

# **ANÁLISIS Y TRATAMIENTO DE DATOS EN SPSS**

**Autores**

**Eduardo Escalante Gómez**

**Alberto Caro Martín**

**Colaborador**

**Luis Barahona Carvajal**

**Autores**

**Eduardo Escalante Gómez**

**Alberto Caro Martín**

**Colaborador**

**Luis Barahona Carvajal**

**ANÁLISIS Y TRATAMIENTO DE DATOS EN SPSS**

**UNIVERSIDAD DE PLAYA ANCHA / VALPARAÍSO**

**Ediciones**

Universidad de Playa Ancha  
de Ciencias de la Educación.

Casilla 34-V / Fax: 285041

Valparaíso.

Se terminó de imprimir esta PRIMERA EDICIÓN  
en el mes de diciembre del año 2002

Portada: Alejandro Torres

Diagramación: Osvaldo Moraga

Impresión: Imprenta de la Universidad de Playa  
Ancha de Ciencias de la Educación, Valparaíso.

**IMPRESO EN CHILE / PRINTED IN CHILE**

## Sobre los autores

Eduardo Escalante Gómez

Profesor de Estado en Castellano, Universidad de Chile; Magister en Ciencias Sociales, Universidad de Gales (Gran Bretaña).

Alberto Caro Martín

Profesor de Estado en Matemática y Estadística, Universidad Técnica del Estado; Licenciado en Bioestadística, Universidad de Chile y Magister en Evaluación, Universidad de Playa Ancha.

Colaborador

Luis Barahona Carvajal

Licenciado en Estadística, Universidad de Valparaíso.



# ÍNDICE

<b>Introducción .....</b>	<b>7</b>
 <b>Cap. I: Conceptos Básicos para el Procesamiento y Análisis de Datos</b>	
Definiciones.....	9
El procesamiento de los datos .....	16
Ejercicio .....	27
 <b>Cap. II: El Proceso de Medición y el Análisis Estadístico</b>	
La Medición .....	29
Instrumentos de medición. Escalas de actitudes .....	44
El análisis de los datos .....	51
Ejercicios .....	60
 <b>Cap. III: El Caso de Estudio</b>	
La escala de autoestima .....	65
Validación de la escala .....	74
Ejemplo de hipótesis .....	75
 <b>Cap. IV: Análisis Descriptivo Univariante</b>	
El análisis gráfico .....	77
El análisis descriptivo .....	83
Ejercicio .....	113
 <b>Cap. V: Comparación de Dos o Más Poblaciones</b>	
Comparación de dos poblaciones independientes .....	115
Poblaciones relacionadas o dependientes .....	121
Comparación de más de dos poblaciones .....	123
 <b>Cap. VI: Análisis Bivariado</b>	
Asociación y prueba de Chi-Cuadrado .....	133
Correlación lineal .....	140
Regresión lineal .....	152

**Cap. VII: Confiabilidad y Validez de una Escala**

Confiabilidad, significado y medición .....	165
Validez, significado y medición .....	173
Análisis factorial, características y tipos .....	174
Pasos en el desarrollo del Análisis Factorial .....	176
Validez factorial de la escala de autoestima .....	181
Ejercicios .....	193
 <b>Anexo 1: Notas sobre Inferencia</b> .....	 197
 <b>Anexo 2: Esquema de algunas pruebas estadísticas</b> .....	 205

## INTRODUCCIÓN

---

Este libro se origina en la experiencia obtenida en la enseñanza en cursos de posgrado. Los investigadores y analistas a menudo se encuentran ante grandes volúmenes de información que deben procesar numéricamente. Algunos realizan los procesos de cálculos en planillas electrónicas como Excel mediante la activación del módulo estadístico disponible en la versión 5 (análisis de datos), y también se emplean las tablas dinámicas. Pero cuando se trata de muestras relativamente grandes, los procedimientos se hacen más complejos y se requiere una mayor velocidad de procesamiento y herramientas de análisis más sofisticadas. Por este motivo, en los cursos de posgrado siempre optamos por la elección de programas estadísticos más especializados.

Últimamente, hemos venido trabajando con el programa SPSS versión 10 para ambiente Windows. Este es un paquete estadístico de análisis de datos con más de 20 años de aplicación en la investigación de las ciencias sociales y económicas. Contiene programas capaces de realizar desde un simple análisis descriptivo hasta diferentes tipos de análisis multivariantes de datos, como pueden ser: análisis factorial, análisis discriminante, análisis de regresión, cluster, análisis de varianza, etc. También permite analizar series temporales, obtener tablas de frecuencias, realizar gráficos. Esto lo constituye en una excelente herramienta para el investigador.

A través del desarrollo de este libro, hemos querido ilustrar sobre el uso de algunas de las herramientas disponibles en el programa, pero por sobre todo clarificar su aplicación en vista a los objetivos e hipótesis del investigador. Hemos dividido el trabajo en dos grandes tratamientos: el análisis unidimensional y el análisis

bidimensional, dejando para otro trabajo el análisis multidimensional. Sin embargo, hemos incluido dos secciones de bastante importancia para el investigador: el análisis de confiabilidad y el análisis factorial de un instrumento.

Nuestra intención es ayudar al investigador en su toma de decisiones cuando debe enfrentar su plan de investigación (planteamiento del problema, objetivos e hipótesis) y el plan estadístico asociado a él. Esto nos ha llevado a elegir un caso de estudio como eje transversal del libro en lo que se refiere a la ilustración de procedimientos estadísticos, pero también hemos incluido algunos ejemplos específicos que permitan ilustrar un procedimiento numérico. También hemos considerado conveniente incluir después de algunos capítulos, ejercicios y preguntas que permitan reflexionar sobre los temas que se van tratando en el libro.

También con el propósito anterior, se incluyen dos Apéndices: “Notas sobre Inferencia” y “Esquema de Pruebas Estadísticas”. En el primero, se recuerdan algunos conceptos estadísticos sobre Inferencia; en el segundo, se resumen las características fundamentales de algunas de las pruebas que se presentan en el texto, en un lenguaje no técnico.

Los autores

Diciembre 2.002



## CAPÍTULO I

---

### CONCEPTOS BÁSICOS PARA EL PROCESAMIENTO Y ANÁLISIS DE DATOS

Cualquier libro que aborde el tratamiento de datos derivados de la investigación científica, se enfrenta con algo más que la explicación de un procedimiento numérico destinado a reducir información, y obtener así un conjunto más reducido de ella, pero capaz de explicar la característica de un grupo de sujetos investigados.

En este capítulo se presentan algunas definiciones sobre conceptos básicos relativos a la investigación y la Estadística, así como la descripción de algunos aspectos relevantes del procesamiento de los datos.

#### DEFINICIONES

##### *Investigación*

Es una indagación disciplinada. “La investigación disciplinada que difiere de la opinión y creencia, se conduce y maneja de manera tal que el argumento puede examinarse con mucho cuidado. El informe no depende de la elocuencia del escritor o de alguna justificación ‘superficial’”. (Cronbach y Supples, 1969:15, como lo citaron McMillan y Schumacher<sup>1</sup>, 1989:9). Los principios pueden variar de una a otra disciplina.

---

<sup>1</sup> McMillan, J. y Schumacher, S. (1989). *Research in Education: A Conceptual Introduction* (segunda edición). Glenview, Il. Scott, Foresman y Co.

## ***Estudio***

Es la búsqueda de conocimiento, empleando métodos reconocidos para la recopilación, el análisis e interpretación de observaciones. Como tal, Kerlinger (1973)<sup>2</sup> sostiene que el estudio es una de las numerosas formas de *conocimiento*. Otras formas de éste incluyen la revelación, la tradición, la autoridad, la lógica, la intuición y el pacto social.

## ***Estudio Científico / Investigación Científica***

“Es una investigación crítica, sistemática, controlada, empírica sobre preposiciones hipotéticas acerca de las supuestas relaciones entre fenómenos naturales”. (Kerlinger, 1973:11). De acuerdo con Borg y Gall (1983)<sup>3</sup> los objetivos de la ciencia y de la investigación científica son: “1) describir, 2) predecir, 3) controlar, y 4) explicar” la naturaleza y las relaciones entre fenómenos. El estudio científico puede considerarse como una creación inductiva o deductiva, y ambas son consideradas como parte de la tradición filosófica *positivista* (Borg y Gall, 1983:26).

## ***Estudio Social / Investigación Social***

Es la investigación sistemática, empírica y analítica de un fenómeno social (es decir, incluye individuos y grupos, hechos y procesos). En los ámbitos educativos, los fenómenos de interés comparten “creencias comunes, prácticas, artefactos, conocimientos tradicionales y comportamientos” (Goetz y

---

<sup>2</sup> Kerlinger, F. (1973). *Foundations of Behavioral Research* (segunda edición), Nueva York: Holt, Reinhart y Winston, Inc.

<sup>3</sup> Borg, W. (1987). *Applying Educational Research: A Practical Guide for Teachers* (segunda edición), Nueva York: Longman. Borg, W. y Gall, M. (1983). *Educational Research: An Introduction* (cuarta edición), Nueva York: Longman Inc.

LeCompte, 1984:2-3, citado en McMillan y Schumacher, 1989:386). Hay una serie de métodos de investigación utilizados por varias ciencias sociales. Por ejemplo, el estudio social/investigación social puede tomar un modelo inductivo o deductivo (p.e., etnografías y estudios de caso sujeto-único, respectivamente), e involucrar la utilización de estrategias interactivas y/o no interactivas. Como resultado del estudio/investigación social, se puede tomar un modelo positivista (es decir, como ciencia social), u otro modelo al cual recurren numerosas áreas afines etiquetadas como naturales, artísticas, ilustrativas y subjetivas (Borg y Gall, 1983:27; McMillan y Schumacher, 1989:385).

### ***Estudio Educativo / Investigación Educativa***

Es el área particular del estudio disciplinado enfocado sobre el estudio de un fenómeno educativo. “La educación es un campo de interrogantes donde los fenómenos, hechos, gente, procesos e instituciones constituyen los elementos en vivo para hacer preguntas de diversa índole. Las perspectivas y los métodos (han sido) tomados para apuntalar las preguntas que provienen de la educación y que son inherentes a la educación” (McMillan y Schumacher, 1989:9).

### ***Estadística***

Cuerpo de conceptos y métodos usados para recolectar e interpretar datos referentes a un área particular de investigación y extraer conclusiones en situaciones en que la incertidumbre y la variación están presentes. También se puede entender como la ciencia que transforma datos en información.

### ***Clasificación de la Estadística***

*Estadística Descriptiva:* Se encarga de recoger y resumir las

características de una población o muestra, deduciendo de esta descripción conclusiones sobre su estructura, además de las relaciones existentes entre otros colectivos distintos con los cuales se compara.

*Estadística Inferencial:* Basándose en los resultados del análisis de la muestra de la población, induce o estima las leyes generales de comportamiento de la población.

### ***Componentes de la Investigación Estadística***

*Población:* Es un conjunto de elementos al que está referida la investigación y de la que se extraen los datos.

*Individuo o unidad estadística:* Cada uno de los elementos que componen la población. Es un ente observable que no tiene por qué ser una persona, puede ser algo abstracto.

*Muestra:* Subconjunto de elementos de la población, a partir del cual se realiza el estudio estadístico en caso de que no sea posible recopilar toda la información de la población.

*Parámetro:* Característica de una población, resumida para su estudio. Se considera como un valor verdadero de la característica estudiada.

*Variable o serie estadística:* Es una representación numérica de los valores, o una función que asigna un valor.

*Variables discretas:* Pueden tomar un número finito o infinito de valores.

*Variables continuas:* Pueden tomar un número infinito de valores.

Ejemplo:

Población: Alumnos de la Universidad de Playa Ancha

Individuo: Alumno

Muestra: Subconjunto de los alumnos de la Universidad de Playa Ancha

Variables:

Carrera: (Cualitativa). Valores que puede tomar esta variable:

- Ingeniería.
- Matemáticas.
- Física.
- Pedagogía....

Año de nacimiento (Cuantitativa discreta)

Nº de Hermanos (Cuantitativa discreta)

Lugar de nacimiento (Cualitativa)

Altura (Cuantitativa continua)

Tabla 1.1. Tipos de Diseños

**Panorámica de las características clave de diseños cuantitativos no experimentales Tipo y características claves****Estudios Descriptivos**

- a) Se usan para investigar la incidencia, frecuencia y/o la distribución de características en una o más poblaciones identificadas y discretas;
- b) se emplean métodos establecidos para seleccionar una o más muestras de cada población;
- c) se utilizan técnicas estandarizadas para recopilar de cada individuo de cada muestra (p.e. instrumentos lápiz y papel, entrevistas, observaciones; y
- d) se aplican estadísticas descriptivas y/o comparativas para analizar los datos para pautas postuladas de diferencia y/o relación.

**Experimental**

- a) Se emplea para investigar pautas y secuencias de crecimiento y/o cambio en una población como función del tiempo;
- b) se aplican métodos establecidos para seleccionar una o más muestras de población;
- c) se usan técnicas estadarizadas para recopilar datos de cada individuo de cada muestra (p.e. instrumentos lápiz-y-papel, entrevistas, observaciones); y
- d) se traducen datos a forma numérica, cuando sea necesario; y se aplican procedimientos de estadísticas descriptivas y/o comparativas para analizar los datos para pautas postuladas de diferencia y/o relación.

### **Correlacional**

- a) Se usa para investigar la naturaleza y la fuerza de las relaciones entre dos o más variables;
- b) se emplea métodos establecidos para seleccionar una o más muestras de individuos;
- c) se aplican técnicas estandarizadas para recopilar datos de cada miembro de cada muestra (p.e. instrumento, lápiz y papel, entrevistas, protocolos, observaciones);
- d) se emplean procedimientos para la traducción de datos de forma numérica cuando sea necesario; y
- e) se aplican estadísticas correctivas, adecuadas para el tipo de datos recopilado, para analizarlos y generar un coeficiente de correlación: es decir un índice numérico de la relación entre cada par de variables investigadas.

### **Predictivo**

- a) Se emplean para investigar los alcances de una o más variables de predicción útiles en la elección de un criterio variable específico (es decir un comportamiento común, serie de comportamientos o calidad de grupo); o para construir teorías acerca de las posibles determinantes de un criterio variable;
- b) se usan métodos establecidos para una o más muestras de individuos;
- c) se aplican técnicas estandarizadas para recopilar datos de cada miembro de cada muestra sobre predicción y criterios variables (por ejemplo, instrumentos, lápiz y papel, protocolo de entrevistas);
- d) se aplican procedimientos de traducción de datos de forma numérica cuando sea necesario; y
- e) se ejecutan procedimientos de correlación, correlación parcial y/o estadística de correlación múltiple (p.e. regresión múltiple) que aporte un índice numérico de la capacidad de las variables de predicción a considerar por o para explicar la variabilidad en el criterio.

## EL PROCESAMIENTO DE LOS DATOS

Resulta conveniente hacer una síntesis del proceso de investigación que se inicia con el planteamiento de un problema de investigación y que implica responder a interrogantes situadas en la frontera de los conocimientos existentes. Se recorre luego un largo camino donde es preciso atender, simultáneamente, a dos niveles diferentes: el *teórico* y el *empírico*. El primero de ellos se resume en la construcción del llamado marco teórico que, convenientemente llevado a la práctica u bien operacionalizado, indicará en definitiva qué datos es preciso obtener para satisfacer las interrogantes iniciales. En cuanto al nivel *empírico*, la aproximación a la realidad quedará concretada en un cierto tipo de diseño, implementado a través de las técnicas de recolección que permitan obtener los datos requeridos. El instrumento de recolección de datos sintetizará ambos planos y será la herramienta concreta que proporcione la información que se necesita.

Se abre, a partir de ese momento, la etapa propiamente dicha de la *recolección de datos*, que puede tener mayor o menor complejidad y duración. En ella se aplican los instrumentos a los objetos de estudio, se completan las lecturas, y se procede a realizar efectivamente los experimentos, encuestas o estudios que sea menester.

Finalizadas las tareas de recolección, el investigador quedará en posesión de un cierto número de datos, a partir de los cuales será posible sacar las conclusiones generales que apunten a esclarecer el problema formulado en los inicios del trabajo. Pero esa masa de datos, por sí sola, no dirá en principio nada, no permitirá alcanzar ninguna conclusión si, previamente, no se ejerce sobre ella una serie de actividades tendientes a organizarla, a poner orden en todo ese multiforme conjunto. Estas acciones son las que integran el llamado *tratamiento o procesamiento de los datos*.



## ***Información Cuantitativa y Cualitativa***

Lo primero que suele hacerse con el conjunto de los datos obtenidos es dividirlos de acuerdo a un criterio bien elemental, separando de un lado la información que es de tipo numérica de la información que se expresa verbalmente o mediante palabras. Los datos numéricos quedarán como tales, cualquiera sea su naturaleza, y se procesarán luego, para exponerlos en forma clara y fácilmente asimilable. El objetivo final será construir con ellos cuadros estadísticos, promedios generales y gráficos ilustrativos, de tal modo que se sinteticen sus valores y se pueda extraer, a partir de su análisis, enunciados teóricos de alcance más general.

Los datos que se presentan de forma verbal podrán sufrir dos destinos diferentes: se convertirán en datos numéricos o quedarán como información no cuantificada, cualitativa. Así, si se ha realizado una serie de entrevistas en profundidad, de tipo no estructurado, pueden cuantificarse algunos de los datos obtenidos, el número de entrevistados por sexo y edad, la respuesta a alguna pregunta específica, en tanto que otras informaciones se mantendrán en su forma verbal, como paradigmas o ejemplos de las opiniones de la muestra frente a ciertos problemas de interés.

El criterio por adoptar dependerá de circunstancias concretas: de los objetivos del trabajo, de la posibilidad de cuantificar cada variable, del tamaño del universo o muestra considerados y de varios otros criterios. En todo caso, se debe tener presente que, siempre que tenga sentido, es preferible cuantificar en lo posible los datos obtenidos. No existe, como pretenden algunos metodólogos, ninguna oposición entre datos cuantitativos y datos cualitativos: lo que existe es un diferente grado de precisión y de complejidad en cada una de las variables que se miden, de modo tal, que siempre es posible abstraer de ellas algunos aspectos cuantitativos en tanto que otros, por su complejidad, tendrán que permanecer como cualitativos.

Una vez adoptado un criterio frente a cada categoría o

subconjunto de datos se podrá abordar ya, con mayor claridad los objetivos del investigador, y las tareas básicas del procesamiento. En primer lugar, será preciso hacer una revisión detallada de todos los datos obtenidos, atendiendo en especial a su coherencia. Si se trata de cuestionarios o de pautas de observación, de registros de tests, o de cualquier otro instrumento de recolección, habrá que examinar cada uno de ellos para analizarlos internamente, buscando sus posibles incongruencias, omisiones o errores, y enmendándolos si cabe hacerlo. Es decir, se revisará sistemáticamente toda la información disponible, juzgando su calidad y el grado de confianza que merece, para determinar qué parte puede incluirse en el informe de investigación, cual debe corregirse o modificarse en algo (acudiendo, llegado el caso, a una nueva recolección de datos) y cual, por sus graves deficiencias, deberá ser excluidas sin más. Por cada tipo de datos se procederá entonces como sigue:

- Los datos numéricos se procesarán, agrupándolos en intervalos y se tabularán. Luego se construirán con ellos cuadros estadísticos, calculándose además las medidas de tendencia central, de dispersión o de correlación que resulten necesarias.
- Los datos verbales que se desee presentar como numéricos sufrirán una primera operación que se denomina *codificación*. De allí en adelante se trabajarán al igual que los otros datos numéricos, mediante la tabulación y el procesamiento en cuadros estadísticos.

### ***La Codificación de la Información Verbal***

La codificación es un procedimiento que tiene por objeto agrupar numéricamente los datos que se expresan en forma verbal para poder luego operar con ellos como si se tratara, simplemente, de datos cuantitativos. Corresponde al Análisis de Contenido. Para lograrlo, se habrá de partir de un cúmulo de informaciones que tengan una mínima homogeneidad, condición necesaria para poder

integrarlas. Pueden tratarse de cientos de respuestas a una misma pregunta o de una variedad de posibles situaciones observadas mediante un mismo ítem de una pauta de observación: en ambos casos, existirá una determinada variedad de declaraciones o de observaciones que presenten las respuestas o los comportamientos de los objetos de estudio.

El primer paso a dar frente a todos estos datos, es realizar una revisión atenta de un subgrupo reducido de ellos con el objeto de encontrar una tipología de respuestas posibles, en concordancia, por otra parte, con las formulaciones teóricas que guían la investigación y con los criterios adoptados en la etapa de operacionalización. A cada categoría de respuestas se les asigna un código particular, un número o letra diferente, que servirá para agrupar tras de sí a todas las respuestas u observaciones que sean idénticas o que, al menos, aparezcan como equivalentes. Luego se procede a señalar a cada uno de los cuestionarios u hojas de observación con el código que le corresponde a cada respuesta.

El procesamiento de los datos codificados se facilita grandemente cuando se realiza mediante computadoras; en este libro se emplea el programa SPSS versión 10 para ambiente Windows (Otras versiones tienen características muy semejantes, por lo que también son aplicables en ellas). Este programa tiene implementados un elevado número de análisis estadísticos, la mayoría de ellos bajo el epígrafe *Análisis* del menú principal. Al igual que en el apartado relativo a los gráficos, un conocimiento profundo de los datos y una idea clara de las preguntas que se desean responder facilitarán el correcto uso y el máximo aprovechamiento de los análisis estadísticos disponibles.

Supongamos que hemos preguntado, por medio de entrevistas estructuradas hechas a una muestra, la opinión que tienen las personas respecto a las Naciones Unidas y su posición ante los conflictos del Medio Oriente. Si la pregunta ha sido abierta, cada sujeto habrá expuesto sus opiniones en algunas breves frases. La codificación permitirá agrupar sus respuestas, para poder

evaluar cuáles son las opiniones más salientes al respecto y posteriormente permitirá crear sistema categoriales.

Por supuesto que podrían elaborarse otros códigos, categorizando más o menos detalladamente las respuestas. Casi siempre que se encuentran ideas similares pero que se expresan a través de palabras diferentes, se procede a unificarlas en un mismo código. Excepcionalmente, cuando no sólo interesa la idea básica sino además los aspectos lingüísticos con que ésta se manifiesta, es necesario buscar un código para cada respuesta tomada textualmente. Los casos difíciles de ubicar, o de respuestas que pudieran situarse en más de una categoría, sólo pueden ser resueltos, consultando con el marco teórico planteado inicialmente. Cuando aparecen respuestas ambiguas, extrañas o simplemente anómalas, es conveniente agruparlas en el código “otras respuestas”, para no abrir demasiadas categorías a una sola pregunta pues eso puede dificultar el análisis. Se debe evitar que esta categoría quede con una frecuencia alta.

### ***El Libro de Códigos***

Una de las herramientas básicas para un procesamiento adecuado de los datos es disponer de lo que se denomina *libro de códigos* que es una matriz que incluye todas las definiciones asociadas a los datos que se recopilarán o se recopilaron. Algunos analistas emplean el libro de códigos para la caracterización de las variables (variable, categorías, subcategorías, códigos). Esta es una versión que apunta más que nada a identificar la manera como se recogen los datos. Otros analistas incluyen el plan de análisis estadístico asociado a las variables en estudio. Finalmente, algunos optamos por la elaboración de una matriz que incluya las fases del proceso investigativo y el tratamiento de los datos. De este modo, se obtiene una visión completa del proceso investigativo y se puede controlar de manera más consistente el tratamiento de la información a procesar.

Independientemente del formato que se elija para elaborar el libro de códigos, el análisis de datos se inicia con su preparación, que mínimamente debería incluir una lista completa de todas las variables, identificando el nombre de la variable, los valores que se han asignado, y una completa descripción de cómo ha sido operacionalizada la variable. En el libro de códigos, se puede identificar el tipo y escala de medición de cada variable. Finalmente, en función del proyecto de investigación, se identifica el *plan de análisis 1*: estadística descriptiva y el *plan de análisis 2*: estadística inferencial. El libro de códigos es fundamental para interpretar los datos y no perderse con los números.

Supongamos que se tiene un cuestionario que en una de sus partes incluye los siguientes aspectos:

*Cuestionario*

Items	1 (TD)	2 (D)	3 (I)	4 (A)	5 (TA)
1. Controlo las circunstancias de mi vida					
2. Estoy satisfecho con mi vida					

3. Sexo:

1. Hombre 2. Mujer

4. Rango Académico

1. Instructor 2. Asistente 3. Asociado 4. Titular

Un libro de códigos básicos incluiría lo siguiente:

**Casos** = 30 (N=30)

**Variables** = 5

**Nombre de las Variables:**

Control

SatisVid

Sexo

RanAcad

**Descripción de las Variables:**

Control: Control personal

SatisVid: Satisfacción con la vida de uno

Sexo: Sexo del Encuestado

RanAcad: Rango Académico del encuestado

**Nombre de la Variables/Atributos de la Variables (Códigos)**

Control

1 = Totalmente en Desacuerdo

2 = En Desacuerdo

3 = Indeciso

4 = Acuerdo

5 = Totalmente de Acuerdo

SatisVid

1 = Totalmente en Desacuerdo

2 = En Desacuerdo

3 = Indeciso

4 = Acuerdo

5 = Totalmente de Acuerdo

Sexo

1 = Hombre

2 = Mujer

RanAcad

1 = Instructor

2 = Asistente

3 = Asociado

4 = Titular

El investigador ha recogido los siguientes datos a partir del cuestionario, el correspondiente libro de códigos y los ha vaciado en la siguiente matriz:

Tabla 1.1. Datos de Prueba del Cuestionario ( $N=30$ )

<i>Caso</i>	SatisVid	Control	Sexo	RenAcad
1	4	3	2	4
2	5	4	1	2
3	3	3	1	1
4	3	2	1	1
5	4	3	2	2
6	5	4	2	3
7	5	5	2	3
8	3	3	1	1
9	2	2	2	1
10	3	3	2	1
11	3	3	1	2
12	4	5	1	3
13	3	3	1	2
14	4	5	2	3
15	5	4	1	3
16	4	3	2	4
17	2	1	1	2
18	3	2	2	1
19	2	3	2	1
20	4	4	1	3
21	4	3	1	2
22	5	5	1	3
23	4	4	2	3
24	3	2	1	2
25	2	2	1	1
26	4	3	2	1
27	5	4	2	4
28	4	3	2	2
29	3	2	1	1
30	5	4	1	2

La elaboración del libro de códigos le permitirá procesar los datos y definir las operaciones estadísticas que desea realizar. Por ejemplo, producirá una tabla de frecuencia para la variable *Sexo*, un gráfico de barras para la variable *RenAcad*, solicitará de SPSS la estadística descriptiva para las variables *Control* y *SatisVid*, aunque previamente solicitará el histograma para ambas variables, en orden a examinar su distribución. Además, se pregunta si existe relación entre el rango académico y el grado de satisfacción con la vida de uno.

### ***La Tabulación***

Una vez definidos los códigos, y ya marcados los cuestionarios o pautas con los mismos, se está en condiciones de proceder a la subsiguiente etapa, la de tabulación.

La palabra *tabulación* deriva etimológicamente del latín *tabula*, y significa hacer tablas, listados de datos que los muestren agrupados y contabilizados. Para ello es preciso contar cada una de las respuestas que aparecen, distribuyéndolas de acuerdo a las categorías o códigos previamente definidos. Cada una de las preguntas de un cuestionario o de una hoja de observación tendrá que ser tabulada independientemente, por lo que es preciso hacer previamente un plan de tabulación que prepare adecuadamente la tarea por realizar.

Esta etapa del trabajo puede desarrollarse manualmente, mediante el procedimiento que se esbozará a continuación, cuando se trate de un número relativamente pequeño de datos; cuando esto no sea así, resultará mejor acudir al procesamiento electrónico de datos, para lo cual existen ya adecuados paquetes de computación que facilitan grandemente esta tarea. Para ello, es necesario preparar los instrumentos de recolección con casilleros especiales de registro para que las máquinas los procesen.

Para realizar una tabulación manual deberán confeccionarse



planillas u hojas de tabulación donde figuren los códigos sobre la base de los cuales se habrán de distribuir los datos. También deben dejarse espacios para señalar, mediante signos convencionales, las unidades que se van contabilizando. En la columna de códigos, se anotarán cada uno de los que se hayan establecido en el proceso de codificación de la variable, si ésta no se presentaba inicialmente en forma cuantificada. Si no es así, irán directamente los intervalos establecidos para la distribución. En el espacio reservado a las respuestas se colocará, por cada cuestionario que se tabule, un signo en la categoría que corresponda según la respuesta que en él aparezca. Una vez tabulados todos los cuestionarios, se contarán las respuestas anotadas en cada casillero, expresándose en números en la columna del total.

El resultado de la tabulación será este cómputo ordenado de las respuestas. Sin embargo, las tablas así obtenidas no serán aptas todavía para presentarse al público, pues deberán sufrir algunas transformaciones de forma que permitan su mejor comprensión. Las tablas que resultan de la tabulación deben considerarse, por lo tanto, como un material de trabajo, como un producto preliminar todavía no completamente terminado.

### ***Los Cuadros Estadísticos***

Luego de terminar con la tabulación de toda la información contenida en los instrumentos de recolección de datos, es preciso presentar los resultados de modo tal que éstos se hagan fácilmente inteligibles aún para los lectores no especializados. Para lograrlo, es preciso presentarlos en la forma más clara posible, haciendo explícito cualquier elemento que pueda dar origen a confusiones o dobles interpretaciones y ordenando toda la información de la manera más rigurosa. Este criterio se manifiesta, en primer lugar, en el correcto titulado de cada cuadro. Ello significa que el título debe contener todas las características de la información que se presenta, en forma concreta, específica y no difusa. Debería

responder a las preguntas ¿Qué?, ¿Cómo?, ¿Cuándo? y ¿Dónde?

Cuando en un cuadro aparecen dos variables, deberá mencionarse primero aquella que suponemos dependiente y luego la independiente. Así quedaría, v.g.: “Opinión de los entrevistados frente al control de la natalidad, según su edad”. Si existen más variables, se respetará la misma regla, anotando para el caso mencionado: “Opinión de los entrevistados frente al control de la natalidad, según sexo y edad”.

Después de encontrar un título adecuado para el cuadro, la tarea siguiente es convertir en porcentajes las cifras reales (llamadas absolutas) que hemos obtenido de la tabulación. Ello permite, por una parte, tener una mejor idea de lo que puede representar una parte frente al todo. Es muy claro lo que significa que el 23% de la gente prefiere X y no Y. En cambio, no se percibe tan claramente el valor relativo de 171 respuestas favorables a X sobre el total de 743 casos. De la misma manera, la práctica de calcular porcentajes permite, de una rápida ojeada, percibir diferencias y similitudes, apreciar variaciones y tendencias y hacer comparaciones con otras investigaciones. Por ello su uso es constante prácticamente en todos los cuadros estadísticos.

Puede ocurrir que, en un cierto cuadro, aparezcan diferentes bases, algunas de las cuales puede llegar a ser menores que las cifras fijadas como mínimas.

Las cifras presentadas deben tender a redondearse en lo posible, pues se introduce cierta confusión al mencionar que hay 32,87% de una cierta categoría y 67,13% de la contraria. Mucho más simple es hablar de 32,8 % y 67,1 por ciento. Por supuesto que, cuando la muestra se mayor, más aceptable será incluir decimales en las cifras presentadas. En cambio, si las bases son pequeñas hay que evitar dar la impresión de una exactitud que la información, en realidad, no posee.

### ***Ejercicio***

Elabore un libro de códigos para un proyecto de estudio sobre la Adaptación a la Modernidad, que se realizará en Profesores , y que considera 3 variables de clasificación: sexo, edad, y origen (urbano-rural) y una escala de actitudes de 15 ítems con 5 alternativas de respuestas (muy en desacuerdo, en desacuerdo, no opina, de acuerdo y muy de acuerdo) sobre el proceso de adaptación a la modernidad.



## CAPÍTULO II

---

### EL PROCESO DE MEDICIÓN Y EL ANÁLISIS ESTADÍSTICO

#### LA MEDICIÓN

Campbell (1928)<sup>4</sup> concibe la medición como la asignación de números para representar propiedades de sistemas materiales no numéricos, en virtud de leyes que gobiernan esas propiedades. Este autor asigna a la medición el papel de establecer el isomorfismo entre la idea de cantidad y las magnitudes de las propiedades a ser medidas: Para que la medida sea posible, es necesario que haya una correspondencia unívoca (isomorfismo) entre los números y los objetos, de manera que se puedan realizar las mismas operaciones con éstos que con los números. Campbell reduce la medición al uso de escalas cuantitativas de intervalo o de razón, a la vez que considera que su sistema no es extensible a la medición de atributos psicológicos. El argumento utilizado para ello fue el de que tales atributos, al ser intrínsecos o no observables, son incapaces de ser medidos.

Para este autor, las propiedades de los números que son más importantes para la medición son tres:

---

<sup>4</sup> *An Account of the Principles of Measurement and Calculation*, Londres, Longman.

### 1) Identidad

- a es igual a b o diferente de b (no hay tercera alternativa)
- Si  $a=b$ , entonces  $b=a$  (simetría o reflexibilidad)
- Si  $a=b$  y  $b=c$ , entonces  $a=c$  (transitividad)

### 2) Orden jerárquico

- Si  $a>b$  entonces  $b<a$  (asimetría)
- Si  $a>b$  y  $b>c$  entonces  $a>c$  (transitividad)

### 3) Aditividad

- Si  $a=p$  y  $b>0$  entonces  $a+b>p$  (aditividad)
- $a+b=b+a$  (conmutatividad)
- Si  $a=p$  y  $b=q$  entonces  $a+b=p+q$  (objetos idénticos pueden sustituirse)
- $(a+b)+c=a+(b+c)$  (asociatividad)

Stevens (1951)<sup>5</sup>, por su parte, se refiere a la medición como la “asignación de números a objetos o eventos según ciertas reglas”, mientras que Torgerson (1958)<sup>6</sup> entiende por tal, “la asignación de números a sistemas que representan la propiedad objeto de medición”. Para Stevens, la escala de medida determina enteramente qué tipo de análisis estadísticos son los apropiados.

Supples y Zinnes (1963)<sup>7</sup> consideran que el proceso de medición implica la conversión de la información acerca de un sistema empírico en un sistema numérico. Posteriormente, Lord y

---

<sup>5</sup> “Mathematichs, measurement and psychophysics”, en S.S. Stevens (ed.) *Handbook of Experimental Psychology*, Nueva York, Willey, p. 1-49

<sup>6</sup> *Theory and Method of Scaling*, New York, Wiley.

<sup>7</sup> “Basic measurement theory”, en R.D. Luce, R. R. Bush y E. Galanter (eds.), *Handbook of Mathematical Psychology*, vol. I, págs. 1-76, Nueva York, Wiley.

Novick (1968)<sup>8</sup> la definen como “un procedimiento para asignar números a propiedades específicas de unidades experimentales de forma que caracteriza y preserva las relaciones específicas en el dominio conductual”<sup>9</sup>.

Entre las definiciones más recientes, se destacan las de Coombs, Dawes y Tversky (1981)<sup>6</sup> según la cual, la medición es el proceso por medio del cual se representan propiedades por medio de números.

### ***El Procedimiento de Medición***

Los teóricos coinciden en que la medida comienza con un procedimiento para identificar los elementos del mundo real con los elementos o constructos del sistema lógico, por medio de una definición semántica precisa; una vez hecho esto, puede procederse a la medición, para lo cual deben seguirse los siguientes pasos: 1) identificar el objeto o sujeto a medir; 2) identificar la propiedad o conducta que se quiere medir en el objeto o sujeto y el instrumento; 3) aplicar una regla de asignación numérica por la que se asigna un número a la propiedad de la unidad experimental (objeto o sujeto) que se está midiendo.

Entonces, en sentido amplio, la medición es un procedimiento por el cual asignamos una determinada categoría o valor a una variable, y para un determinado sujeto, muestra o población. Cuando a Juan (sujeto) le asignamos el valor 1,70 metros para la variable “estatura”, hemos hecho una medición. También hacemos una medición cuando a Beatriz le asignamos la categoría “soltera” para la variable “estado civil”, lo cual muestra que estamos tomando la idea de medición en sentido amplio, ya que no sólo abarca mediciones cuantitativas (1,70 metros) sino también cualitativas (soltera). Esta ampliación del concepto se hace

---

<sup>8</sup> *Statistical theories of Mental Test Scores*, Reading, Mass., Addison-Wiley.

<sup>9</sup> *Introducción a la psicología matemática*, Madrid, Alianza

especialmente importante en las ciencias sociales, porque en este ámbito se manejan mayor cantidad de variables cualitativas que en ciencias como la física o la química, donde prevalecen las variables cuantitativas y de donde hemos sacado la creencia de que sólo se puede medir con reglas, balanzas y decibelímetros.

Las mediciones individuales recaen sobre sujetos, mientras que las mediciones estadísticas además recaen sobre muestras o poblaciones. Así por ejemplo decimos que medimos una muestra A cuando le asignamos la media aritmética 120 respecto de la variable inteligencia. Obviamente, las medidas estadísticas se derivan de las individuales, pues, no se puede obtener un promedio si desconocemos los valores individuales de los sujetos de la muestra.

Para llevar a cabo una medición, necesitamos obligatoriamente tres cosas: a) un instrumento de medición, b) un sistema de medición (o escala), ya que hay diferentes grados de precisión en la medición, y c) un procedimiento para hacer la medición.

Un instrumento de medición es un artificio usado para medir. Una balanza es un instrumento para medir peso, un test mide personalidad, inteligencia, aptitud vocacional, etc., un cuestionario mide conocimientos, etc., pero sea cual fuere el instrumento utilizado, para que sea eficaz debe reunir dos requisitos: *confiabilidad y validez*.

Si la misma balanza hoy informa que peso 70 kg, mañana 80 kg y pasado 63 kg, es un instrumento poco confiable. Un test puede también ser poco confiable si, respecto del mismo sujeto, arroja resultados dispares en diferentes momentos de administración. Confiabilidad significa, entonces, estabilidad o constancia de los resultados.

Esta constancia no es nunca perfecta. El hecho de que puedan existir ligeras variaciones debidas al azar no compromete la confiabilidad del instrumento de medida, siempre que esas variaciones sean poco significativas. También debemos tener en



cuenta los cambios propios del sujeto cuya propiedad medimos. Si el sujeto aumenta de peso o mejora su rendimiento intelectual por un efecto madurativo habrá una variabilidad en los resultados, lo cual no implica necesariamente que el instrumento no sea confiable. En la confiabilidad, como aquí la entendemos, la constancia o la variabilidad en los resultados depende de la habilidad del instrumento mismo, no del sujeto. Por ejemplo, si el instrumento tiene una consigna de aplicación ambigua, podrá ser aplicado de manera distinta antes y después, y los resultados podrán ser diferentes.

Un test o cualquier otro instrumento puede ser confiable pero no válido: es válido cuando mide lo que efectivamente pretende medir, lo cual puede parecer una perogrullada pero que no es tal si nos damos cuenta de los casos de instrumentos que querían medir una propiedad y medían otra. Si con un determinado test pretendemos medir memoria pero sólo medimos ansiedad, no es válido. La validez de un test puede establecerse prospectivamente: nos percatamos que un test de aptitud para la ingeniería es válido cuando, andando el tiempo, constatamos que quienes obtuvieron altos puntajes hoy son buenos ingenieros, y quienes sacaron bajos puntajes no lo son.

Siempre estamos haciendo mediciones. Por ejemplo, cuando decimos que tal o cual individuo es católico (medición nominal), o que tal sujeto tiene más mérito que otro (medición ordinal), o cuando decimos que tal otro sujeto tiene 30 años (medición de razón). Pero la importancia de la medición no reside simplemente en el hecho de que siempre estamos haciendo mediciones, sino en que el acto de medir nos permite conocer cómo varían los fenómenos de la realidad tanto en cantidad como en calidad, lo cual a su vez, es muy importante porque es gracias a que existen unas variaciones que dependen causalmente de otras, que podemos organizar nuestro conocimiento del mundo.

### ***Las Escalas Derivadas***

Las escalas de representación derivadas de los valores originales sirven para mejorar la interpretación y la comprensión de los datos, pero además posibilitan la comparación de resultados de variables distintas expresadas en unidades diferentes. Es el caso de los percentiles y las puntuaciones típicas .

- *Percentiles*: Corresponden al punto de la distribución por encima y por debajo del cual se sitúa un determinado tanto por ciento del grupo. Los 99 puntos percentiles dividen la distribución en 100 partes y cada una de ellas contiene el 1% de los casos.
- *Puntajes estandarizados (z)*: Emplean la estandarización de la variable, la cual es una manera de transformar el puntaje bruto original de modo que cada puntaje se especifica en términos de magnitud y de dirección de su desviación de la media, medida en unidades de desviación estándar. La estandarización se usa para hacer los datos más interpretables. Por ejemplo, estandarizar las variables permite hacer comparaciones entre individuos que tomaron una prueba en diferentes tiempos, y fueron medidos usando diferentes instrumentos. Para calcular el puntaje Z, esto es para estandarizar un valor, se sustrae el puntaje bruto de la media de la distribución y luego se divide este puntaje por la desviación estándar. La estandarización produce una transformación de los puntajes originales de modo que la nueva distribución tendrá una media que es igual a cero y una desviación estándar que es uno..

A partir del puntaje Z se obtienen escalas derivadas (T, CEEB) las que permiten una mejor presentación de los puntajes Z.

- *Notas T*: la media equivale a 50 y la desviación estándar a 10.

- Notas CEEB: la media equivale a 500 y la desviación estándar equivalente a 100.

### ***Escala de Medición y su Clasificación***

La introducción del concepto de escalas de medida en psicología y educación se remonta al trabajo desarrollado por Stevens, demostrando que en las ciencias del comportamiento son posibles diversos niveles de medición, mediante la elaboración de diversas axiomatologías.

Este modelo de medición constituye y ha constituido una herramienta formal poderosa en la catalogación de las pruebas estadísticas. Y si atiende a esta clasificación prevista por Stevens, los datos psicológicos y educacionales que pueden ser medidos en una escala de intervalos no son en principio abundantes. Es decir, la mayoría de las variables psicológicas y educacionales estarían dentro del modelo de escala ordinal (por ejemplo, los tests de inteligencia o los cuestionarios de personalidad).

La medición supone que se puede establecer una relación homomórfica entre un sistema relacional empírico y otro formal. Esto permite introducir la distinción entre niveles de medición, o escalas de medición.

Es interesante la apreciación de Kerlinger quien indica que a pesar de que el nivel de medición de la mayoría de las variables psicológicas es sólo ordinal, el tratamiento estadístico que se hace con ellas es asimilable al de los otros niveles pues “funcionan bien” simplemente. Esto está de acuerdo con la idea que los niveles de medición no condicionan el procesamiento estadístico pues ambos tienen supuestos diferentes.

Se han ofrecido muchas clasificaciones de los tipos de escala, teniendo en cuenta las propiedades de los sistemas numéricos:

Campbell (1928): escalas en ordinales y de razón.

Stevens (1951): nominales, ordinales, de intervalo y de razón

Tanto las escalas de medición como las de clasificación, implican la formulación y evaluación de reglas. Estas reglas son usadas para medir atributos de los objetos, por lo general, pero no de manera exclusiva, sino también de personas. Es importante recordar que se pueden medir sólo atributos de los objetos, no a los objetos mismos. Entre las características de las buenas reglas están la receptibilidad (confiabilidad) y de modo más importante, la validez en los sentidos que se describan. La estandarización es un objetivo importante de la medición debido a que facilita la objetividad, cuantificación, comunicación, economía y generalización científica.

Un aspecto tradicionalmente importante, pero controvertido, de las escalas de medición, es que éstas, por lo general, caen en uno de cuatro niveles (han sido sugeridos otros): nominal, ordinal, de intervalos y de razón. Estos cuatro niveles representan de manera progresiva reglas mejor articuladas, por ejemplo, las escalas nominales tan sólo definen si dos (2) objetos son equivalentes o no, entre sí, con respecto a un atributo crítico, pero las escalas ordinales determinan si un objeto que no es equivalente a otro es mayor o menor que el otro; son posibles resultados más consistentes con niveles superiores de medición. El concepto de invarianza es básico para estos niveles de medición, el cual tiene que ver con lo que permanece igual cuando se realizan cambios permisibles en la escala. Las escalas de niveles superiores están más restringidas en cuanto a la manera en que pueden ser transformadas y pueden aún conservar invarianzas claves.

La idea de medición es intrínsecamente comparativa. Medir algo, en el caso más sencillo, es determinar cuantas veces una cierta unidad o patrón de medida, cabe en el objeto a medir. Para medir la longitud de un objeto físico, nosotros desplazamos una regla o cinta graduada sobre el mismo, observando cuantas unidades (en este caso centímetros o metros) abarca el objeto en cuestión. Es decir que comparamos el objeto con nuestro patrón

de medición para determinar cuántas unidades y fracciones del mismo incluye.

Una *escala* puede concebirse como un continuo de valores ordenados correlativamente que admite un punto inicial y otro final. Si evaluamos el rendimiento académico de estudiantes, podemos asignar el valor cero al mínimo rendimiento imaginable y un valor de 7, 10, 20 ó 100 puntos, según convenga, al mayor rendimiento posible. Con estos dos valores tendríamos ya marcados los límites de nuestra escala: para concluir de confeccionarla, será necesario asignar a los posibles rendimientos intermedios puntajes también intermedios. Con ello obtendremos una escala capaz de medir la variable rendimiento académico a través de los indicadores concretos de los trabajos presentados por los estudiantes, de sus exámenes, pruebas y otras formas de evaluación posibles.

Existen diferentes tipos de escalas que se distinguen de acuerdo a la rigurosidad con que han sido construidas y al propio comportamiento de las variables que miden. Se acostumbra a clasificarlas en cuatro tipos generales que son los siguientes: escalas nominales, ordinales, de intervalos iguales y de cocientes o razones.

<b>Tipos de variable</b>	<b>Clasificación</b>	<b>Escala de Medición</b>
Cualitativa	Discreta	Nominal
		Ordinal
Cuantitativa	Discreta	Intervalo
		Razón
	Continua	Intervalo
		Razón

Tabla No.2.1. Tipos de variables y escalas de medición

Tabla No. 2.2: Escalas y operaciones empíricas

Escalas	Operaciones empíricas básicas	Ejemplo
Nominal Ordinal	Relación de igualdades Relación de mayor o menor	Sexo La clase social
Intervalo	Determinación de la igualdad de la diferencia de intervalos	El rendimiento académico medido en el rango de 0 a 7
Razón	Existencia de un cero real	El tiempo de reacción

Escalas *nominales* son aquéllas en que sólo se establece una equivalencia entre la escala y los diferentes puntos o valores que asume la variable. Son una simple lista de las diferentes posiciones que pueda adoptar la variable, pero sin que en ella se defina ningún tipo de orden o de relación. Si en una investigación educativa al estudiar la ficha escolar, se quiere determinar la estructura de la familia de los estudiantes, tendremos una variable que se designará como “condición familiar”. Los distintos valores que esa variable reconoce serán, concretamente: monoparental, biparental, etc. Entre estos valores no cabe obviamente ninguna jerarquía, no se puede trazar ningún ordenamiento. Sin embargo, a la enunciación explícita de todas estas posibilidades la consideramos como una escala, pues de algún modo es útil para medir el comportamiento de la variable, indicándonos en qué posición se halla en cada caso.

Las escalas *ordinales* distinguen los diferentes valores de la variable, jerarquizándolos simplemente de acuerdo a un rango. Establecen que existe una gradación entre uno y otro valor de la escala, de tal modo que cualquiera de ellos es mayor que el

precedente y menor que el que le sigue a continuación. Sin embargo, la distancia entre un valor y otro queda indeterminada. En otras palabras, tales escalas nos esclarecen solamente el rango que las distintas posiciones guardan entre sí. Un ejemplo de escala ordinal es el que suele usarse para medir la variable “grado de escolaridad”: podemos decir que una persona que ha tenido 2 años de instrucción escolar ha recibido más instrucción que quien sólo tiene un año y menos que quien posee tres. Sin embargo, no puede afirmarse válidamente que la diferencia entre quien posee 2 años de instrucción y quien ha recibido un año es igual a la diferencia entre quienes han recibido 16 y 17 años de educación formal. Por tanto, como no podemos determinar la equivalencia entre las distancias que separan un valor de otro, debemos concluir que la escala pertenece a la categoría ordinal.

Las escalas de *intervalos iguales*, además de poseer la equivalencia de categorías y el ordenamiento interno entre ellas, como en el caso de las ordinales, tienen la característica de que la distancia entre sus intervalos está claramente determinada y que éstos son iguales entre sí. Un ejemplo típico de las escalas de intervalos iguales está dado por las escalas termométricas. Entre 23 y 24 grados centígrados, por ejemplo, existe la misma diferencia que hay entre 45 y 46 grados. Muchas otras escalas, como las que se utilizan en los test psicológicos y de rendimiento, pertenecen a este tipo. La limitación que poseen es que no definen un cero absoluto, un valor límite que exprese realmente la ausencia completa de la cualidad medida. Por ello no se pueden establecer equivalencias matemáticas como las de proporcionalidad: no puede afirmarse que  $24^{\circ}\text{C}$  es el doble de temperatura que  $12^{\circ}\text{C}$ , porque el cero de la escala es un valor arbitrario y no se corresponde con la ausencia absoluta de la variable que se mide.

Por último, están las escalas de *cocientes o razones*. En ellas se conservan todas las propiedades de los casos anteriores, pero además se añade la existencia de un valor cero real, con lo que se hacen posibles ciertas operaciones matemáticas, tales como la

obtención de proporciones y cocientes. Esto quiere decir que un valor de 20 en una escala de este tipo es el doble de un valor de 10, o de las dos terceras partes de un valor de 30. Son escalas de cocientes las que miden el peso, los ingresos, etc. Difícilmente, las variables que intervienen en las ciencias sociales son medidas con escalas de razones, pues son contados los casos en que dichas variables pueden ser definidas con la exactitud y precisión necesarias. La economía y la demografía son, entre estas disciplinas, las que más utilizan escalas de razones.

En SPSS v.10 se pueden definir las variables en *Vista de Variables* columna Medida. El programa admite las definiciones Nominal, Ordinal (que corresponden a las discretas) y de Escala (las que corresponden a continuas o discretas con más de 24 valores distintos):

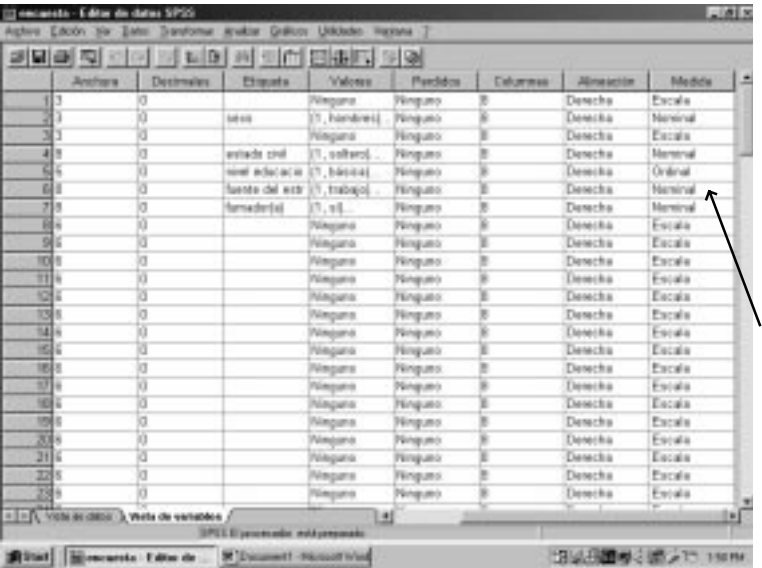


Figura 2.1. Hoja de ingreso de las variables y sus respectivas especificaciones



Para que las observaciones de la realidad sean útiles, tienen que llevarse a cabo en condiciones de rigor metodológico: el objeto de medición ha de ser observable empíricamente (directa o indirectamente), debe presentar cierta variabilidad, y para efectuar las mediciones, es necesario disponer de instrumentos de medidas precisos. Este último aspecto de las condiciones de bondad de un instrumento de medida acapara una amplia zona de estudio de la psicometría clásica.

### *Escalas de Medición y Análisis Estadístico*

Lord (1953)<sup>10</sup>, Anderson (1961)<sup>11</sup> o Prytulak (1975)<sup>12</sup> sostienen que la medición y el tratamiento estadístico son dominios independientes, y como consecuencia de ello, las consideraciones de la medición no influyen en las técnicas estadísticas. Al respecto, existe controversia entre los estadísticos, los experimentalistas y psicómetras respecto a la aplicación de los conceptos de escalas de medición y su asociación con las pruebas estadísticas. Así nos podemos encontrar con que mientras unos investigadores, fuertemente impregnados por las ideas de Stevens, han preconizado que los datos deberán ser analizados con técnicas estadísticas que contemplen el tipo de medida empleada, otros investigadores consideran que los niveles de medida no juegan un papel esencial en la inferencia estadística, ya que ambos campos constituyen dominio independientes (en la línea de Lord). De esta forma, la escala de medida es determinada por el análisis, asumiendo que la restricción de los estadísticos, atendiendo a la escala de medida utilizada es una idea errónea (y además es ignorada en la mayoría

---

<sup>10</sup> "On the statistical treatment of football numbers", The American Psychologist, 8

<sup>11</sup> "Scales and statistics: parametric and nonparametric", Psychological Bulletin, 58.

<sup>12</sup> "Critique of S.S. Stevens' Theory of Measurement Scale Classification", Perceptual and Motor Skills, 41.

de los análisis estadísticos reales). Compartimos la postura de Muñiz cuando señala que “aunque Stevens al desarrollar la idea de las escalas prestó un servicio a la teoría de la medición, su noción condujo a un malentendido que ha sido difícil de eliminar” (2001:302)<sup>13</sup>.

Cuando se inicia el procesamiento de los datos se observa primero los tipos de variables que existen, tratando de identificar aquéllas que pueden influir en el resto de la variables como por ejemplo sexo, edad, estado civil, tipo de colegio, etc..

### ***Fuentes de Variación en las Mediciones***

Cuando tenemos un conjunto de mediciones, las variaciones se deben frecuentemente a dos motivos:

1. A la variación del fenómeno en sí. El fenómeno varía en relación a los demás fenómenos y en relación a sí mismo: “entre” e “intra”.
2. Al error producido en la medición del fenómeno: el error es debido al observador, al instrumento utilizado o a la situación donde se efectúa la medición.

### ***Tipos de Errores de Medición***

El tipo de error cometido puede ser:

- 1.- *Aleatorio*: es el producido por el sistema de realización de la medición. Su valor no afecta al valor real ni al promedio.
- 2.- *Sistemático*: es el error de redondeo que se lleva a cabo en cada una de las mediciones que se efectúan. Es el llamado sesgo.

---

<sup>13</sup> Teoría Clásica de los tests, Pirámide, Madrid

Los dos tipos de errores pueden darse conjuntamente. Es muy importante conocer la cantidad de error que se está cometiendo.

### ***Fuentes de Error en la Medición***

El empleo de instrumentos de medición produce dos tipos de errores: el error de medición y el error de muestreo. El primero se asocia a las diferencias que se producen al medir a un individuo en diversas ocasiones; el segundo, al emplear muestras de individuos para estimar características de la población. Un problema que se presenta al estudiar los resultados de las mediciones y el efecto de los errores, es decidir qué es un resultado habitual y qué no. Una forma de resolverlo es tomar como referencia la distribución Normal en la cual aproximadamente el 95% de los valores se encuentran en el intervalo definido por la media menos dos desviaciones estándar y la media más dos desviaciones estándar. Los valores que caen en ese intervalo se consideran “habituales” mientras que los que caen fuera de ese intervalos se consideran “no habituales”.

Cuando se asume que las variables tienen un distribución normal, se supone que lo habitual o no sigue el patrón indicado, llegándose a considerar como “normales” los resultados incluidos en el intervalo indicado y como “no normales” los que caen fuera. El supuesto de normalidad es necesario en muchos procedimientos estadísticos.

El problema de aplicar este criterio es que determina que la normalidad es uniforme, se asume que todas las variables son normales, pero no siempre es un buen modelo, ni siquiera como modelo estadístico; es más aplicable para algunas variables y por ello se deberá probar este supuesto cuando corresponda.

Existen formas de medición más robustas que ésta; cuando hablamos de rango de normalidad no se habla de la curva normal sino que se está determinando la proporción de personas que están dentro del percentil 97,5 (normalidad) y 2,5 (anormalidad).

Los percentiles no dependen de la forma concreta de la

distribución; son independientes del hecho de que la curva sea de distribución normal o no. Este criterio también es estadístico y se emplea cuando no se tienen otros criterios.

## INSTRUMENTOS DE MEDICIÓN: ESCALAS DE ACTITUDES

### *Escala Tipo Likert*

En la recolección de los datos, es fundamental que se defina el tipo de instrumento de medición que se empleará para recogerlos. Existen muchos tipos de instrumentos: escalas, pruebas, cuestionarios, pautas de observación. Entre las escalas, unas de las más usadas en la investigación para medir actitudes es la escala tipo Likert, pero existen otras escalas como el Diferencial Semántico, la de Thurstone y la de Guttman,.

En esta sección nos referiremos a las escalas tipo Likert que son las que se aplicaron en nuestro caso de estudio.

### *Las Actitudes*

*Likert* señala que las actitudes pueden medirse a través de manifestaciones verbales y asume la posibilidad de estudiar dimensiones de actitud a partir de un conjunto de enunciados que operen como reactivos para los sujetos. Éstos pueden situarse en la variable de actitud desde el polo más favorable al más desfavorable, de tal manera que la variación de las respuestas será debida a diferencias individuales de los sujetos.

Allport (1935:81)<sup>14</sup> concibe las actitudes como, “*un estado mental y nervioso de disposición, adquirido a través de la experiencia, que ejerce una influencia directiva o dinámica sobre las respuestas del individuo*”. Esta definición pone el acento en que las actitudes son disposiciones de comportamiento, por tanto,

---

<sup>14</sup> *Psicología de la personalidad*. Paidós, Buenos Aires

no conductas actuales y, además, predisposiciones habituales que tienen un fundamento fisiológico en conexiones nerviosas determinadas y que se adquieren por la experiencia.

La actitud representa una relación entre un objeto, en esto reside su importancia, y para completar el estudio de su naturaleza, a continuación, nos ocuparemos de analizar los elementos o componentes que las configuran.

### ***Escalas Sumativas***

Este tipo de escala de puntuaciones sumadas está formada por un conjunto de sentencias o proposiciones actitudinales de idéntico valor, a cada una de las cuales los sujetos deben responder, matizando el grado de acuerdo o desacuerdo que personalmente tienen con ellas. Al sumar todas las puntuaciones dadas de todas las preguntas de la escala, se obtiene la puntuación de la actitud del sujeto concreto o de todos los sujetos frente a un determinado tópico.

Estas escalas deben tener las propiedades de un instrumento psicométrico que se caracteriza porque intenta representar el constructo por medio de un puntaje numérico derivado de la aplicación de un conjunto de reactivos o estímulos a los sujetos de interés. En su forma más usual, el instrumento está compuesto por una serie de ítems o preguntas, cada uno de los cuales es calificado o respondido por el individuo de acuerdo con una cierta escala de medición. En muchas ocasiones, las opciones de respuesta se pueden categorizar en dos tipos, correcto o incorrecto, como en el caso de las pruebas de conocimientos y habilidades intelectuales. En otras ocasiones interesa medir más bien sentimientos o dimensiones afectivas como motivación o actitud. En ese contexto generalmente se usan categorías ordinales para las opciones de respuesta (una de las más comunes es la escala Likert de cinco puntos). El puntaje total en el instrumento se genera, sumando las puntuaciones individuales para cada ítem y

es el indicador que representa el nivel que toma el constructo de interés en cada uno de los sujetos estudiados.

Determinar el grado de calidad métrica de un instrumento es de mucha importancia. Al respecto, se puede pensar en las consecuencias derivadas de los resultados de las pruebas estandarizadas de conocimientos que realizan en muchos países los estudiantes en educación básica y media; las pruebas de aptitud académica para ingreso a universidades y los instrumentos psicométricos para medir factores de personalidad que se aplican frecuentemente en la selección de personal para puestos de trabajo.

### ***Algunos Tipos de Escala Likert***

Algunos de las categorías que se emplean en las escala tipo Likert son:

Nunca / Rara vez/ A veces/ A menudo / Siempre  
 Totalmente de Acuerdo/ De acuerdo/ Cerca de 50-50/  
 En desacuerdo /Totalmente en desacuerdo  
 Aprueba decididamente/ Aprueba/Necesita más  
 información/ Desaprueba/ Desaprueba decididamente

A continuación, se presentan algunos ejemplos de escalas tipo Likert más recomendadas.

#### **Escala de 4 puntos**

Esta escala recibe puntajes altos por discriminación y confiabilidad, ejemplo, “Por favor, indique cuán bien la Institución atiende sus requerimientos”

Excede	Los cumple	Casi los cumple	No los aborda
4	3	2	1

### Escala de 5 puntos

También recibe altos puntajes en términos de discriminación y confiabilidad, ejemplo, “Califique el desempeño de su organización”

Significan- temente por sobre	Sobre	Satisface condiciones	Por debajo	Significan- temente por debajo
5	4	3	2	1

### *Ejemplo de Escala Tipo Likert*

La siguiente escala tipo Likert es parte de una batería de escalas para medir el impacto de las nuevas tecnologías como parte del proyecto para medir las actitudes y los docentes hacia las nuevas tecnologías. En nota a pie de página se indica el sitio electrónico donde se puede encontrar más información sobre este proyecto<sup>15</sup>.

Lea cada uno de los enunciados y encierre en un círculo el número correspondiente a la opción que refleja mejor su opinión al respecto.

Totalmente de Acuerdo=TA

De Acuerdo=A

Indeciso=I

En Desacuerdo=D

Totalmente en Desacuerdo =TD

---

<sup>15</sup> Morales, Cesareo, et. al., Actitudes de los estudiantes y docentes hacia la computadora y los medios para el aprendizaje,

<http://investigacion.ilce.edu.mx/dice/proyectos/actitudes/marcoactitudes.htm>

<b>Factor 1 Uso de Correo Electrónico</b>					
<b>No. Reactivo</b>	<b>TD</b>	<b>D</b>	<b>I</b>	<b>A</b>	<b>TA</b>
1. El correo electrónico es un medio efectivo para la divulgación de información para el grupo tareas	1	2	3	4	5
2. Prefiero el correo electrónico a las tradicionales clases informativas como medio de divulgación	1	2	3	4	5
3. Más cursos debieron haber utilizado el correo electrónico para proporcionar información de la clase	1	2	3	4	5
4. El correo electrónico permite un mayor contacto con el asesor	1	2	3	4	5
5. El uso del correo electrónico propicia una mayor interacción entre los estudiantes de cada curso	1	2	3	4	5
6. El uso del correo electrónico propicia una mayor interacción entre estudiante e instructor	1	2	3	4	5
7. El uso del correo electrónico aumenta la motivación para un curso	1	2	3	4	5
8. El uso del correo electrónico hace más interesante un curso	1	2	3	4	5
9. El uso del correo electrónico hace que los estudiantes se sientan más involucrados	1	2	3	4	5
10. El uso del correo electrónico ayuda a los estudiantes a aprender más	1	2	3	4	5
11. El uso del correo electrónico ayuda a proporcionar una mejor experiencia de aprendizaje	1	2	3	4	5





### ***Análisis de Instrumentos: Teoría de Respuesta al Ítem***

Existe en la actualidad un enfoque más poderoso que la Teoría Clásica para el análisis de la calidad técnica de un instrumento de medición psicométrico. Este se llama Teoría de Respuesta a los Ítemes o TRI (“Item Response Theory”, IRT, en inglés).

En cuanto a los requisitos que se deben cumplir para ajustar un modelo de TRI, es necesario mencionar que debe contarse con los datos derivados de la aplicación del instrumento en muestras relativamente grandes de sujetos (como mínimo 200 personas). En general, las respuestas a los ítemes deben ser dicotomizadas, es decir, clasificadas usando los códigos 0 ó 1. Esta codificación dicotómica resulta obvia cuando se están analizando los resultados de pruebas para medir conocimientos o habilidades intelectuales, pues la mayoría de ellas utilizan ítemes de escogencia única (“multiple choice” en inglés). Pero es también posible dicotomizar las respuestas a ítemes que se contestan en una escala ordinal, tipo Likert, por ejemplo, asignando el 1 a las categorías que representen mayor intensidad o nivel en el constructo de interés, y el 0 a las que representen niveles más bajos. Así, por ejemplo, en una prueba psicométrica para medir motivación, se asignará un 1 para las respuestas que indiquen mayores niveles de motivación y 0 para las que indiquen niveles más bajos.

Una de las principales nociones en Teoría de Respuesta a los Ítemes es la Curva Característica del Ítem o CCI. La CCI es el ajuste de una función matemática al comportamiento del ítem, partiendo de las respuestas suministradas por los examinados en el instrumento. Cada punto en la CCI representa la probabilidad de que un sujeto con un nivel específico en el constructo de interés responda correctamente al ítem. La escala que se mide en el eje horizontal, generalmente se ajusta para que, al igual que la distribución normal estándar, tenga un promedio de 0 y una desviación estándar igual a 1. Por tanto, la gran mayoría de sus valores estarán incluidos en el intervalo  $[-3, 3]$ .

## EL ANÁLISIS DE LOS DATOS

### *Características del Análisis de Datos*

La información procesada tiene un valor inestimable: de ella dependerá, por cierto, que puedan o no resolverse las preguntas iniciales formuladas por el investigador. Pero, no obstante, esa información no nos “habla” por sí misma, no es capaz por sí sola de darnos las respuestas deseadas hasta tanto no se efectúe sobre ella un trabajo de análisis e interpretación.

Desde un punto de vista lógico, *analizar* significa descomponer un todo en su partes constitutivas para su más concienzudo examen. La actividad opuesta y complementaria a ésta es la *síntesis*, que consiste en explorar las relaciones entre las partes estudiadas y proceder a reconstruir la totalidad inicial. Lo dicho tiene aplicación directa en la metodología de la investigación: si el objeto es siempre un conjunto coherente, por más que también pueda decirse que es parte de un todo mayor con sus propias leyes y su propia estructura interior, los datos, en tal sentido, no son más que sus elementos integrantes, las partes aisladas que sólo cobran sentido por la síntesis que pueda integrarlos. El procesamiento implica ya un agrupamiento de los mismos en unidades coherentes, pero estas unidades necesitarán de un estudio minucioso de sus significados y de sus relaciones para que luego puedan ser sintetizadas en una globalidad mayor. Estas tareas constituyen, por lo tanto, las últimas y necesarias etapas del trabajo. Resultan fundamentales, por cuanto sin ellas sería imposible encontrar un sentido a toda la labor previamente realizada.

El análisis de los datos no es una tarea que se improvisa, como si recién se comenzara a pensar en él luego de procesar todos los datos. Por el contrario, el análisis surge más del marco teórico trazado que de los datos concretos obtenidos y todo investigador que domine su tema y trabaje con rigurosidad, deberá tener una idea precisa de cuáles serán los lineamientos principales del análisis que habrá de efectuar antes de comenzar a recolectar datos. Se podrá definir así, con suficiente antelación, qué datos

serán capaces de rechazar o afirmar una hipótesis, qué resultados indicarán una u otra conclusión. Esta actividad, llamada por algunos autores *análisis anticipado* es fundamental para evitar sorpresas lamentables, como por ejemplo la de encontrar que no tenemos suficientes datos al final del procesamiento, o de que los que poseemos no nos sirven en realidad para mucho.

Para desarrollar la tarea analítica, hay que tomar cada uno de los datos o conjuntos homogéneos de datos obtenidos, e interrogarnos acerca de su significado, explorándolos y examinándolos mediante todos los métodos conocidos, en un trabajo que para obtener los mejores frutos, debe ser paciente y minucioso. De acuerdo al tipo de datos que se estén analizando, se procederá de un modo u otro, según técnicas y procedimientos que examinaremos en este libro.

### ***El Análisis Cuantitativo***

Este tipo de operación se efectúa, naturalmente, con toda la información numérica resultante de la investigación. Esta, luego del procesamiento que ya se le habrá hecho, se nos presentará como un conjunto de cuadros, tablas y medidas, a las cuales se les han calculado sus porcentajes y presentado convenientemente.

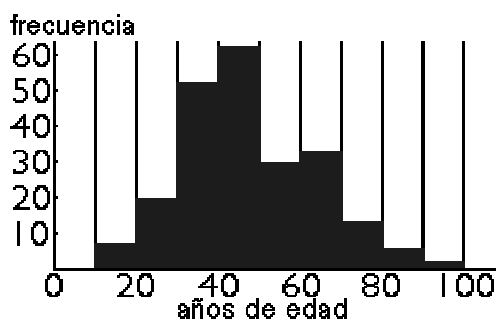
Para cada cuadro que se haya obtenido, será preciso evaluar el comportamiento de las variables que aparezcan en él, precisando la forma en que actúan individualmente. Luego se observarán las relaciones que pueden percibirse entre una y otra variable, si el cuadro es de doble entrada, tratando de precisar la forma en que una afecta a la otra. Si se trata de un cuadro de tres variables, será conveniente examinar primero los valores totales, pues en ellos se ve el funcionamiento global de cada variable operando de modo independiente, para luego pasar a confrontar, por pares, las variables, tratando de detectar las influencias que existan entre ellas.

Siguiendo con el proceso analítico, diremos que es preciso calcular, cuando se trata de datos obtenidos a partir de muestras, los niveles de significación de las diferencias que aparecen entre

porcentajes o entre promedios. Pueden aparecer diferencias entre porcentajes que son inferiores a los errores muestrales, por lo que antes de afirmar alguna conclusión, es preciso conocer los límites dentro de los cuales estas diferencias son estadísticamente significativas. También será necesario, según los casos, calcular determinados coeficientes de correlación, estudiar las variaciones sistemáticas de la información registrada, buscar la estacionalidad y las tendencias que manifiestan las variables y muchas otras operaciones semejantes que permiten sacar conclusiones apropiadas sobre la base de los datos obtenidos.

Cuando ya se hayan hecho todas estas actividades, el análisis de las cifras, en un sentido estricto, habrá concluido. Antes de pasar a la labor de síntesis, es necesario registrar por escrito todos los hallazgos del análisis, para cada cuadro examinado. A partir de ellos es que se podrán extraer conclusiones más generales, por lo que no conviene pasarlos por alto o confiarlos a la memoria. Cabe advertir, para terminar, que no conviene reproducir en palabras *todo* lo que aparece en un cuadro determinado: esa es más una tarea de transcripción verbal que de análisis propiamente dicho y su resultado hace innecesariamente fatigosa la lectura del informe de investigación.

A manera de ejemplo, si se quiere presentar de un modo simple de presentar una distribución de valores es mostrar cada valor como un punto en una escala. Si hay un gran número de valores, puede ser mejor clasificarlos primero y entonces presentar la *frecuencia* de cada clase como un *histograma*.



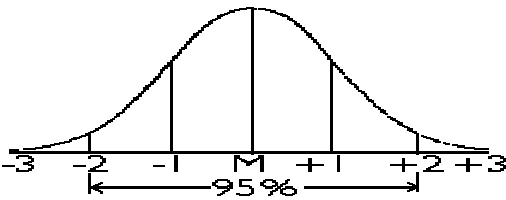


Figura 2.2. Histograma

Si nuestros estudios tienen que ver con personas, ocurrirá con bastante frecuencia que nuestras mediciones estarán distribuidas de acuerdo con la *distribución normal*.

A veces, se quiere poner el énfasis no en la distribución *absoluta*, sino en la *proporcional* o de porcentaje. Un diagrama apropiado para esto es el gráfico de *sectores*.

Por otra parte, el investigador puede estar interesado en examinar la distribuciones desde el punto de vista de las medidas de tendencia central y de las variaciones. El investigador puede elegir habitualmente aquella que muestra mejor el valor típico de la variable. La media aritmética es el más popular, pero puede ofrecer un cuadro equivocado por ejemplo en datos que incluyen un valor que difiere en gran medida de los otros.

A continuación se presentan los estadísticos descriptivos de la variable total autoestima (puntaje total de la escala con la que se midió la autoestima de los sujetos) de nuestro caso de estudio.

Cuadro 2.1. Estadísticos Descriptivos

Estadísticos descriptivos

	N		Asimetría		Curtosis	
	Estadístico	Error típico	Estadístico	Error típico	Estadístico	Error típico
total autoestima	436		-.876	.117	.061	.233
N válido (según lista)	436					

Al examinar los resultados, se informan dos valores relacionados con la *asimetría* y dos con la *curtosis*. Para interpretar la significación de la asimetría, se divide la simetría por el error estándar de la simetría ( $-0,876/0,117 = -7,487$ ). Esto crea un puntaje estandarizado. Es importante calcular este valor como el grado de simetría y es el criterio que se usa para seleccionar la prueba estadística apropiada. Si el valor calculado es menos que  $\pm 2$ , la distribución puede ser interpretada como no sesgada. Esto se puede confirmar, examinando el histograma. La curtosis si es mayor que cero, hay un apuntamiento mayor que la curva normal; si es negativa es más achatada que la normal; si es igual a cero hay el mismo apuntamiento que la normal. SE pueden inerpretar los resultados entregados por SPSS para la curtosis, en forma análoga a los de asimetría.

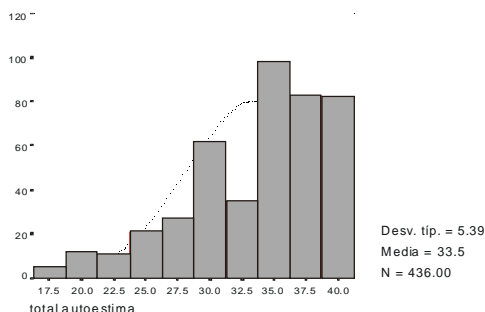


Figura 2.4 Curva sesgada

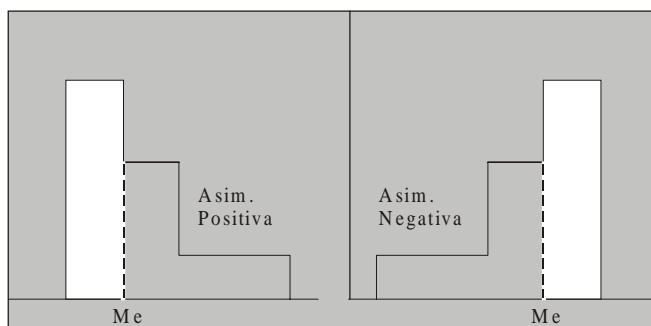


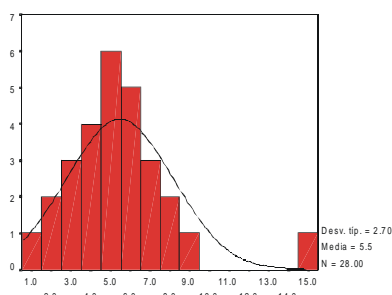
Figura 2.5 Asimetría positiva y negativa

La distribución de la variable que nos interesa comparar también es importante a la hora de establecer la prueba estadística que nos permita evaluar la hipótesis nula de igualdad, de forma que si la distribución es normal y cumple algunas otras condiciones, utilizaremos los denominados *test paramétricos* y si no cumple dichos criterios, tenemos la opción de usar los llamados *test no paramétricos* o de distribución libre, que se basan en los rangos, u otras características de la distribución de la variable. Las pruebas no paramétricas son más exigentes al rechazar la hipótesis nula y por tanto tienen menos posibilidades de acertar cuando no la rechazan (más posibilidades de cometer un error tipo beta). Por otro lado, la potencia aumenta con el tamaño de la muestra; de esta forma, para conseguir la misma potencia con una prueba no paramétrica, bastará con aumentar el tamaño de la muestra en una cantidad determinada. Así, la potencia de una prueba no paramétrica suele calcularse en referencia a su homóloga paramétrica. Por ejemplo, la U de Mann-Whitney tiene una potencia relativa del 95% con respecto a la prueba paramétrica t de Student (significa que con una muestra de 100 sujetos, se consigue la misma potencia con la U que con 95 sujetos en la prueba t).

Una manera de corregir los resultados es eliminar los valores extremos; esta opción es justificable cuando hay uno o dos valores. Otra opción es reemplazar los valores extremos con puntajes más apropiados.



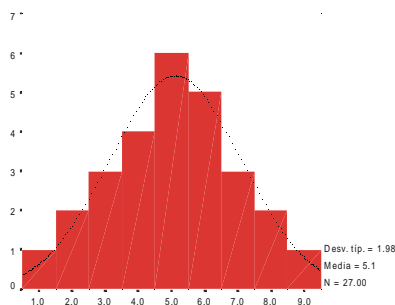
### Ejemplo



Puntaje Original

Se obtuvieron los puntajes (1 a 10) de la aplicación de un test de conocimientos de 28 estudiantes; se detectó un valor extremo y se consideró eliminarlo ya que correspondía a un error en la determinación del puntaje.

Se observa en el histograma un valor extremo que podría provenir de un cálculo erróneo del puntaje total o un error de digitación.



Puntaje corregido

Una vez eliminado el valor extremo, se produce un corrimiento de la media de las observaciones y una disminución en su variabilidad.

La eliminación de valores extremos no siempre es adecuada; en muchos casos, el investigador necesitará

de análisis más exhaustivos de los valores extremos para determinar las causas de éstos y el método de acción a seguir para su análisis.

### ***El Análisis Cualitativo***

Se refiere al que procedemos a hacer con la información de tipo verbal que se ha recogido mediante fichas de uno u otro tipo. Una vez clasificadas éstas, es preciso tomar cada uno de los grupos que hemos así formado para proceder a analizarlos. El análisis se efectúa cotejando los datos que se refieren a un mismo aspecto y tratando de evaluar la fiabilidad de cada información.

Si los datos, al ser comparados, no arrojan ninguna discrepancia seria, y si cubren todos los aspectos previamente requeridos, habrá que tratar de expresar lo que de ellos se infiere, redactando una pequeña nota donde se sinteticen los hallazgos. Si, en cambio, las fichas aportan ideas o datos divergentes, será preciso primeramente determinar, mediante la revisión del material, si se ha cometido algún error en la recolección. Si esto no es así, será necesario ver si la discrepancia se origina en un problema de opiniones o posiciones contrapuestas o si, por el contrario, obedece a alguna manera diferente de categorizar los datos o a errores de las propias fuentes con que estamos trabajando. En todo caso, será conveniente evaluar el grado de confianza que merece cada fuente, teniendo en cuenta su seriedad, sus antecedentes y referencias y toda otra información que pueda resultar de valor al respecto. Ya adoptada una posición frente a las discrepancias encontradas, será el momento de pasar a registrar los hallazgos, nuestras opiniones personales y conclusiones parciales, con lo cual estaremos en condiciones de pasar al siguiente paso, el de la síntesis.

### ***La Síntesis***

Con la síntesis e interpretación final de todos los datos ya analizados, puede decirse que concluye la investigación, aunque teniendo en cuenta que la misma, considerada como intento de obtención de conocimientos, es siempre una tarea inacabada, que debe continuar por fuerza en otras investigaciones concretas.

Sintetizar es recomponer lo que el análisis ha separado, integrar todas las conclusiones y análisis parciales en un conjunto coherente que cobra sentido pleno, precisamente, al integrarse como un todo único. La síntesis es, pues, la conclusión final, el resultado aparentemente simple pero que engloba dentro de sí a todo el cúmulo de apreciaciones que se han venido haciendo a lo largo del trabajo. Las conclusiones finales sólo resultan pertinentes para responder al problema de investigación planteado cuando, en la recolección, procesamiento y análisis de los datos, se han seguido los lineamientos que surgen del marco teórico.

Para alcanzar este resultado, se deben tomar en consideración todas las informaciones analizadas, utilizando para ello las notas ya elaboradas, donde se habrán registrado los hallazgos parciales que hemos hecho. En el caso de cuadros estadísticos, se procederá a comparar los hallazgos de cada cuadro con los otros que tienen relación con el mismo. Así se irá avanzando hacia conclusiones cada vez más generales, menos parciales. Se podrán confeccionar cuadros-resumen, que sinteticen la información más importante que se halla dispersa en otros, para poder presentar un panorama más claro a nuestros lectores. Se procederá, sólo entonces, a extraer las conclusiones finales, que reflejen el comportamiento global de las variables de interés. En función de ellas, redactaremos nuestra síntesis, lo que conviene hacer primero escuetamente, anotando sólo lo esencial. Esta primera síntesis debe ser ordenada y precisa, para lo cual es conveniente numerar nuestras conclusiones correlativamente, teniendo presente el planteamiento inicial del trabajo.

Para el caso de los datos secundarios, se ha de proceder como si se tratara de componer o montar el trabajo general a partir de los elementos parciales de que disponemos. Es una labor eminentemente constructiva, que ha de hacerse parcialmente, para cada punto o capítulo, viendo qué se puede afirmar en cada caso, de qué elementos de apoyo disponemos y cuáles son las conclusiones del caso.

## ***Ejercicios***

### *Datos, variables y escalas de medición*

1. En los siguientes ejemplos, seleccione primeramente la opción que representa el tipo de dato empleado y luego la escala de medición más apropiada para los datos de cada ejemplo:

Tipos de datos

- a. categórico
- b. numérico discreto
- c. numérico continuo

Escalas de medición

- a. nominal
- b. ordinal
- c. intervalar
- d. de razón

1. La edad (en años) de 100 estudiantes
2. Las respuestas de 80 estudiantes egresados de Enseñanza Media que se les preguntó si rindieron la P.A.A.
3. El puntaje obtenido por un grupo de estudiantes en una Prueba de Razonamiento Matemático
4. El CI de un grupo de 30 sujetos
5. El nivel de satisfacción con un trabajo (medido en escala Likert) de 300 empleados
6. El sexo de 45 sujetos
7. El peso de 50 damas de un curso de educación física
8. El Nivel de satisfacción de los estudiantes de sociología con el programa estadístico utilizado
9. Notas obtenidas por un curso de 30 estudiante en la asignatura de de estadística

2. La Universidad desea determinar cuál de dos libros de metodología de la investigación se usará para la formación de pregrado en el área de las ciencias sociales. Para llegar a una decisión se ha decidido seleccionar 80 estudiantes y asignarlos a dos grupos diferentes 40 a cada uno. Cada grupo utilizará un libro diferente. La información que se recogerá de cada estudiante será: sexo, edad (expresada en años), índice de rendimiento académico (expresado con decimales), libro utilizado en clase y puntuación obtenida en la prueba final del curso (0-100, expresada en decimales).

Basándose de esta situación, hay que seleccionar la alternativa correcta en los *ejercicios siguientes*:

1. ¿Qué dos variables son imprescindibles para poder realizar el estudio?
  - a. sexo y edad
  - b. índice de rendimiento académico y sexo
  - c. índice de rendimiento académico y puntuación en la prueba final
  - d. puntuación en la prueba final y libro utilizado en la clase
2. ¿Qué variable es discreta?
  - a. sexo
  - b. índice de rendimiento académico
  - c. puntuación en la prueba final
  - d. edad
3. ¿Cuáles variables son continuas?
  - a. sexo y edad
  - b. índice de rendimiento académico y sexo
  - c. índice de rendimiento académico y puntuación en la prueba final
  - d. puntuación en la prueba final y libro utilizado en la clase

4. ¿Qué variables son categóricas?
- sexo y edad
  - índice de rendimiento académico y sexo
  - índice de rendimiento académico y puntuación en la prueba final
  - sexo y libro utilizado en la clase
5. ¿Qué variable es dependiente?
- sexo
  - índice de rendimiento académico
  - puntuación en la prueba final
  - libro utilizado en la clase
6. Los grupos de estudiantes seleccionados se llaman
- poblaciones
  - muestras
  - estadísticas
  - parámetros
3. A continuación aparece una lista de variables. Utilizando marcas de cotejo (X) indique lo siguiente: si es cualitativa o (N)umérica; si es numérica indica si es (D)iscreta o (Cont)ínua; por último identifica cuál es la escala o nivel de medición de la variable, esto es si es (N)ominal, (O)rdinal, (I)ntervalos o (R)azón:

Variable	Cualitativa	Numérica		Nivel de Medición			
		D	C	N	O	I	R
Años de servicio en la institución							
Prestigio profesional Grado de satisfacción de una persona							
Puntaje obtenido en la pruebas de habilidades matemáticas Nivel socioeconómico							
Máximo de instrucción alcanzado Puntaje obtenido en una escala para medir Liderazgo							
El coeficiente de inteligencia de una persona(CI)							

4. Elabore el plan de preguntas y análisis estadístico para los datos del ejemplo del Libro de Códigos del Capítulo I
5. Use esos datos para construir una base de datos y ejecutar los procedimientos estadísticos con SPSS (descripción, resumen, etc.)





## CAPÍTULO III

---

### EL CASO DE ESTUDIO

#### LA ESCALA DE AUTOESTIMA

En este capítulo se presenta un caso de estudio basado en la aplicación de varias escalas (autoestima, sentimientos negativos, sentimientos positivos, percepción del estrés, nivel de satisfacción de vida, etc.), cuyos resultados se han procesado en SPSS y servirán como base para los análisis estadísticos que se abordan en este libro. Una de las escalas es la *Escala de Autoestima de Rosenberg*.

La autoestima es la valoración que tiene una persona acerca de sí misma y, en mayor o menor medida, esta valoración estará menos o más ajustada a la realidad. De esta forma, la autoestima puede implicar una valoración aproximadamente realista, una sobrevaloración o una infravaloración de sí mismo.

Más específicamente, se debe considerar a la autoestima como una actitud de autovaloración que la persona adopta hacia sí misma. Si se la considera como una actitud, la autoestima abarcará tres manifestaciones o dimensiones distintas: cognitiva, afectiva y conductual. Precisamente son estos tres aspectos los que deberán considerarse en el momento de evaluarla.

Existen diferentes pruebas para evaluar la autoestima, como por ejemplo la Escala de Autoestima de Rosenberg (1965). Ésta es una escala tipo Likert, compuesta por 10 ítemes, los que tienen cuatro alternativas de respuesta, desde totalmente de acuerdo a totalmente en desacuerdo.

Los datos provienen de un estudio que involucró a 439 sujetos. Se contemplaron las siguientes variables de clasificación: *sexo*, *edad*, *estado civil* y *nivel educacional*. Otras variables consideradas fueron: fuente de estrés, fumador/no fumador, los ítems y el puntaje total obtenido en una escala de optimismo, de sentimiento positivo y sentimiento negativo. La edad de los sujetos fue dividida en 3 grupos: 18 a 29, 30 a 44, y 45 y más.

En la *Escala de Autoestima* se asignan los valores siguientes a los ítems:

- Para los ítems 1,2,4,6,7: Fuertemente de acuerdo = 3, De acuerdo = 2, En desacuerdo = 1, y Fuertemente en desacuerdo = 0.
- Para los ítems 3,5,8,9,10 (que tienen una valoración inversa, y se destacan con \*\*): Fuertemente de acuerdo = 0, De acuerdo = 1, En desacuerdo = 2, y Fuertemente en desacuerdo = 3.

La escala tiene una puntuación que va desde 0 a un máximo de 30 puntos. En estas escalas otros puntajes son posibles. Por ejemplo, se puede asignar valores 1-4 puntos en lugar de 0-3 puntos. Algunos investigadores usan una escala Likert de 5 ó 7 puntos.

Este caso de estudio servirá de base para los diferentes análisis univariados y bivariados que se exponen en los capítulos siguientes. Además nos servirá para el análisis de confiabilidad y análisis factorial de la escala.

La Base de Datos del caso de estudio se puede solicitar al correo: [alberto.carro@uv.cl](mailto:alberto.carro@uv.cl)

Tabla No. 3.1. Escala de autoestima de Rosenberg

	<b>ÍTEMES</b>	<b>1 Totalmente en acuerdo</b>	<b>2 De acuerdo</b>	<b>3 En desacuerdo</b>	<b>4 Totalmente en desacuerdo</b>
1.	Siento que soy una persona que valgo, al menos en un plano similar con otros.	TA	A	D	TD
2.	Siento que tengo buenas cualidades.	TA	A	D	TD
3.	Me inclino a pensar, al final, que soy un fracaso.**	TA	A	D	TD
4.	Soy capaz de hacer las cosas tan bien como otras personas.	TA	A	D	TD
5.	Siento que no tengo mucho de que enorgullecerme.**	TA	A	D	TD
6.	Tengo una actitud positiva hacia mi mismo.	TA	A	D	TD
7.	En general, estoy satisfecho conmigo mismo.	TA	A	D	TD
8.	Yo quisiera tener más respeto por mi mismo.**				
9.	A veces me siento inútil.**	TA	A	D	TD
10.	A veces pienso que en general no soy bueno (a).**	TA	A	D	TD

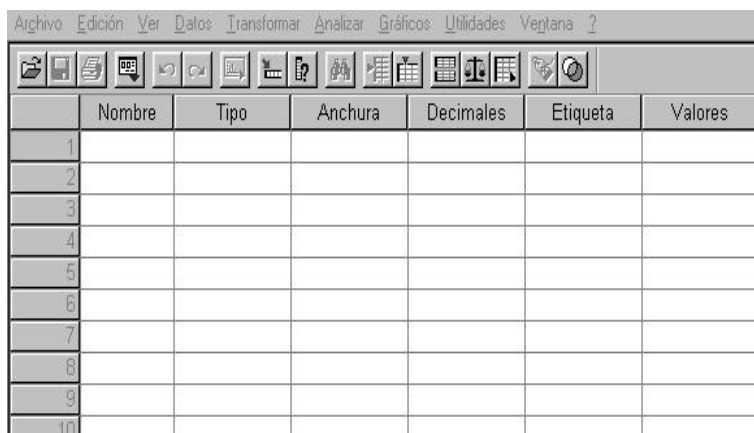
A continuación, se incluye un ejemplo abreviado del libro de códigos con algunas de las variables que se consideraron en el caso de estudio. Incluiremos las más relevantes para el tratamiento de los datos.

Tabla No.3.2. Libro de Códigos

<b>Código</b>	<b>Descripción</b>	<b>Tipo</b>	<b>Escala de Medición</b>	<b>Plan de Análisis 1</b>	<b>Plan de Análisis 2</b>
Edad	Edad del encuestado	Cuantitativa	Escala	Estadística descriptiva	
Fuente	La situación o agentes de estrés	Cualitativa, categórica	Nominal	Tabla de Frecuencia	
Edadgp3	Grupo de edades (3)	Cualitativa	Ordinal	Tabla de Frecuencia, algunos estadísticos descriptivos	
Totautoe	Puntaje total obtenido en la escala de autoestima	Cuantitativa	Escala	Estadística descriptiva	Edadgp3 /Totautoe ANOVA
Tsatvid	Puntaje total obtenido en la escala de satisfacción de vida	Cuantitativa	Escala descriptiva	Estadística	Sexo /Tsatvid t Student Muestra independientes

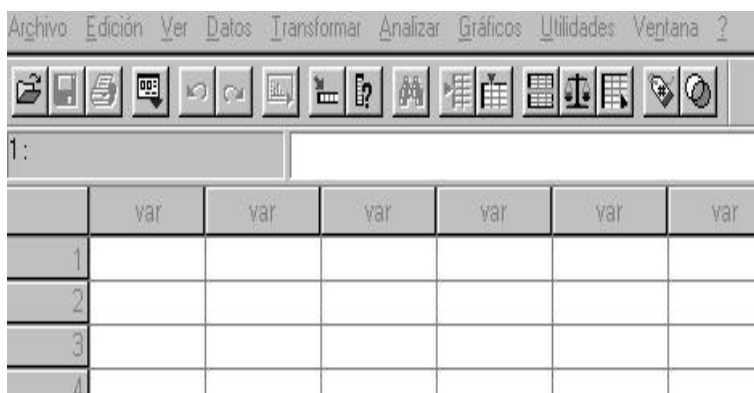
El caso de estudio al que estamos aludiendo contiene 93 variables, de allí que hayamos optado por presentar una versión abreviada de esta herramienta analítica.

El ingreso de los datos a SPSS versión 10 se realiza a través de dos pantallas. La primera, que aparece en la figura 3.1, es para ingresar las variables y sus correspondientes definiciones: nombre, tipo, valores, escala de medición, etc. En la figura 3.2 aparece parte de las variables de nuestro caso de estudio.



	Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						

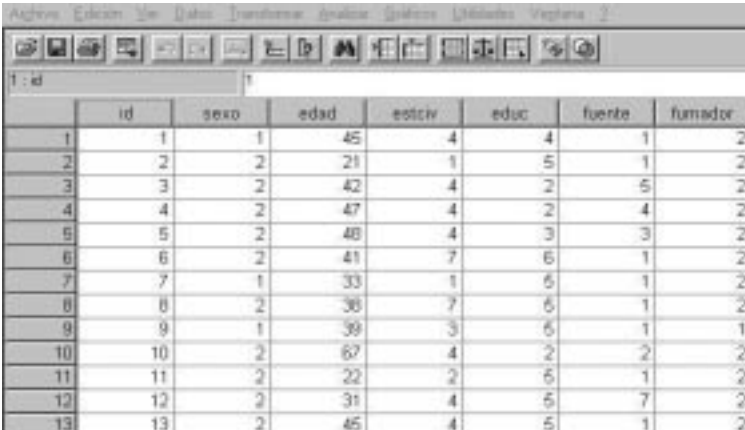
Figura 3.1 Vista de Variables



	var	var	var	var	var	var
1						
2						
3						
4						

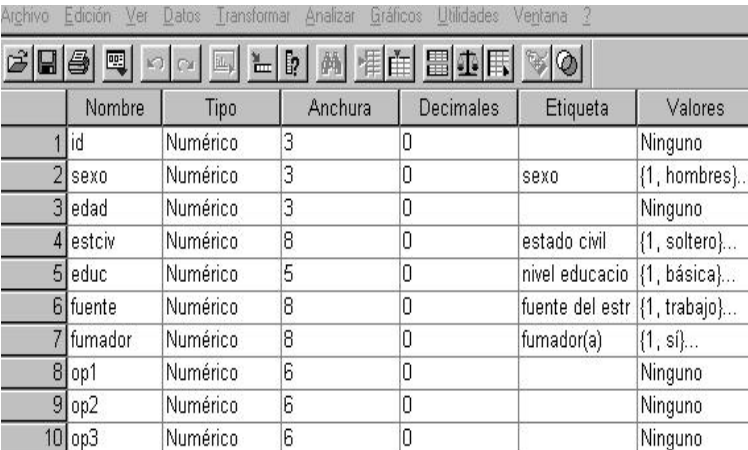
Figura 3.2 Visión parcial de las variables de nuestro caso de estudio

La segunda corresponde al área de ingreso de los datos, según aparece en la figura 3.3. En esta matriz se digitan los valores, como se puede observar en la figura 3.5



	id	sexo	edad	estciv	educ	fuentes	fumador
1	1	1	45	4	4	1	2
2	2	2	21	1	5	1	2
3	3	2	42	4	2	5	2
4	4	2	47	4	2	4	2
5	5	2	48	4	3	3	2
6	6	2	41	7	6	1	2
7	7	1	33	1	6	1	2
8	8	2	36	7	6	1	2
9	9	1	39	3	6	1	1
10	10	2	62	4	2	2	2
11	11	2	22	2	6	1	2
12	12	2	31	4	6	7	2
13	13	2	45	4	5	1	2

Figura 3.3 Área de ingreso de los datos



	Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores
1	id	Númérico	3	0		Ninguno
2	sexo	Númérico	3	0	sexo	{1, hombres}...
3	edad	Númérico	3	0		Ninguno
4	estciv	Númérico	8	0	estado civil	{1, soltero}...
5	educ	Númérico	5	0	nivel educacio	{1, básica}...
6	fuentes	Númérico	8	0	fuentes del estr	{1, trabajo}...
7	fumador	Númérico	8	0	fumador(a)	{1, sí}...
8	op1	Númérico	6	0		Ninguno
9	op2	Númérico	6	0		Ninguno
10	op3	Númérico	6	0		Ninguno

Figura 3.4 Muestra de los datos de nuestro caso de estudio

Los códigos que se asignan a las diferentes variables se definen en *Vista de variables*, columna valores:

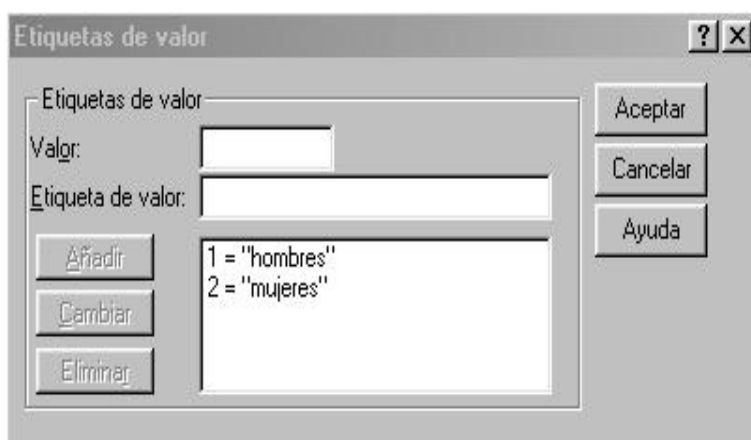


Figura 3.5 Cuadro de diálogo para definición de los códigos de la variable sexo

La definición de las escalas de medición también se definen en la *Vista de Variables*, columna Medida:

Columnas	Alineación	Medida
8	Derecha	Escala
8	Derecha	Nominal
8	Derecha	Escala
8	Derecha	Nominal
8	Derecha	Ordinal
8	Derecha	Nominal
8	Derecha	Nominal
8	Derecha	Escala

Figura 3.6. Columna de defición de las escalas de medición

En relación a nuestro caso de estudio, nos hemos formulado algunas de las siguientes interrogantes:

- (a) ¿Cuál es la edad promedio de la muestra? ¿Cuál es el rango de edad de la muestra (valores mínimo y máximo)?
- (b) ¿Qué porcentaje de la muestra son fumadores?
- (c) ¿Son normales las distribuciones de la edad y de los puntajes totales de cada una de las escalas empleadas en el estudio?
- (d) ¿Son confiables las diferentes escalas (autoestima, sentimientos positivos, sentimientos negativos, percepción del estrés, nivel de satisfacción de vida) empleadas en el estudio?
- (e) ¿Existen diferencias de género estadísticamente significativas en la percepción de la autoestima?.
- (f) ¿Existe una relación estadísticamente significativa entre la autoestima (medida por la escala anterior) y la percepción del estrés (medida por una escala específica)?
- (g) ¿Cuál es el efecto del sexo y la edad en los puntajes de autoestima?
- (h) ¿Hay un efecto de la edad que pueda ser una variable que puede estar afectando los resultados? ¿cómo puede controlarla en su análisis?
- (i) ¿Cuál es mejor predictor de los sentimientos negativos: el optimismo o la autoestima?
- (j) ¿Son las personas más jóvenes (18-29 años) más proclives a ser fumadores que las personas mayores (30-44 años o 45 y más)?
- (k) ¿Cuál es la estructura factorial de la escala de autoestima?
- (l) ¿Quiénes tienen mayor autoestima: los hombres o las mujeres?
- (m) ¿Qué ocurre con la percepción del estrés en los diferentes grupos de edades?



Algunas de las hipótesis que hemos establecido a partir de la interrogantes anteriores son por ejemplo:

1. A menor autoestima, más alta es la percepción del estrés.
2. Existe una relación lineal entre la autoestima y la percepción del estrés.
3. La autoestima es un buen predictor de los sentimientos negativos de las personas.
4. Las personas más jóvenes tienden a ser más fumadoras que las personas mayores.
5. Las personas optimistas tienen un mayor nivel de satisfacción de vida
6. Los hombres son más fumadores que las mujeres. Y además presentan un mayor nivel de estrés.

En orden a examinar estas hipótesis deberemos aplicar diferentes pruebas estadísticas que iremos comentando en el desarrollo de los siguientes capítulos. Por ejemplo, pruebas de normalidad (Kolmogorov Smirnov), de correlaciones, de análisis de regresión lineal, de comparación de medias, análisis de varianza, análisis factorial, etc.

Por ejemplo, se puede establecer una matriz de correlaciones para verificar las intercorrelaciones entre las siguientes variables:

- (a) edad
- (b) percepción del estrés (ttestres)
- (c) sentimiento positivos (ttspos)
- (d) sentimiento negativos (ttsneg)
- (e) satisfacción de vida (tsatv)

El investigador puede explorar el impacto de numerosas variables en la experiencia de sentimientos positivos (ttspos), sentimiento negativos (ttsneg), y estrés percibido (ttestres). En particular, puede estar interesado en la edad, autoestima, y optimismo.

También puede compararse el patrón de correlaciones en mujeres y hombres dividiendo el archivo.

## VALIDACIÓN DE LA ESCALA

Los objetivos específicos son:

- a) Estimar la confiabilidad del instrumento.
- b) Estudiar la validez de constructo del instrumento (estructura factorial).

### *Confiabilidad de la Escala*

El estudio psicométrico de la escala de autoestima permite estimar la *confiabilidad* del instrumento en 0,8541, mediante Alfa de Cronbach. Es decir, el instrumento tiene consistencia interna, es confiable. El coeficiente de discriminación, esto es, el grado en que un ítem sirve para distinguir entre los individuos que obtienen puntuaciones altas y los que las obtienen bajas, varía entre 0,44 y 0,71 (ver detalles de los resultados del análisis de confiabilidad en el capítulo VII). Lo anterior ratifica la idea de la consistencia del instrumento.

El análisis de la confiabilidad también informa sobre el efecto de cada ítem, señalando la confiabilidad si el ítem es eliminado. Se concluye que no hay ítems que tengan un gran efecto sobre la escala, pues la confiabilidad, eliminando un ítem, varía entre 0,8280 (para el ítem SEST6) y 0,8458 (para el ítem SEST1). Tampoco hay un gran efecto sobre la media y la varianza de la escala si se elimina un ítem. La media varía entre 29,83 (si se elimina el ítem SEST1) y 30,54 (si se elimina el ítem SEST7). La varianza varía entre 23,09 (si se elimina el ítem SEST9) y 26,83 (si se elimina el ítem SEST2). En general, no se aprecia

algún ítem que consistentemente influya en la escala.

Finalmente, el alfa de Cronbach de 0,8541 resultó ser significativo, al nivel de significación del 5%.

### ***Validez Factorial de la Escala***

La *validez de constructo*, estudiada a través del análisis factorial con rotación Varimax, permiten identificar la existencia de *dos factores*, que explican el 57,1 % de la variabilidad total. En el primer factor (que explica el 45,4% de la variabilidad) influyen fuertemente los ítems SEST5, SEST7, SEST9 y SEST10, que se asocian con la *escala de sentimiento nrgativo*. El segundo factor está influenciado por los ítems SEST1, SEST2, SEST4, los que se asocian con la *escala de sentimiento positivo*.

En resumen, los resultados obtenidos han mostrado que la escala de autoestima tiene una buena consistencia interna y una adecuada validez de constructo.

El estudio de la estructura factorial ha dado cuenta que el modelo bifactorial permitiría no sólo una puntuación global de la autoestima, si no que también puntuaciones en las dos subescalas, identificadas como «cargas positivas” y ”cargas negativas», incrementando así su valor evaluativo.

### **EJEMPLO DE HIPÓTESIS**

En relación a este caso de estudio, se plantearon diversas hipótesis y se aplicaron las pruebas estadísticas respectivas en SPSS. A modo de ejemplo examinaremos la siguiente hipótesis de investigación: ¿Existen diferencias de género estadísticamente significativa en la percepción de la autoestima? se traduce en las siguientes hipótesis estadísticas.

Hipótesis Nula:  $m_H = m_M$  (Los promedios de puntajes en la

Escala de Autoestima de Rosenberg para hombres y mujeres son iguales)

Hipótesis Alternativa:  $m_H \neq m_M$  (Los promedios de puntajes en la Escala de Autoestima de Rosenberg para hombres y mujeres no son iguales).

Para probar esta hipótesis se empleará la prueba de diferencia de medias para muestras independientes. (Se puede suponer varianzas iguales y desconocidas dado que se trabaja con la misma escala de puntajes. Ver Capítulo V).

La aplicación de la prueba entrega los resultados siguientes:

Prueba de muestras independientes

	Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias							
	F	Sig.	t	gl.	. (bilateral)	media diferencial	error típ. diferencial	% Intervalo de confianza para la diferencia		
								inferior	superior	
total autoestima	1.506	.062	1.622	434	.105	.85	.52	-.18	1.87	
Se han asumido varianzas iguales										
No se han asumido varianzas iguales			1.661	2.349	.098	.85	.51	-.16	1.85	

Se comprueba el supuesto de igualdad de varianzas con la prueba de Levene de lo cual se concluye lo siguiente: No se rechaza la hipótesis de igualdad de varianzas a un nivel de significancia del 5% (valor  $p = 0,062$ ). Con esto se asume varianzas iguales de los puntajes logrados por hombres y mujeres en la escala de autoestima.

No se rechaza la hipótesis de igualdad de medias de los puntajes de la escala de Autoestima para hombres y mujeres, a un nivel de significación del 5% (valor  $p=0,105$ ). Por lo tanto, existe evidencia para concluir que las medias de los puntajes en la escala de Autoestima para hombres y mujeres son iguales.

## CAPÍTULO IV

---

### ANÁLISIS DESCRIPTIVO UNIVARIANTE

Cuando se dispone de datos de una población, y antes de abordar análisis estadísticos más complejos, un primer paso consiste en presentar esa información de forma que ésta se pueda visualizar de una manera más sistemática y resumida. Los datos que nos interesan, dependen, en cada caso, del tipo de variables que estemos manejando. Esto se realiza empleando gráficos o mediante la obtención de medidas descriptivas de resumen de la información. Los procedimientos más importantes, implementados en SPSS, se presentan a continuación.

#### EL ANÁLISIS GRÁFICO

##### *Gráficos de Sectores y de Barras*

*Para variables categóricas*, como el sexo, profesión, etc., se quiere conocer la frecuencia y el porcentaje del total de casos que *caen* en cada categoría. Una forma muy sencilla de representar gráficamente estos resultados, es mediante diagramas de barras o diagramas de sectores (más conocido como diagrama de *torta*).

En los gráficos de sectores, se divide un círculo en tantas porciones como clases tenga la variable, de modo que a cada clase le corresponde un sector del círculo proporcional a su frecuencia absoluta o relativa. Si el número de categorías es excesivamente grande, la imagen proporcionada por el gráfico de sectores no es lo suficientemente clara y por lo tanto la situación ideal es cuando

hay alrededor de tres categorías. En este caso, se pueden apreciar con claridad dichos subgrupos.

Los diagramas de barras son similares a los gráficos de sectores. Se representan tantas barras como categorías tiene la variable, de modo que la altura de cada una de ellas sea proporcional a la frecuencia o porcentaje de casos en cada clase. Estos mismos gráficos pueden utilizarse también para describir *variables numéricas discretas* que toman pocos valores (por ejemplo, número de hijos).

### ***Gráfico de Tallo y Hojas***

Un método para iniciar el análisis exploratorio de los datos, previo al uso de los métodos estadísticos tradicionales, y que además proporciona información rápida, visual y es relativamente nueva, es la *representación gráfica de tallo y hoja*. Esta representación se basa en la ordenación de los datos a manera de gráfico, pero sin llegar a ello, utilizando las decenas y las unidades.

Esta técnica se puede encontrar en el libro de Freud y Simon<sup>16</sup>, pero comentaremos su uso a través del siguiente ejemplo que contiene las calificaciones obtenidas en una prueba de matemáticas:

78	93	61	100	70	83	88	74	97	72
66	73	76	81	83	64	91	70	77	86

Ahora pensaremos en cada uno de los datos, separando las decenas de las unidades, es decir, el número 51 se verá como 5 | 1. De esta manera las decenas se pondrán en una columna, en forma vertical, y las unidades a su derecha:

---

<sup>16</sup> Freund, John E. y Gary A. Simon. *Estadística elemental*. Prentice-Hall Hispanoamericana, SA. México, 1994. (8ª edición.)

---

6	1	6	4					
7	8	0	4	2	3	6	0	7
8	1	3	3	6	8			
9	1	3	7					
10	0							

Cuadro 4.1. Diagrama de Tallo y Hoja.

Para entenderle un poco más, hay que señalar que el primer renglón que dice 6 | 1 6 4 quiere decir que entre la lista de datos se encuentran los valores 61, 66 y 64.

Esta es la representación gráfica tallo y hoja, donde cada renglón es una posición de tronco y cada dígito de la derecha es una hoja. El procedimiento para realizarla es primero empezar con los troncos, es decir la columna de la izquierda, y después dato por dato ir llenando las hojas a la derecha de la línea vertical, en el tronco correspondiente.

Además, si se desean tener los datos ordenados, y hay gente que lo prefiere así, se pueden ordenar las hojas en cada renglón para que la representación quede como sigue:

6	1	4	6					
7	0	0	2	3	4	6	7	8
8	1	3	3	6	8			
9	1	3	7					
10	0							

Cuadro 4.2 Datos ordenados.

En realidad, una representación de tallo y hojas presenta la misma información que la lista original de datos, pero de una manera mucho más compacta (especialmente si la lista de datos es más grande) y manejable. Sin embargo, una información más compleja resulta un poco más difícil de manejar, por lo que en

ocasiones conviene redondear los datos, ignorar sus partes decimales o utilizar las centenas u otras posiciones de los números para las troncos. En cada uno de esos casos, conviene hacer alguna anotación, o poner una nota, a fin de que los lectores puedan identificar las adecuaciones realizadas y así poder interpretar lo que se quiere transmitir.

Para mostrar la información de manera más clara, es posible modificar el número de posiciones del tallo, aumentándola o disminuyéndola de acuerdo a las necesidades particulares de cada problema. Por ejemplo, con los datos anteriores, se pueden dividir en dos, cada posición del tallo, utilizando la primera posición para disponer las hojas 0, 1, 2, 3 y 4, y la segunda posición para las hojas restantes. De esta manera, se obtiene la *representación gráfica de doble tallo*:

6-	1	4
6+	6	
7-	0	0 2 3 4
7+	6	7 8
8-	1	3 3
8+	6	8
9-	1	3
9+	7	
10-	0	

Cuadro 4.3. Representación gráfica de doble tronco.

Con esto se han duplicado el número de posiciones del tallo, con la intención de buscar una mayor claridad en la presentación.

### ***Histograma***

Para *variables numéricas continuas*, tales como la edad, el puntaje en un pretest o en un postest, el tipo de gráfico más utilizado es el histograma. Para construir un gráfico de este tipo, se divide el rango de valores de la variable en intervalos de igual amplitud,



representando sobre cada intervalo, un rectángulo que tiene a este segmento como base. El criterio para calcular la altura de cada rectángulo es el de mantener la proporcionalidad entre las frecuencias absolutas (o relativas) de los datos en cada intervalo y el área de los rectángulos.

### ***Diagrama de Caja y Gráficos para Normalidad***

Otro modo habitual, y muy útil, de resumir una variable de tipo numérico es utilizando el concepto de percentiles, mediante *diagramas de cajas*. La caja central indica el rango en el que se concentra el 50% central de los datos. Sus extremos son, por lo tanto, el 1<sup>er</sup> y 3<sup>er</sup> cuartil de la distribución. La línea central en la caja es la mediana. De este modo, si la variable es simétrica, dicha línea se encontrará en el centro de la caja. Los extremos de los “bigotes” que salen de la caja son los valores que delimitan el 95% central de los datos, aunque en ocasiones coinciden con los valores extremos de la distribución. Se suelen también representar aquellas observaciones que caen fuera de este rango (outliers o valores extremos). Esto resulta especialmente útil para comprobar, gráficamente, posibles errores en nuestros datos. En general, los diagramas de cajas resultan más apropiados para representar *variables que presenten una gran desviación de la distribución normal*.

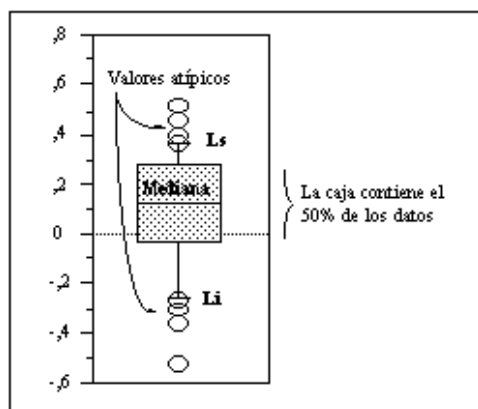


Figura 4.1. Diagrama de caja.

Por último, y en lo que respecta a la descripción de los datos, suele ser necesario, para posteriores análisis, comprobar la normalidad de alguna de las variables numéricas de las que se dispone. Un diagrama de cajas o un histograma son gráficos sencillos que permiten comprobar, de un modo puramente visual, la simetría y el «apuntamiento» de la distribución de una variable y, por lo tanto, valorar su desviación de la normalidad. Existen otros métodos gráficos específicos para este propósito, como son *los gráficos P-P o Q-Q*. En los primeros, se confrontan las proporciones acumuladas de una variable con las de una distribución normal. Si la variable seleccionada coincide con la distribución de prueba, los puntos se concentran en torno a una línea recta. Los gráficos Q-Q se obtienen de modo análogo, esta vez representando los cuantiles de distribución de la variable respecto a los cuantiles de la distribución normal.

### Gráfico Q-Q normal de EDAD

Para SEXO= hombres

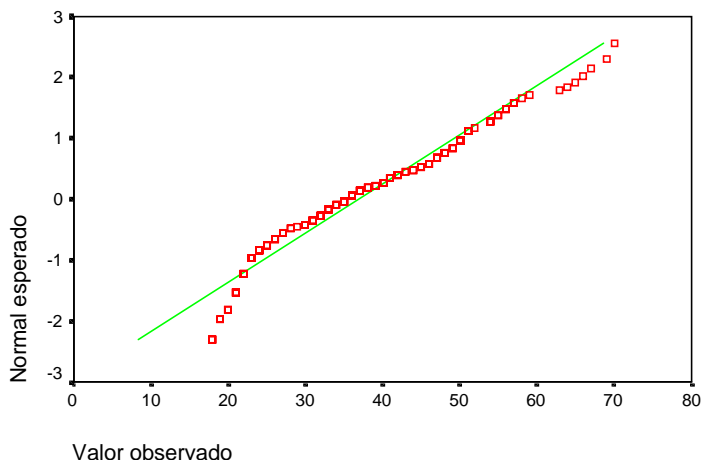


Gráfico 4.1. Gráfico Q-Q.

## EL ANÁLISIS DESCRIPTIVO

Para variables numéricas, en las que puede haber un gran número de valores observados distintos, se intentan las siguientes preguntas:

- a. ¿Alrededor de qué valor se agrupan los datos?
- b. Supuesto que se agrupan alrededor de un número, ¿Cómo lo hacen? ¿Muy concentrados? ¿Muy dispersos?

La exploración y análisis de los datos obtenidos en una investigación comienza necesariamente en el nivel univariado. El análisis univariado de los datos consiste en hacer una revisión del valor de un solo indicador, por ejemplo, la edad, nivel de instrucción o el sexo de los sujetos. Suele usarse cuando interesa conocer en profundidad qué significan, por separado, algunos indicadores y es utilizado con fines descriptivos en las primeras etapas de los estudios realizados. El análisis de los datos tiene como objetivo el responder a las preguntas que se hicieron los investigadores, pero para llegar a ese punto, primero se deben describir las variables o datos que se recogieron durante el estudio.

Este análisis permitirá controlar la presencia de posibles errores en la fase de introducción de los datos, es decir, detectaremos con él, valores fuera de rango (p. ej. un peso de 197 Kg es claramente un dato extraño), o la presencia de valores perdidos. Este análisis inicial también proporcionará una idea de la *forma* que tienen los datos: su posible distribución de probabilidad con sus parámetros de centralización; media, mediana y moda; así como sus parámetros de dispersión; varianza, desviación típica, etc.

Como se verá más adelante, SPSS proporciona varias herramientas para realizar esta descripción, todas ellas bajo el menú *Analizar* y, dentro de él, la opción *Estadísticos descriptivos*. Estos procedimientos son: Frecuencias, Descriptivos, Explorar y

Tablas de Contingencia. En la siguiente pantalla se muestra la ventana para el cálculo de las frecuencias.

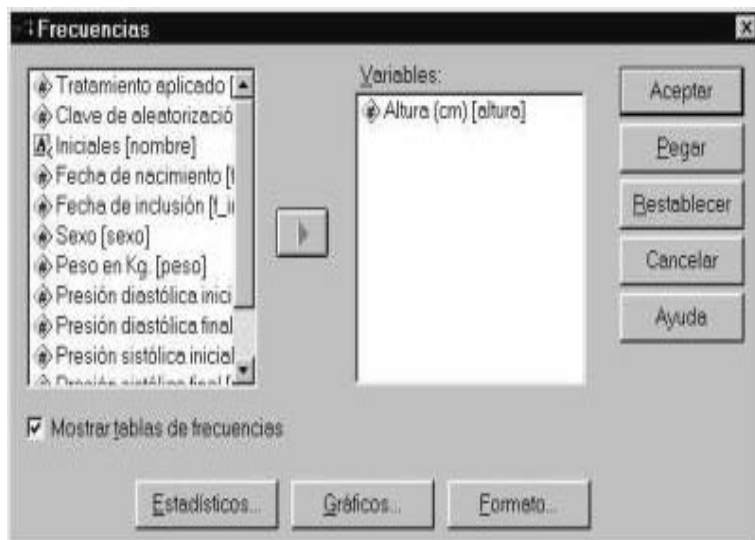


Figura 4.2. Definición de Tabla de Frecuencia.

Hay un error frecuente que es equiparar estadística univariada y estadística descriptiva, si se puede considerar que en el análisis univariado, hay un nivel descriptivo y un nivel inferencial. El primero incluye el resumen de los datos mediante estadísticos descriptivos y la representación gráfica de los mismos. El segundo implica el contraste de hipótesis acerca del valor de los parámetros y del tipo de distribución poblacional (análisis inferencial). Si tenemos una muestra aleatoria de 1.000 estudiantes universitarios entre 17 y 26 años, la interpretación descriptiva dirá por ejemplo que al 58% le gusta la vida nocturna. Se pasa al nivel inferencial si se dice que estos datos son compatibles con los resultados de una investigación según los cuales, el 60% de los jóvenes universitarios de esa edad les gusta la vida nocturna.

Tablas 4.1. Análisis univariado descriptivo e inferencial y el nivel de medición de las variables.

Tipo de Variable	Estadísticos descriptivos	Representación gráfica
Categorica	Frecuencia y porcentajes	Gráfico de barras o sectores
Ordinal	o sectores percentiles	Gráfico de barras
Cuantitativa (*)	Media, desviación estándar, índices de asimetría y curtosis	Histograma

(\*) Si la distribución es muy asimétrica, es aconsejable tratar la variable como ordinal.

Algunos de los contrastes de hipótesis más utilizados en función del tipo de variables son:

Tabla 4.2. Tipos de variables y tipos de contrastes

Tipo de Variable	Contrastes de bondad de ajuste	Contrastes paramétricos
Categorica	Chi Cuadrado	
Ordinal	Chi Cuadrado	
Cuantitativa (*)	Prueba de Kolmogorov-Smirnov para poner a prueba la hipótesis de distribución normal.	Acerca de los valores de la media y varianza poblacional

Pero, además, podrían incluirse muchos otros contrastes o pruebas: proporción, correlaciones, etc.

### ***Análisis Descriptivo con SPSS***

El primer paso en el análisis de cualquier archivo de datos es la obtención de la estadística descriptiva para cada variable. Esto puede usarse también para limpiar los datos (eliminar datos fuera de rango), explorar la distribución de los puntajes, y describir la muestra en estudio, por ejemplo, se pueden examinar preguntas del siguiente tipo: ¿Cuál es el promedio de edad de la muestra?, ¿Cuál es el rango de edad de la muestra (valores mínimos y máximos)?, ¿Cuál es el porcentaje de hombres y mujeres en la muestra?, ¿Hay sujetos que no haya respondido esta pregunta?, ¿Cuán normal es la distribución?.

Gran parte de la utilidad que tiene la Estadística Descriptiva es la de proporcionar un medio para informar basado en los datos recopilados. La eficacia con que se pueda realizar tal proceso de información dependerá de la presentación de los datos, siendo la forma gráfica uno de los más rápidos y eficientes, aunque también uno de los que más pueden ser manipulados o ser malinterpretados si no se tienen algunas precauciones básicas al realizar las gráficas. Existen también varios *tipos de gráficas*, o *representaciones gráficas*, utilizándose cada uno de ellos de acuerdo al tipo de información que se está usando y los objetivos que se persiguen al presentar la información.

En SPSS, el procedimiento *Explorar* nos ofrece las posibilidades de representar gráficamente los datos, examinar visualmente las distribuciones para varios grupos de datos, y realizar pruebas de normalidad y homogeneidad sobre los mismos. En este caso, estamos interesados en examinar la variable cuantitativa edad (dependiente) y la variable categórica sexo (factor) como se puede apreciar en la figura siguiente:

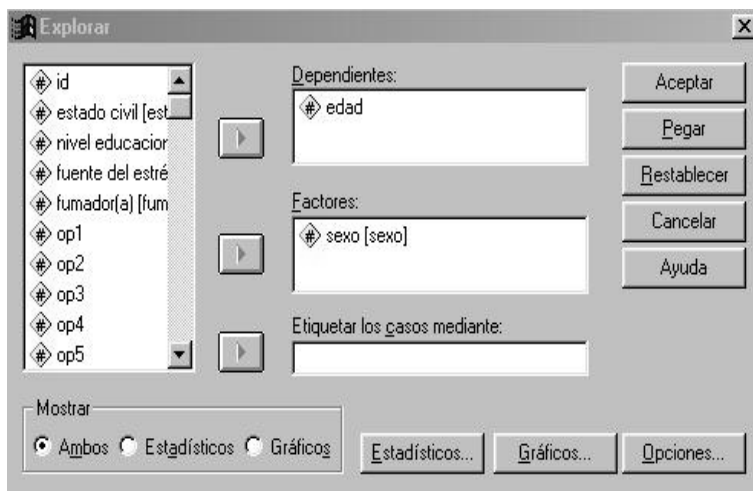


Figura 4.3. Definición de variables.

Cuando se activa la opción Estadísticos se abre un cuadro de diálogo con los siguientes campos y opciones:

*Dependientes:* en este campo, se introduce la variable o variables cuantitativas que contienen los datos a analizar y de las que se quiere obtener los gráficos y/o estadísticos.

*Factores:* aquí se indica la variable que sirve para dividir en grupos los datos. Puede ser numérica o carácter de cadena corta.

*Etiquetar los casos mediante:* este otro campo se utiliza para etiquetar aquellos valores atípicos en los diagramas de caja. La variable que se usa puede ser numérica o carácter.

*Gráficos:* activando este campo y pulsando el botón Gráficos, se abre una ventana para seleccionar el gráfico a representar. Entre éstos se encuentran: Diagramas de Caja; Descriptivos (Gráficos de Tallo y Hoja, e Histogramas); Gráficos con prueba de Normalidad; y Gráficos de Dispersión por nivel con Prueba de Levene.

*Estadísticos:* con este campo y su respectiva opción, se abre una ventana donde se pueden elegir una serie de medidas de tendencia central y de dispersión. Así como una serie de estimadores robustos centrales, valores atípicos de los datos y percentiles.

*Ambos:* en este punto se pueden seleccionar las opciones tanto de Gráficos como de Estadísticos.

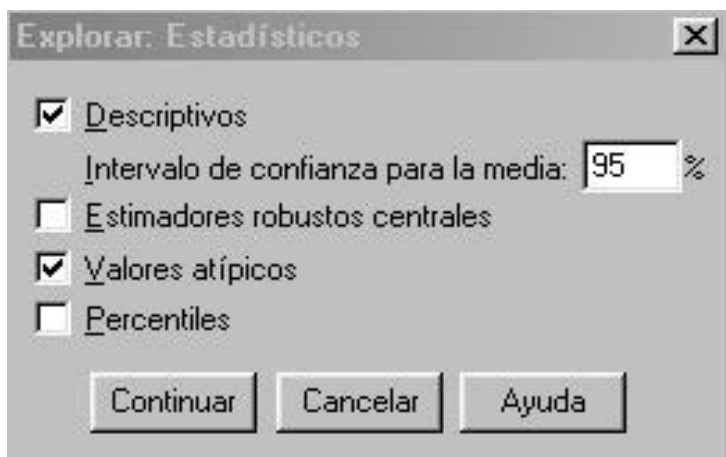


Figura 4.4. Explorar: Estadísticos.

*Los gráficos con pruebas de normalidad* muestran los diagramas de probabilidad normal, donde se representan los valores correspondientes a una distribución normal teórica mediante una recta, y los puntos que se corresponden a las diferentes puntuaciones observadas de los sujetos. También los gráficos de probabilidad sin tendencia, que recogen las desviaciones de los sujetos respecto de la recta normal. Calcula también los estadísticos de Kolmogorov-Smirnov y de Shapiro-Wilks para muestras con 50 o menos observaciones, para contrastar la normalidad. En este caso, se eligió el gráfico de tallos y hojas, el histograma, y gráficos con prueba de normalidad.



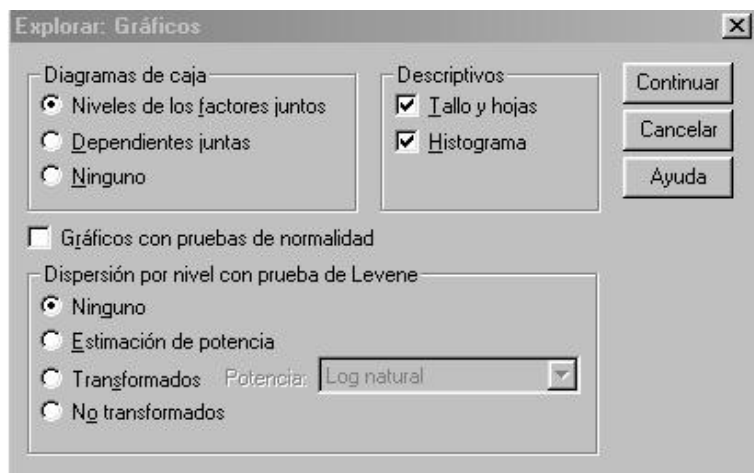


Figura 4.5.8 Explorar: Gráficos.

*La opción dispersión por nivel con prueba de Levene* sirve para representar los diagramas de dispersión por nivel, en el caso de seleccionar algún factor. Junto con los mismos, muestra la pendiente de la recta de regresión y realiza la prueba de Levene sobre la homogeneidad de varianzas. Si se realiza alguna transformación sobre los datos, las pruebas de Levene se llevarán a cabo con los datos transformados. Dentro de este apartado se tienen las siguientes opciones:

*Estimación de potencia:* Representa un gráfico de los logaritmos naturales de las amplitudes intercuartil, respecto a los logaritmos naturales de las medianas de todas las casillas, así como una estimación de la transformación de potencia necesaria para conseguir varianzas iguales de los grupos.

*Transformados:* Genera gráficos de los datos transformados cuando se selecciona una de las alternativas de potencia.

*Aplicación al Caso en Estudio*

*a) Análisis descriptivo de la variable edad*

A continuación, se presenta el procesamiento de la variable *edad* de nuestro caso de estudio (dependiente en el lenguaje del cuadro de diálogo de SPSS) y la variable *sexo* (factor según este programa).

La secuencia de opciones seleccionadas es:

Analizar

Estadísticos descriptivos

Explorar

Edad (dependiente), Sexo (factor)

Gráficos: Tallo y Hoja, Histograma y Gráficos con prueba de Normalidad.

Resumen del procesamiento de los casos

		Casos					
		Válidos		Perdidos		Total	
		N	Porcentaje	N	Porcentaje	N	Porcentaje
EDAD	hombres	185	100.0%	0	.0%	185	100.0%
	mujeres	254	100.0%	0	.0%	254	100.0%

## Descriptivos

SEXO			Estadístico	Error típ.
EDAD	hombres	Media	36.95	.91
		Intervalo de confianza para la media al 95%	35.14	
		Límite inferior	38.75	
		Límite superior		
		Media recortada al 5%	36.40	
		Mediana	36.00	
		Varianza	154.758	
		Desv. típ.	12.44	
		Mínimo	18	
		Máximo	70	
		Rango	52	
		Amplitud intercuartil	21.00	
		Asimetría	.462	.179
		Curtosis	-.596	.355
	mujeres	Media	37.80	.86
		Intervalo de confianza para la media al 95%	36.10	
		Límite inferior	39.49	
		Límite superior		
		Media recortada al 5%	37.00	
		Mediana	36.00	
		Varianza	188.907	
		Desv. típ.	13.74	
		Mínimo	18	
		Máximo	82	
		Rango	64	
		Amplitud intercuartil	21.50	
		Asimetría	.671	.153
		Curtosis	-.072	.304

Cuadro 4.4. Resultados opción Explorar.

Se observa en la tabla que la *edad media* de los hombres es de aproximadamente 37,0 años y 37,8 años en las mujeres, con una desviación estándar (des. típ.) de 12,44 años en hombres y 13,74 años en mujeres.

El *intervalo de confianza* del 95% para la media poblacional de la edad viene dado para los hombres por [35,14 años a 38,75 años] y [36,10 años a 39,49 años] para las mujeres.

El coeficiente de *asimetría* nos dice que existe una leve asimetría a la derecha tanto en hombres como en mujeres, es decir, las frecuencias más altas para la variable edad se encuentran al

lado izquierdo de la media.

El coeficiente de *curtosis* nos muestra que la distribución de edad en los hombres y mujeres presenta una bajo grado de concentración alrededor de las medidas de tendencia central, denominándose a esta distribución como *platicúrtica*, es decir, más achatada que una distribución normal.

La *media recortada* nos indica que eliminando el 5% de los datos extremos (más altos y más bajos), la media del 95% de los casos restantes es de 36,4 años en los hombres y de 37,0 años en las mujeres. La *mediana* indica que el 50% de los sujetos en estudio, hombres o mujeres, tienen una edad superior a los 36 años (o, 50% de los hombres o mujeres tienen una edad inferior a los 36 años). La dispersión determinada a través de la *amplitud intercuartil* (diferencia entre el tercer y primer cuartil, esto es, el 50% central de los casos) es de 21,0 años para hombres y 21,5 años para las mujeres.

Valores extremos

sexo				Número del caso	Valor
EDAD	hombres	Mayores	1	167	70
			2	327	69
			3	408	67
			4	159	66
			5	199	65
		Menores	1	262	18
			2	213	18
			3	412	18
			4	363	19
			5	293	19
	mujeres	Mayores	1	295	82
			2	221	78
			3	298	75
			4	198	74
			5	104	74
		Menores	1	196	18
			2	55	18
			3	407	18
			4	49	19
			5	57	19 <sup>a</sup>

a. En la tabla de valores extremos menores sólo se muestra una lista parcial de los casos con el valor 19.

Tabla 4.5. Valores extremos.

En la tabla de valores extremos, se aprecian las 5 edades máximas y mínimas, tanto para hombres como para mujeres con el número del caso para su ubicación en la ventana de datos.

### *Histogramas y Gráficos de Tallo y Hoja*

Los Histogramas de la variable edad, para cada sexo, son los siguientes:

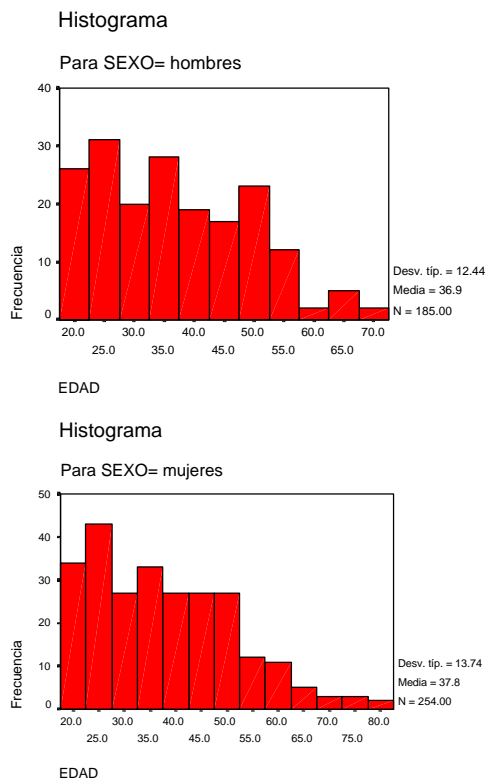


Gráfico 4.2. Histogramas.

En los Histogramas puede observar la asimetría de la variable edad con respecto al sexo y su nivel de apuntamiento (curtosis) claramente achatada y desviada a la izquierda con respecto a la media.

EDAD Stem-and-Leaf Plot for  
SEXO= hombres

Frequency Stem & Leaf

5.00	1 .	88899
34.00	2 .	00111111112222222222333333344444
21.00	2 .	5555666666667777777889
28.00	3 .	000111111122222223333333444
22.00	3 .	555566666666667777788999
18.00	4 .	000001111222223344
22.00	4 .	555666666777788888999
19.00	5 .	0000000001111244444
9.00	5 .	556667789
2.00	6 .	34
4.00	6 .	5679
1.00	7 .	0

Stem width: 10

Each leaf: 1 case (s)

Cuadro 4.6. Gráfico de Tallo y Hoja, variable edad para sexo (hombres)

EDAD Stem-and-Leaf Plot for  
SEXO= mujeres

Frequency Stem & Leaf

6.00	1 .	8 9
49.00	2 .	0011111111222233334444444
34.00	2 .	555566667777888899
23.00	3 .	00111123344
35.00	3 .	5555666677788899
27.00	4 .	0111122234444
35.00	4 .	555666778888999
13.00	5 .	001233
10.00	5 .	5568&
11.00	6 .	00123
5.00	6 .	8&
3.00	7 .	4&
2.00	7 .	&

1.00 Extremes (>=82)

Stem width: 10

Each leaf: 2 case(s)

& denotes fractional leaves.

Cuadro 4.7. Gráfico de Tallo y Hoja, variable edad para sexo (mujeres)

En el diagrama de tallo y hoja, la longitud de cada fila corresponde al número de casos que hay dentro de cada intervalo. Cada caso se representa dentro de la fila con un valor numérico. Para ello, se divide cada valor observado en dos componentes: el tallo (dígito/s inicial/es) y la hoja (dígito/s final/es). La primera y última fila se utilizan para representar los casos extremos. Como se puede observar en el intervalo 9, frecuencia 9 hombres, los casos que caen dentro del mismo, tienen los valores de la edad iguales a: 55, 55, 56, 56, 56, 57, 57, 58, 59. Es importante fijarse en el número de casos para cada hoja (leaf). En mujeres, cada hoja representa dos casos.

### *Diagramas de Caja*

Los diagramas de caja proporcionan para la variable edad indican las mismas característica ya señaladas. Se aprecia la asimetría de las distribuciones, tanto para hombres como para mujeres. Se aprecia, además, que las distribuciones son muy parecidas: coinciden sus medianas, así como sus recorridos intercuartílicos (ancho de la caja)

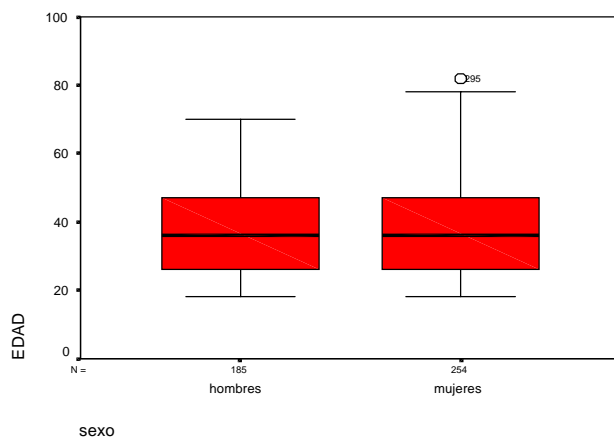


Gráfico 4.3. Diagrama de Caja.

### ***Gráficos de Probabilidad Normal***

Los *gráficos de probabilidad normal* constituyen otra importante herramienta gráfica para comprobar si un conjunto de datos puede considerarse o no procedente de una distribución normal. La idea básica consiste en enfrentar, en un mismo gráfico, los datos que han sido observados frente a los datos teóricos que se obtendrían de una distribución gaussiana. Si la distribución de la variable coincide con la normal, los puntos se concentrarán en torno a una línea recta, aunque conviene tener en cuenta que siempre tenderá a observarse mayor variabilidad en los extremos.

En los *gráficos Q-Q* se representan los cuantiles respecto a los cuantiles de la distribución normal. Además de permitir valorar la desviación de la normalidad, los gráficos de probabilidad permiten conocer la causa de esa desviación. Una curva en forma de «U» o con alguna curvatura significa que la distribución es asimétrica con respecto a la normal, mientras que un gráfico en forma de «S» significará que la distribución tiene colas mayores o menores que la normal, esto es, que existen pocas o demasiadas observaciones en las colas de la distribución.

Parece lógico que cada uno de estos métodos se complemente con procedimientos de análisis que cuantifiquen de un modo más exacto las desviaciones de la distribución normal.

Existen distintos tests o pruebas estadísticas que podemos utilizar para este propósito. El *test de Kolmogorov-Smirnov* es el más extendido en la práctica. Se basa en la idea de comparar la función de distribución acumulada de los datos observados con la de una distribución normal, midiendo la máxima distancia entre ambas curvas.

Como en cualquier test de hipótesis, la hipótesis nula se rechaza cuando el valor del estadístico supera un cierto valor crítico que se obtiene de una tabla de probabilidad.

Dado que en la mayoría de los paquetes estadísticos, como el SPSS, aparece programado dicho procedimiento, y proporciona



tanto el valor del estadístico de prueba como el valor-p correspondiente, no nos detendremos más en explicar su cálculo.

Otro procedimiento muy extendido es el test chi-cuadrado de bondad de ajuste. No obstante, este tipo de procedimientos deben ser utilizados con precaución. Cuando se dispone de un número suficiente de datos, cualquier test será capaz de detectar diferencias pequeñas aún cuando estas no sean relevantes para la mayor parte de los propósitos. El test de Kolmogorov-Smirnov, en este sentido, otorga un peso menor a las observaciones extremas y, por lo tanto, es menos sensible a las desviaciones que normalmente se producen en estos tramos.

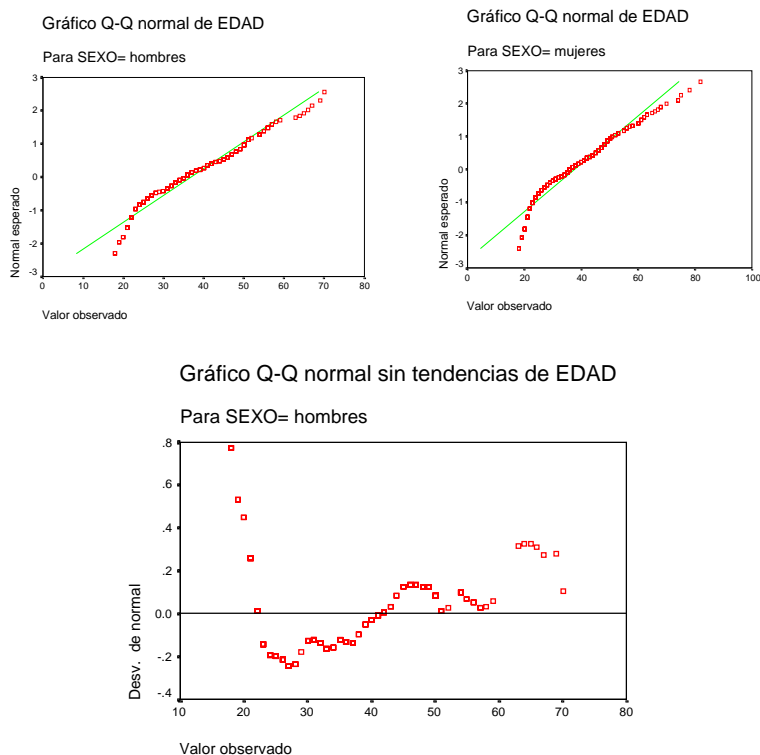


Gráfico 4.4. Gráfico Q-Q variable sexo (hombres).

En el gráfico Q-Q de los puntajes normalizados (obtenidos de la distribución normal acumulada) versus los valores observados (en este caso, los residuos), muestra una línea recta que representa una distribución normal perfecta y los puntos son las observaciones.

La distribución normal de los datos no se ve, ya que existen grandes desviaciones de la línea recta y la distribución en el histograma tiene relativamente una asimetría alrededor de la media. Entonces, proponer un supuesto de normalidad de la distribución de la *edad* con respecto al *sexo* no sería aconsejable.

### Gráfico Q-Q normal sin tendencias de EDA

Para SEXO= mujeres

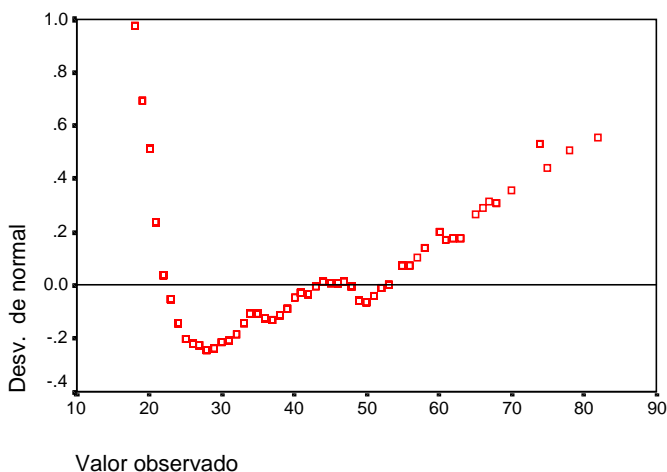
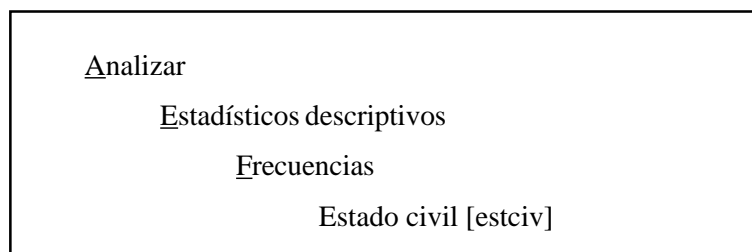


Gráfico 4.5. Gráfico Q-Q variable sexo (mujeres).

En el gráfico de probabilidad normal sin tendencias, se puede deducir que en la distribución de la edad en hombres y mujeres, no existiría tendencia a la normalidad.

### ***b) Análisis Descriptivo de la Variable Estado Civil***

Ahora examinaremos *estado civil* de nuestro caso de estudio. Puesto que se trata de una variable categórica, para describir la muestra en función de la misma, lo único que se puede hacer es ver cuántos sujetos (y qué porcentaje) están en cada nivel. Para elaborar la tabla de frecuencia, se selecciona en la barra de menús de SPSS la secuencia:



El investigador debe definir la(s) variable(s) que desea procesar:

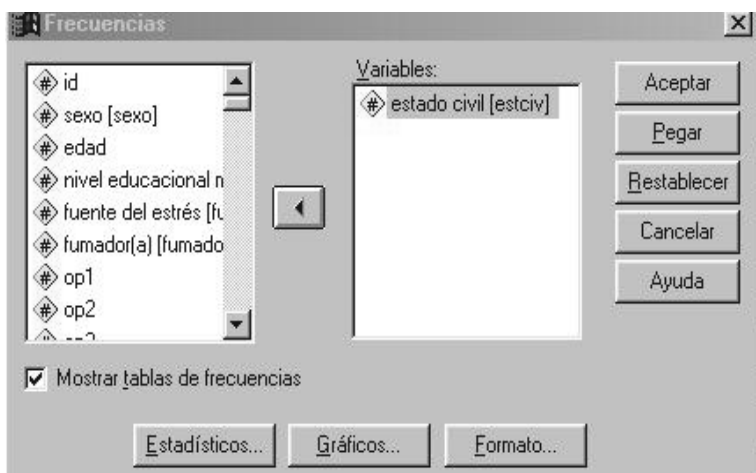


Figura 4.6. Definición de Tabla de Frecuencia para Variable estciv.

estado civil					
		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	soltero	105	23.9	23.9	23.9
	relación esporádica	37	8.4	8.4	32.3
	vive con la pareja	37	8.4	8.4	40.8
	casado por primera vez	189	43.1	43.1	83.8
	casado por segunda vez	30	6.8	6.8	90.7
	separado	10	2.3	2.3	92.9
	divorciado	24	5.5	5.5	98.4
	viudo	7	1.6	1.6	100.0
	Total	439	100.0	100.0	

Cuadro 4.8. Tabla de Frecuencia variable estciv.

Estado civil (Estciv)

Gráfico de Sectores

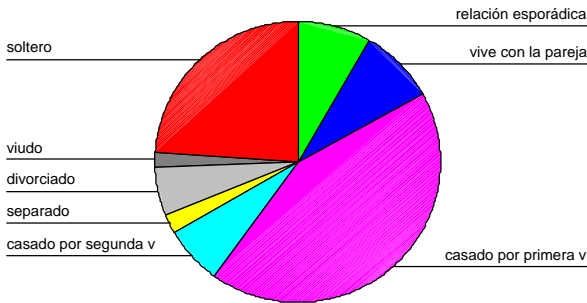


Gráfico 4.6. Gráfico de Sectores variable Estciv.

Una vez descrita la muestra en función de cada variable, se puede poner a prueba el ajuste entre la distribución de frecuencias empírica y una determinada distribución poblacional. Se pueden dar dos casos: (a) que se considere la muestra como representativa de la población y usemos la distribución de frecuencias de la muestra para comprobar la hipótesis de una determinada distribución poblacional, generalmente la distribución uniforme;

y (b) que se conozca la distribución poblacional y se quiera determinar si la muestra es o no representativa de la población.

Supongamos que queremos ver si la muestra en estudio representa adecuadamente a la población, en la que sabemos que el 20% son solteros, el 50% son casados y el 25% viven en pareja. Para ello, se usa la prueba chi-cuadrado (pruebas no paramétricas) que compara las frecuencias empíricas de cada nivel con las que cabría esperar en función de los porcentajes que se dan en la población. Por ejemplo, si en la población, el 50% son casados cabría esperar que la mitad de los sujetos de la muestra, es decir, 30, fueran casados.

Una vez seleccionada la prueba, se especifica la variable que se quiere contrastar, en nuestro caso, el estado civil, y “Valores esperados” se cambia la opción “Todas las categorías iguales” (que pondría a prueba la distribución uniforme) por la de “Valores”, donde se especifican los porcentajes poblacionales de los 8 niveles de la variable, y en el mismo orden en que se habían considerado en la matriz de datos (20, 5, 5, 50, 5, 5, 5, 5) es decir, en el mismo orden que se muestra en la ventana de *etiquetas de valor* en la hoja *vistas de variables*.

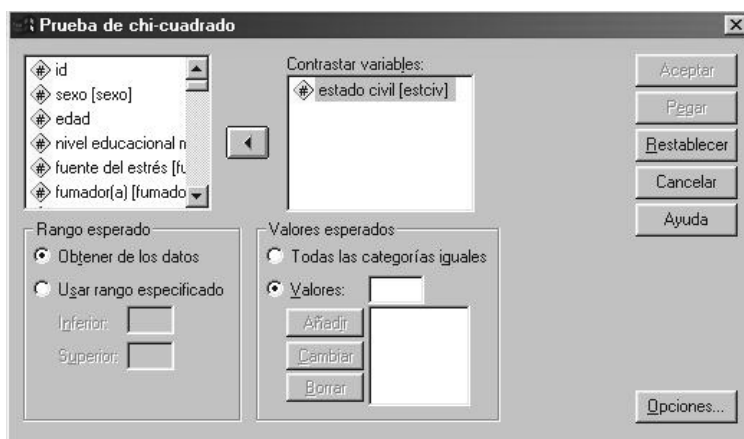


Figura 4.7. Menú Chi-cuadrado.

Lo que hacemos con ello es someter a prueba la hipótesis nula de que cualquier diferencia entre los porcentajes obtenidos en la muestra para cada nivel y los que se dan en la población (y en este caso conocemos) puede deberse al azar. Si la probabilidad asociada a la hipótesis nula es igual o mayor a 0,05 ( $p > 0,05$ ), tendremos que mantenerla; si es menor ( $p < 0,05$ ), se rechaza la hipótesis nula (con un nivel de significación del 5%). Los resultados son los siguientes:

### estado civil

	N observado	N esperado	Residual
soltero	105	87.8	17.2
relación esporádica	37	22.0	15.1
vive con la pareja	37	22.0	15.1
casado por primera vez	189	219.5	-30.5
casado por segunda vez	30	22.0	8.1
separado	10	22.0	-12.0
divorciado	24	22.0	2.1
viudo	7	22.0	-15.0
Total	439		

### Estadísticos de contraste

	estado civil
Chi-cuadrado <sup>a</sup>	48.077
gl	7
Sig. asintót.	.000

a. 0 casillas (.0%) tienen frecuencias esperadas menores que 5. La frecuencia de casilla esperada mínima es 22.0.

Cuadro 4.9. Resultados variable Estciv y su correspondiente estadístico de contraste.

Se rechaza la hipótesis nula dado que la significación (sig. asintót.) indicada en la tabla es 0,000, esto se interpreta como una probabilidad muy baja ( $p < 0,001$ ), mucho menor que 0,05. Por lo tanto, la muestra proviene de una población que no tiene la estructura de la población de referencia. Se aprecia que hay un sesgo, pues los solteros están excesivamente representados.

### ***c) Análisis Descriptivo de la Variable Nivel Educacional***

Examinaremos la variable *nivel educacional* de los sujetos de la muestra de nuestro caso de estudio. Esta variable ha sido considerada como ordinal, puesto que sus distintos niveles pueden ordenarse de menor a mayor. En este caso, tiene sentido obtener los porcentajes acumulados y la mediana (nivel correspondiente al sujeto que deja por debajo de sí al 50% de los casos). La secuencia de pasos es la siguiente:

#### **Analizar**

#### **Estadísticos descriptivos**

#### **Frecuencias**

#### **Nivel educacional [educ]**

#### **Estadísticos la moda y la media**

#### **Gráficos, el gráfico de barra**

El cuadro de diálogo en el que se seleccionan los procedimientos estadísticos seleccionados aparece en la figura siguiente:



Figura 4.8. Tabla de Frecuencias y Estadísticos.

Los resultados son los siguientes:

N	Válido	439
	Perdidos	0
Media		4.00
Moda		5

nivel educacional más alto y completo

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos básica	2	.5	.5	.5
media incompleta	53	12.1	12.1	12.5
media completa	85	19.4	19.4	31.9
capacitación técnica	120	27.3	27.3	59.2
universitaria completa	123	28.0	28.0	87.2
posgrado	56	12.8	12.8	100.0
Total	439	100.0	100.0	

Cuadro 4.10 Medidas de Tendencia Central y  
Tabla de Frecuencias.

Para obtener un gráfico de barras, éste se selecciona en el menú siguiente

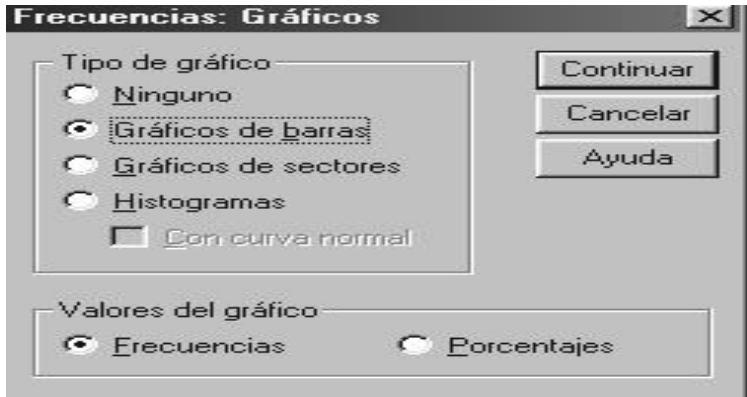


Figura 4.9. Cuadro para definición de Gráfico  
en opción Frecuencias.



En la figura siguiente se presenta el gráfico respectivo.

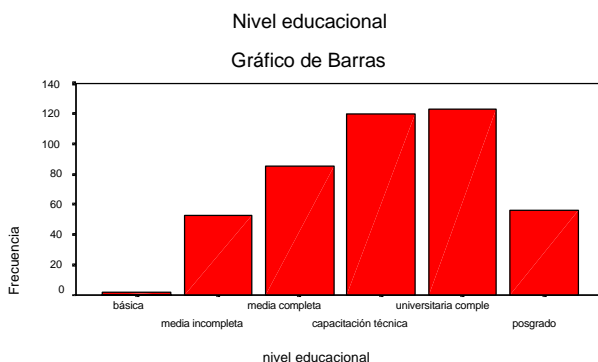


Gráfico 4.7. Gráfico de Barras variable Nivel educacional

A continuación, comprobaremos, en primer lugar, el ajuste de los datos a la distribución uniforme, es decir se pondrá a prueba la hipótesis nula de que en la población los sujetos se reparten de forma uniforme entre los distintos niveles educacionales, esto es, en cada nivel existe la misma proporción de individuos. Para ello se ha seleccionado la opción Chi-cuadrado dentro de *pruebas no paramétricas* y se deja la opción (por defecto) de *todas las categorías iguales*.

Nivel educacional			
	N observado	N esperado	Residual
básica	2	73.2	-71.2
media incompleta	53	73.2	-20.2
media completa	85	73.2	11.8
capacitación técnica	120	73.2	46.8
universitaria completa	123	73.2	49.8
posgrado	56	73.2	-17.2
Total	439		

Estadísticos de contraste

	Nivel educacional
Chi-cuadrado <sup>a</sup>	144.640
gl	5
Sig. asintót.	.000

a. 0 casillas (.0%) tienen frecuencias esperadas menores que 5. La frecuencia de casilla esperada mínima es 73.2.

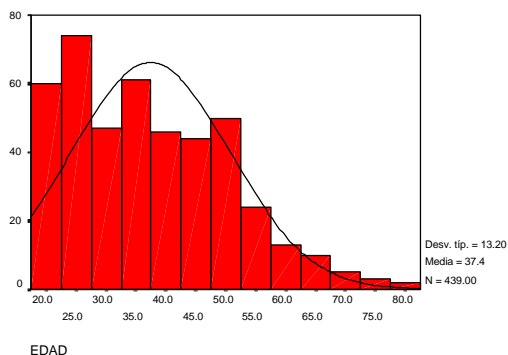
Cuadro 4.11. Prueba de Chi-cuadrado para variable Nivel educacional.

La baja probabilidad ( $p=0,000$ ) asociada a la hipótesis nula lleva a rechazarla; entonces, los sujetos de la población a la que pertenece la muestra no se reparten uniformemente entre los distintos niveles educacionales del caso de estudio.

#### ***d) Análisis Descriptivo de la Variable Edad***

El archivo de datos del caso de estudio contiene también la variable *edad* que es tratada como cuantitativa, por lo que en principio, podríamos calcular todos los estadísticos para describirla. Sin embargo, cuando la distribución es muy asimétrica, además de la media, conviene usar la mediana como estadístico de tendencia central, y, como estadístico de variabilidad, la amplitud intercuartil (diferencia entre el percentil 75 y el percentil 25) en lugar de la desviación típica. La mediana es menos sensible que la media a la presencia de puntuaciones extremas, y por lo tanto, representa mejor la tendencia central de los datos.

Si el cociente entre el estadístico de asimetría y su error típico (en valor absoluto) es mayor que 1,96; podemos considerar la distribución como muy asimétrica, como ocurre en este caso, pues  $0,606/0,117=5,18$ . Tenemos una clara asimetría positiva: hay muchos sujetos con puntuaciones bajas y pocos con puntuaciones altas. Ello se puede observar en el histograma obtenido mediante la activación de la secuencia Gráfico/Histograma e indicándole que muestre la *curva normal*.



**Gráfico 4.8. Histograma variable *Edad***

Puesto que la variable es asimétrica, podemos usar la mediana como índice de tendencia central y la amplitud intercuartil como índice de variabilidad.

### Estadísticos

#### EDAD

N	Válidos	439
	Perdidos	0
Mediana		36.00
Percentiles	25	26.00
	50	36.00
	75	47.00

Cuadro 4.12. Estadísticos variable *Edad*.

Así, la mediana es 36,0 años y la amplitud intercuartil  $47 - 26 = 21$  años. La interpretación es: el 50% o menos de los casos tiene 36 años o menos, y, por otra parte, entre 26 y 47 (es decir, 21 años) se encuentra comprendido el 50% central de los sujetos de la muestra

Para examinar la normalidad de una variable, se puede usar la prueba de Kolmogorov-Smirnov (K-S). La secuencia que se activa es la siguiente:

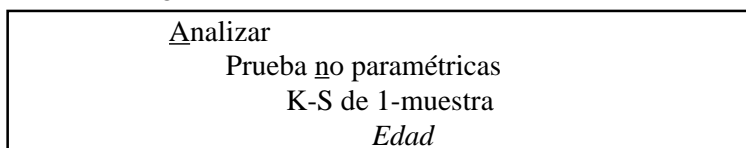


Figura 4.10 Cuadro de diálogo para definición de Prueba de Kolmogorov-Smirnov.

El resultado revela que la significación que nos da el K-S es menor que 0,05 ( $p=0,001$ ), por lo cual tenemos que rechazar la hipótesis nula de que esta variable tiene una distribución normal en la población.

#### Prueba de Kolmogorov-Smirnov para una muestra

		EDAD
N		439
Parámetros normales <sup>a,b</sup>	Media	37.44
	Desviación típica	13.20
Diferencias más extremas	Absoluta	.091
	Positiva	.091
	Negativa	-.070
Z de Kolmogorov-Smirnov		1.899
Sig. asintót. (bilateral)		.001

a. La distribución de contraste es la Normal.

b. Se han calculado a partir de los datos.

#### Prueba de Kolmogorov-Smirnov para una muestra

		EDAD
N		439
Parámetros normales <sup>a,b</sup>	Media	37.44
	Desviación típica	13.20
Diferencias más extremas	Absoluta	.091
	Positiva	.091
	Negativa	-.070
Z de Kolmogorov-Smirnov		1.899
Sig. asintót. (bilateral)		.001

a. La distribución de contraste es la Normal.

b. Se han calculado a partir de los datos.

Cuadro 4.13. Resultados K-S.

### ***Creación de Submuestras para el Análisis con SPSS***

Una vez realizado un estudio descriptivo univariante de cada una de las variables sobre las que se quiere realizar el estudio, se pasa a una segunda fase consistente en realizar un estudio descriptivo de varias variables en conjunto.

En esta primera fase descriptiva podemos hacer dos cosas:

1. Calcular coeficientes descriptivos de las variables considerándolas en grupos de dos (coeficientes de correlación, asociación, tablas de contingencia, etc., se verán más adelante).
2. Si la naturaleza del problema lo permite, se puede estudiar el comportamiento de una variable numérica en distintos subconjuntos de la muestra. Este no es un análisis multivariante propiamente dicho ya que no se involucra el estudio de varias variables en conjunto, pero se puede decir que la partición realizada por la variable de clasificación «genera» nuevas variables, por lo que lo se puede considerar como un análisis de varias variables.

Por ejemplo, supongamos que se dispone de información sobre la variable estatura de una muestra. Si además se cuenta con la información del sexo de los mismos individuos, se pueden crear dos nuevas variables o submuestras (altura de los hombres y altura de las mujeres) a las que se las puede someter tanto a estudios descriptivos por separado como a comparaciones entre las dos submuestras.

En SPSS se utiliza la opción *Datos / Seleccionar casos* en orden a procesar solamente los sujetos de sexo masculino de la base de datos de nuestro caso de estudio. Después de activar las dos opciones anteriores aparece, el siguiente cuadro de diálogo. Se puede observar que el programa está trabajando con *Todos los casos*. En este punto seleccionaremos la opción *Si se satisface la condición* para indicarle al programa que solamente trabajaremos con los sujetos de *sexo masculino*.

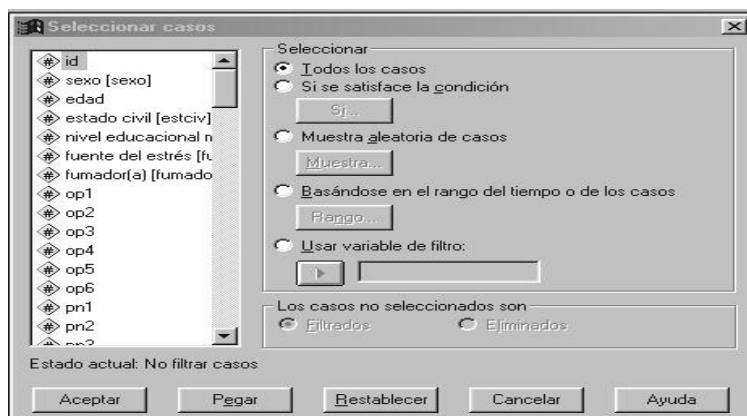


Figura 4.11. Seleccionar casos.

Esto se realiza en la siguiente pantalla:

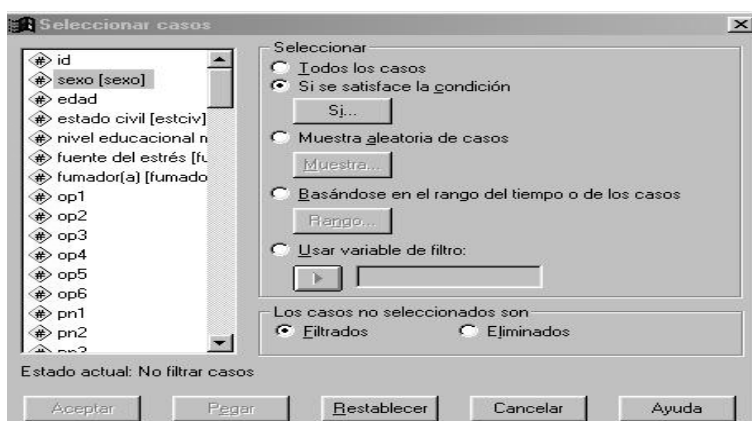


Figura 4.12. Cuadro de diálogo para establecer la condición.

En seguida se indica la condición selectiva: Sexo = 1 (código de los hombres)

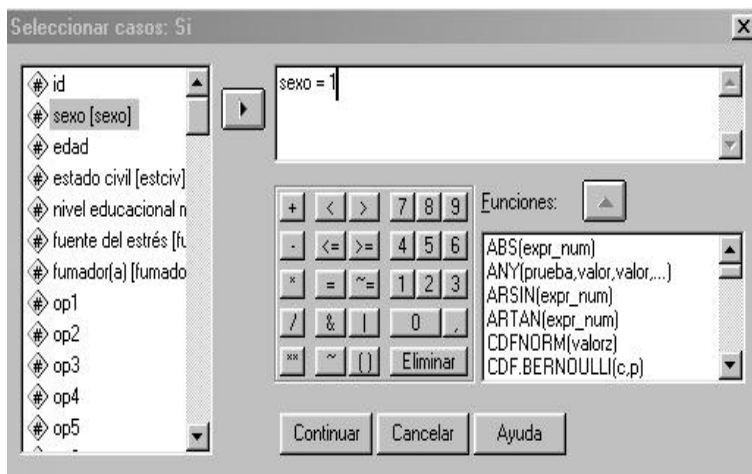


Figura 4.13. Condición para la variable sexo

Al activar esta opción, el cuadro de diálogo nos informa de esto:

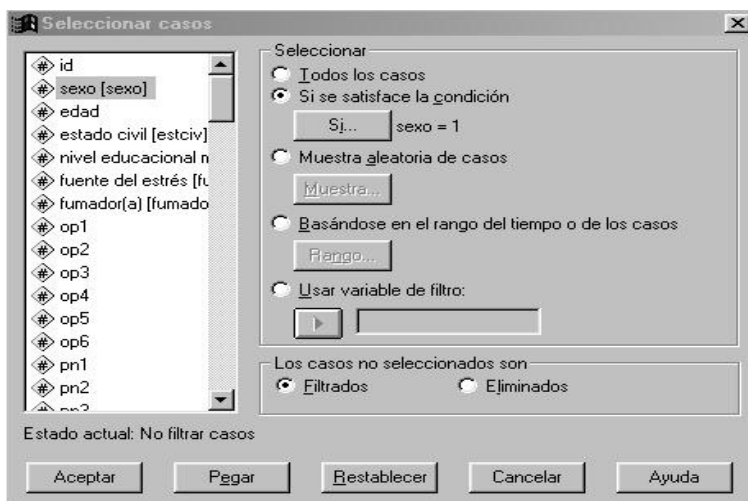


Figura 4.14. Condición establecida: sexo=1.

Como resultado de esta operación, en el área de *Vista de datos* se nos indica los sujetos no seleccionados, marcando el número del sujeto como se indica en la siguiente pantalla y en la *Vista de Variables* se crea una nueva variable filtro que permanecerá activa hasta que decidamos eliminarla.

	id	sexo	edad	estciv	educ
4	4	2	47	4	2
5	5	2	48	4	3
6	6	2	41	7	6
7	7	1	33	1	5
8	8	2	38	7	5
9	9	1	39	3	5
10	10	2	67	4	2
11	11	2	22	2	5
12	12	2	31	4	5
13	13	2	45	4	5
14	14	1	26	4	5
15	15	1	51	5	2
16	16	2	37	5	3
17	17	1	33	1	4
18	18	2	61	8	3
19	19	2	33	1	4
20	20	2	42	4	4
21	21	2	45	4	6

	Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores
90	tsatvida	Numérico	8	0	total satisfacci	Ninguno
91	testres	Numérico	8	0	total estrés per	Ninguno
92	totautoe	Numérico	8	0	total autoestim	Ninguno
93	edadgp3	Numérico	8	0	edad 3 grupos	{1, 18-29}...
94	edadgp5	Numérico	8	0	edad 5 grupos	{1, 18-24}...
95	filter_\$	Numérico	1	0	sexo = 1 (FILT	{0, No seleccio

Figura 4.15. Variable filtro y demarcación de los casos.

Para restablecer la base de datos, o se elimina la variable o en el menú *Datos / Seleccionar casos* se activa *Todos los casos* y se elimina la condición previamente establecida.



## ***Ejercicios***

### Medidas de localización y dispersión

1. Una medida de tendencia central es una puntuación que describe:
  - a. todos los otros valores de la distribución
  - b. el valor más importante de la distribución
  - c. el centro de la desviación estándar
  - d. el centro alrededor del cual tiende a localizarse la distribución
  
2. La media NO describe adecuadamente una distribución cuando:
  - a. la distribución es simétrica
  - b. el valor más repetido de la distribución es muy alto
  - c. la distribución no tiene moda
  - d. hay valores atípicos en la distribución
  
3. Cuando se predice una puntuación individual de la distribución utilizando la media nos aseguramos que:
  - a. el error total es cero
  - b. no hay error en la predicción
  - c. no se necesita ninguna otra información
  - d. la media coincide con la mediana y la moda
  
4. ¿Qué medida se afecta menos con los valores atípicos?
  - a. media
  - b. amplitud
  - c. amplitud intercuartil
  - d. desviación estándar

5. La medida de localización que más se afecta con los valores atípicos es la:
- media
  - mediana
  - moda
  - primer cuartil
6. Si todos los valores de una muestra son iguales entonces la desviación estándar:
- es positiva
  - es igual a cero
  - es negativa
  - no se puede calcular
7. Las medidas de dispersión tienen como propósito indicar:
- la variación en un conjunto de datos
  - el promedio de las desviaciones con respecto a la media
  - la diferencia entre la media, la moda y la mediana
  - la diferencia entre el centro y los extremos de la distribución
8. Si en una distribución se eliminan los valores atípicos entonces, ¿qué le sucede a la desviación estándar ?
- No se afecta
  - Aumenta
  - Disminuye
  - No se puede determinar como va a cambiar

## CAPÍTULO V

---

### COMPARACIÓN DE DOS O MÁS POBLACIONES

La prueba de hipótesis sobre medias poblacionales es una de las aplicaciones más usadas por los investigadores en el campo de las Ciencias Sociales. El uso de software estadístico, como es el caso del SPSS, ofrece una variedad de procedimientos capaces de realizar este tipo de análisis.

El SPSS, como la mayoría de los programas estadísticos, trabaja inicialmente sobre arreglos rectangulares de datos (matrices de datos), donde las filas corresponden a los casos, mientras las columnas denotan las variables.

De esta forma, la pregunta esencial que debemos resolver cuando vamos a comparar dos o más medias poblacionales, empleando muestras aleatorias de las mismas, es si estas representan a procesos de muestreos independientes o relacionados.

La segunda pregunta que deberemos resolver es si se quiere comparar las medias de dos o más de dos grupos de casos.

#### *COMPARACIÓN DE DOS POBLACIONES INDEPENDIENTES*

A continuación se presentarán varias pruebas que tienen la finalidad de comparar dos poblaciones y que corresponden al procedimiento *Analizar/Comparar Medias* de SPSS.

Las poblaciones pueden ser independientes o dependientes y esta distinción se empleará en la presentación.

### ***Muestras Independientes***

Para la comparación de las medias de dos poblaciones independientes respecto a una variable de interés empleando muestras aleatorias de las mismas, por ejemplo, la comparación de hombres y mujeres respecto a los resultados en el test de autoestima, una de las pruebas estadísticas más populares es *la prueba T para muestras independientes*, que está implementada en SPSS.

La prueba T es una prueba paramétrica, o sea que sólo sirve para comparar variables numéricas de distribución normal. También, permite comparar las medias de dos categorías dentro de una misma variable. En caso de tener que analizar variables numéricas de distribución no normal, se debe utilizar otro tipo de pruebas, del grupo de las no paramétricas, como la prueba *U de Mann – Whitney*, la cual emplea rangos de las observaciones muestrales.

La matriz de datos debe estar configurada como es habitual, es decir, debe existir una columna para los datos de la variable de interés y una segunda columna con los códigos. También, pueden compararse dos variables ubicadas en diferentes columnas en la matriz de datos.

### ***Ejemplo***

Asumamos que se está estudiando los efectos del involucramiento de los padres (variable independiente) en los puntajes de un test de los estudiantes (variable dependiente). La mitad de los alumnos de octavo año fueron asignados aleatoriamente al grupo en que se involucró a los padres. El profesor contactó a los padres de estos estudiantes a través del

año y les contó sobre los objetivos educativos que se perseguían. La otra mitad de los estudiantes del curso se asignaron al grupo que no involucró a los padres. Los puntajes de la primera prueba fueron tabulados para todos los niños y se procedió a su procesamiento.

La hipótesis nula establece que las medias de los puntajes, en el test, de alumnos con padres involucrados es igual a la media de los puntajes de los alumnos con padres no involucrados.

	parpad	puntaje
1	1	78.6
2	1	64.9
3	1	100.0
4	1	83.7
5	1	94.0
6	1	78.2
7	1	76.9
8	1	82.0
9	2	81.0
10	2	69.5
11	2	73.8
12	2	66.7
13	2	54.8
14	2	69.3
15	2	73.5
16	2	79.4



Figura 5.1. Etiquetas de valores.

Después de activar Analizar/Comparar medias/Pruueba T para muestras independientes se definió las variables:



Figura 5.2. Prueba T para muestras independientes

El resultado es el siguiente:

**Estadísticos de grupo**

Involucramiento de los padres	N	Media	Desviación típ.	Error típ. de la media
PUNTAJE involucrado	8	82.288	10.788	3.814
no involucrado	8	71.000	8.200	2.899

**Prueba de muestras independientes**

	Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias						
	F	Sig.	t	gl.	Sig. (bilateral)	Diferencia de medias	Error típ. de la diferencia	95% Intervalo de confianza para la diferencia	
								inferior	superior
PUNTAJE Se han asumido varianzas iguales	.337	.571	2.356	14	.034	11.287	4.791	1.012	21.563
No se han asumido varianzas iguales			2.356	13.065	.035	11.287	4.791	.943	21.632

**Cuadro 5.1. Resultados prueba T para muestras independientes.**

La segunda tabla presenta la prueba de Levene para varianzas iguales, que evalúa el supuesto básico de la prueba T de que la varianzas de las dos poblaciones de las que provienen las muestras son aproximadamente iguales (hipótesis nula de homogeneidad de varianzas). En este caso, no se ha violado el supuesto, pues estadístico  $F=0,337$  tiene un valor  $p=0,571$  (mayor que 0,05) y debería usarse la  $t$  de 2,356 con 14 grados de libertad y el nivel de significación de 0,034. Por lo tanto, se rechaza la hipótesis nula sobre igualdad de puntajes medir poblacionales (pues valor  $p=0,034$  es menor a 0,05).

En conclusión, los datos muestran que el involucramiento de los padres tiene un efecto significativo en los puntajes obtenidos por los estudiantes.

### ***Aplicación al Caso en Estudio***

Volviendo al caso en estudio, se estudiará la relación entre el sexo de los sujetos (variable categórica) y el puntaje total obtenido en la escala de autoestima (variable numérica). La secuencia de pasos en SPSS es:

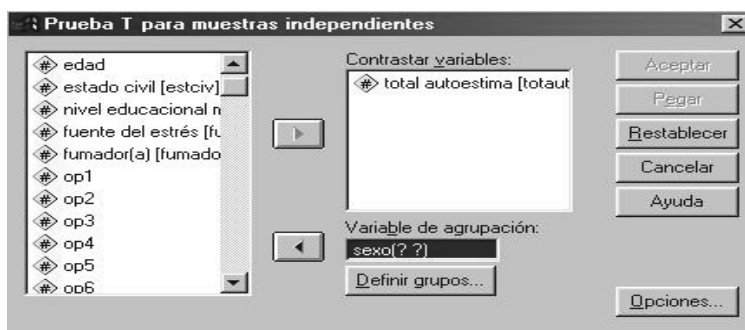
**Analizar****Compare Medias****Prueba T para muestras independientes****Total autoestima Sexo 1 2**

Figura 5.3. Cuadro de diálogo prueba T para muestras independientes.

En la ventana Prueba T para muestras independientes, se selecciona las variables *total autoestima* y se coloca en el espacio Contrastar variables. En seguida, se selecciona la variable *sexo* como Variable de agrupación. Se activa Definir grupos. Aparecerá la siguiente ventana:

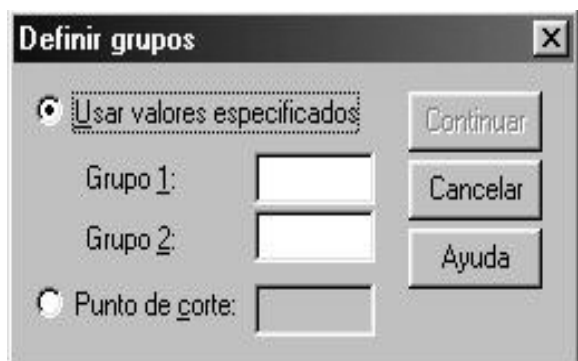


Figura 5.4. Cuadro de definición de grupos.

En la ventana *Definir grupos* se selecciona Usar valores especificados y en Grupo 1 se escribe 1 (hombre) y en Grupo 2, se escribe 2 (mujer). Se activa *Continuar* para regresar a la ventana Prueba T para muestras independientes. En esta ventana se selecciona Opciones y En Intervalo de Confianza se escribe 95. Se activa *Continuar* para regresar a la ventana Prueba T para muestras independientes.

El resultado es el siguiente:

**Estadísticos de grupo**

sexo		N	Media	Desviación típ.	Error típ. de la media
total autoestima	hombres	184	34.02	4.91	.36
	mujeres	252	33.17	5.71	.36

**Prueba de muestras independientes**

		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias						
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error típ. de la diferencia	95% Intervalo de confianza para la diferencia	
									Interior	Superior
total autoestima	Se han asumido varianzas iguales	3.506	.062	1.622	434	.105	.85	.52	-.18	1.87
	No se han asumido varianzas iguales			1.661	422.348	.098	.85	.51	-.16	1.85

**Cuadro 5.2. Resultados descriptivos y prueba t Student.**

La salida del SPSS nos ofrece en primer lugar, las medias de ambos grupos y las desviaciones típicas. También da el resultado de la *prueba de Levene* para contrastar la hipótesis de igualdad de varianzas. Si la probabilidad asociada a esa prueba es 0,05 o menos, se tiene que rechazar la hipótesis nula de que las varianzas son iguales; en caso contrario, se mantiene. En este caso, no podemos rechazar la hipótesis de igualdad de varianzas pues el valor p (sig.) es 0,062 y es mayor al 5%.

Por lo anterior, para contrastar la diferencia de medias, se elige la alternativa con el supuesto de varianzas iguales, cuyo



valor  $t=1,622$  tiene una probabilidad asociada (valor  $p$ , sig. Bilateral) de 0,105 la cual, otra vez, al ser mayor que 0,05 hace no rechazar la hipótesis nula de que las medias poblacionales son iguales; luego, no hay diferencia significativa entre las medias de ambos grupos. Por lo tanto, el nivel medio de autoestima es igual para hombres y mujeres.

## POBLACIONES RELACIONADAS O DEPENDIENTES

Existe una segunda alternativa para contrastar dos medias. Ésta se refiere al supuesto caso en el que las dos poblaciones no sean independientes, es decir, el caso en el que se trate de poblaciones relacionadas. Esta situación se encuentra, por ejemplo, en los diseños apareados, diseños en los que los mismos individuos son observados antes y después de una determinada intervención o en los diseños en los que las muestras son emparejadas de acuerdo a una serie de variables para controlar su efecto (como por ejemplo en los diseños de casos y controles). La prueba  $t$  para muestras relacionadas se utiliza para comparar las medias de un mismo grupo en diferentes etapas, como por ejemplo pre y post tratamiento, o para comparar muestras emparejadas.

Para la realización de este análisis, las dos muestras deben estar en dos variables distintas de la matriz de datos y debe formarse la pareja de muestras antes de poder añadirse la comparación a la lista de variables relacionadas del cuadro de diálogo.

La estructuración de datos es ahora diferente. Lo que representamos es una variable analizada en dos momentos distintos, y la forma de expresarlo es como si tuviésemos dos variables medidas para cada caso. Por ejemplo, podemos llamarlas *Prueba1* y *Prueba2*, resultando una estructuración de datos semejante a la siguiente:

Prueba1	Prueba2
4,50	4,90
4,70	5,20
4,30	6,20
4,60	5,00
5,60	7,30
6,00	8,30

El analista selecciona:

**Analizar**

**Comparar medias**

**Prueba T para muestras relacionadas  
pretest (prueba1) posttest (prueba2)**

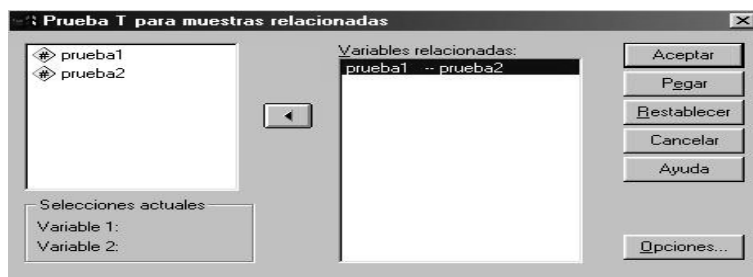


Figura 5.5. Cuadro prueba T para muestras relacionadas.

En seguida indica el intervalo de confianza, activando Opciones:

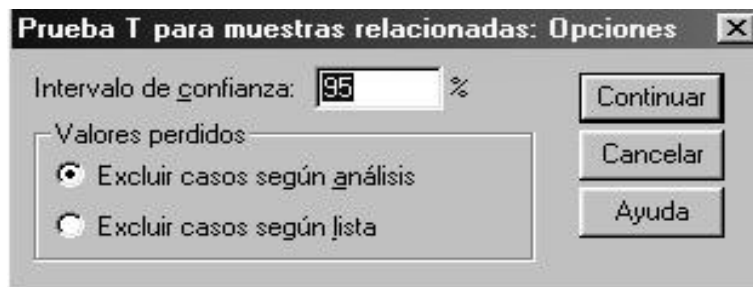


Figura 5.6. Cuadro para fijar el intervalo de confianza.

Los resultados serán los siguientes:

#### Estadísticos de muestras relacionadas

		Media	N	Desviación típ.	Error típ. de la media
Par 1	PRUEBA1	4.9500	6	.6834	.2790
	PRUEBA2	6.1500	6	1.3953	.5696

#### Correlaciones de muestras relacionadas

		N	Correlación	Sig.
Par 1	PRUEBA1 y PRUEBA2	6	.876	.022

#### Prueba de muestras relacionadas

		Diferencias relacionadas				t	gl	Sig. (bilateral)	
		Media	Desviación típ.	Error típ. de la media	95% Intervalo de confianza para la diferencia				
					Inferior				Superior
Par 1	PRUEBA1 - PRUEBA2	-1.2000	.8626	.3521	-2.1052	-.2948	-3.408	5	.019

Cuadro 5.3. Resultado estadísticos prueba T Student para muestras relacionadas.

El promedio de las diferencias es  $-1,2$  y tiene un  $t = -3,408$  el cual está asociado con un valor  $p$  (sig.) de  $0,019$  el cual es menor que el  $5\%$ . Por lo anterior, se rechaza la hipótesis nula que establece la igualdad de las medias poblacionales.

Basados en estos resultados, podríamos decir que el tratamiento produjo en el grupo una diferencia estadísticamente significativa entre la *prueba1* y la *prueba2*, siendo la media del postest (*prueba2*) superior a la media del pretest (*prueba1*).

## COMPARACIÓN DE MÁS DE DOS POBLACIONES

### Análisis de varianza (ANOVA)

El análisis de varianza (ANOVA) es un método de análisis que puede aplicarse cuando el investigador está tratando con una variable dependiente, medida en forma en escala de razón o

intervalo, y una o más variables independientes medidas en forma nominal u ordinal. Este tipo de análisis es la base de los diseños experimentales tradicionales. El procedimiento permite examinar medias de dos o más grupos y establecer la existencia de diferencias estadísticamente significativas. Su objetivo es comparar medias y la comparación se realiza mediante un cociente entre varianzas. La varianza se puede descomponer en varianza entre grupos y varianza dentro de grupos. La varianza entre grupos mide la dispersión atribuible al efecto de la variable experimental. La variable experimental es la variable independiente. La varianza entre grupos es la variación de la media de que cada grupo respecto a la media total. Esta variación se produce en virtud de que cada grupo está sometido a un tratamiento distinto. Los distintos tratamientos generan diferencias en las puntuaciones de la variable dependiente.

El ANOVA es una prueba semejante a la *prueba t Student*, en cuanto a la práctica, pero la comparación entre grupos no es a través de la media y su desviación estándar, sino a través de la varianza de la variable numérica “y”, en cada grupo de la variable categórica “x”.

Básicamente el *Análisis de Varianza*, se utiliza para corroborar si la significación de diferencias entre medias de dos o más grupos, son o no debidas al azar. El estadístico de prueba empleado en el Anova es la razón F. Suponiendo que se analizan 2 grupos, el ANOVA analiza las variaciones entre los dos grupos (inter-grupal) y la compara con la variación dentro de cada grupo (intra-grupal), para obtener mediante el cociente de sumas de cuadrados el valor de F.

Si las diferencias de varianza entre cada grupo son mayores que las intra-grupales, seguramente existen diferencias significativas entre los grupos que no son debidas al azar. Los grupos se definen como en la prueba t, eligiendo una variable categórica. La variable por analizar debe ser numérica y de distribución normal.

También existe un modelo de análisis de varianza multivariado, llamado MANOVA, en el cual se comparan más de una variable numérica en dos o más grupos.

En caso de tener que analizar las medias de variables numéricas cuya distribución no es normal, se debe apelar a las pruebas no paramétricas siendo la *prueba de Kruskal – Wallis* una de las más utilizadas.

La lógica del razonamiento es que cuando la dispersión producida por efecto de los tratamientos de la variable experimental es menor que la dispersión producida por los efectos ajenos a la variable experimental, cabe deducir que los tratamientos producen un efecto similar y las medias son iguales. En sentido contrario, si el efecto de los tratamientos de la variable experimental produce una dispersión mayor que la debida a efectos ajenos a la variable experimental, cabe deducir que los tratamientos producen efectos distintos y las medias son diferentes.

### ***Ejemplo***

Se requiere determinar si al aplicar 3 métodos de enseñanza de metodología de la investigación en la Universidad, se producen resultados diferentes que se comprueban aplicando una prueba final

Método 1. El profesor no desarrolla ni asigna problemas.

Método 2. El profesor desarrolla y asigna problemas.

Método 3. El profesor desarrolla y asigna problemas. Además los estudiantes deben desarrollar un proyecto basado en los temas explicados y comentados en clase.

El mismo profesor le enseña a 3 secciones diferentes de estudiantes, y usa uno de estos 3 métodos en clase. Los estudiantes son asignados aleatoriamente a las 3 secciones. Solamente hay 12 estudiantes en el experimento 4 en cada una de las diferentes secciones. La matriz de datos es la siguiente:

Estudiante	Método 1	Método 2	Método 3
1	16	19	24
2	21	20	21
3	18	21	22
4	13	20	25
Total	68	80	92

El formato de ingreso de los datos en SPSS es:

Método	Puntaje
1	16
1	21
1	18
1	13
2	19
2	20
2	21
2	20
3	24
3	21
3	22
3	25

En SPSS, se debe ingresar los datos en columnas por método de enseñanza y por puntaje. Se puede señalar que el método de enseñanza es la variable independiente (VI) con tres tratamientos como valores; mientras que el puntaje final es la variable dependiente (VD) con los puntajes de los estudiantes como valores. La secuencia que se activa es la siguiente:

**Analizar**

**Comparar medias**

**ANOVA de un factor**

**Factor: método de enseñanza**

**Dependientes: puntajes**

Los resultados entregados por SPSS son:

### Estadísticos Descriptivos

		N	Media	Desv. Est.	Error Est.	95% Intervalo de Confianza al 95%		Mínimo	Máximo
						Límite Superior	Límite Inferior		
Puntaje	Métodos	1	4	17.00	3.37	1.68	11.64 22.36	13	21
		2	4	20.00	.82	.41	18.70 21.30	19	21
		3	4	23.00	1.83	.91	20.09 25.91	21	25
		Total	12	20.00	3.28	.95	17.92 22.08	13	25

### ANOVA

		Suma de los Cuadrados	gl	Media Cuadrática	F	Sig.
Puntaje	Entre Grupos	72.000	2	36.000	7.043	.014
	Intra Grupos	46.000	9	5.111		
	Total	118.000	11			

Cuadro 5.4. ANOVA

La hipótesis nula  $H_0$  señala que los 3 métodos no son significativamente diferentes, medidos a través de los puntajes medios en la prueba final; dicho de otro modo ellos, son igualmente efectivos en mejorar el desempeño de los estudiantes. La hipótesis alterna  $H_a$  señala que el puntaje promedio es significativamente diferente y que la diferencia entre ellos se debe al diferente grado de efectividad de los tratamientos aplicados (métodos de enseñanza). El nivel de significación empleado es  $\alpha = 0,05$ . Comparando el valor de probabilidad (valor p) que en SPSS aparece como (sig) de 0,014 con respecto a 0,05, se observa que es menor, lo que significa que la prueba es fuertemente significativa al 5%; entonces,  $H_0$  debe ser rechazada con ese nivel de significación.

Basados en los resultados experimentales pareciera que los 3 métodos de enseñanza no tendrían el mismo efecto en el rendimiento final de los estudiantes. El tercer método de enseñanza aparentemente es superior. Esto se puede verificar mediante la realización del procedimiento de *comparaciones múltiples*, como se presenta en el ejemplo siguiente, correspondiente al caso en estudio sobre autoestima.

### ***Aplicación al Caso en Estudio***

A continuación, se analizará la hipótesis nula que no hay diferencia en el puntaje total de la autoestima (variable numérica, dependiente) entre los grupos de edades a los que pertenecen los sujetos (variable cualitativa, independiente). Se selecciona:

#### **Analizar**

#### **Comparar Medias**

#### **ANOVA de un factor**

**Factor: edad 3 grupos Dependientes:  
Total autoestima**

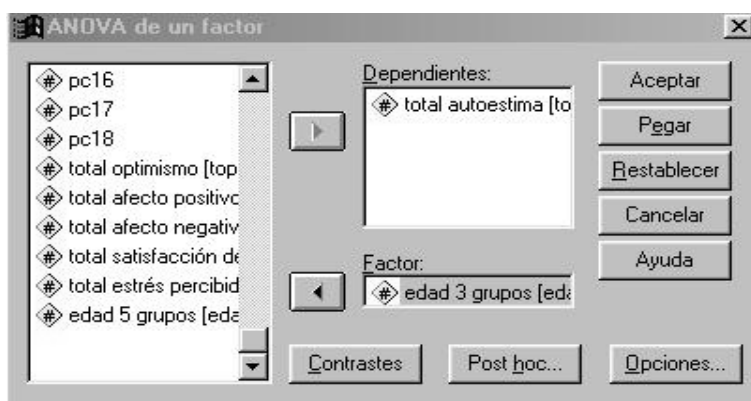


Figura 5.7. Cuadro de diálogo para definir la ANOVA de un factor.



En la ventana *ANOVA de un factor* se selecciona la variable *edadgp3* (tres grupos de edades) y se coloca en el espacio para Factor. Se selecciona la variable *total autoestima* como *Dependiente*. Se activa la opción *Post hoc*. Aparecerá la siguiente ventana.

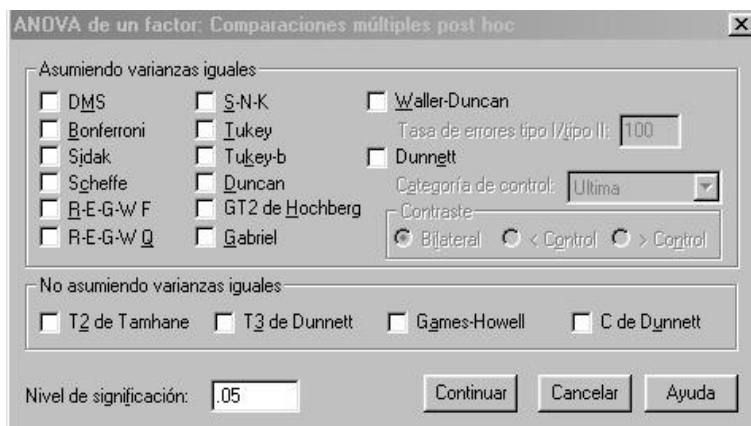


Figura 5.8.. Cuadro definición comparaciones post hoc.

En la ventana *ANOVA de un factor: Post Hoc Comparaciones Múltiples* se selecciona *Tukey* en la sección *Asumiendo varianzas iguales*. En nivel de significación, seleccionar *.05*. Se activa *Continuar* para regresar a la ventana *\_ANOVA de un factor*. En seguida se activa *Opciones* y aparece el siguiente recuadro:

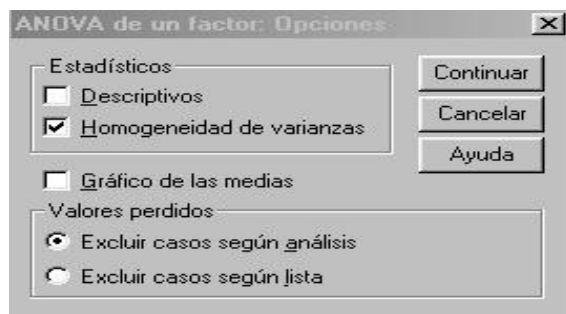


Figura 5.9. Cuadro de definición de varianzas.

En la ventana *ANOVA de un factor: Opciones*, en la sección Estadísticos, seleccionar Homogeneidad de varianzas y en la sección Valores Perdidos seleccionar Excluir casos según análisis. Se activa Continuar para regresar a la ventana *ANOVA de un factor*. El resultado es el siguiente:

**Prueba de homogeneidad de varianzas**

total autoestima

Estadístico de Levene	gl1	gl2	Sig.
1.303	2	433	.273

ANOVA

total autoestima

	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Inter-grupos	258.075	2	129.038	4.505	.012
Intra-grupos	12402.475	433	28.643		
Total	12660.550	435			

Comparaciones múltiples

Variable dependiente: total autoestima  
HSD de Tukey

(I) edad 3 grupos	(J) edad 3 grupos	Diferencia de medias (I-J)	Error típico	Sig.	Intervalo de confianza al 95%	
					Límite inferior	Límite superior
18-29	18-29					
	30-44	-.99	.62	.245	-2.43	.46
	45+	-1.91*	.64	.008	-3.40	-.42
30-44	18-29	.99	.62	.245	-.46	2.43
	30-44					
	45+	-.92	.63	.315	-2.40	.57
45+	18-29	1.91*	.64	.008	.42	3.40
	30-44	.92	.63	.315	-.57	2.40
	45+					

\*. La diferencia entre las medias es significativa al nivel .05.

total autoestima

HSD de Tukey

a,b

edad 3 grupos	N	Subconjunto para alfa = .05	
		1	2
18-29	149	32.60	
30-44	152	33.59	33.59
45+	135		34.50
Sig.		.258	.310

Se muestran las medias para los grupos en los subconjuntos homogéneos.

- a. Usa el tamaño muestral de la media armónica = 144.943.
- b. Los tamaños de los grupos no son iguales. Se utilizará la media armónica de los tamaños de los grupos. Los niveles de error de tipo I no están garantizados.

Cuadro 5.5. Resultados análisis de varianza en SPSS.

Los resultados anteriores permiten no rechazar la hipótesis de igualdad de varianzas de los puntajes de autoestima, entre los grupos de edad (Levene=1,303 con valor  $p=0,273$ , mayor que 0,05). Los resultados de ANOVA permiten rechazar la hipótesis que los promedios de puntajes en los distintos niveles de edad son iguales ( $F=4,505$  con valor  $p=0,012$ , menor que 0,05).

Los resultados de las comparaciones múltiples, según el método de Tukey, señalan que sólo hay una diferencia significativa y es entre los puntajes medios de autoestima para los grupos de 45 y más años y 18 a 29 años ( $\text{dif}=-1,91$  con valor  $p=0,008$ , menor que 0,05). La diferencia  $-1,91$  indica que el grupo de 45 y más años tiene un puntaje medio superior al de 18 a 29 años.

A continuación se presentan otros métodos de comparaciones múltiples post-hoc. Los más usados son los de Scheffé, Tukey, y Newman-Keuls (S-N-K). Todos ellos corrigen el error tipo I cuando se realizan múltiples comparaciones (disminuyen la probabilidad de que aparezca como significativa una diferencia que no lo es). La mayor corrección de este error se consigue, utilizando el contraste de Bonferroni.

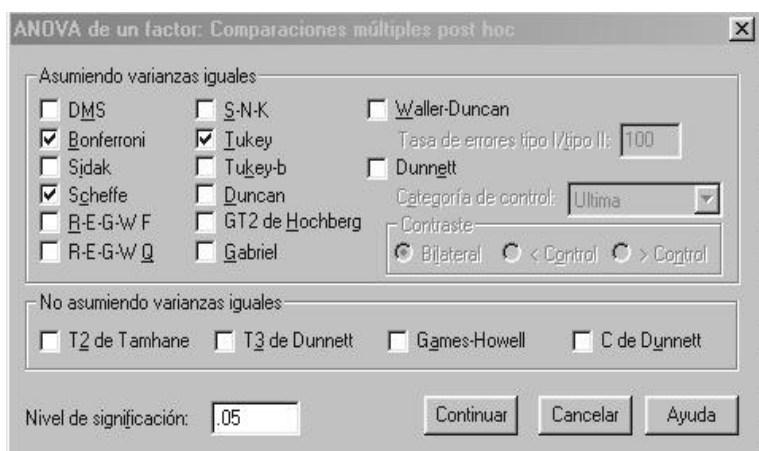


Figura 5.10. Contrastes de Bonferroni, Scheffe y Tukey.

Comparaciones múltiples

Variable dependiente: total autoestima

		Diferencia de medias (I-J)	Error típico	Sig.	Intervalo de confianza al 95%	
(I) edad 3 grupos	(J) edad 3 grupos				Límite inferior	Límite superior
HSD de Tukey	18-29	18-29				
		30-44	-.99	.62	.245	-.243
		45+	-1.91*	.64	.008	-.340
	30-44	18-29	.99	.62	.245	-.46
		30-44				2.43
		45+	-.92	.63	.315	-.240
	45+	18-29	1.91*	.64	.008	.42
		30-44	.92	.63	.315	-.57
Scheffé	18-29	18-29				
		30-44	-.99	.62	.278	-.250
		45+	-1.91*	.64	.012	-.347
	30-44	18-29	.99	.62	.278	-.53
		30-44				2.50
		45+	-.92	.63	.350	-.247
	45+	18-29	1.91*	.64	.012	.34
		30-44	.92	.63	.350	-.64
Bonferroni	18-29	18-29				
		30-44	-.99	.62	.330	-.247
		45+	-1.91*	.64	.009	-.343
	30-44	18-29	.99	.62	.330	-.49
		30-44				2.47
		45+	-.92	.63	.443	-.244
	45+	18-29	1.91*	.64	.009	.38
		30-44	.92	.63	.443	-.60

\*. La diferencia entre las medias es significativa al nivel .05.

total autoestima

edad 3 grupos		N	Subconjunto para alfa = .05	
			1	2
HSD de Tukey <sup>a,t</sup>	18-29	149	32.60	
	30-44	152	33.59	33.59
	45+	135		34.50
	Sig.		.258	.310
Scheffé <sup>a,b</sup>	18-29	149	32.60	
	30-44	152	33.59	33.59
	45+	135		34.50
	Sig.		.292	.345

Se muestran las medias para los grupos en los subconjuntos homogéneos.

- a. Usa el tamaño muestral de la media armónica = 144.943.
- b. Los tamaños de los grupos no son iguales. Se utilizará la media armónica de los tamaños de los grupos. Los niveles de error de tipo I no están garantizados.

Cuadro 5.6. Resultados de la ejecución de las pruebas de contraste.

En el cuadro anterior se puede observar que, con cualquiera de los contrastes post-hoc utilizados, las diferencias significativas son las mismas indicadas antes.

## CAPÍTULO VI

---

### ANÁLISIS BIVARIADO

La estadística bivariada analiza la relación entre dos variables. Estas dos variables pueden estar al mismo nivel, o bien una de ellas puede ser categórica y la otra cuantitativa.

#### *ASOCIACIÓN Y PRUEBA DE CHI-CUADRADO*

Cuando es necesario describir la relación entre dos variables cualitativas, el formato preferido de presentación de la información es la «tabla de contingencia, bivariable o bidimensional». Hay tres aspectos relacionados con el proceso de describir la asociación entre dos variables. Primero, el concepto de asociación mismo y las características de una asociación. Segundo, el proceso de construcción y diseño de tablas cruzadas. Tercero, el análisis de tablas cruzadas y la necesidad de desarrollar una medida que describa la asociación entre las variables tabuladas y sus características.

#### *Concepto de Asociación*

Se dice que existe asociación entre dos variables cuando las distribuciones de una variable no son iguales a lo largo de las distribuciones de la otra variable.

Existen, por lo menos, tres formas de determinar si hay asociación en la tabla. Por supuesto, todas ellas están relacionadas entre sí. Cuando se trabaja con dos variables a la vez, se dice que

se está trabajando con «distribuciones de frecuencias condicionales». Con esto se quiere decir que una tabla nos permite examinar no sólo la distribución general de una variable cualitativa, sino que también nos permite ver las condiciones que afectan como esa variable está influida por otra.

### ***Características de la Asociación***

Una asociación puede tener las siguientes características: *existencia, intensidad, dirección y forma*. Examinemos cada una de éstas.

1. **Existencia** – Se dice que existe asociación si las distribuciones de una variable son o no iguales a lo largo de las distribuciones de la otra variable. Se utilizan para establecer asociación: el porcentual y el modelo de independencia o no asociación.
2. **Intensidad** - Se dice que la asociación en la tabla es fuerte cuando la variable dependiente se distribuye de manera diferente entre las distintas condiciones de la variable independiente. Cuando esto no ocurre, esto es, cuando los valores de la variable dependiente se distribuyen aproximadamente igual entre los valores de la variable independiente, la asociación es débil. En el caso de los datos que se examinan previamente, el género de la persona no influye sobre la preferencia religiosa.
3. **Dirección de la asociación** - Cuando las variables en la tabla son por lo menos ordinales, podemos hablar de la «dirección de la asociación». Cuando las frecuencias de las categorías altas de la variable independiente se asocian con las frecuencias altas (o las bajas con las bajas) de la variable dependiente, se dice que la asociación es positiva. Cuando las frecuencias bajas se asocian con las altas, se dice que la asociación es negativa. Por ejemplo, normalmente esperamos que a más alto es

la educación de una persona, mayor sea su ingreso. Esta es una relación positiva. Por el otro lado, también esperamos que el tamaño de la familia decrezca con el nivel de ingreso. A mayor ingreso, menos hijos se tienen. Esta es una relación negativa.

- 4. Forma de la asociación** - La asociación en una tabla puede ser de dos tipos: lineal o curvilínea. En el primer caso, se observa un patrón escalonado de progresión en la concentración de casos en la variable dependiente, según nos movemos a las categorías más altas de la variable independiente. En la relación curvilínea, se observa primero una progresión escalonada según nos movemos a las categorías más altas de la variable independiente y luego un descenso escalonado en el patrón de los casos. En ambos casos, sería posible tirar una línea a lo largo de las categorías de la tabla para detectar el patrón. La mejor forma de hacer esto es subrayando los por cientos más altos en cada fila de la variable dependiente y ver que dirección toman.

A menudo, hay que trabajar con variables categóricas y se puede estimar necesaria hacer preguntas que relacionen las variables. Así por ejemplo, se puede preguntar si hay diferencias de autoestima en personas de diferente origen social. En una encuesta se puede examinar diferentes opiniones hacia el aborto y examinar si existen diferencias entre hombres y mujeres. Para examinar este tipo de relación se emplea la prueba de Chi-cuadrado y se aplican los siguientes supuestos:

1. Se categorizan sujetos con respecto a una sola variable que tiene dos o más categorías.
2. Cada sujeto aparece una sola vez y en una sola categoría.
3. Cada asignación a una categoría es independiente de cualquier otra asignación. (El que un sujeto se asigne a

una categoría no tiene nada que ver con cómo se asigna otro sujeto).

- 4. Los cálculos se hacen con todos los sujetos del estudio.
- 5. La frecuencia esperada ( $f_e$ ) en cada celda es igual o mayor de 5.

**Ejemplo**

Se desea estudiar si la preferencia religiosa de una persona (variable dependiente) está afectada por su género. La teoría que se sustentaría sería que los hombres tienen preferencias religiosas distintas a las mujeres. Veamos los siguientes datos de prueba para 50 sujetos a los que se le ha preguntado su religión (c=católico, p=protestante) y su género (m=mujer, h=hombre). La secuencia que se activa es:

**Analizar**

**Estadística descriptiva**

**Tablas de contingencia**

**Sexo (columnas) Religiosidad (filas)**

**Casillas Columna**

**Estadístico Chi-cuadrado**

Tabla de contingencia Religiosidad \* Sexo del sujeto

			Sexo del sujeto		Total
			h	m	
Religiosidad	c	Recuento	9	13	22
		% de Sexo del sujeto	47.4%	41.9%	44.0%
	p	Recuento	10	18	28
		% de Sexo del sujeto	52.6%	58.1%	56.0%
Total	Recuento		19	31	50
	% de Sexo del sujeto		100.0%	100.0%	100.0%



Pruebas de chi-cuadrado

	Valor	gl	Sig. asintótica (bilateral)	Sig. exacta (bilateral)	Sig. exacta (unilateral)
Chi-cuadrado de Pearson	.141 <sup>b</sup>	1	.707		
Corrección por continuidad <sup>a</sup>	.007	1	.935		
Razón de verosimilitud	.141	1	.707		
Estadístico exacto de Fisher				.774	.466
Asociación lineal por lineal					
N de casos válidos	50				

a. Calculado sólo para una tabla de 2x2.

b. 0 casillas (.0%) tienen una frecuencia esperada inferior a 5. La frecuencia mínima esperada es 8.36.

### Cuadro 6.1 Pruebas de Chi-cuadrado.

El analista llega a la conclusión de que para esta muestra no se podría sustentar que los hombres tienen preferencias religiosas distintas a las mujeres.

### *Aplicación al Caso en Estudio*

En el caso del archivo de nuestro caso de estudio se puede relacionar la variable *sexo* con la variable dicotómica *fuma/no fuma*. Entonces, supongamos que queremos determinar si existe una relación estadísticamente significativa entre el sexo del sujeto y si fuma o no fuma ya que en el libro de códigos, el investigador se ha planteado la hipótesis alterna que vincula estas dos variables. Ambas variables son categóricas.



Figura 6.1. Cuadro de diálogo de SPSS para Tablas de Contingencia.

Tabla de contingencia sexo \* fumador(a)

Recuento		fumador(a)		Total
		sí	no	
sexo	hombres	33	151	184
	mujeres	52	200	252
Total		85	351	436

Cuadro 6.2. Resultado Tabla de Contingencia:  
sexo\*fumador(a).

La relación entre las dos variables se realiza aplicando una prueba de asociación para variables categóricas. La más usado es la chi-cuadrado y la transformación de este en el *coeficiente de contingencia*. El coeficiente de contingencias se rige por las mismas reglas de la correlación y las mediciones del índice correspondiente fluctúan entre + 1 y - 1, pasando por el cero, donde éste último significa que no hay correlación entre las variables estudiadas y los dos primeros la correlación máxima.

La prueba chi-cuadrado es del tipo no paramétrica y por lo tanto, no requiere de presupuestos acerca de la distribución poblacional (aceptan distribuciones no normales). Si las variables son dicotómicas puede utilizarse también el coeficiente de correlación phi. El estadístico chi-cuadrado no debe utilizarse si hay alguna frecuencia esperada menor que 1, o si el 20% de las celdas tiene frecuencias esperadas menores que 5.

Pruebas de chi-cuadrado

	Valor	gl	Sig. asintótica (bilateral)	Sig. exacta (bilateral)	Sig. exacta (unilateral)
Chi-cuadrado de Pearson	.494 <sup>a</sup>	1	.482		
Corrección por continuidad <sup>b</sup>	.337	1	.562		
Razón de verosimilitud	.497	1	.481		
Estadístico exacto de Fisher				.541	.282
Asociación lineal por lineal	.493	1	.483		
N de casos válidos	436				

a. Calculado sólo para una tabla de 2x2.

b. 0 casillas (.0%) tienen una frecuencia esperada inferior a 5. La frecuencia mínima esperada es 35.87.

Cuadro 6.3. Resultados Pruebas de chi-cuadrado.

Medidas simétricas

		Valor	Error tí. asint. <sup>a</sup>	T aproximada <sup>b</sup>	Sig. aproximada
Nominal por nominal	Coefficiente de contingencia	.034			.482
N de casos válidos		436			

a. Asumiendo la hipótesis alternativa.

b. Empleando el error típico asintótico basado en la hipótesis nula.

### Cuadro 6.4. Medidas simétricas.

Si el valor de chi-cuadrado resulta estadísticamente significativo ( $p < 0,05$ ), se puede afirmar con un nivel de confianza del 5% que las dos variables están relacionadas. En el ejemplo, el valor de chi-cuadrado es 0,494 y no es significativo pues su valor  $p = 0,482$  es mayor que 0,05. Por lo tanto, no se rechaza la hipótesis nula: ambas variables no están relacionadas. Esto significa que las personas fumen o no, no depende del sexo.

Ahora examinaremos si existe una relación entre el *sexo* del sujeto (variable categórica) y la *fuerza de estrés* (variable categórica):

Tabla de contingencia sexo \* fuerza del estrés

Recuento		fuente del estrés								Total
		trabajo	esposa	relaciones	niños	familia	salud/enfermedad	vida en general	8	9
sexo	hombres	110	3	5	5	6	6	13	24	7
	mujeres	114	9	7	20	21	14	19	30	9
Total		224	12	12	25	27	20	32	54	16
										422

Pruebas de chi-cuadrado

	Valor	gl	Sig. asintótica (bilateral)
Chi-cuadrado de Pearson	16.657 <sup>a</sup>	8	.034
Corrección por continuidad			
Razón de verosimilitud	17.607	8	.024
Asociación lineal por lineal	2.556	1	.110
N de casos válidos	422		

a. 0 casillas (.0%) tienen una frecuencia esperada inferior a 5. La frecuencia mínima esperada es 5.09.

Medidas simétricas

		Valor	Error tí. asint. <sup>a</sup>	T aproximada <sup>b</sup>	Sig. aproximada
Nominal por nominal	Coefficiente de contingencia	.195			.034
N de casos válidos		422			

a. Asumiendo la hipótesis alternativa.

b. Empleando el error típico asintótico basado en la hipótesis nula.

### Cuadro 6.5. Resultados relaciones entre variables.

El valor de chi-cuadrado resulta estadísticamente significativo ( $p < 0,05$ ) a un nivel del 5%. Por lo tanto, se rechaza la hipótesis nula, ambas variables están relacionadas, la fuente de estrés está relacionada con el sexo.

En términos de precauciones con respecto al uso de esta prueba, habría que señalar que hay que contar con un tamaño de muestra lo bastante grande para garantizar la semejanza entre la distribución teóricamente correcta y la distribución de muestreo. Cuando las frecuencias esperadas resultan demasiado pequeñas, el valor de la chi-cuadrado será sobrestimado y dará origen a rechazos de la hipótesis nula.

## CORRELACIÓN LINEAL

Cuando se tienen observaciones formadas por una pareja de valores a partir de dos variables, surge la pregunta o necesidad de conocer acerca de si las variables estarán o no relacionadas y qué tan fuerte es esa relación. Para saber esto, generalmente se piensa en obtener un coeficiente de correlación que nos indique el grado de relación lineal entre las variables, pero debe tenerse cuidado de no interpretarlo como una medida de la relación causal entre las variables y también tomar en cuenta que si la relación no es lineal, la correlación no detecta la relación existente entre las variables.

En la estadística paramétrica, se puede calcular el coeficiente de correlación de Pearson, que se puede aplicar a variables que se miden en escala de intervalo o razón, pero bajo el supuesto de que los datos se distribuyen en base a una distribución normal bivariada, pero cuando esto no se cumple, o cuando la escala de medida solamente es ordinal, es preferible usar una medida de asociación de las que se tienen en la estadística no paramétrica.

El coeficiente de correlación lineal mide el grado de relación lineal que existe entre dos variables. Supongamos que queremos

estudiar la correlación existente entre peso y altura de un grupo de personas tomadas al azar. Sometemos los datos recogidos de peso y altura al análisis de correlación y encontramos el coeficiente de correlación entre ambas, que se representa con la letra  $r$ . Suponiendo que  $r = 0.78$ . Esto significa que a mayor altura, correspondería mayor peso.

El coeficiente de correlación de Pearson ( $r$ ) es un índice que mide la magnitud de la relación lineal entre 2 variables cuantitativas, así como el sentido, positivo o negativo, de dicha relación. Indica en qué grado, 2 variables  $X$  e  $Y$  fluctúan simultáneamente, es decir cuánto aumenta  $X$  al aumentar  $Y$  (*correlación positiva*), o cuánto aumenta  $X$  al disminuir  $Y$  (*correlación negativa*). A diferencia de la regresión lineal, el coeficiente de correlación no presupone dependencia de una variable respecto a la otra;  $X$  e  $Y$  se sitúan a un mismo nivel. Asimismo, la existencia de correlación lineal entre 2 variables no implica necesariamente una relación causal entre ellas, sino que se limita a explicar su covariación.

Los coeficientes de correlación  $r$  siempre oscilan entre valores de  $-1$  y  $1$ . El valor cero ( $0$ ) significa que las variables son independientes, esto es, no existe relación lineal entre ellas. Un valor positivo indica que a incrementos en la variable  $X$ , se producen incrementos proporcionales en  $Y$ , mientras que un valor negativo indica que al aumentar  $X$ , la variable  $Y$  tiende a disminuir.

Para interpretar el coeficiente de correlación, *Colton* ha dado los siguientes lineamientos generales:

- Valor de  $r$  de  $0$  a  $0.25$  implica que no existe correlación entre ambas variables.
- Valor de  $r$  de  $0.25$  a  $0.50$  implica una correlación baja a moderada.
- Valor de  $r$  de  $0.50$  a  $0.75$  implica correlación moderada a buena.

- Valor de  $r$  de 0.75 o mayor, implica una muy buena a excelente correlación.
- Estos rangos de valores se pueden extrapolar a correlaciones negativas también.

Se debe tener cuidado al analizar la correlación entre dos variables, de que ambas varíen juntas permanentemente. Esto parece redundante, pero es importante. Por ejemplo, si correlacionamos edad y altura. La altura irá aumentando con la edad hasta un determinado punto en donde ya no aumentará más.

Un coeficiente de correlación que se basa en rangos y que es muy utilizado, es el de *Spearman*. Spearman desarrolló un trabajo en 1940 donde presentó este coeficiente que en lugar de utilizar los valores de las variables, utilizaba los rangos asociados a ellas, mediante éste se tiene una medida de asociación y además permite probar hipótesis; el único supuesto que tiene, es que la escala de medida de la variable es al menos ordinal.

Además de este coeficiente que nos permite medir la asociación entre dos variables, hay otras medidas de asociación para aquellos casos en los que la escala con la que se miden las variables es de otro tipo.

### ***Aplicación al caso en Estudio***

Para analizar la relación lineal entre dos variables cuantitativas, el estadístico a usar es el de correlación de Pearson; previamente hay que comprobar que corresponden a distribuciones normales (para lo cual aplicamos la prueba K-S, de Kolmogorov-Smirnov). Por ejemplo, este procedimiento se aplicará si estamos interesados en examinar si existe relación entre el puntaje total obtenido en la escala de estrés y el puntaje total obtenido en la escala para medir el afecto negativo. Previamente veamos sus distribuciones, elaborando los histogramas respectivos.

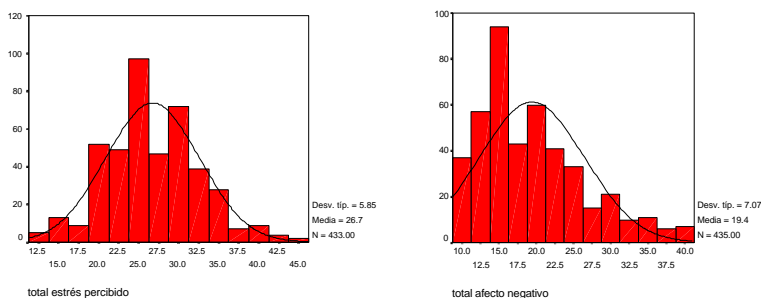


Gráfico 5.1. Histogramas variables cuantitativas

Podemos ver en el gráfico 5.1 que una de las variables tiene cierto grado de sesgo y por lo tanto aplicamos la prueba K-S para examinar la distribución. La hipótesis nula es que la distribución de la población de la cual se ha extraído la muestra es normal.

Prueba de Kolmogorov-Smirnov para una muestra

		total estrés percibido	total afecto negativo
N		433	435
Parámetros normales <sup>a,b</sup>	Media	26.73	19.40
	Desviación típica	5.85	7.07
Diferencias más extremas	Absoluta	.069	.123
	Positiva	.069	.123
	Negativa	-.033	-.092
Z de Kolmogorov-Smirnov		1.439	2.560
Sig. asintót. (bilateral)		.032	.000

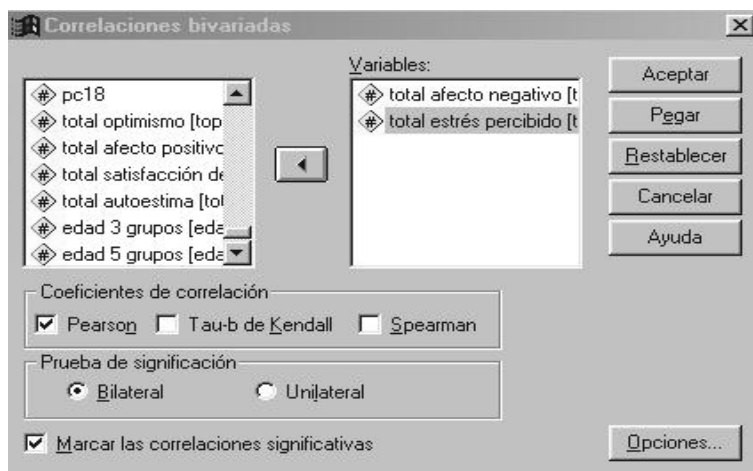
a. La distribución de contraste es la Normal.

b. Se han calculado a partir de los datos.

Cuadro 6.6. Resultados pruebas K-S.

El resultado revela que la significación que nos da el K-S es menor que 0,05 en ambas variables, por lo cual tenemos que rechazar la hipótesis nula. Por lo que, no se cumpliría uno de los supuestos paramétricos para calcular la correlación de Pearson, esto es, la normalidad de las variable.

A pesar de los resultados anteriores, para ejemplificar, se examinará la relación entre el estrés percibido y el afecto negativo, puesto que en nuestro plan de análisis, se considera el estudio de esta hipótesis. Para este efecto, la secuencia de pasos es la siguiente:

**Analizar,****Correlaciones****Bivariadas***Total afecto negativo. Total estrés percibido***Pearson**

Cuadro 6.7. Resultados correlaciones bivariadas.

Este resultado indica que existe una correlación positiva y fuerte que es estadísticamente significativa al nivel 0,01 con una prueba bilateral.

***Diagrama de Dispersión***

El analista también explora los datos a través de un diagrama de dispersión. Selecciona *Gráfico, Dispersión* y aparecerá la siguiente pantalla:



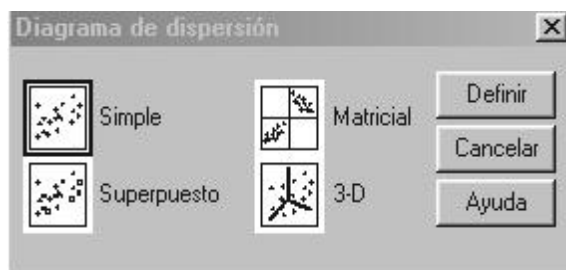


Figura 6.3. Cuadro definición gráfico de dispersión.

En esta ventana se selecciona Simple/Definir y se procede a elegir las variables a considerar en el diagrama, en este caso, el total de la escala de optimismo, el total de la escala de autoestima, y el sexo de los sujetos.



Figura 6.4. Cuadro de diálogo Diagrama de dispersión simple.

El paso siguiente es seleccionar en esta ventana Títulos:

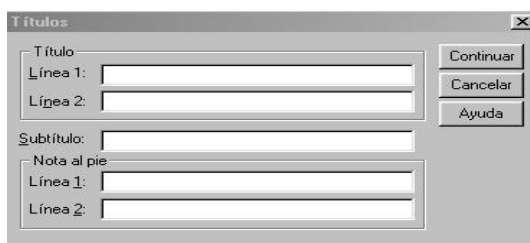


Figura 6.5. Cuadro definición de títulos.

El resultado será el siguiente:



Figura 6.6. Cuadro definición de marcador.

### Diagrama de Dispersión

Total Escala Autoestima - Total Escala Optimismo

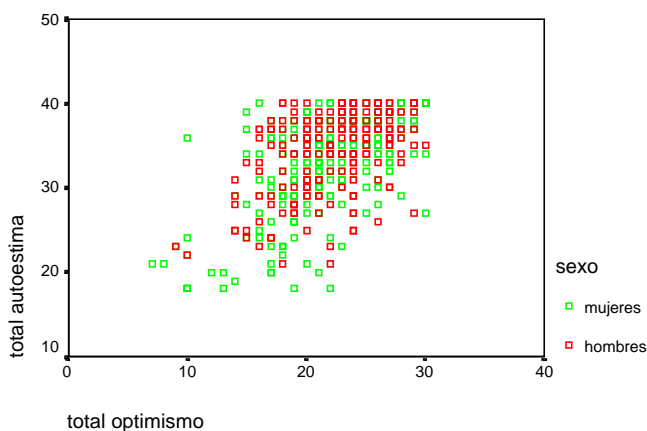


Gráfico 6.2. Gráfico de Dispersión.

Si se desea editar el gráfico, se oprime el ratón dos veces dentro del recuadro donde se encuentra la gráfica. Aparecerá una nueva ventana: el editor de gráficos que permite modificar la gráfica.

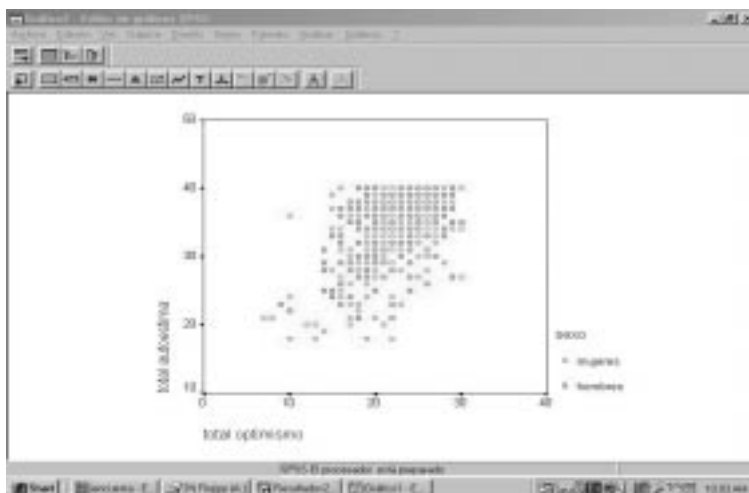


Figura 6.7. Editor de gráficos.

Se selecciona la marca correspondiente al sexo femenino, y se activa Formato y Marcadores. Aparecerá la siguiente ventana:

Se selecciona el *marcador requerido para cada uno de las categorías que identifican la variable sexo*. Para lograr este efecto, se activa una categoría específica, por ejemplo mujeres. El analista puede en seguida seleccionar un color específico del menú Formato/Color:



Figura 6.8. Cuadro definición de colores.

Se aplica el mismo procedimiento para la categoría hombres (seleccionar marcador y color específico). El resultado será el siguiente:

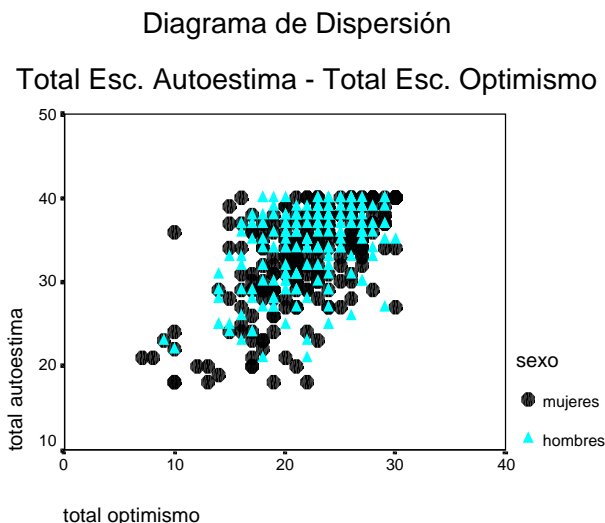


Gráfico 6.3. Gráfico de dispersión con marcadores nuevos.

### ***Coefficientes de Correlación por Rangos***

Existen dos métodos para calcular el coeficiente de correlación de los rangos: uno señalado por Spearman y otro por Kendall. El coeficiente de Spearman es más fácil de calcular que el de Kendall. El coeficiente de correlación de Spearman es exactamente el mismo que el coeficiente de correlación de Pearson calculado sobre el rango de observaciones. En definitiva, la correlación estimada entre X e Y se halla calculado el coeficiente de correlación de Pearson para el conjunto de rangos apareados. El coeficiente de correlación de Spearman es recomendable utilizarlo cuando los datos presentan valores extremos ya que

dichos valores afectan mucho el coeficiente de correlación de Pearson, o ante distribuciones no normales.

El índice Tau-b de Kendall, al igual que el coeficiente de correlación de Spearman, está indicado en el caso en que se quiera medir la asociación existente en dos variables X e Y categóricas, con categorías codificadas de forma ordinal. Sin embargo, mide dicha asociación de distinta manera a como lo hace el coeficiente de correlación de Spearman. El índice Tau-a de Kendall está basado en la noción de concordancia, teniendo en cuenta que no se realiza corrección por empates. Su interpretación es la diferencia entre la proporción de pares concordantes y de pares discordantes.

En nuestro caso de estudio, se agrupó la variable edad en 3 grupos de edades. A nivel de exploración de los datos, interesaba examinar la relación entre los grupos de edades de los sujetos (3 grupos) y su condición de fumador(a) (1=Sí; 2=No). Para realizar este tipo de análisis, se calculó los coeficientes Tau-b de Kendall y Spearman en la opción Correlaciones.

La secuencia de pasos en SPSS es la siguiente:

**Analizar**

**Correlaciones**

**Bivariadas**

***Edad 3 grupos [edadgr3] Fumador(a) [fumador]***



Figura 6.9. Cuadro definición de Correlación Bivariada.

Correlaciones				
			edad 3 grupos	fumador(a)
Tau_b de Kendall	edad 3 grupos	Coefficiente de correlación	1.000	.097*
		Sig. (bilateral)	.	.031
		N	439	436
	fumador(a)	Coefficiente de correlación	.097*	1.000
		Sig. (bilateral)	.031	.
		N	436	436
Rho de Spearman	edad 3 grupos	Coefficiente de correlación	1.000	.103*
		Sig. (bilateral)	.	.031
		N	439	436
	fumador(a)	Coefficiente de correlación	.103*	1.000
		Sig. (bilateral)	.031	.
		N	436	436

\*. La correlación es significativa al nivel 0,05 (bilateral).

Cuadro 6.8. Resultados correlaciones entre variables.

Puesto que para todos los grupos las correlaciones son significativas ( $p < 0,05$ ), se puede afirmar que existe relación estadísticamente significativa entre ambas variables. Se aplicaron otros coeficientes para analizar la situación. Se recurrió al Chi-Cuadrado de Mantel-Haenszel y al coeficiente Gamma de Goodman y Kruskal. Se han utilizado estos estadísticos de prueba porque se desea contrastar la hipótesis de independencia entre el grupo de edad y su condición de fumador y además las variables utilizadas poseen pocas categorías.

**Tabla de contingencia edad 3 grupos \* fumador(a)****Recuento**

		fumador(a)		Total
		sí	no	
edad 3	18-29	34	114	148
grupos	30-44	34	119	153
	45+	17	118	135
Total		85	351	436

**Pruebas de chi-cuadrado**

	Valor	gl	Sig. asintótica (bilateral)
Chi-cuadrado de Pearson	5.964 <sup>a</sup>	2	.051
Corrección por continuidad			
Razón de verosimilitud	6.349	2	.042
Asociación lineal por lineal	4.707	1	.030
N de casos válidos	436		

a. 0 casillas (.0%) tienen una frecuencia esperada inferior a 5. La frecuencia mínima esperada es 26.32.

**Medidas simétricas**

	Valor	Error típ. asint <sup>a</sup>	aproximada <sup>b</sup>	Sig. aproximada
Ordinal por ordir Gamma	.213	.092	2.248	.025
N de casos válidos	436			

a. Asumiendo la hipótesis alternativa.

b. Empleando el error típico asintótico basado en la hipótesis nula.

**Cuadro 6.9. Resultados Correlaciones.**

A partir de estos resultados, y con un nivel de significación del 5%, se puede afirmar que existe relación lineal entre ambas variables (Asociación lineal por lineal = 4,707,  $p = 0,03$  y Gamma = 0,213 con  $p = 0,025$ ).

## REGRESIÓN LINEAL

Muchas veces en un estudio no sólo basta con conocer si existe o no una relación entre dos variables sino que también interesa conocer el tipo de relación existente y la predicción de una con respecto a la otra.

Representamos en un gráfico los pares de valores de una distribución bidimensional: la variable «x» en el eje horizontal o eje de abscisas, y la variable «y» en el eje vertical, o eje de ordenadas.

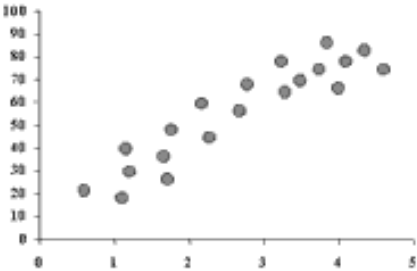


Gráfico 6.4. Gráfico de dispersión.

El *coeficiente de correlación lineal* nos permite determinar si, efectivamente, existe relación entre las dos variables. Una vez que se concluye que sí existe relación, la *regresión* nos permite definir el modelo lineal simple (recta) que mejor se ajusta a esta nube de puntos.

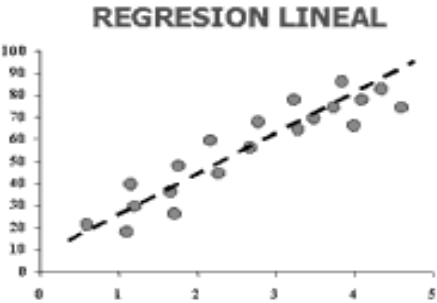


Gráfico 6.5. Gráfico de regresión lineal.



Una recta viene definida por la siguiente fórmula:

$$y = a + bx$$

Donde «y» sería la variable dependiente, es decir, aquella que viene definida a partir de la otra variable «x» (variable independiente). Para definir la recta, hay que estimar los valores de los parámetros «a» y «b»:

El *parámetro* «a» es el “intercepto” valor que toma la variable dependiente «y», cuando la variable independiente «x» vale 0, y es el punto donde la recta cruza el eje vertical.

El *parámetro* «b» determina la pendiente de la recta, su grado de inclinación.

El Procedimiento *Analizar/Regresión/Lineal* de SPSS nos permite estimar el valor de estos dos parámetros, definiendo la recta que mejor se ajusta a esta nube de puntos.

### ***Ejemplo***

En el siguiente ejemplo se desea estudiar que influencia tiene la “edad” en la “resistencia al ejercicio físico”, tomada una muestra de 20 sujetos a los cuales se les aplicó un test de resistencia física, midiendo el tiempo (segundos) que duraban sin fatigarse.

Sujeto	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Edad	19	21	21	22	22	22	23	23	24	24	24	25	26	27	27	31	32	34	35	35
Resistencia	41	43	42	45	42	40	43	44	41	42	45	50	55	47	49	49	51	60	53	56

La primera pregunta que debemos formularnos es ¿Están relacionadas las variables de edad y resistencia física?. Para responder a esta interrogante se debe realizar una análisis del Diagrama de dispersión, seleccionando:

**Gráficos**

**Interactivos**

**Diagramas de Dispersión**

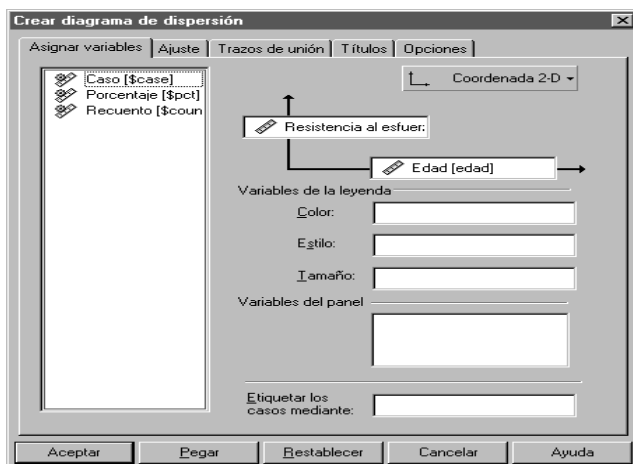


Figura 6.9. Cuadro de diálogo para Crear diagrama de dispersión.

Se ingresa la variable “Edad” en el eje horizontal y la variable “Resistencia al esfuerzo físico” en el eje vertical.

Seleccionado en la opción *Ajuste*, el método de regresión el cual nos muestra la recta ajustada:

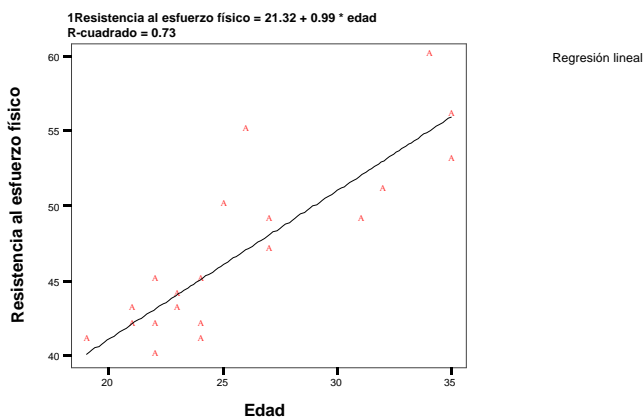


Gráfico 6.6. Diagrama con la recta ajustada.

El diagrama nos induce a pensar que podría haber una relación entre edad y resistencia física y que ésta debiera ser lineal.

Por lo tanto, se ajustará un modelo lineal simple, en el que además se comprobarán los supuestos para su aplicación, los cuales son:

- Aleatoriedad de la muestra
- Linealidad
- Homogeneidad de varianzas
- Normalidad

Para efectuar la regresión lineal, seleccionar *Analizar/Regresión/Lineal* y *aparecerá la siguiente pantalla*:



Figura 6.10. Cuadro Regresión lineal.

Se ingresa como variable dependiente “Resistencia al esfuerzo físico” y como independiente la variable “Edad”. Seleccionar *Gráficos* que entrega la siguiente ventana:



Figura 6.11. Cuadro gráfico para la regresión lineal.

Seleccionar la opción “Gráfico de probabilidad normal” y como eje horizontal (X) \*ZPRED (estimaciones estandarizados) y en el eje vertical (Y) \*ZRESID (residuos estandarizados), dando como resultado del ajuste:

Resumen del modelo<sup>b</sup>

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	.853 <sup>a</sup>	.728	.713	3.08

a. Variables predictoras: (Constante), Edad

b. Variable dependiente: Resistencia al esfuerzo físico

Cuadro 6.10. Resumen modelo de regresión

El coeficiente de correlación (R), nos confirma que existe una relación lineal entre ambas variables y que esta es fuerte y positiva.

El coeficiente de determinación (R cuadrado) muestra que el 73% de la variación de la variable “Resistencia al esfuerzo físico” está explicada por el modelo de regresión lineal, es decir existe un buen ajuste.

ANOVA<sup>b</sup>

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	456.900	1	456.900	48.123	.000 <sup>a</sup>
	Residual	170.900	18	9.494		
	Total	627.800	19			

a. Variables predictoras: (Constante), Edad

b. Variable dependiente: Resistencia al esfuerzo físico

### Cuadro 6.11. ANOVA.

De la tabla de análisis de varianza (ANOVA), se observa un valor de probabilidad asociada tendiente a cero, lo cual nos lleva a rechazar la hipótesis nula, y por lo tanto, el modelo propuesto da una buena explicación de la variabilidad de la resistencia al esfuerzo físico.

Coeficientes<sup>a</sup>

Modelo		Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.
		B	Error típ.	Beta		
1	(Constante)	21.319	3.751		5.683	.000
	Edad	.990	.143	.853	6.937	.000

a. Variable dependiente: Resistencia al esfuerzo físico

### Cuadro 6.12. Coeficientes.

Con respecto a los estimadores de los coeficientes del modelo de regresión, a(constante) y b (de la edad), los valores p asociados tienden a cero. Por lo tanto, se rechazan las hipótesis nulas, con un nivel de significación del 5%, que los coeficientes son cero, en la población.

Es decir el modelo propuesto es el siguiente:

$$(\text{Resistencia al esfuerzo físico})_{\text{ESTIMADO}} = 21,319 + 0,99 * \text{Edad}$$

Pero a pesar de esto, se deben verificar los supuestos para la aplicación de la regresión.

El supuesto de aleatoriedad de la muestra queda asegurado con una adecuada obtención de las observaciones.

El diagrama que sigue, permite verificar la linealidad y la homogeneidad de varianzas, en el cual se ve la resistencia física al esfuerzo estimada con los residuos, ambos tipificados. Si la nube de puntos se mantiene paralela al eje horizontal y presenta un ancho homogéneo, se cumplirían los supuestos de linealidad e igualdad de varianzas respectivamente; en caso contrario, alguno de ellos está siendo violado.

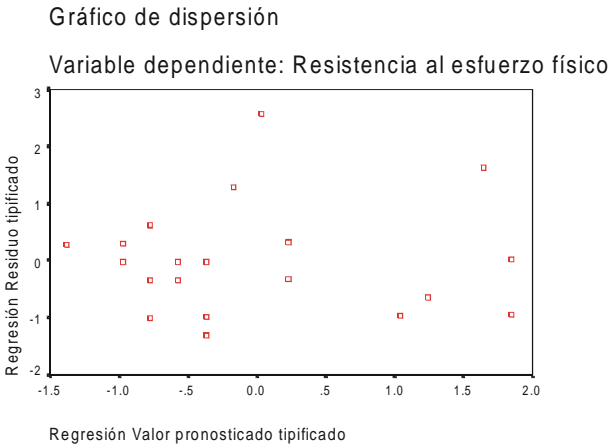


Gráfico 6.7. Residuos

En nuestro ejemplo, se estaría cumpliendo el supuesto de linealidad, pero la anchura de la nube de puntos no es muy homogénea, aunque no exageradamente.

Otros de los supuestos es el de normalidad y para ello se utiliza el gráfico de probabilidad normal, si las observaciones forman una línea recta ajustada a la presentada, existirá normalidad.

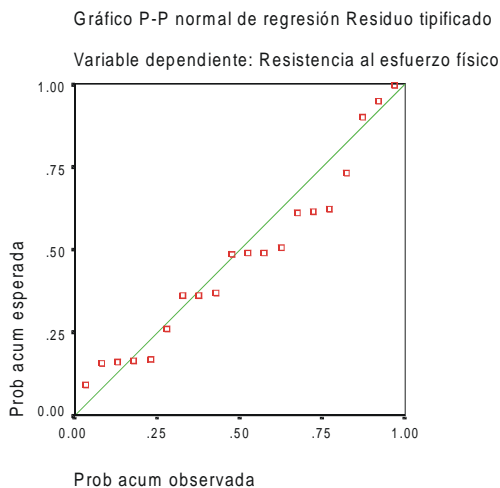


Gráfico 6.8. Gráfico de probabilidad de normalidad.

Se observa una distorsión que pondría en duda el supuesto de normalidad, recomendándose aplicar una prueba como Kolmogorov-Smirnov para verificar normalidad, usando la variable residuos.

Cuando se examina la relación lineal entre dos variables y se concluye que ésta no existe, podemos decir que no existe correlación lineal entre ellas, pero se puede explorar otro tipo de relación o la posibilidad de que esta relación no lineal sea linealizable, por ejemplo si consideramos que la nube de puntos nos induce a pensar que existe una relación exponencial.

### *Aplicación al Caso en Estudio*

Así, en el caso de estudio sobre autoestima, si pensamos que podría haber una relación lineal entre el puntaje total de la escala

de optimismo (variable dependiente) y la edad (variable independiente), entonces al ejecutar la regresión lineal, podemos observar que el coeficiente de correlación de Pearson es bastante bajo, lo que indica una baja correlación lineal entre el puntaje total de la escala de autoestima y la edad.

#### Resumen del modelo<sup>b</sup>

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	.201 <sup>a</sup>	.040	.038	4.34

a. Variables predictoras: (Constante), EDAD

b. Variable dependiente: total optimismo

Cuadro 6.13. Resumen del modelo.

En la tabla de análisis de varianza, se puede observar un valor  $F=18,171$  con un valor  $p=0,000$  que es menor a 0,05. Por lo tanto, se acepta que el modelo propuesto da una buena explicación de la variabilidad de los puntajes de la escala de optimismo.

#### ANOVA<sup>b</sup>

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	342.862	1	342.862	18.171	.000 <sup>a</sup>
	Residual	8170.159	433	18.869		
	Total	8513.021	434			

a. Variables predictoras: (Constante), EDAD

b. Variable dependiente: total optimismo

Cuadro 6.14. ANOVA.

#### Coefficientes<sup>a</sup>

Modelo		Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.
		B	Error típ.	Beta		
1	(Constante)	19.593	.628		31.216	.000
	EDAD	6.747E-02	.016	.201	4.263	.000

a. Variable dependiente: total optimismo

Cuadro 6.15. Coeficientes.



Con respecto a los estimadores de los coeficientes a (constante) y b (de la edad) los valores p asociados son menores que 0,05 y por lo tanto, se rechaza la hipótesis nula de que estos son iguales a cero, con un nivel de significancia del 5%.

Es decir el modelo propuesto es el siguiente:

$$(\text{Puntaje Total de Optimismo})_{\text{ESTIMADO}} = 19,593 + 0,0067 * \text{Edad}$$

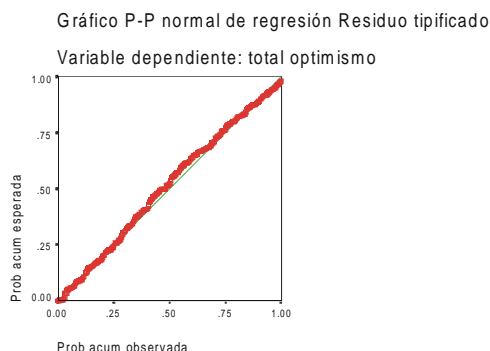


Gráfico 6.9. Gráfico P-P.

El gráfico de probabilidad normal nos muestra que no hay muchos alejamientos de la línea recta, por lo tanto, podría asumirse normalidad.

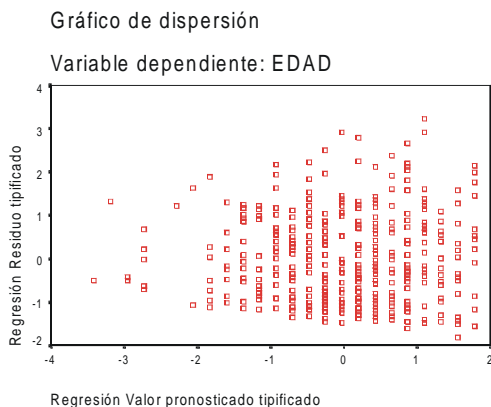


Gráfico 6.10. Gráfico de dispersión.

En el gráfico de los residuos tipificados versus los puntajes de autoestima pronosticados y tipificados, no se observa homogeneidad en la nube de puntos, se recomienda utilizar algún test de homogeneidad de varianzas para confirmarlo.

Sin embargo, a pesar de que se encontró un modelo lineal para ajustar los datos, esto no significa que sea bueno; no hay que olvidarse que el coeficiente de determinación ( $R^2$ ) indica que el 4% de la variación está explicada por el modelo, es decir, el modelo no es de lo mejor.

### Ejercicios

En la página Web <http://members.tripod.com/experimentum/spssds.html#6> aparecen los siguientes datos y las pruebas estadísticas que se aplicaron. Examine la propiedad de los procedimientos aplicados y las interpretaciones que se hace de ellos.

#### ETIQUETAS:

**edocivil** (Estado civil): 1= «casado» , 0= «soltero».

**Antigued** (Antigüedad en el puesto, en años)

**antig2** (Antigüedad en el puesto ): 1= «5 años o menos» ,  
2= «6- 11 años» , 3= «11-15 años» , 4= «16-20 años».

**satlab1** (Satisfacción laboral en el puesto 1) 1= «nada satisfecho» ,  
2= «algo satisfecho» , 3= «bastante satisfecho» ,  
4= «totalmente satisfecho».

**satlab1b** (Satisfacción laboral en el puesto 1): 1= «alta» ,  
0= «baja».

**satlab2** (Satisfacción laboral en el puesto 2): 1= «nada satisfecho» ,  
2= «algo  
satisfecho» , 3= «bastante satisfecho» ,  
4= «totalmente satisfecho».

**DATOS:**

<b>sujeto</b>	<b>edocivil</b>	<b>antigued</b>	<b>antig2</b>	<b>satlab1</b>	<b>satlab1b</b>	<b>satlab2</b>
1	1	10	2	4	1	3
2	1	7	2	3	1	2
3	1	15	3	4	1	3
4	0	2	1	1	0	2
5	1	5	1	3	1	4
6	1	10	2	2	0	3
7	0	1	1	1	0	1
8	0	2	1	2	0	3
9	0	5	1	1	0	2
10	1	6	2	3	1	1
11	0	2	1	2	1	3
12	1	20	4	4	1	3
13	0	3	1	2	0	4
14	0	2	1	2	1	4
15	0	3	1	1	0	2
16	0	4	1	2	0	3
17	1	8	2	3	1	4
18	1	15	3	3	0	2
19	0	1	1	1	0	2
20	0	9	2	3	0	1
21	1	4	1	1	1	2
22	0	20	4	4	1	3
23	0	3	1	1	0	2
24	1	8	2	3	1	3
25	1	5	1	4	1	4
26	0	2	1	1	0	1
27	0	2	1	2	1	3
28	0	6	2	4	0	2
29	1	6	2	3	1	3
30	1	4	1	2	0	1

## Resultados:

Nro.	Variables usadas.	Coeficiente	Valor		Probabilidad	H.nula
1.a	satlab1, edocivil	Chi-cuadrado	11,58		0,009*	rechazo
	satlab1b, edocivil	Coeficiente Phi	0,535		0,003*	rechazo
	satlab1, edocivil	Coef. de contingencia	0,528		0,009*	rechazo
	satlab1, edocivil	V de Cramer	0,621		0,009*	rechazo
	satlab1, edocivil	Lambda	Simet.	0,389	0,01*	rechazo
			satlab (VD)	0,273	0,021*	rechazo
			edocivil (VD)	0,x71	0,02*	rechazo
1.b	antig2, satlab2	Gamma	0,144		0,516	acepto
	antig2, satlab2	Tau-b de Kendall	0,096		0,516	acepto
	antig2, satlab2	Eta	antig2 (VD)	0,467	—	—
			satlab2 (VD)	0,146		
		Correlación de Spearman	0,108		0,57	acepto
1.c	satlab1, satlab2	Kappa	-0,03		0,777	acepto

## CAPÍTULO VII

---

### CONFIABILIDAD Y VALIDEZ DE UNA ESCALA

Un test, como cualquier otro instrumento, no proporciona una medida perfecta, sino que existe un elemento de error e imperfección que resulta imprescindible conocer para saber si mide lo que pretende medir y la precisión con que lo hace. Por ello es importante elegir los instrumentos de medida con la máxima *confiabilidad* y *validez* posibles.

La *confiabilidad* (o *fiabilidad*) se refiere a la exactitud (consistencia y estabilidad) de la medición de un test. Por lo tanto, la fiabilidad nos informa de la cuantía de error de un instrumento de medida, por lo que a mayor fiabilidad, menor error, y por tanto más exacto o preciso será el test elegido. SPSS ofrece distintos estadísticos para calcular la confiabilidad de un test, de los cuales será presentado el denominado “Alfa de Cronbach”.

Además de buscar la estabilidad, un instrumento debe medir la dimensión real que se supone está midiendo, es decir, debe tener *validez*. La validez no es una propiedad del test o de la evaluación como tal, sino más bien del significado de las puntuaciones. Por lo tanto, cuando el investigador elija un test, deberá tener en cuenta sus características psicométricas de confiabilidad y validez con el fin de que contenga el mínimo error posible y esté midiendo el constructo que pretenda medir.

### CONFIABILIDAD, SIGNIFICADO Y MEDICIÓN

Examinaremos la confiabilidad desde el punto de vista de la consistencia interna. Al examinar la escala de autoestima que

hemos estado analizando, se espera que los ítemes estén balanceados para intentar romper la aquiescencia (tendencia a estar de acuerdo con el enunciado); si las preguntas tienen un sentido positivo, la aquiescencia es mayor. Como la aquiescencia aumenta si los ítemes son en sentido positivo, se deben ir intercalando con ítemes negativos. Se deben utilizar muchos ítemes para llegar a un solo valor. Esta es la forma de determinar la confiabilidad de la medida.

La consistencia interna, es el aspecto común que tienen todos los ítemes, en el ejemplo que estamos viendo es el grado de autoestima que mide cada ítem. Para determinarla, se emplea el coeficiente “a” ( Alfa de Cronbach) la cual es una medida de la confiabilidad de consistencia interna. Su definición es la siguiente:

$$a = (n/(n-1)) * (1 - SS_i^2 / S_t^2)$$

Siendo n=número de ítemes;  $SS_i^2$  Suma de las varianzas de los “n” ítemes y  $S_t^2$ = varianza de los puntajes totales.

Las mediciones que realiza un investigador deben ser *confiables*: si las mediciones se repiten, y se obtienen siempre resultados aproximados, se dice que el instrumento con los que se recogieron los datos es confiable. Si el instrumento es mutidimensional o multifacético no es correcto medir la consistencia interna de todos, sino de las diferentes facetas por separado

Si se usa una escala o una medida estandarizada en la investigación (esto es común en la investigación psicológica y educacional), es importante evaluar la confiabilidad (consistencia interna) de los puntajes de la misma.

Es conveniente identificar los ítemes de la escala que serán analizados así como caracterizarlos, lo que se hace empleando algunos coeficientes o índices, así como medidas descriptivas de resumen (media y desviación estándar, cuando corresponda).

En algunos tests se habla del índice de dificultad, pero se

refiere a aquellos instrumentos cuyos ítemes pueden clasificarse correcta o incorrectamente respondidos, así el índice de dificultad puede ser expresado en forma general como:

$$\text{Índice de Dificultad} = \frac{\text{Número de sujetos que aciertan el ítem}}{\text{Número de sujetos que han intentado resolver el ítem}}$$

### ***Ejemplo***

En la evaluación que hemos realizado de un instrumento para medir los roles de los directivos, apoyados en la teoría de H. Mintzberg, aplicamos el análisis de ítemes desde el punto de vista de su confiabilidad, correlación entre los ítemes y los índices de homogeneidad (estos últimos corresponden a lo que en la teoría clásica se denominan “índice de discriminación”). El resultado fue el siguiente cuando analizamos el rol del directivo como “cabeza visible”.

Escala : “Cabeza visible”

Nº de ítemes : 7

Matriz de correlaciones

Ítem	1	2	3	4	5	6	7
1	1.00						
2	.18	1.00					
3	.11	-.01	1.00				
4	.16	.21	-.06	1.00			
5	.13	.47	.17	.12	1.00		
6	.05	-.07	.20	.18	.03	1.00	
7	.26*	.10	-.06	.17	.19	-.12	1.00

\*correlaciones significativas.

Los coeficientes de correlación entre los ítemes son bajos, detectándose escasas correlaciones significativas(\*).

Resumen de la escala:  
Mean=29.9351 Std.Dv.=2.60737 Valid N:77  
Cronbach alpha: .469787  
Standardized alpha: .474733  
Average inter-ítem corr.: .116831

Relación entre cada ítem y el test

Ítem	1 Mean if deleted	2 Var.if deleted	3 StDv.if deleted	4 Itm-Totl Correl.	5 Squared Multp.R	6 Alpha if deleted
1	25.37662	5.689323	2.385230	.300414	.119598	.412095
2	25.72727	5.211335	2.282835	.307952	.274047	.392725
3	25.40260	5.902851	2.429578	.096121	.100083	.483319
4	26.03896	5.024457	2.241530	.255392	.127944	.413857
5	25.83117	4.529938	2.128365	.378327	.275279	.341798
6	25.38961	5.926126	2.434364	.079639	.114174	.491189
7	25.84416	5.274414	2.296609	.169467	.143011	.460553

El *índice de homogeneidad corregido* (columna nº4), es decir, la correlación entre el ítem y la suma de los ítemes restantes nos dice que para correlaciones del índice bajo 0,2 se dirá que no hay mucha relación entre el ítem seleccionado y los restantes como lo es el ítem 3, 6 y 7. El resto de los índices son moderadamente altos.

La correlación múltiple al cuadrado (columna nº5), es decir el grado en que se puede predecir la puntuación en el ítem a partir de las puntuaciones en los restantes ítemes, nos indica que las correlaciones son adecuadas.

El coeficiente Alfa de Cronbach si es eliminado un ítem determinado (columna nº6), nos indica que si eliminamos el ítem nº6 el Alfa de Cronbach aumentará a 0,4912. El coeficiente Alfa de Cronbach inicial es de 0,469787.

Se propone eliminar aquellos ítemes que poseen los índices



de homogeneidad más bajos, es decir los ítems 3 , 6 y 7. El coeficiente Alfa de Cronbach quedará en 0,55, lo que mejora la consistencia interna del instrumento, aunque aún sin alcanzar el valor mínimo deseable en ciencias sociales: 0,67.

### *Aplicación al caso en Estudio*

Para analizar la Escala de Autoestima a la que nos hemos venido refiriendo en este libro, en SPSS se selecciona la ruta:

**Analizar**  
**Escalas**  
**Análisis de fiabilidad**  
**Modelo Alfa**  
**Estadístico**  
**Elemento**  
**Escala si se elimina elemento**

En la ventana Análisis de Fiabilidad se selecciona los ítems que serán sometidos a escrutinio; en este caso, los 10 ítems de la escala de autoestima:

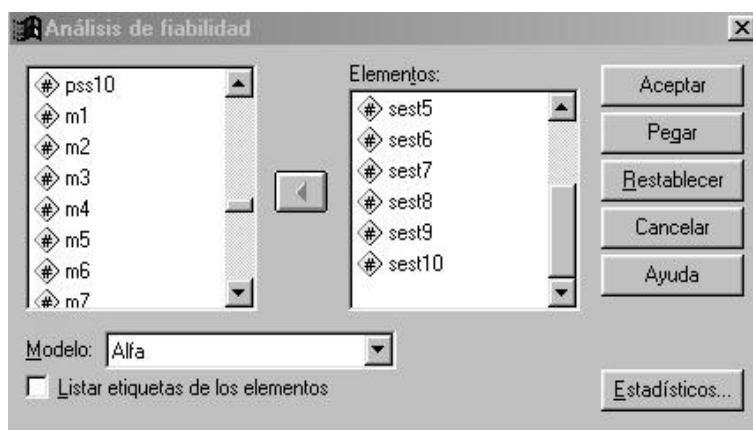


Figura 7.1. Cuadro análisis de fiabilidad

El analista solicita la estadística descriptiva para cada elemento y lo que ocurrirá con la escala si se elimina el ítem:

**Análisis de fiabilidad: Estadísticos**

Descriptivos para:

- ☒ Elemento
- ☐ Escala
- ☒ Escala si se elimina elemento

Entre-elementos:

- ☐ Correlaciones
- ☐ Covarianzas

Resúmenes:

- ☐ Medias
- ☐ Varianzas
- ☐ Covarianzas
- ☐ Correlaciones

Tabla de ANOVA:

- ☒ Ninguno
- ☐ Prueba F
- ☐ Chi-cuadrado de Friedman
- ☐ Chi-cuadrado de Cochran

☐ T-cuadrado de Hotelling

☐ Prueba de aditividad de Tukey

☐ Coeficiente de correlación intraclass

Modelo: **Dos factores, efectos mixtos** Tipo: **Consistencia**

Intervalo de confianza: **95** % Valor de prueba: **0**

Figura 7.2. Cuadro definición de estadísticos

El resultado presenta la siguiente apariencia:

#### RELIABILITY ANALYSIS - SCALE (ALPHA)

	Mean	Std	Dev	Cases
1. SEST1	3.6720	.5874	436.0	
2. SEST2	3.7454	.5180	436.0	
3. SEST3	3.6239	.6820	436.0	
4. SEST4	3.4083	.7448	436.0	
5. SEST5	3.4037	.9334	436.0	
6. SEST6	3.2775	.7928	436.0	
7. SEST7	2.9564	1.0787	436.0	
8. SEST8	3.1697	.8269	436.0	
9. SEST9	2.9610	1.0152	436.0	
10. SEST10	3.2867	.9698	436.0	

---

 Ítem-total Statistics

	Scale Mean if Ítem Deleted	Scale Varianc if Ítem Deleted	Corrected Ítem- Total Correlation	Alpha if Ítem Deleted
SEST1	29.8326	26.6179	.5098	.8458
SEST2	29.7592	26.8269	.5511	.8447
SEST3	29.8807	25.4110	.6075	.8378
SEST4	30.0963	25.9217	.4715	.8474
SEST5	30.1009	25.0289	.4447	.8517
SEST6	30.2271	23.9506	.7054	.8280
SEST7	30.5482	22.7218	.5997	.8384
SEST8	30.3349	23.9060	.6756	.8300
SEST9	30.5436	23.0946	.6075	.8366
SEST10	30.2179	23.8949	.5502	.8420

## Reliability Coefficients

N of Cases = 436.0

N of Ítemes = 10

Alpha = .8541

El Alfa de Cronbach es de 0,8541, lo que indica que el instrumento tiene consistencia interna, es decir, es confiable. El coeficiente de discriminación, esto es, el grado en que un ítem sirve para distinguir entre los individuos que obtienen puntuaciones altas y los que las obtienen bajas, viene dado por la columna correlación total-ítem corregido

En la tabla de resultados “Ítem-total Statistics”, se observa la relación de cada ítem y el test; si es eliminado este ítem, por ejemplo si eliminamos el ítem SEST5, implicaría que la media

del test fuera 30,1009 y su varianza 25,0289. El índice de discriminación, esto es, el grado en que un ítem sirve para distinguir entre los individuos que obtienen puntuaciones altas y los que las obtienen bajas, viene dado por la columna correlación total-ítem corregido y en nuestro problema están en el rango 0,44 a 0,70, lo que indica que son aceptables y que nos permite afirmar que existe homogeneidad entre los ítems que conforman la escala de “Autoestima”. Se recomendaría eliminar el ítem si alguna de estas correlaciones fuera menor que 0,2. La última columna entrega el Alfa de Cronbach si el ítem es eliminado, nótese que eliminando el ítem SEST5 el Alfa de Cronbach será 0.8517 el cual es menor que Alfa de Cronbach con todos los ítems.

### ***Inferencia sobre el Coeficiente de Confiabilidad***

Se pueden emplear los resultados entregados por SPSS para complementar el estudio de la confiabilidad mediante la aplicación de pruebas o del cálculo de intervalos de confianza respectivos.

Al diseñar un instrumento, interesa analizar si efectivamente la confiabilidad observada en la muestra, es compatible con algún valor determinado en la población. Para estudiar este punto, se pueden aplicar los resultados presentados por Muñiz (2001) desarrollados por Kristof (1963) y Feld (1965)) sobre la distribución del estadístico de prueba  $F = (1 - a) / (1 - a_0)$  el cual tiene una Distribución F con  $(N-1)$  y  $(N-1)(n-1)$  grados de libertad, siendo  $N$  el número de sujetos en la muestra,  $n$  el número de ítems del test,  $a$  el valor poblacional postulado y  $a_0$  el valor obtenido en la muestra (No confundir la notación de este Alfa de Cronbach con el empleado para el nivel de significación).

Con respecto a la Escala de Autoestima, dado que interesa analizar si la confiabilidad es significativa, puede plantearse la hipótesis nula que la confiabilidad poblacional es 0 con la hipótesis alternativa que es mayor que 0. Esto es  $H_0: a \leq 0$  versus  $H_1: a > 0$ .

Empleando los resultados entregados por SPSS, ya vistos:

#### Reliability Coefficients

N of Cases = 436.0

N of Ítemes = 10

Alpha = .8541

El estadístico tendrá el valor

$$F=(1-0)/(1-0,8541)=1/(0,1459)=6,854.$$

Si se emplea un nivel de significación del 5%, se compara el valor antes calculado con el entregado por la tabla F con (436-1) y (10-1)(436-1) grados de libertad, esto es 435 y 3.915 grados de libertad. El valor más cercano en la tabla F para  $F(0,95; 435; 3915)$  es  $F(0,95; \infty; \infty)=1$ .

Por lo tanto, dado que el valor calculado para  $F=6,854$  es mayor que 1, se concluye que la hipótesis nula es rechazada y por lo tanto, se acepta  $H_1: a > 0$  o sea, la confiabilidad de la escala es significativa.

## VALIDEZ, SIGNIFICADO Y MEDICIÓN

La validez de un instrumento de medición consiste en la exactitud con que pueden hacerse medidas significativas y adecuadas con él, en el sentido que midan realmente los rasgos que se pretenden medir.

“Un test es una muestra de conducta a partir de la cual se pretenden hacer ciertas inferencias. La validez se refiere al conjunto de pruebas y datos que han de recogerse para garantizar la pertinencia de tales inferencias. Más que el test, lo que se validan son las inferencias” (Muñiz, 2001).

Se reconocen tres grandes tipos de validez: de contenido, predictiva y de constructo. Sin embargo, no se consideran como categorías independientes ni exhaustivas. Ellas están relacionadas y se influyen.

La validez de contenido es la representatividad o adecuación muestral del contenido esto es, la materia o tema del instrumento de medición. Consiste esencialmente en el juicio. Para asegurar su existencia deben estudiarse los ítemes de un test y ponderarlos de acuerdo con su representatividad supuesta del universo. Para ello se emplea el “criterio de jueces”, esto es, la opinión de diversos especialistas en el tema que determinan en que grado el ítem es representativo del universo respectivo.

“La validez predictiva de un test se refiere al grado de eficacia con el que se puede predecir o pronosticar una variable de interés (criterio) a partir de las puntuaciones de ese test. Se operacionaliza mediante el coeficiente de validez, que es la correlación entre el test y el criterio” (Muñiz, 2001).

La validez de constructo trata de responder preguntas como: ¿Qué es lo que este test mide en realidad?, o ¿Este instrumento mide en forma adecuada el concepto abstracto en estudio?. Este tipo de validez se relaciona más con el atributo subyacente que con los puntajes que produce el instrumento. Según Cronbach y Meehl (1955), citados por Muñiz (2001), “se refiere a la recogida de evidencia empírica que garantice la existencia de un constructo psicológico en las condiciones exigibles a cualquier otro modelo o teoría científica”.

Los procedimientos metodológicos empleados con mayor frecuencia, en Psicología, para obtener información acerca de la validez de constructos psicológicos son el análisis factorial y la matriz multirrasgo-multimétodo. Estos, se denominan, respectivamente, *validez factorial* y *validez convergente-discriminante*.

## ANÁLISIS FACTORIAL, CARACTERÍSTICAS Y TIPOS

El Análisis Factorial es un procedimiento estadístico-matemático mediante el cual podemos expresar un gran número

de variables mediante un número menor de dimensiones o constructos comunes a todas ellas (llamados factores), trabajando con sus intercorrelaciones.

El modelo de análisis factorial asume que el rendimiento de las personas en cualquier variable, como por ejemplo una aptitud, está en función de tres fuentes de variación: varianza de factores comunes a los test; varianza de factores específicos de cada test y varianza error.

Pese al interés que puedan tener los factores específicos, la labor científica se centra en los factores comunes existentes entre variables. La puntuación que obtiene una persona al serle aplicado un test, refleja una combinación lineal de su nivel en los distintos factores comunes (aptitudes comunes a todas las tareas), en el componente o factor específico del test (aquella parte de la puntuación que no es atribuible ni a los factores comunes ni al error de medida) y en el componente de error de medida. Es muy importante distinguir la puntuación de una persona en un test, lo cual indica como se ha comportado dicha persona al ser evaluada con ese test, de la puntuación en un factor, que nos indica el grado en que ese individuo posee un determinado atributo. Los factores comunes son los representantes estadísticos de las aptitudes comunes que se buscan.

Hay dos tipos de análisis factorial: el exploratorio y el confirmatorio. En el análisis exploratorio no se conocen los factores “a priori”, si no que son el resultado del proceso. Se utiliza cuando se intentan explorar, por ejemplo, las aptitudes intelectuales y crear hipótesis nuevas. En el análisis factorial confirmatorio se propone un modelo, según el cual hay unos factores (cuyo número es menor al de las variables), que representan a las variables originales. Se utilizará cuando se busca la comprobación de hipótesis ya existentes.

## PASOS EN EL DESARROLLO DEL ANÁLISIS FACTORIAL

### ***Primer Paso: Matriz de datos***

El análisis factorial comienza por obtener una matriz ( $N \times n$ ) que representa las puntuaciones normalizadas obtenidas por los  $N$  sujetos en las  $n$  distintas variables (preguntas del test). A continuación, se presenta la matriz de datos típica:

	Preg.1	Preg.2	Preg.3	Preg.4	.....	Preg.n
suj.1	.	.	.	.	.	
suj.2	.	.	.	.	.	
suj.3	.	.	.	.	.	
suj.4	.	.	.	.	.	
.....	.	.	.	.	.	
suj.N						

Figura 7.3. Matriz de puntuaciones.

### ***Segundo Paso: Matriz de Correlaciones.***

En el segundo paso, se obtiene (mediante el empleo de un software como SPSS) la matriz de correlaciones entre las distintas variables ( $n \times n$ ) que representa los datos de referencia sobre los que situaremos los factores. También, se obtienen las estadísticas descriptivas (medias y desviaciones estándar) de las variables.

Un requisito importante para que tenga sentido el análisis factorial es que las variables están altamente correlacionadas. Para verificar esta condición, SPSS ofrece un conjunto de procedimientos, a saber: valor del determinante, test de esfericidad de Bartlett, el índice KMO, la correlación anti-imagen y la medida de adecuación muestral. Su empleo se presenta más adelante.

La relación encontrada entre las variables en este paso no cambia nunca a lo largo del proceso.



	Preg.1	Preg.2	Preg.3	Preg.4	.....
Preg.1	1,0000	0,2400	0,3000	0,1200	.
Preg.2	0,2400	1,0000	0,6900	0,8300	.
Preg.3	0,3000	0,6900	1,0000	0,1500	.
Preg.4	0,1200	0,8300	0,1500	1,0000	.
.....	.	.	.	.	.

Figura 7.4. Matriz de correlaciones.

Se puede observar que la correlación de una pregunta consigo misma es exactamente igual a 1,000 que es la máxima correlación posible. Observar también que la tabla es simétrica diagonalmente por lo cual se suele representar sólo la diagonal superior izquierda porque los datos se repiten debajo.

### ***Tercer Paso: Aplicación del Análisis Factorial.***

Se construye una matriz factorial inicial sin rotar ( $n \times k$ ) de  $n$  variables por  $k$  factores.

Un factor es un constructo operativamente definido por sus pesos factoriales. Cada factor es una combinación lineal de las variables iniciales que lo componen. Cada una de estas variables contribuye en distinto grado a la formación del factor. Lo ideal es que cada factor esté formado por dos o tres variables. La existencia de un factor garantiza que éste es un constructo útil que puede dar cuenta de las diferencias individuales que pueden observarse en las puntuaciones de los test.

Para obtener los factores hay que tener en cuenta las intenciones del experimentador, pues si quiere comprobar una hipótesis utilizará un análisis factorial confirmatorio (extrayendo el número de factores de acuerdo a lo planteado por la hipótesis).

Si el experimentador quiere explorar en el resumen de la información, utilizará un análisis factorial exploratorio (experimentando con la extracción de un número variable de factores).

Existen bastantes métodos de extracción de factores, muchos de los cuales están implementados en SPSS. Se distinguen dos grupos según el principio empleado: los que intentan el mejor ajuste a la matriz de correlaciones iniciales (máxima verosimilitud, mínimos cuadrados), y los que intentan explicar la mayor proporción de la variabilidad de los datos (componentes principales, factorización de ejes principales, factorización imagen, método alfa, etc.).

El método de los componentes principales es uno de los procedimientos más usados. Sin embargo, algunos autores consideran dicho método como una técnica diferente del análisis factorial. Por ello, recomiendan utilizar el método de máxima verosimilitud o el de mínimos cuadrados, mientras no se tenga una comprensión completa del método que se vaya a utilizar.

Un ejemplo de la estructura de la matriz factorial inicial es el siguiente:

Variables	FACTOR 1	FACTOR 2
Preg.1	0,71	-0,03
Preg.2	0,22	0,65
Preg.3	0,59	-0,16
Preg.4	0,03	0,53

Figura 7.5. Análisis factorial.

Cada factor está compuesto por una serie de saturaciones, o cargas, en cada variable. Estas saturaciones representan la correlación entre cada factor y las distintas variables. Cuanto mayor sea la saturación factorial de una variable en un factor más

cerca estaremos de medir realmente el factor. El valor máximo de las cargas factoriales es 1. En este caso, el factor explicaría completamente la variabilidad de la variable.

En la tabla puede verse que las variables que saturan en el factor I, son las preguntas 1 y la 3. En el factor II, las variables más relacionadas con él son las preguntas 2 y 4.

Una propiedad importante de las cargas o saturaciones factoriales es que si se eleva al cuadrado cada una de las cargas factoriales, dentro de un factor, y se suman dichos cuadrados, esta suma corresponde a la varianza explicada por dicho factor. En el ejemplo anterior, para el primer factor se tiene:  $(0,71)^2 + (0,22)^2 + (0,59)^2 + (0,03)^2 = 0,9015$ . Por lo tanto, el primer factor explica una varianza de 0,9015.

#### ***Cuarto Paso: Rotación de la Matriz Factorial.***

La rotación de la matriz factorial inicial es un proceso de multiplicación de la matriz inicial por una matriz de transformación.

Gráficamente consiste en girar los factores sobre los vectores (que representan las variables) hasta conseguir la mejor agrupación de variables por factor. La rotación puede llevarse a cabo bajo dos condiciones diferentes que decide el experimentador según sus propósitos:

- **Rotación ortogonal:** Se mantiene la independencia de los factores, es decir, no pueden estar relacionados y por tanto, no medirán alguna característica global común a ellos. Algunos tipos son: rotación varimax (el más empleado), quartimax, equamax etc.
- **Rotación oblicua:** No se mantiene la independencia de los factores, los que pueden estar correlacionados. Algunos tipos son: rotación oblimin directo (el más conocido), oblimax, promax, quartimin, covarimin, etc.

Cuando realizamos una rotación ortogonal, obtenemos una matriz rotada similar a la anterior sin rotar ( $n \times k$ ) de variables por factores. Los resultados son muy claros pero tienen un inconveniente: en la naturaleza hay muy pocas variables que no interactúen entre sí y sean exactamente independientes.

Por esta razón, es preferible la rotación oblicua ya que deja más libertad a los datos. Esto es a partir de una rotación oblicua puede obtenerse una estructura ortogonal; sin embargo, con una rotación ortogonal es imposible obtener estructuras factoriales oblicuas.

Cuando se realiza una rotación oblicua nos encontramos como resultado dos matrices: la de *estructura factorial* y la de *patrón factorial*.

- La matriz de estructura factorial presenta los pesos factoriales, las correlaciones de cada variable con el factor, es decir cuanto de cerca está esa variable de ese factor.
- La matriz de patrón factorial presenta los pesos de las variables, que indican la importancia de cada variable para predecir un factor determinado.

	ESTRUCTURA FACTORIAL		PATRÓN FACTORIAL	
	FACTOR I	FACTOR II	FACTOR I	FACTOR II
Preg.1	0,69	0,37	0,71	0,50
Preg.2	0,10	0,89	0,17	0,82
Preg.3	0,84	0,05	0,81	0,23
Preg.4	-0,11	0,81	0,18	0,78

Figura 7.6. Estructura y patrón factorial.

Se puede observar que, en este ejemplo, ambas matrices, la de estructura factorial y la de patrón factorial son muy similares y su interpretación apenas diferirá.

El hecho de utilizar la rotación oblicua permite encontrar una vez llegados a este punto, correlaciones entre los propios factores con las que se puede realizar un nuevo análisis factorial. De este nuevo análisis factorial se obtendrían factores de segundo orden, esto es factores más generales que comprenden otros factores. En el ejemplo de la tabla anterior, no es posible hacer un nuevo análisis factorial dado el escaso número de variables y factores que se ha utilizado.

Hay que recordar una vez más que el análisis factorial tiene como objetivo la descripción de las relaciones de una serie de variables gracias a la formación de un número menor de otras variables comunes a las primeras definibles matemáticamente, los factores.

## VALIDEZ FACTORIAL DE LA ESCALA DE AUTOESTIMA

En la Escala de Autoestima que hemos venido comentando, el sujeto elige una graduación de respuestas que miden la frecuencia de ciertas conductas, utilizando las expresiones *totalmente de acuerdo*, *de acuerdo*, *en desacuerdo*, y *totalmente en desacuerdo*. De los diez ítems, cinco de ellos están redactados de forma negativa, y aparecen destacados con dos asteriscos al final de los mismos. La escala fue aplicada a 439 sujetos.

Nuestro objetivo de partida consiste en averiguar si estos diez ítems se pueden resumir de alguna forma, si existe algo en común entre ellos. Mediante el análisis factorial (por el método de los componentes principales), intentaremos agrupar estos reactivos en factores o componentes que puedan explicar la mayoría de la varianza observada en las respuestas ofrecidas por los sujetos de la muestra. Dado que hay dos grupos de ítems:

aquellos que conforman la escala de *sentimiento positivo* (ítemes 1, 2, 4, 6 y 7), y aquellos que conforman la escala de *sentimiento negativo* (ítemes 3, 5, 8, 9 y 10); se intentará un análisis factorial confirmatorio, considerando dos factores para extraer.

El primer paso es obtener los estadísticos descriptivos, esto es, la media, la desviación típica, el número de casos que se han utilizado en el análisis y el número que se han perdido, es decir, los que no han respondido a ítem concreto. En la tabla 8.1 se presenta estos resultados. En el caso del ítem dos, la desviación típica es 0,52, pero ¿cómo sabemos si este valor es bajo o alto? Para responder a esta pregunta podemos calcular el *coeficiente de variación* respectivo, que indica las veces en las que la media contiene a la desviación típica. Para ello, basta con hallar el cociente entre la desviación típica y la media, por lo que para el ítem 2 obtenemos:  $0,52/3,75 = 0,138$ .

Esto significa que, aproximadamente, la desviación estándar es un 14% de la media. Aquellos ítemes que poseen menor coeficiente de variación se dice que son más homogéneos con respecto a coeficientes de variación mayores.

**Estadísticos descriptivos**

	Media	Desviación típica	N del análisis
SEST1	3.67	.59	436
SEST2	3.75	.52	436
SEST3	3.62	.68	436
SEST4	3.41	.74	436
SEST5	3.40	.93	436
SEST6	3.28	.79	436
SEST7	2.96	1.08	436
SEST8	3.17	.83	436
SEST9	2.96	1.02	436
SEST10	3.29	.97	436

Tabla 7.1. Estadísticos descriptivos.

Estos datos son interesantes puesto que nos dan cierta idea de la calidad de nuestros ítems, sobre todo aquéllos, cuyos valores de la media y la desviación típica son muy altos o muy bajos.

El paso siguiente es calcular la matriz de correlaciones entre las variables y el determinante de dicha matriz. Si las relaciones entre los ítems son pequeñas, es poco probable que se encuentren factores comunes o dimensiones. En la tabla 2, se ha calculado la correlación de cada ítem consigo mismo y con el resto de los ítems que forman parte de la escala. Si se estudia detenidamente esta matriz, se observa que los ítems están correlacionados a un nivel de significancia del 5%, aunque esta correlación es, en general, baja. Por lo tanto, tendría sentido realizar el análisis factorial.

Matriz de correlaciones <sup>a</sup>											
Correlación	SEST1	SEST2	SEST3	SEST4	SEST5	SEST6	SEST7	SEST8	SEST9	SEST10	
SEST1	1.000	.624	.368	.359	.246	.443	.308	.432	.271	.250	
SEST2	.624	1.000	.457	.419	.275	.491	.326	.428	.261	.255	
SEST3	.368	.457	1.000	.357	.477	.415	.381	.464	.397	.382	
SEST4	.359	.419	.357	1.000	.248	.407	.303	.391	.289	.239	
SEST5	.246	.275	.477	.248	1.000	.296	.276	.379	.325	.260	
SEST6	.443	.491	.415	.407	.296	1.000	.587	.643	.490	.425	
SEST7	.308	.326	.381	.303	.276	.587	1.000	.506	.471	.418	
SEST8	.432	.428	.464	.391	.379	.643	.506	1.000	.424	.392	
SEST9	.271	.261	.397	.289	.325	.490	.471	.424	1.000	.607	
SEST10	.250	.255	.382	.239	.260	.425	.418	.392	.607	1.000	
Sig. (Unilateral)	SEST1	.000	.000	.000	.000	.000	.000	.000	.000	.000	
	SEST2	.000	.000	.000	.000	.000	.000	.000	.000	.000	
	SEST3	.000	.000	.000	.000	.000	.000	.000	.000	.000	
	SEST4	.000	.000	.000	.000	.000	.000	.000	.000	.000	
	SEST5	.000	.000	.000	.000	.000	.000	.000	.000	.000	
	SEST6	.000	.000	.000	.000	.000	.000	.000	.000	.000	
	SEST7	.000	.000	.000	.000	.000	.000	.000	.000	.000	
	SEST8	.000	.000	.000	.000	.000	.000	.000	.000	.000	
	SEST9	.000	.000	.000	.000	.000	.000	.000	.000	.000	
	SEST10	.000	.000	.000	.000	.000	.000	.000	.000	.000	

<sup>a</sup>. Determinante = 2.257E-02

Tabla 7.2. Matriz de correlaciones.

Con la matriz de correlaciones podríamos hacernos una idea de la agrupación de los ítems en los factores, puesto que aquellos ítems que presenten una correlación alta entre ellos, es probable que aparezcan dentro de un mismo factor o componente. En este caso, los resultados no muestran una tendencia claramente definida. Un determinante de la matriz muy bajo supone la existencia de variables con correlaciones entre sí muy elevadas. En la tabla anterior, el valor es 0,02257, es decir, muy bajo, aunque

sin llegar al valor cero, por tanto, es posible la realización del análisis factorial.

Para la aplicación del análisis factorial, son necesarios una serie de requisitos. Uno de ellos es que las correlaciones entre las variables han de ser altas. La prueba de esfericidad de Bartlett se emplea para probar la hipótesis nula de que la matriz de correlaciones es una *matriz identidad* (correlación igual a uno entre los ítems). En la tabla 2 se observa que el valor de probabilidad asociado (sig.) tiende a cero, por lo que se rechaza la hipótesis nula a un nivel de significación del 5% y por lo tanto, tiene sentido realizar el análisis factorial.

Otro dato que nos indica si es posible el análisis factorial es la *medida de adecuación muestral KMO*, de Kaiser-Meyer-Olkin. Este es un índice que permite comparar las magnitudes de los coeficientes de correlación obtenidos con las magnitudes de correlación parcial. Cuando este índice toma un valor bajo, se desaconseja la aplicación del análisis factorial, puesto que las correlaciones entre pares de ítems no pueden explicarse a partir de otros ítems. La interpretación se elabora a partir de los siguientes valores:

$0,90 < KMO \leq 1$	Muy bueno
$0,80 < KMO \leq 0,90$	Meritorio
$0,70 < KMO \leq 0,80$	Mediano
$0,60 < KMO \leq 0,70$	Mediocre
$0,50 < KMO \leq 0,60$	Bajo
$KMO \leq 0,50$	Inaceptable

Según la tabla 3, el valor de KMO es de 0,876, lo que supone que la matriz es adecuada para realizar el análisis factorial.



### KMO y prueba de Bartlett

Medida de adecuación muestral de Kaiser-Meyer-Olkin.		.876
Prueba de esfericidad de Bartlett	Chi-cuadrado aproximado	1633.361
	gl	45
	Sig.	.000

Tabla 7.3. KMO y prueba de Bartlett.

Hasta este punto de análisis ¿cómo hemos obtenido estos resultados en SPSS?.

La secuencia de pasos en SPSS es:

#### **Analizar**

#### **Reducción de datos**

#### **Análisis factorial**

#### **Descriptivos**

#### **Descriptivos univariados Coeficientes KMO y prueba de esfericidad de Bartlett**

#### **Extracción**

#### **Extraer Número de factores: 2**



Para obtener los descriptivos univariados, la matriz de correlaciones, los niveles de significación, el determinante y la prueba KMO, se activa Descriptivos.

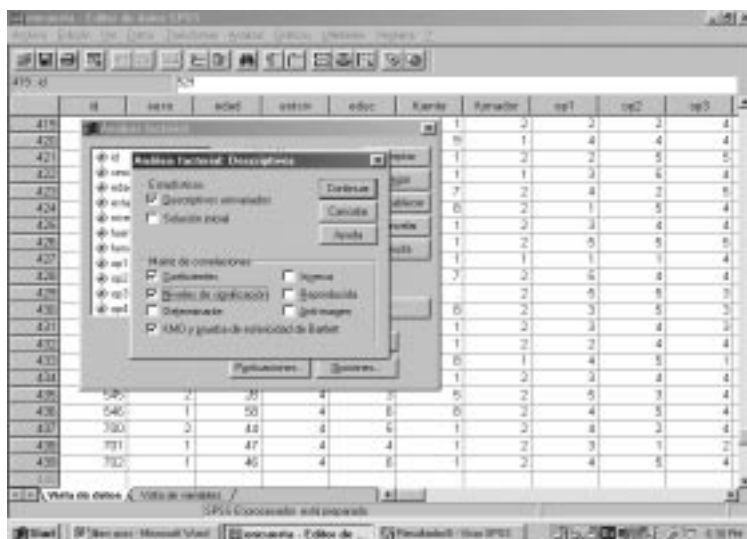


Figura 7.9. Estadísticos y correlaciones.

Otro dato es la *correlación anti-imagen*, que es el negativo del coeficiente de correlación parcial. Deberá haber pocos coeficientes de correlación anti-imagen altos para que sea razonable aplicar el Análisis Factorial. Además en la diagonal de la matriz de correlaciones anti-imagen, se muestra la *Medida de Adecuación de la Muestra (MSA)* para cada ítem, cuyo valor si es inferior a 0,5 indicará que el ítem no se ajusta con la estructura de los demás ítemes, debiendo considerarse la posibilidad de eliminarlo del análisis. En la pantalla anterior se selecciona Anti-imagen, el resultado es:

Matrices anti-imagen											
		SEST1	SEST2	SEST3	SEST4	SEST5	SEST6	SEST7	SEST8	SEST9	SEST10
Covarianza anti-imagen	SEST1	.569	-.248	-1.039E-02	-4.205E-02	-6.765E-03	-2.989E-02	-3.620E-04	-6.388E-02	-1.217E-02	-4.186E-03
	SEST2	-.248	.504	-.112	-9.721E-02	-5.818E-03	-8.655E-02	5.849E-03	-3.609E-04	3.649E-02	8.085E-04
	SEST3	-1.039E-02	-.112	.581	-5.772E-02	-.202	1.188E-02	-3.970E-02	-6.437E-02	-4.141E-02	-6.761E-02
	SEST4	-4.205E-02	-9.721E-02	-5.772E-02	.736	-2.218E-02	-5.304E-02	-1.214E-02	-5.375E-02	-3.327E-02	1.011E-02
	SEST5	-6.765E-03	-5.818E-03	-.202	-2.218E-02	.729	1.867E-02	-8.448E-03	-8.370E-02	-6.539E-02	8.412E-03
	SEST6	-2.989E-02	-8.655E-02	1.188E-02	-5.304E-02	1.867E-02	.428	-.149	-.164	-7.375E-02	-2.595E-02
	SEST7	-3.620E-04	5.849E-03	-3.970E-02	-1.214E-02	-8.448E-03	-.149	.583	-7.105E-02	-7.964E-02	-5.451E-02
	SEST8	-6.388E-02	-3.609E-04	-6.437E-02	-5.375E-02	-8.370E-02	-.164	-7.105E-02	.493	-7.861E-03	-3.009E-02
	SEST9	-1.217E-02	3.649E-02	-4.141E-02	-3.327E-02	-6.539E-02	-7.375E-02	-7.964E-02	-7.861E-03	.531	-.248
	SEST10	-4.186E-03	8.085E-04	-6.761E-02	1.011E-02	8.412E-03	-2.595E-02	-5.451E-02	-3.009E-02	-.248	.588
Correlación anti-imagen	SEST1	.849 <sup>a</sup>	-.463	-1.807E-02	-6.497E-02	-1.051E-02	-6.056E-02	-6.287E-04	-.121	-2.213E-02	-7.240E-03
	SEST2	-.463	.824 <sup>a</sup>	-.207	-.160	-9.596E-03	-.186	1.079E-02	-7.235E-04	7.050E-02	1.485E-03
	SEST3	-1.807E-02	-.207	.892 <sup>a</sup>	-8.823E-02	-.310	2.381E-02	-6.820E-02	-.120	-7.453E-02	-.116
	SEST4	-6.497E-02	-.160	-8.823E-02	.947 <sup>a</sup>	-3.027E-02	-9.446E-02	-1.853E-02	-8.918E-02	-5.320E-02	1.536E-02
	SEST5	-1.051E-02	-9.596E-03	-.310	-3.027E-02	.875 <sup>a</sup>	3.341E-02	-1.296E-02	-.140	-.105	1.285E-02
	SEST6	-6.056E-02	-.186	2.381E-02	-9.446E-02	3.341E-02	.875 <sup>a</sup>	-.298	-.357	-.155	-5.171E-02
	SEST7	-6.287E-04	1.079E-02	-6.820E-02	-1.853E-02	-1.296E-02	-.298	.915 <sup>a</sup>	-.132	-.143	-9.309E-02
	SEST8	-.121	-7.235E-04	-.120	-8.918E-02	-.140	-.357	-.132	.902 <sup>a</sup>	-1.536E-02	-5.585E-02
	SEST9	-2.213E-02	7.050E-02	-7.453E-02	-5.320E-02	-.105	-.155	-.143	-1.536E-02	.849 <sup>a</sup>	-.444
	SEST10	-7.240E-03	1.485E-03	-.116	1.536E-02	1.285E-02	-5.171E-02	-9.309E-02	-5.585E-02	-.444	.850 <sup>a</sup>

<sup>a</sup>. Medida de adecuación muestral

Tabla 7.4 Matrices anti-imagen.

En nuestro problema, pocos valores de la matriz de correlaciones anti-imagen son altos y ninguna medida de adecuación de la muestra es inferior a 0,5. Por lo tanto, es razonable aplicar el análisis factorial.

Los datos después de aplicar el test de Bartlett, el índice KMO y los coeficientes recogidos en la matriz de correlaciones anti-imagen nos permiten comprobar que el análisis es posible. Se espera poder resumir los diez ítems de la escala en factores comunes o dimensiones.

Se procede a la determinación de las *comunalidades*, esto es a la proporción de varianza explicada por los factores comunes. Cuando se emplea el método de componentes principales para la extracción de los factores, las comunales iniciales son siempre iguales a uno. Los resultados aparecen en la tabla 7.5. Así por ejemplo, la comunalidad del ítem 1 vale 0,656; la del ítem 2 es 0,727, etc. En general, podemos afirmar que los diez ítems son explicados por los componentes, puesto que no hay valores bajos, próximos a cero, siendo el ítem más bajo el 5°.

## Comunalidades

	Extracción
SEST1	.656
SEST2	.727
SEST3	.489
SEST4	.431
SEST5	.298
SEST6	.628
SEST7	.542
SEST8	.590
SEST9	.698
SEST10	.647

Método de extracción: Análisis de Componentes principales.

Tabla 7.5 Comunalidades.

La tabla 7.6 nos entrega los autovalores y el total de varianza explicada por las componentes cuyos autovalores son mayores que 1. Se observa que la primera componente explica el 45,4% aproximadamente de la varianza y la segunda componente el 11,7%, explicando ambas el 57,1% de la varianza, siendo este porcentaje considerado razonable.

## Varianza total explicada

Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	4.537	45.372	45.372	4.537	45.372	45.372
2	1.168	11.684	57.055	1.168	11.684	57.055
3	.896	8.957	66.012			
4	.697	6.969	72.982			
5	.675	6.751	79.733			
6	.501	5.007	84.740			
7	.466	4.661	89.401			
8	.384	3.842	93.243			
9	.376	3.763	97.006			
10	.299	2.994	100.000			

Método de extracción: Análisis de Componentes principales.

Tabla 7.6. Varianza total explicada.

La interpretación de los factores es una de las etapas del análisis factorial más importante, y ésta se realiza a través de las *cargas factoriales*, las que corresponden a las correlaciones entre el ítem y el factor o componente no rotado. Estas se muestran en la tabla 7.7.

En nuestro caso, se usará el valor 0,35 para determinar si una carga factorial indica una asociación del ítem con la componente; sin embargo, este criterio está asociado al tamaño muestral y mientras menor es el tamaño muestral, mayor es la dificultad de que un ítem se asocie a una componente o factor. En Opciones, del menú principal de Análisis factorial, puede seleccionarse la opción *Suprimir valores absolutos menores que*, y escribir 0,35 para que SPSS elimine de la matriz de componentes los valores menores a 0,35. De esta forma, se facilita la interpretación.

Matriz de componentes<sup>a</sup>

	Componente	
	1	2
SEST3	.698	4.152E-02
SEST4	.586	.295
SEST5	.543	-5.28E-02
SEST6	.792	-2.75E-02
SEST7	.691	-.255
SEST8	.768	3.075E-03
SEST9	.676	-.490
SEST10	.627	-.503
SEST1	.639	.498
SEST2	.676	.519

Método de extracción: Análisis de componentes principales.

a. 2 componentes extraídos

Tabla 7.7. Matriz de componentes.

En nuestro ejemplo la primera componente está representada por todos los ítems (cargas factoriales positivas), la segunda componente por los ítems SEST4, SEST1, SEST2 (cargas factoriales positivas) y SEST7, SEST9 SEST10 (cargas factoriales negativas).

Como se espera que cada factor posea una correlación alta con un grupo de ítems y baja con otros, se hará una rotación de los factores debiéndose además, examinar las características de los ítems para su interpretación e intentar darles un nombre específico.

Se probará un rotación varimax de los factores y para esto se debe ir a la ventana de análisis factorial nuevamente y seleccionar *Rotación*.

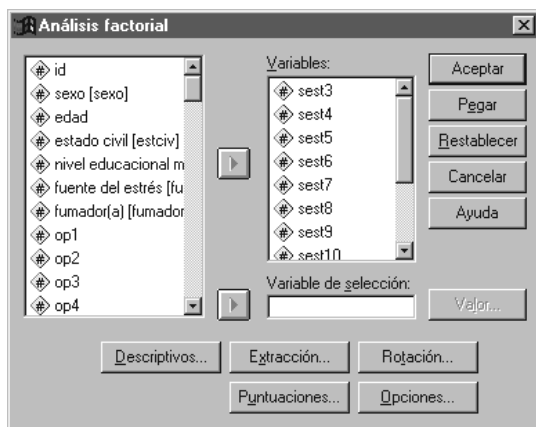


Figura 7.10. Análisis factorial.

Donde se muestra la siguiente ventana :

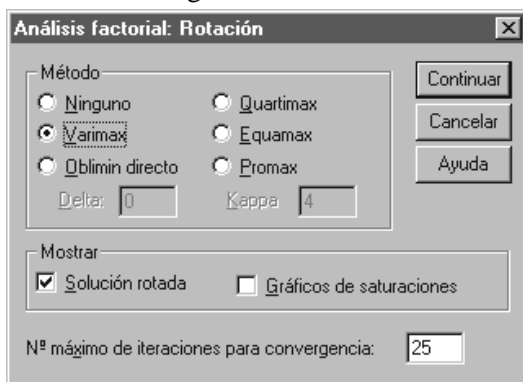


Figura 7.11. Rotación

Al ejecutar la rotación varimax, varían los porcentajes de varianza explicada por cada componente, pero el porcentaje de varianza explicado por las dos componentes sigue siendo de un 57,1%.

Varianza total explicada									
Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción			Suma de las saturaciones al cuadrado de la rotación		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	4.537	45.372	45.372	4.537	45.372	45.372	2.937	29.367	29.367
2	1.168	11.684	57.055	1.168	11.684	57.055	2.769	27.689	57.055
3	.896	8.957	66.012						
4	.697	6.969	72.982						
5	.675	6.751	79.733						
6	.501	5.007	84.740						
7	.466	4.661	89.401						
8	.384	3.842	93.243						
9	.376	3.763	97.006						
10	.299	2.994	100.000						

Método de extracción: Análisis de Componentes principales.

Tabla 7.8. Componentes y varianza total explicada.

En la tabla 7.9, se observa que: la primera componente se asocia ahora con los ítems SEST5, SEST6, SEST7, SEST8, SEST9 y SEST10, la segunda componente con los ítems, SEST1 SEST2, SEST3 y SEST4; nótese que algunos ítems (SES3, SES6, SES8) están asociados a las dos componentes; estos se denominan ítems “complejos” ya que saturan a más de una componente y no son usadas para nombrar las componentes. Así la primera componente deberá ser interpretada con los ítems SES5, SES7, SEST9 y SEST10; la segunda componente con los ítems SEST1, SEST2 y SES4. Según la definición de los ítems, el primer factor (o componente) se asocia con el *sentimiento positivo* mientras que el segundo factor, con el *sentimiento negativo*. Lo anterior, validaría la construcción de la escala de autoestima.



Matriz de componentes rotados<sup>a</sup>

	Componente	
	1	2
SEST3	.477	.511
SEST4	.222	.618
SEST5	.430	.336
SEST6	.593	.526
SEST7	.676	.291
SEST8	.554	.532
SEST9	.828	.111
SEST10	.801	6.773E-02
SEST1	.119	.801
SEST2	.132	.842

Método de extracción: Análisis de componentes principales.

Método de rotación: Normalización Varimax con Kaiser.

<sup>a</sup>. La rotación ha convergido en 3 iteraciones.

Tabla 7.9. Matriz de componentes rotados

### Ejercicios

1. Un equipo de psicólogos diseña una prueba para medir inteligencia espacial. Para ello seleccionan una muestra de siete personas y les aplican la prueba. Transcurridos dos meses, vuelve a aplicar el mismo test a los mismos sujetos. Las puntuaciones de las dos aplicaciones fueron las siguientes:

SUJETO	1ª aplicación	2ª aplicación
A	10	12
B	8	6
C	12	10
D	6	8
E	2	2
F	8	10
G	4	6

Obtenga un indicador de la fiabilidad del test e interprételo (Use un coeficiente de correlación).

2. A continuación se presentan las puntuaciones obtenidas por 4 personas en un test compuesto por 4 elementos.

SUJETO	Ítem1	Ítem2	Ítem3	Ítem4	Total
A	2	3	4	1	10
B	1	3	5	4	13
C	4	1	3	0	8
D	1	3	2	0	6
Varianzas	2,00	1,00	1,67	3,58	8,92

Obtenga un indicador de la fiabilidad del test (Use Alfa de Cronbach, empleando la fórmula y empleando SPSS)

3. Un psicólogo de un centro de educación especial ha medido la actitud hacia las matemáticas con un test de 3 ítems, valorados en una escala de 1 a 5, que ha pasado a 5 sujetos. Los resultados obtenidos han sido los siguientes:

Sujeto	Item1	Item2	Item3	Total
A	2	3	3	8
B	3	3	1	7
C	5	5	3	13
D	1	2	1	4
E	4	2	1	7

A la vista de los resultados, el psicólogo le pide que le ayude a seleccionar los dos ítems que más contribuyen a la consistencia interna del test (Use Alfa de Cronbach, empleando la fórmula y empleando SPSS)

4. Se lleva a cabo un análisis factorial sobre las puntuaciones en un cuestionario. Haciendo esto, se está examinando ...
- Su fiabilidad como consistencia interna
  - Su fiabilidad como estabilidad temporal
  - Su validez referida a un criterio
  - Su validez de constructo
5. Un psicólogo diseña cuatro pruebas para medir *Extroversión* (E1, E2, E3 y E4), tres para medir *Neuroticismo* (N1, N2 y N3) y dos para medir *Psicoticismo* (P1 y P2). Aplica las pruebas a una muestra de 100 sujetos y realiza un análisis factorial a partir de la matriz de correlaciones entre las distintas pruebas. A continuación se ofrece la matriz factorial rotada:

PRUEBA	F I	F II	F III
E1	0,877	0,068	0,090
E2	0,849	0,055	0,201
E3	0,791	0,003	0,076
E4	0,514	0,061	0,257
N1	0,014	0,853	0,090
N2	0,106	0,765	0,012
N3	0,082	0,643	0,114
P1	0,360	0,028	0,720
P2	0,243	0,066	0,687

- ¿Podemos afirmar que las pruebas diseñadas por el psicólogo poseen una adecuada validez factorial?
- ¿Qué cantidad de varianza explica el primer factor?



## **ANEXO I**

---

### **NOTAS SOBRE INFERENCIA**

La Estadística Inferencial es el área de la Estadística que se preocupa de la relación entre muestras y la población de la cual se han seleccionado, para emplear los resultados obtenidos en una muestra con el objeto de “inferir” o “tener una idea” acerca de los valores de las características de la población.

La Estadística Inferencial pretende resolver dos problemas fundamentales: la estimación de parámetros poblacionales a partir de estadísticas muestrales conocidas (Estimación Puntual y Estimación por Intervalos), y la toma de decisiones estadísticas acerca de hipótesis establecidas sobre la población, también con base en el conocimiento y los resultados obtenidos de una muestra de dicha población (Docimasia o Prueba de Hipótesis).

#### **ESTIMACIÓN POR INTERVALOS DE PARÁMETROS POBLACIONALES**

La estimación puntual de un parámetro poblacional desconocido consiste en emplear un estimador apropiado para realizar la estimación. Así por ejemplo, luego de tomar una muestra de 10 alumnos de la carrera de Pedagogía en Educación General Básica, se calcula la media aritmética de las edades y se encuentra que es 20,7 años. Basado en ese resultado, y considerando que la Teoría Estadística avala a la media aritmética como un excelente estimador de la Media Poblacional, podría estimarse la edad media de todos los alumnos de esa carrera también en 20,7 años.

Esta situación puede mejorarse, ofreciendo un conjunto (o intervalo) de valores posibles para el parámetro poblacional. Así por ejemplo, podría estimarse que la edad media de los alumnos de la carrera fluctúa entre 19,2 y 22,2 años (esto es  $20,7 \pm 1,5$  años). A la media aritmética (20,7 años) de la muestra se le ha agregado y restado la cantidad de 1,5 años, denominado *error de estimación*, para determinar el intervalo de posibles valores de la edad media de todos los alumnos de la carrera.

## PRUEBA DE HIPÓTESIS SOBRE PARÁMETROS POBLACIONALES

Se mencionó que las dos grandes áreas de la Inferencia son la estimación de parámetros y el someter a prueba (docimar, contrastar) hipótesis estadísticas. El empleo de estos procedimientos puede depender del conocimiento existente sobre una población. Si no hay información, posiblemente el interés se centre en estimar alguna característica de la misma. Si existe alguna información referente a la población de interés, el objetivo puede ser el aprovecharla para formular hipótesis.

### ***Hipótesis de Investigación***

Las hipótesis de investigación son las hipótesis sustantivas que guían la investigación, por estar fundamentadas teóricamente. Deben estar siempre expresadas como posibles contestaciones a una o varias preguntas de investigación. Además, se formulan en términos de las relaciones que existen entre las variables del estudio.

Por ejemplo, en un estudio sobre las edades de los estudiantes de Educación General Básica (EGB), podríamos preguntarnos (esta sería nuestra pregunta de investigación): ¿Cómo es el nivel de edades de los estudiantes de EGB respecto al resto de estudiantes de la universidad?

Una posible respuesta a esta pregunta de investigación (la hipótesis) sería postular, luego de analizar los fundamentos teóricos así como la información de otros estudios o de organismos relacionados con el tema, que la edad media (la variable de interés) de los estudiantes de EGB es menor que la edad media al resto de estudiantes de la universidad. Esta declaración no es otra cosa que nuestra hipótesis de investigación y constituye, en principio, nuestra respuesta tentativa a la pregunta de investigación. Pero también, los estudios relacionados o la información disponible, nos podrían llevar a postular que la edad media de los estudiantes de EGB es mayor que la edad media del resto de estudiantes de la universidad, o, si no hay seguridad en cuanto a la comparación, que sólo es diferente a la edad media de los otros estudiantes.

### ***Hipótesis Estadísticas***

Las Hipótesis Estadísticas corresponden a la traducción de la Hipótesis de Investigación en términos de aseveraciones o conjeturas acerca de la distribución de una o más variables. Específicamente, las hipótesis estadísticas pueden referirse a parámetros poblacionales (Medias, Proporciones, Varianzas etc.) o a la forma de la distribución (Normal, Uniforme etc.).

Operativamente se consideran la Hipótesis Nula ( $H_0$ ) y la Hipótesis Alternativa ( $H_1$ ). La primera, generalmente, establece lo conocido o lo conservador, mientras que la segunda, que es complementaria a la anterior, refleja lo novedoso o lo que usualmente interesa al investigador. Ambas hipótesis agotan las posibles alternativas, en cuanto a valores de parámetros o formas de distribución.

Siguiendo con el ejemplo, supongamos que se conoce la edad media de los estudiantes de la universidad, y es igual a 22,8 años. Ese valor es una característica de la población lo que se denota como  $m=22,8$  años. Entonces, las hipótesis sobre la edad media de los estudiantes de EGB se formularán respecto a este valor.

Las distintas situaciones señaladas permiten traducir la Hipótesis de Investigación en las siguientes Hipótesis Estadísticas, considerando que lo conocido debe formar parte de la Hipótesis Nula.

$$\begin{array}{lll} \text{a) } H_0: m = 22,8 & \text{b) } H_0: m \leq 22,8 & \text{c) } H_0: m \geq 22,8 \\ H_1: m \neq 22,8 & H_1: m > 22,8 & H_1: m < 22,8 \end{array}$$

Las hipótesis anteriores se denominan (a) bilateral, bidireccional o de dos colas, (b) y (c) se denominan unilaterales, unidireccionales o de una cola. Conviene tener presente que ante una Hipótesis de Investigación sólo una de las alternativas se debe considerar.

### ***Prueba de Hipótesis***

La Prueba de Hipótesis es el procedimiento o la regla de decisión que se emplea para decidir sobre la aceptación o no de la Hipótesis Nula. Se denomina genéricamente *prueba* en SPSS.

La Prueba de Hipótesis consiste en suponer verdadera la Hipótesis Nula y derivar las consecuencias que ello implica. Finalmente, se contrastan las consecuencias mencionadas con la información que aporta la muestra de la población, resumida en los estimadores respectivos que se denominan *estadísticos de prueba*. Si los resultados muestrales son coherentes con las consecuencias, ello se interpreta como que la Hipótesis Nula efectivamente es verdadera y se decidirá aceptarla. Por el contrario, si los resultados muestrales no son coherentes con las consecuencias, ello se interpreta como que la Hipótesis Nula no es verdadera y se decidirá no aceptarla (y se aceptará su complemento, la Hipótesis Alternativa). La decisión, en cualquier caso, involucra riesgos de equivocarse.



Continuando con el ejemplo, supongamos que las hipótesis de interés corresponden a :

$$H_0: m = 22,8 \text{ años}$$

$$H_1: m \neq 22,8 \text{ años}$$

Como se emplea una muestra para decidir, y si se supone que en verdad la edad media de los estudiantes de la Carrera de EGB es 22,8 años, las muestras de alumnos seleccionados pueden diferir de ese valor. El problema es determinar hasta que punto se pueden considerar diferencias, dado que la edad media realmente es 22,8. El monto del error admisible es denominado *nivel de significación* y simbolizado por  $\alpha$ . En todo caso, si en las muestras se obtuvieran valores demasiado bajos respecto a 22,8 o demasiado altos, ambos casos serían indicios de que la edad media de la carrera es diferente a 22,8. Por ello se denomina *bilateral* a la prueba respectiva.

Si se supone que la Hipótesis Nula es verdadera (se puede sólo suponer pues es imposible saberlo con certeza) y es rechazada por la prueba, se cometería el *error tipo I* cuya magnitud es  $\alpha$  y es el *nivel de significación*. Si se supone que la Hipótesis Nula no es verdadera y no es rechazada por la prueba, se cometería el *error tipo II* cuya magnitud es  $1 - b$  (el valor  $1 - b$  se denomina *potencia de la dócima*).

En las aplicaciones, generalmente se emplean niveles de significación del 5% o del 1% para tomar decisiones. Sin embargo, SPSS entrega los denominados *valores-p* (p-value) usualmente como *valores de significación* (Sign. Exacta, aproximada, asintótica etc.) los que facilitan mucho la decisión de aceptar o no la Hipótesis Nula, lo que en definitiva se realiza, comparando dicho valor con el nivel de significación decidido por el investigador ( por ejemplo 5% o 1%).

### ***Elementos de una Prueba de Hipótesis:***

A continuación, se detallan los principales elementos involucrados en una Prueba de Hipótesis:

- a) **Prueba:** Es el procedimiento o la regla de decisión que se emplea para decidir sobre la aceptación o no de la Hipótesis Nula. La selección de la prueba se realiza, considerando la Hipótesis de la Investigación y la respectivas Hipótesis Estadísticas, considerando las características de la muestra y los supuestos necesarios para desarrollar la prueba (p.ej. normalidad, aleatoriedad de la muestra, independencia etc.)
- b) **Hipótesis Nula ( $H_0$ ):** Es la hipótesis que señala lo tradicional y se supone que es verdadera para desarrollar el procedimiento de la prueba.
- c) **Hipótesis Alternativa ( $H_1$ ):** Es complementaria a  $H_0$ , refleja lo novedoso o lo que usualmente interesa al investigador. Ambas hipótesis agotan las posibles alternativas, en cuanto a valores de parámetros o formas de distribución.
- d) **Nivel de significación ( $\alpha$ ):** Corresponde a la probabilidad de rechazar una hipótesis que puede ser verdadera (se denomina también error tipo I). Valores usuales son 5% y 1% ( $\alpha=0,05$  o  $\alpha=0,01$ ). Es fijado por el investigador.
- e) **Estadística de Prueba:** Es la estadística muestral empleada en la prueba. Tiene una distribución en el muestreo con características que se asocian con el parámetro que se estima. Este valor lo entrega SPSS.
- f) **Región Crítica (RC):** Corresponde al conjunto de valores de la *estadística de prueba* que llevan al rechazo de  $H_0$ . Dependen del nivel de significación fijado y de la Hipótesis alternativa. Puede corresponder a un área en

una cola de la distribución de la *estadística de prueba* (en dúcimas unilaterales) o a dos áreas en las colas de esa distribución (en dúcimas bilaterales). Cuando se conoce el valor-p (valor de significación), no es necesario indicar la *región crítica* como ocurre en las pruebas de SPSS.

- g) Región de Aceptación (RA):** Es el complemento de la región crítica. No es necesario indicarla pues toda la prueba se realiza en función de la Hipótesis Nula y su región crítica.



## ANEXO II

---

### ESQUEMA DE ALGUNAS PRUEBAS ESTADÍSTICAS

Con el objeto de complementar la explicación del texto así como para facilitar su aplicación, se presentan en orden alfabético los resúmenes de algunas pruebas mencionadas en el texto. Se indica el propósito de la misma, la hipótesis nula general y la regla para decidir sobre la hipótesis nula basada en los valores “p” (o de significación) entregados por SPSS. Se complementa con un ejemplo de aplicación, empleando los formatos de presentación de resultados y nomenclatura de SPSS para facilitar su interpretación. Los detalles técnicos se pueden encontrar en los textos de Estadística.

#### 1. ALFA DE CRONBACH ( $\alpha$ )

Se pueden emplear los resultados entregados por SPSS para complementar el estudio de la confiabilidad mediante la aplicación de pruebas o cálculo de intervalos de confianza respectivos. Para esto se emplea el estadístico de prueba  $F = (1 - \alpha) / (1 - \alpha_0)$  el cual tiene una Distribución F con  $(N - 1)$  y  $(N - 1)(n - 1)$  grados de libertad, siendo  $N$  el número de sujetos en la muestra,  $n$  el número de ítems del test,  $\alpha$  el valor poblacional postulado y  $\alpha_0$  el valor obtenido en la muestra. (No confundir la notación de este alfa de Cronbach con el empleado para el nivel de significación).

**Hip. Nula:**  $\alpha = 0$

Se pueden plantear pruebas unilaterales o bilaterales

**Decisión:** Si  $F$  observado es mayor al  $F$  de tabla con  $(N-1)$  y  $(N-1)(n-1)$  grados de libertad, se rechaza la Hipótesis Nula.

**Ejemplo:**

Para un test de inteligencia espacial, con 50 ítemes, se plantea la Hipótesis que el Coeficiente Alfa de Cronbach es mayor que cero.

$$H_0 : \alpha \leq 0 \quad \text{versus} \quad H_1 : \alpha > 0$$

Los resultados de la aplicación a una muestra de 40 personas presentan un valor de Alfa de Cronbach  $\alpha_0 = 0,75$ . Por lo tanto:

$$F_0 = (1-0)/(1-0,75) = 4$$

Si se emplea un nivel de significación del 5%, se compara el valor antes calculado con el entregado por la tabla  $F$  con  $(40-1)$  y  $(10-1)(40-1)$  grados de libertad, esto es 39 y 351 grados de libertad. El valor más cercano en la tabla  $F$  para  $F(0,95; 39; 351)$  es 1,39 (para 40 e infinitos grados de libertad).

Luego, dado que el valor calculado para  $F_0 = 4$  es mayor que 1,39 se concluye que la Hipótesis Nula es rechazada y por lo tanto, se acepta que  $\alpha > 0$  esto es, la confiabilidad de la escala es significativa.

## 2. ANÁLISIS DE VARIANZA DE UN FACTOR

Esta prueba tiene como objetivo estudiar si las Medias de  $K$  poblaciones independientes ( $m_1, m_2, \dots, m_k$ ) son iguales, o si al menos una, es diferente de las otras.

**Hip. Nula:**  $m_1 = m_2 = \dots = m_k$

**Supuestos:** Las poblaciones tienen distribución normal. Las

varianzas poblacionales son iguales. Las observaciones son independientes.

**Decisión:** Si  $p$  observado es menor que nivel fijado (p.ej. 0,05 o 5%) se rechaza la Hip. Nula.

**Ejemplo:**

Se prueba si existe diferencias entre edades medias ( $m_i$ ) de estudiantes de 7 carreras de la universidad

$$H_0 : m_1 = m_2 = \dots = m_7$$

En tabla ANOVA respectiva en la Línea indicada como Entre Grupos:

$$F=7,752 \quad \text{Sig.}=0,000$$

**Decisión:** Se rechaza la Hipótesis Nula. Existen diferencias entre las edades medias de los estudiantes de las carreras (al menos un grupo es diferentes de los otros)

Empleando Comparaciones múltiples (Scheffé, Bonferroni, etc.) se pueden detectar diferencias significativas. Serán las que presenten valores de significación (Sign.) menores al fijado (5% por ejemplo).

### 3. ASIMETRIA

Indica el grado de simetría o asimetría de la distribución

Interpretación del Coeficiente (Estadístico):

<1: Distribución asimétrica negativa (cola más larga hacia la izquierda)

=0: Distribución simétrica

>1: Distribución asimétrica positiva (cola más larga hacia la derecha)

**Error Tipico:** Indica la estimación de la variabilidad del coeficiente

**Coefficiente estandarizado**= Coeficiente/Error Típico (no lo entrega SPSS)

Para  $n > 150$  se distribuye aprox. Como normal.

Se puede emplear para probar si la muestra proviene de una población simétrica (o asimétrica positiva o negativa).

**Hip. Nula:** Distribución es simétrica

**Decisión:** Si  $p$  observado es menor que nivel fijado (p.ej 0,05 o 5%) se rechaza Hip. Nula

**Ejemplo:**

Se analiza la simetría de una distribución:

Coeficiente= 0,462    Error Típico= 0,179

Coeficiente Estandarizado =  $0,462/0,179 = 2,58$

**Decisión:** Como el coeficiente estandarizado es mayor que 1,96 (valor de distribución normal asociado con una significación del 5%) se rechaza Hip. Nula a ese nivel de significación. Por lo tanto, la distribución es asimétrica positiva.

#### 4. CHI-CUADRADO (Bondad de Ajuste)

Esta prueba tiene como objetivo comparar las posibles diferencias entre las frecuencias observadas y las esperadas en una distribución, de acuerdo a una determinada hipótesis (por ejemplo que distribución es Normal etc.).

**Hip. Nula:** Distribución empírica se ajusta a una distribución supuesta (Normal, uniforme etc.)

**Decisión:** Si  $p$  observado es menor que nivel fijado (p.ej 0,05 o 5%) se rechaza la Hip. Nula

**Ejemplo:**

Se desea probar si la distribución de frecuencias de una variable nominal, con 6 categorías, es uniforme (todas las



categorías tienen las mismas frecuencias).

**Hip. Nula:** Distribución empírica se ajusta a distribución uniforme.

Chi-Cuadrado= con 118 g.l=5  $p=0,0000$ .

**Decisión:** Se rechaza Hip. Nula. La distribución empírica no tiene una distribución uniforme (hay diferencias entre las frecuencias de las categorías).

## 5. CHI CUADRADO (Independencia)

En procedimiento *Estadística descriptiva/Tablas de Contingencia*, esta prueba tiene como objetivo analizar la independencia (o falta de asociación) de las variables que componen la tabla de contingencia.

**Hip. Nula:** Las variables son independientes.

**Decisión:** Si  $p$  observado es menor que nivel fijado (p.ej 0,05 o 5%) se rechaza la Hip. Nula.

**Ejemplo:**

Se desea probar si las clasificaciones por sexo (hombre y mujer) y capacitación (capacitado, no capacitado) están relacionadas.

Chi-Cuadrado de Pearson= 0,141 g.l=1  $p=0,707$ .

**Decisión:** No se rechaza la Hip. Nula. No existe asociación entre las variables.

## 6. COEFICIENTE DE CORRELACIÓN DE PEARSON ( $r$ )

Esta prueba tiene como objetivo estudiar si la correlación entre dos variables, que tienen una distribución Normal bivariada, es diferente de 0. En caso de ser 0 indicaría que las variables son independientes.

**Hip. Nula:** Las variables son independientes ( $r = 0$ ).

**Decisión:** Si  $p$  observado es menor que nivel fijado (p.ej 0,05 o 5%) se rechaza la Hip. Nula (En SPSS se puede indicar el nivel y pedir que marque correlaciones significativas, esto es, aquellas cuyo valor  $p$  sea menor o igual al fijado).

**Ejemplo:**

Se analiza la correlación entre dos pruebas “Total Stress Percibido” y “Total de Afecto Negativo”.

Correlación de Pearson= 0,674  $p=0,000$  (bilateral).

Decisión: Se rechaza Hip. Nula. Existe asociación entre variables.

## 7. CURTOSIS

Indica el grado de apuntamiento de la distribución en comparación con la distribución Normal (con su misma Media y Desv. Típica)

Interpretación del Coeficiente (Estadístico):

$<1$  : Distribución más aplastada que la Normal (platicúrtica).

$=0$  : Distribución igualmente apuntada que la Normal (mesocúrtica).

$>1$  : Distribución más apuntada que la Normal (leptocúrtica).

**Error Típico:**

Indica la estimación de la variabilidad del coeficiente

**Coeficiente estandarizado**= Coeficiente/Error Típico (no lo entrega SPSS)

Para  $n>150$  se distribuye aproximadamente como normal.

Se puede emplear para determinar si efectivamente no hay apuntamiento (o si éste existe).

**Hip. Nula:** Distribución es igualmente apuntada que la Normal.

**Decisión:** Si  $p$  observado es menor que nivel fijado (p.ej 0,05 o 5%) se rechaza Hip. Nula.

**Ejemplo:**

Se analiza la curtosis de una distribución:

Coficiente= -0,596    Error Típico= 0,355.

Coficiente Estandarizado =  $-0,596/0,355 = -1,67$ .

Decisión: Como coef. Estandarizado está entre  $-1,96$  y  $1,96$  (valor de distribución normal asociado con una significación del 5%) no se rechaza Hip. Nula a ese nivel de significación. Por lo tanto, se acepta que la distribución tiene una apuntamiento semejante al de su Distribución Normal asociada.

## 8. KAPPA

Se emplea para medir acuerdo entre jueces que califican los mismos objetos o sujetos. Toma el valor 1 para el acuerdo perfecto y 0 para desacuerdo completo. Para valores intermedios se proponen los siguientes conceptos:

Menos de 0,40            Acuerdo pobre.

0,40 a 0,75            Buen nivel de acuerdo.

Mayor de 0,75 Excelente nivel de acuerdo.

**Hip. Nula:**  $k=0$

**Decisión:** Si  $p$  observado es menor que nivel fijado (p.ej 0,05 o 5%) se rechaza la Hip. Nula

**Ejemplo:**

En la evaluación de candidatos a un cargo realizadas por 2 jueces, se obtienen los siguientes resultados:

Medida de acuerdo  $Kappa=0,4728$   $p=0,000$ .

**Decisión:** Se rechaza Hip. Nula. Existe acuerdo (moderado) entre jueces.

**9. KOLMOGOROV-SMIRNOV**

Esta prueba se emplea para comparar distribuciones:

- Como (K-S Lilliefors) en Procedimiento *Explorar/Gráficos* se emplea para determinar si una distribución empírica tiene una distribución Normal.

- Como (K-S Z) en Procedimiento *Estadística No Paramétrica* se emplea para comparar una distribución empírica con otras distribuciones (Normal, Uniforme, Poisson) o para comparar si dos muestras tienen una misma distribución.

**Hip. Nula:** Distribución empírica se ajusta a distribución Normal (con media y desv. estándar estimadas desde la muestra).

**Decisión:** Si p observado es menor que nivel fijado (p.ej 0,05 o 5%) se rechaza Hip. Nula.

**Ejemplo:**

Se prueba si una distribución empírica se ajusta a la Normal:

K-S (Lilliefors)=0,1343 con 118 g.l  $p=0,0000$ .

**Decisión:** Se rechaza Hip. Nula, distribución que se prueba no se ajusta a la Normal.

**10. LEVENE**

Esta prueba tiene por objetivo analizar si las varianzas de

varias poblaciones son o no iguales.

**Hip. Nula:** Varianzas de Poblaciones definidas por Factores son homogéneas (iguales).

**Decisión:** Si  $p$  observado es menor que nivel fijado (p.ej 0,05 o 5%) se rechaza Hip. Nula.

**Ejemplo:**

Se comparan varianzas de cinco grupos:

Levene=1,2205 con 4 y 110 g.l.  $p=0,3063$

Decisión: No se rechaza Hip. Nula. Las varianzas de los distintos grupos son homogéneas (iguales).

## 11. PRUEBA T PARA UNA MUESTRA

Esta prueba tiene como objetivo estudiar si la media de una población ( $m$ ) tiene un valor conocido  $m_0$ .

**Hip. Nula:**  $m = m_0$

Se consideran pruebas unilaterales o bilaterales.

**Supuestos:** La población tiene distribución normal.

**Decisión:** Si  $p$  observado es menor que nivel fijado (p.ej 0,05 o 5%) se rechaza la Hip. Nula.

**Ejemplo:**

Se prueba si puntaje medio poblacional en un test de Matemáticas es 100 puntos. Se aplicó el test a 50 alumnos. Puntaje medio en la muestra fue 108,4 puntos.

$$H_0: m = 100 \text{ vs. } H_1: m \neq 100$$

TEST MATEMAT  $t=5,482$  g.l.=49 Sig. (bilateral)=0,000

**Decisión:** Se rechaza Hip. Nula. Puntaje medio es diferente a 100 puntos.

## 12. PRUEBA T PARA DOS MUESTRAS INDEPENDIENTES

Esta prueba tiene como objetivo estudiar si las medias de dos poblaciones independientes ( $m_1$  y  $m_2$ ) son iguales, o si una es mayor o menor que la otra.

**Hip. Nula:**  $m_1 = m_2$

Se consideran pruebas unilaterales o bilaterales.

**Supuestos:** Las poblaciones tienen distribución normal. Las varianzas poblacionales puede suponerse que sean iguales o no.

**Decisión:** Si  $p$  observado es menor que nivel fijado (p.ej 0,05 o 5%) se rechaza la Hip. Nula.

**Ejemplo:**

Se prueba si puntaje medio poblacional en un test de Matemáticas es igual para hombres ( $m_1$ ) y para mujeres ( $m_2$ ). Se emplearon muestras de 12 hombres y 16 mujeres

$$H_0 : m_1 = m_2 \quad \text{vs} \quad H_1 : m_1 \neq m_2$$

TEST MATEMAT Se han asumido varianzas iguales

$t = -0,501$  g.l.=26 Sig. (bilateral)=0,621.

No se han asumido varianzas iguales.

$t = -0,497$  g.l.=23,2 Sig. (bilateral)=0,624.

**Decisión:** En ambos casos, no se rechaza Hip. Nula. Luego, se acepta que puntajes medios para hombres y mujeres son iguales.

## 13. PRUEBA T PARA DOS MUESTRAS RELACIONADAS

Esta prueba tiene como objetivo estudiar si las medias de dos poblaciones no necesariamente independientes ( $m_1$  y  $m_2$ ) son

iguales, o si una es mayor o menor que la otra, considerando pares de valores muestrales ( $X_i$ ,  $Y_i$ ) denominados “datos apareados”, de los cuales se emplean las diferencias.

**Hip. Nula:**  $m_1 = m_2$  (ó  $m_1 - m_2 = 0$ ).

Se pueden plantear pruebas unilaterales o bilaterales.

**Supuestos:** Las población de diferencias tienen distribución normal.

**Decisión:** Si  $p$  observado es menor que nivel fijado (p.ej 0,05 o 5%) se rechaza la Hip. Nula.

**Ejemplo:**

Se prueba si es efectivo un taller para reducir el estrés. Se empleó una muestra de 10 sujetos. Se midió con un test a cada uno antes ( $m_1$ ) y después ( $m_2$ ) de realizar el taller.

$$H_0: m_1 = m_2 \quad \text{vs} \quad H_1: m_1 \neq m_2$$

PAR ANTES-DESPUES: Media de la diferencia= 1.

$$t = 1,399 \quad \text{g.l.} = 9 \quad \text{Sig. (bilateral)} = 0,195$$

**Decisión:** No se rechaza Hip. Nula. Luego, se acepta que puntajes medios antes y después son iguales, esto es que el taller no fue efectivo para reducir estrés.

## 14. REGRESIÓN LINEAL : PRUEBA F TABLA ANOVA

Esta prueba tiene por objetivo estudiar si el modelo de Regresión Lineal (recta  $Y = a + bX$ ) permite explicar la relación entre las variables  $X$  (dependiente) e  $Y$  (independiente), esto es, si la pendiente  $b$  es diferente de 0.

**Hip. Nula:** El modelo de Regresión Lineal no explica la relación entre las variables ( $b = 0$ ).

**Decisión:** Si  $p$  observado es menor que nivel fijado (p.ej 0,05 o 5%) se rechaza la Hip. Nula.

**Ejemplo:**

Se estudió la relación entre las variables “Resistencia al esfuerzo” (var. Dependiente) y “Edad” (var. Independiente).

En tabla ANOVA respectiva en Línea indicada como Regresión:

$F=48,123$  Sig.=0,000.

**Decisión:** Se rechaza Hip. Nula. Existe relación lineal entre las variables.

## 15. REGRESIÓN LINEAL : PRUEBA $t$ PARA COEFICIENTES

Esta prueba tiene por objetivo estudiar si parámetros  $a$  y  $b$  del modelo de Regresión Lineal (recta  $Y = a + bX$ ) siendo variables  $X$  (dependiente) e  $Y$  (independiente) son o no iguales 0. En caso de ser sólo  $a=0$  indicaría que la recta pasa por el origen. En caso de ser  $b=0$  indicaría que no hay relación entre las variables.

En el caso de Regresión Lineal, esta prueba es equivalente a la Prueba  $F$ . Se puede verificar la siguiente relación entre los estadísticos de prueba:  $t^2 = F$

**Hip. Nula 1:** El intercepto del modelo de Regresión Lineal es cero ( $a=0$ )

**Hip. Nula 2:** La pendiente del modelo de Regresión Lineal es cero ( $b=0$ )

**Decisión:** Si  $p$  observado es menor que nivel fijado (p.ej 0,05 o 5%) se rechaza la Hip. Nula.



**Ejemplo:**

Se estudió de la relación entre “Resistencia al esfuerzo” (var. Dependiente) y “Edad” (var. Independiente).

a) En tabla de COEFICIENTES respectiva en Línea indicada como (Constante):

$t=5,683$  Sig.=0,000.

**Decisión:** Se rechaza Hip. Nula. El intercepto es diferente a 0.

b) En la misma tabla de COEFICIENTES en Línea indicada como Edad:

$t=6,937$  Sig.=0,000.

**Decisión:** Se rechaza Hip. Nula. La pendiente es diferente a 0.

**Observación:** Notar que  $t^2=6,937^2=48,123=F$ .

**16. REGRESIÓN MÚLTIPLE**

En el caso de Regresión Múltiple, la interpretación de los resultados del ANOVA así como la tabla de COEFICIENTES respectivos es semejante a los que se indican en Regresión Lineal.

**17. SHAPIRO-WILKS**

Es similar a Kolmogorov-Smirnov para probar normalidad.

I.S.B.N. 956-291-627-8

Registro de Propiedad Intelectual N° 130.367

