se trata de un individuo que fue atacado con un cuchillo recibiendo puñaladas por varias partes de cuerpo, seguramente la causa es una venganza.

## BIBLIOGRAFIA

### a) Libro

1. Roberto Albarracín, CRIMINALISTICA, Editorial Policial, 1957, pp. 12-25

### b) Libro

2. Insp. Luis Alfredo Velasco Davila, TEORIA DE LA INVESTIGACION CRIMINAL, 1963, pp. 83-139

### c) Libro con edición

3. Hair Anderson, Tatham Black, MULTIVARIATE DATA ANALYSIS, (5ta. Edición 1994) Mc Graw Hill