ESTADÍSTICA PARA NO ESTADÍSTICOS

UNA GUÍA BÁSICA SOBRE LA METODOLOGÍA CUANTITATIVA DE TRABAJOS ACADÉMICOS

Héctor Galindo-Domínguez

Economía, Organización y Ciencias Sociales

3ciencias

ESTADÍSTICA PARA NO ESTADÍSTICOS

UNA GUÍA BÁSICA SOBRE LA METODOLOGÍA CUANTITATIVA DE TRABAJOS ACADÉMICOS

Héctor Galindo-Domínguez

Editorial Área de Innovación y Desarrollo,S.L.

ÁREA DE INNOVACIÓN Y DESARROLLO, S.L.

C/Alzamora, 17- 03802- ALCOY (ALICANTE) info@3ciencias.com

Primera edición: **marzo 2020**

ISBN: **978-84-121459-3-9**

DOI: **https://doi.org/10.17993/EcoOrgyCso.2020.59**

ÍNDICE DE CONTENIDOS

ÍNDICE DE FIGURAS

ÍNDICE DE TABLAS

PRÓLOGO

Los seres humanos, desde sus comienzos con el descubrimiento del fuego, siempre han sido seres curiosos, seres que han querido saber más, seres que cuestionaban, seres que se preguntaban. Tal vez, esta curiosidad por querer saber más acerca del mundo que nos rodea ha sido el motor por el que hoy en día cada una de nuestras culturas es como es.

Esta tarea, en una sociedad tan compleja como la actual, en la que la información es muy variada, e incluso en ocasiones excesiva, la investigación se muestra como la herramienta imprescindible para acercarse cada vez más al conocimiento de la realidad.

Esta herramienta, no es una herramienta con normas arbitrarias, sino más bien todo lo contrario. Se trata de un campo muy amplio en el que las pautas y las reglas de actuación están muy especificadas.

No obstante, se nos presenta la dificultad de que, en ocasiones, estas pautas y reglas sobre cómo actuar para conocer una determinada realidad pueden resultar excesivamente complejas para personas ajenas al ámbito de la investigación cuantitativa, de la estadística y de las matemáticas. Es precisamente este el aliciente por el que se presenta el presente trabajo.

El objetivo primordial de este libro no es otro que el de acercar al lector que se inicia en el ámbito de la investigación cuantitativa los pilares teóricos esenciales y las herramientas prácticas básicas para poder llevar a cabo estudios de índole estadística.

El presente trabajo, se espera que pueda ser de utilidad, especialmente, para aquellos estudiantes, que por los motivos que fueren, se hallan en proceso de desarrollo de cualquier trabajo académico oficial, como puede ser el caso de un trabajo de fin de grado, un trabajo de fin de máster o una tesis doctoral. Además, se presenta un trabajo de gran utilidad no solo para estudiantes, sino también para docentes universitarios e investigadores que quieren comenzar a desarrollar investigaciones en las que el dominio de la estadística se muestra como esencial. En vista de estos dos grupos, se presenta un trabajo en el que se ha adaptado la complejidad que puede haber detrás de las matemáticas que hay en la estadística, para que personas ajenas al ámbito, de lo que comúnmente se conoce como ciencias puras y/o personas con conocimientos muy básicos de estadística, puedan introducirse al desarrollo de estudios de carácter cuantitativo.

La estructura de este trabajo está dividida en 10 capítulos, pensados de manera estratégica, y siguiendo siempre un hilo conductor, yendo de lo más simple a lo más complejo, partiendo de la teórica, para posteriormente ver la práctica y empleando como estrategia principal para complementar las explicaciones, las ejemplificaciones, pues estas ayudan en muchas situaciones, a comprender mejor un determinado concepto.

El primer capítulo, llamado *diseño de estudios cuantitativos*, sirve para situar al lector en la organización y el proceso que debe seguir la amplia mayoría de investigaciones en las que se use este enfoque. A medida que el capítulo vaya avanzando, se va poniendo el foco de atención en aquellos conceptos y procedimientos que pueden resultar de mayor utilidad para el lector.

Tras situar al lector en el enfoque de los estudios cuantitativos, en el segundo capítulo, llamado *el contraste de hipótesis y la decisión estadística*, se le explica al lector el proceso más esencial de los procesos estadísticos, por el que se presentan y se contrastar unas determinadas hipótesis para obtener unos determinados resultados.

Ya, conociendo un poco más sobre este proceso de contraste de hipótesis, es hora de entrar en materia, llegando al capítulo III, llamado *Introducción al programa estadístico SPSS Statistics*. En este capítulo, se explica de manera sucinta las partes más esenciales y prácticas de uno de los programas estadísticos más empleados a nivel mundial: SPSS Statistics. Se espera que el lector, desarrolle prácticas de manera complementaria a medida que lee este capítulo, para profundizar más en las funciones mencionadas en este capítulo.

Posteriormente, en el capítulo IV, *llamado características y consideraciones en el uso de instrumentos de medida documental,* tras conocer lo básico sobre el funcionamiento del programa estadístico que se suele emplear, se le presenta al lector la posibilidad de profundizar en las características y las consideraciones más significativas de los instrumentos de medición documental, que serán, principalmente, los tipos de instrumentos que se emplearán en la mayoría de estudios de enfoque cuantitativo.

Y es en este momento, en el que tras esperar haber obtenido una base lo suficientemente sólida para comprender la esencia que subyace a los procedimientos estadísticos, en los capítulos V, VI y VII, llamados *los análisis descriptivos*, *los análisis bivariados* y *los análisis multivariados*, respectivamente, se presentan los principales análisis estadísticos comúnmente empleados en un alto porcentaje de estudios cuantitativos.

A modo de ampliación para los lectores más curiosos, también se presenta el capítulo VIII, llamado *introducción a otros análisis multivariantes*, una serie de análisis multivariantes no recogidos en el capítulo VII, pero que pueden ser de utilidad conocerlos, especialmente para aquellos casos en los que se pretenda realizar investigaciones de mayor calidad. El fin de este capítulo, no es el de profundizar en exceso en estos tipos de análisis, pues este fin no iría acorde al objetivo del libro en sí, sino el de situar al lector en dónde se encuentra y hacerle ver que existen todavía más caminos que se pueden recorrer.

Finalmente, el libro concluye con dos capítulos más. Por una parte, el capítulo IX recoge la bibliografía empleada a lo largo de todas las páginas, y queda a disposición del lector para ser estudiada de manera independiente si así lo considerase. Por otra parte, el capítulo X, llamado *glosario*, recoge aquellas palabras clave que han ido saliendo a lo largo de las páginas y que pueden ser de utilidad para comprobar, una vez finalizada la lectura del libro, el nivel de adquisición de estos conceptos.

Se espera que esta guía sea un trabajo de ayuda en los quehaceres académicos de estudiantes y docentes.

CAPÍTULO I: DISEÑO DE ESTUDIOS CUANTITATIVOS

1.1. Introducción

El proceso de investigación que se lleva a cabo a través del paradigma cuantitativo es un proceso caracterizado por una serie de etapas específicas, que son necesarias conocer antes de empezar directamente con "la acción".

En esta línea, diferenciamos ocho pasos que tenemos que desarrollar de manera ordenada, una tras otra, y mediante las cuales nuestro estudio irá paulatinamente cobrando vida. A continuación describimos en detalle cada uno de los pasos a seguir en cualquier estudio de índole cuantitativa.

1.2. Tema de investigación

El primer paso para llevar a cabo un estudio cuantitativo es determinar cuál será el tema de la investigación.

Para conocer el tema sobre el que basará nuestro estudio podemos partir de diferentes partes, como cuáles son mis intereses en base a experiencias previas o qué creo que puede resultar de utilidad para la sociedad.

Otras fuentes, para poder elegir el tema de investigación pueden ser:

- Materiales escritos, como libros, revistas o periódicos, en los cuales se recoja ya futuras líneas de investigación.
- Teorías que necesiten ser verificadas.
- Observaciones sobre hechos específicos. Estas observaciones nos generarán algunas creencias que deberán ser comprobadas.
- Petición expresa de autoridades académicas, científicas, políticas o administrativas. En este caso, la institución en sí se encargará de presentarnos algunas pinceladas o ideas muy superficiales de lo que se pretende investigar.

Algunas de las preguntas que podemos hacernos para decidir el tema pueden ser (Blaxter, Hughes y Tight, 2008):

- ¿Cuáles son mis intereses y motivaciones?
- ¿Qué deseo estudiar? ¿Con qué tema o área más amplia se relacionaría la investigación?

- ¿Cuáles serían los principales conceptos?
- ¿De qué tiempo dispongo?
- ¿Qué recursos están a mi alcance?
- ¿Qué apoyos necesitaría?

Esta etapa es una de las decisiones más importantes de todo el proceso, pues en un mundo en el que se busca cada vez más la eficiencia, resultaría inútil realizar una investigación sobre un tema que está excesivamente investigado ya. Es por este motivo, que es necesario que el tema que elijamos cumpla con algunas características como originalidad, interés social, viabilidad en su desarrollo y utilidad.

A modo de ilustrar un poco el camino, se muestra a modo de ejemplo los siguientes temas de investigación:

- Efectividad del aprendizaje cooperativo en la mejora del clima social en las aulas de Educación Primaria.
- La orientación motivacional en estudiantes universitarios.
- Relación entre la inteligencia emocional y la resiliencia en estudiantes de FP.

1.3. Formulación de preguntas de investigación y objetivos

Una vez definido el tema, el siguiente paso es el de concretar cuáles van a ser los objetivos que se pretenden buscar.

Una buena manera de definir los objetivos es comenzar planteándose preguntas de investigación, sobre como investigador, qué querría saber más. A pesar de que este primer paso, seguramente, nos arroje unas respuestas amplias, posteriormente ya habrá tiempo para ir especificándolos más.

Las preguntas de investigación deben servirnos para percatarnos de qué es lo que queremos conseguir. Generalmente, suele formularse una pregunta de investigación general, precedida por una serie de preguntas de investigación específicas. Unos ejemplos sobre preguntas de investigación pueden ser las siguientes:

- ¿Sacan las chicas mejores notas que los chicos en ciencias?
- ¿El uso de ordenadores puede reducir la ansiedad hacia las matemáticas?

- ¿Qué metodología activa es más efectiva para desarrollar el pensamiento crítico en estudiantes universitarios?

Estas **preguntas** deben ser la base para **formular los objetivos** que guiarán el trabajo completo. Estos objetivos tienen que cumplir con una serie de características:

- Los objetivos tienen que estar redactados de forma clara y específica. No puede caber lugar a dudas sobre qué es lo que se pretende investigar. Por ejemplo, un objetivo mal redactado sería: "Conocer la efectividad de la metodología ABP". En este caso, la palabra efectividad, a pesar de que se entiende, no deja claro qué es lo que persigue. ¿Qué entendemos por efectividad en este caso? Sería necesario concretar mucho más, por ejemplo, diciendo que entendiendo la efectividad como mejora de la motivación intrínseca del alumnado de Educación Primaria.

- Los objetivos tienen que ser relevantes. No vale con elegir cualquier cosa. Los objetivos deben ser imprescindibles para dar respuesta al tema de investigación planteado previamente.

- Los objetivos tienen que poder responderse. Especialmente en investigadores novatos, es común plantear objetivos muy ambiciosos. En esta línea, cabe recordar que nuestro estudio únicamente realizará una pequeña aportación al mundo científico y que no debemos ir buscando que recoja datos de 50 variables diferentes, pues en última instancia, el estudio aparte de resultar lioso para el lector será mucho más complejo de que sea viable y se pueda finalizar. Es por esto por lo que es mejor presentar unos objetivos sencillos a los que se les pueda dar respuesta que unos objetivos muy abiertos que se nos escapen de las manos por falta de recursos materiales, personales o funcionales (tiempo, dinero…).

- Los objetivos deben estar relacionados con el tema de investigación y deben guiar a este.

En esta línea, algunos objetivos, a modo de ejemplo, pueden ser los siguientes:

- Conocer si existen diferencias significativas en las calificaciones de ciencias en base al género.

- Conocer si el uso de ordenadores permite reducir los niveles de ansiedad ante las matemáticas.

- Conocer si existen diferencias significativas en la efectividad entre la metodología X y la metodología Y en el desarrollo del pensamiento crítico.
-

1.4. Revisión de la literatura

Una vez conocidos los objetivos que guiarán el estudio, llega el momento de conocer qué es lo que se sabe sobre ese tema específicamente; qué es lo que se sabe. Debemos asegurarnos de que todas las preguntas de investigación han sido estudiadas en esta fase de revisión.

Realizar una buena **revisión de la literatura** allanará considerablemente el camino en cuanto que:

- Nos permitirá ver cómo se han planteado los objetivos de otro estudio similar al nuestro. Esto nos dará la opción de modificar levemente algunos matices de nuestros objetivos, tras contrastar ambas opciones.
- Nos permitirá conocer cuál puede ser una posible manera de dar solución a los objetivos. Esta información nos ayudará a conocer y comparar cuáles pueden ser las metodologías, diseños y pruebas estadísticas más apropiadas para resolver los enigmas.
- Finalmente, nos permitirá contrastar los resultados de nuestro estudio, una vez concluido, con los resultados e ideas más significativas de otros estudios similares.

Realizar una buena revisión de la literatura depende en gran parte del proceso de búsqueda que se siga. Ya no es solo, coger cualquier cosa que veamos, sino que resulta necesario evaluar si realmente esa información es útil y de calidad. Principalmente, realizaremos nuestra búsqueda bibliográfica a través de dos fuentes:

- Fuentes manuales: A través de bibliotecas, que recogen libros, manuales, tesis doctorales… de índole muy diversa. Esta opción tiene la desventaja de que es un proceso de búsqueda mucho más lento que si se realiza una búsqueda por internet, pero tiene la ventaja de que puede darse el caso de que estas bibliotecas tengan artículos y documentos no accesibles por internet. Generalmente, estos centros pagan mensual o anualmente una cuota para recibir libros, manuales, artículos… que no son accesibles a través del internet ordinario, si no se paga.

- Fuentes digitales: A través de diferentes buscadores, como Web of Science, Scopus, Google Scholar, Dialnet, etc. Emplear este tipo de fuentes cuenta con la ventaja de que agiliza mucho el proceso de búsqueda, pero no nos asegura que siempre tengamos acceso a todo el contenido, pues existen ciertas revistas en las que se necesita de una suscripción.

Finalmente, unos consejos que pueden resultarnos de utilidad a la hora de buscar bibliografía para nuestro marco teórico son:

- Evaluar la calidad científica del artículo. En caso de tener que decidir, optar principalmente por aquel que posee un mejor diseño, indexado en una mejor revista, etc.
- Acotar lo máximo posible la búsqueda a nuestros objetivos de investigación, dándole prioridad a las publicaciones más recientes.
- No olvidarse de buscar fuentes bibliográficas en otros idiomas, como en inglés.
- Fijarse bien, a través de la lectura de varios artículos sobre el mismo tema, cuáles son las referencias bibliográficas más constantes sobre ese tema. De este modo, podremos identificar aquellas citas que pueden resultar de más interés para nuestro trabajo.
- Debe seguir un modo específico de ser citada: Generalmente, se suele emplear el modo propuesto por la Asociación Americana de Psicología (APA), aunque existe una decena de modos diferentes. Este criterio suele venir dado de fuera, por la propia universidad, en caso de realizar un TFG, TFM o tesis, o por la propia revista o editorial en caso de querer publicar un artículo o libro.

1.5. Diseño metodológico

El diseño de la metodología supone planificar cómo se va a recoger y analizar la información para responder a las preguntas de investigación y lograr los objetivos de investigación planteados.

Generalmente, dentro de este epígrafe es necesario comentar varias categorías que deben incluirse imprescindiblemente. Estamos hablando del diseño del estudio, de la muestra y su proceso de selección, los instrumentos de recogida de información, y del procedimiento de recogida de datos. Detallamos cada una de estas categorías a continuación.

1.5.1. Diseño de la investigación

Existen diseños de investigación de diferente índole, cada uno con sus características, como el modo de selección de la muestra, la cantidad de grupos que toman parte o la cantidad de puntos de recogida de datos que tendrá el estudio. Dentro de los diseños de investigación cuantitativos podemos destacar principalmente tres tipos de diseños: Los diseños descriptivos, los diseños experimentales y los diseños cuasiexperimentales.

1.5.1.1. Diseños descriptivos

Comenzando por los **diseños descriptivos**, son aquellos estudios que tratan de conocer las características de una población determinada. Este tipo de diseño se centra en conocer qué es lo que pasa, pero sin llegar a profundizar en el por qué ocurre eso. Aquellos estudios que buscan aplicar únicamente estadísticos descriptivos, como la media aritmética, la desviación estándar, la moda o conocer cómo se reparten las frecuencias entrarían en este grupo. Es el caso de un investigador que ante una muestra específica trata de conocer algunas variables específicas, como el gusto por la informática, cantidad de hijos, género, años, experiencia laboral… sin la realización de análisis más profundos, que supongan conocer cómo se relacionan estas variables o si alguna de estas variables es capaz de ser predicha por otra.

Los diseños descriptivos pueden ser transversales o longitudinales. Por una parte, los **diseños transversales** son aquellos que tratan de medir unas determinadas características de una muestra en un momento concreto únicamente. Por ejemplo, un investigador que quiere conocer el grado de desarrollo de la competencia de pensamiento crítico en sus estudiantes pasa un cuestionario que mida esta competencia en un momento determinado del curso, y ya.

Por otra parte, los **diseños longitudinales** son aquellos en los que, respetando una misma muestra, se toma mediciones de una determinada característica en dos o más de dos momentos de la historia. Por ejemplo, un investigador que quiere conocer si el pensamiento crítico de sus estudiantes ha variado a lo largo del curso, pasa el cuestionario al inicio del curso, posteriormente a mediados del curso, y finalmente al finalizar el curso. Hay estudios longitudinales de incluso años, donde los mismos sujetos han sido estudiados para conocer la evolución de una determinada característica.

1.5.1.2. Diseños experimentales

Siguiendo con los **estudios experimentales**, estos son tipos de investigaciones en las que el investigador tiene un control total de todas las variables. Suelen tratarse de investigaciones más científicas, más de laboratorio, en las que se desarrolla el experimento. En el ámbito de las ciencias sociales resulta casi imposible tener un control total de todas las variables que rodean a un sujeto, por lo que este tipo de estudios resultan casi imposible de realizar. Este tipo de diseño tiene la ventaja de que, al tener un control total las variables se pueden asegurar con mayor precisión los resultados obtenidos a causa de un tratamiento. Es decir, existe mayor validez y fiabilidad de que una determinada intervención permita predecir cuál será el resultado incluso antes de realizar el experimento. Algún ejemplo de este tipo de estudio puede ser aquellos que se realizan en un laboratorio de química, en el que mezclando determinados reactivos se pretende conseguir un determinado producto.

1.5.1.3. Diseños cuasiexperimentales

Finalmente, los **estudios cuasiexperimentales**, son los diseños por excelencia en el ámbito de las ciencias sociales. Este tipo de estudios son muy similares a los estudios experimentales, pero con la desventaja de que se asume de antemano que el investigador no tiene el control total de todas las variables que rodean al sujeto. Se tratan de casi (cuasi) experimentos. Mientras que en los estudios experimentales, por ejemplo, se puede predecir cómo actuará una bacteria al aplicarle más o menos temperatura, en el caso de las ciencias sociales, podremos intuir en mayor o menor grado como actuará una determinada persona, pero no podremos confirmar tan rotundamente que esa actuación se deba completamente a nuestra intervención, pues las personas, mientras desarrollaban el estudio se han visto en interacción con otros medios, como el entorno cercano, la cultura en la que vive, etc. En esta línea, los resultados se podrían haber "contaminado" del momento y el lugar en el que vive esa persona. Dentro de los estudios cuasiexperimentales diferenciamos tres tipos de investigaciones: Investigaciones pre-post con un único grupo; investigaciones pre-post con grupo control; e investigaciones post.

Las **investigaciones pre-post con un único grupo** son un tipo de diseño en el que investigador mide una determinada variable al inicio de una intervención; posteriormente desarrolla o pone en marcha una intervención mediante la cual se quiere conocer su "efectividad", para finalmente, volver a medir esa determinada variable y comparar los resultados del inicio y los del final. Este tipo de diseño tiene la ventaja de que nos permite conocer si la intervención, en el grupo aplicado, ha sido efectiva. No obstante, tenemos que ser conscientes de que esos resultados solo

han sido efectivos para un determinado grupo, y que por lo tanto, esa intervención aplicada a otro grupo diferente tal vez no obtendría los mismos resultados.

Esta limitación la solventa en gran parte los estudios con **diseño pre-post con grupo control.** En este tipo de diseños se nos presentan dos grupos: uno experimental, que va a someterse a cabo a un tratamiento específico, y uno control, que no se someterá al tratamiento del grupo experimental. El disponer de dos grupos nos va a permitir no solo saber si a lo largo del tiempo los valores de la variable que queramos medir han fluctuado, sino también saber si nuestra intervención con el grupo experimental ha sido efectiva, al poder contrastar el progreso con el del grupo control. Al igual que en cualquier estudio pre-post, en este caso también, se toman mediciones de los sujetos, tanto antes como después de la intervención.

Finalmente, los **estudios solo post**, tratan de medir una determinada variable al finalizar una intervención específica. Este tipo de diseño es posiblemente, el que más limitaciones presente, pues no nos permitirá asumir de manera más o menos clara como en el caso anterior, la causalidad de los resultados. Es decir, en este caso, podremos obtener un valor final ante una determinada variable, pero desconoceremos de dónde partían los sujetos y si esos resultados se deben exclusivamente a la intervención.

1.5.2. La muestra y el proceso de selección muestral

Para comprender el concepto de **muestra** es necesario que definamos primero qué es una **población**.

Una población, también llamada universo, es el conjunto de todos los casos que concuerdan con una serie de especificaciones o características. Cuando cogemos una parte de esta población no hemos hecho otra cosa que seleccionar una muestra. Por ejemplo, si estamos estudiando la motivación del alumnado de educación primaria de España, en general, la población serán todos los estudiantes de educación primaria de España; mientras que la muestra quedará conformada por una pequeña parte de toda esa población.

Volviendo a la parte cuantitativa, en el apartado respectivo a la muestra, en los estudios cuantitativos se presenta principalmente una serie de datos que nos permiten conocer mejor qué muestra existe, qué características tiene y cómo se formó.

Será necesario aportar la cantidad total de casos por las que está consolidada la muestra, generalmente, acompañado de un estadístico de utilidad, como la media

de edad de la muestra, y la desviación típica de este valor. Esta información no suele ir aislada y se suele proporcionar datos aún más concretos a través de variables que puedan resultar importantes en ese estudio. Algunas de las variables que pueden usarse en este momento pueden ser: Reparto de la muestra por género, reparto de la muestra por años de experiencia, reparto de la muestra por área del conocimiento, reparto de la muestra por zona geográfica…

En un segundo momento, será necesario aportar información sobre qué tipo de muestra es y mediante qué proceso se consolidó la muestra. En el proceso de consolidar una muestra existen diferentes maneras de crearla. Dependiendo del modo de elegir a los sujetos que formarán parte de nuestro estudio hablaremos de **diseños muestrales probabilísticos** o **diseños muestrales no probabilísticos**. En aquellos casos en los que cada sujeto tenga una determinada posibilidad de ser elegido hablaremos de diseños muestrales probabilísticos y, por el contrario, en aquellos casos, en los que investigador, elija por conveniencia, cercanía o interés una determinada muestra, hablaremos de diseños muestrales no probabilísticos.

Dentro de los diseños muestrales probabilísticos, en los que cada sujeto tiene una determinada probabilidad de ser seleccionado, nos encontramos con diferentes maneras de seleccionar a la muestra. Entre los más usados, nos encontramos con los diseños de muestreo aleatorio simple, muestreo estratificado, muestreo por conglomerados y muestreo sistemático.

1.5.2.1. El muestreo aleatorio simple

Es la manera más sencilla de formar una muestra que existe. En este modo, dentro de una población, vamos eligiendo al azar uno a uno a cada participante para nuestra muestra. En aquellos casos en los que el sujeto vuelva a formar parte del proceso de selección otorgándole la posibilidad de volver a ser elegido, diremos que se trata de un **muestreo con reemplazamiento**. Por el contrario, en aquellos casos en los que el sujeto, una vez elegido no puede volver a formar parte del proceso de selección, diremos que se trata de un **muestreo sin reemplazamiento**.

Para calcular la probabilidad de ser elegido dividiremos el número total de muestra que queremos formar, entre el número total del que está formada la población. Si por ejemplo, en una población de 5000 sujetos, queremos formar una muestra de 100 sujetos elegidos de manera aleatoria simple, la probabilidad de que salga una determinada persona será de 100/5000 = 0.02 (2% pasado a forma porcentual).

1.5.2.2. El muestreo estratificado

En este modo de selección, el investigador decide de antemano una variable específica para crear diferentes segmentos, llamados estratos. La idea es dividir la muestra total en estratos homogéneos, que posean una determinada característica, para después combinar esta información. Por ejemplo, el investigador puede crear un estrato llamado género. Por una parte tomaría información solo de muestra masculina, y por otra parte, tomaría información solo de muestra femenina, para finalmente, juntar ambos grupos.

En este tipo de muestreo, tenemos que decidir qué porcentaje de muestra queremos para cada estrato (lo que se llama **Afijación**). Dependiendo de qué nos resulte más interesante podremos repartir de manera igualitaria en total de la muestra entre el número de estratos. Por ejemplo, si estamos llevando a cabo un estudio a nivel español, tal vez interese que ambos estratos sean uniformes, 50% del total de la muestra para el género masculino, y el 50% restante para el género femenino. No obstante, si estamos llevando a cabo un estudio para conocer la actitud del alumnado sobre la universidad, tal vez nos interesa que la afijación sea proporcional. Sabiendo que existe un mayor número de chicas que de chicos en la universidad, tal vez, podremos realizar un reparto de 60% de la muestra para chicas y 40% de la muestra para chicos, por ejemplo.

1.5.2.3 Muestreo por conglomerados

La muestra se va eligiendo a través de grupos creados de manera natural. Estos grupos son conocidos como **conglomerados** y son seleccionados de manera aleatoria. Por ejemplo, podemos considerar un conglomerado a las casas que componen una manzana. En este caso, podríamos seleccionar al azar los conglomerados (todas las casas de una determinada manzana). Una variación de este método de selección es el **muestreo por conglomerados en dos etapas**, en el que primero, el investigador elige una muestra aleatoria entre los conglomerados disponibles y posteriormente, una muestra aleatoria de los elementos dentro de cada conglomerado. Siguiendo con el ejemplo anterior, en este caso, elegiríamos al azar una determinada manzana, y posteriormente elegiríamos al azar una determinada casa; a diferencia de si realizásemos únicamente un muestreo por conglomerado donde tendríamos que elegir únicamente al azar todas las casas de una determinada manzana.

1.5.2.4. Muestreo sistemático

En este modo se ordena al azar a todos los posibles sujetos y se va seleccionada cada X casos al siguiente sujeto. Por ejemplo, en una población de 500 casos, podemos

formar una muestra de 100 sujetos eligiendo a cada sujeto cada 5 casos. Para ello, podemos crear una lista en el que se ordenen al azar, o según algún criterio establecido (como orden alfabético) a todos los sujetos y cada 5 casos iremos introduciendo un sujeto más a la muestra final.

1.5.2.5. Representatividad muestral

Otro aspecto que comentar acerca de la muestra es su **representatividad**. La representatividad es un concepto muy a tener en cuenta cuando queremos que las conclusiones de nuestro estudio sean más concluyentes y **extrapolables**. Que una muestra sea o no representativa de una población dependerá de 4 variables diferentes que definimos a continuación:

El tamaño total de la población: Una población mayor supondrá mayor muestra.

Heterogeneidad: Es la diversidad que existe en la población. Generalmente, cuando se desconoce este parámetro se asume que este valor es del 50%.

Margen de error: Asumir márgenes de errores más pequeños van a requerir conocer más casos de la población y por lo tanto, aumentar la muestra. Generalmente, se suele asumir un margen de error del 5%. Es decir, las afirmaciones que hagamos de nuestro estudio contarán con una probabilidad de que sean falsas en el 5% de los casos.

Nivel de confianza: Es el valor contrario al margen de error. En este caso, cuanto mayor sea nuestro nivel de confianza, mayor tendrá que ser la muestra. Generalmente, se suele asumir un nivel de confianza del 95%. Es decir, se suele decir que las conclusiones que saquemos serán ciertas para el 95% de los casos. En el caso de la fórmula matemática empleada el nivel de confianza viene expresado por un valor Z, de modo que, por ejemplo, un nivel de confianza del 88% equivale a Z = 1.56; un nivel de confianza del 98% a Z = 2.32, etc.

Para calcular el tamaño de la muestra se unen estas cuatro variables en una fórmula matemática. En internet existen miles de páginas que nos permitirán calcular este valor jugando con estas variables.

Muy en contra de lo que puede pensar mucha gente, el objetivo de las muestras no es conseguir miles y miles de casos, pues los resultados, con muestras mucho más pequeñas, aunque representativas, son muy probables que obtengan unos resultados muy similares a muestras extremadamente grandes. Es por ese motivo que es importante que nuestras muestras sean representativas, pero tampoco necesario que realicemos estudios con 100.000 casos.

1.5.3. Instrumentos

En este apartado dentro de la metodología se debe aportar información sobre cuáles fueron los materiales empleados para realizar las mediciones oportunas. En estudios de ciencias sociales, estos instrumentos suelen ser instrumentos documentales (escalas, cuestionarios, test…), por lo que no bastará únicamente con indicar el nombre del instrumento, sino que además será necesario aportar información más específica.

Aunque este apartado es más libre y depende en gran medida de cómo el investigador planifique sus estudios, en muchos estudios cuantitativos, en el epígrafe respectivo a los instrumentos, además del nombre del instrumento, se suele aportar la siguiente información:

Número de ítems del instrumento.

Modo de medición del instrumento. Comúnmente los instrumentos en ciencias sociales suelen emplear la escala Likert de 5 puntos (Totalmente en desacuerdo, desacuerdo, neutro, de acuerdo y totalmente de acuerdo), aunque puede que un determinado instrumento emplee otro tipo de escala, como la escala de Thurstone, el escalograma de Guttman, o el diferencial semántico de Osgood.

Nombre de cada dimensión, en caso de tratarse de escalas multidimensionales.

Fiabilidad total de la escala y fiabilidad de cada dimensión, generalmente, proporcionando los valores de Alfa de Cronbach.

1.5.4. Procedimiento

En este último epígrafe dentro del diseño de cualquier estudio cuantitativo se debe clarificar cuál ha sido el proceso completo de recolección de datos. Este proceso suele contemplarse desde el momento en que comienza a decidirse sobre los cuestionarios, pasando por el modo de cómo y cuándo se recogió la información y finalizando en cómo y a través de qué programa informático se analizaron todos los datos.

En aquellos estudios que hayan llevado a cabo una intervención, como los estudios pre-post, en este punto se debe profundizar más en aras de describir especialmente cómo fue ese momento.

1.6. Resultados

Seguidamente, tras definir cuál va a ser la metodología que empleemos, el siguiente paso es conocer, en base a nuestros objetivos cuáles han sido los resultados.

A lo largo del presente trabajo, el objetivo principal de la mayoría de los capítulos no será otro que acercar al lector diferentes tipos de análisis estadísticos que pueda llevar a cabo en esta sección de sus trabajos académicos.

1.7. Discusión

Finalmente, una vez conocidos cuáles han sido los resultados de nuestro estudio cuantitativo, el último paso será interpretar estos datos.

En este momento, se suele realizar un recorrido rápido sobre cuáles eran los objetivos de la investigación y se van interpretando los resultados o valores obtenidos.

De igual modo, en este punto no sirve únicamente con realizar una interpretación de nuestros datos, sino que también será necesario que el investigador, tras el trabajo teórico realizado en pasos anteriores, sea capaz de contrastar sus datos y encontrar similitudes y diferencias con los resultados de otros estudios. Esto dará consistencia y solidez al trabajo, al relacionar el trabajo de uno mismo con el de otros investigadores.

Seguidamente, en este epígrafe se escriben también unas líneas sobre las posibles **implicaciones** de los resultados. Es decir, es necesario recalcar los resultados obtenidos para qué sirven o qué utilidad teórica y/o práctica tienen.

La discusión finaliza recordando cuáles han sido las **limitaciones más significativas que tiene el estudio (por ejemplo, muestra muy pequeña, instrumento no validado, teoría poco sólida…), así como señalando** posibles temáticas de investigación para estudios futuros; esto es, partiendo del estudio presentado, lo que se conoce como **prospectiva.**

1.8. Difusión de los resultados

Tras haber acabado de realizar nuestra investigación cuantitativa será necesario reflexionar sobre cómo vamos a dar a conocer a la comunidad científica las aportaciones de nuestro trabajo.

En el caso de trabajos académicos, este paso suele darlo la universidad, añadiendo los TFG, TFM y tesis doctorales en sus repositorios institucionales. De igual modo,

es común que la propia universidad sea quien se encargue de publicar y difundir información sobre las defensas de tesis doctorales.

En el caso de trabajos de índole más científicos, la difusión suele realizarse a través de diversos medios. Destacan principalmente los medios de divulgación y los medios científicos y académicos. Se detallan a continuación.

1.8.1. Medios de divulgación

Este tipo de medios de difusión, en su mayoría, carecen del mismo proceso seguido por los medios científicos y académicos para la publicación de trabajos. En los medios de divulgación, lo que interesa es más el llegar a la máxima audiencia posible, sin profundizar tanto en el proceso de planificación y elaboración del artículo. Interesan las conclusiones novedosas.

Entre los medios de divulgación más comunes nos encontramos con los periódicos, la televisión, las revistas de divulgación, las redes sociales, los foros de opinión, etc.

1.8.2. Medios científicos y académicos del área

Dentro del mundo universitario, estos medios suelen ser los más comunes. Cuando hablamos de medios científicos y académicos hacemos referencia a la difusión de los resultados obtenidos en un estudio a través de cualquier medio que siga los procedimientos propios de cualquier publicación científica. Los medios científicos y académicos más comunes en los que suelen publicar los trabajos los investigadores son a través de artículos en revistas científicas, a través de libros o capítulos de libros y a través de congresos o simposios científicos.

1.8.2.1. Las revistas de impacto

A día de hoy, gran parte de los investigadores suelen optar por publicar sus trabajos en revistas científicas, especialmente, en aquellas revistas conocidas como revistas de impacto. Se espera que una revista de impacto sea aquella que recoge los estudios más significativos del ámbito, que ayuden a avanzar en el conocimiento. Publicar en revistas de impacto implica la creación de trabajos originales con metodologías rigurosas.

Las dos principales bases de datos, al menos en España, que recoge las principales revistas de impacto mundiales para cualquier área del conocimiento son la **Web of Science**, propiedad de Clarivate Analytics, y **Scopus**, propiedad de la empresa Elsevier. Dentro de estas bases de datos se encuentran las revistas más prestigiosas de cada área del conocimiento divididas en **cuartiles** (Q1; Q2; Q3 y Q4). El impacto

de la revista es mayor mientras menor sea el cuartil, por lo que publicar en una revista situada en el cuartil 1 (Q1), implica un mayor prestigio de revista que una situada en el cuartil 4 (Q4). Cabe señalar, que simplemente el hecho de conseguir publicar en estas bases de datos, independientemente del cuartil, como investigador ya es un paso significativo. Existen muchas otras revistas, que no están indexadas (recogidas) en ninguna de estas dos bases de datos y por lo tanto, su impacto podrá ser significativamente inferior.

Aparte de la Web of Science y Scopus, existen otras muchas bases de datos, de prestigios muy variados, como **Dialnet**, **FECYT**, **SJR**, **Latindex**...

1.8.2.2. Proceso de evaluación de trabajos

El proceso de publicación científico no es directo, como más o menos podría serlo quien escribe en un blog, o en una página web. En este proceso se suele respetar, mayoritariamente una **revisión por pares,** es decir, un proceso en el que desde el anonimato, dos revisores comprueban la calidad del trabajo del autor. Las decisiones de estos pueden ir desde el rechazo inicial, pasando por la aceptación con cambios, hasta la aceptación inicial.

Es importante pensar y preparar un buen trabajo previamente a ser enviado, pues el periodo de revisión para la mayoría de las revistas oscila entre los 3 meses y el año de espera incluso.

La persona encargada de mediar y realizar todos los trámites, como enviar el artículo a los revisores, enviar el artículo para maquetar y/o incluso enviar el artículo para mejorar el lenguaje empleado (en el caso de artículos en inglés), es el **editor.**

1.8.2.3. Causas más comunes de rechazo de trabajos

La editorial Springer, una de las más conocidas a nivel mundial por estar en posesión de muchas de las revistas más prestigiosas de todos los campos del conocimiento, mencionan los motivos más comunes por los que se rechazan los artículos científicos enviados. Los dividen en razones técnicas y en razones editoriales.

Por una parte, las razones técnicas son aquellas razones en las que el trabajo no está mal, pero requiere de más trabajo para poder publicarse. Principalmente destacan:

- Datos incompletos, como un tamaño de muestra demasiado pequeño, o controles no existentes o deficientes.
- Análisis deficiente como el uso de pruebas estadísticas inadecuadas o la falta de estadísticas por completo.

- Uso de metodología inapropiada para confirmar la hipótesis del estudio o de una metodología antigua, que ha sido superada por métodos más nuevos y poderosos que proporcionan resultados más sólidos;
- Motivo de investigación débil donde la hipótesis no está clara o tampoco es científicamente válida, o los datos no responden a la pregunta planteada.
- Conclusiones imprecisas sobre suposiciones que no son apoyadas por los datos proporcionados.

Por otra parte, las razones editoriales hacen referencia a motivos ajenos al propio artículo científico, más relacionado con la falta de adecuación a las pautas dadas por cada revista para estructurar cada artículo. Las razones principales son:

- Fuera del alcance de la revista: La temática del artículo no tiene nada que ver con los artículos que publica esa revista.
- No hay suficiente avance o impacto para la revista.
- Se ha ignorado la ética en la investigación, como el consentimiento de los pacientes o la aprobación de un comité de ética para la investigación con animales.
- Falta de una estructura adecuada o el hecho de no cumplir los requisitos de formato de la revista.
- Falta de los detalles necesarios para que los lectores comprendan y repitan el análisis y los experimentos de los autores.
- falta de referencias actualizadas o referencias que contienen una alta proporción de autocitas.
- Tiene una calidad lingüística deficiente que no puede ser entendida por los lectores.
- Dificultad para seguir la lógica o datos mal presentados; y
- Violación de la ética de la publicación: No escribir a autores que han tomado parte significativamente del trabajo, no mencionar si se tienen intereses ocultos, no mencionar si el trabajo es subvencionado, etc.

CAPÍTULO II: EL CONTRASTE DE HIPÓTESIS Y LA DECISIÓN ESTADÍSTICA

2.1. Introducción

Siempre que realicemos cualquier tipo de análisis estadístico que nos permita conocer si el fenómeno que estamos estudiando ocurre en la muestra o no, como, por ejemplo, ver si existen diferencias entre chicos y chicas ante la ansiedad de los exámenes o ver si existen diferencias de salario entre quienes tienen estudios superiores, secundarios y primarios, vamos a tener que tomar una decisión que nos ayude a quedarnos con una solución: Sí hay diferencias o no hay diferencias.

La **decisión estadística** que tomemos es fruto de un proceso de cuatro pasos llamado **prueba de hipótesis** que explicaremos a continuación. Estos cuatro pasos que seguiremos son los siguientes:

1) Establecer la hipótesis nula y la hipótesis alterna.

2) Seleccionar el nivel de significación.

3) Seleccionar la prueba estadística.

4) Procesar la información y tomar una decisión.

2.2. Establecer la hipótesis nula y la hipótesis alterna

Antes de nada, una hipótesis es una idea o suposición que tiene el propio investigador sobre un fenómeno. Cuando realicemos una prueba de hipótesis, una hipótesis inicial va a dar dos posibles soluciones, de las cuales una va a resultar favorable y otra negativa.

Estas dos posibles soluciones se llaman **Hipótesis Nula** o H_0 e **Hipótesis del investigador**, hipótesis alterna o H1. La hipótesis Nula es siempre la que afirma que no hay diferencias entre los grupos que hemos analizado mientras que la hipótesis alterna es la que afirma que sí hay diferencias entre los grupos que hemos analizado.

En todos los casos comenzaremos eligiendo la hipótesis nula, es decir, afirmando que no existen diferencias entre los grupos, y posteriormente, lo que se hace es ver si la hipótesis nula es cierta y por lo tanto nos quedaríamos con ella, o, por el contrario, si es falsa, rechazarla y quedarnos con la hipótesis alterna.

Es importante recalcar, de igual modo, que cuando hablamos de que hay diferencias, lo que queremos decir es que esas diferencias son significativas estadísticamente hablando. Es decir, si queremos ver si hay diferencias entre chicos y chicas en un examen de matemáticas cuya calificación va del 0 a 10, y vemos que las chicas han tenido 0,03 puntos más de nota, pues sí que hay diferencias en las notas, pero probablemente no sean unas diferencias estadísticamente significativas como para afirmar que las chicas sacan mejores notas que los chicos.

Veamos todo esto con el ejemplo de antes:

Supongamos que quieres saber si existen diferencias de sueldo entre una muestra de hombres y mujeres que has consolidado. Esta sería la hipótesis, pues es la idea cuya respuesta quieres conocer.

Esta hipótesis, como te habrás podido dar cuenta, tiene dos posibles soluciones, que son las siguientes:

H_0: No existen diferencias significativas de sueldo entre hombres y mujeres.

H_1: Sí existen diferencias significativas de sueldo entre hombres y mujeres.

Una vez delimitadas las dos hipótesis pasamos al siguiente paso, establecer un nivel de significación.

2.3. Seleccionar el nivel de significación

Veníamos diciendo que siempre comenzaríamos eligiendo la hipótesis nula y que dependiendo de los resultados nos quedaríamos con ella o por el contrario la rechazaríamos y nos quedaríamos con la alterna.

En este punto se nos plantea la cuestión de ¿Y en qué momento sabemos que ya tenemos que rechazar una para quedarnos con la otra? ¿Dónde ponemos el límite?

El sistema más común para establecer el límite matemático entre una hipótesis y la otra es la del **nivel de significación**.

Este nivel de significación está muy relacionado con el porcentaje de confundirnos que estamos dispuesto a asumir y el nivel de confianza que tenemos en la afirmación escogida. Varía entre 0 y 100% y uno es complementario del otro, es decir, si trabajamos con un margen de confianza (certeza de afirmar que la opción elegida sea la correcta) del 90% significa que tenemos un 10% de margen de error (certeza de afirmar que la opción elegida sea falsa en el 10% de los casos).

Lo más común en investigaciones en ciencias sociales es establecer el nivel de significación en el 5% (0.05); es decir, en caso de aceptar la hipótesis alterna vamos a afirmar que va a ser cierta con un margen de error del 5%, lo que supone un nivel de confianza del 95%.

Estos porcentajes, se emplean siempre en forma decimal de modo que el 100% equivaldría a 1.00 y el 0% equivaldría a 0.000. A modo ilustrativo se presenta la Figura 1.

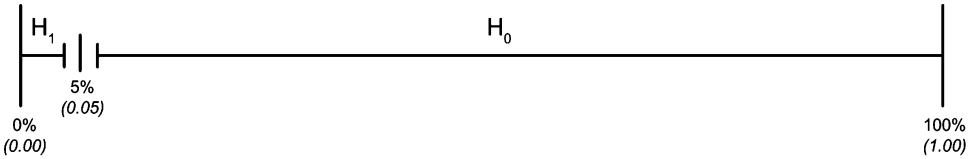


Figura 1. Límite entre Hipótesis y Nivel de significación.
Fuente: elaboración propia.

Lo ideal sería asumir el menor margen de error posible, pues esto nos aseguraría que los resultados serían casi certeros en su totalidad. No obstante, elegir márgenes de error muy bajos o muy grandes puede suponer que asumamos conclusiones que no son ciertas. Veamos esto con un ejemplo.

Supongamos que eres juez y tienes delante de ti a una persona que ha sido culpado de asesinato. Las opciones que existen entre la realidad y tu decisión son 4:

1) Que el culpable sea realmente inocente y le dejes en libertad, lo cual supondría una decisión correcta.

2) Que el culpable sea realmente inocente y lo condenes, lo cual supondría un error (Has asumido un margen de error tan pequeño que has condenado a un inocente).

3) Que el culpable sea realmente culpable y lo condenes, lo cual supondría una decisión correcta.

4) Que el culpable sea realmente culpable y lo dejes en libertad, lo cual supondría un error (Has asumido un margen error tan grande que has dejado en libertad a un culpable).

En la opción 2, por ejemplo, tal vez tienes 80 pruebas totalmente contundentes que te hacen pensar que la persona es inocente, pero aún tienes 1 prueba que no ha quedado demostrada, a pesar de ser una prueba que no tiene demasiada importancia. Ante esta situación, tú decides condenarle. En este caso es probable

que hayas asumido un margen de error tan pequeño en tu decisión (No quieres confundirte por nada del mundo), que has acabado condenando a un inocente.

Por el contrario, en la opción 4 pasa al revés. Tienes pruebas sólidas para condenar a la persona de asesinato, pero como aún tienes un par de pruebas que no están muy claras eres menos crítico en tu decisión (No quieres que entre en prisión alguien que no es culpable al 100%) y acabas dejando en libertad a un culpable.

De este modo, establece un margen de error lo más adecuado posible para tu investigación, y ante la duda toma 5%, o lo que es lo mismo en forma decimal 0.005.

2.4. Selección de la prueba estadística

Existen muchísimas pruebas estadísticas, cada una pensada para resolver una duda diferente. Tras fijar el nivel de significación es cuando debemos escoger qué prueba vamos a emplear para conseguir nuestro objetivo.

Todas las pruebas que existen se clasifican en dos grandes grupos, llamados **pruebas paramétricas** y **pruebas no paramétricas**.

Las **pruebas paramétricas** son pruebas mucho más potentes y robustas, pero cuentan con el problema de que requiere de que se cumplan ciertas condiciones. Estas condiciones suelen ser que la muestra siga una **distribución normal** y tenga homogeneidad de varianzas (**homocedasticidad**). En caso contrario se debería emplear las **pruebas no paramétricas.**

2.4.1. La distribución normal

La distribución normal, también conocida como distribución de Gauss, hace referencia a cómo se distribuyen los datos. Según este tipo de distribución, los datos formarían una campana perfecta, fenómeno que nos permitiría modelizar lo que estemos estudiando, tal y como vemos en la Figura 2.

En muchas características de la sociedad esta distribución se cumple. Por ejemplo, si nos fijamos en la altura, hay muy pocas personas que midan por debajo de 1.50 metros, hay algunas más personas entre el 1.50 y el 1.60 metros, pero la mayoría de las personas entran en el grupo de altura comprendida entre el 1.60 y el 1.80 metros. Dentro del intervalo 1.80 metros hasta el 1.90 entran algunas personas, y por encima del 1.90 hay muy pocas personas.

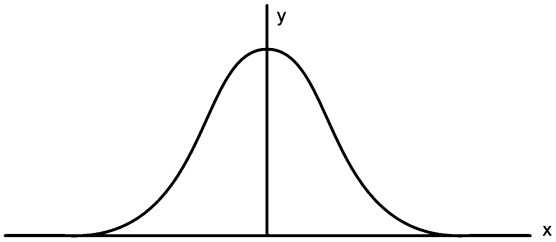


Figura 2. Ejemplo de distribución normal.
Fuente: elaboración propia.

En la Figura 2, acorde al ejemplo que hemos visto, en el eje X irían las diferentes alturas de las personas, y en el eje Y iría el recuente de personas que entran en cada categoría.

La manera que emplearemos para conocer si nuestra muestra sigue con la distribución normal es aplicando una prueba estadística. Deberemos elegir entre aplicar la prueba de **Kolmogorov-Smirnov** o la prueba de **Saphiro-Wilk**. ¿Cuándo aplicar cada una?

En caso de que nuestra muestra sea muy pequeña (< 50 casos) haremos uso de Saphiro-Wilk. Por el contrario, si la muestra es superior a 50 casos haremos uso de Kolmogorov-Smirnov.

En SPSS Statistics las pruebas de normalidad las encontraremos en la ruta Analizar > Explorar. Posteriormente introducimos las variables que queramos estudiar su normalidad y elegimos Gráficos. Una vez dentro marcamos la casilla Gráficos de Normalidad con Pruebas y obtendremos una Tabla como la Tabla 1.

Tabla 1. Ejemplo de prueba de normalidad.

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
Variable	.134	50	.026	.947	50	.025

Fuente: elaboración propia.

Llegados a este punto tendremos que volver a recordar las hipótesis que en este caso serían:

H_0: No hay diferencias significativas en la distribución. La muestra sigue la distribución normal.

H_1: Sí hay diferencias significativas en la distribución. La muestra no sigue la distribución normal.

Como veníamos diciendo el nivel de significación que usaremos será de 0.05 (margen de error del 5%). En este caso, mi muestra estaba formada por 70 estudiantes, por lo que me fijaré en las columnas de Kolmogorov-Smirnov. Más concretamente, lo único que nos interesa en este caso es la columna de *Sig.* El valor dio 0.026, que al ser un valor inferior a 0.05 que era nuestro límite, tenemos que rechazar H0 y quedarnos con H_1 afirmando que la distribución de nuestra muestra no sigue la normal.

2.4.2. La homocedasticidad

El otro criterio que tenemos que analizar para ver si podemos usar o no pruebas paramétricas es el de estudiar su **homocedasticidad** o **heterocedasticidad**.

La homocedasticidad, también conocida como homogeneidad de la varianza, hace referencia a cómo los resultados obtenidos son constantes en un modelo. Se muestra un ejemplo en la Figura 3.

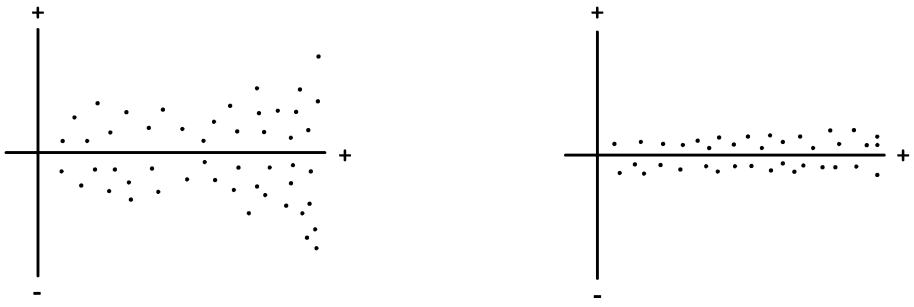


Figura 3. Heterocedasticidad (Izquierda) y Homocedasticidad (Derecha).
Fuente: elaboración propia.

Esta característica es muy importante especialmente cuando queremos realizar predicciones de una característica de la muestra, por ejemplo, a través de los modelos de regresión lineal que veremos más adelante.

Imagínate que en la Figura 3 cada punto hace referencia a un sujeto; el eje X hace referencia a la nota académica y el eje Y hace referencia a la predisposición a trabajar en equipo que tiene cada sujeto.

En el caso de la heterocedasticidad (gráfico a la izquierda) observamos que no vamos a poder predecir claramente la predisposición a trabajar en equipo, debido a que, si cogemos cualquier punto del eje X, referente a la nota académica, vemos que hay varios valores que difieren mucho entre sí en el eje Y, hecho que supone que no todos los estudiantes con una nota de 9.5 (sobresaliente), por ejemplo, tendrían la misma predisposición hacia el trabajo en equipo.

Sin embargo, en el caso de homocedasticidad (gráfico a la derecha) observamos lo contrario. Esto significaría que independientemente de la nota académica siempre existe un mismo o muy parecido grado de predisposición hacia el trabajo en equipo.

En *SPSS Statistics*, es posible realizar una prueba de homocedasticidad a través de un análisis o prueba conocida como **Test de Levene**. Este test se puede realizar a través de varias vías. Aquí hay dos diferentes:

Si estás comparando 2 grupos, por ejemplo, chicos y chicas, vamos a *Analizar > Comparar Medias > Prueba T* e introduces la variable de agrupación (en este caso sería la variable *Género)* y la variable cuya homocedasticidad quieres analizar. En este caso no está defina la prueba T que se ha elegido para no complicar más las cosas. Lo importante aquí es que una vez que conozcas las pruebas T vuelves a este paso. Las diferentes pruebas T las veremos en el capítulo VI.

Una vez seguida esta ruta obtendrás una Tabla algo más larga que la Tabla 2, pues se le han suprimido algunas columnas que no se explicarán aquí.

Tabla 2. Prueba de Levene y Prueba T.

		Prueba de Levene	
		F	**Sig.**
Variable a Analizar	**Se asumen varianzas iguales**	.652	.420
	No se asumen varianzas iguales		

Fuente: elaboración propia.

Para ver si hay homocedasticidad o no nos fijaremos en el *Sig.* de la columna de la prueba de Levene. Aquí formularemos de nuevo nuestras hipótesis.

H_0: No hay diferencias significativas en las varianzas (Homocedasticidad).

H_1: Sí hay diferencias significativas en las varianzas (Heterocedasticidad).

Si trabajamos con un límite de error de 0.05, la significancia del test de Levene es muy superior a este valor (0.420), por lo que nos tenemos que quedar con H0 y aceptar que existe homogeneidad de varianzas (se asumen varianzas iguales).

Otra ruta para realizar la prueba de Levene es cuando tenemos más de dos grupos que queremos comparar, por ejemplo, cada uno de los cursos de Educación Primaria. En este caso iremos a *Analizar > Comparar medias > ANOVA de un factor*. Igual que en el caso anterior introducimos el factor o la variable de agrupación, que en este caso es el curso, y en la sección de variables dependientes, las variables que

queramos analizar. La prueba de Levene se encuentra en el botón *Opciones > Prueba de Homogeneidad de las varianzas*. La Tabla resultante será como la Tabla 3.

Tabla 3. Prueba de Levene a través de una ANOVA.

Estadístico de Levene	gl1	gl2	Sig.
.483	2	181	.618

Fuente: elaboración propia.

En este caso, la manera de estudiar la homocedasticidad es la misma que en el caso anterior, con las mismas hipótesis. Observamos que nuestra significancia (*Sig.*) es superior a 0.05 y por lo tanto asumimos que existe homogeneidad en las varianzas.

Una vez llegados a este punto, con la información recabada ya nos permitirá conocer si podremos utilizar las pruebas paramétricas o por el contrario tendremos que emplear pruebas no paramétricas.

2.5. Procesar la información y tomar una decisión

Ya hemos elegido la prueba estadística y le hemos dicho al software estadístico que empleemos que, en base a la base de datos que le hemos proporcionado nos haga el análisis para ese objetivo.

El software nos proporcionará, entre muchos más datos, un nivel de significancia, que lo escribirá como Sig., que irá entre 0 y 1.

Si hemos supuesto que el margen de error que estamos dispuestos a asumir es de 0.05 (5%), entonces si el resultado obtenido de la significancia en el análisis va de 0.05 a 1.00 nos quedaremos con la hipótesis nula, la que afirmaba que no habría diferencias significativas en esa muestra, y si por el contrario el resultado obtenido en el análisis va de 0.00 a 0.049 rechazaremos la hipótesis nula y nos quedaremos con la hipótesis alterna, lo que supondría que sí habría diferencias significativas en la muestra.

Esta parte del procesamiento de la información y la toma de decisiones la veremos con más detalle en los próximos capítulos.

CAPÍTULO III: INTRODUCCIÓN AL PROGRAMA ESTADÍSTICO SPSS STATISTICS

3.1. Introducción

SPSS Statistics es un programa de análisis estadístico de datos formado por una interfaz gráfica considerablemente intuitiva que hace que el programa, una vez conocidos los pilares de la estadística básica, resulte fácil de manejar.

De igual modo, *SPSS Statistics* es un programa que permite a investigadores trabajar con bases de datos formadas por miles de variables, y aunque es cierto que existen procedimientos estadísticos que no son recogidos por este programa, cabe mencionar que los análisis estadísticos más comunes se incluyen, y que por lo tanto, para nuestras investigaciones a nivel principiante será un recurso imprescindible.

En el presente capítulo se tratará de realizar un recorrido por la interfaz de este programa resaltando aquellos aspectos que pudiesen ser de mayor significancia para cualquier persona que quiere iniciarse en el mundo de los análisis cuantitativos.

3.2. Sistema de ventanas de SPSS Statistics

Una vez descargado e instalado *SPSS Statistics*, como investigadores, nuestros datos los organizaremos principalmente en dos ventanas: En la ventana de variables y en la ventana de datos. Podremos cambiar de una a otra ventana las veces que queramos. Las describimos a continuación.

3.2.1. Ventana de variables

Es la ventana en la se recogen todas las variables que vamos a estudiar en nuestro estudio. Cuando tengamos en mente una variable y queramos pasarla al *SPSS Statistics*, tendremos que aportar información referente al tipo de variable que es. Como se observa en la Figura 4, cada fila representa a una variable diferente. Más concretamente, *SPSS Statistics* nos pide la siguiente información:

- Nombre: El nombre que queremos ponerle a la variable. Si es una variable que pertenece al total de algo más grande, suele ir acompañada de un número. Por ejemplo, si necesitamos 10 variables (ítems) para medir el pensamiento crítico, podremos numerar las variables como PC01, PC02, PC03 y así sucesivamente. Otras variables, como el género, que no forman parte de ningún constructo psicológico, se pueden escribir tal cual.

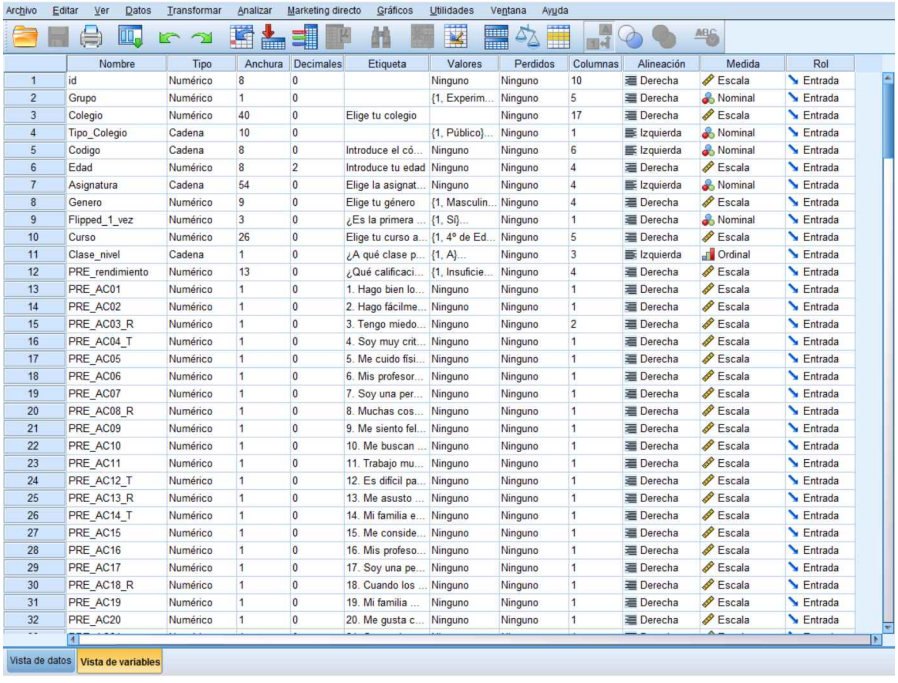
- Tipo: Indica qué tipo de dato es el que estamos introduciendo. Se recomienda, inicialmente, mantener el tipo de dato en numérico, pues otro tipo de datos puede causar problemas a la hora de desarrollar determinados análisis.

- Anchura: Hace referencia a la cantidad de dígitos utilizados para guardar la variable. Si podemos predecir que una determinada variable estará formada por 8 dígitos, pues pondremos 8. No pasa nada si le indicamos más anchura, aunque luego no la usemos.

- Decimales: Número de decimales que queremos que nos muestre sobre las variables SPSS Statistics.

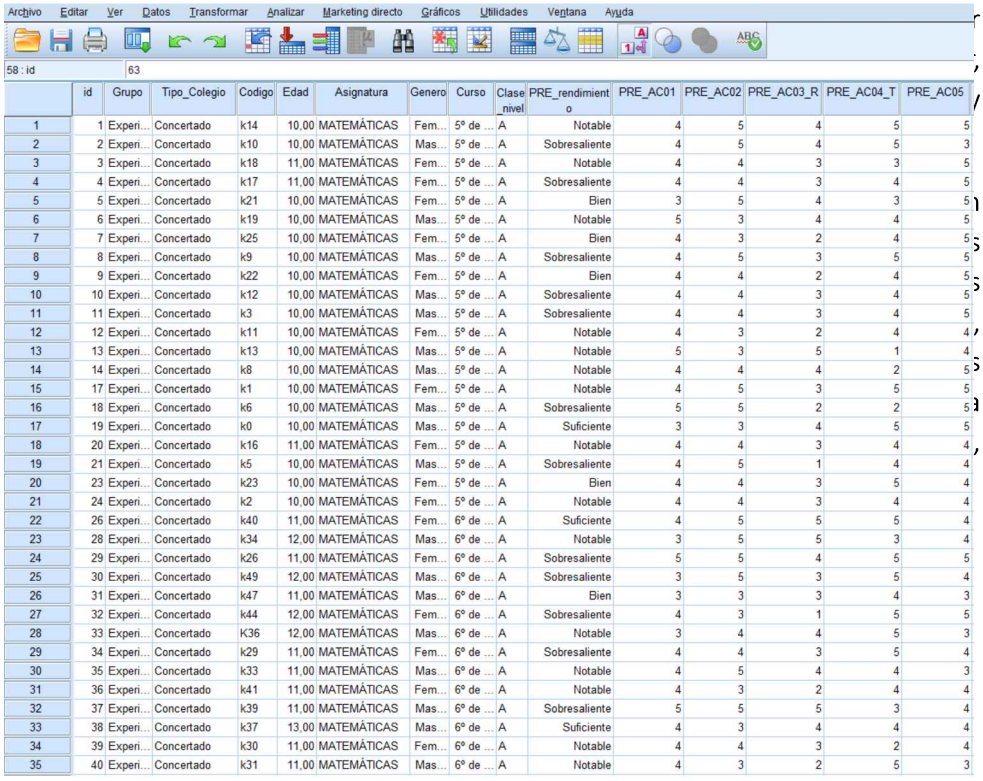
- Etiqueta: En caso de tener un nombre de variable no muy identificable, en la etiqueta podemos aportar una pequeña descripción de dicha variable.

- Valores: En caso de que tengamos en mente usar variables categóricas, en esta sección asignaremos un valor a cada categoría. Por ejemplo, codificando la variable género, podríamos usar el valor 1 para hombres y el valor 2 para mujeres. De este modo, en la ventana de datos no sería necesario escribir cada categoría, y con escribir el número que hayamos codificado sería suficiente.



	Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores	Perdidos	Columnas	Alineación	Medida	Rol
1	id	Numérico	8	0		Ninguno	Ninguno	10	Derecha	Escala	Entrada
2	Grupo	Numérico	1	0		{1, Experim...	Ninguno	5	Derecha	Nominal	Entrada
3	Colegio	Numérico	40	0	Elige tu colegio		Ninguno	17	Derecha	Escala	Entrada
4	Tipo_Colegio	Cadena	10	0		{1, Público}...	Ninguno	1	Izquierda	Nominal	Entrada
5	Codigo	Cadena	8	0	Introduce el có...	Ninguno	Ninguno	6	Izquierda	Nominal	Entrada
6	Edad	Numérico	8	2	Introduce tu edad	Ninguno	Ninguno	4	Derecha	Escala	Entrada
7	Asignatura	Cadena	54	0	Elige la asignat...	Ninguno	Ninguno	4	Izquierda	Nominal	Entrada
8	Genero	Numérico	9	0	Elige tu género	{1, Masculin...	Ninguno	4	Derecha	Escala	Entrada
9	Flipped_1_vez	Numérico	3	0	¿Es la primera ...	{1, Sí}...	Ninguno	1	Derecha	Nominal	Entrada
10	Curso	Numérico	26	0	Elige tu curso a...	{1, 4º de Ed...	Ninguno	5	Derecha	Escala	Entrada
11	Clase_nivel	Cadena	1	0	¿A qué clase p...	{1, A}...	Ninguno	3	Izquierda	Ordinal	Entrada
12	PRE_rendimiento	Numérico	13	0	¿Qué calificaci...	{1, Insuficie...	Ninguno	4	Derecha	Escala	Entrada
13	PRE_AC01	Numérico	1	0	1. Hago bien lo...	Ninguno	Ninguno	1	Derecha	Escala	Entrada
14	PRE_AC02	Numérico	1	0	2. Hago fácilme...	Ninguno	Ninguno	1	Derecha	Escala	Entrada
15	PRE_AC03_R	Numérico	1	0	3. Tengo miedo...	Ninguno	Ninguno	2	Derecha	Escala	Entrada
16	PRE_AC04_T	Numérico	1	0	4. Soy muy crit...	Ninguno	Ninguno	1	Derecha	Escala	Entrada
17	PRE_AC05	Numérico	1	0	5. Me cuido físi...	Ninguno	Ninguno	1	Derecha	Escala	Entrada
18	PRE_AC06	Numérico	1	0	6. Mis profesor...	Ninguno	Ninguno	1	Derecha	Escala	Entrada
19	PRE_AC07	Numérico	1	0	7. Soy una per...	Ninguno	Ninguno	1	Derecha	Escala	Entrada
20	PRE_AC08_R	Numérico	1	0	8. Muchas cos...	Ninguno	Ninguno	1	Derecha	Escala	Entrada
21	PRE_AC09	Numérico	1	0	9. Me siento fel...	Ninguno	Ninguno	1	Derecha	Escala	Entrada
22	PRE_AC10	Numérico	1	0	10. Me buscan ...	Ninguno	Ninguno	1	Derecha	Escala	Entrada
23	PRE_AC11	Numérico	1	0	11. Trabajo mu...	Ninguno	Ninguno	1	Derecha	Escala	Entrada
24	PRE_AC12_T	Numérico	1	0	12. Es difícil pa...	Ninguno	Ninguno	1	Derecha	Escala	Entrada
25	PRE_AC13_R	Numérico	1	0	13. Me asusto ...	Ninguno	Ninguno	1	Derecha	Escala	Entrada
26	PRE_AC14_T	Numérico	1	0	14. Mi familia e...	Ninguno	Ninguno	1	Derecha	Escala	Entrada
27	PRE_AC15	Numérico	1	0	15. Me conside...	Ninguno	Ninguno	1	Derecha	Escala	Entrada
28	PRE_AC16	Numérico	1	0	16. Mis profeso...	Ninguno	Ninguno	1	Derecha	Escala	Entrada
29	PRE_AC17	Numérico	1	0	17. Soy una pe...	Ninguno	Ninguno	1	Derecha	Escala	Entrada
30	PRE_AC18_R	Numérico	1	0	18. Cuando los ...	Ninguno	Ninguno	1	Derecha	Escala	Entrada
31	PRE_AC19	Numérico	1	0	19. Mi familia ...	Ninguno	Ninguno	1	Derecha	Escala	Entrada
32	PRE_AC20	Numérico	1	0	20. Me gusta c...	Ninguno	Ninguno	1	Derecha	Escala	Entrada

Figura 4. Ventana de Variables en Spss Statistics 23.0.
Fuente: elaboración propia.



Archivo Editar Ver Datos Transformar Analizar Marketing directo Gráficos Utilidades Ventana Ayuda

58 : id 63

	id	Grupo	Tipo_Colegio	Codigo	Edad	Asignatura	Genero	Curso	Clase_nivel	PRE_rendimiento	PRE_AC01	PRE_AC02	PRE_AC03_R	PRE_AC04_T	PRE_AC05
1	1	Experi...	Concertado	k14	10,00	MATEMÁTICAS	Fem...	5º de ...	A	Notable	4	5	4	5	5
2	2	Experi...	Concertado	k10	10,00	MATEMÁTICAS	Mas...	5º de ...	A	Sobresaliente	4	5	4	5	3
3	3	Experi...	Concertado	k18	11,00	MATEMÁTICAS	Fem...	5º de ...	A	Notable	4	4	3	3	5
4	4	Experi...	Concertado	k17	11,00	MATEMÁTICAS	Fem...	5º de ...	A	Sobresaliente	4	4	3	4	5
5	5	Experi...	Concertado	k21	10,00	MATEMÁTICAS	Fem...	5º de ...	A	Bien	3	5	4	3	5
6	6	Experi...	Concertado	k19	10,00	MATEMÁTICAS	Mas...	5º de ...	A	Notable	5	3	4	4	5
7	7	Experi...	Concertado	k25	10,00	MATEMÁTICAS	Fem...	5º de ...	A	Bien	4	3	2	4	5
8	8	Experi...	Concertado	k9	10,00	MATEMÁTICAS	Mas...	5º de ...	A	Sobresaliente	4	5	3	5	5
9	9	Experi...	Concertado	k22	10,00	MATEMÁTICAS	Fem...	5º de ...	A	Bien	4	4	2	4	5
10	10	Experi...	Concertado	k12	10,00	MATEMÁTICAS	Mas...	5º de ...	A	Sobresaliente	4	4	3	4	5
11	11	Experi...	Concertado	k3	10,00	MATEMÁTICAS	Mas...	5º de ...	A	Sobresaliente	4	4	3	4	5
12	12	Experi...	Concertado	k11	10,00	MATEMÁTICAS	Fem...	5º de ...	A	Notable	4	3	2	4	4
13	13	Experi...	Concertado	k13	10,00	MATEMÁTICAS	Mas...	5º de ...	A	Notable	5	3	5	1	4
14	14	Experi...	Concertado	k8	10,00	MATEMÁTICAS	Mas...	5º de ...	A	Notable	4	4	4	2	5
15	17	Experi...	Concertado	k1	10,00	MATEMÁTICAS	Fem...	5º de ...	A	Notable	4	5	3	5	5
16	18	Experi...	Concertado	k6	10,00	MATEMÁTICAS	Mas...	5º de ...	A	Sobresaliente	5	5	2	2	5
17	19	Experi...	Concertado	k0	10,00	MATEMÁTICAS	Mas...	5º de ...	A	Suficiente	3	3	4	5	5
18	20	Experi...	Concertado	k16	11,00	MATEMÁTICAS	Fem...	5º de ...	A	Notable	4	4	3	4	4
19	21	Experi...	Concertado	k5	10,00	MATEMÁTICAS	Mas...	5º de ...	A	Sobresaliente	4	5	1	4	4
20	23	Experi...	Concertado	k23	10,00	MATEMÁTICAS	Fem...	5º de ...	A	Bien	4	4	3	5	4
21	24	Experi...	Concertado	k2	10,00	MATEMÁTICAS	Fem...	5º de ...	A	Notable	4	4	3	4	4
22	26	Experi...	Concertado	k40	11,00	MATEMÁTICAS	Fem...	6º de ...	A	Suficiente	4	5	5	5	4
23	28	Experi...	Concertado	k34	12,00	MATEMÁTICAS	Mas...	6º de ...	A	Notable	3	5	5	3	4
24	29	Experi...	Concertado	k26	11,00	MATEMÁTICAS	Fem...	6º de ...	A	Sobresaliente	5	5	4	5	5
25	30	Experi...	Concertado	k49	12,00	MATEMÁTICAS	Mas...	6º de ...	A	Sobresaliente	3	5	3	5	4
26	31	Experi...	Concertado	k47	11,00	MATEMÁTICAS	Mas...	6º de ...	A	Bien	3	3	3	4	3
27	32	Experi...	Concertado	k44	12,00	MATEMÁTICAS	Fem...	6º de ...	A	Sobresaliente	4	3	1	5	5
28	33	Experi...	Concertado	K36	12,00	MATEMÁTICAS	Mas...	6º de ...	A	Notable	3	4	4	5	3
29	34	Experi...	Concertado	k29	11,00	MATEMÁTICAS	Fem...	6º de ...	A	Sobresaliente	4	4	3	5	4
30	35	Experi...	Concertado	k33	11,00	MATEMÁTICAS	Mas...	6º de ...	A	Notable	4	5	4	4	3
31	36	Experi...	Concertado	k41	11,00	MATEMÁTICAS	Fem...	6º de ...	A	Notable	4	3	2	4	4
32	37	Experi...	Concertado	k39	11,00	MATEMÁTICAS	Mas...	6º de ...	A	Sobresaliente	5	5	5	3	4
33	38	Experi...	Concertado	k37	13,00	MATEMÁTICAS	Mas...	6º de ...	A	Suficiente	4	3	4	4	4
34	39	Experi...	Concertado	k30	11,00	MATEMÁTICAS	Fem...	6º de ...	A	Notable	4	4	3	2	4
35	40	Experi...	Concertado	k31	11,00	MATEMÁTICAS	Mas...	6º de ...	A	Notable	4	3	2	5	3

Figura 5. Ventana de Datos en Spss Statistics 23.0.
Fuente: elaboración propia.

3.3. Barra de herramientas en SPSS Statistics

Una vez dominadas estas dos ventanas, también es importante mencionar que en *SPSS Statistics*, tenemos dos barras de herramientas, ilustradas en la Figura 6: La barra de herramientas del menú y la barra de herramientas del editor de datos.

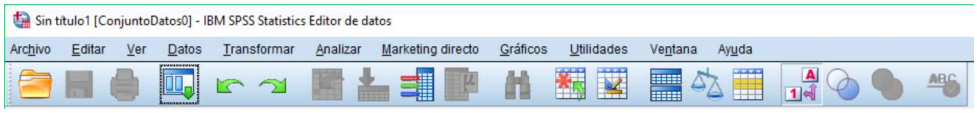


Figura 6. Barra de Herramientas en Spss Statistics 23.0.
Fuente: elaboración propia.

3.3.1. La barra de herramientas del menú

Dentro de la barra de herramientas del menú, nos encontramos con las siguientes opciones:

Archivo: Entre las opciones más significativas están la de crear una nueva base de datos, abrir una base de datos existente, o imprimir una base de datos.

Editar: Nos permite cortar, copiar o pegar información del área de la base de datos que queramos. Nos da la opción también de deshacer o rehacer la última acción realizada, así como de insertar una nueva variable o caso, dependiendo de la ventana en la que estemos.

Datos: Lo más llamativo de este apartado es la opción de "Seleccionar casos", que la usaremos en aquellos momentos en lo que queramos establecer una condición para seleccionar una parte de la muestra.

Transformar: Las opciones más significativas de este apartado son la de "**Calcular variable**", que nos permitirá calcular el valor de una variable basándonos en las operaciones matemáticas que consideremos oportunas (por ejemplo, CPtotal = CP01 + CP02 + CP03 + CP04); y la opción de "**recodificar en distintas variables**" o "**recodificar en las mismas variables**". En esta segunda opción, *SPSS Statistics* nos da la posibilidad de recodificar una variable sobrescribiendo la información de esa misma, o por el contrario, creando una nueva variable. Este tipo de acción es muy útil cuando queremos invertir los valores de un ítem (por ejemplo, si queremos invertir los datos de una variable Likert de 5 puntos, la recodificaremos de modo que en 5, tendrá que ser el 1; el 4 tendrá que ser el 2; el 3 se mantendrá igual; el 2 tendrá que ser el 4 y el 1 tendrá que ser el 5) o cuando queremos realizar agrupaciones (En una escala de 6 puntos, si queremos convertirla en 3 puntos, recodificar el 1 y el 2 para el nuevo valor 1, el 3 y el 4 para el nuevo valor 2, y el 5 y el 6 para el nuevo valor 3).

Analizar: A través de este apartado será donde realizaremos todos los análisis estadísticos que se trabajan en el presente trabajo.

Gráficos: A través de esta ventana, *SPSS Statistics* nos facilita la creación de todo tipo de gráficos cuantitativos.

3.3.2. La barra de herramientas de datos

Por otra parte, respecto a la barra de herramientas de datos, las opciones más significativas son las siguientes:

Recuperar los cuadros de diálogos recientes: Este recuadro nos permitirá acceder rápidamente a las acciones que hemos realizado recientemente.

Botones de deshacer y rehacer.

Etiquetas de valor: Nos permite ir jugando con el valor de las variables. Si está seleccionado, nos mostrará la información con texto a cada categoría para cada variable, mientras que si no está seleccionada, nos mostrará la información numérica para cada variable. Esta opción es muy útil si hemos empleado la sección de "Valores" de la ventana de variables.

CAPÍTULO IV: CARACTERÍSTICAS Y CONSIDERACIONES EN EL USO DE INSTRUMENTOS DE MEDIDA DOCUMENTAL

4.1. Introducción

El mundo que nos rodea está lleno de instrumentos de medición para casi todas y cada una de las variables que te puedas imaginar. Para medir el peso empleamos una báscula, para medir la presión atmosférica empleamos un barómetro, para medir la temperatura un termómetro, y así podríamos continuar. Todos estos instrumentos de medición son instrumentos mecánicos que nos permiten calcular diversos aspectos de la realidad de manera más o menos exacta.

Sin embargo, hay otras variables destacables por su esencia que deben ser estudiadas y para las cuales, no existen instrumentos de medición mecánicos. Una peculiaridad de estas variables es que son variables latentes, es decir, variables que no se pueden observar directamente, sino que deben ser inferidas principalmente a través de algún **instrumento de medición documental**. Algunos ejemplos de estas variables, podrían ser la ansiedad hacia cierto trabajo, la motivación hacia coger el transporte público o el autoconcepto emocional de un estudiante de ingeniería. La ansiedad, la motivación y el autoconcepto son variables intangibles que no se pueden observar a simple vista. No se pueden cuantificar tan sencillamente.

Para todas estas variables es necesario planificar, diseñar, construir y probar unos instrumentos que nos permitan determinar qué sujetos tienen dicha característica y en qué grado.

A pesar de que generalmente se empleen de manera indistinta, existen principalmente 3 tipos de instrumentos de medición documental: Los cuestionarios, los inventarios y las escalas, que procederemos a explicar en las siguientes páginas.

4.2. Los instrumentos de medición documental

El cuestionario es el primer instrumento de medición documental. Este se puede identificar como un instrumento formado por preguntas abiertas y/o cerradas cuyo último fin suele ser el de tomar las respuestas como clasificador generalmente dicotómico (dos grupos).

Un ejemplo de cuestionario es un examen. Un examen nos permite clasificar a los estudiantes en quienes han aprobado (los que han sacado más de la mitad) y en

quienes han suspendido (los que no han sacado más de la mitad) al enfrentarse a la misma prueba.

Por otra parte, los inventarios son menos empleados, pero de gran utilidad en ciertos casos en los que queremos clasificar a cada persona en uno de los grupos. A diferencia del cuestionario que solamente clasifica los casos en 2, los inventarios clasifican a los individuos en tantos grupos como se desee, sin un orden de importancia como en las escalas.

Un ejemplo de inventario podría ser un test de las inteligencias múltiples, que permite clasificar a cada persona en base a los resultados obtenidos en una de las 8 inteligencias que propone esta teoría. Se puede apreciar por lo tanto que existen más de 2 categorías y que no existe un orden, siendo las categorías nominales.

Finalmente, nos encontramos con las escalas. Las escalas son instrumentos que nos permiten conocer el grado o el nivel en el que se encuentra una persona respecto a una variable determinada. Las escalas suelen estar formadas por un conjunto de ítems, denominados **reactivos**, medidos en escala Likert, cuya suma nos da un valor que nos permite graduar el nivel en el que dicha persona se encuentra.

La **escala Likert** es una escala que mide el grado de acuerdo o desacuerdo hacia un reactivo y suele tomar por lo general entre 5 a 9 puntos. Por ejemplo, una persona que está midiendo tu autoconcepto físico, podría formular el reactivo "*Me gusta como soy físicamente*"; a lo que tú podrías dar una valoración entre totalmente de acuerdo a totalmente en desacuerdo. Juntando tu valoración a varios reactivos obtendrías un total que podría colocar como una persona con autoconcepto físico bajo, medio o alto.

4.3. La dimensionalidad

Una peculiaridad de las variables latentes es que son ciertamente complejas de medir, pues se debe saber con rigurosidad qué es lo que se está calculando.

Es por ello, que es muy común que cada variable latente esté formada a su vez por varias dimensiones (también llamados **factores**) encargadas de medir cada una de las diferentes partes de las que las variables latentes están conFiguradas.

Por ejemplo, para medir el autoconcepto no podemos entenderlo como tal, como algo general, sino que hay que especificar más; hay que dividirlo en partes más pequeñas. Es por ello, que un buen instrumento para medir el autoconcepto debería englobar diferentes facetas del mismo, como el autoconcepto físico, el autoconcepto

familiar, el autoconcepto emocional, el autoconcepto académico y el autoconcepto social, siendo la suma de todos ellos más una parte que no hemos sido capaces de explicar, la valoración que se tenga de lo que se denomina autoconcepto. Se muestra un ejemplo ilustrativo en la Figura 7.

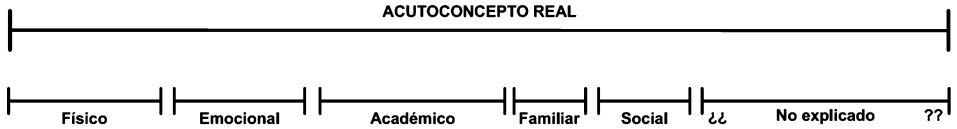


Figura 7. Relación entre una variable latente y sus dimensiones.
Fuente: elaboración propia.

De este modo, aunque el objetivo de la escala sea medir el autoconcepto, como es el caso, realmente lo que estamos midiendo son los diferentes tipos de autoconcepto, cuya suma darían el autoconcepto general.

Es de destacar que en cualquier instrumento va a resultar casi imposible explicar el 100% de la variable latente, pues siempre se nos van a "escapar" algunas dimensiones que no hemos considerado, lo que se denomina como **varianza no explicada**. Lo que sí vamos a poder ver es en qué porcentaje contribuye a la explicación de la variable latente, normalmente a través de un análisis factorial (Véase capítulo VII).

En algunos casos podría considerarse variables latentes formadas únicamente por una dimensión (**unidimensionalidad**). Entre los criterios más empleados para saber si una única dimensión podría consolidar una variable latente está el criterio de Carmines y Zeller (1979) mediante el que se dice que, si una dimensión explica más del 40% de la varianza de la variable latente, podría considerarse unidimensional. Otra pista podría ser la de ver si el primer factor explica más de 5 veces la varianza del segundo factor (Lord, 1980), en cuyo caso podría considerarse una variable latente unidimensional.

Para saber qué porcentaje de la varianza explica cada factor, será necesario llevar a cabo un análisis factorial exploratorio, procedimiento que explicaremos más adelante.

4.4. Los ítems invertidos

Cuando estemos delante de una escala habrá ocasiones en las que nos encontremos con que no todos los ítems llevan la misma dirección. Es decir, no en todos los ítems la puntuación más alta es la más favorable y no en todos los ítems la puntuación más baja es la menos favorable.

En el caso que se muestra a continuación estamos tratando de analizar la intención de abandono laboral en base a estos dos ítems improvisados:

- Estoy pensando en dejar este trabajo.
- Pretendo permanecer en este trabajo durante mucho tiempo.

Si recordamos, la escala Likert suele emplearse en una valoración ascendente, donde 1 indica un bajo nivel de acuerdo con la oración, y el 5 indica un alto nivel de acuerdo con la oración.

Si pensamos en una persona con una intención de abandono laboral alto, la primera pregunta la respondería con un 5, mostrándose totalmente de acuerdo con el hecho de estar pensando en dejar su trabajo. No obstante, a la segunda afirmación, esta misma persona no sería común que respondiese con un 5 también, pues la idea es precisamente la contraria. En este segundo caso, un 5 en el segundo ítem indicaría que está totalmente de acuerdo con el hecho de permanecer en este trabajo durante mucho tiempo. ¿Se aprecia cómo los ítems se miden en direcciones contrarias?

Lo más recomendable sería codificar todos los ítems en positivo de manera ascendiente (en dirección del constructo), donde el 1 sería la valoración más baja y poco a poco fuese aumentando el grado de positivismo. Para detectar qué ítems están invertidos un truco podría ser el de pensar en una persona con las puntuaciones más favorables a todos los ítems y preguntarse a uno mismo ¿En este ítem qué contestaría, con un 1 o con un 5?

En el caso de arriba, si seguimos este mismo truco, una persona con una intención de abandono muy alta daría un 5 al primer ítem (totalmente de acuerdo), y un 1 al segundo ítem (totalmente en desacuerdo).

Pero ¿por qué se emplean los ítems invertidos? El uso principal de este tipo de ítems es para evitar la **aquiescencia**. La aquiescencia es la tendencia de los sujetos a responder de forma afirmativa con independencia del contenido (Tomás, Sancho, Oliver, Galiana, Meléndez, 2012). En otras palabras, si todos ítems están en positivo; es muy probable que nuestro cerebro, acostumbrado a contestar con una valoración alta a todos los ítems, conteste a todas las demás cuestiones restantes en esta línea, sin leer si quiera el contenido del ítem. Podríamos catalogarlo, como un truco que usan los investigadores para contrarrestar la monotonía de la escala y mantener la mente del participante activa.

La cantidad de ítems en una dirección y en la otra por escala es un poco orientativa, aunque Likert (1932) recomienda que aproximadamente la mitad de los ítems estén en positivo y la otra mitad invertidos.

En *SPSS Statistics*, invertir las categorías de un ítem es muy sencillo y lo podemos preparar tanto *a priori*, como *a posteriori*.

Para ello, vamos a ir a la pestaña Transformar; y aquí, vamos a elegir entre *Recodificar en las mismas variables* o *Recodificar en distintas variables*.

La diferencia entre ambas es que, mientras la primera nos va a recodificar el ítem invertido en la misma variable (es decir, nos va a sustituir el ítem invertido, por el ítem positivo en el mismo lugar), para la segunda opción, se va a crear una variable nueva con las mismas propiedades y se escribirá ahí, manteniendo tanto la variable original, como la nueva. Ante la duda, lo mejor es mantener ambas, añadiéndole el sufijo "_R" (de recodificada, por ejemplo) a la segunda variable.

Una vez dentro, pasaremos al cuadro de *variable de entrada > Variable de salida* la variable que queramos recodificar. Posteriormente, presionaremos el botón *Valores antiguos y nuevos* y en la nueva ventana que se nos abra, colocaremos en la parte izquierda (relativa al valor antiguo) las categorías de la variable de origen que queramos cambiar. Entre las opciones, las tres más significativas son:

- Valor: Simplemente se elige un único valor a transformar de la variable de origen.
- Rango: Se selecciona un rango entre dos valores para transformarlo a un único valor.
- Todos los demás valores: Se selecciona para transformar los valores no seleccionados.

Tras elegir una de las opciones arriba indicada, en la parte derecha, respectiva al valor nuevo, indicaremos el valor de la recodificación. Pulsamos en el botón *Añadir* y *Continuar*.

Nos llevará de nuevo a la ventana donde lo habíamos dejado. A la derecha del todo, en la variable de salida, indicamos el nombre de la nueva variable y si lo quisiésemos su etiqueta. Finalmente, pulsamos en aceptar y nuestra variable invertida se habrá recodificado correctamente.

4.5. La validez

Según Pérez (1986), la validez es la capacidad de un instrumento para medir lo que dice medir. No obstante, el concepto de **validez** como tal es un concepto bastante relativo (Gil, 2015) y más complejo de medir que la fiabilidad. Esto es debido a que mientras que la fiabilidad es analizada experimentalmente, la validez incluye la explicación de elementos teóricos, con la complicación que esto conlleva para su aceptación o rechazo (Argibay, 2006).

La validez como característica general de un instrumento está constituida por diversos tipos de validez. Es por esto, que cuando hablamos de validez, comúnmente, debemos mencionar mínimo, a la **validez de contenido**, a la **validez de constructo** y a la **validez de criterio**, teniendo presente que existen más tipos de validez como la validez predictiva, la validez concurrente o la validez aparente, entre otras.

4.5.1. Validez de contenido

Por su parte, la **validez de contenido** es la primera fase en cualquier diseño de instrumentos y corresponde a la fase cualitativa del mismo. La validez de contenido hace referencia a si un instrumento recoge de manera adecuada todas las partes de las que puede estar formada la variable que se pretende medir.

Un ejemplo de mala validez de contenido podría ser unas oposiciones al cuerpo de maestros y maestras de educación primaria, en el que todas las preguntas están únicamente ligadas con las matemáticas y no con cada una de las diferentes materias de Educación Primaria. En este caso, quienes aprueben no habrán demostrado que sean mejores profesores de educación primaria (que sería el principal objetivo de estas oposiciones), sino que habrán demostrado ser únicamente mejores profesores de matemáticas de educación primaria. Las preguntas que se han realizado no recogen la variedad de los posibles contenidos.

Existen principalmente tres modos de conocer la validez de contenido: Mediante la **validez de respuesta**, mediante la **validez racional**, y mediante la **validez de jueces**.

4.5.1.1. Validez de respuesta

La validez de respuesta conlleva realizar entrevistas en profundidad a la población objetivo con el objetivo de conocer qué factores se deberían tener en cuenta para establecer la dimensionalidad de la variable que se quiere medir.

4.5.1.2. Validez racional

La validez racional supone realizar una revisión de la literatura (libros, artículos, informes…) de la variable que estemos analizando con el fin de asegurar la mejor representatividad de los ítems. Esta revisión documental nos permitirá sustentar el instrumento en teorías o modelos previos.

4.5.1.3. Validez de jueces

Uno de los métodos más conocidos para determinar la validez de contenido es realizar una validación por jueces o expertos. Esta validación conlleva formar un grupo de expertos o jueces que evalúen cada una de las diferentes propiedades (pertinencia, coherencia, claridad, suficiencia…) de cada ítem para realizar las modificaciones oportunas en aquellos lugares con mayor problemática en el instrumento.

En esta línea, para analizar el grado de acuerdo o desacuerdo de los jueces o expertos sobre cada ítem, se suelen emplear diversos estadísticos como el índice **Kappa de Cohen** (u otros como el coeficiente **W de Kendall)**. Este estadístico oscila entre 0 y 1, suponiendo 0 la no concordancia máxima y 1 la concordancia máxima.

Tabla 4. Interpretación de la Kappa de Cohen (McHugh, 2012).

Valor de Kappa	0 – 0.19	0.20 – 0.39	0.40 – 0.59	0.60 – 0.79	0.80 – 0.90	> 0.90
Nivel de acuerdo	Nada	Mínimo	Débil	Moderado	Fuerte	Casi perfecto
% de que sea fiable	0 – 4%	4 – 15%	15 – 35%	35 – 63%	64 – 81%	82 – 100%

Fuente: elaboración propia.

En *SPSS Statistics* para conocer el grado de acuerdo entre dos jueces es necesario acceder al estadístico Kappa, a través de *Analizar > Estadísticos Descriptivos > Tablas Cruzadas*. Ahí colocamos al juez 1 en filas, y al juez 2 en columnas; presionamos el botón *Estadísticos* y seleccionamos *Kappa.*

Es importante recordar, que este tipo de procedimientos se deben ir haciendo en Tablas de contingencia de 2x2, pues para conocer el nivel de acuerdo en Tablas de más de 2x2 existen otros procedimientos diferentes.

Pongamos el ejemplo de dos evaluadores que están analizando una clase en búsqueda de estudiantes con TDAH. Los resultados obtenidos se ilustran en la Tabla 5.

Tabla 5. Ejemplo de Tabla de contingencia entre evaluadores.

		Evaluador 1		**Total**
		Con TDAH	**Sin TDAH**	
Evaluador 2	**Con TDAH**	38	6	44
	Sin TDAH	12	36	48

	Evaluador 1		Total
	Con TDAH	Sin TDAH	
Total	50	42	92

Fuente: elaboración propia.

Tal y como se muestra en la Tabla 5, de los 92 casos totales, los evaluadores están de acuerdo en 74 (la suma de la primera diagonal). Ambos concuerdan en que 38 estudiantes tienen TDAH y que 36 estudiantes no tienen TDAH. Sin embargo, hay 18 casos (12 + 6), en los que no hay acuerdo.

Si analizamos la siguiente Tabla (Tabla 6) y nos fijamos en el estadístico más importante de la misma, el Valor Kappa (.610), podemos apreciar en ayuda de la interpretación que se hacía en la Tabla 4, que la concordancia entre ambos evaluadores es moderada, lo cual puede suponer que sea necesario analizar dichos casos y/o emplear otros métodos de diagnóstico.

Tabla 6. Ejemplo de cómo se muestra el estadístico Kappa.

	Valor	Error tip. Asint.	T	Sig.
Medida de acuerdo (Kappa)	.610	.082	5.903	.000

Fuente: elaboración propia.

4.5.2. Validez de constructo

Uno de los problemas cuando diseñamos instrumentos documentales es que aquello que queremos medir no es observable directamente. Ante esta problemática, la solución está en proponer una serie de ítems que traten de medir precisamente, aquello que mediante otros métodos no sería posible medir.

Un **constructo**, en el campo de la psicología, es una construcción teórica que se sabe que existe, pero resulta complicada de explicar y/o definir. Algunos ejemplos de constructo son la inteligencia, la personalidad, la creatividad, la ansiedad, etc.

En este hilo, la **validez de constructo** trata de confirmar que aquello que estamos preguntando y aquello que estamos midiendo van acorde.

En la validez de constructo se trata de establecer grupos de ítems que cuando se unan formen dimensiones. Entre las técnicas más empleadas para este fin se hallan el análisis factorial exploratorio, si desconocemos la teoría y queremos ver qué nos dice la estadística, y el análisis factorial confirmatorio, si conocemos el modelo teórico y queremos comprobar lo bien o mal que está construido (Pérez-Gil, Chacón y Moreno, 2000). Profundizaremos en este aspecto más adelante.

En estos análisis se trata de conocer si los ítems elegidos para la explicación de un constructo son apropiados para tal fin. Para ello se espera que existan ítems que tengan que correlacionar, así como ítems que no correlacionen.

Por ejemplo, si establecemos un modelo para estudiar el constructo de la felicidad, y decimos que la felicidad está formada por las dimensiones de satisfacción, alegría y placer por vivir, se espera que los ítems de estas tres dimensiones correlacionen positivamente entre sí, y correlacionen negativamente, por ejemplo, con otras dimensiones como la tristeza, la desesperación, la depresión, la ansiedad, etc. Positiva o negativamente, lo importante es que correlacionen. Diremos que un constructo presenta **validez convergente**, cuando aquellas dimensiones que se espera que correlacionen entre sí, efectivamente lo hagan.

Siguiendo con otro ejemplo, supongamos que estamos midiendo el constructo de la autoestima. Se puede esperar que este constructo, tenga poco que ver, por ejemplo, con la inteligencia. En esta línea, diremos que un constructo presenta **validez discriminante** o **validez divergente** cuando aquellas dimensiones que se espera que no correlacionen entre sí, efectivamente, no lo hagan.

4.5.3. Validez de criterio

La **validez de criterio** es el tipo de validez en el que comparamos nuestro propio test con otras variables ajenas al test, denominadas criterios. Generalmente, el test que empleemos como criterio suele ser el mejor test de referencia (generalmente, por haber sido el test con mejores propiedades psicométricas) para cada constructo. El autoconcepto tendrá un mejor test de referencia, la resiliencia tendrá otro mejor test de referencia, la motivación tendrá otro mejor test de referencia y así sucesivamente. A estos test de referencia con las mejores características como instrumento de medición se les denomina **Gold Standard**.

El procedimiento para comparar nuestro test con el Gold Standard se realiza, generalmente, a través de la correlación entre ambos test obteniendo un **coeficiente de validez**.

Como resulta lógico, en aquellos casos en los que se está construyendo una escala para un constructo que carece de escalas predecesoras, la validez de criterio simplemente no se contempla en la validación de la escala (Zavando, Suazo y Manterola, 2010).

4.6. La fiabilidad

El concepto de **fiabilidad** expresa el grado de precisión de la medida. Morales (2007) señala que una fiabilidad alta, los sujetos medidos con el mismo instrumento en ocasiones sucesivas hubieran quedado ordenados de manera semejante. Si baja la fiabilidad, sube el error, haciendo que los resultados variasen más de una medición a otra. En definitiva, cuando hablamos de fiabilidad nos referimos a que un instrumento permita obtener resultados similares en diferentes mediciones, demostrando que los resultados obtenidos en los diferentes momentos no han sido por motivos de azar.

La fiabilidad es una característica propia de una muestra específica, más que una característica propia de un instrumento. Al obtener unos valores de fiabilidad óptimos, nos estamos refiriendo a que los datos obtenidos son precisos en la muestra en la que se ha aplicado un determinado instrumento. No obstante, un instrumento que para una muestra fue fiable, para otra muestra puede no serlo.

Entre los métodos de calcular la fiabilidad nos encontramos con especialmente 3 maneras diferentes: El **test-retest**, los **test paralelos** y los **coeficientes de consistencia internos**, aunque de estos 3 métodos seguramente, como investigador nobel, solamente vayas a hacer uso de los coeficientes de consistencia interna.

4.6.1. Método test-retest

Respecto al test-retest, como bien dice el propio nombre, se basa en que un mismo test sea respondido en momentos diferentes por los mismos sujetos. El intervalo de tiempo que se debe dejar entre el primer momento y el segundo puede ir desde unos días hasta meses; aunque no es muy recomendable extenderlo en exceso, debido a que puede que los sujetos cambien sus maneras de pensar.

4.6.2. Método de los test paralelos

Por otra parte, el método de test o pruebas paralelas se utiliza en las situaciones en las que tenemos dos versiones de un mismo instrumento que miden la misma variable latente. Se aplica primero un test y al de un pequeño periodo de tiempo se aplica el siguiente. Posteriormente se analiza el grado de correlación entre los ítems que trataban de medir una misma dimensión y se obtiene un valor conocido como **coeficiente de equivalencia**, que se relaciona directamente con la fiabilidad obtenida.

Debido a que ambas pruebas tratan de medir la misma variable latente, en el caso de pruebas fiables, ambas pruebas deberían presentar medias y varianzas iguales en las puntuaciones. También es de destacar que ambas pruebas tienen que presentar ítems referidos a una misma variable latente, redactados de manera diferente.

No se considerará pruebas paralelas aquellos casos en los que la única diferencia entre un test y el otro sea la variación en el orden de los ítems o el orden de las alternativas.

4.6.3. Métodos de consistencia interna

4.6.3.1. Alfa de Cronbach

Dejando estos dos métodos de lado, tal vez el procedimiento destacado para conocer la fiabilidad de un instrumento es el de aplicar coeficientes de consistencia interna, para cuyo mismo fin existen dos estadísticos principales: el estadístico **KR-20** (Kuder Richardson) y el estadístico α **(alfa) de Cronbach.** La única diferente que existe entre ambos es que, mientras que KR-20 se aplica con ítems dicotómicos (En cuestionarios cuyas respuestas son solamente sí y no, por ejemplo), alfa de Cronbach se emplea para ítems continuos. La fórmula para ambos casos es exactamente la misma.

Es por ello por lo que cuando tengamos en frente un instrumento al que tengamos que calcularle su fiabilidad interna, en *SPSS Statistics* solamente tendremos que ir a *Analizar > Escala > Análisis de fiabilidad*. Pasamos las variables cuya consistencia interna queramos conocer a la parte de *elementos* y mantenemos el modelo en *Alfa*. En este punto es pertinente explicar dos nuevas ideas.

La primera es que cuando analicemos un instrumento, es una buena práctica aportar tanto el valor alfa de Cronbach general (pasando a la columna *elementos* todos los ítems del instrumento), como el valor alfa de Cronbach de cada dimensión (pasando a la columna *elementos* únicamente los ítems que forman una dimensión).

A continuación, en la Tabla 7, se muestra un ejemplo, en la misma línea del autoconcepto.

Tabla 7. Valores alfa de Cronbach generales y dimensionales del autoconcepto.

Dimensión	Alfa de Cronbach
General	.883
Autoconcepto físico	.792
Autoconcepto académico	.821
Autoconcepto familiar	.800

Dimensión	Alfa de Cronbach
Autoconcepto emocional	.623
Autoconcepto social	.921

Fuente: elaboración propia.

La segunda es que el alfa de Cronbach toma valores desde 0 a 1, siendo 0 la fiabilidad nula y 1 la fiabilidad absoluta. Según George y Mallery (2003), un valor por debajo de 0.5 resulta inaceptable, un valor entre 0.5 y 0.6 resulta pobre, un valor entre 0.6 y 0.7 resulta cuestionable, un valor entre 0.7 y 0.8 resulta aceptable, un valor entre 0.8 y 0.9 resulta bueno y un valor entre 0.9 y 1.0 resulta excelente. Como regla general, suele considerarse una dimensión o instrumentos fiables aquellos que poseen un valor alfa de Cronbach superior a .70.

Finalmente, cabe subrayar que la fiabilidad se ve influenciada por dos aspectos importantes: la cantidad de ítems y la variabilidad de las respuestas a dichos ítems.

La cantidad de ítems que engloban el instrumento completo o cada una de las dimensiones del instrumento son clave para asegurar la fiabilidad. Esto se debe a que en pocos ítems es complicado realizar una discriminación perfecta entre los sujetos; es decir, escalas de pocos ítems suelen estar relacionadas con mayor dificultad para encontrar variación en los datos y sin variación en los datos no hay fiabilidad. Por ejemplo, sería muy complicado medir el autoconcepto general de una persona solo a través de dos ítems, en los que, además, todos los sujetos han contestado a ambas preguntas con la misma valoración. Este fenómeno no nos permitiría discriminar qué persona tendría un alto autoconcepto y qué persona tendría un bajo autoconcepto. Cuando ocurre esto se dice que no hay **varianza** en los datos (Los datos no varían).

En esta línea, Comrey (1985) recomienda que se asignen mínimo 5 ítems por dimensión o factor. Es decir, un potencial instrumento para medir el autoconcepto según el modelo descrito en la Tabla 4 formado por 5 dimensiones debería tener como mínimo 25 ítems.

Lo ideal es buscar el punto intermedio óptimo en el que con la menor cantidad de ítems se obtenga la máxima discriminación posible entre los sujetos.

4.6.3.2. Método de las dos mitades

En este método se coge el instrumento que tengamos, se divide en dos partes, (generalmente números pares por una parte y números impares por otra) y se correlacionan las dos mitades. El resultado obtenido de esta correlación será considerado como un índice de fiabilidad del test.

En SPSS Statistics es posible analizar nuestro instrumento por este método a través de Analizar > Escala > Análisis de Fiabilidad; en Modelo seleccionamos Dos mitades e introducimos a la columna de elementos todas las variables del instrumento. El resultado será una Tabla similar a la que nos encontramos en la Tabla 8.

Tabla 8. Ejemplo de Coeficiente de Spearman-Brown.

Alfa de Cronbach	**Parte 1**	**Valor**	.743
		N de elementos	8
	Parte 2	**Valor**	.776
		N de elementos	8
	N total de elementos		16
Coeficiente de Spearman-Brown		**Longitud igual**	.750
		Longitud desigual	.750

Fuente: elaboración propia.

En este caso, *SPSS Statistics* nos proporciona los valores de alfa de Cronbach de ambas partes, así como el número de ítems de cada parte. Al ser un instrumento de 16 ítems, para conocer el coeficiente de Spearman-Brown tendremos que fijarnos en la fila de Longitud Igual. Por el contrario, si el instrumento tuviese una cantidad de números impares nos fijaríamos en Longitud desigual. Los resultados obtenidos por el método de las dos mitades apuntan a una fiabilidad aceptable.

4.6.3.3. El coeficiente Omega

Dejando estos métodos de lado, cabe hacer una pequeña mención al estadístico omega (ω). Este estadístico se presenta como una alternativa al estadístico alfa de Cronbach y puede sernos útil en aquellos casos en los que no queremos que nuestra fiabilidad dependa del número de ítems (McDonald, 1999). Omega es un buen estadístico para calcular la fiabilidad cuando realicemos análisis factoriales, pues este valor está basado en las cargas factoriales (Gerbing y Anderson, 1988). Es decir, el valor de omega no se va a ver influenciado por el número de ítems, si no que va a estar basado en la calidad de los autovalores de una dimensión.

El modo de interpretar este valor es el mismo que el de Alfa, aceptando valores entre .70 y 1 como buenos (Campo-Arias y Oviedo, 2008).

SPSS Statistics no trae consigo ningún comando para calcular el valor omega y debe ser calculado, por lo tanto, fuera de este de manera manual.

4.7. Principales diferencias entre validez y fiabilidad

Debemos tener cuidado, pues fiabilidad no es equiparable a validez. Cuando hablamos de validez a grandes rasgos podríamos centrarnos en que este concepto hace referencia a cómo un instrumento mide aquello que queremos medir, mientras que la fiabilidad hace alusión más al hecho de aquello que se esté midiendo (sea correcto o no) pueda perdurar en el tiempo y los resultados obtenidos no haya sido efecto del azar.

Como vemos en la Figura 8, la validez y la fiabilidad se pueden representar a través de una diana. Podemos imaginarnos que la diana es la variable latente que queremos medir, y el arma es el instrumento que vamos a usar para medir dicha variable. En la diana de la izquierda del todo, el arma no ha sido ni válida, ni fiable, pues a pesar de apuntar siempre al medio, cada caso ha ido a un sitio muy diferente, incluso fuera de la diana. En la segunda diana, tendríamos validez, pues hemos dado todos en la diana (hemos medido lo que queríamos medir), pero no ha sido muy fiable, pues cada tiro ha impactado en posiciones muy diferentes. En la tercera diana pasaría lo contrario. A pesar de haber apuntado a la diana, todos los disparos fueron fuera, por lo que no habría validez (estaríamos midiendo otra cosa), pero tendríamos fiabilidad (los resultados perduran y son consistentes). Finalmente, en la cuarta diana, apuntamos al medio y todas las balas dan en el centro, resultado de una validez y una fiabilidad buena.

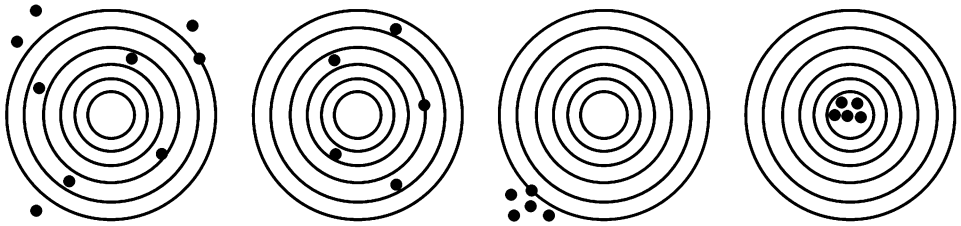


Figura 8. Validez y Fiabilidad a través de dianas.
Fuente: elaboración propia.

Como ya comentábamos, cabe destacar que la fiabilidad no es una característica de los instrumentos, sino más bien una característica del contexto en el que se ha aplicado. Es por ello por lo que siempre que apliquemos cualquier instrumento cuantitativo, es necesario que se aporte información sobre el contexto en el que se llevó a cabo el experimento, pues el proporcionar esa información, aparte de que es un indicio de buena práctica científica, ayudará a futuros investigadores a replantearse usar el mismo instrumento en muestras de características similares. Puede llegar a darse el caso, incluso, de que un instrumento sea fiable en un contexto y no fiable en otro.

4.8. La triangulación

Aunque la triangulación es un procedimiento más aplicado en estudios de índole cualitativa, también es necesario hacer cierta mención a este proceso en determinados estudios cuantitativos.

Cuando hablamos de triangulación, hablamos de una técnica compleja de comparación de diferentes tipos de análisis de datos, teorías, investigadores… con un mismo objetivo. Este proceso ayuda a otorgar mayor validez a unos datos y permiten potenciar las conclusiones que de ellos derivan (Rodríguez, Pozo y Gutiérrez, 2006).

En otras palabras, el objetivo cuando triangulamos no es otro que el de recabar la máxima información posible, a partir de distintos puntos de vista (instrumentos, observadores, entrevistas, teorías, métodos de investigación…) con el fin de poder contrastar toda la información obtenida y sacar conclusiones más sólidas.

Dependiendo de qué triangulemos podemos destacar, principalmente entre tres tipos de triangulación: Triangulación de investigador, triangulación teórica y triangulación metodológica (Arias, 2000).

4.8.1. Triangulación de investigador

Mediante esta técnica se emplean a múltiples observadores para investigar un fenómeno, en contraposición a un único investigador. Al triangular investigadores, se facilita la eliminación de posibles sesgos provenientes de una única persona y se asegura que las observaciones sean más confiables. Desde la investigación cuantitativa, ante un mismo suceso, distintos investigadores pueden preparar un instrumento para diferentes grupos de sujetos, totalmente heterogéneos, con el fin de comprobar si los datos obtenidos de sujetos asignados a una determinada zona son análogos a los datos obtenidos de sujetos de otra zona totalmente diferente.

4.8.2. Triangulación teórica

Mediante esta técnica se trata de probar teóricas o hipótesis contrarias. En investigación cuantitativa, este tipo de triangulación puede emplearse cuando tenemos dos ideas que van en distintas líneas y se pretende conocer cuál de las dos es cierta y cuál falsa, o si el resultado final es fruto de la unión de ambas hipótesis.

4.8.3. Triangulación metodológica

Mediante esta técnica se trata de combinar diferentes métodos de investigación ante un mismo suceso. Desde la investigación cuantitativa, se puede emplear esta

técnica para pasar unas escalas a unos sujetos, y contrastar los resultados con unas entrevistas, o unas observaciones, por ejemplo.

4.9. Consideraciones éticas

Aunque puede sonar absurdo, conviene recordar que cuando aplicamos en cualquier contexto determinado en ciencias sociales los instrumentos de medida debemos tener en cuenta que los sujetos que participan en nuestro estudio son personas.

Este hecho implica que el investigador debe hacer uso de unas normas de carácter innato a la propia ética profesional en aras de proteger al participante en todo momento.

En palabras de Silva (2002), la ética profesional está íntimamente ligada con el propio investigador e implica en todo momento entrega vocacional, responsabilidad, honestidad intelectual y práctica, siendo esencialmente un compromiso ineludible de hacer bien las cosas que no puede dejarse de hacer pues es precisamente esta propia ética la que nos hace más o menos, nos hace mejores o peores, nos enriquece o nos empobrece…

En esta línea y en lo que respecta al uso de instrumentos de medida documentales existen algunas consideraciones que podrían servir de utilidad al lector. Entre ellas destacamos las siguientes:

- Bajo ninguna circunstancia se debe alterar la información proporcionada por el participante. Esto incluye, no modificar ni falsificar ninguno de los resultados proporcionados por el sujeto.
- No se debe obligar al participante a que proporcione información que no desea. En esta línea, lo más adecuado sería establecer los ítems optativos (en caso de realizar la recogida de datos online).
- Se le debe informar al participante que la participación en el estudio es totalmente voluntaria, así como que su no participación no conlleva la pérdida de ningún tipo de beneficio.
- Los datos recogidos deben ser obligatoriamente anónimos. Esto supone la incapacidad de poder relacionar a una persona determinada con sus propios datos, por lo que se recomienda incluso no pedir si quiera las iniciales de nombres y apellidos de los estudiantes.

- Los datos recogidos deben ser obligatoriamente privados, siendo necesario comunicar al participante de antemano quién tendrá acceso a los mismos. De igual modo, los datos se deben interpretar como generalizaciones y no como particularidades.
- Se le debe comunicar de antemano al participante qué se va a hacer con la información que proporcione: Uso para artículos científicos, congresos… así como dónde se albergará la información y cuándo y cómo se eliminará la información.
- Antes de poder decidir si participar o no, se le debe hacer llegar al participante cuál va a ser el objeto de estudio.

En caso de trabajar con personas menores de edad, se deberá crear un consentimiento informado para las familias o tutores legales de las y los menores, como el recogido en la Figura 9. En caso de haber un mediador entre el investigador y las familias, como puede ser alguna institución o centro educativo, es altamente recomendable que sea la propia institución la que haga llegar este documento a las familias. De igual modo, es imprescindible indicar si se les tomará fotos y vídeos durante el proceso. Con el fin de que sirva de ayuda, se presenta a continuación un modelo de consentimiento informado, fácilmente adaptable a cualquier situación con menores.

Estimado padre o madre,
Soy (nombre y apellidos del investigador), (rango y procedencia: empresa, universidad…). Actualmente me hallo desarrollando un trabajo que tiene como fin (objetivo del estudio). Es por ello que a través de este documento pido su permiso para la participación de su hija/o. La relevancia de este proyecto reside en (indicar relevancia científica y social del estudio).

La tarea a desarrollar por los estudiantes será (quehacer de los participantes). Los cuestionarios están validados en población de edad igual a su hija/o, por lo que las preguntas que se le hagan están explicadas en términos que pueda entender y su hija/o solamente participará si él o ella está dispuesto/a a hacerlo. Solamente (personas que tendrán acceso a la información) tendré/tendremos acceso a la información de su hija/o. No se le tomará ningún tipo de imagen ni vídeo como parte de la investigación. Las conclusiones en el estudio serán indicadas en términos generales de grupo y no como particularidades de cada estudiante. Al finalizar el estudio me comprometo a hacer entrega del documento que recoja los resultados grupales a todos los padres y todas las madres interesados/as. Este documento será proporcionado al centro para que (objetivo por el que se le proporcionará el documento al centro). Para más información, póngase en contacto con la dirección del centro una vez finalizado el estudio.

La participación en este estudio es voluntaria. La participación del estudiante en este estudio no conllevará la pérdida de ningún beneficio al que tenga derecho. Incluso si le otorga permiso a su hijo/a para participar, él/ella tiene toda la libertad para rechazar la participación. Si su hijo/a acepta participar, tiene toda la libertad de terminar con su participación en cualquier momento. Ni usted, ni su hijo/a perderán ningún derecho o recurso legal debido a la participación de su hijo/a en esta investigación. Cualquier información que se obtenga del estudio será recogida de manera totalmente anónima y confidencial a través de (método que se seguirá para garantizar el anonimato). Esta información se almacenará (digitalmente o en papel) en un espacio personal e inaccesible para cualquier otra persona y se eliminará (momento en el que se eliminará la información. Por ejemplo: Cuando la vida útil del proyecto haya llegado a su fin). Las conclusiones de este estudio pueden ser usadas para (indicar si se usará la información para artículos, congresos o cualquier otro tipo de medio de divulgación) siempre salvaguardando todos los aspectos éticos recogidos en este documento.

En caso de tener cualquier otra duda o necesitar información más detallada, por favor póngase en contacto conmigo a través del correo electrónico (indicar correo electrónico). Una vez completado este documento, por favor, devuélvaselo al docente de su hijo/a en la mayor brevedad posible.

Atentamente,
(Nombre y apellidos del investigador)

Por favor, indique si estaría o no dispuesto a permitir que su hijo/a participe en este proyecto marcando una X en alguna de las dos oraciones siguientes. En la parte inferior, indique su nombre, su firma y el nombre de su hijo/a.

____ Acepto que mi hijo/a participe en el proyecto.
____ No acepto que mi hijo/a participe en el proyecto.

Nombre y Firma del padre/madre Nombre del estudiante

Figura 9. Modelo de consentimiento informado para investigaciones.

4.10. La reactividad psicológica

Cuando hablamos de **reactividad psicológica**, nos estamos refiriendo al fenómeno por el que los sujetos de un estudio modifican consciente o inconscientemente su comportamiento o conducta al sospechar que están siendo estudiados. Estos cambios, dependiendo de la situación, podrán ser positivos o negativos, aunque en todos los casos resulta desfavorable la presencia de este fenómeno, pues hace que el estudio pierda validez en cierto grado. Existen diferentes formas de manifestarse la reactividad psicológica. Entre las más significativas destacan las siguientes: El efecto Hawthorne, El efecto John Henry, El efecto Pigmalión y la Deseabilidad social.

4.10.1. El efecto Hawthorne

El efecto Hawthorne ocurre cuando los sujetos de un estudio conocen que están siendo estudiados y por consiguiente, alteran su conducta por esta causa y no en respuesta, por ejemplo, a ningún tipo de manipulación contemplada en el estudio experimental.

En investigación cuantitativa, este efecto puede darse en aquellos casos que se les comente a los sujetos que van a ser estudiados para un determinado fin. En ciencias sociales, es obvio que los sujetos serán conscientes de que una determinada escala, por ejemplo, sirve para obtener datos, y por ende, sus opiniones serán recogidas para ser estudiadas en conjunto. No obstante, es recomendable no profundizar excesivamente en el tema, y dar a los sujetos únicamente unas pinceladas de los objetivos y tareas a realizar.

4.10.2. El efecto John Henry

El efecto John Henry es una variación del efecto Hawthorne; pero en este caso, los sujetos, al conocer que forman parte de un grupo control, modifican su comportamiento para obtener unos resultados más favorables que los del grupo experimental.

Este término se comenzó a usar cuando un trabajador del acero, Gary Saretsky, en 1870 se enteró de que su producción sería comparada con la obtenida con un taladro de valor. Al enterarse, trabajo tan duramente que murió en este proceso.

En investigación cuantitativa, este efecto puede darse en aquellos estudios con grupo control, a quienes no se les hará nada, y experimental, a quienes tomarán parte de una intervención; de modo que el grupo control, al enterarse que sus resultados serán contrastados con los del grupo experimental, modificarán su conducta para tratar de mejorar a estos.

4.10.3. La deseabilidad social

La deseabilidad social hace referencia a cómo un individuo es capaz de modificar su conducta en base a lo que la sociedad considera como "lo mejor", con el fin de quedar bien con el investigador, profesor... favoreciendo que se dé el resultado experimental que se quería conseguir. Es decir, mediante la deseabilidad social, el sujeto trata de engañar al investigador dando unas respuestas, generalmente, socialmente muy aceptadas, que no son reales con su situación.

En el mundo de la investigación cuantitativa, si hacemos un pase de cuestionarios a alumnos, y les preguntamos sobre si son felices con ellos mismos, si se sienten a gusto con sus familias... aunque la realidad sea la contraria (poca felicidad, apatía hacia la familia...), es posible que el sujeto, por miedo o por querer complacer, conteste en vista de lo que la sociedad quiere escuchar (hay que ser felices, todas las familias tienen que ser maravillosas...).

Para controlar este fenómeno, existen instrumentos de medición, llamados escalas de sinceridad, que suelen llevar en ocasiones ítems falsos o ítems trampa, con la finalidad de medir la tendencia del sujeto a dar una imagen favorable de sí mismo en el test o a falsificar la contestación.

CAPÍTULO V: LOS ANÁLISIS DESCRIPTIVOS

5.1. Introducción

Todo buen estudio comienza por la estructuración de una buena base estadística. Generalmente, antes de entrar en análisis más complejos, es ciertamente común que se suelan presentar datos referentes a recuentos y valores descriptivos de un instrumento, una muestra…

Siempre que realicemos cualquier tipo de trabajo que involucre análisis de datos, el análisis descriptivo de nuestra muestra ha de ser el primer paso que deberemos dar. El análisis descriptivo nos va a ayudar a conocer mejor dónde se coloca y cómo se dispersan los datos recolectados.

Estos procedimientos básicos que deben aportarse adaptado a cualquier estudio nos permiten describir las propiedades de las distribuciones de determinadas variables. Es aquí, pues, donde hablamos de los procedimientos de Frecuencias y Descriptivos.

Generalmente, la diferencia entre uno y otro estriba en que, mientras que el procedimiento **Frecuencias** se emplea con variables categóricas, el procedimiento **Descriptivos** se emplea con variables cuantitativas.

5.2. El procedimiento Frecuencias

Supongamos que nuestra variable a describir es el nivel educativo de los participantes. En este caso, lo que nos interesa es conocer cómo se distribuyen los participantes en base a su nivel educativo. Por ello, lo que más podría interesarnos sería una Tabla de frecuencias, un gráfico de barras o un gráfico circular, por ejemplo.

Siguiendo este ejemplo, podríamos obtener una Tabla y un gráfico de barras similar al que se muestra en la Tabla 9.

Tabla 9. Ejemplo de Tabla de Frecuencias.

	Frecuencia	Porcentaje	Porcentaje Válido	Porcentaje Acumulado
Sin Estudios	12	30 %	30 %	30 %
Primarios	16	40 %	40 %	70 %
Secundarios	8	20 %	20 %	90 %
Superiores	4	10 %	10 %	100 %
Total	40	100 %	100 %	-

Fuente: elaboración propia.

La Tabla 9 nos proporciona 4 datos interesantes para cada categoría (sin estudios, primarios, secundarios y superiores):

- En el caso de la frecuencia, hace referencia al recuento total de casos para cada categoría. De este modo, en nuestra muestra hay 12 personas sin estudios, 16 con estudios primarios, 8 con estudios secundarios y 4 con estudios superiores. En total 40 personas.

- En el caso del porcentaje, nos indica qué porcentaje en relación al total (de casos válidos y perdidos) abarca la categoría indicada. Como es el caso, el 30% de la muestra no tiene estudios, el 40% tiene estudios primarios, y así sucesivamente.

- La siguiente columna es la del porcentaje válido. El porcentaje válido hace alusión a la frecuencia porcentual como en el caso de la columna porcentaje, pero a diferente de esta, el porcentaje válido calcula el porcentaje para cada categoría eliminando los valores perdidos. Algunos casos de valores perdidos pueden ser: 1) Aquellas personas que no respondieron porque o bien, no sabían o porque se negaban a responder o 2) Aquellas personas que no necesitaban responder a una pregunta (Por ejemplo, preguntarle a un soltero, cuántos años lleva casado).

- Finalmente, el porcentaje acumulado nos proporciona información, como bien indica el propio nombre, sobre qué cantidad de porcentaje se ha acumulado hasta esa misma categoría, sumando consigo las categorías anteriores. Es así, como el 30% de la muestra carece de estudios, el 70% de la muestra carece de estudios o tiene estudios primarios, el 90% de la muestra carece de estudios o tiene estudios primarios o secundarios…

Es posible acceder a este procedimiento en *SPSS Statistics* a través de *Analizar* > *Estadísticos Descriptivos* > *Frecuencias*. Haciendo clic en la pestaña *Gráficos*, nos permite generar para cada variable un **gráfico de barras**, un **gráfico circular** o un **histograma**, según el caso.

Esta opción de *Frecuencias* también tiene una opción que puede resultar interesante según qué caso. Estamos hablando de los valores percentil, a los cuales se accede dándole en el botón de *Estadísticos*. El **percentil** es una medida de posición comúnmente usada en estadística que nos indica qué porcentaje de datos, ordenados de menor a mayor, hay por debajo de determinado percentil. Existe en total un máximo de 99 percentiles.

Si nosotros tras hacer un examen al alumnado, por ejemplo, decimos que el valor del percentil 25 es 5,00 nos estamos refiriendo a que el 25% de nuestro alumnado ha sacado un 5,00 o menos. Si decimos que el valor del percentil 70 es 8,00 nos estamos refiriendo a que el 70% del alumnado ha sacado un 8,00 o menos.

Estos percentiles permiten hacer agrupaciones más grandes, llamadas **cuartiles**. Existen 3 cuartiles (Q1, Q2 y Q3), que dividen a una muestra total en 4 partes iguales, recogiendo cada grupo el 25% de la muestra. El percentil 25 (P25) equivale a Q1; el percentil 50 (P50) equivale a Q2, que a su vez este valor será exactamente igual que el valor de la mediana; y finalmente, el percentil 75 (P75) será equivalente a Q3. El total de los percentiles, se pueden clasificar también en **deciles**. Existen 9 deciles que dividen el total de la muestra en 9 partes diferentes. En la Figura 10, ilustrada a continuación se muestra una equivalencia gráfica entre cuartiles y deciles.

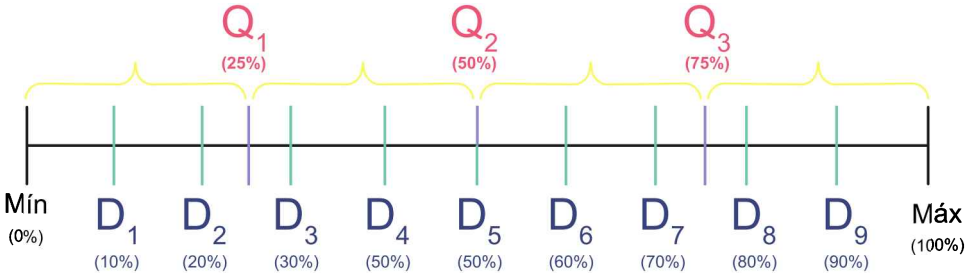


Figura 10. Equivalencia entre deciles (D) y cuartiles (Q).
Fuente: elaboración propia.

5.3. El procedimiento Descriptivos

Por otra parte, supongamos que nuestra variable a analizar, en vez de ser categórica es cuantitativa. En este caso, el procedimiento a seguir deberá ser el de Descriptivos, y no el de Frecuencias. Dentro de los estadísticos descriptivos, los grupos de estadísticos más significativos se agrupan del siguiente modo: Descriptivos de tendencial central, de dispersión y de distribución.

5.3.1. Descriptivos de tendencia central

Los descriptivos de tendencia central indican sobre qué punto se encuentra ubicado el conjunto de los datos, ayudando a resumir el total de los datos. Los más comunes son la media, la mediana y la moda.

5.3.1.1. La media aritmética

La **media aritmética** es el promedio aritmético obtenido de dividir la suma de los valores entre el número de casos. Se emplea para variables cuantitativas principalmente.

5.3.1.2. La mediana

La **mediana** indica el punto medio según el cual se encuentra, tanto por encima como por debajo el 50% de los casos. Cuando el número de casos es par, la mediana es el promedio de los dos casos centrales. Se suele emplear con variables ordinales.

5.3.1.3. La moda

La **moda** es el valor que ocurre con mayor frecuencia. Es posible que exista más de 1 moda en cada variable. Sirve para todo tipo de variables, aunque suele usarse más en variables categóricas.

5.3.2. Descriptivos de dispersión

Los descriptivos de dispersión nos sirven para cuantificar cuánto se separan o se dispersan los datos. Los más comunes son la desviación estándar o típica, la varianza y el rango.

5.3.2.1. La desviación típica

La **desviación típica** es uno de los estadísticos de dispersión más usados. Nos indica cuántos puntos por encima o por debajo se alejan los casos de la media. Se emplea para variables cuantitativas principalmente. Una desviación típica baja indica que los datos se agrupan de manera cercana a la media aritmética. Por el contrario, una desviación típica alta indica que los datos se agrupan de manera lejana a la media aritmética.

Pongamos por caso, en el que estudiando el sueldo medio de un país, todos los ciudadanos cobran en torno a los 1000 €. En este caso la desviación típica de los datos será baja. En el otro extremo, pongamos por caso un país en el que la mayoría de los ciudadanos cobran menos de 100 €, pero existen algunos casos de ciudadanos que cobran extremadamente mucho más que este grupo. Para ambos casos, la media aritmética (el sueldo medio) de ambos países será muy similar; no obstante, la desviación de los datos es totalmente diferente. Mientras que en el primer caso, existe una desviación típica baja, pues están todos los datos agrupados de manera cercana; para el segundo caso, existe una desviación típica alta, al estar los datos mucho más dispersos que en el primer caso.

5.3.2.2. La varianza

La **varianza** es el cuadrado de la Desviación Estándar. Se obtiene al calcular la suma de los cuadrados de las desviaciones respecto a la media dividida por el número de casos menos 1. Se emplea para variables cuantitativas.

5.3.2.3. El rango

Finalmente, el **rango**, también conocido como Recorrido o amplitud, es la diferencia entre el valor más grande (también llamado **m**áximo) y el valor más pequeño (también llamado **Mínimo**). Se emplea para variables cuantitativas, principalmente, aunque en los estudios cuantitativos, en pocos casos suele emplearse este estadístico.

5.3.2.4. Los valores atípicos

Dentro de cómo se dispersan los datos de una muestra dada, resulta interesante estudiar los **valores atípicos** (en inglés, *outlier*), pues por motivo de estos, podemos encontrarnos con unos resultados que realmente, eliminando estos casos, son totalmente diferentes a los que obtendríamos si se respetasen.

Un outlier es un caso, sujeto… que puntúa muy distintamente del resto de sujetos. Tener outliers en nuestra muestra puede ser un riesgo, pues estos casos pueden indicar que esos sujetos pertenecen a otra población diferente de la muestra que estamos inspeccionando. Es por este motivo, que es muy recomendable, antes de comenzar a realizar análisis más profundos, eliminarlos.

Imaginemos que estamos estudiando la motivación de un grupo de sujetos, y vemos que la amplia mayoría de personas obtiene entre un 4 y un 5 de puntuación final en la motivación (Uno 4,25; otro 4,65…), pero hay dos sujetos que se muestran extraños a estos valores, y puntúan, uno con un 12, y otro con un 0. En este caso, es posible que estos dos casos, alteren mucho la media aritmética y la desviación típica de los resultados, por lo que para ser lo más preciso posible con los datos reales, una posible opción sería eliminarles de la muestra.

En el caso de análisis con 1 única variable cuantitativa, lo más fácil para detectar estos casos atípicos es a través de gráficos de caja (en inglés, plotbox). La ruta a seguir en SPSS Statistics es Analizar > Estadísticos Descriptivos > Explorar. En la opción de Estadísticos, marcar la opción Valores atípicos. Se muestra un ejemplo en la Figura 11.

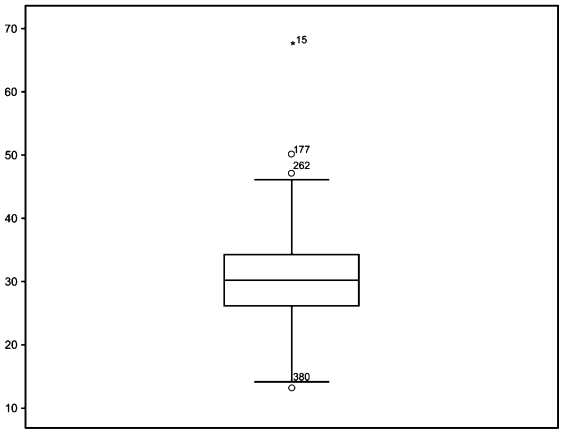


Figura 11. Gráfico de caja con valores atípicos.
Fuente: elaboración propia.

En este caso, tras inspeccionar el gráfico o diagrama de caja, observamos que para nuestra variable, tenemos cuatro casos atípicos: tres por encima, y uno por debajo. Es de observar, como dependiendo del tipo de caso atípico que sea, el símbolo empleado es uno u otro. Así, por ejemplo, para los casos atípicos leves, se ven representados a través de círculos, mientras que los casos atípicos extremos, se ven representados a través de asteriscos.

Por otra parte, en aquellos casos en los que queramos evaluar los casos atípicos en análisis con dos variables, o bien, analizamos las variables por separado a través de los diagramas de caja, o las analizamos en conjunto a través de gráficos de dispersión. En este segundo caso, podremos detectar un caso atípico, al obtener un gráfico similar al que se recoge en la Figura 12.

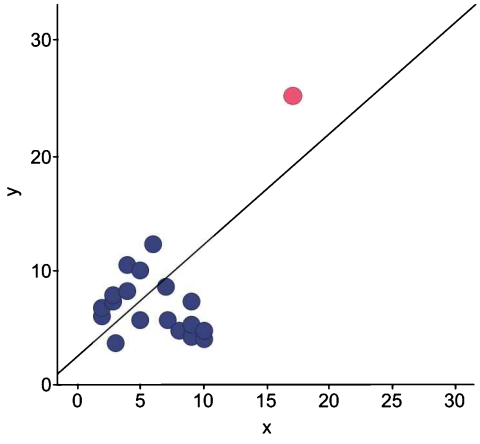


Figura 12. Gráfico de dispersión con un caso atípico.
Fuente: elaboración propia.

5.3.3. Descriptivos de distribución

Los descriptivos de distribución sirven para comparar la representación gráfica de nuestra muestra con la distribución normal. Esta distribución se compara a través de dos medidas: La curtosis y la asimetría.

5.3.3.1. La curtosis

Por una parte, la curtosis mide el grado de datos que se agrupan en torno a la moda. En otras palabras, la curtosis hace referencia a cuán de pronunciado es el pico de la curva. De este modo nos encontraremos con una curva leptocúrtica cuando sea más pronunciado que la curva de la distribución normal; con una curva mesocúrtica cuando sea igual que la curva de la distribución normal; o con una curva platicúrtica cuando sea menor que la curva de la distribución normal. Ilustramos cada una de estas curvas, en la Figura 13.

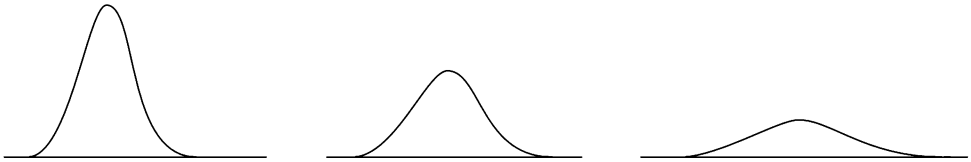


Figura 13. Tipos de curvas según la curtosis.
Fuente: elaboración propia.

5.3.3.2. La asimetría

Por otra parte, nos encontramos con la asimetría. Este concepto hace referencia a cuánto se mueve la curva hacia la izquierda o hacia la derecha. Si la mayoría de los casos se encuentran en la derecha, diremos que existe asimetría hacia la izquierda, y si por el contrario la mayoría de los casos se encuentran en la izquierda, diremos que existe asimetría hacia la derecha. En el caso de que la media, la mediana y la moda coincidan, habrá simetría. La representación gráfica de la asimetría se ilustra en la Figura 14.

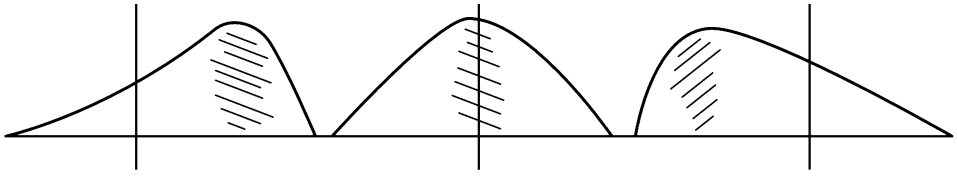


Figura 14. Tipos de curvas según la asimetría.
Fuente: elaboración propia.

El procedimiento para acceder a estos estadísticos en *SPSS Statistics* es a través de la ruta *Analizar* > *Estadísticos Descriptivos* > *Descriptivos*. Nótese que haciendo clic en

el botón *Opciones*, nos permite complementar nuestro análisis descriptivo con los principales estadísticos comentados.

5.4. Gráfica para análisis descriptivos

Los gráficos que realizaremos dependerán en función de qué variable estemos analizando, dependiendo de si estamos tratando con una variable cualitativa, cuantitativa discreta o cuantitativa continua.

Para el caso de las variables cualitativas, los gráficos principales que se suelen emplear son 2: Los diagramas de barras y los diagramas de sectores circulares. Para el caso de las variables cuantitativas discretas, como por ejemplo el número de hermanos, se suelen emplear diagramas de líneas y diagramas de barras; y para el caso de las variables cuantitativas continuas, como por ejemplo la altura, se suelen emplear los histogramas y polígonos de frecuencias. Pasamos a describir cada tipo de gráfico a continuación:

5.4.1. Los diagramas de barras

En los diagramas de barras, en el eje x se representan los datos ordenados en clases mientras que en el eje y se pueden representar frecuencias absolutas o relativas. A continuación se muestra un pequeño diagrama de barras a modo de ejemplo.

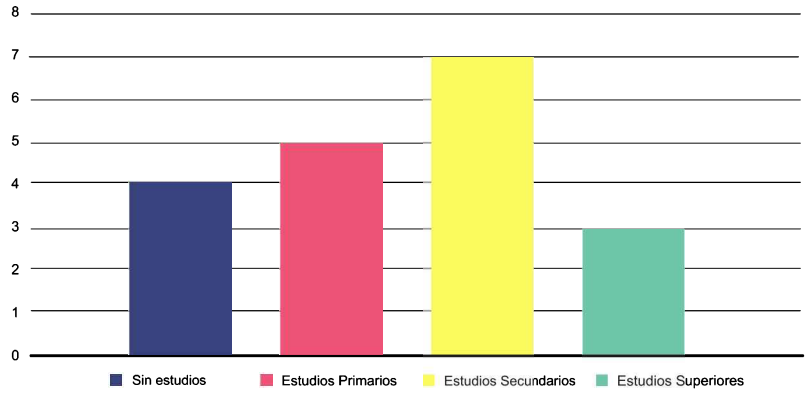


Figura 15. Ejemplo de Diagrama de barras.
Fuente: elaboración propia.

5.4.2. Los diagramas de sectores circulares

Los diagramas de sectores circulares se pueden considerar una Figura geométrica en la que la información se distribuye dentro de la Figura como puede ser un anillo en el que cada porción dentro de la Figura representa la información porcentual del total de datos.

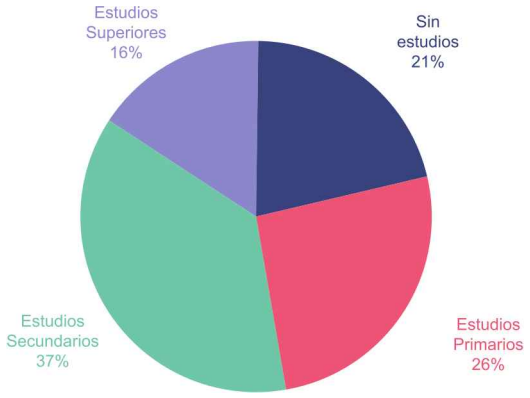


Figura 16. Ejemplo de Diagrama de sectores circulares.
Fuente: elaboración propia.

5.4.3. Los diagramas de líneas

Se emplea para mostrar cambios en una o más variables que se relacionan con una segunda variable, tal como el tiempo.

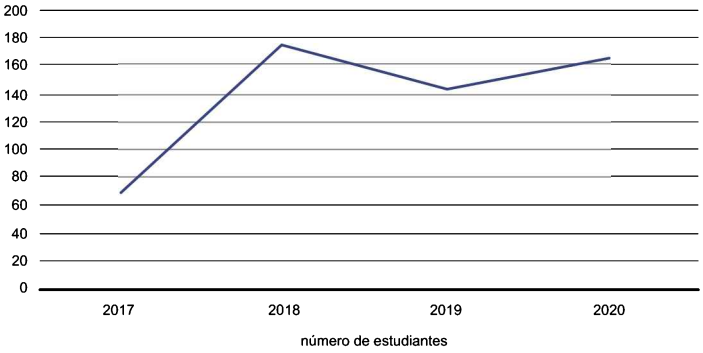


Figura 17. Ejemplo de Diagrama de líneas.
Fuente: elaboración propia.

5.4.4. Los histogramas

Los histogramas son gráficas que representan un conjunto de datos que se emplean para representar datos de una variable cuantitativa. En el eje horizontal se representan los valores tomados por la variable, en el caso de que los valores considerados sean continuos, la forma de representar los valores es mediante intervalos de un mismo tamaño llamados clases. En el eje vertical se representan los valores de las frecuencias de los datos. Las barras que se levantan sobre la horizontal y hasta una altura representan la frecuencia.

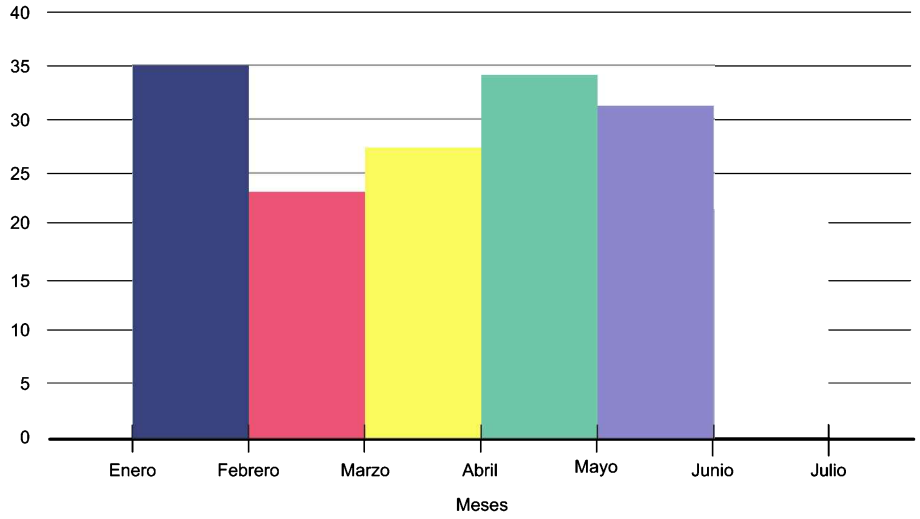


Figura 18. Ejemplo de Histograma.
Fuente: elaboración propia.

5.4.5. Los polígonos de frecuencias

Estos se construyen a partir de los puntos medios de cada clase. Se construye uniendo los puntos medios de cada clase localizados en las tapas superiores de los rectángulos utilizados en los histogramas de las gráficas.

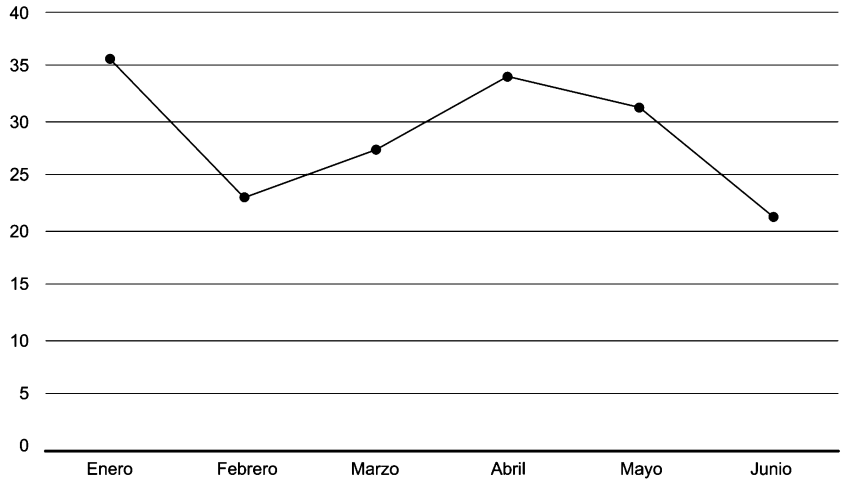


Figura 19. Ejemplo de Polígono de frecuencias.
Fuente: elaboración propia.

CAPÍTULO VI: LOS ANÁLISIS BIVARIADOS

6.1. Introducción

El análisis bivariado hace referencia a las diversas acciones que se pueden realizar entre dos variables. Entre las acciones más comunes nos encontramos con comparar las medias de dos variables o conocer cuál de unidas o desunidas están dos variables (nivel de correlación).

Dependiendo de cuál sea nuestro objetivo y cómo sean las condiciones de nuestra muestra, el análisis que elijamos será totalmente distinto. Es por este motivo, que a lo largo de este capítulo haremos una introducción a las pruebas bivariadas más comunes en el análisis de datos.

6.2. Comparación de medias

Una práctica muy extendido en el ámbito de la investigación cuantitativa es la de comparar las medias de dos grupos y conocer si entre ambas medias hay o no diferencias. En este tipo de análisis, en el que se involucran dos variables, vamos a tener siempre una variable independiente, y otra variable dependiente.

- Por una parte, la variable independiente en este tipo de análisis será siempre la variable de agrupación, es decir, la que separe a un grupo de otro. Esta variable puede contener 1, 2 o más de 2 categorías o grupos. Se le llama variable independiente porque esta variable no depende de nada, y por ende, se piensa que es innata al sujeto.

- Por ejemplo, estamos estudiando si los chicos tienen un mayor autoconcepto académico que las chicas. En este ejemplo, la variable independiente es el género, que a su vez está compuesto por dos grupos o categorías: Ser chico o ser chica. Se dice que esta variable independiente es una variable dicotómica pues posee dos categorías diferentes (chico/chica). Por el contrario, si tuviese más de 2 categorías, por ejemplo, el nivel de estudios de una persona (sin estudios, primarios, secundarios o superiores) diríamos que es una variable politómica. En estas pruebas de comparación de medias, la variable independiente va a ser siempre una variable categórica, ya sea dicotómica o politómica.

- Por otro lado, la variable dependiente es aquella variable que, como bien dice la propia palabra, se piensa que depende de la variable independiente.

En el ejemplo anterior, la variable dependiente sería el autoconcepto académico, pues pensamos que en función del género (variable independiente), el autoconcepto podrá oscilar según cual sea el valor del género. En estos análisis de comparación de medias, la variable dependiente será siempre una variable cuantitativa.

Otro aspecto a tener en cuenta en los análisis bivariados es si las categorías están relacionadas o por el contrario son independientes, pues las pruebas a emplear para ambos casos son diferentes.

Afirmaremos que los grupos están relacionados cuando existe una vinculación directa entre ambos grupos. Si, por ejemplo, aplicamos una prueba de lectura antes de comenzar el curso a un grupo de estudiantes y posteriormente, al finalizar el curso volvemos a realizar otra prueba de lectura a los mismos estudiantes para comparar los resultados, tendremos dos grupos: El grupo inicial y el grupo final. No obstante, son grupos que están relacionados, pues el grupo final y el grupo final ha estado consolidado por las mismas personas.

Por otra parte, si los grupos no tienen nada que ver hablaremos de grupos independientes. Si en un estudio queremos analizar la actitud hacia las matemáticas en función del género, tendremos dos grupos: chicos y chicas. No obstante, cada grupo es diferente del otro, por lo que hablaríamos de dos grupos independientes. Pero también, siguiendo este mismo ejemplo, si en otro estudio queremos analizar la actitud hacia las matemáticas en función del oficio, tendremos muchos grupos no relacionados entre ellos: El grupo de mineros, el de albañiles, el de fontaneros… Hablaríamos de más de dos grupos independientes.

En base a estas dos condiciones (número de categorías de la variable independiente y relación o no de esta misma variable), podemos diferenciar las siguientes pruebas estadísticas:

6.2.1. Prueba T de Student para muestras independientes

Es la prueba paramétrica encargada de medir diferencias de medias en los casos en que exista una variable independiente de 2 categorías (por ejemplo, chico/chica), y estos grupos sean independientes. Un ejemplo en el que se podría usar esta prueba estadística sería cuando queremos conocer si existen diferencias significativas entre chicos y chicas en el rendimiento académico. En *SPSS Statistics*, se accede a esta prueba a través de Analizar > Comparar medias > Prueba T para muestras independientes. Tras seleccionar cuál es la variable de prueba (variable

dependiente), y la variable de agrupación (variable independiente), obtendremos una Tabla dividida en dos grandes partes.

En un primer momento, nos muestra los resultados para la prueba de Levene, que recordamos que se basa en conocer si existe o no homocedasticidad en la muestra. Como ya hemos explicado, establecemos la hipótesis nula (No hay diferencias significativas en las varianzas, por lo tanto se asume que las varianzas son iguales) y la hipótesis alterna (Sí hay diferencias significativas en las varianzas, por lo tanto se asume que no las varianzas no son iguales). Al observar que el valor Sig. es p > .005, podemos concluir que, efectivamente, las varianzas son iguales. Llegados a este punto, seguiremos la línea de "Se asumen varianzas iguales" y nos olvidaremos de la línea "No se asumen varianzas iguales".

Tabla 10. Prueba de muestras independientes I.

		Prueba de Levene de igualdad de varianzas	
		F	Sig.
Rendimiento Académico	Se asumen varianzas iguales	1,191	,275
	No se asumen varianzas iguales		

Fuente: elaboración propia.

Ahora ya sí que se puede continuar analizando la prueba T de Student. La segunda parte de la tabla será algo similar, con un poco más de información a la siguiente tabla:

Tabla 11. Prueba de muestras independientes II.

		Prueba t para la igualdad de medias		
		t	gl	Sig. (bilateral)
Rendimiento académico	Se asumen varianzas iguales	-,192	1057	,848

Fuente: elaboración propia.

De la siguiente tabla, volvemos a fijarnos en el p-valor y establecemos nuestras dos hipótesis. Las hipótesis serán las siguientes:

- H_0: No hay diferencias significativas en el rendimiento académico en función del género.
- H_1: Sí hay diferencias significativas en el rendimiento académico en función del género.

En este caso, al ser un p-valor muy por encima del nivel crítico, generalmente, establecido en p = .005 (p = .848), no podemos descartar la hipótesis nula, por lo

que podemos confirmar que no existirían diferencias significativas en el rendimiento académico en función del género.

6.2.2. Prueba U de Mann Whitney

Para los casos en que no se cumplan los supuestos de normalidad, existe la prueba estadística análoga a la T de Student para muestras independientes, llamada prueba **U de Mann Whitney**. Esta prueba se accede en *SPSS Statistics* a través de Analizar > Pruebas no paramétricas > Cuadros de diálogos antiguos > 2 muestras independientes. En este caso, obtendremos una Tabla similar a la que se muestra a continuación:

Tabla 12. Prueba de muestras independientes.

	Rendimiento Académico
U de Mann-Whitney	139275,500
W de Wilcoxon	285345,500
Z	-,184
Sig. asintótica (bilateral)	,854

Fuente: elaboración propia.

El modo de interpretar la información es el mismo. Nos encontramos con un valor de la prueba estadística (U = 139.275,5) y un p-valor (p = .854), que en este caso es muy similar al valor realizado en la prueba paramétrica. Este p-valor al estar por encima del valor crítico, nos mantenemos con la hipótesis nula de igualdad de medias.

6.2.3. Prueba T de Student para muestras relacionadas

Es la prueba paramétrica encargada de medir diferencias de medias en los casos en que exista una variable independiente de 2 categorías relacionadas (por ejemplo, rendimiento académico antes y rendimiento académico después). Un ejemplo en el que se podría usar esta prueba estadística sería cuando queremos conocer si existen diferencias significativas entre los estudiantes de 4º de Educación Primaria en el rendimiento académico antes y después de haber llevado a cabo una intervención educativa. En *SPSS Statistics*, se accede a esta prueba a través de Analizar > Comparar medias > Prueba T para muestras relacionadas. En el cuadro resultante, introduciremos cuál es la variable inicial, y cuál es la variable contigua, para cada par de variables. Tras este procedimiento se obtendrá una Tabla similar con algo más de información a la que se muestra a continuación:

Tabla 13. Prueba de muestras emparejadas.

		t	gl	Sig. (bilateral)
Par 1	Pre_Rendimiento - Post_Rendimiento	1,930	192	,055

Fuente: elaboración propia.

Volvemos a establecer la hipótesis nula de igualdad de medias y la hipótesis alterna de diferencia de medias. Observamos que el p-valor es por poco superior al valor crítico (p = .055), por lo que es necesario quedarse con la hipótesis nula, afirmando que no existen diferencias significativas entre la fase pre y la fase post en el rendimiento académico de los estudiantes de 4º de Educación Primaria, tras haberse sometido a la intervención.

6.2.4. Prueba T de Wilcoxon

Para los casos en que no se cumplan los supuestos de normalidad, existe la prueba estadística análoga a la T de Student para muestras relacionada, llamada **prueba T de Wilcoxon**. Esta prueba se accede en *SPSS Statistics* a través de Analizar > Pruebas no paramétricas > Cuadros de diálogos antiguos > 2 muestras relacionadas.

En este caso, obtendremos una Tabla similar a la que se muestra a continuación:

Tabla 14. Prueba T de Wilcoxon.

	POST_Rendimiento – Pre_Rendimiento
Z	-1,817
Sig. asintótica (bilateral)	,069

Fuente: elaboración propia.

El modo de interpretar la información es el mismo. Nos encontramos con un valor de la prueba estadística (Z =-1.817) y un p-valor (p = .069). Este p-valor al estar por encima del valor crítico, nos mantenemos con la hipótesis nula de igualdad de medias, concluyendo que no existen diferencias significativas en el rendimiento académico.

6.2.5. Prueba ANOVA de un factor

Es la prueba paramétrica encargada de medir diferencias de medias en los casos en que exista una variable independiente de más de 2 categorías independientes (por ejemplo, oficios). Un ejemplo en el que se podría usar esta prueba estadística sería cuando queremos conocer si existen diferencias significativas en el autoconcepto académico de los estudiantes de varios colegios de la zona. En *SPSS Statistics*, se accede a esta prueba a través de Analizar > Comparar medias > ANOVA de un factor. En el cuadro resultante, introduciremos la variable dependiente en la lista de dependientes (en nuestro caso el autoconcepto, que pensamos que este depende

del colegio) y en el factor introduciremos nuestra variable independiente o variable de agrupación (en nuestro caso, colegio). Tras este procedimiento se obtendrá una Tabla similar con algo más de información a la que se muestra a continuación:

Tabla 15. Tabla de ANOVA.

	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Entre grupos	96,187	17	5,658	7,070	,000
Dentro de grupos	833,067	1041	,800		
Total	929,254	1058			

Fuente: elaboración propia.

Establecemos la hipótesis nula de igualdad de medias ($p > .005$) y alterna de diferencia de medias ($p < .005$). Observando el p-valor de la ANOVA de un factor, se aprecia cómo es un valor inferior al valor crítico $p = .005$ ($p = .000$), por lo que en este caso, es necesario rechazar la hipótesis nula, y quedarse con la hipótesis alterna, confirmando por lo tanto que en este caso existen diferencias en el autoconcepto del alumnado en función del colegio.

6.2.5.1. Las pruebas Post-Hoc

Pero esto no es todo, pues en caso de diferencias de medias en la ANOVA es necesario dar un paso más: Sabemos que hay diferencias, ¿pero entre qué grupos hay diferencias? Para dar respuesta a este interrogante se presentan las **pruebas Post-Hoc**, encargadas de comparar las medias de todas las posibilidades posibles: El grupo 1 con el 2, el grupo 1 con el 3, el grupo 2 con el 3…

Existen más de una decena de pruebas a elegir. Ante la duda, puede usarse a modo introductorio la prueba **post-hoc de Tukey**, o la prueba **post-hoc de Scheffe**. A estas pruebas, como al resto de pruebas Post-Hoc se acceden a través del mismo panel que en el que hemos realizado la ANOVA, haciendo clic en el botón Post-Hoc y seleccionando la prueba necesaria. Posteriormente obtendremos una Tabla similar a la que se muestra a continuación:

Tabla 16. Comparaciones múltiples.

(I) Colegio	(J) Colegio	Diferencia de medias (I-J)	Error estándar	Sig.	Intervalo de confianza al 95%	
					Límite inferior	Límite superior
Colegio 1	Colegio 2	,011	,155	,997	-,36	,38
	Colegio 3	,533*	,200	,023	,06	1,01
Colegio 2	Colegio 1	-,011	,155	,997	-,38	,36
	Colegio 3	,522*	,175	,009	,11	,93

(I) Colegio	(J) Colegio	Diferencia de medias (I-J)	Error estándar	Sig.	Intervalo de confianza al 95%	
					Límite inferior	Límite superior
Colegio 3	Colegio 1	-,533*	,200	,023	-1,01	-,06
	Colegio 2	-,522*	,175	,009	-,93	-,11

Fuente: elaboración propia.

En esta Tabla observamos todas las comparaciones posibles, y una vez más nos presenta un p-valor (Sig,) para cada comparación. Tras establecer nuevamente la hipótesis nula de igualdad de medias, y la hipótesis alterna de diferencia de medias, vemos, que en el presente caso, la hipótesis nula de igualdad de medias se da en las comparaciones del colegio 1 y 2; y la hipótesis alterna de diferencia de medias se da en las comparaciones del colegio 1 y 3, y 2 y 3. En otras palabras, el colegio 1 y 2 presentan similar autoconcepto, pero los colegios 1 y 3, y 2 y 3 presentan diferentes valores en el autoconcepto.

Para poder rematar este análisis, se muestra un análisis de subconjuntos homogéneos, encargado de intentar agrupar los colegios que han participado en el análisis en macrogrupos con el mismo autoconcepto, en este caso. Un ejemplo se muestra a continuación:

Tabla 17. Subconjuntos homogéneos.

Elige tu colegio	Subconjunto para alfa = 0.05	
	1	2
Colegio 3	3,49	
Colegio 2		4,01
Colegio 1		4,02
Sig.	1,000	,998

Fuente: elaboración propia.

Este análisis nos muestra que en el subconjunto 1 se halla únicamente el alumnado del colegio 3, y en el subconjunto 2 se halla el alumnado de los colegios 1 y 2, mostrándonos la media aritmética de cada colegio.

6.2.6. Prueba H de Kruskal-Wallis

Para los casos en que no se cumplan los supuestos de normalidad, existe la prueba estadística análoga a la ANOVA de un factor, llamada **prueba H de Kruskal-Wallis**. Esta prueba se accede en *SPSS Statistics* a través de Analizar > Pruebas no paramétricas > Cuadros de diálogos antiguos > k muestras independientes. En este caso, obtendremos una Tabla similar a la que se muestra a continuación:

Tabla 18. Prueba H de Kruskal Wallis.

	Autoconcepto
Chi-cuadrado	11,863
gl	2
Sig. asintótica	,003

Fuente: elaboración propia.

Nuevamente, establecemos nuestras dos hipótesis, nula y alterna, y observamos el p-valor (Sig. asintótica) del análisis.

En este caso, el valor obtenido (p = .003) es inferior al valor crítico permitido, por lo que rechazamos la hipótesis nula y nos quedamos con la hipótesis alterna, confirmando que existen diferencias significativas en los autoconceptos de los diferentes colegios estudiados.

6.2.7. Prueba ANOVA de medidas repetidas

Es la prueba paramétrica encargada de medir diferencias de medias en los casos en que exista una variable independiente, como puede ser una intervención (en dos etapas o más etapas) y una variable dependiente, como puede ser (el valor del rendimiento académico de un examen). Un ejemplo en el que se podría usar esta prueba estadística sería cuando queremos conocer si existen diferencias significativas en el clima social del aula de los estudiantes en diferentes puntos del curso. En *SPSS Statistics*, se accede a esta prueba a través de *Analizar* > *Modelo Lineal General* > *Medidas repetidas*. Para poder utilizar este procedimiento es necesario que tengamos bien preparada la base de datos que vayamos a utilizar. Para ello, en este análisis es necesario que se tenga una variable para cada momento de recogida y para cada caso. Es decir, si queremos conocer el clima social del aula de un grupo, tendremos que medir este valor a todo el alumnado en los diferentes puntos que queramos (Inicio de curso, final del 1º trimestre, final del 2º trimestre y final de curso, por ejemplo). Volviendo al *SPSS Statistics*, en la ventana resultante, nos pedirá nombrar el factor intra-sujetos (en nuestro caso lo llamaremos tiempo), y el número de niveles. El número de niveles hace referencia a la cantidad de momentos de medición que va a tener nuestra variable tiempo, que en este caso son 4 momentos. En la nueva ventana resultante arrastramos nuestras variables a cada uno de los diferentes momentos, como se muestra en la Figura que se muestra a continuación:

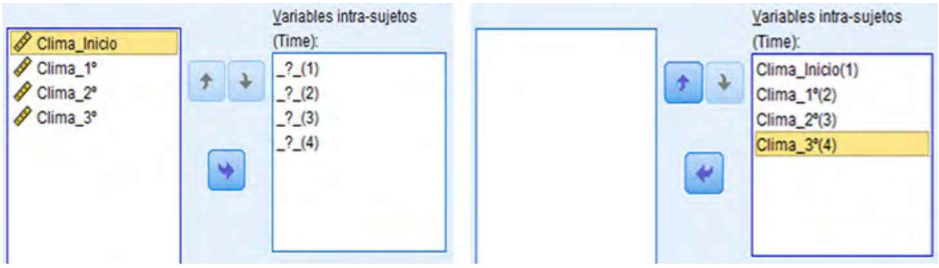


Figura 20. Establecimiento tiempos en una ANOVA de medidas repetidas.
Fuente: elaboración propia.

En esta ventana, la opción más interesante aparece en el botón de gráficos. Para crear el gráfico añadimos la variable tiempo al eje horizontal.

Posteriormente, llevando a cabo el análisis nos fijaremos en la Tabla de las pruebas multivariantes, que tendrá un aspecto similar a este:

Tabla 19. Pruebas multivariante en una ANOVA de medidas repetidas.

Efecto		Valor	F	Gl de hipótesis	gl de error	Sig.
Time	Traza de Pillai	,988	56,270	3,000	2,000	,018
	Lambda de Wilks	,012	56,270	3,000	2,000	,018
	Traza de Hotelling	84,405	56,270	3,000	2,000	,018
	Raíz mayor de Roy	84,405	56,270	3,000	2,000	,018

Fuente: elaboración propia.

Esta Tabla nos muestra cuatro pruebas diferentes para contrastar nuestras hipótesis. En cualquier caso, todas las pruebas son significativas, por lo que tendremos que rechazar la hipótesis nula y quedarnos con la hipótesis alterna, confirmando que sí que existen diferencias significativas en el clima social del aula en los diferentes puntos estudiados. De aquí, obtendremos un gráfico similar al que se muestra a continuación:

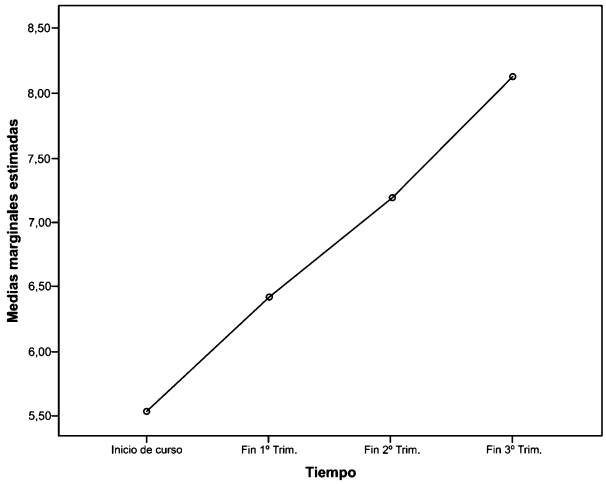


Figura 21. Ejemplo de gráfico a partir de una ANOVA de medidas repetidas.
Fuente: elaboración propia.

En esta línea de los análisis de medidas repetidas, podemos encontrarnos el caso de que no queremos únicamente medir diferentes momentos en un único grupo, sino que además, queremos comparar los resultados entre más de 1 grupo. Este tipo de análisis es muy común en aquellos estudios que se quiere realizar un pre-post, con un grupo control y un grupo experimental. Nosotros, siguiendo con el ejemplo, trataremos de conocer si hay diferencias significativas en el clima social del aula al inicio de curso entre los estudiantes de 4º y 5º de Educación Primaria.

En este caso, volviendo a *SPSS Statistics*, seguimos el mismo procedimiento, con una excepción: cuando lleguemos al momento de introducir las variables, en la sección de Factores inter-sujetos tendremos que introducir cuál es la variable de separación de grupos (en nuestro caso el grupo aula, que estará formado por los estudiantes de 4.º y 5.º).

Tras proceder con el análisis, accedemos a la Tabla de pruebas de contrastes dentro de sujetos, siendo esta Tabla algo similar a la que se muestra a continuación:

Tabla 20. Prueba de contraste intrasujetos en una ANOVA de medidas repetidas.

Origen	Time	Tipo III de suma de cuadrados	gl	Media cuadrática	F	Sig.
Tiempo	**Lineal**	14,078	1	14,078	217,865	,000
Tiempo * Grupo	**Lineal**	4,122	1	4,122	63,793	,000
Error(Time)	**Lineal**	,517	8	,065		

Fuente: elaboración propia.

Para este caso, resulta interesante fijarse en el p-valor de la fila Tiempo, y en el p-valor de la fila Tiempo*Grupo. Estas dos filas se diferencian en que mientras la fila Tiempo nos muestra el p-valor que nos ayudará a conocer si ha habido diferencias significativas en el clima social del aula en el tiempo sin tener en cuenta el grupo de pertenencia, la fila Tiempo*Grupo nos muestra el p-valor que nos ayudará a conocer si ha habido diferencias significativas en el tiempo en función del grupo de pertenencia. Para este tipo de análisis, este p-valor suele tener especial interés, pues nos ayudará a conocer si nuestro tratamiento, intervención… ha sido más efectivo que el de otro grupo.

En caso de existir diferencias, si la variable independiente posee más de 2 categorías, *SPSS Statistics* nos permitirá también llevar a cabo pruebas post-hoc. En nuestro caso como solo hemos tenido en cuenta 2 momentos, el inicio y el final del curso, no podremos llevar a cabo Post-Hoc.

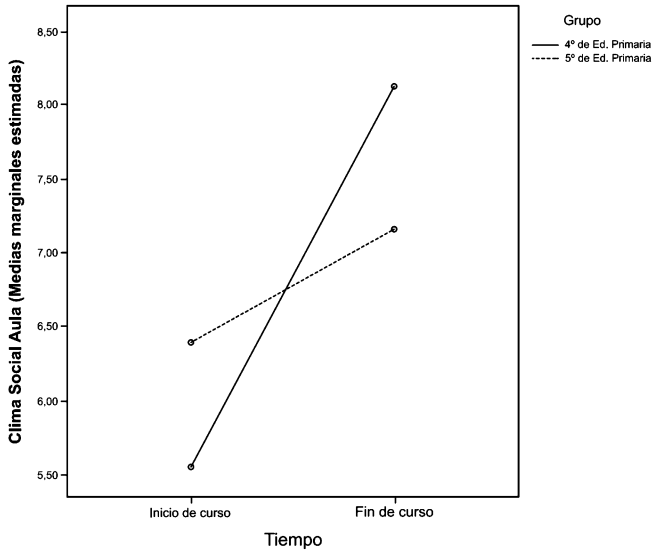


Figura 22. Ejemplo de gráfico pre-post con grupo control y experimental.
Fuente: elaboración propia.

Según los datos del p-valor obtenido, podemos rechazar para ambos casos las hipótesis nulas y quedarnos con las hipótesis alternas. Por lo tanto

- Sí que existen diferencias significativas en el clima social del aula con el paso del tiempo, independientemente del grupo de pertenencia.
- Sí que existen diferencias significativas en el clima social del aula con el paso del tiempo en función del grupo, siendo el grupo de 4º de Educación Primaria

quien mejora más significativamente, que el grupo de 5º de Educación Primaria, quien mejora, pero más lentamente.

Se puede observar esta información en el gráfico recogido en la Figura 22, proporcionado por *SPSS Statistics*.

Para formar este tipo de gráficos se accede al panel principal de Medidas Repetidas, y en la pestaña de gráficos, en el eje horizontal podemos la variable Tiempo, y en Las líneas separadas introducimos la variable de agrupación.

6.3. Tamaño del efecto

Tan importante como conocer si existen diferencias significativas entre nuestros grupos es analizar cuán de grandes son dichas diferencias.

El valor de la significancia solamente nos aporta información sobre si los grupos son iguales o no. No obstante, para conocer el tamaño de dichas diferencias, lo que se suele conocer como el **tamaño del efecto**, es necesario ir un paso más lejos.

Al igual que cada objetivo tiene sus pruebas estadísticas, con el tamaño del efecto ocurre lo mismo: Cada prueba estadística tiene su tamaño del efecto (Domínguez-Lara, 2017). Los principales estadísticos asignados al tamaño del efecto se muestran en la Tabla 4 y se explican a continuación.

Tabla 21. Principales estadísticos para calcular el tamaño del efecto.

Grupos	Prueba estadística	Tamaño del Efecto	Interpretación
2 grupos	T de Student	d	.20: pequeña, .50: mediana, .80: grande.
	U de Mann Whitney	r	.20: pequeña, .30: mediana, .50: grande
	T de Wilcoxon		
>2 grupos	ANOVA	η2	.01: pequeña, .06 mediana, .14: grande.
	H de Kruskal-Wallis		

Fuente: elaboración propia.

6.3.1. D de Cohen

Explicando un poco cada estadístico, el más común entre todos los aplicados en pruebas paramétricas de 2 grupos es la ***d* de Cohen**, aunque no el único, pues nos encontramos también con otros como la Δ de Glass o la *g* de Hedges.

La *d* de Cohen relaciona las medias y las desviaciones típica de cada grupo y se calcula a través de la siguiente fórmula matemática, donde $\overline{x}_1$ y $\overline{x}_2$ es la media aritmética de cada grupo, la *DTPond* es la Desviación Típica ponderada, n_1 y n_2 hacen referencia al

número de casos o sujetos de cada grupo y DT_1 y DT_2 es la Desviación Típica de cada grupo:

$$d = \frac{\bar{x}_1 - \bar{x}_2}{DT_{Pond}}$$

$$DT_{Pond} = \sqrt{\frac{(n_1 - 1) \cdot DT_1^2 + (n_2 - 1) \cdot DT_2^2}{n_1 + n_2}}$$

SPSS Statistics no nos lo calcula directamente por lo que imaginemos que hemos obtenido los datos que aparecen en la Tabla 22 tras obtener que existen diferencias significativas entre dos grupos en una prueba T de Student:

Tabla 22. Caso práctico para calcular la d de Cohen.

Grupo	Casos - Sujetos	Media aritmética (x⃐)	Desviación Típica (DT)
Grupo 1	105	4.014	1.671
Grupo 2	79	3.493	1.785

Fuente: elaboración propia.

Comenzaríamos calculando la DT_{Pond}, que quedaría de la siguiente manera:

$$DT_{Pond} = \sqrt{\frac{(105 - 1) \cdot 1.671^2 + (79 - 1) \cdot 1.785^2}{105 + 79}}$$

El resultado de la DT_{Pond} sería de 2.92. Ahora siguiendo con la siguiente fórmula, sustituimos y quedaría:

$$d = \frac{4.014 - 3.493}{2.92} = 0.178$$

El valor de la *d* de Cohen es de 0.178, por lo que, aunque sí existen diferencias significativas entre los grupos, podemos concluir que el tamaño de las diferencias es pequeño.

Aquellos investigadores que no quieran calcular este valor manualmente, en internet se puede encontrar una cantidad variada de calculadoras que introduciéndoles los datos proporcionarán el valor directamente.

La *d* de Cohen puede ser negativa dependiendo de los valores totales de las medias de ambos grupos. Más concretamente, la d de Cohen será negativa siempre que la media aritmética del primer grupo sea menor que la media aritmética del segundo grupo, en cuyo caso supondrá un tamaño del efecto favorable para el segundo grupo.

6.3.2. R de Rossenthal

Otro estadístico para 2 grupos empleado en pruebas no paramétricas es la *r* de Rossenthal, que se calcula con la siguiente fórmula, donde el valor *Z* es el valor que nos aporta *SPSS Statistics* de una prueba U de Mann Whitney o T de Wilcoxon y *N* es la cantidad total de gente con la que se ha hecho el análisis:

$$r = \frac{Z}{\sqrt{N}}$$

6.3.3. Eta Cuadrada

En el caso de que tengamos más de 2 grupos, para el caso de la ANOVA y H de Kruskal Wallis, el estadístico para calcular el tamaño del efecto será principalmente η^2 (eta cuadrada), aunque también hay otros estadísticos como ω^2 y ε^2 (Moncada-Jiménez, Solera-Herrera, y Salazar-Rojas, 2002).

A diferencia de la d de Cohen, *SPSS Statistics* sí que proporciona el valor de η2. Para ello, accedemos a *Analizar* > *Comparar medias* > *Medias*. Aquí, en el botón *Opciones*, seleccionamos la opción *Tabla de ANOVA y eta* y obtendremos una Tabla similar a la siguiente:

Tabla 23. Eta y Eta Cuadrada.

	Eta	Eta cuadrada
Autoconcepto * Colegio	,215	,046

Fuente: elaboración propia.

De esta Tabla, nos resulta interesante el valor de eta cuadrado (η^2 = .046), que conociendo la interpretación de este valor, nos indica que el tamaño de las diferencias existentes es entre pequeñas y medianas.

6.4. Correlación de variables

Una de las múltiples opciones que podemos aprovechar en un análisis bivariado es el de conocer el grado de correlación entre dos variables cuantitativas.

En esta línea, los coeficientes de correlación numéricos más comunes son el Coeficiente de **Correlación de Pearson**, para pruebas paramétricas, y el Coeficiente de **Correlación de Spearman**, para pruebas no paramétricas.

Imaginemos que queremos ver si la nota académica del grado universitario de los estudiantes, medida de 0 a 10 puntos, tiene relación con la cantidad de ingresos que estos tienen.

Como podemos apreciar tenemos dos variables continuas, la nota académica y la cantidad de ingresos, por lo que se procedería a aplicar el coeficiente de correlación de Pearson o de Spearman, dependiendo si cumplen con las condiciones paramétricas o no.

La ruta en *SPSS Statistics* para llegar a estos coeficientes de correlación es a través de *Analizar > Correlacionar > Bivariadas*.

Una vez introducidas las variables que queremos analizar obtendremos una Tabla similar a la Tabla 24.

Tabla 24. Ejemplo de coeficiente de correlación de Pearson y Spearman.

		Variable 1	Variable 2
Variable 1	Correlación	1	.609
	Sig. (bilateral)		.000
Variable 2	Correlación	.609	1
	Sig. (bilateral)	.000	

Fuente: elaboración propia,

Estos dos coeficientes de correlación oscilan entre -1 y 1. Mientras más tienda a 0 implicará menor correlación y mientras más tienda a 1 o -1 implicará mayor correlación. Una interpretación aproximada del valor r de Pearson o ρ de Spearman podría ser la que se muestra en la Tabla 25.

Tabla 25. Interpretación de los valores de correlación r de Pearson y ρ de Spearman.

	Muy Alta	Alta	Media	Baja	Muy baja
Positiva	$1 \leq r < 0.80$	$+0.80 \leq +0.60$	$+0.60 \leq +0.40$	$+0.40 \leq +0.20$	$+0.20 \leq 0$
Negativa	$-1 \leq r < -0.80$	$-0.80 \leq -0.60$	$-0.60 \leq -0.40$	$-0.40 \leq -0.20$	$-0.20 \leq 0$

Fuente: elaboración propia.

Un valor de **correlación positiva** supone una relación directa entre las variables (Las dos aumentan gradualmente), mientras que una **correlación negativa** supone una relación inversa entre las variables (A medida que aumenta una, decrece la otra).

En la Tabla 4, podemos apreciar los números destacados, que nos indican por una parte que el coeficiente de correlación es entre moderado y fuerte (.609), que además tiene una correlación positiva (observable en el signo de la correlación) y que, finalmente, es un valor significativo (.000) al estar por debajo de nuestro nivel de significación (.05); hecho que nos permitiese afirmar que la *Variable 1* y la *Variable 2* están correlacionadas.

Estos tipos de análisis suelen ir acompañados de un gráfico específico conocido como gráfico de dispersión, que une cada eje con una variable cuantitativa diferente. Se muestra un ejemplo en la Figura 23. Mientras más recta sea la línea que forman los puntos, mayor será la correlación entre los mismos.

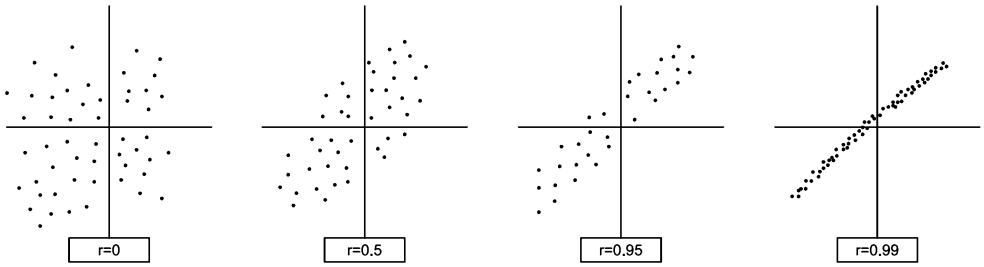


Figura 23. Gráficos de dispersión y coeficientes de correlación.
Fuente: elaboración propia.

En *SPSS Statistics* se puede acceder a crear este tipo de gráficos de dispersión a través de *Gráficos* > *Generador de Gráficos*.

6.5. La Kappa de Cohen

El índice **Kappa de Cohen** es un estadístico que nos permite conocer el grado de concordancia o acuerdo entre dos evaluadores, jueces o instrumentos de medición ante un mismo fenómeno.

Este procedimiento, suele emplearse en investigación cuantitativa en áreas como la educación o la psicología en el análisis de contenido de un instrumento. Mediante este procedimiento se pretende conocer el grado de acuerdo o desacuerdo que hay entre dos evaluadores ante la pertinencia de un determinado ítem para ser introducido en una escala o cuestionario. El índice de Kappa es un estadístico que oscila entre 0 y 1, y que mientras más cercano a 1, implica que el grado de concordancia es mayor.

Para desarrollar este análisis inicialmente establecemos las siguientes hipótesis:

H_0: No existe concordancia entre los evaluadores.

H_1: Sí existe concordancia entre los evaluadores.

Tras la ordenación de datos, de una manera similar a la que se muestra a continuación, en la que cada caso representa a un ítem en concreto y cada evaluador contempla la pertinencia o no de cada ítem se organiza la base de datos del siguiente modo:

	Evaluador_1	Evaluador_2
1	No pertinente	No pertinente
2	Pertinente	Pertinente
3	No pertinente	No pertinente
4	Pertinente	Pertinente
5	Pertinente	Pertinente
6	No pertinente	Pertinente
7	Pertinente	No pertinente
8	Pertinente	Pertinente
9	No pertinente	Pertinente
10	Pertinente	No pertinente
11	Pertinente	Pertinente
12	Pertinente	Pertinente
13	No pertinente	No pertinente
14	Pertinente	Pertinente
15	Pertinente	No pertinente
16	No pertinente	Pertinente
17	No pertinente	No pertinente
18	Pertinente	Pertinente
19	Pertinente	Pertinente
20	No pertinente	No pertinente

Figura 24. Base de datos para realizar un análisis con la Kappa de Cohen .
Fuente: elaboración propia.

En *SPSS Statistics*, el acceso al índice Kappa de Cohen se realiza a través de *Analizar* > *Estadísticos Descriptivos* > *Tablas cruzadas*. Introducimos a cada evaluador en las filas y columnas y en el botón *Estadísticos* seleccionamos la opción Kappa. En el ejemplo que se muestra a continuación, se presentan los resultados de una Tabla cruzada en la que se intenta conocer el grado de acuerdo de unos evaluadores en el proceso de consolidación de una escala en fase inicial de 20 ítems.

Tabla 26. Ejemplo de Tabla de contingencia en un análisis con Kappa de Cohen.

		Evaluador 2		Total
		No pertinente	Pertinente	
Evaluador 1	No pertinente	5	3	8
	Pertinente	3	9	12
Total		8	12	20

Fuente: elaboración propia.

En esta escala vemos que hay 5 ítems en los que ambos evaluadores están de acuerdo en que no es pertinente incluirlo, y 9 ítems que ambos evaluadores están de acuerdo en incluirlo. No obstante, hay 6 (3 + 3) casos en total de desacuerdo, en los que un evaluador considera pertinente introducir un determinado ítem y el

otro evaluador no. Para conocer el grado de acuerdo continuamos interpretando el p-valor del estadístico Kappa, recogido en la Tabla que se muestra a continuación.

Tabla 27. Ejemplo del análisis de Kappa de Cohen.

	Valor	Error estandarizado asintótico	T aproximada	Significación aproximada
Kappa	,375	,211	1,677	,094

Fuente: elaboración propia.

En este caso, podemos observar que el p-valor (p = .094) es superior al nivel crítico establecido de antemano, que suele ser de p = .005, por lo que esto implica que es necesario mantener la hipótesis nula que suponía la no concordancia existente entre los evaluadores. Esta no concordancia puede verse reflejada en el bajo valor de Kappa (k = .375), siguiendo la siguiente interpretación:

Tabla 28. Interpretación del índice Kappa de Cohen.

Índice Kappa (k)	Interpretación
0.00 – 0.20	Muy baja concordancia
0.21 – 0.40	Baja concordancia
0.41 – 0.60	Moderada concordancia
0.61 – 0.80	Alta concordancia
0.81 – 1.00	Muy alta concordancia

Fuente: elaboración propia.

6.6. Las Tablas de contingencia y la prueba Chi-Cuadrado

Las **Tablas de contingencia**, también conocidas como **Tablas cruzadas**, son Tablas en la que la información recogida queda agrupada en una matriz en la que se trata de conocer si dos variables nominales (por ejemplo, tipo de universidad) u ordinales (por ejemplo, estudios alcanzados) se relacionan. Un ejemplo de lo que es una Tabla de contingencia se muestra a continuación. En ella tratamos de colocar los datos del tipo de universidad al que van hombres y mujeres:

Tabla 29. Ejemplo de Tabla de contingencia.

		Tipo de Universidad		Total
		Pública	Privada	
Género	Mujer	157	81	238
	Hombre	97	49	146
Total		254	130	384

Fuente: elaboración propia.

Aunque en este caso, la Tabla sea de 2x2 también es posible realizar Tablas de todas las opciones posibles: 2x3, 3x3, 4x5…

Este tipo de análisis se puede realizar con más de 2 variables a través de los procedimientos de capa, pero es un proceso más complejo y que no merece la pena explicar en un primer acercamiento a la investigación cuantitativa. Es por ello que este epígrafe estará centrado en el análisis bivariado de variables nominales u ordinales.

Una vez que hemos realizado nuestra Tabla de contingencia ya nos permite conocer un poco más acerca de la información. No obstante, si lo que nos interesa es conocer si las dos variables analizadas están o no relacionadas es necesario dar un paso más.

Llegados a este punto surgen dos conceptos interesantes: El **recuento observado** y el **recuento esperado**.

El recuento observado es la frecuencia que nos ha mostrado la Tabla de contingencia. En el ejemplo, vemos que el recuento observado de mujeres que estudian en la universidad pública es 157. El recuento observado es el valor real, lo que se tiene. Por otra parte, tenemos el recuento esperado, que como bien dice su nombre, es el valor que se esperaría que tuviese una determinada casilla en caso de que todo fuese normal y no existiese ninguna diferencia.

Tabla 30. Tabla de contingencia con frecuencia esperada.

			Tipo de Universidad		Total
			Pública	Privada	
Género	Mujer	Recuento	157	81	238
		Recuento esperado	157,4	80,6	238,0
	Hombre	Recuento	97	49	146
		Recuento esperado	96,6	49,4	146,0
Total		Recuento	254	130	384
Recuento esperado		254,0	130,0	384,0	

Fuente: elaboración propia.

Por ejemplo, en la Tabla 30, podemos intuir a simple vista que no existen grandes diferencias entre lo que existe y lo que se esperaba que existiese. Las mujeres que van a la pública son 157, y para que existiese una proporcionalidad perfecta sería necesario que existiesen 157,4 mujeres en la pública. En el caso de mujeres en la privada igual: existen 81, pero se esperaba 80,6 en condiciones de perfecta proporcionalidad…

Bajo esta premisa de jugar con el recuento esperado y observado de las diferentes casillas de una Tabla nace una nueva prueba estadística llamada prueba de Chi-Cuadrado o ji-cuadrado, que nos indicará en general, si las variables de una

determinada Tabla de contingencia están o no relacionadas. A un valor de chi-cuadrado, le acompaña un p-valor, para su fácil interpretación. En este caso y para el ejemplo que estamos mostrando, establecemos nuestra hipótesis nula de igualdad en las variables y nuestra hipótesis alterna de diferencia en las variables.

Tabla 31. Ejemplo de coeficiente de correlación de Pearson y Spearman.

	Valor	gl	Significación asintótica (bilateral)
Chi-cuadrado de Pearson	,009	1	,924

Fuente: elaboración propia.

Interpretando el p-valor observamos que es un valor muy por encima al nivel crítico p=.005 (p = .924), por lo que es necesario quedarse con la hipótesis nula, afirmando que las variables tipo de universidad y género no dependen la una de la otra. Es decir, no hay ciertas preferencias entre hombres por estudiar en un tipo de universidad, ni cierta preferencia entre mujeres por estudiar en un tipo de universidad.

Yendo a *SPSS Statistics*, para realizar una Tabla de contingencia accedemos a *Analizar* > *Estadísticos descriptivos* > *Tablas cruzadas*. En el botón *Casillas* podemos marcar, además del recuento observado, el recuento esperado. Finalmente, para que el programa nos facilite la prueba Chi-Cuadrado, vamos al botón *Estadísticos* y seleccionamos *Chi-Cuadrado*.

6.7. El análisis de correspondencia simple

El **análisis de correspondencia** es un análisis descriptivo cuyo objetivo es resumir una gran cantidad de datos en un número reducido de dimensiones, con la menor pérdida de información posible.

Tabla 32. Ejemplo de Tabla de correspondencias.

Grado Universitario	Deporte						
	Fútbol	Baloncesto	Tenis	Rugby	Ping-Pong	Atletismo	Total
Ed. Primaria	39	40	20	26	27	4	156
Ed. Social	35	25	13	55	19	5	152
Ed. Infantil	28	24	7	14	17	3	93
Periodismo	13	7	5	5	6	4	40
Medicina	6	5	4	4	7	2	28
Enfermería	13	10	6	5	14	2	50
Odontología	1	1	2	3	2	0	9
Psicología	15	11	10	4	9	1	50
Biología	6	6	3	2	2	0	19
Geología	5	4	4	5	3	0	21

Grado Universitario	Deporte						
	Fútbol	Baloncesto	Tenis	Rugby	Ping-Pong	Atletismo	Total
Física	12	12	2	7	9	1	43
Química	3	2	0	1	2	0	8
Total	176	147	76	131	117	22	669

Fuente: elaboración propia.

Este tipo de análisis se utiliza principalmente cuando nos encontramos con Tablas de contingencia formadas por dos variables nominales u ordinales. Un ejemplo en el que podría utilizarse este análisis sería el de conocer la preferencia del alumnado de los distintos grados universitarios sobre determinados deportes. De este análisis podríamos obtener una Tabla similar a la siguiente.

Esta Tabla, por sí mismo no nos aporta gran información, simplemente nos deja recogida la preferencia de cada estudiante de grado hacia cierto deporte.

Es necesario, pues continuar con el análisis de correspondencia y conocer si la información presente en esta Tabla resulta significativa o no. Para ello, se nos presenta una Tabla resumen (Véase Tabla 33), que recoge 2 conceptos interesantes: La **significancia** y la **proporción de inercia explicada**.

Tabla 33. Ejemplo del análisis de correspondencia.

Dimensión	Valor singular	Inercia	Chi cuadrado	Sig.	Proporción de inercia	
					Contabilizado	Acumulado
1	,247	,061			,577	,577
2	,132	,017			,164	,741
3	,129	,017			,156	,897
4	,094	,009			,084	,981
5	,045	,002			,019	1,000
Total		**,106**	**70,903**	**,073**	**1,000**	**1,000**

Fuente: elaboración propia.

Por una parte, la significancia (Sig.), aunque no es un concepto que no haya salido, nos indica si el modelo propuesto es significativo o no. Para ello, nuevamente, tenemos que plantear nuestras dos hipótesis de igualdad y diferencias. En este caso, el modelo presenta un p-valor algo superior al valor crítico (p = .073) por lo que es necesario mantener la hipótesis nula y confirmar que no hay diferencias significativas entre el grado universitario y el deporte de preferencia. No obstante, a los valores por debajo de p = .100 se les dice que son **valores tendenciales**, es decir, que sin llegar a ser valores estadísticamente significativos, son valores que presentan cierta tendencia a que existan diferencias.

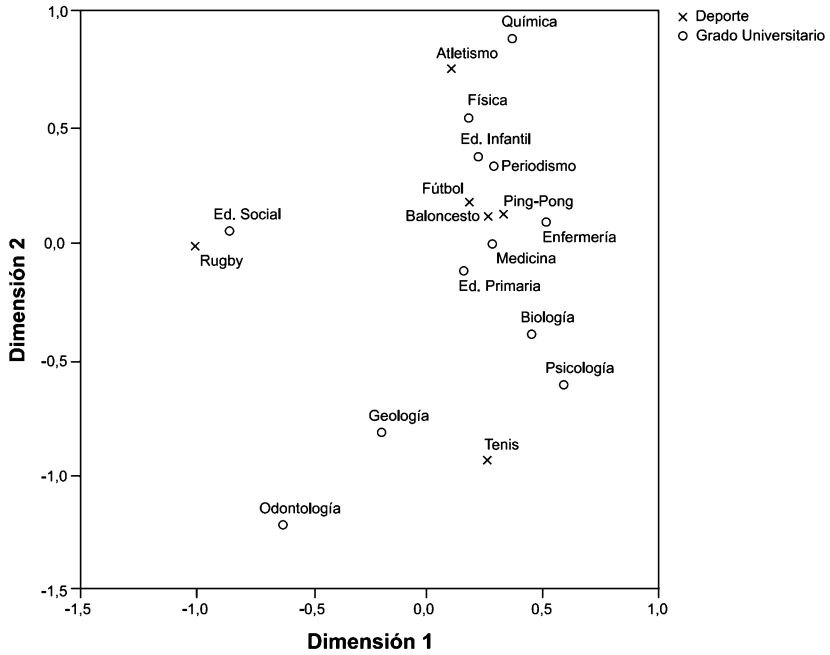


Figura 25. Gráfico de dispersión biespacial.
Fuente: elaboración propia.

Por otra parte, en la Tabla resumen (Véase Tabla 33) se nos introduce el concepto de Proporción de Inercia. La **Inercia** en los análisis de correspondencia es el equivalente a la varianza explicada en los análisis factoriales. Es decir, ambos valores nos indican qué porcentaje de los datos se han podido explicar gracias al modelo propuesto.

Generalmente, los análisis de correspondencias suelen estar formados por 2 dimensiones, por lo que la inercia explicada en esas 2 dimensiones es del 74.1% (o en forma decimal .741). Un valor muy bueno, pues el modelo de dos dimensiones ha permitido explicar el 74.1% de los datos. Finalmente, gracias a la interpretación gráfica de este análisis, se podrá conocer la cercanía o lejanía de las categorías de ambas variables cualitativas.

En el ejemplo que se muestra en la Figura 25, se puede observar que el alumnado de Educación Social está más asociado al Rugby que a cualquier otro deporte; o por ejemplo, se observa como el alumnado de Física y Química tienen cierta preferencia al Atletismo.

En *SPSS Statistics*, para poder realizar este tipo de análisis se debe preparar correctamente una base de datos específica con 3 variables definidas:

- La variable categórica 1: En nuestro caso puede ser los distintos grados universitarios, aunque dependiendo del estudio variará.
- La variable categórica 2: En nuestro caso puede ser los distintos deportes, aunque dependiendo del estudio variará.
- La frecuencia: Es necesario crear una tercera variable que recoja la frecuencia para cada posible cruce de las variables categóricas. En nuestro caso es necesario indicar la frecuencia de estudiantes de Educación Primaria que prefieren cada deporte, la frecuencia de estudiantes de Educación Social que prefieren cada deporte, y así con todas las posibilidades posibles, quedando una base de datos similar a la que se muestra en la Figura 26.

	Grado	Deporte	frecuencia
1	Ed. Primaria	Fútbol	39
2	Ed. Primaria	Baloncesto	40
3	Ed. Primaria	Tenis	20
4	Ed. Primaria	Rugby	26
5	Ed. Primaria	Ping-Pong	27
6	Ed. Primaria	Atletismo	4
7	Ed. Social	Fútbol	35
8	Ed. Social	Baloncesto	25
9	Ed. Social	Tenis	13
10	Ed. Social	Rugby	55
11	Ed. Social	Ping-Pong	19
12	Ed. Social	Atletismo	5
13	Ed. Infantil	Fútbol	28
14	Ed. Infantil	Baloncesto	24
15	Ed. Infantil	Tenis	7
16	Ed. Infantil	Rugby	14
17	Ed. Infantil	Ping-Pong	17
18	Ed. Infantil	Atletismo	3
19	Periodismo	Fútbol	13
20	Periodismo	Baloncesto	7
21	Periodismo	Tenis	5
22	Periodismo	Rugby	5

Figura 26. Base de datos para formar un análisis de correspondencias.
Fuente: elaboración propia.

Una condición importante para hacer este análisis es que los datos tienen que ser ponderados. Para ello accederemos a *Datos* > *Ponderar Datos*, y aquí ponderaremos los casos según la variable Frecuencia. Posteriormente, para realizar el análisis de correspondencia lo haremos a través de la siguiente ruta: *Analizar* > *Reducción de dimensiones* > *Análisis de correspondencias*. Moveremos las dos variables

categóricas, una a las filas y otra a las columnas y definiremos los rangos (número de categorías de cada variable).

CAPÍTULO VII: LOS ANÁLISIS MULTIVARIADOS

7.1. Introducción

En el capítulo VI hemos observado las posibles operaciones a las que podemos someter dos variables. Sin embargo, los análisis bivariados nos limitan en que únicamente nos permiten jugar con dos variables. Es por ello, que otro grupo de pruebas tratan de juntar en un mismo análisis más de dos variables. Se trata de los análisis multivariados.

A lo largo de este capítulo, vamos a hacer una revisión de aquellos tipos de análisis que podrían resultar más de interés para una persona que comienza a introducirse en el mundo del análisis de datos, sin olvidarnos que la clasificación de pruebas multivariantes se extiende mucho más que los análisis que se comentan.

Más concretamente, en este capítulo nos centraremos en 5 tipos de análisis multivariantes: El análisis factorial, el análisis clúster, el análisis discriminante, el análisis de regresión lineal y el análisis de segmentación.

7.2. Análisis factorial (exploratorio)

Una de las pruebas más comunes para conocer las diferentes partes (dimensiones) de las que está compuesto un instrumento es el análisis factorial.

Según el modelo *The Big Five*, por ejemplo, sabemos que la personalidad como tal, engloba diferentes dimensiones: la inestabilidad emocional, la responsabilidad, la apertura a la experiencia, la extraversión y la amabilidad.

Imagínate que tenemos un instrumento de 40 ítems para medir la personalidad de los sujetos. Si realmente el modelo fuese cierto, el análisis factorial nos ayudaría a conocer qué ítem pertenece a qué categoría o dimensión, clasificando así todos y cada uno de los ítems en los grupos que le correspondiese. Es decir, un ítem que mide la responsabilidad se juntaría con el resto de los ítems que miden la responsabilidad, un ítem de amabilidad se juntaría con el resto de los ítems de amabilidad y así sucesivamente.

Tabla 34. Ejemplo de Tabla de correlaciones para análisis factorial.

	i01	i02	i03	i04	i05	i06
i01	1.00	**0.96$_1$**	0.03	-0.12	**0.89$_1$**	-0.12
i02	-	1.00	-0.10	0.05	**0.92$_1$**	0.16
i03	-	-	1.00	**0.87$_2$**	0.21	**0.94$_2$**

	i01	i02	i03	i04	i05	i06
i04	-	-	-	1.00	0.08	**0.85$_2$**
i05	-	-	-	-	1.00	0.12
i06	-	-	-	-	-	1.00

Fuente: elaboración propia.

Para conocer a qué grupo pertenece cada ítem se aplica el coeficiente de correlación de Pearson entre todas las diferentes posibilidades: El ítem 1 con el ítem 2, el ítem 1 con el ítem 3… el ítem 2 con el ítem 1, el ítem 2 con el ítem 2… así hasta acabar con todas las correlaciones. Aplicándolo con números lo veríamos como aparece en la Tabla 34.

Recordamos que 1 es la correlación máxima. Las Tablas aparecen cortadas a la mitad debido a que la segunda mitad es repetición de la primera, pero a la inversa por lo que no aporta más información.

En la Tabla 34 se han destacado a negrita las correlaciones más altas, y se pueden observar cómo han salido dos grupos: Por una parte, el primer grupo queda constituido los ítems i01, i02 e i05, pues has correlacionado muy bien entre ellos, pero estos ítems no han correlacionado con el resto. Por otra parte, el segundo grupo queda constituido por los ítems i03, i04 e i06, como se puede observar en la Figura 27.

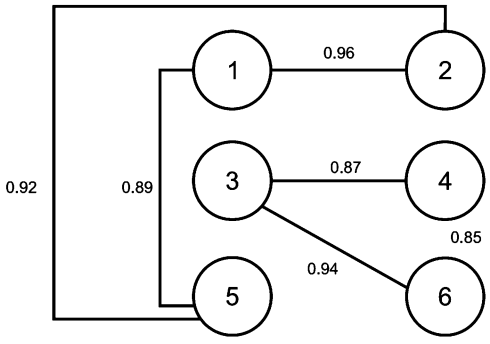


Figura 27. Ejemplo de clasificación de los ítems según sus correlaciones.
Fuente: elaboración propia.

Aunque este ha sido un mero ejemplo, en la práctica veremos que hay que tener cuidado con ítems que correlacionan entre sí muy alto, pues podemos estar corriendo el riesgo de que realmente ambos ítems estén midiendo lo mismo.

Cabe recordar que a cada uno de los grupos que salen del análisis factorial se le denomina dimensión o factor. En este caso existirían dos factores o dimensiones.

Una peculiaridad del análisis factorial es que únicamente se puede emplear con variables medidas en escala continua o con variables de muchas categorías; 12 categorías o más (Comrey, 1985). Además, es recomendable que se aplique únicamente en muestras con más de 100 casos o mínimo 5 veces el número de variables (Hair, 1999).

Para acceder en *SPSS Statistics* al análisis factorial vamos a *Analizar > Reducción de dimensiones > Factor.*

Antes de nada, debemos conocer si nuestros datos son apropiados para poder realizar un análisis factorial. Entre los procedimientos más comunes para conocer su pertinencia se encuentra la **prueba de Kaiser-Meyer-Olkin** (KMO) y la **prueba de esfericidad de Bartlett**. Estas pruebas se encuentran en esa misma ruta, en el botón de *Estadísticos*.

Tras pasar todas las variables del instrumento a la columna de variables, obtendremos una Tabla similar a la Tabla 35.

Tabla 35. Prueba KMO y esfericidad de Bartlett.

Medida KMO de adecuación de muestreo		.799
Prueba de esfericidad de Bartlett	**Aprox. Chi-Cuadrado**	7987.750
	gl	861
	Sig.	.000

Fuente: elaboración propia.

La confirmación para poder realizar un análisis factorial vendrá asignado a un valor de KMO por encima de 0.7, y un valor estadísticamente significativo (<0.05 o el nivel de significancia definido). En este caso KMO está por encima de 0.7 y la esfericidad de Bartlett es significativa, por lo que aceptamos que podemos realizar el análisis factorial.

Posteriormente, obtendremos una Tabla con la cantidad de factores o dimensiones que se han encontrado en el análisis. La Tabla será similar a la que se muestra en la Tabla 36.

Tabla 36. Varianza total explicada de un análisis factorial.

Componente	Total	% de varianza	% acumulado
1	5.052	31.574	31.574
2	1.948	12.177	43.751
3	1.710	10.689	54.440
4	1.300	8.128	62.568
5	.0926	5.789	68.357

Componente	Total	% de varianza	% acumulado
6	.889	5.554	73.911
...	...	...	...
16	.183	1.142	100

Fuente: elaboración propia.

En esta Tabla se observa todos los posibles factores, a lo que llama componente. El máximo de factores es el total de ítems introducidos, siendo 1 factor para cada ítem. No obstante, el problema que tiene asignar un factor a cada ítem es que no vamos a poder formar grupos de ítems, y por lo tanto resulta inútil realizar un análisis factorial si no es para agrupar.

Uno de los criterios más empleados para conocer cuántos factores (componentes) hay que escoger es el **criterio de la raíz latente**. Este criterio nos indica que debemos escoger todos los factores que tengan un autovalor (Lo que en *SPSS Statistics* corresponde a la columna de '*Total*') superior a 1. En el caso de la Tabla 36, observamos que hay 4 factores por encima del 1, por lo que esos serán los factores que cojamos.

Sin embargo, si deseas cambiar el valor 1 por otro valor o incluso indicarle a *SPSS Statistics* si deseas obtener un número previamente determinado de factores, lo puedes realizar desde la misma ruta del análisis factorial, en el botón *Extracción*.

Otro detalle significativo de la Tabla 36, es que observamos que estos 4 factores explican el 62.568 % de la varianza (El factor 1 es el más importante porque explica el 31.574% de la varianza). Es decir, que, si estuviésemos midiendo la ansiedad hacia las matemáticas, por ejemplo, a través de este instrumento estaríamos midiendo la ansiedad hacia las matemáticas en un 62.568%. El resto hasta 100% son factores que se nos escapan y que no hemos analizado.

Seguidamente, otra Tabla que obtendremos del análisis factorial será la Tabla de matriz de componente, definida en la Tabla 37.

Esta Tabla nos indica qué ítem va con qué factor. Aquellas variables que correlacionen más alto con el factor será un claro indicio de que pertenecen a ese grupo. Por ejemplo, el ítem i01 tiene una correlación de .620 con el factor 1, una correlación de .021 con el factor 2, una correlación de .201 con el factor 3, y una correlación de .110 con el factor 4, por lo que el ítem i01 pertenece al factor 1.

Tabla 37. Matriz de componente de un análisis factorial.

	1	2	3	4
i01	**.620**	.021	.201	.110
I02	**.740**	.102	.086	.067
i03	**.820**	.021	.201	.001
i04	**.910**	.031	.022	.017
i05	.310	**.821**	.203	.288
i06	.291	**.943**	.194	.195
i07	.172	**.786**	.042	.016
i08	.031	**.680**	.105	.264
i09	.029	.210	**.787**	.195
i10	.111	.194	**.888**	.123
i11	.219	.215	**.934**	.301
i12	.101	.054	**.796**	.105
i13	.210	.105	.103	**.895**
i14	.043	.013	.017	**.846**
i15	.304	.210	.192	**.921**
i16	.106	.017	.160	**.944**

Fuente: elaboración propia.

En este ejemplo, está muy claro observar qué ítem va con qué factor. No obstante, luego en la práctica hay ocasiones que puede resultar más complejo conocer esto mismo, pues suele darse el caso en que un mismo ítem puede entrar en 2 grupos diferentes.

Con este objetivo se realiza una acción extra denominada **rotación**, que trata de realzar los valores de un ítem para un factor determinado y reducir el resto de las correlaciones para ese mismo ítem.

Existen diversos métodos de rotación (*Varimax*, *Quartimax*, *Equamax*, *Promax*…), accesibles a través de la misma ruta del análisis factorial, en el botón *Rotación*. El más común de estos métodos de rotación es el método *Varimax*, que trata de reducir el número de variables que tienen valores altos en un factor.

Llegados a este punto en el que ya están todas las variables asignadas a cada factor, el último paso sería únicamente asignarle un nombre a cada factor, preguntándonos ¿Qué es lo que están midiendo este conjunto de variables?

Para concluir este epígrafe, se quería recordar al lector que existen dos tipos de análisis factoriales: El **análisis factorial exploratorio**, que ha sido el explicado a lo

largo de estas líneas y que nos ha permitido crear grupos a partir de las variables importadas; y el **análisis factorial confirmatorio**.

7.3. Análisis clúster o conglomerados

El análisis clúster, también conocido como análisis de conglomerados, es una copia del análisis factorial. Es decir, el objetivo de ambos es el de crear grupos lo más homogéneos posible a partir de la información que proporcionemos.

La principal diferencia entre uno y otro estriba en que mientras que para el análisis factorial solamente podemos emplear variables continuas, en el caso del análisis clúster estas variables pueden ser de cualquier tipo. El análisis clúster es más subjetivo que el análisis factorial, pues la cantidad de grupos dependerá de cómo interpretemos nosotros los resultados.

A la hora de crear los grupos podemos encontrarnos con que ya tenemos en mente cuántos grupos queremos formar (**métodos no jerárquicos**), o, por el contrario, si veremos primero la información y luego elegiremos la cantidad de grupos (**métodos jerárquicos**).

Para acceder a los métodos no jerárquicos, iremos en *SPSS Statistics* a *Analizar* > *Clasificar* > *Clúster de K-Medias* (*K-Medias* es la técnica más común de los métodos no jerárquicos, aunque existen más técnicas como *Quick Cluster*, *Taxmap*, *Fortin*, *Block Clustering*, etc).

En esta ruta, introduciremos en la columna *Variables* los ítems que queramos agrupar, y en la casilla de *Número de clústeres*, introduciremos el número de grupos que queremos que se formen.

De este análisis, pondremos especial atención a la Tabla de *centros de clústeres finales*. Esta Tabla nos proporcionará información sobre cómo se han unido los clúster.

Tabla 38. Ejemplo de Tabla de centros de clústeres finales.

	Clúster		
	1	***2***	***3***
Puntuación talento expresivo	**33**	29	27
Puntuación talento inventivo	39	**44**	37
Puntuación lectoescritura	30	**35**	26
Puntuación orientación espacial	20	19	**22**
Puntuación habilidades manuales	38	39	**44**
Puntuación Habilidades artísticas	8	**13**	11

	Clúster		
	1	*2*	*3*
Puntuación habilidades tecnológicas	33	27	28

Fuente: elaboración propia.

En el ejemplo recogido en la Tabla 38, se realizó un análisis clúster de K-Medias en el que se trataron de clasificar en 3 clústeres las puntuaciones para una serie de talentos y habilidades.

En esta Tabla, se pueden observar las características y las puntuaciones medias para pertenecer a cada grupo: en el primer clúster, al tener las puntuaciones más altas, el talento expresivo se ha unido bastante bien con las habilidades tecnológicas por lo que la gente que entre aquí se caracterizaría por altos valores en estas dos habilidades; el segundo clúster se ha formado con el talento inventivo, la lectoescritura y las habilidades artísticas; y el clúster 3, ha quedado formado por las variables de orientación espacial y habilidades manuales.

Es cierto que puede resultar en ocasiones bastante complejo clasificar a ciertas personas, pues puede darse el caso de que haya gente que puntúe alto para pertenecer a un determinado clúster, pero bajo en otra habilidad de ese mismo clúster, haciendo más compleja su clasificación.

Según este análisis, el centro para pertenecer a cada clúster vendría dado para cada una de las puntuaciones indicadas a cada talento y habilidad. Es decir, una persona que ha obtenido exactamente la misma puntuación en todas las variables a la columna de Clúster 3, sería introducida automáticamente a este grupo.

Para observar cuántos sujetos pertenecen a cada clúster, se presenta una última Tabla, similar a la que se muestra en la Tabla 39.

Tabla 39. Ejemplo de Tabla de número de casos en cada clúster.

Clúster	**1**	204,000
	2	121,000
	3	155,000
Válidos		480,000
Perdidos		,000

Fuente: elaboración propia.

Yendo con el otro tipo de método, los métodos jerárquicos, recordamos que nos proporcionan primero la información del análisis clúster para posteriormente decidir cuántos grupos queremos formar. Este tipo de análisis se desarrolla en SPSS Statistics

a través de Analizar > Clasificar > Clúster jerárquico. En este caso, podemos realizar la agrupación, según nos interesa, agrupando variables, o agrupando sujetos (casos). Para ello, en el menú principal, en Clúster, seleccionaremos Casos o Variables.

Una vez confirmado el análisis, la primera Tabla que obtendremos es el historial de conglomeración, recogido en la Tabla 40. Para este ejemplo, cabe mencionar que se han introducido seis variables diferentes.

Tabla 40. Historial de conglomeración en un análisis clúster.

Etapa	Clúster Combinado		Coeficientes	Primera aparición del clúster de etapa		Etapa Siguiente
	Clúster 1	Clúster 2		Clúster 1	Clúster 2	
1	1	4	1467,542	0	0	4
2	2	3	3869,300	0	0	4
3	5	6	7000,111	0	0	5
4	1	2	13254,127	1	2	5
5	1	5	18783,008	4	3	0

Fuente: elaboración propia.

Aunque la interpretación de esta Tabla puede resultar algo complejo, la idea es que se parte de la Etapa 1, momento en el que se juntan las variables 1 y 4. En este punto se cuenta por lo tanto con cinco grupos, por una parte, el grupo 1, formado por las variables 1 y 4, y por otra parte, cada una de las variables que no se han unido (variable 2, 3, 4 y 6). No obstante, si no quedamos conformes con la clasificación hecha, podemos optar por dar un paso más y reorganizar las variables en el siguiente paso. Tras la etapa 1, observamos que la Etapa siguiente, viendo la última columna, es la etapa 4. Si procedemos a esta etapa, vemos que la variable 1 se ha unido con la variable 2, por lo que en este momento, tras 2 etapas, tendríamos 4 grupos, por una parte, el grupo 1, formado por las variables 1, 2 y 4 (Al estar el grupo anterior formado por las variables 1 y 4, se respeta este grupo), y por otra parte, las variables que no se han agrupado, que son las variables, 3, 5 y 6. Así podríamos seguir sucesivamente hasta donde el investigador considere oportuno.

En los análisis de conglomerados, es común interpretarlos de manera visual, a través de lo que se conoce como **dendrograma**. Un dendrograma es un gráfico bidimensional en el que en el eje y se muestran las variables o casos que hayamos introducido al análisis y en el eje x se muestra la distancia a la que se encuentra una variable de otra. Mientras esta distancia sea más corta, significará que la similitud entre ambas variables es más cercana.

Para observar el dendrograma a un análisis clúster es necesario ir a *Analizar* > *Clasificar* > *Clúster jerárquico*; y en la opción *Gráficos*, marcar la casilla *Dendrograma*.

En el ejemplo que se muestra a continuación, ilustrado en la Figura 28, se realizó un análisis clúster en el que se le pedía a 50 sujetos su opinión acerca del consumo de drogas al salir de fiesta y su rendimiento académico.

En este ejemplo, puede apreciarse como existen cuatro grupos principales, coloreados de distintas tonalidades. Es cierto, que es de destacar que existe un grupo con únicamente 2 casos (el sujeto 22 y el sujeto 25), pero hay que observar que la unión con el grupo situado a su izquierda está a una distancia muy lejana, por lo que se optó por mantener este grupo independiente.

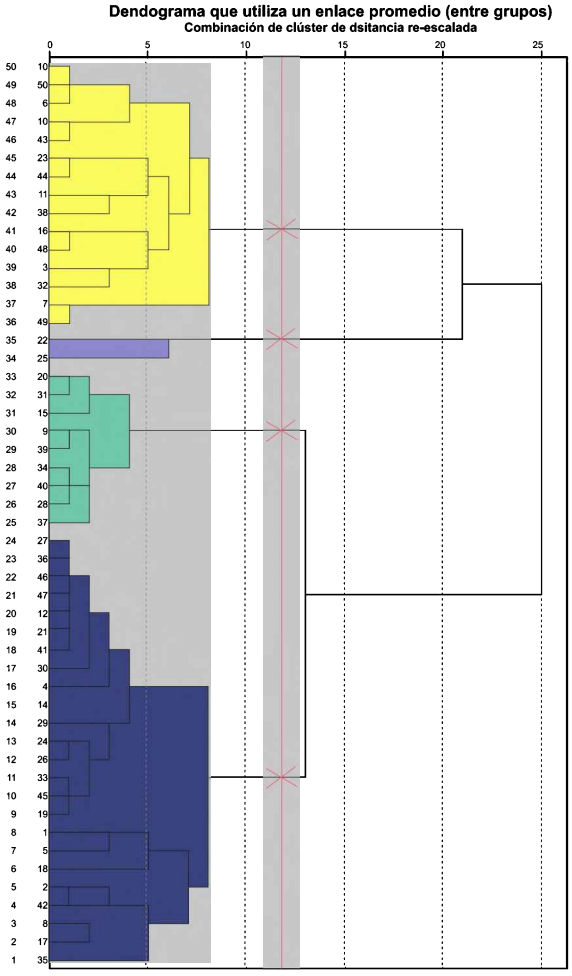


Figura 28. Ejemplo de dendrograma.
Fuente: elaboración propia.

Generalmente, el análisis clúster nos permite decir qué casos pertenecen a qué grupo, pero no nos da información sobre qué características tiene cada grupo. Para

ello es necesario complementar este análisis, con un análisis de segmentación. Este lo veremos más adelante.

A modo de orientación para el lector, podríamos sacar en conclusión tras realizar ambos análisis, clúster y segmentación, que los grupos poseen las siguientes características:

- Primer grupo (formado por los sujetos 22 y 25): Gente con rendimiento académico muy bajo y alta frecuencia de salir de fiesta (grupo minoritario).
- Segundo grupo (formado por los sujetos 13,50, 6, 10…): Gente con rendimiento académico medio, entre el aprobado y el suspenso. La mayoría suele salir de fiesta y algunos de ellos, además, fuman.
- Tercer Grupo (formado por los sujetos 27, 36, 46, 47…): Son personas con notas aceptables, que van desde el bien hasta el notable, y que nunca o rara vez salen de fiestas, además de nunca o rara vez fumar.
- Cuarto grupo (formado por los sujetos 20, 31, 15, 9, 39…): Son personas que tienen notas excelentes, por encima del 7,50 y nunca salen de fiesta.

7.4. Análisis discriminante

El análisis discriminante es un tipo de análisis explicativo que nos va a servir para conocer las características que diferencian a una cantidad específica de grupos; por ejemplo, conocer qué características diferencian a los estudiantes que aprueban selectividad, de aquellos que no aprueban selectividad.

Esto nos permitirá, además de conocer las causas que diferencian a los estudiantes que aprueban y no aprueban selectividad, realizar predicciones sobre quién va a ser más propenso a aprobar, antes de que ocurra el acontecimiento.

Para poder realizar un análisis discriminante es necesario tener una variable dependiente de cuantas categorías queramos y tantas variables independientes continuas como queramos.

En *SPSS Statistics*, para acceder al análisis discriminante iremos a *Analizar > Clasificar > Discriminante*. En *variable de agrupación* introduciremos nuestra variable dependiente (La que recoge la clasificación por grupos) y en *Independientes* las variables continuas que queramos estudiar.

Imaginemos que queremos predecir si un estudiante va a aprobar o no selectividad en función de sus notas obtenidas en matemáticas, lengua y física y química en bachiller.

En este ejemplo y volviendo a *SPSS Statistics*, en *variable de agrupación* introducimos nuestra variable dependiente, que sería una variable categórica en la que se recoge si aprueba o no selectividad. Por otra parte, en las *variables independientes* introducimos las calificaciones de los estudiantes en bachiller en las tres asignaturas mencionadas.

En este análisis será importante que en el botón *Estadísticos*, marquemos que queremos observar los *coeficientes de la función de Fisher*. En este análisis encontramos algunas Tablas importantes, que explicamos a continuación.

La primera Tabla importante que hay que mencionar es la Tabla de Autovalores, recogida en la Tabla 41. Esta Tabla, nos indica en resumen cuan de bueno es nuestro modelo para predecir si van a aprobar o no selectividad. Para conocer esta información, la Tabla nos aporta la correlación canónica, y aunque en un primer momento no nos diga mucho, para convertir este valor a porcentaje, lo que tenemos que hacer es elevarlo al cuadrado y luego pasarlo a forma porcentual multiplicándolo por 10. En nuestro caso, el modelo propuesto explica el 78.4% de la varianza (0.8862 x 100), lo cual es un valor muy bueno, pues es capaz de clasificar gran porcentaje de los datos únicamente con las 3 variables independientes que le hemos dado.

Tabla 41. Autovalores en un análisis discriminante.

Función	Autovalor	% de varianza	% acumulado	Correlación canónica
1	3,655	100,0	100,0	,886

Fuente: elaboración propia.

Una vez conocido esto, la siguiente Tabla es la de *Lambda de Wilks*. De esta Tabla, solamente es necesario comentar que nos indicará si la discriminación de los grupos del modelo es significativa o no.

Tabla 42. Lambda de Wilks en un análisis discriminante.

Prueba de funciones	Lambda de Wilks	Chi-cuadrado	gl	Sig.
1	,215	25,378	3	,000

Fuente: elaboración propia.

Lo que interesa es que los grupos (quienes aprueban y quienes no aprueban) estén lo más discriminados posible, para que no se mezclen. Para conocer si lo están o no, simplemente establecemos nuestra hipótesis nula, que recoge que no hay diferencias

entre los grupos, y nuestra hipótesis alterna, que recoge que sí hay diferencias entre los grupos. En este caso el p-valor (p = .000) es inferior al valor crítico (p = .005), por lo que podemos rechazar la hipótesis nula y confirmar la hipótesis alterna, comentando que los dos grupos existentes, de aprobados y suspendidos, están significativamente separados.

Posteriormente, la Tabla de *coeficientes de función discriminante canónica estandarizados*, recogida en la Tabla 43, nos aporta información sobre qué variables han sido más importantes a la hora de crear el modelo (función predictiva).

Tabla 43. Construcción de función predictiva.

	Función 1
mate_bachiller	,275
lengua_bachiller	,513
FyQ_bachiller	,394

Fuente: elaboración propia.

Estos valores, al estar estandarizados, van de 0 a 1, indicando que mientras más cercano esté a 1 más importante es. En este caso, se puede observar cómo las variables de mayor a menor importancia para predecir si alguien aprueba o no selectividad serían: En primer lugar lengua, posteriormente, física y química, y posteriormente matemáticas.

Finalmente, vamos con la Tabla de los *coeficientes de la función de clasificación de Fisher*. Esta Tabla probablemente sea una de las más importantes del análisis de discriminación pues nos proporciona información crucial para poder crear una fórmula matemática que nos ayude a clasificar a cualquier estudiante que queramos en las categorías establecidas. Esta Tabla tendrá una pinta similar a la que se muestra a continuación:

Tabla 44. Coeficientes de función de clasificación.

	Selectividad	
	Aprobado	Suspendido
mate_bachiller	1,788	1,037
lengua_bachiller	1,980	,590
FyQ_bachiller	2,172	1,137
(Constante)	-23,409	-5,631

Fuente: elaboración propia.

En vista de esta Tabla, tendremos que crear dos fórmulas: Una para aquellos que aprueban selectividad, y otra para aquellos que suspenden selectividad. La manera de crear cada fórmula será la siguiente:

*Constante + Variable1 * (Valorobtenidoendichavariable) + Variable 2 * (Valorobtenidoendichavariable) + Variable 3 * (Valorobtenidoendichavariable)*

En nuestro caso, siguiendo este patrón, la fórmula para predecir si alguien aprueba o no selectividad quedará del siguiente modo:

- Aprobado =-23.409 + 1.788 * NotaMatemáticas + 1.980 * NotaLengua + 2.172 * NotaFisyQuim.
- Suspendido = -5.631 + 1.037 * NotaMatemáticas + 0.590 * NotaLengua + 1.137 * NotaFisyQuim.

Para clasificar a un estudiante tendremos que probar sus calificaciones en ambas fórmulas, y observar en cuál de las dos obtiene un valor superior. En aquella fórmula que haya obtenido un valor superior, ahí se clasificará.

Imaginemos que un estudiante ha sacado un 5.2 en matemáticas, un 4.95 en lengua y un 6.20 en física y química. En este caso, para la fórmula de aprobado su puntuación será de 9.15 y para la fórmula de suspendido su puntuación será de 9.73. Por lo que según esta fórmula, al ser la puntuación de suspendido, mayor a la de aprobado, es bastante posible que suspenda.

7.5. Análisis de regresión lineal

El análisis de regresión lineal es un tipo de análisis en el que se trata de estudiar la relación y la capacidad de predicción entre una o unas variables independientes continuas con otra variable dependiente continua. Pongamos un ejemplo para ver esto más claro.

Imagínate que estamos tratando de conocer qué nota media de selectividad (del 1 al 10) sacarán los estudiantes en función de su autoconcepto académico, social, físico y autoestima personal (Medidos del 1 al 5). Es decir, queremos saber cómo la nota de selectividad está influenciada y puede venir predicha por distintas dimensiones del autoconcepto. En este caso, el problema respondería a la siguiente cuestión.

Tabla 45. Ejemplo de variables dependiente e independiente en análisis de regresión.

Variable Dependiente		Variables Independientes
Nota de Selectividad	=	A. académico + A. Social + A. Físico + A. Personal

Fuente: elaboración propia.

Existen dos tipos principales de análisis de regresión lineal, el **análisis de regresión simple** y el **análisis de regresión** múltiple. La diferencia entre ambos es que en el

análisis de regresión lineal simple únicamente se emplea una variable independiente para crear la ecuación y en análisis de regresión lineal múltiple se emplean dos o más variables independientes. El caso del ejemplo de arriba se trataría de un análisis de regresión lineal múltiple debido a que cuenta con 3 variables independientes.

Condición indispensable para poder aplicar un análisis de regresión, al igual que con las pruebas paramétricas que vimos en el capítulo anterior, es que las variables sigan una distribución normal y que exista homocedasticidad.

En caso de que existan ambas condiciones, en *SPSS Statistics* iremos a *Analizar > Regresión > Lineales. En Dependientes* introducimos nuestra variable dependiente y en *Independientes* introducimos nuestra o nuestras variables independientes.

Antes de continuar vamos a definir dos de los métodos de introducción de variables más comunes, casilla que se sitúa debajo de donde introducimos las variables independientes. De manera predeterminada viene el método *Intro*. Si dejamos elegido este método, la manera de crear la ecuación va a ser introducir a la vez todas las variables, es decir, nos va a dar toda la información de golpe y nosotros vamos a tener que decir qué variables formarán parte de la ecuación y qué variables no formarán parte de la ecuación. Si tienes dudas sobre qué método introducir es recomendable dejar este como predeterminado.

El otro método es el *método por pasos* o *stepwise*. Este método va introduciendo las variables a la ecuación de la más importante a la menos importante, y aquellas que considere que no son aceptables, las rechaza de la ecuación.

En ambos casos, cuando realicemos el análisis de regresión obtendremos unas Tablas muy similares. La primera de ellas será el resumen del modelo, que tendrá una apariencia parecida a la Tabla 46.

Tabla 46. Resumen del modelo de un análisis de regresión lineal.

Modelo	R	R cuadrado	R Cuadrado Ajustado	Error estándar de la estimación
1	.561	.315	.311	.7587

Fuente: elaboración propia.

De esta Tabla el valor más importante es el valor de **R^2** (Se le suele llamar también **coeficiente de determinación**) o en su defecto el **R^2 Ajustado**, el que prefiramos. Este valor nos dirá cuán de correcto es el modelo para poder predecir lo que queremos predecir. En pocas palabras, es un indicador de la calidad del modelo (ecuación).

Dependiendo de en qué ámbito estemos trabajando, los valores de R2 pueden verse enormemente afectados. En esta línea, Anderson, Sweeney y Williams (2001)

matizan que en el caso de las ciencias sociales se considerarían útiles valores de R2 por encima de 0.25, en las ciencias naturales valores entre 0.60 y 0.90 y en el caso de la aplicación de negocios los valores de R2 son muy variantes. En este caso, el modelo tiene un valor R2 de .315, o lo que es lo mismo el modelo puede explicar el 31,5 % de la variación de los resultados. En Ciencias Sociales es difícil, por no decir imposible, encontrar modelos de regresión lineales con unos valores de R^2 superiores al 80%, por lo que, por lo general, un modelo de estas cualidades ya nos ayudaría en cierto modo a poder encontrar alguna correlación interesante. Un indicio para conocer si nuestro modelo es válido o no es la información que proporciona la Tabla de la *ANOVA* (Véase Tabla 47).

Tabla 47. Prueba ANOVA de un análisis de regresión lineal.

Modelo		Suma de Cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	171.711	4	42.928	74.570	.000
	Residuo	373.609	649	.576		
	Total	545.320	653			

Fuente: elaboración propia.

De la Tabla de *ANOVA*, nos fijaremos que el modelo sea significativo por debajo de nuestro nivel de significancia admitido (generalmente .05). En este caso, al serlo, es una pista que nos confirma que las variables que incluye el modelo han pasado el filtro de calidad.

En último lugar, tal vez la Tabla más importante, que es la que nos va a permitir construir el modelo, es la Tabla de coeficientes, que será muy similar a la que se recoge en la Tabla 48.

En esta Tabla, antes de nada, tenemos que elegir aquellas variables importantes; las que son capaces de predecir la nota de selectividad, y rechazar aquellas variables que no son importantes en este caso.

Tabla 48. Coeficientes de un análisis de regresión lineal.

Modelo		Coeficientes no estandarizados		Coeficientes Estandarizados	t	Sig.
		β	Error estándar	Beta		
1	(Constante)	5.851	.271		21.632	.000
	Aut. Físico	-.110	.057	-.068	-1.908	.057
	Aut. Personal	-.076	.051	-.054	-1.494	.136
	Aut. Académico	.755	.044	.577	17.263	.000
	Aut. Social	-.114	.059	-.070	-1.939	.053

Fuente: elaboración propia.

Para ello primero nos fijamos en la columna de significancia. Lo interesante para nosotros es seleccionar aquellas variables que sean significativas (<.05) para el modelo. En este caso, la única variable significativa para el modelo va a ser el autoconcepto académico, pues el resto de las variables obtuvieron puntuaciones superiores a nuestro valor de significancia y por lo tanto deben ser rechazadas.

Posteriormente, nos queda formar el modelo, para el cual contamos con una constante (siempre va a ser significante) y una variable independiente (autoconcepto académico). La manera de formar la ecuación va a ser la siguiente:

$$Nota\ de\ Selectividad = Constante + (\beta\ Aut.\ Académico * Puntuación\ Aut.\ Académico)$$

O lo que es lo mismo:

$$Nota\ de\ Selectividad = 5.851 + (.755 * Puntuación\ Aut.\ Académico)$$

Y listo, este sería nuestro modelo de regresión lineal. Mediante este modelo, se estima que una persona que tenga un autoconcepto académico de 4.5 sobre 5 podría sacar en selectividad la siguiente nota:

$$Nota\ de\ Selectividad = 5.851 + (.755 * 4.5) = 9.24$$

Comenzábamos el epígrafe haciendo referencia a que las variables independientes únicamente podían ser variables continuas, y aunque esto es cierto, podría ser de ayuda hacer alusión a las **variables dummy**. Las variables dummy o variables ficticias son un método de transformación de variables en el que una variable ordinal o nominal se convierte en una variable continua. Por ejemplo, el género es una variable nominal formada por 2 categorías: hombres y mujeres; no obstante, esta se puede transformar en una variable continua asignándole el valor 0 a mujeres y 1 a hombres o viceversa. De este modo, variables que no podrían entrar a formar parte de un análisis de regresión, sería posible incluirlas.

Para poder realizar esta acción tendremos que volver a codificar las variables que queramos modificar yendo a *Transformar > Recodificar* en las mismas variables. En la nueva ventana, introducimos en el cuadro *Variables de cadena* todas las variables que queramos modificar y pulsamos el botón *Valores antiguos y nuevos.* Una vez llegado aquí, en el valor antiguo introduciremos cada categoría por separado (Por ejemplo, si h es hombre y m mujer, introduciremos primero h y luego m) y en el valor nuevo le asignamos un número (Se suele empezar por el 0).

Tras realizar este procedimiento, ya tendremos nuestra variable dummy preparada para introducir al análisis de regresión lineal.

7.6. Análisis de segmentación

El **análisis de segmentación** es un tipo de análisis estadístico que divide la muestra total en grupos homogéneos entre los casos de cada grupo y heterogéneos con el resto de los grupos utilizando un proceso secuencial descendente. Este análisis en *SPSS Statistics* se lleva a cabo principalmente con la ayuda de un algoritmo conocido como **CHAID** (Chi-Squared Automatic Interaction Detection).

El análisis de segmentación posee la característica de que puede realizarse con cualquier tipo de variable, cualitativa como cuantitativa y permite crear tantos grupos como salgan o por el contrario, definir los niveles de antemano nosotros.

Pongamos por ejemplo que queremos conocer cómo clasificar al alumnado de un colegio en función de su rendimiento académico total, teniendo en cuenta únicamente su edad, género, curso y asignatura favorita.

Tras realizar el análisis de segmentación obtenemos un árbol similar al que se muestra a continuación.

Es de comentar, que el orden de aparición de las variables, en este tipo de análisis, sí que es importante, siendo necesario considerar como variables más importantes las que están más arriba, y variables menos importantes las que están más abajo.

A cada variable le acompaña unos datos básicos como son la media y la desviación típica de ese grupo, así como el número de casos y el porcentaje que representa del total cada grupo.

En el ejemplo que se muestra en la parte inferior, se puede observar cómo se han creado un total de 5 grupos para clasificar al alumnado en función de su rendimiento académico.

En el cuadro superior, se recogen todos los casos, 822 casos. En la primera división, la variable más significativa es el curso ($p = .000$). Esta variable divide inicialmente al alumnado en aquellos que cursan 1º y 2º por una parte (308 casos; 37.5%), y mayores de 2º por otra parte (514 casos; 62.5%). Este segundo grupo, es un grupo definitivo, pues no subyacen más divisiones al mismo. No obstante, entre los estudiantes de 1º y 2º aún continúa la división en nuevos subgrupos.

Cogiendo en este momento, únicamente al alumnado de 1.º y 2.º, se puede apreciar como la siguiente variable significativa (p = .013) para realizar divisiones grupales es la asignatura favorita, surgiendo de esta división 3 nuevos grupos: aquellos que tienen de asignatura favorita matemáticas, informática o educación física (107 casos; 13%); aquellos que tienen de asignatura favorita lengua castellana, ciencias sociales u otras lenguas (118 casos; 14.4%); y aquellos que tienen de asignatura favorita las ciencias naturales (83; 10.1%). De estos 3 grupos, 2 de ellos son definitivos, pues no presentan nuevas subdivisiones. No obstante, el primer grupo, el que tiene como asignatura favorita las matemáticas, la informática o la educación física, posee una nueva variable significativa: el género (p = .047). Esta nueva subdivisión crea dos nuevos grupos: El grupo de chicos (54 casos; 6.6%) y el grupo de chicas (53 casos; 6.4%). En este caso el género se interpreta en el gráfico a través de las categorías 1 para chicos y 2 para chicas.

En total, si quisiésemos clasificar al alumnado en función de su rendimiento académico total tendríamos los siguientes grupos:

Alumnado mayor al 3.º curso.

Alumnado de 2.º curso o menor, que tiene como asignatura favorita las matemáticas, la informática o la educación física y es chico.

Alumnado de 2.º curso o menor, que tiene como asignatura favorita las matemáticas, la informática o la educación física y es chica.

Alumnado de 2.º curso o menor, que tiene como asignatura favorita las ciencias naturales.

Alumnado de 2.º curso o menor, que tiene como asignatura favorita la lengua castellana, las ciencias sociales u otras lenguas.

En *SPSS Statistics* este análisis se realiza a través de Analizar > Clasificar > Árbol. En la casilla de variable dependiente introducimos cuál es la variable que creemos que podemos dividir, y en las variables independientes, introducimos las variables que creemos que han podido tener influencia para que la variable dependiente varíe en un sentido u otro.

Cabe señalar, que depende de las variables que introduzcamos, el *SPSS Statistics* nos indicará una Tabla llamada Clasificación. Esta Tabla nos indicará un valor llamado Porcentaje correcto, que no es otra cosa que el porcentaje de casos correctos que ha conseguido clasificar a través del algoritmo empleado.

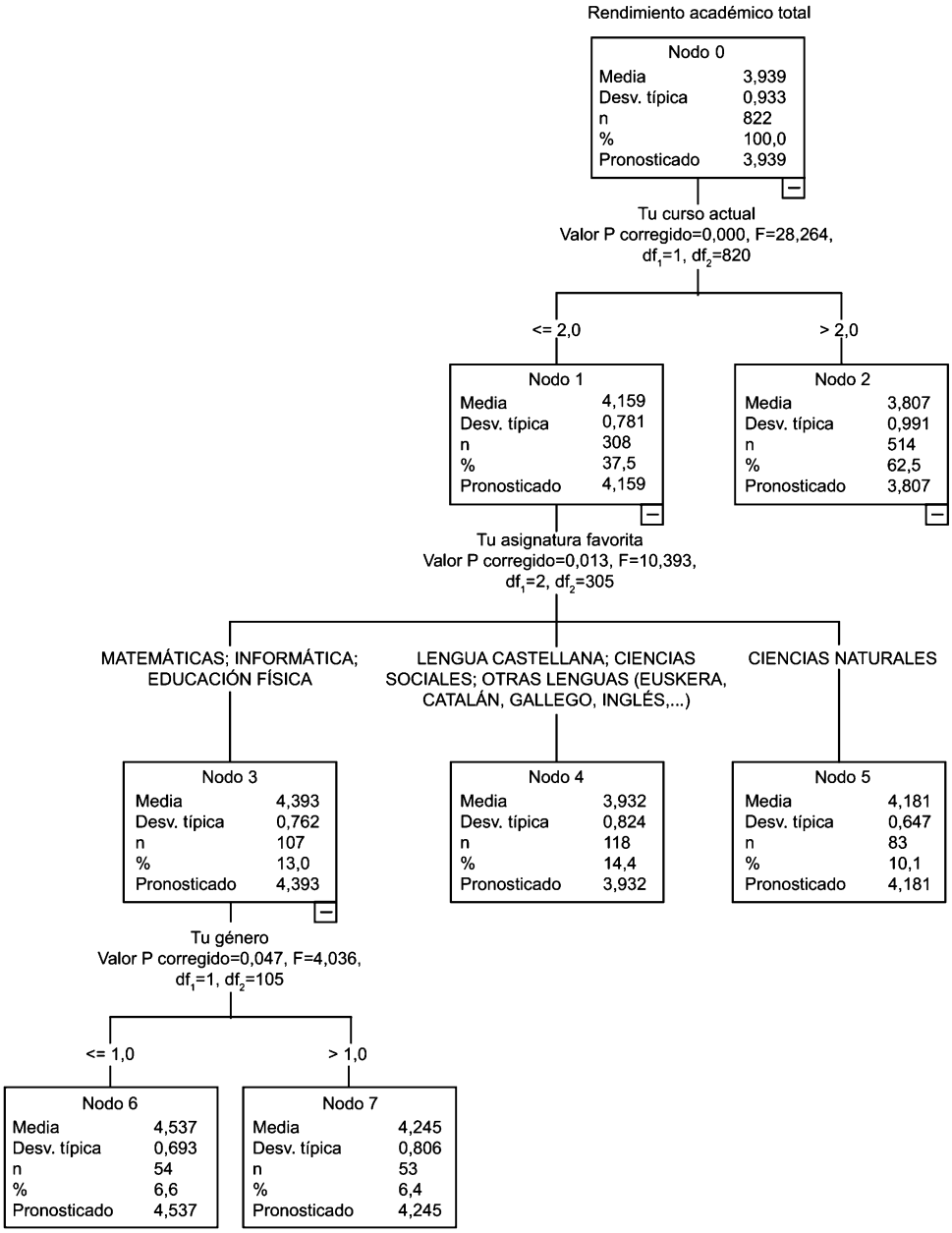


Figura 29. Ejemplo de análisis de segmentación.
Fuente: elaboración propia.

CAPÍTULO VIII: INTRODUCCIÓN A OTROS ANÁLISIS MULTIVARIANTES

8.1. Introducción

Llegados a este punto, los análisis mostrados hasta el momento son más que suficientes para que cualquier persona, estudiante o profesor, pueda llevar a cabo los análisis cuantitativos más comúnmente empleados en los estudios científicos del área de las ciencias sociales principalmente. Estos análisis, aunque pueda existir alguno más complejo que otro, en general, son análisis relativamente usados y de una complejidad investigadora media.

No obstante, una vez interiorizados los procedimientos para realizar los análisis mencionados hasta el momento, se presenta a continuación una tirada de nuevos tipos de análisis, que a pesar de requerir un conocimiento fecundo en investigación cuantitativa, pueden servir a modo introductorio.

En este capítulo, cobran especial interés los análisis factoriales confirmatorios y los modelos de ecuaciones estructurales al ser unos análisis que se realizan con programas independientes al que hemos trabajado hasta el momento, y que además guardan cierta importancia a la hora de conocer cuán de bueno o malo ha podido ser un modelo, siendo unos análisis de los más comunes a la hora de validar escalas y cuestionarios.

8.2. El AFC y el SEM

8.2.1. Definición de ambos procedimientos

A diferencia del **análisis factorial exploratorio** (AFE), que partiendo de unas variables tratamos de observar cómo se agrupan, con el **análisis factorial confirmatorio** (AFC) pretendemos evaluar lo buen o malo que es el modelo que ha resultado de ese análisis factorial exploratorio.

Respecto a los **modelos de ecuaciones estructurales** (del inglés Structural Equation Model, SEM) podemos definirlos como un modo más avanzado de realizar modelos de regresión, como los estudiados hasta el momento.

La principal diferencia entre estos dos tipos de procedimientos estriba en la causalidad de las variables. Mientras que en los AFC las dimensiones (variables observadas, como veremos algo más abajo) simplemente covarían o correlacionan unas con otras, en el caso, de los SEM se espera que sean modelos más complejos donde una

dimensión prediga otra. Es decir, en este último caso, habrá algunas dimensiones que funcionarán como variable dependiente y otras como variable independiente. Con esto no queremos decir que en los AFC no haya ningún rastro de análisis de regresión, pues como veremos a continuación, sí que existen, lo que ocurre que este se da desde los ítems (variables observadas) hacia las dimensiones (**variables latente**).

8.2.2. Convenciones del AFC y del SEM

Innato a la psicometría de cualquier instrumento, y como ya hemos visto a lo largo de este trabajo, los ítems se agrupan en dimensiones. Es decir, estos ítems son los que hacen que se conformen una dimensión. Pero a su vez, la unión de varias dimensiones, que se relacionan teórica o prácticamente entre sí, hace que se forme un constructo. Véase que vamos de lo simple, lo que se puede observar y medir, como es el caso de los ítems, a lo complejo y abstracto, como es el caso de un constructo psicológico.

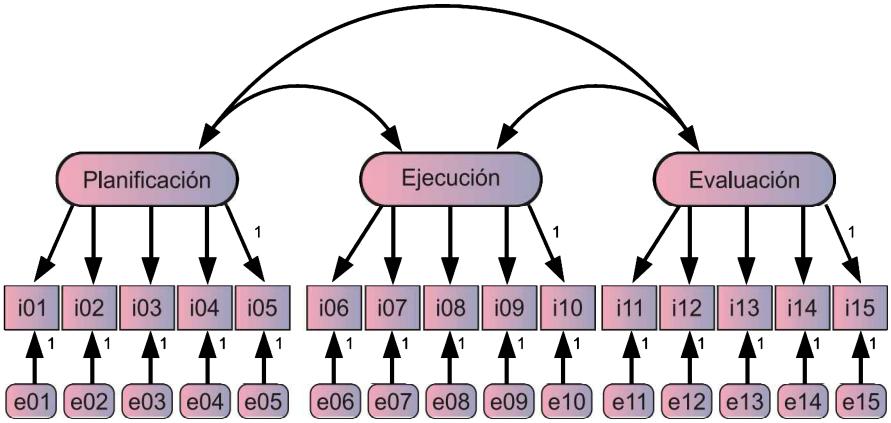


Figura 30. Ejemplo de un modelo teórico, apto para realizar un AFC.
Fuente: elaboración propia.

En la Figura 30 se muestra un ejemplo inventado de un modelo propuesto para realizar un AFC. Este modelo tiene como fin evaluar la calidad del aprendizaje de los docentes en diferentes momentos de su práctica. Se piensa, por el análisis factorial exploratorio previo y por diferentes teorías que la calidad puede ser evaluada en tres momentos o dimensiones diferentes: Planificación, ejecución y evaluación. Puede observarse como cada dimensión a su vez, está formada por 5 ítems.

En este ejemplo hay diferentes reglas necesarias de explicar. Por una parte, podemos darnos cuenta de que cada parte está representada por diferentes formas. Tanto en los AFC, como en los SEM, a los ítems de un cuestionario se les conoce como variables observadas, pues son indicadores que miden a los sujetos. La convención

para este tipo de variables es la de representarlas encerradas en un cuadrado o rectángulo. Por otra parte, a las variables que no se pueden observar, se las conoce como variables latentes y se representan en los AFC y en los SEM a través de círculos u óvalos.

Otro aspecto que destacar es el de las flechas. En esta línea, las flechas con punta en los dos sentidos significan correlación y **covarianza**, dependiendo si entendemos los valores estandarizados o no. En el modelo arriba propuesto, en la Figura 30, se puede observar cómo las 3 diferentes dimensiones correlacionan entre ellas. Sin embargo, las flechas en un único sentido implican causalidad. Es decir, en el origen de la recta de la flecha se encuentra la **variable predictora** o independiente, y al final, en la punta de la flecha, se encuentra la variable dependiente. En el ejemplo ilustrado en la Figura 31, podríamos decir que el rendimiento académico, depende o es predicho por el clima de aula, la motivación y el autoconcepto del alumnado. Estas últimas tres variables funcionan como variables predictoras.

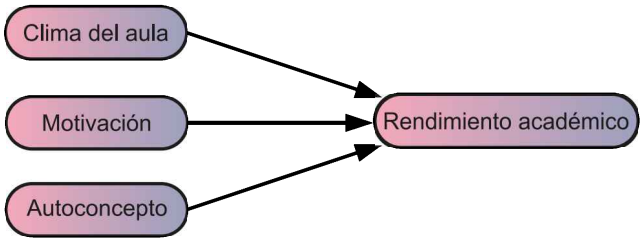


Figura 31. Ejemplo predictivo simple en un Modelo de Ecuaciones Estructurales.
Fuente: elaboración propia.

A las variables que afectan a otra variable y no recibe efecto de ninguna variable se les llama **variables exógenas**. En el ejemplo de arriba, el clima de aula, motivación y autoconcepto funcionarían como variables exógenas. Por el contrario, a las variables que reciben el efecto de otra variable, se las conoce como variable endógena. En el caso del ejemplo de arriba, el rendimiento académico funcionaría como **variable endógena**.

¿Pero qué pasa cuando esto se complica un poco más? Por ejemplo, en el caso de la Figura 32, podemos observar cómo existen dos variables, el autoconcepto académico y la motivación, que funciona como variable endógena y exógena a la vez. La interpretación podría ser la de: El autoconcepto académico, predice la motivación, pero a su vez, la motivación también predice el autoconcepto académico. A su vez, ambas variables son predictoras del rendimiento académico. Repetimos: Este es un simple ejemplo ilustrativo e inventado, y en la práctica este modelo probablemente

carecería de valor estadístico. Profundizaremos un poco más en este tipo de variables endógenas y exógenas en los análisis de mediación.

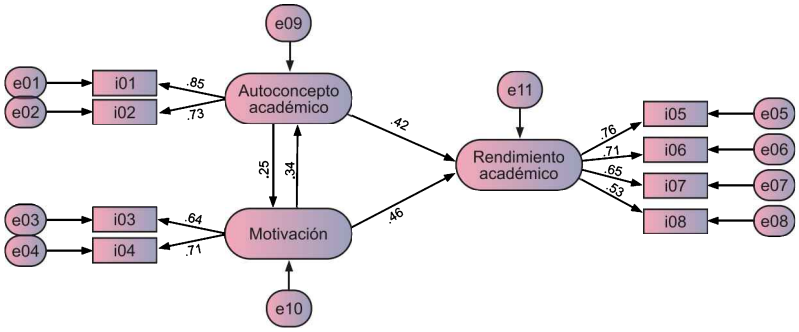


Figura 32. Ejemplo predictivo complejo en un Modelo de Ecuaciones Estructurales.
Fuente: elaboración propia.

Finalmente, tanto en los AFC, como en los SEM, se añade un nuevo tipo de variable conocida como **variable error**. La variable error recoge la probabilidad de error que se produce al realizar una predicción. Es decir, en la Figura 32, vemos como tanto cada ítem como cada variable endógena van asociadas a una variable error. Esto es así porque se espera que cuando tratamos de hacer una predicción, siempre haya una probabilidad de confundirse en aquello que estamos midiendo. Cada ítem y/o variable endógena va a tener un margen de error, más grande o pequeño, dependiendo de la calidad predictiva del ítem, que va a afectar a cómo el ítem observable contribuye a la variable no observada. Este error es necesario recogerlo en el modelo. Generalmente, vienen representadas por la letra e y recogidas en círculos, aunque la convención es representar los mismos sin círculos ni rectángulos (Ruiz, Pardo y San Martín, 2011). En los AFC únicamente se les pone variables error a los ítems, mientras que en los SEM, como el de la Figura 32, en el que existe causalidad entre las variables latente, las variables error aparecen tanto en los ítems como en las propias variables endógenas.

8.2.3. Evaluación de modelos AFC y SEM a través de la bondad de ajuste

Para conocer lo buen o mal modelo que se propone tanto en un AFC, como en un SEM, en estadística se emplea el término bondad de ajuste. Un buen modelo, podremos decir que tiene buena bondad de ajuste, y un mal modelo diremos que tiene mala bondad de ajuste. Este término hace referencia al ajuste que tiene con el modelo ideal. Es decir, si tenemos buena bondad de ajuste, nos referiremos a que el modelo propuesto se parece a grandes rasgos con el modelo perfecto que se podría hacer en cada caso. La evaluación de la bondad de ajuste debe ser un objetivo

primordial tanto de cualquier análisis factorial confirmatorio, como de cualquier modelo de ecuaciones estructurales.

Para conocer la bondad de ajuste de un modelo, se suele emplear una serie de estadísticos que comúnmente se agrupan en tres grandes categorías: Los estadísticos de ajuste absoluto, los estadísticos de ajuste incremental y los estadísticos de ajuste de parsimonia. Se explican a continuación.

8.2.3.1. Estadísticos de ajuste absoluto

Estos índices no usan un modelo alternativo como base para comparar lo buen o mal modelo que es, sino que más bien, solo observa en qué nivel se predice la matriz de correlaciones. Los índices de bondad de ajuste absoluto más comunes en este bloque son: **X^2/gl** (Chi-Cuadrado entre los grados de libertad), **SRMR** (Standardized Root Mean Square Residual), y **RMSEA** (Root Mean Square Error of Aproximation). Cada uno de estos índices tiene un punto de corte, que separa un modelo bueno de un modelo mediocre. Es por ello, que en la literatura se recomienda usar los siguientes puntos de corte para evaluar la bondad de ajuste absoluta: En el caso del X2/gl se considerará un buen modelo a aquellos modelos con valores por debajo de 5 (Wheaton, Muthen, Alwin y Summers, 1977), aunque este valor puede estar muy influenciado por el tamaño de la muestra (Bentler y Bonnet, 1980; Jöreskog y Sörbom, 1993). En el caso del SRMR se espera que sea un valor inferior a .05 para tener buena bondad de ajuste absoluto (Diamantopoulos y Siguaw, 2000). Finalmente, se espera que el RMSEA tenga un valor inferior a .08 para un buen ajuste, y valores entre .08 y .10 para un ajuste mediocre (MacCallum, Browne y Sugawara, 1996).

8.2.3.2. Estadísticos de ajuste incremental

Estos índices comparan el modelo propuesto con algún otro existente, llamado generalmente modelo nulo (Escobedo, Hernández, Estebané y Martínez, 2016). Los índices de bondad de ajuste absoluto más comunes en este bloque son: **NFI** (Normed Fit Index), **IFI** (Bollen's Incremental Fix Index), **TLI** (Tucker-Lewis Index) y **CFI** (Benter's Comparative Fit Index). Se espera que todos estos índices, para tener una buena bondad de ajuste incremental estén por encima de .90 (Hooper, Coughlan y Millen, 2008).

8.2.3.3. Estadísticos de ajuste de parsimonia

Estos índices tratan de conocer el punto óptimo sobre la complejidad del modelo. Modelos más complejos tendrán un peor ajuste de parsimonia que modelos teóricos más simples. En este grupo destaca el índice **AIC** (Akaike Information Criterion). Este índice es complejo de interpretar pues no se compara con ningún valor que se tenga

de antemano, sino que se compara entre varios modelos que hayamos construido nosotros, de modo que aquel modelo con un valor de AIC más cercano a 0 tendrá una mejor bondad de ajuste de parsimonia (Hooper, Coughlan y Millen, 2008).

8.2.4. Toma de decisiones en un AFC y SEM

Una vez evaluado lo bueno o mal modelo pueden darse dos casos: Puede que el modelo propuesto presente una buena bondad de ajuste inicial y que no sea necesario conocer información más profunda sobre dónde falla el modelo debido a que los estadísticos arriba mencionados están en un buen nivel; o por el contrario, puede darse el caso de que un modelo sea malo y se necesite reajustar por motivo de que los estadísticos están fuera de los rangos indicados.

En este segundo caso, existe un modo de conocer donde se hallan los fallos a través del procedimiento de los índices de modificación. El índice de modificación (*modification index, M.I.*) es un valor que nos va a ayudar a conocer qué variables encajan mal en donde están. Comúnmente, se espera que cada par de variables no tengan un índice de modificación superior a 10.0 (Cole, Motivala, Khanna, Lee, Paulus y Irwin, 2005).

En el ejemplo ilustrado en la Figura 33, se muestra un fragmento de un AFC. En este ejemplo, fijándonos en la columna de M.I. se puede sacar en claro varios aspectos.

			M.I.	Par Change
e30	<-->	Acad.	8,934	,035
e29	<-->	Acad.	19,343	-,045
e29	<-->	e30	5,413	-,053
e28	<-->	Famil	15,457	,034
e28	<-->	e30	8,859	,065
e28	<-->	e29	12,107	-,066
e27	<-->	Soc.	7,695	-,059
e27	<-->	Acad.	9,356	,040
e27	<-->	e30	48,724	,204
e27	<-->	e29	20,141	-,115
e26	<-->	Soc.	4,762	,050
e26	<-->	Acad.	5,681	-,033
e26	<-->	e30	7,598	-,086
e26	<-->	e29	104,806	,280

Figura 33. Ejemplo de índices de modificación en SPSS Statistics Amos 23.
Fuente: elaboración propia.

En primer lugar, se puede observar cómo los programas estadísticos muestran la relación del error para cada par de ítems (e29 <--> e28, por ejemplo), pero también

muestra la relación del error de un determinado ítem con cada dimensión (e29 <--> Acad, por ejemplo). Comentar en esta línea, que para evaluar qué ítems están bien y mal situados, únicamente nos fijaremos en los M.I. para cada par de ítems.

En segundo lugar, es necesario comentar que lo que se indica en el ejemplo, son los errores de cada ítem, y no los ítem en sí. Esto quiere decir, que si hemos decidido llamar "e01" al error que subyace al ítem "i01" no habrá dificultades en saber qué error va asociado a qué ítem, pero en el caso de que el error "e01" esté asociado al ítem "i1e29", las acciones de mejorar el modelo habrá que realizarlas sobre el ítem en cuestión y no sobre la variable error.

En tercer lugar, es importante saber que solamente tendremos que fijarnos en los índices de modificación de cada par de ítems que haya dentro de cada dimensión. Es decir, si un modelo bidimensiones, en el que la dimensión 1 está formada por los ítems 1 a 4, y la dimensión 2 está formada por los ítems 5 a 8, el índice de modificación e01<-->e05 tendremos que ignorarlo, pues cada variable pertenece a una dimensión diferente.

Por lo tanto, en el ejemplo que se recoge arriba, podemos concluir que existe un problema en los ítems 28<-->29 (M.I. = 12.10), 27<-->30 (M.I. = 48.72), 27<-->29 (M.I. 20.14) y 26<-->29 (M.I. = 104.80). Todas estas correlaciones están por encima de un índice de modificación por encima de 10, lo cual apunta a que ambos ítems han sido interpretados de manera similar. Mientras mayor sea el índice de modificación, se espera que cada par de ítems lo hayan interpretado los participantes de manera muy similar.

En este punto las acciones que se pueden llevar a cabo pueden ser muy variadas, desde establecer una correlación entre ambos ítems, hasta eliminar el ítem. Se recogen las principales decisiones a continuación.

8.2.4.1. Correlación de ambos ítems

Esta opción consiste en establecer una correlación entre ambos ítems, para indicar que ambos ítems están relacionados. Esta correlación se muestra a través de las flechas de doble punta y se relaciona a través de los errores de ambos ítems. En el ejemplo mostrado en la Figura 34, se observa cómo tras analizar los índices de modificación, los ítems 1, 3 y 6 covarían/correlacionan entre sí, por lo que se ha establecido la respectiva flecha.

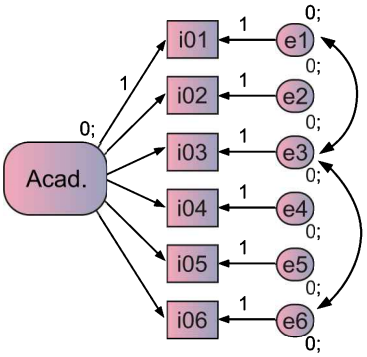


Figura 34. Ejemplo de covarianza de ítems en un AFC.
Fuente: elaboración propia.

8.2.4.2. Eliminar uno de los dos ítems

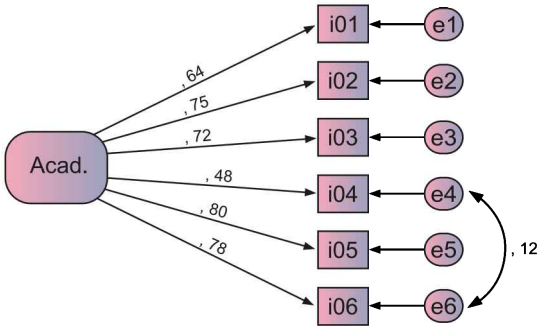


Figura 35. Ejemplo visual para la eliminación de un ítem en un AFC.
Fuente: elaboración propia.

Al ver que ambos ítems podrían estar midiendo cosas muy similares, puede optarse en algunos casos por eliminar uno de los dos ítems. Esta decisión de cuál de ambos ítems eliminar vendrá motivada por diferentes fuentes. Desde la vertiente teórica, podemos optar por mantener el ítem que más se acerque a una determinada teoría, aunque desde la vertiente estadística, podemos optar por mantener aquel ítem que mejor es capaz de predecir, y por ende, que más aporta a una determinada dimensión. En este segundo caso tendremos que fijarnos en las cargas factoriales de cada ítem. En el ejemplo de la Figura 35, podemos observar como la **carga factorial** del ítem 4 (λ = .48) es considerablemente menor a la carga factorial del ítem 6 (λ = .78). En este caso, si decidimos eliminar uno de los dos ítems eliminaríamos el ítem 4. Es de comentar, además que los ítems que poseen una carga factorial inferior a .50 dentro de su dimensión es un indicio para reconsiderar tal ítem (Lloret, Ferreres, Hernández y Tomás, 2017).

8.2.4.3. Combinar ambos ítems en uno nuevo

Este procedimiento puede ser considerado cuando ambos ítems correlacionan y son buenos, pero cada uno de ellos aporta un matiz diferente. En este caso, puede optarse por reformular ambos ítems y crear uno nuevo que recoja los matices de ambos ítems. Este caso requiere de volver a hacer un nuevo pase de cuestionarios para conocer nuevamente lo bueno o mal ítem que es el nuevo ítem.

Tras llevar a cabo estos ajustes, sería necesario volver a evaluar la calidad del modelo a través de la bondad de ajuste. Se espera que si se han seguido estos pasos, el modelo mejore.

8.2.5. La invarianza factorial en estudios multigrupo

En muchas ocasiones puede darse el caso de que un modelo puede ser bueno, pero solo para un determinado grupo de la muestra. Por ejemplo, puede haber un modelo que tenga buena bondad de ajuste para los chicos, pero una mala bondad de ajuste para las chicas; o un modelo con buena bondad de ajuste para los estudiantes de primaria, pero una mala bondad de ajuste para los estudiantes de secundaria. En esta línea, nace un nuevo concepto: El concepto de la **invarianza factorial**.

Para poder evaluar la invarianza factorial es necesario disponer de al menos dos grupos diferenciados (Chicos y chicas/Estudiantes de Infantil, Primaria y ESO…) y un modelo en común para todos, que es el modelo que se pretende evaluar si es válido para todos los casos.

El procedimiento es medianamente sencillo. Se coge cada grupo por separado y se le plantea cada uno el modelo propuesto, si las diferencias obtenidas en los diferentes estadísticos de bondad de ajuste no son muy grandes, significa que el modelo es invariante en función de esa determinada variable.

Generalmente, se tiende a comparar el modelo en función de los grupos en 4 momentos diferentes. Estos momentos son:

Momento configural: En este momento, el modelo no tiene ningún tipo de restricción. Se compara el modelo para cada grupo tal cual está.

Momento métrico: En este momento, se evalúa además de todo lo anterior, las cargas factoriales para cada grupo. Se espera que para cada grupo, el modelo obtenga unas cargas factoriales similares, y no que lo que para un grupo es una buena variable predictora, para otro grupo no sea una variable significativa.

Momento escalar: En este momento, se evalúa además de todo lo anterior, lo que en estadística se conoce como intercepto. Se espera que ambos grupos tengan interceptos similares para ser un modelo invariante. Tras superar satisfactoriamente estas tres etapas, se puede considerar que el modelo ya posee invarianza factorial. No obstante, se puede aportar información de un cuarto momento.

Momento estricto: En este momento, se evalúa además de todo lo anterior, la varianza y la covarianza de los errores de cada ítem. Se espera que estas covarianzas sean similares para presentar un modelo invariante.

Los estadísticos más comunes que se suelen evaluar en la invarianza factorial son el χ^2/gl, el CFI, el TLI, el RMSEA y el AIC, y la manera de presentar los datos es similar a la manera mostrada en la Tabla 49.

Tabla 49. Ejemplo de análisis de la invarianza factorial en una prueba multigrupo.

	χ^2/gl	$\Delta\chi^2/gl$	CFI	ΔCFI	TLI	ΔTLI	RMSEA	ΔRMSEA	AIC	ΔAIC
Configural	2.174	-	.851	-	.835	.	.033	-	3154.999	-
Métrico	2.176	.002	.842	.009	.833	.002	.034	.001	3188.036	33.03
Escalar	2.187	.011	.832	.010	.831	.002	.034	.000	3212.152	24.11
Estricto	2.376	.189	.792	.040	.804	.027	.036	.002	3491.152	279.0

Fuente: elaboración propia.

En esta Tabla, se recoge tanto el valor de cada estadístico para cada momento, pero también se aporta información de cómo incrementa o se reduce cada estadístico con el paso de una fase a otra. Esto se muestra a través del símbolo delta (Δ), que en estadística significa incremento o variación, dependiendo del contexto. Para poder entender que un momento es invariante, se deben observar estos valores de Δ para cada estadístico, de modo que podremos considerar que un momento cumple hasta ese punto con la invarianza factorial si su ΔCFI y su ΔTLI no varían más de .01 de una fase a otra. Además se espera también que los valores de RMSEA y AIC no fluctúen en exceso de un paso a otro.

En el ejemplo mostrado, se puede observar como el modelo es invariante hasta el momento escalar, pues en el momento estricto, su ΔCFI y ΔTLI son superiores a .01 (Cheung y Rensvold, 2002). Además, su ΔAIC varía mucho con respecto a los dos momentos anteriores. En este caso diríamos que el modelo es invariante, pues recordamos que la invarianza se asume al cumplir, como mínimo, con los tres primeros momentos, pero que no se muestra invariante respecto a las variaciones y covariaciones de los errores (momento estricto).

8.2.6. AFC y SEM en la práctica

Para llevar a cabo toda la teoría recogida en este punto es necesario emplear otro tipo de programas alejados del *SPSS Statistics*. Estamos hablando de programas como *SAS*, *EQS*, *MPlus*, o tal vez, uno de los más usados, *SPSS AMOS*.

En este último caso, nada más entrar al programa observaremos un menú a la izquierda tal y como el que se recoge en la Figura 36:

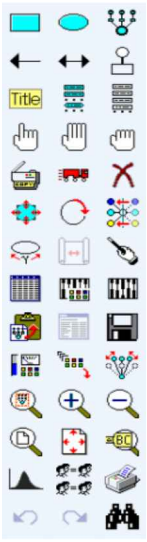


Figura 36. Menú de herramientas de SPSS AMOS.
Fuente: elaboración propia.

Especialmente, para estructurar nuestro modelo, emplearemos las dos primeras líneas de dibujos (rectángulos, círculos, flechas…). Posteriormente, para seleccionar o deseleccionar, emplearemos la cuarta fila, a través de las manos dibujadas. Para mover y borrar emplearemos el camión y la cruz roja, respectivamente. Para redimensionar un círculo o cuadrado y para rotar una Figura, emplearemos el cuadrado con flechas azul y el círculo con giro, respectivamente. También, en las últimas líneas nos muestra algunas herramientas para inspeccionar con el Zoom, por si deseamos acercarnos o alejarnos del dibujo.

No obstante, los botones más importantes son los tres botones situados en la octava fila. El procedimiento para usar estos tres botones sería: Inicialmente, con el botón de la izquierda, seleccionamos dónde tenemos la base de datos con la que vamos a trabajar. Posteriormente, en el botón del medio, seleccionamos las propiedades del análisis. Dentro de esta ventana, en la pestaña *Estimation* se recomienda marcar la opción de estimar medias e interceptos. De igual modo, se recomienda, en la pestaña *Output* pedir a *SPSS AMOS* que nos proporcione información sobre las estimaciones

estandarizadas, así como sobre los índices de modificación, para poder evaluar los ítems del modelo. Finalmente, con el botón de más a la derecha correríamos el análisis.

Para mostrar los resultados sobre el diagrama haremos click sobre la flecha roja hacia arriba situada al lado del panel principal. Seguidamente en la misma columna de donde se halla la flecha roja, marcaremos la opción de *Standardized estimates*. Esto hará que *SPSS AMOS* nos muestre los valores estandarizados en nuestro modelo, de mayor fácil interpretación. Estos valores que aparecerán dibujados en el modelo representarán el valor de las correlaciones de Pearson, y en el caso de los análisis de regresión, indicará las cargas factoriales.

Respecto a los estadísticos para evaluar la bondad de ajuste, estos se hallan en el botón central de la novena fila, *Amos Output*, en la sección *Model Fit*. Para evaluar la bondad de ajuste nos fijaremos en la línea en la que aparece el modelo por defecto (*Default model*). De igual modo, si hemos seleccionado que nos muestre los índices de modificación, esta información aparece en la sección *Modification índices*.

8.3. El análisis de mediación

Está claro que cuando estamos en casa y vemos que comienza a llover, lo primero que pasa por nuestra mente es coger un paraguas y es ya en ese momento cuando salimos a la calle. Visto de esta manera tenemos una condición (Si llueve), que conlleva a una acción (Usar paraguas). No obstante, aunque no caigamos en ello, en un punto intermedio se encuentra una tercera variable escondida: El deseo de no mojarse. Es decir, cuando llueve, como no queremos mojarnos usamos paraguas. Precisamente es aquí donde los análisis de mediación cobran importancia.

Un análisis de mediación es un tipo de análisis que nos permite buscar explicaciones teóricas que ayudan a entender por qué una determinada variable (variable independiente, en nuestro caso, llover no depende de nada) influye en otra (variable dependiente, en nuestro caso, usar paraguas dependerá de si llueve o no). En este caso, entre ambas variables existe la presencia de una variable mediadora que explica por qué cuando llueve usamos paraguas. Veamos otro caso.

Los análisis de mediación se suelen representar como la imagen recogida en la Figura 37. En este análisis de mediación se piensa que los niveles de autoeficacia de un profesor tendrán efecto en lo quemado que se sentirá en su trabajo (burnout). Esta acción se piensa que está mediada por la capacidad que tenga el docente para hacer frente a los problemas que se le presenten.

En este diagrama, se espera que la autoeficacia predecirá de manera positiva la resiliencia (la letra a será positiva). Es decir, a mayor autoeficacia, se espera que la resiliencia del docente sea también mayor. Por otra parte, se piensa que mientras más autoeficacia tenga un docente, menos burnout presentará (la letra c' será negativa), y finalmente, se espera que mientras más resiliencia tenga un docente, menos quemado se sentirá, pues sabrá gestionarse y autorregularse mejor (la letra b será negativa).

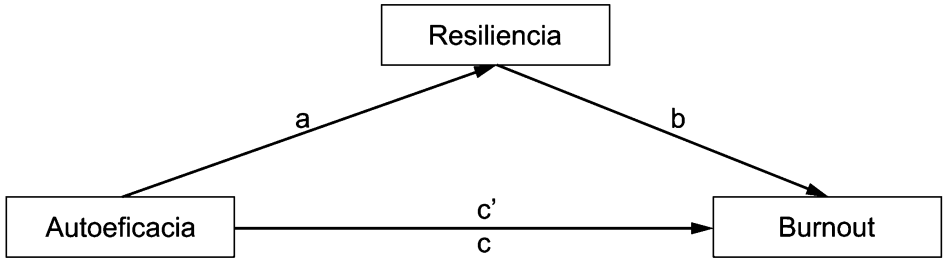


Figura 37. Gráfico conceptual del análisis de mediación.
Fuente: elaboración propia.

En palabras más técnicas a las letras a, b y c' por separadas se las conoce como el **efecto directo** del modelo. Es decir, el efecto que tiene una variable independiente sobre otra dependiente.

Por otra parte, los análisis de mediación tratan de medir también el impacto que tiene la variable mediadora en el modelo sin tener en cuenta el efecto directo (sin tener en cuenta que la autoeficacia puede predecir el burnout, es decir, la c'). Esto se consigue a través de unir dos análisis de regresión, por una parte, el análisis *a* y por otra parte, el análisis de regresión *b*, dando como resultado del efecto indirecto la ruta *ab*. Tal vez, lo más importante de cualquier análisis de mediación es este **efecto indirecto**. Para conocer si esta variable es significativa o no, se calcula los límites superiores (LS) e inferiores (LI) a un intervalo de confianza del 95% y si este intervalo no cruza el valor cero, entonces diremos que es estadísticamente significativo. En este ejemplo, se aporta más información sobre el análisis de mediación. Para este caso, recogido en la Tabla 50, el límite superior (-.1856) y el límite inferior (-.1047) no cruzan el valor cero, por lo que la variable mediadora sí que es estadísticamente significativa. De igual modo, para conocer si la relación entre la variable independiente (autoeficacia) y la dependiente (burnout) se ve significativamente reducida al introducir la variable mediadora, se utiliza el **test de Sobel**. Es decir, este test evalúa si la mediación es significante o no. Se espera, para que la variable mediadora sea buena, que este test sea significativo, es decir, que su p-valor esté por debajo de .05.

Tabla 50. Estadísticos del análisis de mediación.

				(95% IC)	
Efecto	**Ruta**	**Coeficiente**	**SE**	**LS**	**LI**
Efecto directo de la autoeficacia en la Resiliencia	a	.2428***	.0212	.2012	.2845
Efecto directo de la resiliencia en el Burnout	b	-.5889***	.0631	-.7129	-,4650
Efecto directo de la autoeficacia en el Burnout	c'	-.0613*	.0303	-.1209	-.0018
Efecto total de la autoeficacia en el Burnout	c	-.2043***	.0289	-.2612	-.1475
Efecto Indirecto	ab	-.1430 (sig.)	.0206	-.1856	-.1047
Test de Sobel	z	-7.2348***	0.019		
Efecto del modelo total de Burnout (R^2 = .28)					

* p < .05; ** p < .01; *** p < .00; (sig.), significante. 10,000 muestras de bootstrap.
Fuente: elaboración propia.

Finalmente, el **efecto total** (c) es la suma del efecto directo y el efecto indirecto.

La calidad total del modelo es evaluada como cualquier análisis de regresión lineal, a través de su capacidad de predecir lo que ocurrirá, a través del estadístico R^2. En este ejemplo, el modelo propuesto es capaz de predecir correctamente un 28% (R2 = .28) de la varianza. Este valor es un buen valor para tratarse de un área de las ciencias sociales. En otras áreas como la física o la química, se espera que la capacidad predictora sea mucho más superior a esta, en torno al 90-95%.

En este tipo de análisis, suele emplearse una técnica conocida como bootstrapping. El bootstrap es una técnica que pretende reducir el sesgo ocasionado por la muestra lo máximo posible a través del remuestreo. Es decir, mediante esta técnica se pretende aumentar la muestra de un experimento a través de elegir aleatoriamente e introducir a la muestra total uno de todos los casos. Así por ejemplo, si tenemos un experimento con 20 personas y queremos hacer bootstrap de 10000 casos, lo que hará el ordenador será entre esos 20 casos, elegir 1 al azar y lo introducirá con la muestra total. En este momento, con 21 casos en total, volverá a elegir 1 caso al azar y lo volverá a introducir a la muestra total. Nuevamente, seleccionará uno de los 22 casos al azar y lo volverá a introducir a la muestra total, y así sucesivamente hasta crear 10.000 muestras virtuales. En caso de utilizarse esta técnica, es necesario que se indique, como se ha indicado en la Tabla 50.

Este tipo de análisis se pueden realizar a través del software estadístico *SPSS AMOS*, aunque requiere de unos conocimientos bastante avanzados en estadística, objetivo lejano al fin de este libro. No obstante, en caso de ser de interés del lector, a través de una macro llamada PROCESS (https://processmacro.org/) los investigadores pueden realizar este tipo de análisis estadístico a través del programa *SPSS Statistics*.

8.4. El análisis de moderación

El análisis de moderación es un procedimiento en el que una variable independiente que trata de predecir una variable dependiente tiene la influencia de una variable intermedia, conocida como variable moderadora, que hace que el signo y la fuerza con la que la variable independiente predice la variable dependiente pueda cambiar (Galindo-Domínguez, 2019).

En el ejemplo que se muestra a continuación, basado en el estudio de Martínez, García y Maya (2001), se puede observar la presencia de tres variables: La depresión, el estrés, el apoyo social percibido. Se piensa que el estrés es predictor de la depresión. No obstante, se piensa que esta interacción entre el estrés y la depresión se puede ver afectado por el apoyo social que reciba una persona, de modo que una persona que reciba un alto apoyo social ayudará especialmente a aquellos que poseen unos valores de estrés altos, muchísimo más de lo esperado. La manera conceptual y estadística de representar este análisis de moderación es el que se muestra en la Figura 38. A la izquierda se halla la manera conceptual y a la derecha la manera estadística.

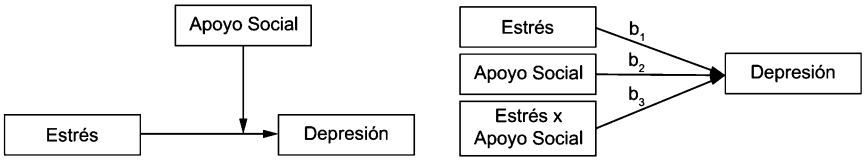


Figura 38. Gráfico conceptual y estadístico del análisis de moderación.
Fuente: elaboración propia.

En los análisis de moderación, al efecto que tiene la variable independiente sobre la dependiente (b_1) y la variable moderadora sobre la variable dependiente (b2), se le conoce como **efecto condicional**, y de igual modo que en el análisis de mediación, en el análisis de moderación sabremos que este efecto condicional (estos análisis de regresión lineal, por separado) es significativo cuando el intervalo de confianza para el límite superior e inferior no cruce el valor cero. Es decir, si obtuviésemos que el límite inferior (LI) fuese-.3307 y el límite superior (LS) fuese .3465, este intervalo al cruzar el valor cero no sería significativo pues el valor cero está en este intervalo. Por otra parte, y tal vez lo más importante del análisis de moderación es conocer la interacción de la variable independiente y moderadora sobre la variable dependiente (b3). En este caso, igual que con los efectos condicionales, analizaremos tanto el límite superior como el límite inferior para conocer si la variable moderadora es significativa o no. En caso de no estar el cero incluido en el intervalo de confianza podremos confirmar que la variable moderadora, efectivamente cumple su función.

Volviendo con el ejemplo, gráficamente, en el caso de que el apoyo social no fuese una variable moderadora (si su intervalo de confianza incluyese el valor cero) esperaríamos un gráfico similar al que se muestra a continuación, en la Figura 39. En este caso, el estrés predice perfectamente la depresión, independientemente del apoyo social percibido.

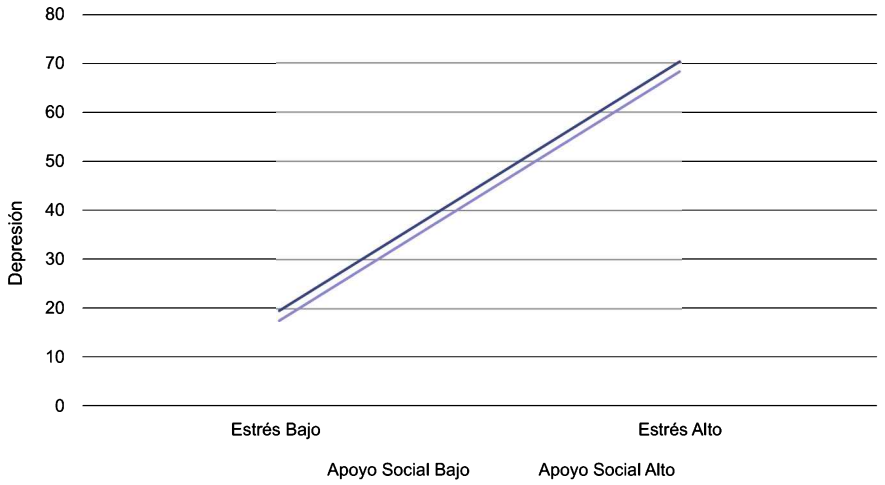


Figura 39. Gráfico de un análisis de moderación sin variable moderadora.
Fuente: elaboración propia.

No obstante, nos encontramos, que el modelo de predicción varía en función de la variable moderadora, que influye en la fuerza con la que el estrés predice la depresión. En este caso, si analizásemos el intervalo de confianza para los límites superior e inferior, observaríamos que el valor cero no está incluido en el mismo.

En la Figura 40, observamos como en aquellas personas con unos valores de estrés altos, el recibir un apoyo social alto ayuda a reducir la depresión mucho más que en aquellos casos que no la reciben. Este apoyo social es especialmente importante en muestras con alto estrés, pues como se aprecia, en aquellos casos con bajo estrés, el hecho de recibir apoyo social apenas varía.

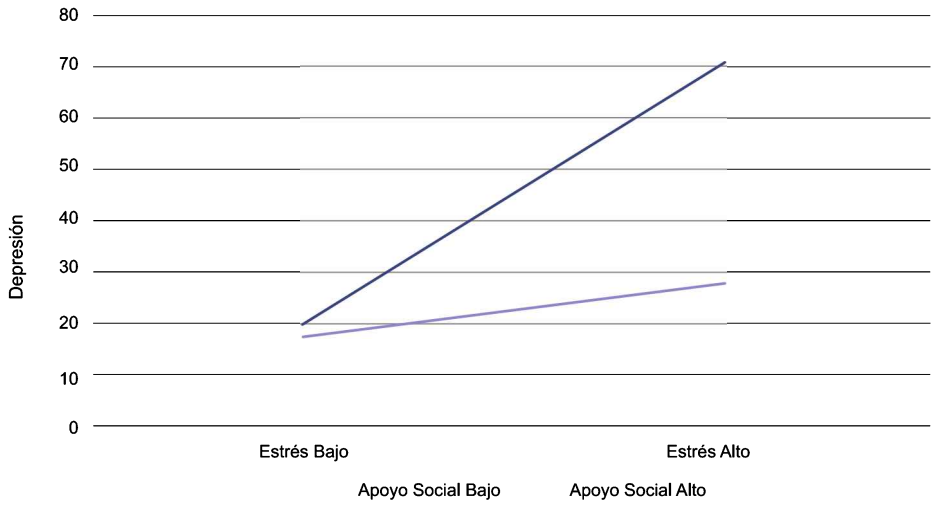


Figura 40. Gráfico de un análisis de moderación con variable moderadora.
Fuente: elaboración propia.

Al igual que en el caso del análisis de mediación, este tipo de análisis no viene incluido como algo básico en el programa *SPSS Statistics*, por lo que es necesario o bien emplear algún tipo de programa más complejo, como *SPSS AMOS*, o por el contrario, instalar en *SPSS Statistics* la macro PROCESS (https://processmacro.org/).

REFERENCIAS BIBLIOGRÁFICAS

Anderson, D. R., Sweeney, D. J., y Williams, T. A. (2001). *Estadística para administración y economía*. Thomson.

Argibay, J. C. (2006). Técnicas psicométricas. Cuestiones de validez y confiabilidad. *Subjetividad y procesos cognitivos*, *8*, 15-33. http://dspace.uces.edu.ar:8180/xmlui/handle/123456789/765

Arias, M. M. (2000). La triangulación metodológica: Sus principios, alcances y limitaciones. *Investigación y educación en enfermería, 18*(1), 13-26. https://www.redalyc.org/articulo.oa?id=105218294001

Bentler, P. M., y Bonnet, D. C. (1980). Significance Tests and Goodness of Fit in the Analysis of Covariance Structures. *Psychological Bulletin, 88*(3), 588-606. https://www.researchgate.net/publication/232518840_Significance_Tests_and_Goodness-of-Fit_in_Analysis_of_Covariance_Structures

Berlanga, V., y Rubio, M. J. (2012). Clasificación de pruebas no paramétricas. Como aplicarlas en *SPSS Statistics*. *REIRE. Revista d'Innovació i Recerca en Educació, 5*(2), 101-113. https://core.ac.uk/download/pdf/39101714.pdf

Blaxter, L., Hugues, C., y Tight, M. (2008). *Cómo se investiga*. Editorial Grao.

Campo-Arias, A., y Oviedo, H. C. (2008). Propiedades psicométricas de una escala: la consistencia interna. *Revista de Salud Pública, 10*(5), 831-839. https://scielosp.org/article/rsap/2008.v10n5/831-839/es/

Carmines, E. G., y Zeller, R. A. (1979). *Reliability and validity assessment*. Sage.

Cheung, G. W., y Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*, 233-255. https://doi.org/10.1207/S15328007SEM0902_5

Cole, J. C., Motivala, S. J., Khanna, D., Lee, J. Y., Paulus, H. E., y Irwin, M. R. (2005). Validation of single-factor structure and scoring protocol for the health assessment questionnaire-disability index. *Arthritis & Rheumatism, 53*(4), 536-542. https://doi.org/10.1002/art.21325

Comrey, A. L. (1985). *Manual de Análisis Factorial*. Cátedra.

Diamantopoulos, A., y Siguaw, J.A. (2000). *Introducing LISREL*. Sage Publications.

Domínguez-Lara, S. (2017). Magnitud del efecto, una guía rápida. *Educación Médica*, *19*(4), 251-254. https://doi.org/10.1016/j.edumed.2017.07.002

Escobedo, M. T., Hernández, J. A., Estebané, V., y Martínez, G. (2016). Modelos de ecuaciones estructurales: Características, fases, construcción, aplicación y resultados. *Ciencia y Trabajo, 55*, 16-22. https://scielo.conicyt.cl/pdf/cyt/v18n55/art04.pdf

Galindo-Domínguez, H. (2019). El análisis de moderación en el ámbito socioeducativo a través de la macro Process en SPSS Statistics. *REIRE. Revista d'Innovació i Recerca en Educació, 12*(1), 1-11. https://dialnet.unirioja.es/servlet/articulo?codigo=6749279

George, D., y Mallery, P. (2003). *SPSS Statistics for Windows step by step: A Simple Guide and Reference. 11.0 Update*. Allyn & Bacon.

Gerbing, D. W., y Anderson J. C. (1988). An update paradigm for scale development incorporating unidimentionality and its assessment. *Journal of Marketing Research, 25*(2), 186-192. https://www.jstor.org/stable/3172650?seq=1

Gil, J. A. (2015). *Metodología Cuantitativa en Educación.* Editorial UNED.

Hair, J. F., Anderson, R. E., y Tatham, R. L. (1999). *Análisis multivariante.* Prentice-**Hall.**

Hooper, D., Coughlan, J., y Mullen, M. R. (2008). Structural equation modelling: Guidelines for determining model fit. *Journal of Business Research Methods, 6*(1), 53-60. https://www.researchgate.net/publication/254742561_Structural_Equation_Modeling_Guidelines_for_Determining_Model_Fit

Jöreskog, K., y Sörbom, D. (1993). *LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language.* Scientific Software International Inc

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology, 140,* 150. (Traducción al castellano en C. H. Wainerman (Comp.) (1976), *Escalas de medición en ciencias sociales,* 199-260. Nueva visión.

Lloret, S., Ferreres, A., Hernández, A., y Tomás, I. (2017). The exploratory factor analysis of ítems: guided analysis based on empirical data and software. *Anales de psicología, 33*(2), 417-432. https://doi.org/10.6018/analesps.33.2.270211

Lord, F. M. (1980). *Applications of item response theory to practical testing problems.* Erlbaum.

MacCallum, R. C., Browne, M. W., y Sugawara, H. M. (1996). Power Analysis and Determination of Sample Size for Covariance Structure Modeling. *Psychological Methods, 1*(2), 130-49. http://ww.w.statpower.net/Content/312/Handout/MacCallumBrowneSugawara96.pdf

Martinez, M. F., García, M., y Maya, I. (2001). El efecto amortiguador del apoyo social sobre la depresión en un colectivo de inmigrantes. *Psicothema, 13*(4), 605-610. http://www.psicothema.com/psicothema.asp?id=486

McDonald, R. P. (1999). *Test theory: An unified treatment.* Lawrence Erlbaum Associates, Inc.

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica, 22*(3), 276-282. https://www.ncbi.nlm.nih.gov/pubmed/23092060

Moncada-Jiménez, J., Solera-Herrera, A., y Salazar-Rojas, W. (2002). Fuentes de varianza e índices de varianza explicada en las ciencias del movimiento humano. *Revista de Ciencias del ejercicio y la salud, 2*(2), 70-74. https://www.researchgate.net/publication/234139405_Fuentes_de_varianza_e_indices_de_varianza_explicada_en_las_ciencias_del_movimiento_humano

Morales, P. (2007). *La fiabilidad de los tests y escalas.* http://web.upcomillas.es/personal/peter/estadisticabasica/Fiabilidad.pdf

Pérez, R. (1986). *Pedagogía experimental. La medida en educación.* UNED.

Pérez-Gil, J. A., Chacón, S., y Moreno, R. (2000). Validez de constructo: el uso de análisis factorial exploratorio-confirmatorio para obtener evidencias de validez. *Psicothema, 12*(2), 442-446. http://www.psicothema.com/psicothema.asp?id=601

Rodríguez, C., Pozo, T., y Gutiérrez, J. (2006). La triangulación analítica como recurso para la validación de estudios de encuesta recurrentes e investigaciones de réplica en Educación Superior. *RELIEVE. Revista electrónica de Investigación y Evaluación Educativa, 12*(2), 289-305. https://ojs.uv.es/index.php/RELIEVE/article/view/4231/3838

Ruiz, M. A., Pardo, A., y San Martín, R. (2010). Modelos de ecuaciones estructurales. *Papeles del psicólogo, 3*(1), 34-45. http://www.papelesdelpsicologo.es/pdf/1794.pdf

Silva, J. M. (2002). ¿Qué es eso de la ética profesional? *Revista Contaduría y Administración, 205*, 5-11. https://www.redalyc.org/pdf/395/39520502.pdf

Tomás, J. M., Sancho, P., Oliver, A., Galiana, L., y Meléndez, J. C. (2012). Efectos de métodos asociados a ítems invertidos vs. Ítems en negativo. *Revista Mexicana de Psicología, 29*(2), 105-115. https://www.redalyc.org/pdf/2430/Resumenes/Resumen_243030190001_1.pdf

Wheaton, B., Muthen, B., Alwin, D., F., y Summers, G. (1977). Assessing Reliability and Stability in Panel Models. *Sociological Methodology, 8*(1), 84-136. https://www.jstor.org/stable/pdf/270754.pdf?seq=1

Zavando, D., Suazo, I., y Manterola, C. (2010). Validez en la investigación imaginológica. *Revista Chilena de Radiología, 16*(2), 75-79. http://dx.doi.org/10.4067/S0717-93082010000200007

GLOSARIO

Afijación
AIC
Alfa de Cronbach
Algoritmo CHAID
Análisis clúster
Análisis de conglomerados
Análisis de correspondencia
Análisis de mediación
Análisis de moderación
Análisis de regresión
Análisis de regresión múltiple
Análisis de regresión simple
Análisis de segmentación
Análisis discriminante
Análisis factorial confirmatorio
Análisis factorial exploratorio
ANOVA de medidas repetidas
ANOVA de un factor
APA
Aquiescencia
Asimetría
Asociación Americana de Psicología
Bootstrapping
Calcular Variable
Carga factorial
CFI
Coeficiente de determinación
Coeficiente de equivalencia
Coeficiente de validez
Conglomerado
Conglomerado en dos etapas
Consistencia Interna
Constructo
Correlación
Correlación canónica
Correlación de Pearson
Correlación de Spearman
Correlación negativa
Correlación positiva
Covarianza
Criterio de la raíz latente
Cuartil
Curtosis
Curva leptocúrtica
Curva mesocúrtica
Curva platicúrtica
D de Cohen
Decil
Decisión estadística
Dendrograma
Descriptivos
Descriptivos de dispersión
Descriptivos de distribución
Descriptivos de tendencia central
Deseabilidad Social
Desviación típica
Diagrama de líneas
Dialnet
Dimensionalidad
Diseño cuasi-experimental
Diseño descriptivo
Diseño experimental
Diseño longitudinal
Diseño muestral no probabilístico
Diseño muestral probabilístico
Diseño transversal
Distribución normal
Dos mitades
Editor
Efecto condicional
Efecto directo
Efecto Hawthorne
Efecto indirecto
Efecto John Henry
Efecto total
Equamax
Escala Likert
Eta Cuadrada
Extrapolable
Factor
FECYT
Fiabilidad
Formulación de objetivos
Frecuencias
Fuentes digitales de búsqueda
Fuentes manuales de búsqueda
Gold Standard
Grados de libertad
Gráfico circular
Gráfico de barras
Gráfico de dispersión
H de Kruskal-Wallis
Heterocedasticidad
Hipótesis alterna
Hipótesis del investigador
Hipótesis nula
Histograma
Homocedasticidad
IFI
Implicaciones
Índice de modificación
Inercia
Instrumento de medición
documental
Invarianza factorial
Ítem
Ítem invertido
Kaiser-Meyer-Olkin
Kappa de Cohen
KMO
Kolmogorov-Smirnov
KR-20
Latindex
Limitaciones
Máximo
Media aritmética
Mediana
Método de K-medias
Métodos jerárquicos
Métodos no jerárquicos
Mínimo
Moda
Modelo de ecuaciones estructurales
Momento conFigural
Momento escalar
Momento estricto
Momento métrico
Muestra
Muestreo con reemplazamiento
Muestreo sin reemplazamiento
Multidimensionalidad
NFI
Nivel de significación
Omega
Outlier
Percentil
Población
Polígono de frecuencias
Porcentaje
Porcentaje acumulado
Porcentaje válido
Post-Hoc
Post-Hoc de Scheffe
Post-Hoc de Tukey
Pregunta de investigación
Pre-Post con grupo control
Pre-Post con un único grupo
Promax
Proporción de inercia explicada
Prospectiva
Prueba de esfericidad de Bartlett
Prueba de hipótesis
Prueba no paramétrica
Prueba paramétrica
Quartimax
R de Rossenthal
Rango
Reactividad psicológica
Reactivo
Recodificar en distintas variables
Recodificar en las mismas variables
Recuento esperado
Recuento observado
Representatividad
Revisión de la literatura
Revisión por pares
RMSEA
Rotación
Saphiro-Wilk
Scopus
Significancia
SJR
Solo Post
SRMR
T de Student
T de Wilcoxon
Tabla cruzada
Tabla de contingencia
Tamaño del efecto
Test de Levene
Test de Sobel
Test paralelos
Test-retest
TLI
Triangulación
Triangulación de investigador
Triangulación metodológica
Triangulación teórica
U de Mann Whitney
Unidimensionalidad
Validez
Validez convergente
Validez de constructo
Validez de contenido
Validez de criterio
Validez de jueces
Validez de respuesta
Validez discriminante
Validez divergente
Validez racional
Valor atípico
Valor tendencial
Variable
Variable categórica
Variable dependiente
Variable dicotómica
Variable dummy
Variable en escala
Variable endógena
Variable error
Variable exógena
Variable independiente
Variable latente
Variable nominal
Variable ordinal
Variable politómica
Variable predictora
Varianza
Varianza no explicada
Varimax
W de Kendall
Web of Science
X2/gl

Economía, Organización y Ciencias Sociales