



ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO
FACULTAD DE CIENCIAS
CARRERA MATEMÁTICA

**ANÁLISIS Y PROGRAMACIÓN DE LA CUASI-IMPLICACIÓN DE
GRAS ENTRE VARIABLES**

Trabajo de Integración Curricular

Tipo: Proyecto de Investigación

Presentado para optar al grado académico de:

MATEMÁTICO

AUTOR: BRYAN FRANCISCO ILBAY CHAVEZ

DIRECTOR: Dr. RUBÉN PAZMIÑO MAJI, PhD

Riobamba – Ecuador

2023

©2023, Bryan Francisco Ilbay Chavez

Se autoriza la reproducción total o parcial, con fines académicos, por cualquier medio o procedimiento, incluyendo la cita bibliográfica del documento, siempre y cuando se reconozca el Derecho de Autor.

Yo, BRYAN FRANCISCO ILBAY CHAVEZ, declaro que el presente Trabajo de Integración Curricular es de mi autoría y los resultados del mismo son auténticos. Los textos en el documento que provienen de otras fuentes están debidamente citados y referenciados.

Como autor asumo la responsabilidad legal y académica de los contenidos de este Trabajo de Integración Curricular; El patrimonio intelectual pertenece a la Escuela Superior Politécnica de Chimborazo.

Riobamba, 14 de Abril de 2023

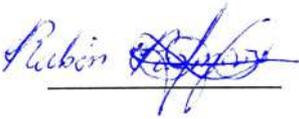


Bryan Francisco Ilbay Chavez

060453566-6

ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO
FACULTAD DE CIENCIAS
CARRERA MATEMÁTICA

El Tribunal del Trabajo de Integración Curricular certifica que: El Trabajo de Integración Curricular; Tipo: Proyecto de Investigación, **ANÁLISIS Y PROGRAMACIÓN DE LA CUASI-IMPLICACIÓN DE GRAS ENTRE VARIABLES**, realizado por el señor: **BRYAN FRANCISCO ILBAY CHAVEZ**, ha sido minuciosamente revisado por los Miembros del Tribunal del Trabajo de Integración Curricular, el mismo que cumple con los requisitos científicos, técnicos, legales, en tal virtud el Tribunal Autoriza su presentación.

	FIRMA	FECHA
Ing. María de Lourdes Palacios Robalino PRESIDENTE DEL TRIBUNAL		2023-04-14
Dr. Rubén Antonio Pazmiño Maji, Phd DIRECTOR DEL TRABAJO DE INTEGRACIÓN CURRICULAR		2023-04-14
Dr. Jorge Washington Congacha Aushay, MSc ASESOR DEL TRABAJO DE INTEGRACIÓN CURRICULAR		2023-04-14

DEDICATORIA

Dedico este trabajo de titulación con mucha fé, amor y respeto a Dios ya que, con su cuidado, fortaleza y sabiduría me ha sabido encaminar en cada momento de mi vida. A mis padres Víctor Ilbay y Anita Chavez por ser mis consejeros de vida, sobre todo por ser mi ejemplo de constancia, dedicación y fortaleza diaria; también quiero agradecerles por inculcar valores en mí, los cuales han hecho de mí una persona de bien y siendo ellos el eje principal de mi vida. A mi hermano Jhon Paúl Ilbay que desde el cielo me ha sabido guiar y acompañar en cada ámbito de mi vida, con su bendición cumplir todas las promesas propuestas. A mis tíos paternos y maternos por el cariño y apoyo que me han sabido demostrar en todos estos años, demostrándome que cada esfuerzo tiene su recompensa. A mis tías paternas y maternas por cada uno de los consejos brindados y por el amor demostrado haciéndome sentir un hijo más en su familia. A mis primos Adrián, Josué y Nataly por ser mis compañeros de vida y por demostrarme su lealtad en todo momento. A mi hermano de corazón Sebastián por estar siempre pendiente de mí y con sus travesuras alegrarme la vida. A Anabel Córdova por ser mi mejor amiga, confidente, consejera y acompañante en todo el trayecto de mi vida universitaria. A Andrea Nata, por haber sido mi consejera en todo el tiempo que compartimos en la universidad que desde el cielo seguirá siendo una gran amiga.

Bryan

AGRADECIMIENTO

Un profundo agradecimiento a Dios, quién me ha brindado salud y sabiduría para poder culminar con éxito la tan anhelada carrera, a mi familia por el esfuerzo y apoyo incondicional para poder cumplir esta gran meta propuesta.

A la Escuela Superior Politécnica de Chimborazo por haberme formado éticamente con los mas altos estándares académicos, para poder servir a la sociedad de forma responsable, de la misma manera a los docentes quienes se han esforzado para brindarnos una educación de calidad.

De manera especial, brindar un sincero agradecimiento al Dr. Rubén Pazmiño Maji, tutor del presente trabajo de titulación, por su tiempo y ayuda absoluta al impartir los conocimientos teóricos y metodológicos necesarios de una manera excepcional.

Bryan

ÍNDICE DE CONTENIDOS

ÍNDICE DE TABLAS	ix
ÍNDICE DE ILUSTRACIONES	x
RESUMEN	xii
ABSTRACT	xiii
RESUMEN	xiv
ABSTRACT	xv
INTRODUCCIÓN	1

CAPÍTULO I

1. PROBLEMA DE INVESTIGACIÓN	2
1.1. Planteamiento del Problema	2
1.2. Objetivos	3
1.2.1. <i>Objetivo General</i>	3
1.2.2. <i>Objetivo Específicos</i>	3
1.3. Justificación	3

CAPÍTULO II

2. MARCO TEÓRICO	4
2.1. Antecedentes de investigación	4
2.1.1. <i>Análisis Estadístico Implicativo</i>	4
2.1.2. <i>La cuasi-implicación de Regis Gras</i>	4
2.1.3. <i>Ánalysis cohesivo</i>	6
2.1.4. <i>Nodos significativos</i>	6
2.1.5. <i>Otros métodos no ASI similares a la Cuasi-Implicación</i>	7
2.1.6. <i>Formulación Matemática</i>	8
2.1.7. <i>Ejemplo detallado de aplicación</i>	12
2.2. La programación en MATLAB	23
2.2.1. <i>Variables</i>	24
2.2.2. <i>Funciones</i>	24

2.2.3.	<i>Comandos de programación</i>	25
2.2.4.	<i>Programación en objetos en MATLAB</i>	26

CAPÍTULO III

3.	MARCO METODOLÓGICO	28
3.1.	Enfoque de investigación	28
3.2.	Nivel de investigación	28
3.3.	Diseño de la investigación	28
3.3.1.	<i>Según la manipulación o no de la variable independiente</i>	28
3.3.2.	<i>Según las intervenciones en el trabajo de campo</i>	29
3.4.	Tipo de estudio	29
3.5.	Población y Planificación, selección y cálculo del tamaño de la muestra	29
3.6.	Métodos, técnicas e instrumentos de investigación	29

CAPÍTULO IV

4.	MARCO DE ANÁLISIS E INTERPRETACIÓN DE RESULTADOS	30
4.1.	Procesamiento, análisis e interpretación de resultados	30
4.1.1.	<i>Elaborar el seudocódigo de la cuasi-implicación de Gras</i>	30
4.1.2.	<i>Automatización en R de una Base de Datos. Análisis Estadístico Implicativo</i>	34
4.1.3.	<i>Programación en MATLAB. Algoritmo del Análisi Estadístico Implicativo</i>	39
4.1.4.	<i>Programación en MATLAB. Algoritmo del Análisi Estadístico Implicativo para una base de 7 variables y 20 filas.</i>	39
4.1.5.	<i>Programación en MATLAB. Algoritmo del Análisi Estadístico Implicativo para una base de 7 variables y 200 filas.</i>	44
4.1.6.	<i>Programación en MATLAB. Algoritmo del Análisi Estadístico Implicativo para una base aleatoria de m variables y n filas.</i>	48
4.2.	Discusión	51
4.2.1.	<i>Estadísticos descriptivos: Matriz Cohesión (M-Chic y R-Chic), Matriz Cohesión (M-Chic y Chic), Gráfica (R-Chic y Chic), Nodos (M-Chic y R-Chic), Nodos (M-Chic y Chic)</i>	51
4.2.2.	<i>Gráfica de intervalos de Matriz Cohesión (M-Chic y R-Chic), Matriz Cohesión (M-Chic y Chic), Gráfica (R-Chic y Chic), Nodos (M-Chic y R-Chic), Nodos (M-Chic y Chic)</i>	74

4.3.	Comprobación de hipótesis	79
4.3.1.	<i>T de una muestra: Matriz Cohesión (M-Chic y R-Chic), Matriz Cohesión (M-Chic y Chic), Gráfica (R-Chic y Chic), Nodos (M-Chic y R-Chic), Nodos (M-Chic y Chic)</i>	79

CAPÍTULO V

5.	CONCLUSIONES Y RECOMENDACIONES	92
5.1.	CONCLUSIONES	92
5.1.1.	<i>Relación con objetivos, hipótesis y problema</i>	92
5.1.2.	<i>Sobre los objetivos específicos</i>	92
5.1.3.	<i>Sobre la hipótesis</i>	94
5.1.4.	<i>Sobre el problema</i>	95
5.2.	RECOMENDACIONES	95

BIBLIOGRAFÍA

ÍNDICE DE TABLAS

Tabla 1-2: Matriz de datos binarios	13
Tabla 2-2: Ejemplo del Valor de la Probabilidad de éxito	14
Tabla 3-2: Valores de la probabilidad de éxito	15
Tabla 4-2: Ejemplo de la $Card(LM, GA)$	16
Tabla 5-2: Valores de la Cardinalidad	17
Tabla 6-2: Ejemplo de la distribución Binomial	18
Tabla 7-2: Valores de la distribución Binomial	18
Tabla 8-2: Ejemplo de la Intensidad Implicativa	19
Tabla 9-2: Valores de la intensidad Implicativa	19
Tabla 10-2: Valores de Entropía	21
Tabla 11-2: Valores de la Matriz Cohesión de nivel 0	22
Tabla 12-2: Ejemplo de la Cohesión de nivel 0	22
Tabla 13-2: Valores de la Matriz Cohesión de nivel 1	23
Tabla 1-4: Comparaciones Estadísticas 1	51
Tabla 2-4: Comparaciones Estadísticas 2	52
Tabla 3-4: Comparaciones Estadísticas 3	53
Tabla 4-4: Comparaciones Estadísticas descriptivas	79
Tabla 5-4: Prueba	79
Tabla 6-4: Comparaciones Estadísticas descriptivas: μ media entre las Gráficas de Rchic-Chic.	88

ÍNDICE DE ILUSTRACIONES

Ilustración 1-2: Grafo Implicativo	20
Ilustración 2-2: Árbol de Cohesión	23
Ilustración 1-4: Diagrama de Flujo Pt.1.	31
Ilustración 2-4: Diagrama de Flujo Pt.2	32
Ilustración 3-4: Diagrama de Flujo Pt.3.	33
Ilustración 4-4: Paquetes en R del Análisis Estaístico Implicativo.	34
Ilustración 5-4: Script en el entorno de R.	35
Ilustración 6-4: Ventana de RChic.	35
Ilustración 7-4: Ventana Selección de Archivo.	36
Ilustración 8-4: Ventana Implicative Graph.	36
Ilustración 9-4: Ventana Grafo Implicativo con Valores.	37
Ilustración 10-4: Ventana de variables.	37
Ilustración 11-4: Ventana Árbol de Cohesión.	38
Ilustración 12-4: Ventana Environment.	38
Ilustración 13-4: Ventana Matriz de Cohesión.	39
Ilustración 14-4: Algoritmo de una base aleatoria de 7 variables y 20 filas.	42
Ilustración 15-4: Grafo Implicativo MATLAB de una base aleatoria de 7 variables y 20 filas.	43
Ilustración 16-4: Grafo Implicativo R de una base aleatoria de 7 variables y 20 filas.	43
Ilustración 17-4: Algoritmo de una base aleatoria de 7 variables y 200 filas.	47
Ilustración 18-4: Grafo Implicativo MATLAB de una base aleatoria de 7 variables y 200 filas.	47
Ilustración 19-4: Grafo Implicativo R de una base aleatoria de 7 variables y 200 filas.	48
Ilustración 20-4: Algoritmo de una base aleatoria de m variables y n filas.	50
Ilustración 21-4: Grafo Implicativo R de una base aleatoria de m variables y n filas.	51
Ilustración 22-4: Histograma Matriz de Cohesión M-Chic y R-Chic	54
Ilustración 23-4: Histograma Matriz de Cohesión M-Chic y Chic.	55
Ilustración 24-4: Histograma Gráficas M-Chic y R-Chic.	56
Ilustración 25-4: Histograma Nodos de R-Chic y Chic.	57
Ilustración 26-4: Histograma Nodos de M-Chic y Chic.	58
Ilustración 27-4: Histograma (con curva normal) de Matriz de Cohesión M-Chic y R-Chic.	59
Ilustración 28-4: Histograma (con curva normal) Matriz de Cohesión M-Chic y Chic.	60
Ilustración 29-4: Histograma (con curva normal) de Gráfica R-Chic y Chic.	61

Ilustración 30-4:	Histograma (con curva normal) de Nodos M-Chic y R-Chic.	62
Ilustración 31-4:	Histograma (con curva normal) de Nodos M-Chic y Chic.	63
Ilustración 32-4:	Gráfica de valores individuales de la Matriz Cohesión M-Chic y R-Chic.	64
Ilustración 33-4:	Gráfica de valores individuales de la Matriz Cohesión M-Chic y Chic. .	65
Ilustración 34-4:	Gráfica de valores individuales de las Gráficas R-Chic y Chic.	66
Ilustración 35-4:	Gráfica de valores individuales de los Nodos M-Chic y R-Chic.	67
Ilustración 36-4:	Gráfica de valores individuales de los Nodos M-Chic y Chic.	68
Ilustración 37-4:	Gráfica de caja de la Matriz de Cohesión M-Chic y R-Chic.	69
Ilustración 38-4:	Gráfica de caja de la Matriz de Cohesión M-Chic y Chic.	70
Ilustración 39-4:	Gráfica de caja de los Gráficos R-Chic y Chic.	71
Ilustración 40-4:	Gráfica de caja de los Nodos M-Chic y R-Chic.	72
Ilustración 41-4:	Gráfica de caja de los Nodos M-Chic y Chic.	73
Ilustración 42-4:	Gráfica de intervalos de Matriz Cohesión M-Chic y R-Chic.	74
Ilustración 43-4:	Gráfica de intervalos de Matriz Cohesión M-Chic y Chic.	75
Ilustración 44-4:	Gráfica de intervalos de los Gráficos ón R-Chic y Chic.	76
Ilustración 45-4:	Gráfica de intervalos de los Nodos M-Chic y R-Chic.	77
Ilustración 46-4:	Gráfica de intervalos de los Nodos M-Chic y Chic.	78
Ilustración 47-4:	Histograma de la Matriz de Cohesión M-Chic y R-Chic.	80
Ilustración 48-4:	Histograma de la Matriz de Cohesión M-Chic y Chic.	81
Ilustración 49-4:	Histograma de Nodos M-Chic y R-Chic.	82
Ilustración 50-4:	Valores individuales de la Matriz de Cohesión M-Chic y R-Chic.	83
Ilustración 51-4:	Gráfica de los valores individuales de la Matriz de Cohesión M-Chic y Chic.	84
Ilustración 52-4:	Gráfica de los valores individuales de los Nodos M-Chic y R-hic.	85
Ilustración 53-4:	Gráfica de caja de la Matriz de Cohesión M-Chic y R-Chic	86
Ilustración 54-4:	Gráfica de caja de la Matriz de Cohesión M-Chic y Chic.	87
Ilustración 55-4:	Gráfica de caja de los Nodos M-Chic y R-Chic.	88
Ilustración 56-4:	Histograma de las Gráficas R-Chic y Chic.	89
Ilustración 57-4:	Gráfica de los valores individuales de la Gráfica R-Chic y Chic.	90
Ilustración 58-4:	Gráfica de caja de las Gráficas R-Chic y Chic.	91

RESUMEN

El objetivo de esta investigación es realizar un análisis y programación de la Cuasi-implicación de Régis Gras en el software Matlab. La Cuasi-Implicación fue creado por Régis Gras y es uno de los procesos usados en el Análisis Estadístico Implicativo. La Cuasi-implicación nos permite determinar relaciones de causa y efecto (con un número pequeño de excepciones). La aplicación del análisis jerárquico de similitudes, busca construir particiones cada vez mas finas, en el conjunto de variables, construidas de forma ascendente en un árbol cohesivo, mediante el criterio de similitud entre las mismas. La técnica de clustering es usada para explorar dentro de un conjunto de datos y agruparlos en cluster o grupos más pequeños y que tengan objetos o individuos con características similares. El concepto central del clustering y en el que se basa todo es la llamada similaridad entre los objetos que se están agrupando. Esta investigación es de tipo cuantitativo ya que para realizar la programación primero se hizo una análisis de la formulación matemática de la Cuasi-implicación de Régis Gras, luego se procedió ha hacer la programación la cual dio paso al análisis de las 383 bases aleatorias de datos binarios, en la salida del programa se presentan las matrices de similaridad y sus diferentes niveles, también muestra los nodos significativos.. El trabajo tiene un tipo de diseño pre-experimental pues ya existe una programación hecha tanto en en software CHIC y software R-Chic, en las cuales nos basamos para crear y comparar la programación Mchic. Como resultados se obtuvo una similaridad mayor al 90 %, es decir que la programación que se elaboró en Matlab (Mchic) puede ser una herramienta favorable con la cual pueda trabajar el ASI, por otro lado, los usuarios que estén familiarizados con Matlab pueden usar Mchic sin problema alguno.

Palabras clave: <MATEMÁTICA>, <ANÁLISIS ESTADÍSTICO>, <CUASI IMPLICACIÓN>, <PROGRAMACIÓN>, <MATLAB>, <DATOS BINARIOS>, <DIDÁCTICA>.

0869-DBRA-UPT-2023



SUMMARY/ABSTRACT

The aim of this research is to perform an analysis and programming of Régis Gras' Quasi-implication in Matlab software. Quasi - implication was created by Régis Gras and is one of the processes used in Implicative Statistical Analysis. Quasi-implication allows us to determine cause and effect relationships (with a small number of exceptions). The application of hierarchical similarity analysis seeks to construct increasingly finer partitions in the set of variables, built upwards in a cohesive tree, by means of similarity criteria between them. The clustering technique is used to explore inside a data set and group it into clusters or smaller groups that have similar targets or individuals with similar characteristics. The central concept of clustering and what it is all based on is the so-called similarity between the objects that are being grouped. This research is of quantitative type since to carry out the programming first an analysis of the mathematical formulation of the quasi-implication of Régis Gras was made, then we proceeded to make the programming which gave way to the analysis of the 383 random bases of binary data, in the output of the program the similarity matrices and their different levels are presented, it also shows the significant nodes. The work has a type of pre-experimental design because there is already a programming done in both CHIC and R-Chic software, on which we based ourselves to create and compare the Mchic programming. As results we obtained a similarity greater than 90 %, that is to say that the programming that was elaborated in Matlab (Mchic) can be a favorable tool with which the user can work, on the other hand, the users who are familiar with Matlab can use Mchic without any problem.

Keywords: <MATHEMATICS>, <STATISTICAL ANALYSIS>, <QUASI IMPLICATION>, <PROGRAMMING>, <MATLAB>, <BINARY DATA>, <DIDACTICS>.

Dra. Nanci M. Inca Ch. Mgs.

0602926719

RESUMEN

El objetivo de esta investigación es realizar un análisis y programación de la Cuasi-implicación de Régis Gras en el software Matlab. La Cuasi-Implicación fue creado por Régis Gras y es uno de los procesos usados en el Análisis Estadístico Implicativo. La Cuasi-implicación nos permite determinar relaciones de causa y efecto (con un número pequeño de excepciones). La aplicación del análisis jerárquico de similitudes, busca construir particiones cada vez mas finas, en el conjunto de variables, construidas de forma ascendente en un árbol cohesivo, mediante el criterio de similitud entre las mismas. La técnica de clustering es usada para explorar dentro de un conjunto de datos y agruparlos en cluster o grupos más pequeños y que tengan objetos o individuos con características similares. El concepto central del clustering y en el que se basa todo es la llamada similaridad entre los objetos que se están agrupando. Esta investigación es de tipo cuantitativo ya que para realizar la programación primero se hizo una análisis de la formulación matemática de la Cuasi-implicación de Régis Gras, luego se procedió a hacer la programación la cual dio paso al análisis de las 383 bases aleatorias de datos binarios, en la salida del programa se presentan las matrices de similaridad y sus diferentes niveles, también muestra los nodos significativos.. El trabajo tiene un tipo de diseño pre-experimental pues ya existe una programación hecha tanto en software CHIC y software R-Chic, en las cuales nos basamos para crear y comparar la programación Mchic. Como resultados se obtuvo una similaridad mayor al 90 %, es decir que la programación que se elaboró en Matlab (Mchic) puede ser una herramienta favorable con la cual pueda trabajar el ASI, por otro lado, los usuarios que estén familiarizados con Matlab pueden usar Mchic sin problema alguno.

Palabras clave: <MATEMÁTICA>, <ANÁLISIS ESTADÍSTICO>, <CUASI IMPLICACIÓN>, <PROGRAMACIÓN>, <MATLAB>, <DATOS BINARIOS>, <DIDÁCTICA>.

0869-DBRA-UPT-2023

SUMMARY/ABSTRACT

The aim of this research is to perform an analysis and programming of Régis Gras' Quasi-implication in Matlab software. Quasi - implication was created by Régis Gras and is one of the processes used in Implicative Statistical Analysis. Quasi-implication allows us to determine cause and effect relationships (with a small number of exceptions). The application of hierarchical similarity analysis seeks to construct increasingly finer partitions in the set of variables, built upwards in a cohesive tree, by means of similarity criteria between them. The clustering technique is used to explore inside a data set and group it into clusters or smaller groups that have similar targets or individuals with similar characteristics. The central concept of clustering and what it is all based on is the so-called similarity between the objects that are being grouped. This research is of quantitative type since to carry out the programming first an analysis of the mathematical formulation of the quasi-implication of Régis Gras was made, then we proceeded to make the programming which gave way to the analysis of the 383 random bases of binary data, in the output of the program the similarity matrices and their different levels are presented, it also shows the significant nodes. The work has a type of pre-experimental design because there is already a programming done in both CHIC and R-Chic software, on which we based ourselves to create and compare the Mchic programming. As results we obtained a similarity greater than 90 %, that is to say that the programming that was elaborated in Matlab (Mchic) can be a favorable tool with which the user can work, on the other hand, the users who are familiar with Matlab can use Mchic without any problem.

Keywords: <MATHEMATICS>, <STATISTICAL ANALYSIS>, <QUASI IMPLICATION>, <PROGRAMMING>, <MATLAB>, <BINARY DATA>, <DIDACTICS>.

Dra. Nanci M. Inca Ch. Mgs.

0602926719

INTRODUCCIÓN

Para comenzar, tengamos en cuenta que la sociedad genera datos de manera masiva todos los días y en consecuencia de esto, la matemática en relación con la estadística, busca la manera de analizar dichos datos y obtener información comprensible y manejable para las personas.

Así mismo, se ve la necesidad de trabajar con metodologías tecnológicas capaces de trabajar, con grandes bases de datos que están fuera del alcance de la capacidad humana, por lo que, dentro del método de la Minería de Datos nos encontramos con la herramienta denominada Análisis Estadístico Implicativo, una herramienta creada por Régis Gras hace 48 años y es conocido como ASI por *Analyse Statistique Implicative*, en francés. La herramienta tiene sus investigaciones desarrolladas en el software CHIC y a la vez tiene una única plataforma libre llamada RCHIC.

CHIC Y RCHIC contienen funciones comunes como: Árbol de Similitud, Gráfico Implicativo, Árbol de Cohesión y Reducción que son parte del ASI. La cuasi-implicación se basa en un índice estadístico que expresa la implicación entre variables del trabajo propuesto por Régis Gras y mide la intensidad de implicación con un grafo implicativo donde se represente la relación entre variables.

Por lo que, en este documento se presenta una propuesta de trabajo de titulación sobre la implementación del software Matlab como otra plataforma para ASI, debido a que Matlab tiene una capacidad impresionante para trabajar con matrices, buscando así desarrollar un programa que trabaje de manera automática la Cuasi-Implicación de Gras y sea mas amigable con el usuario.

CAPÍTULO I

1. PROBLEMA DE INVESTIGACIÓN

1.1. Planteamiento del Problema

La sociedad actual está generando datos a un ritmo exponencial, requiriendo herramientas, procesos y nuevos métodos analíticos. El conocimiento subyacente de estos datos se extrae y se adapta a modelos comprensible para el ser humano, para que la informaciónn futura se pueda predecir a partir de la información observada previamente. El Análisis Estadístico Implícitivo ha surgido como un conjunto de herramientas que se pueden ensamblar a través de la minería de datos, con el objetivo de descubrir significados entre eventos y variables en conjuntos de datos para la Didáctica de la Matemática.

El Análisis Estadístico Implicativo tiene su versión en el software R con el paquete Chic, tiene una utilización de manera estadística y no es muy compatible con los métodos de programación que son utilizados hoy en día. Por lo cual se implementara una programación en Matlab que nos permita estar a la mano con los nuevos métodos de programación ya que el Analisis Estadístico Implicativo está siendo utilizado en varios ámbitos de las Ingenierías y la Medicina, por otro lado, para muchos se les hace más fácil manejar el Software Matlab ya que se lo ha venido usando por muchos años pero en Matlab no se encuentra automatizado el ASI por lo que muchos prefieren buscar otro método de análisis de datos binarios. Así que en este trabajo pretendemos presentar la programación de la Similaridad de Lerman que viene siendo una parte del ASI y que tenga una similaridad del 70 %.

En este proyecto de investigación sólo vamos a analizar y programar la Cuasi-Implicación de Gras que es parte del ASI. Para saber si la programación funciona de manera correcta para cualquier rango de variables se va a trabajar con un colectivo de estudio conformado por 100 000 bases aleatorias de datos binarios. Y solamente se hará el análisis con datos binarios pues se pretende hacer la programación exclusivamente para que ese tipo de datos. Los datos estarán en archivos con extensión CSV debido a que como nos basaremos en la programación de RChic pues este trabaja con archivos de ese tipo de extensión y como el Software Matlab puede leer tranquilamente este tipo de archivos entonces mejoraría la experiencia de los usuarios al momento de usar la programación.

1.2. Objetivos

1.2.1. *Objetivo General*

Programar en Matlab la Cuasi Implicación de Régis Gras entre variables binarias y validarla mediante la similaridad.

1.2.2. *Objetivo Específicos*

- Analizar la Cuasi Implicación de Régis Gras.
- Determinar la formulación matemática la Cuasi Implicación de Régis Gras.
- Elaborar el seudocódigo de la Cuasi Implicación de Régis Gras.
- Programar la Cuasi Implicación de Régis Gras.
- Estudiar la similaridad entre los resultados del programa Matlab (MChic) y el paquete RChic.

1.3. Justificación

La necesidad de trabajar con metodologías tecnológicas capaces de trabajar, con grandes bases de datos que están fuera del alcance de la capacidad humana, por lo que, dentro del método de la Minería de Datos nos encontramos con la herramienta denominada Análisis Estadístico Implicativo, una herramienta creada por Régis Gras hace 48 años y es conocido como ASI por *Analyse Statistique Implicative*, en francés. La herramienta tiene sus investigaciones desarrolladas en el software CHIC y a la vez tiene una única plataforma libre llamada RCHIC.

La investigación que se propone es de tipo cuantitativa, considerando que el objetivo del trabajo de titulación es estudio y generación de un seudocódigo de la cuasi-implicación de Gras. La investigación cuantitativa ya que nos permite recolectar y analizar datos numéricos, la investigación transversal nos ayuda a analizar datos de variables recopiladas en un período de tiempo sobre una población predefinida y la investigación pre experimental de la forma RGXO1, donde RG es el grupo de base de datos estadísticas aleatorias obtenidas mediante un muestreo aleatorio simple. X es el tratamiento (la ejecución) y O1 es es la correlación del programa diseñado en el respectivo de RCHIC, por lo tanto se tomara en cuenta el paquete RCHIC para lograr una programación similar en MATLAB.

CAPÍTULO II

2. MARCO TEÓRICO

2.1. Antecedentes de investigación

2.1.1. *Análisis Estadístico Implicativo*

En 1980, el francés Régis Gras crea el Análisis Estadístico Implicativo ASI (Analyse Statistique Implicative), es un método de análisis de datos asimétricos que permite la extracción y construcción de conocimiento a partir de conjuntos de datos que conectan un conjunto de objetos u objetos con un conjunto de variables en forma de normas y reglas generalizadas, ya partir de la contingencia de estas reglas, la interpretación y, por tanto, cierta predicción sobre los campos del conocimiento (Gras, & Kuntz, 2009).

El Análisis Estadístico Implícitivo es una herramienta de minería de datos basada en técnicas estadísticas multivariadas, teoría de cuasi-implicación, inteligencia artificial y álgebra booleana para modelar la cuasi-implicación entre eventos y variables en un conjunto de datos (Gras, & Kuntz, 2009). El trabajo de Régis Gras es ampliamente considerado como una importante contribución al estudio de los criterios de evaluación. Su origen es la modelización estadística de la cuasi-implicación: cuando se observa una variable a o una combinación de a en una población, a menudo b también está presente (Gras, & Kuntz, 2009).

El ASI se originó a partir de la didáctica de las matemáticas para resolver problemas cognitivos a través de la clasificación, por lo que uno de sus usos es detectar posibles causas como el aprendizaje y otros procesos. La cuasi-implicación de que difiere de los métodos simétricos basados en la distancia o en la correlación radica principalmente en la asimetría y el conjunto de reglas resultante que es una aproximación a la hipótesis de causa (Gras, & Kuntz, 2009).

2.1.2. *La cuasi-implicación de Regis Gras*

El Análisis Estadístico Implicativo comienza observando el conjunto I está formado por n individuos y el conjunto A está formado por las características p , $A = \{a_1, a_2, \dots, a_p\}$ también se considera que:

$$A_i = \{x \in I : a_i(x) = 1\}, \text{Card}(I) = n, \text{Card}(A_i) = n_{a_i} \text{ y } \text{Card}(\bar{A}_i) = n_{\bar{a}_i}$$

El análisis se realiza calculando los índices de implicación y de cohesión. En un análisis cohesivo puede especificar la tipicidad y la proporción de participación de sujetos y variables y nodos importantes en la formación de clases. Los índices de implicación y la cohesión se define en términos de probabilidad (Gras, & Kuntz, 2009).

La Intensidad de la implicación entre variables

En lógica matemática, la regla “ a_i entonces a_j ” es verdadera si para todo x , $a_j(x)$ es nulo solo si $a_i(x)$ también es nulo; es decir, si se establece el conjunto A de los x para los cuales $a_i(x) = 1$ están en el conjunto B de los x donde $a_j(x) = 1$ (Gras, & Kuntz, 2009). Esta inclusión estricta se observa en casos excepcionales de la vida cotidiana. Por ejemplo en los casos de prueba de conocimiento se pueden ver que algunos estudiantes le convence el ítem a_i pero no es de su agrado el ítem a_j , sin contradecir al ítem a_j cuando se eligió a a_i . La formalización matemática de esta situación está representada por la cuasi-implicación $a_i \rightarrow a_j$ (Gras, & Kuntz, 2009).

La cuasi-implicación $a_i \rightarrow a_j$ significa que cuando a_i esta presente entonces generalmente a_j está también presente.

La medida de relación implícita clasifica a partir según la inverosimilitud de la aparición datos, del número de casos que la invalidan, es decir, cuantifica el asombro del experto ante un número extremadamente pequeño de contraejemplos, a los observados en el caso de una ausencia de relación independiente (Gras, & Kuntz, 2009).

En una jerarquía dirigida, el árbol generado no está necesariamente vinculado, lo que significa que esta jerarquía contiene solo clases importantes de acuerdo con las medidas de calidad adoptada (Gras, & Kuntz, 2009).

Los nodos internos del árbol dirigido, representan la jerarquía dirigida, que describe relaciones de implicación complejas, conocidas como **R-reglas**, entre propiedades de A : (Gras, & Kuntz, 2009).

1. Cuando $R \rightarrow a_i$ puede entenderse como **una consecuencia de R**.
2. R-regla $a_i \rightarrow R$ significa que una R-regla R **puede ser deducida de la observación de a_i**
3. La regla $R' \rightarrow R''$ significa que la propiedad R'' **es una consecuencia de una propiedad predefinida de R'** .

El propósito de la Intensidad Implicativa es expresar la imposibilidad del asombro del número de ejemplos que contradicen la regla $a_i \rightarrow a_j$, para lo que se compara el número de contraejemplos observados con el número de contraejemplos bajo la hipótesis de ausencia relación. Supongamos que elegimos aleatoriamente, de I , dos subconjuntos U y V con n_{a_i} y n_{a_j} elementos respectivamente. Sea $X_{a_i \wedge \bar{a}_j} = Card(U \cap \bar{V})$ es una variable aleatoria asociada, en relación con el número de

contraejemplos en este modelo de aleatorio (Gras, & Kuntz, 2009).

2.1.3. *Ánalysis cohesivo*

La cohesión surge como una medida de la calidad implicativa de la R-regla. El objetivo es descubrir R-reglas $R' \rightarrow R''$ con una fuerte relación implicativa entre las componentes de R' y las de R'' , las relaciones implicativas son suficientemente significativas (Gras, & Kuntz, 2009).

Intuitivamente esto significa que se debe contrastar con el desorden de una experiencia aleatoria, y se plantea que la entropía es bastante conveniente para medir ese desorden (Gras, & Kuntz, 2009).

Se parte de considerar una regla $a_i \rightarrow a_j$ de orden 1 y definir la variable aleatoria Y como indicadora del evento: $X_{a_i \wedge \bar{a}_j} \geq n_{a_i \wedge \bar{a}_j}$ luego:

$$Pr[Y = 1] = Pr[X_{a_i \wedge \bar{a}_j} \geq n_{a_i \wedge \bar{a}_j}]$$

y

$$Pr[Y = 0] = 1 - \varphi(a_i, a_j)$$

2.1.4. *Nodos significativos*

Los nodos significativos de un árbol jerárquico son los nodos correspondientes a una clasificación compatible lo mejor posible con los valores y la calidad de los valores de cohesión (Gras, & Kuntz, 2009).

Definición 2.1. Se llama preorden inicial y global Ω sobre $A \times A$, al preorden inducido por la aplicación Coh (cohesión) sobre $A \times A$.

$$G_{Coh}(\Omega) = \{(a, b); (c, d) : Coh(a, b) < Coh(c, d)\}$$

Sea $S\Pi_k$ el conjunto de pares separados al nivel k y $R\Pi_k$ el conjunto de pares que ya se han reunido hasta este nivel k . $G_{Coh}(\Omega) \cap [S\Pi_k \times R\Pi_k]$ está formado por los pares de parejas que al nivel k respetan el preorden inicial. Por ejemplo, si se tiene: $Coh(e, f) < Coh(a, b)$, entonces $((e, f); (a, b)) \in G_{Coh}(\Omega)$ y si al nivel k , e y f están aún separados, mientras que a y b se reúnen en la clase formada, la pareja $((e, f); (a, b)) \in G_{Coh}(\Omega) \cap [S\Pi_k \times R\Pi_k]$ (Gras, & Kuntz, 2009).

Ahora bien, el cardinal de este último conjunto es función de este nivel k , y es un indicador del acuerdo entre el preorden inicial Ω y el preorden Π_k inducido (Gras, & Kuntz, 2009).

Al cardinal de $G_{Coh}(\Omega) \cap [S\Pi_k \times R\Pi_k]$ se le asocia el índice aleatorio $G_{Coh}(\Omega^*) \cap [S\Pi_k \times R\Pi_k]$, donde Ω^* es un preorden aleatorio en general, provisto de una probabilidad uniforme, de todos los

preordenes del mismo tipo cardinal que Ω (Gras, & Kuntz, 2009).

Este índice tiene:

- **Por esperanza:** $\frac{1}{2}s_k r_k$.
- **Por varianza:** $\frac{s_k r_k (s_k + r_k + 1)}{12}$

Siendo: $Card[S_k] = s_k$ y $Card[R_k] = r_k$. El índice centrado se define como:

$$Coh(\Omega, k) = \frac{Card[G_{Coh}(\Omega) \cap [S\Pi_k \times R\Pi_k]] - \frac{1}{2}s_k r_k}{\sqrt{\frac{s_k r_k (s_k + r_k + 1)}{12}}}$$

Este índice sirve de **estadística global de los niveles**. Sus variaciones son consideradas para significar la constitución de un nivel significativo, ya que este **indicador de acuerdo** entre el preorden inicial y el definido por la división Π_k se vuelve máximo cuando se alcanza un determinado acuerdo, acuerdo que no puede ser sino provisional a causa de la evolución ascendente de la jerarquía.

Definición 2.2. Se llama **nivel significativo** a todo nivel que corresponde a un máximo local de $Coh(\Omega, k)$ durante la construcción de la jerarquía. Se dirá que la división Π_k esta en rasonancia parcial con Ω . Si, además, $Coh(\Omega) \cap [S\Pi_k \times R\Pi_k] = [S\Pi_k \times R\Pi_k]$, se dirá que la división Π_k esta en una rasonancia total con Ω (Gras, & Kuntz, 2009).

Se llama **nodo significativo** cualquier nodo formado a un nivel que corresponde a un máximo local de $v(\Omega, k)$, donde:

$$v(\Omega, k) = Coh(\Omega, k) - Coh(\Omega, k - 1).$$

2.1.5. Otros métodos no ASI similares a la Cuasi-Implicación

En general, cualquiera que sea el problema a resolver, no existe una única técnica para solucionarlo, sino que puede ser abordado siguiendo aproximaciones distintas. El número de técnicas es muy grande y solo puede crecer en el futuro (Aluja, 2001).

- **Market Basket Analysis o Análisis de la cesta de la compra:** Te permite determinar qué productos se compran juntos, permitiéndote incluir variables técnicas que ayuden a la interpretación como día de la semana, ubicación, forma de pago. Y también se puede utilizar en entornos distintos a los supermercados, especialmente en el comercio electrónico e incluye el factor tiempo (Aluja, 2001).
- **Redes bayesianas:** Consiste en representar todos los posibles eventos que nos interesen utilizando el programa de posibles transiciones entre eventos. Se puede clasificar en base al conocimiento experto a partir de datos. Te permite hacer conexiones causales y hacer predicciones (Aluja, 2001).

- **Árboles de decisión:** Nos permite encontrar y ver las reglas de decisión de manera visual sobre las que actúan los clientes a partir de los datos históricos almacenados. Su principal ventaja es la facilidad de uso (Aluja, 2001).

2.1.6. Formulación Matemática

Definición 2.3. La regla $a_i \Rightarrow a_j$ se dice admisible al nivel de confianza $1 - \alpha$ si:

$$P[X_{a_i \wedge \tilde{a}_j} \leq n_{a_i \wedge \tilde{a}_j}] \leq \alpha$$

donde $n_{a_i \wedge \tilde{a}_j} = \text{Card}(A_i \cap \tilde{A}_j)$ es el número de contraejemplos a la regla $a_i \Rightarrow a_j$ observados en la muestra.

Esto significa que el número de contraejemplos observados es pequeño. La distribución de $X_{a_i \wedge \tilde{a}_j}$ depende del esquema de la selección aleatoria aplicado a elegir los conjuntos, estos pueden ser Binomial, Hipergeométrica o Poisson. Para simplificar la notación durante el desarrollo posterior, las variables o características a_i y a_j se llamarán a y b , y los conjuntos: $A_i = \{x \in I : a_i(x) = 1\}$ y $A_j = \{x \in I : a_j(x) = 1\}$ como A y B respectivamente (Gras, & Kuntz, 2009).

Primer Modelo de la hipótesis de ausencia de relación

En el primer modelo explicaremos sobre los siguientes temas: Forma de extraer los individuos y su Modelo Probabilístico en cada uno de sus casos (Gras, & Kuntz, 2009).

1. **Forma de extraer los individuos:** Se observan n_a y n_b individuos bajo el supuesto de que en I existen otros individuos que poseen las característica a y b (Gras, & Kuntz, 2009).
2. **Modelo Probabilístico:** De I se extraen de forma independiente, dos conjuntos X e Y de tamaños n_a y n_b respectivamente. Sean $\Omega = \{(X, Y) : X \subset I, Y \subset I, \text{Card}(X) = n_a, \text{Card}(Y) = n_b\}$, $K_0 = \text{Card}(A \cap \tilde{B})$ y $K = \text{Card}(X \cap \tilde{Y})$ la variable aleatoria correspondiente (Gras, & Kuntz, 2009). Según el modelo probabilístico asumido, tenemos $C_{n_a}^n C_{n_b}^n$ son formas diferentes de extraer estos pares de conjuntos $(X, Y) \in \Omega$, y de ellos solo $C_{n_b}^{n_b} C_{n_a - K_0}^{n_b} C_{K_0}^{n - n_b}$ son favorables a la ocurrencia del evento $K = K_0$, por lo tanto (Gras, & Kuntz, 2009):

$$P(K = K_0) = \frac{C_{n_b}^n C_{n_a - K_0}^{n_b} C_{K_0}^{n - n_b}}{C_{n_a}^n C_{n_b}^n} = \frac{C_{n_a - K_0}^{n_b} C_{K_0}^{n - n_b}}{C_{n_a}^n}$$

lo cual muestra que bajo la hipótesis de ausencia de relación, K sigue la ley hiopergeométrica de parámetros n , n_a y $n - n_b$, $K \sim H(n, n_a, n - n_b)$ (Gras, & Kuntz, 2009).

Los modelos siguientes surgen bajo el supuesto de que el conjunto I de individuos haya sido

seleccionado de un conjunto mayor de individuos, que denotaremos por ζ (Gras, & Kuntz, 2009).

Segundo Modelo de la hipótesis de independencia entre a y b

En el Segundo modelo explicaremos sobre los siguientes temas: Forma de extraer los individuos y su Modelo Probabilístico (Gras, & Kuntz, 2009).

1. Caso 1: El cardinal de ζ es un valor finito N

a) **Forma de extraer los individuos:** Se extraen n individuos de una población ζ de tamaño N , donde N y n se consideran conocidos, mientras que n_a, n_b y $n_{a\wedge\bar{b}}$ son valores observados (Gras, & Kuntz, 2009).

b) **Modelo Probabilístico:** Del conjunto ζ de tamaño N , se extrae una muestra I de tamaño n . En esta muestra I , se observan los individuos que poseen las características $a(n_a)$ y $b(n_b)$ y los que poseen a y no $b(n_{a\wedge\bar{b}})$ (Gras, & Kuntz, 2009).

Existen C_n^N formas posibles de extraer de ζ una muestra I de tamaño n , y de ellos los que verifican que $K = K_0$ son: $C_{K_0}^{N_{a\wedge\bar{b}}} C_{a-K_0}^{N-N_{a\wedge\bar{b}}}$, donde (Gras, & Kuntz, 2009):

$$N_{a\wedge\bar{b}} = \{x \in \zeta : a(x) = 1 \wedge b(x) = 0\}$$

Por lo tanto, $P(K = K_0) = \frac{C_{K_0}^{N_{a\wedge\bar{b}}} C_{a-K_0}^{N-N_{a\wedge\bar{b}}}}{C_n^N}$ es decir, bajo este modelo K sigue la ley hipergeométrica de parámetros N, n y $N_{a\wedge\bar{b}}$, $K \sim H(N, n, N_{a\wedge\bar{b}})$, donde, bajo la hipótesis de independencia se tiene que (Gras, & Kuntz, 2009):

$$\frac{N_{a\wedge\bar{b}}}{N} = \frac{N_a}{N} * \frac{N_{\bar{b}}}{N}$$

2. Caso 2: El cardinal de ζ es un valor infinito.

En este caso se pueden extraer infinitas muestras de tamaño n de la población ζ , y en cada una de estas extracciones, la probabilidad de obtener $K = K_0$ es la misma. Para una muestra de tamaño n esta probabilidad se estimaría a través de $p = \frac{n_{a\wedge\bar{b}}}{n}$ la cual, bajo la hipótesis de independencia, se puede escribir como $p = \frac{n_a}{n} * \frac{n_{\bar{b}}}{n} = \hat{p}(a)\hat{p}(\bar{b})$ (Gras, & Kuntz, 2009).

Por lo tanto,

$$P(K = K_0) = C_{K_0}^n p^{K_0} (1-p)^{n-K_0}$$

y K sigue la ley binomial de parámetros n y p , $K \sim B(n, p)$.

3. **Caso 3:** El cardinal de ζ es un valor infinito y el cardinal de I es indeterminado.

En este caso se trata de extraer de la población ζ de tamaño infinito, una muestra I , cuyo tamaño es aleatorio.

Denotemos por M el tamaño de la muestra I , donde M es una magnitud aleatoria, entonces (Gras, & Kuntz, 2009):

$$P(K = K_0) = \sum_{n=0}^{\infty} P(M = n) * P(K = K_0 : M = n) \quad (2.1)$$

para que $K = K_0$ debe cumplirse que $n \geq K_0$, por lo que la expresión anterior puede reescribirse como (Gras, & Kuntz, 2009):

$$P(K = K_0) = \sum_{n \geq K_0}^{\infty} P(M = n) * P(K = K_0 : M = n) \quad (2.2)$$

Supongamos que la variable aleatoria M sigue una distribución Poisson de parámetro $\lambda = n$, valor observado de M (Gras, & Kuntz, 2009).

$$P(K = K_0 : M = n) = C_{K_0}^n p^{K_0} (1-p)^{n-K_0}, \text{ con } p = \frac{n_a n_{\bar{b}}}{n} \quad (2.3)$$

Sustituyendo (2) en (1) y asumiendo que $M \sim P(n)$, obtenemos:

$$P(K = K_0) = \sum_{n \geq K_0}^{\infty} \frac{e^{-n} n^n}{n!} * C_{K_0}^n p^{K_0} (1-p)^{n-K_0}$$

$$P(K = K_0) = \frac{e^{-np} n p^{K_0}}{K_0!} \sum_{n \geq K_0}^{\infty} \frac{e^{-n(1-p)} [n(1-p)]^{n-K_0}}{(n-K_0)!}$$

donde el segundo factor vale 1, por lo que:

$$P(K = K_0) = \frac{e^{-np} n p^{K_0}}{K_0!}$$

es decir, bajo este caso K sigue la ley de Poisson de media np , $K \sim \text{Poisson}(np)$, siendo $np = \frac{n_a n_{\bar{b}}}{n}$ bajo la hipótesis de independencia.

Bajo los modelos considerados se demuestra que:

$$E(K) = \frac{n_a n_{\bar{b}}}{n} \quad (2.4)$$

Definición 2.4. La Intensidad Implicativa de la regla $a_i \Rightarrow a_j$ se define como (Gras, & Kuntz, 2009):

$$\varphi(a_i, a_j) = 1 - P[K \leq n_{a_i \wedge \bar{a}_j}] \text{ si } n_{a_i} \neq n$$

en caso contrario, $\varphi(a_i, a_j) = 0$.

La regla es retenida para un α dado si: $\varphi(a_i, a_j) \geq 1 - \alpha$, lo cual es equivalente a decir que la regla es admisible, con confianza $1 - \alpha$ (Gras, & Kuntz, 2009).

Definición 2.5. El índice de Implicación de la regla $a_i \Rightarrow a_j$ se define como (Gras, & Kuntz, 2009):

$$q(a_i, \tilde{a}_j) = \frac{n_{a_i \wedge \tilde{a}_j} - \frac{n_{a_i} n_{\tilde{a}_j}}{n}}{\sqrt{\frac{n_{a_i} n_{\tilde{a}_j}}{n}}}$$

Esto demuestra que: $p(a_i, a_j) = -q(a_i, \tilde{a}_j) \sqrt{\frac{n}{n_{\tilde{a}_j} n_{a_j}}}$ (Gras, & Kuntz, 2009).

Definición 2.6. Entropía: La podemos interpretar como la cantidad de medida de información que nos reporta una fuente binaria, con probabilidad de éxito igual a la intensidad implicativa de la regla $a_i \rightarrow a_j$, o el valor medio de la incertidumbre de un observador antes de conocer la salida de una fuente binaria en la que puede o no ocurrir el evento $a_i \rightarrow a_j$ obteniendo la siguiente fórmula (Gras, & Kuntz, 2009):

$$E = -p \log_2 p - (1-p) \log_2 (1-p), \text{ con } p = \varphi(a_i, a_j)$$

Definición 2.7. La cohesión de una R-regla $a_i \rightarrow a_j$ de grado 1 es (Gras, & Kuntz, 2009):

$$Coh(a_i, a_j) = \begin{cases} \sqrt{1-E^2} & \text{si } p \geq 0.5 \\ 0 & \text{en caso contrario} \end{cases}$$

Si $\varphi(a_i, a_j) = 0.5$, la regla $a_i \rightarrow a_j$ es neutral en el sentido de que da lugar al número de contraejemplos esperado bajo independencia, es decir (Gras, & Kuntz, 2009),

$$n_{a_i \wedge \tilde{a}_j} = E[X_{a_i \wedge \tilde{a}_j}] = \frac{n_{a_i} n_{\tilde{a}_j}}{n}$$

Definición 2.8. La Cohesión de la clase de variables $R = \{a_i, \dots, a_k\}$ es la medida geométrica de las cohesiones de los pares de variables que la conforman (Gras, & Kuntz, 2009),

$$Coh(R) = \left\{ \prod_{\substack{i,j \\ j>1}} C(a_i, a_j) \right\}^{\frac{2}{k(k-1)}}$$

Definición 2.9. La intensidad de implicación de una clase A sobre una clase B es:

$$\psi(A, B) = \left\{ \sup_{\substack{i=1, \dots, k \\ j=1, \dots, s}} \varphi(a_i, a_j) \right\}^{rs} [Coh(A)Coh(B)]^{\frac{1}{2}}$$

Cuando se obtiene las cohesiones de nivel cero, comienza el proceso de acumulación, que incluye el hecho de que en cada nuevo nivel o paso del proceso, las variables individuales o las capas creadas se combinan con otras capas, o las variables aisladas tienen cohesiones como máximas.

La implementación del algoritmo propuesto se va a realizar en el software MATLAB ya que es un software que está orientado a la programación con estructura vectoriales y matriciales apropiados para este tipo de aplicaciones matemáticas y además nos permite la visualización grafica de los resultados que deseamos obtener.

2.1.7. Ejemplo detallado de aplicación

Consideraremos el siguiente ejemplo: Se considera una muestra aleatoria de un grupo de $n = 20$ estudiantes de la carrera de Matemática de la ESPOCH, dieron su aceptación o no en agrado a las siguientes asignaturas: Lógica Matemática (LM), Geometría Analítica (GA), Topología (To), Ecuaciones Diferenciales (ED), Análisisi Funcional (AF), Historia de la Matemática (HM) y Álgebra (AL).

Asumiremos que la variable aleatoria $K = \text{Card}(X \cap \bar{Y})$ sigue una distribución Binomial de parámetros n y $p = \frac{n_a}{n} \frac{n_b}{n}$ y trabajaremos con las variables LM, GA, To, ED, AF, HM, AL. En las tablas siguientes se mostraran, respectivamente, la sub-tabla con los datos para las variables seleccionadas para la ejemplificación, y algunas de las implicaciones que se pueden formar, así como el cardinal, el valor de p y el número de contraejemplos de la regla $a_i \rightarrow a_j$ (Gras, & Kuntz, 2009).

Tabla 1-2: Matriz de datos binarios

Tabla de datos							
	LM	GA	To	ED	AF	HM	AL
est1	1	1	0	1	1	0	1
est2	1	1	0	0	0	1	1
est3	1	0	1	1	0	1	1
est4	0	0	1	1	1	0	0
est5	0	0	1	1	1	0	1
est6	0	1	0	1	0	0	1
est7	0	0	0	0	0	0	0
est8	0	1	0	0	1	1	1
est9	0	0	1	1	0	1	0
est10	1	1	0	0	1	1	0
est11	0	1	0	1	1	0	0
est12	1	0	0	1	1	1	1
est13	0	0	1	0	0	1	1
est14	1	1	1	1	0	1	0
est15	0	0	1	0	1	1	1
est16	1	1	1	0	0	0	0
est17	0	1	0	1	0	0	0
est18	1	1	0	1	1	1	1
est19	0	0	1	0	1	0	1
est20	0	1	1	1	1	0	1
n_{a_i}	8	11	10	13	11	10	12
$n_{\bar{a}_j}$	12	9	10	7	9	10	8
n	20	20	20	20	20	20	20

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

Para empezar debemos tener en cuenta que los n_{a_i} son todos los alumnos que les agrade la asignatura que esta representado con el número 1 y el $n_{\bar{a}_j}$ son todos los alumnos que no comprendan la asignatura que esta representado con el número 0.

Tenemos la variable LM que en el proceso de nuestro estudio será a_i y la variable GA que sera a_j . Para obtener la probabilidad de éxito, debemos contabilizar en la variable LM el total de n_{a_i} , entonces para la variable LM es $n_{a_i} = 8$, ahora para la variable GA contabilizamos el total de $n_{\bar{a}_j}$, para el caso de la variable GA es $n_{\bar{a}_j} = 9$, sabiendo esto aplicaremos la fórmula de la probabilidad de éxito que es la siguiente:

$$p = \frac{n_{a_i}}{n} * \frac{n_{\bar{a}_j}}{n}$$

Teniendo en cuenta que nuestro $n = 20$ es el total de individuos de nuestra población estudiada

$$p = \frac{8}{20} * \frac{9}{20} = 0.18$$

$$p = \frac{8}{20} * \frac{8}{20} = 0.16$$

Tabla 2-2: Ejemplo del Valor de la Probabilidad de éxito

Valor de la Probabilidad de éxito				
a_i	a_j	n_{a_i}	$n_{\bar{a}_i}$	p
LM	GA	8	9	0.18
LM	AL	8	8	0,16
To	HM	10	10	0,25
GA	HM	11	10	0,28
AF	LM	11	12	0,33
AF	AI	11	8	0,22
ED	AI	13	8	0,26

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

Como podemos observar en el ejemplo del valor de la probabilidad de éxito para las variables (LM,GA) y (LM,AL) se puede calcular de la misma forma con las demas variables restantes.

Tabla 3-2: Valores de la probabilidad de éxito

Valores de la Probabilidad de éxito							
	LM	GA	To	ED	AF	HM	AI
LM	0	0.1800	0.2000	0.1600	0.1800	0.2000	0.1600
GA	0.3300	0	0.2750	0.2200	0.2475	0.2750	0.2200
To	0.3000	0.2250	0	0.2000	0.2250	0.2500	0.2000
ED	0.3600	0.2700	0.3000	0	0.2700	0.3000	0.24000
AF	0.3300	0.2475	0.2750	0.2200	0	0.2750	0.2200
HM	0.3000	0.2250	0.2500	0.2000	0.2250	0	0.2000
AL	0.3600	0.2700	0.3000	0.2400	0.2700	0.3000	0

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

Para obtener el valor de la cardinalidad que está representado por $Card(A_i \cap \bar{A}_j)$ se debe tener en cuenta la cantidad de individuos que poseen la característica a_i , pero no la característica a_j

Tabla 4-2: Ejemplo de la $Card(LM, GA)$

$Card(A_i \cap \bar{A}_j)$			
	LM	GA	
est1	1	1	0
est2	1	1	0
est3	1	0	1
est4	0	0	0
est5	0	0	0
est6	0	1	0
est7	0	0	0
est8	0	1	0
est9	0	0	0
est10	1	1	0
est11	0	1	0
est12	1	0	1
est13	0	0	0
est14	1	1	0
est15	0	0	0
est16	1	1	0
est17	0	1	0
est18	1	1	0
est19	0	0	0
est20	0	1	0
$Card(LM, GA)$			2

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

Observando el ejemplo debemos contabilizar el total con las variables LM y GA que cumplen esta condición $Card(LM \cap GA) = 2$, así debemos desarrollar para cada una de las variables, para obtener nuestra tabla de Cardinalidad con los datos que nos brinda la Tabla de Datos de las variables Binarias.

Tabla 5-2: Valores de la Cardinalidad

Valores de la Cardinalidad							
	LM	GA	To	ED	AF	HM	AL
LM	0	2	5	3	4	2	3
GA	5	0	8	4	5	6	5
To	7	7	0	4	5	5	4
ED	7	5	6	0	5	7	5
AF	7	5	6	4	0	6	3
HM	4	5	5	5	5	0	3
AL	7	6	6	5	4	5	0

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

Para obtener los valores de la Distribución Binomial entre dos variables utilizaremos la fórmula:

$$P(K = K_0) = C_{k_0}^n p^{k_0} (1 - p)^{n - k_0}$$

Donde:

n es el número total de individuos de nuestra población.

k_0 es el número de éxitos $Card(A_i \cap \bar{A}_j)$.

p es el número de la probabilidad de éxito.

$1 - p$ es el número de la probabilidad de fracaso. El número combinatorio

$$C_{k_0}^n = \frac{n!}{k!(n - k)!}$$

$$\begin{aligned} P(LM, GA) &= C_2^{20} (0.18)^2 (0.82)^{20-2} \\ &= \frac{20!}{2!(20-2)!} (0.18)^2 (0.82)^{20-2} \\ &= \frac{20!}{2!(18)!} (0.18)^2 (0.82)^{18} \\ &= 0.2748 \end{aligned}$$

$$\begin{aligned} P(LM, AL) &= C_3^{20} (0.16)^3 (0.84)^{20-3} \\ &= \frac{20!}{3!(20-3)!} (0.16)^3 (0.84)^{20-3} \\ &= \frac{20!}{3!(17)!} (0.16)^3 (0.84)^{17} \\ &= 0.5990 \end{aligned}$$

Tabla 6-2: Ejemplo de la distribución Binomial

Distribución Binomial	
LM-GA	0.2748
LM-AL	0.5990

Fuente: Elaboración propia.
Realizado por: Ilbay, Bryan, 2023.

Como podemos observar en el ejemplo del valor de la distribución Binomial para las variables (LM,GA) y (LM,AL) se puede calcular de la misma forma con las demas variables restantes.

Tabla 7-2: Valores de la distribución Binomial

Valores de la Distribución Binomial							
	LM	GA	To	ED	AF	HM	AI
LM	0	0.2748	0.8042	0.5990	0.7151	0.2061	0.5990
GA	0.3083	0	0.9291	0.5420	0.6273	0.7012	0.7343
To	0.7723	0.9397	0	0.6296	0.7156	0.6172	0.6296
ED	0.5639	0.5357	0.6080	0	0.5357	0.7723	0.6573
AF	0.6732	0.6273	0.7012	0.5420	0	0.7012	0.3289
HM	0.2375	0.7156	0.6172	0.8042	0.7156	0	0.4114
AI	0.5639	0.7190	0.6680	0.6573	0.3375	0.4164	0

Fuente: Elaboración propia.
Realizado por: Ilbay, Bryan, 2023.

Para calcular la intensidad implicativa utilizaremos la siguiente fórmula:

$$\varphi(a_i, a_j) = 1 - P[K \leq n_{a_i \wedge \bar{a}_j}]$$

Tabla 8-2: Ejemplo de la Intensidad

Implicativa

Intensidad Implicativa		
$\varphi(LM;GA)$	$1 - P[k \leq 2]$	$1-0.2748=0.7252$
$\varphi(LM;AL)$	$1 - P[k \leq 3]$	$1-0.5990=0.4010$

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

Como podemos observar en el ejemplo del Valor de la Intensidad Implicativa para las variables (LM,GA) y (LM,AL) se puede calcular de la misma forma con las demas variables restantes.

Tabla 9-2: Valores de la intensidad Implicativa

Valores de la intensidad Implicativa							
	LM	GA	To	ED	AF	HM	AL
LM	0	0.7252	0.1958	0.4010	0.2849	0.7939	0.4010
GA	0.6917	0	0.0709	0.4580	0.3727	0.2988	0.2657
To	0.2277	0.0603	0	0.3704	0.2844	0.3828	0.3704
ED	0.4361	0.4643	0.3920	0	0.4643	0.2277	0.3427
AF	0.3268	0.3727	0.2988	0.4580	0	0.2988	0.5711
HM	0.7625	0.2844	0.3828	0.1958	0.2844	0	0.5886
AL	0.4361	0.2810	0.3920	0.3427	0.6625	0.5836	0

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

CHIC muestra los índices de implicación entre variables, en por ciento. En la figura se muestra el grafo implicativo que se forma a partir de los valores de intensidad implicativa obtenidos entre las variables analizadas.

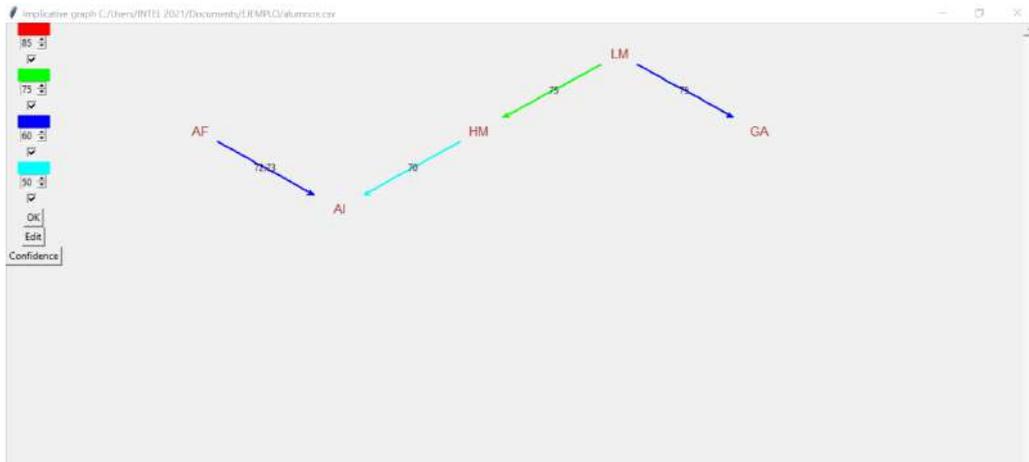


Ilustración 1-2: Grafo Implicativo

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

Para calcular la Entropía utilizaremos la siguiente fórmula:

$$E = -p \log_2 p - (1 - p) \log_2 (1 - p), \text{ con } p = \varphi(a_i, a_j)$$

Siendo p el valor de la intensidad implicativa.

Por ejemplo la Entropía entre las variables (LM,GA) y (LM,AI) es el siguiente:

$$E(LM, GA) = -0.7252 \log_2 (0.7252) - (1 - 0.7252) \log_2 (1 - 0.7252) = 0.8483$$

$$E(LM, AI) = -0.4010 \log_2 (0.4010) - (1 - 0.4010) \log_2 (1 - 0.4010) = 0.9715$$

Como podemos observar en el ejemplo del Valor de la Entropía para las variables (LM,GA) y (LM,AI) se puede calcular de la misma forma con las demas variables restantes.

Tabla 10-2: Valores de Entropía

Matriz de Entropía							
	LM	GA	To	ED	AF	HM	AL
LM	0	0,8483	0,7134	0,9715	0,8620	0,7339	0,9715
GA	0,8911	0	0,3693	0,9949	0,9527	0,8798	0,8353
To	0,7740	0,3286	0	0,9509	0,8614	0,9600	0,9509
ED	0,9882	0,9963	0,9661	0	0,9963	0,7740	0,9274
AF	0,9116	0,9527	0,8798	0,9949	0	0,8798	0,9138
HM	0,7909	0,8614	0,9600	0,7134	0,8614	0	0,9773
AI	0,9882	0,8566	0,9661	0,9274	0,9224	0,9797	0

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

Para calcular el índice de cohesión utilizaremos la siguiente fórmula:

$$Coh(a_i, a_j) = \begin{cases} \sqrt{1 - E^2} & \text{si } p \geq 0.5 \\ 0 & \text{en caso contrario} \end{cases}$$

Por ejemplo el Índice de Cohesión para las variable (LM,GA) y (LM,Al) es el siguiente:

$$Coh(LM, GA) = \sqrt{1 - (0.8483)^2} = 0.5296$$

$$Coh(LM, Al) = \sqrt{1 - (0.9715)^2} = 0.2370$$

Este caso la $Coh(LM, Al) \leq 0.5$ por lo tanto tomara el valor de 0.

Como podemos observar en el ejemplo del Valor del Índice de Cohesión para las variables (LM,GA) y (LM,Al) se puede calcular de la misma forma con las demas variables restantes.

Tabla 11-2: Valores de la Matriz Cohesión de nivel 0

Matriz de Cohesión de nivel 0							
	LM	GA	To	ED	AF	HM	AL
LM	0	0.5296	0	0	0	0.6792	0
GA	0.4537	0	0	0	0	0	0
To	0	0	0	0	0	0	0
ED	0	0	0	0	0	0	0
AF	0	0	0	0	0	0	0.4061
HM	0.6120	0	0	0	0	0	0.2121
AI	0	0	0	0	0.3862	0.2003	0

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

Para calcular las cohesiones nivel 1 utilizaremos la siguiente fórmula:

$$\psi(A,B) = \left\{ \sup_{\substack{i=1,\dots,k \\ j=1,\dots,s}} \varphi(a_i, a_j) \right\}^{r \times s} [Coh(A)Coh(B)]^{\frac{1}{2}}$$

Para este ejemplo utilizaremos el máximo de la tabla de las cohesiones de nivel cero que se encuentra en $(LM, HM) = 0.6792$ y también utilizaremos los valores de la tabla de Intensidad Implícita con la variable GA por lo cual utilizaremos los valores de (LM, GA) y (HM, GA)

Tabla 12-2: Ejemplo de la Cohesión de nivel 0

Intensidad Implícita					
$A = \{LM, HM\}$	$\varphi(a, b)$	GA	Coh(A)	Coh(B)	$\psi(A, B)$
$B = GA$	LM	0.7252	0.6742	1	0.4334
	HM	0.2844			

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

Como podemos observar en el ejemplo del Valor del Índice de Cohesión de nivel 1 para las variables (LM, HM) con la variable GA se puede calcular de la misma forma con las demás variables restantes.

Tabla 13-2: Valores de la Matriz Cohesión de nivel 1

Matriz de Cohesión de nivel 1						
Nivel 1	(LM,HM)	GA	To	ED	AF	AL
(LM,HM)	0	0.4334	0.1208	0.1325	0.0669	0.2855
GA	0.3944	0	0	0	0	0
To	0.1208	0	0	0	0	0
ED	0.1567	0	0	0	0	0
AF	0.0880	0	0	0	0	0.4061
AI	0.2807	0	0	0	0,3862	0

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

Como el ejemplo la tabla Valores de la Matriz Cohesión de nivel 1 no tiene cantidades mayores 0.5 por lo tanto nuestro ejemplo solo tiene hasta el nivel 1 de Cohesión.



Ilustración 2-2: Árbol de Cohesión

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

2.2. La programación en MATLAB

En esta sección abordaremos los siguientes temas relacionados a Matlab: Variables, Funciones, Comandos de programación, Programación en objetos.

2.2.1. Variables

Una variable consiste en una ubicación en el sistema de almacenamiento o en la memoria principal de la computadora (llamada espacio de trabajo en MATLAB) y un nombre simbólico (identificador) asociado con esa ubicación. Este espacio contiene cantidades conocidas o desconocidas o información, es decir valores. Los nombres de variables son una forma común de referirse a valores almacenados: esta separación de nombre y contenido permite que el nombre se use independientemente de la información exacta que representa. Un identificador en el código fuente de la computadora se puede asociar con un valor en tiempo de ejecución, por lo que el valor de una variable puede cambiar durante la ejecución del programa (Arellano, 2013).

El concepto de variable en programación puede no corresponder directamente al concepto de variable en matemáticas. El valor de una variable en programación no tiene que ser parte de una ecuación o fórmula como lo es en matemáticas. En programación, una variable se puede usar en un proceso iterativo: se le puede asignar un valor en un lugar, luego usarse en otro lugar, luego asignarse un nuevo valor y luego usarse nuevamente. Este tipo de proceso se conoce como iteración (Arellano, 2013).

En la programación informática, las variables suelen recibir nombres largos para que sean relativamente descriptivas cuando se utilizan, mientras que en matemáticas las variables suelen recibir nombres cortos que constan de uno o dos caracteres para que sean más descriptivas cuando se reescriben y manipulan (Arellano, 2013).

En MATLAB, las variables se pueden crear usando código fuente a partir de operaciones que usan constantes, otras variables o funciones; y directamente a través de la ventana de comandos (Arellano, 2013).

2.2.2. Funciones

Las funciones provistas por MATLAB (funciones internas) pueden ser:

1. Funciones de archivo M, aquellas que son implementadas como archivos M.
2. Funciones built-ins, aquellas que son programas ejecutables precompilados.

Muchas funciones internas de MATLAB están sobrecargadas para trabajar diferentes tipos de datos son válidos. Las funciones específicas de MATLAB se encuentran en una subcarpeta de la carpeta del kit de herramientas de Matlab (Arellano, 2013). Lista de funciones en cada subcarpeta (categorías) y documentación de acceso para cada característica dentro de ellos, digítalos con doc o help seguido

del nombre del subdirectorio en el indicador de la ventana Ordenar (Arellano, 2013).

Observación: A diferencia de las funciones de los archivos M, las funciones (built-in) integradas no le permiten ver su código fuente, pero la mayoría de estas funciones tienen un archivo M asociado, que simplemente contiene la documentación de ayuda para la función (Arellano, 2013).

2.2.3. Comandos de programación

En Matlab los comandos mas utilizados son los siguientes:

1. **For:** Un bloque for en cada iteración asigna a la variable la columna *i*-ésima de la expresión y ejecuta las órdenes. En la práctica las expresiones suelen ser del tipo escalar (Porras, 2018):

```
for variable=expresión
    Orden
    Orden
    ...
    Orden
Fin
```

a) La sentencia **for** sirve para recorrer dato a dato ya sea un vector o una matriz (Porras, 2018).

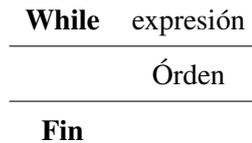
2. **If:** Un bloque **if**, lo que hace es evaluar una expresión lógica y si es cierta ejecuta las órdenes que encuentre antes del **end**. Puede escribirse de varias maneras (Porras, 2018):

```
if expresión
    Órdenes evaluadas si la expresión es verdadera
Fin
```

Puede que nos interese que en caso de no ejecutar dicha orden ejecute otra distinta. Esto se lo indicaremos usando **else** dentro del bloque (Porras, 2018).

```
if expresión
    Órdenes evaluadas si la expresión es verdadera
else
    Órdenes evaluadas si la expresión es falsa
Fin
```

3. **While:** Un bloque **while** ejecuta las órdenes mientras todos los elementos de la expresión sean verdaderos (Porras, 2018).



- a) El comando **while** muestra la sentencia, que mientras una condición como la del ejemplo: $a \leq 5$ se cumpla muestre los valores que sean menores al número indicado (Porras, 2018).
- b) **Continue:** La sentencia **continue** hace que se pase inmediatamente a la siguiente iteración del bucle **for** o del bucle **while** saltando todas las órdenes que hay entre el **continue** y el **fin** del bucle en la iteración actual (Porras, 2018).

2.2.4. Programación en objetos en MATLAB

Un algoritmo es una secuencia de pasos no ambigua, finita y ordenada que nos conduce a la solución de un problema. Se representan mediante Diagramas de Flujo o Pseudocódigo. La programación es la implementación (conversión) de un algoritmo, a través de un determinado lenguaje de programación, en un software. Los programas suelen subdividirse en partes menores (módulos), de modo que la complejidad algorítmica de cada una de las partes sea menor que la del programa completo, lo cual ayuda al desarrollo del programa.

a) Los Lenguajes de Programación

Un lenguaje de programación es un lenguaje que puede ser utilizado para controlar el comportamiento de una máquina, particularmente una computadora. Consiste en un conjunto de símbolos y reglas sintácticas y semánticas que definen su estructura y el significado de sus elementos y expresiones (Arellano, 2013).

Un lenguaje de programación permite a uno o más programadores especificar de manera precisa: sobre qué datos una computadora debe operar, cómo deben ser estos almacenados, transmitidos y qué acciones debe tomar bajo una variada gama de circunstancias. Todo esto, a través de un lenguaje que intenta estar relativamente próximo al lenguaje humano o natural (Arellano, 2013).

Una característica relevante de los lenguajes de programación es precisamente que más de un programador puedan tener un conjunto común de instrucciones que puedan ser comprendidas entre ellos para realizar la construcción del programa de forma colaborativa. Los procesadores usados en las computadoras son capaces de entender y actuar según lo indican programas escritos en un lenguaje fijo llamado lenguaje de máquina. Todo programa escrito en otro lenguaje puede ser ejecutado de dos maneras (Arellano, 2013):

- 1) Mediante un programa que va adaptando las instrucciones conforme son encontradas. A este proceso se le llama interpretar y a los programas que lo hacen se los conoce como intérpretes (Arellano, 2013).
- 2) Traduciendo este programa al programa equivalente escrito en lenguaje de máquina. A ese proceso se le llama compilar y al traductor se le conoce como compilador (Arellano, 2013).

b) Clasificación de los Lenguajes de Programación

- 1) Por el nivel de abstracción
 - a'* Lenguajes de bajo nivel: Aquellos que mas se asemejan al lenguaje de una computadora (lenguaje de máquina)
 - b'* Lenguajes de mediano nivel: Aquellos conformados por nemónicos convertibles en forma directa a lenguaje máquina.
 - c'* Lenguajes de alto nivel: Aquellos que están conformados por elementos del lenguaje humano.
- 2) Por la forma de ejecución
 - a'* Compilados: Aquellos que convierten todo un programa a lenguaje máquina para su ejecución.
 - b'* Interpretados: Aquellos que van convirtiendo sentencias de un programa a lenguaje máquina conforme vaya siendo necesario durante su ejecución (proceso de datos).
- 3) Por el paradigma de programación Un paradigma de programación es la filosofía utilizada en la construcción del software, podemos mencionar entre ellos a los paradigmas:
 - a'* Imperativo.
 - b'* Funcional.
 - c'* Lógico.
 - d'* Orientado a Objetos.
 - e'* Paralelo.

El Lenguaje de Programación MATLAB es:

- a'* Un Lenguaje de Programación de Alto Nivel.
- b'* Un Lenguaje de Programación Compilador e Interpretador.
- c'* Un Lenguaje de Programación Imperativo. Orientado a Objetos y Paralelo.

CAPÍTULO III

3. MARCO METODOLÓGICO

3.1. Enfoque de investigación

El tipo de investigación que vamos a utilizar en nuestro trabajo de titulación es la investigación cuantitativa ya que nos permite recolectar y analizar datos numéricos, la investigación transversal nos ayuda a analizar datos de variables recopiladas en un período de tiempo sobre una población predefinida

3.2. Nivel de investigación

Es una investigación correlacional, ya que es un método de investigación pre-experimental y nos permite evaluar dos variables por lo tanto evaluar la relación estadística que existe entre ella, en este caso la correlación entre los resultados de los programas R-Chic y M-Chic.

3.3. Diseño de la investigación

la investigación pre experimental de la forma RGXO1, donde RG es el grupo de base de datos estadísticas aleatorias obtenidas mediante un muestreo aleatorio simple. X es el tratamiento (la ejecución) y O1 es es la correlación del programa diseñado en el respectivo de RCHIC, por lo tanto se tomara en cuenta el paquete RCHIC para lograr una programación similar en MATLAB.

3.3.1. *Según la manipulación o no de la variable independiente*

La permanencia de la manipulación de las variables que intervienen en la investigación, la investigación es analizada sin modificaciones ni alteraciones, por lo que existe grandes niveles de validez de los resultados obtenidos; ya que la variable independiente es las bases de datos analizar en este caso son 383 evaluadas una a una. Para después comprobar la similaridad de los resultados de cada una de las bases evaluadas.

3.3.2. Según las intervenciones en el trabajo de campo

La población son los resultados de la correlación del programa diseñado con el respectivo de RCHIC sobre la muestra de estudio, tal como: matriz de similaridad, dendograma, copresencias, nodos significativos, nombre del archivo, número de filas que conforman la base de datos, número de columnas que conforman la base de datos, el total de datos.

3.4. Tipo de estudio

La investigación es de tipo Teórica ya que aportará con conocimiento a la parte Matemática en la similaridad de los programas R-Chic y M-Chic.

3.5. Población y Planificación, selección y cálculo del tamaño de la muestra

El colectivo de estudio lo conforman las 100000 bases de datos aleatorias formadas por un máximo de 1000 observaciones y 100 variables, por la amplitud es un estudio de muestreo conformado por 383 bases de datos aleatorias binarias.

El cálculo aproximado del tamaño de la población se muestra en forma detallada en el Apéndice. Por el gran tamaño de la población, se escogió una muestra utilizando el método de muestreo aleatorio simple con parámetro de interés la media, se consideró la fórmula para el cálculo de la muestra

$$n = \frac{S^2}{\frac{E^2}{Z_{\frac{\alpha}{2}}^2} + \frac{S^2}{N}}$$

Para aplicar la fórmula se utilizaron los parámetros desviación estándar = 1; $\alpha = 5\%$; $Z = 1,96$; $E = 10,01\%$; $N = 100000$ y se generó un tamaño de la muestra de 383,2 que redondeado es 383.

3.6. Métodos, técnicas e instrumentos de investigación

Al obtener las 383 bases a evaluar, se medirá el nivel de similaridad de cada una de ellas con la comparación del programa M-Chic con el R-Chic. Para esta comparación se utilizará el software Matlab en la versión R2001a y el software R en la versión R 3.6.3.

CAPÍTULO IV

4. MARCO DE ANÁLISIS E INTERPRETACIÓN DE RESULTADOS

4.1. Procesamiento, análisis e interpretación de resultados

El seudocódigo desarrollado está basado en el conjunto de herramientas y operaciones de Matlab, las cuales nos permitieron programar la Cuasi-Implicación de Grass para variables binarias en el entorno de trabajo de Matlab. Esta filosofía de programación permite a las personas aumentar el número de herramientas disponibles para el estudio del Análisis Estadístico Implicativo. Las personas, ahora contarán con una herramienta que facilita los cálculos sobre la Cuasi-Implicación de Grass, los cuales están de una manera mas detallada y con una interfaz mas accesible al momento de utilizar el seudocódigo.

4.1.1. *Elaborar el seudocódigo de la cuasi-implicación de Gras*

Para la elaboración del seudocódigo debemos tener en cuentas todas operaciones y cálculos que requiere la Cuasi-Implicación de Gras, para ello elaboramos un Diagrama de Flujo el cual nos permite tener una guía clara sobre lo que vamos a programar

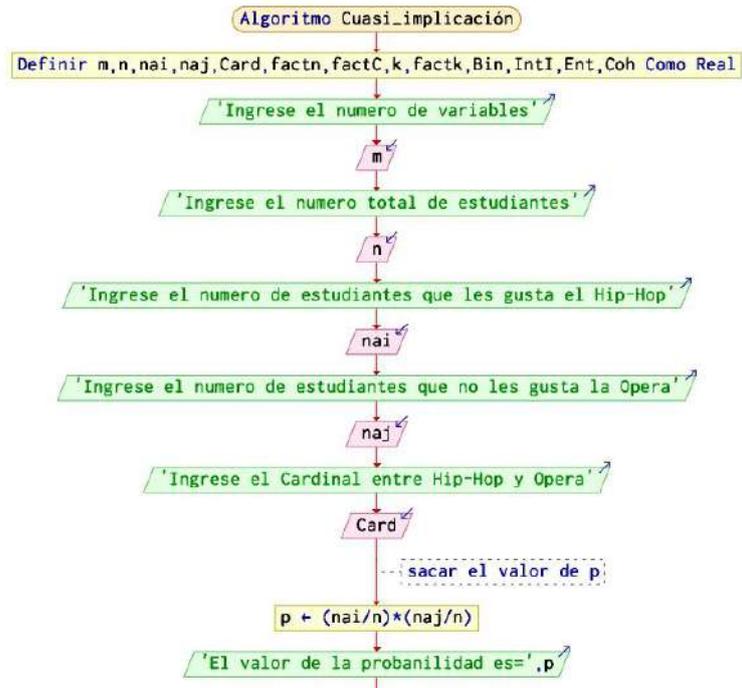


Ilustración 1-4: Diagrama de Flujo Pt.1.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

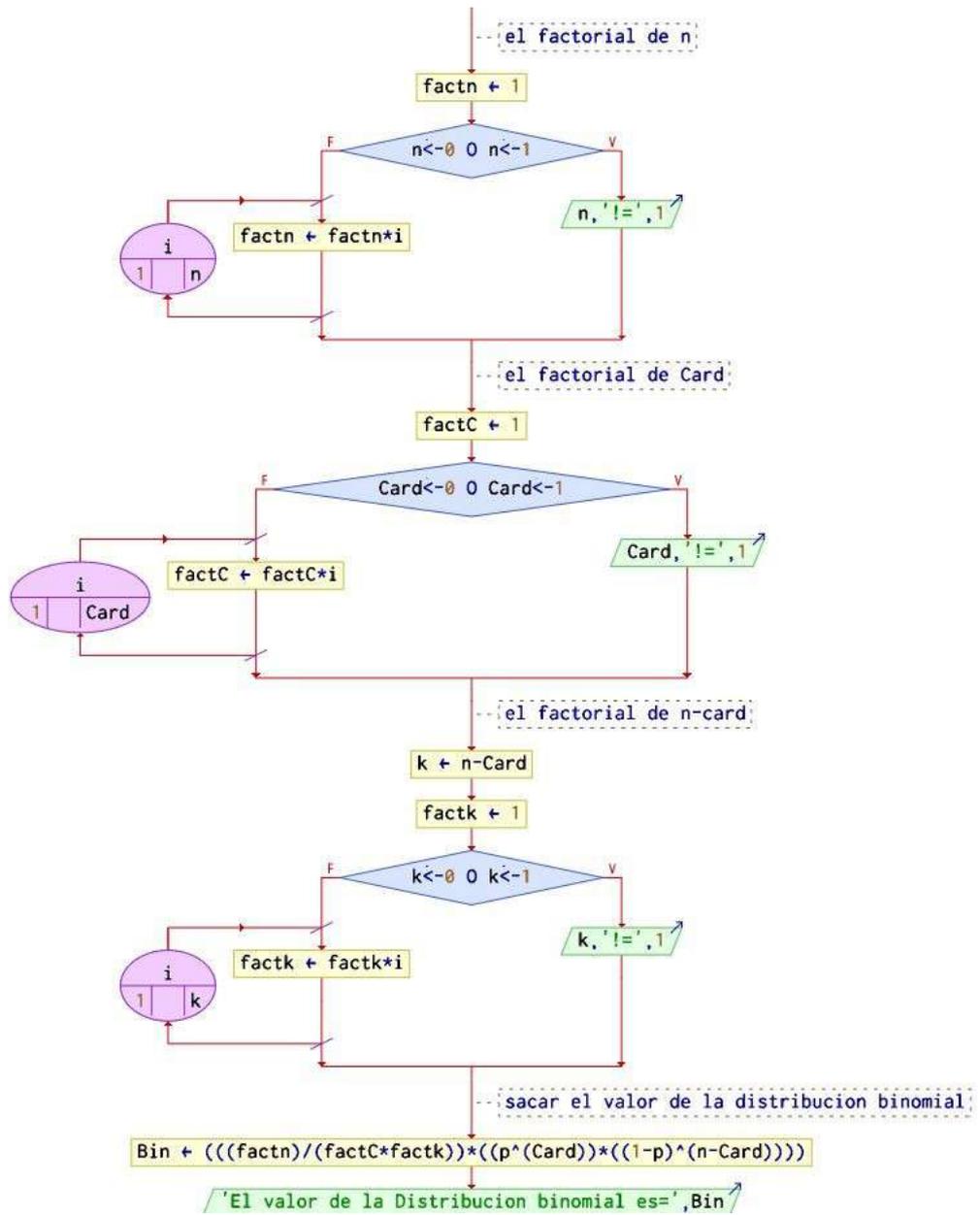


Ilustración 2-4: Diagrama de Flujo Pt.2

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

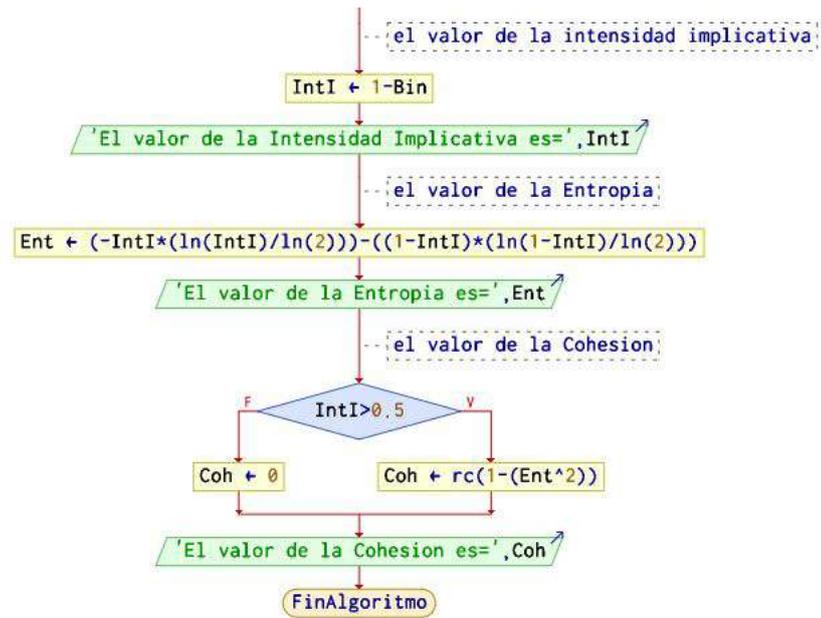


Ilustración 3-4: Diagrama de Flujo Pt.3.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

En el cual encontramos las siguientes asignaciones y operaciones:

- n número de filas (estudiantes o personas).
- m número de columnas (variables).
- n_{a_i} número de personas que les agrada (tipo de música, asignaturas, etc) y en otras palabras sería la celda que contiene al número 1.
- n_{a_j} número de personas que **no** les agrada (tipo de música, asignaturas, etc) y en otras palabras sería la celda que contiene al número 0.
- $Card(n_{a_i} \cap n_{\bar{a}_j}) = k_0$ número de coincidencias entre 1 y 0.
- p valor de la probabilidad de éxito. Su fórmula es la siguiente:

$$p = \frac{n_{a_i}}{n} * \frac{n_{\bar{a}_j}}{n}$$

- $P(K = K_0)$ valor de la distribución binomial. Su fórmula es la siguiente:

$$P(K = K_0) = C_{k_0}^n p^{k_0} (1 - p)^{n - k_0}$$

- $1 - P[K \leq K_0]$ valor de la Intensidad Implicativa. Su fórmula es la siguiente:

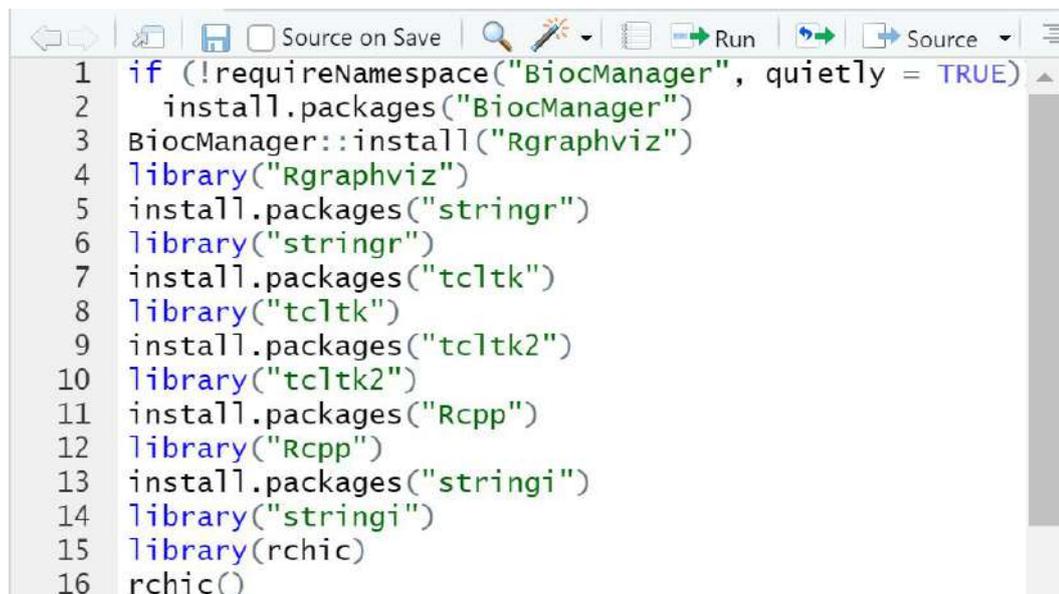
$$1 - P[K \leq K_0] = 1 - C_{k_0}^n p^{k_0} (1 - p)^{n - k_0}$$

4.1.2. Automatización en R de una Base de Datos. Análisis Estadístico Implicativo

Debemos tener en cuenta varios aspectos y conocimiento en R como es: la instalación de paquetes tanto internamente del programa, como utilizando archivos zip que se encuentran en internet. El paquete que no va ayudar en R para la automatización de una base de Datos es RChic.

Para la automatización en R debemos importar los datos de Excel en un archivo csv, extraemos la tabla de datos como Numeric Matrix y cada una de las variables con sus datos con Column Vectors. Utilizamos los siguientes pasos:

- Debemos tener previamente instalados algunos paquetes y para el cual vamos a utilizar el siguiente script:



```
1 if (!requireNamespace("BiocManager", quietly = TRUE))
2   install.packages("BiocManager")
3 BiocManager::install("Rgraphviz")
4 library("Rgraphviz")
5 install.packages("stringr")
6 library("stringr")
7 install.packages("tcltk")
8 library("tcltk")
9 install.packages("tcltk2")
10 library("tcltk2")
11 install.packages("Rcpp")
12 library("Rcpp")
13 install.packages("stringi")
14 library("stringi")
15 library(rchic)
16 rchic()
```

Ilustración 4-4: Paquetes en R del Análisis Estadístico Implicativo.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

Ya instalados estos paquetes, podemos llamar al paquete RChic y ejecutarlo

```
library(rchic)
```

```
rchic()
```

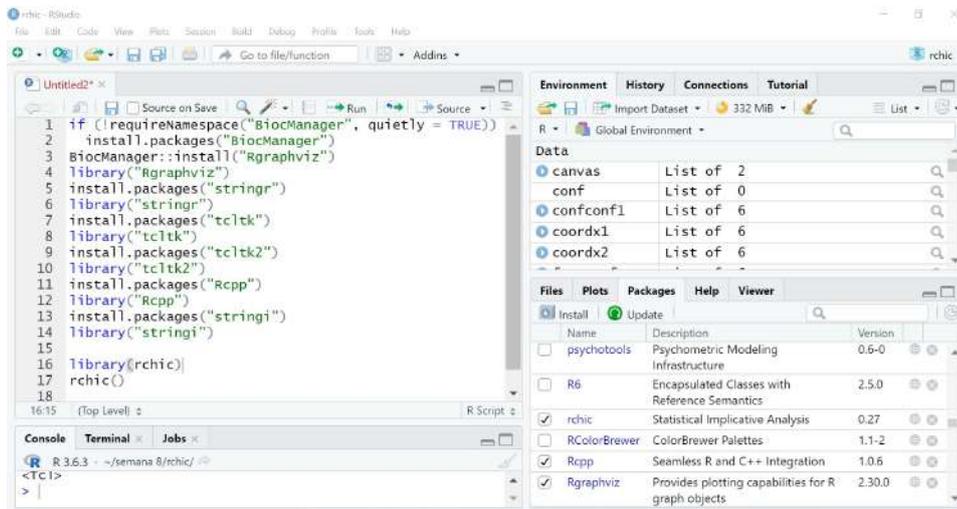


Ilustración 5-4: Script en el entorno de R.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

Nos aparecerá una ventana la cual nos permitirá realizar la automatización de la cualquier base de Datos en los temas: Árbol de similaridad, Árbol Jerárquico y Grafo Implicativo.

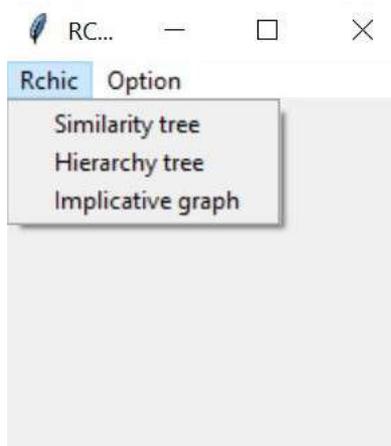


Ilustración 6-4: Ventana de

RChic.

Fuente:Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

Escogemos el tema en el que vamos a trabajar, en mi caso me toco realizar la automatización de una base de Datos en Análisis Estadístico Implicativo para el cual utilizaremos la opción Implicative Graph (Grafo Implicativo). Para esto escogemos la base de Datos que debe estar en archivo csv.

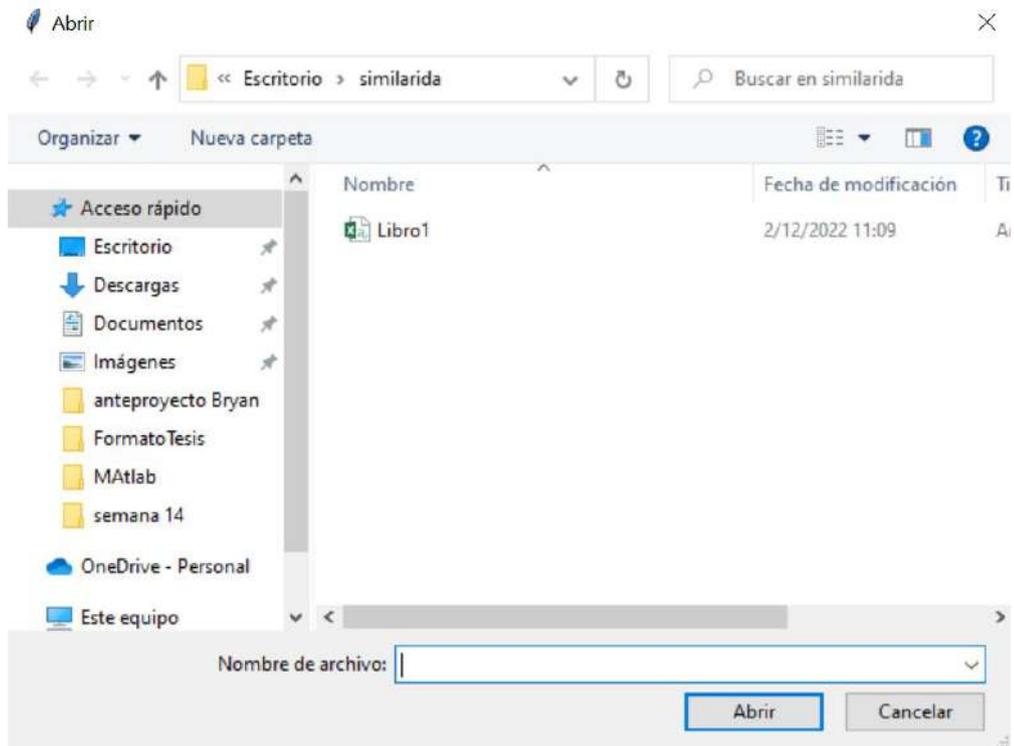


Ilustración 7-4: Ventana Selección de Archivo.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

Al escoger el archivo en csv se desplegará la ventana del grafo Implicativo con cada una de las variables

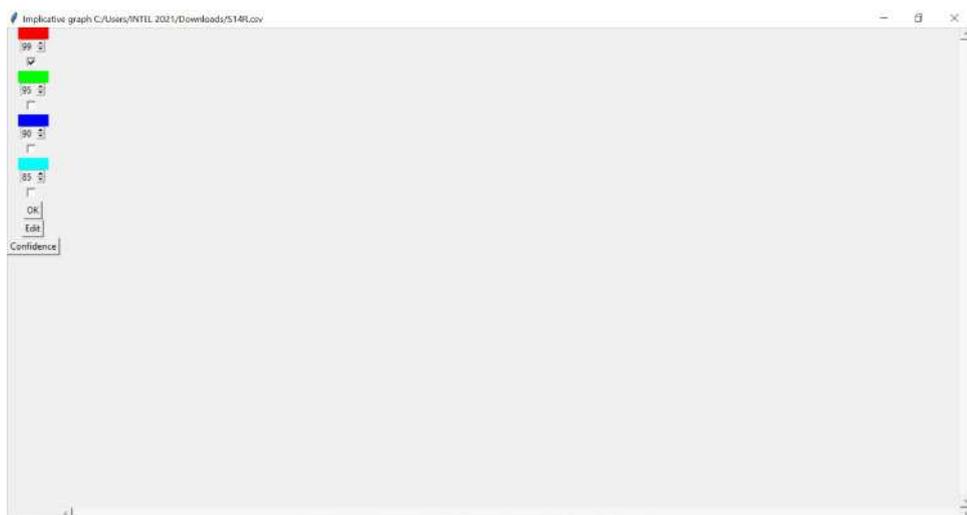


Ilustración 8-4: Ventana Implicative Graph.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

Y después de combinar las intensidades de implicación que se encuentran al lado izquierdo de la ventana de color rojo, verde, azul y celeste; obtenemos el gráfico que deseamos.

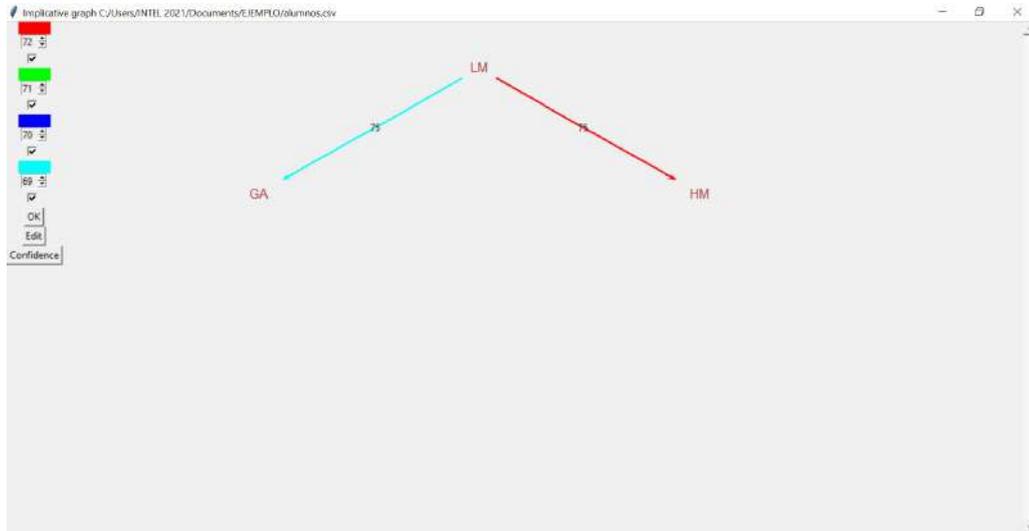


Ilustración 9-4: Ventana Grafo Implicativo con Valores.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

Como vamos a evaluar la similaridad con los niveles de cohesión así que seguimos el mismo proceso del Grafo Implicativo pero esta vez escogiendo la opción de Hierarchy tree (Árbol de Cohesión). Siguiendo los mismos pasos del Grafo Implicativo mostrará la siguientes ventanas.

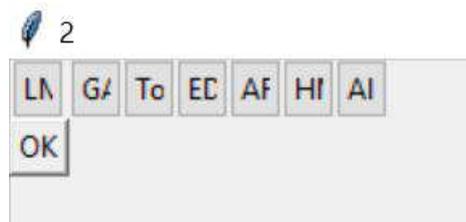


Ilustración 10-4: Ventana de variables.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

	LM	GA	To	ED	AF	HM	AI
LM	0.0000000	0.4663074	0	0	0.0000000	0.6106872	0.0000000
GA	0.3461684	0.0000000	0	0	0.0000000	0.0000000	0.0000000
To	0.0000000	0.0000000	0	0	0.0000000	0.0000000	0.0000000
ED	0.0000000	0.0000000	0	0	0.0000000	0.0000000	0.0000000
AF	0.0000000	0.0000000	0	0	0.0000000	0.0000000	0.3349839
HM	0.5064838	0.0000000	0	0	0.0000000	0.0000000	0.1595450
AI	0.0000000	0.0000000	0	0	0.3024025	0.1303408	0.0000000

Showing 1 to 7 of 7 entries, 7 total columns

Ilustración 13-4: Ventana Matriz de Cohesión.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

4.1.3. Programación en MATLAB. Algoritmo del Análisi Estadístico Implicativo

4.1.4. Programación en MATLAB. Algoritmo del Análisi Estadístico Implicativo para una base de 7 variables y 20 filas.

Para generar este algoritmo nos ayudamos con una base ya creada en Excel que contienen 7 variables y 20 filas y en ella trabajamos todas las operaciones que intervienen para obtener la Intensidad Implicativa, en la cual debemos obtener:

- a) La dimensión de la matriz (nxm).
- b) Número total de 1 (ani).
- c) Número total de 0 (anj).
- d) Tabla de Probabilidad de éxito(p1).
- e) Tabla de Cardinales (Card).
- f) Tabla de Distribución Binomial (Bin)
- g) Tabla de la Intensidad Implicativa (Int)
- h) Gráfico Implicativo (Grafo Implicativo)

Para esto hemos generado el siguiente algoritmo:

```
clc

A=xlsread('Libro1')

% el valor total de 1 en cada variable

ani=sum(A)

[n,m] = size(A)

%elvalordeOencadavariabile

anjHip = n - ani(1,1);

anjOpe = n - ani(1,2);

anjFla = n - ani(1,3);

anjHea = n - ani(1,4);

anjReg = n - ani(1,5);

anjRap = n - ani(1,6);

anjPun = n - ani(1,7);

%elvalordelaprobabilidadexito

pHIPOPE = (ani(1,1)/n) * (anjOpe/n);

pHIPPUN = (ani(1,1)/n) * (anjPun/n);

pFLARAP = (ani(1,3)/n) * (anjRap/n);

pOPERAP = (ani(1,2)/n) * (anjRap/n);

pREGHIP = (ani(1,5)/n) * (anjHip/n);

pREGPUN = (ani(1,5)/n) * (anjPun/n);

pHEAPUN = (ani(1,4)/n) * (anjPun/n);

%Cardinal

M = A(:,1);

N = A(:,2);

O = A(:,3);

P = A(:,4);

Q = A(:,5);

R = A(:,6);

S = A(:,7);

L = (M > 0) + (N == 0);

L1 = (M > 0) + (S == 0);

L2 = (O > 0) + (R == 0);

L3 = (N > 0) + (R == 0);
```

```

L4 = (Q > 0) + (M == 0);

L5 = (Q > 0) + (S == 0);

L6 = (P > 0) + (S == 0);

CardHIPOPE = sum(L > 1);

CardHIPPUN = sum(L1 > 1);

CardFLARAP = sum(L2 > 1);

CardOPERAP = sum(L3 > 1);

CardREGHIP = sum(L4 > 1);

CardREGPUN = sum(L5 > 1);

CardHEAPUN = sum(L6 > 1);

%tabla

ai = ['HIP','HIP','FLA','OPE','REG','REG','HEA'];

aj = ['OPE','PUN','RAP','RAP','HIP','PUN','PUN'];

Nai = [ani(1,1);ani(1,1);ani(1,3);ani(1,2);ani(1,5);ani(1,5);ani(1,4)];

Naj = [anjOpe;anjPun;anjRap;anjRap;anjHip;anjPun;anjPun];

p = [pHIPOPE;pHIPPUN;pFLARAP;pOPERAP;pREGHIP;pREGPUN;pHEAPUN];

Card = [CardHIPOPE;CardHIPPUN;CardFLARAP;CardOPERAP;CardREGHIP;CardREGPUN;CardHEAPUN];

T = table(ai,aj,Nai,Naj,p,Card)

%DISTRIBUCIONBINOMIAL

BINOHIPOPE = binocdf(CardHIPOPE,n,pHIPOPE);

BINOHIPPUN = binocdf(CardHIPPUN,n,pHIPPUN);

BINOFLARAP = binocdf(CardFLARAP,n,pFLARAP);

BINOOPERAP = binocdf(CardOPERAP,n,pOPERAP);

BINOREGHIP = binocdf(CardREGHIP,n,pREGHIP);

BINOREGPUN = binocdf(CardREGPUN,n,pREGPUN);

BINOHEAPUN = binocdf(CardHEAPUN,n,pHEAPUN);

DISTRIBUCION – BINOMIAL = [BINOHIPOPE;BINOHIPPUN;BINOFLARAP;BINOOPERAP;BINOREGHIP;BINOREGPUN;BINOHEAPUN];

T1 = table(DISTRIBUCIONBINOMIAL)

%INTENSIDADIMPLICATIVA

IHIPOPE = 1 - BINOHIPOPE;

IHIPPUN = 1 - BINOHIPPUN;

IFLARAP = 1 - BINOFLARAP;

IOOPERAP = 1 - BINOOPERAP;

IREGHIP = 1 - BINOREGHIP;

```

```

IREGPUN = 1 - BINOREGPUN;

IHEAPUN = 1 - BINOHEAPUN;

I = [(HIP,OPE)';(HIP,PUN)';(FLA,RAP)';(OPE,RAP)';(REG,HIP)';(REG,PUN)';(HEA,PUN)];

INTENSIDAD - IMPLICATIVA = [IHIPOPE;IHIPPUN;IFLARAP;IOPERAP;IREGHIP;IREGPUN;IHEAPUN];

T2 = table(I,INTENSIDAD - IMPLICATIVA)

HIP = 1;

OPE = 2;

FLA = 3;

HEA = 4;

REG = 5;

RAP = 6;

PUN = 7;

s = [OPE OPE FLA REG REG REG HIP HEA];

t = [FLARAP RAP HIP HEA PUN PUN PUN];

weights = [0 IOPERAP IFLARAP IREGHIP 0 IREGPUN IHIPPUN IHEAPUN];

names = {'HIP' 'OPE' 'FLA' 'HEA' 'REG' 'RAP' 'PUN'};

G = graph(s,t,weights,names)

plot(G,'EdgeLabel',G.Edges.Weight)

```

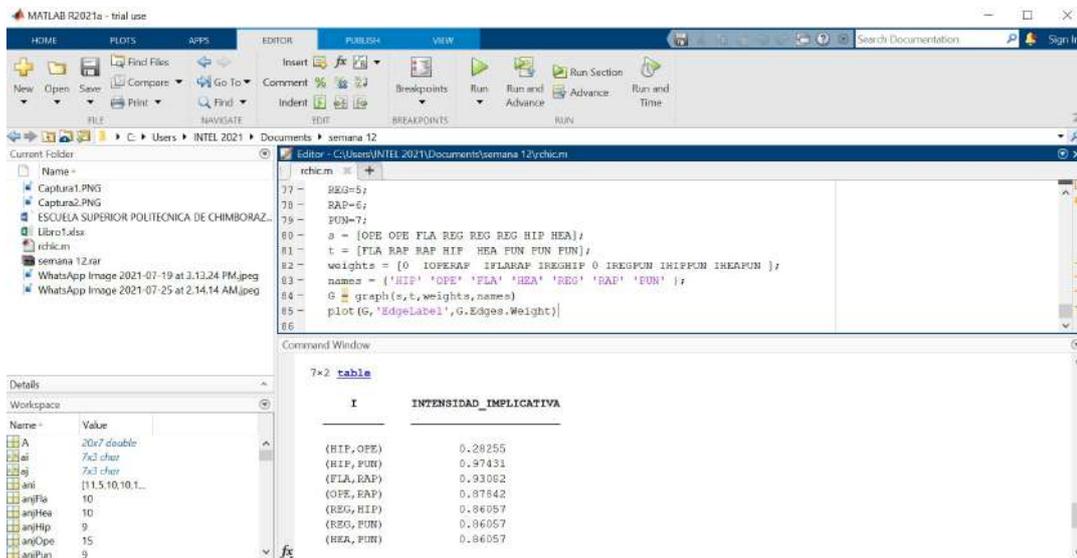


Ilustración 14-4: Algoritmo de una base aleatoria de 7 variables y 20 filas.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

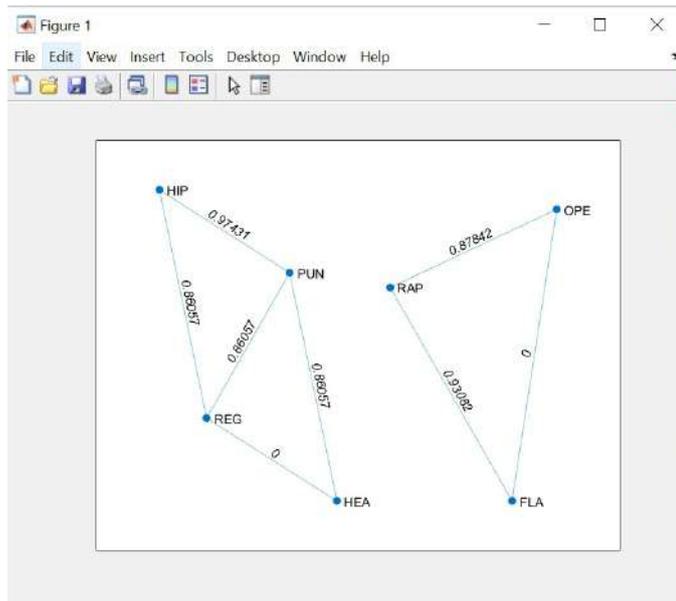


Ilustración 15-4: Grafo Implicativo MATLAB de una base aleatoria de 7 variables y 20 filas.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

Ahora vamos a comparar nuestro Grafo Implicativo de Matlab, con el que nos genera R con la paquete de RChic.

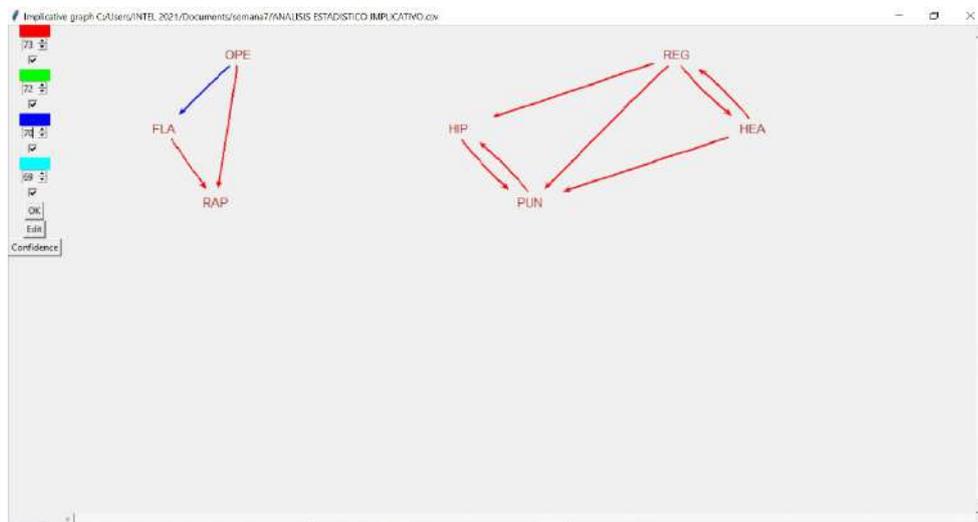


Ilustración 16-4: Grafo Implicativo R de una base aleatoria de 7 variables y 20 filas.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

EL cuál nos comprueba que nuestras operaciones en Matlab están bien elaboradas ya que coinciden con los datos y gráfica en R.

4.1.5. Programación en MATLAB. Algoritmo del Análisis Estadístico Implicativo para una base de 7 variables y 200 filas.

Para generar este algoritmo nos ayudamos con una base ya creada en Excel que contienen 7 variables y 200 filas y en ella trabajamos todas las operaciones que intervienen para obtener la Intensidad Implicativa, en la cual debemos obtener:

- a) La dimensión de la matriz (nxm).
- b) Número total de 1 (ani).
- c) Número total de 0 (anj).
- d) Tabla de Probabilidad de éxito(p1).
- e) Tabla de Cardinales (Card).
- f) Tabla de Distribución Binomial (Bin)
- g) Tabla de la Intensidad Implicativa (Int)
- h) Gráfico Implicativo (Grafo Implicativo)

Para esto hemos generado el siguiente algoritmo:

```
clc
A=xlsread('Libro1')
% el valor total de 1 en cada variable
ani=sum(A)
[n,m] = size(A)
%elvalordeOencadavariabile
anjHip = n - ani(1,1);
anjOpe = n - ani(1,2);
anjFla = n - ani(1,3);
anjHea = n - ani(1,4);
anjReg = n - ani(1,5);
anjRap = n - ani(1,6);
anjPun = n - ani(1,7);
%elvalordelaprobabilidaddexito
pHIPOPE = (ani(1,1)/n)*(anjOpe/n);
pHIPPUN = (ani(1,1)/n)*(anjPun/n);
pFLARAP = (ani(1,3)/n)*(anjRap/n);
pOPERAP = (ani(1,2)/n)*(anjRap/n);
```

```

pREGHIP = (ani(1,5)/n)*(anjHip/n);

pREGPUN = (ani(1,5)/n)*(anjPun/n);

pHEAPUN = (ani(1,4)/n)*(anjPun/n);

%Cardinal

M = A(:, 1);

N = A(:, 2);

O = A(:, 3);

P = A(:, 4);

Q = A(:, 5);

R = A(:, 6);

S = A(:, 7);

L = (M > 0) + (N == 0);

L1 = (M > 0) + (S == 0);

L2 = (O > 0) + (R == 0);

L3 = (N > 0) + (R == 0);

LA = (Q > 0) + (M == 0);

L5 = (Q > 0) + (S == 0);

L6 = (P > 0) + (S == 0);

CardHIPOPE = sum(L > 1);

CardHIPPUN = sum(L1 > 1);

CardFLARAP = sum(L2 > 1);

CardOPERAP = sum(L3 > 1);

CardREGHIP = sum(LA > 1);

CardREGPUN = sum(L5 > 1);

CardHEAPUN = sum(L6 > 1);

%tabla

ai = ['HIP' ; 'HIP' ; 'FLA' ; 'OPE' ; 'REG' ; 'REG' ; 'HEA'];

aj = ['OPE' ; 'PUN' ; 'RAP' ; 'RAP' ; 'HIP' ; 'PUN' ; 'PUN'];

Nai = [ani(1, 1); ani(1, 1); ani(1, 3); ani(1, 2); ani(1, 5); ani(1, 5); ani(1, 4)];

Naj = [anjOpe; anjPun; anjRap; anjRap; anjHip; anjPun; anjPun];

p = [pHIPOPE; pHIPPUN; pFLARAP; pOPERAP; pREGHIP; pREGPUN; pHEAPUN];

Card = [CardHIPOPE; CardHIPPUN; CardFLARAP; CardOPERAP; CardREGHIP; CardREGPUN; CardHEAPUN];

T = table(ai, aj, Nai, Naj, p, Card)

%DISTRIBUCINBINOMIAL

```

```

BINOHIPOPE = binocdf(CardHIPOPE,n,pHIPOPE);

BINOHIPPUN = binocdf(CardHIPPUN,n,pHIPPUN);

BINOFLARAP = binocdf(CardFLARAP,n,pFLARAP);

BINOOPERAP = binocdf(CardOPERAP,n,pOPERAP);

BINOREGHIP = binocdf(CardREGHIP,n,pREGHIP);

BINOREGPUN = binocdf(CardREGPUN,n,pREGPUN);

BINOHEAPUN = binocdf(CardHEAPUN,n,pHEAPUN);

DISTRIBUCION – BINOMIAL = [BINOHIPOPE;BINOHIPPUN;BINOFLARAP;BINOOPERAP;BINOREGHIP;BINOREGPUN;BINOHEAPUN];

DISTRIBUCION – BINOMIAL = [BINOHIPOPE;BINOHIPPUN;BINOFLARAP;BINOOPERAP;BINOREGHIP;BINOREGPUN;BINOHEAPUN];

T1 = table(DISTRIBUCION_BINOMIAL)

%INTENSIDADIMPLICATIVA

IHIPOPE = 1 – BINOHIPOPE;

IHIPPUN = 1 – BINOHIPPUN;

IFLARAP = 1 – BINOFLARAP;

IOOPERAP = 1 – BINOOPERAP;

IREGHIP = 1 – BINOREGHIP;

IREGPUN = 1 – BINOREGPUN;

IHEAPUN = 1 – BINOHEAPUN;

I = [(HIP,OPE)';(HIP,PUN)';(FLA,RAP)';(OPE,RAP)';(REG,HIP)';(REG,PUN)';(HEA,PUN)'];

INTENSIDAD – IMPLICATIVA = [IHIPOPE;IHIPPUN;IFLARAP;IOOPERAP;IREGHIP;IREGPUN;IHEAPUN];

T2 = table(I,INTENSIDAD – IMPLICATIVA)

HIP = 1;

OPE = 2;

FLA = 3;

HEA = 4;

REG = 5;

RAP = 6;

PUN = 7;

s = [OPE OPE FLAREGREGREGHIPHEA];

t = [FLARAPRAPHIPHEAPUNPUNPUN];

weights = [0 IOOPERAP IFLARAP IREGHIP 0 IREGPUN IHIPPUN IHEAPUN];

names = {'HIP' 'OPE' 'FLA' 'HEA' 'REG' 'RAP' 'PUN'};

G = graph(s,t,weights,names)

plot(G,'EdgeLabel',G.Edges.Weight)

```

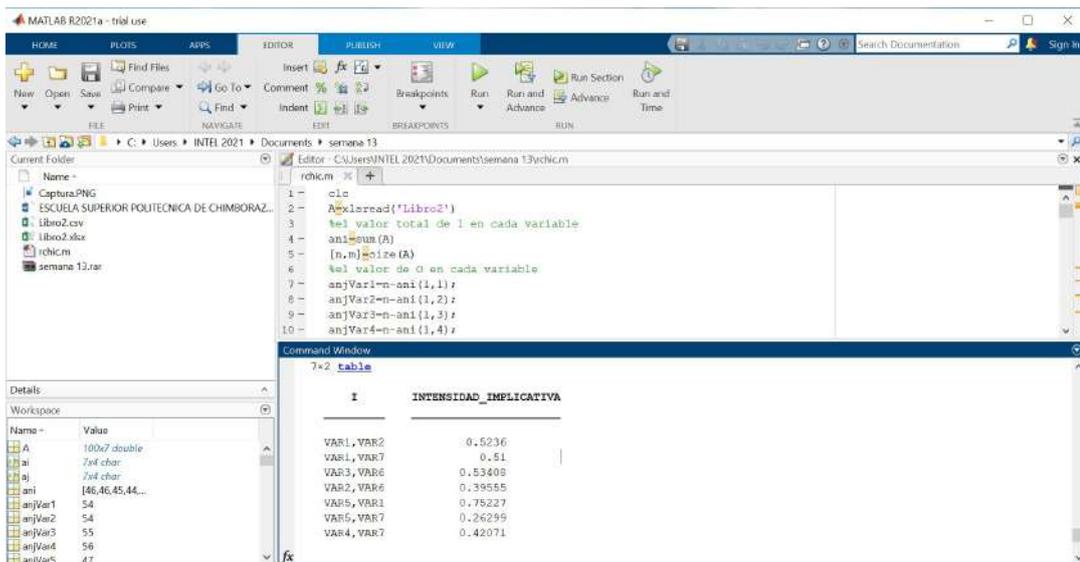


Ilustración 17-4: Algoritmo de una base aleatoria de 7 variables y 200 filas.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

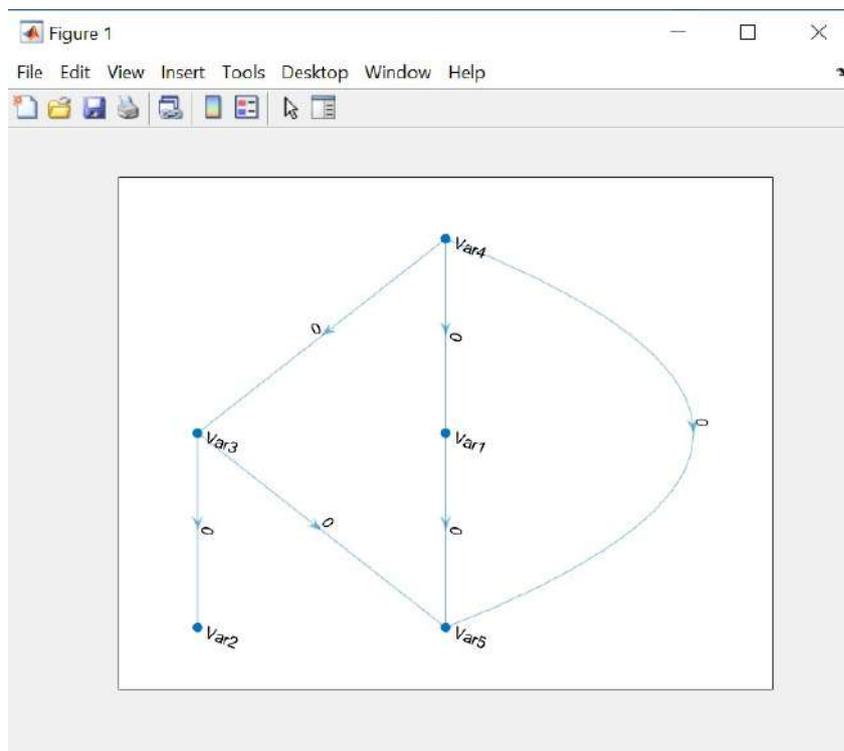


Ilustración 18-4: Grafo Implicativo MATLAB de una base aleatoria de 7 variables y 200 filas.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

Ahora vamos a comparar nuestro Grafo Implicativo de Matlab, con el que nos genera R con la paquete de RChic.

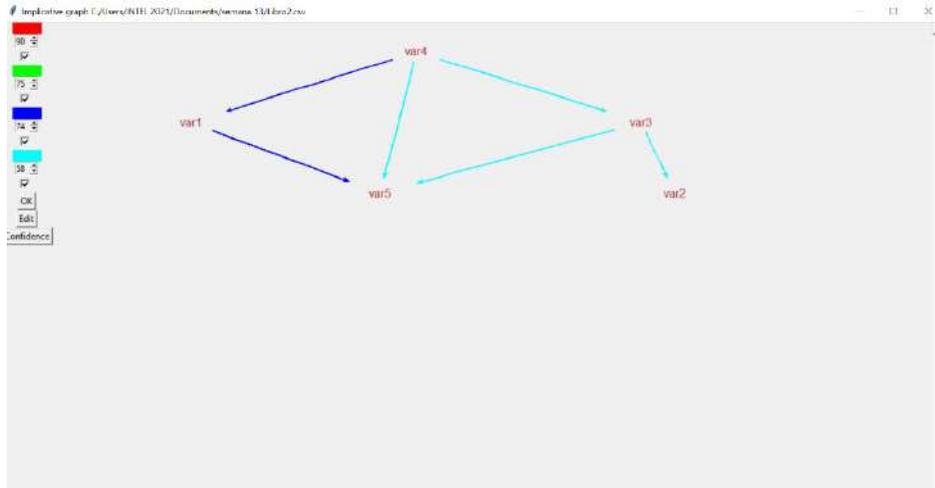


Ilustración 19-4: Grafo Implicativo R de una base aleatoria de 7 variables y 200 filas.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

EL cuál nos comprueba que nuestras operaciones en Matlab están bien elaboradas ya que coinciden con los datos y gráfica en R.

4.1.6. Programación en MATLAB. Algoritmo del Análisi Estadístico Implicativo para una base aleatoria de m variables y n filas.

Para generar este algoritmo nos ayudamos con una base aleatoria creada en Matlab que contienen m variables y n filas y en ella trabajamos todas las operaciones que intervienen para obtener la Intensidad Implicativa de una manera general, en la cual debemos obtener:

- a) La dimensión de la matriz ($n \times m$).
- b) Número total de 1 (a_{ni}).
- c) Número total de 0 (a_{nj}).
- d) Tabla de Probabilidad de éxito (p_1).
- e) Tabla de Cardinales (Card).
- f) Tabla de Distribución Binomial (Bin).

g) Tabla de la Intensidad Implicativa (Int).

h) Gráfico (Grafo Implicativo).

Para esto hemos generado el siguiente algoritmo:

```
clc

n=input('Ingrese el numero de filas=');

m=input('Ingrese el numero de variables=');

A=randi([0,1],n,m);

A

[n,m] = size(A)

%elvalortotalde1encadavariabile

ani = sum(A)

%elvalorde0encadavariabile

anj = n - ani(1,:);

%elvalordeprobabilidaddexito

for i = 1 : m

for j = 1 : m

if i = j

p(i,j) = (ani(:,i)/n).*(anj(:,j)/n);

p1(i,j) = p(i,j);

end

end

end

p1

%Cardinal

for i = 1 : m

for j = 1 : m

if i = j

Card(i,j) = sum((A(:,i) > 0).*(A(:,j) == 0));

end

end

end

Card

%DISTRIBUCIONBINOMIAL

for i = 1 : m
```

```

for j = 1 : m

    if i = j

        Bin(i, j) = binocdf(Card(i, j), n, p1(i, j));

    end

end

end

Bin

%INTENSIDADIMPLICATIVA

for i = 1 : m

    for j = 1 : m

        if i = j

            Int1(i, j) = 1 - Bin(i, j);

        end

    end

end

Int1

```

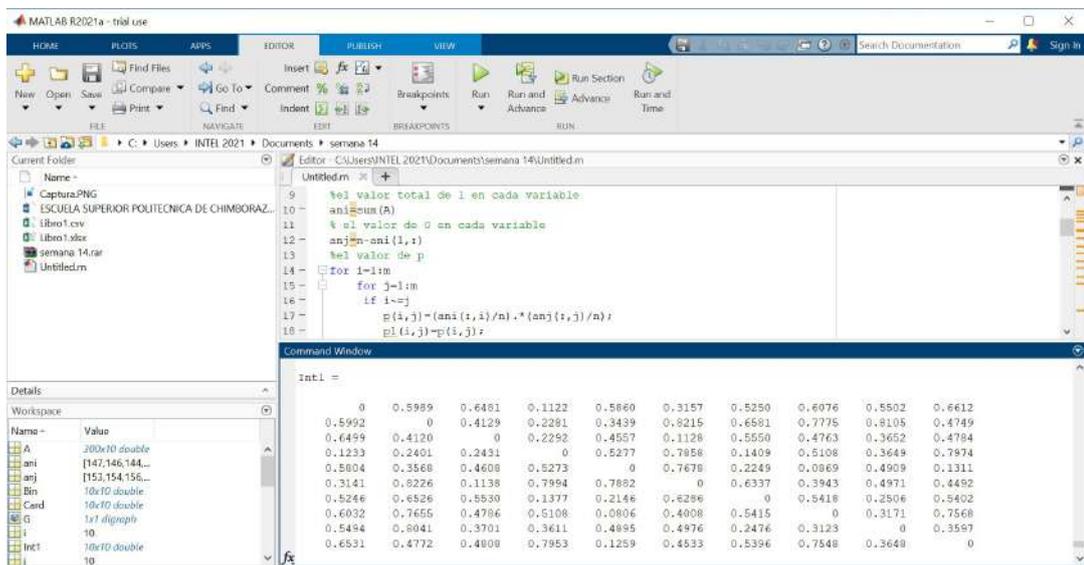


Ilustración 20-4: Algoritmo de una base aleatoria de m variables y n filas.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

Ahora vamos a comparar nuestro Grafo Implicativo de Matlab, con el que nos genera R con la paquete de RChic.

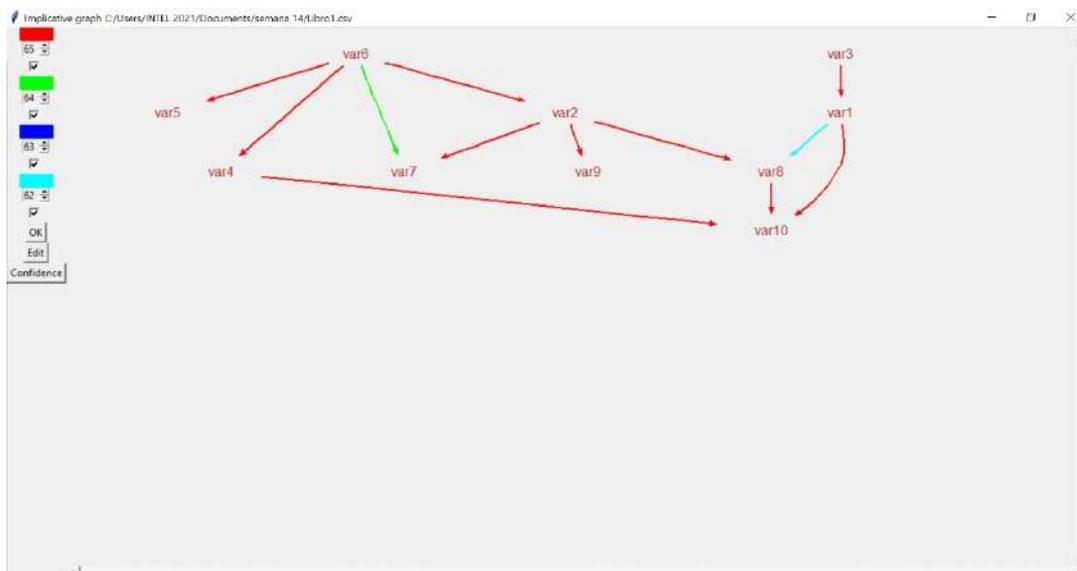


Ilustración 21-4: Grafo Implicativo R de una base aleatoria de m variables y n filas.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

EL cuál nos comprueba que nuestras operaciones en Matlab están bien elaboradas ya que coinciden con los datos y gráfica en R.

4.2. Discusión

4.2.1. Estadísticos descriptivos: Matriz Cohesión (M-Chic y R-Chic), Matriz Cohesión (M-Chic y Chic), Gráfica (R-Chic y Chic), Nodos (M-Chic y R-Chic), Nodos (M-Chic y Chic)

Tabla 1-4: Comparaciones Estadísticas 1

Estadísticas							
Variable	Conteo total	N	N*	Porcentaje	PrcAcum	Media	Desv.Est.
Matriz Cohesión (M-Chic y R-Chic)	383	383	0	100	100	0.99760	0.01295
Matriz Cohesión (M-Chic y Chic)	383	383	0	100	100	0.99315	0.03147
Gráfica (R-Chic y Chic)	383	383	0	100	100	0.83341	0.04579
Nodos (M-Chic y R-Chic)	383	383	0	100	100	0.93284	0.03182
Nodos (M-Chic y Chic)	383	383	0	100	100	0.92880	0.03626

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

El análisis que utiliza un promedio mayor es el de la Matriz Cohesión (M-Chic y R-Chic) con un valor igual al 0.99760 con un conteo total de 383 bases evaluadas siendo el 100% de las bases totales de la muestra, y obteniendo una desviación estándar del 0.01295. El análisis que utiliza un promedio menor es el de la Gráfica (R-Chic y Chic) con un valor igual al 0.83341 con el conteo total de 383 bases evaluadas siendo el 100% de las bases totales de la muestra, y obteniendo una desviación estándar del 0.04579. Para el análisis Matriz Cohesión (M-Chic y Chic) con un conteo total de 383 bases evaluadas siendo el 100% de las bases totales de la muestra, obteniendo un promedio igual a 0.99315 y una desviación estándar del 0.03147. El análisis de los Nodos (M-Chic y R-Chic) con un conteo total de 383 bases evaluadas siendo el 100% de las bases totales de la muestra, obteniendo un promedio igual a 0.93284 y una desviación estándar del 0.03182. Y como último análisis tenemos en de los Nodos (M-Chic y Chic) con un conteo total de 383 bases evaluadas siendo el 100% de las bases totales de la muestra, obteniendo un promedio igual a 0.92880 y una desviación estándar del 0.03626.

Tabla 2-4: Comparaciones Estadísticas 2

Estadísticas							
Variable	Varianza	CoefVar	Mínimo	Q1	Mediana	Q3	Máximo
Matriz Cohesión (M-Chic y R-Chic)	0.000168	1.30	0.82227	1.00000	1.00000	1.00000	1.00000
Matriz Cohesión (M-Chic y Chic)	0.00099	3.17	0.47193	1.00000	1.00000	1.00000	1.00000
Gráfica (R-Chic y Chic)	0.00210	5.49	0.70000	0.79400	0.83500	0.87100	0.93000
Nodos (M-Chic y R-Chic)	0.00101	3.41	0.88000	0.90400	0.93000	0.96000	0.99000
Nodos (M-Chic y Chic)	0.00132	3.90	0.87000	0.89700	0.92400	0.96400	0.99000

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

El análisis de la Matriz Cohesión (M-Chic y R-Chic) obtiene una varianza igual a 0.00168, su coeficiente de variación igual a 1.30, siendo el mínimo 0.82227 y su máximo igual a 1, el primer cuartil muestra que el 25% del análisis es menor o igual a 1 y tercer cuartil nos representa que el 75% del análisis es mayor o igual a 1 dando a entender que las bases evaluadas son simétricas. Para el análisis de la Matriz Cohesión (M-Chic y Chic) obtiene una varianza igual a 0.00099, su coeficiente de varianza igual a 3.17, siendo el mínimo 0.47193 y su máximo igual a 1, el primer cuartil muestra que el 25% del análisis es menor o igual a 1 y tercer cuartil nos representa que el 75% del análisis es mayor o igual a 1 dando a entender que las bases evaluadas son simétricas. El análisis de la Gráfica (R-Chic y Chic) teniendo una varianza igual a

Tabla 3-4: Comparaciones Estadísticas 3

Variable	Estadísticas				
	IQR	Modo	N para moda	Asimetría	Curtosis
Matriz Cohesión (M-Chic y R-Chic)	0.000000	1	339	-9.71	112.40
Matriz Cohesión (M-Chic y Chic)	0.000000	1	296	-12.70	200.28
Gráfica (R-Chic y Chic)	0.077000	0.882	8	-0.01	-1.01
Nodos (M-Chic y R-Chic)	0.056000	0.956	9	0.04	-1.22
Nodos (M-Chic y Chic)	0.067000	0.897	8	0.10	-1.23

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

0.00210 con un coeficiente de variación igual a 5.49, teniendo como mínimo igual a 0.70000 y su máximo 0.93000, el primer cuartil representa que el 25 % del análisis es menor o igual a 0.79400 y el tercer cuartil representa el 75 % del análisis es mayor o igual 0.87100. El análisis de los Nodos (M-Chic y R-Chic) teniendo una varianza igual a 0.00101 con un coeficiente de variación igual a 3.41, teniendo como mínimo igual a 0.88000 y su máximo 0.99000, el primer cuartil representa que el 25 % del análisis es menor o igual a 0.90400 y el tercer cuartil representa el 75 % del análisis es mayor o igual 0.96000. Y por último el análisis de los Nodos (M-Chic y Chic) teniendo una varianza igual a 0.00132 con un coeficiente de variación igual a 3.90, teniendo como mínimo igual a 0.87000 y su máximo 0.99000, el primer cuartil representa que el 25 % del análisis es menor o igual a 0.89700 y el tercer cuartil representa el 75 % del análisis es mayor o igual 0.96400.

El análisis de la Matriz Cohesión (M-Chic y R-Chic) al no haber dispersión de datos el rango intercuartil (IQR) es igual a 0, la moda es igual a 1 permitiendo notar que los datos obtenidos son simétricos, evaluando 339 bases que son el 88 % de las bases evaluadas tienen una asimetría del -9.71 y su curtosis es igual a 112.40 este resultado se encuentra por encima de la curva normal. El análisis de la Matriz Cohesión (M-Chic y Chic) al no haber dispersión de datos el rango intercuartil (IQR) es igual a 0, la moda es igual a 1 permitiendo notar que los datos obtenidos son simétricos, evaluando 296 bases que son el 77 % de las bases evaluadas tienen una asimetría del -12.70 y su curtosis es igual a 200.28 este resultado se encuentra por encima de la curva normal. Para el análisis de la Gráfica (R-Chic y Chic) al obtener una dispersión de datos el rango intercuartil (IQR) es igual a 0.07700, la moda es igual a 0.882, evaluando 8 bases

que son el 2% de las bases evaluadas tienen una asimetría del -0.01 y su curtosis es igual a -1.01 este resultado se encuentra por debajo de la curva normal. El análisis de los Nodos (M-Chic y R-Chic) al obtener una dispersión de datos el rango intercuartil (IQR) es igual a 0.05600, la moda es igual a 0.956, evaluando 9 bases que son el 2% de las bases evaluadas tienen una asimetría del 0.04 y su curtosis es igual a -1.22 este resultado se encuentra por debajo de la curva normal. Y por último el análisis de los Nodos (M-Chic y Chic) al obtener una dispersión de datos el rango intercuartil (IQR) es igual a 0.06700, la moda es igual a 0.897, evaluando 8 bases que son el 2% de las bases evaluadas tienen una asimetría del 0.10 y su curtosis es igual a -1.23 este resultado se encuentra por debajo de la curva normal.

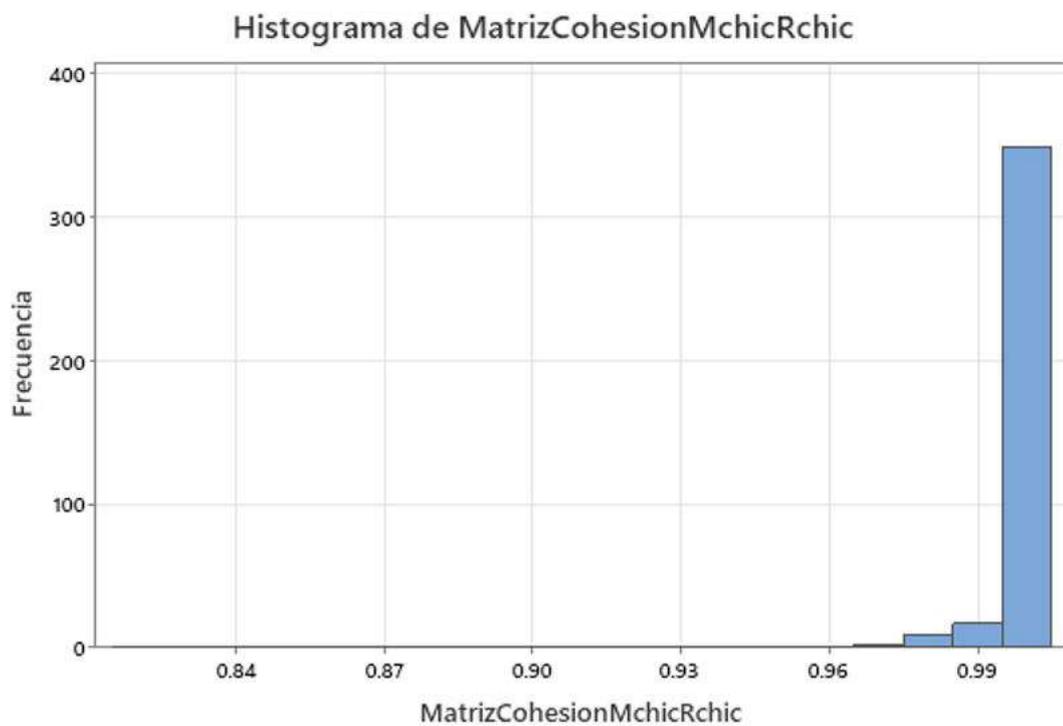


Ilustración 22-4: Histograma Matriz de Cohesión M-Chic y R-Chic

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

En la Ilustración 4-24 muestra el análisis para la Matriz Cohesión (M-Chic y R-Chic), cuya gráfica es un histograma donde se puede notar que existe medias y varianzas distintas, en el intervalo de 0.96 a 1 representa que la mayoría de datos aparentemente representan una distribución normal, por lo tanto al no existir datos aislados las bases evaluadas son simétricas pero su asimetría es negativa.

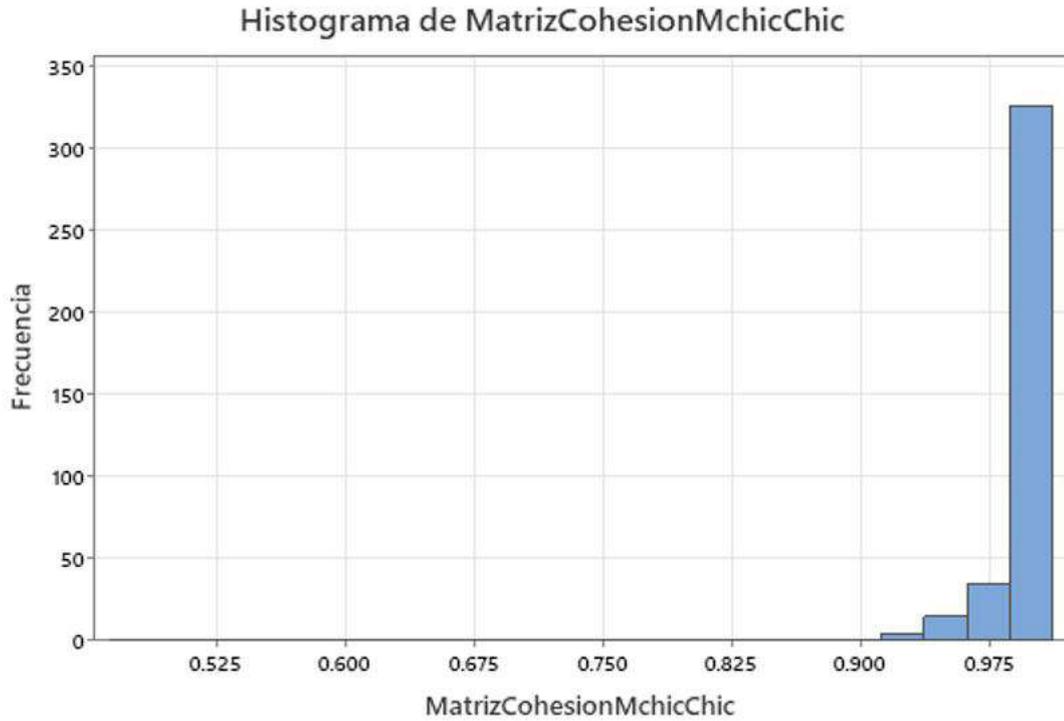


Ilustración 23-4: Histograma Matriz de Cohesión M-Chic y Chic.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

En la Ilustración 4-25 muestra el análisis para la Matriz Cohesión (M-Chic y Chic), cuya gráfica es un histograma donde se puede notar que existe medias y varianzas distintas, en el intervalo de 0.90 a 1 representa que la mayoría de datos aparentemente representan una distribución normal, por lo tanto al no existir datos aislados las bases evaluadas son simétricas pero su asimetría es negativa.

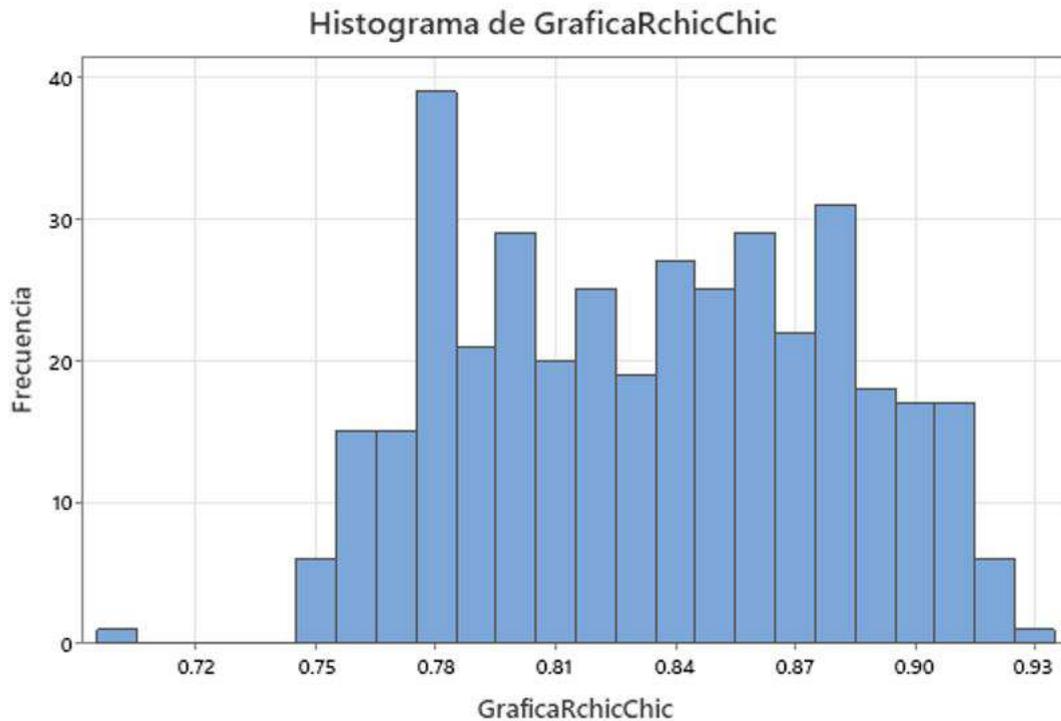


Ilustración 24-4: Histograma Gráficas M-Chic y R-Chic.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

En la Ilustración 4-26 muestra el análisis para las Gráficas M-Chic y R-Chic, cuya gráfica es un histograma donde se puede notar que existe medias y varianzas distintas, en el intervalo de 0.50 a 0.72 representa uno datos estan aislados los cuales son las bases con variables 0 a 15 ya que el programa trabaja con mayor presición cuando el número de variables es mayor a 20, en el intervalo de 0.74 a 1 se encuentra la mayor cantidad de datos con distribuciones diferentes y por lo tanto su asimetría es negativa.

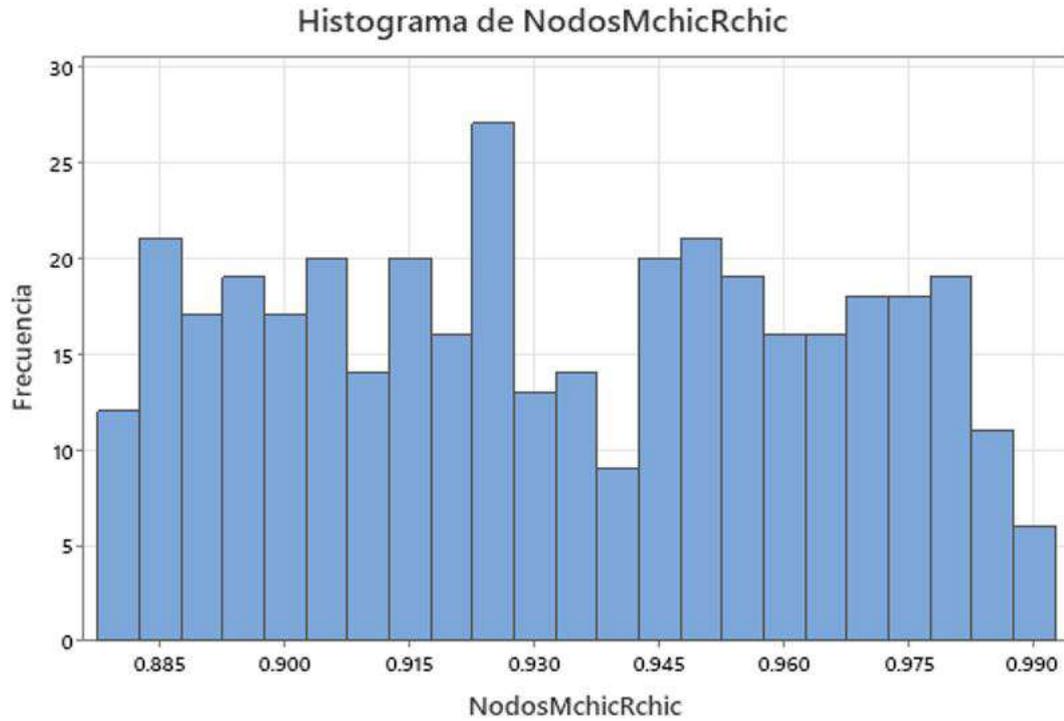


Ilustración 25-4: Histograma Nodos de R-Chic y Chic.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

En la Ilustración 4-27 muestra el análisis para las Nodos de R-Chic y Chic, cuya gráfica es un histograma donde se puede notar que existe medias y varianzas distintas, en el intervalo de 0.85 a 1 se encuentra la mayor cantidad de datos con distribuciones diferentes, pero se asemeja a una distribución normal y por lo tanto su asimetría es positiva.

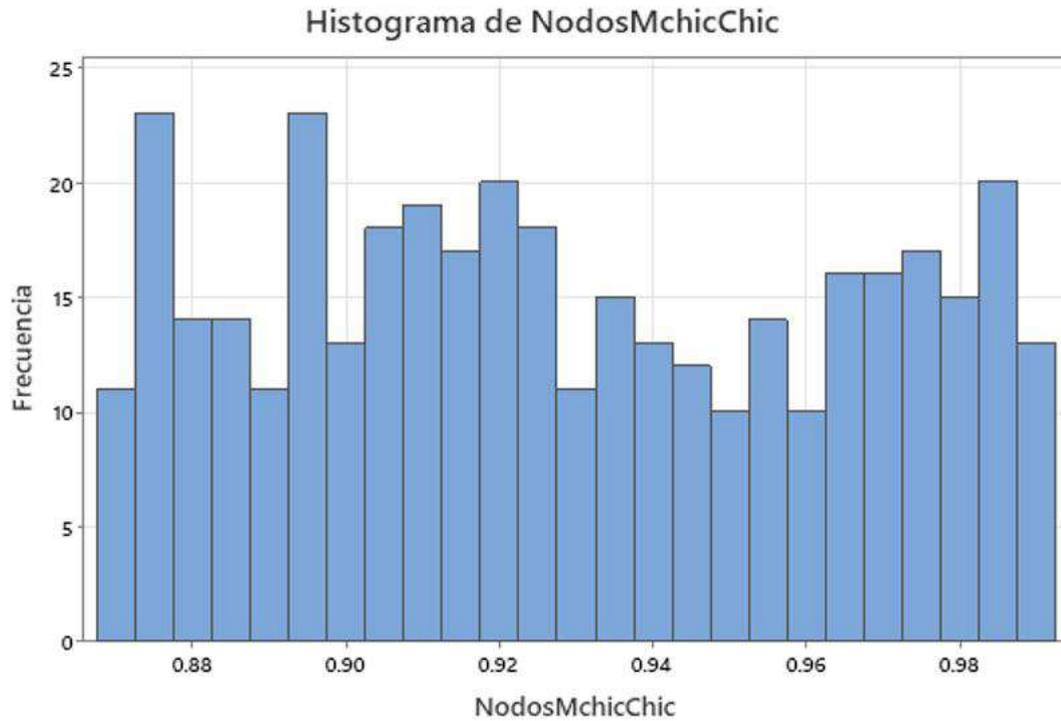


Ilustración 26-4: Histograma Nodos de M-Chic y Chic.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

En la Ilustración 4-28 muestra el análisis para las Nodos de R-Chic y Chic, cuya gráfica es un histograma donde se puede notar que existe medias y varianzas distintas, en el intervalo de 0.80 a 1 se encuentra la mayor cantidad de datos con distribuciones diferentes, pero se asemeja a una distribución normal y por lo tanto su asimetría es positiva.

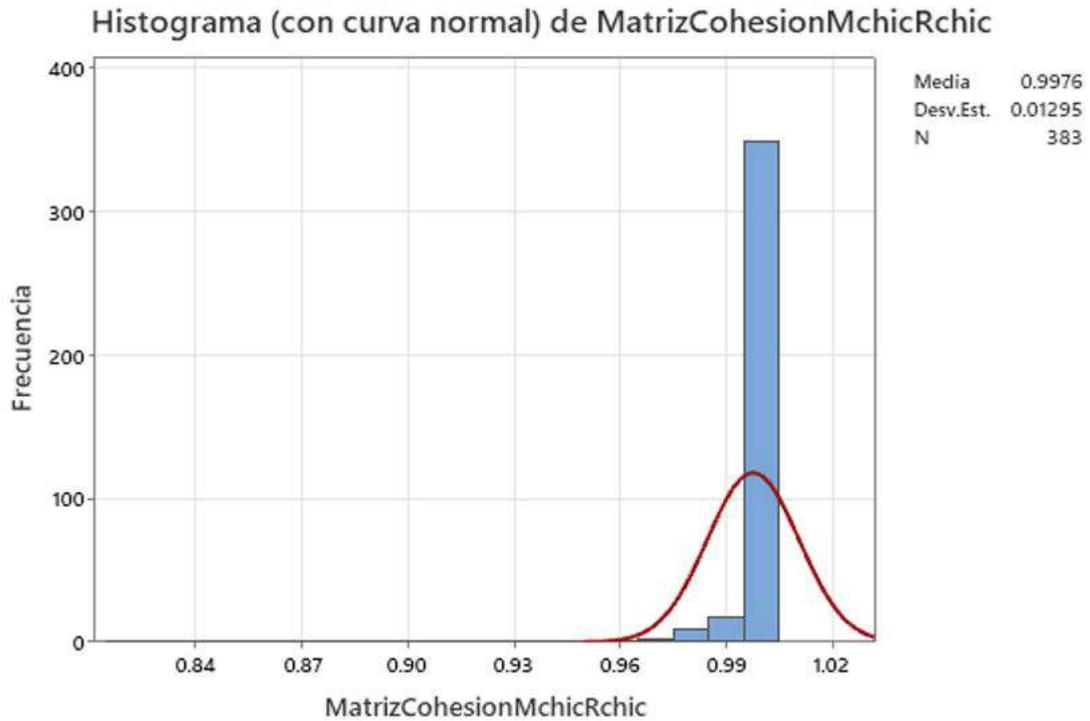


Ilustración 27-4: Histograma (con curva normal) de Matriz de Cohesión M-Chic y R-Chic.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

En la Ilustración 4-29 muestra el análisis para la Matriz Cohesión (M-Chic y R-Chic), cuya gráfica es un histograma con curva normal donde se puede notar que existe media igual a 0.9976 y una desviación estándar de 0.01295, en el intervalo de 0.96 a 1 representa que la mayoría de datos aparentemente representan una distribución normal pero sobrepasa en el pico más alto, por lo tanto al no existir datos aislados las bases evaluadas son simétricas pero su asimetría es negativa.

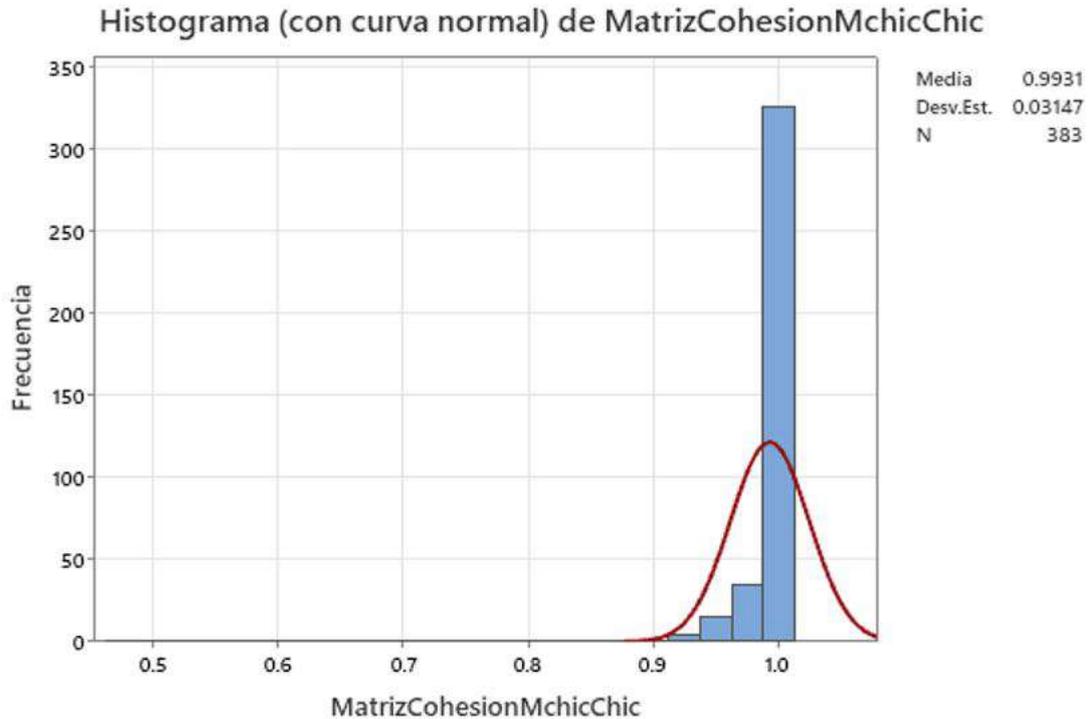


Ilustración 28-4: Histograma (con curva normal) Matriz de Cohesión M-Chic y Chic.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

En la Ilustración 4-30 muestra el análisis para la Matriz Cohesión (M-Chic y Chic), cuya gráfica es un histograma con curva normal donde se puede notar una media igual a 0.9931 y una desviación estándar de 0.03147, en el intervalo de 0.90 a 1 representa que la mayoría de datos aparentemente representan una distribución normal pero sobrepasa en el pico más alto, por lo tanto al no existir datos aislados las bases evaluadas son simétricas pero su asimetría es negativa.

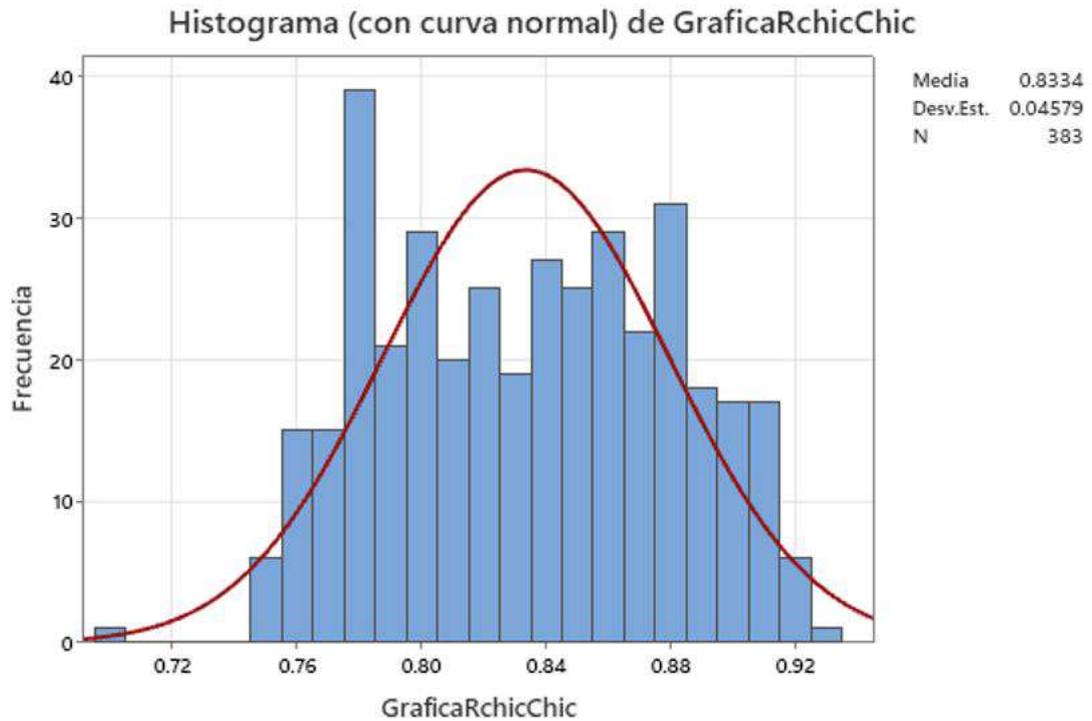


Ilustración 29-4: Histograma (con curva normal) de Gráfica R-Chic y Chic.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

En la Ilustración 4-31 muestra el análisis para las Gráficas M-Chic y R-Chic, cuya gráfica es un histograma con curva normal donde se puede notar una media igual a 0.8334 y una desviación estándar de 0.04579, en el intervalo de 0.50 a 0.72 representa uno datos estan aislados los cuales son las bases con variables 0 a 15 ya que el programa trabaja con mayor precisión cuando el número de variables es mayor a 20, en el intervalo de 0.74 a 1 se encuentra la mayor cantidad de datos con distribuciones diferentes sobrepasando la curva normal y por lo tanto su asimetría es negativa.

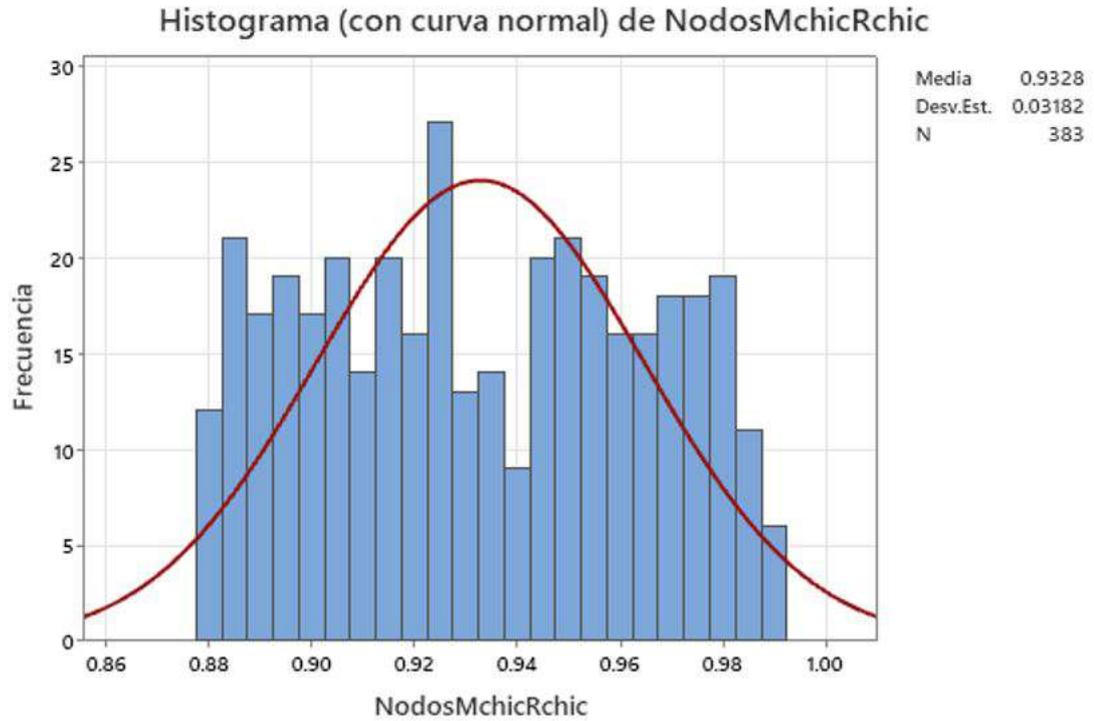


Ilustración 30-4: Histograma (con curva normal) de Nodos M-Chic y R-Chic.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

En la Ilustración 4-32 muestra el análisis para las Nodos de M-Chic y R-Chic, cuya gráfica es un histograma con curval normal donde se puede notar una media igual a 0.9328 y una desviación estándar de 0.3182, en el intervalo de 0.85 a 1 se encuentra la mayor cantidad de datos con distribuciones diferentes sobrepasando la curva normal, pero en la mayor de las distribuciones estan dentro de la curva normal y por lo tanto su asimetría es positiva.

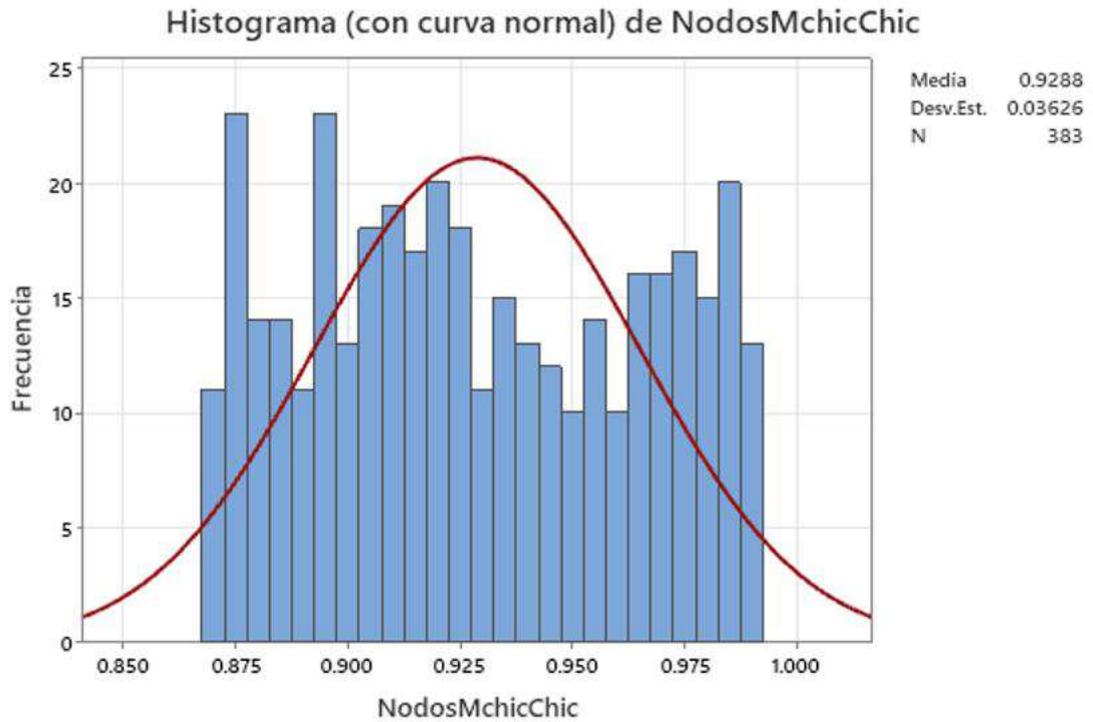


Ilustración 31-4: Histograma (con curva normal) de Nodos M-Chic y Chic.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

En la Ilustración 4-33 muestra el análisis para las Nodos de M-Chic y Chic, cuya gráfica es un histograma con curval normal donde se puede notar una media igual a 0.9288 y una desviación estándar de 0.03626, en el intervalo de 0.80 a 1 se encuentra la mayor cantidad de datos con distribuciones diferentes sobrepasando la curva normal, pero en la mayor de las distribuciones estan dentro de la curva normal y por lo tanto su asimetría es positiva.

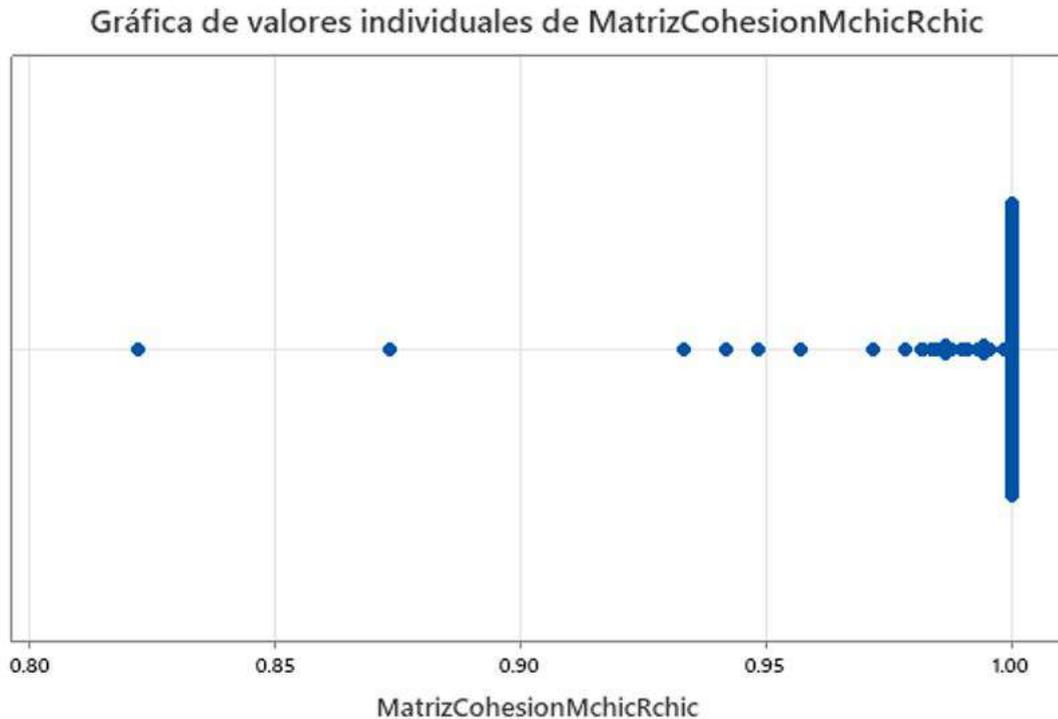


Ilustración 32-4: Gráfica de valores individuales de la Matriz Cohesión M-Chic y R-Chic.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

En la Ilustración 4-34 muestra el análisis para la Matriz Cohesión (M-Chic y R-Chic), cuya gráfica representa los valores individuales donde en los intervalos 0.80 a 0.85 y 0.85 a 0.90 se muestra un valor, en el intervalo de 0.90 a 0.95 se muestra tres valores y en intervalo de 0.95 a 1 se encuentran el resto de valores y la mayoría se encuentra en 1, por lo cual podemos concluir que las bases evaluadas son simétricas.

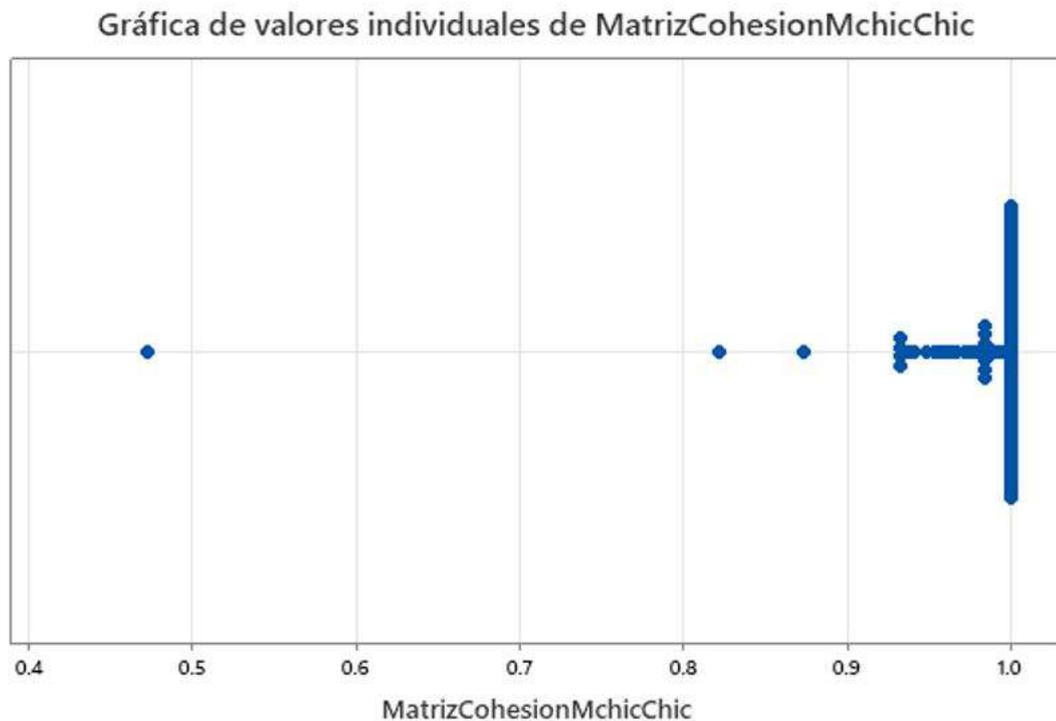


Ilustración 33-4: Gráfica de valores individuales de la Matriz Cohesión M-Chic y Chic.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

En la Ilustración 4-35 muestra el análisis para las Gráficas (M-Chic y R-Chic), cuya gráfica representa los valores individuales donde en el intervalo 0.40 a 0.50 se muestra un valor es el intervalo mas aislado de la gráfica ya que contiene una base menor a 15 variables, en el intervalo de 0.80 a 0.90 se muestra dos valores y en intervalo de 0.90 a 1 se encuentran el resto de valores y la mayoría se encuentra en 1, por lo cual podemos concluir que las bases evaluadas son simétricas.

Gráfica de valores individuales de GraficaRchicChic

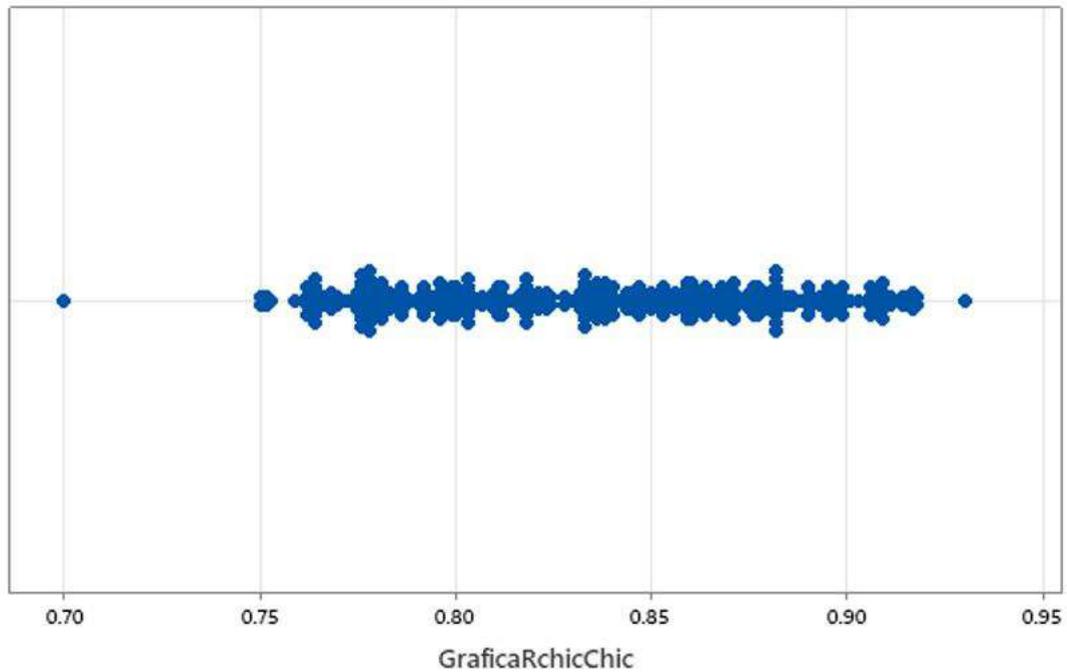


Ilustración 34-4: Gráfica de valores individuales de las Gráficas R-Chic y Chic.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

En la Ilustración 4-36 muestra el análisis para las Gráficas R-Chic y Chic cuya gráfica representa los valores individuales donde en el intervalo 0.70 a 0.75 se muestra un valor, en los intervalos 0.75 a 0.80, 0.80 a 0.85 y de 0.85 a 0.90 se muestra una distribución de datos similar por ende es donde se encuentran la mayor cantidad de datos evaluados y por último el intervalo 0.90 a 0.95 se muestra una pequeña parte de datos evaluados.

Gráfica de valores individuales de NodosMchicRchic

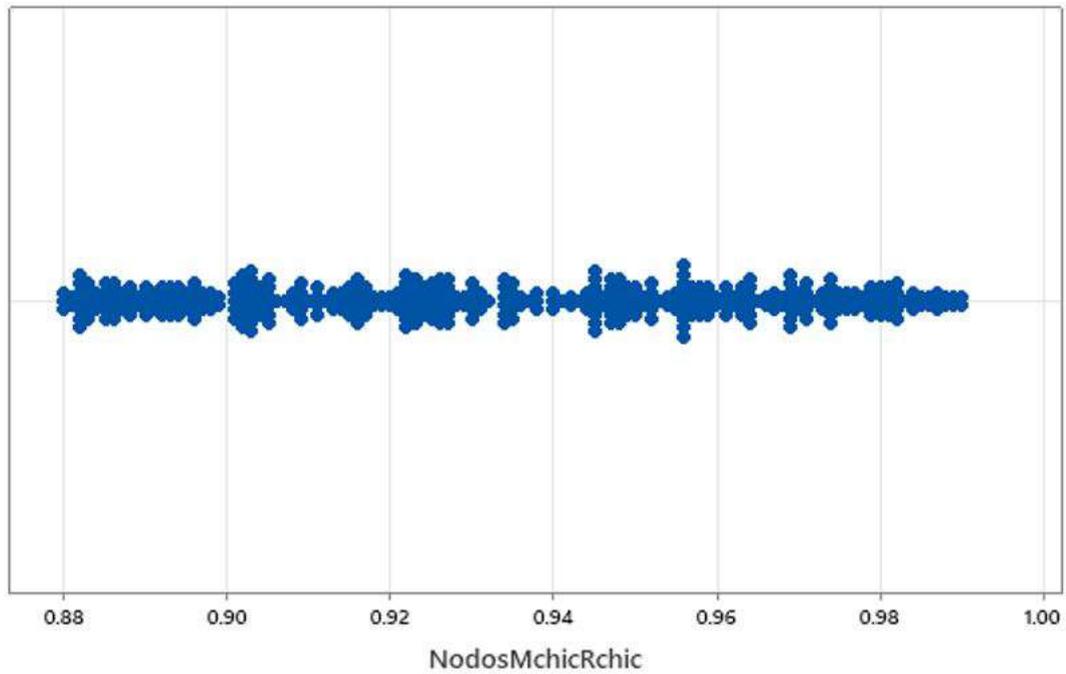


Ilustración 35-4: Gráfica de valores individuales de los Nodos M-Chic y R-Chic.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

En la Ilustración 4-37 muestra el análisis para los Nodos M-Chic y R-Chic cuya gráfica representa los valores individuales donde en el intervalo de 0.88 a 1 se muestra una distribución de datos similar por ende es donde se encuentran la mayor cantidad de datos evaluados.

Gráfica de valores individuales de NodosMchicChic

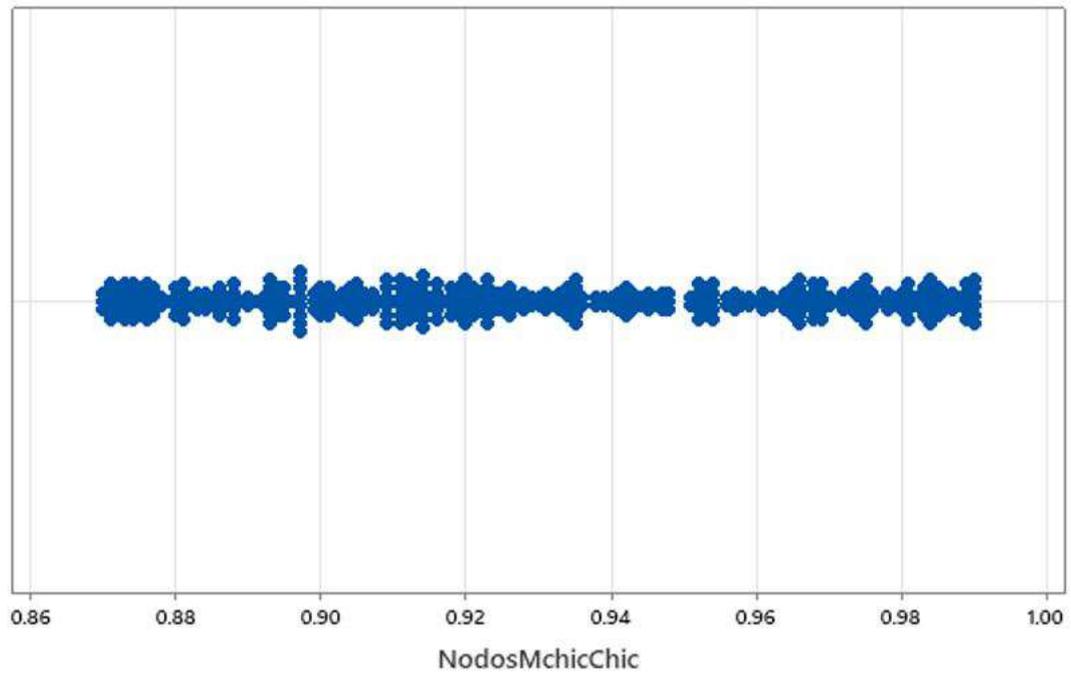


Ilustración 36-4: Gráfica de valores individuales de los Nodos M-Chic y Chic.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

En la Ilustración 4-38 muestra el análisis para los Nodos M-Chic y Chic cuya gráfica representa los valores individuales donde en el intervalo de 0.86 a 1 se muestra una distribución de datos similar por ende es donde se encuentran la mayor cantidad de datos evaluados.



Ilustración 37-4: Gráfica de caja de la Matriz de Cohesión M-Chic y R-Chic.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

En la Ilustración 4-39 muestra el análisis para la Matriz (M-Chic y R-Chic), cuya gráfica de caja se representa los intervalos 0.80 a 0.85 y 0.85 a 0.90 se muestra un valor, en el intervalo de 0.90 a 0.95 se muestra tres valores y por ende son bases con variables menores a 15 y en intervalo de 0.95 a 1 se encuentran el resto de valores y la mayoría se encuentra en 1, por lo cual podemos concluir que las bases evaluadas son simétricas.



Ilustración 38-4: Gráfica de caja de la Matriz de Cohesión M-Chic y Chic.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

En la Ilustración 4-40 muestra el análisis para la Matriz (M-Chic y Chic), cuya gráfica de caja donde se representa en el intervalo 0.40 a 0.50 se muestra un valor es el intervalo mas aislado de la gráfica ya que contiene una base menor a 15 variables, en el intervalo de 0.80 a 0.90 se muestra dos valores y en intervalo de 0.90 a 1 se encuentran el resto de valores y la mayoría se encuentra en 1, por lo cual podemos concluir que las bases evaluadas son simétricas.

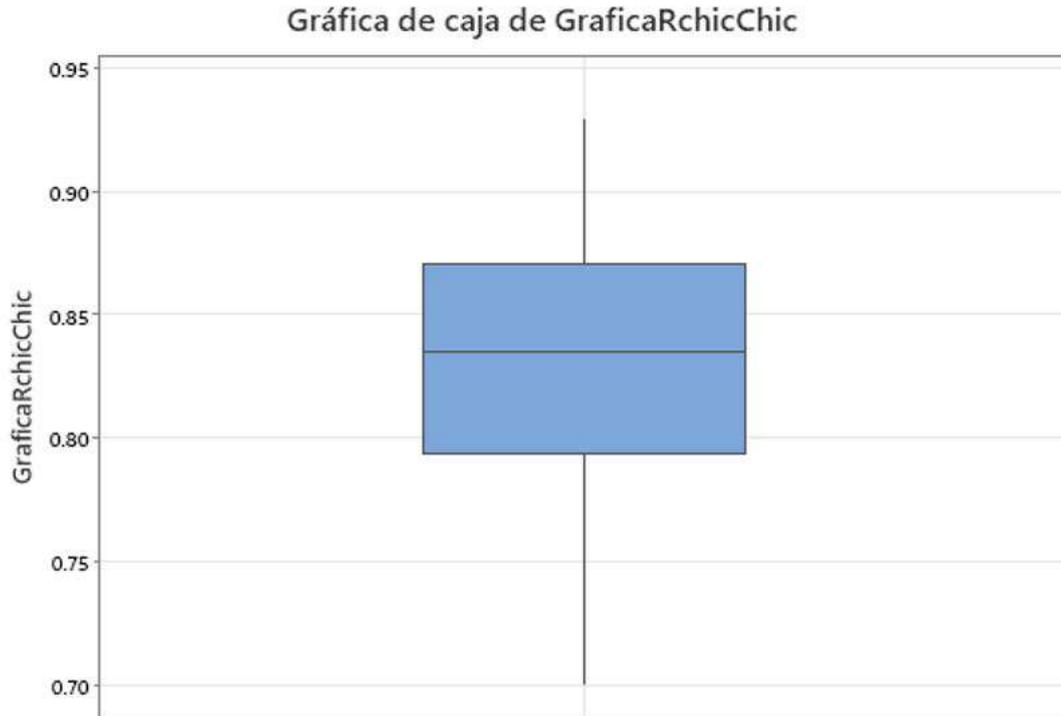


Ilustración 39-4: Gráfica de caja de los Gráficos R-Chic y Chic.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

En la Ilustración 4-41 muestra el análisis para las Gráficas R-Chic y Chic cuya gráfica de caja que representa en los intervalos 0.80 a 0.85 y de 0.85 a 0.90 se muestra una distribución de datos similar por ende es donde se encuentran la mayor cantidad de datos evaluados.

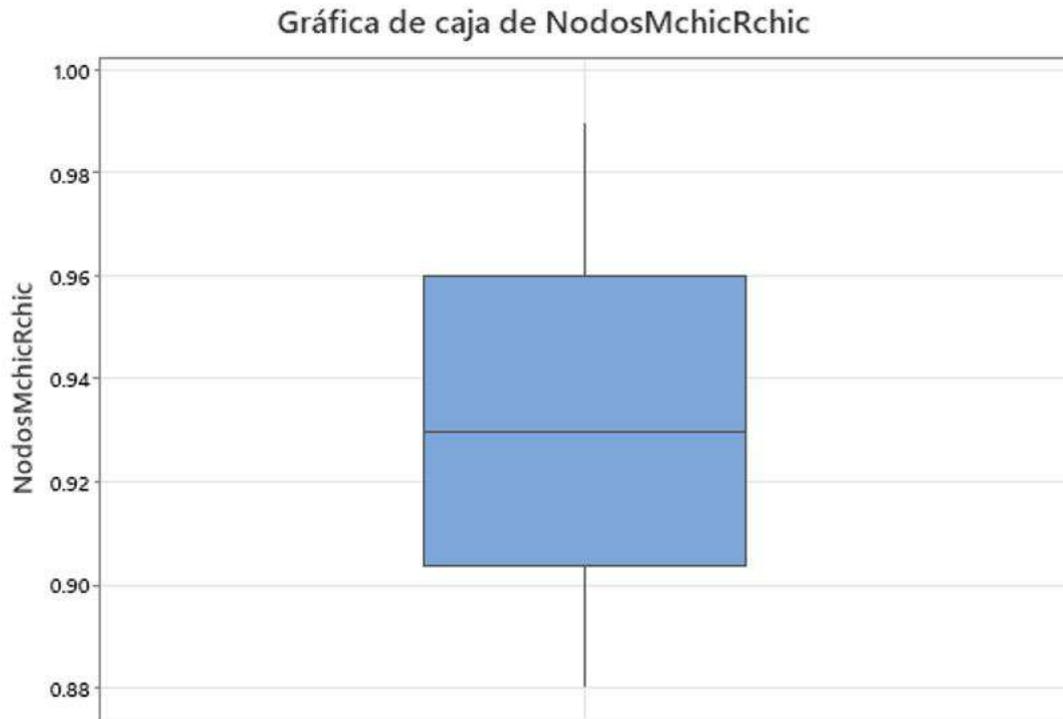


Ilustración 40-4: Gráfica de caja de los Nodos M-Chic y R-Chic.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

En la Ilustración 4-42 muestra el análisis para los Nodos M-Chic y R-Chic cuya gráfica de caja pues representa los valores individuales donde en el intervalo de 0.90 a 0.93 se muestra relativamente una distribución de datos similar a la del intervalo 0.93 a 0.96 por ende es donde se encuentran la mayor cantidad de datos evaluados.

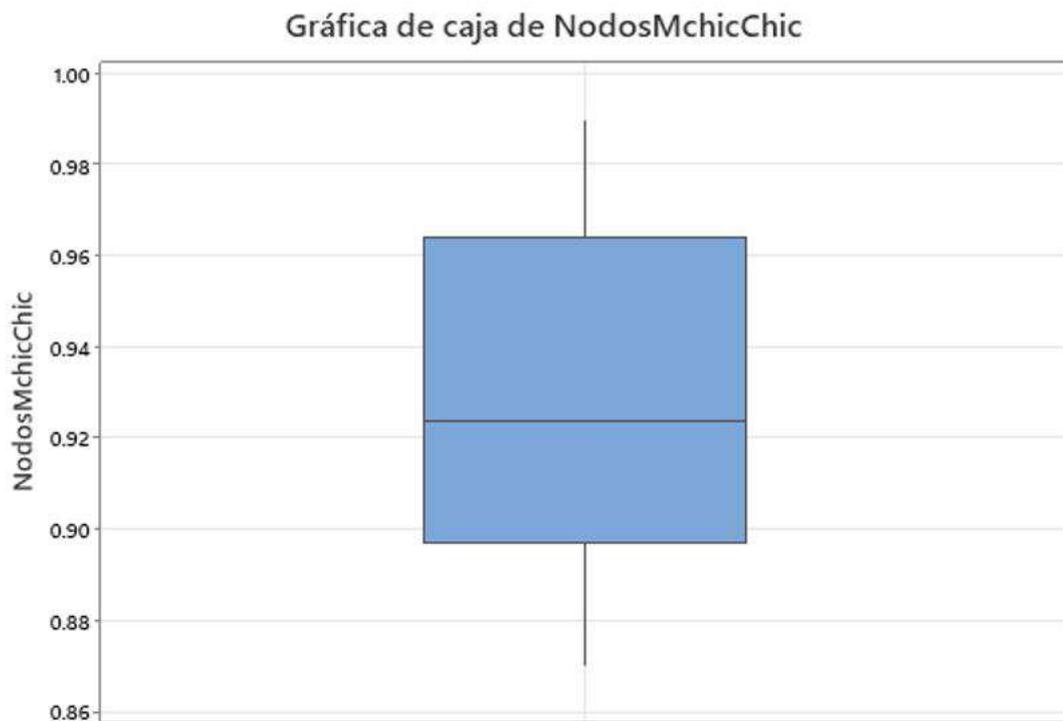


Ilustración 41-4: Gráfica de caja de los Nodos M-Chic y Chic.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

En la Ilustración 4-43 muestra el análisis para los Nodos M-Chic y Chic cuya gráfica de caja donde representa en el intervalo de 0.90 a 0.92 se muestra relativamente una distribución de datos menor a la del intervalo 0.92 a 0.96 por ende es donde se encuentran la mayor cantidad de datos evaluados en estos intervalos.

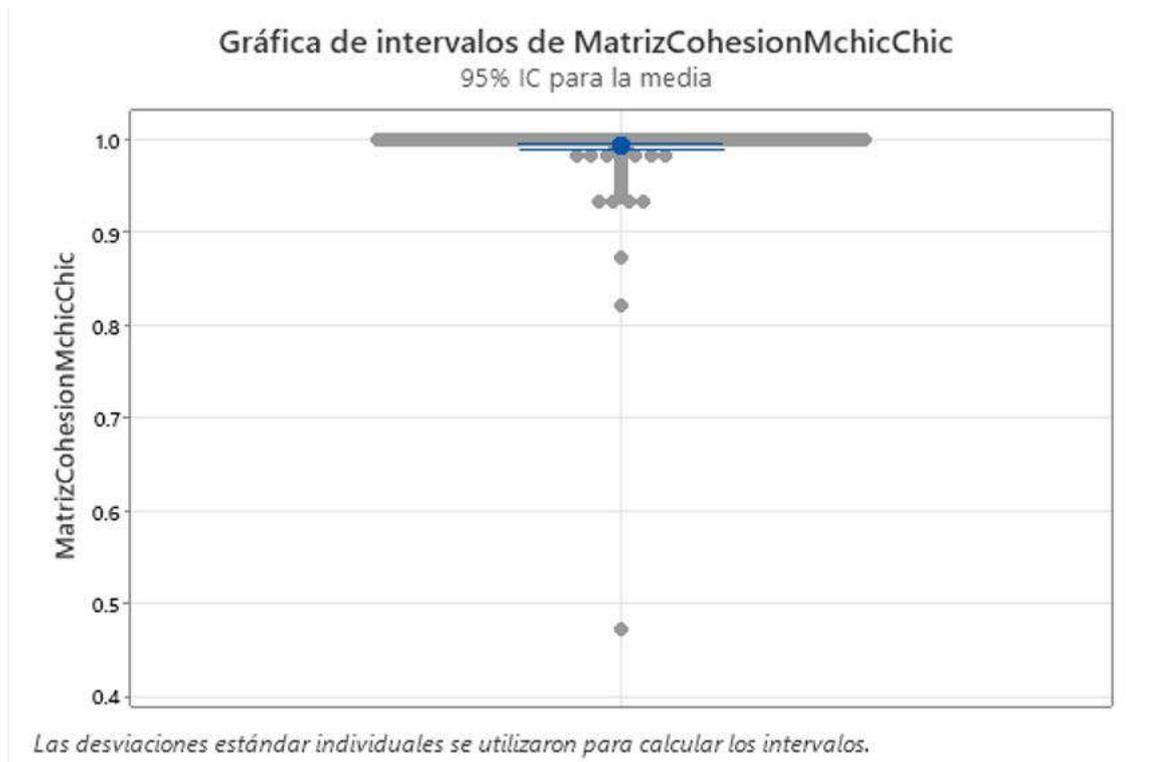
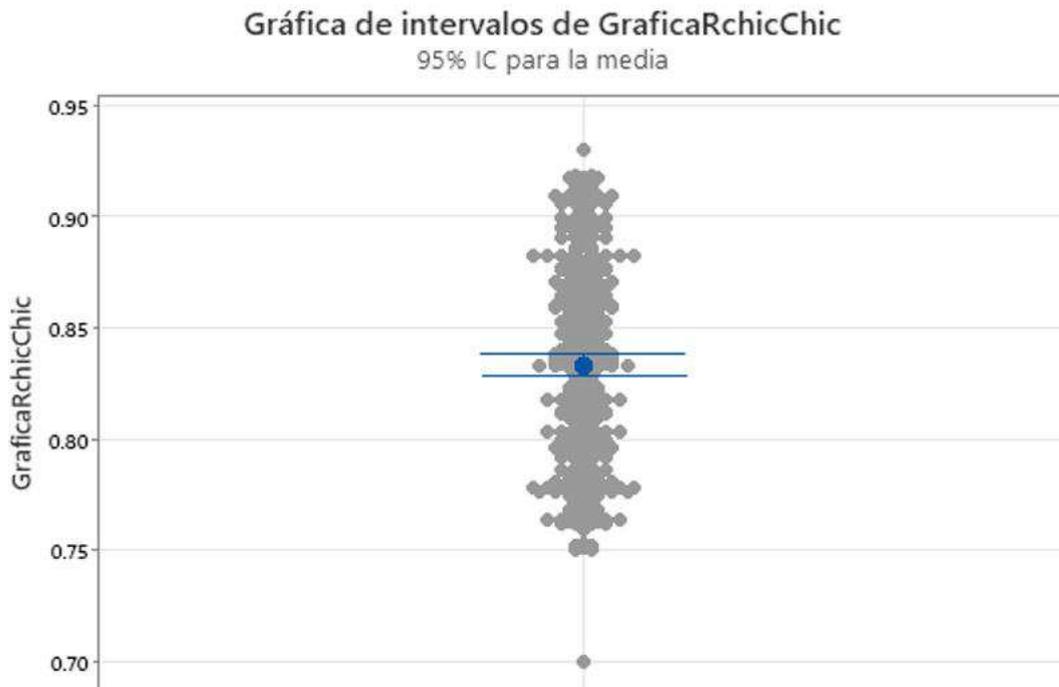


Ilustración 43-4: Gráfica de intervalos de Matriz Cohesión M-Chic y Chic.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

En la Ilustración 4-45 muestra el análisis para la Matriz Cohesión M-Chic y Chic cuya gráfica de intervalos nos muestra la cantidad de datos que se encuentran en la región de la media igual a 0.99315 siendo el 95 % de datos evaluados, pero existe un dato que esta demasiado aislado y esto se debe a una base de datos menor a 15 variables.



Las desviaciones estándar individuales se utilizaron para calcular los intervalos.

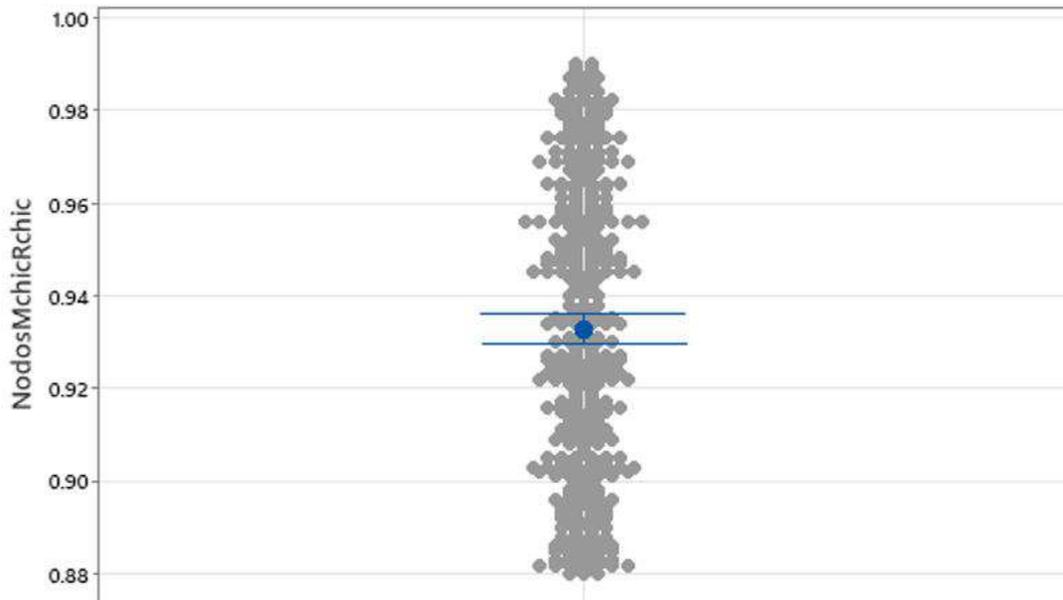
Ilustración 44-4: Gráfica de intervalos de los Gráficos ón R-Chic y Chic.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

En la Ilustración 4-46 muestra el análisis para los Gráficos ón R-Chic y Chic cuya gráfica de intervalos nos muestra la cantidad de datos que se encuentran en la región de la media igual a 0.83341 teniendo una distribución simétrica en el intervalo 0.75 a 0.95.

Gráfica de intervalos de NodosMchicRchic
95% IC para la media



Las desviaciones estándar individuales se utilizaron para calcular los intervalos.

Ilustración 45-4: Gráfica de intervalos de los Nodos M-Chic y R-Chic.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

En la Ilustración 4-47 muestra el análisis para los Nodos M-Chic y R-Chic cuya gráfica de intervalos nos muestra la cantidad de datos que se encuentran en la región de la media igual a 0.93284 teniendo una distribución simétrica en el intervalo 0.88 a 1.

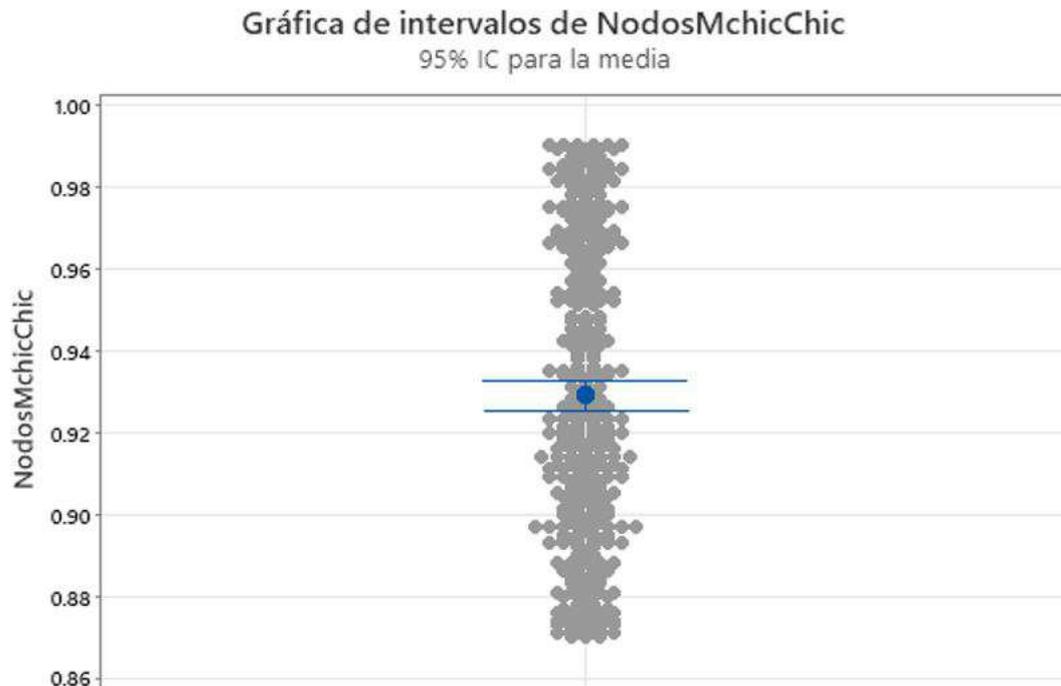


Ilustración 46-4: Gráfica de intervalos de los Nodos M-Chic y Chic.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

En la Ilustración 4-48 muestra el análisis para los Nodos M-Chic y Chic cuya gráfica de intervalos nos muestra la cantidad de datos que se encuentran en la región de la media igual a 0.92400 teniendo una distribución simétrica en el intervalo 0.86 a 1.

4.3. Comprobación de hipótesis

4.3.1. *T* de una muestra: *Matriz Cohesión (M-Chic y R-Chic)*, *Matriz Cohesión (M-Chic y Chic)*, *Gráfica (R-Chic y Chic)*, *Nodos (M-Chic y R-Chic)*, *Nodos (M-Chic y Chic)*

Tabla 4-4: Comparaciones Estadísticas descriptivas

Estadísticas descriptivas					
Muestra	N	Media	Desv.Est.	Error estándar de la media	Límite inferior de 95 % para μ
Matriz Cohesión (M-Chic y R-Chic)	383	0.997597	0.012948	0.000662	0.996507
Matriz Cohesión (M-Chic y Chic)	383	0.99315	0.03147	0.00161	0.99049
Nodos (M-Chic y R-Chic)	383	0.93284	0.03182	0.00163	0.93015

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

Prueba

Hipótesis nula $H_0 : \mu = 0.9$

Hipótesis alterna $H_1 : \mu \leq 0.9$

Tabla 5-4: Prueba

Muestra	Valor T	Valor p
Matriz Cohesión (M-Chic y R-Chic)	147.52	0.000
Matriz Cohesión (M-Chic y Chic)	57.93	0.000
Nodos (M-Chic y R-Chic)	20.20	0.000

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

El análisis de la Matriz Cohesión M-Chic y R-Chic nos muestra que al término de evaluar las 383 base tenemos los siguientes valores una media igual 0.997597, una desviación estándar igual al 0.012948, teniendo un error estándar en la media de 0.000662 y su límite inferior del 95 % para μ es igual a 0.996507 y el valor de la diferencia en relación con la variación en los datos de la muestra es 147.52. El análisis de la Matriz Cohesión M-Chic y Chic nos muestra que al término de evaluar las 383 base tenemos los siguientes valores una media igual 0.99315, una desviación estándar igual al 0.03147, teniendo un error estándar en la media de 0.000161 y su

límite inferior del 95 % para μ es igual a 0.99049 y el valor de la diferencia en relación con la variación en los datos de la muestra es 57.93. El análisis de los Nodos M-Chic y R-Chic nos muestra que al termino de evaluar las 383 base tenemos los siguientes valores una media igual 0.93284, una desviación estándar igual al 0.03182, teniendo un error estándar en la media de 0.00163 y su límite inferior del 95 % para μ es igual a 0.93015 y el valor de la diferencia en relación con la variación en los datos de la muestra es 20.20.

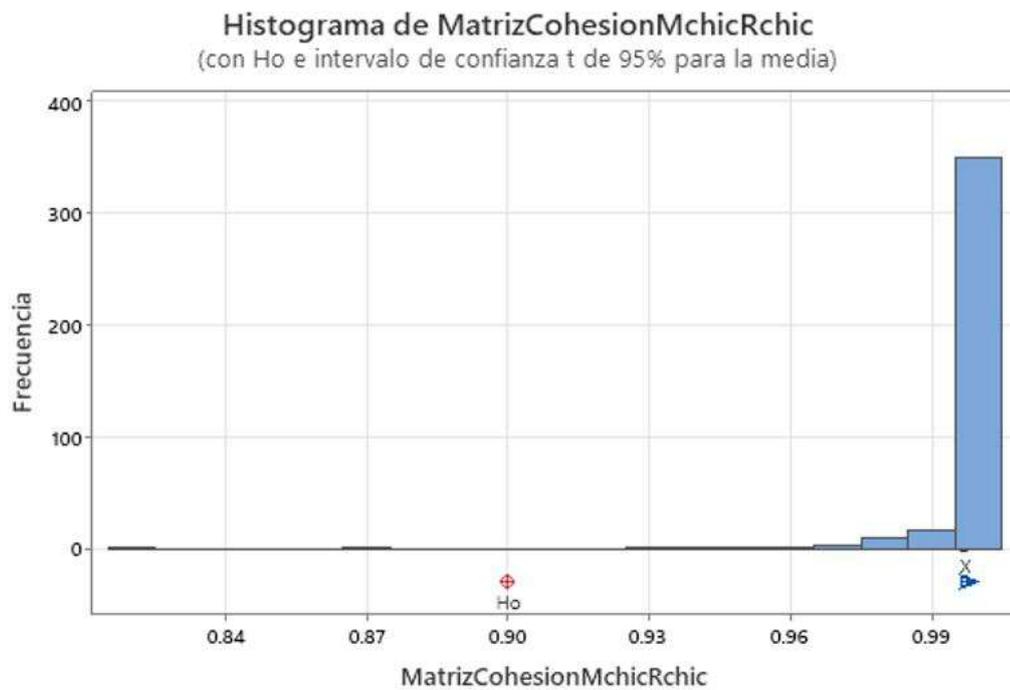


Ilustración 47-4: Histograma de la Matriz de Cohesión M-Chic y R-Chic.

Fuente: Elaboración propia.

Realizado por: Ilibay, Bryan, 2023.

En la Ilustración 4-49 muestra el análisis para la Matriz Cohesión (M-Chic y R-Chic), cuya gráfica es un histograma con H_0 y el intervalo de confianza del 95% para la media igual a 0.997597 donde se puede notar que existe medias y varianzas distintas, en el intervalo de 0.96 a 1 representa que la mayoría de datos aparentemente representan una distribución normal, por lo tanto al no existir datos aislados las bases evaluadas son simétricas pero su asimetría es negativa.

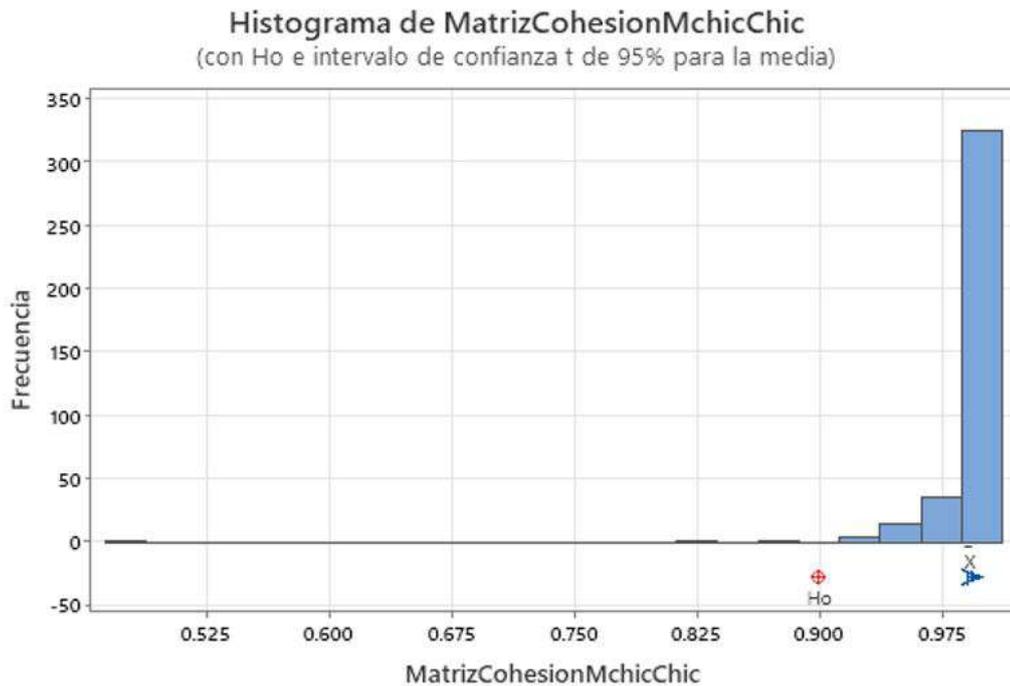


Ilustración 48-4: Histograma de la Matriz de Cohesión M-Chic y Chic.

Fuente: Elaboración propia.

Realizado por: Ibay, Bryan, 2023.

En la Ilustración 4-50 muestra el análisis para la Matriz Cohesión (M-Chic y Chic), cuya gráfica es un histograma con H_0 y el intervalo de confianza del 95 % para la media igual a 0.99315 donde se puede notar que existe medias y varianzas distintas, en el intervalo de 0.90 a 1 representa que la mayoría de datos aparentemente representan una distribución normal normal, por lo tanto al no existir datos aislados las bases evaluadas son simétricas pero su asimetría es negativa.

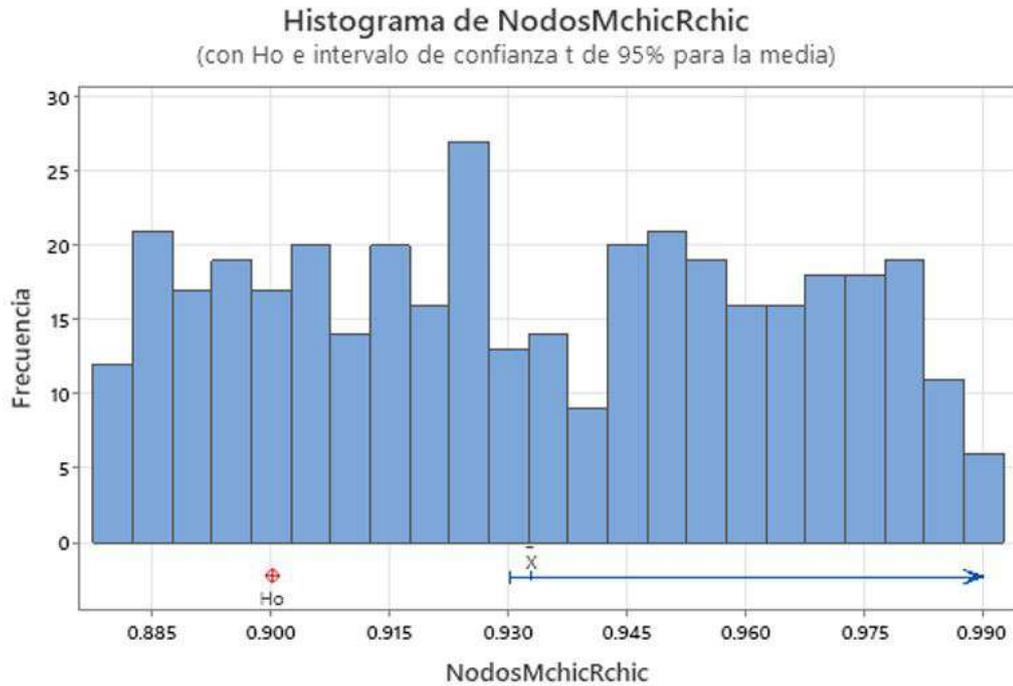


Ilustración 49-4: Histograma de Nodos M-Chic y R-Chic.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

En la Ilustración 4-51 muestra el análisis para las Nodos de R-Chic y Chic, cuya gráfica es un histograma con H_0 y el intervalo de confianza del 95 % para la media igual a 0.93284 donde se puede notar que existe medias y varianzas distintas, en el intervalo de 0.80 a 1 se encuentra la mayor cantidad de datos con distribuciones diferentes, pero se asemeja a una distribución normal y por lo tanto su asimetría es positiva.

Gráfica de valores individuales de MatrizCohesionMchicRchic
(con H_0 e intervalo de confianza t de 95% para la media)

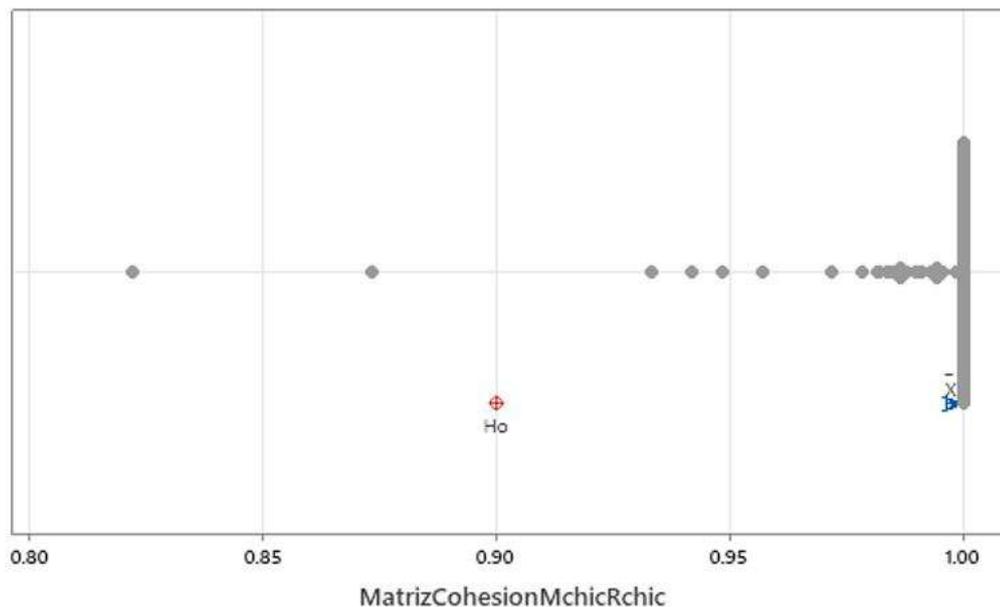


Ilustración 50-4: Valores individuales de la Matriz de Cohesión M-Chic y R-Chic.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

En la Ilustración 4-52 muestra el análisis para la Matriz (M-Chic y R-Chic), cuya gráfica representa los valores individuales con H_0 y el intervalo de confianza del 95% para la media igual a 0.997597 donde los intervalos 0.80 a 0.85 y 0.85 a 0.90 se muestra un valores aislados de la gráfica ya que contiene una base menor a 15 variables, en el intervalo de 0.80 a 0.90 se muestra dos valores y en intervalo de 0.90 a 1 se encuentran el resto de valores y la mayoría se encuentra en 1, por lo cual podemos concluir que las bases evaluadas son simétricas.

Gráfica de valores individuales de NodosMchicRchic
(con H_0 e intervalo de confianza t de 95% para la media)

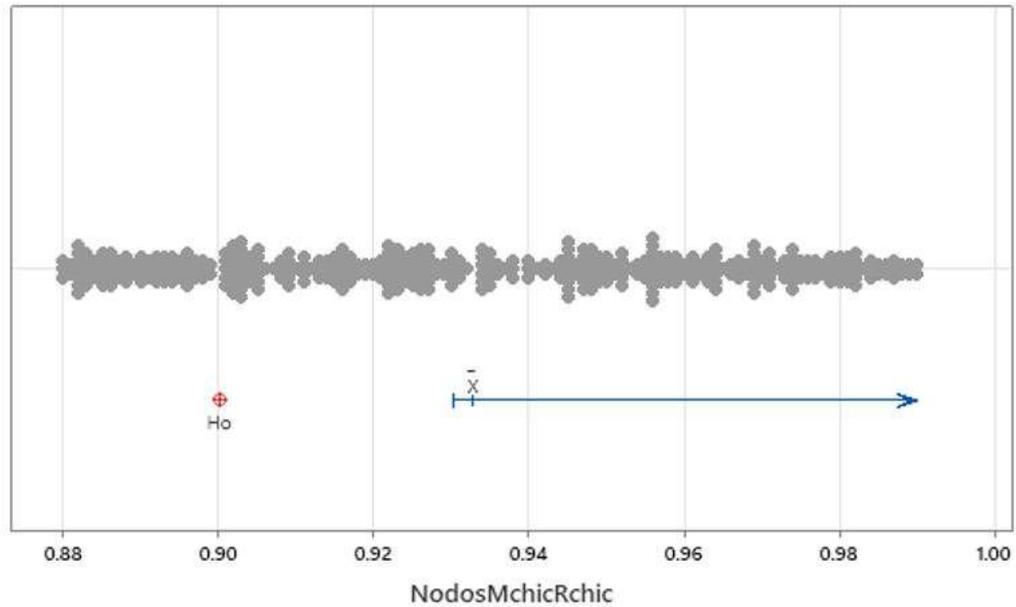


Ilustración 52-4: Gráfica de los valores individuales de los Nodos M-Chic y R-Chic.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

En la Ilustración 4-54 muestra el análisis para los Nodos M-Chic y R-Chic cuya gráfica representa los valores individuales con H_0 y el intervalo de confianza del 95 % para la media igual a 0.93284 donde en el intervalo de 0.88 a 1 se muestra una distribución de datos similar por ende es donde se encuentran la mayor cantidad de datos evaluados.

Gráfica de caja de MatrizCohesionMchicRchic
 (con H_0 e intervalo de confianza t de 95% para la media)

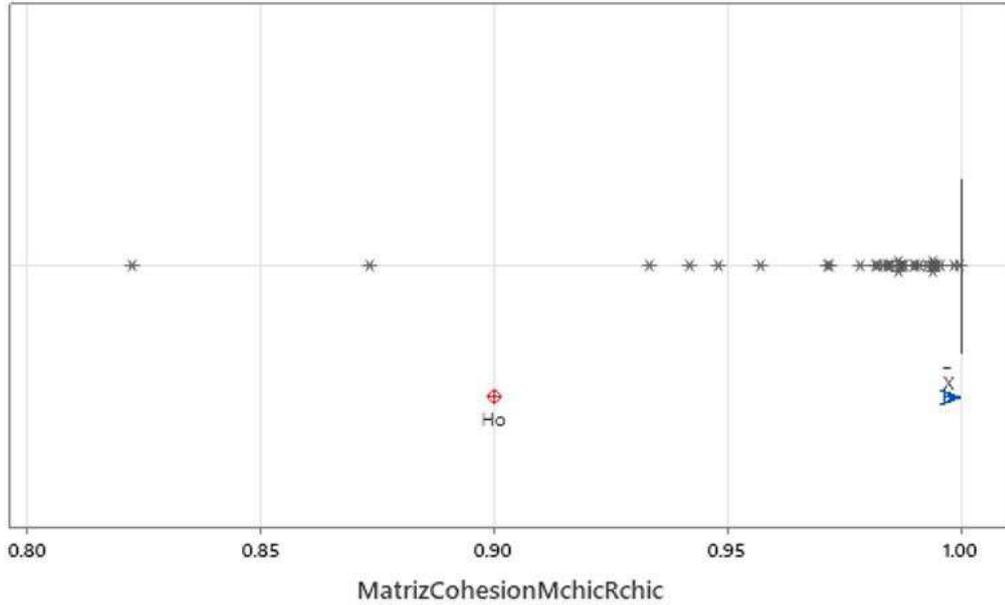


Ilustración 53-4: Gráfica de caja de la Matriz de Cohesión M-Chic y R-Chic

Fuente: Elaboración propia.

Realizado por: Ibay, Bryan, 2023.

En la Ilustración 4-55 muestra el análisis para la Matriz (M-Chic y R-Chic), cuya gráfica de caja con H_0 y el intervalo de confianza del 95% para la media igual a 0.997597 donde se representa los intervalos 0.80 a 0.85 y 0.85 a 0.90 se muestra un valor, en el intervalo de 0.90 a 0.95 se muestra tres valores y por ende son bases con variables menores a 15 y en intervalo de 0.95 a 1 se encuentran el resto de valores y la mayoría se encuentra en 1, por lo cual podemos concluir que las bases evaluadas son simétricas.

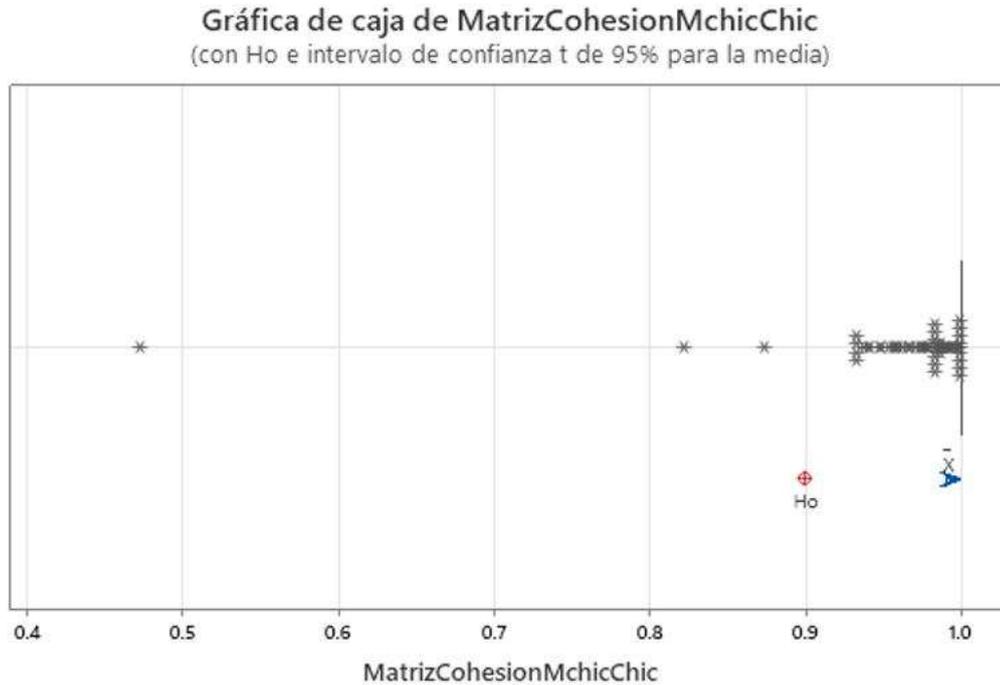


Ilustración 54-4: Gráfica de caja de la Matriz de Cohesión M-Chic y Chic.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

En la Ilustración 4-56 muestra el análisis para la Matriz (M-Chic y Chic), cuya gráfica de caja con H_0 y el intervalo de confianza del 95 % para la media igual a 0.99315 donde se representa en el intervalo 0.40 a 0.50 se muestra un valor es el intervalo mas aislado de la gráfica ya que contiene una base menor a 15 variables, en el intervalo de 0.80 a 0.90 se muestra dos valores y en intervalo de 0.90 a 1 se encuentran el resto de valores y la mayoría se encuentra en 1, por lo cual podemos concluir que las bases evaluadas son simétricas.

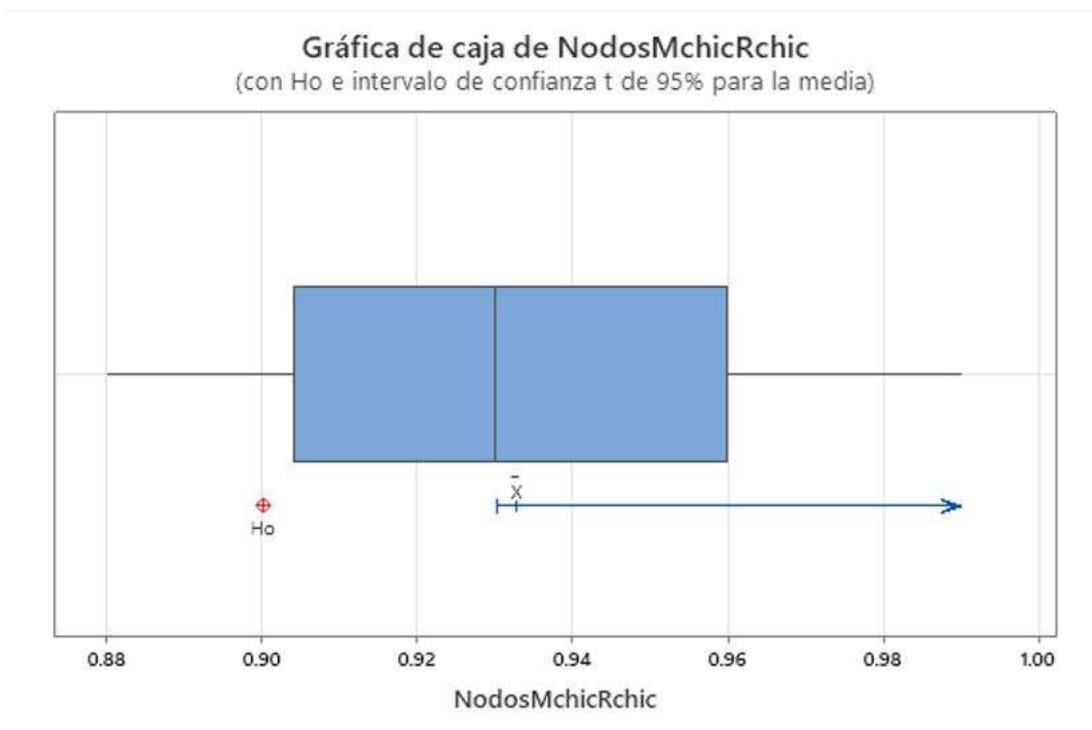


Ilustración 55-4: Gráfica de caja de los Nodos M-Chic y R-Chic.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

En la Ilustración 4-57 muestra el análisis para los Nodos M-Chic y R-Chic cuya gráfica de caja con H_0 y el intervalo de confianza del 95 % para la media igual a 0.93284 donde se representa los valores individuales donde en el intervalo de 0.90 a 0.93 se muestra relativamente una distribución de datos similar a la del intervalo 0.93 a 0.96 por ende es donde se encuentran la mayor cantidad de datos evaluados.

T de una muestra: Gráfica R Chic-Chic

Tabla 6-4: Comparaciones Estadísticas descriptivas: μ media entre las Gráficas de Rchic-Chic.

Estadísticas descriptivas					
Muestra	N	Media	Desv.Est.	Error estándar de la media	Límite inferior de 95 % para μ
Gráfica (R-Chic y Chic)	383	0.83341	0.04579	0.00234	0.82955

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

Prueba

Hipótesis nula $H_0 : \mu = 0.8$

Hipótesis alterna $H_1 : \mu \leq 0.8$

Valor T	Valor P
14,28	0

El análisis de las Gráficas R-Chic y Chic nos muestra que al termino de evaluar las 383 base tenemos los siguientes valores una media igual 0.83341, una desviación estándar igual al 0.04579, teniendo un error estándar en la media de 0.00234 y su límite inferior del 95 % para μ es igual a 0.82955 y el valor de la diferencia en relación con la variación en los datos de la muestra es igual a 28.46 y el Valor p es 0 ya que la prueba de hipótesis se realizo al 0.80.

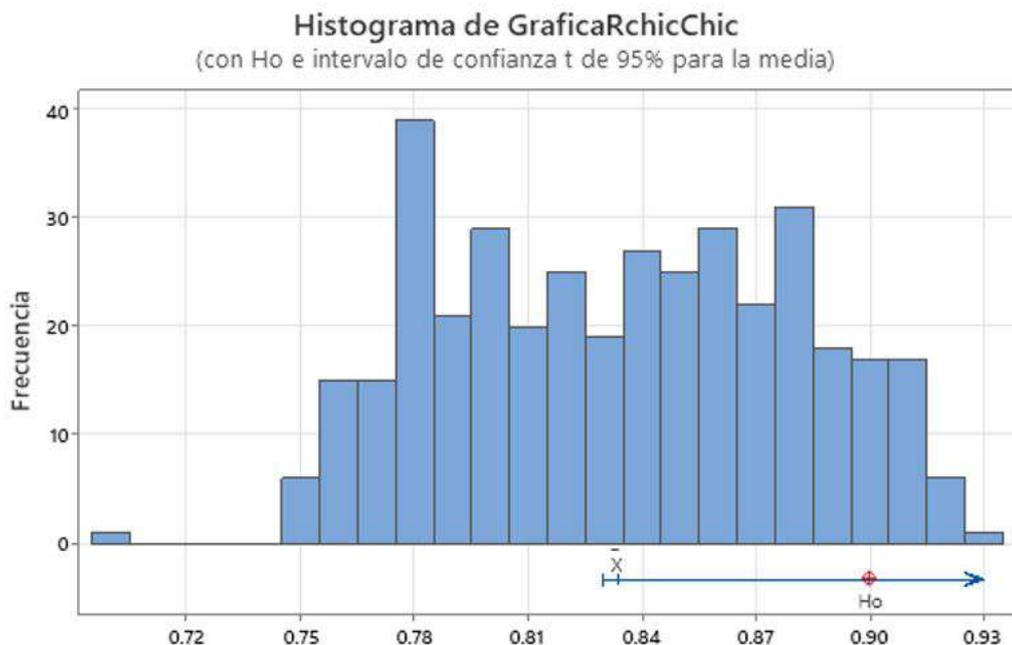


Ilustración 56-4: Histograma de las Gráficas R-Chic y Chic.

Fuente: Elaboración propia.

Realizado por: Ibay, Bryan, 2023.

En la Ilustración 4-58 muestra el análisis para las Gráficas M-Chic y R-Chic, cuya gráfica es un histograma con H_0 y el intervalo de confianza del 95 % para la media igual a 0.83341 donde se puede notar que existe medias y varianzas distintas, en el intervalo de 0.50 a 0.72 representa uno datos estan aislados los cuales son las bases con variables 0 a 15 ya que el programa trabaja con mayor presición cuando el número de variables es mayor a 20, en el intervalo de 0.74 a 1 se

encuentra la mayor cantidad de datos con distribuciones diferentes y por lo tanto su asimetría es negativa

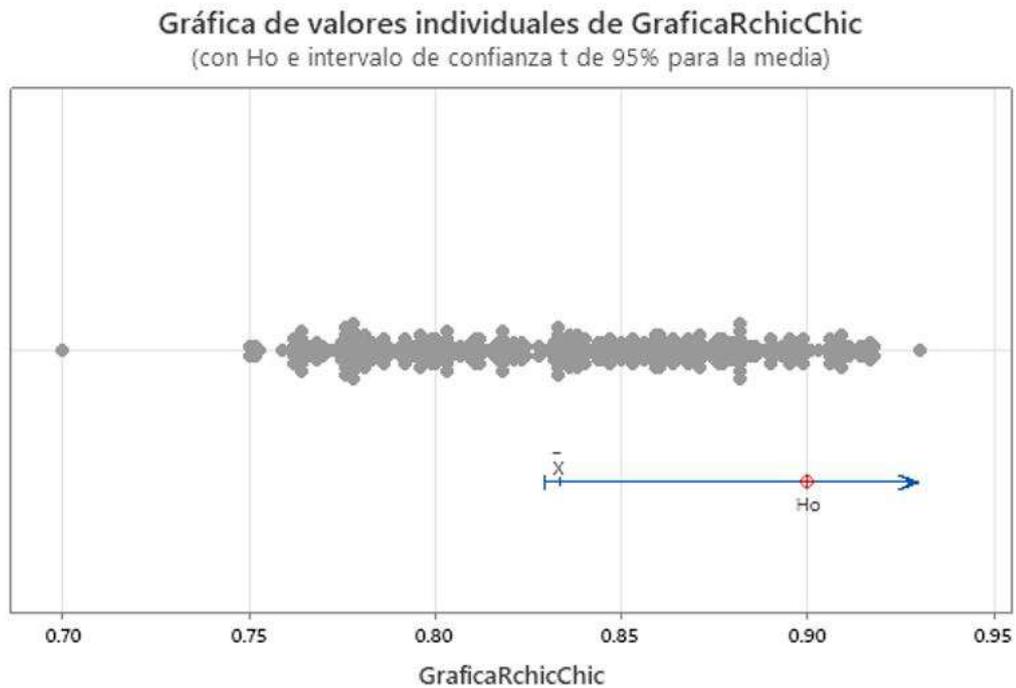


Ilustración 57-4: Gráfica de los valores individuales de la Gráfica R-Chic y Chic.

Fuente: Elaboración propia.

Realizado por: Ilbay, Bryan, 2023.

En la Ilustración 4-59 muestra el análisis para las Gráficas R-Chic y Chic cuya gráfica representa los valores individuales con H_0 y el intervalo de confianza del 95 % para la media igual a 0.83341 donde en el intervalo 0.70 a 0.75 se muestra un valor, en los intervalos 0.75 a 0.80, 0.80 a 0.85 y de 0.85 a 0.90 se muestra una distribución de datos similar por ende es donde se encuentran la mayor cantidad de datos evaluados y por último el intervalo 0.90 a 0.95 se muestra una pequeña parte de datos evaluados.

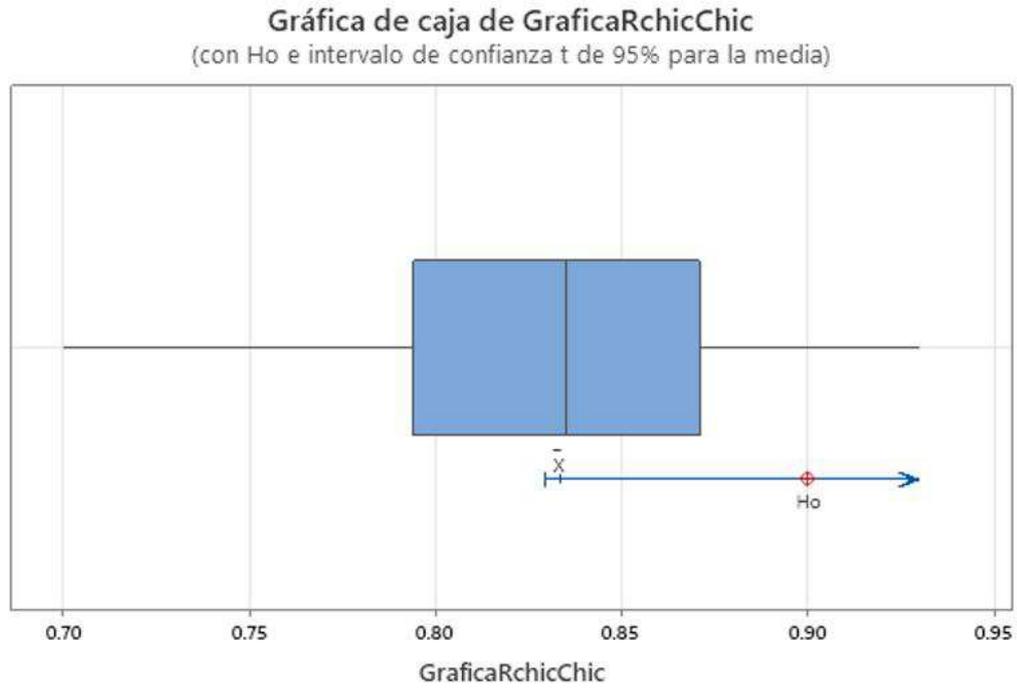


Ilustración 58-4: Gráfica de caja de las Gráficas R-Chic y Chic.

Fuente: Elaboración propia.

Realizado por: Ibay, Bryan, 2023.

En la Ilustración 4-60 muestra el análisis para las Gráficas R-Chic y Chic cuya gráfica de caja con H_0 y el intervalo de confianza del 95% para la media igual a 0.83341 donde se representa en los intervalos 0.80 a 0.85 y de 0.85 a 0.90 se muestra una distribución de datos similar por ende es donde se encuentran la mayor cantidad de datos evaluados.

CAPÍTULO V

5. CONCLUSIONES Y RECOMENDACIONES

5.1. CONCLUSIONES

En este trabajo de trabajo de Integración Curricular tuvo como objetivo fue programar la Cuasi-Implicación de Régis Gras entre variables binarias en Matlab. Para ello fue necesario estudiar las fórmulas matemáticas del Análisis Estadístico Implicativo enfocado más en la Cuasi-Implicación.

Las cuales nos permitieron la elaboración de pseudocódigo de la Cuasi Implicación de Régis Gras en Matlab denominado (M-Chic), al término del estudio de la similitud con los paquetes R-Chic y Chic los resultados obtenidos se puede notar una similitud en los datos.

Se tomó al inicio una hipótesis comprendida al 75% pero al término de la programación y del estudio de la similitud de las bases, se obtuvo una similitud el 90% en los resultados obtenidos con los paquetes anteriormente mencionados.

Esta programación permite aumentar el número de herramientas disponibles sobre el Análisis Estadístico Implicativo; quien utilice el pseudocódigo podrá notar los cálculos de una manera detalla sobre Cuasi-Implicación y con una interfaz mas accesible.

5.1.1. *Relación con objetivos, hipótesis y problema*

Con el propósito de generar un pseudocódigo en el software Matlab para la Cuasi-Implicación de Régis Gras que esta dentro del estudio del Análisis Estadístico Implicativo, tras el estudio y el empleo de las fórmulas matemáticas existentes, logrando la programación y su comparación con los paquetes existentes que son R-Chic en el software R y el paquete CHIC, para el análisis se utilizó un diseño pre-experimental del tipo un solo grupo aleatorio de la forma RGX01. Esta sección tiene el propósito de verificar el cumplimiento y la relación entre los objetivos específicos, objetivo general, hipótesis y problema de este trabajo.

5.1.2. *Sobre los objetivos específicos*

Se propusieron 5 objetivos específicos en el Capítulo.- Problema de Investigación Sección 1.2.2
Objetivos específicos.

- El primer objetivo específico fue “Analizar la Cuasi-Implicación de Régis Gras” que se cumplió en el Capítulo 2.- Marco Teórico, específicamente en la Sección 2.2 La Cuasi-Implicación de Régis Gras.
- El segundo objetivo específico fue “Determinar la formulación matemática de la Cuasi-Implicación de Régis Gras” que se cumplió en el Capítulo 2.- Marco Teórico, específicamente en la Sección 2.5 Formulación matemática.
- El tercer objetivo específico fue “Elaborar el pseudocódigo de la Cuasi-Implicación de Régis Gras” que se cumplió en el Capítulo 4.- Marco de Análisis e Interpretación de resultados, específicamente en la Sección 4.1.1 Elaborar el pseudocódigo de la Cuasi-Implicación de Gras.
- El cuarto objetivo específico fue “Programar la Cuasi-Implicación de Régis Gras” que se cumplió en el Capítulo 4.- Marco de Análisis e Interpretación de resultados, específicamente en la Sección 4.1.3 Programación en MATLAB. Algoritmo del Análisis Estadístico Implicativo, Sección 4.1.4 Programación en MATLAB. Algoritmo del Análisis Estadístico Implicativo para una base de 7 variables y 20 filas, Sección 4.1.5 Programación en MATLAB. Algoritmo del Análisis Estadístico Implicativo para una base de 7 variables y 200 filas y la Sección 4.1.6 Programación en MATLAB, Algoritmo del Análisis Estadístico Implicativo para una base aleatoria de m variables y n filas.
- El quinto objetivo específico fue “Estudiar la similaridad entre los resultados del programa Matlab (MChic) y RChic” que se cumplió en el Capítulo 4.- Marco de Análisis e Interpretación de resultados, específicamente en la Sección 4.2.1 Estadísticos descriptivos: Matriz Cohesión (M-Chic y R-Chic), Matriz Cohesión (M-Chic y Chic), Gráfica (R-Chic y Chic), Nodos (M-Chic y R-Chic), Nodos (M-Chic y Chic), la Sección 4.2.2 Gráfica de intervalos de Matriz Cohesión (M-Chic y R-Chic), Matriz Cohesión (M-Chic y Chic), Gráfica (R-Chic y Chic), Nodos (M-Chic y R-Chic), Nodos (M-Chic y Chic) y la Sección 4.3.1 T de una muestra: Matriz Cohesión (M-Chic y R-Chic), Matriz Cohesión (M-Chic y Chic), Gráfica (R-Chic y Chic), Nodos (M-Chic y R-Chic), Nodos (M-Chic y Chic).

Sobre el objetivo general El objetivo general planteado fue “Programar en Matlab la Cuasi Implicación de Régis Gras entre variables binarias y validarlas mediante la similaridad”. La programación de la Cuasi Implicación y su validación se encuentran en las secciones 4.1.6, 4.2.1, 4.2.2 y 4.3.1.

- La Sección 4.1.6 Programación en MATLAB, Algoritmo del Análisis Estadístico Implicativo para una base aleatoria de m variables y n filas, es la programación general de la Cuasi Implicación.

- La Sección 4.2.1 Estadísticos descriptivos: Matriz Cohesión (M-Chic y R-Chic), Matriz Cohesión (M-Chic y Chic), Gráfica (R-Chic y Chic), Nodos (M-Chic y R-Chic), Nodos (M-Chic y Chic); siendo los análisis de las bases obtenidas en M-Chic, R-Chic y Chic.
- La Sección 4.2.2 Gráfica de intervalos de Matriz Cohesión (M-Chic y R-Chic), Matriz Cohesión (M-Chic y Chic), Gráfica (R-Chic y Chic), Nodos (M-Chic y R-Chic), Nodos (M-Chic y Chic);siendo los análisis de las bases obtenidas en M-Chic, R-Chic y Chic.
- la Sección 4.3.1 T de una muestra: Matriz Cohesión (M-Chic y R-Chic), Matriz Cohesión (M-Chic y Chic), Gráfica (R-Chic y Chic), Nodos (M-Chic y R-Chic), Nodos (M-Chic y Chic); siendo los análisis de las bases obtenidas en M-Chic, R-Chic y Chic.

5.1.3. Sobre la hipótesis

La hipótesis planteada fue “Programar en Matlab la Cuasi Implicación de Régis Gras entre variables binarias y validarlas mediante la similaridad al 75%”, se puede comprobar que la similaridad de los resultados sobrepasan el 75% y comprobando que el programa M-Chic responde correctamente en las secciones:

- La Sección 4.1.6 Programación en MATLAB, Algoritmo del Análisis Estadístico Implicativo para una base aleatoria de m variables y n filas.
- La Sección 4.2.1 Estadísticos descriptivos: Matriz Cohesión (M-Chic y R-Chic), Matriz Cohesión (M-Chic y Chic), Gráfica (R-Chic y Chic), Nodos (M-Chic y R-Chic), Nodos (M-Chic y Chic).
- La Sección 4.2.1 Estadísticos descriptivos: Matriz Cohesión (M-Chic y R-Chic), Matriz Cohesión (M-Chic y Chic), Gráfica (R-Chic y Chic), Nodos (M-Chic y R-Chic), Nodos (M-Chic y Chic).
- La Sección 4.2.2 Gráfica de intervalos de Matriz Cohesión (M-Chic y R-Chic), Matriz Cohesión (M-Chic y Chic), Gráfica (R-Chic y Chic), Nodos (M-Chic y R-Chic), Nodos (M-Chic y Chic).
- La Sección 4.3.1 T de una muestra: Matriz Cohesión (M-Chic y R-Chic), Matriz Cohesión (M-Chic y Chic), Gráfica (R-Chic y Chic), Nodos (M-Chic y R-Chic), Nodos (M-Chic y Chic).
- La Sección 4.3.1 T de una muestra: Matriz Cohesión (M-Chic y R-Chic), Matriz Cohesión (M-Chic y Chic), Gráfica (R-Chic y Chic), Nodos (M-Chic y R-Chic), Nodos (M-Chic y Chic).

5.1.4. Sobre el problema

El problema planteado fue “Se puede programar la Cuasi Implicación de Régis Gras entre variables binarias en Matlab”, se comprobó la programación de la Cuasi Implicación en las secciones:

- La Sección 4.1.6 Programación en MATLAB, Algoritmo del Análisis Estadístico Implicativo para una base aleatoria de m variables y n filas.
- La Sección 4.2.1 Estadísticos descriptivos: Matriz Cohesión (M-Chic y R-Chic), Matriz Cohesión (M-Chic y Chic), Gráfica (R-Chic y Chic), Nodos (M-Chic y R-Chic), Nodos (M-Chic y Chic).
- La Sección 4.2.1 Estadísticos descriptivos: Matriz Cohesión (M-Chic y R-Chic), Matriz Cohesión (M-Chic y Chic), Gráfica (R-Chic y Chic), Nodos (M-Chic y R-Chic), Nodos (M-Chic y Chic).
- La Sección 4.2.2 Gráfica de intervalos de Matriz Cohesión (M-Chic y R-Chic), Matriz Cohesión (M-Chic y Chic), Gráfica (R-Chic y Chic), Nodos (M-Chic y R-Chic), Nodos (M-Chic y Chic).
- la Sección 4.3.1 T de una muestra: Matriz Cohesión (M-Chic y R-Chic), Matriz Cohesión (M-Chic y Chic), Gráfica (R-Chic y Chic), Nodos (M-Chic y R-Chic), Nodos (M-Chic y Chic).

Pero al no existir un cluster que gráfique la Matriz de Cohesión en Matlab no se pudo comprobar la similaridad de Gráficas con R-Chic.

5.2. RECOMENDACIONES

Con base de los resultados obtenidos en el presente proyecto de investigación, a continuación se enumeraran algunas recomendaciones.

- Trabajar en la mejora de la programación de la Cuasi-Implicación de Régis Gras, hecha para esta tesis, con el fin de que al momento de que se muestran los datos de las matrices de los diferentes niveles de cohesión, los decimales tengan un porcentaje de similaridad aún más alto al momento de comparar con el programa original CHIC.
- Buscar una manera de implementar a la programación de la Cuasi-Implicación de Régis Gras el dendograma (Árbol jerárquico cohesivo), el cual debe alimentarse con los datos de los

nodos significativos.

- Se sabe que la programación de la Cuasi-Implicación de Régis Gras tiene un alto porcentaje de similaridad con el programa CHIC, los dos trabajan con la ley de Poisson y el tipo de implicación según la teoría clásica, pero se recomienda al usuario saber exactamente con qué distribución va a trabajar para que obtenga mejores resultados ya sea en M-Chic, en R-Chic o Chic.

BIBLIOGRAFÍA

- BARRAGAN-PAZMIÑO, B.M.; PAZMIÑO-MAJI, R.A.** "Literatura Científica sobre Análisis Estadístico Implicativo: Un mapeo sistemático de la década que transcurre". *CIENCIA DIGITAL* [en línea], 2018, vol. 2. [Consulta: 23 abril 2022]. ISSN 2602-8085. Disponible en: <http://cienciadigital.org/revistacienciadigital2/index.php/CienciaDigital/issue/view/13>
- COUTRIER, R., et al.** "Statistical implicative analysis for educational data sets: 2 analysis with RCHIC" [en línea], 2015. [consulta: 1 diciembre 2022]. ISBN 978-84-608-3627-8. Disponible en: <https://gedos.usal.es/handle/10366/127757>
- COUTURIER, R.; PAZMIÑO, R.** "Use of Statistical Implicative Analysis in Complement of Item Analysis". *International Journal of Information and Education Technology*, vol. 6, no. 1 (2016), ISSN 2010-3689.
- GRAS, R., et al.** "El Análisis Estadístico Implicativo (ASI) en respuesta a problemas que le dieron origen". 2009 pp. 3-50.
- GREGORI HUERTA, P., et al.** "On the probability distribution of the classical Gras implication index between two binary random variables". 2014.
- IURATO, G.** "The Implicative Statistical Analysis: An Interdisciplinary Paradigm". [en línea], 2012. [consulta: 3 noviembre 2021]. Disponible en: <https://hal.archives-ouvertes.fr/hal-00750049>
- LERMAN, CI.** "Foundations and Methods in Combinatorial and Statistical Data Analysis and Clustering". [en línea], 2016. [consulta: 17 enero 2022]. Disponible en: https://books.google.com/books/about/Foundations_and_Methods_in_Combinatorial.html?hl=es&id=rJzWCwAAQBAJ
- LÓPEZ-ROLDÁN, P.; FACHELLI, S.** 2021. "Análisis de clasificación". [en línea], 2016. [consulta: 9 noviembre 2021]. Disponible en: <https://www.mdx.cat/handle/10503/113211>
- MARTÍNEZ GÓMEZ, M.; MARÍ BENLLOCH, M.D.** "Distribución Binomial". 2010.
- MEDEROS, Y.G., et al.** 2015. "Análisis estadístico implicativo en la identificación de factores de riesgo en pacientes con cáncer de pulmón". *MediSan* [en línea], 2015, vol. 19(08), [Consulta: 7 noviembre 2021]. ISSN 1029-3019. Disponible en: <https://www.medigraphic.com/cgi-bin/new/resumen.cgi?IDARTICULO=60754>
- ORUS, P., et al.** "Teoría y aplicaciones del Análisis Estadístico Implicativo: primera aproximación en lengua hispana". 2009.

PAZMIÑO-MAJI, R.A., et al. “El Analisis Estadistico Implicativo y la mejora del aprendizaje en el marco de las Analiticas de Aprendizaje: Un Mapeo Sistemático de Literatura”. 2018.

PEREZ, M.G., et al. “Cuasi-implicacion estadística y determinación automática de clases de equivalencia en imágenes de resonancia magnética de cerebro”. *Revista Politecnica* [en línea], 2014, vol. 34(1), [Consulta: 15 diciembre 2021]. ISSN 2477-8990. Disponible en: https://revistapolitecnica.epn.edu.ec/ojs2/index.php/revista_politecnica2/article/view/260

SAGARO DEL CAMPO, N.M., et al. . “¿Por que emplear el analisis estadístico implicativo en los estudios de causalidad en salud?” *Revista Cubana de Informatica Medica* [en línea], 2019, vol. 11(1), [Consulta: 6 noviembre 2021]. ISSN 1684-1859. Disponible en: http://scielo.sld.cu/scielo.php?script=sci_abstract&pid=S1684-18592019000100088&lng=es&nrm=iso&tlng=en

VALLS PLA, X., et al. “Diseño de un paquete R para el analisis estadístico implicativo”. [en línea], 2014. [consulta: 15 diciembre 2021]. Disponible en: <http://repositori.uji.es/xmlui/handle/10234/107441>

ZAMORA, L., et al. “Conceptos fundamentales del Analisis Estadístico Implicativo (ASI) y su soporte computacional CHIC”. *Contribuciones al ASI*. [en línea], 2009, vol. 4. [Consulta: 1 noviembre 2021]. Disponible en: http://repositori.uji.es/xmlui/bitstream/handle/10234/125568/asi4esp_v18_libro.pdf





epoch

**Dirección de Bibliotecas y
Recursos del Aprendizaje**

**UNIDAD DE PROCESOS TÉCNICOS Y ANÁLISIS BIBLIOGRÁFICO Y
DOCUMENTAL**

REVISIÓN DE NORMAS TÉCNICAS, RESUMEN Y BIBLIOGRAFÍA

Fecha de entrega: 23 / 05 / 2023

INFORMACIÓN DEL AUTOR/A (S)
Nombres – Apellidos: Bryan Francisco Ilbay Chavez
INFORMACIÓN INSTITUCIONAL
Facultad: Ciencias
Carrera: Matemática
Título a optar: Matemático
f. Analista de Biblioteca responsable: Ing. Rafael Inty Salto Hidalgo

0869-DBRA-UTP-2023

