

PREDICCIONES BIOCLIMATICAS PARA LA FRECUENCIA DEL ASMA A TRAVES DE GRAFICOS DE MOSAICOS DE TABLAS DE CONTINGENCIA

Vivian Sistachs Vega¹ y Gladys Linares Fleites, Departamento de Matemática Aplicada,
Facultad de Matemática y Computación, Universidad de La Habana

RESUMEN

Las relaciones existentes entre la frecuencia de ocurrencia de la enfermedad ASMA y algunas variables climáticas son formuladas estadísticamente a través de tablas de contingencia con el objetivo final de poder pronosticar la frecuencia de ocurrencia de la enfermedad a través de modelos logitos. Dichos pronósticos pueden representarse gráficamente a través de mosaicos lo que esclarece la comprensión para los médicos y meteorólogos. Ocho gráficos ilustran los pronósticos de la enfermedad ASMA, en niños y adultos, en los cuatro períodos del año 1992 en el que se enmarca el estudio realizado en diferentes hospitales de la Ciudad de la Habana, Cuba.

Key words: loglineal, predicción, tablas de contingencia.

ABSTRACT

The existent relationships between the frequency of occurrence of the illness ASTHMA and some climatic variables are formulated statistically through contingency table with the final objective of being able to predict the frequency of occurrence of the illness through model logitos. This prediction can be represented graphically through mosaics what clarifies the understanding for the doctors and meteorologists. Eight graphics illustrate the prediction of the illness ASTHMA, in children and adults, in the four periods of the year 1992 in the one that the study is framed carried out in different hospitals of the City of the Havana, Cuba.

Key words: loglinear, prediction, contingency tables.

MSC: 62P10

1. INTRODUCCION

Para corroborar el hecho de que la salud y la capacidad de trabajo del hombre dependen en gran medida de las condiciones climáticas de su lugar de residencia, además para poder prever la disponibilidad de recursos ante determinadas situaciones se realizaron búsquedas de pronósticos bioclimáticos a diferentes enfermedades de alta incidencia en nuestra ciudad, en particular la enfermedad ASMA.

Al realizar los pronósticos bioclimáticos existen diferentes enfoques, de hecho V. Sistachs (1996) propone una metodología de trabajo para estos fines, pero no es única y existen otras formas de abordar el problema como por ejemplo a través de series de tiempo o utilizando los gráficos de mosaicos.

Para el pronóstico de la enfermedad ASMA tanto en pacientes adultos como en niños, se tomaron las frecuencias diarias de ocurrencia de la enfermedad durante el año 1992. Los datos provienen de los controles de asistencia al cuerpo de guardia del Departamento de Estadística de los Hospitales Calixto García, Nacional y Emergencia, en el caso de los adultos, y de los Hospitales Pediátricos Centro Habana, Willian Soler y Marianao en el caso de los niños. Todos radican en la Ciudad de La Habana, Cuba.

También se utilizó el meteorograma mensual de la estación meteorológica de Casablanca que contiene la información de 10 variables climáticas a saber:

TMAX: Temperatura máxima.

TMED: Temperatura media.

TMIN: Temperatura mínima.

TVA: Tensión de vapor de agua.

¹E-mail: vivan@matcom.uh.cu

DOA: Densidad de oxígeno de la atmósfera
 HR: Humedad relativa.
 PPP: Presión atmosférica.
 TSS: Tipo de situación sinóptica.
 ELTD: Estado del tiempo diario.

Además con la colaboración de los especialistas del Instituto de Meteorología se definieron los siguientes cuatro periodos del año sobre los que se debían realizar los pronósticos:

- Período I: (INV) Invierno (diciembre, enero, febrero y marzo).
- Período II: (PRI) Transición del invierno al verano (abril y mayo).
- Período III: (VER) Verano (junio, julio, agosto y septiembre).
- Período IV: (OTO) Transición del verano al invierno (octubre y noviembre).

Las variables fueron codificadas solo en dos categorías cada una la primera categoría representa siempre la ocurrencia de bajos resultados de la variable y la segunda categoría la ocurrencia de altos resultados. El pronóstico se realiza haciendo uso de los gráficos de mosaicos.

El trabajo se estructuró en tres partes. En la primera parte se aborda el problema de las relaciones que se obtienen entre la frecuencia de ocurrencia de la enfermedad y las variables climáticas. En la segunda parte se trata el problema de la construcción del mosaico y por último en la parte tres se ilustra, en el caso del Período I para los niños, como se conforma el pronóstico haciendo uso del mosaico.

2. RELACIONES EXISTENTES ENTRE EL ASMA Y LAS VARIABLES CLIMATICAS

Un estudio exploratorio de datos, basados fundamentalmente en Análisis de Correspondencia simple y múltiple, nos permitió establecer relaciones entre la enfermedad ASMA y algunas variables climáticas que se muestran en el cuadro 1. (Ver Linares, G. y Sistachs, V. 1996)

Por ejemplo en el período INV, tanto para el caso de los adultos como para el de los niños, se evidencia relación de la frecuencia de ocurrencia de la enfermedad con la TVA. En el caso del período PRI, en los niños, hay relación de la enfermedad con la DOA. Para el período VER, en los adultos, las relaciones que se obtienen son del TSS con la frecuencia de ocurrencia de la enfermedad. Por último, en el período OTO, en los adultos, hay relación de la enfermedad con el TSS y el ELTD y en los niños la relación de la enfermedad es con TVA.

Con estos resultados se puede realizar una síntesis de cuáles son las variables más importantes. Obsérvese que, tanto en el caso de niños, como en el de los adultos, solo quedan tres variables en cada Período. El cuadro 1 contiene la información sobre cuales son las tablas de contingencia con que se trabaja.

Cuadro1. Resumen de las TC por periodo

	INVIERNO		PRIMAVERA		VERANO		OTOÑO	
ETIQUETA	ADULTO	NIÑO	ADULTO	NIÑO	ADULTO	NIÑO	ADULTO	NIÑO
A	ASMA	ASMAP	ASMA	ASMAP	ASMA	ASMAP	ASMA	ASMAP
B	TVA	TVA	TMED	TMED	DOA	ELTD	ELTD	TVA
C	TSS	TSS	DOA	DOA	TSS	TVA	TSS	DOA

El paso siguiente consiste en ajustar el modelo log-lineal de independencia a esas tablas de contingencia en las que el estimador del valor esperado de las frecuencias observadas es el cociente del producto de los totales marginales sobre el total de la muestra. (Ver A. Agresti 1996)

3. ALGORITMO DE CONSTRUCCION DE GRAFICOS DE MOSAICOS PARA TABLAS DE CONTINGENCIA DE 3 ENTRADAS

Cuando se tiene la tabla de contingencia se puede construir el mosaico asociado a ella. M. Friendly (1994) propone un algoritmo general para poder realizar dicha construcción.

En el caso particular de una tabla de contingencia de orden 3 (es decir IJK) los pasos a seguir serían los siguientes:

La base del área del gráfico de mosaico se divide en partes proporcionales al total marginal de la primera variable, digamos A, es decir proporcional a $n_{i..}$.

Las partes son las bases del rectángulo con una altura proporcional a la frecuencia condicional de la segunda variable, digamos B, es decir, la altura es proporcional a $n_{ij.} / n_{i..}$.

Finalmente el rectángulo IJ se divide verticalmente obteniéndose los nuevos rectángulos con base proporcional a $n_{ijk} / n_{ij.}$.

La construcción culmina con el dibujo y sombreado del gráfico. Esto se hace para un mejor impacto visual, ya que permite observar que los tamaños relativos de los rectángulos indican desviaciones de la independencia; para aumentar este impacto visual del mosaico, se utiliza un sombreado que refleja el tamaño de los residuos (se utilizan los residuos estandarizados).

Las celdas con desviaciones positivas se dibujan con un contorno de líneas continuas, mientras que las desviaciones negativas, su contorno se dibuja con líneas discontinuas. Este tipo de gráfico es más efectivo cuando aparece en colores.

El sombreado se emplea cuando los residuos estandarizados son en valor absoluto mayor ó igual que 2, y se hace más intenso el sombreado cuando es mayor ó igual que 4.

Cuando las Tablas de Contingencia son de orden superior la construcción del gráfico resulta algo engorroso e incluso puede resultar algo difícil su interpretación, además según la literatura consultada en el sistema SAS aparece implementado la construcción de este tipo de gráfico.

4. MOSAICO PARA LA PREDICCIÓN DE LA FRECUENCIA DE OCURRENCIA DEL ASMA EN EL PERIODO I (INV)

Para ilustrar como hacer la predicción empleando mosaicos observe la Tabla 3.1 que muestra la tabla de contingencia asociada a los niños en el período I. Como se planteó en la parte 2 las variables con que se trabajan son: variable A, frecuencia de ocurrencia de la enfermedad ASMA, variable B, tensión de vapor de agua (TVA) y variable C, tipo de situación sinóptica (TSS) cuyas categorías se plantearán más adelante.

La Tabla 3.2 muestra en el modelo de independencia ($[A][B][C]$) cuales son las frecuencias esperadas y los residuos estandarizados.

Tabla 3.1. Tabla de frecuencias observadas en el Período I

B	A	C	
		1	2
1	1	3	17
	2	8	2
2	1	24	45
	2	13	6

Tabla 3.2. Tabla de frecuencias esperadas, y en los paréntesis los residuos estandarizados del modelo de independencia ($[A][B][C]$)

B	A	C	
		1	2
1	1	9.2 (-2.04)	13.4 (0.94)
	2	3 (2.88)	4.4 (-1.13)
2	1	27 (-0.58)	39.4 (0.89)
	2	8.8 (1.42)	12.8 (-1.91)

Con la información anterior puede obtenerse las tablas marginales por filas y columnas y así poder construir el mosaico.

Dados los datos anteriores en la Tabla 3.3 puede observarse las partes proporcionales a los totales marginales y basados en ellos podemos construir el mosaico en el cual la base del rectángulo lo representan los totales marginales de la variable B, la altura de ese rectángulo lo representa la variable A en su proporcionalidad a $n_{ij.}/n_{i..}$ y la tercera variable C interviene dividiendo el rectángulo de forma proporcional a $n_{ijk}/n_{ij.}$ obteniéndose el gráfico el cual se sombrea según la información de la Tabla 3.2.

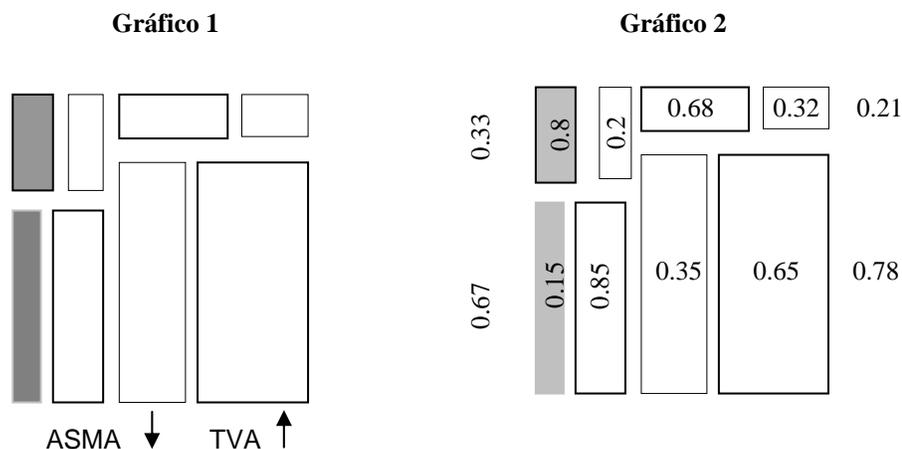
Tabla 3.3. Tabla de totales marginales por filas y sus proporciones

A	B	C		n_{ij}
		1	2	
1	1	0.15	0.85	0.67
	2	0.8	0.2	0.33
2	1	0.35	0.65	0.68
	2	0.68	0.32	0.21

$n_{1..} = 30$
 $n_{2..} = 88$

Para ganar claridad en la construcción e interpretación del mosaico deben observarse los tamaños de los rectángulos que se forman para ver si hay grandes desviaciones; pues de no haberlas entonces estamos ante un modelo de independencia.

Obsérvese que también deben estar claro en un problema las categorías de las variables para la interpretación. En nuestro ejemplo la variable A (ASMA) toma valor 1 cuando la frecuencia de ocurrencia de la enfermedad es baja y 2 cuando es alta, la variable B (TVA) tiene el valor 1 cuando es baja y 2 cuando es alta y en el caso de la variable C (TSS) tiene el valor 1 cuando es típico de verano y hay presencia de anticiclón y el valor 2 cuando es típico de invierno. Los valores de la Tabla 3.3 me permiten construir el Gráfico 1



De ese gráfico resulta entonces evidente que la presencia de una baja frecuencia de ocurrencia de la enfermedad ASMA esta asociada a un comportamiento alto de la TVA y la TSS categoría 1 que significa situación típica del verano e influencia de anticiclón.

Los resultados que se obtienen a través de esta técnica gráfica son similares a los resultados obtenidos por V. Sistachs (1996) utilizando técnicas analíticas con el modelo logito, donde se obtiene que el logito crece al aumentar la categoría de la TVA y se puede concluir, que una disminución de la frecuencia de ocurrencia de la enfermedad ASMA es más probable cuando hay una alta TVA.

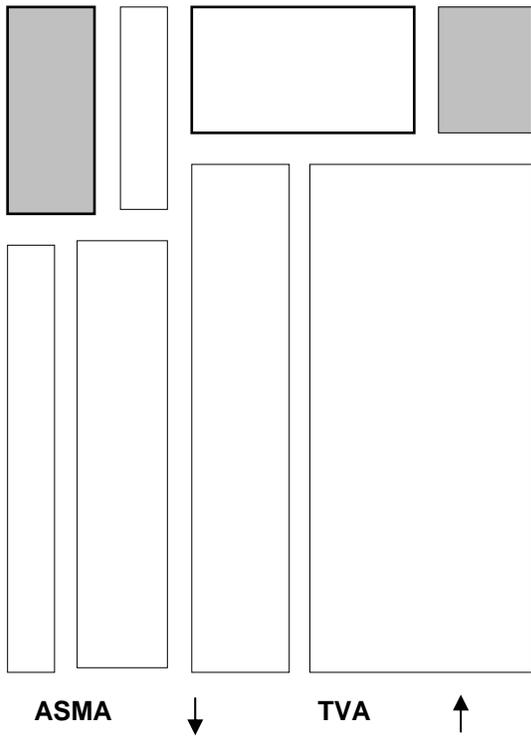
Si se observa el gráfico de la PRIMAVERA ADULTOS y el de VERANO NIÑOS, que aparece en el anexo 2, puede decirse que hay independencia en esas tablas debido a que no hay desviaciones notables como en los otros gráficos.

5. CONCLUSIONES

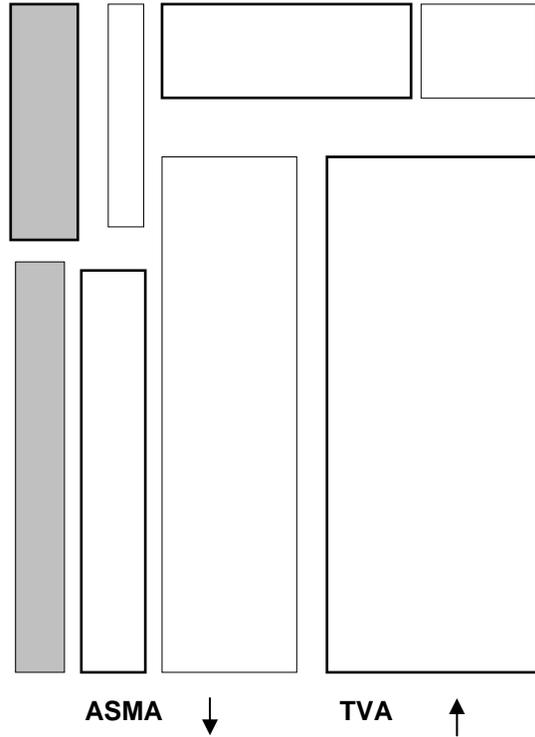
Los gráficos de mosaico brindan la posibilidad de exponer resultados de predicción de forma más evidente y asequible para el especialista (usuario no estadístico) que a través de técnicas analíticas más sofisticadas y precisas como son los modelos logitos. Además del ejemplo desarrollado, se realizan mosaicos y análisis similares en otras 7 situaciones que abarcan los casos estudiados, a saber niños (PRI, VER, OTO) y adultos (INV, PRI, VER, OTO), los cuales aparecen en el Anexo 1 y 2.

ASMA

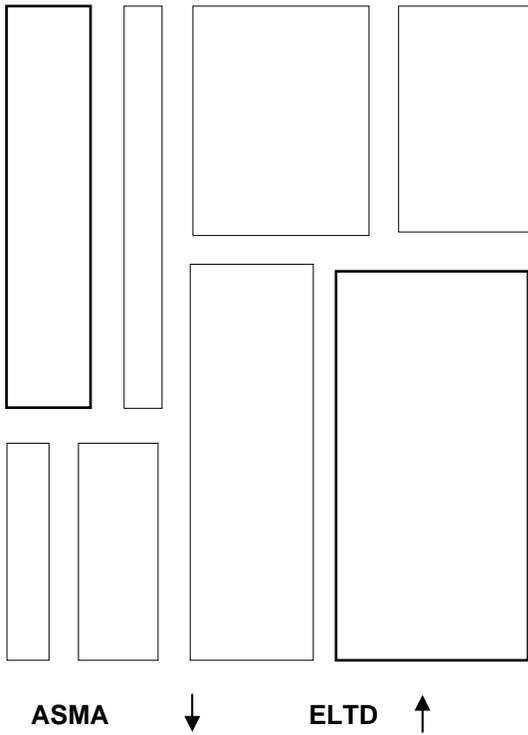
INVIERNO ADULTOS



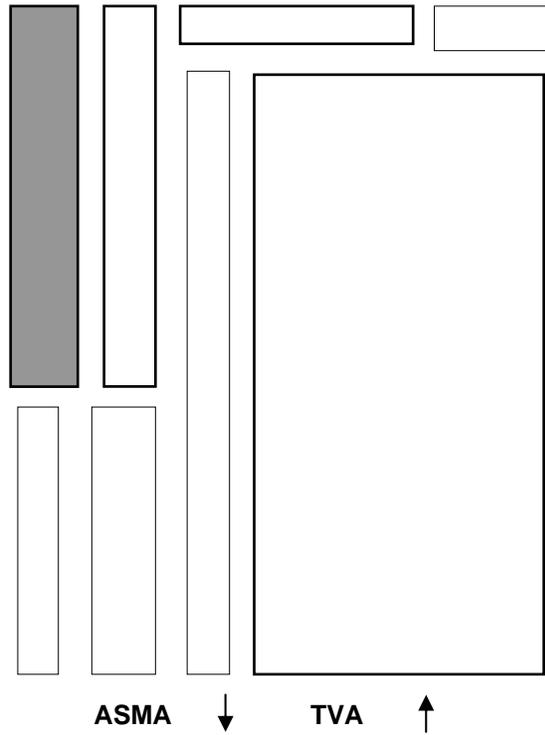
INVIERNO NIÑOS



OTOÑO ADULTOS

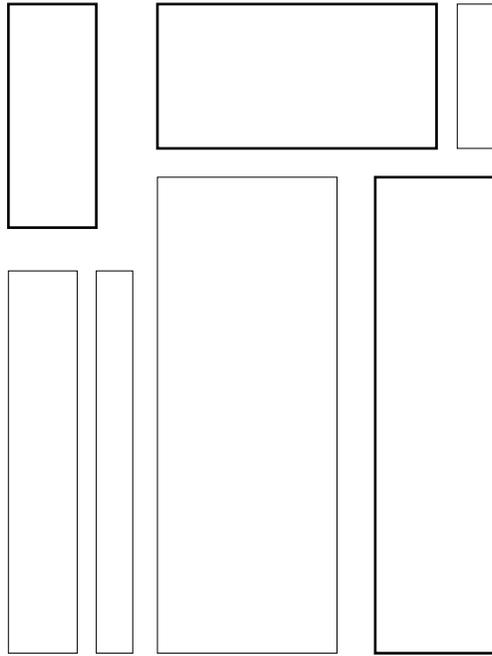


OTOÑO NIÑOS



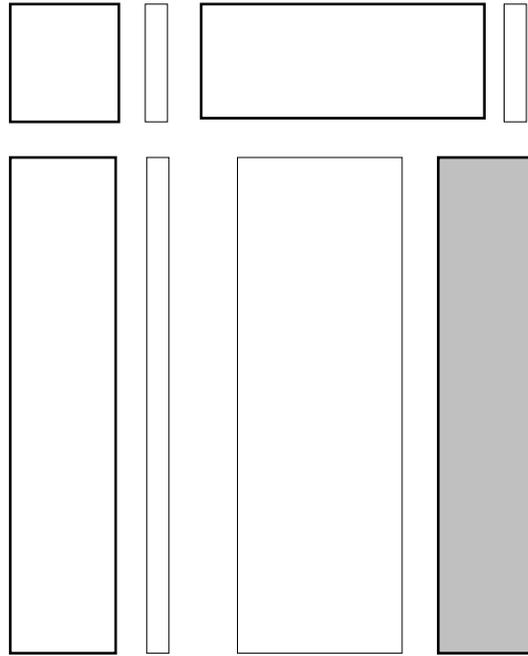
ASMA

PRIMAVERA ADULTOS



INDEPENDENCIA (ASMA)

PRIMAVERA NIÑOS



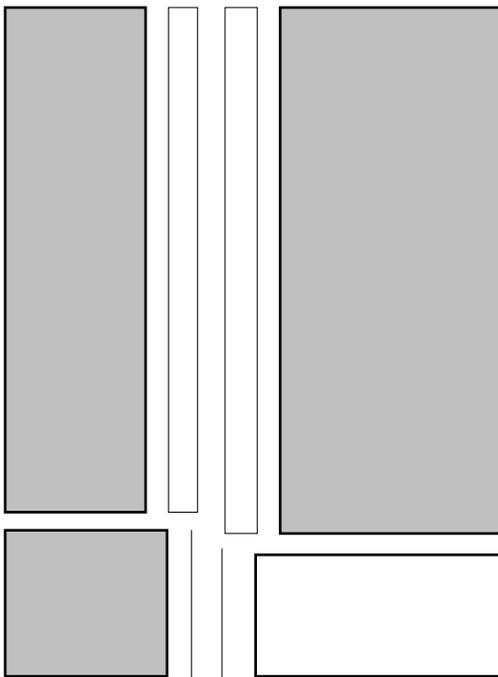
ASMA



DOA



VERANO ADULTOS



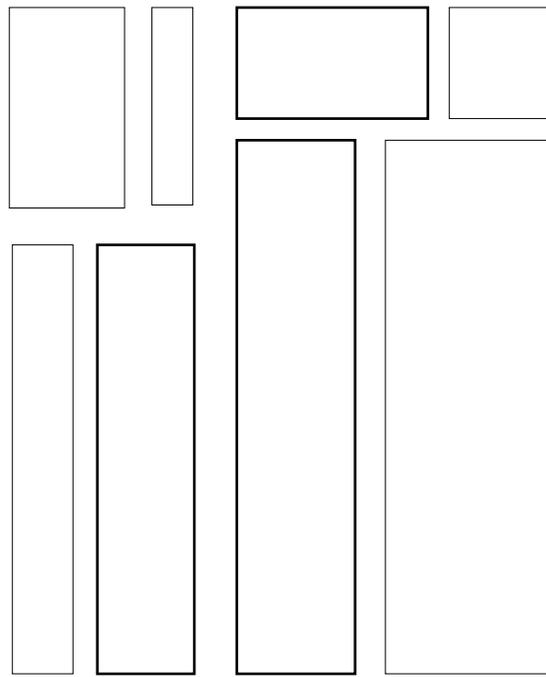
ASMA



TSS



VERANO NIÑOS



INDEPENDENCIA (ASMA)

REFERENCIAS

- AGRESTI, A. (1990): **Categorical Data Analysis**, John Wiley. New York.
- FRIENDLY, M. (1994): "Mosaic Display for Multi-Way Contingency Tables", **Journal of the American Statistical Association**, 89(425), Theory and Methods.
- HARTIGAN, J.A. and B. KLEINER (1981): **Mosaics for Contingency Table**. Computer Science and Statistics Proceedings of the 12th Symposs.
- HEYDEN, P.G.M. VAN der and J. LEEUW (1981): "De Correspondence Analysis and complementary to loglinear analysis", **Psychometrika** 59, 429-337.
- LINARES, G. and V. SISTACHS (1996): "Methodology for making-up bioclimatic predictions", **Revista de Investigación Operacional**, 17(1-3).
- SISTACHS, V. (1996): **Pronósticos bioclimáticos a través de técnicas estadísticas con datos categóricos**. Tesis de Maestría. Facultad de Matemática y Computación, Universidad de La Habana. Cuba.