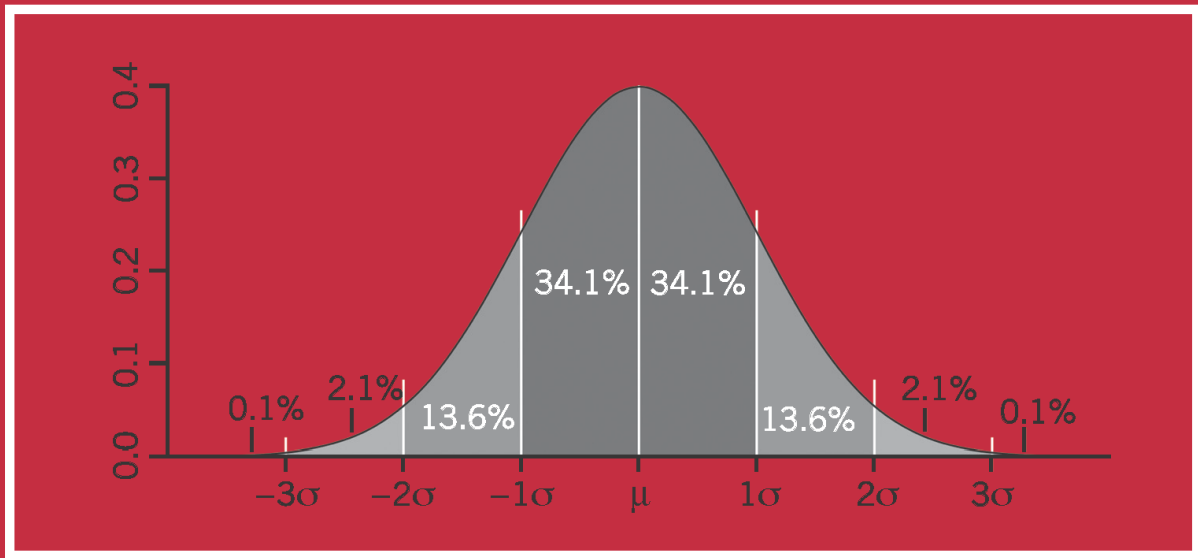


Editores:

Fuensanta Hernández Pina
Javier J. Maquilón Sánchez
José D. Cuesta Sáez de Tejada
Tomás Izquierdo Rus

Investigación y análisis de datos para la realización de TFG, TFM y tesis doctoral



$$\sqrt{2^{1+2 \cdot 3} \cdot ((1-2)+3)} \\ 5(2+2) \quad 101_2 = 5_{10}$$

**INVESTIGACIÓN Y ANÁLISIS DE DATOS
PARA LA REALIZACIÓN DE
TFG, TFM Y TESIS DOCTORAL**

INVESTIGACIÓN Y ANÁLISIS DE DATOS PARA LA REALIZACIÓN DE TFG, TFM Y TESIS DOCTORAL

Editores:

Fuensanta Hernández Pina

Javier J. Maquilón Sánchez

José D. Cuesta Sáez de Tejada

Tomás Izquierdo Rus

1ª Edición, 2015

Reservados todos los derechos. De acuerdo con la legislación vigente, y bajo las sanciones en ella previstas, queda totalmente prohibida la reproducción y/o transmisión parcial o total de este libro, por procedimientos mecánicos o electrónicos, incluyendo fotocopia, grabación magnética, óptica o cualesquiera otros procedimientos que la técnica permita o pueda permitir en el futuro, sin la expresa autorización por escrito de los propietarios del *copyright*.

© Los autores:

Fuensanta Hernández Pina

Javier J. Maquilón Sánchez

José D. Cuesta Sáez de Tejada

Tomás Izquierdo Rus

Fuensanta Monroy Hernández

ISBN: 978-84-942655-6-3

Depósito Legal: MU 481-2015

Diseño e impresión: Compobell, S.L. (Murcia)

Edición a cargo de Compobell, S.L. (Murcia)

Impreso en España - Printed in Spain

ÍNDICE

Introducción	9
<i>Fuensanta Hernández Pina</i>	
Capítulo 1. EL PROCESO DE INVESTIGACIÓN CIENTÍFICA	11
<i>Fuensanta Hernández Pina y Javier J. Maquilón Sánchez</i>	
Capítulo 2. INSTALACIÓN DEL PROGRAMA MYSTAT EN EL SISTEMA OPERATIVO MICROSOFT WINDOWS.....	35
<i>Tomás Izquierdo Rus y Fuensanta Monroy Hernández</i>	
Capítulo 3. MANEJO DEL EDITOR DE DATOS DEL PROGRAMA MYSTAT 12	41
<i>José D. Cuesta Sáez de Tejada y Tomás Izquierdo Rus</i>	
Capítulo 4. REPRESENTACIONES GRÁFICAS DE DATOS CUANTITATIVOS	65
<i>José D. Cuesta Sáez de Tejada y Fuensanta Monroy Hernández</i>	
Capítulo 5. ESTADÍSTICA DESCRIPTIVA.....	85
<i>Fuensanta Hernández Pina y Tomás Izquierdo Rus</i>	
Capítulo 6. SIGNIFICACIÓN DE LA DIFERENCIA ENTRE DOS MUESTRAS	105
<i>Javier J. Maquilón Sánchez y José D. Cuesta Sáez de Tejada</i>	
Capítulo 7. ANÁLISIS DE LA VARIANZA (ANOVA).....	117
<i>Javier J. Maquilón Sánchez y Tomás Izquierdo Rus</i>	
Capítulo 8. CORRELACIÓN Y REGRESIÓN.....	131
<i>Fuensanta Monroy Hernández y Javier J. Maquilón Sánchez</i>	

Capítulo 9. VALIDEZ Y FIABILIDAD DE LOS INSTRUMENTOS DE MEDIDA	157
<i>Fuensanta Hernández Pina y José D. Cuesta Sáez de Tejada</i>	
Capítulo 10. ESTADÍSTICA NO PARAMÉTRICA	173
<i>José D. Cuesta Sáez de Tejada y Javier J. Maquilón Sánchez</i>	
Capítulo 11. ANÁLISIS DE DATOS CUALITATIVOS CON ATLAS.TI	197
<i>Tomás Izquierdo Rus y Gustavo Montealegre Lynett</i>	
Capítulo 12. INFORME ESCRITO DE LA INVESTIGACIÓN	209
<i>Fuensanta Monroy Hernández y Fuensanta Hernández Pina</i>	
Capítulo 13. PRESENTACIÓN Y DEFENSA ORAL DE LA INVESTIGACIÓN	229
<i>Fuensanta Monroy Hernández y Fuensanta Hernández Pina</i>	

INTRODUCCIÓN

El propósito de este libro es introducir y apoyar a los estudiantes de Ciencias Sociales y de la Educación en los procedimientos de la metodología de la investigación y los contenidos estadísticos necesarios para comprender las investigaciones llevadas a cabo dentro de su campo de estudio, así como la planificación y ejecución de sus propias investigaciones.

Las competencias para el desarrollo de un trabajo investigación se han convertido en tema esencial en la formación de los estudiantes de grado, máster y doctorado. Estas competencias incluyen la adquisición del conocimiento científico y de la terminología relacionada con la investigación, esto es, la competencia del saber. Asimismo, los estudiantes deben adquirir la competencia del saber hacer, que se traduce en la adquisición de habilidades y procedimientos para la ejecución de las propias investigaciones y la elaboración de informes de investigación aplicados a los estudios cualitativos y cuantitativos. Y por último, saber comunicar por escrito y oralmente los informes de la investigación a la comunidad científica de la manera más convincente y ética, es decir, las competencias de ser y estar del investigador, así como el compromiso y la implicación con la investigación para el desarrollo y la innovación.

El libro que presentamos tiene el nivel adecuado para los que se inician en la investigación y para aquellos que se encuentran en un nivel más avanzado. Cada lector o estudiante puede utilizarlo de acuerdo con su nivel de estudios. Para los estudiantes de **grado**, el libro permite iniciarse en el proceso y las técnicas de la investigación básica. Para los de **máster**, nivel más avanzado, este libro es un apoyo al trabajo que deben realizar de su Trabajo de Fin de Máster. Y para los doctorandos, este texto les ayudará a profundizar y aplicar cada uno de sus capítulos en el logro de los requisitos establecidos en su proceso de formación doctoral (lectura de artículos relacionados con sus tesis, presentación de comunicaciones a congresos, elaboración de artículos, etc.), así como las competencias propias en la realización de sus tesis doctoral, como por ejemplo las competencias técnico-académicas: conocimiento y aplicación de los pasos del proceso de la investigación, capacidad de

identificación de las técnicas más apropiadas al problema de investigación, capacidad de análisis e interpretación de los datos de la investigación, presentación escrita y oral de la investigación, entre otras.

Para alcanzar estas competencias el libro está dividido en trece capítulos. En el capítulo primero se presenta el proceso general de la investigación y se explican los pasos, desde el inicio hasta la presentación oral de los resultados de la investigación.

En los capítulos dos y tres se introduce el paquete estadístico MYSTAT. Primero, la instalación del programa en el sistema operativo Microsoft Windows y, a continuación, el manejo del editor de datos con la versión más reciente del programa.

Los capítulos cuatro y cinco se centran en la estadística descriptiva y representaciones gráficas de datos cuantitativos. Los dos siguientes capítulos, seis y siete, abordan los módulos dedicados a la diferencia entre medias y el ANOVA, y el capítulo ocho introduce la correlación y la regresión.

En esta nueva edición se han añadido capítulos nuevos como el nueve, dedicado al cálculo de la validez y fiabilidad de los instrumentos de recogida de datos de la investigación, y el capítulo once, que aborda el análisis de datos cualitativos e introduce uno de los programas más utilizados para el análisis cualitativo, el ATLAS.ti. El capítulo diez se centra en las técnicas de la estadística no paramétrica, de amplio uso y aplicación en investigación educativa. Por último, se ha introducido el capítulo doce y trece, que tratan la presentación de informes escritos y la exposición y defensa oral de los resultados de investigación.

La presente edición es el resultado de varios años de experiencia docente con estudiantes de grado, máster y doctorado en la aplicación del programa MYSTAT. El material se ha adaptado a la nueva versión del programa MYSTAT y se han introducido las experiencias acumuladas tanto en las clases teóricas como prácticas. Cada uno de los capítulos desmenuza paso a paso las temáticas abordadas, lo que permite el uso individual por parte del estudiante así como del profesor.

Fuensanta Hernández Pina

Capítulo 1.

EL PROCESO DE INVESTIGACIÓN CIENTÍFICA

Fuensanta Hernández Pina
Javier J. Maquilón Sánchez

INTRODUCCIÓN

En este primer capítulo abordamos los aspectos básicos para la planificación y ejecución de una investigación. Primero se destaca la importancia de la investigación como el proceso científico para la generación de conocimientos válidos y fiables. También se justifica la necesidad de investigar, como el modo adecuado para dar respuesta a los problemas educativos y sociales en base a evidencias y resultados científicamente contrastados.

Se explican con detalle los pasos que los investigadores han de realizar cuando llevan a cabo una investigación empírica. En cada uno de los apartados se desarrollan y ejemplifican las tareas a realizar, además de identificar las metodológicas más apropiadas para algunos de los problemas de investigación planteados.

¿Qué tienen en común todas las metodologías de investigación? Todas ellas siguen el método científico, siendo comúnmente compartido por todas las áreas de conocimiento y por los investigadores, independientemente del ámbito en el que desarrollen su actividad.

En este libro, como material docente, se pretende iniciar al estudiante en la adquisición de la competencia investigativa. Por ello consideramos que este primer capítulo le puede capacitar para:

1. Analizar la importancia de la metodología de investigación para dar respuesta a las preguntas de investigación.
2. Identificar, describir y aplicar los pasos y acciones del proceso de la investigación.

3. Seleccionar la metodología de investigación más adecuada al problema planteado.

De modo gráfico, el proceso de investigación puede representarse y sintetizarse (Figura 1.1) en torno a los conceptos clave que se incluyen a continuación y que serán abordados a lo largo de este primer capítulo (Hernández-Pina y Maquilón, 2009).

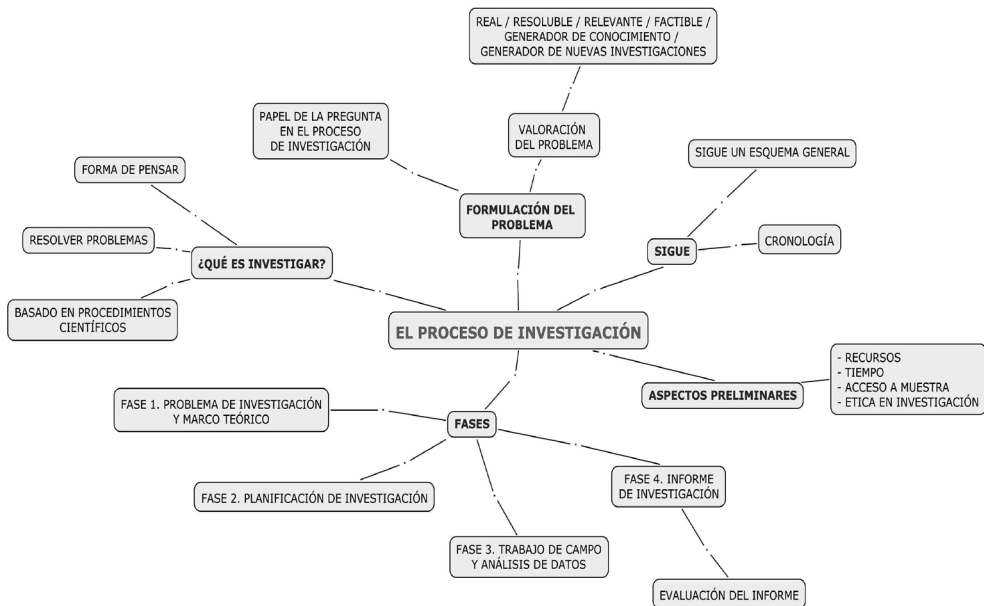


Figura 1.1. Representación gráfica del proceso de investigación.

1. CONCEPTUALIZACIÓN DE LA INVESTIGACIÓN

Resulta necesario, antes de profundizar en el proceso de la investigación científica, que se clarifiquen los conceptos básicos y que se delimite la función que la *pregunta de investigación* tiene dentro de dicho proceso. Investigar podría definirse como una forma de pensar y resolver problemas, ya que la investigación permite buscar posibles soluciones a través de procedimientos científicos y no a través de la elucubración, pero es un concepto que debe abordarse más profundamente.

1.1. Concepto de investigación

Entre las diversas definiciones de lo que es investigación, consideramos que puede ser entendida como «el estudio de los métodos, los procedimientos y las técnicas utilizados para obtener conocimientos, una explicación y una compren-

sión científicos de los fenómenos educativos, así como también para “solucionar” los problemas educativos y sociales» (Hernández Pina, 1995, p. 80). Esta definición proporciona una gran importancia a los métodos de investigación como base disciplinar, amplía su radio de acción y da cabida a todos los planteamientos metodológicos, ya sean cuantitativos, cualitativos o mixtos. Investigar podría interpretarse como el proceso mediante el cual intentamos conseguir información sistemática y rigurosa que nos permita dar respuestas a preguntas, solucionar problemas o lograr un mejor conocimiento y comprensión del mundo que nos rodea; siempre basándonos en evidencias.

Desde esta perspectiva, una definición de investigación debe hacer referencia al fenómeno u objeto de estudio, a los métodos, procedimientos y técnicas adecuados (métodos y metodología), y a la finalidad de la misma (creación y acumulación del conocimiento o solución de problemas). En este capítulo nos centramos exclusivamente en los aspectos metodológicos relacionados con el proceso de la investigación, ya que en otros capítulos se abordan las técnicas cuantitativas y cualitativas de análisis de datos, la elaboración de informes, la presentación y difusión de resultados, etc.

1.2. Problemas y preguntas de investigación

Una pregunta de investigación es el desencadenante de cualquier investigación que siempre se enmarca bajo un paradigma de investigación y, por lo tanto, condiciona el método y de la metodología que se debe adoptar para resolver un problema. Los métodos de investigación son formas alternativas de responder a preguntas de investigación, desde diferentes modos de investigar, dependiendo siempre de los objetivos que se hayan planteado que condicionarán la elección del método. El investigador no decide a priori qué método y metodología va a seguir en su investigación, sino que es la formulación de su pregunta la que le llevará al método y a la metodología que debe seguir. En la figura 1.2, se observa la relación que se establece entre la pregunta de investigación y la elección del método y la metodología.

La necesidad de recurrir a métodos diversos para estudiar un tema educativo no es una invención reciente. Ya Aristóteles reconocía esta multiplicidad metodológica en su tratado *De Anima*, cuando indicaba que si no existe un método general único para resolver las cuestiones de esencia, para cada materia diferente habremos de determinar el proceso apropiado de investigación (Hernández Pina, 1995, p. 83).

Toda investigación comienza con la clarificación o comprensión del problema, además de la delimitación del mismo. Decidir qué pregunta o preguntas formular y seleccionar el modo de indagación disciplinada más idóneo para responder a cada pregunta, ya sea inductivo o deductivo (Figura 1.3). Para dar respuesta a las múltiples preguntas que una investigación puede generar es necesario tener un conocimiento amplio de la diversidad de métodos de investigación existente para realizar la elección más apropiada.



Figura 1.2. Papel de la pregunta en el proceso de investigación (Buendía, Colás y Hernández-Pina, 1997).

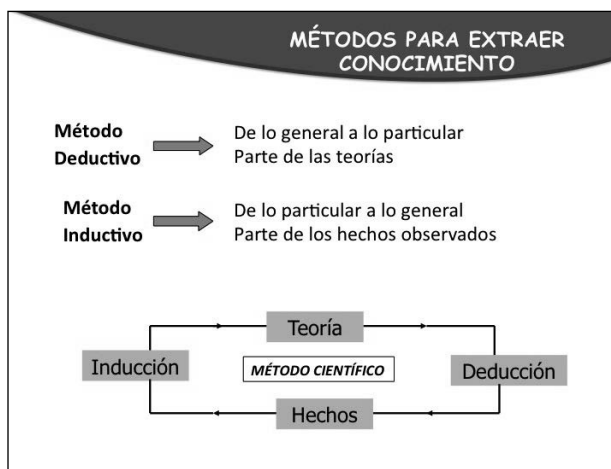


Figura 1.3. Métodos científicos para extraer conocimiento.

2. DESARROLLO DE UNA INVESTIGACIÓN. FASES DEL PROCESO

Como ya hemos anticipado, el proceso de investigación sigue una serie de fases que aparecen representadas en el siguiente esquema conceptual (Figura 1.4). Abordar y clarificar cada una de estas fases conlleva detallar todo el proceso de investigación; para ello comenzamos por los aspectos preliminares, continuando con la cronología o distribución temporal del planteamiento de la investigación. El resto de fases, dada su extensión y complejidad se abordan en el apartado 3 en el que se explican cada una de las cuatro fases principales de una investigación y las etapas y acciones vinculadas a las mismas.



Figura 1.4. Esquema del proceso de investigación.

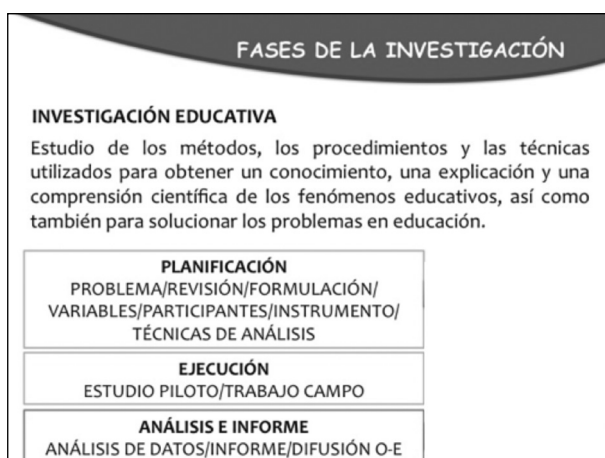


Figura 1.5. Fases de la investigación.

Básicamente (Figura 1.5), podemos anticipar que las acciones o componentes relacionadas con la planificación de la investigación deben materializarse en la planificación de investigación (lo que tenemos previsto hacer). A continuación se realizará la investigación, lo que supone la realización del estudio piloto y el trabajo de campo, para concluir con el análisis la elaboración del informe de investigación, que integra los análisis, de los resultados, conclusiones y discusión de los datos obtenidos en la fase de ejecución de la investigación.

2.1. Aspectos preliminares de la investigación

Antes de ejecutar una investigación han de tenerse en cuenta una serie de aspectos que deben ser valorados positivamente por parte del investigador. En caso de no superar esta valoración, debería volverse a planificar la investigación. Tales aspectos son:

- *Disponibilidad de recursos.* Que el investigador disponga de los recursos materiales y humanos necesarios para poder llevar a cabo la investigación con las suficientes garantías de éxito.
- *Tiempo disponible.* Realizar una estimación lo más real posible del tiempo que hay que dedicarle a cada fase y acciones de la investigación, constatando que realmente se dispone del mismo para poder realizarla.
- *Acceso a la muestra/participantes.* Valorar si el investigador va a poder recoger la información de los participantes en la investigación como fuentes de datos. Por ejemplo, si se va a realizar una investigación sobre las “causas que han llevado a los jóvenes menores de 18 años, a consumir internet”, tenemos que constatar que el investigador podrá acceder a jóvenes menores de 18 años, y que se le va a permitir recoger la información aportada directamente por los jóvenes objeto de estudio.
- *Consideraciones éticas.* El investigador ha de ser honesto y sincero a lo largo de toda la investigación, no falsear datos, no condicionar la respuesta de los participantes, no facilitar información personal de los participantes, etc., es decir, respetar todas las consideraciones éticas de cualquier investigación.

2.2. Planificación temporal de un trabajo de investigación

Este apartado supone la realización de un calendario de fases y sus tareas asociadas, para la adecuada y planificada realización de un trabajo de investigación. En este calendario se deben incluir dos aspectos principales (-1- Tiempo estimado de dedicación o fecha aproximada de la realización y, -2- Tareas a realizar) para cada uno de los tópicos principales en el desarrollo de cualquier trabajo de investigación. En la figura 1.6 se incluye una propuesta de planificación para que el investigador cumplimente con los datos de su propio trabajo o, al menos, le sirva de referente para la elaboración de su propio cronograma de investigación.

CRONOLOGÍA DE LA INVESTIGACIÓN		
Tópico	Tiempo/fecha	Tareas a realizar
Elaboración del proyecto		
Revisión de la literatura		
Actualización de conocimientos sobre metodología, informática, análisis de datos, etc.		
Preparación de los materiales		
Trabajo de campo.		
Análisis de los datos. Recursos y ayudas si fuese necesario		
Elaboración y redacción del trabajo.		
Preparación para la presentación escrita, encuadernación, etc. y temas administrativos. Los plazos.		
Disposición de todo lo necesario para la presentación oral y defensa.		

Figura 1.6. Planificación temporal de una investigación.

3. FASES, ETAPAS Y ACCIONES EN EL DESARROLLO DE LA INVESTIGACIÓN

3.1. Fase primera. El problema de investigación y el marco teórico

ETAPA 1. PRIMERA REVISIÓN DE LA BIBLIOGRAFÍA

La concreción del problema de investigación se hace a partir de una primera revisión de la literatura conceptual y de investigación, específica del área de conocimiento a la que pertenece el problema que vamos a investigar. Esta primera revisión (Figura 1.7)

REVISIÓN DE LA BIBLIOGRAFÍA	
REVISIÓN DE LA BIBLIOGRAFÍA	
⊙	Conocimiento general del tema
⊙	Delimitar el área problemática
⊙	Constatar novedad y originalidad
⊙	Orientar para la valoración
⊙	Colaborar en la formulación del problema
⊙	Realizar un marco teórico

Figura 1.7. Primera revisión bibliográfica.

permitiría conocer qué han hecho otros investigadores en ese tema, qué metodología han utilizado, qué instrumentos han empleado, a qué conclusiones han llegado, etc. Nos permite saber de qué partimos y hacia dónde vamos. Todas estas acciones culminarán con la delimitación del tema de investigación que se va a realizar.

ETAPA 2. LA FORMULACIÓN DEL PROBLEMA DE INVESTIGACIÓN Y LA ELECCIÓN DEL MÉTODO

Los problemas de investigación se formulan inicialmente de forma muy general. Es importante que los problemas se especifiquen, delimiten y concreten lo máximo posible. Para ello hay que restringir el campo de estudio y formular una pregunta general y varias específicas. Como se observa en la figura 1.8 antes de llegar a la formulación definitiva del problema hay que elegir el área de investigación (área problemática), tras lo cual debemos identificar y delimitar el problema para hacerlo abordable con los recursos de que dispongamos (humanos y materiales), seguidamente se debe valorar el problema comprobando que es real, resoluble, relevante, factible, generador de nuevos problemas y generador de nuevos conocimientos (Figura 1.9). En el caso de obtener una valoración positiva en todos los aspectos indicados estamos en situación de formular definitivamente el problema de investigación.



Figura 1.8. Proceso para la formulación definitiva del problema de investigación.

Un problema de investigación resulta significativo e interesante cuando tiene relación con una teoría, conocimiento o práctica en su desarrollo. La importancia puede partir en uno o varios de estos criterios:

- Conocimiento acerca de una práctica permanente.
- Comprobación de una teoría.
- Generabilidad.

- Ampliación de la comprensión de un tema.
- Avance metodológico.
- Temas actuales.
- Evaluación de una práctica o política específicas en un lugar determinado.
- Investigación exploratoria.

Como se anticipaba, existen una serie de características que han de valorarse antes de formular definitivamente el problema (Figura 1.9). Si el investigador que, tras el estudio previo del área problemática y la revisión de la literatura, da una respuesta positiva a todas y cada una de ellas, ya se encuentra en disposición de formular definitivamente el problema y ejecutar el proyecto de investigación, realizando un estudio piloto y el trabajo de campo. Para ello ha de valorarse que el problema de investigación sea:

- *Real*. Que esté basado en un hecho real, contrastado y no en supuestos o creencias.
- *Resoluble*. Que el problema planteado tenga al menos una posible solución, no tratándose de algo utópico e inalcanzable.
- *Relevante*. Que los resultados aportados tras la investigación vayan a tener importancia para la comunidad científica.
- *Factible*. Que el problema pueda resolverse con los recursos que se cuentan. Que se podrá acceder a la muestra de investigación y obtener información de la misma.
- *Generador de nuevos conocimientos*. Que los resultados de la investigación constituirán una aportación novedosa al campo científico objeto de estudio.

CARACTERÍSTICAS DESEABLES (VALORACIÓN)			
Al valorar un problema se debe responder si es:			
Real, Resoluble, Relevante, Factible, Generador de nuevos conocimientos, Generador de nuevos problemas			
Un problema se debe expresar en modo de pregunta y ha de incluir la descripción, asociación o intervención entre las variables más relevantes de la investigación.			
Tipos	Origen	Hipótesis	Finalidad
Abierto	Realidad	Inductiva	Generar teorías
Cerrado	Teoría	Deductiva	Comprobar teorías

Figura 1.9. Valoración de problemas de investigación.

- *Generador de nuevos problemas de investigación.* Que las conclusiones de la investigación sean el desencadenante de nuevos problemas a resolver, aportando novedosas implicaciones de investigación al área de estudio.

A modo de conclusión se puede decir que un problema de investigación es adecuado cuando contribuye a aumentar el cuerpo de conocimientos, tanto teórico como práctico, de una disciplina; conduce a nuevos problemas de investigación, pudiéndose desarrollar mediante procedimientos empíricos y se ajusta a las posibilidades del investigador como por ejemplo, a sus líneas de investigación, a su preparación, a los recursos disponibles y al tiempo real con el que cuenta para hacer su trabajo.

Para profundizar remitimos al lector a otro manual (Buendía, Colás y Hernández-Pina, 1997) en el que se ejemplifica cómo la pregunta de investigación se convierte en el desencadenante de la investigación, su relación con los paradigmas de investigación desde el que plantear la investigación, seleccionar el método y la metodología más adecuados para responder a la pregunta formulada.

3.2. Fase segunda. Planificación de la investigación

En esta segunda fase se desarrollan las etapas y acciones relacionadas con la planificación de la investigación, integra desde la etapa 3 a la etapa 7, permitiendo hacer efectivo el trabajo de campo que se explica en la etapa 8 que corresponde a la tercera fase de la investigación.

ETAPA 3. REVISIÓN DE LAS FUENTES BIBLIOGRÁFICAS

La revisión de las fuentes, una vez definido el problema de investigación, sirve para conocer las informaciones que orientan a el estudio del problema que se está investigando, conocer cómo otros autores han abordado temas similares, evitar enfoques estériles repitiendo lo ya conocido, interpretar mejor los resultados, buscar nuevos enfoques en la forma de abordar los problemas, aportar más resultados al campo de estudio, etc. El análisis y la síntesis de las fuentes consultadas permitirán escribir el marco teórico de la investigación.

Las fuentes consultadas se clasifican en: *Referencias generales o preliminares, fuentes primarias y fuentes secundarias.*

ETAPA 4. LA FORMULACIÓN DE LOS OBJETIVOS Y LAS HIPÓTESIS

Para que el problema de investigación sea abordable y resoluble, el investigador debe concretar qué es lo que se va a realizar de un modo más específico y operativo; para ello formula los objetivos y las hipótesis (Figura 1.10). Los objetivos tienen un carácter más descriptivo y las hipótesis buscan la relación entre las variables. Las hipótesis se definen como conjeturas, soluciones tentativas, proposiciones o especulaciones que el investigador ofrece como posible respuesta a su problema de inves-

tigación. Se definen también como generalizaciones o suposiciones comprobables empíricamente que se presentan como la respuesta al problema de investigación.

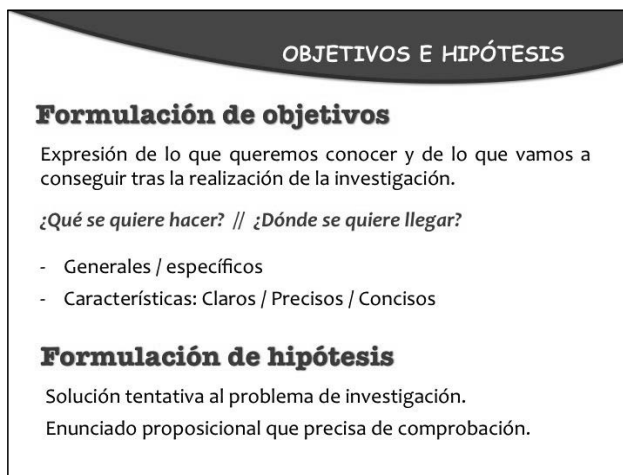


Figura 1.10. Formulación de objetivos e hipótesis.

Los objetivos, por su parte, se definen como la expresión de las descripciones que el investigador desea hacer en torno a la realidad que está estudiando. Los objetivos y las hipótesis son elementos básicos en cualquier investigación, ya que son el hilo conductor de la investigación, convirtiéndose en el puente de unión entre el problema de investigación y la solución del mismo.

A modo de sugerencia se incluyen una serie de indicaciones para redactar de modo correcto los objetivos y las hipótesis de investigación. Se enumeran seguidamente:

1. Concretar los objetivos y las hipótesis después de haber realizado la primera o la segunda revisión de la bibliografía.
2. En la redacción del informe de investigación, las hipótesis y los objetivos deben aparecer dentro del marco empírico de la investigación y antes del apartado dedicado a la metodología.
3. Redactar las hipótesis y los objetivos en forma expositiva y en afirmativa. Evitar las preguntas y la doble negación.
4. Redactar una hipótesis u objetivo general y varios específicos.
5. No redactar las hipótesis de investigación en forma de hipótesis nula.
6. Las hipótesis deben establecer diferencias o relaciones entre dos o más variables.
7. Definir claramente las variables que se incluyan en las hipótesis y los objetivos.
8. Es necesario que cada hipótesis y objetivo sea comprobable empíricamente.
9. En las hipótesis y objetivos deben aparecer de forma implícita:
 - Identificar las VI (variables independientes) y las VD (variables dependientes).

- Las medidas a utilizar (modo de valorar o medir cada variable).
- Sobre quién se va a hacer la investigación (participantes/muestra).
- El tipo de análisis a realizar (de modo orientativo, no indicando la prueba estadística a emplear).

ETAPA 5. DEFINICIÓN Y CATEGORIZACIÓN DE LAS VARIABLES

Una vez formulados los objetivos y/o las hipótesis el investigador debe definir las variables implicadas en sus formulaciones, así como los posibles valores (categorías) que pueden obtener dichas variables.

Una *variable* es la manifestación externa de un constructo o concepto. Es una característica o propiedad que adopta diferentes valores y que puede ser medida con alguna escala que varía en diferentes grados de precisión. La imagen, la percepción o el concepto que somos capaces de medir es lo que llamamos variables. En investigación diferenciamos entre:

- *Concepto* que expresa una abstracción formada por generalizaciones extraídas de casos particulares. Es una imagen mental.
- *Constructo* que es un concepto. Es un invento consciente y deliberado para un propósito científico particular.
- *Indicador* es la serie de criterios que el investigador selecciona como reflejo de un concepto.
- *Variable* es la manifestación externa de un constructo o concepto. Un concepto susceptible de medida que se convierte en variable.

Existen diferentes tipos de variables en función de la clasificación que se haga de ellas. Remitimos al lector a un texto más especializado para profundizar sobre este tema.

ETAPA 6. EL DISEÑO DE LA INVESTIGACIÓN Y LA ELECCIÓN DE LA MUESTRA

Siguiendo con el proceso de la investigación abordamos el apartado correspondiente al diseño y a la elección de la muestra o los participantes en la investigación. Ambas acciones están íntimamente relacionadas ya que el diseño que se seleccione condicionará la elección de la población/muestra o los participantes del estudio.

Un diseño de investigación puede definirse como el puente de unión entre el problema de investigación y la posible solución al mismo. En el diseño de la investigación se integran o establecen relaciones de los diferentes componentes de la investigación (Figura 1.11), permitiendo identificar ¿qué se investiga?, ¿a quién?, ¿dónde?, ¿por qué?, ¿cómo? y ¿cuándo?

A continuación se abordan los aspectos más básicos a tener en cuenta a la hora de elegir la muestra. Si necesita profundizar en este tema se sugiere consultar manuales específicos de metodología de investigación y técnicas de muestreo.



Figura 1.11. Diseño de investigación.

Los elementos, ya sean personas, documentos (escritos/gráficos) o fenómenos, constituyen la muestra de una investigación. Existe una relación jerárquica de conceptos en torno al proceso de muestreo que debe clarificarse:

- *Universo* es la serie real o hipotética de elementos que comparten unas características relacionadas con el problema de investigación.
- *Población* es un conjunto definido, limitado y accesible del universo que forma el referente para poder seleccionar la muestra. Es el grupo al que se intenta generalizar los resultados.
- *Muestra* es el conjunto de individuos extraídos de la población a partir de algún procedimiento específico de muestreo.
- *Elemento o individuo (muestral)* es la unidad más pequeña en la que podemos descomponer la muestra, la población o el universo. Esta unidad puede ser una persona, un grupo, una institución, un centro, un documento, un fenómeno, etc. La identificación de este elemento se realiza en función del problema de investigación.

Todos estos conceptos deben ser definidos en cada investigación. En la figura 1.12 se representa cada uno de ellos y su relación cualitativa.

Otro de los aspectos que es necesario clarificar en este apartado, dentro de la etapa sexta de la investigación, es la identificación de los distintos tipos de muestra (Figura 1.13) y que se definen como:

- *Muestra invitada*. Compuesta por los sujetos de la población a quienes les solicitamos participar en la investigación.
- *Muestra participante*. Integrada por los sujetos que aceptan formar parte de la investigación.

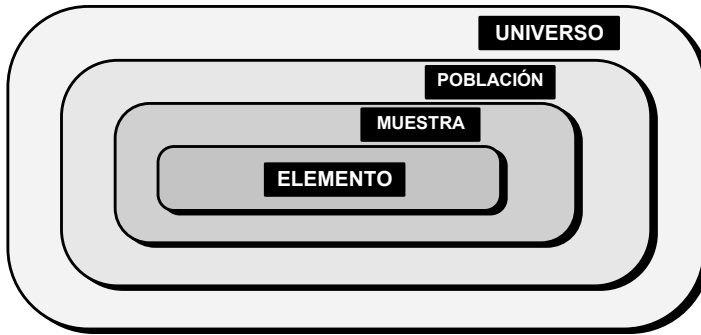


Figura 1.12. Relación entre los conceptos básicos de muestreo (Buendía, Colás y Hernández-Pina, 1997).

- *Muestra real o productora.* Es aquella parte de la muestra participante que aporta la información (datos), que servirán para realizar el trabajo de investigación, respondiendo a los objetivos e hipótesis, obteniendo conclusiones, etc. No toda la muestra participante llega a ser real, ya que puede descartarse información por diversos motivos.



Figura 1.13. Clasificación de las muestras de investigación (adaptado de Buendía, Colás y Hernández-Pina, 1997).

Por último, la elección de los participantes debe realizarse utilizando alguna técnica de muestreo y luego describirla. Por motivos de espacio mencionamos dichas técnicas y remitimos a textos más especializados. Las técnicas se clasifican en probabilísticas que tienen su base en la equiprobabilidad (ejemplo: azar simple, aleatorio sistemático, aleatorio estratificado, etc.); muestreo no probabilístico que tienen su base en nuestros deliberados, accidentales, casuales, intencionados, voluntarios, etc.; y por último muestreos mixtos que combinan los dos anteriores.

ETAPA 7. LOS INSTRUMENTOS DE RECOGIDA DE DATOS

La siguiente acción en el proceso de la investigación es la elección o construcción de los instrumentos para la recogida de la información (datos). Entendemos por dato la información recogida de los sujetos, objetos o características (grupos, programas, contexto, etc.) de la investigación. Estos datos pueden adoptar formas diversas, como por ejemplo la edad, sexo, ingresos económicos, nivel de estudios, respuestas a un cuestionario, diarios, anecdóticos, escalas de observación, escalas de actitudes, etc.

La decisión que el investigador debe tomar en esta fase de planificación es qué clase de datos necesita recoger y qué instrumento o técnica empleará. La instrumentación implica no solo la recogida o diseño de los instrumentos sino las condiciones en que se aplican los mismos, como puede ser el dispositivo que utilice (papel y lápiz, cuestionario, escala, plataforma informática, etc.). De este modo se plantean una serie de interrogantes al investigador que debe responder para poder continuar (Figura 1.14):

- ¿Cuándo recogerá los datos? *Determinar el momento más oportuno.*
- ¿Dónde recoger los datos? *Determinar el lugar.*
- ¿Con qué frecuencia se recogerán los datos? *Determinar cuántas veces deberán los participantes suministrar información.*
- ¿Quién recogerá la información? *Determinar quién aplicará los instrumentos.*
- ¿Con qué instrumentos se recogerá la información? *El investigador elabora el instrumento o emplea alguno previamente estandarizado.*

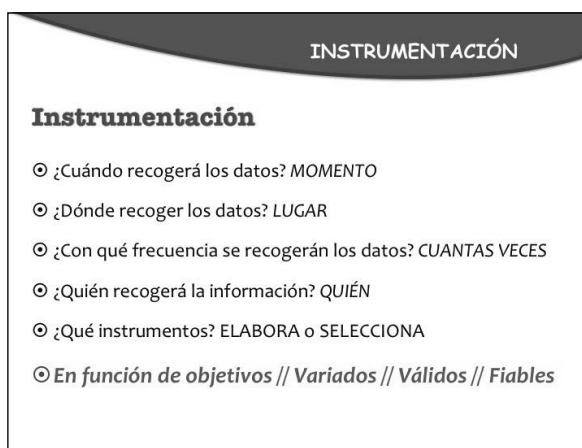


Figura 1.14. Instrumentación en la investigación.

El investigador se puede encontrar ante la disyuntiva de seleccionar los instrumentos de recogida de información de entre alguno ya existente o tener que elaborarlos. En cualquiera de estas situaciones deben tenerse unos amplios conocimientos sobre el tema, ya que seleccionar el instrumento de los ya existentes requiere conocer los "posibles", previamente habiendo identificado las características, potencialidades

y limitaciones, y así realizar la elección más adecuada a los objetivos de la investigación. Por el contrario, si el investigador decide elaborar los instrumentos de recogida de información, es necesario disponer de la formación adecuada en el diseño y validación de instrumentos y en técnicas de recogida de datos.

3.3. Fase tercera. Trabajo de campo y análisis de los datos

ETAPA 8. TRABAJO DE CAMPO

El trabajo de campo consiste en llevar a cabo el proyecto de investigación y realizar todas las tareas planificadas con el fin de cumplir los objetivos/hipótesis de la investigación. De acuerdo con el esquema general del proceso de la investigación esta fase se inicia con el estudio piloto (Figura 1.15), muy necesario en muchas investigaciones ya que puede evitar errores irreparables que conlleven el tener que abandonar la investigación.

En esta etapa también se incluye la explicación del procedimiento de investigación y los términos en los que debe ser redactada dentro del informe de investigación.

A) Estudio piloto

Al finalizar el planteamiento de la investigación, es recomendable llevar a cabo este tipo de estudio. Consiste en realizar una investigación a pequeña escala de la que está prevista ejecutarse.

Su finalidad es detectar fallos o desajustes en la planificación de la investigación, que se materializa en el proyecto de investigación, como son revisión de las fuentes, formulación de objetivos, muestreo, desajustes en los instrumentos de recogida de datos, elección del diseño de investigación, o cualquier otro fallo que podría acarrear grandes costos de tiempo y materiales que en determinadas situaciones podrían no llegar a ser asumibles, teniendo que cancelar la investigación. El estudio piloto no es preceptivo, pero si se duda de la viabilidad de la investigación a gran escala se hace muy recomendable su realización.

El estudio piloto se realiza siempre con un número reducido de sujetos que tienen las mismas características que los sujetos de la población de la que posteriormente se extraerá la muestra, teniendo por finalidad mejorar la forma en que se van a recoger los datos, depurar los instrumentos de recogida de los datos, corregir posible fallos en el procedimiento, familiarizarse el investigador con esta fase de procedimiento, planificar las técnicas de análisis de datos más apropiadas, recibir feedback de los sujetos para mejorar el procedimiento, la aplicación de los instrumentos, etc.

B) Procedimiento

El procedimiento se define como aquellas tareas que el investigador realiza con los participantes o las actividades que se realizan durante todas las fases de la investigación que finalmente produzcan datos.

Redactar el procedimiento consiste en detallar la secuencia de acontecimientos que tienen lugar al llevar a cabo la investigación o experimento. Se debe especificar claramente la secuencia de tareas o actividades que se harán con los sujetos o harán los sujetos (Figura 1.15). En un estudio de laboratorio sería la explicación detallada de cada tarea con aparatos o materiales que los sujetos tienen que manipular. En un aula serán las actividades que los sujetos realizarán, bien sea el método al que se han sometido o las tareas que ha tenido que realizar para producir los datos que el investigador necesita recoger. En la publicación de artículos e investigaciones este apartado es importante porque permite realizar réplicas de la investigación a otros investigadores.



Figura 1.15. Estudio piloto y procedimiento de la investigación.

En el procedimiento también resulta muy interesante, sobre todo para poder llevar a cabo réplicas de la investigación en otros contextos, clarificar cuáles han sido los problemas con los que el investigador se ha encontrado a lo largo del proceso y, lo más importante, cómo estos inconvenientes han sido resueltos, facilitando así el trabajo a futuros investigadores.

ETAPA 9. ANÁLISIS DE LOS DATOS. ESQUEMAS DE DECISIÓN

Esta fase tiene un carácter teórico-práctico. En ella se incluyen las etapas en las que hay que diseñar el plan de análisis de datos más apropiado al problema e hipótesis/objetivos planteados, la obtención de los datos y su análisis, la comprobación de las hipótesis o la descripción de los resultados de los objetivos.

En la etapa 7 hemos explicado los instrumentos de recogida de los datos. La información recogida con ellos va a condicionar las técnicas de análisis de datos que serán empleadas en la investigación.

A la hora de plasmar la información obtenida, el investigador tiene dos modos de hacerlo: de forma numérica o a través de palabras, en forma de texto o narrativa. Por tanto en investigación podemos encontrarnos con dos tipos de datos (o variables): datos cualitativos y datos cuantitativos.

Los datos cuantitativos se obtienen cuando las variables estudiadas se miden a lo largo de una escala que indica la cantidad presente de la variable medida. Su representación se hace de forma numérica. En investigación tenemos múltiples ejemplos de este tipo de datos: puntuaciones en un test, valores de rendimiento, respuestas a una escala de estimación numérica, etc.

Los datos categóricos o cualitativos expresan cuántos elementos puede haber dentro de una determinada categoría. La forma de representarlos es a través de frecuencias o porcentajes. Los datos categóricos expresan el número de objetos, elementos, individuos, acontecimiento, opiniones, etc., dentro de una categoría determinada. El investigador puede informar del número o porcentaje de sujetos que cumplen una determinada condición, de la frecuencia o el porcentaje de sujetos que expresan una respuesta en un cuestionario, etc.

El análisis estadístico se asocia normalmente a la investigación cuantitativa, especialmente con la investigación de corte experimental, quasi-experimental, tipo encuesta, etc. En cambio, el análisis en la investigación cualitativa descansa más en la reducción, la inducción y en la descripción.

El término estadística (hoy se ha optado más por hablar de análisis de datos) tiene múltiples significados; sin embargo, la mayoría se refiere como la teoría, procedimiento y metodología por medio de la cual los datos son analizados. La estadística no debemos entenderla como una acumulación de datos sino como el modo de describir y dar sentido a grandes cantidades de datos. La estadística tiene un valor heurístico de gran importancia si el investigador conoce todas sus posibilidades.

La elección de la técnica de análisis más apropiada va a depender de dos aspectos: primero, de qué tipo de datos disponemos y, segundo, qué tipo de análisis es el apropiado a los objetivos o a las hipótesis. En la figura 1.16 presentamos tres organigramas que facilitan la identificación del tipo de análisis a utilizar. Para ello, primero hay que partir del tipo de datos: nominales, ordinales o de intervalos. A continuación, comprobar qué es lo que los objetivos/hipótesis demandan, es decir, describir, establecer relación entre las variables o establecer hipótesis de la diferencia entre grupos.

En cuanto a la investigación cualitativa se define como aquella en la que los investigadores recopilan los datos en situaciones reales mediante la interacción con las personas que participarán, siendo fuente de los datos necesarios para llevar a cabo la investigación. Estos diseños se basan en una filosofía constructivista que asume la realidad socio-educativa como una experiencia heterogénea, emergente, interactiva y socialmente compartida que es interpretada individualmente por cada persona en función de sus percepciones y creencias. Es por ello que este tipo de investigación analiza y describe las conductas sociales, las opiniones, pensamientos y las percepciones de las personas.

Los resultados de investigaciones cualitativas son relevantes para la generación de los fundamentos teóricos de los temas abordados, el desarrollo de normas, la mejora de la práctica educativa y para poder dar respuesta a problemáticas socio-educativas. Durante la recogida de información los participantes pueden llegar a incluir sus propios sentimientos, sus creencias, sus ideas, sus pensamientos y sus conductas, lo cual complica enormemente el poder presentar los resultados en términos globales.

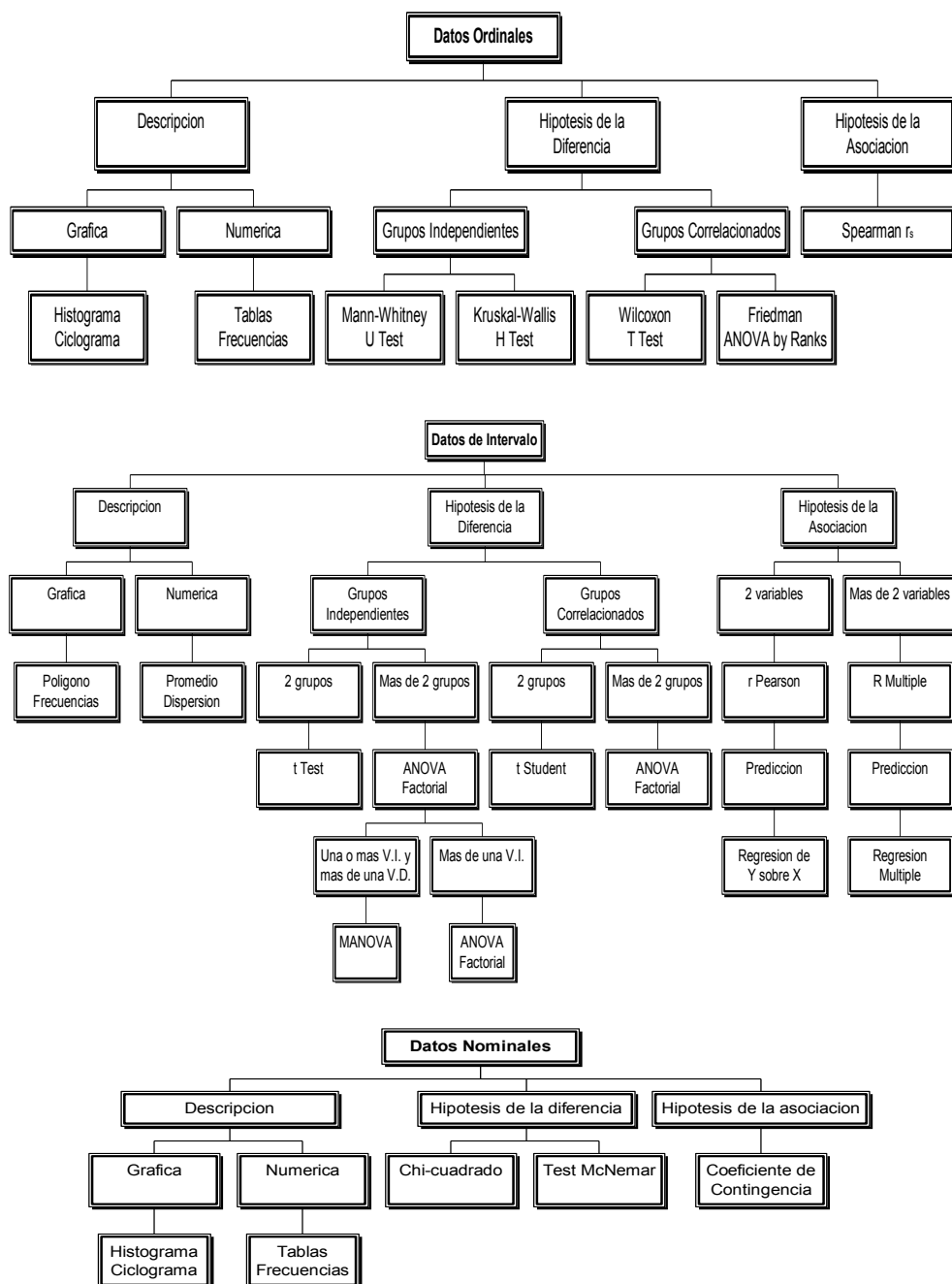


Figura 1.16. Organigramas para la elección de la prueba estadística (Buendía, Colás y Hernández-Pina, 1997).

Las técnicas de recogida de información toman un papel muy relevante en este tipo de diseño ya que de ellas va a depender el resultado de la investigación. Generalmente aunque se establecen los instrumentos a priori suele ser necesaria una revisión/ampliación de los mismos durante la investigación. Entre los más usuales podemos destacar las entrevistas, la observación participante, etc.

El grado de implicación del investigador también puede variar, ya que hay situaciones en las que se mantiene totalmente al margen, de un modo imparcial, y en otras en las que ha de implicarse y participar en el desarrollo de la investigación. Las metodologías de investigación de corte cualitativo necesitan de actividades que implican un conjunto de manipulaciones, transformaciones, reflexiones y comprobaciones realizadas a partir de los datos con el fin de extraer significados relevantes para un problema de investigación.

El procedimiento que se sigue para el análisis de la información está regulado por las tareas, actividades y operaciones que se integran dentro del proceso general de análisis de datos cualitativos (PGADC). Las fases principales son:

- Reducción de la Información
- Disposición y presentación de la información
- Obtención de resultados / Verificación de conclusiones

Cada una de estas tres fases se componen de una serie de actividades y operaciones que permiten realizar de uno modo científico el análisis de la información cualitativa con la que poder dar respuesta a los objetivos de las investigaciones que se puedan estar desarrollando bajo este paradigma de investigación.

3.4. Fase cuarta. El informe de la investigación

ETAPA 10. EL INFORME ESCRITO DE LA INVESTIGACIÓN. MODELO GENERAL

En este apartado ofrecemos una guía breve para la elaboración del informe de investigación ya que en otro capítulo de este manual se aborda en profundidad este tema. Los subapartados que se integran siguen las pautas establecidas por la Asociación Americana de Psicología (APA) en lo que respecta a la realización y presentación de informes de investigación.

PARTES DEL INFORME DE INVESTIGACIÓN

En esencia un informe de investigación consta de las siguientes partes:

PÁGINAS INICIALES (PORTADA/AGRADECIMIENTOS/ÍNDICE).

- 1. RESUMEN Y ABSTRACT (PALABRAS CLAVE).**
- 2. INTRODUCCIÓN TEÓRICA.**
- 3. METODOLOGÍA.**
 - a) Población y muestra. Participantes.
 - b) Diseño.

- c) Variables.
- d) Materiales e instrumentos de medida y de recogida de los datos.
- e) Procedimiento.

4. ANÁLISIS Y RESULTADOS.

5. DISCUSIÓN Y CONCLUSIONES.

6. REFERENCIAS.

ANEXOS Y APÉNDICES.

Explicamos brevemente cada una de las partes:

PORTADA

En la portada deberá aparecer el título del trabajo y el autor o autores del mismo. Es conveniente que en la portada también se incluya el nombre de la persona o personas que han encargado el trabajo, la materia, empresa, o institución en la que se presenta el trabajo, la entidad que ha financiado el proyecto, la persona que lo ha solicitado o dirigido, la editorial que lo publica y la fecha de entrega o defensa del mismo. También puede incluirse, de un modo muy discreto, alguna imagen o gráfico relacionado con el trabajo con la intención de hacerlo más vistoso y atractivo. Como recomendación es conveniente seguir las pautas específicas para dos diferentes modelos de informes, ya que la portada de un informe de investigación no es igual que la portada de un trabajo fin de máster (TFM) o una tesis doctoral.

ÍNDICE

En el índice o sumario se indicará cada una de las partes en las que se ha dividido el informe –introducción, capítulos y apartados de cada capítulo, referencias, etc.–, reseñando la página en la que se encuentra (*sólo la página en la que empieza*) cada una de dichas partes.

Después del índice, y aunque no suele ser obligatorio, pueden incluirse los agradecimientos o dedicatorias antes de comenzar con el contenido en sí del informe, TFM o tesis doctoral.

1. RESUMEN O ABSTRACT

Un abstract o resumen es la síntesis de la investigación que aparece al comienzo de un artículo o trabajo de investigación. Su finalidad es facilitar una información rápida del contenido de los resultados de una investigación. En un par de minutos el lector sabe de qué va el trabajo y si se relaciona o no con su tema de interés/estudio. Cuando un investigador está buscando fuentes de documentación escritas, la lectura de este resumen le permite decidir si el documento le es válido para su investigación o de lo contrario, descartarlo. Junto al resumen

y abstract se incluyen algunas palabras clave (normalmente cuatro) con las que indexar el trabajo en las bases de datos y los catálogos.

2. INTRODUCCIÓN

Esta sección debe incluir dos grandes bloques. Uno dedicado a la revisión de la literatura sobre el tema, que sustentará teóricamente la investigación, y otro que engloba el propósito de la investigación, así como los objetivos o las hipótesis.

En un informe de investigación en general, y en una TFM o tesis doctoral en particular, la introducción da paso al MARCO TEÓRICO del trabajo, incluyéndose en él toda la información teórica y de investigación relevante sobre el tema, estructurada en apartados y subapartados, justificándose así el estudio planteado. Este apartado debe concluir con la delimitación del problema a investigar y la formulación de la pregunta de investigación.

Seguidamente comenzará el segundo gran bloque del informe de investigación o de la Tesis Doctoral, nos referimos al MARCO EMPÍRICO, que se inicia con la formulación de los objetivos e hipótesis de la investigación, anticipando el siguiente apartado centrado en el método de investigación y la metodología a seguir, que se explica a continuación.

3. METODOLOGÍA

En esta sección el investigador informa sobre las acciones llevadas a cabo durante la realización de la investigación. La información incluida en este apartado debe ser lo suficientemente amplia y explícita como para que otro investigador pueda hacer una *réplica* del mismo en un contexto diferente. Dicha información variará dependiendo del método de investigación utilizado. En este gran apartado se incluyen y explican ampliamente los siguientes puntos:

- a) Población y muestra. Participantes.
- b) Diseño.
- c) Variables.
- d) Materiales e instrumentos de medida y de recogida de los datos.
- e) Procedimiento.

4. ANÁLISIS Y RESULTADOS

El investigador debe informar sobre las técnicas estadísticas utilizadas, justificando su elección. En cualquier caso, las técnicas empleadas para el análisis e interpretación de los datos responderán a las hipótesis u objetivos y nunca a la novedad de la técnica.

En este apartado debe responderse uno a uno a todos los objetivos y/o hipótesis, de un modo claro y ordenado, ya que muchos investigadores, *erróneamente*, una vez que disponen de los datos informatizados en

cualquier programa estadístico, comienzan a realizar análisis indiscriminadamente olvidando lo realmente relevante, que ya había quedado establecido en los objetivos e hipótesis al comienzo de la investigación. Recordemos que los objetivos hacen referencia a lo que consideramos más relevante descubrir, analizar, describir, etc.; así que, si llegado el momento creemos que podría realizarse algún análisis no previsto en los objetivos, convendría revisar los objetivos ya que, posiblemente, éstos no hayan sido formulados adecuadamente.

5. DISCUSIÓN Y CONCLUSIONES

En esta sección se señalan las implicaciones y utilidad de los hallazgos, efectuando una interpretación desde varias perspectivas, siempre desde el marco de los objetivos e hipótesis planteados. Es decir, al menos ha de incluirse un párrafo donde se sintetice cada uno de los objetivos previstos en la investigación, pero con una redacción diferente a la que se ha realizado en el apartado de análisis de datos.

6. REFERENCIAS

Este apartado incluye un listado ordenado alfabéticamente de aquellos trabajos que han sido citados en el informe (no los consultados, que habrán sido más). En cuanto al formato de presentación se debe seguir la normativa APA (Asociación Americana de Psicología).

ANEXOS Y APÉNDICES

En los anexos o apéndices se debe incluir aquella información relevante para la investigación pero que no ha sido incluida a lo largo del mismo, por ejemplo, hay investigadores que recogen el instrumento en el “cuerpo del trabajo” y otros que lo colocan como anexo. También debe incluirse aquella información relevante para la investigación considerada como difícilmente accesible para la persona o personas que tengan que evaluar el informe de investigación, TFM o la Tesis Doctoral. Por ejemplo, un documento al que hemos tenido acceso por una determinada razón y que no es accesible al público en general, etc.

ETAPA 11. PRESENTACIÓN ORAL DE UN INFORME DE INVESTIGACIÓN. EL PÓSTER

Resulta usual que los investigadores realicen la presentación de los resultados de su investigación oralmente. Esta presentación puede adoptar formas diversas como son la conferencia, la ponencia, la comunicación, la defensa de una tesis, de una TFM, el poster, etc.

Dada la importancia de esta etapa se incluye un capítulo específico sobre la presentación oral de trabajos de investigación y la realización de presentaciones visuales.

REFERENCIAS

- Colás-Bravo, P., Buendía Eisman, L. y Hernández-Pina, F. (2009). *Competencias científicas para la realización de una tesis doctoral*. Barcelona: Davinci.
- Buendía-Eisman, L., Colás-Bravo, P. y Hernández-Pina, F. (1997). *Métodos de investigación en psicopedagogía*. Madrid: McGraw-Hill.
- Hernández-Pina, F. (1995). *Bases metodológicas de la investigación educativa. Fundamentos*. Murcia: DM.
- Hernández Pina, F. y Maquilón Sánchez, J.J. (2009). El proceso de investigación. Del problema al informe de investigación. En P. Colás-Bravo, L. Buendía-Eisman, y F. Hernández Pina, *Competencias científicas para la realización de una tesis doctoral* (pp. 31-62). Barcelona: Davinci.

Capítulo 2.

INSTALACIÓN DEL PROGRAMA MYSTAT EN EL SISTEMA OPERATIVO MICROSOFT WINDOWS

*Tomás Izquierdo Rus
Fuensanta Monroy Hernández*

INTRODUCCIÓN

MYSTAT 12 para Windows es un programa estadístico basado en SYSTAT, Inc. Este programa ha sido diseñado para facilitar la enseñanza y el aprendizaje de la estadística para aquellos alumnos que durante sus estudios universitarios cursan esta disciplina, ya sea en grados o postgrados.

El uso masivo de los diferentes sistemas operativos de Microsoft Windows ha hecho posible que MYSTAT 12 se convierta en un programa estadístico de fácil manejo para los estudiantes que necesiten analizar datos en sus trabajos y proyectos de investigación tanto a lo largo de la carrera universitaria como en su posterior vida profesional.

Mystat 12 tiene por objetivo fundamental ofrecer un software que nos permita realizar la totalidad de los análisis estadísticos habitualmente utilizados en ciencias sociales. Es un software muy completo y en este manual, veremos sólo una parte muy débil de sus posibilidades. Existen muchos otros softwares como Systat o SPSS que permiten lograr los mismos fines, es decir hacer análisis estadísticos.

Podemos descargarnos gratuitamente Mystat 12 en la siguiente dirección: **<http://www.systat.com/MystatProducts.aspx>**

1. OBJETIVOS

Los objetivos que pretendemos con el presente capítulo son:

- Conocer diferentes procesos de instalación de un programa bajo cualquier sistema operativo de Windows.
- Instalar el programa Mystat en su ordenador de trabajo desde cualquier unidad de almacenamiento.
- Valorar positivamente la sencillez del proceso de instalación de un programa bajo Windows.

2. CONTENIDOS

Como hemos indicado en la introducción, Mystat 12 se podrá emplear exclusivamente en sistemas operativos de Microsoft Windows.

La instalación se puede realizar desde cualquier sistema de almacenamiento: disco duro, cd-rom, dvd, flash drive, etc., tan sólo debemos saber claramente la letra asignada a la unidad de almacenamiento en la que se encuentra el programa de instalación (Figura 2.1).

Uno de los procesos más cómodos para instalar cualquier programa es a través del Menú Inicio, el Explorador de Windows o Equipo (por defecto Windows instala un ac-

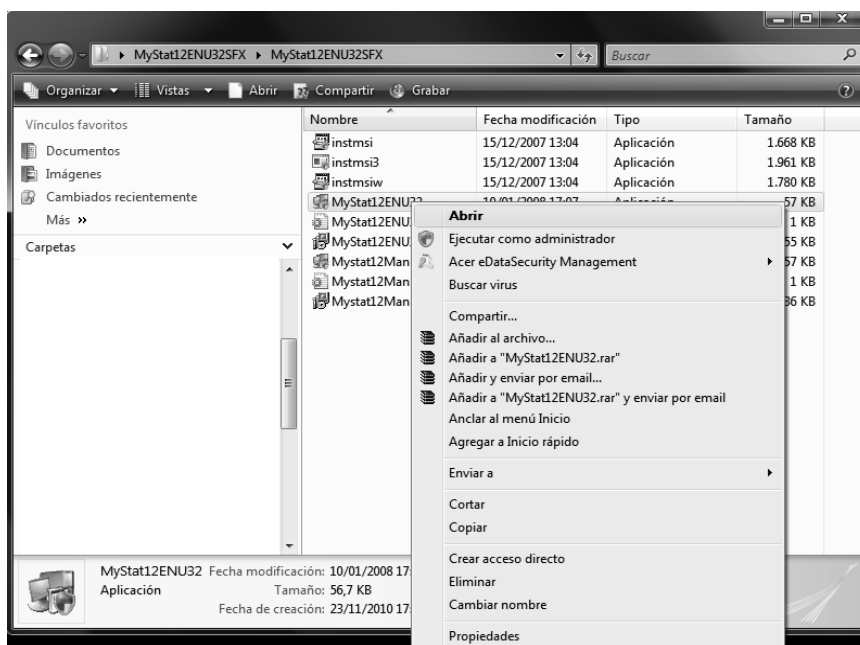


Figura 2.1. Localización del fichero para instalar Mystat.

ceso a este programa en el escritorio), tan solo ha de repicar sobre el icono y en la ventana que aparece localizar la unidad en la que se encuentra el programa para instalar.

Después de la unidad tenemos que localizar el fichero que dará comienzo la instalación del programa.

Otro modo de proceder a la hora de instalar un programa es a través del comando **Ejecutar**, al cual podemos acceder desde el botón **INICIO** que se encuentra en la esquina inferior izquierda de la pantalla. En la figura 2.2 puede observarse la ubicación exacta de dicho comando que nos permitirá abrir un programa, una carpeta, un documento o un sitio web.

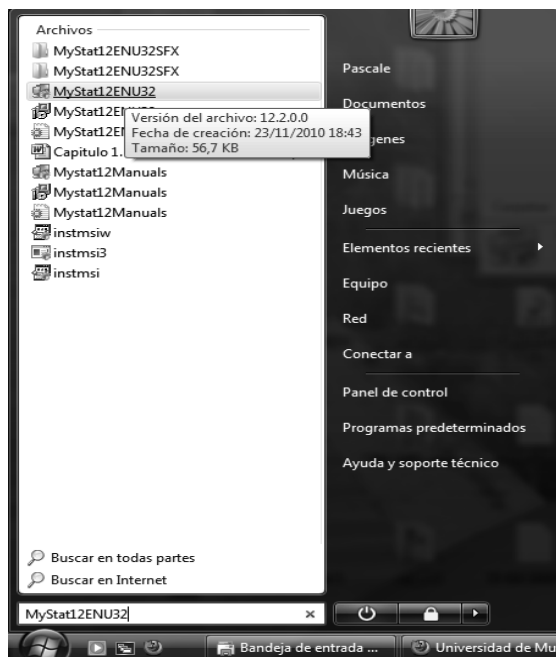


Figura 2.2. Ubicación del comando Ejecutar en Windows.

El procedimiento completo para instalar un programa a través de este comando comienza escribiendo el nombre del programa ejecutable en el espacio que aparece en la parte superior del icono de **Inicio de Windows**. En el menú desplegable nos aparecerán todos los archivos que, con nombres parecidos seleccionamos solamente el archivo ejecutable denominado **Mystat12ENU32**.

En el momento que se acepte ejecutar el fichero de instalación de Mystat comenzará dicha instalación. Aparece entonces la primera **pantalla del MYSTAT 12** (Figura 2.3). Posteriormente, nos saldrá una pantalla donde nos preguntará si queremos que el programa se instale en la carpeta por defecto que utiliza Windows, **C:\ARCHIVOS DE PROGRAMA\MYSTATW** o bien seleccionamos personalmente la ruta de instalación (Browse) (Figura 2.4).



Figura 2.3. Ventana instalación Mystat 12.

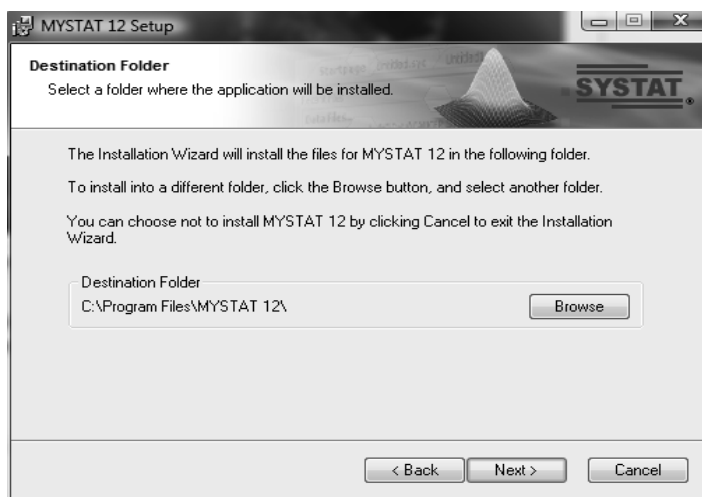


Figura 2.4 Selección carpeta instalación Mystat 12.

Una vez proporcionada toda la información necesaria (unidad y carpeta de instalación), pulse el botón **Next** y proseguirá la instalación del programa Mystat (Figura 2.5), y al concluir se creará el acceso directo al programa Mystat. Podremos acceder a él picando en **Inicio**, después **Todos los programas** y finalmente la carpeta **MYSTAT** (Figura 2.6).

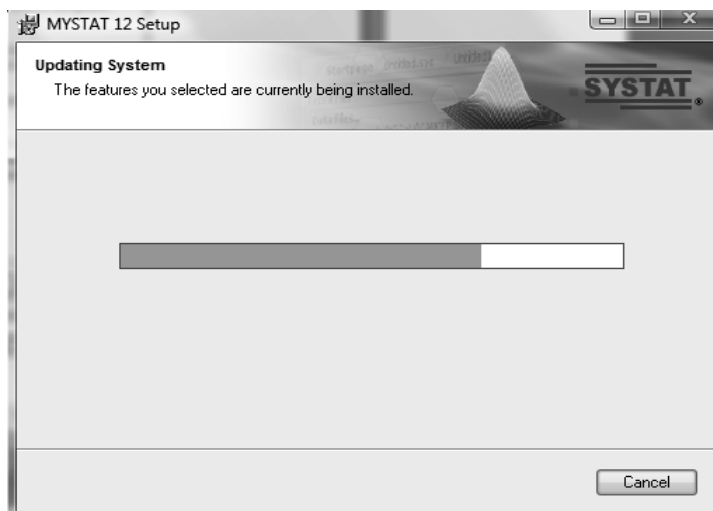


Figura 2.5. Proceso de instalación de Mystat.

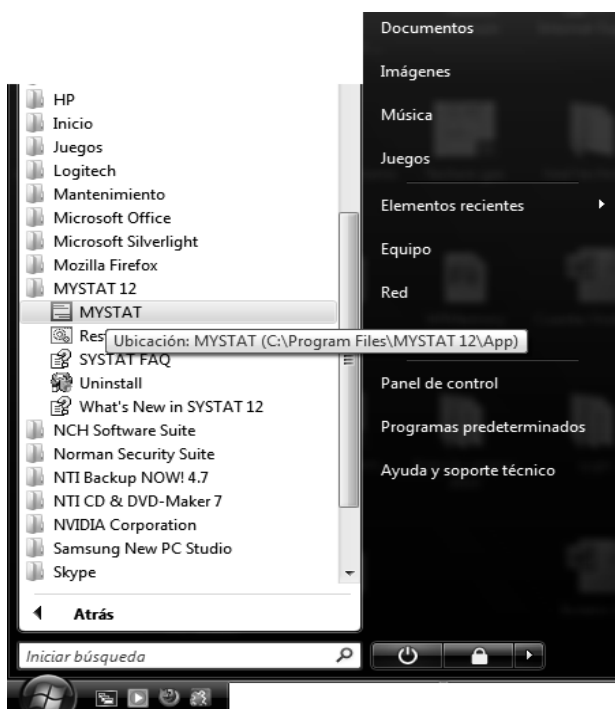


Figura 2.6. Ubicación del programa Mystat tras su instalación.

3. EJERCICIOS PROPUESTOS

A modo de ejercicio propuesto, en este capítulo planteamos una actividad de recuperación del programa Mystat. Como cualquier otro programa informático es susceptible de un mal funcionamiento, ya sea a causa del propio sistema o por un error humano. Por esta razón planteamos esta actividad, que consiste en reinstalar en programa para corregir posibles errores.

Al reinstalar el programa no se eliminan ninguno de los ficheros existentes, ya sean del programa o ficheros de datos propios, tan solo se sobrescriben los existentes corrigiendo posibles errores.

Para ello, inicie el proceso de instalación del programa tal y como hemos explicado anteriormente, es recomendable que reinstale el programa siguiendo un proceso diferente. Es decir, que si inicialmente se instaló desde el **Explorador de Windows** (Figura 2.7), debería realizarlo desde el proceso **Ejecutar** (Figura 2.2) siguiendo los pasos explicados en apartados anteriores.

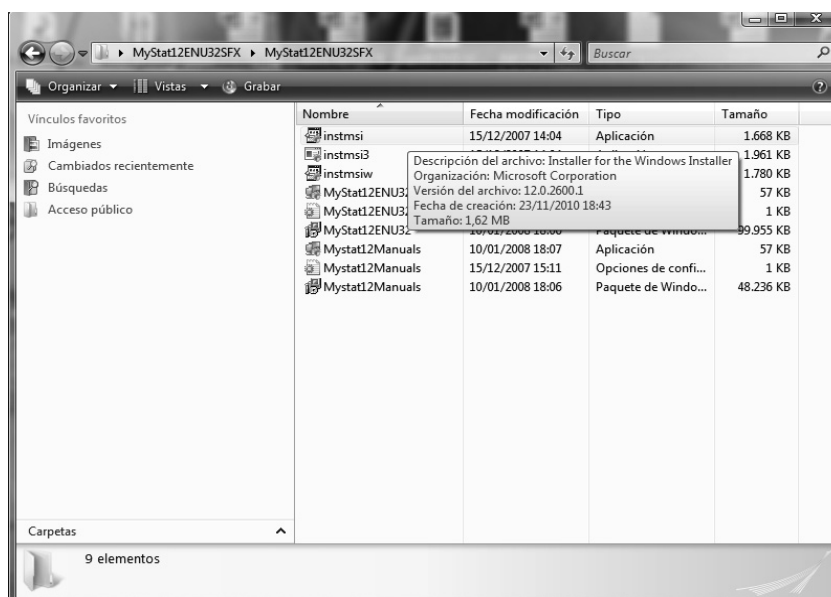


Figura 2.7. Instalación de Mystat desde el Explorador de Windows.

Recomendamos que cree una carpeta en la que guarde los ficheros creados con el programa y que si ha de eliminar o modificar algún fichero confirme que corresponde a uno creado por usted. La razón es que los ficheros de datos creados por Mystat tienen la extensión **.sys** y el sistema operativo Windows asocia esta extensión a ficheros del propio sistema existiendo muchos de éstos en diferentes carpetas que **NUNCA DEBEN SER ELIMINADOS**, ya que influirán en el correcto funcionamiento del propio sistema o de algún programa instalado.

Capítulo 3.

MANEJO DEL EDITOR DE DATOS DEL PROGRAMA MYSTAT 12

*José D. Cuesta Sáez de Tejada
Tomás Izquierdo Rus*

INTRODUCCIÓN

En este capítulo se tratarán todos los aspectos relacionados con la introducción de datos en el programa Mypstat, la codificación de variables, introducción y modificación de datos, la recodificación de variables, el almacenamiento y recuperación de ficheros, etc.

1. OBJETIVOS

Los objetivos que pretendemos con el presente capítulo son:

- Conocer la nomenclatura especial de las variables en el programa Mypstat.
- Conocer la diferencia entre variable cuantitativa y cualitativa y la codificación numérica o alfanumérica.
- Conocer el proceso de recodificación de datos en un fichero.
- Conocer la utilidad de transformar puntuaciones directas en ordinales.
- Aprender a encontrar y seleccionar casos que cumplan una o varias condiciones.
- Aprender a modificar el formato de los datos decimales.
- Dominar el proceso de guardado y recuperación de ficheros desde la hoja de datos.
- Valorar la utilidad del módulo Mypstat Worksheet dentro del programa estadístico Mypstat.

2. CONTENIDOS

Al ejecutar el icono del programa estadístico Mymstat 12, lo primero que aparece en pantalla es el editor de datos para introducir los datos (Figura 3.1), quedando en la parte superior todos los módulos del programa.

El editor de datos es una matriz de doble entrada; en la primera columna puede observarse una serie de números que indican el número de casos que contiene el fichero. En la primera fila aparecen las celdillas donde se irá introduciendo el nombre de las variables. Este programa al tratarse de una versión para uso educativo tiene una limitación impuesta en el número de variables en lo referente a las columnas (100), pero sin límite en el número de filas. Otras versiones similares de este programa, pero con carácter profesional, no tienen más limitación que la impuesta por el propio equipo informático.

Cuando Mymstat 12 se inicia, aparece una ventana principal (Startpage) desde la cual podremos abrir una base de datos ya existente, o bien crear una nueva. En esta nueva ventana, justo encima del logotipo de Mymstat 12 que aparece en fondo verde, observamos tres pestañas: *Startpage*, que es la ventana de inicio, *Untitled.syo*, es la ventana donde se mostrarán los resultados del análisis realizado, y *Untitled .syz* o editor de datos es una ventana muy parecida a una hoja Excel, en la que aparecen unas rejas donde estarán ubicados los datos (Figura 3.1).

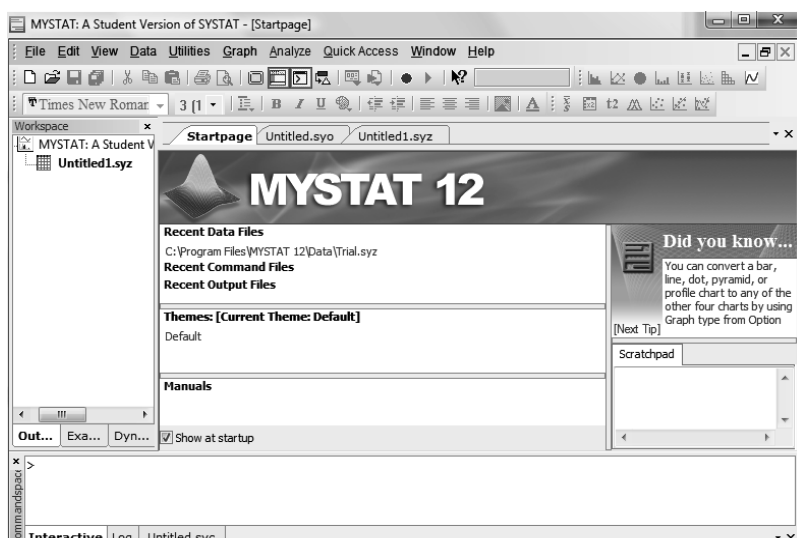


Figura 3.1. Editor de datos.

2.1. Asignar nombre a las variables

La primera tarea a realizar antes de comenzar a introducir datos será la de nombrar las variables e introducirlas en los espacios correspondientes de la primera fila.

Para iniciar el proceso de introducir el nombre de las variables debe hacer click en una pestaña situada en la parte inferior del editor de datos denominada “Variable”. Aparecerá un nuevo editor de datos y escribiremos el nombre de la variable en la primera casilla en blanco que aparece en dicho editor y que se corresponde con la primera fila, por tanto habrá tantas filas como variables tenga nuestra investigación. Una vez introducido el nombre, por defecto nos aparecerán rellenas el resto de casillas que conforman las características de las variables. A continuación, pasamos a describir cada una de estas características:

VARIABLE NAME: Nombre de la variable.

- El nombre de la variable debe ser único.
- Máximo de 8 caracteres.
- Los caracteres pueden ser letras, cifras, un punto o símbolos, _ * \$.
(todos los demás símbolos están prohibidos).
- Al final del nombre no puede colocarse un punto.
- No podemos dejar espacios vacíos, no son posibles.
- No hay diferencia entre mayúsculas y minúsculas en la denominación de una variable.
- Tratar de no escribir palabras acentuadas si es posible.

VARIABLE LABEL: Etiqueta o descripción de la variable.

VALUE LABELS: valores de la variable.

- p.e.: 1 = Hombre, 2 = Mujer.

TYPE: tipo de variable (numérica o cadena “string”). Se puede cambiar en la celda haciendo doble clic en la parte de la izquierda.

CATEGORICAL: Tenemos la opción “YES” o “NO”, al igual que la opción anterior, y nos permite indicar si la variable es alfanumérica o presenta categorías, es decir cuando la variable es cualitativa. Recordemos que estas variables son aquellas del tipo de respuesta “si” o “no” o bien de carácter ordinal (de acuerdo, en desacuerdo, indiferente, a veces, a menudo...)

CHARACTERS: Ancho de la columna en la vista de datos. Podemos aumentar o disminuir el valor, pero por defecto marcaremos “8”. Con este valor, serán visibles todos los caracteres de la columna en la vista de datos.

DECIMALS: Número de decimales.

DISPLAY: Esta opción es especialmente útil en el caso de variables cuantitativas y tenemos tres opciones: *Normal*, si la variable presenta valores numéricos normales, con decimales incluidos, *Exponential notation* (si se trata de valores numéricos exponenciales) y *Date and time*, se emplea cuando queremos expresar datos en formato de fecha hora. Cuando seleccionamos esta opción podemos observar como, al situar el cursor encima de la columna situada a la izquierda **DATA/TIME FORMAT**, nos ofrece la posibilidad de dar diferentes formatos a la fecha y hora.

Una vez que han sido nombradas todas las variables, ya puede empezar a introducir los datos de cada sujeto para cada variable. En el proceso de introducción de datos siempre es recomendable hacerlo por filas, es decir para cada sujeto debe escribirse un valor para todas las variables, ya que si se hace por columnas es posible que el programa se bloquee y se pierda el trabajo no guardado.

En caso de no querer asignarle valor a una variable debe colocarse un punto (.), este símbolo indicará al programa que el sujeto no ha asignado valor a dicha variable. Nunca es conveniente asignarle el valor cero (0) sobre todo a las variables numéricas, ya que esto puede generar conflictos en el sistema al realizar determinados cálculos.

Con el fin de evitar problemas, antes de proceder a la introducción de las variables, planifique cual va a ser el nombre con el que designará cada una de las variables y el orden de las mismas, ya que las modificaciones posteriores son muy costosas y en algunos casos imposible.

2.2. Vaciado de datos en MYSTAT 12

Una vez nombradas las variables, sitúe el cursor en la celdilla que se encuentra debajo de la primera variable (Figura 3.2). Esta posición corresponde al primer sujeto (recuerde haber numerado el instrumento empleado para recoger datos en función de los sujetos), verificando que los instrumentos cumplimentados por los participantes en el trabajo están correctamente cumplimentados y pueden ser vaciados en el programa estadístico, ya que una vez numerados pueden surgir problemas a la hora de introducir los datos porque no es conveniente dejar filas sin datos.

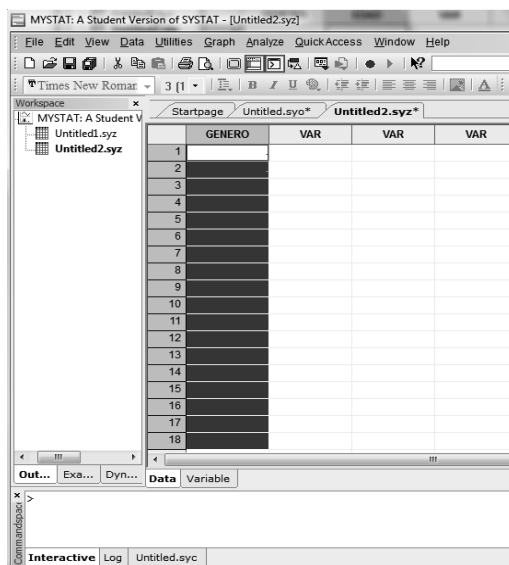


Figura 3.2. Introducción del primer dato en el editor de datos.

Debe tener en cuenta que las variables numéricas admitirán números y las alfanuméricas números y letras. Si comete un error en este punto recibirá el mensaje de advertencia que incluimos en la figura 3.3.

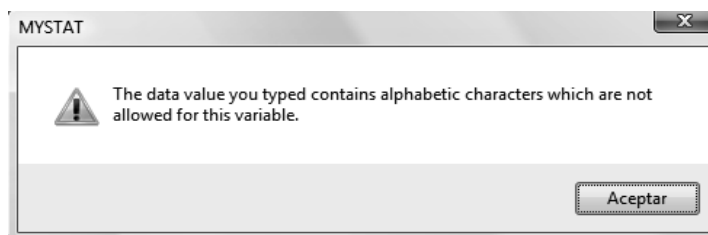


Figura 3.3. Error por no estar acorde el tipo de variable con los datos.

2.3. Modificar datos ya introducidos

Si ha cometido algún error al introducir el nombre de las variables o alguno de los datos, puede modificarlos situándose justamente encima de la celdilla a rectificar y haciendo clic con el botón del derecho del ratón seleccionamos la opción “Eliminar” (Figura 3.4) e introducimos el nuevo valor.

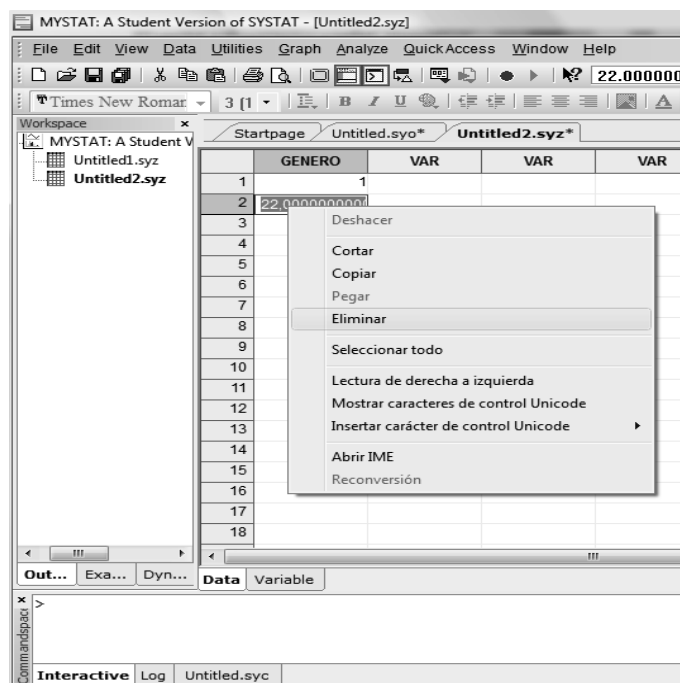


Figura 3.4. Modificación de valores de una variable en el editor de datos.

Es muy recomendable una planificación previa en papel de las variables antes de proceder a introducir las variables en el editor de datos de MYSTAT 12.

No olvide, sin embargo, que para cambiar el nombre de una variable siempre deberemos acceder a la hoja de datos correspondiente a la pestaña “variable” (Figura 3.5). Nunca podremos cambiar los nombres de las variables directamente desde el editor de datos principal.

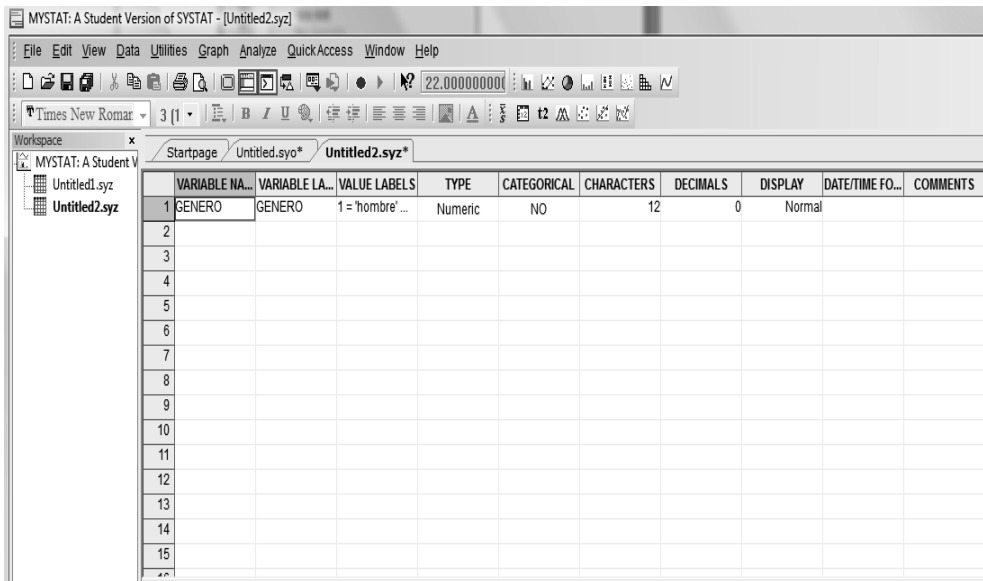


Figura 3.5. Cambiar nombre de variable desde la opción “Variable”.

Como norma general, cuando realice alguna modificación de los datos debe guardar el fichero antes de continuar trabajando, para evitar olvidos o errores al almacenar los ficheros. Más adelante explicamos el proceso de guardado y recuperación de ficheros con Mystat.

2.4. Encontrar un caso dentro del editor de datos

Desde la hoja de trabajo seleccione la opción EDIT/DATA/GO TO... (Figura 3.6) de la barra de menú de Mystat.

Esta opción permite encontrar el primer caso que cumple la condición que haya especificado en la hoja de datos (Figura 3.7).

En esta ventana de diálogo seleccione la variable, pulse el botón **Add** y a continuación seleccione el caso que desea encontrar. El programa le llevará directamente al caso seleccionado.

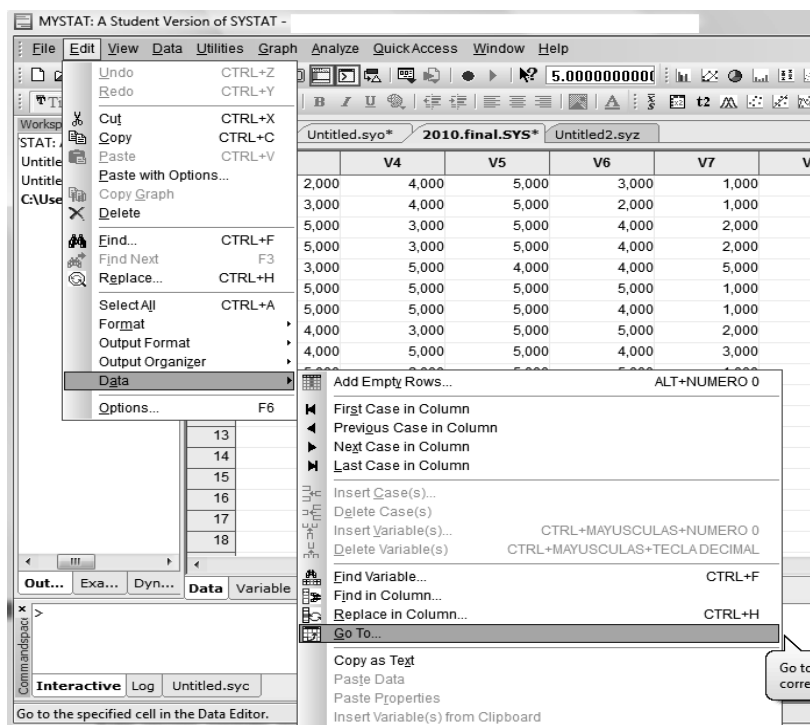


Figura 3.6. Opción Edit/Data/Go To.

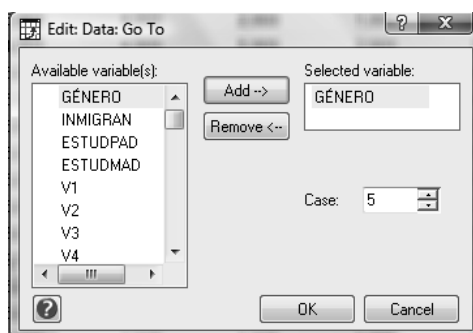


Figura 3.7. Ventana Editor Go To.

2.5. Crear nuevas variables a partir de existentes

Para acceder a esta opción de la hoja de datos realice la siguiente secuencia DATA/TRANSFORM/LET (Figura 3.8).

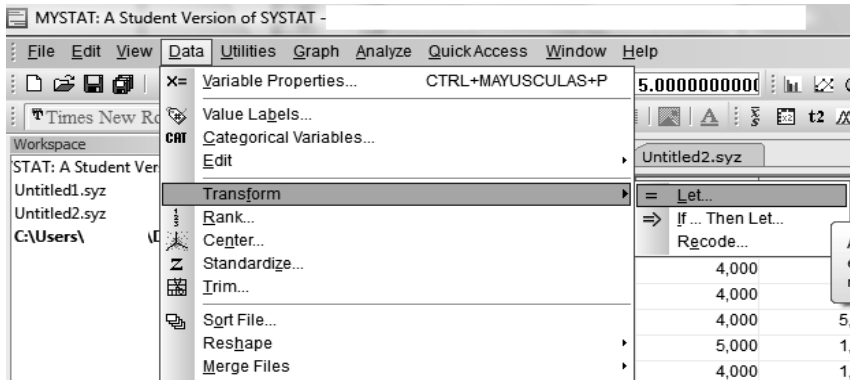


Figura 3.8. Opción Data/Transform/Let.

Desde la ventana de comandos de Data/Transform/Let (Figura 3.9) podrá crear una nueva variable que sea una operación algebraica de otra variable ya existente. Las variables, tanto las existentes como las que se crean, han de ser numéricas.

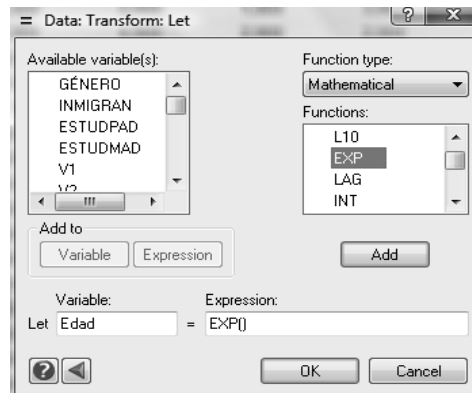


Figura 3.9. Data/Transform/Let.

El proceso se inicia escribiendo en el recuadro que se encuentra a la derecha de la instrucción **Let** el nombre que le vamos a asignar a la nueva variable o seleccionando una variable ya existente. A continuación, en el rectángulo que aparece a la derecha "**expression**", introduzca la operación algebraica a realizar.

Esta operación se selecciona entre las opciones que aparecen en la parte superior derecha de la ventana, se marca y se pulsa el botón **Add** (dentro de este recuadro encontrará una selección de todas las operaciones que puede realizar, como por ejemplo la raíz cuadrada, coseno, etc.) finalmente se escribe, entre paréntesis, el nombre de la variable original. Para que la operación planteada se efectúe y se cree la nueva variable con el cálculo que le hemos indicado debe pulsarse el botón **Ok**.

2.6. Recodificar una variable existente

Para acceder a este comando de la hoja de datos debe realizar la siguiente secuencia DATA/TRANSFORM/RECODE (Figura 3.10).

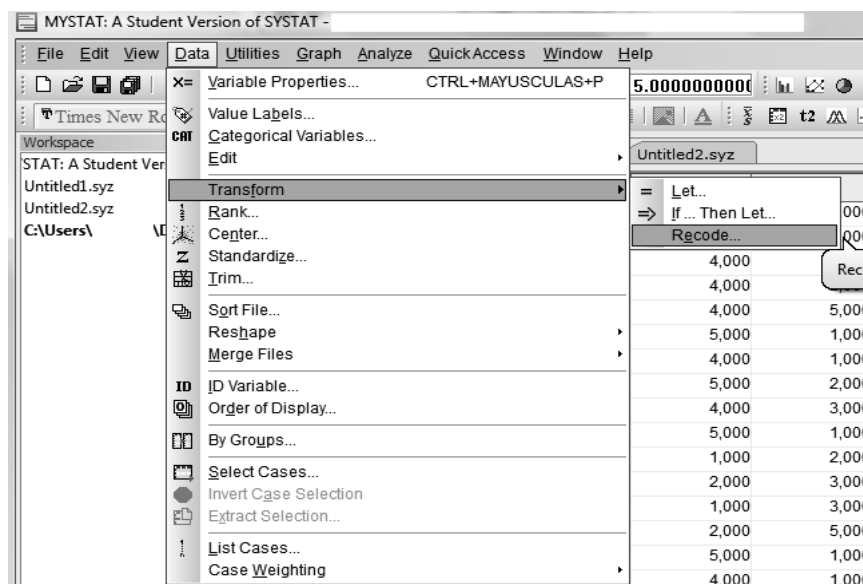


Figura 3.10. Opción Data/Transform/Recode.

La recodificación (recode) es una opción que encontrará en la barra de menú Transform. Esta función permite cambiar los códigos asignados a los valores de una variable. Su utilidad radica en agrupar, en un único valor, diferentes valores de una variable y para realizar cambios de una variable con formato de cadena a una con formato numérico.

Como se observa en la figura 3.11, se introduce el nombre de la variable que contiene los datos originales (GENERO) en la casilla Recode From y el nombre de la nueva variable en la casilla Recode To. En la parte inferior se aplica la condición de recodificación, empleando para ello la casilla Old Value para indicar el antiguo valor y la casilla New value contendrá el nuevo valor que sustituirá al anterior. En este ejemplo que se presenta, pasamos de una variable numérica GÉNERO a una variable alfanumérica GÉNERO\$, aunque también puede hacerse a la inversa. Hemos de prestar especial atención en este tipo de transformaciones a las comillas (" ") que hay que asignar a las categorías de las variables alfanuméricas, ya que si no se colocan NO darán un error visible, aunque la nueva variable no será creada. Es usual que se olviden con bastante frecuencia, para evitar este error tan sólo hay que revisar la correcta realización de la recodificación después de haber pulsado el botón **Ok**.

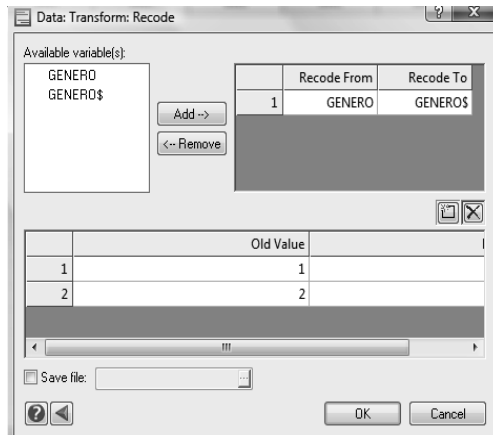


Figura 3.11. Ventana Editor / Recode.

2.7. Seleccionar un caso o casos condicionados

El acceso a este comando se realiza siguiendo la secuencia DATA/SELECT CASES (Figura 3.12).

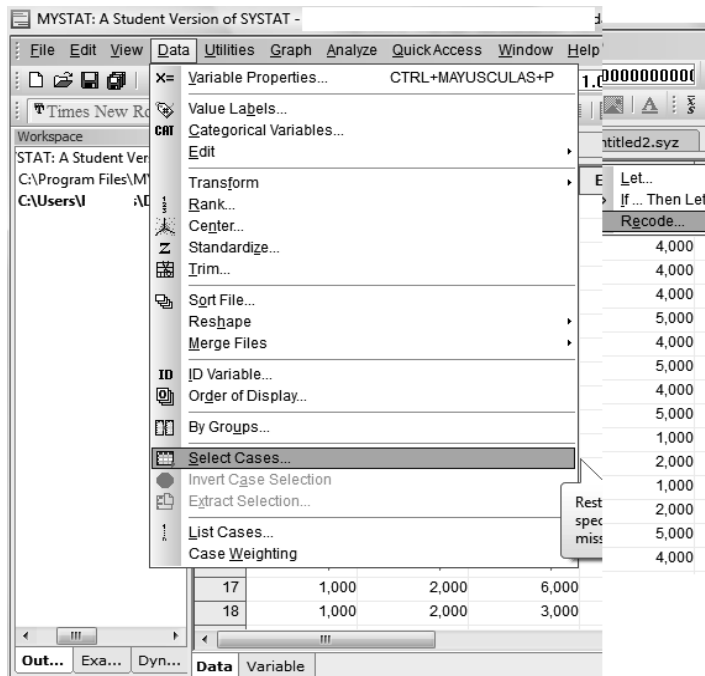


Figura 3.12. Opción Data/Select cases.

Esta opción permite seleccionar los casos que cumplen una determinada condición. En la ventana de diálogo que aparece (Figura 3.13) se introduce la condición que debe cumplir la variable o variables de un sujeto para ser seleccionado/a.

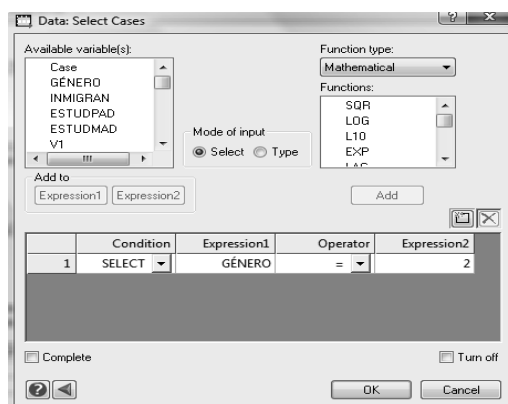


Figura 3.13. Ventana Data / Select case.

En la figura 3.13 ponemos como ejemplo que el programa seleccione a todos los sujetos que cumplan la condición de tener asignado como género el valor 2, es decir que según nuestra codificación sean mujeres.

Al pulsar el botón **Ok**, obtenemos como resultado la hoja de datos con una serie de sujetos a los que se les ha asignado un punto negro (Figura 3.14), este es el modo que tiene el programa de indicar que dichos sujetos cumplen la condición o condiciones establecidas en la figura 3.13.

Startpage Untitled.syo* Untitled1.syz*			
	GENERO	VAR	VAR
1	1		
2	1		
3	1		
✓ 4	2		
5	2		
6	2		
7	2		
8	1		
9	1		
10	1		
11	2		
12	2		
13	2		
14	2		
15	2		
16	1		
17	1		
18	1		

Figura 3.14. Resultado de la selección de sujetos efectuada.

Una vez realizada cualquier selección debe proceder con cuidado, ya que si a continuación realiza algún tipo de análisis éste se efectuará únicamente con los sujetos seleccionados.

Como indicábamos más arriba es posible realizar la selección de un grupo de sujetos en función de más de una condición, para ello debe unir las distintas condiciones con la conjunción AND (Figura 3.15).

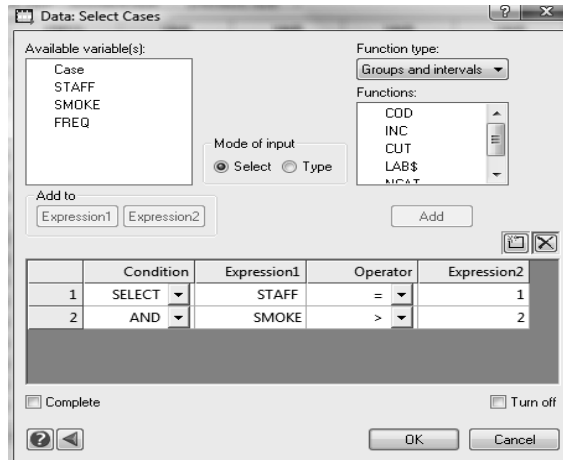


Figura 3.15. Ventana Data / Select Cases con dos condiciones (AND).

El proceso de redacción de esta condición es sumamente sencillo, en el recuadro que se encuentra a la derecha de **Select** escriba el nombre de la primera variable seguido del signo = y el criterio de selección que desee, escribimos la conjunción **AND** (que significa Y) o la conjunción **OR** (que significa O), a continuación escribimos la segunda variable seguida del signo > y el segundo criterio de selección que deseemos.

Si se emplea la conjunción AND, para que un caso sea seleccionado tiene que cumplir simultáneamente las dos condiciones establecidas. En cambio si empleamos la conjunción OR, los sujetos serán seleccionados cuando cumplan, al menos, una de las condiciones.

Con la secuencia que hemos puesto a modo de ejemplo en la figura 3.15 se seleccionarán los casos existentes en el fichero que cumplan ambas condiciones, que pertenezcan a la FACULTAD (codificada 2) y que tengan una EDAD igual a 23 años. Para establecer las condiciones deben emplearse los signos igual (=), mayor (>), menor (<), mayor o igual (>=), menor o igual (<=) y distinto de (<>).

Otra utilidad de este comando (Select cases) es el poder fragmentar el archivo original en tantos archivos como condiciones queramos establecer. Es decir, que tras una selección como la que acabamos de realizar, es posible crear un nuevo fichero que contendrá únicamente los sujetos que cumplen la condición o condiciones que hayamos establecido.

Esta opción se activa mediante la opción EDITOR / EXTRACT SELECTION (Figura 3.16) y una vez seleccionada nos pedirá la ubicación y el nombre del nuevo fichero que queremos crear con los datos extraídos (Figura 3.17).

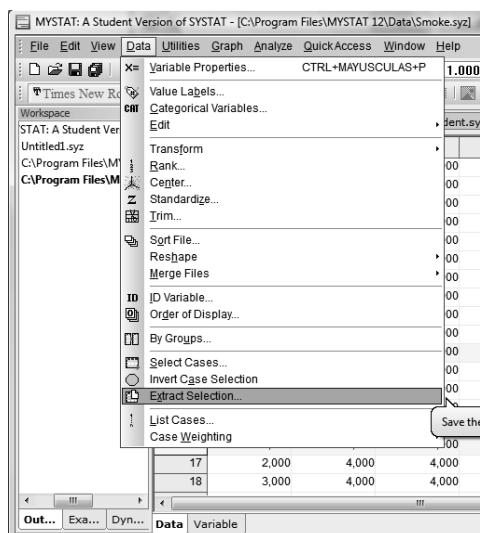


Figura 3.16. Opción Data / Extract Selection.

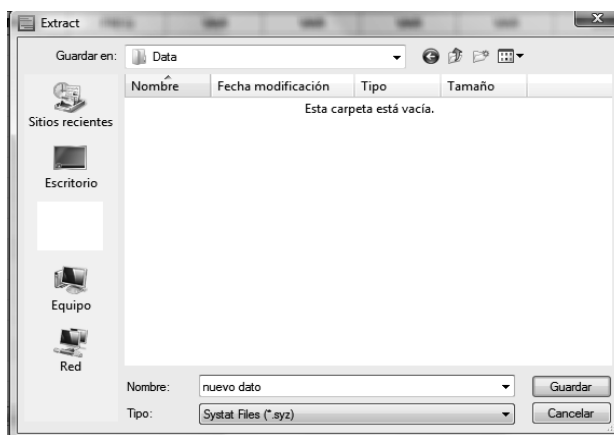


Figura 3.17. Creación de un nuevo fichero a partir de una selección.

2.8. Transformación de variables continuas en ordinales

Esta opción está disponible desde el comando DATA/RANK (Figura 3.18) y permite transformar las variables continuas en variables ordinales.

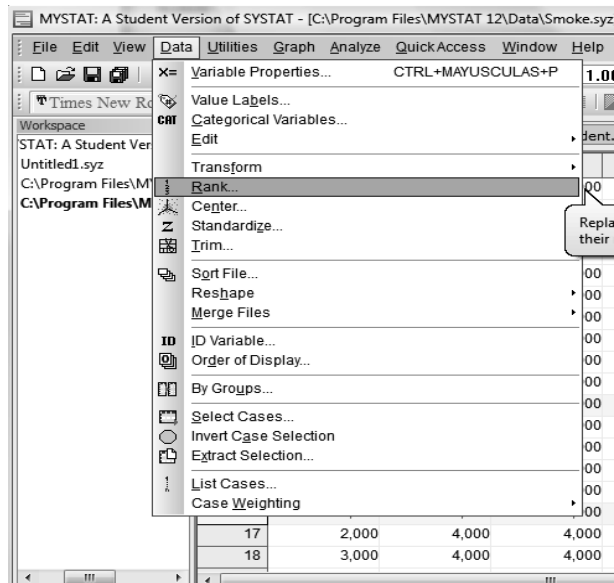


Figura 3.18. Opción Data / Rank.

En la figura 3.19 se han seleccionado las variables continuas VARIABLE1 y está pendiente de selección VARIABLE2, tras seleccionarla y pulsar el botón **Ok** aparecerá una nueva ventana de diálogo que le invitará a guardar en un nuevo fichero las variables transformadas, con el fin de que los datos originales no se vean alterados.

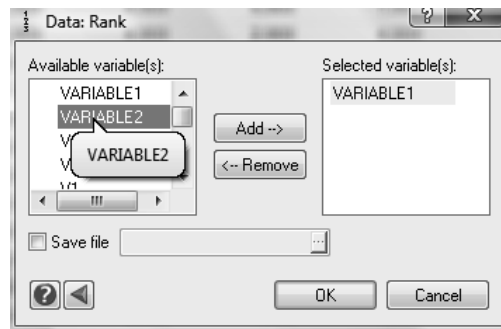


Figura 3.19. Ventana selección variables continuas Rank.

Una de las utilidades más comunes de este comando radica en poder calcular el **coeficiente de correlación de Spearman** a partir de variables continuas, ya que este coeficiente únicamente puede calcularse con variables continuas y transformar “a mano” una serie elevada de datos puede llegar a ser prácticamente imposible.

2.9. Ordenar un fichero en función de una variable

A este módulo se accede desde la secuencia DATA/SORT FILE (Figura 3.20) y permite ordenar el contenido de un fichero en función de una variable que necesariamente ha de ser numérica.

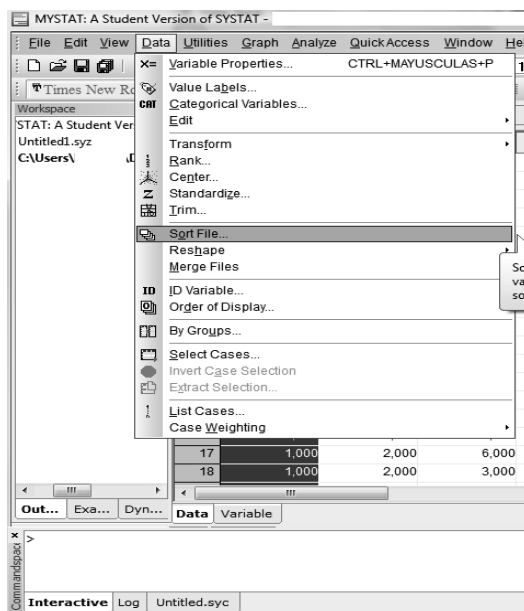


Figura 3.20. Opción Data / Sort File.

Al seleccionar esta opción, accedemos a la ventana Sort que nos permite seleccionar la variable o variables que serán el criterio de ordenación del fichero original (Figura 3.21). Como ejemplo hemos seleccionado la variable VAR2 que servirá de criterio para ordenar el fichero de trabajo.

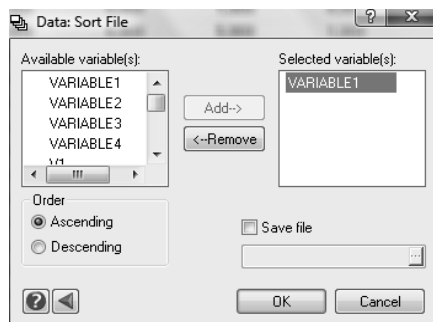


Figura 3.21. Ventana Sort File.

Esta opción permite seleccionar el número de decimales con el que presentar los datos en el editor de datos. Como puede observar en la figura 3.24, el número de decimales oscila entre cero y catorce.

CATEGORICAL	CHARACTERS	DECIMALS	DISPLAY
NO	12	0	Normal
NO	12	0	Normal
NO	12	1	Normal
NO	12	2	Normal
NO	12	3	Normal
NO	12	4	Normal
NO	12	5	Normal
NO	12	6	Normal
NO	12	7	Normal
NO	12	8	Normal
NO	12	9	Normal
NO	12	10	Normal
NO	12	11	Normal
NO	12	12	Normal
NO	12	13	Normal
NO	12	14	Normal
NO	12	3	Normal
NO	12	3	Normal

Figura 3.24. Ventana Variable / Decimals.

2.11. Salvar un fichero de datos por primera vez

Una vez que se han introducido unos pocos datos debe proceder a **salvar el fichero** por primera vez, de tal manera que se le asigne un nombre y una ruta para poder guardarlo. Después de esta primera vez, debe ir guardando periódicamente para evitar pérdidas de datos por razones imprevisibles como puede ser un corte de corriente eléctrica, etc.

Para guardar un fichero por primera vez debe proceder del siguiente modo:

- 1°. En editor de datos pulse sobre la opción **File** de la barra de menú y posteriormente en **Save Active As** (Figura 3.25).

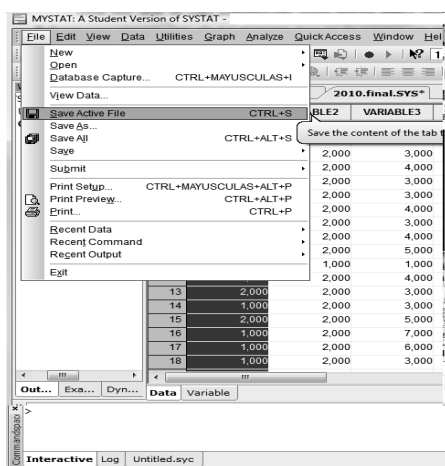


Figura 3.25. Opción File / Save Active As.

- 2°. En la ventana de diálogo que aparece (Figura 3.26) elija el nombre, la unidad y la ruta del fichero; pique posteriormente en el botón **Save as** para confirmar la acción. Los ficheros de datos tienen extensión .SYS; no olvides añadir esta extensión al nombre del fichero que contiene los datos que has introducido. Recuerde que la extensión máxima del nombre del fichero debe ser 8 caracteres.

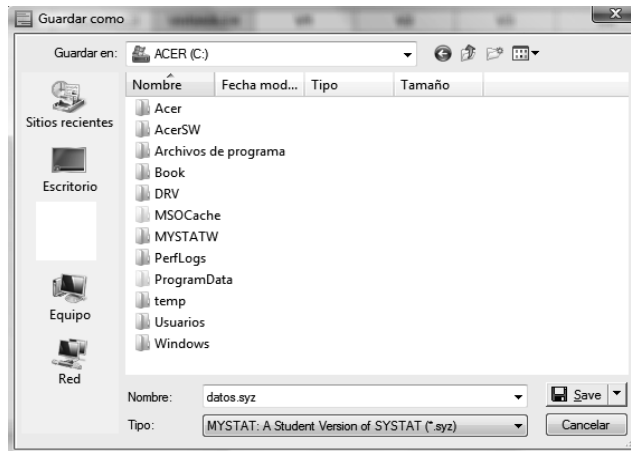


Figura 3.26. Ventana Save a File.

2.12. Salvar un fichero previamente guardado

Para salvar un fichero ya almacenado y que ha sido modificado, se realiza pican-do en el icono con aspecto de disco en la barra de menú y posteriormente en **Save As** (Figura 3.27). Este modo de salvar hace innecesario añadir de nuevo el nombre, unidad y ruta del fichero, ya que el fichero se almacenará con el nombre y en la ubicación en la que se encontraba previamente.

Si desea modificar el nombre o la ubicación del fichero, o simplemente duplicar el fichero en diferentes directorios, tendrá que guardarlo con la opción **Save As** que explicábamos en el apartado anterior.

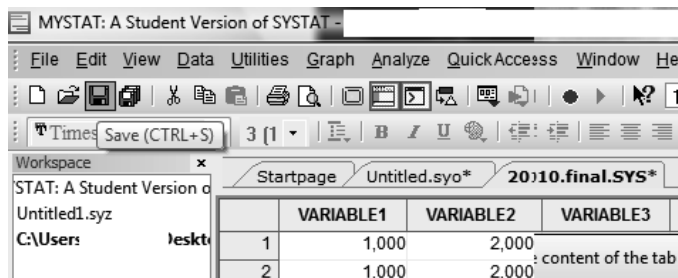


Figura 3.27. Ventana Menú / Save.

2.13. Abrir un fichero previamente almacenado

La recuperación de un fichero se puede hacer desde el menú de inicio. En ambos casos debe picar en **File** de la barra de menú, **Open** y luego **Data**. A continuación indique la unidad, la ruta y el nombre del fichero a recuperar. Para confirmar pulse **Abrir** (Figura 3.28).

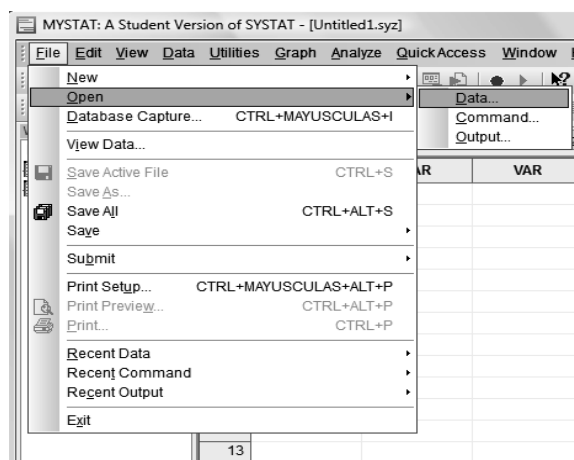


Figura 3.28. Ventana Open Data.

Esta acción abrirá el editor de datos con toda la información que contiene el fichero. Para comenzar a analizar los datos emplearemos las opciones situadas en el menú principal de la parte superior del editor de datos.

Hemos de hacer un comentario más respecto a este apartado, ya que solemos abrir los ficheros de trabajo de cualquier otro programa simplemente repicando sobre el nombre del fichero. Automáticamente el programa asociado a él se encarga de abrir el fichero con los datos. En caso de MYSTAT esto no sucede, jamás repicando sobre un fichero de datos de Mystat se abrirá el programa permitiendo acceder a los datos, ya que como hemos comentado en otras ocasiones la extensión de algunos ficheros de datos (.sys) es idéntica a la empleada por el sistema operativo Windows, en sus distintas versiones, en los ficheros que guardan la configuración del sistema. Por ello siempre deben abrirse los ficheros una vez que el programa ha sido ejecutado y siguiendo los pasos que se explican en este apartado.

2.14. Guardar los análisis de datos en MAIN

Cuando haya realizado los análisis de datos que necesite, debe o bien tomar nota en papel de los resultados, imprimirlos o almacenarlos en un fichero. Estas dos últimas operaciones debe hacerlas desde **editor de datos**, pincando en **File / Print** si desea imprimir los análisis por impresora o bien **File / Save Active File**, si desea

almacenar el fichero. Para ello debe escribir el nombre del fichero con la extensión .syo y después pulse el botón **Save**. Por ejemplo, si desea almacenar el fichero de resultados con el nombre **prueba**, el nombre completo deberá ser **prueba.syo** (Figura 3.29) y pulse el botón **Guardar** para confirmar el guardado.



Figura 3.29. Ventana Guardar como...

2.15. Guardar los ficheros de representaciones gráficas

En el capítulo dedicado a las representaciones gráficas se ha explicado ampliamente el proceso para almacenar cada uno de los tipos de representaciones gráficas que permite realizar el programa estadístico Mystat.

2.16. Configuración del editor de datos

Desde el **Editor de datos**, y utilizando el menú principal, accedemos a cada uno de los módulos que componen **Mystat 12** y, además, toda una serie de opciones que permiten un manejo más cómodo y útil del programa. Entre las más utilizadas se encuentran las siguientes:

2.16.1. Selección del tipo de letra

Desde el editor de datos realice la siguiente secuencia EDIT/ FORMAT/ FONT, tal y como se observa en la figura 3.30.

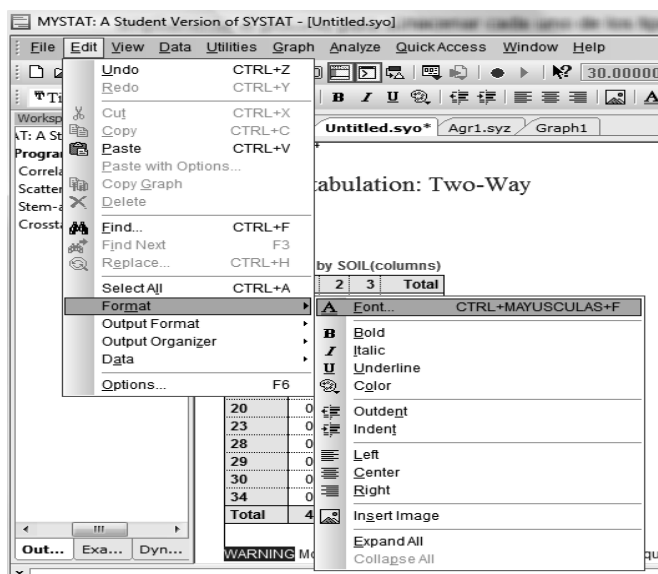


Figura 3.30. Opción Edit/format/font.

Esta opción permite seleccionar el tipo de letra con la que desea presentar las salidas (resultados) en pantalla. La figura 3.31 presenta las diversas opciones **Fuente** (tipo de letra), **Estilo de fuente** (normal, negrita, cursiva, etc), **Tamaño**, **Efectos**, **Muestra** y **Color de la fuente**.

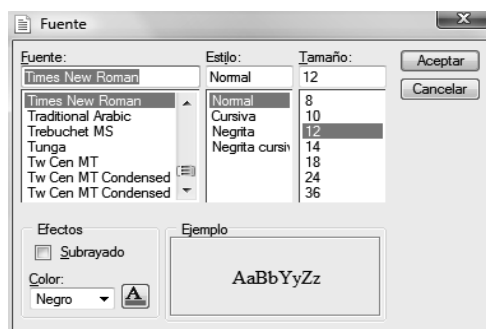


Figura 3.31. Ventana de selección de Fuente.

2.16.2. Seleccionar el número de decimales y organizar los resultados

Desde el editor de datos, a través del menú Edit, accedemos a las opciones Output format y Output Organizer, tal y como se observa en la figura 3.32.

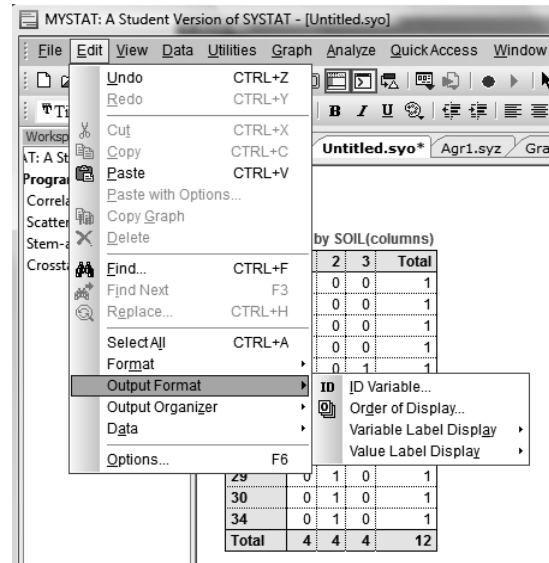


Figura 3.32. Opción Data/Output Formats y Output Organizer.

A través de la opción **Output format** podemos elegir el número de decimales en los resultados de los análisis. Esta misma operación, se puede realizar (como se comentó anteriormente) a través de la opción Variable / Decimals (Figura 3.23). Resulta recomendable utilizar al menos dos decimales.

Un poco más abajo aparece un encabezado con **Output Organizer** con tres opciones. Seleccionando cada una de ellas podemos crear contenidos, agregar anotaciones o bien desplazarnos por el editor de salida sin necesidad de buscar uno a uno los gráficos o los resultados obtenidos que queramos utilizar en ese momento. A través de la opción **Collapse** se crean unas carpetas en el margen izquierdo. Puede abrir y cerrar los iconos de las carpetas haciendo clic en el signo “+” y “-” respectivamente (Figura 3.33).

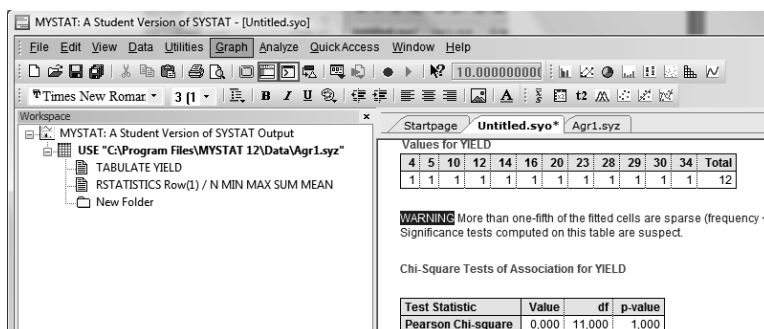


Figura 3.33. Ventana Output Organizer.

2.16.3. Activar la hoja de datos desde MAIN

Desde el menú principal realice la siguiente secuencia VIEW/ DATA EDITOR, tal y como se observa en la figura 3.34.

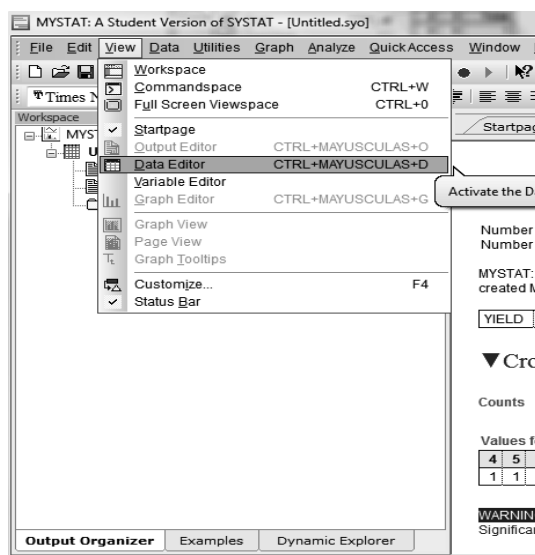


Figura 3.34. Opción Windows/Worksheet.

Este comando permite retomar la hoja de datos principal. También podemos acceder directamente a la hoja de datos haciendo clic sobre la pestaña que contiene el “nombre.syz” de nuestra base de datos.

2.17. Consideraciones a la hora de incluir variables para realizar diferentes análisis estadísticos

PRIMERA observación importante. Cuando los datos que vaya a introducir en la hoja de datos se refieran a la comparación de resultados entre un grupo experimental y un grupo de control debe crear una variable que puede denominar GRUPO, teniendo en cuenta que los datos que incluya en esa columna diferenciarán la pertenencia al grupo de control o al grupo experimental; asignando el código 1 a todos los sujetos del grupo de control y código 2 a todos los sujetos del grupo experimental. La siguiente variable que defina, por ejemplo RENDI si pretende analizar el rendimiento de un grupo de sujetos, será la que recoja los resultados correspondientes al grupo de control y al grupo experimental.

SEGUNDA observación importante. En el caso del ANOVA no olvide que hay que introducir una variable que diferencie la pertenencia a cada uno de los grupos que han intervenido en cada uno de los niveles de la variable independiente o

grupos. Esta diferencia puede realizarla utilizando también códigos distintos. Por ejemplo, la variable puede denominarse FACTOR y estar codificada con 1, 2, 3, etc., según haya intervenido en cada uno de los niveles de la variable independiente, niveles del experimento o grupos.

Capítulo 4.

REPRESENTACIONES GRÁFICAS DE DATOS CUANTITATIVOS

*José D. Cuesta Sáez de Tejada
Fuensanta Monroy Hernández*

INTRODUCCIÓN

La representación gráfica de los datos de investigación tiene la ventaja de mostrar dichos datos de tal modo que con un simple vistazo pueda observarse su estructura, distribución, tendencia, etc. *Se dice que una imagen vale por mil palabras.* En el caso de la investigación esta frase se hace aún más real pues, ¿cómo pueden ser interpretados mil números por muy ordenados que se encuentren en una tabla, si no es a través de una representación visual de la globalidad de los mismos?

La representación gráfica de datos ha adquirido una gran importancia en la investigación educativa a partir de la perspectiva del Análisis Exploratorio de Datos (AED) desarrollada por Tukey entre otros.

MYSTAT dispone de varios módulos para representar gráficamente los resultados de los análisis de una o más variables. Estas representaciones pueden realizarse tanto con variables cualitativas como variables cuantitativas, la única condición es que las categorías de estas variables deben estar codificadas numéricamente.

1. OBJETIVOS

Los objetivos que pretendemos con el presente capítulo son:

- Conocer la utilidad de los diagramas de cajas como representación gráfica de datos.
- Conocer la utilidad de los histogramas de frecuencia como representación gráfica de datos.

- Conocer la utilidad de los diagramas de puntos como representación gráfica de datos.
- Conocer la utilidad de los diagramas de tallos y hojas como representación gráfica de datos.
- Realizar todas las representaciones gráficas de datos que permite el programa Mystal e interpretar sus resultados.
- Valorar la utilidad de las representaciones gráficas de datos como modo de sintetizar y exponer visualmente grandes cantidades de datos.

2. CONTENIDOS

Antes de abordar cada una de las representaciones gráficas anteriormente citadas debe conocerse el funcionamiento del icono gráfico del programa y en cuya ventana aparecerán todas estas representaciones gráficas, nos referimos a la utilidad **Graph**.

En el módulo principal del programa de Mystal 12 podemos encontrar una opción que permite acceder a la ejecución de diferentes tipos de gráficos en función de las necesidades del estudiante. A esta opción se accede picando en **Graph** de la barra de opciones del menú principal (Figura 4.1).

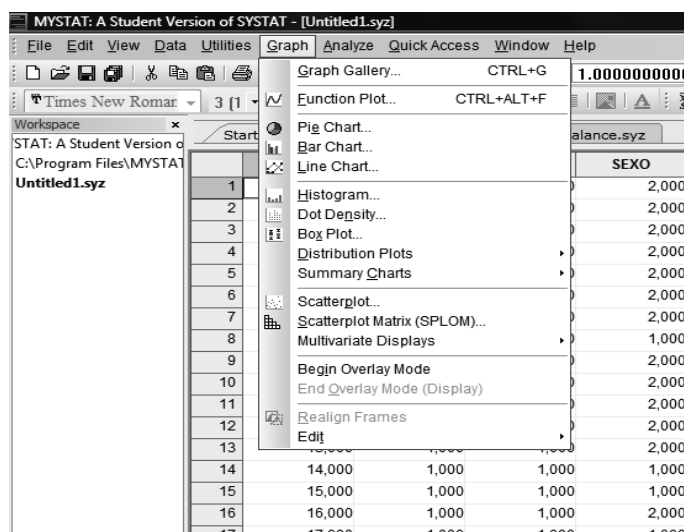


Figura 4.1. Configuración del módulo gráfico de Mystal.

Como se observa en la figura 4.1, al marcar la opción **Graph** aparece un menú desplegable que muestra todas las opciones posibles de gráficos. De tales opciones le sugerimos que marque la opción **Graph Gallery**. Marcando esta opción podrá seleccionar y visualizar directamente el tipo de gráfico que deseamos obtener sin necesidad de acceder al resto de opciones que aparecen en la parte inferior del mismo (Figura 4.2).

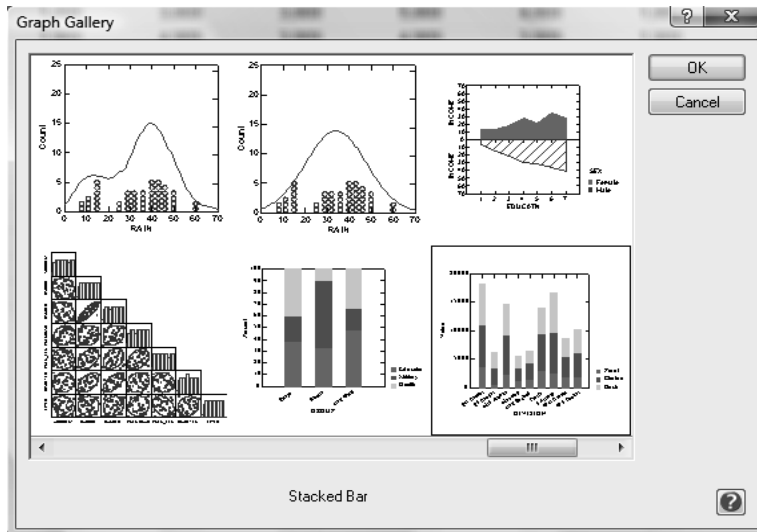


Figura 4.2. Opción de Graph Gallery para visualizar diferentes gráficos.

2.1. Histogramas de frecuencias

El histograma es una representación gráfica que permite representar medidas de intervalo para una única variable. Esta representación se realiza desde el módulo Hist (Figura 4.3).

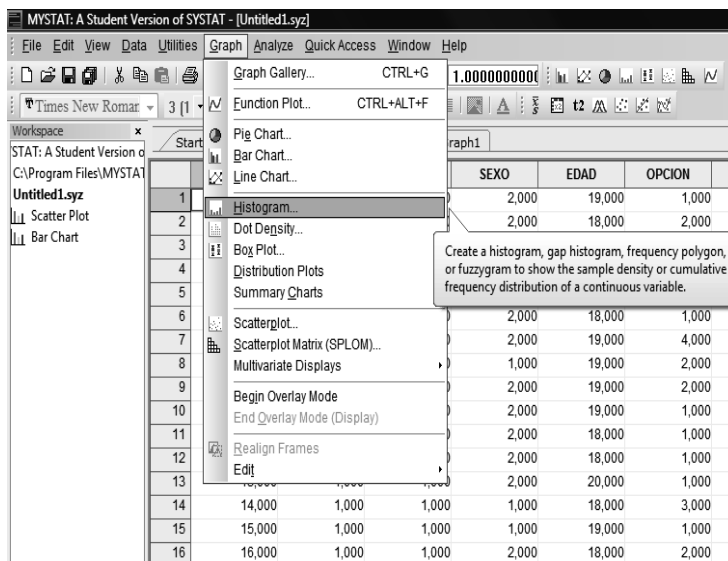


Figura 4.3. Módulo Histograma.

Pasos que debe seguir para realizar un histograma:

1°. Desde el menú principal, cargue el fichero de trabajo picando en File/Open/Data.

Una vez comprobado que éste es el fichero con el que desea trabajar y que su contenido es correcto, puede proceder a cerrarlo, modificarlo, minimizarlo, etc., para continuar con la realización del histograma.

2°. Pique sobre la opción **Histogram...** de la barra que se encuentra en el menú principal y aparece la siguiente ventana (Figura 4.4)

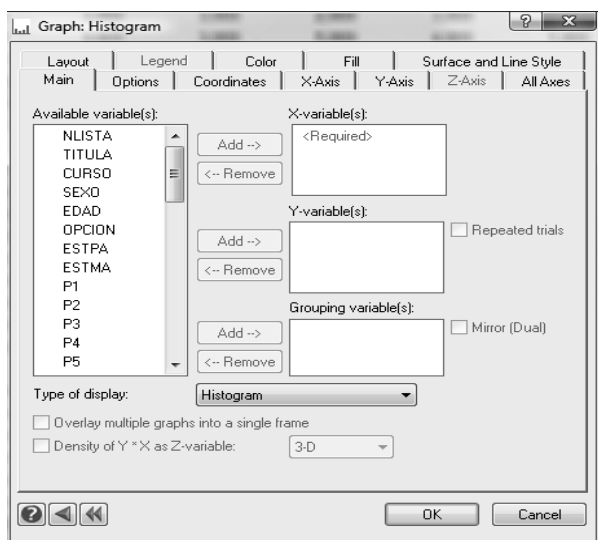


Figura 4.4. Ventana Histograma.

3°. Seleccione la variable o variables que desea representar con un histograma. Esta acción se realiza picando una a una las variables, y a continuación pulsando en el botón **Add**. Cada variable se irá incorporando al rectángulo de la derecha. En caso de equivocación al seleccionar alguna variable, proceda del modo inverso, márkela en el rectángulo de la derecha y posteriormente pulse en el botón **Remove**. Una vez seleccionadas las variables pulse el botón **Ok** para obtener el histograma/as.

4°. Grabar, imprimir o cancelar los resultados obtenidos. Una vez dibujado el histograma de cada una de las variables seleccionadas, si es necesario almacenar los resultados en un fichero tendrá que hacerlo desde la barra principal de la ventana de gráficos, para ello pulse las siguientes opciones **File/Save As...**, (Figura 4.5) y en la ventana de diálogo que aparece (Figura 4.6) seleccione la ubicación donde desea guardar el archivo y el nombre del fichero, sin olvidar añadirle la extensión **.BMP**. En la figura 4.6, puede observar un ejemplo, en el que el fichero se grabará en el Escritorio de Windows, y con el nombre **graph.bmp**.

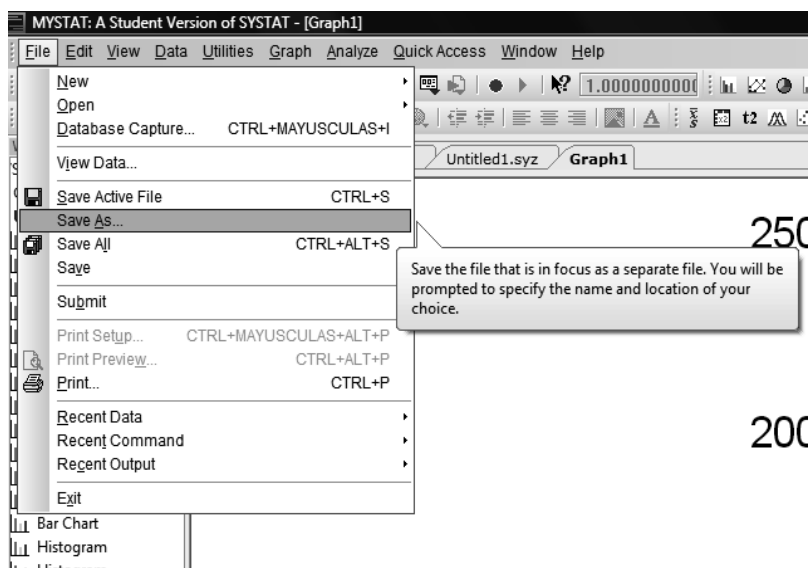


Figura 4.5. Localización de la opción Save As...

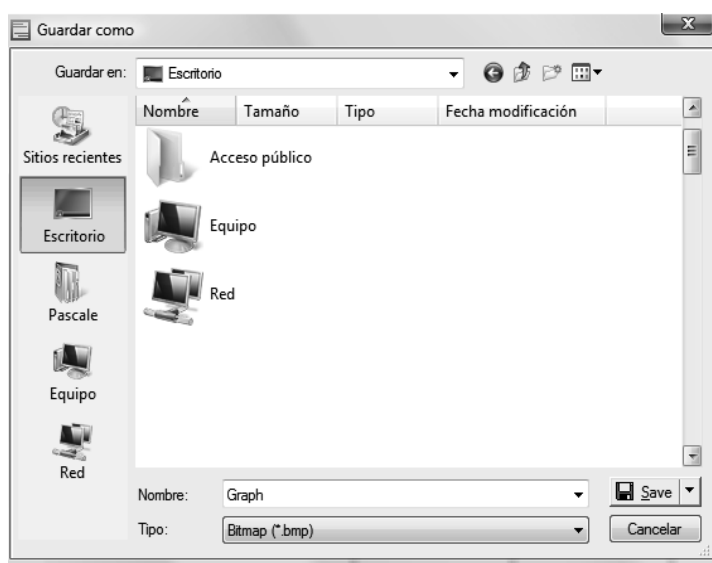


Figura 4.6. Ventana Save Graph As.

El gráfico se puede imprimir directamente pulsando sobre la opción **Print** del menú **File** (Figura 4.7), no siendo esta opción incompatible con la grabación del fichero.

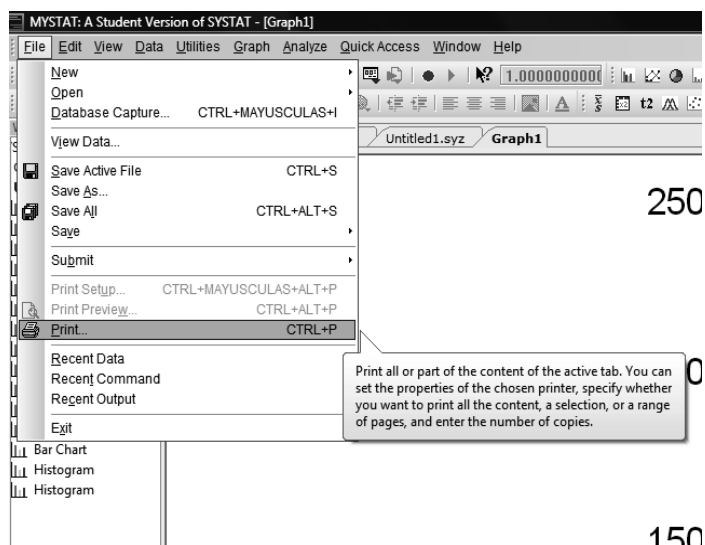


Figura 4.7. Opción Print (imprimir) de Mynstat.

2.2. Gráfico de caja (Box Plot)

Este módulo gráfico muestra una variable en forma de caja. Dicha caja representa el 50% central de los datos de la variable, coincidiendo sus límites extremos con los cuartiles 1 y 3. La línea central de la caja representa a la mediana (Figura 4.8). Los datos extremos se denominan valores fuera del intervalo u **outlier**. Estos casos extremos pueden ser sujetos interesantes de estudio cuando se trabaja con datos de investigación, ya que pueden representar casos especiales sobre los que diseñar y aplicar acciones de mejora, tratamientos experimentales, programas de intervención, etc.

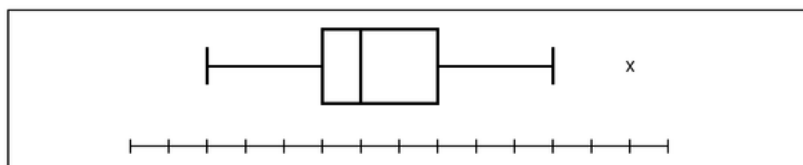


Figura 4.8. Representación gráfica de una variable.

Los pasos a seguir para la realización de esta representación gráfica de los datos de una variable son similares a los explicados en el histograma.

1°. Desde el menú principal, cargue el fichero de trabajo picando en File/Open/Data.

Una vez comprobado que éste es el fichero con el que desea trabajar y que su contenido es correcto, puede proceder a cerrarlo, modificarlo, minimizarlo, etc., para continuar con la realización del diagrama de caja.

2º. Pique en el botón del módulo **Box Plot** (Figura 4.9). Observará una ventana de diálogo que le permite seleccionar las variables que desea representar mediante este tipo de gráfico (Figura 4.10).

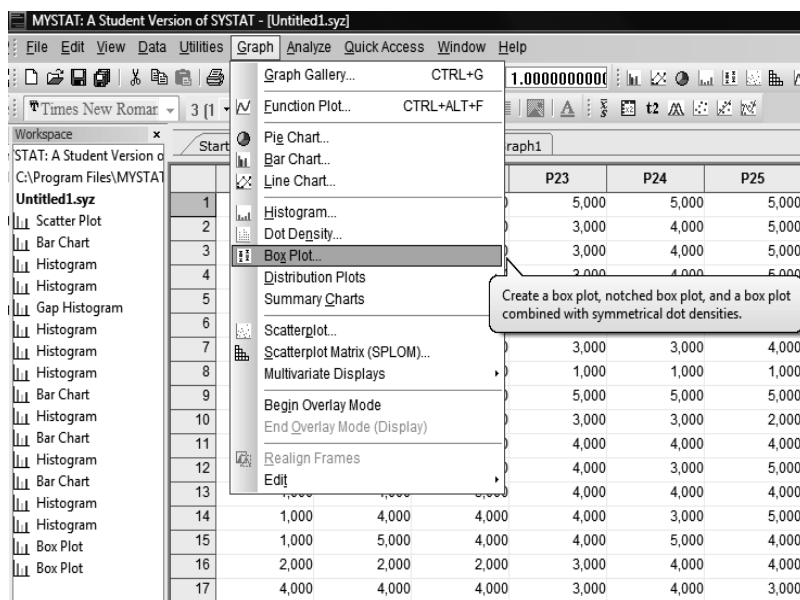


Figura 4.9. Módulo Diagrama de caja (Box Plot).

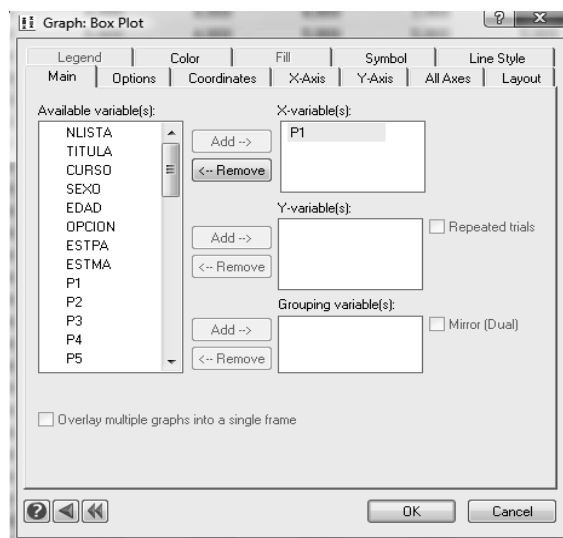


Figura 4.10. Ventana Box Plot.

3°. Seleccione las variables. Esta acción se realiza picando una a una las variables, y a continuación picando en el botón **Add**. Cada variable se irá incorporando al rectángulo de la derecha. En caso de necesitar corregir una selección realizada erróneamente, debe marcarla en el rectángulo de la derecha y posteriormente picar el botón **Remove**. Una vez seleccionadas las variables pulse el botón **Ok** para obtener los diagramas de caja.

4°. Grabar, imprimir o cancelar los resultados obtenidos. Una vez dibujado el histograma de cada una de las variables seleccionadas, si es necesario almacenar los resultados en un fichero tendrá que hacerlo desde la barra principal de la ventana de gráficos, para ello pulse las siguientes opciones **File/Save As...**, (Figura 4.5) y en la ventana de diálogo que aparece (Figura 4.6) seleccione la ubicación donde desea guardar el archivo y el nombre del fichero, sin olvidar añadirle la extensión **.BMP**. En la figura 4.6, puede observar un ejemplo, en el que el fichero se grabará en el Escritorio de Windows, y con el nombre **graph.bmp**.

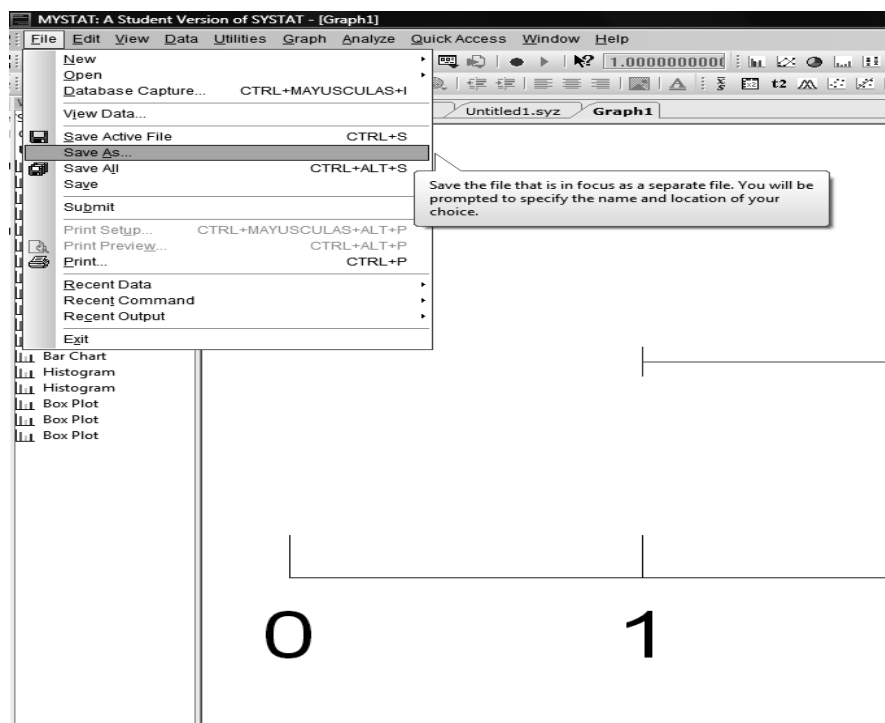


Figura 4.11. Localización de la opción *Save As...*

El gráfico se puede imprimir directamente pulsando sobre la opción **Print** del menú **File** (Figura 4.12), no siendo esta opción incompatible con la grabación del fichero.

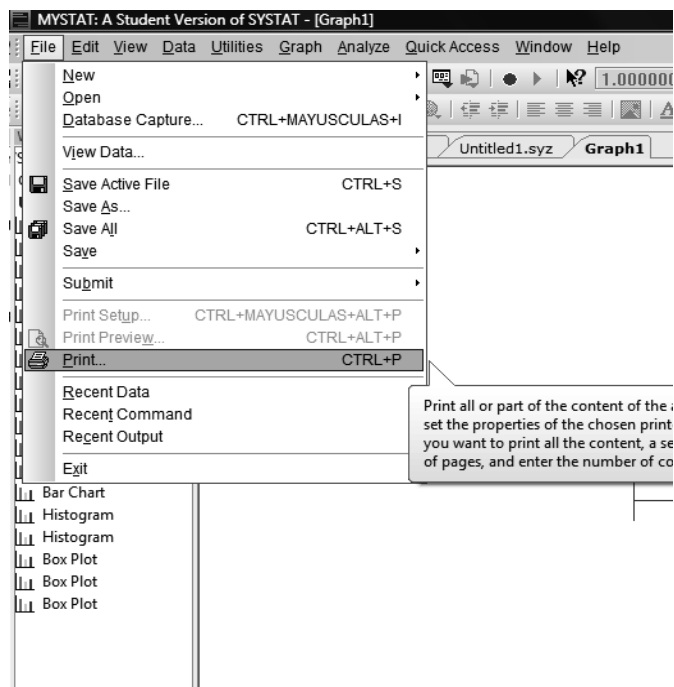


Figura 4.12. Opción Print (imprimir) de Mystat.

2.3. Diagrama de tallos y hojas (stem and leaf)

El diagrama de tallos y hojas es una representación gráfica similar al histograma, con la salvedad de que los datos se representan sobre el eje Y en lugar del eje X empleado por el histograma.

El **tallo** corresponde a todos los dígitos de la serie de datos que se encuentran a la izquierda del dígito que hace referencia a las unidades y las **hojas** hacen referencia al dígito restante, es decir al de las unidades. Es posible que con un ejemplo se comprenda mejor, imaginemos que se dispone de las siguientes puntuaciones de una variable 12, 13, 15, 17, 20, 22, 26, 26, 27, 29, 29, 31, 32, 33, 34, 34, 34, ..., el diagrama de tallos y hojas que puede realizarse tiene una estructura en la que el **1, 2 y 3** sería el tallo y los dígitos que se encuentran a la derecha del tallo indican los casos y número de casos que se encuentran entre un tallo y otro, como si de un intervalo se tratase.

```

1 2357
2 0266799
3 123444
....
    
```

Los pasos a seguir son similares a los explicados en las representaciones gráficas anteriores, el histograma y el diagrama de caja.

1°. Desde el menú principal, cargue el fichero de trabajo picando en File/Open/Edit.

Una vez comprobado que éste es el fichero con el que desea trabajar y que su contenido es correcto, puede proceder a cerrarlo, modificarlo, minimizarlo, etc., para continuar con la realización del diagrama de tallos y hojas.

2°. Tras recuperar el fichero con el que desea trabajar, seleccione el módulo **Stem and Leaf** (Figura 4.13). Observará una ventana de diálogo como la que sigue (Figura 4.14).

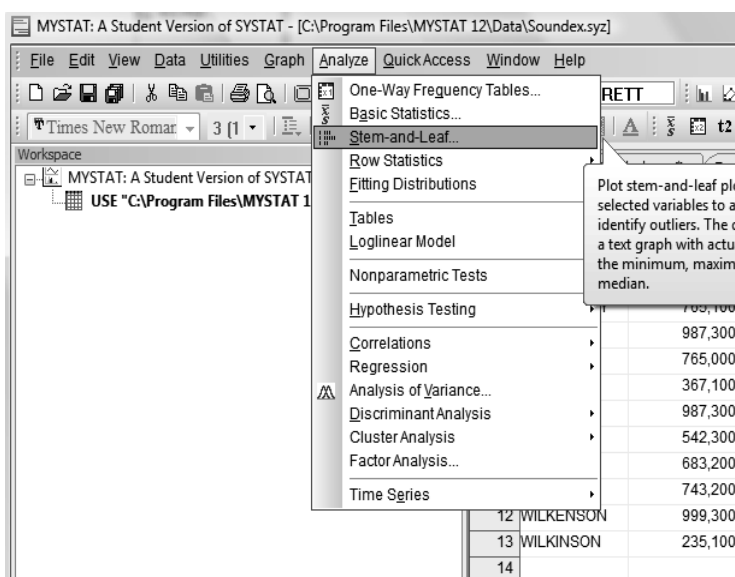


Figura 4.13. Módulo Stem and leaf.

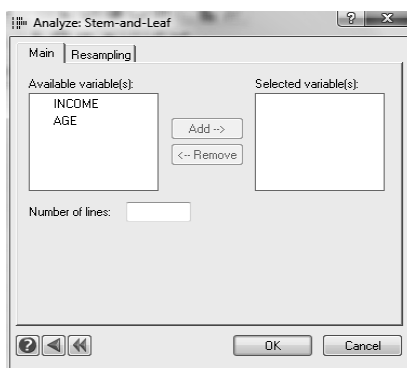


Figura 4.14. Ventana Stem and leaf Plots.

3°. Seleccione las variables. Esta acción se realiza picando una a una las variables, y a continuación picando en el botón **Add**. Cada variable se irá incorporando al rectángulo de la derecha. En caso de necesitar corregir una selección realizada erróneamente, debe marcarla en el rectángulo de la derecha y posteriormente picar el botón **Remove**. Una vez seleccionadas las variables pulse el botón **Ok** para obtener los diagramas de tallos y hojas (Figura 4.15).

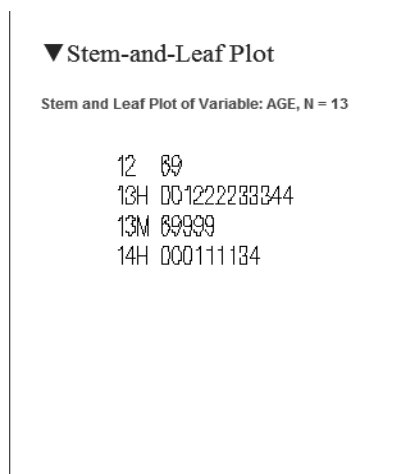


Figura 4.15. Resultado Stem and leaf Plot.

4°. Grabar, imprimir o cancelar los resultados obtenidos. Una vez dibujado el diagrama de tallos y hojas de cada una de las variables seleccionadas, si es necesario almacenar los resultados en un fichero tendrá que hacerlo desde la barra principal de la ventana de gráficos, para ello pulse las siguientes opciones **File/Save Active File**, y en la ventana de diálogo que aparece seleccione la unidad de disco, el directorio, y el nombre del fichero (Figura 4.16).

Observe que en el ejemplo empleado para presentar la figura 4.15 algún tallo puede no ser único, es decir que puede aparecer repetido, como por ejemplo el 13, esto es debido a la ubicación de la mediana (en este caso se encuentra en el valor 135), por esta razón el programa Mystat realiza esta distribución gráfica incluyéndose un tallo repetido. La configuración de la opción Lines cuando se seleccionan las variables a representar, también condicionará el resultado de la representación, es conveniente que se realicen varias aproximaciones hasta que nos quede un gráfico lo suficientemente claro.

En la figura 4.16, puede observar un ejemplo con el que se almacena un diagrama de tallos y hojas en la unidad C, en el subdirectorio Data y con el nombre graph.syo.

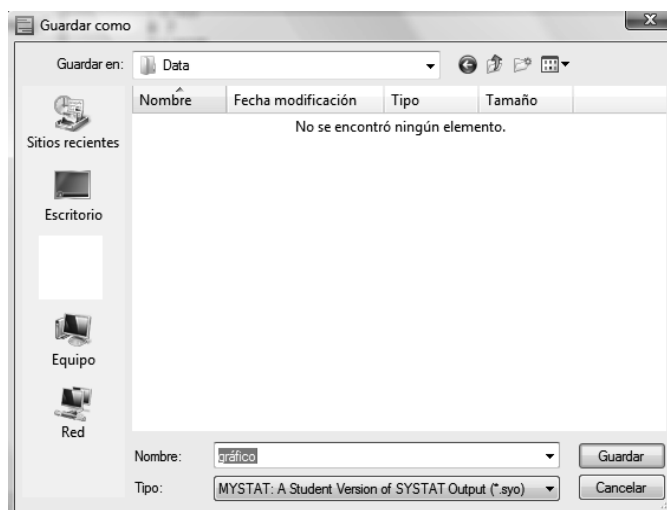


Figura 4.16. Ventana Save Graph As.

El gráfico se puede imprimir directamente pulsando sobre la opción **Print** del menú **File** (Figura 4.17), no siendo esta opción incompatible con la grabación del fichero.

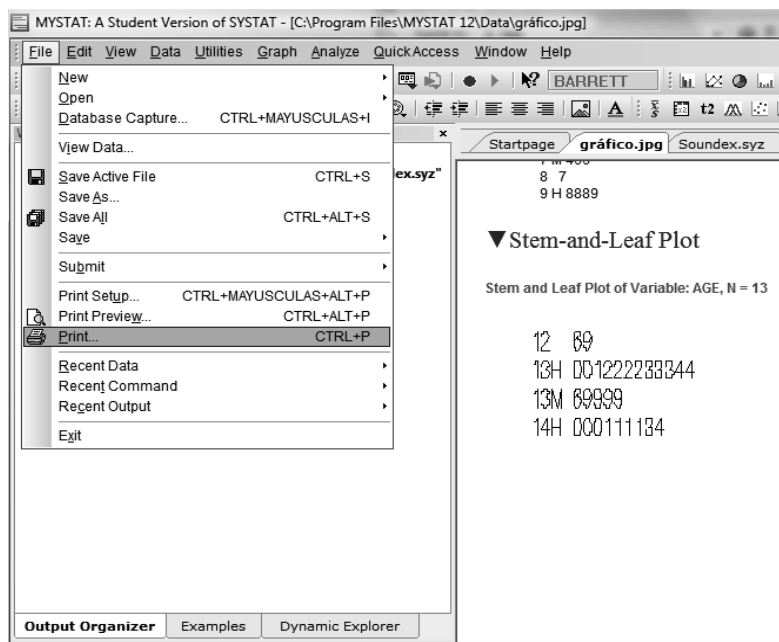


Figura 4.17. Opción Print (imprimir) de Mystat Graph.

2.4. Diagrama de dispersión o diagrama de puntos

Los diagramas de dispersión representan de un modo gráfico la relación existente entre dos variables. Estos diagramas suelen ser el mejor complemento visual del cálculo numérico de la regresión y de la correlación entre dos variables.

Los pasos a seguir para realizar un diagrama de dispersión de dos variables son los siguientes:

1º. Desde menú principal, cargue el fichero de trabajo picando en File/Open/Data.

Una vez comprobado que éste es el fichero con el que desea trabajar y que su contenido es correcto, puede proceder a cerrarlo, modificarlo, minimizarlo, etc., para continuar con la realización del diagrama de dispersión.

2º. Tras recuperar el fichero con el que desea trabajar, pique en el botón del módulo **Plot** (Figura 4.18). Observará una ventana de diálogo como la que sigue (Figura 4.19).

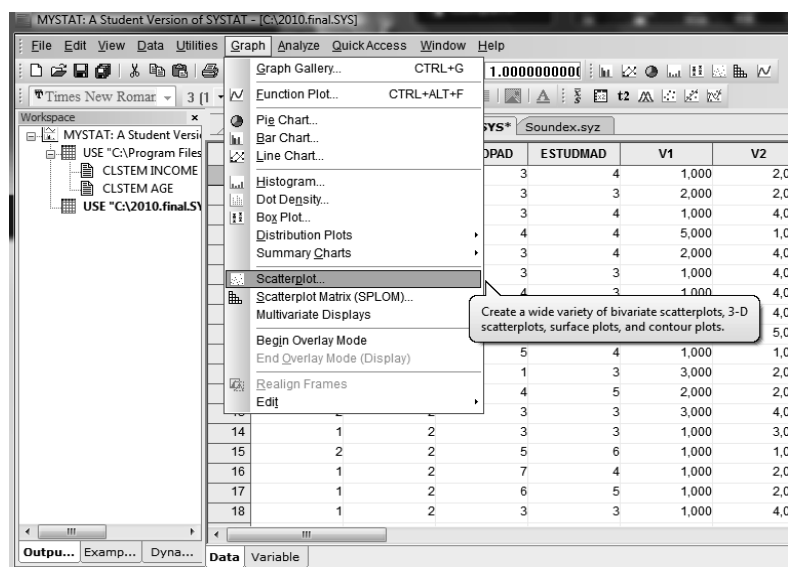


Figura 4.18. Módulo Scatterplot-Diagrama de dispersión.

3º. Seleccione las variables. Esta acción se realiza seleccionando una variable y asociándola a cada uno de los dos ejes (X e Y) picando sobre el botón que hace referencia a cada uno de dichos ejes.

En caso de necesitar corregir una selección realizada erróneamente, debe marcarla en el rectángulo correspondiente de la derecha y picando sobre el botón **Remove** (ya sea el del eje X o el Y). Una vez seleccionadas las variables pulse el botón **Ok** para obtener el diagrama de dispersión de las dos variables seleccionadas.

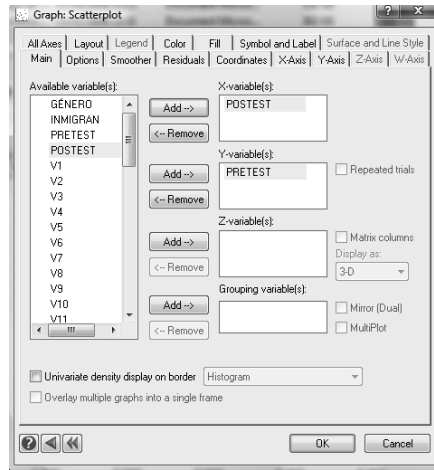


Figura 4.19. Ventana Scatterplot.

Si selecciona más de una variable para el eje X, el programa realizará tantos diagramas como variables haya incluido en el rectángulo destinado a las variables X. En la variable Y no es posible introducir más de una variable cada vez, teniendo que volver a acceder al módulo **Scatterplot** si necesitamos modificar dicha variable.

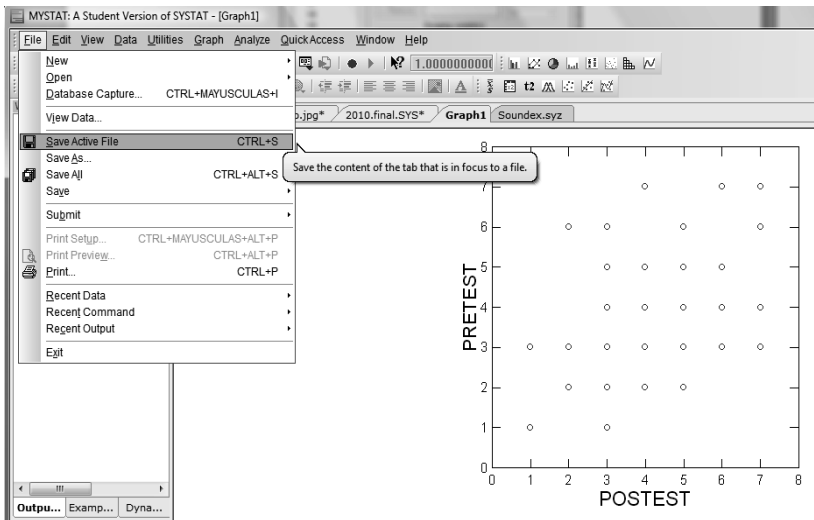


Figura 4.20. Localización de la opción Save Active File.

4°. Grabar, imprimir o cancelar los resultados obtenidos. Una vez dibujado el diagrama de dispersión de cada una de las variables seleccionadas, si es necesario almacenar los resultados en un fichero tendrá que hacerlo desde la barra principal

de la ventana de gráficos, para ello pulse las siguientes opciones **File/Save Active File**, (Figura 4.20) y en la ventana de diálogo que aparece (Figura 4.21) seleccione la unidad de disco (**Drives**), el directorio (**Directories**), y el nombre del fichero, sin olvidar añadirle la extensión **.BMP**.

En la figura 4.21, puede observar un ejemplo con el que se almacena un diagrama de dispersión en la unidad C, en el subdirectorio EQUIPO, y con el nombre **graph.bmp**.

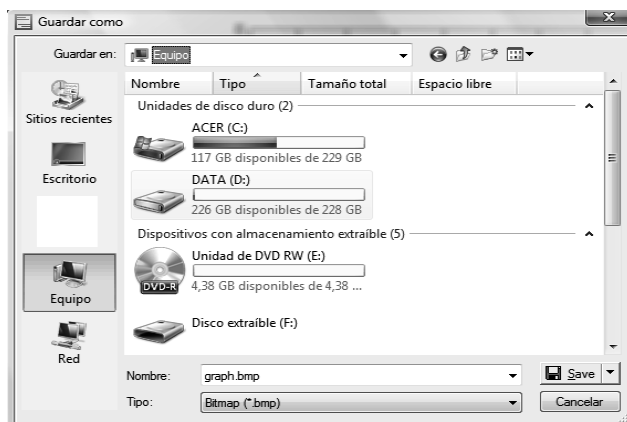


Figura 4.21. Ventana Save Graph As.

El gráfico se puede imprimir directamente pulsando sobre la opción **Print** del menú **File** (Figura 4.22) no siendo esta opción incompatible con la grabación del fichero.

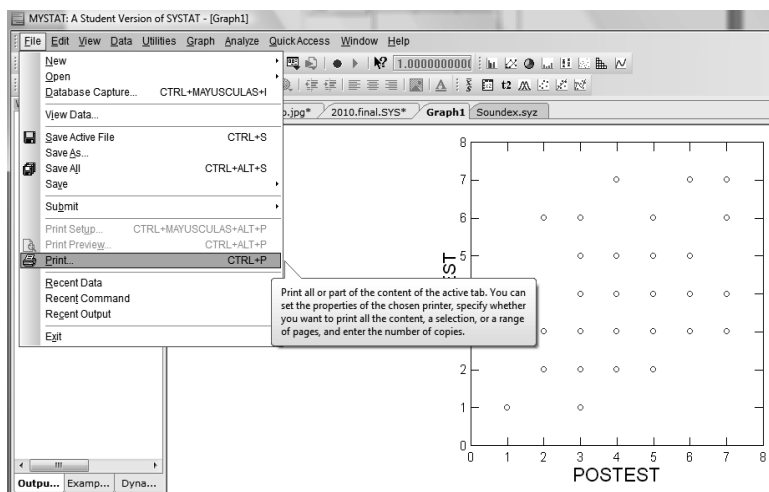


Figura 4.22. Opción Print (imprimir).

3. APLICACIÓN PRÁCTICA CON ORDENADOR

3.1. Histogramas de frecuencias

Representar gráficamente los siguientes datos referidos a una variable X: 4,13,9,2,8,5,5,15,31,6,4,8,8,15,10,22,6,8,2,11,4,6.

1°. Desde el menú principal, cargue el fichero de trabajo picando en File/Open/Data.

Una vez comprobado que éste es el fichero con el que desea trabajar y que su contenido es correcto, puede proceder a cerrarlo, modificarlo, minimizarlo, etc., para continuar con la realización del histograma.

2°. Pique sobre el icono del módulo **Histogram** de la barra que se encuentra en el menú principal y aparece la siguiente ventana (Figura 4.23)

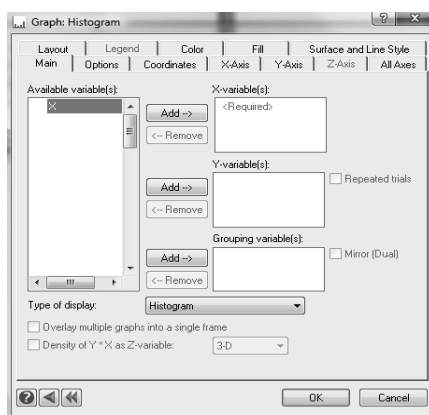


Figura 4.23. Ventana Histogram.

3°. Seleccione la variable y a continuación pulsando en el botón **Add**.

4°. En la figura 4.24, puede observar los resultados de nuestro ejemplo,

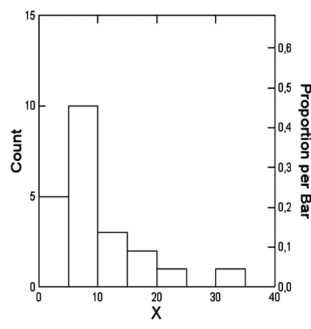


Figura 4.24. Histograma.

3.2. Gráfico de caja (Box Plot)

1°. Desde el menú principal, cargue el fichero de trabajo picando en File/Open/Data.

2°. Pique en el botón del módulo **Box Plot** (Figura 4.9). Observará una ventana de diálogo que le permite seleccionar las variables que desea representar mediante este tipo de gráfico (Figura 4.25).

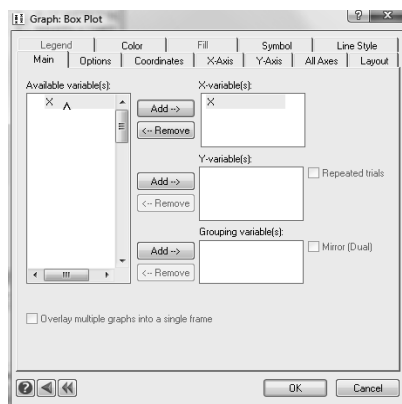


Figura 4.25. Ventana Box Plot.

3°. Seleccione la variable y a continuación pulsando en el botón **Add**.

4°. En la figura 4.26, puede observar los resultados de nuestro ejemplo,

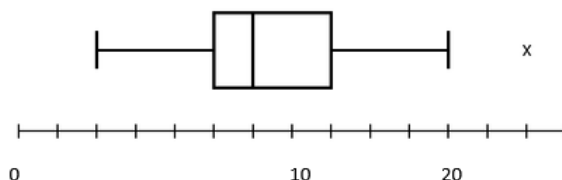


Figura 4.26. Box Plot.

En esta representación podemos observar claramente los outliers, valores como el 22 y el 31.

3.3. Diagrama de tallos y hojas. (Stem and leaf)

1°. Desde el menú principal, cargue el fichero de trabajo picando en File/Open/Data.

2°. Tras recuperar el fichero con el que desea trabajar, pique en el botón del módulo **Stem and Leaf** y observará una ventana de diálogo como la que sigue (Figura 4.27).

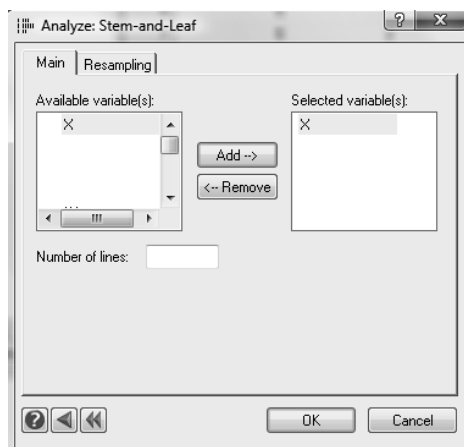


Figura 4.27. Ventana Stem and leaf Plots.

- 3°. Seleccione la variable y a continuación pulsando en el botón **Add**.
- 4°. En la figura 4.28, puede observar los resultados de nuestro ejemplo,

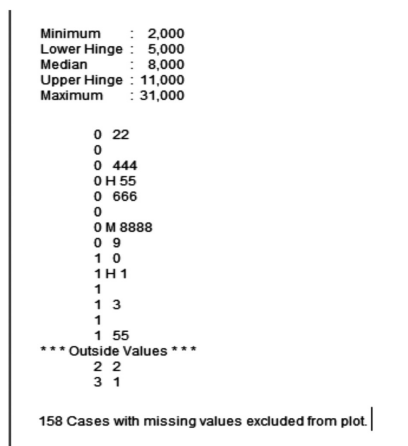


Figura 4.28. Diagrama Box Plot.

3.4. Diagramas de dispersión o diagramas de puntos

Debemos recordar que para realizar un diagrama de dispersión es necesario tener dos variables, por tanto, a la variable X de nuestro ejemplo podemos añadir una variable Y:

X: 4,13,9,2,8,5,5,15,31,6,4,8,8,15,10,22,6,8,2,11,4,6.

Y: 8,15,12,14,15,14,12,20,21,22,23,4,14,16,17,9,9,8,7,5,21,22

1°. Desde el menú principal cargue el fichero de trabajo picando en File/Open/Data.

2°. Tras recuperar el fichero con el que desea trabajar, pique en el botón del módulo **Scatterplot** (Figura 4.29).

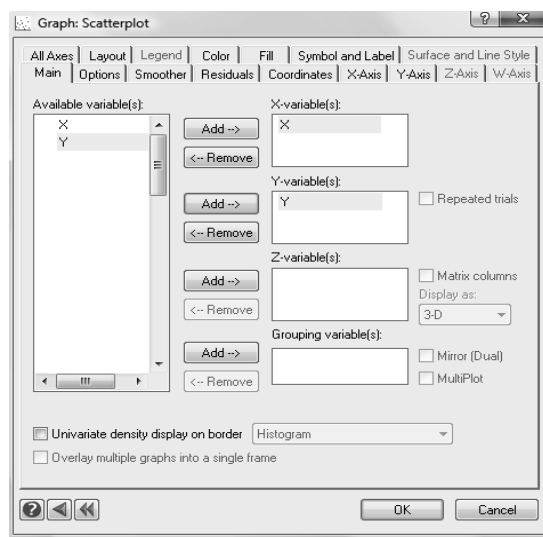


Figura 4.29. Ventana Plot.

3°. Seleccione la variable y a continuación pulsando en el botón **Add**.

4°. En la figura 4.30, puede observar los resultados de nuestro ejemplo.

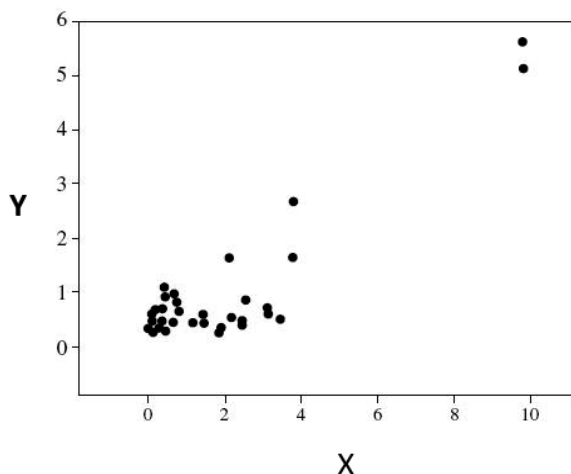


Figura 4.30. Gráfico de dispersión con los valores X e Y.

4. EJERCICIOS PROPUESTOS

4.1. Dados los valores de X relacionados con el número de hijos e Y con el gasto en euros en material escolar:

X	Y
3	100
4	112
5	150
7	210
2	60
3	85
2	77

Elaborar:

- a) Un histograma
- b) Un gráfico de caja
- c) Un diagrama de tallos y hojas
- d) Un diagrama de dispersión

Capítulo 5.

ESTADÍSTICA DESCRIPTIVA

Fuensanta Hernández Pina
Tomás Izquierdo Rus

INTRODUCCIÓN

Uno de los objetivos de la estadística descriptiva es poner orden en los datos obtenidos como resultado de la evaluación de los estudiantes, de una investigación, de información necesaria para tomar decisiones, etc.

La estadística descriptiva, al resumir y describir agrupaciones de datos, pone en orden los datos y permite extraer información válida y fiable sobre los mismos. De ahí, que al iniciar cualquier análisis de datos el investigador o evaluador intente hacer cálculos que le permitan ver el comportamiento y la estructura existente en dichos datos.

En capítulos anteriores, hemos comprobado gráficamente cómo una determinada cantidad de datos se distribuyen y muestran una estructura fácil de analizar visualmente. Con los análisis descriptivos las informaciones que se obtienen complementan la información ofrecida por los gráficos y permiten, además, saber más acerca de tales datos.

1. OBJETIVOS

Los objetivos que pretendemos con el presente capítulo son:

- Conocer el cálculo de los estadísticos descriptivos más usuales.
- Interpretar los resultados del cálculo de los estadísticos descriptivos.
- Valorar la estadística descriptiva como herramienta útil y válida de cara a obtener resultados y conclusiones de investigación educativa y/o social.

2. CONTENIDOS

El módulo básico para el cálculo de estadísticos descriptivos con Mystat 12 es el módulo **Basic Statistic**, aunque antes de proceder al cálculo con ordenador es conveniente clarificar conceptualmente algunos de estos estadísticos:

Mean (Media). Es el promedio aritmético de todos los datos. También se le conoce como el centro de gravedad de la distribución de datos. Su cálculo es muy sencillo, se suman todas las puntuaciones y se dividen por el número de datos.

$$\bar{X} = \frac{\sum X}{N}$$

Una característica importante de la media es que al restarla a cada una de las puntuaciones la suma de las diferencias, denominadas desviaciones, es igual a cero.

$$\sum (X - \bar{X}) = 0$$

De todos los promedios, la media es la única que toma en consideración todos los datos, es la más estable y consistente y su uso está muy generalizado.

Median (Mediana). Se define como el dato que deja por debajo de sí al 50% de los datos y por encima al otro 50%. Su cálculo también es muy sencillo.

La mediana es considerada el valor central de todos los datos ordenados. Cuando los datos presentan una distribución normal, la media y mediana coinciden. En caso de sesgo la media puede desplazarse a la derecha o a la izquierda produciendo asimetría.

Moda (moda). Es el dato o valor que más se repite en una serie de números. Aunque este promedio no viene calculado por MYSTAT es fácil deducirlo a través de los gráficos que se realicen.

El concepto de variabilidad es muy importante en investigación educativa. Este concepto se refiere al grado de dispersión o separación que los datos presentan entre sí. Cuando existe total homogeneidad en unos datos la dispersión es cero. En cambio, cuanto más difieran los datos entre sí la variabilidad será mayor. El concepto de varianza tiene importantes aplicaciones en investigación educativa y social y en el análisis estadístico complejo.

El estudio de la dispersión o variación de los datos hemos dicho que se basa en la desigualdad que las puntuaciones presentan entre sí. Comparar todas las puntuaciones entre sí para extraer una conclusión acerca de su variabilidad es un trabajo tedioso y poco útil, de ahí que para llevar a cabo el estudio de la dispersión o variabilidad de un grupo se haga tomando como referencia alguno de los estadísticos centrales o promedios, la mayoría de las veces la media. Así pues la desviación típica o estándar, que es la que describe la variabilidad de un grupo, se hace viendo cuanto se aleja cada puntuación de la media del grupo al que pertenece.

El módulo **Basic-Statistic** ofrece varias medidas de dispersión. A saber:

Range (rango). Es la medida de dispersión más sencilla, su cálculo se efectúa restando del valor más alto de la serie el valor más bajo. Aunque su utilidad es muy escasa, nos informa de la amplitud en la distribución de los datos.

Variance (varianza). Este estadístico tiene una gran utilidad en el cálculo de otros análisis más complejos, la varianza es la media de las desviaciones al cuadrado de cada puntuación respecto a la media. Es decir, la diferencia de cada puntuación respecto a la media se eleva al cuadrado, todas estas desviaciones al cuadrado se suman y el resultado se divide por el número de datos. Dicha media es lo que denominamos **varianza**. Que viene expresada por:

$$\text{Varianza} = \frac{\sum X^2}{N - 1} \text{ siendo } X \text{ la diferencia entre } X_i \text{ y } \bar{X}.$$

En algunos textos viene representada por:

$$S^2 = \frac{\sum (X - \bar{X})^2}{N - 1}$$

SD (desviación típica o estandar). Esta es la medida de dispersión por excelencia. Una vez conocido el proceso de cálculo de la varianza su cálculo es muy sencillo, ya que podemos obtenerla haciendo la raíz cuadrada de la varianza. Este proceso de extracción de la raíz cuadrada se hace para llevar los datos de la varianza a las puntuaciones originales y poder hacer operaciones y aplicaciones posteriores.

Maximum (puntuación máxima). Informa del valor más alto de la serie de datos.

Minimum (puntuación mínima). Informa del valor más bajo de la serie de datos.

Sum(suma). Es la suma de todas las puntuaciones.

Kurtosis (curtosis). Informa de la distribución de los datos respecto a la media. Cuando los datos varían poco, en torno al promedio, se produce una curva puntiaguda conocida como **leptocúrtica**. Si hay mucha dispersión, el efecto es un aplanamiento que se conoce con el nombre de **platicúrtica**. En ambos casos, la comparación de estas dos curvas se hace con lo que se denomina curva normal o **mesocúrtica**.

Skewness (sesgo o asimetría). Es la representación que una serie de datos puede ofrecer respecto a la distribución normal. A veces, por diversas razones, la distribución presenta desviaciones acumuladas de los datos a la derecha o a la izquierda. En estadística, estas desviaciones tienen un nombre y una interpretación por parte del investigador. Si la acumulación de las puntuaciones es hacia la izquierda la asimetría es positiva; si lo hace a la derecha, es negativa. El índice que calcula el programa hace referencia a estos dos tipos de sesgo. Cuanto más grande sea el valor, más asimétrica será la curva.

SEM (error estándar de medida). Este estadígrafo es una desviación que hace referencia a la distribución muestral de medias. Se utiliza para calcular la prueba **t** de Student. Su cálculo es sencillo, sólo hay que dividir la desviación típica (**s**) de la muestra por la raíz cuadrada de **n**.

$$S_x = s / \sqrt{n}$$

Dadas las características más informáticas de éste manual es recomendable que consulte algún otro libro de estadística para el cálculo no informático de los estadísticos descriptivos.

Una vez que a nivel conceptual se conocen la mayoría de los estadísticos descriptivos y que los resultados que se obtengan pueden ser interpretados, pasamos al cálculo de los mismos a través de Mynstat. Éste cálculo se realiza picando en el módulo **Basic-Statistic** (Figura 5.1) y los pasos a seguir para obtener los resultados son los siguientes:

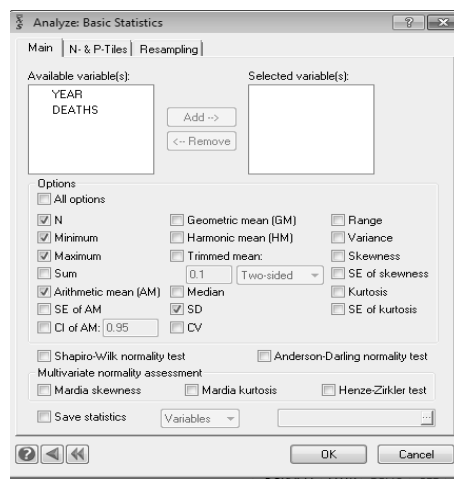


Figura 5.1. Módulo Basic-Statistics.

Los pasos a seguir para realizar los estadísticos descriptivos son los siguientes:

1°. Desde la barra de herramientas de Mynstat 12, cargue el fichero de trabajo picando en File/Open/Data.

Una vez comprobado que éste es el fichero con el que desea trabajar y que su contenido es correcto, puede proceder a cerrarlo, modificarlo, minimizarlo, etc., para continuar con la realización de los estadísticos.

2°. Pique sobre la opción **Basic-Statistic**, desde la barra de herramientas, y luego la opción Analyze (Figura 5.2).

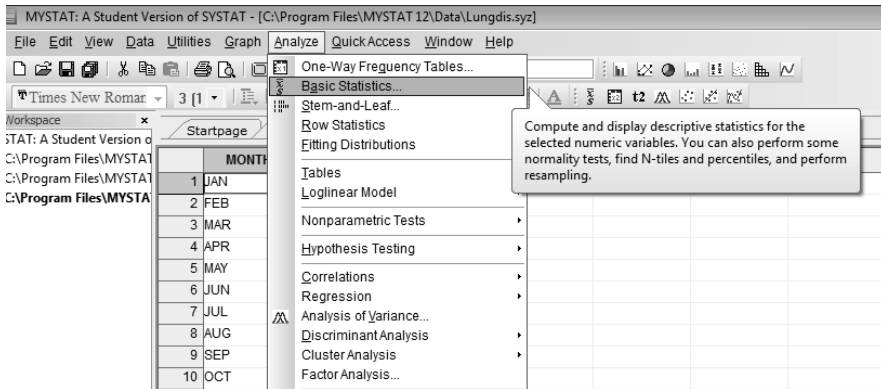


Figura 5.2. Ventana Basic-Statistics.

3°. Seleccione la variable o variables a las que desea realizar el cálculo de sus estadísticos descriptivos. Esta acción se realiza picando una a una las variables y a continuación, pulsando en el botón **Add**. Cada variable se irá incorporando al rectángulo de la derecha. En caso de equivocación al seleccionar alguna variable, proceda del modo inverso, márkuela en el rectángulo de la derecha y posteriormente pulse en el botón **Remove**.

A continuación seleccione (picando sobre su nombre) los estadísticos que desea calcular, por defecto aparecen unos marcados pero usted, picando sobre ellos, puede marcar o desmarcar. Lo más rápido suele ser marcar la opción **All options**, ya que de este modo realizará todos los estadísticos posibles y posteriormente se interpretarán solamente aquellos que necesite.

Una vez seleccionadas las variables y los estadísticos a calcular (Figura 5.3), pulse el botón **Ok** para obtener los resultados.

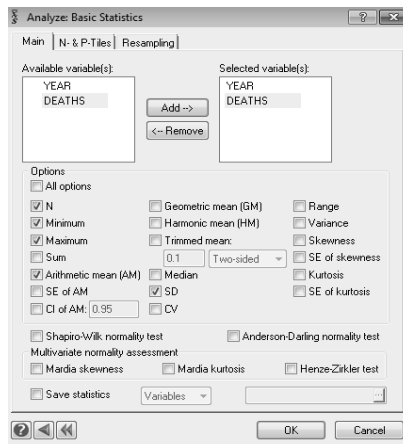


Figura 5.3. Ventana Basic Statistics con variables y estadísticos.

4º. Grabar, imprimir o cancelar los resultados obtenidos. Para grabar los resultados debe proceder desde **File/Save As**. Seleccione la unidad, ruta y nombre de archivo, sin olvidar añadir la extensión .syo ya que, como hemos indicado en otros capítulos, los datos almacenados es texto que puede ser editado con cualquier procesador de textos (Figura 5.4).

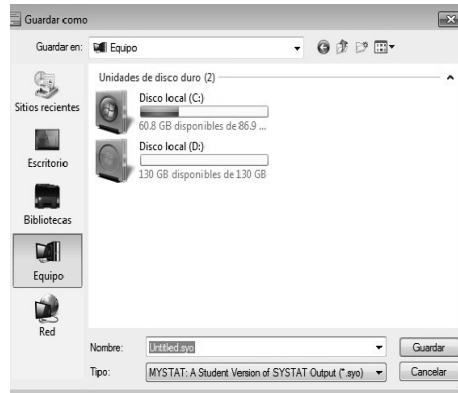


Figura 5.4. Ventana Save As.

2.1. Análisis descriptivo por subgrupos

Si desea realizar algún análisis estadístico por subgrupos, es decir seleccionar un subgrupo de la totalidad de sujetos, por ejemplo sólo las mujeres o únicamente los que tengan más de una determinada edad, etc., pues este tipo de análisis no puede realizarse directamente ya que el módulo Basic-Statistic no dispone de ninguna opción para la selección de sujetos en función de variables. Por lo tanto debemos proceder de otro modo que pasamos a explicar:

1º. Una vez seleccionado el fichero con el que desea trabajar, pase a la hoja de datos y, con ésta en pantalla, seleccione la opción **Data/Select Cases** (Figura 5.5). Este modo de seleccionar sujetos ya ha sido abordado en otro capítulo anterior cuando explicábamos las posibilidades que ofrece la hoja de datos, pero para facilitar el trabajo al lector volvemos a reproducir.

2º. En la hoja de diálogo que aparece (Figura 5.5), seleccione la variable que empleará como criterio para seleccionar los sujetos a los que quiere realizar el estudio descriptivo, en nuestro ejemplo es la variable GENERO, y pulse el botón **Add**.

3º. Observará que en una línea inferior, denominada **Select**, aparece la variable seleccionada y el criterio de selección, en este caso seleccionamos los que tienen asignado GENERO=1, es decir a los hombres.

4º. Pulse el botón **Ok** y minimice la ventana, sin cerrarla.

5º. A continuación vuelva a ejecutar el módulo **Basic-Statistics** y en la ventana de diálogo específica de este módulo (Figura 5.6) seleccione la variable o variables a analizar y los estadísticos a calcular.

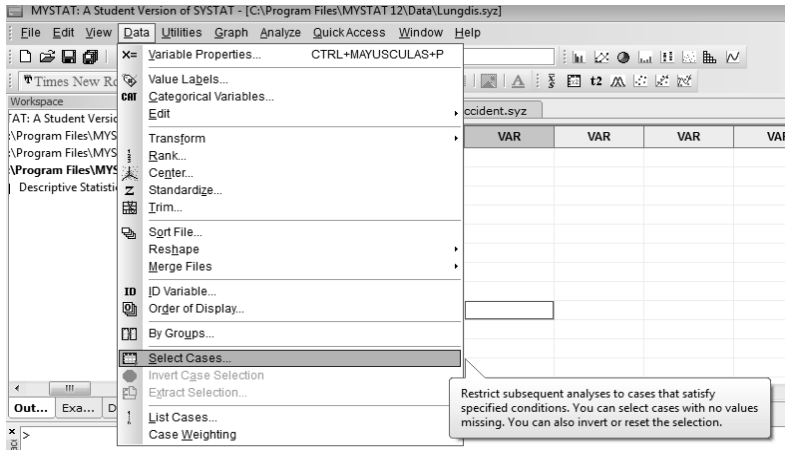


Figura 5.5. Opción Data / Select Cases.

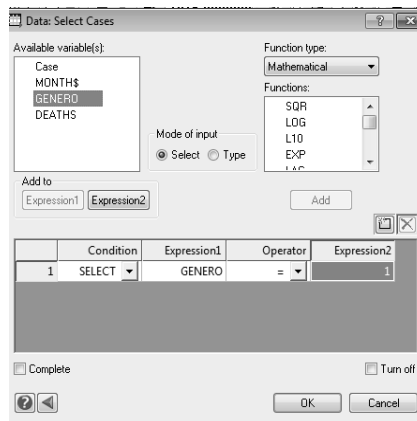


Figura 5.6. Ventana Data/Select Cases de Basic-Statistics.

6°. Normalmente, cuando se hace una selección por subgrupo lo que se pretende es comparar determinados estadísticos de ciertas variables entre los diferentes subgrupos que componen una variable.

Por ejemplo, si observa las variables seleccionadas en la figura 5.7 puede observar que estamos calculando los estadísticos descriptivos de las variables INGRESOS y CI, al realizar este estudio por subgrupos, en concreto en función de las categorías de la variable GENERO, lo que pretendemos hacer es comparar los ingresos y el coeficiente de inteligencia de los sujetos en función de su género, es decir diferenciando y comparando entre hombres de mujeres.

Como la selección que hemos realizado corresponde únicamente a los hombres, tendríamos que volver a repetir la selección pero esta vez asignando GENERO=2,

es decir seleccionando a las mujeres, para ello volveríamos a realizar todo el proceso desde el apartado 3 antes de pasar al último apartado, el 7º.

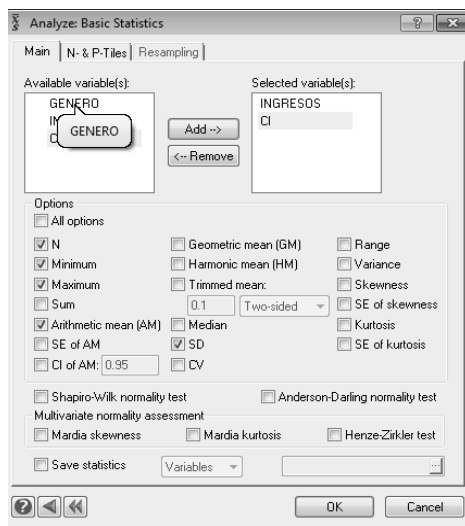


Figura 5.7. Ventana Basic-Statistics con variables y estadísticos.

7º. Grabar, imprimir o cancelar los resultados obtenidos. Para grabar los resultados de ambos subgrupos debe proceder desde **File/Save As**. Seleccione la unidad, ruta y nombre de archivo, sin olvidar añadir la extensión .syo ya que, como hemos indicado en otros capítulos, los datos almacenados es texto que puede ser editado con cualquier procesador de textos (Figura 5.8).

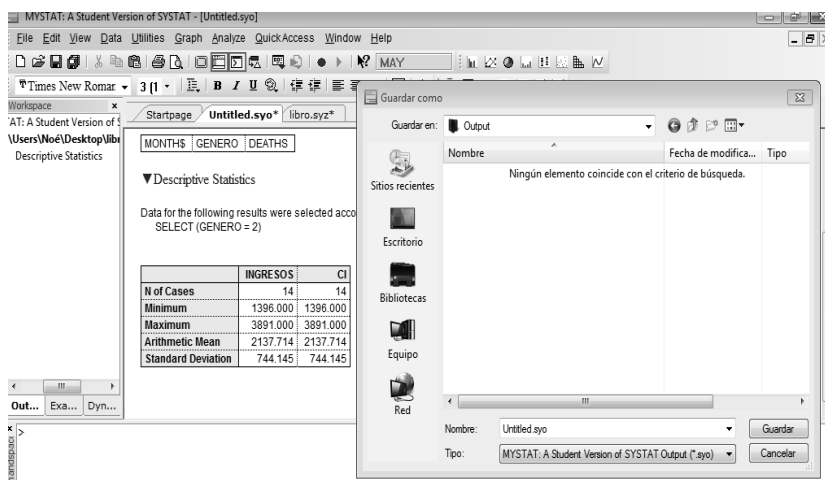


Figura 5.8. Ventana Save As.

2.2. Cálculo de frecuencias y porcentajes con mystat

Este módulo se emplea para calcular frecuencias (número de sujetos) y los porcentajes que representan, para una, dos o más variables.

Partimos, para emplear el módulo, de unos datos que hemos introducido, por ejemplo los que aparecen en la figura 5.9.

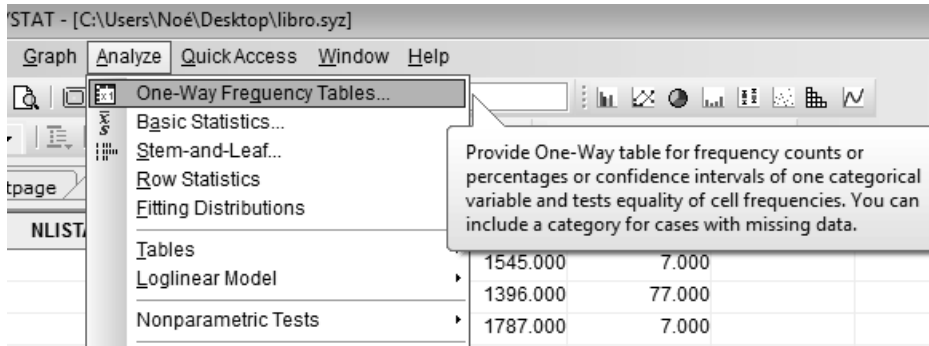


Figura 5.9. Módulo Tables.

Partimos de un fichero con algunas variables, la primera que hace referencia al número de cuestionario, seguida de la edad, el género, la pregunta 1, pregunta 2, pregunta 3...

Al pulsar sobre el botón que ejecuta este módulo aparece la ventana siguiente (Figura 5.10). Observe que debe quedar marcada la opción **Counts**.

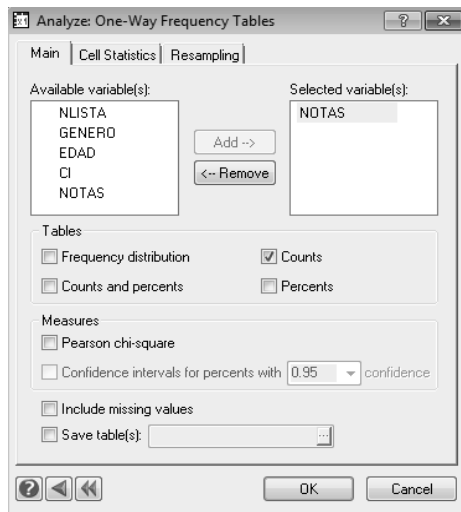


Figura 5.10. Ventana del módulo Tables.

Para realizar el cálculo de los porcentajes de una variable, como, por ejemplo, la variable P1 hace referencia a la lectura de revistas de investigación educativa y ha sido codificada como:

1. Cuando la persona NUNCA lee estas revistas.
2. Cuando la persona lee A VECES estas revistas.
3. Cuando la persona lee A MENUDO estas revistas.
4. Cuando la persona lee FRECUENTEMENTE estas revistas.
5. Cuando la persona lee SIEMPRE (CASI SIEMPRE) estas revistas.

Si pretendemos calcular cuantas personas nunca leen estas revistas (1), cuantas las leen a veces (2), cuantas las leen a menudo (3)... etc., y el porcentaje que representan sobre el total de sujetos debemos proceder del siguiente modo:

- 1) Seleccionamos la variable P1 de la ventana de la izquierda y picamos en el botón **Add**. De esta manera pasará a la ventana de la derecha, como se observa en la figura 5.11.
- 2) Una vez seleccionada, picamos en el botón **Ok** y ya está realizado el cálculo. En la ventana de resultados de Mystal aparecen los datos referidos a los porcentajes (entre paréntesis) y los datos de las frecuencias (número de sujetos que han respondido cada una de las opciones 1, 2, 3....) y otra indicando el porcentaje que representan cada grupo de sujetos en función del total (Figura 5.12).
- 3) Recuerde que estamos haciendo cálculos para todos los sujetos (el total de sujetos), sin separarlos por subgrupos mediante una variable como puede ser el GÉNERO. En el apartado anterior, ya indicábamos como hacerlo en el cálculo de los estadísticos descriptivos, es decir separando hombres de mujeres para ver que grado de lectura de las revistas de investigación tienen ellas frente a ellos. Pasamos a explicar cómo debe realizarse este cálculo con el módulo **Tables**.

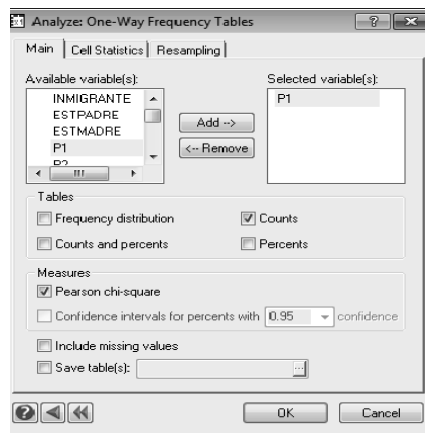


Figura 5.11. Selección de la variable P1.

▼ Crosstabulation: One-Way

Table of Counts and Percents

Values for ITEM1

NUNCA	ALGUNAS VECES	A MENUDO	FRECUENTEMENTE	SIEMPRE	Total
71(39.444%)	58(32.222%)	14(7.778%)	11(6.111%)	26(14.444%)	180(100.000%)

Figura 5.12. Frecuencias y porcentajes de la P1.

2.2.1. Cálculo de frecuencias y porcentajes por subgrupos

Para realizar el cálculo de los porcentajes de una variable en función de otra, como por ejemplo, la variable P1 hace referencia a la lectura de revistas de investigación. Si queremos analizar los datos en función del GÉNERO (siendo hombres o mujeres los que han respondido). Recuerde que la variable P1 ha sido codificada como:

1. Cuando la persona NUNCA lee estas revistas
2. Cuando la persona lee A VECES estas revistas
3. Cuando la persona lee A MENUDO estas revistas
4. Cuando la persona lee FRECUENTEMENTE estas revistas
5. Cuando la persona lee SIEMPRE (CASI SIEMPRE) estas revistas

Y la variable GÉNERO: (1) Hombre y (2) Mujer.

Para calcular la cantidad de hombres y de mujeres que NUNCA leen revistas de investigación (1), cuantas leen A VECES (2), cuantas las leen A MENUDO (3), etc., y el porcentaje que representan sobre el total de hombres, por un lado y mujeres por otro lado, debemos proceder del siguiente modo:

- 1) Desde el menú herramientas de Mystat, seleccionamos la opción **Analyze—Tables—Two-Way**
- 2) Seleccionamos la variable P1 de la ventana de la izquierda y le picamos al botón **Add**. De esta manera pasará a la ventana de la derecha (**Row Variable**). Después seleccione la variable GÉNERO y se añade a la casilla **Column Variable**, como se observa en la figura 5.13.
- 3) Una vez seleccionadas, picamos en el botón **Ok** y ya está realizado el cálculo. En la ventana de resultados de **Mystat** aparece la tabla con los resultados, donde se referencia a la Frecuencia (número de sujetos que han respondido cada una de las opciones 1, 2, 3..., pero en función del género (hombre-mujer) en las columnas), y otra indicando el porcentaje que representan cada grupo de sujetos en función de la variable GÉNERO (Figura 5.14).
- 4) Recuerde que estamos haciendo cálculos para todos los sujetos pero separados en dos subgrupos, hombres frente a mujeres, lo mismo puede hacerse con cualquier otra variable, incluso pueden incluirse simultáneamente más de una variable para establecer subgrupos.

- 5) Observe que para realizar correctamente este cálculo debe estar marcado en la ventana del módulo **Tables-Two-Way** las opciones **Counts and percents**.

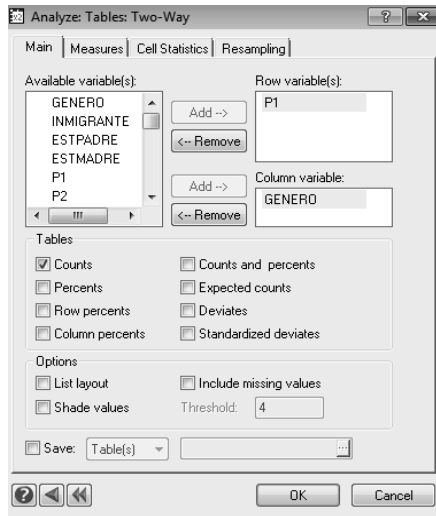


Figura 5.13. Ventana seleccionando P1 en función del GENERO.

Table of Counts and Percents

ITEM1(rows) by GENERO(columns)

	HOMBRE	MUJER	Total
NUNCA	39(21.667%)	32(17.778%)	71(39.444%)
ALGUNAS VECES	31(17.222%)	27(15.000%)	58(32.222%)
A MENUDO	5(2.778%)	9(5.000%)	14(7.778%)
FRECUENTEMENTE	9(5.000%)	2(1.111%)	11(6.111%)
SIEMPRE	16(8.889%)	10(5.556%)	26(14.444%)
Total	100(55.556%)	80(44.444%)	180(100.0%)

Figura 5.14. Resultados del análisis de la P1 en función del GENERO.

Para interpretar los resultados obtenidos en la figura 5.14 hemos de tener en cuenta lo siguiente:

- 1) ROWS – FILAS (opciones de la variable P1)
- 2) COLUMN – COLUMNAS (opciones de la variable GENERO)
- 3) Observe que el total de las columnas en porcentajes suma 100%, esto quiere decir que independientemente del número de sujetos que sean hombres o mujeres, son el total de hombres y el total de mujeres.

3. APLICACIÓN PRÁCTICA CON ORDENADOR

Los siguientes datos indican la distribución de la variable peso de una población de 19 personas: 60 90 100 95 100 50 80 93 86 60 50 91 80 70 60 68. Calcule la moda, media, mediana, varianza y desviación típica

1°. Desde la barra de herramientas de Mynstat, cargue el fichero de trabajo pican-do en File/Open/Data.

2°. Pique sobre el icono del módulo **Analyze—Basic-Statistic** de la barra que se encuentra en Mynstat Main y aparece la siguiente ventana (Figura 5.15):

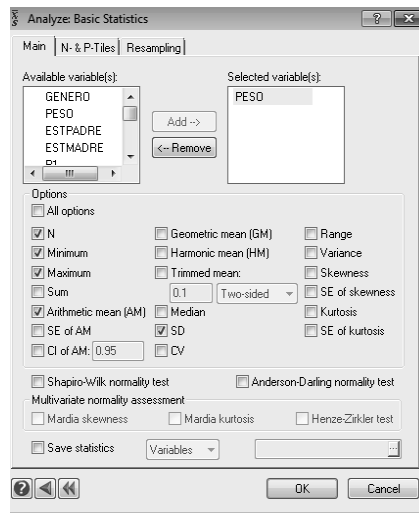


Figura 5.15. Ventana Basic Statistics.

3°. Seleccionamos la variable sobre la cual vamos a realizar el cálculo de sus estadísticos descriptivos y, a continuación, pulsamos el botón **Add**. Una vez seleccionada la variable y los estadísticos a calcular, pulsamos el botón **Ok** para obtener los resultados.

▼ Descriptive Statistics

	HORASESTUDIO
N of Cases	53
Minimum	3.000
Maximum	15.000
Median	7.000
Arithmetic Mean	7.396
Standard Deviation	2.575
Variance	6.628

Figura 5.16. Tabla resultados Basic-Statistics.

La tabla anterior nos indica: horas máxima de estudio, (15), y mínimo, (3), la media de toda la muestra, (7.396), la varianza, (6.628) y su correspondiente desviación típica, (2.575), la cual nos indica cuanto se desvían e número de horas con respecto al valor medio, (7.396), y la mediana, (7), es decir es el valor que divide en dos la muestra de datos.

Imaginemos que deseamos obtener los mismos estadísticos descriptivos, pero esta vez teniendo en cuenta una variable más, el género:

1º. Una vez seleccionado el fichero con el que desea trabajar, pase a la hoja de datos y con ésta en pantalla seleccione la opción **Data/Select Cases**. Este modo de seleccionar sujetos ya ha sido abordado en otro apartado anterior cuando explicábamos las posibilidades que ofrece la hoja de datos, pero para facilitar el trabajo al lector volvemos a reproducir.

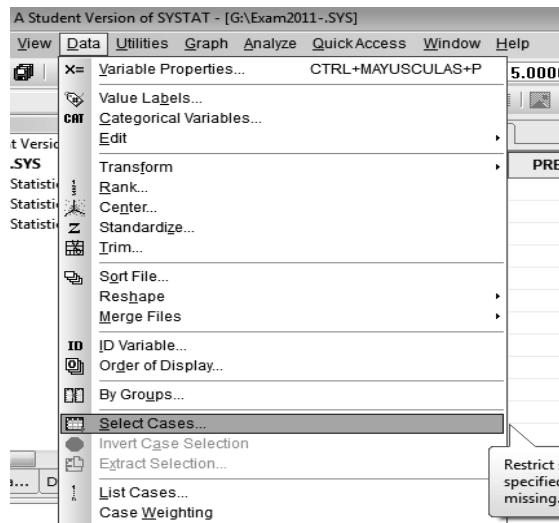


Figura 5.17. Opción Data / Select Cases.

2º. En la hoja de diálogo que aparece (Figura 5.18), seleccione la variable que empleará como criterio para seleccionar los sujetos a los que quiere realizar el estudio descriptivo, en nuestro ejemplo es la variable GÉNERO y pulse el botón **Add**.

3º. Observará que en una línea inferior, denominada **Select**, aparece la variable seleccionada y el criterio de selección, en este caso seleccionamos los que tienen asignado GÉNERO=1, es decir a los hombres.

4º. Pulse el botón **Ok** y a continuación vuelva a ejecutar el módulo **Stats** y, en la ventana de diálogo específica de este módulo, seleccione la variable o variables a analizar y los estadísticos a calcular, en este caso la misma variable que el ejemplo anterior, (el peso), pero además sólo nos interesa el de los hombres, genero=1 (Figura 5.19).

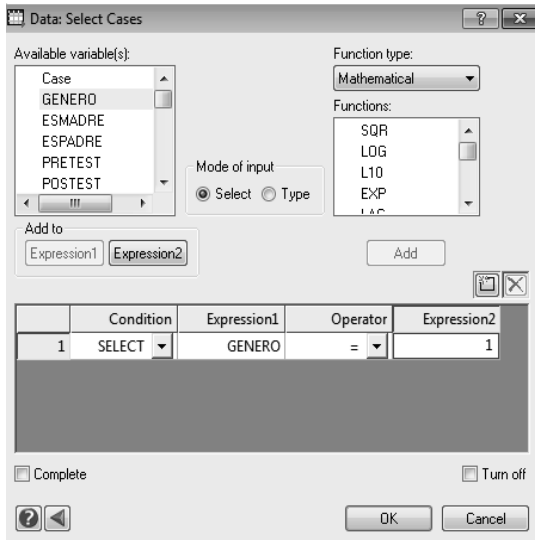


Figura 5.18. Ventana Select Cases.

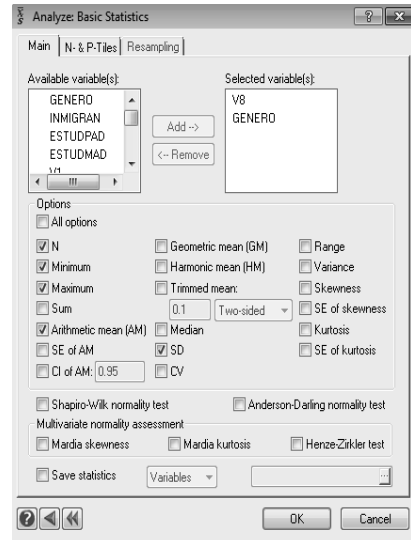


Figura 5.19. Ventana Basic Statistics con variables y estadísticos.

▼ Descriptive Statistics

Data for the following results were selected according to
SELECT (GENERO = 1) / COMPLETE

	V8	GENERO
N of Cases	99	99
Minimum	1.000	1.000
Maximum	5.000	1.000
Arithmetic Mean	2.535	1.000
Standard Deviation	1.534	0.000

Figura 5.20. Ventana Save Text in Main Window.

5°. Pulse OK y obtendrá los resultados (Figura 5.20).

Esta ventana nos indica el número de casos, 99, y los estadísticos descriptivos para la variable genero/hombre. Estos datos, a diferencia de los anteriores nos hacen ver los valores máximo y mínimo que pertenecen al grupo de hombres y el valor medio del hombre (2.535) es inferior al de la media teniendo en cuenta tanto hombres como mujeres (7.39).

De la misma manera, podemos proceder a realizar el cálculo de frecuencias y porcentajes con Mystat con el mismo ejemplo propuesto.

Al pulsar sobre el botón que ejecuta este módulo aparece la ventana siguiente (Figura 5.21). Observe que debe quedar marcada la opción **Counts and percents**.

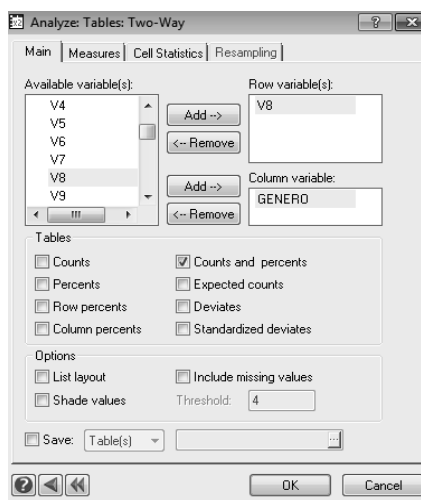


Figura 5.21. Ventana del módulo Tables Two-Way.

Para realizar el cálculo de los porcentajes de una variable, seleccionamos la variable situada en la ventana de la izquierda, picamos en el botón **Add** y de esta manera pasará a la ventana de la derecha.

Una vez seleccionada, picamos en el botón **Ok** y ya está realizado el cálculo. En la ventana de resultados aparecen dos tablas con los resultados, una que hace referencia a la Frecuencia y otra indicando el porcentaje que representan cada grupo de sujetos en función del total (Figura 5.22).

Table of Counts and Percents

Values for V8

1	2	3	4	5	Total
70(39.106%)	36(20.112%)	15(8.380%)	20(11.173%)	38(21.229%)	179(100.000%)

Figura 5.22. Frecuencias y porcentajes de la V8.

La tabla de contingencia de la figura nos muestra en la parte inferior las frecuencias y porcentajes, por ejemplo comprobamos que hay 36 sujetos con valores de 2, 15 con valor 3 y al lado de las frecuencias observamos los porcentajes, un 39% con valores de 1, un 21.22% con valores de 5, etc.

También podemos proceder a realizar el cálculo de frecuencias y porcentajes de una variable en función de otra. La variable peso podemos analizarla en función del GENERO (siendo hombres o mujeres los que han respondido).

1. Seleccione la variable peso y pique el botón **Add**. De esta manera pasará a la ventana de la derecha. Después seleccione la variable GENERO y se añade a la primera casilla **Column variable**, como se observa en la figura 5.23.

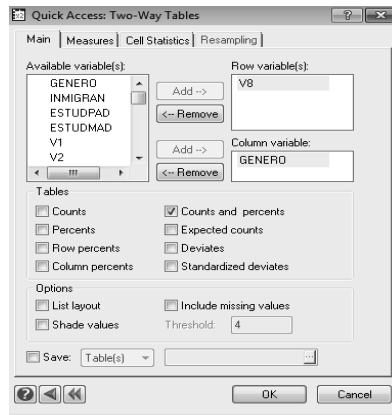


Figura 5.23. Modelo Two-way Tables para dos variables.

- Una vez seleccionadas, pique en el botón **Ok** y ya está realizado el cálculo. En la ventana de resultados aparecen tres columnas con los resultados para cada opción de respuesta, una que hace referencia a la Frecuencia (número de sujetos que han respondido cada una de las opciones 1, 2, 3..., pero en función del género (hombre-mujer) en las columnas) y otra indicando el porcentaje que representan cada grupo de sujetos en función de la variable GENERO (Figura 5.24).

Table of Counts and Percents

V8(rows) by GENERO(columns)

	hombre	mujer	Total
1	38(21.229%)	32(17.877%)	70(39.106%)
2	19(10.615%)	17(9.497%)	36(20.112%)
3	9(5.028%)	6(3.352%)	15(8.380%)
4	17(9.497%)	3(1.676%)	20(11.173%)
5	16(8.939%)	22(12.291%)	38(21.229%)
Total	99(55.307%)	80(44.693%)	179(100.0%)

Figura 5.24. Resultados del análisis de la V8 en función del GENERO.

La interpretación de esta tabla quedaría de la siguiente manera: un 21.22% de hombres con valores de 1, 17.8% para mujeres, que suponen un 39.10% del total. Las columnas nos indican que hay un 55.30% de hombres y 44.69% de mujeres, 99 y 80 respectivamente.

4. APLICACIÓN PRÁCTICA SIN ORDENADOR

Los siguientes datos indican la distribución de la variable peso de una población de 19 personas: 60 90 100 95 100 50 80 93 86 60 50 91 80 70 60 68. Calcule la moda, media, mediana, varianza y desviación típica.

a. Cálculo de la media

$$\bar{x} = \frac{\sum_i x_i n_i}{n}$$

$$\bar{x} = \frac{1233}{16} = 77.06$$

b. Cálculo de la moda

El valor más repetido es 60

c. Cálculo de la mediana

Una vez ordenados los datos, observamos que la mediana es 80

50 50 60 60 60 68 70 **80 80** 90 86 91 93 95 100 100

d. Varianza y desviación típica

$$s^2 = \frac{\sum_i (x_i - \bar{x})^2 n_i}{n}$$

$$S^2 = \frac{(50-77,06)^2 + (50-77,06)^2 + (60-77,06)^2}{16} = 303.79$$

La desviación típica S: es la raíz cuadrada de la varianza, por tanto $\sqrt{303.79} = 17.04$.

5. EJERCICIOS PROPUESTOS

5.1. Esta tabla contiene datos referentes a la distribución por edades en una provincia española, unos datos referentes a la capital y el resto relacionados con el resto de municipios:

EDAD	CAPITAL (%)	RESTO DE MUNICIPIOS (%)
0-15	6	3
15-30	9	5
30-50	16	8
50-70	12	7
70-90	7	7

Elabora los histogramas correspondientes a ambas series, calcular el valor de la mediana y de la media correspondiente.

5.2. Calcular, media, mediana y desviación típica de las siguientes calificaciones en 22 alumnos: 5 6 7 8 9 10 11 12 13 14 16 1 3 1 4 3 9 4 1 2 1 2.

5.3. Se estudian los ingresos (en euros) de algunas familias en diversos municipios de una determinada comunidad autónoma española

INGRESOS	700-900	900-1100	1100-1300	1300-1500	1500-1700
NÚMERO DE FAMILIAS	13	10	4	6	4

- ¿Cuál es el número de familias cuyas rentas se incluyen entre 700 y 900?
- ¿Cuál es la proporción de familias cuyas rentas se incluyen entre 900 y 1500?
- ¿Cuál es la media de las rentas? (Precisar la fórmula utilizada)
- ¿Cuál es la desviación típica de las rentas? (Precisar la fórmula utilizada)
- ¿Qué miden la media y la variación?
- Hacer el histograma correspondiente

5.4. Dados los datos demográficos de países europeos en millones de habitantes (UE15, población en 2002): Bélgica 10310, Dinamarca 5368, Alemania 82440, Grecia 10598, España 40409, Francia 59344, Irlanda 3883, Italia 58018, Luxemburgo 444, Países Bajos 16105, Austria 8139, Portugal 10336, Finlandia 5195, Suecia 8909, Reino Unido 60114.

- Calcule media, mediana y desviación típica.
- Calcule el segundo cuartil.
- Describa los datos mediante una representación gráfica.
- Calcule la media y la mediana suprimiendo Alemania. ¿Qué conclusiones puede extraer?

Capítulo 6.

SIGNIFICACIÓN DE LA DIFERENCIA ENTRE DOS MUESTRAS

Javier J. Maquilón Sánchez
José D. Cuesta Sáez de Tejada

INTRODUCCIÓN

MYSTAT cuenta con un módulo que permite comprobar o contrastar la significación de la diferencia entre dos medias procedentes de dos grupos. La prueba t de Student se utiliza para la comparación de dos medias poblacionales.

Cuando en una investigación se comparan los resultados obtenidos por un grupo experimental con los obtenidos por un grupo de control; o una media obtenida en la variable dependiente en la situación de pretest con la obtenida en el posttest después de un tratamiento, la prueba t es la más apropiada.

1. OBJETIVOS

Los objetivos que pretendemos con el presente capítulo son:

- Conocer la utilidad de la prueba t -test para comparar grupos tanto dependientes como independientes.
- Diferenciar si un grupo es dependiente o independiente.
- Emplear la prueba t -test para comprobar la existencia de diferencias significativas entre grupos.
- Valorar la utilidad de la prueba t -test en los resultados de trabajos de investigación.

2. CONTENIDOS

En investigación educativa, al comparar grupos, éstos están relacionados cuando entre ellos existe algún elemento común. Por ejemplo, si a un grupo le aplicamos un pretest para medir la variable dependiente, un tratamiento y un posttest para medir de nuevo la variable dependiente, la comparación de los resultados obtenidos entre el pretest y el posttest debe hacerse con una prueba de significación de diferencias para grupos relacionados. La fórmula sería la siguiente:

$$t = \frac{\bar{D}}{\sqrt{\frac{\sum D^2 - \frac{(\sum D)^2}{N}}{N(N-1)}}$$

donde:

$\bar{D}$: es la media de la diferencia de todos los pares de puntuaciones

$\sum D^2$: sumatorio de los cuadrados de las diferencias

$\sum (D)^2$: es el cuadrado de la suma de las diferencias

N : número de pares de puntuaciones

$N-1$: grados de libertad (número de pares de puntuaciones menos uno)

Dos grupos son independientes cuando los sujetos han sido elegidos al azar o la asignación al grupo experimental y al grupo de control se ha hecho de forma aleatoria. Por ejemplo, si un grupo lo utilizamos como grupo experimental y su equivalente como grupo de control, la comparación de sus medias en el posttest hemos de hacerla con una prueba de significación para muestras independientes.

$$t_0 = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2 \times \left[\frac{1}{n_1} + \frac{1}{n_2} \right]}}$$

donde:

1. " t_0 " tiene probabilidad de ocurrencia en la tabla de distribución *t* de *Student* con $n_1 + n_2 - 2$ grados de libertad.

2. $s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$ es la varianza común estimada.

3. $\bar{x}_1$ es el promedio de la muestra 1 y $\bar{x}_2$ es el promedio de la muestra 2.

4. n_1 y n_2 son los tamaños de las respectivas muestras.

El error típico de la diferencia de las medias se estima a partir de las varianzas de cada distribución, y se hace a través de la siguiente fórmula:

$$s = \sqrt{\frac{\sum x_1^2 + \sum x_2^2}{gl_1 + gl_2}}$$

En cualquiera de los dos casos MYSTAT hará los análisis con la opción **t-test**.

2.1. Prueba t-test para grupos independientes

Se utiliza para comparar las medias de una variable entre dos poblaciones diferentes. Los pasos a seguir son:

1°. Desde Starpape, cargue el fichero de trabajo picando en File/Open/Data.

Una vez abierto el fichero, tendrá que activar la hoja de trabajo, picando desde la ventana principal (Starpape) la opción Untitled1.syz.

Una vez comprobado que éste es el fichero con el que desea trabajar y que su contenido es correcto, puede proceder a cerrarlo, modificarlo, minimizarlo, etc., para continuar con la realización del histograma.

2°. Desde la hoja de trabajo, seleccione la opción Analyze/Hypothesis Testing/Mean/Two-Sample T-Test y aparece la siguiente ventana (Figura 6.1). A continuación seleccione la/s variable/es que se pretendan comparar (VD).

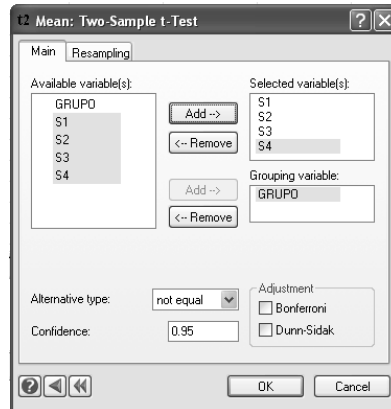


Figura 6.1. Ventana t-test.

Para añadir las variables dependientes (VD), se marcan con el puntero del ratón y se pulsa sobre el botón **Add**, pasando dichas variables al rectángulo de la derecha.

3°. Tras las variables dependientes, se selecciona una única variable independiente (VI), en nuestro ejemplo la hemos consignado como pertenencia al grupo (GRUPO) y está codificada con valores 1 y 2.

4°. Pulse **ok** y en la ventana Untitled.syo aparecerán los análisis de la variable/s dependiente/s en función de la variable independiente consignada (Figura 6.2).

5°. Es conveniente que los resultados obtenidos sean almacenados en un fichero para su posterior tratamiento, interpretación, etc., en un fichero. Los ficheros grabados van a tener la extensión .SYO. En nuestro caso hemos decidido como nombre de fichero **t-test.syo** y pretendemos guardarlo en el disco duro C, y en la carpeta **MYSTAT 12**. Los resultados obtenidos aparecerán en un cuadro de diálogo semejante al de nuestro ejemplo (figura 6.2).

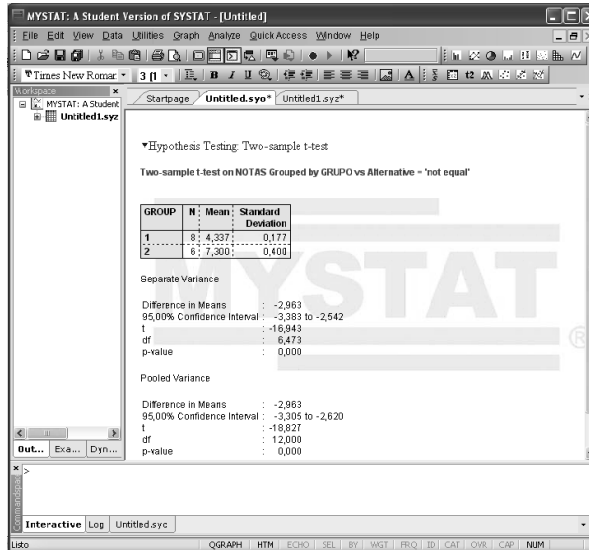


Figura 6.2 Resultados de la prueba t para dos muestras independientes.

En nuestro ejemplo, el p-valor asociado al estadístico de contraste (.000) es inferior que α (.05), por tanto se rechaza la hipótesis nula al nivel de significación de α , y aceptamos la hipótesis alterna. Existen diferencias significativas entre ambos grupos.

2.2. Prueba t-test para grupos dependientes

Se usa para comparar las medias de dos muestras relacionadas. La prueba se emplea en diseños pretest y posttest, sobre los mismos individuos o unidades muestrales. Los pasos que debe seguir para esta prueba son:

- 1º. Desde Starpage, cargue el fichero de trabajo picando en File/Open/Data. Una vez abierto el fichero, tendrá que activar la hoja de trabajo, picando desde la ventana principal (Starpage) la opción Untitled1.syz. Una vez comprobado que éste es el fichero con el que desea trabajar y que su contenido es correcto, puede proceder a cerrarlo, modificarlo, minimizarlo, etc., para continuar con la realización del histograma.
- 2º. Desde la hoja de trabajo, seleccione la opción Analyze/Hypothesis Testing/Mean/Paired T-Test y aparece la siguiente ventana (Figura 6.3). A continuación seleccione las variables dependientes, en este caso corresponden al pretest y posttest, que se pretendan comparar (VD). Para añadir las variables, se marcan con el puntero del ratón y se pulsa sobre el botón **Add**, pasando dichas variables al rectángulo de la derecha.
- 3º. Pulse **ok** y en la ventana Untitled.syo aparecerán los análisis de las variables dependientes (Figura 6.4).

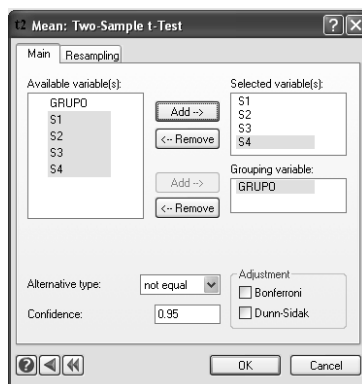


Figura 6.3. Ventana t-test con variables dependientes.

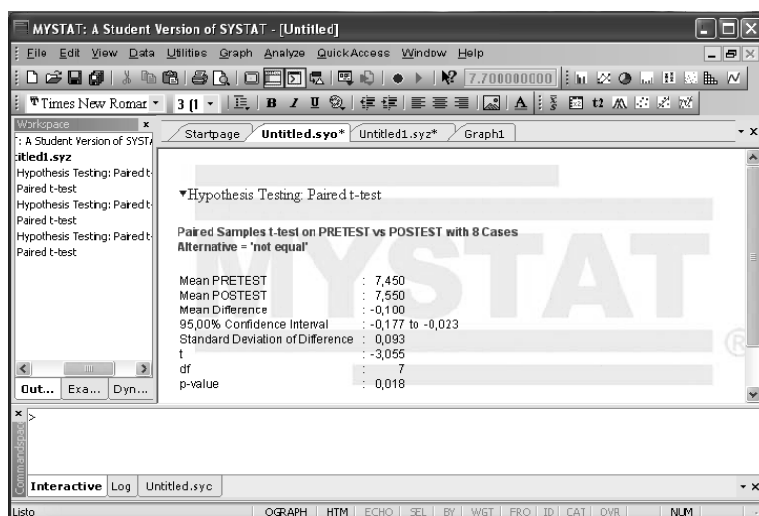


Figura 6.4 Resultados de la prueba t para dos muestras relacionadas.

4º. Es conveniente que los resultados obtenidos sean almacenados en un fichero para su posterior tratamiento, interpretación, etc., en un fichero. Los ficheros grabados van a tener la extensión .SYO. En nuestro caso hemos decidido como nombre de fichero **t-test-D.syo** y pretendemos guardarlo en el disco duro C, y en la carpeta **MYSTAT 12**. Los resultados obtenidos aparecerán en un cuadro de diálogo semejante al de nuestro ejemplo (Figura 6.4).

En este caso, el valor p-asociado al estadístico de contraste (.018) es menor que α (.05), luego, al nivel de significación de .05, se rechazará la hipótesis nula. La diferencia entre lo observado en la muestra y lo esperado bajo la hipótesis nula es estadísticamente significativa.

3. APLICACIÓN PRÁCTICA SIN ORDENADOR

3.1. Para muestras dependientes

Imaginemos que tenemos un grupo formado por seis estudiantes. Le aplicamos una prueba de rendimiento antes de comenzar el curso (pretest) y obtenemos los siguientes resultados:

Sujetos	Pretest
1	8
2	2
3	4
4	3
5	3
6	2

Tabla 1. Puntuaciones Pretest.

Finalizado el curso, le volvemos aplicar la prueba de rendimiento y obtenemos los siguientes resultados:

Sujetos	Posttest
1	9
2	7
3	9
4	9
5	6
6	8

Tabla 2. Puntuaciones Posttest.

Sujetos	Pretest	Posttest	D	D ²
1	8	9	1	1
2	2	7	5	10
3	4	9	5	10
4	3	9	6	24
5	3	6	3	9
6	2	8	<u>3</u>	<u>36</u>
$\Sigma D=23$		$\Sigma D^2=90$		

Tabla 3. Cálculo de las diferencias entre Pretest y Posttest.

Ahora, comenzamos a realizar las operaciones necesarias para poder hacer las correspondientes sustituciones en la fórmula:

$$\bar{D} = \frac{23}{6} = 3.83$$

$$(\Sigma D)^2 = 529$$

$$t = \frac{3.83}{\sqrt{\frac{\frac{529}{6}}{90 - \frac{6}{6(6-1)}}}} = \frac{3.83}{\sqrt{\frac{1.9}{30}}} = \frac{3.83}{.25} = 15.32$$

El estadístico t calculado (15.32) se localiza en los valores críticos de la tabla t con sus correspondientes grados de libertad, en este caso, $N-1$, 6-1. El resultado obtenido es que las medias obtenidas en el pretest y postest son claramente diferentes entre sí, siendo este resultado estadísticamente significativo con una $p < .05$.

3.2. Para muestras independientes

GRUPO 1	GRUPO 2
$x_1 = 10$	$x_2 = 5$
$n_1 = 20$	$n_2 = 20$
$\Sigma x_1^2 = 156$	$\Sigma x_2^2 = 130$

$$S = \frac{156 - 130}{19 + 19} = .92 \quad S_{\bar{x}_1 - \bar{x}_2} = .92 \sqrt{\frac{1}{20} + \frac{1}{20}}$$

$$S_{\bar{x}_1 - \bar{x}_2} = .92 \sqrt{\frac{1}{10}} = .92$$

$$t = \frac{10 - 5}{.92} = 5.43$$

Tabla 4. Cálculo de las diferencias para muestras independientes.

Una vez que tenemos el estadístico t (5.43), buscamos en la tabla de distribución t en función de los grados de libertad estimados en nuestro caso (categorías -1 del primer grupo + categorías -1 del segundo). En este ejemplo, el estadístico t de 5.43 con 28 gl es significativo con un $p < .05$.

Aunque las fórmulas son diferentes en ambos casos, la interpretación y presentación de los resultados en la *prueba t* es la misma. El valor t calculado es un número de tres o cuatro dígitos con dos decimales, como 2.65;1.35; 4.54;2.34. para determi-

nar el nivel de significación, el investigador compara este número con los valores teóricos t en una tabla, denominada tabla de distribución para la prueba t (Fisher y Yates, 1974)¹. En esta tabla se localizan dos números: los grados de libertad (gl) y el nivel de significación estimado. Los grados de libertad son el número de observaciones libres de variar menos uno para cada uno de los grupos. Para cada prueba hay un número determinado de grados de libertad, que después emplearemos para calcular la significación estadística de la prueba. En una tabla t , la intersección de la fila de los grados de libertad y la columna del nivel de significación es el valor de teórico t . Si esta t crítica es menor que el valor t calculado por la ecuación de la prueba t , esto significa que la diferencia observada en las medias es mayor que lo que podía haberse esperado bajo la hipótesis nula, de modo que la hipótesis puede ser rechazada con ese nivel de significación (McMillan y Schumacher, 2005).

4. APLICACIÓN PRÁCTICA CON ORDENADOR

Ahora procederemos a realizar uno de los ejemplos anteriores empleando el programa Mýstat 12:

- 1º. Desde Startpage, cargue el fichero de trabajo picando en File/Open/Data. Una vez abierto el fichero, active la hoja de trabajo picando la ventana donde figura el nombre del archivo con la extensión .SYZ (Figura 6.5)

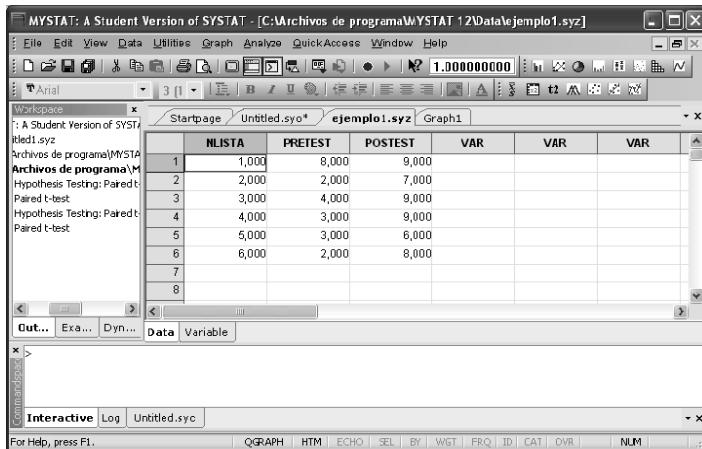


Figura 6.5 Hoja de datos.

- 2º. Seleccione la opción Analyze/Hypothesis Testing/Mean/Paired T-Test y aparece la ventana de la Prueba-T. Desplace al recuadro de la derecha las variables dependientes, en este caso corresponden a un pretest y a un posttest (Figura 6.6), que pretendemos comparar (VD).

¹ Fisher y Yates (1974), Statical Tables for Biological, Agricultural and Medical Research, tabla III, 6ª edición, Londres: Longman Group UK Ltd. (publicado previamente por Oliver and Boyd Ltd. Edimburgo)

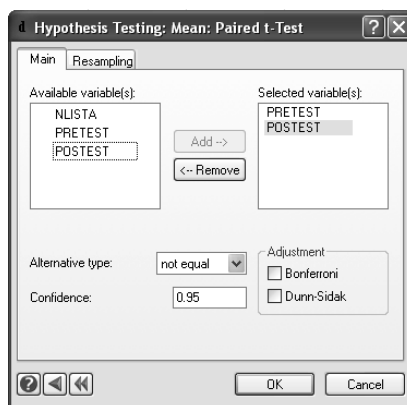


Figura 6.6. Prueba t-test con Mystat 12.

3º. Pulse **Ok** y en la pantalla *Untitled.syo* aparecerán los análisis de las variables dependientes (Figura 6.7).

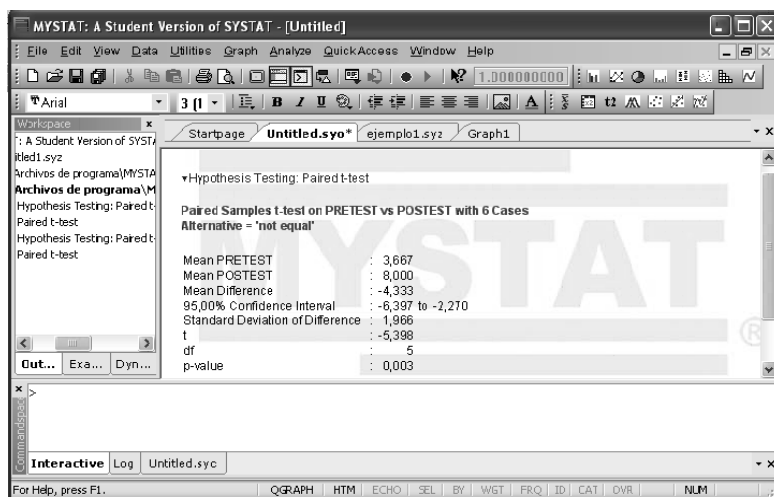


Figura 6.7. Mystat T-Test.

El resultado obtenido, al igual que el ejemplo realizado manualmente nos indica que las medias obtenidas en el pretest y postest son claramente diferentes entre sí ($.003 < .05$), siendo este resultado estadísticamente significativo.

5. EJERCICIOS PROPUESTOS

5.1. La creciente preocupación de una comunidad autónoma española sobre el escaso interés y rendimiento en el aprendizaje de la Física entre los alumnos de

secundaria, llevó a esta comunidad a establecer un plan destinado a fomentar su aprendizaje. El plan consistió en visitar durante dos meses, un día a la semana, un parque de exposiciones de carácter científico, en el cual se improvisaba alguna de las clases que tenía lugar en el centro educativo. Se esperaba que el carácter lúdico de la actividad tuviera resultados positivos en el aprendizaje de la Física. Para ellos se escogió a 10 alumnos, y estos fueron los resultados:

	Previo a la visita	V P		Rendimiento Final
		V	P	
Sujeto 1	2	I	A	3
Sujeto 2	4	S	R	8
Sujeto 3	3	I	Q	8
Sujeto 4	7	T	U	7
Sujeto 5	4	A	E	5
Sujeto 6	2			4
Sujeto 7	7	A		8
Sujeto 8	3	L		5

Tabla 5. Puntuaciones antes y después de la visita al parque.

- Selecciona la prueba estadística adecuada.
- ¿Se ha producido un incremento real en el interés y rendimiento de los alumnos en la asignatura de Física tras su visita al parque de exposiciones?
- Interpreta la decisión que has tomado.

5.2. Se trata de comparar el número de horas dedicadas al estudio entre un centro educativo de la provincia de Murcia y el número de horas que se dedica a la lectura en el resto de las provincias españolas. Para ello se extrae al azar un grupo de 17 alumnos de la provincia de Murcia, y se extrae la media semanal que cada uno de ellos dedica a la lectura. Se procede de la misma manera con otros sujetos de diferentes provincias españolas. Estos son los resultados:

Centro de Murcia: 2, 3, 4, 2, 4, 5, 5, 2, 2, 2, 2, 2, 2, 4, 4, 5, 6

Resto de centros: 1, 1, 2, 3, 3, 3, 3, 5, 5, 4, 4, 4, 5, 5, 4, 3, 2

- Selecciona la prueba estadística adecuada.
- ¿Existen diferencias significativas entre las horas dedicadas a la lectura en la provincia de Murcia y el resto de provincias españolas?
- Interpreta la decisión que has tomado

5.3. Las investigaciones sobre la producción de palabras revelaron numerosos factores que afectan a la velocidad y a la precisión con la cual se denominan algunas imágenes de objetos. Seis personas se propusieron participar en una experiencia

donde se trataba de decir oralmente los nombres de objetos representados por imágenes en blanco y negro. Se presentaron 30 imágenes de objetos, cuyos nombres tendrían que reproducir. Para una imagen dada, se tuvo en cuenta el tiempo pasado entre la presentación de la imagen y el principio de la enunciación del nombre del objeto representado.

Condición 1: 15 de los nombres de imágenes presentadas tienen una elevada frecuencia de uso en nuestra lengua, lo que significa que se encuentran estos nombres relativamente a menudo.

Condición 2: 15 de los nombres de imágenes presentadas tienen una frecuencia escasa de uso, lo que significa que estos nombres raramente podemos encontrarlos.

En la siguiente tabla aparece el tiempo transcurrido entre la visión de la imagen y la reproducción del nombre para cada una de las condiciones.

SUJETOS	1	2
1	574	602
2	624	620
3	626	654
4	572	593
5	499	653
6	645	673

Tabla 6. Visión de la imagen y reproducción del nombre.

¿Se puede concluir, que a un nivel de significación de .05, los tiempos medios de producción son diferentes?

5.4. Dos grupos de alumnos de la titulación de Grado en Pedagogía deben aprender unos conceptos matemáticos con la ayuda de un programa informático. Todos efectúan los aprendizajes propuestos, pero en tiempo diferentes (medido en minutos).

Grupo A : (n = 11)

4; 5; 7; 8; 9; 10; 11; 12; 13; 17; 40

Grupo B : (n = 11)

3; 6; 14; 15; 16; 18; 19; 20; 21; 25; 35

¿Se puede decir que el grupo A difiere del grupo B en cuanto al tiempo de aprendizaje?

Nivel de significación=.05

5.5. Deseamos conocer las habilidades de comunicación de los padres de niños con refuerzo educativo. Los valores de estas habilidades de comunicación oscilan entre 1 (la peor comunicación posible) hasta 10 (la mejor comunicación). Los dos grupos son: A=Padres de niños sin refuerzo, y B=Padres de niños con refuerzo.

A= 8, 4, 6, 3, 1, 4, 4, 6, 4, 2, 2, 1, 1, 4, 3, 3, 2, 6, 3, 4

B= 2, 1, 1, 3, 2, 7, 2, 1, 3, 1, 0, 2, 4, 2, 3, 3, 0, 1, 2, 2

¿Podemos afirmar con toda seguridad que en el 95% de los casos no existen diferencias significativas entre las habilidades de comunicación de los padres y el hecho de que sus hijos tengan o no refuerzo educativo?

Capítulo 7.

ANÁLISIS DE LA VARIANZA (ANOVA)

Javier J. Maquilón Sánchez
Tomás Izquierdo Rus

INTRODUCCIÓN

El análisis de la varianza, también denominado ANOVA, es una técnica que permite conocer si uno o más variables dependientes (variables a explicar) están en relación con una o varias variables independientes (variables explicativas). A diferencia de la prueba t, se aplica a más de dos grupos para comparar las medias entre ellas. Pueden darse cuatro situaciones:

- a) Que exista una única variable dependiente y una o más variables independientes discretas (**análisis de la varianza o ANOVA de 1 factor**). Las variables independientes se llaman factores. El análisis consiste en probar si las diferencias de variación en cada grupo (o muestra), definidas por los diferentes niveles de variable independiente, se alejan de manera significativa del valor 0.
- b) Que exista una única variable dependiente, pero una o más variables independientes continuas, entonces se recurre a **modelo de regresión lineal simple o múltiple**. La regresión consiste en evaluar los parámetros de una función lineal, es decir estimar los valores de la variable que deben explicarse y probar si las diferencias (los residuos) entre valores medidos y valores estimados se alejan significativamente de 0.
- c) Que existan algunas variables independientes discretas (discontinuas) y otras continuas (**análisis de la covariación**). El modelo consta a la vez de factores y variables explicativas o regresoras. Se busca si las relaciones entre la variable que debe explicarse (dependiente) y las variables explicativas (independientes) continuas siguen siendo las mismas para todos los niveles de las variables explicativas discretas.

- d) Cuando hay varias variables que explicar (dependientes) simultáneamente, en vez de llevar a cabo varios análisis de varianza (uno por cada variable que debe explicarse), se realiza una **análisis factorial de la varianza**. Se tratará de probar si las diferencias de variación en cada grupo definido por los diferentes niveles de las variables explicativas se alejan de manera significativa del valor 0, y eso de manera simultánea en el conjunto de las variables explicativas o independientes (de ahí la diferencia con el análisis de variación simple).

1. OBJETIVOS

Los objetivos que pretendemos con el presente capítulo son:

1. Analizar resultados de investigaciones en las que el número de grupos a comparar sea más de dos.
2. Conocer las diferencias de medias entre poblaciones.
3. Realizar generalizaciones a k poblaciones, de tal manera que el análisis de la varianza es una extensión de la clásica prueba de comparación de medias de dos muestras (T-Test).
4. Diferenciar ANOVA de un factor y ANOVA de dos factores fijos.

2. CONTENIDOS

Antes de proceder a explicar el uso del módulo ANOVA haremos una breve introducción a la lógica general del análisis de la varianza. Este método debe su nombre al hecho que utilice medidas de varianza, con el fin de determinar el carácter significativo, o no, de las diferencias de media.

El análisis de la varianza tiene como objetivo la comparación de las medias poblaciones, a partir de muestras aleatorias e independientes tomadas en cada una de ellas, lo cual nos permitirá definir las fuentes de variabilidad. Hablaremos de ANOVA de un factor fijo o aleatorio

Esta técnica se basa en la demostración matemática de que los datos muestrales (datos procedentes de cada uno de los grupos a comparar) producen dos estimaciones independientes de la varianza poblacional:

1. Varianza dentro de los grupos o intragrupos (error). Esta varianza se basa en cómo cada puntuación se aleja de la media de su grupo.
2. Varianza entre los grupos o intergrupos. Esta varianza se basa en la diferencia que existe entre las medias de cada uno de los grupos y la media total.

En el supuesto de que todas las muestras procediesen de la misma población, es decir no hubiese habido efecto de la variable independiente las varianzas de error y entre grupos serían iguales y su suma igual a la varianza total o varianza poblacional.

2.1. ANOVA de un factor

Se aplica para comparar las medias procedentes de más de dos grupos. Los pasos a seguir en este análisis son:

- 1º. Abra el fichero desde *Startup*, con la secuencia File/Open/Data.
- 2º. Seleccione la opción Analysis of Variance del menú Analyze. En la ventana de diálogo (Figura 7.1) traslade la variable dependiente al cuadro Dependent y la variable GRUPO al cuadro Factors.
- 4º. Pulse Ok.
- 5º. Proceda a grabar los resultados con la secuencia File/Save As introduciendo en la ventana de diálogo la unidad, la ruta y el nombre del fichero.

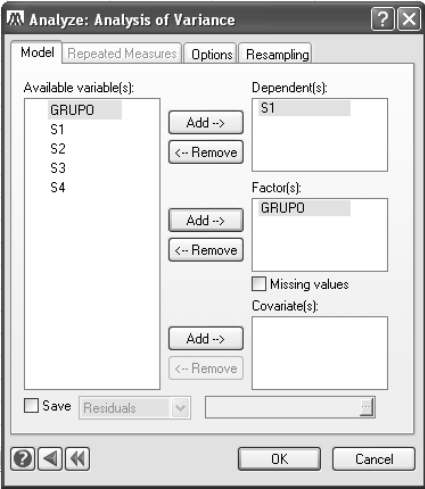


Figura 7.1. ANOVA de un factor.

2.2. ANOVA de dos factores

El análisis de varianza con más de un factor estudia los datos de la variable dependiente desde la perspectiva de más de una variable independiente por separado (efectos principales) y de forma conjunta (efectos de interacción). El modelo de este análisis queda como sigue:

		FACTOR A (4)			
		A1	A2	A3	A4
FACTOR B (3)	B1	A1B1	A2B1	A3B1	A4B1
	B2	A1B2	A2B2	A3B2	A4B2
	B3	A1B3	A2B3	A3B3	A4B3

Tabla 7.1. Distribución de factores.

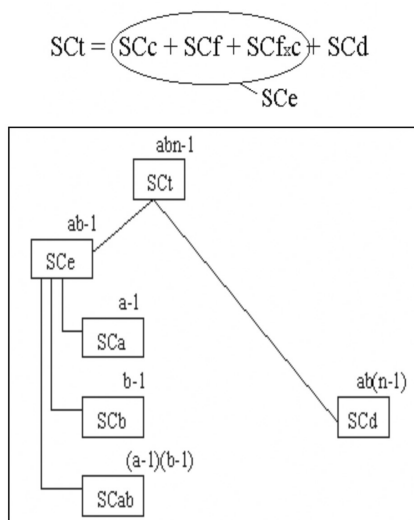


Figura 7.2 Descomposición de la suma de cuadrados total.

Este análisis de la varianza contempla dos variables independientes o factores y una variable dependiente. El proceso de cálculo es similar al de un ANOVA de un factor sólo que en el apartado Factors hemos de incluir las dos variables independientes. En el ejemplo son GRUPO y SEXO (Figura 7.3).

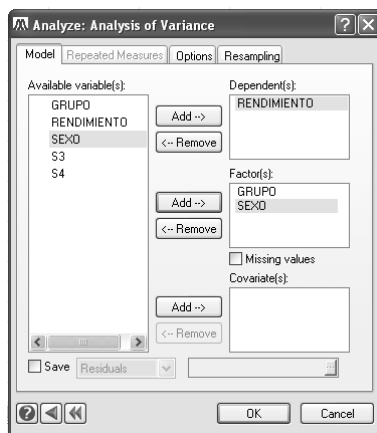


Figura 7.3. ANOVA de dos factores.

Una vez elegidas las variables y pulsado OK, observará la salida por pantalla de los resultados. Proceda a grabar los resultados con la secuencia File/Save As (Figura 7.4) introduciendo en la ventana de diálogo la unidad, la ruta y el nombre del fichero.

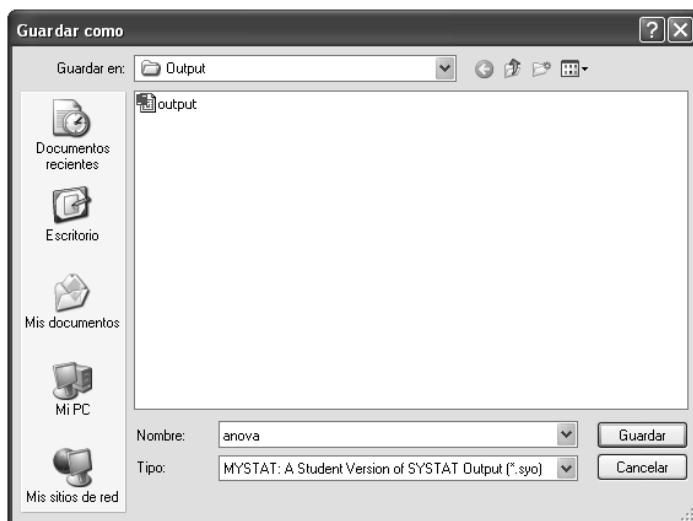


Figura 7.4 Save As en Startpage.

3. APLICACIÓN PRÁCTICA SIN ORDENADOR

3.1. ANOVA de un factor

Observe el siguiente ejemplo: existen 3 grupos a los cuales hemos aplicado diferentes metodologías de enseñanza en la asignatura “Análisis de Datos” y han obtenido los resultados que aparecen reflejados en la tabla 7.2. Con la prueba ANOVA intentamos ver si existe o no diferencia entre los grupos después de haberles aplicado a cada uno de ellos diferentes tratamientos (metodologías de enseñanza). La hipótesis nula supone que las medias de los tres grupos son iguales. La hipótesis alternativa establece que h_0 (hipótesis nula) no es cierta.

GRUPO 1		GRUPO 2		GRUPO 3	
SUJETOS	X	SUJETOS	X	SUJETOS	
1	5	1	13	1	6
2	3	2	9	2	6
3	4	3	10	3	5
4	4	4	8	4	5
5	6	5	9	5	3

Tabla 7.2 Metodologías de enseñanza.

$$x_1 = 4.4$$

$$x_3 = 5$$

$$x_2 = 9.8$$

$$x_{\text{total}} = 6.4$$

El primer paso es el cálculo de la suma de cuadrados: suma de cuadrados total, suma de cuadrados intergrupo y suma de cuadrados intragrupo.

Suma de cuadrados total

$$SC_t = \sum (x - \bar{x})^2$$

$$SC_t = (5-6.4)^2 + (3-6.4)^2 + (4-6.4)^2 + (4-6.4)^2 + (6-6.4)^2 + (13-6.4)^2 + (9-6.4)^2 + (10-6.4)^2 + (8-6.4)^2 \dots$$

Suma de cuadrados intergrupo

$$SC_{inter} = \frac{(\sum x_1)^2}{N_1} + \frac{(\sum x_2)^2}{N_2} + \frac{(\sum x_k)^2}{N_k} - \frac{(\sum x)^2}{N}$$

$(\sum x_i)^2$ = sumatorio de las puntuaciones en un grupo

K= número de grupos

N_i =número de puntuaciones en un grupo

$$SC_{inter} = \frac{(22)^2}{5} + \frac{(49)^2}{5} + \frac{(96)^2}{15}$$

Suma de cuadrados intragrupo

$$SC_{intra} = SC_{inter} - SC_{tot}$$

A continuación procedemos al cálculo de las medias cuadráticas, dividiendo la suma cuadrática intergrupo y la suma cuadrática intragrupo entre sus grados de libertad (gl) respectivos, es decir los grados de libertad intergrupo, cuya fórmula es K-1, siendo K el número de grupos y los grados de libertad intragrupo, N-K, siendo N el número total de observaciones.

$$MC_{inter} = \frac{SC_{inter}}{2}$$

$$MC_{intra} = \frac{SC_{intra}}{12}$$

Una vez calculada las medias cuadráticas, calculamos la razón F:

$$F = \frac{MC_{inter}}{MC_{intra}}$$

Para obtener el valor de F tenemos que consultar la tabla de valores críticos de la distribución F (Tabla 7.3) en la cual encontraremos los grados de libertad inter e intra obtenidos anteriormente, siendo los grados de libertad inter los correspondientes al numerador y los grados de libertad intra los correspondientes al denominador. En nuestro ejemplo, es 2gl para el numerador y 12 gl para el denominador.

A continuación comparamos el valor resultante obtenido del cociente (F) con el valor reflejado en la tabla para los grados de libertad anteriormente obtenidos en la tabla, en nuestro caso sería 3.89. Si el valor resultante del cociente F es menor que el valor reflejado en la tabla aceptamos h_0 , concluyendo que no existen diferencias significativas, por el contrario si el valor es superior rechazamos h_0 , aceptando la hipótesis alternativa sobre la existencia de diferencias significativas.

$$\alpha = 0.05$$

VALORES CRÍTICOS DE LA DISTRIBUCIÓN F

Grados de libertad	Nivel de significación para un extremo de la prueba							
	0.250	0.100	0.050	0.025	0.010	0.005	0.0025	0.0005
	Nivel de significación para los dos extremos de la prueba							
	0.500	0.200	0.100	0.050	0.020	0.010	0.005	0.001
21	0.686	1.323	1.721	2.080	2.518	2.831	3.135	3.819
22	0.686	1.321	1.717	2.074	2.508	2.819	3.119	3.792
23	0.685	1.319	1.714	2.069	2.500	2.807	3.104	3.768
24	0.685	1.318	1.711	2.064	2.492	2.797	3.091	3.745
25	0.684	1.316	1.708	2.062	2.485	2.787	3.078	3.725
26	0.684	1.315	1.706	2.056	2.479	2.779	3.067	3.707
27	0.684	1.314	1.703	2.052	2.473	2.771	3.057	3.690
28	0.683	1.313	1.701	2.048	2.467	2.763	3.047	3.674
29	0.683	1.311	1.699	2.045	2.462	2.756	3.038	3.659
30	0.683	1.310	1.697	2.042	2.457	2.750	3.030	3.646
35	0.682	1.306	1.690	2.030	2.438	2.724	2.996	3.591
40	0.681	1.303	1.684	2.021	2.423	2.704	2.971	3.551
45	0.680	1.301	1.679	2.014	2.412	2.690	2.952	3.520
50	0.679	1.299	1.676	2.009	2.403	2.678	2.937	3.496
55	0.679	1.297	1.673	2.004	2.396	2.668	2.925	3.476
60	0.679	1.296	1.671	2.000	2.390	2.660	2.915	3.460
65	0.678	1.295	1.669	1.997	2.385	2.654	2.906	3.447
70	0.678	1.294	1.667	1.994	2.381	2.648	2.899	3.435
80	0.678	1.292	1.664	1.990	2.374	2.639	2.887	3.416
90	0.677	1.291	1.662	1.987	2.368	2.632	2.878	3.402
100	0.677	1.290	1.660	1.984	2.364	2.626	2.871	3.390
125	0.676	1.288	1.657	1.979	2.357	2.616	2.858	3.370
150	0.676	1.287	1.655	1.976	2.351	2.609	2.849	3.357
200	0.676	1.286	1.653	1.972	2.345	2.601	2.839	3.340
∞	0.6745	1.2816	1.6448	1.9600	2.3267	2.5758	2.8070	3.2905

Tabla 7.3. Valores críticos de la F.

3.2. ANOVA de dos factores

Continuando con el ejemplo anterior, supongamos que queremos conocer los efectos de las tres metodologías de enseñanza y el género en relación a las calificaciones finales obtenidas en la asignatura de “Análisis de Datos”. La hipótesis en este caso sería: diferentes metodologías pueden afectar de manera diferente a hombres y mujeres. Estaríamos ante un doble análisis de la varianza, también denominado diseño factorial 3x2 (Tabla 7.4) porque existen 3 niveles de variable, la independiente (metodología) y dos de la segunda (género).

FACTOR 2	FACTOR 1		
	GRUPO 1	GRUPO 2	GRUPO 3
HOMBRE	5	13	6
	3	9	6
MUJER	4	10	5
	4	8	5
	6	9	3

Tabla 7.4. Diseño factorial 3x2.

El procedimiento de cálculo es muy similar al ANOVA de un factor.

Suma de cuadrados total

$$SC_t = \sum (x - \bar{x})^2$$

$$SC_t = (5-6.4)^2 + (3-6.4)^2 + (4-6.4)^2 + (4-6.4)^2 + (6-6.4)^2 + (13-6.4)^2 + (9-6.4)^2 + (10-6.4)^2 + (8-6.4)^2 + \dots$$

Suma de cuadrados intergrupo

$$SC_{\text{inter}} = \frac{(\sum x_1)^2}{N_1} + \frac{(\sum x_2)^2}{N_2} + \frac{(\sum x_k)^2}{N_k} - \frac{(\sum x)^2}{N}$$

$$SC_{\text{inter}} = \frac{(8)^2}{2} + \frac{(22)^2}{2} + \frac{(96)^2}{15}$$

Suma de cuadrados intragrupo

$$SC_{\text{intra}} = SC_{\text{inter}} - SC_{\text{tot}}$$

Suma de cuadrados para el factor 1

$$SC_{\text{inter}} = \sum \frac{(\text{columnas})^2}{N_{\text{columnas}}} - \frac{(\sum x)^2}{N}$$

$$SC_1 = \frac{(22)^2}{5} + \frac{(49)^2}{5} - \frac{9216}{15}$$

Suma de cuadrados para el factor 2

$$SC_2 = \sum \frac{(\text{filas})^2}{N_{\text{filas}}} - \frac{(\sum x)^2}{N}$$

$$SC_2 = \frac{(42)^2}{6} + \frac{(54)^2}{6} - \frac{9216}{15}$$

Suma de cuadrados para la interacción para el factor 1 y 2

$$SC_{1 \times 2} = SC_{inter} - SC_1 - SC_2$$

A continuación, procedemos al cálculo de las medias cuadráticas, aunque a diferencia del procedimiento anterior calcularemos las medias cuadráticas intragrupo, las medias cuadráticas para las tres metodologías y para el género y las medias cuadráticas de la interacción.

$$MC_{intra} = \frac{SC_{intra}}{gl_{intra}}$$

Los grados de libertad intragrupo se obtienen restando al número total de observaciones, en nuestro ejemplo $N=15$, el número de grupos tanto de hombres como de mujeres repartidos para cada una de las metodologías, en nuestro ejemplo $K=6$.

Media cuadrática para el factor 1

$$MC_1 = \frac{SC_1}{gl_1}$$

Los grados de libertad en este caso se obtienen restando el número de grupos menos 1, en nuestro ejemplo los grados de libertad es igual a 3.

Media cuadrática para el factor 2

$$MC_2 = \frac{SC_2}{gl_2}$$

Los grados de libertad en este caso se obtienen restando el número de grupos menos 1, en nuestro ejemplo los grados de libertad es igual a 2.

Media cuadrática de la interacción del factor 1 y 2

$$MC_{1 \times 2} = \frac{SC_{1 \times 2}}{gl_{1 \times 2}}$$

Para obtener el valor de F , en este caso tendremos que calcularlo tanto para los grupos como para el género y para la interacción, es decir habrá que calcular tres hipótesis, la primera nos permitirá comprobar si existen o no diferencias entre los grupos al aplicar diferentes metodologías, la segunda hipótesis nos permitirá ver si existen diferencias o no en función del género y la tercera hipótesis nos permitirá conocer los efectos de la interacción.

En el caso de la primera F ,

$$F = \frac{MC_1}{MC_{intra}}$$

Los grados de libertad para el numerador ($N-1$), 2 y del denominador 9 ($15-6$)

En el caso de la segunda F ,

$$F = \frac{MC_2}{MC_{intra}}$$

Los grados de libertad para el numerador (N-1), 1 y del denominador 9 (15-6). Y finalmente calculamos F para la interacción:

$$F = \frac{MC_{1 \times 2}}{MC_{intra}}$$

siendo sus grados de libertad correspondientes para el numerador ($gl_{1 \times 2}=2$), obtenidos tras aplicar la siguiente fórmula: número de columnas menos 1 x número de filas menos 1 y para el denominador 9, al igual que los anteriores.

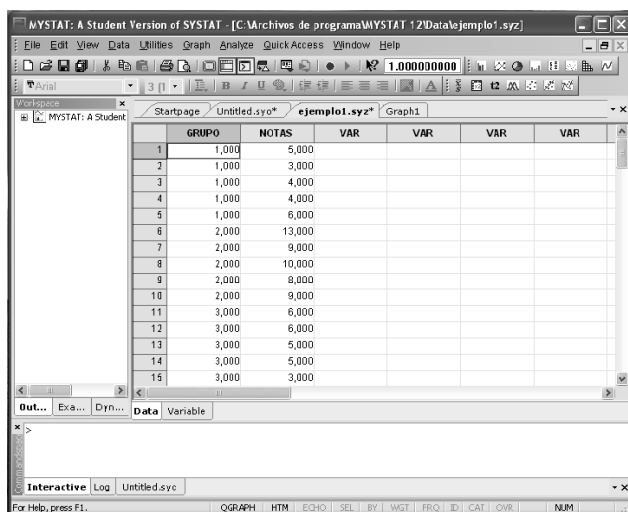
Y finalmente, como ya hicimos en el ANOVA de un factor comparamos el valor resultante obtenido del cociente F de las tres hipótesis con el valor reflejado en la tabla para los grados de libertad anteriormente obtenidos para cada uno de ellos en la tabla 6.3. Si los valores resultantes de los cocientes de F es menor que el valor reflejado en la tabla aceptamos h_0 , concluyendo que no existen diferencias significativas, por el contrario si el valor es superior rechazamos h_0 , aceptando la hipótesis alternativa sobre la existencia de diferencias significativas.

4. APLICACIÓN PRÁCTICA CON ORDENADOR

4.1. ANOVA de un factor

La aplicación con el programa Mymstat 12 del ejemplo anteriormente expuesto quedaría de la siguiente manera:

1°. Abra el fichero desde *Startpage*, con la secuencia File/Open/Data. (Figura 7.5)



	GRUPO	NOTAS	VAR	VAR	VAR	VAR
1	1,000	5,000				
2	1,000	3,000				
3	1,000	4,000				
4	1,000	4,000				
5	1,000	6,000				
6	2,000	13,000				
7	2,000	9,000				
8	2,000	10,000				
9	2,000	8,000				
10	2,000	9,000				
11	3,000	6,000				
12	3,000	6,000				
13	3,000	5,000				
14	3,000	5,000				
15	3,000	3,000				

Figura 7.5. Hoja de datos del fichero.

- 2º. Seleccione la opción Analysis of Variance del menú Analyze. En la ventana de diálogo (Figura 7.6) traslade la variable dependiente al cuadro Dependent y la variable GRUPO al cuadro Factors.

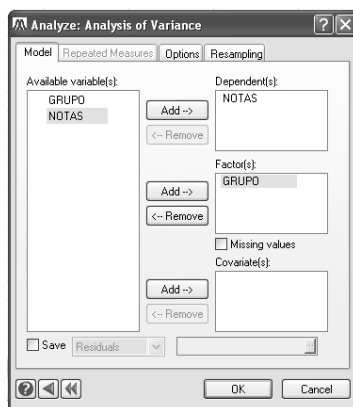


Figura 7.6. Módulo ANOVA.

- 3º. Pulse Ok.
- 4º. Obtenemos los siguientes resultados (Figura 7.7), que según aparecen en nuestra ventana confirman la existencia de diferencias significativas entre los tres grupos ($.000 < .05$).

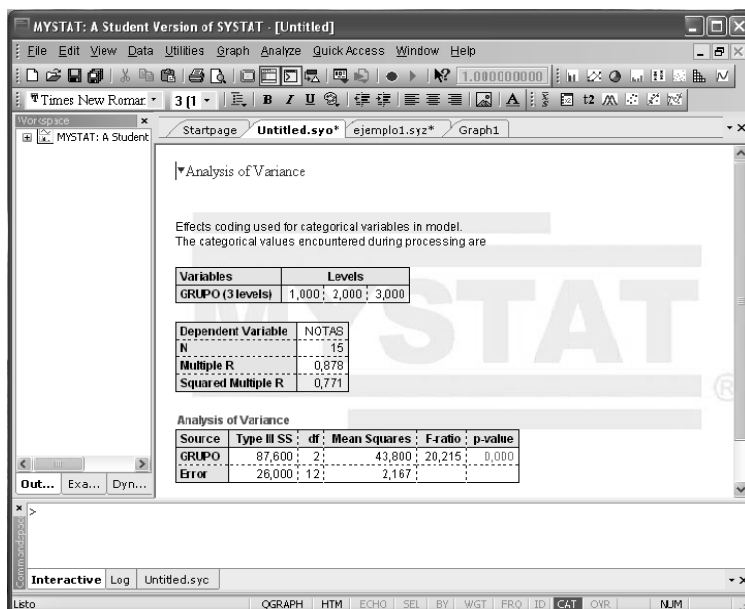
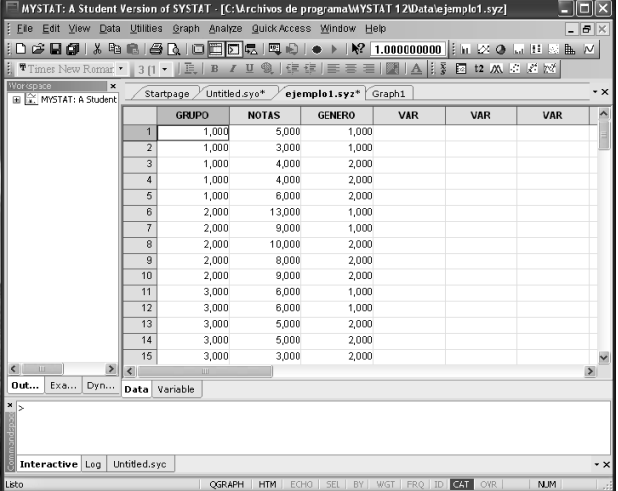


Figura 7.7. Resultados del ANOVA de un factor.

4.2. ANOVA dos factores

1º. Abra el fichero desde *Startpage*, con la secuencia File/Open/Data. (Figura 7.8).



	GRUPO	NOTAS	GENERO	VAR	VAR	VAR
1	1,000	5,000	1,000			
2	1,000	3,000	1,000			
3	1,000	4,000	2,000			
4	1,000	4,000	2,000			
5	1,000	6,000	2,000			
6	2,000	13,000	1,000			
7	2,000	9,000	1,000			
8	2,000	10,000	2,000			
9	2,000	8,000	2,000			
10	2,000	9,000	2,000			
11	3,000	6,000	1,000			
12	3,000	6,000	1,000			
13	3,000	5,000	2,000			
14	3,000	5,000	2,000			
15	3,000	3,000	2,000			

Figura 7.8. Hoja con los datos del fichero.

2º. Seleccione la opción Analysis of Variance del menú Analyze. En la ventana de diálogo (Figura 7.9) traslade la variable dependiente al cuadro Dependent y las variables GRUPO y SEXO al cuadro Factors.

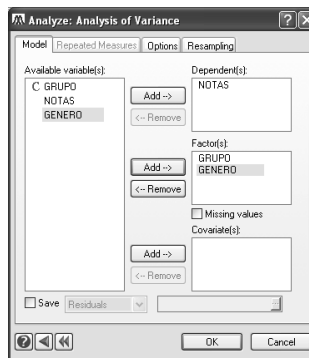


Figura 7.9. Módulo ANOVA.

3º. Pulse Ok.

4º. Obtenemos los siguientes resultados (Figura 7.10), que según aparecen en nuestra ventana confirman nuevamente la existencia de diferencias significativas en cuanto a las calificaciones por grupo ($.000 < .05$), pero no por género ($.205 < .05$), ni tampoco debido a la interacción grupo x género ($.315 < .05$),

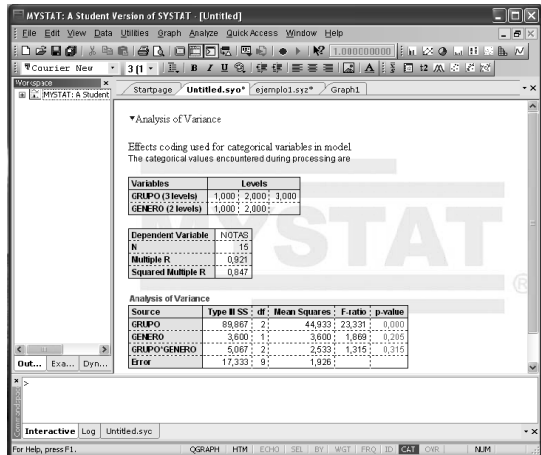


Figura 7.10. Resultados del ANOVA de dos factores.

5. EJERCICIOS PROPUESTOS

5.1. Se dispone de tres muestras que comprenden a 5 individuos, cuyos valores se dan en el siguiente cuadro. Queremos conocer si las diferencias de tiempo de ejecución de tres tipos de test diferentes son significativas o no.

TEST 1	TEST 2	TEST 3
7	3	0
6	3	0
6	3	1
5	4	1
5	5	2

5.2. Se quiere conocer si la cantidad de horas dedicadas al estudio durante el segundo cuatrimestre de la titulación de Grado en Educación Social varía de unas asignaturas a otras. Para ello, tomamos como referencia 10 sujetos (n=10) y 3 asignaturas diferentes (k=3).

Asignatura 1	Asignatura 2	Asignatura 3
50.00	16.00	12.00
52.00	35.00	12.00
12.00	12.00	12.00
10.00	32.00	22.00
20.00	11.00	25.00
25.00	20.00	14.00
22.00	40.00	18.00
22.00	16.00	17.00
30.00	16.00	16.00
22.00	25.00	21.00

5.3. El director de un centro educativo desea saber si los años de experiencia docente de los profesores (A1: menos de dos años; A2: más de dos años) y tres métodos de enseñanza influyen en el rendimiento de una asignatura. 12 alumnos son asignados al azar a las 6 condiciones referidas. Su rendimiento académico al final de curso fue el siguiente:

	B1	B2	B3
A1	4	4	8
	3	3	10
A2	2	5	10
	1	6	10

- ¿Influye la experiencia docente del profesor sobre el rendimiento?
- ¿Influyen los métodos de enseñanza?
- ¿Hay interacción?

5.4. Un profesor de la Facultad de Ciencias de la Actividad Física y del Deporte (CAFD) desea comparar la eficacia de dos tipos de entrenamiento deportivo en 8 estudiantes del Grado en CAFD (A1: entrenamiento físico; A2: entrenamiento físico y psicológico). Asigna al azar 4 sujetos a cada condición. A lo largo de 4 semanas (B1, B2, B3, B4) los 8 deportistas fueron evaluados en un test de rendimiento deportivo (escala de 0 a 15 puntos). Los resultados fueron:

	B1	B2	B3	B4
A1	3	4	7	7
	6	5	8	8
	3	4	7	9
	3	3	6	8
A2	1	2	5	10
	2	3	6	10
	2	4	5	9
	2	3	6	11

Analiza e interpreta los resultados.

5.5. Una muestra elegida al azar de 6 estudiantes de la asignatura *Investigación y TIC* leen 1, 3 y 5 veces una lista de 50 palabras al objeto de memorizarlas. Tras cada lectura se les pasa una tarea de recuerdo. Los resultados (o aciertos) fueron:

- 1 lectura: 15, 17, 14, 18, 18, 16
 3 lecturas: 21, 25, 22, 24, 29, 27
 5 lecturas: 28, 32, 34, 35, 30, 30

- ¿Incrementa el número de lecturas el recuerdo?
- ¿Dónde se dan los mejores y peores resultados?

Capítulo 8.

CORRELACIÓN Y REGRESIÓN

Fuensanta Monroy Hernández
Javier J. Maquilón Sánchez

INTRODUCCIÓN

Este capítulo explica cómo llevar a cabo el análisis de correlación y regresión. En el caso del análisis de correlación, conoceremos la relación entre dos o más variables. Dependiendo del tipo de variables se utilizarán coeficientes diferentes: para variables continuas se usará el coeficiente r de Pearson, para variables ordinales el coeficiente de Spearman y para variables dicotómicas el coeficiente Phi... En los siguientes apartados sólo introduciremos los coeficientes de Pearson y Spearman. En cuanto a la regresión, se utiliza para predecir valores de una de las variables a partir de las otras.

1. OBJETIVOS

Los objetivos que pretendemos con el presente capítulo son:

1. Describir la relación entre dos variables X e Y empleando el coeficiente de Pearson y Spearman.
2. Realizar diagramas de dispersión y rectas de regresión como representación gráfica de una correlación y una regresión respectivamente.
3. Predecir los valores de una variable, dados los valores de otra variable que correlaciona con ella (análisis de regresión).

2. CONTENIDOS

El análisis de correlación permite conocer la relación entre dos variables cuantitativas. Normalmente una de las variables es la causa (x) y la otra variable es la

consecuencia (y) mediante una función f , $y=f(x)$. Cuando se tiene una relación positiva y perfecta todos los puntos se ajustan a una línea recta ($r=1$), si la relación es negativa y perfecta ($r=-1$) y si no hay ninguna relación ($r=0$). Los supuestos más importantes son que para cada valor de (x) existe una población de valores de (y) con distribución normal y que cada una de las distribuciones tiene la misma varianza, lo que se aplica también a (y) con relación a (x).

El análisis de regresión, $y=f(x_1, x_2, \dots, x_p) + \varepsilon$, es un método que estudia la variabilidad de una variable dependiente o criterio en relación con la variabilidad de una variable independiente o predictor (regresión simple), $y = \beta_0 + \beta_1 x + \varepsilon$, o más de una variable independiente (regresión múltiple), $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_p x_p + \varepsilon$.

La pregunta central de este análisis es cómo varía la media de la distribución de la variable dependiente condicionada a los valores de la variable independiente. El análisis de regresión es muy útil para predecir el valor de una variable (y) si se conoce el valor de otra variable (x). Los valores de la variable dependiente son una función de la combinación lineal de los valores de la/s variable/s independientes y sus pesos.

2.1. Coeficiente de Pearson

Este coeficiente viene representado por la letra r y se utiliza cuando las variables vienen expresadas en medidas continuas. Para su cálculo debe proceder del modo siguiente:

- 1º. Abra el fichero desde Startpage/File/Open/Data.
- 2º. Seleccione la opción Correlations del menú Analyze. Marque la opción Simple del submenú desplegable. Aparecerá una ventana de diálogo (figura 8.1); en ella elija las variables que desea correlacionar y después pulse el botón Add.



Figura 8.1. Correlación de Pearson.

- 3°. Debajo de dicha ventana observará dos opciones: Listwise y Pairwise. La primera calcula la correlación entre los pares de variables usando el mismo número de casos y con la segunda se eliminan aquellos casos a los que falta un dato en cualquiera de las dos variables. Por defecto aparece la primera opción, Listwise.
- 4°. Pulse el botón OK y verá la matriz de correlaciones en la ventana Untitled.
- 5°. Proceda a grabar los resultados con la secuencia File/Save As (figura 8.2) introduciendo en la ventana de diálogo la unidad, la ruta y el nombre del fichero.

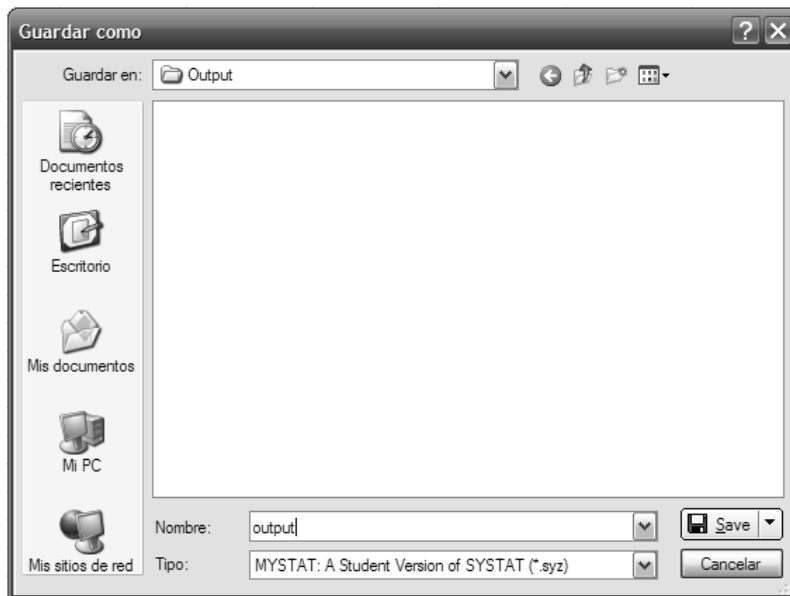


Figura 8.2. Ventana Guardar como en Startpage.

2.2. Coeficiente de Spearman

El coeficiente de Spearman es un coeficiente de correlación para variables ordinales. El proceso de cálculo es muy similar al coeficiente de Pearson.

- 1°. Abra el fichero desde Startpage/File/Open/Data.
- 2°. Seleccione la opción Correlations del menú Analyze y marque la opción Simple del submenú desplegable. En opciones escoja Rank order data: Spearman en la ventana de diálogo que aparece (figura 8.3) y elija las variables que desea correlacionar.
- 3°. Pulse el botón OK.

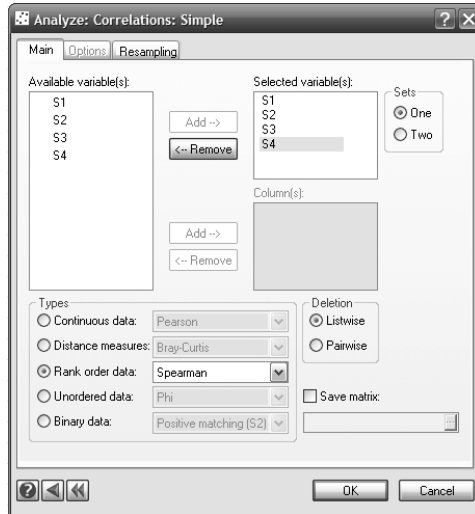


Figura 8.3. Correlación de Spearman.

4º. Grabe los resultados con la secuencia File/Save As (figura 8.4) introduciendo en la ventana de diálogo la unidad, la ruta y el nombre del fichero.

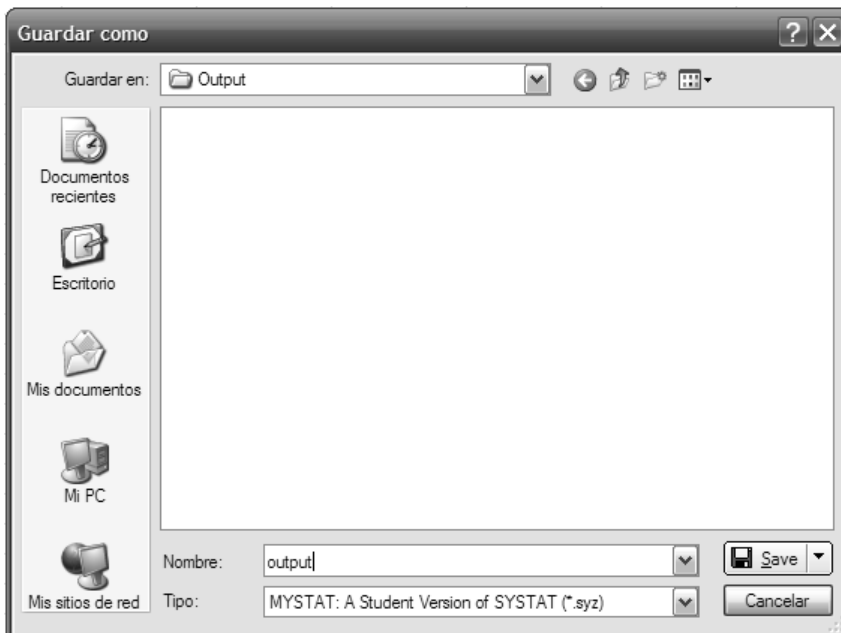


Figura 8.4. Ventana Guardar como en Startpage.

2.3. Diagramas de dispersión

Cada uno de los coeficientes de correlación calculados anteriormente puede ser representado gráficamente a través de un diagrama de dispersión utilizando la opción gráfica Scatterplot.

Para ello proceda del modo siguiente:

- 1°. Abra el fichero con el que vaya a realizar los análisis con la secuencia File/ Open/Data.
- 2°. Cuando seleccione la opción Scatterplot del menú Graph, aparecerá una ventana de diálogo como la observada en la figura 8.5.
- 3°. Marque las variables dependientes y pulse el botón con la inscripción Y-> trasladándolas al rectángulo de la derecha. A continuación marque la variable independiente y pulsa el botón con la inscripción X-> trasladándolas al rectángulo de la derecha.

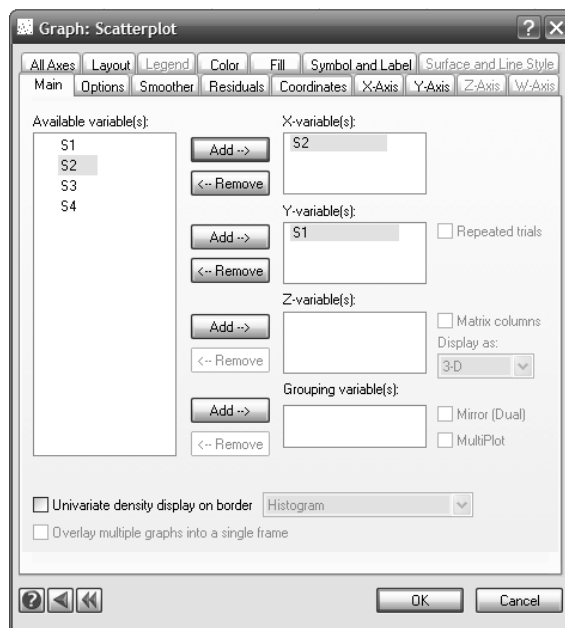


Figura 8.5. Gráfico Scatterplot.

- 4°. Pulse OK y aparecerá gráficamente el diagrama de dispersión de las variables que haya seleccionado.
- 5°. Grabe el fichero desde la ventana File/Save active file en la que se encuentre (figura 8.6).

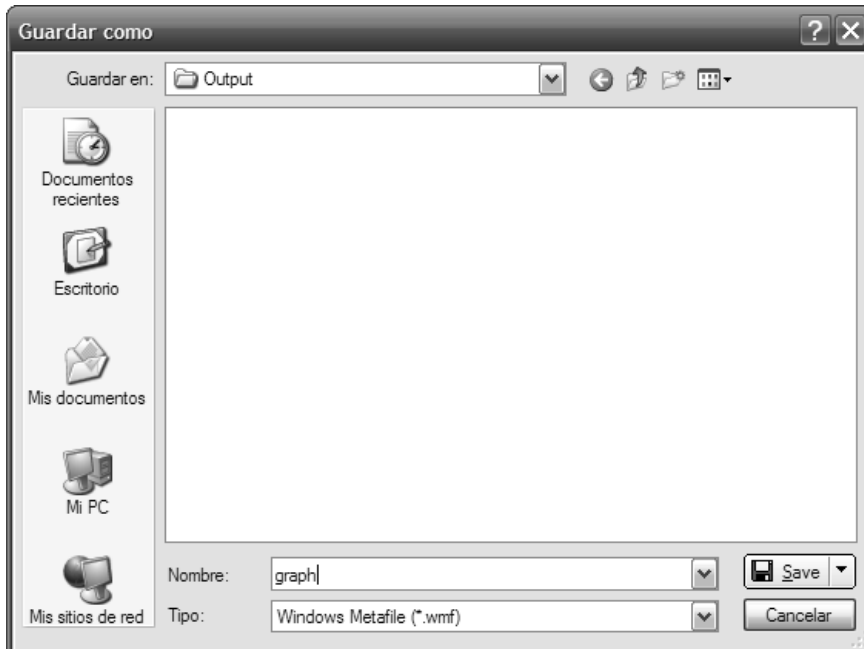


Figura 8.6. Ventana Save active file en Graph.

2.4. Análisis de regresión simple

El análisis de regresión es un método que estudia la variabilidad de una variable dependiente o criterio en relación con la variabilidad de una variable independiente o predictor (regresión simple) o más de una variable independiente (regresión múltiple).

En términos causales, intenta establecer que los cambios producidos en la variable dependiente vienen determinados por los cambios en la variable independiente. Para calcular el coeficiente de regresión debe proceder del modo siguiente:

- 1º. Abra el fichero desde /File/Open/Data.
- 2º. Seleccione la opción Regression del menú Analyze y marque Least Squares en el desplegable. Aparecerá la ventana de diálogo como la de la figura 8.7. En ella seleccione la variable que desea utilizar como independiente y la que será variable dependiente.
- 3º. En la parte inferior izquierda hay dos opciones, asegúrese que está marcada la opción Include constant.
- 4º. Pulse el botón OK, y en la ventana Untitled.syo verá los resultados.
- 5º. Proceda a grabar los resultados con la secuencia File/Save As (figura 8.8) introduciendo en la ventana de diálogo la unidad, la ruta y el nombre del fichero.

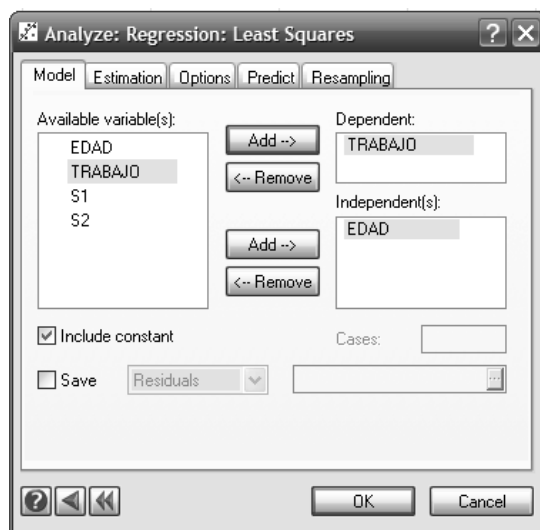


Figura 8.7. Regresión simple.

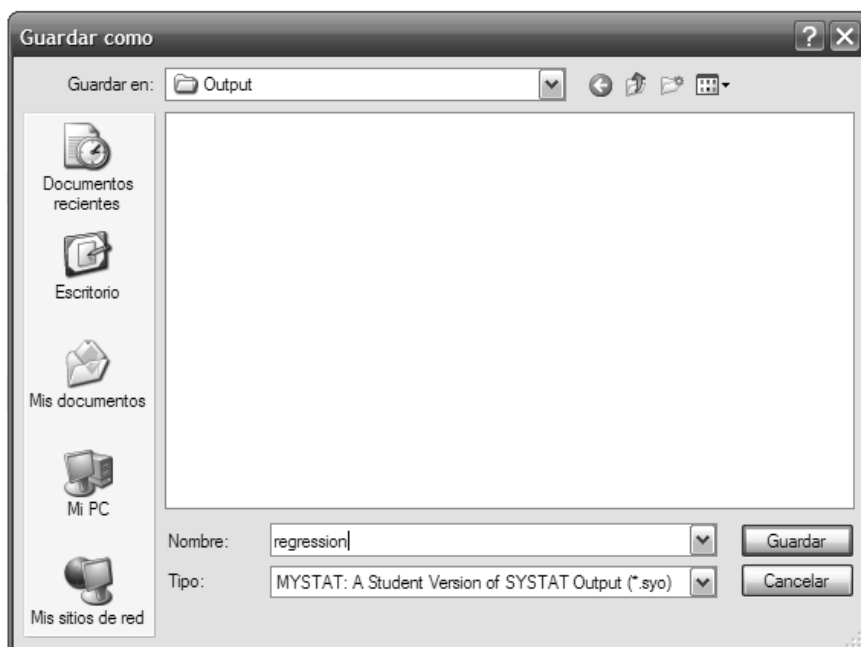


Figura 8.8. Ventana Save as en Untitled.syo.

2.5. Análisis de regresión múltiple

Cuando se incluye más de una variable independiente o predictora para predecir la variable dependiente o criterio, la línea de regresión que hemos calculado en el apartado anterior se convierte en un plano de regresión. Expresa la predicción de la variable dependiente a partir de las variables independientes. Los pasos a seguir son:

- 1°. Abra el fichero desde File/Open/Data.
- 2°. Seleccione la opción Regression del menú Analyze y marque Least Squares en el desplegable. Aparecerá la ventana de diálogo como la de la figura 8.9. En ella se selecciona la variable dependiente.
- 3°. A continuación seleccione la primera variable independiente.
- 4°. Seguidamente seleccione la segunda variable independiente.



Figura 8.9. Regresión múltiple.

- 5°. En la parte inferior izquierda hay dos opciones, asegúrese que está marcada la opción Include constant.
- 6°. Pulse el botón OK, y en la ventana Untitled.syo verá los resultados.
- 7°. Grabe los resultados con la secuencia File/Save As (figura 8.10) introduciendo en la ventana de diálogo la unidad, la ruta y el nombre del fichero.

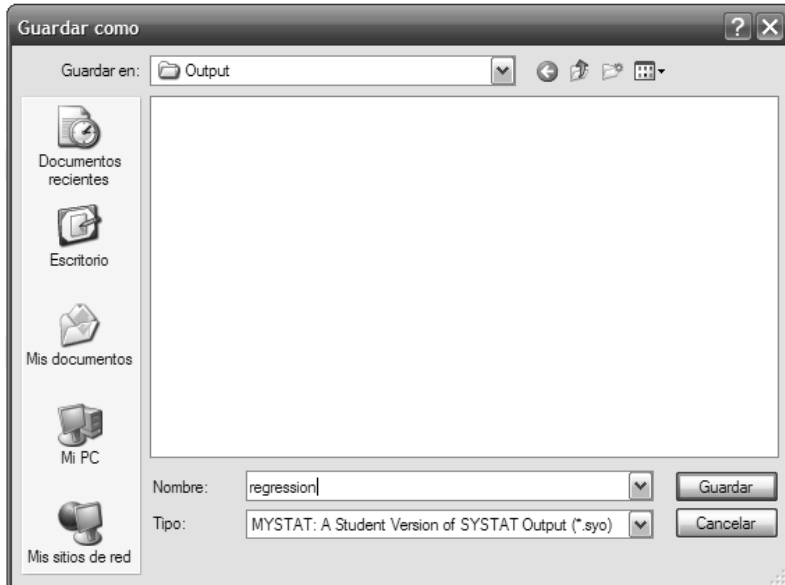


Figura 8.10. Ventana Save as en Untitled.syo.

2.6. Representación gráfica de la recta de regresión

En este apartado vamos a explicar cómo llevar a cabo la representación gráfica de la recta de regresión. Cuando la relación entre dos variables viene determinada por una función, se dice que entre ambas variables existe una relación funcional. En tal caso, todos los datos caen dentro de una línea que representa dicha función. Por ejemplo, si (x) determina el valor de (y) y ésta, a su vez, es la variable dependiente y (x) es la variable independiente, por definición decimos que $Y=f(x)$, siendo f el valor que determina la función.

El diagrama de dispersión permite detectar *casos extremos* en la relación entre las variables que pueden estar afectando al coeficiente de correlación y, por tanto, dar coeficientes inexactos de la relación real entre las variables.

2.6.1. Recta de regresión

En un diagrama de dispersión que representa la relación entre dos variables, la recta de regresión es la línea que mejor predice los valores de la variable dependiente (y) a partir de los valores de la variable independiente (x). Esta línea de regresión es la que mejor se ajusta a la suma de los cuadrados residuales. Su cálculo es como sigue:

- 1º. Abra el fichero con el que va a realizar su análisis con la secuencia File/Open/Data.

- 2°. Seleccione la opción Scatterplot del menú Graph y aparecerá la ventana de diálogo que observa en la figura 8.11.

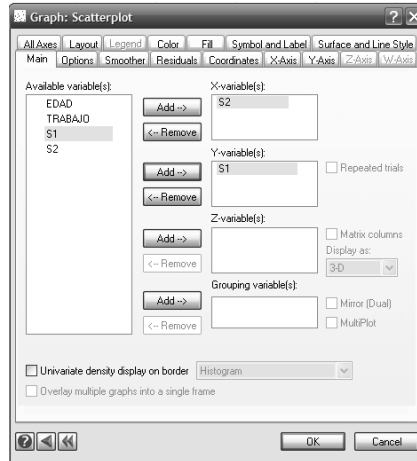


Figura 8.11. Recta de regresión.

- 3°. Seleccione las variables que trasladará a la cuadrícula marcada con y (variable dependiente) y a la cuadrícula marcada con x (variable independiente). La asignación que haga para que una variable sea x o y, va a depender del trabajo de investigación.
- 4°. Pique la pestaña Smoother. Señale la opción Linear.
- 5°. En la ventana que aparece, Confidence Intervals on regression line (Figura 8.12), escriba .95 o .99 según el nivel de confianza que seleccione. Por defecto sale .95.

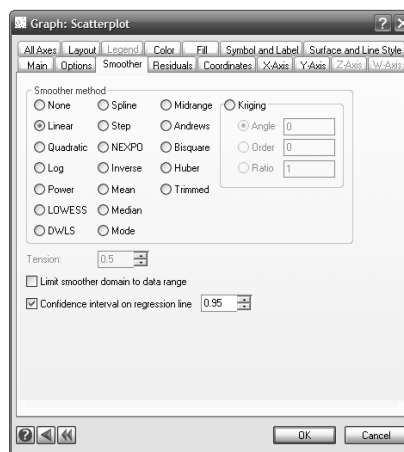


Figura 8.12. Ventana Confidence Intervals.

- 6°. Pique OK para obtener el resultado y verá una línea a través de la nube de puntos y dos líneas hiperbólicas que representan el intervalo de confianza que ha seleccionado. Esta línea central representa la mejor línea para predecir los valores de la variable dependiente a partir de los de la variable independiente. También puede modificar el aspecto del gráfico con las diferentes opciones de Graph Scatterplot.
- 7°. Puede proceder a guardar el gráfico desde la ventana Graph. Para ello siga la ruta File/Save Active File, y en la ventana de diálogo que aparece (figura 8.13) seleccione la unidad de disco (Drives), el directorio (Directories), y el nombre del fichero. En la figura 8.13, puede observar un ejemplo, en el que el fichero se grabará en la unidad C, en el subdirectorio Graph, con el nombre graph.

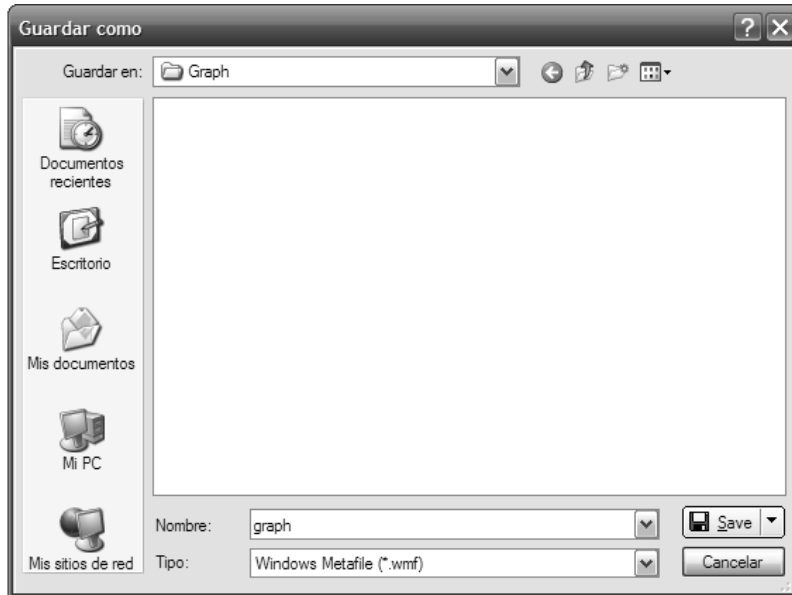


Figura 8.13. Ventana Save Active File en Graph.

El gráfico se puede imprimir directamente pulsando sobre la opción Print del menú File, no siendo esta opción incompatible con la grabación del fichero.

3. APLICACIÓN PRÁCTICA SIN ORDENADOR

3.1. Coeficiente de Pearson

Disponemos de 4 estudiantes de la titulación de Grado en Educación Social que han obtenido una serie de puntuaciones, tras haberles aplicado dos pruebas diferentes. Estos han sido los resultados:

Alumnos	X	y
Alumno 1	7	28
Alumno 2	6	25
Alumno 3	4	19
Alumno 4	3	16

Para calcular la r de Pearson empleamos la siguiente fórmula:

$$r = \frac{\sum XY - \frac{\sum X \sum Y}{N}}{\sqrt{\left(\sum X^2 - \frac{(\sum X)^2}{N}\right)\left(\sum Y^2 - \frac{(\sum Y)^2}{N}\right)}}$$

$\sum XY$: sumatorio de los productos X e Y

$\sum X$: sumatorio de las puntuaciones de X

$\sum y$: sumatorio de las puntuaciones de Y

$\sum X^2$: suma de las puntuaciones de X al cuadrado

$\sum y^2$: suma de las puntuaciones de Y al cuadrado

N: número de pares de puntuaciones

de tal manera, que el cuadro anterior quedaría de la siguiente manera:

X	Y	X ²	Y ²	XY
7	28	49	784	196
6	25	36	1296	150
4	19	16	361	76
3	16	9	256	48
$\sum X: 20$	$\sum y: 88$	$\sum X^2: 110$	$\sum y^2: 2697$	$\sum XY: 470$
$(\sum X)^2: 400$	$(\sum y)^2: 7744$			

A continuación, sustituimos los valores en la formula anteriormente expuesta y obtendremos un resultado comprendido entre -1 y +1. Si el valor es cercano a +1, la relación es positiva; un valor cercano a -1 nos indicará una relación negativa y un valor cercano a 0 nos indica ausencia de relación.

3.2. Coeficiente de Spearman

Como ya se comentó anteriormente el coeficiente de Spearman se emplea con variables ordinales. Imaginemos el caso de un grupo de estudiantes del Grado en Pedagogía a los cuales hemos clasificado en una escala de 1 a 5 tras aplicarle dos pruebas, en una mediamos su autoestima y en otra su motivación

Alumnos	X	Y
Alumno 1	2	5
Alumno 2	4	4
Alumno 3	5	1
Alumno 4	3	2
Alumno 5	1	3

A continuación aplicamos la formula r de Spearman:

$$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

Antes de comenzar a sustituir en la formula, calculamos el valor D y D²

Alumnos	X	Y	D	D ²
Alumno 1	2	5	-3	9
Alumno 2	4	4	0	0
Alumno 3	5	1	4	16
Alumno 4	3	2	1	1
Alumno 5	1	3	-2	4

Al igual que el coeficiente de Pearson, el resultado final tras aplicar la formula oscilará entre -1 y +1.

3.3. Análisis de regresión

La aplicación de la formula se emplea en aquellos casos donde tratamos de predecir o estimar los valores de Y que obtendríamos para distintos valores de X.

Imaginemos que queremos conocer la puntuación y que obtendría un quinto alumno que hubiera obtenido un 5 en la prueba x.

Alumnos	x	y
Alumno 1	7	28
Alumno 2	6	25
Alumno 3	4	19
Alumno 4	3	16

El primer paso, es aplicar la formula de los mínimos cuadrados:

$$y - \bar{y} = \frac{COV(X,Y)}{S_X^2} (x - \bar{x})$$

y: es el valor de la variables que queremos calcular

$\bar{y}$: es la media de las puntuaciones y

Cov(x,y): es la correlación existente entre ambas variables

S_x^2 : desviación típica de las puntuaciones de x

x: es el valor de la variable sobre el cual queremos calcular el valor y

$\bar{x}$: es la media de las puntuaciones y

El primer paso para el cálculo, es determinar la correlación de xy mediante

$$r = \frac{\sum XY - \frac{\sum X \sum Y}{N}}{\sqrt{\left(\sum X^2 - \frac{(\sum X)^2}{N}\right) \left(\sum Y^2 - \frac{(\sum Y)^2}{N}\right)}}$$

Una vez calculada la correlación, el siguiente paso es calcular el coeficiente de correlación lineal R^2 , el cual nos indicará la fiabilidad de la predicción del valor de y si conocemos x. La fórmula sería:

$$R^2 = (r_{xy})^2 \times 100$$

Y finalmente, sustituimos en la fórmula:

$$y - \bar{y} = \frac{COV(X,Y)}{S_X^2} (x - \bar{x})$$

El valor obtenido nos predice el valor de y si x tuviera un valor de 5, con una fiabilidad que dependerá del valor resultante tras calcular R^2 .

4. APLICACIÓN PRÁCTICA CON ORDENADOR

4.1. Coeficiente de Pearson

Para su cálculo, siga los siguientes pasos:

- 1º. Abra el fichero desde File/Open/Data empleando los mismos datos que hemos utilizado en el ejemplo realizado manualmente (Figura 8.14).

	X	Y	VAR	VAR	VAR	VAR
1	7,000	28,000				
2	6,000	25,000				
3	4,000	19,000				
4	3,000	16,000				
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						

Figura 8.14. Hoja de datos.

- 2º. Seleccione la opción Correlations Simple del menú Analyze. Aparecerá una ventana de diálogo (figura 8.15); en ella seleccione las variables que desee correlacionar y después pulse el botón Add.
- 4º. Pulse el botón OK, verá la matriz de correlaciones en la ventana Untitled.syo (figura 8.16)

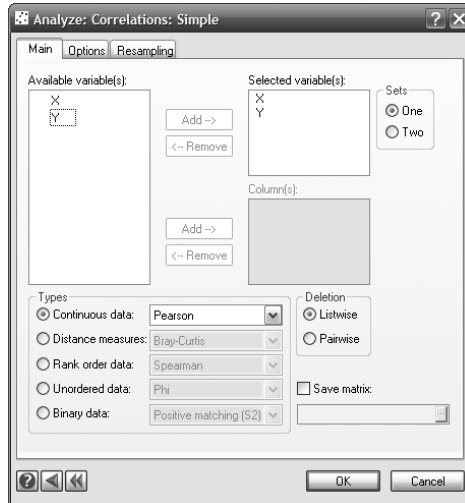


Figura 8.15 Correlación de Pearson.

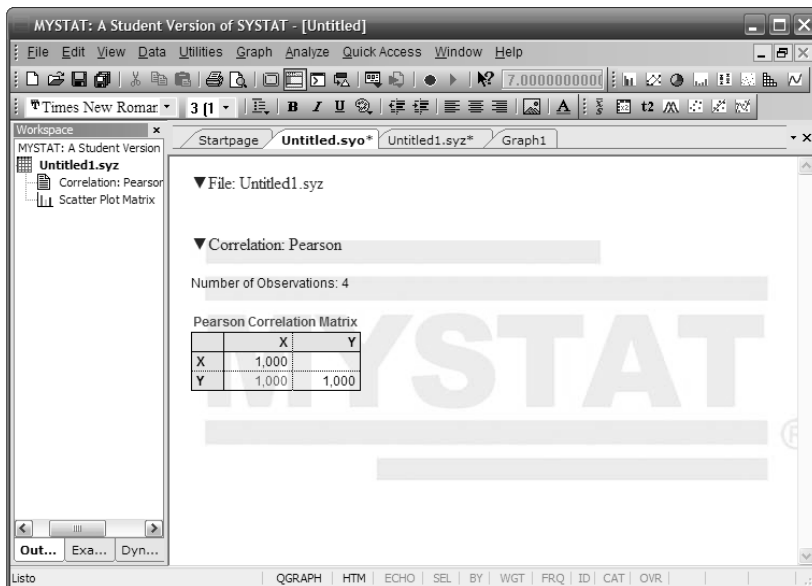


Figura 8.16. Ventana Untitled.syo tras aplicar Correlación de Pearson.

Como podemos observar en la pantalla resultante la correlación X-Y es 1.000, es una relación lineal positiva. A continuación procedemos a realizar la representación gráfica.

- 5°. Seleccione la opción Scatterplot del menú Graph. Marque las variables dependientes y pulse el botón con la inscripción Y-variable(s) trasladándolas al rectángulo de la derecha. A continuación marque la variable independiente y pulse el botón con la inscripción X-variable(s) trasladándolas al rectángulo de la derecha (Figura 8.17).

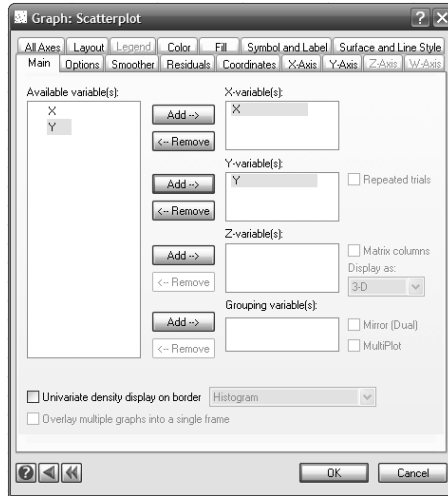


Figura 8.17. Scatterplot de Graph.

- 6°. Pulse Ok y aparecerá el gráfico (figura 8.18).

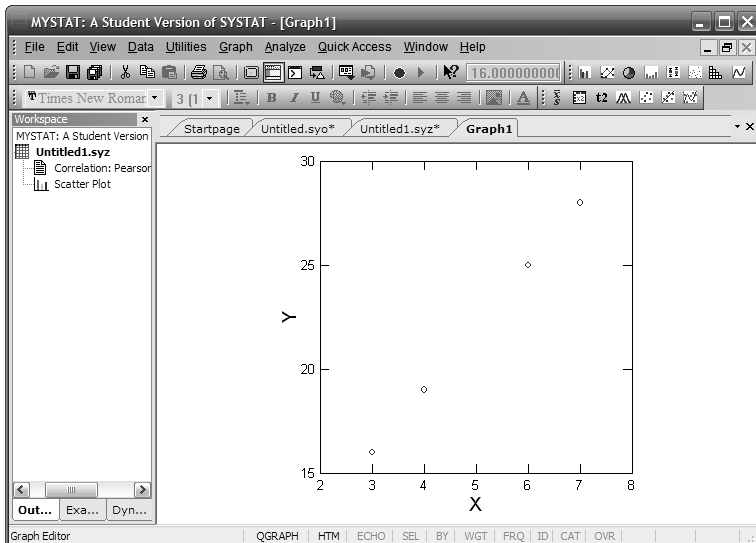


Figura 8.18. Diagrama de dispersión.

Tal y como observamos en nuestro gráfico, a medida que aumenta el valor de x aumenta el valor de y , de ahí que hablemos de una relación lineal positiva.

4.2. Coeficiente de spearman

Seguiremos el ejemplo del grupo de estudiantes de Pedagogía a los cuales hemos clasificado en una escala de 1 a 5 tras aplicarle dos pruebas, en una mediamos su autoestima y en otra su motivación

Alumnos	X	Y
Alumno 1	2	5
Alumno 2	4	4
Alumno 3	5	1
Alumno 4	3	2
Alumno 5	1	3

Los pasos a seguir son:

- 1º. Desde Correlations Simple (abriendo Analyze, de la barra de menú), pique Rank order data, con la ventana de diálogo que aparece (figura 8.19) y seleccione las variables que desea correlacionar.



Figura 8.19. Correlación de Spearman.

- 2º. Pulse el botón OK (figura 8.20)

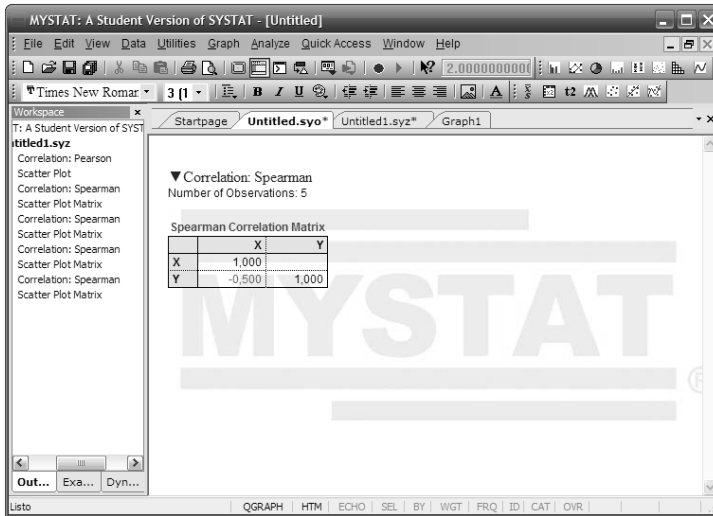


Figura 8.20. Correlación de Spearman.

Como podemos observar en la pantalla resultante la correlación x-y es -.500, siendo una relación lineal negativa moderada. A continuación procedemos a realizar la representación gráfica.

- 3°. Seleccione Scatterplot del menú Graph y marque las variables dependientes. Pulse el botón con la inscripción Y-variable(s) trasladándolas al rectángulo de la derecha. A continuación marque la variable independiente y pulsa el botón con la inscripción X-variable(s) trasladándolas al rectángulo de la derecha (Figura 8.21).

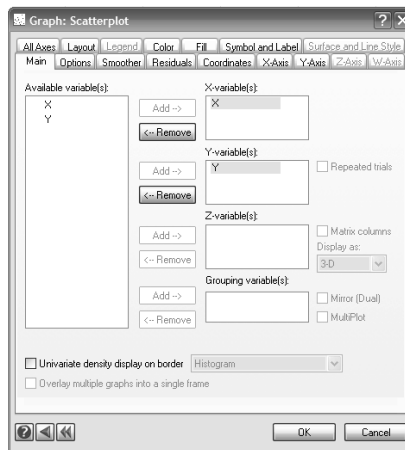


Figura 8.21. Scatterplot en Graph.

4º. Pulse OK y aparecerá el gráfico (Figura 8.22)

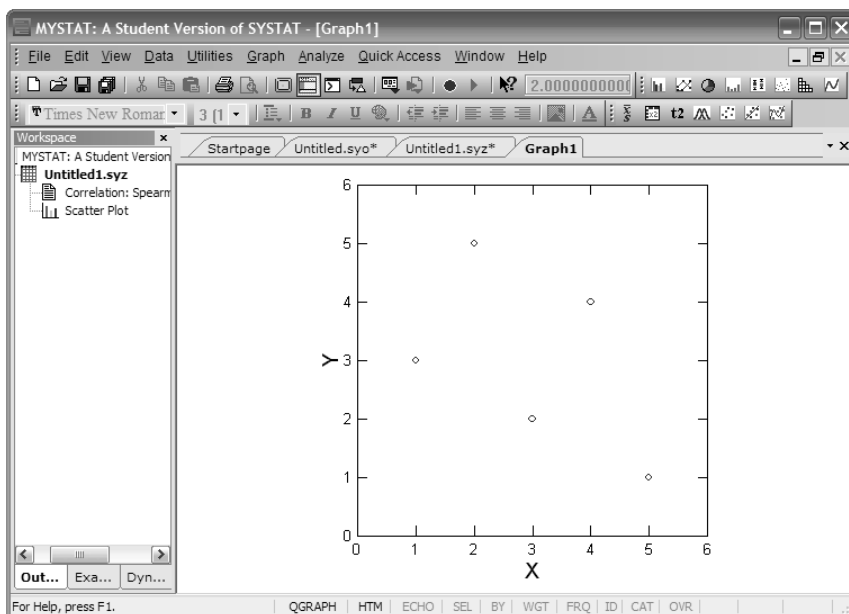


Figura 8.22. Diagrama de dispersión.

Tal y como observamos en el gráfico, a medida que aumenta el valor de una de las variables disminuye el valor de la otra, de ahí que hablemos de una relación lineal negativa.

4.3. Análisis de regresión simple

Siga los siguientes pasos:

- 1º. Abra el fichero desde File/Open/Data al cual hemos añadido algunos datos complementarios a los ya empleados en el ejemplo realizado manualmente.
- 2º. Seleccione la opción Regression del menú Analyce, marcando Least Squares del desplegable. Aparecerá la ventana de diálogo como la de la figura 8.23. En ella seleccione la variable que desea utilizar como independiente y la que será variable dependiente.
- 3º. En la parte inferior izquierda hay dos opciones, asegúrese que está marcada la opción Include constant.
- 4º. Pulse el botón OK y, en la ventana Untitled.syo, verá los resultados (Figura 8.24).

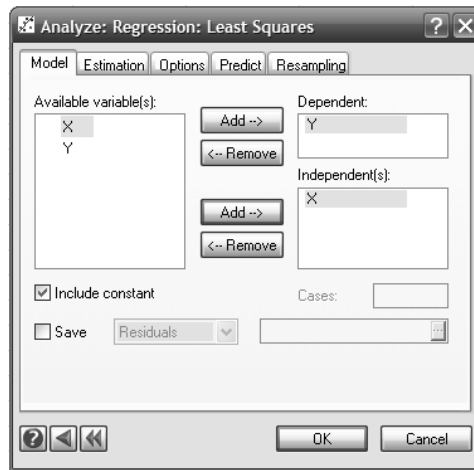


Figura 8.23. Regresión Simple.

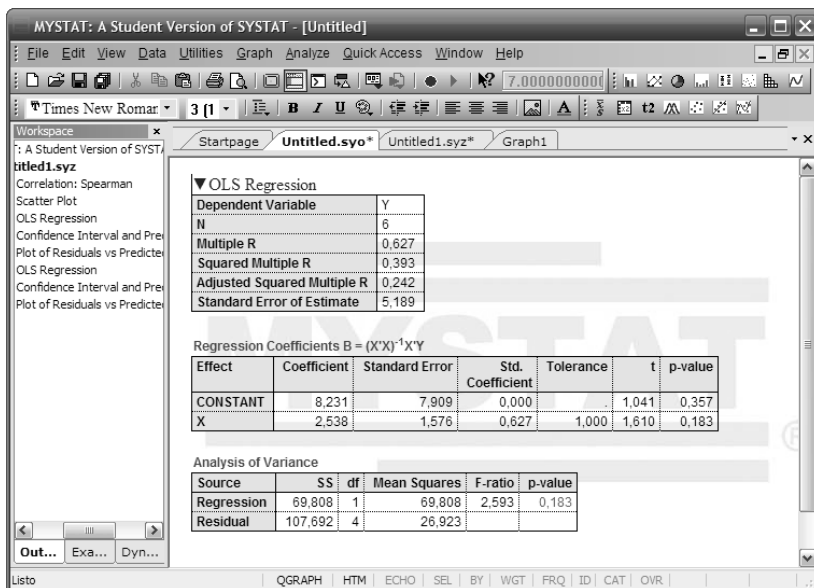


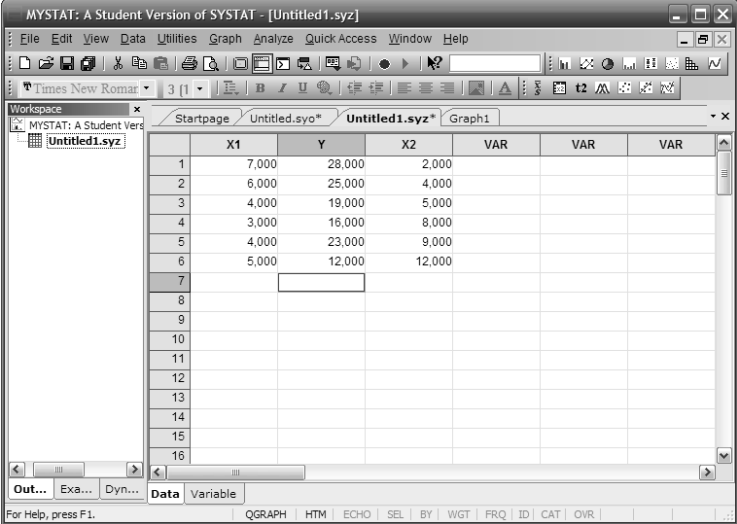
Figura 8.24. Análisis de regresión simple.

Tal y como vemos reflejado en la pantalla, el valor $p=.183$ nos indica que no existe relación entre los valores de x e y , es decir, los cambios de la variable dependiente no están producidos por los cambios en la variable independiente.

4.4. Análisis de regresión múltiple

De la misma manera que con la regresión simple, podríamos realizar un análisis de regresión múltiple. Los pasos son:

- 1°. Abra el fichero desde File/Open/Data, el cual dispondrá ahora dos de variables independientes (x_1 y x_2), tal y como se muestra en la figura 8.25.



	X1	Y	X2	VAR	VAR	VAR
1	7,000	28,000	2,000			
2	6,000	25,000	4,000			
3	4,000	19,000	5,000			
4	3,000	16,000	8,000			
5	4,000	23,000	9,000			
6	5,000	12,000	12,000			
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						

Figura 8.25. Hola de datos para la regresión múltiple.

- 2°. Seleccione la opción Regression del menú Analyze y marque Least Squares en el desplegable. En ella seleccionamos las dos variables independientes y la que será variable dependiente (Figura 8.26).

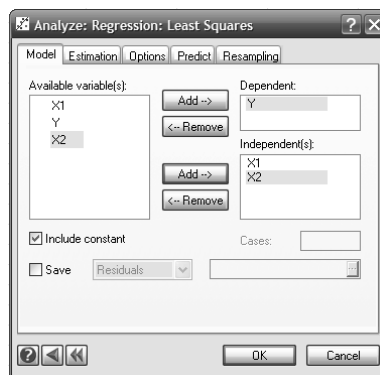


Figura 8.26 Regresión múltiple.

3º. Pulse el botón OK y, en la ventana Untitled.syo, verá los resultados (figura 8.27).

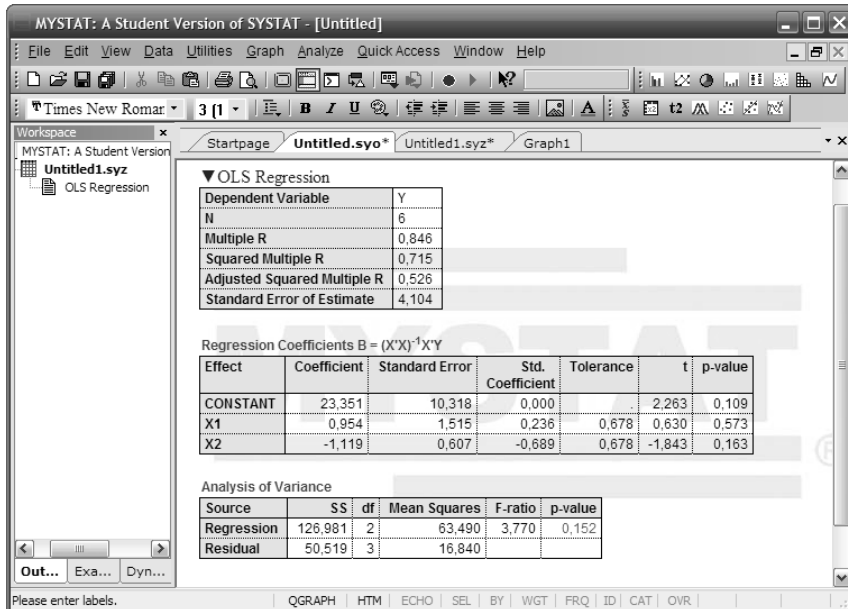


Figura 8.27. Análisis de regresión múltiple.

Tal y como vemos reflejado en la pantalla, el valor $p=0.152$ nos indica que ninguna de las dos variables independientes son buenas predictoras de la variable dependiente.

4.5. Recta de regresión

Para su realización, la secuenciación es la siguiente:

- 1º. Abra el mismo fichero del ejemplo anterior con la secuencia File/Open/Data.
- 2º. Seleccione la opción Scatterplot del menú Graph y aparecerá la ventana de diálogo que observa en la figura 8.28.
- 3º. Seleccione las variables que trasladará a la cuadrícula marcada con Y-variable(s) (variable dependiente) y a la cuadrícula marcada con X-variable(s) (variable independiente).

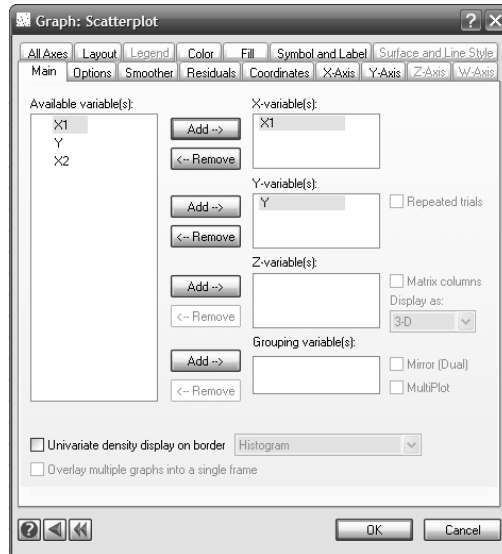


Figura 8.28. Ventana Scatterplot.

- 4°. Pique la pestaña Smoother. Señale la opción Linear.
- 5°. En la ventana que aparece Confidence Intervals on regression line (figura 8.29), escriba .95 o .99 según el nivel de confianza que seleccione. Por defecto sale .95.

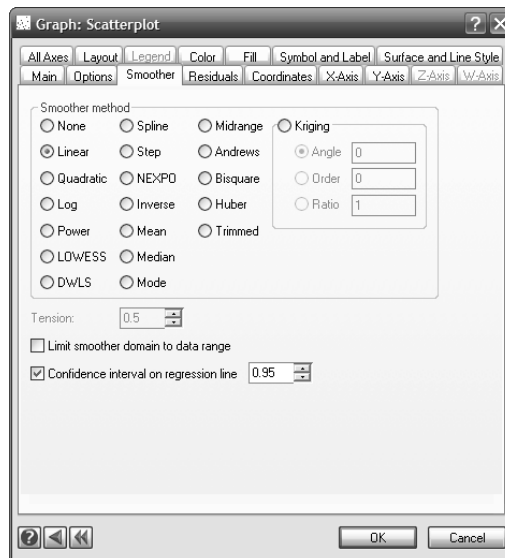


Figura 8.29. Ventana Confidence Intervals.

6°. Pique OK para obtener el resultado y verá una línea a través de la nube de puntos y dos líneas hiperbólicas que representan el intervalo de confianza que ha seleccionado. Esta línea central representa la mejor línea para predecir los valores de la variable dependiente a partir de los de la variable independiente (figura 8.30).

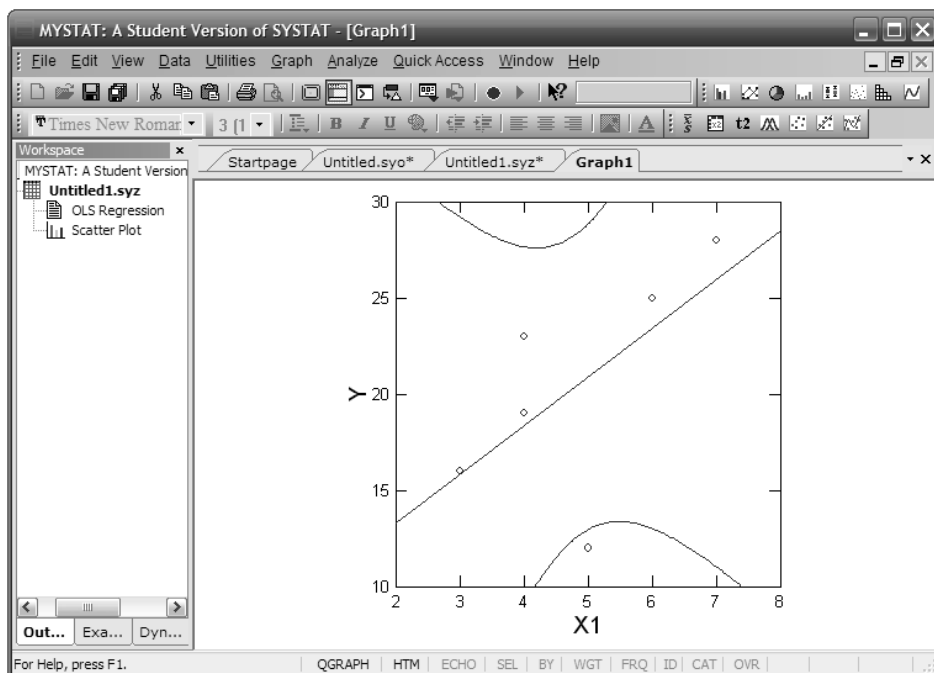


Figura 8.30. Recta de regresión.

5. EJERCICIOS PROPUESTOS

5.1. En un estudio sobre dinámica de grupos en el 2º curso de la titulación de Grado en Pedagogía, el número de intervenciones del grupo se estudió en función del tiempo de duración de cada una de las sesiones. Los datos obtenidos se proporcionan en el cuadro siguiente.

Tiempo	17.4	17.7	18.2	18.0	19.0	19.4	19.5	19.1	19.7	18.5	19.3	19.6	18.4	17.8	18.6
Número de intervenciones	25	41	47	46	58	73	89	79	78	60	85	93	67	37	53

- Estudiar la correlación y la regresión entre las dos variables.
- Representa gráficamente la correlación y regresión.
- Estimar el número de intervenciones en tiempos de 17.5, 19.5 y 15 minutos.

5.2. Se realizó un estudio sobre la relación entre la edad (años) y el tiempo de reacción (segundos) a un estímulo visual en una muestra de 7 estudiantes del Grado en Psicología.

Edad	11	12	13	14	15	16	17
Tiempo	30	90	120	130	150	160	160

- Dar una estimación del tiempo de reacción de un sujeto de 3 años edad.
- ¿Se puede proporcionar una estimación del tiempo de reacción para una edad de 10 años? Justifique su respuesta

5.3. Tras estudiar a dos grupos de 20 estudiantes del Máster de Investigación e Innovación en Educación Infantil y Educación Primaria el rendimiento en Inglés (x) y en Matemáticas (y). Los resultados obtenidos son los siguientes: $\sum X = 906.94$ $\sum X^2 = 66240.326$ $\sum Y = 1808.06$, $\sum Y^2 = 168674.53$, $\sum XY = 92225.015$

- Estudiar la correlación y la regresión entre las dos variables.

5.4. Se dispone de 50 observaciones que asocian el nivel de estudios de los padres (x) y el tiempo destinado a encontrar un trabajo: Los cálculos preliminares dan: $SX = 7701$ $SX^2 = 1187593$ $SY = 20632$ $SY^2 = 8537124$ $SXY = 3182416$

- Estudiar la correlación y la regresión entre las dos variables.

Capítulo 9.

VALIDEZ Y FIABILIDAD DE LOS INSTRUMENTOS DE MEDIDA

Fuensanta Hernández Pina
José D. Cuesta Sáez de Tejada

INTRODUCCIÓN

El concepto de validez, desde la perspectiva clásica, es el grado en que un instrumento indica el grado de exactitud con el que mide el constructo teórico que pretende medir y si se puede utilizar con el objetivo previsto. Es decir, una prueba es válida si “mide lo que dice medir”. Es el grado en que una medición concuerda con un referente (énfasis en resultado, no en instrumento). Se puede decir, que la validez tiene tres grandes componentes:

- validez de contenido
- validez de criterio o criterial
- validez de constructo

Las tres se refieren a aspectos diferentes y la utilización de uno u otro concepto de validez dependen del tipo de prueba. En este capítulo nos vamos a centrar en la la estimación de la validez de constructo mediante el análisis factorial.

El Análisis Factorial es una técnica de reducción de datos que sirve para encontrar grupos homogéneos de variables a partir de un conjunto numeroso de variables. Esos grupos homogéneos se forman con las variables que correlacionan mucho entre sí y procurando, inicialmente, que unos grupos sean independientes de otros. Tiene como finalidad última el obtener un conjunto de variables nuevas a partir de las variables originales. Lo interesante de esta técnica de análisis es que el número de variables nuevo es mucho menor que el número inicial de variables.

El Análisis Factorial distingue dos tipos de variabilidad en la matriz de correlaciones: una variabilidad que es común a todas las variables (conocida como *comunalidad*) y otra variabilidad que es propia de cada variable (conocida como *unicidad*). La finalidad es encontrar un número de variables reducido que permita explicar la variabilidad común en la matriz de correlaciones.

El Análisis Factorial no es propiamente un método de reducción de información, ya que supone que las nuevas variables realmente existen aunque no sean directamente observables. Las puntuaciones de los sujetos en las variables no directamente observables (las que se estiman mediante el análisis) determinan la puntuación en las variables que sí son observables (de las que parte el análisis).

El análisis se realiza en tres pasos principales: (a) determinar si la matriz de correlaciones es adecuada para el análisis; (b) determinar el número óptimo de variables nuevas (factores); y (c) determinar la relación entre cada variable observada y las variables nuevas. A diferencia de lo que ocurre en otras técnicas como el análisis de varianza o el de regresión, en el análisis factorial todas las variables del análisis cumplen el mismo papel: todas ellas son independientes en el sentido de que no existe a priori una dependencia conceptual de unas variables sobre otras.

Para estimar la fiabilidad de un instrumento nos centraremos en el método de la consistencia interna basado en el alfa de Cronbach. Este método permite estimar la fiabilidad de un instrumento de medida a través de un conjunto de ítems que se espera que midan el mismo constructo o dimensión teórica. La medida de la fiabilidad mediante el alfa de Cronbach asume que los ítems (medidos en escala tipo Likert) miden un mismo constructo y que están altamente correlacionados. Cuanto más cerca se encuentre el valor del alfa a 1 mayor es la consistencia interna de los ítems analizados. La fiabilidad de la escala debe obtenerse siempre con los datos de cada muestra para garantizar la medida fiable del constructo en la muestra concreta de investigación.

Como criterio general, presentamos el siguiente esquema Figura 9.1. donde se sugieren las recomendaciones para evaluar los coeficientes de alfa de Cronbach. Mide la fiabilidad entendida como consistencia interna. Este coeficiente analiza la homogeneidad de los ítems que componen una escala. Su valor oscilará entre 0 y 1,

FIABILIDAD				
Muy baja	Baja	Regular	Aceptable	Elevada
0				1
0%				100%
de fiabilidad				de fiabilidad
en la medición				en la medición
(la medición está				(no hay error).
contaminada de				
error).				

Figura 9.1. Esquema para la interpretación de los diferentes valores del coeficiente α de Cronbach.

y en general se suele aceptar que su valor debe ser superior a .7 y que a medida que aumenta mejora la consistencia interna de la escala. No obstante, valores elevados (superiores a .9) pueden ser indicativos de la existencia de ítems redundantes.

1. OBJETIVOS

Los objetivos que pretendemos con el presente capítulo son:

- Conocer las especificaciones mínimas para obtener una solución inicial y cuáles son las opciones disponibles para personalizar la ejecución del procedimiento del Análisis Factorial.
- Conocer el cálculo de la consistencia interna de una prueba mediante el modelo Alfa de Cronbach.
- Interpretar los resultados obtenidos tras el cálculo de la validez y fiabilidad de una prueba con Mystat 12.

2. CONTENIDOS

2.1. Análisis factorial

El módulo para el cálculo del Análisis Factorial con Mystat 12 es el módulo **Factor Analysis**. Para llevar a cabo un Análisis Factorial debemos proceder del modo que pasamos a explicar.

Los pasos a seguir para realizar los estadísticos descriptivos son los siguientes:

- 1º. Desde la barra de herramientas de Mystat 12, cargue el fichero de trabajo picando en File/Open/Data.
Una vez comprobado que éste es el fichero con el que desea trabajar y que su contenido es correcto, puede proceder a cerrarlo, modificarlo, minimizarlo, etc., para continuar con la realización de los estadísticos.
- 2º. Pique sobre la opción **Analyze**, desde la barra de herramientas, y luego la opción Factor Analysis. (Figura 9.2)
- 3º. Seleccione la variable o variables a las que desea realizar el cálculo factorial. Esta acción se realiza picando una a una las variables, y a continuación pulsando en el botón **Add**. Cada variable se irá incorporando al rectángulo de la derecha. En caso de equivocación al seleccionar alguna variable, proceda del modo inverso, márquela en el rectángulo de la derecha y posteriormente pulse en el botón **Remove**.

A continuación seleccione (picando sobre su nombre) los estadísticos que desea calcular, por defecto aparecen unos marcados, pero usted, picando sobre ellos, puede marcar o desmarcar. Lo más rápido suele ser marcar la opción **All options**, ya que de este modo realizará todos los estadísticos posibles y posteriormente se interpretarán solamente aquellos que cada uno/a necesite.

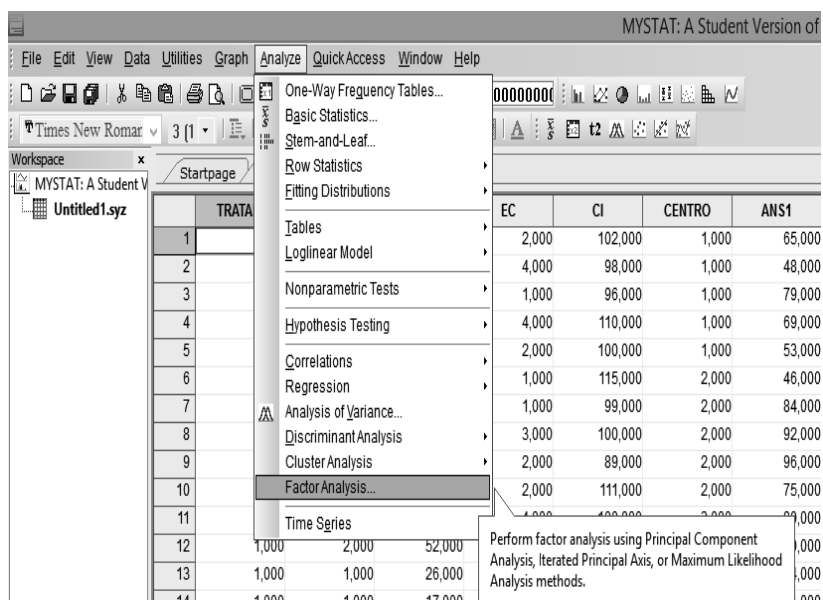


Figura 9.2. Opción *Analyze/Factor Analysis*.

Una vez seleccionadas las variables y los estadísticos a calcular (Figura 9.3), pulse el botón **Ok** para obtener los resultados.

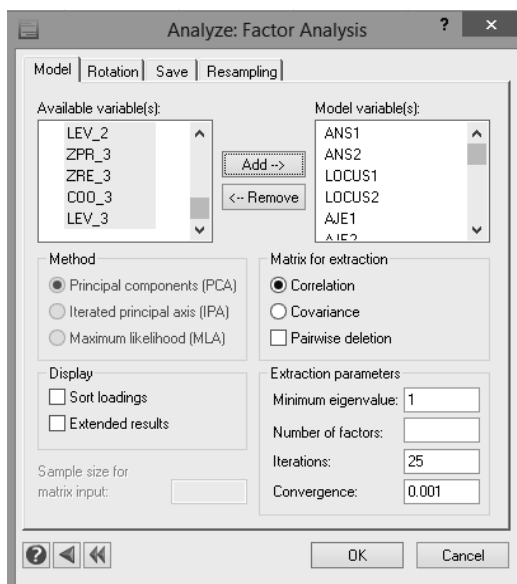


Figura 9.3. Ventana para el *Análisis Factorial* con variables y estadísticos.

4º. Grabar, imprimir o cancelar los resultados obtenidos. Para grabar los resultados debe proceder desde **File/Save As**. Seleccione la unidad, ruta y nombre de archivo, sin olvidar añadir la extensión **.syo**, ya que como hemos indicado en otros capítulos, los datos almacenados es texto que puede ser editado con cualquier procesador de textos (Figura 9.4).

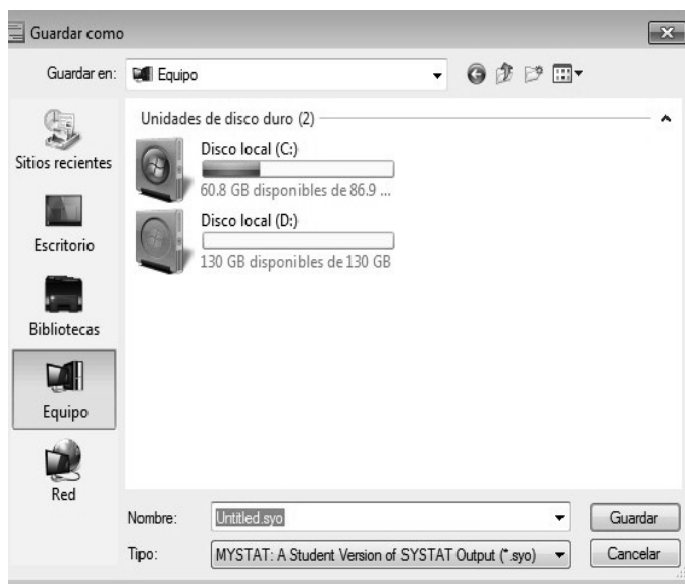


Figura 9.4. Ventana Save As.

2.2. Análisis de fiabilidad: alfa de cronbach

El coeficiente Alfa de Cronbach es un modelo de consistencia Interna, basado en el promedio de las correlaciones entre los ítems. Entre las ventajas de esta medida se encuentra la posibilidad de evaluar cuánto mejoraría (o empeoraría) la fiabilidad de la prueba si se excluyera un determinado ítem.

Para ello, en Mystal 12, el módulo que se emplea para su cálculo es **Correlations**.

1º. Desde la barra de herramientas de Mystal 12, cargue el fichero de trabajo picando en File/Open/Data.

Una vez comprobado que éste es el fichero con el que desea trabajar y que su contenido es correcto, puede proceder a cerrarlo, modificarlo, minimizarlo, etc., para continuar con la realización de los estadísticos.

2º. Pique sobre la opción Analyze, desde la barra de herramientas, y luego la opción Factor Analysis. (Figura 9.5)

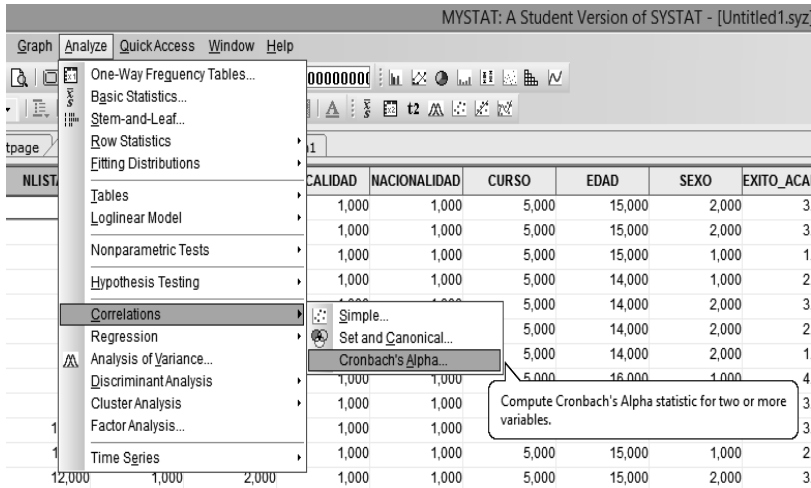


Figura 9.5. Módulo Correlations-Alfa de Cronbach.

3°. Seleccione la variable o variables a las que desea realizar el cálculo factorial. Esta acción se realiza picando una a una las variables, y a continuación pulsando en el botón **Add**. Cada variable se irá incorporando al rectángulo de la derecha. En caso de equivocación al seleccionar alguna variable, proceda del modo inverso, márquela en el rectángulo de la derecha y posteriormente pulse en el botón **Remove**.

A continuación seleccione (picando sobre su nombre) los estadísticos que desea calcular, por defecto aparecen unos marcados, pero usted, picando sobre ellos, puede marcar o desmarcar. Lo más rápido suele ser marcar la opción **All options**, ya que de este modo realizará todos los estadísticos posibles y posteriormente se interpretarán solamente aquellos que cada uno/a necesite.

Una vez seleccionadas las variables y los estadísticos a calcular (Figura 9.6), pulse el botón **Ok** para obtener los resultados.

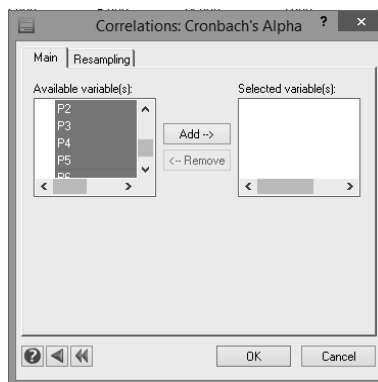


Figura 9.6. Ventana para el Análisis de Fiabilidad Alfa de Cronbach.

4º. Grabar, imprimir o cancelar los resultados obtenidos. Para grabar los resultados debe proceder desde **File/Save As**. Seleccione la unidad, ruta y nombre de archivo, sin olvidar añadir la extensión .syo, ya que como hemos indicado en otros capítulos, los datos almacenados es texto que puede ser editado con cualquier procesador de textos (Figura 9.7).

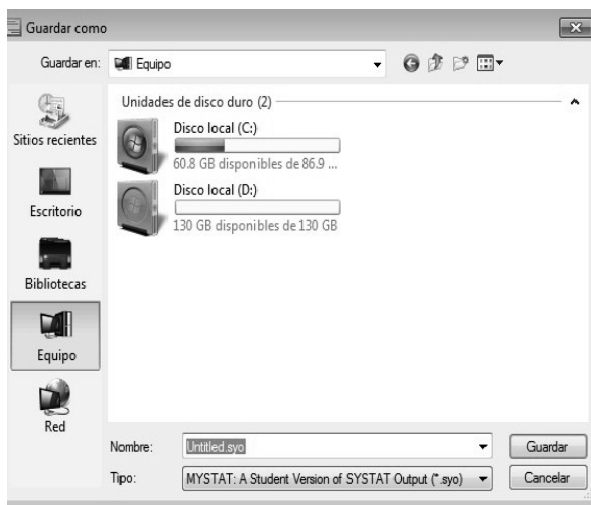


Figura 9.7. Ventana Save As.

3. APLICACIÓN PRÁCTICA CON ORDENADOR DEL ANÁLISIS FACTORIAL Y FIABILIDAD

En nuestro ejemplo hemos tomado una base de datos denominada PCA.syz. En ella se incluye los datos de un estudio sobre los alumnos de un centro educativo. Los datos se refieren a un cuestionario sobre la evaluación de los recursos disponibles (vivienda, dinero, apoyo social, etc.) para los participantes en el estudio. El cuestionario consta de unos 15 (incluyendo datos personales de los estudiantes) ítems, a los que 231 participantes respondieron utilizando una escala tipo Likert.

1º. Una vez seleccionado el fichero con el que desea trabajar, pase a la hoja de datos y con ésta en pantalla seleccione la opción **Analyze/Factor Analysis** (Figura 9.8).

2º. Una vez seleccionada la opción **Factor Analysis** nos aparecerá un cuadro de diálogo como se muestra en la Figura 9.9. Seleccione el conjunto de variables/ factores que desea analizar y trasládelas a la ventana Model variable (s). La lista de variables del archivo de datos contiene un listado de todas las variables del archivo.

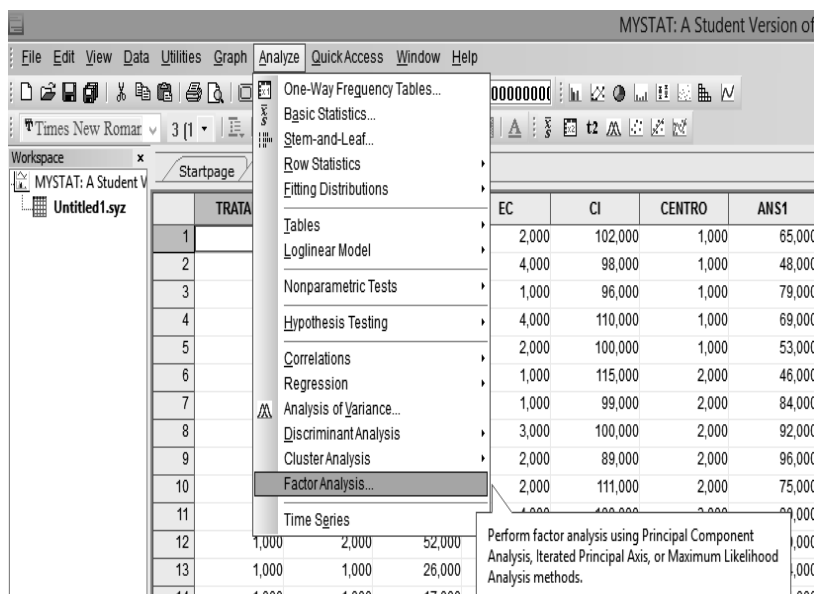


Figura 9.8. Opción *Analyze/Factor Analysis*.

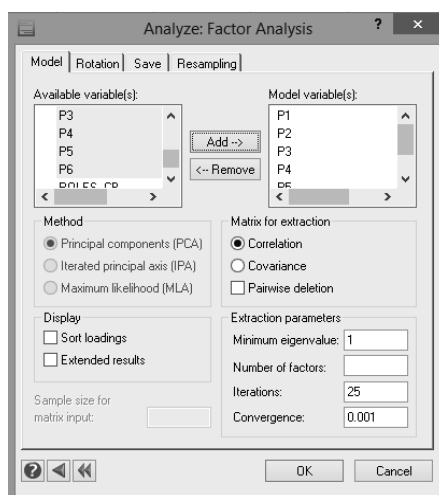


Figura 9.9. Ventana para el Análisis Factorial con variables y estadísticos.

3°. Para facilitar la interpretación de los factores se puede llevar a cabo **una rotación de los factores**. Hay diversos métodos de rotación. Algunos mantienen la independencia de los factores rotados (VARIMAX), y otros permiten la existencia de factores correlacionados (OBLIMIN). La rotación VARIMAX es la que más simplifica la interpretación de los factores.

A continuación seleccione la pestaña **Rotation** y marcamos la opción **Varimax**, que es un método de rotación ortogonal que minimiza el número de variables que tienen saturaciones altas en cada factor. Simplifica la interpretación de los factores optimizando la solución por columna (Figura 9.10).

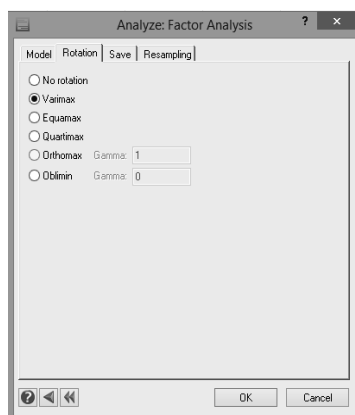


Figura 9.10. Método de Rotación Varimax.

4°. A continuación pulse OK, y en la pestaña Untitled. Syo obtiene los resultados. La primera tabla que nos interesa comentar es la que se denomina **Latent Roots/Eigen Values** (Items Latentes/Valores Propios) y **Variance Explained by Components** (Figura 9.11 y Figura 9.12 respectivamente).

Latent Roots (Eigenvalues)

1	2	3	4	5	6
2,998	1,185	0,995	0,319	0,278	0,226

Figura 9.11. Items Latentes/Valores Propios.

Variance Explained by Components

1	2
2,998	1,185

Figura 9.12. Varianza según los componentes.

Si se fija en los resultados de nuestro de ejemplo, de las seis variables sólo dos obtuvieron valores superiores a 1. La variables 1 y 2. Lo primero que advertimos es el número de factores. En este caso tenemos dos factores (o dos variables hipotéticas) que explican la diversidad de respuestas a la escala original. Podemos imaginar que estos dos factores indican que los seis ítems podrían reducirse a dos preguntas básicas que resumen todo el cuestionario. Esta interpretación estará más clara en los factores rotados.

6°. En cuanto a la tabla denominada **Component Loadings o Análisis de los Componentes**, (Figura 9.13), se observa las correlaciones de las variables o componentes principales sin rotar. La segunda información de interés al construir una escala o test es verificar las correlaciones de los ítems con este primer factor. Deben ser suficientemente altas (.40 o más) y mayores que las que tienen con otros factores. Estos pesos indican lo que el ítem tiene en común con todo lo que es común a todos los ítems. Se interpreta de manera muy semejante a los valores de la Figura 6. Esta correlación ítem-primer factor aporta una información semejante a la correlación ítem-total aunque las magnitudes suelen ser mayores. A veces se utiliza como criterio en la selección de los ítems; por lo menos aporta una información adicional en la misma línea. Los ítems que tienen en este primer factor una correlación inferior a la que tienen en otros factores

Es la correlación de cada una de las variables con cada uno de los factores. Estos coeficientes se corresponden a los coeficientes de correlación lineal de Pearson entre la componente y la variable implicadas.

Así, el factor 1 está muy relacionado con la variable 3 (.843), con la variable 4 (.849) y la variable 5 (.811). En resto de correlaciones, los valores son inferiores. En este ejemplo las correlaciones con el primer factor son muy altas; esto es normal cuando los ítems son pocos y sobre todo cuando se utiliza el análisis de Componentes Principales, en el que esta correlación está sobrestimada.

Component Loadings

	1	2
Variable 4	0,849	0,350
Variable 3	0,843	0,252
Variable 5	0,811	0,411
Variable 2	0,675	-0,623
Variable 1	0,673	-0,634
Variable 6	0,017	0,200

Figura 9.13. Análisis de los componentes.

7°. En la tabla de abajo vemos la variación porcentual explicada por cada uno de los factores, **Percent of Total Variance Explained** (Figura 9.14).

Percent of Total Variance Explained

1	2
49,974	19,742

Figura 9.14. Porcentaje de Varianza Explicada.

El primer factor sin rotar explica la máxima cantidad de varianza común a todos los ítems, es decir, debida a lo que los ítems tienen en común y que corresponde supuestamente a lo que pretendemos medir. Por eso es importante que la proporción de varianza que explica este primer factor sin rotar sea bastante mayor que la que explican cada uno de los demás factores. Sin embargo el que todos los ítems tengan

en este factor su mayor peso no es prueba de que haya un único factor o de que el instrumento sea realmente unidimensional. El segundo factor explica la mayor proporción de la varianza que va quedando y así sucesivamente. En este caso el primer factor explica una proporción de varianza (49.97%) mucho mayor que el segundo factor (19.74%). A partir del primer factor suele haber una caída en estos valores. En cualquier caso, entre las dos componentes extraídas se acumula el 69% de la variabilidad de las variables originales. En otras palabras, de la totalidad o 100% del fenómeno estudiado, el componente 1 explica dicho fenómeno en un 49.974% y el componente 2 lo explica en un 19.742%, entonces, si unimos ambos componentes, se obtiene que el porcentaje que explican los dos componentes del fenómeno investigado es de 69%, o que la varianza total del fenómeno es explicada en un 69% por los factores 1 y 2.

8°. La siguiente tabla son los resultados de la matriz rotada (Figura 9.15). El procedimiento probablemente más seguido es el de Gutman-Kaiser; programado por defecto en los programas estadísticos: se rotan solamente los factores que en el primer análisis (antes de las rotaciones) tienen un eigenvalue (o raíz característica, o varianza) mayor de 1. En este caso se ha utilizado el método de rotación ortogonal Varimax, que tiende a forzar la diferenciación, los contrastes, y deja más clara la estructura subyacente a toda la escala. La primera tabla que debemos comentar es la de la proporción de varianza total explicada por los factores rotados, que es la misma que ya conocemos por el análisis previo. Lo que sucede es que la misma proporción total de varianza aparece ahora repartida de manera diferente, y esto nos permite valorar qué factores son más importantes. En este caso la diferencia entre los dos factores es pequeña.

Rotated Loading Matrix (VARIMAX, Gamma = 1.000000)

	1	2
item4	0,911	0,119
item5	0,908	0,047
item3	0,857	0,200
item1	0,268	0,884
item2	0,276	0,876
item6	0,115	-0,165

Figura 9.15. Varianza de los factores rotados (Varimax).

Los ítems aparecen ordenados de manera descendiente. Los valores de la tabla son los coeficientes de los factores después de la rotación. Observamos que los ítems con valores más altos para el factor 1 son el ítem 4, ítem 5 y el ítem 3, y en el factor 2, el ítem 1 y el ítem 2. Las variables que más se parecen son las que acaban asociándose a cada uno de los factores.

En la figura 9.16 vemos que la proporción de varianza total explicada por los factores rotados es la misma que ya conocemos por el análisis previo. Lo que sucede es que la misma proporción total de varianza aparece ahora repartida de manera diferente, y esto nos permite valorar qué factores son más importantes. En este caso la diferencia entre los dos factores es un poco más pequeña que la tabla anterior sin rotar.

Frecuentemente nos encontraremos con una proporción total de varianza explicada por los factores que oscila entre el 50 y 60%. Con todos los *factores posibles* llegaríamos al 100% de la varianza, pero muchos de estos factores serían de hecho irrelevantes para la interpretación.

Percent of Total Variance Explained	
1	2
42,499	27,217

Figura 9.16. Proporción de varianza explicada por los diversos factores rotados.

9°. Para reducir el número de factores de manera no arbitraria se emplea una sencilla gráfica denominada **Scree Plot** (Figura 9.17). El Mystat 12 nos ofrece este gráfico en el que aparecen el número de factores o componentes (en el eje X; el número de componentes coincide con el número de ítems) y los *eigenvalues* [eje Y] o varianza explicada por cada factor. El *punto de corte* para establecer el número de factores que se van a rotar se sitúa en el punto de inflexión de la línea descendente que va uniendo los diversos *eigenvalues*.

En este ejemplo los factores (Componentes Principales) con un *eigenvalue* superior a 1 son 2 (son los que se seleccionarían con el procedimiento de Gutman-Kaiser programado por defecto en el Mystat 12). Vemos que a partir del factor 2 hay un punto de inflexión. Dos factores pueden explicar menos varianza que tres, cuatro o cinco factores, pero la estructura será más clara e interpretable, algo habitual en la interpretación de muchas escalas.

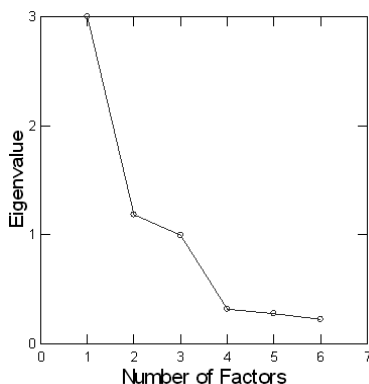


Figura 9.17. Scree Plot.

10°. En el siguiente gráfico (Figura 9.18), el gráfico de Factores o **Factor Loading Plots**, observe que aparecen los dos factores en cada uno de los ejes. Los puntos representan las variables. Este gráfico es útil para ver agrupamientos, para ver clusters. En nuestro ejemplo, vemos que las variables de la derecha, son las variables que se cargan en el factor 1, y las variables de arriba son las que se cargan en el factor 2.

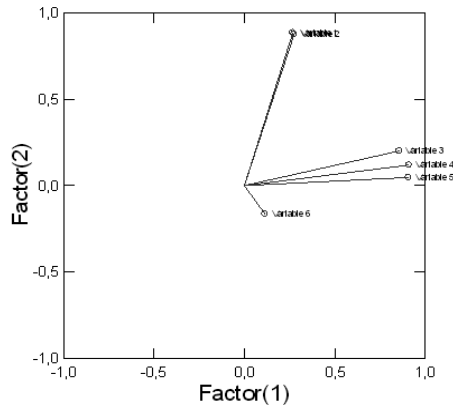


Figura 9.18. Factor Loadings Plot.

2.2. Análisis de fiabilidad: Alfa de Cronbach

El coeficiente Alfa de Cronbach es un modelo de consistencia Interna, basado en el promedio de las correlaciones entre los ítems. Entre las ventajas de esta medida se encuentra la posibilidad de evaluar cuánto mejoraría (o empeoraría) la fiabilidad de la prueba si se excluyera un determinado ítem.

Para ello, en Mystat 12, el módulo que se emplea para su cálculo es **Correlations** (Figura 9.19).

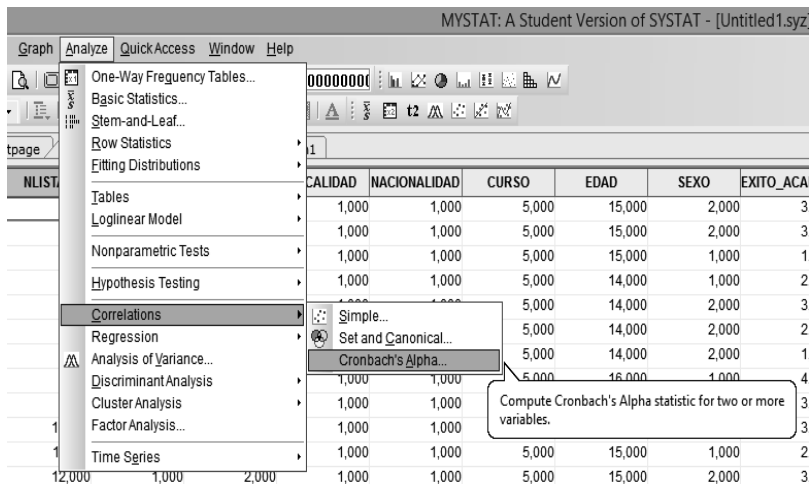


Figura 9.19. Módulo Correlations-Alfa de Cronbach.

1º. En primer lugar seleccione todas las variables y pinche la opción “Add” para pasar las variables que vamos a analizar a la casilla “Select variable (s)” (Figura 9.20).

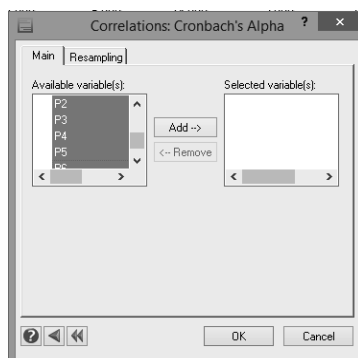


Figura 9.20. Cuadro de diálogo Alfa de Cronbach.

2º. Los resultados aparecen en la Figura 9.21. El valor obtenido en nuestro ejemplo para las seis variables de la prueba analizada es de .664 laartimos de un fichero con algunas variables, la primera que hace referencia al número de cuestionario, seguida de la edad, el género, la pregunta 1, pregunta 2, pregunta 3...

▼ Cronbach's Alpha

Number of Variables : 6
Number of Cases : 749

Variables

P1	P2	P3	P4	P5	P6
----	----	----	----	----	----

Cronbach's Alpha : 0,664

Figura 9.21. Valor Alfa de Cronbach.

Como criterio general, George y Mallery (2003, p. 231) sugieren las recomendaciones siguientes para evaluar los coeficientes de alfa de Cronbach:

- Coeficiente alfa $>.9$ es excelente
- Coeficiente alfa $>.8$ es bueno
- Coeficiente alfa $>.7$ es aceptable
- Coeficiente alfa $>.6$ es cuestionable
- Coeficiente alfa $>.5$ es pobre
- Coeficiente alfa $<.5$ es inaceptable

4. EJERCICIOS PROPUESTOS

4.1. Se desea construir una batería de pruebas que permita medir la orientación vocacional, que hace referencia a un conjunto de practics destinadas al esclarecimiento de la problemática vocacional. Se decide incluir una prueba conocimientos

adquiridos (CA), otra de aptitudes (AP), una prueba de inteligencia (I), y otra de actitudes (AC). Se aplican las pruebas a veinte sujetos. Se va a efectuar un Análisis Factorial para reducir el conjunto de variables medidas a un conjunto menor de factores interpretando esos factores. Realiza un análisis factorial siguiendo el método de los componentes principales en Mystal 12 teniendo en cuenta los datos adjuntos.

Sujetos	CA	AP	I	AC
1	17	15	16	16
2	16	15	16	15
3	18	16	17	17
4	17	16	15	15
5	18	14	15	17
6	18	15	15	17
7	19	15	15	18
8	17	14	15	16
9	16	15	15	16
10	17	15	16	15
11	16	16	15	15
12	16	15	16	15
13	18	16	17	17
14	17	16	15	15
15	18	14	15	17
16	18	15	15	17
17	19	15	15	18
18	17	14	15	16
19	16	15	15	16
20	17	15	16	15

4.2. Una empresa de productos precocinados está interesada en incrementar sus ventas para lo cual, diseña una campaña de publicidad basándose en los resultados de una encuesta. La encuesta constaba una serie de afirmaciones a las que los sujetos contestaban desde 1 (muy en desacuerdo) hasta 5 (muy de acuerdo). Las cuestiones planteadas en la encuesta fueron:

- Item 1: Compro este tipo de comida cuando invito a los amigos a casa
- Item 2: Me gusta comer estos alimentos cuando estoy solo/a en casa
- Item 3: Son comidas rápidas y fáciles de preparar
- Item 4: Me gusta comprar siempre el mismo tipo de alimento
- Item 5: A los supermercados donde voy compro siempre la misma marca
- Item 6: Cambio del tipo de alimentos y de marca para valorar diferentes calidades del producto
- Item 7: Los alimentos precocinados sólo me gusta consumirlos cuando no tengo tiempo de cocinar
- Item 8: Uno de mis productos favoritos es el pescado
- Item 9: Principalmente mezclo productos precocinados y naturales

	F_1	F_2	F_3
Item 1	.71	.01	-.02
Item 2	.70	.06	.06
Item 3	.50	.3	.06
Item 4	.01	.8	.01
Item 5	.02	.7	.08
Item 6	-.06	-.55	.05
Item 7	.02	.01	.7
Item 8	.002	.1	.75
Item 9	.25	.2	.75

Calcula, a partir de la matriz factorial, los siguientes valores:

- Comunalidad para cada variable.
- Porcentaje de varianza explicada por cada factor y porcentaje de varianza explicada por el conjunto de factores.
- Representa gráficamente por los factores F_1 y F_2 por un lado y F_2 y F_3 y la solución factorial obtenida.
- Interpreta y nombre los factores obtenidos. ¿En qué debería incidir la campaña publicitaria en base a los resultados obtenidos con el análisis factorial?

4.3. Tenemos un cuestionario para evaluar el grado de satisfacción de un curso con tres preguntas y se desea saber si los datos que se obtienen a partir de esta herramienta, son fiables. Para evaluar la fiabilidad de este cuestionario, este último se aplicó a 10 jueces. Las preguntas y los resultados se muestra a continuación:

Item 1: ¿El curso ha respondido a sus expectativas?

Item 2: ¿Los docentes dominaban la materia?

Item 3: ¿Se desarrolló de acuerdo a lo programado?

Jueces	ítem 1	ítem 2	ítem 3	Total fila
1	4	2	4	10
2	2	1	3	6
3	3	1	2	6
4	2	2	2	6
5	2	1	4	7
6	1	1	3	5
7	2	4	4	10
8	3	3	4	10
9	4	4	3	11
10	1	1	1	3
Total Columna	24	20	30	74
Media	2.4	2	3	7.4
Desv. Típica	1.0749677	1.24721913	1.05409255	2.67498702

Capítulo 10. ESTADÍSTICA NO PARAMÉTRICA

*José D. Cuesta Sáez de Tejada
Javier J. Maquilón Sánchez*

INTRODUCCIÓN

Las pruebas no paramétricas deben emplearse al darse las siguientes condiciones: 1) pretendemos analizar datos cualitativos, 2) las muestras provienen de poblaciones que no se distribuyen normalmente y/o 3) simplemente deseamos emplear una prueba más sencilla que la prueba *t* o ANOVA.

El programa MYSTAT 12 incluye como estadística no paramétrica lo siguiente: *Tables* (para analizar tablas de contingencia unidimensional y bidimensional), *Prueba de los signos*, *Prueba de Wilcoxon*, *Prueba de Friedman*, etc. A continuación ofrecemos una tabla resumen para explicar el uso de unas pruebas u otras:

	PRUEBAS PARAMÉTRICAS		PRUEBAS NO PARAMÉTRICAS	
	COMPARACIÓN DE DOS MUESTRAS	COMPARACIÓN DE K MUESTRAS	COMPARACIÓN DE DOS MUESTRAS	COMPARACIÓN DE K MUESTRAS
	PRUEBA χ^2			
GRUPOS INDEPENDIENTES	T-STUDENT	ANOVA	Mann-Whitney	Kruskal-Wallis
GRUPOS RELACIONADOS	T-STUDENT	ANOVA	Wilcoxon	Friedman

Tabla 10.1. Técnicas inferenciales paramétricas y no paramétricas.

1. OBJETIVOS

Los objetivos que pretendemos con el presente capítulo son:

- Analizar datos categoriales, en el caso de una, dos o más variables independientes (Prueba Ji-Cuadrado)
- Comparar datos de un mismo grupo en situación de pretest y postest (prueba de los signos).
- Comparar datos para muestras correlacionadas basándonos en la intensidad de la diferencia entre los valores de pretest y postest (prueba de Wilcoxon).

2. CONTENIDOS

2.1. Prueba Ji-Cuadrado

Se emplea esta prueba cuando los datos aparecen distribuidos en frecuencias. Dichas frecuencias, denominadas frecuencias observadas, deben compararse con las frecuencias esperadas, y así podremos determinar si la población de la cual proviene la muestra está dividida en partes iguales. Esta comparación permitirá aceptar o rechazar la hipótesis nula. La prueba ji-cuadrado puede aplicarse cuando existe una sola variable y dos o más variables. La fórmula sería la siguiente:

$$\chi^2_{(F-1)(C-1)} = \sum_{i=1}^F \sum_{j=1}^C \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

χ^2 PARA TABLAS UNIDIMENSIONALES

- 1º. Abra el fichero desde File/Open/Data.
- 2º. Seleccione la opción Tables: Two-Way del menú Analyze.
- 3º. En la ventana de diálogo (Figura 10.1) seleccione la/s variable/es. Con el botón Add las incorpora a la cuadrícula de la derecha.
- 4º. En la parte inferior derecha verá que están marcados unos cuadros con varias opciones. Algunas vienen marcadas con una cruz por defecto. Otras pueden seleccionarse picando sobre ellas con el botón izquierdo del ratón, o anularlas picando sobre las que están marcadas.
- 5º. Pulse OK. Los resultados que verá en la pantalla principal no incluyen el análisis χ^2 para tablas unidimensionales. Esto lo puede hacer a partir de la fórmula conocida:

Para calcular el valor p para comprobar la significación debe proceder del modo siguiente: Data/Transform/Let e irá a la ventana de diálogo que observa en la figura 10.2.

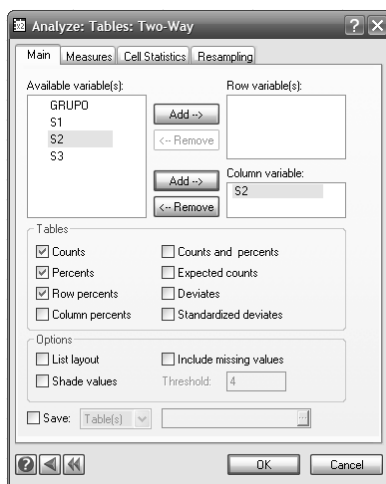


Figura 10.1. Ventana Tables: Two-Ways.

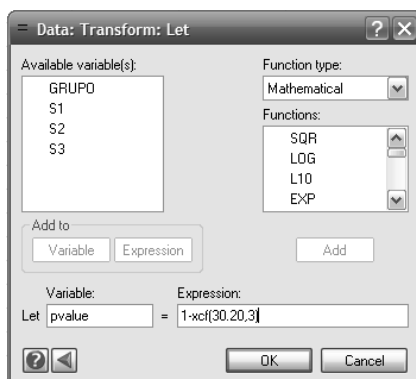


Figura 10.2. Ventana Data:Transform:Let.

En el cuadrado junto a la instrucción *Let*, escriba la variable a la que va a asignar un determinado valor. Este valor viene calculado por la función que se encuentra en el cuadro seguido del signo =. En este espacio escriba la siguiente secuencia:

1-xcf (valor de la χ^2 , grados de libertad)

Fíjese en el ejemplo de la figura 10.3.

Variable	Variable or expression
Let pvalue	= 1-xcf{30.20,3}

Figura 10.3. Ventana de función.

A continuación pulse OK. Observará junto a la columna de la variable la proporción con la que cada frecuencia observada es significativa.

χ^2 PARA TABLAS BIDIMENSIONALES

Para el análisis de la χ^2 en tablas de dos o más variables independientes debe proceder del modo siguiente:

1°. Abra el fichero desde File/Open/Data.

2°. Seleccione Tables: Two-Way. En la ventana de diálogo (Figura 10.4) elija la/las variable/es dependientes y con el botón Add la/las inserta en la cuadrícula superior derecha. Seguidamente seleccione la variable independiente y llévela a la casilla correspondiente. Recuerde que todas las variables insertadas en la primera casilla serán analizadas en función de esta última variable. La posición de las variables en uno y otro espacio va a depender de los objetivos de la investigación.

3°. Seguidamente elija las opciones que le interesen: porcentaje, porcentaje por fila, etc.

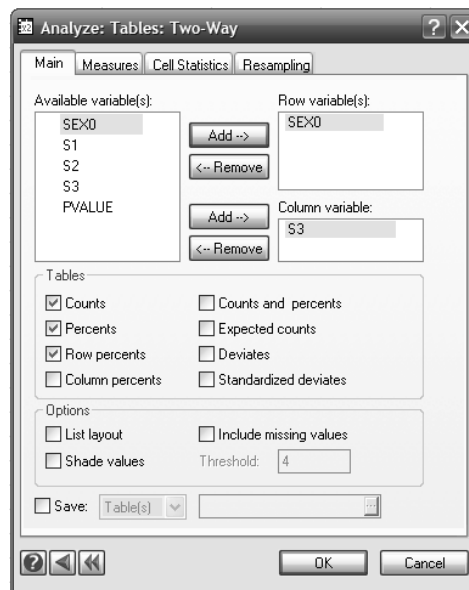


Figura 10.4. Ventana Tables.

4°. Pulse OK y proceda a grabar los resultados con la secuencia File/Save As (Figura 10.5) introduciendo en la ventana de diálogo la unidad, la ruta y el nombre del fichero.

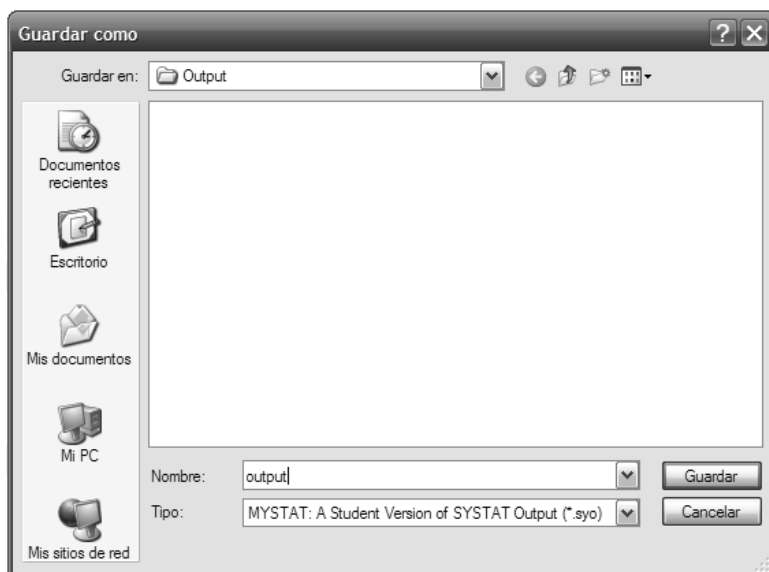


Figura 10.5. Ventana Save as en File.

En la salida por pantalla verá que después de la tabla de contingencia se relacionan una serie de datos que le van a permitir interpretar los resultados alcanzados. En el primer apartado, aparece la prueba de significación estadística χ^2 de Pearson, los grados de libertad y la probabilidad. En un segundo apartado, aparecen los coeficientes que indican la fuerza de la asociación entre las variables. Deberá elegir el más apropiado a la tabla que haya analizado:

- a) *Phi* es el apropiado para tablas 2x2. Este coeficiente es idéntico a la *r* de Pearson. Su valor oscila entre -1 y +1.
- b) *V de Cramer* es una modificación de *Phi*. Su valor puede oscilar entre 0 y +1. Cuanto mayor es este valor, la asociación será más intensa.
- c) *Contingencia*. Este coeficiente oscila entre 0 y, puede sobrepasar +1. Este límite dependerá del tamaño de la tabla así que se recomienda comparar coeficientes de tablas del mismo tamaño. Cuanto mayor sea este coeficiente mayor será la asociación entre las variables.
- d) Otros coeficientes como el de *Goodman-Kruskal Gamma*, *Kendal Tau-b*, *Stuart Tau-c*, *Gamma Rho*, *Sommers D*... son apropiados para interpretar variables ordinales.

2.2. Prueba de los signos

Esta prueba es útil para comparar datos de un mismo grupo en situación de pretest y postest. La base del análisis de esta prueba está en el signo, al comparar

las puntuaciones entre el pretest y el posttest, obteniéndose un nuevo valor denominado D. Se eliminan aquellos valores que nos den el valor cero como resultado de comparar pretest y posttest.

$$P(X = x) = nCx * p^x q^{n-x}$$

$$nCx = \frac{n!}{x!(n-x)!}$$

Los pasos a seguir en Mynstat son:

- 1º. Abra el fichero desde File/Open/Data.
- 2º. Seleccione la opción Nonparametrics Tests: Sign del menú Analyze.
- 3º. En la ventana de diálogo (Figura 10.6) seleccione la/s variable/es y pulse el botón Add para trasladarlas a la cuadrícula de la derecha.

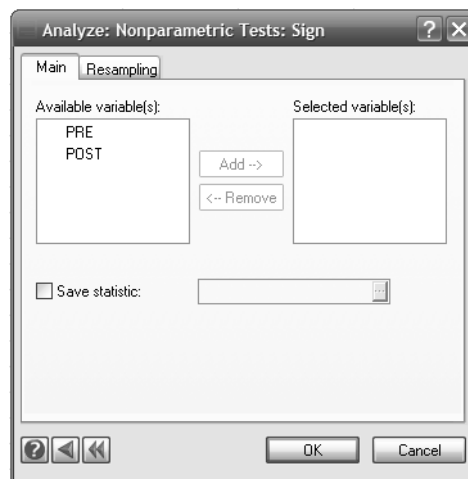


Figura 10.6. Ventana Tests Sign.

4º. Pulse OK y proceda a grabar los resultados con la secuencia File/Save As (Figura 10.7) introduciendo en la ventana de diálogo la unidad, la ruta y el nombre del fichero.

Los resultados aparecen especificados en dos apartados. El primero hace referencia a las frecuencias de las diferencias y el segundo a su probabilidad. El primer apartado expresa cuantas veces la variable fila (POST) tiene valores más grandes que la variable columna (PRE). El segundo apartado, en la intersección fila (POST) y columna (PRE), verá la proporción con la que la diferencia entre el PRE y el POST es significativa.

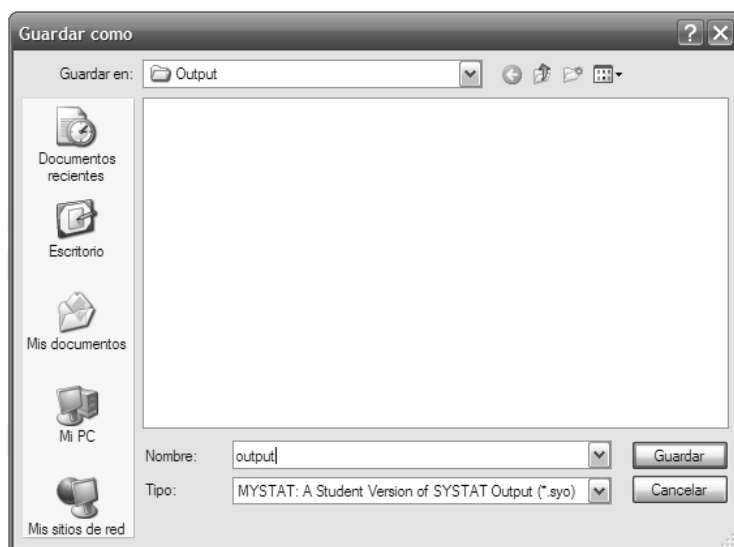


Figura 10.7. Ventana Save as en File.

2.3. Prueba de Wilcoxon

Es similar a la prueba de los signos. Se emplea con muestras dependientes y tratará de determinar la intensidad de las diferencias entre dos pares de puntuaciones. Al igual que la prueba de los signos, debemos calcular el valor D. Posteriormente se eliminan los valores que den cero resultantes de aplicar la formula $D=X_1-X_2$ y se ordenan de menor a mayor hasta obtener una serie de valores ordinales. Se ordenan los valores absolutos de las diferencias y se asignan los rangos. Posteriormente, a los rangos se le pone el signo de la diferencia (positivo o negativo). Finalmente se calcula la T, una positiva y otra negativa, sumando los rangos con signo bien positivo o negativo. Las formulas empleadas para el cálculo, que aplicaremos más adelante, son:

$$T_E = \frac{N(N-1)}{4} \quad \sigma T_m = \sqrt{\frac{(2N+1)T_e}{24}} \quad Z = \frac{T_1 - T_e}{\sqrt{\frac{(2N+1)T_e}{\sigma T_m}}}$$

Siga los pasos siguientes:

- 1°. Abra el fichero con la secuencia File/Open/ Data.
- 2°. Seleccione la opción Nonparametrics Tests: Wilcoxon del menú Analyze.
- 3°. En la ventana de diálogo (figura 10.8) seleccione la/s variable/es y pulse el botón Add para trasladarlas a la cuadrícula de la derecha.

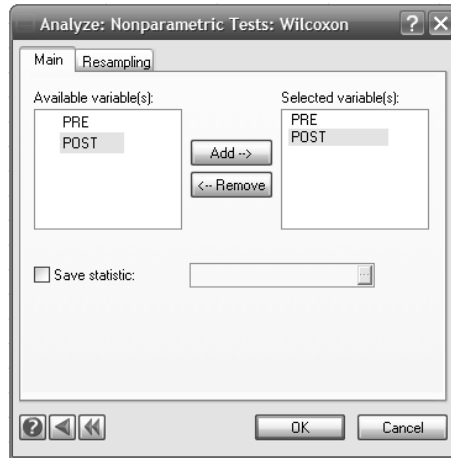


Figura 10.8. Ventana Tests Wilcoxon.

4º. Pulse OK y proceda a grabar los resultados con la secuencia File/Save As (Figura 10.9) introduciendo en la ventana de diálogo la unidad, la ruta y el nombre del fichero.

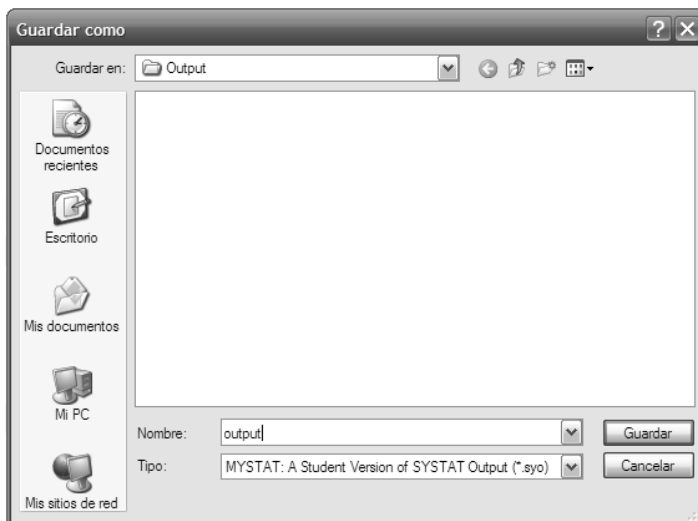


Figura 10.9. Ventana Save as en File.

2.4. Prueba de Friedman

Se emplea para comparar K muestras. La prueba de análisis de varianza de Friedman prueba la hipótesis nula que k variables apareadas proceden de la misma

población. Se clasifica a cada individuo sobre cada una de las variables k (filas de 1 a k). La prueba se basa en estas filas. Las fórmulas que debemos emplear para el cálculo manual son:

Ecuación 1:

$$X_F^2 = \frac{12N}{k(k+1)} \left(\sum_j R_j^2 - \frac{k(k+1)^2}{4} \right)$$

$$R_j = \frac{1}{N} \sum_i r_i^j$$

Ecuación 2:

$$X_{AF}^2 = \frac{(k-1) \sum_{j=1}^k R_j^2 - (kN^2/4)(kN+1)^2}{kN(kN+1)(2kN+1)/6 - (1/k) \sum_{i=1}^N R_i^2}$$

Ecuación 3:

$$S_{ij} = Q_i \left(R_i^j - \frac{k+1}{2} \right)$$

Ecuación 4:

$$W_{ij} = Q_i(R_i^j)$$

Ecuación 5:

$$S_j = \sum_{i=1}^N S_{ij}$$

Los pasos a seguir son los siguientes:

- 1°. Abra el fichero File/Open/ Data.
- 2°. Seleccione la opción Nonparametrics Tests: Friedman del menú Analyze.
- 3°. En la ventana de diálogo (Figura 10.10) seleccione la/s variable/es y pulse el botón Add para trasladarlas a la cuadrícula de la derecha.
- 4°. Pulse OK y proceda a grabar los resultados con la secuencia File/Save As (Figura 10.11) introduciendo en la ventana de diálogo la unidad, la ruta y el nombre del fichero.

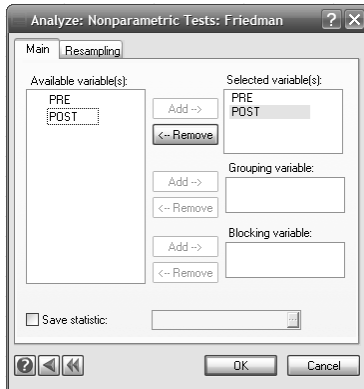


Figura 10.10. Ventana Tests Friedman.

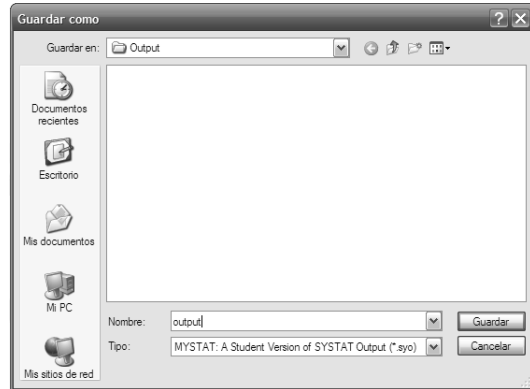


Figura 10.11. Ventana Save as en File.

En los resultados que aparecen en pantalla verá una columna con el nombre de las variables y a su lado la suma de sus rangos (no olvide que esta prueba se basa en datos ordinales). A continuación se muestra el coeficiente de Friedman y el coeficiente de concordancia de Kendall. En la última línea aparece la probabilidad de la significación del coeficiente de Friedman y sus grados de libertad.

3. APLICACIÓN PRÁCTICA SIN ORDENADOR

3.1. Prueba Ji-Cuadrado con una sola variable

Imaginemos que un profesor, antes de realizar los exámenes, pregunta a sus alumnos ($N=48$) las preferencias a la hora de realizar la tipología de la prueba, según sea:

- Una prueba oral
- Una prueba escrita de desarrollo
- Una prueba objetiva con tres opciones de respuestas

Tras recoger la información, obtiene los siguientes resultados (frecuencias observadas):

Opciones	Frecuencias observadas (f_o)	Frecuencias esperadas (f_e)	$(f_o - f_e)^2$	$\frac{[f_o - f_e]^2}{f_e}$
Prueba oral	18	16	4	.25
Prueba escrita	19	16	9	.56
Prueba objetiva	11	16	25	.56

Tabla. 10.2

El siguiente paso consiste en obtener las frecuencias esperadas, dividiendo el número total de alumnos entre las opciones (en este caso 3). Posteriormente, se procede a realizar las operaciones $(f_o - f_e)^2$ y dividir cada valor entre cada una de sus frecuencias esperadas y sustituirlo en la fórmula:

$$\chi^2_{(F-1)(C-1)} = \sum_{i=1}^F \sum_{j=1}^C \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Una vez calculado el valor χ^2 hemos de fijarnos en la tabla de los valores críticos para Ji-cuadrado (Tabla 10.3). Para interpretar χ^2 , al igual que hacíamos con la prueba T, debemos tener en cuenta los g/l pero, a diferencia de la prueba ji-cuadrado, éstos están basados en categorías y no en número de observaciones.

v/p	0,001	0,0025	0,005	0,01	0,025	0,05	0,1	0,15	0,2	0,25	0,3	0,35	0,4	0,45	0,5
1	10,8274	9,1404	7,8794	6,6349	5,0239	3,8415	2,7055	2,0722	1,6424	1,3233	1,0742	0,8735	0,7083	0,5707	0,4549
2	13,8150	11,9827	10,5965	9,2104	7,3778	5,9915	4,6052	3,7942	3,2189	2,7726	2,4079	2,0996	1,8326	1,5970	1,3863
3	16,2660	14,3202	12,8381	11,3449	9,3484	7,8147	6,2514	5,3170	4,6416	4,1083	3,6649	3,2831	2,9462	2,6430	2,3660
4	18,4662	16,4238	14,8602	13,2767	11,1433	9,4877	7,7794	6,7449	5,9886	5,3853	4,8784	4,4377	4,0446	3,6871	3,3567
5	20,5147	18,3854	16,7496	15,0863	12,8325	11,0705	9,2363	8,1152	7,2893	6,6257	6,0644	5,5731	5,1319	4,7278	4,3515
6	22,4575	20,2491	18,5475	16,8119	14,4494	12,5916	10,6446	9,4461	8,5581	7,8408	7,2311	6,6948	6,2108	5,7652	5,3481
7	24,3213	22,0402	20,2777	18,4753	16,0128	14,0671	12,0170	10,7479	9,8032	9,0371	8,3834	7,8061	7,2832	6,8000	6,3485
8	26,1239	23,7742	21,9549	20,0902	17,5345	15,5073	13,3616	12,0271	11,0301	10,2189	9,5245	8,9094	8,3505	7,8325	7,3441
9	27,8767	25,4625	23,5893	21,6660	19,0228	16,9190	14,6837	13,2880	12,2421	11,3887	10,6564	10,0060	9,4136	8,8632	8,3428
10	29,5879	27,1119	25,1881	23,2093	20,4832	18,3070	15,9872	14,5339	13,4420	12,5489	11,7807	11,0971	10,4732	9,8922	9,3418
11	31,2635	28,7291	26,7569	24,7250	21,9200	19,6752	17,2750	15,7671	14,6314	13,7007	12,8987	12,1836	11,5298	10,9199	10,3410
12	32,9092	30,3182	28,2997	26,2170	23,3367	21,0261	18,5493	16,9893	15,8120	14,8454	14,0111	13,2661	12,5838	11,9463	11,3403
13	34,5274	31,8830	29,8193	27,6882	24,7356	22,3620	19,8119	18,2020	16,9848	15,9839	15,1187	14,3451	13,6356	12,9717	12,3398
14	36,1239	33,4262	31,3194	29,1412	26,1189	23,6848	21,0641	19,4062	18,1508	17,1169	16,2221	15,4209	14,6853	13,9961	13,3393
15	37,6978	34,9494	32,8015	30,5780	27,4884	24,9958	22,3071	20,6030	19,3107	18,2451	17,3217	16,4940	15,7332	15,0197	14,3389
16	39,2518	36,4555	34,2671	31,9999	28,8453	26,2962	23,5418	21,7931	20,4651	19,3689	18,4179	17,5646	16,7795	16,0425	15,3385
17	40,7911	37,9462	35,7184	33,4087	30,1910	27,5871	24,7690	22,9770	21,6146	20,4887	19,5110	18,6330	17,8244	17,0646	16,3382
18	42,3119	39,4220	37,1564	34,8052	31,5264	28,8693	25,9894	24,1555	22,7595	21,6049	20,6014	19,6993	18,8679	18,0860	17,3379
19	43,8194	40,8847	38,5821	36,1908	32,8523	30,1435	27,2036	25,3289	23,9004	22,7178	21,6891	20,7638	19,9102	19,1069	18,3376
20	45,3142	42,3358	39,9969	37,5663	34,1696	31,4104	28,4120	26,4976	25,0375	23,8277	22,7745	21,8265	20,9514	20,1272	19,3374
21	46,7963	43,7749	41,4009	38,9322	35,4789	32,6706	29,6151	27,6620	26,1711	24,9348	23,8578	22,8876	21,9915	21,1470	20,3372
22	48,2676	45,2041	42,7957	40,2894	36,7807	33,9245	30,8133	28,8224	27,3015	26,0393	24,9390	23,9473	23,0307	22,1663	21,3370
23	49,7276	46,6231	44,1814	41,6383	38,0756	35,1725	32,0069	29,9792	28,4288	27,1413	26,0184	25,0055	24,0689	23,1852	22,3369
24	51,1790	48,0336	45,5584	42,9798	39,3641	36,4150	33,1962	31,1325	29,5533	28,2412	27,0960	26,0625	25,1064	24,2037	23,3367
25	52,6187	49,4351	46,9280	44,3140	40,6465	37,6525	34,3816	32,2825	30,6752	29,3388	28,1719	27,1183	26,1430	25,2218	24,3366
26	54,0511	50,8291	48,2898	45,6416	41,9231	38,8851	35,5632	33,4295	31,7946	30,4346	29,2463	28,1730	27,1789	26,2395	25,3365
27	55,4751	52,2152	49,6450	46,9618	43,1945	40,1133	36,7412	34,5736	32,9117	31,5284	30,3193	29,2266	28,2141	27,2569	26,3363
28	56,8918	53,5939	50,9936	48,2782	44,4608	41,3372	37,9169	35,7150	34,0266	32,6205	31,3909	30,2791	29,2486	28,2740	27,3362
29	58,3006	54,9662	52,3355	49,5878	45,7223	42,5569	39,0875	36,8538	35,1394	33,7109	32,4612	31,3308	30,2825	29,2908	28,3361

Tabla 10.3. Valores críticos de la distribución Ji-Cuadrado.

En nuestro ejemplo los grados de libertad (K-1) es igual a 2, por tanto si el valor resultante de aplicar la formula de ji-cuadrado es inferior al valor de la tabla para 2 g/l y un nivel de .05 (5.99) aceptamos la hipótesis nula, las frecuencias en la población son iguales.

3.2. Prueba Ji-Cuadrado con dos variables

El mismo profesor decide preguntar a sus alumnos (N=48) por las preferencias a la hora de realizar la tipología de la prueba, sin embargo ahora va a tener en cuenta el género, es decir las opciones preferidas por los hombres y por las mujeres. Se realiza la misma tabla, pero se añade la variable género.

	HOMBRES	MUJERES	TOTAL
Prueba oral	8	6	14
Prueba escrita	9	6	15
Prueba objetiva	11	8	19
TOTAL	28	20	48

Tabla 10.4. Distribución Hombres/Mujeres por prueba.

Para obtener las frecuencias esperadas, éstas se calculan a través del producto de los totales de filas y columnas dividido por el número total de casos, $28 \times 14 / 48$, tal y como se muestra en la tabla 10.5:

	HOMBRES	MUJERES	TOTAL
Prueba oral	8 (8.16)	6(5.8)	14
Prueba escrita	9 (8.75)	6(6.25)	15
Prueba objetiva	11 (11)	8(7.91)	19
TOTAL	28	20	48

Tabla 10.5. Frecuencias esperadas.

Frecuencias observadas (f_o)	Frecuencias esperadas (f_e)	$(f_o - f_e)^2$	$\frac{[f_o - f_e]^2}{f_e}$
8	8.16	.0256	.003
9	8.75	.062	.007
11	11	0	0
6	5.8	.04	.006
6	6.25	.062	.009
8	7.91	.09	.011

Tabla 10.6. Fórmula χ^2 .

Posteriormente se aplica la fórmula de χ^2 y se obtiene un valor que hemos de comparar con el de la tabla 10.3, pero esta vez teniendo en cuenta que los grados de libertad vendrán determinados por: (filas-1) $\times$ (columnas-1). En el ejemplo los grados de libertad es igual a 2. Del mismo modo que con una sola variable, si el valor de la tabla es mayor que el de χ^2 se acepta la hipótesis nula (Tabla 6).

3.3. Prueba de los signos

Supongamos el caso del profesor que desea que sus alumnos valoren de una escala de 1 a 5 metodologías de enseñanza diferentes. La muestra consta de 6 sujetos y cada uno de ellos experimenta dichas metodologías. Los resultados son los siguientes:

X_1	X_2	$X_1 - X_2$
5	4	1
3	5	-2
5	5	0
5	3	2
4	2	2
5	1	4

Tabla 10.7

Nos encontramos con 4 puntuaciones positivas, una negativa y otra puntuación que no debemos tener en cuenta ya que su valor es cero:

$$\chi^2 = (4-1)^2/5 = 1.8$$

Cuando se compara este valor con el de la tabla 1, para un grado de libertad (3.841) se aprecia que es superior a 1.8, por tanto se acepta la hipótesis nula sobre la no existencia de diferencias significativas en la preferencia de una metodología u otra.

3.4. Prueba de Wilcoxon

Un profesor desea comprobar la eficacia de un programa de entrenamiento virtual con un grupo de 30 sujetos. Antes de aplicar el programa realiza una primera medición para, después, poder comprobar la eficacia del programa en una segunda medición. Los resultados fueron:

X_1	X_2	$X_1 - X_2$	R_D	R+	R-
15	14	1	1.5	1.5	
10	15	-5	7.5		7.5
9	5	4	6	6	
4	3	1	1.5	1.5	
8	2	6	9.5	9.5	
5	11	-6	9.5		9.5
7	5	2	3.5	3.5	
5	8	-3	5		5
12	7	5	7.5	7.5	
9	7	2	3.5	3.5	
N=30	..	...	...	$\Sigma \dots$	$\Sigma \dots$

Tabla 10.8

Posteriormente se ordenan los valores de menor a mayor, calculando la mediana en los valores repetidos (R_D). Se calcula después el sumatorio de R_D y se sustituye en la fórmula:

$$T_E = N(N-1)/4$$

De esta manera obtenemos T_e , siendo N el número total de observaciones. El siguiente paso consiste en calcular $R+(T_1)$ y $R-(T_2)$, es decir se separan los valores positivos de los negativos, se mantiene el mismo orden y se calcula el sumatorio de ambos valores. A continuación se calcula el valor de la desviación estándar de la T de Wilcoxon:

$$\sigma T_m = \sqrt{\frac{(2N+1)T_e}{24}}$$

Una vez determinada la desviación estándar se calcula el valor Z :

$$Z = \frac{T - m_T}{\sigma_T}$$

El valor Z obtenido se ha de comparar con los valores de la tabla en función del nivel de significación y de la potencia a un nivel de .05. Si dicho valor es superior al valor de Z se acepta la hipótesis nula, por tanto no existirían diferencias significativas entre las puntuaciones primera y segunda tras haber aplicado el nuevo programa.

Potencia	.05	.025	.01	.005
	Prueba bilateral (α)			
	.10	.05	.02	.01
.25	.97	1.29	1.65	1.90
.50	1.64	1.96	2.33	2.58
.60	1.90	2.21	2.58	2.83
.67	2.08	2.39	2.76	3.01
.70	2.17	2.48	2.85	3.10
.75	2.32	2.63	3.00	3.25

Tabla 10.9. Tabla en función del nivel de significación y de la potencia.

3.5. Prueba de Friedman

Un investigador dispone de 4 grupos a los cuales ha aplicado tres pruebas diferentes. Los resultados aparecen en la siguiente tabla:

	PRUEBA 1	PRUEBA 2	PRUEBA 3
GRUPO A	3 (2)	1 (1)	4 (3)
GRUPO B	4 (2)	2 (1)	6 (3)
GRUPO C	7 (3)	4 (2)	2 (1)
GRUPO D	5 (2)	6 (3)	3 (1)

Tabla 10.10. Prueba de Friedman.

El primer paso consiste en ordenar el grupo fila, de mayor a menor, de tal manera que el nuevo orden aparece entre paréntesis. Se aplica la formula:

$$\chi^2 = \frac{12}{Nk(k+1)}$$

Donde N= 4 (número de grupos), K=3 (número de pruebas). Una vez obtenido el valor de χ^2 , se localiza en la tabla 2 (valores para ji-cuadrado) para un $\alpha=.05$ y 2 grados de libertad (K-1). Si el valor de la tabla es superior al calculado anteriormente se acepta la hipótesis nula, en caso contrario se rechaza.

4. APLICACIÓN PRÁCTICA CON ORDENADOR

4.1. Prueba Ji-Cuadrado con una sola variable

Empleando el mismo ejemplo realizado manualmente:

- 1º. Abra el fichero desde File/Open/Data.
- 2º. Seleccione la opción Tables: Two-Way del menú Analyze.
- 3º. En la ventana de diálogo (figura 10.12) seleccione la/s variable/es. Con el botón Add las incorpora a la cuadrícula de la derecha.
- 4º. En la parte inferior derecha verá que están marcados unos cuadros con varias opciones. Algunas vienen marcadas con una cruz por defecto. Otras pueden seleccionarse picando sobre ellas con el botón izquierdo del ratón, o anularlas picando sobre las que están marcadas.

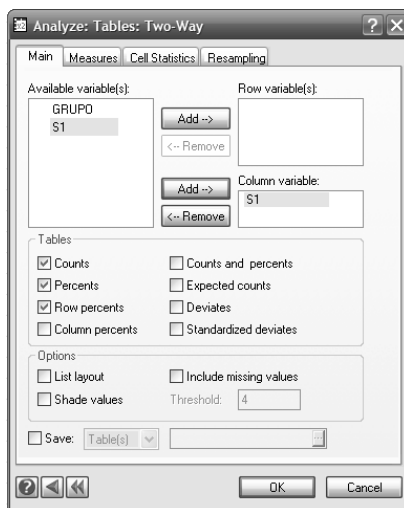


Figura 10.12. Ventana Tables: Two-Ways.

5°. Pulse OK. Los resultados que verá en la pantalla principal no incluyen el análisis χ^2 para tablas unidimensionales (figura 10.13). Esto lo puede hacer a partir de la fórmula conocida:

$$\chi^2_{(F-1)(C-1)} = \sum_{i=1}^F \sum_{j=1}^C \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

MYSTAT: A Student Version of SYSTAT - [Untitled]

File Edit View Data Utilities Graph Analyze Quick Access Window Help

Workspace: Startpage Untitled.syo* Untitled1.syz*

Perents of Total Count

S1(rows) by S1(columns)

	11	16	18	19	Total	N
11	16,667	0,000	0,000	0,000	16,667	1,000
16	0,000	50,000	0,000	0,000	50,000	3,000
18	0,000	0,000	16,667	0,000	16,667	1,000
19	0,000	0,000	0,000	16,667	16,667	1,000
Total	16,667	50,000	16,667	16,667	100,000	6,000
N	1,000	3,000	1,000	1,000		

Row Percents

S1(rows) by S1(columns)

	11	16	18	19	Total	N
11	100,000	0,000	0,000	0,000	100,000	1,000
16	0,000	100,000	0,000	0,000	100,000	3,000
18	0,000	0,000	100,000	0,000	100,000	1,000
19	0,000	0,000	0,000	100,000	100,000	1,000
Total	16,667	50,000	16,667	16,667	100,000	
N	1,000	3,000	1,000	1,000		

Figura 10.13. Resultados χ^2 tablas unidimensionales.

Para calcular el valor p para comprobar la significación debe proceder del modo siguiente: Data/Transform/Let e irá a la ventana de diálogo que observa en la figura 10.14.

Data: Transform: Let

Available variable(s):

- GRUPO
- S1

Function type: Mathematical

Functions:

- SQRT
- LOG
- L10
- EXP

Add to:

Variable: Let pvalu

Expression: = 1-ccf(1,4,2)

OK Cancel

Figura 10.14. Ventana Data:Transform:Let.

6°. Pulse OK. Observará junto a la columna de la variable la proporción con la que cada frecuencia observada es significativa (Figura 10.15).

	S1	PVALU	VAR	VAR	VAR	VAR	VAR
1	18,000	0.497					
2	19,000	0.497					
3	11,000	0.497					
4	16,000	0.497					
5	16,000	0.497					
6	16,000	0.497					
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							
17							

Figura 10.15. Proporción con la que cada frecuencia observada es significativa.

4.2. χ^2 para tablas bidimensionales

Siga los pasos siguientes:

1º. Abra el fichero desde File/Open/Data.

2º. Seleccione Tables: Two-Way. En la ventana de diálogo (figura 10.16) elija la/las variable/es dependientes y con el botón Add la/las inserta en la cuadrícula superior derecha. Seguidamente seleccione la variable independiente y llévela a la casilla correspondiente. Recuerde que todas las variables insertadas en la primera casilla serán analizadas en función de esta última variable. La posición de las variables en uno y otro espacio va a depender de los objetivos de la investigación.

3º. Seguidamente elija las opciones que le interesen: porcentaje, porcentaje por fila, etc.

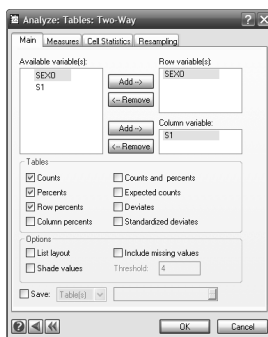


Figura 10.16. Ventana Tables: Two-Way.

4º. Pulse Ok y verá como se relacionan una serie de datos que permitirán interpretar los resultados alcanzados. En este ejemplo, hemos extraído una de las tablas de contingencia:

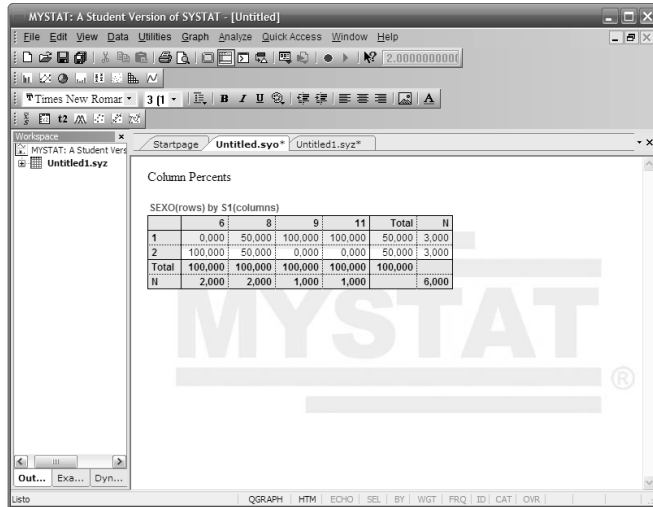


Figura 10.17. Tabla de contingencia χ^2 bidimensional.

4.3. Prueba de los signos

Continuando con el ejemplo anterior, los pasos a seguir en Mystat son:

1º. Abra el fichero desde File/Open/Data (Figura 10.18).

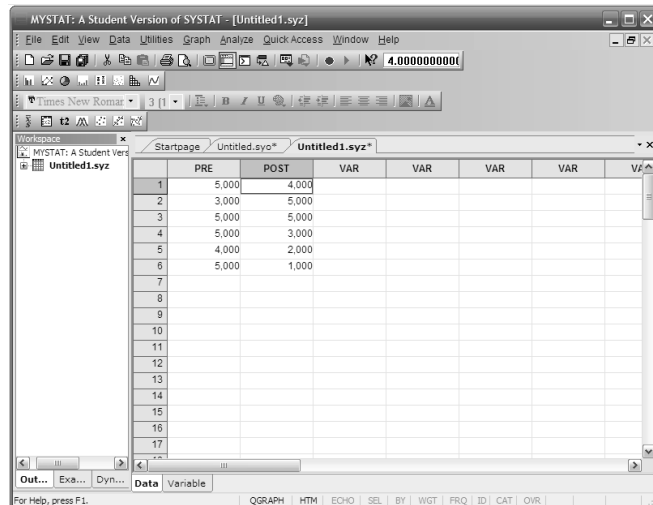


Figura 10.18. Hoja de datos.

2°. Seleccione la opción Nonparametrics Tests: Sign del menú Analyze. En la ventana de diálogo (figura 10.19) seleccione la/s variable/es y pulse el botón Add para trasladarlas a la cuadrícula de la derecha.

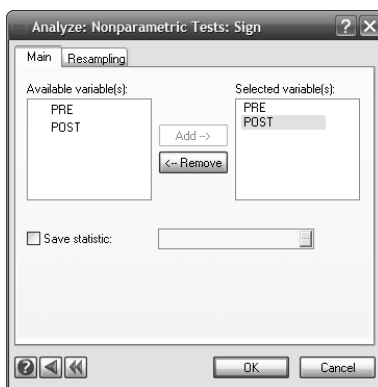


Figura 10.19. Ventana Tests Sign.

3°. Pulse OK y obtendrá la proporción con la que la diferencia entre el PRE y el POST es significativa (Figura 10.20). En el ejemplo el valor obtenido es de .375.

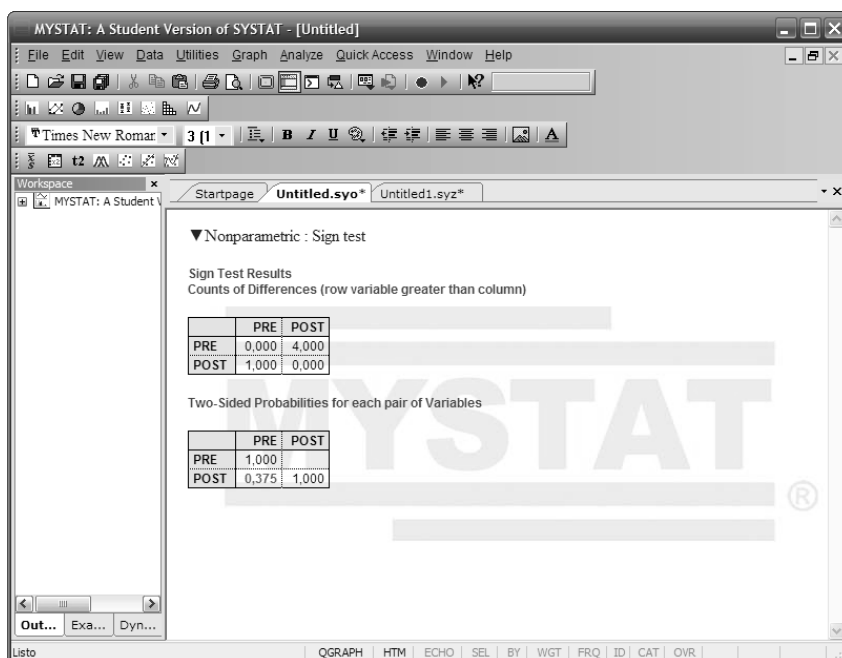


Figura 10.20. Ventana de resultados Test Sign.

4.4. Prueba de Wilcoxon

Los pasos a seguir son:

- 1°. Abra el fichero con la secuencia File/Open/ Data.
- 2°. Seleccione la opción Nonparametrics Tests: Wilcoxon del menú Analyze.
- 3°. En la ventana de diálogo (figura 10.21) seleccione la/s variable/es y pulse el botón Add para trasladarlas a la cuadrícula de la derecha.

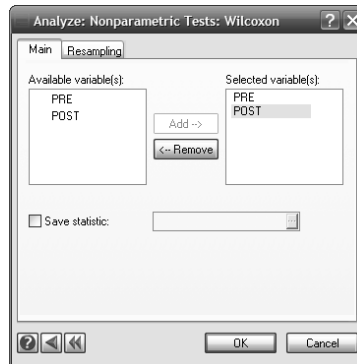


Figura 10.21. Ventana Tests Wilcoxon.

4°. Pulse OK y se obtienen los resultados (figura 10.22) a partir de los datos mostrados en el ejemplo. Su interpretación es similar a la prueba de los signos, el valor .5474 nos indica la intensidad de la relación.

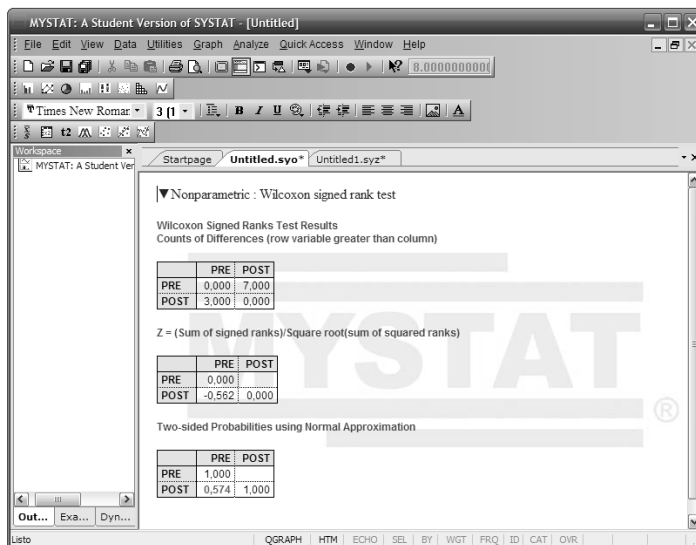


Figura 10.22. Ventana de resultados Test Wilcoxon.

4.5. Prueba de Friedman

Los pasos son los siguientes:

- 1°. Abra el fichero File/Open/ Data.
- 2°. Seleccione la opción Nonparametrics Tests: Friedman del menú Analyze.
- 3°. En la ventana de diálogo (figura 10.23) seleccione la/s variable/es y pulse el botón Add para trasladarlas a la cuadrícula de la derecha.

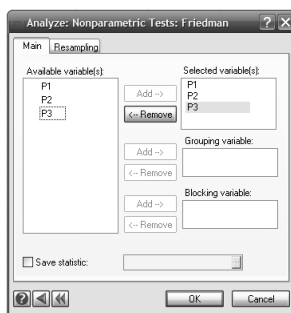


Figura 10.23. Ventana Tests Friedman.

- 4°. Pulse OK y obtendrá los resultados mostrados en la figura 10.24.

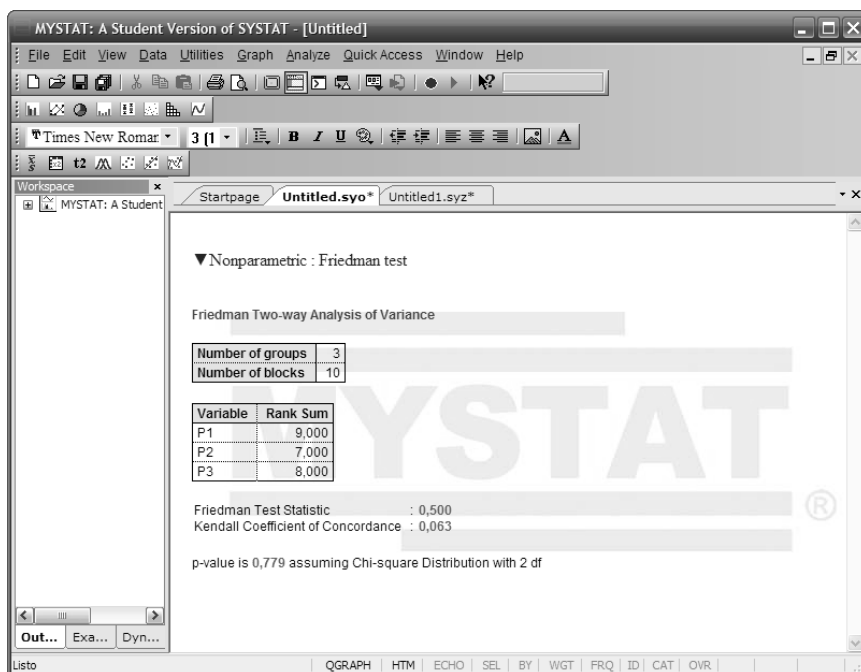


Figura 10.24. Venta de resultados Tests Friedman.

En los resultados que aparecen en pantalla observará una columna con el nombre de las variables y a su lado la suma de sus rangos. A continuación aparece el coeficiente de Friedman (.500) y el coeficiente de concordancia de Kendall (.063); en la última línea verá que la probabilidad de la significación es de .779 para dos grados de libertad.

5. EJERCICIOS PROPUESTOS

5.1. En un estudio sobre género se contrastaron las variables sexo y nivel de ingresos. Los resultados obtenidos, en una muestra de 528 individuos, se presentan en la tabla 10.11:

	ALTO	MEDIO	BAJO
HOMBRE	40	30	32
MUJER	34	23	26

Tabla 10.11. Género y nivel de ingresos.

Determine si el nivel de ingresos está condicionado por el género. Interprete los resultados.

5.2. Se evalúa el rendimiento de doce estudiantes universitarios en una determinada materia, durante el primer cuatrimestre, a través de una prueba objetiva. Después, se evalúa nuevamente a estos alumnos, en el segundo cuatrimestre, en la misma materia tras haberse empleado una metodología diferente. ¿El primer método empleado ofrece resultados significativamente diferentes al utilizado en el segundo caso?

H_0 no existen diferencias significativas entre el primer y segundo método

H_1 existen diferencias significativas entre el primer y segundo método

$\alpha = .05$

PRUEBA A	PRUEBA B	DIFERENCIAS
20.4	21.7	-1.3
25.4	26.3	-0.9
25.6	26.8	-1.2
25.6	28.1	-2.5
26.6	26.2	0.4
28.6	27.3	1.3
28.7	29.5	-0.8
29.0	32.0	-3.0
29.8	30.9	-1.1
30.5	32.3	-1.8
30.9	32.3	-1.4
31.1	31.7	-0.6

Tabla 10.12. Rendimiento en 3 pruebas.

5.3. Un investigador quiere determinar si la visión de una película sobre delincuencia juvenil cambia las opiniones de los miembros de un grupo sobre la severidad de las sanciones hacia los jóvenes. Extrae aleatoriamente una muestra de 100 adultos y les pide pronunciarse sobre la mayor o menor severidad de los castigos. A continuación les presenta la película y obtiene los resultados siguientes:

		Opinión antes de la película	
		Menos	Más
Opinión después de la película	Más	59	7
	Menos	8	26

Tabla 10.13

H_0 no existen diferencias significativas entre el primer y segundo método

H_1 existen diferencias significativas entre el primer y segundo método

$\alpha = .05$

¿Existen diferencias estadísticamente significativas?

5.4. Un psicopedagogo quiere analizar la comprensión social de los alumnos de una escuela infantil. Evalúa esta dimensión a partir de las respuestas que los alumnos ofrecen a una serie de cuestiones referentes a imágenes que representan distintas situaciones sociales. Cada alumno obtiene, de esta forma, un resultado comprendido entre 0 y 100. El psicopedagogo no puede afirmar que las diferencias observadas entre los resultados son numéricamente exactas (no puede decir que un resultado de 60 es el doble de un resultado de 30, ni que la diferencia entre 60 y 40 son el doble exactamente de la diferencia entre 40 y 30). Para probar los efectos que tiene la asistencia a la escuela sobre la comprensión social de los alumnos utiliza 8 pares de gemelos. De forma aleatoria, uno de los gemelos va a la escuela mientras que el otro permanece en la casa durante un trimestre. Al final del trimestre, se vuelve a estimar la comprensión social de cada uno de los alumnos.

H_0 : no hay diferencias entre la comprensión social de los alumnos que han permanecido en casa y la de los alumnos que están en la escuela.

Mostrados los resultados en la siguiente tabla, analice si existen diferencias significativas.

Pares	Niños escolarizados	Niños no escolarizados
a	82	63
b	69	42
c	73	74
d	43	37
e	58	51
f	56	43
g	76	80
h	65	62

Tabla 10.14. Resultados niños escolarizados.

5.5. A un grupo de estudiantes se les convoca, en cinco momentos diferentes de una mañana ($K=5$), en una clase de un centro. Se aplica un test (calificaciones comprendidas entre 0 y 10) en cada uno de los momentos. Los resultados son los siguientes:

SUJETOS	PRUEBA 1	PRUEBA 2	PRUEBA 3	PRUEBA 4	PRUEBA 5
A	10	6	8	9	7
B	6	3	4	5	2
C	7	6	5	8	4
D	4	1	3	5	2

Tabla 10.15

¿Existen diferencias entre los datos a un nivel de significación de .05?

Capítulo 11.

ANÁLISIS DE DATOS CUALITATIVOS CON ATLAS.ti

*Tomás Izquierdo Rus
Gustavo Montealegre Lynett*

INTRODUCCIÓN

En los últimos años se ha producido un aumento considerable de las investigaciones cualitativas que ha ido acompañado de la aparición de los programas de análisis de datos, que facilitan enormemente el trabajo. Es importante tener presente que el programa de análisis es una herramienta que no sustituye al investigador, por tanto se convierte en una ayuda para la realización del análisis de datos. En el ámbito de las Ciencias Sociales el programa ATLAS.ti es uno de los softwares cualitativos más utilizados, por la gran cantidad de ventajas que presenta respecto a otros programas.

Basado en la teoría fundamentada (Grounded Theory) que plantea una construcción inductiva de la teoría, ATLAS.ti permite acciones como la segmentación del texto en citas, la codificación, la realización de comentarios y anotaciones y toda una serie de actividades que, de no disponer del programa, realizaríamos sirviéndonos de otras herramientas como lápices de colores, tijeras, fichas, papel...

ATLAS.ti es una herramienta en el ámbito de la investigación cualitativa que ayuda al investigador, sobre todo cuando maneja grandes volúmenes de material investigativo en formato de texto, audio, gráfico y video. Además facilita el trabajo de equipos de investigación porque permite que varios autores trabajen en un mismo proyecto de investigación, plasmando las tareas realizadas por cada uno de ellos e identificando las respectivas fuentes de ideas.

La nueva versión de ATLAS.ti introduce como novedades la posibilidad de trabajar con documentos en formato PDF. Esta versión permite trabajar con imágenes bajo el formato de Google Earth, efectuar análisis con datos de encuestas en línea, igualmente se ha incluido la edición de documentos en vivo con actualizaciones dinámicas multiproyecto.

El presente capítulo está diseñado con la versión 7.5.2 que es la más actual en el momento de elaboración de esta guía. No obstante los conceptos básicos y la forma de trabajar con el programa es la misma en las diferentes versiones. El programa se puede encontrar en la página de internet **www.atlasti.com**

1. OBJETIVOS

Los objetivos perseguidos con el presente capítulo son los siguientes:

- 1.1. Proporcionar una orientación inicial para trabajar con el programa ATLAS.ti.
- 1.2. Mostrar las fases típicas de un proyecto con ATLAS.ti.
- 1.3. Conocer las funciones principales y alcance del programa ATLAS.ti en el desarrollo de una investigación de tipo cualitativo.
- 1.4. Generar los informes necesarios para la realización de trabajos de investigación de tipo cualitativo.

2. CONTENIDOS

1. INICIO DEL PROGRAMA

Para ejecutar el programa ATLAS.ti, una vez que lo haya instalado en su ordenador, debe proceder de la siguiente forma: todos los programas, busque la carpeta Scientific Software y elija la opción ATLAS.ti, tal y como se muestra en la figura siguiente (Figura 11.1).

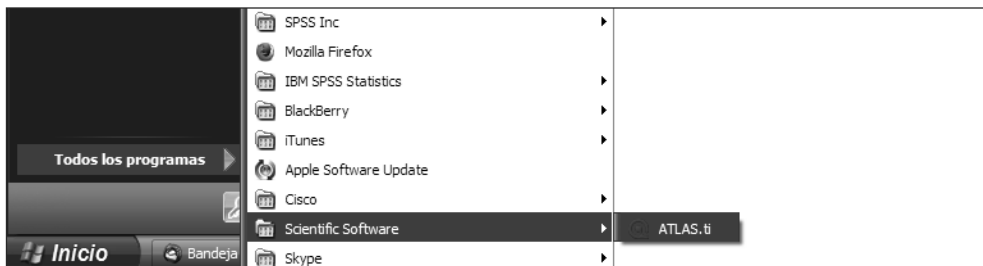


Figura 11.1. Inicio del programa ATLAS.ti

2. PROCESO GENERAL DE ANÁLISIS

Instalado el programa, se recomienda seguir el siguiente esquema del proceso general de análisis de datos con ATLAS.ti (Figura 11.2), en cualquier proyecto de investigación.

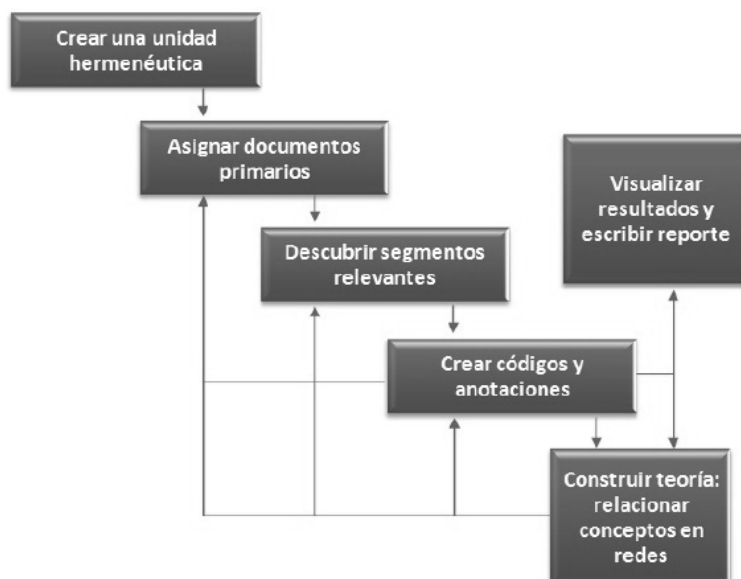


Figura 11.2. Proceso general de análisis (Muhr y Frieze, 2004)¹.

En los siguientes apartados se detallan cada una de las etapas del proceso general de análisis de datos.

3. HU- UNIDAD HERMENÉUTICA

La Unidad Hermenéutica (HU) proporciona la estructura de datos para cada proyecto en particular con ATLAS.ti. Los documentos primarios que representan a los datos, códigos, enlaces conceptuales y memos formarán parte de la HU. Al activar una HU con un solo archivo estamos activando todos los materiales asociados automáticamente. Para crear una HU hemos de dirigirnos al menú Project y seleccionar la opción New Hermeneutic (Figura 11.3).

Seleccione un nombre para identificar la UH creada y, posteriormente, grabe la misma con las opciones Project y Save As. Le indicará la ruta donde guardar el fichero, por defecto aparecerá: Mis documentos - Scientific Software - ATLAS.ti-TextBank.

Puede crear un comentario con la opción *Edit Comment* con el fin de introducir un primer comentario o reflexión sobre la HU creada. Conocer el autor de cada elemento es fundamental con el objeto de identificar el trabajo de diferentes autores en un mismo proyecto. Para ello seleccionamos la opción *Documents, Miscellaneous y Change Autor*, y así figurará el nombre establecido por el usuario.

¹ Muhr, T. & Frieze, S. (2004). *User's manual for ATLAS.ti 5.0* (2ª. Ed.). Berlin: Scientific Software Development.

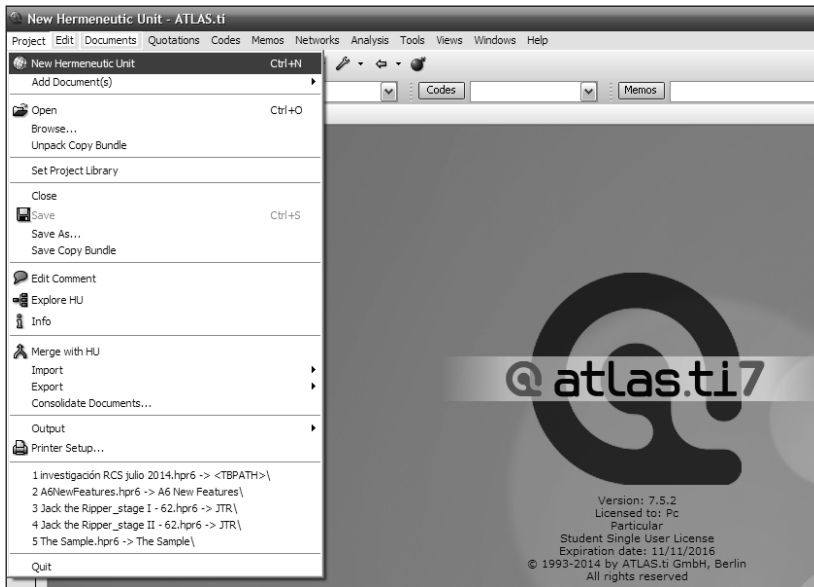


Figura 11.3. Unidad hermenéutica.

4. PDocs- DOCUMENTOS PRIMARIOS

Los documentos primarios representan a los materiales de texto, gráficos, de audio, de foto o video que deseamos analizar. El contenido de los *PDocs* se almacena en archivos de datos en el ordenador, en la ruta especificada en la UH (Figura 11.4).

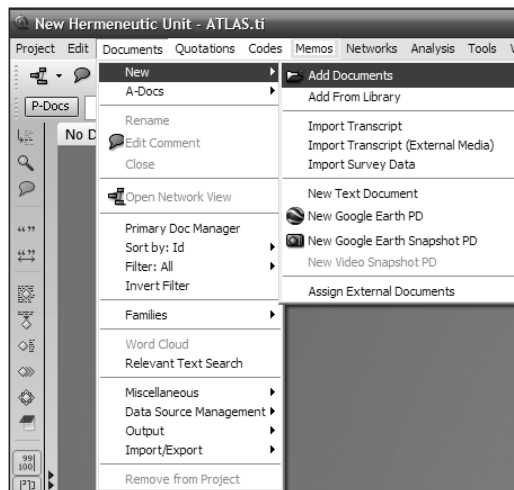


Figura 11.4. Asignar documentos.

Los textos (en formato word y RTF) asignados a una HU como documentos primarios se pueden editar en cualquier momento del proceso analítico sin importar que el documento haya sido codificado o no, ya que existe una adecuada sincronización entre ellos. Los documentos primarios pueden ser datos de texto, gráficos, audio y video. También se pueden editar presentaciones de powerpoint o tablas de excell.

Vincular los documentos primarios correspondientes a la HU es la siguiente tarea a realizar (Figura 11.5). Se aconseja que se cree una carpeta determinada (ubicada dentro de la carpeta TEXT BANK, que queda instalada con el programa ATLAS.ti) para almacenar los documentos a analizar.

Compruebe el *manager de documentos primarios* y observará los documentos vinculados.



Figura 11.5. Manager de documentos.

5. CREACIÓN DE CITAS (Quotations)

Es un segmento de un PDocs que tiene interés o relevancia para el investigador. En los documentos de texto una cita puede comprender desde un único carácter, una palabra, una frase o párrafo. Las citas las crea el investigador o investigadores del proyecto. Las citas pueden ser textuales; frases, turnos de palabras; gráficas: zona seleccionada dentro de un PDocs gráfico; o multimedia: fragmento de audio o video con una extensión en minutos, segundos y milisegundos.

Al crear una cita es importante abrir el PDocs sobre el que vamos a trabajar (Figura 11.6), haciendo click sobre el archivo de la lista de archivos ubicada en el manager de PDocs.

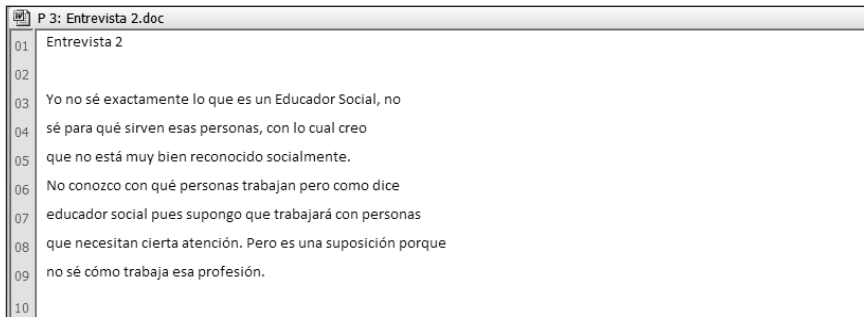


Figura 11.6. Documento primario.

Seleccione el texto que considere importante y escoja la opción *Quotations, Create Free Quotations*. Visualizará una línea vertical en el lado derecho, que indica que tiene creada una cita. Al abrir el manager de citas en *Quotations, Quotations Manager* encontrará las citas que ha creado (Figura 11.7).

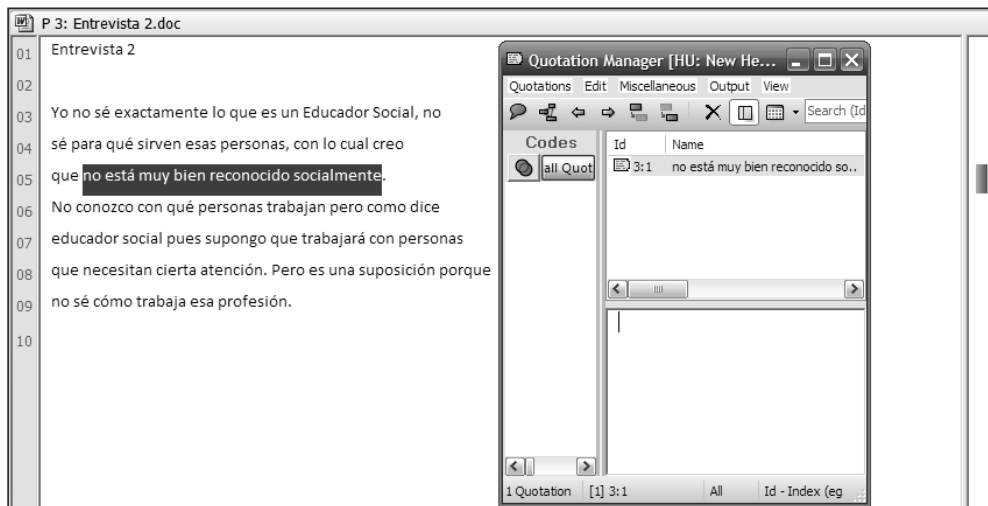


Figura 11.7. Creación de citas.

Creadas las citas, se pueden establecer relaciones entre citas bien dentro del mismo documento primario o en otro documento. El proceso a seguir es el siguiente: identifique la cita y despliegue el menú *Quotations*, elija la opción *Hiper-link, Create Link Source*, luego identifique la otra cita y siga el mismo proceso anterior pero marcando *Create Link Target*. Se nos desplegará un menú con varias posibilidades para denominar la relación, marcamos la opción y finalizamos la relación (Figura 11.8).

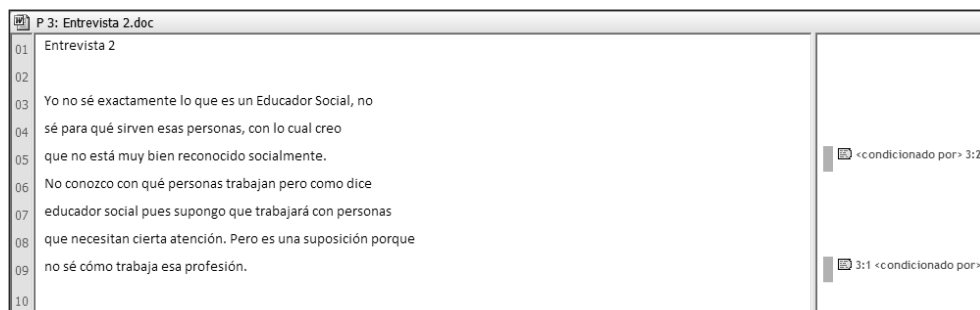


Figura 11.8. Relaciones entre citas.

Para visualizar de forma gráfica las relaciones creadas entre citas procederemos de la siguiente forma: primeramente seleccionamos la cita y luego marcamos del menú Quotations la opción *Open Network View* (Figura 11.9).

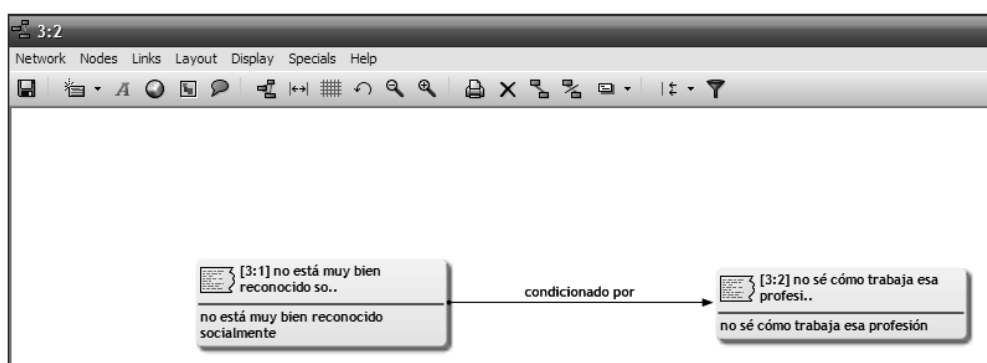


Figura 11.9. Visualización gráfica de relaciones entre citas.

6. CODIFICACIÓN (Codes)

Es el procedimiento de asociar palabras clave con selecciones de texto. Dentro del programa ATLAS.ti la codificación es la asociación que se establece entre una cita y un código. El programa permite cuatro procedimientos de codificación: codificación abierta, codificación en vivo, codificación por lista y codificación rápida.

La codificación conlleva un proceso que consta de los siguientes pasos: 1) identificar la cita en el texto, 2) desplegar el menú *Codes* y elegir el procedimiento que nos interese del desplegable *Coding*, 3) asignamos el nombre al código en caso de haber elegido la opción *Open Coding* (Figura 11.10) y aparecerá en el margen derecho de la pantalla.

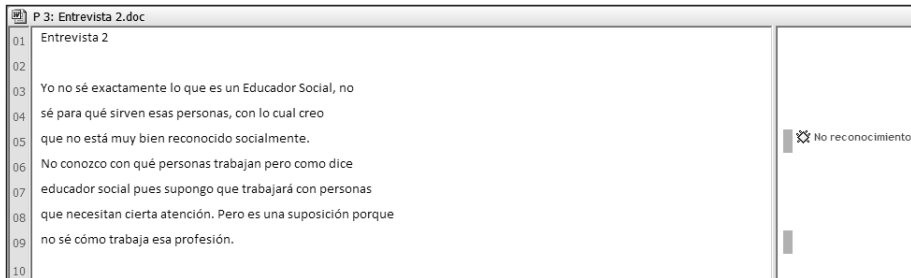


Figura 11.10. Establecimiento de códigos.

De igual modo que procedimos con las citas, podemos visualizar los códigos desde la opción *Codes Manager*, seleccionamos el código y luego marcamos *Open Network View*. Se genera una red semántica con el código seleccionado (Figura 11.11). Podemos importar cuantos objetos consideremos (citas, códigos...) escogiendo la tarea *Import Nodes* del menú *Nodes*.

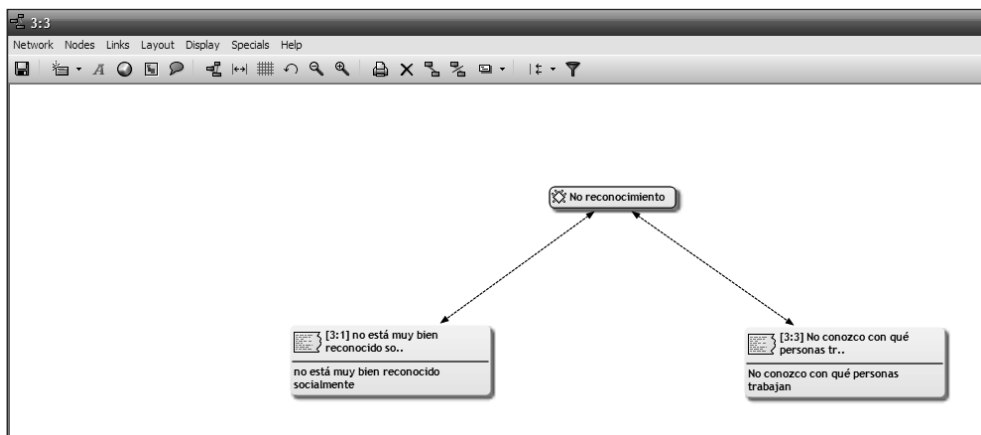


Figura 11.11. Visualización gráfica de códigos.

7. FAMILIAS (*Families*)

Las familias se emplean para agrupar un grupo de objetos (documentos primarios, códigos, memos...) con el objeto de facilitar el análisis en la tarea investigadora. Se puede emplear una familia de códigos para agrupar a códigos que comparten ciertas características. Esta tarea reduce considerablemente la cantidad de objetos a los que tiene que prestar atención el investigador, lo que aumenta el rendimiento y productividad del análisis.

En el menú *Codes*, *Edit Families*, *Open Family Manager* elegimos la opción *New Family* y asignamos el nombre de la familia. A continuación seleccionamos los códigos que forman parte de esta familia (Figura 11.12).



Figura 11.12. Creación de familias.

Al seleccionar la familia podemos visualizarla desde *Code Family Manager*, *Open Network View* e importando objetos que considere pertinentes (citas, códigos...). De esta forma se crea una red semántica que facilita el proceso analítico (Figura 11.13).

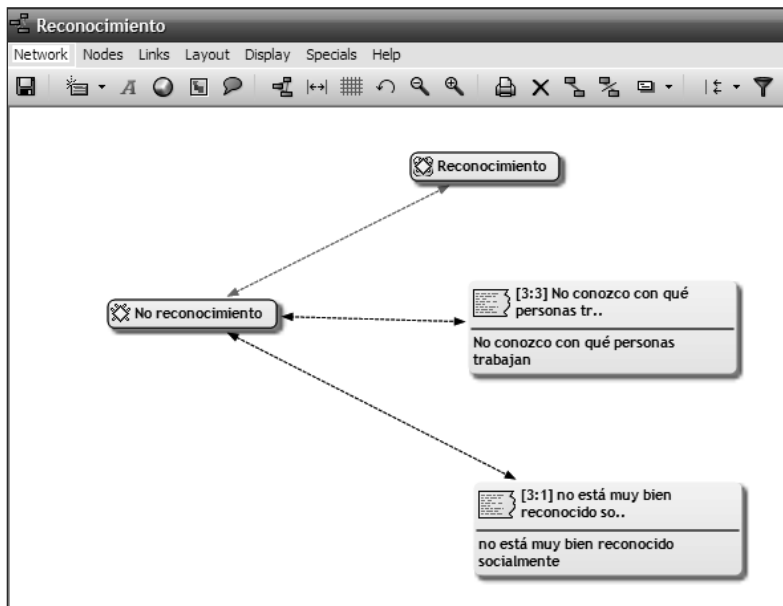


Figura 11.13. Visualización gráfica de familias.

8. HERRAMIENTA DE CONSULTA (Query Tool)

La herramienta de consulta se utiliza para recuperar citas empleando los códigos a los que se asociaron durante el proceso de codificación. Difiere de una búsqueda textual para buscar ocurrencias de texto que cumplan un determinado patrón, Query Tools permite recuperar combinaciones de códigos.

Esta herramienta se ejecuta seleccionando, del menú principal de la HU, la ventana *Analysis, Query Tools* y eligiendo, con doble clic, los códigos o familias a relacionar. El programa ofrece tres conjuntos de operadores: booleanos, semánticos y de proximidad. Los operadores booleanos (Figura 11.14) son los más utilizados y nos permiten combinaciones de palabras clave de acuerdo a un conjunto de operaciones.

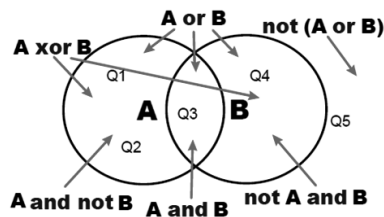


Figura 11.14. Consulta booleana representada en diagramas de Venn.

Los operadores booleanos se encuentran cerca del margen izquierdo de la ventana Query Tools (Figura 11.15). Hay cuatro operadores disponibles en el programa: OR (o), XOR (uno u otro), AND (y) y NOT (no). En la casilla inferior derecha se muestra el resultado de las relaciones efectuadas.

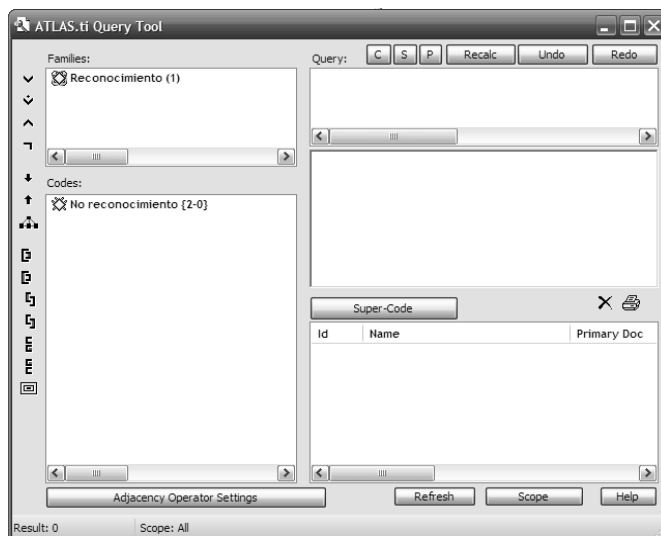


Figura 11.15. Herramienta de consulta.

9. INFORMES DE RESULTADOS

El programa ATLAS.ti posibilita algunas funciones para visualizar los resultados tras el análisis de los PDocs. Uno de los listados proporcionado por ATLAS.ti es la tabla de Códigos-Documentos Primarios. Este listado muestra la frecuencia de todos los códigos (o los seleccionados) por documentos (o grupos de documentos).

El proceso a seguir es el siguiente: despliegue el menú *Code Manager*, elija la opción *Output, Codes-Primary-Documents-Table* (Figura 11.16 y Figura 11.17), luego marcamos Excel. Se nos desplegará una tabla con las frecuencias de los códigos en los documentos analizados.

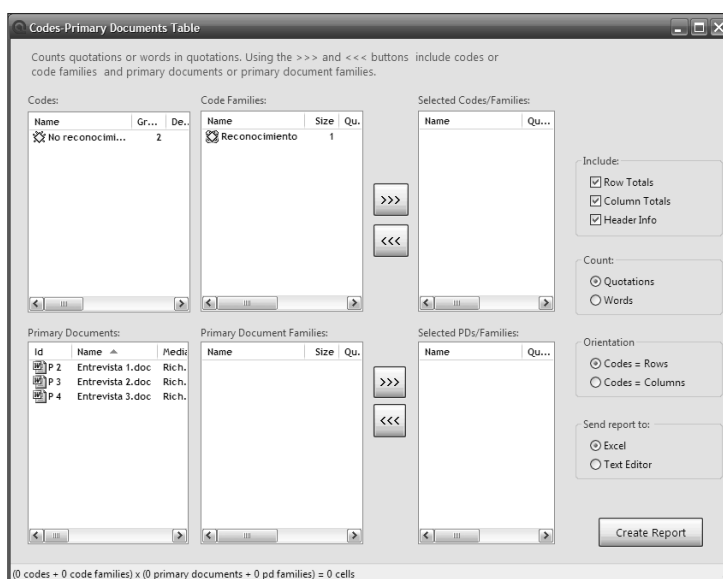


Figura 11.16. Códigos-Documentos Primarios.

CODES-PRIMARY-DOCUMENTS-TABLE			
Código-filtro: Todos [1]			
DP-Filtro: Todos [1]			
Cita-filtro: Todos [2]			
	P 2:	TOTALES:	
	Entrevista1.doc		
No			
Reconocimiento	2		2
TOTALES:	2		2

Figura 11.17. Frecuencias de códigos.

Esta información es especialmente importante a la hora de difundir los trabajos de investigación en sus diferentes modalidades, especialmente en la presentación e interpretación de los resultados y las conclusiones extraídas, en consonancia con los objetivos planteados.

Para finalizar este capítulo, las diversas funciones presentadas con ATLAS.ti nos han permitido hacernos una idea de lo productivo que puede llegar a ser este programa en la realización de trabajos de investigación cotidianos. El propósito central de este capítulo ha sido mostrar las principales funciones que permiten este programa y la elección de las mismas acorde a las diferentes necesidades de investigación.

3. EJERCICIOS PROPUESTOS

3.1. Realice una breve entrevista a sus compañeros que analice las impresiones sobre la formación adquirida en el periodo de prácticas en cualquiera de las titulaciones oficiales de la Universidad de Murcia. A continuación lleve a cabo las siguientes tareas:

- a) Transcriba la entrevista.
- b) Redacte un listado de categorías.
- c) Elabore una Unidad Hermenéutica en ATLAS.ti

3.2. Analice la entrevista anterior con los recursos proporcionados por ATLAS.ti a través de los siguientes pasos:

- d) Segmente y codifique la entrevista
- e) Introduzca anotaciones (comentarios, autoría, fechas, etc.)
- f) Elabore familias de códigos.
- g) Establezca relaciones entre citas y códigos.

3.3. Represente gráficamente, a través de una network, los elementos creados y las relaciones establecidas en el ejercicio anterior.

3.4. Presente un breve informe del análisis de la entrevista incluyendo la networks.

Capítulo 12.

INFORME ESCRITO DE LA INVESTIGACIÓN

Fuensanta Monroy Hernández
Fuensanta Hernández Pina

INTRODUCCIÓN

Esta edición incluye dos nuevos capítulos que abordan de forma específica la elaboración y presentación del informe escrito de investigación, así como la presentación y defensa de la misma. En el Capítulo 0 (El proceso de la investigación educativa) el informe escrito fue presentado brevemente como cuarta fase del proceso general en una investigación. En el presente capítulo se desarrollará dicha fase y se abordarán cuestiones concretas en relación al desarrollo del Trabajo de Fin de Grado (TFG) y del Trabajo de Fin de Máster (TFM). A modo de guía se empleará la normativa vigente de la Universidad de Murcia que regula los TFG en la Facultad de Educación (Universidad de Murcia, 2014b), y la normativa del Máster Universitario en Formación del Profesorado de Educación Secundaria Obligatoria y Bachillerato, Formación Profesional, Enseñanzas de Idiomas y Enseñanzas Artísticas (Universidad de Murcia, 2014c) que presenta el TFM, si bien el estudiante deberá remitirse siempre a la normativa específica de su titulación. Dichas normativas establecen las directrices generales relacionadas con la definición, elaboración, orientación, presentación, defensa, evaluación y gestión administrativa de los Trabajos de Fin de Grado y de Máster de titulaciones concretas en la Facultad de Educación de la Universidad de Murcia.

Tanto el TFG como el TFM forman parte de los planes de estudio de los grados y másteres respectivamente y suponen la elaboración de una memoria final escrita que recoja los aspectos más relevantes del trabajo realizado. Dicha memoria será un trabajo personal y autónomo del estudiante, de elaboración individual, y que tendrá un carácter integrador (Universidad de Murcia, 2014a, b, c). Asimismo, este trabajo permite al estudiante mostrar la adquisición de las competencias asociadas a su titulación.

1. OBJETIVOS

Los objetivos que el presente capítulo pretende alcanzar son:

- Guiar al estudiante en la redacción del informe del Trabajo de Fin de Grado y Trabajo de Fin de Máster para que tenga en cuenta cuestiones sobre el contenido, formato, estructura, estilo y aspectos lingüísticos
- Informar sobre la normativa aceptada para citar fuentes bibliográficas correctamente y evitar así incurrir en plagio académico
- Señalar una serie de recomendaciones que aseguren una adecuada elaboración del informe o memoria

2. CONTENIDOS

La redacción del TFG y TFM debe tener en cuenta tres grandes aspectos: el formato, estructura, estilo académico, expresión escrita, y contenido. A continuación se detallará cada uno de estos puntos.

2.1. Formato

Las especificaciones de formato del TFG de la Facultad de Educación están detalladas en los manuales de Trabajo de Fin de Grado específicos de los diferentes títulos de grado, si bien para el presente capítulo se referenciará el Manual del Trabajo de Fin de Grado del Título de Educación Primaria (Universidad de Murcia, 2014a) a modo de guía (Manual de TFG, en adelante). Dicho manual está sujeto a modificaciones y deberá revisarse en el momento de la realización del trabajo. La versión actualmente en vigor (junio y julio 2014) establece que la extensión mínima de la memoria será 20 páginas y máximo 30 páginas sin incluir los anexos. El Manual también presenta las siguientes pautas que permitirán dar homogeneidad al trabajo y guiar al estudiante:

- Tipo de letra para el cuerpo del trabajo: Arial de 11 puntos o Times New Roman de 12 puntos
- Interlineado: 1 línea
- Márgenes: 2,5 cm por cada lado
- Formato de cita bibliográfica: Normativa APA
- Tablas y Figuras: Se mantendrá el tipo y tamaño de letra. Las figuras, gráficos, ilustraciones, tablas o cuadros que sean incorporados deben ir numerados y tener un título. Además, es preciso hacer referencia a los mismos en el cuerpo del texto. Véase apartado 2.2.2 para más detalle
- Se debe evitar dibujos y decoraciones innecesarias pues se trata de un trabajo académico
- Se debe revisar cuidadosamente la ortografía

Además, habrá que añadir los números de página, que podrán ir centrados en la parte inferior de la página.

La normativa de TFM (Universidad de Murcia, 2014c) establece que el informe impreso del mismo debe tener una extensión de entre 20 y 40 páginas que no incluirá los anexos pero sí el resto de apartados. El tamaño de letra debe ser de 12 puntos con un interlineado de 1,5. El TFM debe ser presentado en un fichero PDF a través de la aplicación TF de la Universidad de Murcia (<https://tf.um.es>) en los plazos previstos. Otras especificaciones de formato de TFM quedan a criterio del estudiante, pero una posible guía podría ser la presentada anteriormente para el TFG.

Para muchos estudiantes el TFG y TFM supone el primer gran trabajo académico al que se enfrentan. Esto significa que es posible que se desconozcan diversas cuestiones de formato. Por este motivo es aconsejable recurrir a videos o tutoriales de ayuda en formato papel u online que ilustren, por ejemplo, cómo generar índices o tablas de contenidos, cómo insertar números de página por secciones, cómo aplicar sangrías, etc.

2.2. Estructura

2.2.1. Estructura general del TFG y TFM

A continuación se presenta a modo de ejemplo la estructura que puede tener el TFG de la Facultad de Educación según la guía tomada como ejemplo en el Manual de TFG (Universidad de Murcia, 2014a). La memoria de TFG incluirá los siguientes apartados, si bien se insta a los estudiantes a referirse a la modalidad concreta de trabajo y línea que elijan a la hora de desarrollar los siguientes apartados:

- Portada (Figura 12.1)
- Índice paginado
- Resumen (150-200 palabras) en español y en otro idioma, preferentemente inglés o francés. El resumen es un sumario breve del contenido del trabajo que debe incluir la temática que se ha tratado, los instrumentos de recogida de información utilizados, etc.
- Palabras clave, tanto en español como en otro idioma, preferentemente el inglés o francés
- Introducción y justificación de por qué se ha elegido la temática objeto de estudio
- El desarrollo del trabajo
- Conclusiones, consecuencias e implicaciones
- Referencias. Estas se redactarán siguiendo fielmente los criterios establecidos en las normas de la APA (2010) (véase apartado 2.3.3)
- Anexos (si los hubiere)

		FACULTAD DE EDUCACIÓN
TRABAJO DE FIN DE GRADO GRADO EN EDUCACIÓN PRIMARIA		
TÍTULO DEL TFG EN ESPAÑOL		
NOMBRE Y APELLIDOS DEL/DE LA ALUMNO/A DNI		
LÍNEA DEL TFG		
CURSO ACADÉMICO 20... / 20... CONVOCATORIA DE		

Figura 12.1 Ejemplo de portada del TFG (Universidad de Murcia, 2014a).

En el caso del TFM, la normativa del Máster que se está empleando como ejemplo aquí (Universidad de Murcia, 2014c) propone un esquema con carácter orientativo que a continuación se detalla y amplía:

0. Páginas iniciales

- Portada (título del TFM, tutor(es), autor, lugar y fecha de presentación)
- Agradecimientos (opcional)
- Índice de contenidos con la numeración de la página
- Resumen de aproximadamente 200 palabras en español e inglés
- Cinco palabras clave en español e inglés extraídas del Tesauro de la UNESCO (véase <http://databases.unesco.org/thesp/>)

I. Marco teórico

Esta sección se puede denominar de diferentes maneras: *Marco teórico*, *Revisión teórica*, *Planteamiento teórico*, *Fundamentación teórica del problema de investigación*. En el Marco Teórico se identificarán y fundamentarán los problemas de investigación que darán paso al estudio empírico. Esta sección puede iniciarse con una introducción y a continuación se abordará el problema de investigación, incluyendo el origen y la justificación, los antecedentes y el estado actual de la investigación en el ámbito objeto de estudio.

II. Marco empírico

Este apartado también se puede denominar *Metodología*, o *Fases y diseño de la investigación*. Esta amplia sección contiene los siguientes subapartados:

1. Objetivos y/o hipótesis de la investigación
2. Diseño de investigación
3. Participantes (muestra) y contexto
4. Propuesta de intervención (si la investigación lo requiere)
5. Variables (descripción de las mismas si el diseño empleado lo requiere)
6. Recogida de información: instrumento(s), procedimiento
7. Plan de análisis de la información (estadísticos y/o procedimiento general de análisis de información cualitativa)

III. Presentación y discusión de los resultados

Esta fase puede presentarse en tres secciones separadas (*Resultados*; *Discusión*; *Conclusiones*) o bien solo dos (*Resultados y Discusión*; *Conclusiones e Implicaciones*). En cualquier caso, los contenidos serán los siguientes:

1. Análisis y descripción de los resultados
2. Discusión e interpretación de los resultados
3. Conclusiones
4. Implicaciones prácticas socioeducativas y de investigación. En este apartado tienen cabida las reflexiones e implicaciones personales con los hallazgos y la experiencia realizada

IV. Referencias

Las referencias bibliográficas consultadas para la realización de la investigación se redactarán siguiendo las normas establecidas por la American Psychological Association (APA, 2010) (véase apartado 2.3.3).

V. Anexos

En los anexos se introducirá toda información adicional que no sea fundamental para entender el texto. Esta sección comprende datos recabados durante la investigación, tablas de resultados, ficheros SPSS o similares, materiales e instrumentos utilizados en el trabajo empleados, etc. También en esta sección se incluirá aquella información que, aun siendo relevante, es demasiado extensa y dificulta la lectura del texto principal. En el caso de datos imprescindibles o de absoluta relevancia para entender el texto, es conveniente presentar estos en el apartado correspondiente. Los anexos se numerarán correlativamente con números arábi-

gos (Anexo 1, Anexo 2, Anexo 3), y se mencionarán en el texto para guiar al lector. La sección Anexos se ubicará al final del documento detrás de la bibliografía.

Es preciso recordar que las estructuras de trabajo aquí expuestas aquí deben servir como guía. El estudiante deberá remitirse al manual o la normativa vigente de su titulación concreta.

2.2.2. Tablas y figuras

Las tablas y figuras permiten presentar gran cantidad de datos e información de forma esquemática que facilita su comprensión. Aunque a veces sea difícil distinguir las tablas de las figuras, la principal diferencia entre estas es que las primeras muestran valores numéricos o información textual ordenada en filas y columnas (ej. datos estadísticos o descriptivos), mientras que las figuras muestran gráficos, imágenes, dibujos, fotografías, ilustraciones o cualquier otro tipo de representación (APA, 2010).

Según la normativa APA (2010) sobre la presentación de tablas y figuras en un informe, se deberá hacer referencia a las mismas en el texto siguiendo la numeración arábica (Tabla 1, Tabla 2, Tabla 3) y no con letras (incorrecto: Tabla 1a, Tabla 1b, Tabla 1c) en el orden en el que se mencionan en el texto. Además, se aplicará siempre el mismo formato y no se usarán colores. Cada tabla o figura debe ir acompañada de un título que se situará encima o debajo de la tabla según la normativa APA vigente en el momento de redacción del informe.

Finalmente, un informe bien redactado evitará repetir en el texto la información ya presentada en una tabla o figura, y hará alusión únicamente a los datos más relevantes o llamativos. Asimismo, se debe evitar introducir imágenes personales sin el consentimiento de las partes implicadas, y tampoco se debe olvidar referenciar las fuentes originales como por ejemplo libros o revistas científicas.

2.3. El estilo académico y la expresión escrita

El TFG, TFM y otros informes como las tesis doctorales son trabajos académicos que tienen un estilo muy particular. El estilo académico implica el uso de lenguaje preciso, objetivo, y gramaticalmente correcto que permita expresar las ideas con claridad. Esto significa que el autor debe ser riguroso en cuestiones tanto de contenido como de forma (expresión escrita).

El estilo académico se caracteriza por ser un estilo más formal que el empleado en otro tipo de textos o en el lenguaje hablado, pero debe ser claro y comprensible. El tono será impersonal y objetivo, y nunca coloquial o informal. El lenguaje que se emplee en la redacción deberá ser simple, con frases cortas y sin ambigüedad.

Asimismo, el autor del TFG/TFM debe demostrar dominio de la competencia de la expresión escrita. Esto se traduce en una redacción correcta, un texto estructurado y organizado, y una progresión lógica de ideas que permita al lector seguir los diferentes pasos del proyecto. Las frases deben estar correctamente construidas, y no serán demasiado largas ni complejas. A continuación se presentan algunas pautas que ayudarán a conseguir producir una redacción aceptable.

2.3.1. Ser objetivo

La escritura académica no debe dejar entrever la implicación personal del autor. Esto se consigue empleando un estilo objetivo e impersonal con el cual no se introduzcan pronombres personales en primera persona (yo, nosotros, me, nos) (McMillan & Weyers, 2010). Así, se evitarán frases como *“En este proyecto daré a los niños unas actividades para mejorar su comprensión lectora”* y se escribirá preferiblemente *“En este proyecto los estudiantes realizarán unas actividades para mejorar su comprensión lectora”*.

Otra forma de introducir objetividad en el texto puede ser sustituyendo los verbos por sustantivos relacionados (McMillan & Weyers, 2010). Así, la frase *“Evaluaré/evaluaremos a los niños al final del ejercicio”*, se podría sustituir por *“La evaluación de los estudiantes tendrá lugar al final del ejercicio”*, o bien *“Se evaluará a los estudiantes al término del ejercicio”*. Estas recomendaciones podrán tomarse de forma más flexible en la sección final de reflexiones e implicaciones personales, si bien el estilo debe ser siempre formal.

2.3.2. Uso adecuado del vocabulario, ortografía, gramática y puntuación

a) Vocabulario

Uno de los criterios de evaluación que tendrá en cuenta el tribunal de TFG y TFM hace referencia a aspectos formales del trabajo, concretamente la claridad de expresión, y la corrección lingüística y ortográfica (véase apartado 3 de este capítulo).

El TFG y TFM son productos de un proceso de lectura y maduración por parte del estudiante, así que se espera que este se familiarice con los distintos documentos consultados (revistas científicas, libros y manuales, informes, ensayos, etc.) y asimile el estilo de redacción propio de estos textos. El autor, por tanto, empleará variedad de vocabulario y procurará ser preciso en su expresión. Se debe evitar el lenguaje coloquial propio del lenguaje oral, y también las jergas y el lenguaje demasiado complejo. Una recomendación al estudiante en proceso de lectura es que, a medida que se vaya leyendo la documentación seleccionada tras la búsqueda bibliográfica, se reflexione sobre lo leído, se tomen apuntes, se planteen preguntas, y se comience a redactar el texto sin esperar a haber terminado de leer toda la literatura. Esto ayudará a ordenar las ideas y estructurar el texto, pero también a que no se olviden temas o puntos importantes, y sobre todo a que se pierda el miedo a escribir.

Finalmente, se recomienda evitar abreviaturas, y en caso de usar acrónimos (ej. UNESCO, OCDE), estos deben ser detallados la primera vez que se mencionen. Así, por ejemplo, se debe escribir *“Entre los objetivos de la Organización para la Cooperación y el Desarrollo Económicos (OCDE) se encuentra... [...]. Los representantes de los países miembros en OCDE tienen...”*.

b) Ortografía

Como futuro profesional en Educación, el estudiante de TFG/TFM debe prestar una especial atención a la ortografía, así que el diccionario de la Real Academia Española (RAE) y otros manuales serán herramientas de consulta fundamentales. La corrección ortográfica permitirá además evitar penalización por parte del tribunal evaluador del trabajo. Las consultas sobre reglas ortográficas son particularmente importantes, pues la RAE realiza actualizaciones y recomendaciones que conviene observar. Así, por ejemplo

“la palabra *solo*, tanto cuando es adverbio y equivale a solamente (*Solo llevaba un par de monedas en el bolsillo*) como cuando es adjetivo (*No me gusta estar solo*), así como los demostrativos *este*, *ese* y *aquel*, con sus femeninos y plurales, funcionen como pronombres (*Este es tonto*; *Quiero aquella*) o como determinantes (*aquellos tipos*, *la chica esa*), no deben llevar tilde según las reglas generales de acentuación, bien por tratarse de palabras bisílabas llanas terminadas en vocal o en -s, bien, en el caso de *aquel*, por ser aguda y acabar en consonante distinta de -n o -s” (RAE, n.d.).

c) Gramática

El TFG y el TFM deben también estar exentos de errores gramaticales y de palabras y expresiones no recogidas en el diccionario. Durante la redacción del informe pueden surgir dudas sobre el uso correcto de determinadas expresiones incorrectas muy comunes en el lenguaje oral, como por ejemplo “deber + infinitivo” y “deber de + infinitivo”. Así, “deber” seguido de infinitivo y sin preposición significa “tener obligación de”, mientras que “deber de” seguido de infinitivo significa “ser probable” o “suponerse” (Mesanza, 2009). Un ejemplo concreto del ámbito del trabajo académico que se está tratando en este capítulo sería “*La solución al problema debe ir acompañada del planteamiento*” (es obligado que la solución también contenga el planteamiento), y “*En este grupo debe de haber igual número de niñas y de niños*” (existe la probabilidad de que/es posible que en el grupo la proporción de niños y niñas sea similar).

La concordancia es otro aspecto que se debe respetar en la redacción de cualquier escrito académico. La concordancia es la “coincidencia obligada de determinados accidentes gramaticales (género, número y persona) entre distintos elementos variables de la oración” (RAE, 2005), y puede ser verbal y nominal. Para mantener la concordancia verbal, que es la que se establece entre el verbo y su sujeto (RAE, 2005), se dirá por ejemplo “*Jugaba en el patio un grupo de alumnos*”, en lugar de “*JugabaN en el patio el grupo de alumnos*”, pues el verbo debe concordar con el sujeto (un grupo) que es singular. En los casos en los que hubiera varios sustantivos de diferente género, domina el masculino (Mesanza, 2009), así por ejemplo “*Los ejercicios y actividades están planteados...*” (incorrecto: “*están planteadaS*”).

La concordancia nominal es la que “establece el sustantivo con el artículo o los adjetivos que lo acompañan” (RAE, 2005), y se observa en ejemplos como “*Al mediodía se hará una pausa de recreo*” en lugar de “*A LA mediodía se hará una pausa*”, que es incorrecto; otro ejemplo correcto sería “*patio y comedor escolares*” en lugar de “*patio y comedor escolar*” (incorrecto). Además, según la gramática española “se dice o escribe *el, un, algún y ningún* ante palabra que empiece por “a” tónica” (Mesanza, 2009, p. 66). Así por ejemplo, se dirá “*en un aula de educación primaria*” (incorrecto: en UNA aula) y también “*en esa aula hay cinco alumnos*” (incorrecto: en ESE aula).

Asimismo, se debe evitar los errores lingüísticos denominados dequeísmo y queísmo, que consisten en añadir y omitir, respectivamente, elementos de enlace (Mesanza, 2009). Así, por ejemplo, se evitará el dequeísmo en “*Los alumnos piensan DE que la actividad es entretenida*” (incorrecto), y se escribirá “*Los alumnos piensan que la actividad es entretenida...*” (correcto). Se evitará el queísmo en frases como “*A pesar que se hayan programado actividades de ocio*” (incorrecto) y se escribirá “*A pesar DE que se hayan programado actividades*” (correcto). Otro ejemplo de queísmo que se debe evitar: “*Los profesores de primaria están seguros que el programa será un éxito*” (incorrecto), y se escribirá “*Los profesores de primaria están seguros DE que el programa será un éxito*” (correcto). La consulta de las fuentes oficiales de gramática española ayudará a evitar fallos comunes.

d) Puntuación

Finalmente, se deberá observar las reglas de puntuación del español (uso de la coma, dos puntos, y punto y coma, etc.), y también en este caso es recomendable seguir las pautas establecidas por la RAE. Los signos de puntuación sirven para marcar las pausas y la entonación con que deben leerse los enunciados, organizar el discurso y sus diferentes elementos para facilitar su comprensión, evitar posibles ambigüedades en textos que, sin su empleo, podrían tener interpretaciones diferentes, y señalar el carácter especial de determinados fragmentos de texto - citas, incisos, intervenciones de distintos interlocutores en un diálogo, etc. (RAE, 2005).

Un error muy extendido es la introducción de coma entre sujeto y predicado. La RAE (2005) apunta explícitamente que “es incorrecto escribir coma entre el sujeto y el verbo de una oración, incluso cuando el sujeto está compuesto de varios elementos separados por comas”. Así, no se debe escribir “*Los alumnos, deberán realizar juegos*”, sino “*Los niños deberán realizar juegos*” (correcto); tampoco se escribirá “*Los padres, alumnos y profesores del centro, tienen la obligación de respetar las normas*”, sino “*Los padres, alumnos y profesores del centro tienen la obligación de respetar las normas*” (correcto). Tampoco en el caso de un sujeto extenso se debe introducir coma entre este y el verbo (incorrecto: “*Los alumnos de segundo de la E.S.O que no hayan aprobado el examen, tendrán que realizar un trabajo práctico*” – no se coloca coma antes del verbo “tendrán”).

2.3.3. Citación y referencias

El Manual de TFG (Universidad de Murcia, 2014a) presenta como anexo un resumen de la normativa de citación de la American Psychological Association (APA) que servirá como guía para citar correctamente las fuentes consultadas durante la realización del trabajo. Es importante recordar que las normas de la APA son revisadas y actualizadas regularmente, así que será imprescindible consultar la guía docente o manual del TFG/TFM para conocer la edición vigente en el momento de redacción y presentación de la memoria, y así evitar errores de citación.

Las normas de la APA hacen alusión a la forma de citar en el texto, así como al listado de referencias bibliográficas que aparecen al final del informe. A continuación se muestra un par de normas de citación a modo de ejemplo (Universidad de Murcia, 2014a, p. 13):

a) Citar en el texto según las normas de la APA

- Citas textuales: si el texto que se quiere citar tiene menos de 40 palabras se indica autor, año, página y texto entre comillas.

Ejemplo:

El Manual de la APA se publicó por primera vez como un “standard of procedure, to which exceptions would doubtless be necessary, but to which reference might be made in cases of doubt” (Bentley et al., 1929, p. 57).

- Citas no textuales: se indica el autor y el año; el texto no va entre comillas. En las referencias a un trabajo de dos autores se citan ambos siempre (ej. Prosser y Trigwell, 2006). En las referencias a un trabajo de más de dos autores y menos de seis, la primera vez se citan todos. Las demás veces se cita solo el apellido del primer autor seguido de la expresión “et al.” (del latín “et alii”, “y otros”) y del año.

Ejemplo:

Según Stes, De Maeyer, y Van Petegem (2010) los enfoques de enseñanza son [...]. En su análisis Stes et al. (2010) identifican [...].

b) Referencias bibliográficas según las normas de la APA

Para citar las fuentes consultadas, ya sea un libro completo, un capítulo de libro, un artículo de revista científica, una página web, etc., e indicarlas correctamente en el apartado Bibliografía o Referencias, habrá que consultar minuciosamente la normativa APA, pues la forma de referenciar varía según el caso. Se remite al estudiante al Manual de TFG (2014a) o al Manual de la APA (2010) para consultar cada caso en detalle.

2.3.4. Redacción de la memoria en inglés

La normativa de TFG y de TFM de la Facultad de Educación (Universidad de Murcia, 2014b, c) establecen que la memoria puede ser redactada y defendida en idioma español, inglés o francés (según sea el caso). En este caso se debe tener en cuenta las siguientes recomendaciones:

- El nivel de conocimiento del idioma del estudiante debe ser suficientemente avanzado para permitirle producir un trabajo de calidad y defenderlo con fluidez y soltura.
- Se recurrirá a diccionarios para consultar la ortografía de palabras que se desconozcan o en caso de duda. Asimismo, se observarán las normas de puntuación propias del idioma y que en algunos casos pueden ser diferentes de las que rigen el español.
- En los trabajos redactados en idioma inglés, los estudiantes deben seguir las siguientes indicaciones (adaptado de McMillan & Weyers, 2010):
 - o Se evitarán las contracciones de sujeto y verbo como en *"They'll work together in pairs"* (incorrecto) y se preferirá *"They will work together in pairs"*.
 - o Tampoco se empleará la primera persona (I, me, we, us). En su lugar, se recurrirá a la voz pasiva como estrategia para introducir un estilo impersonal (*"Students will be given a text..."*, en lugar de *"I will give students a text..."*).
 - o Será necesario decantarse por un tipo de ortografía concreta, bien inglés americano (AE), bien inglés británico (BE), pues muchas palabras varían. Ejemplos: *colour* (BE) – *color* (AE); *got* (BE) – *gotten* (AE); *travelled* (BE) – *traveled* (AE); *centre* (BE) – *center* (AE). Diccionarios de prestigio como Oxford Dictionaries, o Cambridge Dictionaries señalan la ortografía correcta en cada caso.
 - o Se evitarán expresiones que aludan a un sexo concreto, así que en lugar de *policeman/policewoman*, se preferirá *police officer*. En otros casos se puede recurrir al plural, así que una frase como *"He/she will read three texts"* podría transformarse en *"Students will read three texts"*.

2.4. Contenido

Los contenidos del informe corresponden a las secciones de la estructura general enumeradas en el Manual de TFG (Universidad de Murcia, 2014a) y en la normativa de TFM (Universidad de Murcia, 2014c) (véase apartado 2.2.1). Puesto que el trabajo se desarrollará bajo la supervisión de un orientador (TFG) o tutor (TFM), se contará en todo momento con su asesoramiento sobre cómo enfocar el trabajo y qué pautas se deben seguir.

Tanto el TFG como el TFM tienen diferentes modalidades. En el caso del TFG en la Facultad de Educación, el trabajo estará enmarcado en alguna de las siguientes líneas vinculadas a los distintos ámbitos que forman parte de los planes de Prácticas Externas o de las Prácticas Escolares, según la titulación (Universidad de Murcia, 2014b).

- a) Para el Grado de Educación Social:
 - 1. Actividades de observación y conocimiento de la realidad
 - 2. Actividades de actuación y colaboración con el Tutor/a de la Institución
 - 3. Actividades puntuales con autonomía y protagonismo del estudiante
- b) Para el Grado de Pedagogía:
 - 1. Actividades de observación y conocimiento de la realidad
 - 2. Actividades de actuación y colaboración con el Tutor/a de la Institución
 - 3. Actividades puntuales con autonomía y protagonismo del estudiante
 - 4. Estudios y análisis de temas pedagógicos
- c) Para los Grados de Educación Primaria y Educación Infantil:
 - 1. Actividades de conocimiento del centro escolar y su contexto
 - 2. Actividades de conocimiento de la organización y gestión del aula
 - 3. Actividades de conocimiento de los procesos de enseñanza aprendizaje y actuaciones puntuales
 - 4. Actividades de planificación y puesta en práctica de la acción docente

En el caso del TFM, y más concretamente el Máster en Formación del Profesorado en Educación Secundaria (Universidad de Murcia, 2014c), el trabajo se integrará en alguna de las siguientes modalidades:

- a) Trabajos de investigación de naturaleza básica o aplicada relacionados con la enseñanza
- b) Revisiones bibliográficas que conduzcan a un análisis crítico sobre aspectos disciplinares relevantes sobre la práctica docente
- c) Diseño reflexivo y crítico de actividades de innovación vinculadas a la materia Prácticas de Enseñanza (práctica docente)

2.4.1. El plagio

El plagio, esto es “la acción y efecto de copiar en lo sustancial obras ajenas, dándolas como propias” (Real Academia de la Lengua, n.d.), ha estado presente desde el inicio de la palabra escrita. En el contexto académico el plagio se considera una falta grave, sin embargo, la expansión de Internet ha traído consigo un incremento en el uso de *copiar-pegar* de fuentes online al alcance de todos. El elevado número de trabajos demandados a los estudiantes, así como la extensión de los mismos, motivan a muchos estudiantes a copiar de otros trabajos o fuentes

electrónicas sin referenciar estos adecuadamente. En ocasiones se puede incurrir en plagio por desconocimiento u olvido de citar, si bien esto no exime de la falta. Asimismo, la modificación puntual de términos, la sustitución de ciertas palabras por sinónimos (paráfrasis), o invertir el orden en la frase sin la correspondiente cita de la fuente de origen también se considera plagio (McMillan & Weyers, 2010), así que se deberá poner todos los medios para evitar cometer fraude académico.

A menudo los estudiantes confunden dos actividades bien diferenciadas: por un lado, fundamentar correctamente el trabajo de investigación, y por otro, realizar aportaciones novedosas “de cosecha propia”. Así, los estudiantes piensan que para justificar su trabajo pueden apropiarse de las ideas de otros autores sin citarlos. En un Trabajo de Fin de Grado o Máster se espera que el estudiante desarrolle ambas acciones, es decir, el estudiante debe sustentar su investigación con literatura y argumentos científicamente aceptados que reflejen la importancia de alcanzar los objetivos fijados en el trabajo; esto será desarrollado en el Marco Teórico. Además, el estudiante debe desarrollar su proyecto de forma reflexiva e individual en función de la justificación alcanzada anteriormente; esto quedará reflejado en el Marco Empírico mediante el diseño de la investigación, el enfoque adoptado, la propuesta de intervención (si la hubiere), el análisis, descripción, discusión e interpretación de los resultados, y las implicaciones prácticas socioeducativas y de investigación. Será en esta última sección donde el estudiante podrá reflexionar sobre los hallazgos, conclusiones, y su experiencia personal.

Muchas universidades emplean hoy día motores de búsqueda o adquieren programas de software específicos para detectar la copia literal no referenciada en los trabajos de sus estudiantes. En un intento por combatir el plagio y la deshonestidad académica, la página web de la Biblioteca de la Universidad de Murcia ofrece tutoriales que alertan sobre esta plaga:

<http://www.um.es/web/biblioteca/contenido/formacion/videos>

Finalmente, todo autor o estudiante de TFG/TFM debe observar las siguientes indicaciones para evitar incurrir en plagio (McMillan & Weyers, 2010):

- Evitar copiar información de fuentes online, a menos que se cite la fuente
- Anotar las fuentes de origen completas (incluidas las páginas) cuando se esté tomando apuntes de las lecturas
- Seguir fielmente las normas establecidas por la institución en lo referente a citar fuentes bibliográficas
- Evitar emplear citas en exceso
- Comprobar que las propias ideas son realmente novedosas. Es posible que muchas de las ideas que se tengan provengan de fuentes que se leyeron antaño y que hoy se han olvidado. El autor deberá asegurarse de que se consiguen las fuentes de referencia

3. EVALUACIÓN DEL TRABAJO DE FIN DE GRADO Y FIN DE MÁSTER

El Manual de TFG (Universidad de Murcia, 2014a) presenta los siguientes criterios que el tribunal empleará para evaluar y calificar los TFG:

Aspectos formales:

1. Claridad de expresión, corrección lingüística y ortográfica
2. Buena estructuración de las partes atendiendo a la metodología planteada
3. Fuentes de información pertinentes y correctamente citada según la APA

Contenidos:

4. Capacidad para definir y acotar la temática, problema o ámbito de estudio
5. Conocimiento de la materia
6. Interdisciplinariedad del contenido
7. Originalidad de las ideas y aportaciones personales
8. Corrección metodológica
9. Capacidad de análisis y síntesis
10. Adecuación de la interpretación de resultados y las conclusiones
11. Capacidad de reflexión, de crítica y autocrítica

Aspectos transversales:

12. Aplicación práctica del trabajo al ámbito educativo y social
13. Uso de medios y herramientas apropiadas
14. Inclusión pertinente de materiales ilustrativos

Los estudiantes que hubieran conseguido una calificación igual o mayor a 9 (nueve) podrían voluntariamente exponer y defender su TFG ante el Tribunal Evaluador a fin de poder optar a la mención de Matrícula de Honor. Si se hubiese presentado el TFG en francés o inglés se deberá realizar la exposición y defensa del mismo en el mismo idioma. Véase el Capítulo 13 para más detalles.

Por su parte, el tribunal de Trabajo de Fin de Máster en la Facultad de Educación podrá tener en cuenta los siguientes aspectos en su valoración del trabajo (Universidad de Murcia, 2014c):

Memoria (máximo 7 puntos)

Valoración de los apartados de la memoria:

1. Planteamiento del problema y marco teórico
2. Diseño metodológico (de acuerdo con la modalidad de TFM)
3. Análisis de datos y resultados
4. Conclusiones, consecuencias e implicaciones

Valoración de la calidad de la redacción

Informe del tutor: interés, dedicación, implicación del estudiante y cumplimiento de plazos en la realización del TFM

Exposición (2 puntos máximo) (véase Capítulo 13)

Valoración de la calidad de los recursos y soportes utilizados

Valoración de la calidad expositiva

Defensa del trabajo (1 punto máximo) (véase Capítulo 13)

Valoración de la comprensión de las preguntas y de la capacidad de respuesta

Esta guía es una herramienta muy útil para el estudiante, ya que le ayudará a organizar a nivel temporal las diferentes fases de la investigación y elaboración del informe, así como a distribuir sus esfuerzos y evitar divagar.

Finalmente, se propone una serie de cuestiones que pueden servir como autoevaluación que conviene realizar antes de presentar, entregar o difundir un trabajo de investigación (adaptado de Hernández Pina, Maquilón Sánchez, y Cuesta Sáez de Tejada, 2008, p. 20). Dicha autoevaluación ayudará a valorar diferentes aspectos para poder decidir si el informe está correctamente realizado y puede ser difundido y/o defendido.

Portada: ¿Incluye todos los datos formales necesarios y/o prescriptivos?

Título: ¿Revela con claridad el contenido del trabajo?

Introducción:

- Revisión de la literatura:
 - ¿Es toda ella relevante?
 - ¿Cubre hasta el momento actual?
 - ¿Se incluye el marco teórico, con sus apartados y subapartados?
- Objetivos y planteamiento del problema: ¿Los objetivos están formulados adecuadamente?

Metodología:

¿Es la muestra representativa?

¿Se señalan los procedimientos y método de muestreo?

¿Están descritas las variables de un modo correcto siguiendo un adecuado criterio?

¿Se identifica el diseño empleado, experimental, cuasi-experimental o pre-experimental, observación, investigación evaluativa, estudio de casos, etc.?

¿Permite la descripción de materiales y/o instrumentos utilizados hacer una réplica exacta del trabajo?

¿Se describe el procedimiento seguido en la investigación?

Análisis de los datos y resultados:

¿Las técnicas estadísticas utilizadas son las más apropiadas en relación a los objetivos?

¿Aparecen los resultados bien descritos y claramente presentados por cada objetivo?

¿Son los hallazgos congruentes con las expectativas?

Discusión y conclusiones:

¿Se hace un resumen claro de lo realizado?

¿Se señalan aportaciones y deficiencias?

¿Tienen los resultados obtenidos implicaciones para futuras investigaciones en ese campo?

Bibliografía:

¿Se recogen todas las referencias que aparecen en el texto?

¿Están citadas adecuadamente, por ejemplo, siguiendo la normativa APA?

Apéndices: ¿Incluyen información relevante y necesaria para la comprensión y valoración del trabajo de investigación?

4. RECOMENDACIONES A TENER EN CUENTA

Las siguientes pautas y recomendaciones pretenden ayudar al estudiante a organizar el tiempo disponible y sacar el máximo provecho a sus esfuerzos:

- Asegúrate de que entiendes perfectamente qué es lo que tienes que hacer y cómo lo tienes que hacer. Antes de comenzar a escribir, estudia minuciosamente toda la documentación que se te ha facilitado y que está disponible en la página web de la universidad. En caso de dudas concretas, puedes consultar a tu orientador/tutor sobre la estructura, formato, objetivos u organización del trabajo.
- Intenta comprender perfectamente cuáles son los objetivos que plantea tu trabajo para que el resto de secciones sean congruentes con los mismos. Esto te permitirá producir un trabajo coherente.
- No te acomodes ante la idea de que la entrega de la memoria será dentro de varios meses. El tiempo pasa muy rápido y es importante organizarse correctamente para evitar estrés a medida que se aproxima la fecha final. A menudo emplearemos más tiempo que el planificado para la realización de determinados apartados del trabajo.
- En la organización del tiempo es importante recordar que solo se podrá defender el TFG/TFM cuando se hayan superado todas las asignaturas del plan de estudios. Además, se debe contar con la validación o autorización del orientador/tutor y con el informe correspondiente para que el trabajo pueda ser defendido (Universidad de Murcia, 2014b, c).

- Procura trabajar de forma eficaz y eficiente. Para ello deberías (McMillan & Weyers, 2010):
 - o planificar tu rutina de trabajo, ponerte manos a la obra lo antes posible y evitar postergar el inicio de las tareas;
 - o dar prioridad a las tareas más importantes y evitar las distracciones;
 - o mantener tus materiales, borradores, y lugar de estudio bien organizado; y
 - o realizar descansos que te ayudarán a reflexionar y asentar ideas.
- Gestiona tu tiempo para conseguir la máxima eficiencia. Para ello será muy útil elaborar un cronograma (Figura 12.2) que contemple las diferentes fases del proceso de investigación que fueron presentadas en el capítulo inicial de este libro. Un cronograma es un calendario de trabajo en el que anotarás qué tareas vas a realizar, cuándo las vas a realizar, y las fechas clave a recordar (cita con tu orientador/tutor, fecha estimada de impresión, fecha de entrega del trabajo). Tu calendario puede tener el formato de diario u horario (similar al empleado en la escuela) y el estudiante será el que decida el nivel de detalle que incluirá y que mejor se ajusta a su forma particular de trabajar. Para este fin puedes reciclar una vieja agenda o bien crear tu propio formato en una simple libreta.

Tareas (cualquier tarea que consideres importante)	octubre	noviembre	diciembre	enero	febrero	marzo	abril	mayo	junio	julio
Propuesta de trabajo										
Búsqueda de información y fuentes bibliográficas										
Lectura										
Recogida de datos										
Redacción del borrador del informe (TFG/TFM)										
Corrección del mismo (tutor)										
Inclusión de cambios y edición										
Corrección final								x	x	
Impresión									x	
Ensayos de la exposición									x	
Fecha entrega									x	

Figura 12.2. Cronograma de trabajo (adaptado de Kumar, 1999, p. 185).

- Será de gran ayuda crear una plantilla del informe en la que se introduzcan las secciones de que está compuesto el TFG/TFM siguiendo la estructura presentada anteriormente (apartado 2.2). Esto ayudará a organizar las ideas y visualizar en qué parte del proceso nos encontramos.
- Comienza a redactar tu informe lo antes posible. Con la ayuda de procesadores de textos es fácil manipular el texto y mover párrafos de una sección a otra si se observa que una idea encaja mejor en otra sección.

- Desarrolla una redacción con un estilo propio y académico para plasmar las ideas de otros haciendo siempre alusión a las fuentes consultadas. Recuerda en todo momento la gravedad de incurrir en plagio, aun si es inconsciente, y ten presente las recomendaciones presentadas en este capítulo que te ayudarán a evitarlo.
- No excedas el límite de palabras o páginas fijado por la normativa vigente de tu institución o facultad. Si sobrepasaras la extensión establecida, considera mover algún contenido a los anexos, siempre y cuando sea secundario y no entorpezca la lectura fluida del informe.
- Crea algún tipo de registro de datos (Figura 12.3), como por ejemplo una hoja de cálculo, que te permita anotar las referencias y tomar apuntes breves de las lecturas que vas realizando según desarrollas el marco teórico.

AÑO	AUTOR	TÍTULO Y/O REFERENCIA	PDF	PAPEL	NOTAS	PALABRAS CLAVE	RELACIONA DO CON...	IDEA INTRODUCIDA EN INFORME
1970	Biggs, J. B.	Faculty patterns in study behavior. Australian Journal of Psychology, 22, 161-174.		x	Study of behaviour in arts and science Results: arts SS need to develop strategies appropriate for organising unrelated materials	CONTEXT-DEPEND	Ramsden, 1979:423	S
1977	Entwistle	Strategies of Learning and studying: recent research findings		x	Describes Marton's DA and SA, and Pask's holist and serialist briefly. General account, but too brief	APPROACHES TO LEARNING		N
1994	Prosser, M., Trigwell, K., & Taylor, P.	A phenomenographic study of academics' conceptions of science learning and teaching. Learning and Instruction, 4, 217-231.	x		IMPORTANT!!!!!!!!!!!!	CONCEPTIONS OF LEARNING APPROACHES TO TEACHING		N
1981	Saljo, R.	Learning approaches and outcomes: some empirical observations. Instructional Science, 10, 47-65.		x	Showed relationship between approach to learning and outcome in everyday studies and experimental situations		Gibbs, 1982	S

Figura 12.3. Ejemplo de tabla de registro de lecturas consultadas.

- La revisión y edición del informe escrito forma parte del proceso de redacción. Es una fase crucial y no debe considerarse tiempo perdido. Antes de enviar el documento al orientador/tutor es imprescindible leerlo y releerlo varias veces. Incluso puede resultar útil leerlo en voz alta o que otras personas nos den su opinión. Esta tarea nos permitirá detectar faltas de ortografía, errores, gazapos, así como identificar frases ininteligibles e incoherencias en la argumentación. Si entregamos al orientador/tutor un texto ya repasado, la comprensión del texto será más fácil y se agilizará el tiempo destinado a la corrección.
- El orientador/tutor necesitará tiempo para leer y aportar feedback y recomendaciones sobre cómo mejorar tu trabajo, que luego deberás introducir en el borrador del trabajo. Ten en cuenta esta fase fundamental del proceso cuando planifiques tu cronograma.
- La impresión y encuadernación de tu trabajo será realizada por terceras personas, así que recuerda disponer de tiempo suficiente antes de la fecha de entrega para que este paso pueda realizarse correctamente.

- Planifica con mucho detalle la redacción, lectura, edición y corrección del TFG/TFM. No subestimes el tiempo que requieren todas estas fases del proceso de elaboración del informe de investigación.

REFERENCIAS

- American Psychological Association (2010). *Publication Manual of the American Psychological Association* (6th edition). Washington, DC: American Psychological Association.
- Hernández Pina, F., Maquilón Sánchez, J. J., y Cuesta Sáez de Tejada, J. D. (2008). *El proceso de investigación y el análisis de datos en ciencias sociales*. Murcia: DM.
- Kumar, R. (1999). *Research Methodology. A step-by-step guide for beginners*. London: Sage.
- McMillan, F., & Weyers, J. (2010). *How to write dissertations and project reports*. Essex: Pearson.
- Mesanza, J. (2009). *Hablar y escribir correctamente. Barbarismos, impropiedades y dudas en el español oral y escrito*. Madrid: Wolters Kluwer España.
- Real Academia Española (n.d.). *El adverbio solo y los pronombres demostrativos, sin tilde*. Recuperado de <http://www.rae.es/consultas/el-adverbio-solo-y-los-pronombres-demostrativos-sin-tilde#sthash.LLuJ1a0O.Y98OyXo0.dpuf>
- Real Academia Española y Asociación de Academias de la Lengua Española (2005). *Diccionario panhispánico de dudas*. Madrid: Santillana.
- Universidad de Murcia (2014a). *Manual del Trabajo de Fin de Grado (TFG), Título de Educación Primaria, Convocatoria de junio y junio 2014*. Recuperado de <http://www.um.es/documents/299436/1854327/Manual+de+TFG+EDUCACIO%CC%81N+PRIMARIA+2013-2014.pdf/b7b24b34-8584-42ef-93a3-c5b5f29af60b>
- Universidad de Murcia (2014b). *Normativa por la que se regulan los Trabajos de Fin de Grado (TFG) en la Facultad de Educación* (aprobada en Junta de Facultad el 10 de febrero de 2014). Recuperado de http://www.um.es/documents/299436/319886/NORMATIVA+TFG+APROBADA+JUNTA+CENTRO+10.2.14_definitiva.pdf/d3893b4b-a3aa-46f1-8f7a-63af52529fb5
- Universidad de Murcia (2014c). *Normativa de los Trabajos de Fin de Máster. Instrucciones sobre el procedimiento a seguir para la presentación, defensa y evaluación de los Trabajos de Fin de Máster*. Máster Universitario en Formación del Profesorado de Educación Secundaria Obligatoria y Bachillerato, Formación Profesional, Enseñanzas de Idiomas y Enseñanzas Artísticas. Recuperado de <http://www.um.es/documents/299436/1951373/NORMATIVA+TFM+2014-15DEF.pdf/e1f70b82-7210-4481-a49e-8797daa052c8>

Capítulo 13.

PRESENTACIÓN Y DEFENSA ORAL DE LA INVESTIGACIÓN

Fuensanta Monroy Hernández
Fuensanta Hernández Pina

INTRODUCCIÓN

El presente capítulo orienta sobre la presentación y defensa oral de la investigación llevada a cabo en el Trabajo de Fin de Grado (TFG) y Trabajo de Fin de Máster (TFM). Según la normativa de TFG (Universidad de Murcia, 2014b) y el Manual del Trabajo de Fin de Grado del Título de Educación Primaria (Universidad de Murcia, 2014a) tomado en este capítulo a modo de ejemplo de guía (Manual de TFG, en adelante), el TFG comprende la elaboración y redacción de una memoria final escrita (véase Capítulo 12) que podría ser expuesta oralmente ante tribunal. Esta situación se produciría voluntaria y únicamente en el caso de los TFG que hubieran conseguido una calificación igual o mayor a 9, y que desearan optar a matrícula de honor.

En el caso del Trabajo de Fin de Máster, la normativa vigente tomada como ejemplo en este capítulo corresponde al Máster Universitario en Formación del Profesorado de Educación Secundaria Obligatoria y Bachillerato, Formación Profesional, Enseñanzas de Idiomas y Enseñanzas Artísticas (Universidad de Murcia, 2014c), y establece que el TFM será presentado y defendido en acto público ante un tribunal, empleándose para ello el recurso o apoyo visual que el candidato estime oportuno.

1. OBJETIVOS

Los objetivos que el presente capítulo pretende alcanzar son:

- Guiar al estudiante en la preparación de la presentación y defensa oral de su Trabajo de Fin de Grado (TFG) y Trabajo de Fin de Máster (TFM) para que

tenga en cuenta cuestiones sobre la organización del tiempo, los recursos disponibles, y el protocolo y las habilidades orales y expositivas.

- Detallar la elaboración de una presentación visual con diapositivas o de un póster como herramientas de apoyo durante la presentación oral.
- Presentar una serie de recomendaciones que aseguren una correcta exposición oral de la memoria.

2. CONTENIDOS

La exposición oral del TFG (únicamente en los casos de los estudiantes que desearan optar a la matrícula de honor) y del TFM consiste en una defensa pública en la que el autor resaltarán las principales aportaciones realizadas en la memoria. Para preparar correctamente dicha defensa es imprescindible haber acabado íntegramente la memoria escrita. Únicamente entonces se podrá saber qué aspectos son relevantes y deben ser transmitidos en la exposición oral, pues no se puede dar detalle de todo lo que se ha escrito. Será, por tanto, necesario desarrollar la capacidad de síntesis y priorizar qué información merece la pena desarrollar y qué datos son menos relevantes.

Una vez se ha terminado de redactar la memoria, se puede comenzar a planificar la presentación oral. Para ello hay que tener en cuenta varias cuestiones:

- ¿Qué duración máxima tendrá la presentación oral? ¿Y el acto de defensa en su conjunto?
- ¿Qué ocurrirá durante la presentación y defensa? ¿Me hará preguntas el tribunal?
- ¿Se puede emplear recursos de apoyo durante la exposición? Si es así ¿cuáles puedo emplear? ¿Qué formato han de tener?
- ¿Puedo llevar un guión que me ayude durante la presentación?
- ¿Cómo debo presentar mi trabajo? ¿Qué vestuario es el adecuado?

A continuación se aborda cada una de estas cuestiones que el estudiante deberá tener en cuenta antes y durante la presentación y defensa oral del TFG y TFM.

2.1. Distribución del tiempo

Para preparar correctamente la presentación y defensa será imprescindible saber de antemano cuánto tiempo tendremos asignado para exponer, pues no es lo mismo planificar una exposición de 3 minutos que de 45 minutos. Se debe diferenciar entre la duración de la presentación del estudiante (esto es, el tiempo que dura la exposición de la memoria), y la duración total del acto de defensa, que incluye además la intervención del tutor/cotutores y del tribunal.

Los estudiantes de TFG que deseen exponer su trabajo para optar a matrícula de honor disponen de un tiempo aproximado de veinte minutos. Tras la exposición por parte del estudiante, los miembros del Tribunal pueden realizar las preguntas

que estimen oportunas o debatir sobre el contenido de la memoria. El tiempo total del acto no debe superar los 30 minutos (Universidad de Murcia, 2014b).

Según la normativa del TFM (2014c) el tiempo estipulado para la presentación de la investigación no debe exceder los veinte minutos. Seguidamente el tutor o cotutores pueden realizar una breve intervención, tras la cual los miembros del tribunal tendrán la palabra durante un tiempo máximo de diez minutos para realizar preguntas y observaciones. Tras dicha intervención, el estudiante responderá a las preguntas. Por tanto, el tiempo que el estudiante debe destinar a su presentación oral es 20 minutos como máximo, y debe ajustar los contenidos de la misma para no exceder ese límite.

La distribución del tiempo también alude a la planificación que el estudiante debe hacer para poder preparar su presentación correctamente y sin prisas. Así, este punto también deberá ser tenido en cuenta en el cronograma presentado en el capítulo anterior. Si no se posee experiencia en la realización de presentaciones orales, conviene disponer de al menos dos semanas para estructurar los contenidos en los diferentes apartados y ensayar varias veces para conseguir realizar la presentación en el tiempo estimado.

2.2. Proceso de presentación y defensa

Para prepararse eficazmente para la presentación es recomendable tener presente los criterios de evaluación que observará el tribunal. Esto ayuda al candidato a centrar sus esfuerzos en la preparación. Los criterios de evaluación del TFG y TFM están enumerados en el apartado 3 de este capítulo.

Al inicio del proceso de presentación y defensa el tribunal explicará las características de la misma (protocolo y tiempo disponible). A continuación cederá la palabra al estudiante para que comience su exposición. Tras la misma, el presidente del tribunal dará la palabra a los miembros para que planteen preguntas o hagan comentarios y reflexiones. Se dará al estudiante la posibilidad de responder a las cuestiones que estime oportuno, y a continuación se dará por finalizado el acto.

La presentación deberá contener los principales apartados de la memoria y destacar los aspectos más relevantes. El estudiante deberá encontrar el equilibrio entre los distintos apartados para lograr realizar su exposición en el tiempo asignado. Se debe recordar que el tribunal habrá leído la memoria, por lo que no será necesario explicar en detalle los pormenores del proyecto. Si quedara algún punto por aclarar, esto podría ser explicado al tribunal durante la ronda de preguntas tras la exposición.

Durante la presentación se espera que el estudiante posea una fluidez y claridad expositiva propios de un profesional de la educación. Es, por tanto, imprescindible ensayar las veces que sea necesario para lograr exponer el trabajo correctamente en el margen de tiempo previsto y sin tener que acelerar el discurso.

2.3. Formato de la presentación

Lo primero que se debe averiguar es si está permitido el uso de recursos o herramientas de apoyo durante la exposición oral. Normalmente se puede emplear los medios audiovisuales disponibles en el espacio asignado para dicho acto. Uno de las herramientas más comunes en las exposiciones orales es una presentación con texto esquematizado en diapositivas en el que se puede insertar animaciones de texto e imágenes (por ejemplo Microsoft PowerPoint, Prezi, etc.), si bien el estudiante puede emplear otros recursos como mapas conceptuales, maquetas, pósters, o audiciones y videos (si fueran relevantes).

Muchos estudiantes se preguntan si realmente es necesario hacer una presentación con diapositivas. Si bien no es obligatorio, sí es muy recomendable, ya que sirve como guía de la exposición al estudiante y al tribunal. Cualquier apoyo visual ayudará a concretar la estructura y los contenidos que se están exponiendo.

2.3.1. Recursos y herramientas de apoyo

Al igual que en el caso de la memoria impresa, las herramientas empleadas (presentación visual, póster, etc.) no debe contener errores ortográficos. La letra debe ser de un tamaño y color apropiado y legible. A continuación se presenta cada uno de estos recursos.

a) La presentación con diapositivas y animación

Muchos estudiantes se preguntan con preocupación cuántas diapositivas deben incluir en su presentación. No hay respuesta a esta pregunta, y la única forma de saberlo es ensayar cronometrando el tiempo empleado en presentar cada diapositiva. Por lo general, se suele recomendar una diapositiva (o pantalla) por minuto de exposición, así que la presentación no debería tener más de 20-22 diapositivas, si bien esto es solo una orientación.

Cada estudiante debe personalizar su presentación y decidir qué contenidos y estructura ha de tener. Una posible estructura podría ser la siguiente:

- Una diapositiva para el título del proyecto y el nombre del autor
- Una o varias diapositivas con los aspectos más relevantes del Marco Teórico
- Varias diapositivas para el desarrollo del Marco Empírico. Será imprescindible presentar claramente los objetivos del trabajo, así como otros apartados (diseño, participantes, etc.)
- Varias diapositivas para las conclusiones, discusión de resultados, e implicaciones
- No será necesario incluir el resumen, palabras clave, ni las referencias bibliográficas
- Una diapositiva final de agradecimientos o que indique el final de la presentación

Se debe evitar copiar y pegar frases literalmente de la memoria a las diapositivas, ya que lo que se muestra en la presentación debe ser la información más relevante presentada de forma sintetizada. Una presentación con un exceso de texto puede ser contraproducente y dificultar (en lugar de facilitar) la exposición oral, pues el estudiante tenderá a leer (y no exponer oralmente) los contenidos de la memoria.

Además, se puede dar una serie de recomendaciones que ayudarán a desarrollar con éxito una presentación de diapositivas. En general se aconseja:

- Emplear un tipo de letra de fácil lectura (tipo Arial o Verdana) y evitar usar dos tipografías distintas
- Intentar mantener la coherencia y consistencia de formato en todo el documento. El tipo y tamaño de letra del texto al mismo nivel debe ser siempre el mismo
- No escribir todas las palabras en mayúsculas o negrita. El motivo por el que se emplean mayúsculas, negrita, subrayado, etc. es para destacar o diferenciar una idea del resto. Si todo el texto tiene el mismo formato, no estaremos logrando nuestro objetivo de destacar o diferenciar ideas.
- Emplear gráficos fáciles de comprender e ilustraciones que reflejen lo que el estudiante explica
- No introducir imágenes o gráficos irrelevantes al tema de la presentación. A veces los estudiantes desean embellecer sus diapositivas con fondos o imágenes que no guardan relación con el trabajo académico y simplemente distraen la atención de la audiencia
- No sobrecargar de información las transparencias
- No abusar de la animación y efectos especiales, ya que pueden ralentizar la presentación, y a menudo distraen y resultan innecesarios
- Evitar los colores estridentes y que no armonicen entre sí. Una buena elección es aplicar distintas tonalidades del mismo color
- No leer la presentación pues se pierde el contacto visual con la audiencia y convierte al orador en prescindible
- Asegurarse de que la presentación de diapositivas no contiene errores ortográficos, gramaticales o de puntuación

Finalmente, es recomendable que el estudiante lleve consigo copia de la presentación visual en diferentes soportes (en el caso del PowerPoint) por si un dispositivo (pen drive) no se reconociera en el ordenador o no se abriera correctamente. Asimismo, es aconsejable llevar una copia de la presentación impresa en papel por si hubiera algún fallo informático que impidiera mostrar al tribunal la presentación. En este caso el estudiante debe presentar su trabajo normalmente, mostrando así que tiene pleno conocimiento y dominio de los contenidos, y que la presentación visual es solo un apoyo no imprescindible.

b) El póster

El póster es un medio visual que ayuda a ofrecer una síntesis del trabajo realizado. El póster podría estar compuesto de los siguientes apartados:

- Datos de identificación: título (el mismo que el de la memoria escrita), autor y filiación (titulación, curso académico y convocatoria)
- Resumen
- Cuerpo central: en esta sección se presentará el contenido de cada apartado de la memoria, que variará en función del tipo de trabajo realizado. Es aconsejable que tenga la misma estructura que la presentada en la memoria (objetivos, método, desarrollo, etc.)
- Conclusiones: es recomendable incluir un apartado específico con las conclusiones del trabajo. Si procede, también se puede incluir una breve discusión de los resultados, una interpretación de los mismos, recomendaciones para futuros trabajos, sugerencias, etc.

El póster debe ser visualmente atractivo para que capte la atención de la audiencia y ayude a seguir la exposición. Será por tanto necesario que este tenga una secuencia lógica de principio a fin de acuerdo con el desarrollo de la memoria.

En cuanto al formato, hay que recordar que la información mostrada en el póster debe poder leerse e identificarse con facilidad y sin exigir esfuerzo para ser descifrado. A continuación se presenta una serie de recomendaciones sobre el formato de póster:

- Se debe emplear un tipo de letra de fácil lectura (tipo Arial o Verdana) y evitar usar dos tipografías distintas en el mismo póster
- El título del trabajo debe poder leerse desde una distancia de 1,5-2 metros, así que se aconseja imprimir el título en negrita y con un tamaño de al menos 36 puntos
- El nombre del autor y los encabezamientos de los apartados deben ser de menor tamaño que el título (por ejemplo, negrita 30 puntos)
- Los encabezamientos de niveles inferiores pueden aparecer en negrita y tamaño 24 puntos.
- El texto principal no debe contener negrita y puede ser de un tamaño de 20 puntos o más. Es importante que haya consistencia en el formato empleado en cada apartado
- Las tablas y figuras deben ser esquemáticas, de fácil interpretación, y de un tamaño legible
- El texto, las tablas y las figuras no deben confundirse con el fondo del póster. Se debe intentar encontrar armonía en cuanto a tamaño, tipografía y colores
- Se debe introducir espacios en blanco entre los distintos apartados para lograr equilibrio entre estos y los bloques de texto
- Las representaciones gráficas deben ser de alta calidad para que se vean nítidas en el póster impreso

En cuanto a los contenidos de cada apartado, no se debe ofrecer todos los detalles del proyecto y hay que evitar textos extensos o con demasiados datos. Asimismo, es crucial asegurarse de que el póster no contiene errores ortográficos, gramaticales o de puntuación.

Llegado el momento de la impresión del póster, esta se puede realizar bien de una sola pieza en un establecimiento especializado, bien de varias piezas para luego montar el póster en el momento de la presentación. Una vez en la sala de la presentación oral, el estudiante será el responsable de colocar su póster en el lugar que se le asigne, para lo cual será necesario que lleve consigo el material requerido para la fijación (cinta adhesiva, tijera, masilla para colocar carteles, etc.). Al término de la presentación y defensa, el estudiante retirará su póster.

2.3.2. *Guión de ayuda durante la presentación*

En este punto las opiniones son dispares, así que se recomienda seguir las sugerencias del tutor de TFG/TFM. Por lo general, se aconseja no disponer de notas o apuntes como elemento de apoyo en la presentación, pues los estudiantes suelen recurrir a ellos en exceso y en algunos casos se llega incluso a leer literalmente lo escrito, perdiéndose la esencia de la presentación oral. En otros casos, sin embargo, sí puede ser buena idea llevar un guión con palabras clave que ayuden a progresar de forma fluida en la presentación. En cualquier caso, el estudiante debe ensayar su exposición tantas veces como sea necesario y hablar de manera natural (no de memoria) y fluida, haciendo referencias al apoyo visual (póster, presentación de diapositivas, etc.), lo que guiará al tribunal y mantendrá su atención.

2.3.3. *Presentación oral en otros idiomas*

Según la normativa de TFG y TFM (Universidad de Murcia, 2014b, c), es posible elaborar y defender el TFG en idioma francés o inglés, en cuyo caso el estudiante deberá comunicar esta decisión en el momento de solicitar la línea de trabajo. Las recomendaciones en estos casos son las mismas que las dadas en el capítulo anterior en lo que respecta a dominio avanzado del idioma. Esto permitirá expresarse y pronunciar correctamente, lo que dará claridad y fluidez a la exposición.

2.4. El protocolo y las habilidades orales y expositivas

Por último pero no menos importante, hay que tener en cuenta el protocolo del acto, así como las habilidades orales y expositivas que se espera que posea todo estudiante de TFG/TFM. Puesto que la defensa de este tipo de trabajos es un acto académico que ha de desarrollarse conforme a un protocolo, el estudiante debe vestir de manera neutra y adecuada. No será necesario vestir con traje de chaqueta u oscuro, pero sí de forma correcta. Asimismo, el estudiante se referirá a los miembros del tribunal de manera formal y educada empleando la fórmula “Profesor/a + primer apellido” (ejemplo: Profesor Hernández), o “Doctor/a + primera apellido”

(ejemplo: Doctora García). Aunque exista vínculo con alguno de los miembros del tribunal, el estudiante no debe olvidar que se encuentra en un acto académico y se debe respetar un protocolo y saber estar.

A la hora de exponer su trabajo el estudiante debe expresarse con claridad y seguridad. Debe emplear un lenguaje preciso y la terminología pertinente con soltura. Las oraciones deben estar correctamente construidas y ser comprensibles para la audiencia. Se debe asimismo cuidar la pronunciación. Al mismo tiempo, el estudiante debe intentar resultar ameno en su intervención y mantener el contacto visual con el tribunal en todo momento, para lo cual será imprescindible que levante la vista de la presentación o de las notas que pueda llevar consigo como apoyo. Se debe moderar la longitud del discurso y no exceder del tiempo asignado.

Una vez acabada la presentación oral, el estudiante debe agradecer a los miembros del tribunal su atención. Estos intervendrán con preguntas, objeciones, sugerencias o reflexiones sobre la memoria y/o la presentación oral. El estudiante debe responder al tribunal con calidad intelectual e intentar no polemizar. Es recomendable responder de forma individual a cada miembro del tribunal, y a todas las preguntas planteadas con fórmulas como “En cuanto a la pregunta del Doctor...”. Al término del acto el estudiante debe agradecer a los miembros del tribunal una vez más su asistencia y participación en el acto, así como las recomendaciones y comentarios aportados.

3. EVALUACIÓN DEL TRABAJO DE FIN DE GRADO Y FIN DE MÁSTER

El Manual de TFG (Universidad de Murcia, 2014a) presenta los siguientes criterios que el tribunal empleará para evaluar y calificar la presentación y defensa oral del TFG:

Aspectos formales:

1. Claridad de expresión, corrección lingüística y ortográfica
2. Buena estructuración de las partes atendiendo a la metodología planteada
3. Fuentes de información pertinentes y correctamente citada según APA

Contenidos:

4. Capacidad para definir y acotar la temática, problema o ámbito de estudio
5. Conocimiento de la materia
6. Interdisciplinariedad del contenido
7. Originalidad de las ideas y aportaciones personales
8. Corrección metodológica
9. Capacidad de análisis y síntesis
10. Adecuación de la interpretación de resultados y las conclusiones
11. Capacidad de reflexión, de crítica y autocrítica

Aspectos transversales:

12. Aplicación práctica del trabajo al ámbito educativo y social
13. Uso de medios y herramientas apropiadas
14. Inclusión pertinente de materiales ilustrativos

En relación a la exposición y defensa oral:

El estudiante:

- Es claro, preciso, demuestra riqueza de vocabulario y capta el interés del tribunal durante su exposición
- Su expresión gestual, indumentaria y forma de dirigirse al Tribunal son los que corresponden a este acto académico
- Demuestra dominar el tema tratado y realiza una síntesis de los aspectos más relevantes
- Se ha ajustado al tiempo disponible para la exposición
- Responde con corrección a las preguntas del Tribunal

En relación al diseño y uso de la presentación visual:

- La presentación visual es original, creativa y con colores y elementos visuales adecuados
- El estudiante combina bien los elementos visuales con el discurso oral, incluyendo información relevante en cada diapositiva y con aportaciones más allá de lo señalado en las diapositivas
- El contenido de las diapositivas (tablas, figuras, gráficos, organizadores previos....), la tipografía y el tamaño de letra son legibles, claros y adecuados

Por su parte, el tribunal de Trabajo de Fin de Máster tendrá en cuenta los siguientes aspectos en su valoración de la presentación y defensa oral del TFM (Universidad de Murcia, 2014c):

Memoria (máximo 7 puntos) (véase Capítulo 12)

Valoración de los apartados de la memoria:

1. Planteamiento del problema y marco teórico
2. Diseño metodológico (de acuerdo con la modalidad de TFM)
3. Análisis de datos y resultados
4. Conclusiones, consecuencias e implicaciones

Valoración de la calidad de la redacción

Informe del tutor: interés, dedicación, implicación del estudiante y cumplimiento de plazos en la realización del TFM

Exposición (2 puntos máximo)

Valoración de la calidad de los recursos y soportes utilizados

Valoración de la calidad expositiva

Defensa del trabajo (1 punto máximo)

Valoración de la comprensión de las preguntas y de la capacidad de respuesta

Esta guía es una herramienta muy útil para el estudiante, ya que le ayudará a organizar a nivel temporal las diferentes fases de la investigación y elaboración del informe, así como a distribuir sus esfuerzos y evitar divagar.

Al igual que ya se hiciera el final del capítulo anterior, es aconsejable realizar una autoevaluación con el fin de poder valorar si el informe está correctamente realizado y puede ser difundido y/o defendido.

4. RECOMENDACIONES A TENER EN CUENTA

Las siguientes pautas y recomendaciones pretenden ayudar al estudiante a organizar el tiempo disponible y sacar el máximo provecho a sus esfuerzos:

- La exposición oral es una parte muy importante del TFG y TFM, así que deberás dedicar tiempo a formarte y practicar la presentación oral frente a una audiencia. Puedes recurrir a cursos formativos sobre cómo hablar en público, que te ayudarán no solo en tu defensa del TFG/TFM, sino en otras situaciones similares en tu carrera profesional.
- Anota en tu cronograma de trabajo (véase Capítulo 12) cuándo tienes previsto: 1) preparar la presentación oral, 2) crear el recurso o medios visuales que vayas a emplear en la presentación, y 3) ensayar la exposición.
- Repasa los contenidos del póster o presentación de diapositivas para asegurarte de que no haya errores ortográficos, gramaticales o de puntuación.
- Ensaya delante de otras personas y animales a que te pregunten cuestiones sobre tu trabajo. Esto te ayudará a enfrentarte a la situación y ganar confianza. También puedes recurrir a la grabación en video. Verte y oírte te ayudará a detectar aspectos que pueden ser mejorados, pero también a identificar aquello que haces bien, lo cual contribuirá a hacerte sentir más seguro de ti mismo.
- Elabora una lista con las cosas que no debes olvidar llevar contigo el día de la presentación y defensa oral: el póster y los materiales de fijación del mismo, varias copias de la presentación visual, presentación impresa en papel, agua, otros materiales o recursos que desees mostrar al tribunal durante la exposición, etc.
- El día de la presentación es un día importante pero no por ello debemos hacer cosas diferentes a las que hacemos normalmente. Es importante que sigas

la misma rutina que cualquier otro día (desayunar y prepararte, revisar tu lista para asegurarte que no olvidas nada, etc.) y no improvises.

- Acude con suficiente antelación al lugar de evento y prevé cualquier contratiempo que pueda surgir (atascos, retrasos en los servicios de transporte público, hora punta, etc.).
- Recuerda los puntos fuertes de tu trabajo y mantén la tranquilidad y el optimismo.

REFERENCIAS

- McMillan, F., & Weyers, J. (2010). *How to write dissertations and project reports*. Essex: Pearson.
- Universidad de Murcia (2014a). *Manual del Trabajo de Fin de Grado (TFG), Título de Educación Primaria, Convocatoria de junio y junio 2014*. Recuperado de <http://www.um.es/documents/299436/1854327/Manual+de+TFG+EDUCACIO%CC%81N+PRIMARIA+2013-2014.pdf/b7b24b34-8584-42ef-93a3-c5b5f29af60b>
- Universidad de Murcia (2014b). *Normativa por la que se regulan los Trabajos de Fin de Grado (TFG) en la Facultad de Educación* (aprobada en Junta de Facultad el 10 de febrero de 2014). Recuperado de http://www.um.es/documents/299436/319886/NORMATIVA+TFG+APROBADA+JUNTA+CENTRO+10.2.14_definitiva.pdf/d3893b4b-a3aa-46f1-8f7a-63af-52529fb5
- Universidad de Murcia (2014c). *Normativa de los Trabajos de Fin de Máster. Instrucciones sobre el procedimiento a seguir para la presentación, defensa y evaluación de los Trabajos de Fin de Máster*. Máster Universitario en Formación del Profesorado de Educación Secundaria Obligatoria y Bachillerato, Formación Profesional, Enseñanzas de Idiomas y Enseñanzas Artísticas. Recuperado de <http://www.um.es/documents/299436/1951373/NORMATIVA+TFM+2014-15DEF.pdf/e1f70b82-7210-4481-a49e-8797daa052c8>

$2 > -3$

$0.999... = 1$

∞

$\pi \approx 3.14$

$\sqrt{2}$

$1 + 2 \cdot 3$

$(1 - 2$

$5(2 + 2)$

$101_2 =$

χ

$\alpha \omega$

ISBN 978-84-942655-6-3



9 788494 265563