

# ciencias matemáticas

## DIRECTOR

Dr. José A. Mesejo Chiong

## MIEMBROS

Dra. Marta Álvarez

Dr. Miguel Atencia Ruíz

Dr. Rafael Bello Pérez

Dr. Mijail Borges Quintana

Dr. Julián Bravo Castellero

Dr. Emilio Cerdá Tena

Dr. Jorge Estrada Sarlabous

Dr. Jesús García Molina

Dra. Susana Gómez Gómez

Dr. Gonzalo Joya Caparrós

Dr. Miguel Katrib Mora

Dra. Ángela M. León Mecías

Dr. Mateo Lezcano Brito

Dr. Jean M. Loubes

Dr. Joaquín D. Pina Amargós

Dr. Carlos Sánchez Fernández

Dr. José Ruíz Schucloper

Dra. Vivian Sistachs Vega

Dr. Reinaldo Rodríguez Ramos

Dr. Mariano Rodríguez Ricard

Dra. Rita Roldán Inguanzo

Dr. Alejandro Rosete Suárez

Dr. Pablo Olivares Rieumont

Dr. Humberto Madrid de Vega

## Edición

Aldo R. Gutiérrez Rivera

## Diseño y edición computarizada

Marlene Sardiña Prado



## EDITORIAL ACADEMIA

Calle 20, no. 4110, entre 41 y 47,  
Miramar, Playa.

La Habana, Cuba.

Teléfonos: (537) 202 7920, ext. 129,  
214 4195 (directo)

Web: [www.gecyt.cu](http://www.gecyt.cu)

ISSN 0256-5374

**PUBLICACIÓN SEMESTRAL**

## CONTENIDO

**El particionamiento de tablas e índices como una de las herramientas que se deben utilizar para la gestión de grandes volúmenes de datos / 3**

*William Korin Rodríguez*

**Comparación de métodos de regresión en la predicción de dióxido de carbono / 10**

*Gladys Linares / Adrián Saldaña*

*/Marcela Rivera / Hortensia J. Cervantes*

**Un acercamiento a las nociones de competencias básicas que la asignatura Análisis Matemático debe contribuir a desarrollar en los estudiantes de Ciencia de la computación / 17**

*Idania Urrutia Romaní / Virginia Álvarez Suárez*

**Generalizaciones y caracterizaciones del concepto de métrica / 26**

*Otilio Mederos Anoceto / Rita Roldán Iguanzo*

**Data Warehousing en soluciones de inteligencia de negocios / 39**

*Lucina García Hernández / Lis Velázquez Vidal*

*/ Mijail Veliz Monteagudo*

**Extracción de conocimientos lógicos y estadísticos a partir de tablas de datos binarios: Un estudio con alumnos de la enseñanza media superior de Santiago de Cuba / 54**

*Larisa Zamora Matamoros / Pilar Orús Baguena*

*/ Isabel Alonso Berenguer / Alexander Gorina Sánchez*

*/ Mailyne Álvarez Caneda*

**Polinomios multiortogonales clásicos en la recta real / 73**

*Walter Carballosa Torres / Luis R. Piñeiro Díaz*

# **Extracción de conocimientos lógicos y estadísticos a partir de tablas de datos binarios: Un estudio con alumnos de la enseñanza media superior de Santiago de Cuba**

*Larisa Zamora Matamoros\* / Pilar Orús Baguena\*\* / Isabel Alonso Berenguer\*  
/ Alexander Gorina Sánchez\* / Mailyn Álvarez Caneda\**

*\*Departamento de Matemática.*

*Facultad de Matemática y Computación. Universidad de Oriente, Cuba.*

*\*\*Departamento de Matemáticas.*

*Escuela de Tecnología y Ciencias Experimentales. Universidad Jaume I, España.*

## **Cómo citar este artículo:**

Zamora, A., Orús, P., Alonso, I., Gorina, A. y Álvarez, C. (2011). Extracción de conocimientos lógicos y estadísticos a partir de tablas de datos binarios: Un estudio con alumnos de la enseñanza media superior de Santiago de Cuba. *Ciencias Matemáticas*, 25 (único 2009-2011), 54-72.

# Extracción de conocimientos lógicos y estadísticos a partir de tablas de datos binarios: Un estudio con alumnos de la enseñanza media superior de Santiago de Cuba

Larisa Zamora Matamoros\* / Pilar Orús Baguena\*\* / Isabel Alonso Berenguer\*  
/ Alexander Gorina Sánchez\* / Mailyn Álvarez Caneda\*

\*Departamento de Matemática.

Facultad de Matemática y Computación. Universidad de Oriente, Cuba.

\*\*Departamento de Matemáticas.

Escuela de Tecnología y Ciencias Experimentales. Universidad Jaume I, España.

**Resumen:** en el trabajo se realiza un estudio sobre las habilidades que poseen alumnos de 10<sup>mo</sup> y 11<sup>no</sup> grado de la enseñanza media superior del municipio de Santiago de Cuba, para el trabajo simultáneo de la Lógica, la Estadística Elemental y el Tratamiento de Datos, en la interpretación de la información obtenida en una encuesta y presentada en una tabla de datos binarios, tomando como referente los trabajos de Pitarch (2002) y de Orús (1992), realizados con alumnos de la ESO en España y de primaria en Francia respectivamente. A partir de la aplicación de las técnicas del Análisis Estadístico Implicativo (ASI) se muestra como se relacionan el trabajo lógico y el estadístico a través del tratamiento de datos, se detectan diferencias significativas en las habilidades adquiridas por alumnos de dos centros de enseñanza media superior, tomados como muestra para la experimentación. Por último, se presentan, con ayuda del ASI, un conjunto de *meta reglas* que evidencian las relaciones entre las operaciones lógicas, estadísticas y de tratamiento de datos.

## 1. INTRODUCCIÓN

La Estadística es una rama de la Matemática que provee a los ciudadanos de una cultura específica muy útil a la hora de leer e interpretar, entre otras cosas, tablas y gráficos que cada vez con mayor frecuencia aparecen en los medios de difusión masiva, contribuyendo a lograr una mejor comprensión del mundo que les rodea.

También resulta valiosa para la vida profesional, ya que en muchas profesiones se requieren conocimientos básicos del tema y su estudio contribuye a desarrollar un razonamiento crítico, basado en la valoración de la evidencia objetiva.

Numerosos investigadores se han referido a la importancia de formar una cultura estadística en la sociedad para potenciar su desarrollo, tal es el caso de Batanero (2003), quien ha planteado que la Estadística es una parte de la educación general deseable para los futuros ciudadanos adultos, quienes precisan adquirir la capacidad de lectura e interpretación de tablas y gráficos estadísticos, que con frecuencia aparecen en los medios informativos. También es útil para la vida posterior, ya que en muchas profesiones se precisan conocimientos básicos del tema y su estudio ayuda al desarrollo personal, fomentando un razonamiento crítico, basado en la valoración de la evidencia objetiva.

Todo lo anterior pone de manifiesto el interés de realizar investigaciones encaminadas a perfeccionar el proceso de enseñanza y aprendizaje de la Estadística, en la formación básica de los ciudadanos. Al respecto se han realizado varios estudios y planteado numerosas direcciones de trabajo, como la de Batanero (1999), quien expuso en la *Conferencia Internacional Experiencias y Expectativas de la Enseñanza de la Estadística- Desafíos para el Siglo XXI*, el interés y viabilidad de introducir el Tratamiento de Datos en la enseñanza actual de la Estadística.

Orús (1992) se encamina en esta línea de investigación en Didáctica de la Estadística, con su tesis doctoral, al introducir el trabajo lógico sobre tablas de datos en la enseñanza primaria.

Una investigación más reciente, que continúa esta dirección e integra las dos propuestas anteriores, es la realizada por Pitarch (2002) y dirigida por Orús, la que analiza el razonamiento de los alumnos en la Enseñanza Secundaria Obligatoria (ESO) en España, utilizando simultáneamente el Tratamiento de Datos, la Lógica y la Estadística Elemental.

En Cuba recientemente se han comenzado a dar los primeros pasos para la incorporación de la enseñanza de la Estadística en el currículo de la escuela media (secundaria básica) y media superior (preuniversitaria). Esto, aparejado al creciente interés internacional en el desarrollo de la enseñanza y aprendizaje de la Estadística en los niveles secundarios e incluso en la enseñanza primaria, ha motivado a que un grupo de profesores de la Universidad de Oriente en Santiago de Cuba, comenzaran a investigar sobre las principales deficiencias que existen en la enseñanza de esta ciencia en el nivel preuniversitario. Un primer trabajo encaminado en esta dirección es el de Zamora y Alonso (2006), en el que se presenta una metodología encaminada a orientar a los profesores de la enseñanza media superior en lo referente a la organización y desarrollo del proceso de enseñanza-aprendizaje de la Estadística y las Probabilidades.

Por otro lado, Delgado (1999), aborda una propuesta metodológica-organizativa dirigida a lograr una mayor difusión de la cultura estadística en aquellos que la utilizan o la requieren, en dependencia de las condiciones y características propias del entorno

en el que se desea alcanzar esta cultura, en este caso, según Delgado en las condiciones de Cuba.

En esta misma línea Gorina et al. (2007a) analizan la importancia de la educación estadística en la formación integral de los estudiantes de doctorado en Ciencias Pedagógicas en Cuba y en Gorina et al. (2007b) los autores proponen un sistema de habilidades a desarrollar en los doctorantes con el propósito de facilitar el trabajo de formación de la competencia estadística. En Veiga (2007) se presenta una reflexión acerca de algunos de los obstáculos y dificultades conceptuales con los que se enfrentan los estudiantes de Secundaria en el aprendizaje de la Estadística en Brasil.

En el citado trabajo de Zamora y Alonso (2006) y según el análisis valorativo hecho sobre el actual plan de estudio de Matemática de la enseñanza preuniversitaria cubana y los resultados del diagnóstico realizado por las autoras, se pudo constatar que a partir de los cursos 2004-2005 y 2005-2006 se incorporan contenidos de Estadística a los programas de 10mo y 11no grado respectivamente. A partir de este mismo diagnóstico, se pudo confirmar que dichos programas no contemplan explícitamente el Tratamiento de Datos y que los contenidos de la Estadística Descriptiva se imparten empleando métodos tradicionales, concluyéndose que la enseñanza de la Estadística presenta, desde el punto de vista de su diseño curricular, dificultades similares a las planteadas por Pitarch (2002), lo cual sirvió de motivación para desarrollar un estudio similar al realizado en España con la ESO, pero con alumnos de la enseñanza preuniversitaria, con el objetivo de analizar las habilidades que poseen alumnos de 10mo y 11no grado, para el trabajo simultáneo con aspectos esenciales de Tratamiento de Datos, Lógica y Estadística Elemental, y poder comparar los resultados con los obtenidos en la ESO.

## 2. METODOLOGÍA

Dado el interés de comparar los resultados de la presente investigación con los obtenidos en la ESO, se decidió emplear la misma metodología seguida por Pitarch, con ligeras modificaciones.

Se parte de considerar un cuestionario inicial con el objetivo de obtener información primaria que permitiese modelar el trabajo lógico y estadístico, así

como el tratamiento de datos sobre un instrumento común, la tabla de datos. Se confecciona el cuestionario de la experimentación, similar a los de Orús (1992) y al de Pitarch (2002), cuyo diseño se realizó en el marco de la Teoría de Situaciones de Brousseau, concebido como una *situación fundamental\** que modeliza el funcionamiento del contenido matemático, en este caso, la lógica y la clasificación en situación escolar, mediante *variables* de la situación.

Respecto al análisis de los datos de la experimentación, queremos destacar la aplicación del Análisis Estadístico Implicativo (ASI), el cual es un método de análisis no simétrico de datos que permite la extracción y la estructuración del conocimiento en forma de normas y reglas generalizadas, a partir de un conjunto de datos que interrelaciona una población de sujetos (u objetos) con un conjunto de variables. Su origen es la modelización estadística de la cuasi-implicación: cuando la variable o la conjunción de variables "*a*" es observada en la población, entonces generalmente la variable "*b*" lo es también. Las variables en cuestión pueden ser de distintos tipos: binario, modal, frecuencial o de intervalo. Contrariamente a los métodos de análisis simétricos basados, por ejemplo, en una distancia o en una correlación, los conjuntos de reglas obtenidas pueden conducir a hipótesis de causalidad. Estos conjuntos se estructuran según diferentes características comunes complementarias (grafo implicativo, jerarquía orientada). La determinación cuantitativa de los sujetos o descriptores responsables de estas estructuras viene dada por su contribución o su tipicidad.

Como soporte computacional de las técnicas de este tipo de análisis multivariante de datos, el ASI, se utiliza el sistema informático C.H.I.C. (Classification Hiérarchique Implicative et Cohésitive) en su versión 4.1, el cual es un programa informático estadístico adaptado por Regis Gras a partir de los índices de proximidad o distancia de Lerman, I. (1981) y se caracteriza por su sencillez de manejo, ser innovador en el tratamiento de la información y proporcionar, de

\*Situación fundamental: una situación fundamental correspondiente al conocimiento *C* es un conjunto mínimo de situaciones *a*-didácticas que son específicas de *C* que dan lugar a un campo de problemas suficientemente extenso como para proporcionar una buena representación de *C*, teniendo en cuenta el manejo de los valores de las variables didácticas y la reconstrucción de *C* en la situación didáctica en cuestión.

una forma rápida, gran cantidad de cálculos estadísticos y gráficos: árboles y grafo de clasificación e implicación.

## 2.1 El sistema informático C.H.I.C.

El sistema informático C.H.I.C. permite realizar el análisis de los datos a partir de la definición de variables, que pueden ser principales (binarias, modales, frecuenciales o de intervalo) o suplementarias. Las primeras son las que se tienen en cuenta normalmente por C.H.I.C., y las segundas, son variables binarias o modales que no intervienen en el cálculo de las contribuciones de las categorías, sólo se tienen en cuenta en la búsqueda de la contribución o la tipicidad.

Este sistema informático muestra las implicaciones más relevantes entre los valores de las diversas variables, y entre sus funciones (opciones del programa) está construir:

- El Árbol de Similitud, el cual utiliza el índice de proximidad de I. C. Lerman, para el cálculo de los niveles de agregación de las variables y su clasificación.
- El Árbol Cohesivo, efectuando los cálculos de los índices de cohesión implicativa entre las variables y las clases de variables.
- El Grafo Implicativo, calculando las implicaciones directas entre variables, estructurando originalmente las variables mediante cadenas de implicaciones en el sentido del Análisis Estadístico Implicativo, referenciado en el apartado anterior.

En los tres casos muestra una ventana de resultados numéricos (índices y niveles de estructuración) y otra que presenta el gráfico correspondiente.

La medida de la relación implicativa  $a \Rightarrow b$  se evalúa a partir de la inverosimilitud de la aparición, en los datos, del número de casos que la invalidan, es decir, cuantifica "el asombro" del experto ante el número inverosímilmente pequeño de contraejemplos. Una implicación entre clases de variables solo toma verdaderamente su sentido bajo la condición de que dentro de cada clase de variables, cuya relación se examina con otras, exista una cierta "cohesión" entre las variables que la constituyen. (Gras R., Kuntz P. et Briand H., 2001).

C.H.I.C. permite conocer la *contribución* en cada una de las clases, de cada uno de los sujetos y también de las *variables suplementarias* que se definan para caracterizar a los individuos. Cada una de estas últimas contribuye más o menos a la formación de la clase: lo cual significa que los valores que ellos dan a las variables, se da en el sentido de su similaridad. R. Gras y H. Ratsimba-Rajohn han elaborado un criterio que permite evaluar esta contribución relativa en cada clase.

Ciertos sujetos son “típicos” en el comportamiento de un conjunto de la población en el sentido siguiente: En el estudio de la similaridad, ellos atribuyen al conjunto de variables, valores compatibles con las similaridades constituidas sobre estas variables por la población. Si algunas variables suplementarias han sido definidas por el usuario, se obtendrá así la tipificación de estas variables a partir de aquellas que los individuos las satisfacen.

La utilización de C.H.I.C. en nuestra investigación se realiza en tres momentos:

- En primer lugar, sobre la tabla de datos binarios T, la cual se obtiene a partir de las respuestas dadas por los alumnos al cuestionario inicial. Los resultados obtenidos no se presentan en este artículo, se pueden consultar en Zamora et al. (2007).
- En segundo lugar, sobre la matriz a priori MAP, construida a partir de la operacionalización de cada una de las preguntas del cuestionario de la experimentación, con el objetivo de obtener una clasificación de las preguntas del mismo, atendiendo a las operaciones matemáticas que los alumnos deben realizar para responderlas.
- Por último, se aplica el C.H.I.C. sobre la matriz de datos a posteriori MP. Esta matriz se conforma a partir de las respuestas correctas (1) o incorrectas (0) dadas por los alumnos al cuestionario de la experimentación. De esta forma se realiza el estudio a posteriori del cuestionario. El grafo y los árboles que se obtienen permiten analizar los resultados de los alumnos con respecto a su trabajo matemático, al utilizar una tabla de valores para contestar al cuestionario.

Además del ASI, como método cualitativo de análisis de los datos, se emplearon métodos

cuantitativos, en particular de la Estadística Descriptiva e Inferencial.

## 2.2 Cuestionarios

Se parte de considerar el cuestionario inicial presentado por Pitarch (2002)\*, el cual indagaba acerca de los gustos de los alumnos de la enseñanza secundaria obligatoria (ESO) en España por diferentes tipos de músicas, y su objetivo era obtener información primaria que permitiese modelar el trabajo lógico, estadístico, así como el tratamiento de datos en este tipo de enseñanza. Nuestro cuestionario persigue este mismo objetivo, empleando como excusa y motivación del cuestionario el gusto de los alumnos, de la enseñanza preuniversitaria del municipio de Santiago de Cuba, por diferentes asignaturas y profesiones (carreras universitarias). Cada pregunta del mismo puede ser respondida de forma afirmativa (1) o negativa (0), generándose así una matriz booleana o tabla de datos binarios.

La elaboración del cuestionario de la experimentación se realiza a partir de las respuestas que habían manifestado los alumnos en este cuestionario inicial. Los resultados obtenidos al aplicar las técnicas del ASI a la tabla de datos binarios, con el objetivo de obtener información acerca de semejanzas o proximidades entre los alumnos y entre las preguntas del cuestionario, permitieron realizar algunas modificaciones y obtener así el cuestionario de la experimentación libre de posibles preguntas ambiguas o equivalentes, el cual se muestra en el anexo 1. El trabajo de elaboración y modificación del cuestionario pueden consultarse en Zamora et al. (2007).

## 2.3 Análisis a priori del cuestionario de la experimentación

Para el análisis a priori del cuestionario de la experimentación se procedió, al igual que en la investigación de Pitarch, a realizar una caracterización de las operaciones lógico-matemáticas que debían realizar los alumnos al responder las preguntas del mismo, las cuales se distribuyen en los siguientes bloques y se muestran en la tabla 1.

\* pp. 52-53

- Manejo básico sobre la tabla (TABL, TFIL, TCOL).
- Recuentos obligatorios o alternativos (RROBL, RALT).
- Comparación entre filas y/o columnas (CFIL, CCOL).
- Proximidad /distancia (FDIF, FCOI, FAUS).
- Relación semántica (CSEM).
- Operaciones lógicas (LIMP, LEXI, LUNI, LCON, LDIS, LNEG).
- Operaciones básicas de estadística (EFIM, EFEX, Efre, EPER, EMIM).
- Razonamiento o argumentación de la respuesta (AEGU).

A partir de esta caracterización, se elaboró una nueva matriz booleana, en la cual las filas responden a las operaciones lógico-matemáticas y las columnas a las

preguntas. El valor 1 en una casilla significa que el estudiante debe realizar la operación asociada a la fila para responder la pregunta correspondiente a la columna de la casilla, y un valor cero en el caso contrario. La matriz así obtenida se muestra en el anexo 2 y se denomina matriz a priori (MAP).

Esta matriz permite analizar a priori cuál puede ser el comportamiento esperado en las respuestas dadas por los alumnos, si se tiene en cuenta la proximidad entre las preguntas según su caracterización, lo que a priori sugiere una proximidad entre las respuestas de los alumnos a las preguntas. De la aplicación del C.H.I.C a la matriz a priori MAP se obtuvo el árbol de similitud que se muestra en la figura 1, marcándose los nodos significativos (en trazos más gruesos), siendo el nivel 1 el más significativo.

**Tabla 1.** Caracterización de las preguntas del cuestionario de la experimentación

**TABL:** Utilizar un elemento concreto de la tabla, para responder a la pregunta.

**TFIL:** Utilizar una o varias filas para responder a la pregunta.

**TCOL:** Utilizar una o varias columnas para responder a la pregunta.

**ROBL:** Enumerar de forma obligatoria una cantidad concreta (presencia: 1, ausencia de carácter: 0).

**RALT:** Enumerar de forma alternativa una cantidad concreta (presencia: 1, ausencia de carácter: 0).

**CFIL:** Comparar filas.

**CCOL:** Comparar columnas.

**FDIF:** Tener en cuenta las diferencias entre filas/columnas.

**FCOI:** Fijarse en las coincidencias (tanto las co-presencias: 1-1, como en las co-ausencias: 0-0).

**FPRE:** Fijarse únicamente en las co-presencias.

**FAUS:** Fijarse únicamente en las co-ausencias.

**CSEM:** Comparar entre los criterios de la tabla (filas), a partir de la posible relación semántica que existe entre las asignaturas y entre las carreras.

**LIMP:** Implicación lógica.

**LEXI:** Cuantificador existencial.

**LUNI:** Cuantificador universal.

**LCON:** Conector lógico conjunción.

**LDIS:** Conector lógico disyunción.

**LNEG:** Conector lógico negación.

**EFIM:** Aplicar implícitamente la frecuencia absoluta.

**EFEX:** Aplicar explícitamente la frecuencia absoluta.

**Efre:** Aplicar explícitamente la frecuencia relativa.

**EPER:** Aplicar explícitamente el porcentaje.

**EMIM:** Utilización implícita de la medida de centralización: la moda.

**ARGU:** Razonar en la respuesta.

De dicho árbol se aprecia que se forman las siguientes clases de cuasi equivalencia:

$$C1 = \{P1, P2, P9, P10, P3, P8\},$$

$$C2 = \{P5, P21, P14, P19, P17, P18\},$$

$$C3 = \{P4, P4.1, P6, P7, P7.1\}$$

$$C4 = \{P11, P12, P15, P16, P20, P13\}.$$

La clase C1 contiene tres nodos significativos a los niveles 1, que es el más significativo, y a los niveles 3 y 8. Dicha clase agrupa preguntas que se caracterizan por un recuento obligatorio por filas y columnas, contenidos elementales de Estadística (P9 y P10), y operaciones lógicas de conjunción (P8), por lo que se puede concluir que la clase C1 agrupa preguntas que involucran actividades enumerativas, de manejo de datos y estadísticas.

La clase C2 posee un nodo significativo al nivel 19, que involucra a todas las preguntas de la clase, las que se caracterizan por la operación lógica de conjunción, conjunción con negación e implicación, con excepción de las preguntas P5 y P21 que se caracterizan por exigir un recuento obligatorio por columnas, por lo que dicha clase se caracteriza por las actividades lógicas: conjunción e implicación.

La clase C3 presenta dos nodos significativos, el primero al nivel 5 y agrupa a las preguntas P4, P4.1 y P6, las que se caracterizan por realizar recuentos alternativos y operaciones lógicas empleando los cuantificadores existencial y universal, y el segundo al nivel 13, que agrupa a todas las preguntas de la clase, es decir, a las anteriores y a las P7 y P7.1 que se caracterizan por realizar operaciones lógicas empleando el cuantificador existencial, por lo que se puede concluir que esta clase comprende cuestiones abiertas que requieren del trabajo con toda la tabla y el empleo de diversas estrategias de búsqueda sobre la misma.

En la clase C4 aparece un nodo significativo al nivel 15, que involucra a todas las preguntas de la clase, exceptuando a la pregunta 13, las que se caracterizan por trabajar las coincidencias y diferencias por columnas, de manera que dicha clase queda caracterizada en cuanto al trabajo sobre la proximidad entre los sujetos.

La clasificación en clases obtenidas de acuerdo a la caracterización de las preguntas, será empleada en el análisis a posteriori.

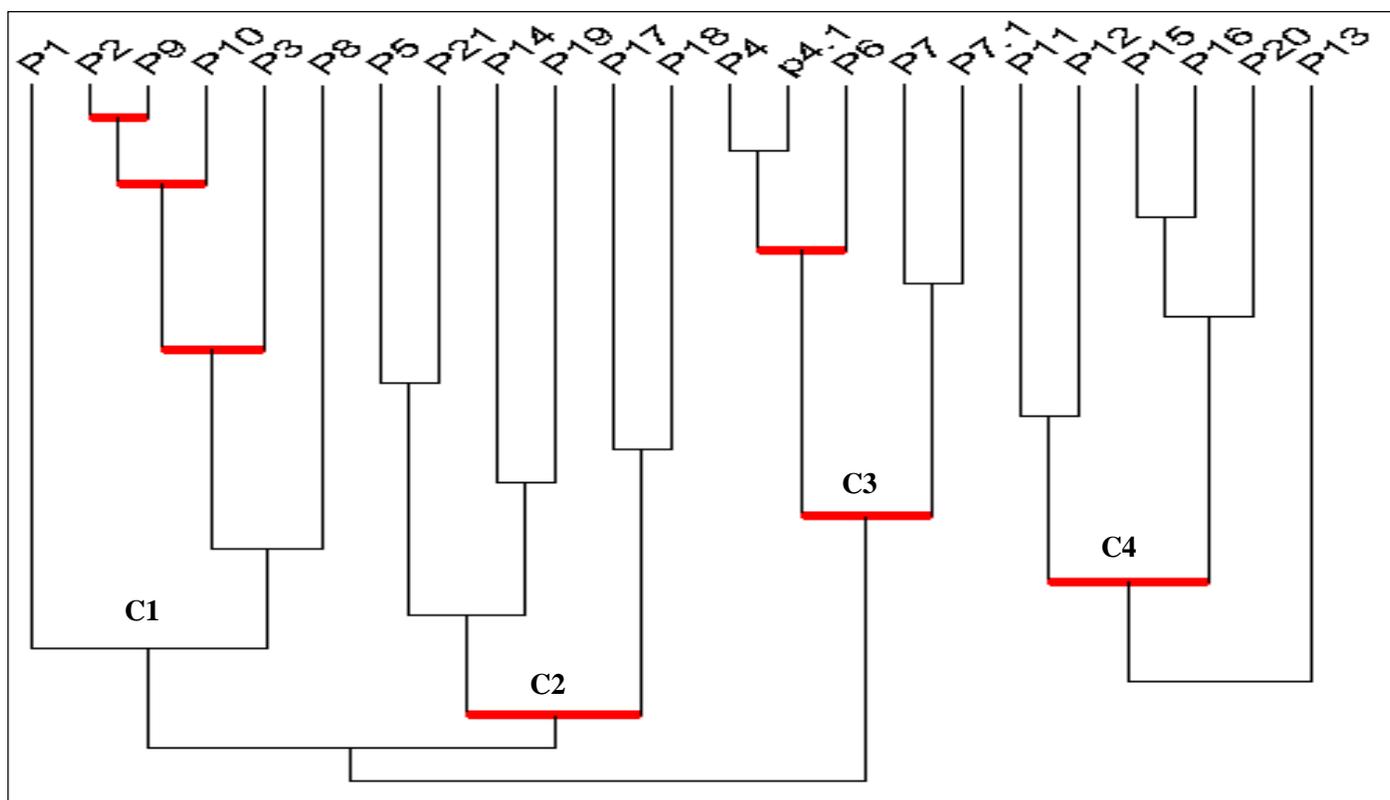


Fig. 1. Árbol de similaridad de la matriz a priori MAP.

### 3. EXPERIMENTACIÓN DEL CUESTIONARIO Y TRATAMIENTO DE DATOS

#### 3.1 Condiciones de la experimentación

Para desarrollar la investigación se seleccionaron dos centros de enseñanza preuniversitaria: Rafael María de Mendive y Cuqui Bosch, por no presentar características educativas especiales, considerándose como una muestra representativa de los preuniversitarios del municipio de Santiago de Cuba.

Fueron seleccionados, en una primera etapa, 108 alumnos de 11<sup>no</sup> grado de ambos centros, mediante el muestreo aleatorio estratificado con afijación proporcional al tamaño de la población (ver tabla 2), cuyas edades oscilan entre 15 y 17 años, a los cuales se les aplicó el cuestionario inicial que indagaba acerca de sus gustos por determinadas asignaturas y carreras.

**Tabla 2.** Tamaño de muestra para la aplicación del cuestionario inicial

Centro de enseñanza preuniversitaria		
Rafael María de Mendive	Cuqui Bosch	Total
55	53	1 082

Del análisis de los resultados del cuestionario se obtuvo la tabla de datos binarios  $T$  (ver anexo 1) de dimensión  $20 \times 21$ , en la cual las filas representan las preguntas y las columnas a los alumnos, los que fueron codificados por  $Ei$ , con  $i = 1, \dots, 21$ .

Para determinar las dimensiones de la tabla se tuvo en cuenta la experiencia de la investigación desarrollada por Pitarch\*, en la cual se obtuvo un 97% de éxito en la lectura y manipulación elemental de la misma, con una dimensión de  $15 \times 20$ . También se consideró el hecho de que la presente experimentación se aplicaría solamente a alumnos de niveles equivalentes al 2do ciclo de la ESO. Por todo lo anterior se aumentó la dimensión de dicha tabla a  $20 \times 21$ .

Para la aplicación del cuestionario de la experimentación fueron seleccionados 102 alumnos de los centros "Rafael María de Mendive" y "Cuqui Bosch", de 10<sup>mo</sup> y 11<sup>no</sup> grados, mediante el muestreo aleatorio estratificado con afijación proporcional al tamaño de la población (ver tabla 3).

**Tabla 3.** Tamaño de muestra para la aplicación del cuestionario de la experimentación

Centros de enseñanza preuniversitaria			
	Rafael María de Mendive	Cuqui Bosch	Total
Décimo grado	26	25	51
Décimo primero	25	26	51
Total	51	51	102

#### 3.2 Análisis de los datos

Para el análisis de los datos se emplearon métodos cuantitativos y cualitativos, dentro de los cuantitativos fueron aplicados los métodos clásicos de la Estadística Descriptiva e Inferencial: el cálculo de porcentajes, los diagramas de barras y las pruebas de hipótesis para proporciones. Como método cualitativo de análisis se empleó el ASI.

##### 3.2.1 Métodos cuantitativos de análisis

Se determinó el porcentaje de respuestas correctas por centro y por grado, para lo cual se empleó la hoja de cálculo electrónica Excel. Con el objetivo de investigar si las diferencias que se observaban entre los centros y entre los grados eran significativas, (para un nivel de confiabilidad del 95%) se procedió a aplicar el software estadístico SPSS.

En la figura 2 se muestran los porcentajes de respuestas correctas por centro, de la cual se puede observar que las preguntas con mayores dificultades fueron P8, P9, P10, P21, relacionadas con actividades enumerativas y dentro de éstas las relacionadas con las estadísticas, las P12, P15, P20, que abordan la proximidad entre sujetos y las P18 y P19 que se vinculan a operaciones lógicas. Justamente es en las citadas preguntas donde se localizan las mayores diferencias entre los centros. Al observar la figura 2 puede comprobarse que en todas estas preguntas los porcentajes más bajos se obtuvieron en el centro "Rafael María de Mendive", exceptuando la pregunta P18.

Al analizar los porcentajes de respuestas correctas por grado, se obtuvo el diagrama de barras que se muestra en la figura 3.

Las diferencias más significativas (de más del 10%) se presentan en las preguntas antes citadas, a las que se adicionan la P5 y P13, que abordan operaciones lógicas y de proximidad, respectivamente. En todas estas preguntas los porcentajes más bajos se obtuvieron en 10mo grado.

\* pp. 52-53, 81.

Al aplicar el software estadístico SPSS se obtuvo que las diferencias observadas entre los centros son significativas para las preguntas P03 y P09 de la clase C1 (actividades enumerativas y estadísticas); P19 de la clase C2 (actividades lógicas); P6 y P7.1 de la clase C3 (trabajo con toda la tabla y varias estrategias de búsqueda en la misma); P12, P15 y P20 de la clase C4 (proximidad entre sujetos), concluyéndose que existen diferencias significativas entre los dos centros investigados en las cuatro clases que se han formado a partir de la proximidad existente en las operaciones

que deben realizar los alumnos para responder a las preguntas del cuestionario.

Por otro lado, en las preguntas P09 y P10 de la clase C1; P05 y P19 de C2 y P11, P12, P13 y P15 de C4, se obtuvo que existen diferencias significativas entre los grados. Como se observa, no existen diferencias significativas entre los grados con respecto a las preguntas que requieren del trabajo con toda la tabla y el empleo de diversas estrategias de búsqueda sobre la misma (clase C3).

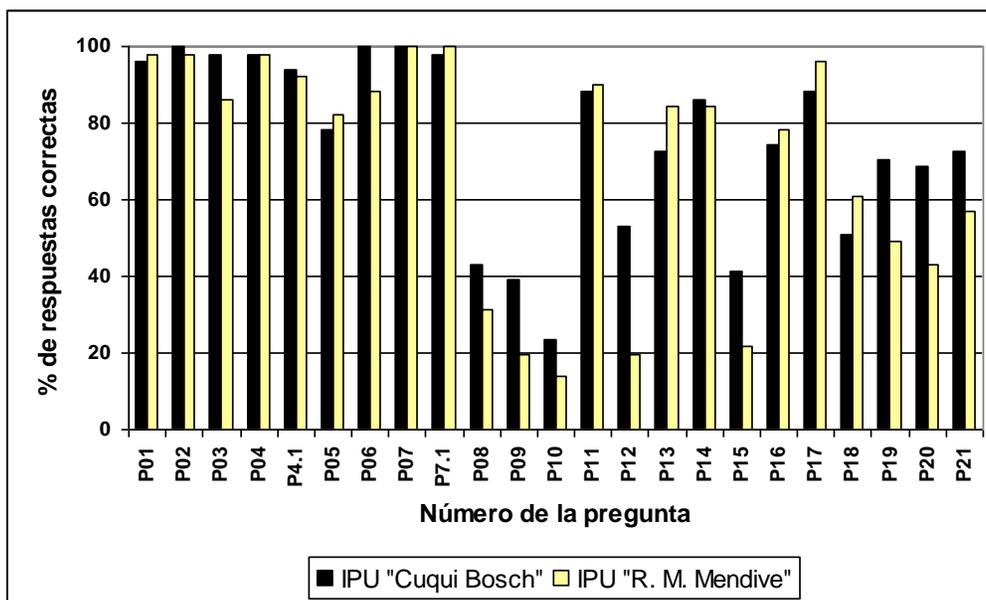


Fig. 2. Porcentajes de respuestas correctas por centros.

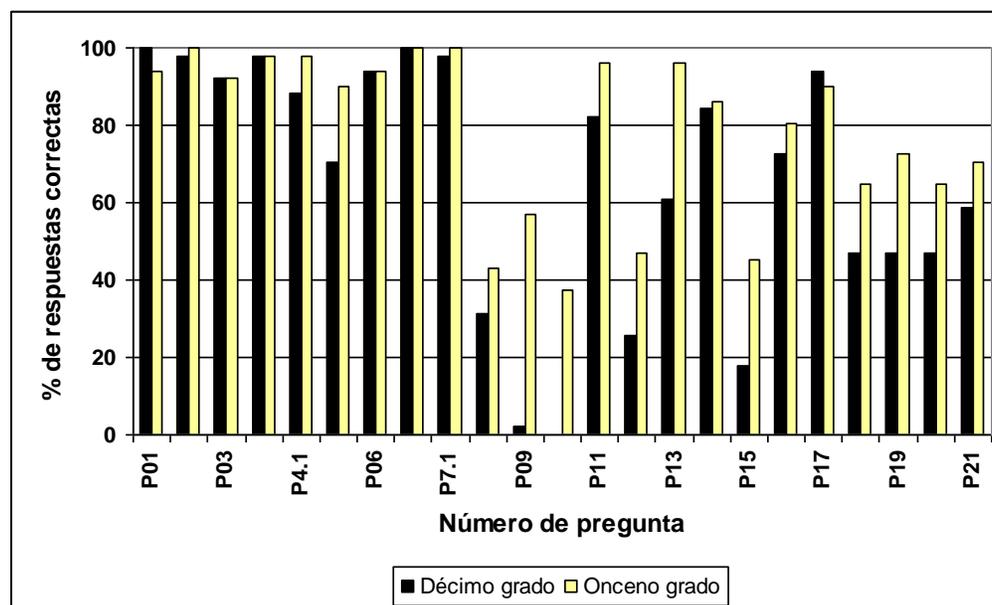


Fig. 3. Porcentajes de respuestas correctas por grados.

### 3.2.2 Método cualitativo de análisis

El análisis cualitativo se basó en la interpretación del árbol de similaridad, el árbol cohesitivo y el grafo implicativo, obtenidos mediante el programa informático C.H.I.C. Las implicaciones fueron evaluadas según el Método Clásico, y se empleó la Ley Binomial\* para la simulación del número aleatorio de contraejemplos a la implicación o a la regla\*\*.

Tras la experimentación se conformó una nueva matriz, la matriz a posteriori (MP), que consta de 102 filas, asociadas a los alumnos encuestados y 23 columnas, correspondientes a la clasificación en 1 (respuesta correcta) ó 0 (respuesta incorrecta) realizadas a las respuestas dadas por los alumnos al cuestionario de la experimentación. Con el objetivo de poder determinar el centro, el grado y la calidad del alumno que eran más típicos y que contribuían más a la formación de las clases y de los caminos, se le adicionaron 8 columnas a la matriz, dos para identificar el centro de procedencia, dos para identificar el grado y las 4 restantes para clasificar al alumno en "Excelente", si respondía correctamente más del 90%

de las preguntas, "Bueno" si respondía correctamente entre el 75 y el 89%, "Regular" si respondía correctamente entre el 60 y el 74% y "Malo" si respondía correctamente menos del 60% de las preguntas del cuestionario. Para esta codificación se utilizó el sistema de evaluación que actualmente se aplica en la enseñanza preuniversitaria en Cuba.

### Árbol de similaridad

El árbol de similaridad sobre la matriz MP se muestra en la figura 4. En él se observa que se forman cuatro clases de cuasi equivalencias, que denotaremos por: D1, D2, D3 y D4.

Las clases D1 y D2 que se forman, a pesar de no ser significativas debido a la alta frecuencia de éxito en las respuestas de los alumnos a las preguntas que la conforman, agrupan cuestiones vinculadas a operaciones de búsqueda en toda la tabla, así como de actividad enumerativa. La formación de estas dos clases muestra que existen diferencias entre las preguntas que la integran, en la primera la búsqueda se realiza fijando las filas (representan las preguntas) y realizando el recuento por columnas (representan a los individuos) y en la segunda al contrario, aspectos estos

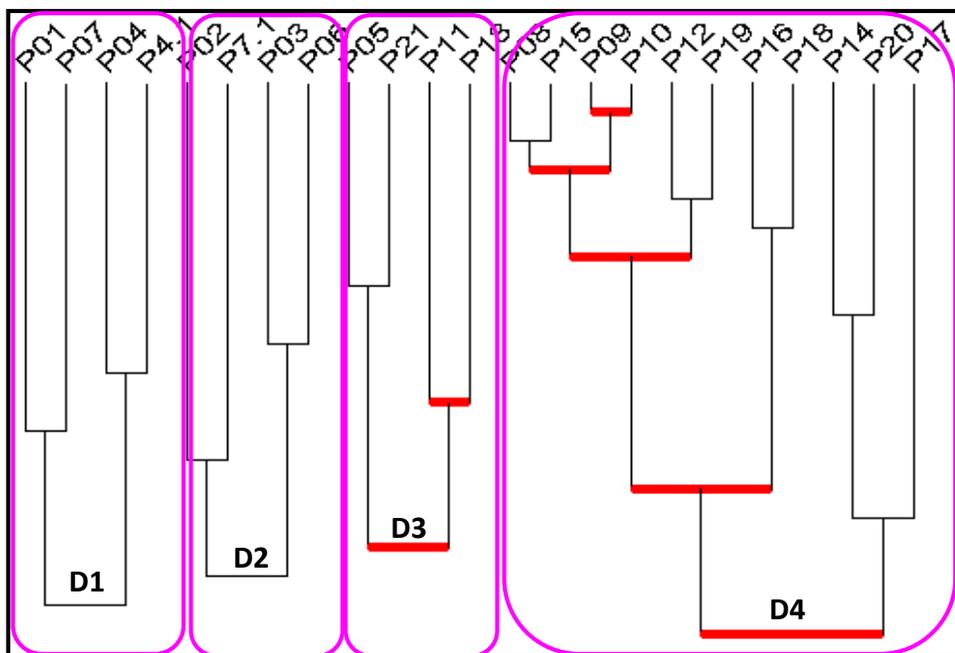


Fig. 4. Árbol de similaridad de la matriz a posteriori MP.

\* Para modelar la inclusión de conjuntos según la teoría de la implicación estadística el C.H.I.C. brinda dos opciones, el Método Clásico y el Método Entrópico. Así mismo, para la modelización del número aleatorio de contraejemplos a la implicación o a la regla, se debe elegir entre la ley Binomial y la de Poisson.

\*\* Versión 4.1 del C.H.I.C. (Ayuda online).

que son comparables matemáticamente. El porcentaje de éxitos de las preguntas muestra que los alumnos poseen menos habilidades en el recuento por filas.

La clase D3 contiene dos nodos significativos a los niveles 11 y 15, éste último agrupando a todas las preguntas de la clase, las cuales se caracterizan por exigir a los alumnos actividades lógicas y de proximidad para sus respuestas. En particular las operaciones de proximidad que aparecen en esta clase son aquellas que exigen la búsqueda de las co-presencias (1,1) y co-ausencias (0,0), que a pesar de que ambas representan las coincidencias, no tienen el mismo sentido.

La clase D4 contiene al nodo más significativo, nivel 6 de la clasificación, que agrupa a aquellas preguntas que suponen operaciones enumerativas para responderlas, en particular las asociadas a las operaciones estadísticas, y de proximidad. Del árbol se aprecia como los nodos significativos marcan el nivel de agregación con clases no significativas. En esta clase las actividades que suponen un trabajo lógico aparecen como diluidas en el trabajo de proximidad o comparación de datos, es decir, no aparecen agrupadas en una misma clase tal y como ocurre en la matriz a priori MAP. Su significación se debe a que, a pesar de la baja frecuencia de éxito de los alumnos en

sus respuestas, existen individuos semejantes en su comportamiento.

En la tabla 4 se muestra, a modo de ejemplo, los índices de similaridad obtenidos para cada uno de los nodos más significativos, así como las preguntas que conforman cada clase.

**Tabla 4.** Índices de similaridad de los nodos significativos de la matriz a posteriori

Nivel	Preguntas que agrupa	Índice de similaridad
1	P9 y P10	0.999813
3	P8, P15, P9y P10	0.948941
6	P8, P15, P9, P10, P12 y P19	0.781703
11	P11 y P13	0.576379
14	P8, P15, P9, P10, P12, P19, P16 y P18	0.412933
16	P5, P21, P11y P13	0.271671
19	P8, P15, P9, P10, P12, P19, P16, P18, P14, P20 y P17	0.0561027

En las tablas 5 y 6 se muestran los porcentajes de alumnos que han contribuido a formar las clases del árbol de similaridad y que son típicos a las mismas.

**Tabla 5\*.** Porcentaje de alumnos que contribuyen a la formación de las clases, por grados y centro

Contribución de los individuos							
	Clase 1	Clase 3	Clase 6	Clase 11	Clase 14	Clase 16	Clase 19
10mo C	0,00	0,00	0,00	40,00	12,00	24,00	12,00
11no C	34,62	15,38	15,38	92,31	53,85	88,46	57,69
10mo M	0,00	0,00	0,00	61,54	0,00	30,77	3,85
11no M	20,00	0,00	0,00	92,00	4,00	44,00	8,00

**Tabla 6.** Porcentaje de alumnos que son típicos de las clases, por grados y centro

Tipicalidad de los individuos							
	Clase 1	Clase 3	Clase 6	Clase 11	Clase 14	Clase 16	Clase 19
10mo C	0,00	0,00	0,00	40,00	0,00	24,00	0,00
11no C	34,62	34,62	34,62	92,31	34,62	88,46	34,62
10mo M	0,00	0,00	0,00	61,54	0,00	38,46	0,00
11no M	20,00	20,00	20,00	92,00	20,00	44,00	20,00

\* 10mo C, M: 10mo grado del centro “Cuqui Bosch” (C) y “Rafael María de Mendive” (M).  
 11no C, M: 11no grado del centro “Cuqui Bosch” (C) y “Rafael María de Mendive” (M).

Se observa que a la formación de las clases 1, 3 y 6, que agrupan las operaciones estadísticas y de proximidad, no contribuyen los alumnos de décimo grado y tampoco existen alumnos típicos de este grado. Los alumnos del centro Rafael María de Mendive sólo contribuyen a la formación de la clase 1, que agrupa operaciones estadísticas. Con respecto a la tipicidad de los individuos, nuevamente los alumnos de Mendive sólo son típicos de la clase 1.

Comparando la razón entre alumnos del Cuqui Bosch y Rafael María de Mendive, que son típicos y contribuyen a la formación de las clases, es de aproximadamente 2/1 para las clases 1, 3 y 6, es decir, por cada 2 alumnos del Cuqui Bosch que son típicos o que contribuyen a la formación de las clases, sólo hay uno que lo hace de Rafael María de Mendive.

De todo lo anterior se desprende que el centro Cuqui Bosch es el que más contribuye a la estructuración del conjunto de datos (con un mayor índice de éxito entre sus alumnos).

### Árbol cohesivo

La figura 5 del árbol cohesivo de la matriz a posteriori MP muestra que sólo se ha conformado una sola clase de preguntas, la cual resulta significativa con un índice de cohesión de 0.80. Esta clase contiene al nodo más significativo, al nivel 11, que es precisamente el nivel al cual se agrupan todas las preguntas de la clase.

A partir de la estructuración que se ha obtenido del árbol cohesivo, se desprende la formación de la *meta regla*\*  $C:C_1 \Rightarrow C_2$ , donde a su vez  $C_1:A \Rightarrow B$  y  $C_2:P \Rightarrow Q$  son metas reglas definidas a partir de reglas o meta reglas, las cuales se detallan a continuación:

- A: "si el estudiante es capaz de calcular la frecuencia relativa (P10) entonces es capaz de calcular la frecuencia absoluta (P09)",
- B: "si el estudiante es capaz de hallar las diferencias entre columnas (P12), será capaz de hallar las coincidencias (co-presencias y co-ausencias) (P13) entre columnas".

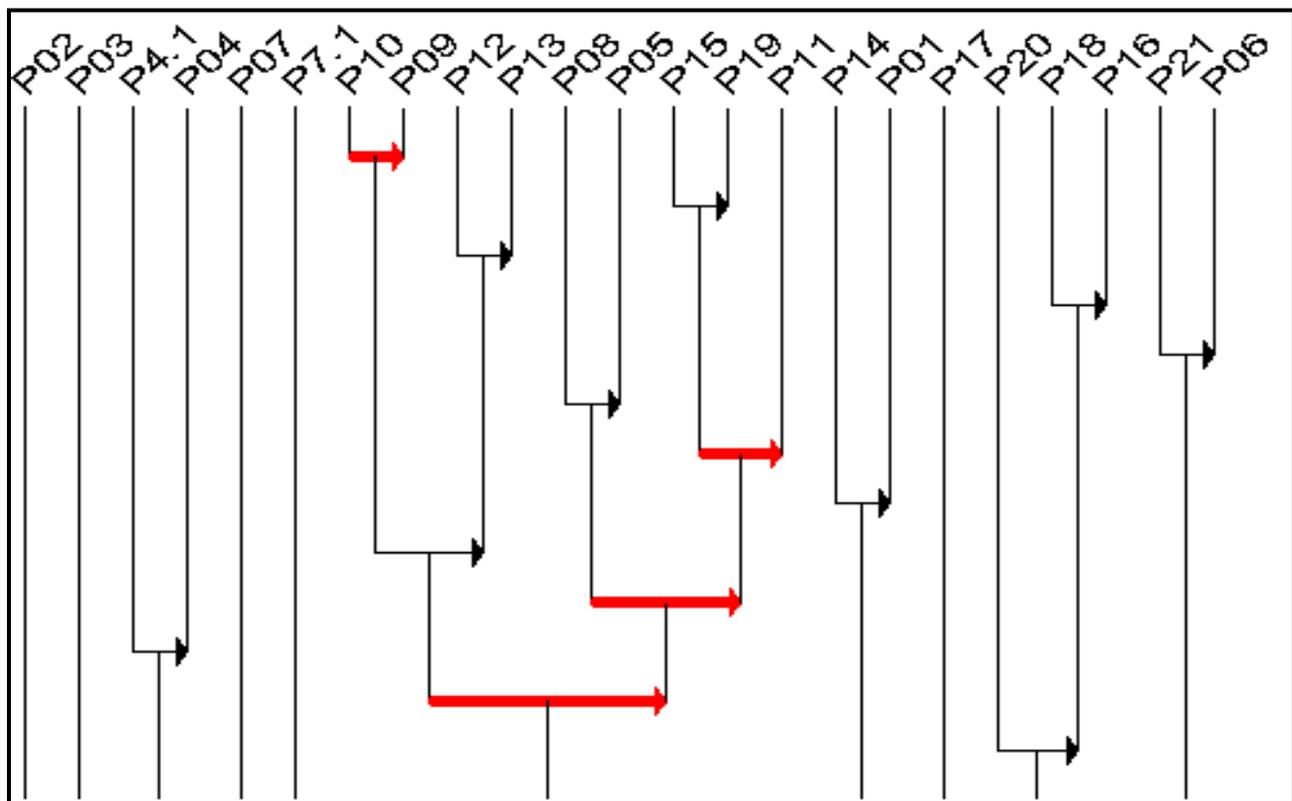


Fig. 5. Árbol cohesivo de la matriz MP.

\* Meta regla: reglas de reglas

Regla: significado semántico de una cuasi-implicación del tipo: "si  $a$  entonces  $b$ ". Es una especie de "teorema" vinculando una premisa y una conclusión (Ayuda online del C.H.I.C.).

P: “si el estudiante es capaz de calcular el porcentaje de alumnos que cumplen una cierta condición (P08), será capaz de realizar actividades enumerativas (P05)”.

Q: “si el estudiante es capaz de buscar las coincidencias (co-presencias y co-ausencias) entre columnas (P15), será capaz de realizar actividades que involucren la operación lógica de conjunción de negaciones (P19) y si es capaz de realizar esta última, será capaz de utilizar las coincidencias (co-presencias y co-ausencias) para identificar comportamientos similares entre columnas (P11)”.

Las operaciones que aparecen involucradas en la meta regla antes enunciada, estaban próximas en el árbol de similaridad con un nivel significativo.

El centro, grado y tipo de estudiante que contribuyen y son típicos a los niveles significativos en que se han formado las reglas con un riesgo menor a 0.261, vuelven a ser los alumnos del “Cuqui Bosch”, del 11<sup>no</sup> grado y los que fueron evaluados de excelente, respectivamente.

En la tabla 7 se muestra, a modo de ejemplo de salida del C.H.I.C, los valores de riesgos de la contribución y tipicidad de cada una de las variables suplementarias definidas, al nodo más significativo del árbol cohesitivo.

**Tabla 7.** Riesgos\* en la contribución y tipicidad de las variables suplementarias al nodo más significativo del árbol cohesitivo

Tipicidad de las var. Supl.			Contribución de las var. Supl.		
card GO 28	p 0.275	1-p 0.725	card GO 17	p 0.167	1-p 0.833
La variable PreC es típica a esta clase con un riesgo de: 0.00603			La variable PreC contribuye a esta clase con un riesgo de: 0.00709		
intersection avec le groupe optimal 22			intersection avec le groupe optimal 16		
La variable PreM es típica a esta clase con un riesgo de: 0.994			La variable PreM contribuye a esta clase con un riesgo de: 0.998		
intersection avec le groupe optimal 6			intersection avec le groupe optimal 1		
La variable Grado10 es típica a esta clase con un riesgo de: 1			La variable Grado10 contribuye a esta clase con un riesgo de: 0.994		
intersection avec le groupe optimal 3			intersection avec le groupe optimal 3		
La variable Grado11 es típica a esta clase con un riesgo de: 0.000279			La variable Grado11 contribuye a esta clase con un riesgo de: 0.036		
intersection avec le groupe optimal 25			intersection avec le groupe optimal 14		
La variable EstExce es típica a esta clase con un riesgo de: 1.27e-006			La variable EstExce contribuye a esta clase con un riesgo de: 2.84e-007		
intersection avec le groupe optimal 15			intersection avec le groupe optimal 13		
La variable EstBue es típica a esta clase con un riesgo de: 0.157			La variable EstBue contribuye a esta clase con un riesgo de: 0.671		
intersection avec le groupe optimal 8			intersection avec le groupe optimal 3		
La variable EstReg es típica a esta clase con un riesgo de: 0.999			La variable EstReg contribuye a esta clase con un riesgo de: 1		
intersection avec le groupe optimal 5			intersection avec le groupe optimal 1		
La variable EsrMalo es típica a esta clase con un riesgo de: 1			La variable EsrMalo contribuye a esta clase con un riesgo de: 1		
intersection avec le groupe optimal 0			intersection avec le groupe optimal 0		
La variable típica a esta clase es EstExce con un riesgo de: 1.27e-006			La variable que contribuye más a esta clase es EstExce con un riesgo de: 2.84e-007		

\* Un riesgo bajo de error, indica un alto grado de confianza en la contribución o tipicidad de los alumnos del grupo optimal.

### Grafo implicativo

El grafo implicativo de la matriz a posteriori MP en la figura 6, muestra las relaciones de implicación entre las preguntas del cuestionario de la experimentación, en el sentido de que el número de contraejemplos en los que cumpliéndose el antecedente se verifica el consecuente es significativamente bajo. En el grafo aparecen las implicaciones con un índice de cuasi implicación, según la teoría del Análisis Estadístico Implicativo, de 0.99 con trazos gruesos y de 0.95 con trazos más finos.

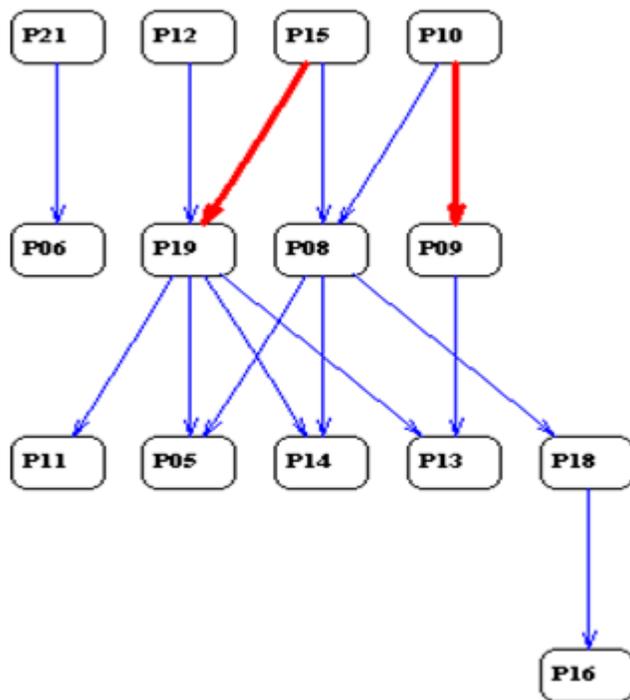


Fig. 6. Grafo implicativo de la matriz a posteriori MP.

Los orígenes y nodos del grafo, que conforman la estructuración de las cadenas de implicaciones, están marcados por preguntas con un porcentaje bajo de respuestas correctas.

Las preguntas que marcan las consecuentes son aquellas que exigen en sus respuestas la realización de operaciones de proximidad, pero todas ellas con un significado coloquial, o exigen de operaciones de búsqueda en una sub-tabla de la tabla binaria cuya respuesta resulta evidente visualmente.

La implicación  $P21 \rightarrow P06$  evidencia que si los alumnos saben realizar operaciones de búsqueda abierta en toda la tabla, sabrán realizar operaciones de proximidad que tengan un significado coloquial.

La cadena marcada por la pregunta P15 está estructurando a una gran parte de las preguntas del cuestionario (ver figura 6a). En ella se evidencia que si los alumnos dominan las operaciones de proximidad: co-presencias y co-ausencias (P15), son capaces de realizar correctamente operaciones lógicas de conjunción de negaciones (P19) o enumerativas que impliquen el cálculo de porcentajes (P08), y si dominan al menos una de estas últimas operaciones, serán capaces de realizar actividades que impliquen operaciones de proximidad que tengan un significado coloquial, o de búsqueda en una subtabla cuya solución resulte evidente visualmente. Por lo que podemos decir que *el lenguaje coloquial es condición suficiente para el trabajo de proximidad*.

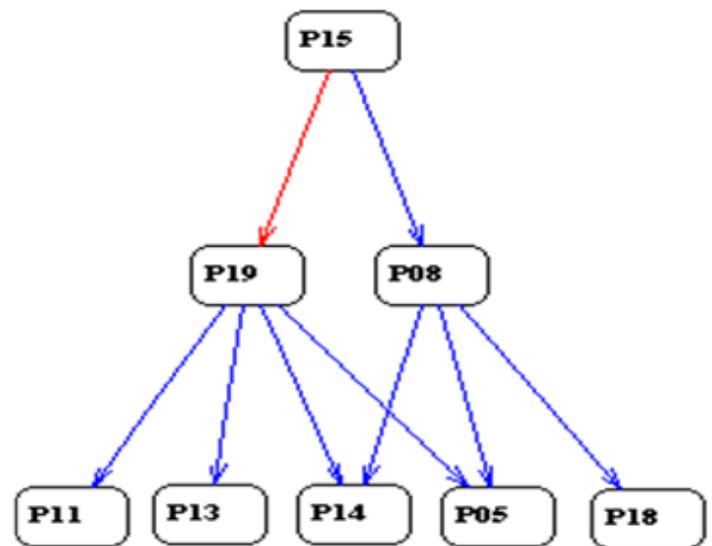


Fig. 6a. Subgrafo del grafo implicativo marcado por P15.

Lo explicado anteriormente pone de manifiesto la existencia de dos clases de alumnos, una que agrupa a los que priorizan el trabajo lógico, determinada por el camino  $P15 \rightarrow P19 \rightarrow \dots$ , y la otra a los que priorizan el trabajo enumerativo, caracterizada por el camino  $P15 \rightarrow P08 \rightarrow \dots$ .

Del grafo se observa además que la cadena de implicaciones con nodo en P19 posee también como antecedente la operación lógica de diferencia (P12), con un nivel de confiabilidad menor (Ver figura 6b). La diferencia de confiabilidad entre las implicaciones de antecedentes P15 y P12, con consecuente común P19, podría interpretarse por la complejidad matemático-estadística de la pregunta P15 (sobre *coincidencias*: co-

presencias y co-ausencias) respecto de la pregunta P12 (sobre *diferencias*) que no exige de un conocimiento matemático específico para poder responderse acertadamente.

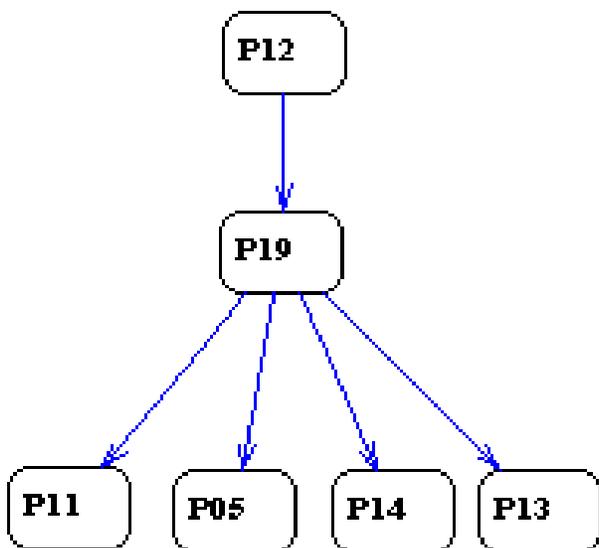


Fig. 6b. Subgrafo del grafo implicativo marcado por P12.

El camino  $P15 \rightarrow P08 \rightarrow P18 \rightarrow P16$  (ver figura 6c), marca que los alumnos que controlan muy bien las operaciones de proximidad: co-presencia y co-ausencia (P15), dominan operaciones de este tipo provistas de un significado coloquial (P16), si son capaces de transitar con éxito por las operaciones que implican el cálculo de porcentaje (P08) y de implicación lógica (P18), respectivamente.

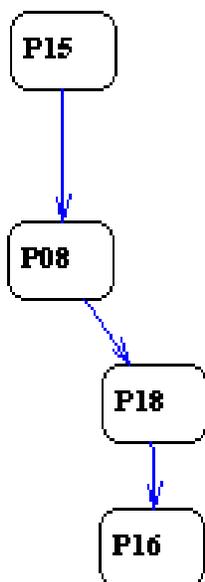


Fig. 6c. Subgrafo del grafo implicativo marcado por P15.

Por último, la cadena  $P10 \rightarrow P09 \rightarrow P13$  (ver figura 6d) está marcando que los alumnos que dominan el cálculo de la frecuencia relativa (P10), tendrán éxitos en el cálculo de la frecuencia absoluta (P09) y si tiene éxito en esta última, también lo tendrán con operaciones que exijan de la búsqueda en toda la tabla, cuya solución sea evidente visualmente (P13). Observar que el nodo caracterizado por la pregunta P08 posee también como antecedente a la pregunta P10, es decir, los alumnos con habilidades en el cálculo de la frecuencia relativa, muestran, aunque con un nivel de confiabilidad menor, habilidades en el cálculo de porcentajes.



Fig. 6d. Subgrafo del grafo implicativo marcado por P10.

Las preguntas P01, P02 y P03, que implican operaciones elementales de manejo de tabla, no aparecen como consecuentes en el grafo, ya que, al tener un alto porcentaje de éxitos en sus respuestas, no resultan significativas para conformar implicaciones a los niveles fijados, 0.99 y 0.95.

Todas las preguntas cuyas implicaciones aparecen en el grafo, ya aparecieron relacionadas en el árbol de similaridad o en el cohesitivo o en ambos.

Al analizar la contribución y la tipicidad de las variables suplementarias a los caminos del grafo implicativo, nuevamente se ha obtenido que el centro, grado y tipo de alumno que contribuyen y son típicos a los niveles significativos en que se han formado los caminos, con un riesgo menor a 0.261, vuelven a ser los alumnos del “Cuqui Bosch”, del 11<sup>no</sup> grado y los que fueron evaluados de excelente, respectivamente.

#### 4. CONCLUSIONES

El estudio realizado sobre las habilidades que poseen alumnos de 10<sup>mo</sup> y 11<sup>no</sup> grado de la enseñanza preuniversitaria del municipio de Santiago de Cuba, para el trabajo simultáneo con aspectos esenciales de Tratamiento de Datos, la Lógica y la Estadística Elemental, arrojó que las preguntas con mayores dificultades fueron las relacionadas con actividades estadísticas (P09 y P10), las que abordan la proximidad entre sujetos (P12, P15 y P20) y entre variables (P21) y las que se vinculan a operaciones lógicas (P18 y P19). Resultados muy próximos a los obtenidos por Pitarch con alumnos de la ESO, en cuanto a la lectura y manipulación elemental de la tabla.

En todas estas preguntas, los porcentajes más bajos de respuestas al cuestionario se obtuvieron en el centro “Rafael María de Mendive” y dentro de éste en el 10<sup>mo</sup> grado y las diferencias, en cuanto al porcentaje de respuestas correctas, resultaron ser significativas con un nivel de confiabilidad del 95%, por centro y/o por grados, para aquellas preguntas vinculadas a las actividades estadísticas (P09 y P10), de proximidad entre sujetos (P11, P12, P13, P15 y P20) y lógicas de conjunción de negaciones (P19).

Con la experimentación realizada se ha evidenciado que los alumnos poseen destrezas en el manejo básico de la tabla binaria, con un porcentaje de éxito mayor al 96%, lo cual muestra que es un instrumento didáctico de trabajo matemático “económico” para el profesor, ya que los alumnos saben trabajar con ella, a pesar de que las habilidades de manejo de la tabla no son homogéneas.

Analizando más finamente las habilidades y conocimientos mostrados por los alumnos se pudo comprobar que éstos poseen menos habilidades en el trabajo por filas que por columnas en la tabla binaria, ya que, pese a que los elementos fila-columna son comparables matemáticamente en una matriz o tabla, poseen diferente sentido para los alumnos: las columnas representan los sujetos (con sus respuestas) y las filas las preguntas del cuestionario (variables). A esta conclusión se puede llegar a través de la interpretación de las agrupaciones (clases D1 y D2) obtenidas en el árbol de similaridad y de las implicaciones del árbol cohesivo, donde las habilidades en la búsqueda por columnas son

condición necesaria para la búsqueda por filas y éstas suficientes para la búsqueda por columna.

Las cuestiones de mayor dificultad, sean estadísticas, lógicas o de proximidad, con un porcentaje de éxito en sus respuestas menor al 60% y que no tienen gran complejidad de resolución, son preguntas de extracción de conocimientos a partir de toda la tabla, y no solamente de submatrices o parejas de variables, lo cual muestra que la tabla de datos binarios se podría utilizar para mejorar las habilidades que los alumnos poseen en estas operaciones, proponiéndose actividades de complejidad creciente de búsqueda y de operaciones a realizar.

De los análisis realizados se ha podido observar que las preguntas que suponen un trabajo lógico no aparecen como determinantes en el dominio de las técnicas estadísticas y de tratamiento de datos, sino que más bien aparecen como diluidas en el trabajo de proximidad o comparación de datos. Como determinante aparecen las preguntas que exigen la búsqueda de diferencias y el mayor dominio de los conceptos estadísticos, no sólo numéricos y de porcentaje, sino también de coincidencias, ya que las co-presencias y co-ausencias son necesarias para la clasificación.

A partir de la aplicación de las técnicas del Análisis Estadístico Implicativo hemos mostrado como se relacionan el trabajo lógico y el estadístico a través del tratamiento de datos, lo cual se evidencia en las siguientes *reglas* y *meta reglas* obtenidas:

- “si el estudiante es capaz de calcular la frecuencia relativa, entonces es capaz de calcular la frecuencia absoluta” (R).
- “si el estudiante es capaz de hallar las diferencias entre columnas, será capaz de hallar las coincidencias (co-presencias y co-ausencias) entre columnas” (R).
- “si el estudiante es capaz de calcular el porcentaje de alumnos que cumplen una cierta condición, será capaz de realizar actividades enumerativas” (R).
- “si el estudiante es capaz de buscar las coincidencias (co-presencias y co-ausencias) entre columnas, será capaz de realizar actividades que involucren la operación lógica de conjunción de negaciones y si es capaz de realizar esta última, será capaz de utilizar las coincidencias (co-presencias y co-ausencias) para identificar comportamientos similares entre columnas (P11)” (MR).

- “si el estudiante domina las operaciones estadísticas, entonces es capaz de realizar las actividades que involucran operaciones de proximidad” (R).

Para la obtención de los resultados de esta investigación, el empleo del sistema informático CHIC, de implementación del Análisis Estadístico Implicativo, resultó ser un instrumento de gran valor, pues facilitó el perfeccionamiento del cuestionario, la caracterización de las preguntas según el tipo de operaciones a realizar y de las respuestas dadas por los alumnos, permitiendo además identificar asociaciones y estructuraciones de los datos (las preguntas, los sujetos, las variables suplementarias, etc.), que posibilitaron profundizar en la interpretación de los resultados y extraer más conocimientos de los datos.

## REFERENCIAS

Batanero, C. (1999): “Taller sobre análisis exploratorio de datos en la enseñanza secundaria”. Actas de la Conferencia Internacional *Experiencias e Expectativas do Ensino de Estatística- Desafios para o Século XXI*. Florianópolis, Santa Catarina, Brasil.

Batanero, C. (2003): “Los retos de la cultura estadística”. Universidad de Granada, España. <http://www.ugr.es/local/batanero>.

Belén, M. y Mónica, L. (2006): “Matemática y computación: desde un marco teórico-epistemológico a la práctica. Consideraciones básicas del entorno clic para docentes no informatizados”. Edeutec. Revista Electrónica de Tecnología Educativa, N°. 21.

Couturier, R., Bodin, A. y Gras, R. (2006): “Clasificación jerárquica, implicativa y cohesiva. Manual de usuario del sistema informático CHIC”.

Couturier, R., Bodin, A. y Gras, R. (2007) : Classification Hiérarchique Implicative et Cohésive, C.H.I.C. Software, versión 4.1.

Delgado, F. M., (1999) : “Contribución a la Difusión de la Cultura Estadística en Cuba”, Atas da Conferência Internacional *Experiências e Expectativas do Ensino de Estatística - Desafios para o Século XXI*, Florianópolis, Santa Catarina, Brasil.

Gorina, A., Alonso, I. y Zamora, L. (2007a): “La formación integral de los doctores en Ciencias Pedagógicas. Una mirada desde la educación estadística”. Actas del V Taller Internacional “*Innovación Educativa-Siglo XXI*” y Primer Simposio de la red de

investigación en Ciencias de la Educación para Iberoamérica. ISBN: 978-959-16-0551-1.

Gorina, A., Alonso, I. y Zamora, L. (2007b): “Un Sistema de Habilidades para la formación de una competencia estadística en los doctores en Ciencias Pedagógicas”, Materiales del V Congreso Internacional Virtual de Enseñanza de las Matemáticas, CVEM 2007.

Gras R., Kuntz P. et Briand H. (2001) : “Les fondements de l'analyse statistique implicative, Mathématiques et Sciences Humaines”, N° 154-155.

Guerrero, C. (1999): “La importancia de la información a priori”, Uno. Revista de Didáctica de las Matemáticas, N° 20, pp. 57-62.

Guerrero, S. (1999): “Tratamiento de la información”, Uno. Revista de Didáctica de las Matemáticas, N° 20, pp. 5-8.

Lerman, I., Gras, R., Rostam, H., (1981) : “Elaboration et évaluation d'un indice d'implication pour des données binaires”, I et II, Mathématiques et Sciences Humaines N° 74, pp. 5-35 et N° 75, pp. 5-47.

Orús, P. (1992) : “Le raisonnement des élèves dans la relation didactique; effets d'une initiation à l'analyse classificatoire dans la scolarité obligatoire”, Thèse présentée à l'Université de Bordeaux-I. (Ed: IREM de Bordeaux).

Orús, P., Pitarch, I. (2000) : “Utilisation didactique des tableaux des données et du logiciel CHIC à l'école élémentaire”, Actes des Journées sur La fouille dans les données par la méthode d'analyse statistique implicative. Caen, pp.85-98.

Pitarch, I. (2002): “Estudio sobre la viabilidad y el interés didáctico del tratamiento de la información en la ESO”, Trabajo de investigación del tercer ciclo presentado en la Universitat Jaume I y dirigido por Pilar Orús Báguena. (Ed: UJI, Castellón).

Zamora, L. y Alonso, I. (2006): “Metodología para la impartición de tópicos de Estadística y Probabilidades en la enseñanza preuniversitaria en Cuba”, Acta del RELME XX. Camagüey, Cuba, pp. 263-269.

Zamora, L., Alonso, I., Gorina, A., Álvarez, M. Orús, P. (2007): “Tratamiento de datos, lógica y estadística elemental en la enseñanza preuniversitaria”. Actas del 4to Encuentro Internacional de Análisis Estadístico Implicativo. Castellón, España, pp. 371-381.

## ANEXO 1: CUESTIONARIO DE LA EXPERIMENTACIÓN

### 1.1 La consigna

Estimado alumno, le solicitamos que observe con detenimiento la “Tabla T” que se le ha entregado y a partir de la información que ella le brinda, conteste, cuidadosamente, el cuestionario que se le presenta.

### 1.2 Tabla T

	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	E20	E21	
ASIGNATURAS	Te gusta la Matemática?	1	1	1	0	0	1	0	0	0	0	1	1	1	1	1	0	1	1	0	0	
	Te gusta el Español y la Literatura?	0	0	0	1	1	1	0	0	1	1	1	0	0	0	1	1	1	0	1	1	
	Te gusta la Historia?	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Te gusta la Geografía?	0	0	0	0	0	1	0	0	0	0	1	1	0	0	1	0	0	1	0	1	0
	Te gusta el Inglés?	1	1	1	1	1	0	0	0	1	1	0	1	1	1	0	1	0	0	1	0	0
	Te gusta la Física?	1	1	1	0	0	0	1	0	1	1	1	1	1	1	1	1	0	1	1	0	0
	Te gusta la Química?	1	1	0	0	0	0	0	0	0	1	1	1	0	1	1	0	0	1	1	0	0
	Te gusta la Biología?	0	0	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1
	Te gusta la Computación?	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Te gusta la Cultura Política?	0	0	0	0	1	0	1	1	0	0	1	1	0	1	1	0	0	1	0	1	0
	Te gusta la Educación Física?	0	1	0	0	0	1	0	1	1	1	1	1	1	1	1	0	0	1	1	1	0
	Te gusta la Preparación Militar?	0	0	0	0	0	1	0	1	1	1	1	1	0	1	1	0	0	1	1	1	0
CARRERAS	Te gustaría estudiar Medicina?	0	0	0	1	0	1	1	0	1	1	1	0	0	1	1	1	1	0	0	1	1
	Te gustaría estudiar Matemática?	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0
	Te gustaría estudiar Informática?	1	1	1	0	0	0	0	0	0	0	0	0	1	0	1	0	0	1	1	1	0
	Te gustaría estudiar Idioma?	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Te gustaría estudiar Historia?	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
	Te gustaría estudiar Cultura Física?	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0
	Te gustaría estudiar una Carrera Militar?	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
	Te gustaría estudiar alguna Ingeniería?	1	1	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	0

### 1.3 Las preguntas

- P1. ¿Al estudiante E2 le gusta la asignatura Matemática?
- P2. ¿A cuántos estudiantes les gusta la asignatura de Matemática?
- P3. ¿Cuántas asignaturas le gustan al estudiante E10?
- P4. ¿Hay alguna asignatura que haya sido elegida por todos los estudiantes? ¿En caso afirmativo, indique cuál?
- P4.1 ¿Hay otras? En caso afirmativo, indique cuáles.
- P5. Los estudiantes entrevistados ¿prefieren la asignatura de Matemática o la de Química?
- P6. ¿Hay algún estudiante a quien le guste todas las asignaturas? En caso afirmativo, indique su código.
- P7. ¿Existe alguna asignatura de mayor preferencia que la Matemática?
- P7.1 Si hay alguna, indíquela y razone su respuesta.
- P8. ¿A qué porcentaje de los estudiantes que les gusta la asignatura de Biología, le gustaría estudiar la carrera de Medicina?
- P9. Si consideramos la asignatura de Física como una variable estadística, ¿qué frecuencia absoluta diría usted que tiene en esta distribución estadística?



**Abstract:** In this work, there is a study on the skills of students from 10th and 11th grades of pre-university for the simultaneous work with essential aspects of Data Treatment, Logics and Elementary Statistics. In the interpretation of the information obtained in a survey and presented in a table of binary data, taking as reference the works of Pitarch (2002) and Orús (1992), carried out with students from the ESO in Spain and primary school in France respectively. By applying the techniques of Implicative Statistical Analysis, this article shows the relationship between the logical and statistical work through the treatment of data, detecting significant differences in the acquired skills for students of two pre-university schools taken as sample for the experimentation. Finally, it is presented, with help of the ISA, a set of rules that show the relations between the logical and statistical operations and the treatment of data.