BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA

FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS

Validez y confiabilidad del cuestionario de calidad de vida SF-36 en mujeres con LUPUS, Puebla

TESIS

PARA OBTENER EL TÍTULO DE
Lic. en Matemáticas Aplicadas

PRESENTA
Guadalupe Santos Sánchez

DIRECTORES DE TESIS
Dra. Gladys Linares Fleites
Dra. Hortensia Josefina Reyes Cervantes

PUEBLA, PUE. ENERO 2017

Gracias a esas personas importantes en mi vida, que siempre estuvieron listas para brindarme toda su ayuda, ahora me toca regresar un poquito de todo lo inmenso que me han otorgado. Con todo mi cariño está tesis se las dedico a ustedes:

Papá: José Edmundo Santos Sánchez

Mamá: Basilia Sánchez Alcalá

Hermanos: Eladia, Filiberto, Juan, Balbina, José Alberto, Seferino, Marcelo, María Julia, María del Roció.

Novio: Antonio Pérez González

Agradecimientos

Le agradezco a dios por haberme acompañado y guiado a lo largo de mi carrera, por ser mi fortaleza en los momentos de debilidad y por brindarme una vida llena de aprendizajes, experiencias y sobre todo felicidad.

Le doy gracias a mis padres por apoyarme en todo momento, gracias a sus consejos y palabras de aliento me han ayudado a crecer como persona y a luchar por lo que quiero, por los valores que me han inculcado y por confiar en mí.

A mis hermanos: Gracias por su apoyo, cariño y por estar en los momentos más importantes de mi vida. Este logro también es de ustedes.

A mi novio: La ayuda que me has brindado ha sido sumamente importante, estuviste a mi lado inclusive en los momentos y situaciones más tormentosas, siempre ayudándome. No fue sencillo culminar con éxito este proyecto, sin embargo siempre fuiste muy motivador y esperanzador, me decías que lo lograría perfectamente. Me ayudaste hasta donde te era posible, incluso más que eso. Muchas gracias, Moris.

A mis directores de tesis: Gladys Linares Fleites y Hortensia Josefina Reyes Cervantes por darme la oportunidad de realizar esta tesis, muchas gracias por la orientación, el seguimiento y la supervisión de la misma, pero sobre todo por la motivación y el apoyo recibido a lo largo de estos años.

A los investigadores del área de la salud de la Benemérita Universidad Autónoma de Puebla. En especial, a la Dra. Socorro Méndez Martínez por todo el apoyo y la facilidad que me fue otorgada, como hacer uso de los datos, para determinar la confiabilidad y validez del cuestionario SF-36.

¡Muchas gracias a todos!

Introducción

El interés fundamental de esta tesis es comprender y desarrollar la información propuesta en la literatura sobre las técnicas estadísticas relacionadas con la confiabilidad y validez de instrumentos de medición.

Antes de entrar propiamente en el tema, podríamos decir que un cuestionario (constructo, escala, subescala, etc.) es un instrumento utilizado para la recolección de información. La problemática, al momento de la recolección de datos en ciertas áreas de investigación donde la medición no puede hacerse directamente, se centra en la construcción de los instrumentos a emplear con esta finalidad, de manera que permitan recabar información válida y confiable.

La confiabilidad consiste en determinar hasta donde las respuestas de un instrumento de medición aplicado a un conjunto de individuos, son estables independientemente del individuo que lo aplique y el tiempo en el que es aplicado. La validez es el grado en el que el instrumento mide lo que queremos medir y el modelo factorial suele proponerse como uno de los métodos de validación de constructo por lo que profundizamos en el mismo: la aplicación de estos conceptos se realiza con el apoyo del instrumento SF-36 que mide la Calidad de Vida Relacionada con la Salud.

El concepto de «calidad de vida» se introduce como un criterio más a considerar cuando se define el estado de salud de una persona. Debido a que la calidad de vida se basa en mediciones con una carga variable de subjetividad, se requieren métodos de evaluación válidos, reproducibles y confiables. El mejor conocimiento de las evaluaciones para medir la calidad de vida permitirá incorporar estos instrumentos en la evaluación integral de individuos, en la conducción de ensayos clínicos y en la investigación de servicios de salud.

Los principales objetivos que se persiguen en esta tesis son los siguientes:

1.- Investigar los conceptos de confiabiliadad y validez de constructo, así como, los procedimientos para estimarlos.

2.- Analizar el análisis factorial exploratorio y el análisis factorial confirmatorio como técnicas fundamentales de la validez de instrumentos de medición.

3.- Determinar la confiabilidad y validez del cuestionario SF-36 mediante el uso del software R-Commander (de distribución libre).

La tesis está estructurada en tres capítulos. El primero se refiere a los conceptos de confiabilidad y validez de instrumentos de medición. En el segundo se desarrolla el análisis factorial en sus dos modalidades: el exploratorio y confirmatorio. En el tercer capítulo se desarrolla el proceso de validación y confiabilidad del cuestionario de calidad de vida SF-36. Finalmente, se presentan las conclusiones y la bibliografía utilizada.

Índice general

Índice de figuras

Índice de cuadros

Capítulo 1

Confiabilidad y validez

Frecuentemente los investigadores necesitan tener seguridad que el instrumento que utilizan para extraer cierta información de cualquier fenómeno mida lo que realmente quieren medir y que sea coherente, para esto todo instrumento de medición debe tener dos importantes características que son la confiabilidad y validez. La confiabilidad nos indica el grado en el que la aplicación repetida del instrumento al mismo sujeto, produzca los mismos resultados y la validez se refiere al grado en el que un instrumento mide lo que se supone que debe medir. Por esta razón es muy importante que el investigador deba averiguar u obtener la confiabilidad y validez del instrumento utilizado en su estudio, ya que si los datos obtenidos no son confiables y válidos, lo resultados merecen poco interés [3] y [16].

El concepto de confiabilidad es distinto del concepto de la validez. En el sentido más usual del término (no el único), un instrumento es válido si comprueba o mide aquello que pretendemos medir. Un instrumento puede ser válido, porque mide lo que decimos que mide y queremos medir, pero lo puede medir con un margen de error grande; con instrumentos parecidos o en mediciones sucesivas hubiéramos obtenido resultados distintos. También puede haber una confiabilidad alta (los sujetos están clasificados, ordenados, con poco margen de error) y a la vez el instrumento puede carecer de validez, porque no mide lo que se pretende o lo que se dice que se está midiendo, [24].

La confiabilidad y la validez de un instrumento no son cualidades completamente independientes. Un instrumento de medición que no sea confiable no puede ser válido, pues si es errático, incongruente e inexacto tampoco medirá con validez el atributo en cuestión. Sin embargo, un instrumento de medición puede ser confiable y no obstante carecer de validez; más aún, un alto grado de confiabilidad no comprueba la validez de un instrumento para determinado propósito, [32].

1.1. Confiabilidad

En este capítulo se estudian los conceptos preliminares básicos de la confiabilidad y se mencionan los procedimientos para su estimación. La confiabilidad, también denominada precisión, corresponde al grado con que los puntajes de una medición se encuentran libres de error de medida. Es decir, al repetir la medición en condiciones constantes estas deberían ser similares. Este concepto se relaciona con la estabilidad del instrumento en sí mismo, independiente del individuo quien lo aplique (observador) y del momento en que es aplicado (tiempo). En principio la confiabilidad expresa el grado de precisión de la medida, una manera de verificar la precisión es medir lo mismo varias veces, o varios observadores independientes miden lo mismo para obtener una media que se estima más precisa que lo que un único observador ha estimado. Otro concepto que nos ayuda a comprender qué entendemos por confiabilidad es el de consistencia o predictibilidad, [24].

Otra manera de aproximarse al concepto de confiabilidad es preguntarse: ¿Hasta dónde los resultados obtenidos con un instrumento de medición constituyen la medida "verdadera" de la propiedad que se pretende medir? Todavía existe otra posibilidad de cómo podemos enfocar la confiabilidad de un instrumento de medición; ella responde a la siguiente cuestión: ¿cuánto error está implícito en la medición de un instrumento? Se entiende que un instrumento es menos confiable en la medida que hay un mayor margen de error implícito en la medición. De acuerdo con esto, la confiabilidad puede ser definida como la ausencia relativa de error de medición en el instrumento, [29].

Finalmente, lo que nos va a decir un coeficiente de confiabilidad es si el instrumento diferencia adecuadamente a los sujetos en aquello que mide el test o escala. Con un test o escala pretendemos diferenciar a los sujetos; establecer quién tiene más o menos del rasgo que medimos.

1.1.1. Procedimientos para estimar la confiabilidad

1.1.1.1. El modelo clásico de la confiabilidad

En la teoría clásica de los test se supone que la puntuación observada de una persona en una prueba está compuesta por una puntuación verdadera más algún error no sistemático de medición. La puntuación verdadera se define como la puntuación obtenida por un individuo si todas las condiciones, tanto internas como externas, estuvieran controladas y el instrumento

de medición fuera “perfecto”. Pero, esta condición es sólo hipotética, ya que como sabemos en toda medición hay un error de medición implícito. El error se refiere al aumento o disminución de la medición, como resultado de diferentes factores que determinan el error de medición; éstos dependen, algunas veces, de una validación inadecuada del instrumento de medida y otras, de las condiciones externas bajo las cuales se realiza la medición, [29].
La puntuación verdadera de una persona en una prueba particular se define como el promedio de las puntuaciones que obtendrá si presentara la prueba un número infinito de veces, por este motivo la puntuación verdadera de una persona no puede medirse exactamente, así que tiene que ser estimada a partir de su puntuación observada en la prueba.

Por consecuencia, la puntuación observada puede ser expresada en los términos del modelo siguiente:

$$x_t = x_v + x_e. \tag{1.1.1}$$

En donde:
$x_t =$ Puntuación observada, puntuación total de un individuo,
$x_v =$ Puntuación verdadera, que representa lo que un individuo realmente sabe o siente (depende de lo que se esté preguntando o midiendo),
$x_e =$ Puntuación debida a errores de medición, que puede tener signo más o signo menos [24] y [29].

La Teoría Clásica de los Test (TCT) opta por describir más detalladamente el componente de error a partir de la inclusión de algunos supuestos sobre su comportamiento a lo largo de un conjunto de mediciones en una población según (Gulliksen, 1950; Lord & Novick, 1968; Nunnally, 1987; Thorndike, 1989,1996; Muñiz, 1996; Martínez, 1996; Herrera et al., 2000). Citados en [6]. Estos son los siguientes:
a) El valor esperado del error aleatorio es igual a cero.
b) El error se distribuye normalmente con media cero y varianza s_e^2.
c) El error aleatorio de medición en una prueba no se encuentra correlacionado con la puntuación verdadera en la prueba, con el error de medición en otra prueba ni con la puntuación verdadera en otra prueba.
d) Las varianzas de las puntuaciones observadas, las puntuaciones verdaderas y del error son finitas y mayores que cero.

La varianza cuantifica todo lo que hay de diferencia entre los sujetos. La fórmula básica de la confiabilidad parte del hecho de que la varianza de las puntuaciones totales de un test podemos descomponerla de la siguiente manera:

$$s_t^2 = s_v^2 + s_e^2. \tag{1.1.2}$$

En donde:

s_t^2 = Varianza total, expresa todo lo que hay de diferente en las puntuaciones totales; unos sujetos tienen puntuaciones totales más altas, otras más bajas, etc, la varianza será mayor si los sujetos difieren mucho entre sí. Si lo que pretendemos con un instrumento de medida es clasificar, detectar diferencias, una varianza grande estará asociada en principio a una mayor confiabilidad.

s_v^2 =Varianza verdadera, expresa todo lo que hay de diferente debido a que los sujetos son distintos en lo que pretendemos medir, o dicho de otra manera, expresa todo lo que hay de diferente debido a lo que los ítems tienen en común, de relación, y que es precisamente lo que queremos medir. El término verdadero no hay que entenderlo en un sentido cuasi filosófico, aquí la varianza verdadera es la que se debe a respuestas coherentes (o respuestas relacionadas), y esta coherencia (o relación verificada) en las respuestas suponemos que se debe a que los ítems miden lo mismo.

s_e^2 = Varianza debida a errores de medición, o debida a que los ítems miden en parte cosas distintas, a lo que no tienen en común. Puede haber otras fuentes de error (respuestas descuidadas, falta de motivación al responder, etc.), pero la fuente de error que controlamos es la debida a falta de relación entre los ítems, que pueden medir cosas distintas o no muy relacionadas. El error aquí viene a ser igual a incoherencia en las respuestas, cualquiera que sea su origen (incoherencia sería aquí responder no cuando se ha respondido sí a un ítem de formulación supuestamente equivalente) [24] y [29].

La confiabilidad enfocada desde el punto de vista de la teoría del error de medición, nos llevaría a establecer una relación inversa con respecto a la confiabilidad, en los términos siguientes: a mayor error implícito en la medición, menor confiabilidad; mientras que a menor error, mayor confiabilidad. En términos prácticos, esto significa que si podemos estimar la varianza de error de una medida, también podemos estimar su confiabilidad. Todo lo cual nos lleva a que la confiabilidad puede ser vista como la proporción de la varianza "verdadera" con respecto a la varianza total:

$$r_{tt} = \frac{s_v^2}{s_t^2} \tag{1.1.3}$$

En términos verbales, la ecuación (1.1.3) se expresa como:

$$\textit{Confiabilidad} = \frac{\textit{Varianza debida a lo coherente en las respuestas}}{\textit{Varianza debida a lo coherente y no coherente en las repuestas}}$$

En la varianza total influye tanto lo que se responde de manera coherente o relacionada, como lo que hay de incoherente o inconsistente (por la causa que sea), la confiabilidad expresa la proporción de consistencia o coherencia empírica. La confiabilidad de una prueba equivale a la razón entre la varianza de las puntuaciones verdaderas y la varianza de los puntajes totales, y en esta medida expresa la proporción de la varianza total en un grupo de puntajes que corresponde o puede ser atribuida a las variaciones entre las puntuaciones libres de error. Esta noción puede observarse en la siguiente ecuación que se deduce de las fórmulas (1.1.2) y (1.1.3) [6],

$$r_{tt} = 1 - \frac{s_e^2}{s_t^2}$$

Existen varias maneras para estimar la confiabilidad de una medida, las más conocidas son: (a) confiabilidad de reaplicación (test-retest); (b) confiabilidad de versiones equivalentes (pruebas paralelas) y (c) confiabilidad de consistencia interna.
Acontinuación nos referiremos a esta última.

1.1.1.2. Confiabilidad de consistencia interna

Este tipo de confiabilidad permite determinar el grado en que los ítems de una prueba están correlacionados entre sí. La confiabilidad de consistencia interna, pone énfasis en las puntuaciones de los sujetos y no en el contenido o el formato de los reactivos. Por lo tanto, si los ítems del instrumento correlacionan positivamente entre sí, éste será homogéneo, independientemente del tipo de contenido que se haya utilizado. Por el contrario, la prueba será heterogénea si los reactivos no tienen una correlación positiva entre sí, aun cuando aparentemente estén midiendo el mismo rasgo. Como se puede comprender, la distinción entre lo homogéneo y lo heterogéneo no es una dicotomía, sino un continuo. Por otra parte, la homogeneidad está relacionada con la característica de unidimensionalidad de una prueba, la cual indica que el instrumento mide una sola variable (un rasgo) en lugar de una combinación de ellas. Si una prueba es homogénea, podemos suponer que todos los ítems miden una característica común, [29].

Existen diferentes procedimientos para estimar la confiabilidad de consistencia interna. Algunos de los más conocidos son los siguientes: (a) Dos mitades, corregido por la fórmula de Spearman-Brown; (b) Kuder-Richardson y (c) Alpha de Cronbach, [29].
A continuación nos referiremos a este último.

Coeficiente Alfa de Cronbach (α)

El coeficiente α fue propuesto en 1951 por Cronbach como un estadístico para estimar la confiabilidad de una prueba, o de cualquier compuesto obtenido a partir de la suma de varias mediciones. Para evaluar la confiabilidad o la homogeneidad de las preguntas o ítems, es común emplear el coeficiente Alfa de Cronbach cuando se trata de alternativas de respuestas policotómicas, como las escalas tipo Likert; la cual puede tomar valores entre 0 y 1, donde: 0 significa confiabilidad nula y 1 representa confiabilidad total [7] y [24].
El coeficiente α de Cronbach puede ser calculado por medio de dos formas:

1.- Mediante la varianza de los ítems y la varianza del puntaje total, [7], [16] y [25].

$$\alpha = \frac{k}{k-1}[1 - \frac{\sum S_i^2}{S_t^2}], \tag{1.1.4}$$

Donde:
α: Coeficiente de confiabilidad de la prueba o cuestionario.
k: Número de ítems del instrumento.
S_t^2: Es la Varianza total del instrumento.
$\sum S_i^2$: Es la Suma de la varianza individual de los ítems, $i = 1, ..., k$

2.- Mediante la matriz de correlación de los ítems.

$$\alpha = \frac{k\bar{p}}{1 + \bar{p}(k-1)}. \tag{1.1.5}$$

Donde
k: Número de ítems.
$\bar{p}$: Promedio de las correlaciones lineales entre cada uno de los ítems, [7] y [24].

1.1.2. Interpretación del coeficiente de confiabilidad

El coeficiente de confiabilidad es un coeficiente de correlación, teóricamente significa correlación del test consigo mismo, [7] y [29].
Sus valores oscilan entre cero (0) y uno (1.00). Una manera práctica de interpretar la magnitud de un coeficiente de confiabilidad puede ser guiada por la escala mostrada en el Cuadro 1.1.

Rangos	Magnitud
0,81 a 1,00	Muy Alta
0,61 a 0,80	Alta
0,41 a 0,60	Moderada
0,21 a 0,40	Baja
0,01 a 0,20	Muy Baja

Fuente: Tomado de Ruiz Bolívar (2002).

Cuadro 1.1: Interpretación de la magnitud del coeficiente de confiabilidad de un instrumento. Fuente: Ruíz Bolivar (2002).

No hay normas para determinar que coeficiente de confiabilidad resulta aceptable, algunas valoraciones pueden encontrarse en libros de texto y por diversos autores, pero son sólo orientadoras. En la práctica cada coeficiente hay que valorarlo en su situación: tipo de instrumento (define un rasgo muy simple o muy complejo), tipo de muestra (homogénea o heterogénea) y uso pretendido del instrumento (mera investigación sobre grupos o toma de decisiones sobre sujetos) [24].
Los valores del coeficiente de confiabilidad oscilan entre 0 y 1, pero ocasionalmente podemos encontrar valores negativos, simplemente porque no se cumplen en un grado apreciable las condiciones de estos modelos; en este caso (valor negativo) podemos interpretar este coeficiente como cero, [24].

El coeficiente de confiabilidad puede interpretarse de la siguiente manera:

1. Expresa la proporción de varianza debida a lo que los ítems tienen de relacionado, un coeficiente de .70 indica que el 70 % de la varianza (diferencias en los totales) se debe a lo que los ítems tienen en común (de coherencia en las respuestas), y un 30 % de la varianza se debe a errores de medición o a lo que de hecho tienen los ítems de no relacionado. De esta interpretación podemos decir que es una interpretación literal, que se desprende directamente de la fórmula (Suma de covarianzas/Varianza total).

2. Es una estimación del coeficiente de correlación que podemos esperar con un test similar, con el mismo número y tipo de ítems. Esta interpretación se deriva directamente del modelo teórico propuesto por Cronbach. De un universo o población de posibles ítems hemos escogido una muestra de ítems que es la que conforma nuestro instrumento. Si la confiabilidad es alta, con otra muestra de ítems de la misma población de ítems obtendríamos unos resultados semejantes (los sujetos quedarían ordenados de manera similar).

3. El coeficiente de confiabilidad nos dice, si un test discrimina adecuadamente, si clasifica bien a los sujetos, si detecta bien las diferencias que existen entre los sujetos de una muestra. Diferencias ¿En qué? En aquello que es común a todos los ítems y que es lo que pretendemos medir. Es más, sin diferencias entre los sujetos no puede haber un coeficiente de confiabilidad alto. La confiabilidad es una característica positiva siempre que interese detectar diferencias que suponemos que existen. Esto sucede cuando medimos rasgos de personalidad, actitudes, etc. medir es, de alguna manera, establecer diferencias.

4. Índice de precisión. Hemos visto que el coeficiente de confiabilidad expresa una proporción, la proporción de varianza verdadera o varianza debida a lo que los ítems tienen en común. También sabemos que un coeficiente de correlación elevado al cuadrado expresa una proporción (la proporción de varianza compartida por dos variables). Es decir que la raíz cuadrada de una proporción equivale a un coeficiente de correlación. En este caso la raíz cuadrada de un coeficiente de confiabilidad equivale al coeficiente de correlación entre las puntuaciones obtenidas (con nuestro instrumento) y las puntuaciones verdaderas (obtenidas con un test ideal que midiera lo mismo). Este coeficiente se denomina índice de precisión [24].

Finalmente la interpretación del coeficiente de confiabilidad se complementa con el cálculo y uso del error típico o margen de error; es la oscilación probable de las puntuaciones si los sujetos hubieran respondido a una serie de tests paralelos; a mayor confiabilidad (a mayor precisión) bajará la magnitud del error probable [24].

1.2. Validez

En el tema anterior, nos enfocamos en determinar hasta donde las respuestas de un instrumento de medición aplicado a un conjunto de individuos, son estables independientemente del individuo que lo aplique y el tiempo en el que es aplicado. Así como también analizamos los diferentes procedimientos para estimar la confiabilidad. A continuación, nos interesa estudiar si el instrumento de medición es válido, esto se refiere a verificar si el instrumento de medición mide lo que realmente queremos medir, para esto empezamos con los conceptos preliminares básicos de validez, después con los procedimientos para estimar, y finalmente la importancia de la validez.

1.2.1. Conceptos preliminares básicos

Tradicionalmente la validez de un cuestionario, se había presentado como la cualidad del instrumento para medir los rasgos o características que se pretenden medir. Por medio de la validación se trata de determinar si realmente el cuestionario mide aquello para lo que fue creado. Últimamente, el concepto de validez se ha modificado considerablemente. Cronbach en 1971 señalaba que la validación es el proceso por medio del cual el investigador que desarrolla cuestionarios obtiene evidencia para sustentar sus inferencias. Este proceso de validación requiere un estudio empírico dirigido a recolectar la evidencia requerida. La validez se ve como una evaluación -más que una característica- de cuán apropiadas y adecuadas son las interpretaciones y los usos que se hacen de los resultados del cuestionario.

La validez de un instrumento tiene que ver con las preguntas siguientes: ¿qué miden los puntajes del test? y ¿qué predicen dichas puntuaciones? (Guilford, 1954; Nunnally, 1967; Anastasi, 1976; Magnusson, 1982) [26].

La validez no es una propiedad del cuestionario; aunque, por costumbre, se sigue hablando de la validez del cuestionario. La validez es una cuestión de grado. No existe en términos absolutos. No se puede decir que el cuestionario es válido o inválido. Aumenta o disminuye dependiendo de la calidad de la evidencia que la sustenta. Nuevas evidencias pueden incrementarla o reducirla.

Hoy día la validación de una inferencia se presenta como el proceso de determinar si la teoría y las evidencias empíricas respaldan esta inferencia. La validez se refiere siempre a un tipo de uso o interpretación específico. No se puede hablar de la validez de un cuestionario sea cual fuere su uso. A veces los usos son muy próximos, pero aun así hay diferencias. La validez es un concepto unitario. No se puede hablar de diferentes tipos de validez (contenido, constructo, criterio). Se habla más bien de un concepto –validez- y de diversos tipos de evidencia. Sin embargo, en la literatura es común encontrar el término: tipos de validez, para referirse a los diferentes procedimientos para estimar la validez [22].

1.2.2. Procedimientos para estimar la validez

La validez así como la confiabilidad de un instrumento, comprende diferentes aspectos y técnicas de evaluación. Cuando investigamos la validez en un instrumento determinado, intentamos responder tres tipos de cuestiones, que aluden a igual número de tipos de validez. Estas cuestiones son:

1. ¿Cuán representativo es el comportamiento elegido como muestra del universo que se intenta representar? (validez de contenido).

2. ¿Qué significado tiene el comportamiento con respecto a los atributos del individuo que son de interés para la medición? (validez de constructo).

3. ¿Hasta dónde se puede predecir el rendimiento del sujeto o su aprendizaje en un programa de entrenamiento (o hasta dónde se puede anticipar su nivel de desempeño en el trabajo), a partir de su ejecución en la prueba? (validez predictiva).

A continuación solo nos referiremos a la validez de constructo [30].

1.2.2.1. Validez de Constructo

La validez de constructo intenta responder la pregunta ¿hasta dónde un instrumento mide realmente un determinado rasgo latente o una característica de las personas y con cuánta eficiencia lo hace? Esta pregunta tiene sentido, particularmente en los instrumentos que se utilizan en la investigación psicoeducativa, en este campo se hacen mediciones indirectas de ciertas variables internas del individuo que denominamos constructos, [3] y [7].
En consecuencia, es necesario que podamos mostrar evidencia de que, efectivamente, el instrumento mide el rasgo que pretende medir [30].

Cronbach (1960) ha sugerido los pasos siguientes para establecer la validez de constructo: (a) identificar las construcciones que pudieran explicar la ejecución en el instrumento; (b) formulación de hipótesis comprobables a partir de la teoría que enmarca a cada construcción; y (c) recopilación de datos para probar estas hipótesis, [9].

El término constructo se usa en psicología para referirse a algo que no es observable, pero que literalmente es construido por el investigador para resumir o explicar las regularidades o relaciones que él observa en la conducta. Por tanto, la mayoría de los nombres de rasgos se refieren a constructos. Para las preguntas acerca de si el instrumento revela algo significativo respecto de las personas, se usa el término validez de constructo. La validez de constructo es la principal de los tipos de validez, y que la validez de constructo es el concepto unificador que integra las consideraciones de validez de contenido y de criterio en un marco común para probar hipótesis acerca de relaciones teóricamente relevantes. Entre los procedimientos o técnicas estadísticas utilizados para la contrastación de la validez de constructo destaca en mayor medida el Análisis Factorial. En general, podemos decir que esta es la técnica por excelencia utilizada para la validación de constructo, [32].

1.2.3. Importancia de la validez

Si comparamos la confiabilidad con la validez, nos damos cuenta que la obtención de la primera puede ser reducida básicamente a una cuestión técnica. Sin embargo, la validez es mucho más que eso. Tiene que ver con el aspecto sustantivo de la ciencia misma. También se relaciona con la epistemología, en tanto que teoría del conocimiento, y con los paradigmas científicos. No obstante, las dificultades prácticas que se presentan para lograr obtener medidas válidas y confiables, dentro del paradigma de la ciencia clásica, en los últimos años se han desarrollado una serie de métodos, técnicas y procedimientos, que facilitan, cada vez más, esta tarea. Pero, más que el manejo de todo este instrumental tecnológico (métodos estadísticos, procedimientos electrónicos, paquetes computarizados, etc.), lo más importante es que el investigador se haga consciente de la necesidad de utilizar instrumentos apropiados, técnicamente bien calibrados, a fin de garantizar la utilidad y significado de los resultados obtenidos. Queda claro entonces que la construcción de instrumentos de medición no se reduce a la simple presentación de un listado de preguntas en un formato determinado. Construir “buenos” instrumentos de medición es, primero que todo, una tarea técnica, que requiere, por parte del investigador, un entrenamiento específico para acometerla con éxito. Construir un instrumento técnicamente bien hecho implica, en sí mismo, una investigación. De allí que cuando se requiera hacer un estudio (trabajo o tesis de grado, trabajo de ascenso, investigación libre) antes de tomar la seria decisión de construir un instrumento de medición, sin ser un especialista en el área, se debería averiguar previamente acerca de la existencia de la disponibilidad comercial de dicho instrumento en el mercado, o a través de otros investigadores. Si después de esta indagación se llega a determinar que el instrumento no existe y que es indispensable trabajar en el desarrollo del mismo, lo más recomendable sería buscar el asesoramiento técnico especializado, [30].

Capítulo 2

Análisis Factorial

En numerosas áreas de Psicología y de Ciencias del Comportamiento no es posible medir directamente las variables que interesan; por ejemplo, los conceptos de inteligencia y de clase social. En estos casos es necesario recoger medidas indirectas que estén relacionadas con los conceptos que interesan. Las variables que interesan reciben el nombre de variables latentes y la metodología que las relaciona con variables observadas recibe el nombre de Análisis Factorial.

El análisis factorial es una técnica utilizada para descubrir agrupaciones de variables de tal forma que las variables de cada grupo están altamente correlacionadas, y los grupos están relativamente incorrelacionados. De este modo se consigue reducir un número de variables intercorrelacionadas a un número inferior de factores no correlacionados, que permiten explicar la mayor parte de variabilidad de cada una de las variables.

2.1. Conceptos preliminares básicos

El problema de encontrar factores que expliquen los datos fué planteado por primera vez por Charles Spearman (1863-1945), que observó que los niños que obtenían buenas puntuaciones en un test de habilidad mental también las obtenían en otros, lo que le llevó a postular que eran debidas a un factor general de inteligencia, el factor g (Spearman, 1904). L. Thurstone (1887-1955) estudió el modelo con varios factores y escribió uno de los primeros textos de análisis factorial (Thurstone, 1947). El análisis factorial fue considerado hasta los años 60 como una técnica psicométrica con poca base estadística, hasta que los trabajos de Lawley y Maxwell (1971) establecieron formalmente la estimación y el contraste del modelo factorial bajo la hipótesis de normalidad. Desde entonces, las aplicaciones del modelo factorial se han extendido a todas las ciencias sociales, [10].

Los constructos teóricos que son objeto de investigación por las ciencias sociales no son directamente medibles y es necesaria la utilización de indicadores manifiestos para su medición. La relación entre el constructo teórico y sus indicadores manifiestos definirá la validez de la medida obtenida.

Existen modelos formales cuyo cometido se centra en el análisis empírico de las relaciones entre variables observadas y variables latentes; entre ellos, el análisis factorial es el modelo más utilizado en la investigación psicométrica. El análisis factorial engloba un conjunto de modelos matemático estadísticos que analizan las relaciones de dependencia entre variables. Su objetivo es explicar la variabilidad contenida en p variables observadas por medio de m variables latentes, es decir, analizar la estructura interna o dimensionalidad de los datos. El análisis factorial (AF) se haya estrechamente unido al estudio de la validez interna de un test o cuestionario y en el ámbito psicométrico es la técnica más utilizada.

2.2. El modelo del análisis factorial

El modelo matemático del análisis factorial (AF) supone que cada una de las p variables observadas es función de un número m factores comunes ($m < p$) más un factor específico o único. Tanto los factores comunes como los específicos no son observables y su determinación e interpretación es el resultado del AF.

Analíticamente, supondremos un total de p variables observables tipificadas, la existencia de m variables latentes llamadas factores comunes y p factores únicos. El modelo se define de la siguiente forma:

$$
\begin{aligned}
x_1 &= \lambda_{11} f_1 + \cdots + \lambda_{1m} f_m + u_1 \\
x_2 &= \lambda_{21} f_1 + \cdots + \lambda_{2m} f_m + u_2 \\
&. \\
&. \\
&. \\
x_p &= \lambda_{p1} f_1 + \cdots + \lambda_{pm} f_m + u_p
\end{aligned}
$$

que podemos expresar de forma matricial como:

$$X = \Lambda F + U \tag{2.2.1}$$

.

$$X = \begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_p \end{pmatrix}, \Lambda = \begin{pmatrix} \lambda_{11} \cdots \lambda_{1m} \\ \cdot \quad \cdot \quad\quad \cdot \\ \cdot \quad\quad \cdot \quad \cdot \\ \cdot \quad\quad \cdot\ \cdot \\ \lambda_{p1} \cdots \lambda_{pm} \end{pmatrix}, F = \begin{pmatrix} f_1 \\ \cdot \\ \cdot \\ \cdot \\ f_m \end{pmatrix}, U = \begin{pmatrix} u_1 \\ \cdot \\ \cdot \\ \cdot \\ u_p \end{pmatrix}$$

donde:

X: es el vector de las variables observadas.
Λ: es la matriz factorial. Recoge las cargas factoriales ó (saturaciones).
λ_{ij}: es la correlación entre la variable i y el factor j.
F: es el vector de factores comunes.
U: es el vector de factores únicos o factores específicos, [20] y [21].

Hipótesis del modelo
1.- Los factores comunes f_j son incorrelacionados e identicamente distribuidos con media 0 y varianza 1, para $j = 1, 2, ..., m$.

2.- Los factores específicos u_i son incorrelacionados y distribuidos con media 0 y varianza ψ_i^2, para $i = 1, 2, ..., p$.

3.- f_j y u_i tienen distribuciones independientes para todas las combinaciones de i y j con $i = 1, 2, ..., p$ y $j = 1, 2, ..., m$.

El modelo del análisis factorial presenta las siguientes propiedades:

1.- La matriz de covarianzas de X es $Cov(X) = \sum = E(XX^t) - E(X)[E(X)]^t$

como $E(X) = 0$, entonces:

$Cov(X) = \sum = E(XX^t) = [\sigma_{ij}]$
donde:

$$\sigma_{ii} = Var(X_i) \qquad \text{y} \qquad \sigma_{ij} = Cov(X_i, X_j)\ .$$

2.- $Cov(F) = E(FF^t) = I_m$ donde I_m es la matriz identidad de orden m.

3.- $Cov(U) = E(UU^t) = \Psi$ donde $\Psi = diag(\psi_1^2, ..., \psi_p^2)$.

4.- $Cov(F,U) = E(FU^t) = 0$, donde 0 es la matriz cero de orden $m \times p$.

5.- El parámetro λ_{ij} es la covarianza entre la variable X_i y el factor común f_j.

$$
\begin{aligned}
Cov(X,F) &= E(XF^t) \\
&= E((\Lambda F + U)F^t) \\
&= E(\Lambda F F^t + UF^t) \\
&= \Lambda E(FF^t) + E(UF^t) \\
&= \Lambda.
\end{aligned}
$$

6.- La matriz de covarianzas coincide con la matriz de correlaciones, $R = \sum$, la razón es que todas las variables observadas están tipificada. Es decir,

$$
\begin{aligned}
\sum &= E(XX^t) \\
&= R \\
&= \begin{pmatrix} 1 & r_{12} & \cdots & r_{1p} \\ r_{21} & 1 & \cdots & r_{2p} \\ . & & . & . \\ . & & . & . \\ . & & . & . \\ r_{p1} & r_{p2} & \cdots & 1 \end{pmatrix}.
\end{aligned}
$$

Un modelo factorial que verifica las hipótesis anteriores, sobre todo que los factores son incorrelacionadas y de varianza uno, se dice que tiene factores ortogonales y recibe el nombre de modelo factorial ortogonal. Sin embargo, los factores comunes están generalmente correlacionados y su matriz de covarianzas, $E[FF^t]$, no es necesariamente diagonal. Por lo tanto, cuando se requiere un sistema no ortogonal, los factores pueden ser rotados a forma oblicua. Un modelo factorial que tiene factores oblicuos, se denomina modelo factorial oblicuo [2] y [5].

Teoremas fundamentales

TEOREMA **1** (Teorema de Thurstone). *Bajo la hipótesis del modelo factorial se verifica:*

$$\sum = \Lambda\Lambda^t + \Psi. \tag{2.2.2}$$

Demostración.

$$\begin{aligned}\sum &= E(XX^t)\\ &= E[(\Lambda F+u)(\Lambda F+U)^t]\\ &= \Lambda E[FF^t]\Lambda^t + E[UU^t] + \Lambda E[FU^t] + E[UF^t]\Lambda^t\\ &= \Lambda\Lambda^t + \Psi\ .\end{aligned}$$

□

Desarrollando el resultado de Thurstone, se tiene que:

$$\begin{aligned}Cov(X_i, X_j) &= r_{ij}\\ &= \textstyle\sum_{k=1}^{m} \lambda_{ik}\lambda_{jk} \qquad i \neq j \qquad i, j : 1, ..., p.\end{aligned}$$

$$\begin{aligned}Var(X_i) &= 1\\ &= \textstyle\sum_{k=1}^{m} \lambda_{ik}^2 + \psi_i^2 \qquad i : 1, ..., p.\end{aligned}$$

$$\sum = R = \Lambda\Lambda^t + \Psi$$

$$\begin{pmatrix} 1 & r_{12} & \cdots & r_{1p}\\ r_{21} & 1 & \cdots & r_{2p}\\ . & & . & .\\ . & & . & .\\ . & & . & .\\ r_{p1} & r_{p2} & \cdots & 1\end{pmatrix} = \begin{pmatrix}\lambda_{11} \cdots \lambda_{1m}\\ . \quad . \quad .\\ . \quad . \quad .\\ . \quad . \quad . \\ \lambda_{p1} \cdots \lambda_{pm}\end{pmatrix}\begin{pmatrix}\lambda_{11} \cdots \lambda_{p1}\\ . \quad . \quad .\\ . \quad . \quad .\\ . \quad . \quad .\\ \lambda_{1m} \cdots \lambda_{pm}\end{pmatrix} + \begin{pmatrix}\psi_1^2 & 0 & \cdots & 0\\ . & . & & .\\ . & & . & .\\ . & & . & .\\ 0 & \cdot\ \cdot\ \cdot & 0 & \psi_p^2\end{pmatrix}.$$

Los elementos de la diagonal de $\Lambda\Lambda^t$ son llamados comunalidades y los elementos de ψ son llamadas varianzas específicas o unicidades, [10] y [11].

TEOREMA **2.** *La varianza de la variable X_j se descompone de la siguiente forma:*

$$1 = h_j^2 + \psi_j^2$$

donde
h_j^2 es la comunalidad, que se define como parte de la varianza que es debida a los factores comunes.
ψ_j^2 es la unicidad, que se define como parte de la varianza que es debida a los factores específicos.

Demostración.

De $R = \Lambda\Lambda^t + \Psi$ se tiene $1 = \lambda_{j1}^2 + \cdots + \lambda_{jk}^2 + \psi_j^2$ para $j = 1, 2, ..., p$. Ahora bien si se denota $h_j^2 = \lambda_{j1}^2 + \cdots + \lambda_{jk}^2$ se tiene $1 = h_j^2 + \psi_j^2$.

$\square$

TEOREMA **3.** *Se puede reproducir la correlación entre las variables observables a partir de las cargas factoriales:*

$$r_{ij} = \lambda_{i1}\lambda_{j1} + \cdots + \lambda_{im}\lambda_{jm} \ .$$

Demostración.

Se deduce directamente de $R = \Lambda\Lambda^t + \Psi$. $\square$

La matriz de correlaciones reducida, R^*, se obtiene de R substituyendo los unos de la diagonal por las comunalidades.

$$R^* = \begin{pmatrix} h_1^2 & r_{12} & \cdots & r_{1p} \\ r_{21} & h_2^2 & \cdots & r_{2p} \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ r_{p1} & r_{p2} & \cdots & h_p^2 \end{pmatrix}.$$

Evidentemente se verifica que $R = R^* + \Psi$, lo cual implica que [11]:

$$R^* = \Lambda\Lambda^t. \tag{2.2.3}$$

TEOREMA **4.** *Se verifica:*
1. El modelo factorial existe si R es la suma de una matriz semidefinida positiva y una matriz diagonal con elementos no negativos.

2. El número m de factores comunes es el rango de la matriz reducida R^*.

3. Las comunalidades son aquellos valores $0 \leq h_i^2 \leq 1$, tales que R^* es matriz semidefinida positiva (tiene m valores propios positivos).

Demostración.

Es una consecuencia de la relación entre R^* y Λ en la ecuación (2.2.3). La matriz factorial Λ es de rango completo, es decir, $rang(\Lambda) = m$. A su vez, la matriz simétrica $\Lambda\Lambda^t$ es también de rango m y semidefinida positiva y, por ello, el número de factores comunes coincide con el rango de la matriz reducida de correlaciones [10]. $\square$

2.3. Análisis Factorial Exploratorio (AFE)

Los ítems tienen un peso especificado distinto según sea su relación con cada factor; por lo general en cada factor hay ítems con pesos grandes y otros próximos a cero, los ítems que más pesan en cada factor son los que lo definen. El AFE se reduce a la búsqueda de estos pesos, de manera que, expliquen toda la varianza presente en las variables originales, [33].
El AFE nos indica cómo tienden a agruparse los ítems o variables, es decir, el grado de relación (correlación) de cada ítem con cada factor. Los factores comunes equivalen a constructos hipotéticos o conceptos subyacentes o latentes (no observables directamente) deducidos de las correlaciones entre las variables.

El proceso que sigue el AFE se puede resumir en los siguientes tres pasos:
1. Determinar si es pertinente realizar un AF.
2. Elegir el método para extraer los factores, esto es, estimar las saturaciones.
3. Rotar la solución a fin de facilitar su interpretación, [25].

2.3.1. Pertinencias del Análisis Factorial Exploratorio

Cuando se pretende analizar la conveniencia de la aplicación del AF a un conjunto de variables, se analizan criterios y se realizan contrastes de hipótesis previos a la extracción de los factores. Entre ellos destacamos los siguientes:

1.- La evaluación de coeficientes de correlación de Pearson de las variables observables altamente significativas.

El cual consiste en observar la matriz de correlaciones entre las variables que entran en el análisis, esto se realiza a partir de la matriz de datos originales. Se trata simplemente de comprobar si existe un gran número de altas correlaciones. Para saber si estas correlaciones son significativas podemos hacer pruebas de hipótesis sobre los coeficientes de correlación. Las hipótesis de la prueba son $H_0 : \rho = 0 \;\; vs \;\; H_1 : \rho \neq 0$.
Usualmente, en los softwares, se proporciona el grado de significación de cada una de estas correlaciones, los valores de p. Se rechazará a la hipótesis nula, $H_0 : \rho = 0$, cuando el valor de p es menor que el valor de significancia α.

2.- Que el determinante de la matriz de correlación sea relativamente bajo, próximo a cero.

3.- Que el cálculo del índice de Kaiser-Meyer-Olkin (KMO) $> 0{,}60$.

Un coeficiente de correlación parcial mide la correlación existente entre dos variables una vez que se han descontado los efectos lineales de otras variables. En un modelo factorial se pueden interpretar esos efectos de otras variables como los correspondientes a factores comunes. Por lo tanto, el coeficiente de correlación parcial entre dos variables sería equivalente al coeficiente de correlación entre los factores específicos de dos variables. De acuerdo con el modelo de AF los coeficientes de correlación teóricos calculados entre cada par de factores específicos o comunes son nulos por hipótesis. Si los coeficientes de correlación parcial constituyen una aproximación a dichos coeficientes teóricos, deben estar próximos a cero, [20].
La medida del estadístico KMO se define como:

$$KMO = \frac{\sum_{i\neq j}\sum_{i\neq j} r_{ij}^2}{\sum_{i\neq j}\sum_{i\neq j} r_{ij}^2+\sum_{i\neq j}\sum_{i\neq j} a_{ij}^2}$$

donde:
r_{ij}: representa el coeficiente de correlación simple entre las variables i- ésima y j- ésima.
a_{ij}: representa la correlación parcial entre las variables i- ésima y j- ésima.

En el caso de que exista una adecuación de los datos a un modelo de AF, el término del denominador que contiene los coeficientes a_{ij}, será pequeño y, consecuentemente, la medida del estadístico KMO estará próxima a 1.
Un valor de la medida KMO de 0.80 a 0.90 es muy bueno, mientras que los valores por debajo de 0.50 no son aceptables.

4.- Que el resultado del test de Bartlett sea significativo.

El objetivo es comprobar que la matriz de correlaciones es significativamente distinta de la matriz identidad, o sea, $H_0 : R = I \;\; vs \;\; H_1 : R \neq I$. Si la matriz de correlaciones fuera la matriz identidad no habrá correlación entre variables y no tendría sentido llevar a cabo un AF. El test de esfericidad de Bartlett nos permite hacer esta comparación. Si se considera que $H_0 : R = I \Leftrightarrow |R| = 1$, las hipótesis se pueden plantear de la siguiente manera: $H_0 : |R| = 1 \;\; vs \;\; H_1 : |R| \neq 1$. La hipótesis alternativa asume que el determinante de R, indicador de la varianza generalizada de dicha matriz, es distinto de uno. Un determinante próximo a cero indica que una o más variables pueden expresarse como combinación lineal de las otras variables. Rechazar H_0 sería indicativo de correlaciones entre las variables observables y tendrá sentido el AF [32].

El estadístico de prueba es:

$$\chi^2 = -\left[n - 1 - \frac{1}{6(2p+5)}\right] ln\,|R|$$

donde:
n: es la dimensión de la muestra.
p: es el número de variables observadas que entran a formar parte de la matriz de correlaciones.
$|R|$: es el determinante de la matriz de correlaciones (observada).

La utilización de este test presupone que los datos provienen de una distribución normal multivariante y bajo este supuesto el estadístico de prueba se distribuye χ^2 con $\frac{1}{2(p^2-p)}$ grados de libertad, [20].

2.3.2. Métodos de estimación para la obtención de factores

Existen diferentes métodos de estimación para obtener los factores comunes, aquí expondremos el de Componentes Principales (CP) y las bases del procedimiento de estimación por máxima verosimilitud (ML), por ser la que utiliza R-Commander para ejecutar el Análisis Factorial.

2.3.2.1. Componentes Principales

Consiste en estimar las puntuaciones factoriales mediante las puntuaciones tipificadas de las primeras k componentes y la matriz de cargas factoriales mediante las correlaciones de las variables originales con dichas componentes.

Recordemos que las componentes principales se expresan en función de las variables originales como:
$C_1 = c_{11}X_1 + \cdots + c_{1p}X_p$
$C_2 = c_{21}X_1 + \cdots + c_{2p}X_p$
.
.
.
$C_p = c_{p1}X_1 + \cdots + c_{pp}X_p$.
Se puede demostrar que este sistema es reversible. De modo, que las variables originales se expresan en función de las componentes

$$X_1 = c_{11}C_1 + \cdots + c_{1p}C_p$$
$$X_2 = c_{21}C_1 + \cdots + c_{2p}C_p$$
$$\vdots$$
$$X_p = c_{p1}C_1 + \cdots + c_{pp}C_p.$$

Una hipótesis que plantea el modelo factorial es que las componentes deben estar tipificadas, ahora bien, las componentes C no están tipificadas, esto se puede resolver usando las componentes tipificadas, las cuales se obtienen al dividir cada componente por su desviación típica, es decir,

$$Z_k = \frac{C_k}{\sqrt{\theta_k}} \qquad \text{para k=1,2,...,p,}$$

despejando C_k se tiene que

$$C_k = \sqrt{\theta_k} Z_k \qquad \text{para k=1,2,...,p.} \tag{2.3.1}$$

Sustituyendo (2.3.1) en las expresiones de las variables originales se tiene,

$$X_j = c_{1j}\sqrt{\theta_1}Z_1 + c_{2j}\sqrt{\theta_2}Z_2 + \cdots + c_{pj}\sqrt{\theta_p}Z_p \qquad \text{para j=1,2,...,p.}$$

Pero teniendo en cuenta que $c_{kj}\sqrt{\theta_k} = r_{kj}$, entonces

$$X_j = r_{1j}Z_1 + r_{2j}Z_2 + \cdots + r_{pj}Z_p \qquad \text{para j=1,2,...,p.} \tag{2.3.2}$$

Si agrupamos los ultimos $(p - m)$ términos en (2.3.2) tenemos

$$X_j = r_{1j}Z_1 + r_{2j}Z_2 + \cdots + r_{mj}Z_m + (r_{m+1,j}Z_{m+1} + \cdots + r_{pj}Z_p). \tag{2.3.3}$$

Por otro lado, la ecuación de la variable j-ésima en el modelo factorial es de la forma

$$X_j = \lambda_{j1}f_1 + \lambda_{j2}f_2 + \cdots + \lambda_{jm}f_m + u_j. \qquad \text{para j=1,2,...,p.} \tag{2.3.4}$$

Comparando (2.3.3) y (2.3.4), podemos ver que los m factores f_k se estiman mediante las m primeras componentes principales tipificadas Z_k y además se tiene las estimaciones de los coeficientes λ_{jk}:

$$\widehat{\lambda}_{jk} = r_{kj}.$$

Dadas las estimaciones anteriores se obtiene la estimación de la comunalidad

$$\widehat{h}_j^2 = \widehat{\lambda}_{j1}^2 + \widehat{\lambda}_{j2}^2 + \cdots + \widehat{\lambda}_{jm}^2,$$

y la estimación del factor único u_j

$$\widehat{u}_j = r_{m+1,j}Z_{m+1} + r_{m+2}Z_{m+2} + \cdots + r_{pj}Z_p.$$

Finalmente, la especificidad o unicidad, es decir, la parte de la varianza debida al factor específico, será:

$$\widehat{\psi}_j^2 = 1 - \widehat{h}_j^2.$$

2.3.2.2. Máxima verosimilitud

Este método está basado en el modelo dado por la ecuación $X = \Lambda F + U$, asumiendo la hipótesis de normalidad multivariante, consiste en aplicar el modelo de máxima verosimilitud. Puede emplearse con la matriz de correlaciones o con la matriz de covarianzas muéstrales. Este método tiene la ventaja de que las estimaciones obtenidas no dependen de la escala de medida de las variables.

Supongamos que las n observaciones de las p variables provienen de una distribución normal con μ y matriz de covarianzas $(\sum)$, es decir, $X = (X_1, ..., X_p)^t \sim N_p(\mu, \sum)$. Entonces, el logaritmo de la función de verosimilitud puede ser expresado como:

$$£\left(\sum\right) = \ln c - \frac{n}{2}\ln\left|\sum\right| + \frac{n}{2}(n-p-1)\ln|S| - \frac{n}{2}tr\left(\sum^{-1}S\right), \quad (2.3.5)$$

en donde:
c: es un término constante.
n: es el número de observaciones.
p: es el número de variables observables.
$\sum$: es la matriz de covarianzas poblacionales.
S: es la matriz de covarianzas muéstrales.

El máximo de la función dada en la ecuación (2.3.5) debe ser, omitiendo funciones constantes de las observaciones, igual a:

$$M = -\frac{n}{2}\left[\ln\left|\sum\right| + tr\left(S\sum^{-1}\right)\right] = -\frac{n}{2}\left[\ln\left|\Lambda\Lambda^t + \Psi\right| + tr(S(\Lambda\Lambda^t + \Psi)^{-1})\right].$$

La estimación de las cargas factoriales se obtiene minimizando la función

$$F\left(\sum\right) = \ln\left|\sum\right| + tr\left(S\sum^{-1}\right) - \ln|S| - p, \quad (2.3.6)$$

con $\sum = \Lambda\Lambda^t + \Psi$. Dicha minimización es equivalente a maximizar la función dada por la ecuación (2.3.5).
Las derivadas respecto de Λ y Ψ son

$$\frac{\partial F}{\partial \Lambda} = 2\sum^{-1}\left(\sum - S\right)\sum^{-1}\Lambda,$$

$$\frac{\partial F}{\partial \Psi} = diag\left(\sum^{-1}\left(\sum - S\right)\sum^{-1}\right).$$

Por lo tanto, las ecuaciones a resolver para obtener estimaciones de Λ y Ψ son:

$$\sum^{-1}\left(\sum - S\right)\sum^{-1} L = 0, \qquad diag\left(\sum^{-1}\left(\sum - S\right)\sum^{-1}\right) = 0$$

con las restricciones, $\sum = \Lambda\Lambda^t + \Psi$ y $\Lambda^t \sum^{-1} \Lambda$ es diagonal.

La última condición es sólo una restricción para concretar una solución, puesto que si Λ es solución, también lo es ΛT, siendo T matriz ortogonal.

Con cualquier método de estimación conseguir una estructura factorial clara y simple depende en buena medida del número de factores. El número de factores que se desea extraer debe estar bien justificado, por lo tanto, exponemos y valoramos otros procedimientos para determinar el número de factores que se deben extraer en el AFE.

1.- En el procedimiento de Kaiser-Gutman se extraen y rotarán solamente los factores que en el primer análisis (antes de las rotaciones) tienen un eigenvalor (o valor propio) mayor de 1. Este procedimiento tiene el inconveniente de que pudieran existir diferencias muy pequeñas (por ejemplo entre un factor con un eigenvalor de 1.01 y otro de 0.99).

2.- Utilizar el gráfico de sedimentación, en el que aparecen en el eje X el número de componentes o factores y en el eje Y los eigenvalores o varianza explicada por cada factor. El punto de corte para establecer el número de factores que se van a rotar se sitúa en el punto de inflexión de la línea descendente que va uniendo los diversos eigenvalores.

3.- Para reducir el número de factores de manera no arbitraria se pueden encontrar varias orientaciones, como eliminar los factores en los que ninguna variable tiene un peso superior a 0.30 [26].

2.3.3. Rotación de factores

La matriz de cargas factoriales tiene un papel importante para interpretar el significado de los factores. Cuando los factores son ortogonales cuantifican el grado y tipo de la relación entre éstos y las variables originales. En la práctica, los métodos de extracción de factores pueden no proporcionar matrices de cargas factoriales adecuadas para la interpretación.

En ocasiones es difícil interpretar el significado de los factores a partir de la matriz de cargas factoriales, sobre todo si aparecen varios factores compartiendo variables. Puede haber distintas variables que muestren correlaciones altas con varios factores haciendo difícil su interpretación. Las rotaciones son transformaciones lineales que facilitan la interpretación sin alterar la proporción de varianza explicada por los factores [25].

La rotación de factores transforma la matriz factorial inicial en otra denominada matriz factorial rotada, de más fácil interpretación. La matriz factorial rotada es una combinación lineal de la primera y explica la misma cantidad de varianza inicial. El objetivo de la rotación es intentar aproximar la matriz factorial al principio de estructura simple (Thurstone 1947, citado en [20]). Según este principio, la matriz factorial debe reunir las siguientes características:

1.- Cada factor debe contener cargas altas y cargas próximas a cero.
2.- Cada variable debe ser explicada por un solo factor.
3.- No deben existir factores con la misma distribución, es decir, los factores distintos deben presentar distribuciones de cargas altas y bajas distintas.

Estas características no suelen lograrse, lo que se trata es alcanzar una solución lo más aproximada posible a ello.

Existen varios métodos de rotación que podemos agrupar en dos grandes tipos: ortogonales y oblicuos. Tanto en las rotaciones ortogonales como en las de tipo oblicuo, la comunalidad de cada variable no se ve modificada; sin embargo, cambia la varianza explicada por cada factor. Los dos tipos de rotación son útiles y suele recomendarse hacer los dos, pero si se tiene que elegir el de interpretación más sencilla, lo habitual es optar por la rotación ortogonal que suele ser de hecho la opción preferida, aunque no necesariamente la mejor, [8].

2.3.3.1. Rotaciones ortogonales

En las rotaciones ortogonales los factores se rotan de tal modo que los ángulos entre ellos sean siempre ángulos rectos. La rotación ortogonal de factores hace variar las cargas factoriales y, por lo tanto, el significado de los factores. Sin embargo, las diferentes soluciones factoriales analíticas son matemáticamente equivalentes ya que explican la misma cantidad de varianza en cada variable y en el conjunto de la matriz. Además los factores rotados reproducen las correlaciones de forma precisa al igual que las soluciones no rotadas, [20].

TEOREMA **5.** *Toda rotación ortogonal de una solución es también solución.*

Demostración.
Si $m > 1$, la matriz factorial no es única, es decir, si existen Λ y Ψ de modo que $R = \Lambda\Lambda^t + \Psi$, entonces para toda matriz ortogonal T:

$$R = (\Lambda T)(T^t\Lambda^t) + \Psi = (\Lambda T)(\Lambda T)^t + \Psi.$$

□

Aunque los elementos ΛT son diferentes de los coeficientes iniciales Λ, su capacidad para generar las covarianzas no cambia. Por lo tanto, toda rotación ortogonal de una solución es también solución, [10], [20] y [21].

Las rotaciones ortogonales se basan en la idea de maximizar la varianza de los cuadrados de las cargas factoriales, con lo que se consigue que los valores se dispersen al máximo, aumentando los mayores y disminuyendo los más pequeños. Dada una matriz factorial Λ, queremos encontrar una matriz ortogonal T tal que la nueva matriz factorial $B = \Lambda T$ defina unos factores que tengan una estructura más simple. Dentro de los métodos de rotación ortogonal se encuentran las rotaciones Quartimax, Varimax y Equamax. El más popular de estos es el procedimiento de rotación Varimax, al que nos referiremos a continuación.

Rotación varimax

El principal objetivo del método Varimax, es eliminar las saturaciones negativas y describir los datos por tan pocas saturaciones como sea posible. Esto se realiza, como veremos, maximizando una cierta función de los cuadrados de las saturaciones a través de un procedimiento iterativo.

Se describirá un procedimiento que es una variante del método Varimax original ideado por Kaiser en 1958. Lo esencial de este procedimiento es que en cada iteración todos los factores se rotan simultáneamente. Denotemos por $Ł_0$ la matriz de $p \times k$ de factores no rotados. La i-ésima fila la de $Ł_0$, denotada por l_i, es un vector de orden k. Sea T de $k \times k$ una matriz de rotación ortogonal cuya r-ésima columna es τ_r con $r = 1, ..., k$. La matriz $\Lambda = [\lambda_{ir}]$ de cargas factoriales rotadas está dada por $\Lambda = Ł_0 T$ y así $\lambda_{ir} = l_i \tau_r$. Definiremos los escalares

$$d_r = \sum_{i=1}^{p} \lambda_{ir}^2$$

$$d_r = \sum_i (l_i \tau_r)^2.$$

Así, d_r es la suma de los cuadrados de las saturaciones de la r-ésima columna de Λ. El criterio χ, maximizado en el método Varimax simultaneo, está dado por

$$\begin{aligned}\chi &= \sum_{r=1}^{k} \left[\sum_{i=1}^{p} \left(l_{ir}^2 - \frac{d_r}{p} \right)^2 \right] \\ &= \sum_r \left[\sum_i \lambda_{ir}^4 - \frac{d_r^2}{p} \right] \\ &= \sum_r \left[\sum_i (l_i \tau_r)^4 - \frac{d_r^2}{p} \right].\end{aligned}$$

De esta definición es claro que χ representa la suma de los cuadrados de las desviaciones de los valores de λ_{ir}, cada uno medido como una desviación de las correspondientes columnas de media $\frac{d_r}{p}$. La maximización es con respecto a los elementos de τ.
Puesto que las columnas de τ satisfacen las condiciones $\tau_r^t \tau_r = 1$ y $\tau_r^t \tau_s = 0$ $(r \neq s)$, debemos igualar a cero la derivada con respecto a los elementos de τ la expresión:

$$y = \chi - 2\sum_r \sum_s a_{rs}\tau_r^t \tau_s$$

donde los coeficientes a_{rs} son multiplicadores indeterminados tales que A_{sr} es idéntico a A_{rs}.

Usando las expresiones anteriores encontramos que la derivada de y con respecto al vector τ_s está dado por:

$$\frac{\partial y}{\partial \tau_s} = 4\sum_i \left[(l_i\tau_s)^3 l_i - \left(4\frac{d_s}{p}\right)\sum_i (l_i\tau_s l_i) - 4\sum_r a_{rs}\tau_r \right]$$

$$\frac{\partial y}{\partial \tau_s} = 4\sum_i \left(c_{is} - d_s\frac{\lambda_{is}}{p} \right) l_i - 4\sum_r A_{rs}\tau_r,$$

donde $c_{ir} = (l_i\tau_r)^3 = (\lambda_{ir})^3$.

La expresión $\frac{\partial y}{\partial \tau_s}$ es la s-ésima columna de la matriz

$$4\left[L_0 C \left(\frac{1}{p}\right) L_0 \Lambda D - \tau A \right],$$

donde C es la matriz $p \times k$ con elementos c_{ir}, A es la matriz simétrica de orden k con elementos A_{rs} y D es la matriz diagonal cuyos elementos son $d_1, ..., d_k$. Así, tomando en cuenta todos los valores de s, tenemos que $\frac{\partial y}{\partial \tau} = 4(B - \tau A)$ con $B = L_0\left[C - \left(\frac{1}{p}\right)\Lambda D\right]$.

La matriz ortogonal τ que maximiza χ satisface así la ecuación $\tau A = B$. Puede demostrarse que si A es definida positiva y simetrica; la ecuación anterior corresponde a un máximo para χ. Entonces, de $\tau A = B$ tenemos $A = \tau^t B = \Lambda^t C - \left(\frac{1}{p}\right)\lambda^t \Lambda D$. Así la matriz del lado derecho es simétrica cuando χ se máximiza. El elemento r-ésimo diagonal de A es:

$$\sum_i \lambda_{ir}^4 - \frac{1}{p}\left(\sum_i \lambda_{ir}^2\right)^2.$$

De la definicion de χ se ve que $tr(A)$ es su valor máximo. Las matrices τ, A y B que satisfacen $\tau A = B$ se encuentran por un procedimiento iterativo. Comenzemos con una aproximacion inicial τ, esta produce una aproximación inicial $L_1 = L_o\tau_1$ para Λ, y de L_1 obtenemos aproximaciones C_1, D_1 y B_1 para C, D y B respectivamente. Ahora se tiene que encontrar una matriz simétrica y definida positiva A y una matriz ortogonal τ_2 que satisfaga la ecuación $\tau_2 A_1 = B_1$. Si multiplicamos cada lado de esta ecuación por su transpuesta tenemos $A_1^2 = B_1^t B_1$.

La matriz $B_1^t B_1$ es una matriz simétrica y definida positiva de orden k. Por lo tanto, podemos expresarla en la forma $U\Delta U$, donde U es ortogonal y Δ es matriz diagonal de elementos positivos. Así, es posible encontrar A_1 en la forma $A_1 = u\Delta U^t$ y de $\tau_2 A_1 = B_1$ se tiene que τ_2 esta dado por $\tau_2 = B_1 A_1^{-1}$.

El procedimiento se repite con τ_3 en lugar de τ_2 y si es necesario, con una sucesión de matrices τ_s. Puede demostrarse que para τ_1 suficientemente cerca de τ, esta sucesión converge a τ. La sucesión de matrices A_s, correspondiente es tal que, los valores de $tr(A_s)$ forman una sucesión ascendente de valores que convergen al máximo del criterio χ. En la práctica una buena aproximación inicial de τ es innecesaria, así que comúnmente se toma $\tau = I$ y $L_1 = L_0$. En general, se requieren pocas iteraciones, [21].

2.4. Análisis Factorial Confirmatorio (AFC)

Cuando el investigador tiene suficientes conocimientos previos para formular hipótesis concretas sobre la relación entre indicadores y dimensiones latentes, su interés se centra en contrastar estas hipótesis. Por ejemplo, al traducir o adaptar cuestionarios ya desarrollados sabemos qué ítems deberían medir qué dimensiones. El modelo de análisis factorial confirmatorio (AFC) corrige las deficiencias inherentes a la perspectiva exploratoria y conduce a una mayor concreción de las hipótesis que deben ser contrastadas. Su especificación difiere de la perspectiva exploratoria en aspectos esenciales como:

– Permitir restricciones en algunas saturaciones. Lo habitual es suponer la validez de cada ítem, es decir, que satura en un único factor. Se delimita así el concepto de factor común a aquel que subyace únicamente a sus indicadores concretos y se evita introducir factores de difícil interpretación.

– Permitir contrastes estadísticos de las hipótesis especificadas.

– Permitir componentes únicas correlacionadas. Aunque es un recurso poco elegante, se justifica por la existencia de otros factores sin interés, como un método de medición común que no se desea explicitar en la especificación

[18].

2.4.1. El modelo del Análisis Factorial Confirmatorio

La estructura de covarianzas del AFC es muy similar a la del AFE, a la que simplemente se le impondrán algunas restricciones. Para introducirnos en el AFC es necesario presentar una serie de convenciones y términos no utilizados hasta el momento. Los ítems o indicadores o variables manifiestas u observables, son aquellas que se miden directamente. Por lo general, se les asignan letras latinas mayúsculas, como X y Y. Los factores o variables latentes, no observables, son aquellas que no pueden ser medidas directamente, se les denota con las letras griegas, como ε y δ, [1].

Desde el punto de vista del AFC, la respuesta de cada sujeto en cada ítem esta generada por unas variables no observadas (factores latentes) que explican la variabilidad de las respuestas en el ítem. Previsiblemente, el factor latente nunca explicará de forma totalmente satisfactoria la variabilidad de las respuestas del ítem. A esta parte no explicada por el factor se le denomina error de medida (también llamado unicidades o factores específicos), [16].

Suponga que se observa un vector X de respuestas, p -variado, de una población que tiene media μ y matriz de varianzas y covarianzas $\sum$ tales que:

$$X_1 = \sum_{j=1}^{m} \lambda_{ij}\xi_j + \delta_i \qquad para \qquad i = 1, 2, ..., p.$$

donde:
$X_1, X_2, ..., X_P$: son las variables observadas.
$\xi_1, \xi_2, ..., \xi_m$: son los factores.
λ_{ij}: es el peso (o contribución) del factor común j en la variable i.

λ_{ij} se llaman cargas, pesos o saturaciones factoriales. En forma matricial el modelo quedaría de la siguiente forma, [1] y [19]:

$$X = \Lambda\xi + \delta \tag{2.4.1}$$

$$
\begin{aligned}
x_1 &= \lambda_{11} f_1 + \cdots + \lambda_{1m}\xi_m + \delta_1 \\
x_2 &= \lambda_{21} f_1 + \cdots + \lambda_{2m}\xi_m + \delta_2 \\
&\ \ \vdots \\
x_p &= \lambda_{p1} f_1 + \cdots + \lambda_{pm}\xi_m + \delta_p.
\end{aligned}
$$

El modelo se suele representar en un diagrama de flujos, acorde con su especificación. Convencionalmente, los rectángulos representan ítems y las elipses, factores comunes. Flechas unidireccionales entre factores comunes e ítems expresan saturaciones. Flechas bidireccionales indican correlaciones entre factores comunes o únicos. La Figura 2.1 muestra los diagramas de dos posibles modelos de AFE y de AFC, [18]. En el modelo de AFC, los factores únicos de las variables v_1 y v_4 que podrían compartir método de medición están correlacionados. Se resalta que v_1, v_2 y v_3 son indicadores exclusivamente de f_1 mientras que v_4, v_5 y v_6 lo son sólo de f_2.
En un principio, los programas para estimar modelos de AFC eran escasos y requerían conocimientos de álgebra matricial. Actualmente, existe una gran variedad de ellos, todos accesibles y sencillos de utilizar (en algunos, el usuario se limita a dibujar el diagrama del modelo) que permiten estimar cualquier modelo de ecuaciones estructurales, [18].

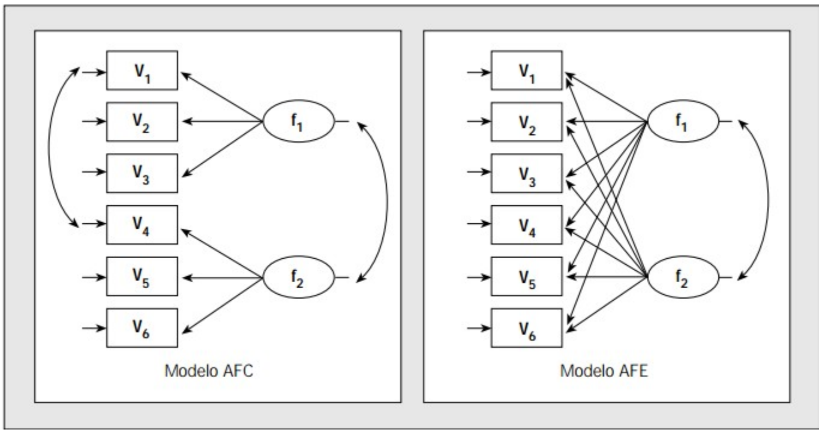


Figura 2.1: Diferencias entre el diagrama de flujos de los modelos de Análisis Factorial Confirmatorio y Análisis Factorial Exploratorio con 6 variables y 2 factores.

2.4.2. El Análisis Confirmatorio frente al Exploratorio

Los análisis factoriales exploratorio y confirmatorio no se distinguen sólo por la intención del investigador (explorar, ver qué sale, o confirmar hipótesis previamente establecidas); suponen también procedimientos (y programas de computadora) distintos. La principal diferencia conceptual entre los dos análisis es que, mientras que el AFE consiste en una búsqueda de relaciones subyacentes, en el AFC el investigador parte de un modelo a priori de dichas relaciones. Explícitamente las principales características distintivas del AFC son: (a) el número de factores es establecido de antemano por el investigador; (b) el investigador decide de antemano las saturaciones de cada variable observable sobre cada factor y (c) la relación entre los factores en el modelo se especifica de antemano, [17] y [19].

En el cuadro 2.1, se presentan de manera resumida las características del AFE y del AFC. La principal particularidad del análisis confirmatorio es que la estructura del modelo es totalmente controlable y manipulable por el investigador, por lo que en el cuadro se presentan aquellas condiciones que son más habituales y que constituyen el paradigma del análisis, [14] y [19].

AFE	AFC
Todos los factores comunes están correlacionados (rotación oblicua) o bien ningún factor correlaciona (rotación ortogonal).	El investigador determina qué factores comunes están correlacionados.
Todas las variables observadas están afectadas por todos los factores comunes.	El investigador determina qué factores comunes afectan a qué variables observadas.
Los errores específicos no están correlacionados entre sí.	Habitualmente no existe correlación entre los términos de error.
Todas las variables observadas están afectadas por un término de error.	Las variables observadas tienen asociado un error de medida. Este error puede ser fijado a un valor nulo.
Todos los factores no están correlacionados con los términos de error.	Habitualmente no existe correlación entre los términos de error y los factores comunes.

Fuente: Tomado de Lévy & Varela .

Cuadro 2.1: Diferencias entre el Análisis Factorial Confirmatorio y Exploratorio. Fuente: Lévy & Varela

Las hipótesis del AFE (clásico) son muy rígidas, y no permiten incorporar elementos importantes del conocimiento sustantivo de los expertos. Las limitaciones del AFE son ampliamente superadas en el AFC al poder especificar suposiciones más realistas, [14].

Una ventaja que se ha señalado frecuentemente a favor del análisis confirmatorio es la posibilidad que tiene el investigador para establecer relaciones entre los factores. En el acercamiento tradicional, o ningún factor correlaciona -rotación ortogonal- o todos lo hacen -rotación oblicua-, mientras que en el acercamiento confirmatorio puede establecerse a priori un conjunto de condiciones más flexibles en torno a la relación entre los factores; por ejemplo, que dos correlacionen entre y otros dos no estén correlacionados, [17].

A pesar de las diferencias observadas entre ambos procedimientos, la estructura factorial obtenida mediante el AFE constituirá una aproximación valida al futuro modelo confirmatorio, y las cargas factoriales extraídas podrán ser empleadas para fijar de antemano algún parámetro a un valor determinado, [19].

De esta forma, el investigador que pretenda medir un constructo debería identificar previamente sus dimensiones subyacentes y establecer una serie de variables observables como indicadores de esas dimensiones latentes. El AFC contrastaría los datos con el modelo teórico presentado y calcularía índices de ajuste que informarán si dicho modelo constituye una representación plausible de la realidad, [19].

2.4.3. Fases del Análisis Factorial Confirmatorio

Como ya se ha mencionado, el AFC tiene como objetivo determinar si un modelo de medida especificado por el investigador, basándose en hipótesis teóricas o en un AFE previo, es consistente con la realidad. Para llegar a obtener alguna conclusión al respecto, es preciso abordar una serie de fases, comunes al conjunto de los procedimientos que operan con los SEM. Las fases esenciales en la ejecución del AFC son:

1. La especificación del modelo.
2. La identificación del modelo.
3. La estimación de parámetros.
4. La evaluación del ajuste del modelo, [19].

2.4.3.1. Especificación del modelo

Para establecer la estructura del modelo el investigador se basa en estudios previos (como el AFE) o en un sólido sustento teórico. Establecer formalmente un modelo implica tomar decisiones respecto a los siguientes aspectos: (1) el número de factores comunes, (2) el número de variables observables, (3) la relación entre las variables observables y los factores comunes, (4) la relación entre los factores comunes, (5) la relación entre los factores específicos y las variables observables, y (6) las relaciones entre factores específicos.

La relación entre las variables observadas y las variables latentes se expresa como:

$$X = \Lambda\xi + \delta \tag{2.4.2}$$

ésta ecuación implica que cada variable observable sería función de la contribución de cada factor común y el error de medida o factores específicos que tiene asociado.

En las ecuaciones y análisis que se presentarían a lo largo de este capítulo se sigue el procedimiento habitual de asumir que todas las variables están desviadas respecto a sus medias, con un valor esperado igual a cero, $E[X] = E[\xi] = E[\delta] = 0$, lo que explica la ausencia de un término constante. Se asume además la independencia entre factores comunes y específicos: $cov[\xi\delta^t] = 0$.

Si denotamos como $\sum$ a la matriz de covarianzas entre las variables observables (X) del modelo (2.4.2), resulta que:

$$\begin{aligned}
\textstyle\sum &= E(XX^t) - E(X)[E(X)]^t \\
&= E\left[(\Lambda\xi + \delta)(\Lambda\xi + \delta)^t\right] \\
&= \Lambda E[\xi\xi^t]\Lambda^t + \Lambda E[\xi\delta^t] + E[\delta\xi^t]\Lambda^t + E[\delta\delta^t].
\end{aligned}$$

Si hacemos $\Phi = E[\xi\xi^t]$ y $\Theta = E[\delta\delta^t]$ entonces $\sum$ puede escribirse como:

$$\sum = \Lambda\Phi\Lambda^t + \Theta. \tag{2.4.3}$$

Es muy importante, para desarrollos posteriores analizar el contenido de la expresión (2.4.3) y centrar nuestra atención en los parámetros: $\sum$, Φ, Θ. De la ecuación (2.4.3), se tiene:

La matriz $\sum = E(XX^t) = [\sigma_{ij}]$ siendo $\sigma_{ii} = Var(X_i)$ y $\sigma_{ij} = Cov(X_i, X_j)$ con $i, j = 1, ...p$. El número de elementos distintos es: $\frac{p(p+1)}{2}^3$.

La matriz Λ contiene las $p \times m$ saturaciones o cargas factoriales, siendo λ_{ij} la saturación de la variable X_i con el factor común ξ_j.

La matriz $\Phi = E\left[\xi\xi^t\right] = [\phi_{rs}]$ siendo $\phi_{rr} = Var\left(\xi_r\right)$ y $\phi_{rs} = Cov\left(\xi_r, \xi_s\right)$ con $r, s = 1, ...m$. El número de elementos distintos: $\frac{m(m+1)}{2}$.

La matriz $\Theta = E\left[\delta\delta^t\right] = [\theta_{ij}]$ siendo $\theta_{ii} = Var\left(\delta_i\right)$ y $\theta_{ij} = Cov\left(\delta_i, \delta_j\right)$ con $i, j = 1, ...p$. El número de elementos distintos: $\frac{p(p+1)}{2}$.

La matriz de coeficientes de estructura $Q = \Lambda\Phi = [q_{ij}]$ siendo $q_{ij} = corr\left(X_i, \xi_j\right)$.

Por lo tanto, la ecuación (2.4.3) expresa los $\frac{p(p+1)}{2}$ elementos distintos de $\sum$ en función de $p \times m + \frac{m(m+1)}{2} + \frac{p(p+1)}{2}$ parámetros desconocidos de las matrices Λ, Φ y Θ. Así, los parámetros que se deberían estimar aparecen vinculados mediante la expresión (2.4.3), a los valores de las varianzas y covarianzas poblacionales de las variables observadas, [1] y [19].

El AFC se reduce, a grandes rasgos, a obtener estimaciones de las matrices Λ, Φ y Θ que hagan que la matriz de varianzas y covarianzas poblacional estimada $\sum$ obtenida a partir de ellas, sea lo más parecida posible a la matriz de varianzas y covarianzas muestral que se obtiene a partir de los valores muéstrales de las variables observadas. Pero, para poder entrar en el procedimiento de estimación, es necesario abordar previamente el problema de la identificación que se plantea en el método del AFC, [1].

2.4.3.2. Identificación del modelo

Hemos visto que en el método de AFC disponemos de información sobre las varianzas y covarianzas muéstrales de las variables observables y con ella hemos de estimar una serie de parámetros (cargas factoriales, varianzas y covarianzas de los factores comunes, y las varianzas y covarianzas de los factores específicos). Al igual que ocurre con un sistema de ecuaciones lineales, podemos disponer en principio de más ecuaciones que incógnitas, del mismo o de mayor número de incógnitas que ecuaciones. Pues bien, la identificación del modelo de AFC hace referencia, precisamente, a la cuestión de si los parámetros del modelo pueden o no ser determinados de forma única. De esta forma, la matriz $\sum$ (o en su caso, la matriz muestral S) es la fuente de la identificación y cada parámetro de las matrices Λ, Φ y Θ corresponderá con un parámetro o una combinación lineal de parámetros de $\sum$, [1] y [19].

A la hora de determinar si un modelo está o no identificado, caben tres soluciones posibles:

1. Que esté exactamente identificado, en cuyo caso se podrá estimar cada parámetro estructural a partir de una única combinación de los elementos de $\sum$, por lo que tendrían una única solución.

2. Que el modelo esté sobre identificado, estando todos los parámetros identificados, al menos un parámetro podrá obtenerse a partir de dos o más ecuaciones diferentes.

3. Que el modelo esté infra identificado, en cuyo caso no sería posible establecer ecuaciones de covarianza para alguno de los parámetros, por lo que no todos podrían ser estimados.

En palabras de Long (1983) citado en [1], si se intenta estimar un modelo que no esté identificado, los resultados que se obtendrían serán estimaciones arbitrarias de los parámetros lo que desembocará en interpretaciones carentes de sentido. Si no se imponen restricciones a los parámetros a estimar, necesariamente habrá un número infinito de soluciones posibles para los mismos. Consideremos un modelo de AF como el planteado en la ecuación (2.4.2).

$$X = \Lambda\xi + \delta.$$

Por otra parte, la matriz $\sum$ que contiene las varianzas y covarianzas de las variables observables puede descomponerse tal y como se mostró en la ecuación (2.4.3),

$$\sum = \Lambda\Phi\Lambda^t + \Theta.$$

Si no se impone ningún tipo de restricción a los parámetros Λ, Φ y Θ, y si existe un conjunto de parámetros que cumplen la ecuación (2.4.3), entonces habrá un infinito número de ellos.
Veámoslo:
Sea M cualquier matriz de orden $s \times s$ no singular, por lo tanto, invertible. Si definimos,

$$\ddot{\Lambda} = \Lambda M^{-1}; \qquad \ddot{\xi} = M\xi$$

entonces,

$$\begin{aligned}
\ddot{\Lambda}\ddot{\xi} + \delta &= \left(\Lambda M^{-1}\right)\left(M\xi\right) + \delta \\
&= \Lambda\left(\left(M^{-1}\right)M\right)\xi + \delta \\
&= \Lambda\xi + \delta
\end{aligned}$$

de tal forma que si X cumple la ecuación (2.4.2), también se cumple que:

$$X = \ddot{\Lambda}\ddot{\xi} + \delta.$$

La matriz de covarianzas de $\ddot{\xi}$ vendrá dada por:

$$\begin{aligned}
\ddot{\Phi} &= E\left[\ddot{\xi}\ddot{\xi}^t\right] \\
&= E\left[\left(M\xi\right)\left(M\xi\right)^t\right] \\
&= ME\left[\xi\xi^t\right]M^t \\
&= M\Phi M^t.
\end{aligned}$$

Si operamos en la ecuación (2.4.3) se obtiene que:

$$\begin{aligned}
\ddot{\Lambda}\ddot{\Phi}\ddot{\Theta} &= \left(\Lambda M^{-1}\right)\left(M\Phi M^t\right)\left(\left(M^t\right)^{-1}\Lambda^t\right) + \Theta \\
&= \Lambda\left(MM^{-1}\right)\Phi\left(M^t\left(M^T\right)^{-1}\right)\Lambda^{-1} + \Theta \\
&= \Lambda\Phi\Lambda^t + \Theta,
\end{aligned}$$

de modo que, si $\sum$ cumple la ecuación (2.4.3), también se cumple que:

$$\sum = \ddot{\Lambda}\ddot{\Phi}\ddot{\Lambda}^t + \Theta.$$

Dado que las matrices marcadas con "¨"solo serían iguales en el caso en que $M = I$, la matriz identidad, existen infinitas matrices M invertibles que dan lugar a infinitas soluciones del modelo. En consecuencia, este modelo se definiría como no identificado.

¿Qué tipo de restricciones pueden imponerse a los parámetros? Por ejemplo, si una carga factorial de la matriz Λ se fija a cero, $\lambda_{ij} = 0$, estaremos indicando que el factor ξ_j no afecta causalmente a la variable observada X_i. Si fijamos a cero un elemento de la matriz Φ, $\phi_{ij} = 0$ estaremos señalando que los factores ξ_i y ξ_j no están correlacionados. Si todos los elementos fuera de la diagonal se fijan a cero, los factores serán ortogonales (como ocurre en el AFE). Restricciones similares se pueden imponer a los elementos de la matriz Θ.

A pesar de que la forma más efectiva de comprobar si un modelo está identificado es demostrando algebraicamente que es posible igualar cada parámetro estructural a una combinación de los elementos de $\sum$, en la práctica esta tarea es tediosa y compleja; por lo que en la literatura se han propuesto una serie de reglas o condiciones necesarias que suelen demostrarse como lo suficientemente exigentes para garantizar la identificación del modelo; estas nos permiten determinar más fácilmente el estatus de identificación de un modelo. En este sentido, el investigador debería centrarse en las siguientes condiciones:

1. Comparar la información disponible (varianzas y covarianzas muéstrales) con el número de parámetros que han de estimarse. El número de parámetros a estimar ha de ser menor o igual que el número de varianzas-covarianzas muéstrales.

2. Establecer una escala para los factores comunes. Esto se consigue fijando la saturación de una de las variables observadas por factor a 1 o la varianza de cada factor a 1.

3. Analizar la relación entre las variables observables y los factores comunes:
- Cuando solo hay un factor, el modelo puede estar identificado si hay al menos tres indicadores con cargas no nulas sobre él.
- Que habiendo al menos tres indicadores por factor, los errores asociados con los indicadores no estén correlacionados entre sí, cada indicador carga solo sobre un factor y los factores pueden covariar entre ellos.
- En el caso de disponer únicamente de dos indicadores sería necesario que exista correlación entre los factores, los errores asociados con cada indicador no están correlacionados y cada indicador carga solo sobre un factor.

4. Fijar arbitrariamente el coeficiente de regresión entre las variables observadas y los términos de error al valor 1.

Cualquier modelo que cumpla las condiciones anteriores estará identificado (o sobre identificado) y se podrá proceder a la estimación de sus parámetros.

La información disponible son siempre las $\frac{p(p+1)}{2}$ varianzas-covarianzas muéstrales. Como el número de parámetros a estimar es $p \times m + \frac{m(m+1)}{2} + \frac{p(p+1)}{2}$, el modelo estará sin identificar si no se imponen, al menos $p \times m + \frac{m(m+1)}{2}$ restricciones. Sólo si hay más varianzas y covarianzas muéstrales que parámetros, el modelo estará sobre identificado, [1] y [19].

En la práctica para posibilitar la identificación y, posteriormente, la estimación de los parámetros del modelo, es preciso imponer restricciones fijando alguno de esos parámetros a una constante. La principal razón de esta fijación a priori, radica en la necesidad de otorgarle una escala de medición a las variables latentes, dado que, al ser constructos que no se miden de forma directa, carecen de métrica. Para ello, se iguala a 1 la saturación (λ_{ij}) de un indicador por factor, de tal forma que el indicador elegido actuará como variable de referencia de ese constructo. Otra opción es estandarizar las variables latentes fijando de antemano su varianza a 1, solución que resultará especialmente útil cuando la métrica de los indicadores de una misma variable latente sea diferente, [19].

2.4.3.3. Estimación de parámetros

El proceso de estimación tiene como objetivo encontrar los valores, a partir de los datos muéstrales, de Λ, Φ y Θ, tal que cumplan las restricciones impuestas en el proceso de identicación y que generen una matriz de covarianzas estimada $\widehat{\sum}$ que sea tan próxima como sea posible a la matriz de covarianzas muestral S. Así, la matriz residual R obtenida de la diferencia entre ambas matrices $\left(S - \widehat{\sum}\right)$ deber a ser próxima a cero, [19].

A partir de lo descrito, el proceso de estimación del AFC puede sintetizarse en los dos pasos siguientes:

1. Dada la matriz de covarianzas muestral S, se estima el modelo hipotetizado: lo que supone encontrar valores para las matrices, Λ, Φ y Θ, que satisfagan la ecuación(2.4.3), pero habrá que rechazar todas aquellas soluciones que no cumplan las restricciones que se han impuesto en la identificación del modelo. Llamemos genéricamente $\widehat{\Lambda}$, $\widehat{\Phi}$ y $\widehat{\Theta}$, a las matrices que si cumplan las restricciones impuestas en el proceso de identificación, éstas generan una

matriz estimada $\widehat{\sum}$.

2. Se determina el ajuste del modelo hipotetizado. Esto es, se determina en qué medida $\widehat{\sum}$ esta próxima a S. Para determinar el grado de proximidad entre ambas matrices es preciso definir una función de ajuste entre ellas: $F\left(S-\widehat{\sum}\right)$, [1] y [19].

Así, el objetivo final es obtener, del conjunto de valores de los parámetros, aquellos que generen una matriz estimada $\widehat{\sum}$ que minimice su función de ajuste con S: $F=\left(S-\widehat{\sum}\right)\approx 0$. El proceso estima las varianzas y covarianzas en cada iteración (que es considerado como un mínimo local) y en el mínimo final de la función de minimización, se calculará el ajuste y todos los estimadores. Esto significa que la matriz de covarianzas estimada (también llamada reproducida) y la matriz de covarianzas observable son próximas y por ello se ha llegado al mínimo. Si la matriz residual es próxima a cero el ajuste es bueno, [19].

Si la función de minimización o ajuste llega a un mínimo final, ello significa que la función converge hacia una solución y que se ha llegado a una cierta correspondencia entre la matriz reproducida y la observada. Si el valor de la función de ajuste es igual a cero, esto supone que después de un cierto número de iteraciones se ha llegado a la matriz observada y que el ajuste es perfecto.
Existen varias funciones de ajuste que difieren según el método de estimación empleado. Los métodos de estimación de parámetros habitualmente utilizados son: mínimos cuadrados no ponderados, mínimos cuadrados generalizados y máxima verosimilitud, [1], [19] y [20].

La estimación por mínimos cuadrados no ponderados (ULS, por sus siglas en inglés) toma como estimadores a los valores que minimizan la siguiente función de ajuste:

$$F_{ULS}\left(S,\widehat{\sum}\right)=\frac{1}{2}tr\left[\left(S-\widehat{\sum}\right)^2\right] \tag{2.4.4}$$

donde:

tr: indica la traza de la matriz.
S: es la matriz de covarianza odservada.
$\widehat{\sum}$: es la matriz de covarianza reproducida por el modelo.

Este método tiene dos limitaciones que hacen que no sea muy utilizado: (1) no existen contrastes estadísticos asociados a este tipo de estimación y, (2) los estimadores dependen de la escala de medida de las variables observables. Sin embargo, una ventaja de este método es que no es necesario asumir ningún tipo de distribución teórica de las variables observadas, frente al supuesto de normalidad multivariada que asumen otros métodos de estimación.

La estimación por mínimos cuadrados generalizados (GLS, por sus siglas en inglés) se basa en ponderar la matriz cuya traza se calcula en la ecuación (2.4.4) mediante la inversa de la matriz de covarianzas muestral, esto es: [20]

$$F_{GLS}\left(S, \widehat{\sum}\right) = \tfrac{1}{2} tr\left[\left(S - \widehat{\sum}\right) S^{-1}\right]^2.$$

La estimación por máxima verosimilitud (ML, por sus siglas en inglés) implica minimizar la función de ajuste:

$$F_{ML}\left(S, \widehat{\sum}\right) = tr\left(S\widehat{\sum}^{-1}\right) + \left[\ln\left|\widehat{\sum}\right| - \ln|S|\right] - p, \qquad (2.4.5)$$

donde:
p: es el número de variables observadas.

Obsérvese que cuanto más se aproximen las matrices, S y $\widehat{\sum}$, más se aproximará $S\widehat{\sum}^{-1}$ a la matriz identidad $p \times p$, como la traza de esa matriz identidad es p, el primer término de la ecuación (2.4.5) se aproximará p cuando las matrices estén próximas, compensándose con el término p de la expresión. Por otra parte, la diferencia de los logaritmos de la diferencia de los determinantes de S y $\widehat{\sum}$ tenderán a cero, dado que, cuando las matrices estén próximas, también lo estarán sus determinantes. Por lo tanto, cuando las matrices sean iguales la función de ajuste sera cero, [20].

El método de estimación más común en los modelos de estructuras de covarianzas es el de ML, que proporciona estimaciones coherentes, eficientes, invariante al tipo de escala y no sesgadas cuando se cumple el supuesto de normalidad multivariada, [19].

Es importante que los datos sometidos al análisis sean los originales ya que en el proceso iterativo se estiman las varianzas y las covarianzas de la matriz reproducida y no las correlaciones, [20].

En la práctica un modelo no está identificado cuando la función de minimización no converge y es incapaz de llegar a un mínimo final y encontrar un estimador para cada parámetro. Cuando el proceso iterativo es exitoso, se podrá proceder a la evaluación del ajuste del modelo y, en el caso de que sea aceptable, a la interpretación de los parámetros finalmente obtenidos, [19].

2.4.3.4. Evaluación del ajuste del modelo

Antes de pasar a interpretar los resultados del AFC que se ha efectuado, es necesario determinar hasta qué punto el modelo asumido se ajusta a los datos muéstrales.

En cuanto a la evaluación de la calidad del modelo, el escalar obtenido como resultado de la función de ajuste empleada, junto con la matriz residual resultante de la diferencia entre matrices observada y predicha por el modelo, serán el punto de partida para la obtención de los índices de bondad de ajuste, índices que informaran de hasta qué punto la estructura definida a través de los parámetros del modelo reproduce la matriz de covarianzas de los datos muéstrales, [19].

En este sentido el modelado mediante estructuras de covarianzas no se sustenta en un único estadístico que describa la adecuación de las predicciones realizadas por el modelo. Es por ello que la evaluación de la bondad de ajuste de un modelo es más un proceso relativo que un criterio absoluto, por lo que se recomienda la evaluación complementaria de tres tipologías de índices de ajuste global:

- Índices de ajuste absoluto: determinan el grado en el que el modelo predice, a partir de los parámetros estimados, la matriz de covarianzas observada. Entre estos índices destacan, el índice χ^2, el índice de bondad de ajuste (GFI: Goodness of Fit Índex), el residuo estandarizado cuadrático medio (SRMR: Standardized Root Mean Square Residual) y el error cuadrático medio de aproximación (RM-SEA: Root Mean Square Residual Error of Aproximation).

- Índices de ajuste incremental: comparan el ajuste global del modelo propuesto con un modelo de referencia, habitualmente un modelo nulo en el que no se especifica ninguna relación entre las variables. El índice del ajuste normado (NFI: Normed Fit Index), el índice de bondad de ajuste comparativo (CFI: Comparative Fit Index), el índice de bondad de ajuste (GFI: Goodness of Fit Index) y el índice de bondad de adecuación ajustado (AGFI: Adjusted Goodness of Fit Index) son algunos ejemplos. Por lo general estos índices son fáciles de interpretar ya que sus valores oscilan entre 0 (ajuste ineficaz del modelo a los datos) y 1 (ajuste perfecto), considerándose habitualmente 0.90 como un inicio de ajuste apropiado.

- Índices de parsimonia: ponen en relación el ajuste alcanzado con el número de parámetros libres del modelo. Entre ellos se destacan el índice de calidad de ajuste de parsimonia (PGFI: Parsimonious Goodness of Fit Index) y el índice de ajuste normado de parsimonia (PNFI: Parsimonious Normed Fit Index). La interpretación de estos índices no se realiza en términos absolutos, sino comparando diferentes modelos con el fin de determinar cuál de ellos goza de una mayor parsimonia. Así, cuanto mayor es el valor del índice mayor es la parsimonia del modelo. La interpretación inversa la recibe otro índice de parsimonia, el criterio de información de Akaike (AIC: Akaike Information Criterion), que informa de una mayor parsimonia a medida que decrece su valor, [19].

En la actualidad, los softwares estadísticos como LISREL (Scientific Software International), AMOS (SPSS) y EQS 6 Structural Equations (Multivariante Software) proporcionan una gran variedad de índices de ajuste, incluso cuando estos ya no se consideran apropiados en la literatura científica (por ejemplo, NFI del EQS; GFI y AGFI en LISREL, etc). Esta abundancia de indicadores genera en ocasiones confusión al investigador, sobre todo cuando alguno de estos índices tienen tendencia a sobrevalorar el ajuste de los modelos, pudiendo llevar a la falsa conclusión de que el modelo es adecuado cuando no lo es, [17].

Estadístico X^2 para el contraste global del modelo.

El índice de ajuste por excelencia en los modelos AFC es χ^2. El punto de partida sería comparar las matrices de covarianzas observada (S) y la de covarianzas reproducida $\left(\widehat{\sum}\right)$, en el caso de que sean iguales, no habrá diferencia entre las dos y no rechazaríamos la hipótesis $H_0 : S = \widehat{\sum}$. Únicamente en este caso el modelo estaría perfectamente identificado y arrojaría un estadístico χ^2 de cero con cero grados de libertad. Por lo tanto, podemos establecer las siguientes hipótesis:

$$H_0 : S = \widehat{\sum} \qquad vs \qquad H_a : S \neq \widehat{\sum}.$$

Para el contraste de estas hipótesis en Blenter y Bonett se propone el estadístico:

$$\chi^2 = N * F^0_{ML} \tag{2.4.6}$$

donde:
N: es el número de datos.
F^0_{ML}: es el valor que toma la función de ajuste al realizar la estimación por máxima verosimilitud.

El estadístico se distribuye, bajo la hipótesis nula como una χ^2 con los siguientes grados de libertad:
$g.l = \frac{p(p+1)}{2} - k$.

Siendo p el número de variables observadas y k el número de parámetros que se han de estimar, asociados a la hipótesis nula, que varían en función de cada modelo, [1].

No rechazamos que $S = \widehat{\sum}$ en el caso de que χ^2 sea suficientemente pequeño (es decir, el valor de p sea superior a, por ejemplo, $\alpha = 0{,}05$).

El estadístico se utiliza, para contrastar la validez del modelo teórico propuesto por el investigador. Sin embargo, este índice rara vez es utilizado como prueba única o concluyente de bondad del ajuste del modelo. En la práctica, interesa más cuantificar el grado de ajuste (o desajuste) del modelo que simplemente rechazar o no la hipótesis nula, [20].
Por lo tanto, se ha optado por complementar el índice de ajuste basado en el estadístico χ^2 con otro conjunto de indicadores. De hecho, en la práctica, si un modelo presenta un buen ajuste a través del CFI y del RMSEA conjuntamente, es muy poco probable que el modelo no sea adecuado a los datos. Estos índices de ajuste son, por tanto, una buena guía en la búsqueda del modelo que mejor se ajusta a los datos, [4] y [17].

Índices comparativos de ajuste.

El software libre R 3.2.4 (paquete R-commander) ofrecen, además del estadístico χ^2 como índice de ajuste global, un segundo estadístico que denominaremos modelo χ^2 independiente (también llamado modelo de referencia). Este estadístico se distribuye también como una χ^2 bajo la hipótesis nula de que existe una completa independencia entre las variables observadas (matriz de correlaciones es la identidad) y tendría tantos grados de libertad como el número de datos menos el número de parámetros independientes (varianzas) que se han de estimar, [1]. Los índices que se proponen son comparativos en el sentido de que comparan el valor del modelo teórico que se evalúa, con el modelo independiente.

Índice CFI

Este índice fue desarrollado por Blenter 1992 citado en [17] a partir del índice previo NFI, que corrige para evitar que tome valores más allá del rango 0-1. El CFI compara el χ^2 de dos modelos: un modelo independiente que mantiene que no existe relación entre las variables del modelo, y el modelo teórico propuesto por el investigador.
Esta comparación se corrige por los grados de libertad de uno y otro modelo del siguiente modo:

$$CFI = \left| \frac{\left(\chi^2_{indep} - gl_{indep}\right) - \left(\chi^2_{teorico} - gl_{teorico}\right)}{\left(\chi^2_{indep} - gl_{indep}\right)} \right|.$$

Conforme el X^2 del modelo teórico propuesto disminuye, el numerador y denominador se igualan, por lo que la situación ideal es que ambos sean equivalentes (CFI =1); esto es que el valor del estadístico χ^2 del modelo teórico sea cero. En general se considera que el CFI debe estar en torno a 0.95 para considerar que el modelo se ajusta adecuadamente a los datos (valor mínimo de buen ajuste es 0.90). Este valor, sin embargo, es relativo ya que, en modelos de gran complejidad el χ^2 siempre se alejara del cero, lo que hace disminuir el CFI. Por lo tanto, la interpretación del índice CFI se debe valorar conjuntamente con otros índices, teniendo en cuenta el tipo de modelo que se está analizando, [1] y [17].

Índice RMSEA

El índice de bondad de ajuste más robusto propuesto a la fecha es el Error Cuadrático Medio de Aproximación (RMSEA). Este índice ha sido desarrollado como una medida absoluta de la diferencia de la estructura de relaciones entre el modelo propuesto y los valores de covarianza en población medida Steiger, 1990. Su cálculo es como sigue:

$$RMSEA = \sqrt{\frac{\widehat{\delta_T}}{(gl_{teorico})(N-1)}}$$

en este caso el término $\widehat{\delta_T} = max\left(\chi^2_{teorico} - gl_{teorico}, 0\right)$.

La importancia de este índice radica en que refleja una diferencia absoluta entre el modelo propuesto y los datos observados, tomando en cuenta el número de estimaciones y el tamaño de la muestra implicada por el modelo bajo prueba. Es muy importante notar que este índice, debido a su origen y propiedades estadísticas, compara el modelo con la estructura de relaciones entre las variables en la población.

El índice RMSEA y su intervalo de confianza, cuando toman valores menores a 0.05 es indicio de que el ajuste entre el modelo y los datos es muy bueno, pero si sus valores resultan entre 0.05 y 0.08 el ajuste del modelo a los datos es razonable; mientras que si sus valores están entre 0.08 y .10 indica un ajuste pobre o mediocre. Ahora bien, el modelo deberá de rechazarse si los valores del índice RMSEA resultan mayores a 0.10. No obstante lo anterior, por razones prácticas debe incorporarse evidencia que refuerce los resultados obtenidos. Para ello se recurre a interpretar otros índices de bondad de ajuste, [4], [15], [17] y [19].

Índice RMR y SRMR.

El último grupo de índices que analizaremos son los basados en los residuos que son un promedio de las diferencias entre las covarianzas muéstrales y las estimadas que se derivan del modelo. Esto es:

$$RMR = \sqrt{\frac{\sum_{i=1}^{p}\sum_{j=1}^{i}(s_{ij}-\widehat{\sigma_{ij}})^2}{\frac{p(p+1)}{2}}},$$

donde p es el número de variables observadas.
Como los residuos sin estandarizar están afectados por la escala en que se mide la variable, se suelen utilizar los residuos estandarizados construyéndo-

se el llamado SRMR que está acotado entre 0 y 1. El ajuste se considera aceptable si el SRMR toma valores inferiores a 0.08 (preferentemente inferior a 0.05).

Se considera por lo tanto aconsejable presentar estos índices de ajuste junto con el χ^2 del modelo teórico propuesto, sus grados de libertad y la probabilidad asociada, [1] y [15].

2.4.4. Interpretación del modelo

Hasta el momento nos hemos centrado en evaluar la bondad de ajuste global del modelo. Ahora el siguiente paso, a la hora de determinar su idoneidad en la descripción de los constructos, es evaluar el ajuste de sus componentes, que en el caso del AFC serán los parámetros del modelo confirmatorio especificado.

En el caso de que tanto el ajuste global como individual sea aceptable, el modelo será consistente con los datos y se podrá proceder a su interpretación y presentación matricial y/o ecuacional. Si el ajuste no es bueno, se procederá a reespecificar el modelo, para lo que será necesario realizar un análisis minucioso de los resultados obtenidos. En concreto, con el fin de mejorar la bondad de ajuste del modelo habrá que examinar, al menos, los siguientes aspectos:

• Prueba de significación de parámetros: El estadístico t informa de la significación estadística de cada parámetro a partir de la razón entre el valor del estimador y su error típico. Para un nivel de significancia $\alpha \leq 0{,}05$, un valor t comprendido entre -1.96 y 1.96, indicará que el parámetro en cuestión no es estadísticamente significativo, por lo que será necesario eliminarlo o dejarlo a un valor determinado.

• Matriz de residuos normalizados: El análisis de los residuos normalizados permitirá identificar errores de predicción entre las matrices $\widehat{\sum}$ y S. Todo residuo cuyo valor este fuera de los límites entre -2.58 y 2.58, para un nivel de significancia $\alpha \leq 0{,}01$, indicará que no se ha podido reproducir convenientemente, a partir de los parámetros del modelo, la covarianza entre el par de variables implicado.

• Índices de Modificación (IM): son calculados por todos los parámetros fijos del modelo, informando del cambio esperado en el valor de χ^2 si se libera un determinado parámetro fijo y reestima de nuevo el modelo manteniendo estables el resto de parámetros. De esta forma, el IM será el valor resultante de la diferencia en el χ^2 entre el modelo que tiene el parámetro fijado y el que lo mantiene libre. Un $IM \geq 3{,}84$ indica que se produce una disminución estadísticamente significativa en el valor de χ^2. Se toma esta referencia por ser el valor teórico de χ^2 para 1 grado de libertad y un nivel de significancia de 0.05, [4] y [19].

Es importante tener presente que la reespecificacion del modelo se ha de llevar a cabo gradualmente y siempre apoyándose en una justificación teórica que sustente el cambio o corrección impuesta. Así, si hay más de un parámetro no significativo se ha de eliminar primero el que tenga una razón critica más baja y proceder a continuación a examinar la bondad de ajuste del nuevo modelo corregido. La razón es que la modificación de un solo parámetro incidirá en la estimacion de los demás, de tal forma que un parámetro que no era significativo en el modelo original puede llegar a serlo en el modelo reespecificado, [19].

Sin embargo, existen muchos problemas que pueden generarse como consecuencia de una reespecificacion poco meditada. Si el investigador cae en la tentación de ir incorporando o eliminando relaciones sin más, hasta lograr un ajuste razonable y no tiene en cuenta si estas modificaciones están o no sustentadas por el marco teórico de su investigación, puede provocarse que el modelo al que se llega no sea generalizable (Mccallumn, Roznowski y Necowitz, 1992), [1].
En este mismo sentido, Padhazur 1982 y Sorbom 1989 citado en [1], afirman que es científicamente incorrecto modificar un modelo simplemente para que mejore su ajuste, ya que el cambio debe ser teóricamente interpretable y el investigador debe ser capaz de justificar cual es el motivo para añadir una relación causal determinada.

Capítulo 3

Validación y confiabilidad del cuestionario SF-36

En este capítulo se presenta la aplicación de los criterios estudiados en esta tesis a un instrumento del área de Salud que se está usando para abordar problemas reales de gran relevancia en la salud.

3.1. Planteamiento del problema

El Lupus Eritematoso Sistémico (LES) es una enfermedad denominada autoinmune, es decir, provoca una alteración en el sistema inmunológico que lo lleva a desconocer el cuerpo del enfermo. Este padecimiento no discrimina: ataca el corazón, riñones, las articulaciones, el cerebro y cualquier otro órgano y tejido por igual. El nivel de daño que provoca en el cuerpo y los órganos que afecta varía en cada paciente. En México, lo padecen 1.5 millones de personas, pero podrían padecerlo un número mayor, ya que el tiempo estimado para un diagnóstico correcto es de dos a cinco años, según la Federación Española de Enfermos de LUPUS. En este lapso la aplicación de medicamentos erróneos puede alterar el curso de la verdadera enfermedad y provocar complicaciones; no hay cifras que señalen el número de casos que son mal diagnosticados a pesar de que en México se reportan de dos a ocho casos por cada 100 mil habitantes al año.
No existe una prueba específica que permita determinar si una persona padece la enfermedad. El diagnóstico se basa en la presencia de al menos cuatro de los once criterios establecidos por la Asociación Americana de Reumatología (ACR); los que presentan un mayor porcentaje son: el aumento de anticuerpos anticelulares, sensibilidad a la luz, dolor de cabeza, dolor en articulaciones y manchas rojizas en mejillas y nariz, conocidas como “alas de mariposa”, [27].

El Lupus Eritematoso Sistémico (LES) es una enfermedad de causa desconocida, aunque la herencia, el entorno y los cambios hormonales juegan un papel importate que afecta a todas las edades, pero con mayor frecuencia a adultos entre los 18 y 50 años con predominio del sexo femenino, en una proporción de un hombre por cada 10 a 12 mujeres. La prevalencia del LES varía en los distintos grupos de población, oscilando entre 300 y 400 pacientes por cada 100.000 habitantes. Es más común en ciertos grupos étnicos, especialmente los afroamericanos, [23].
Aparte de los factores fisiológicos antes mencionados, el LES puede afectar de diversas formas la Calidad de Vida Relacionada con la Salud (CVRS) de quienes padecen esta enfermedad, más concretamente a nivel de la actividad física, sexual, mental y social. Desde el punto de vista subjetivo, la calidad de vida relacionada con la salud es la valoración que realiza una persona, de acuerdo con sus propios criterios del estado físico, emocional y social en el que se encuentra en un momento dado, y refleja el grado de satisfacción con una situación personal a nivel: fisiológico (sintomatología general, discapacidad funcional, situación analítica, sueño, respuesta sexual), emocional (sentimientos de tristeza, miedo, inseguridad, frustración) y social (situación laboral o escolar, interacciones sociales en general, relaciones familiares, amistades, nivel económico, participación en la comunidad, actividades de ocio, entre otras), [31].
La calidad de vida relacionada con la salud en pacientes con LES se ha venido evaluando a través de instrumentos tanto genéricos como específicos. Según [34] en una revisión de literatura sobre instrumentos de CVRS utilizados en estudios con pacientes con LES se ha encontrado que las áreas de la CVRS más afectadas en estos pacientes han sido la percepción de salud, la fatiga, el dolor corporal, la actividad funcional a nivel laboral, la autonomía, las relaciones sociales, familiares y la desesperanza aprendida (respecto a la imprevisibilidad de la enfermedad del LES).

Dado que para las personas con LUPUS en el estado de Puebla no se ha validado ningún instrumento de medición, se tomó el cuestionario SF-36 y se realizó una adaptación cultural. El objetivo de este estudio, realizado por investigadores del área de la Salud de la BUAP, es validar dicho instrumento.

3.1.1. Instrumento de medición

El cuestionario de salud SF-36 fue desarrollado a principios de los noventa, en Estados Unidos, para su uso en el Estudio de los Resultados Médicos (Medical Outcomes Study, MOS). Es una escala genérica que proporciona un perfil del estado de salud y es aplicable tanto a los pacientes como a la población general. Ha resultado útil para evaluar la calidad de vida relacionada con la salud (CVRS) en la población general y en subgrupos específicos, comparar la carga de muy diversas enfermedades, detectar los beneficios en la salud producidos por un amplio rango de tratamientos diferentes y valorar el estado de salud de pacientes individuales. Sus buenas propiedades psicométricas, que han sido evaluadas en más de 400 artículos, y la multitud de estudios ya realizados que permiten la comparación de resultados, lo convierten en uno de los instrumentos con mayor potencial en el campo de la CVRS.

El SF-36 evalúa aspectos de la calidad de vida en poblaciones adultas (mayores de 16 años). El producto de su aplicación es la construcción de ocho conceptos o escalas de salud resultado del promedio de la suma de las preguntas contenidas en el cuestionario. Estos conceptos son: a) función física (FF), b) rol físico (RF), c) dolor corporal (DC), d) salud general (SG), e) vitalidad (VT), f) función social (FS), g) rol emocional (RE) y h) salud mental (SM).
Además de los ocho conceptos de salud, el SF-36 incluye el concepto general de cambios en la percepción del estado de salud actual y en la del año anterior. La respuesta a esta pregunta describe la transición de la percepción respecto al mejoramiento o empeoramiento del estado de salud. En el cuadro I se presenta una descripción de las escalas de salud y sus respectivas interpretaciones de acuerdo con resultados bajos o altos por cada escala, [13].

Tabla1. Contenido de las escalas del SF-36

		Significado de las puntuaciones de 0 a 100	
Dimensión	N.º de ítems	«Peor» puntuación (0)	«Mejor» puntuación (100)
Función física	10	Muy limitado para llevar a cabo todas las actividades físicas, incluido bañarse o ducharse, debido a la salud	Lleva a cabo todo tipo de actividades físicas incluidas las más vigorosas sin ninguna limitación debido a la salud
Rol físico	4	Problemas con el trabajo u otras actividades diarias debido a la salud física	Ningún problema con el trabajo u otras actividades diarias debido a la salud física
Dolor corporal	2	Dolor muy intenso y extremadamente limitante	Ningún dolor ni limitaciones debidas a él
Salud general	5	Evalúa como mala la propia salud y cree posible que empeore	Evalúa la propia salud como excelente
Vitalidad	4	Se siente cansado y exhausto todo el tiempo	Se siente muy dinámico y lleno de energía todo el tiempo
Función social	2	Interferencia extrema y muy frecuente con las actividades sociales normales, debido a problemas físicos o emocionales	Lleva a cabo actividades sociales normales sin ninguna interferencia debido a problemas físicos o emocionales
Rol emocional	3	Problemas con el trabajo y otras actividades diarias debido a problemas emocionales	Ningún problema con el trabajo y otras actividades diarias debido a problemas emocionales
Salud mental	5	Sentimiento de angustia y depresión durante todo el tiempo	Sentimiento de felicidad, tranquilidad y calma durante todo el tiempo
Ítem de Transición de salud	1	Cree que su salud es mucho peor ahora que hace 1 año	Cree que su salud general es mucho mejor ahora que hace 1 año

Cuadro 3.1: Escalas del instrumento SF-36.

Este cuestionario fue aplicado a pacientes de un hospital del estado de Puebla.

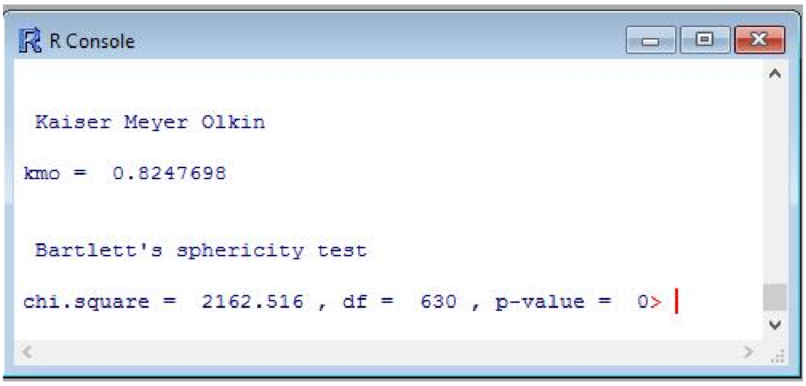
El muestreo fue por conveniencia, se encuestó a 90 mujeres mayores a 18 años, a quienes se les aplicó, de manera individual el cuestionario de Salud SF-36 compuesto por 36 preguntas (ítems) que valoran los estados tanto positivos como negativos de la Calidad de Vida Relacionada con la Salud (CVRS), en mujeres con el Lupus Eritematoso Sistémico (LES) del estado de Puebla (México).

3.2. Análisis de la validez del cuestionario SF-36

A continuación se muestran los resultados obtenidos en el análisis de la validez del cuestionario SF-36 utilizando el paquete R-Commander. En el anexo A se muestra el procedimiento.

3.2.1. Análisis preliminares

Antes de aplicar el Análisis Factorial se debe comprobar si la correlación entre las variables analizadas es lo suficientemente grande como para justificar la factorización de la matriz de coeficientes de correlación. Esta comprobación se realizó mediante el test de Bartlett (1950) y el índice de KMO cuyos resultados, se observan en el Cuadro 3.2.



Cuadro 3.2: KMO y prueba de esferecidad de Bartlett.

El índice de KMO es de 0.82 lo que indica que se puede realizar el AF, por otro lado la prueba de Bartlett es significativa, ya que $P <$ alfa, y por consecuente se rechaza la hipótesis nula de que la matriz de coeficientes de correlación no es significativamente distinta de la matriz identidad. Por lo anterior concluimos que la muestra es adecuada para la realización del AFE.

3.2.2. Análisis Factorial Exploratorio

Esta sección consiste en la obtención de los valores y vectores propios de la matriz de coeficientes de correlación que se obtienen a partir de la matriz de datos.

Para determinar el número de factores a extraer se realizó un análisis multidimensional con el método de Componentes Principales (CP), empleando el software R-Commander. Este método calcula tantas componentes principales como variables originales y, así pues, se reproduce la totalidad de la varianza.

Por la regla de Kaiser-Guttman extraemos tantos factores como autovalores mayores que uno se encuentren: en nuestro caso extraemos nueve componentes. Ver cuadro 3.3

```
Component variances:
     Comp.1      Comp.2      Comp.3      Comp.4      Comp.5      Comp.6      Comp.7
12.37265463  3.09476334  2.14971046  1.76988430  1.75076667  1.40855455  1.34906792
     Comp.8      Comp.9     Comp.10     Comp.11     Comp.12     Comp.13     Comp.14
 1.26166628  1.25191077  0.94342981  0.82229446  0.77111675  0.68624047  0.65360251
    Comp.15     Comp.16     Comp.17     Comp.18     Comp.19     Comp.20     Comp.21
 0.55940031  0.50754457  0.47778893  0.43961923  0.37753153  0.37313790  0.33995998
    Comp.22     Comp.23     Comp.24     Comp.25     Comp.26     Comp.27     Comp.28
 0.31320582  0.30650593  0.26153700  0.24395300  0.24172116  0.21116733  0.18237964
    Comp.29     Comp.30     Comp.31     Comp.32     Comp.33     Comp.34     Comp.35
 0.15531753  0.14973904  0.13033974  0.11578721  0.10770874  0.09223174  0.07251181
    Comp.36
 0.05524893
```

Cuadro 3.3: Componentes principales.

Las siguientes salidas en R-Commander (ver cuadro 3.4) son:

- Standard deviation: es la varianza asociada a cada factor (el cuadrado de las desviaciones estándar) viene expresada por su valor propio o raíz característica de la matriz de coeficientes de correlación o de la matriz de covarianzas.

– Proportion of Variance: es la proporción de la varianza que explica cada componente principal. Su suma es igual a 1.

– Cumulative proportion: es la proporción acumulada, se calcula sumándolas progresivamente.

Por ejemplo: Observamos que las dos primeras componentes agrupan un 42.9 % de la variación, o lo que es lo mismo, hay un 57.1 % de variación que no se explica.

```
Importance of components:
                          Comp.1     Comp.2     Comp.3     Comp.4     Comp.5     Comp.6
Standard deviation     3.5174784 1.75919394 1.46618909 1.33036999 1.32316540 1.18682541
Proportion of Variance 0.3436849 0.08596565 0.05971418 0.04916345 0.04863241 0.03912652
Cumulative Proportion  0.3436849 0.42965050 0.48936468 0.53852813 0.58716054 0.62628705
                           Comp.7     Comp.8    Comp.9    Comp.10    Comp.11    Comp.12
Standard deviation     1.16149383 1.12323919 1.1188882 0.97130315 0.90680453 0.87813254
Proportion of Variance 0.03747411 0.03504629 0.0347753 0.02620638 0.02284151 0.02141991
Cumulative Proportion  0.66376116 0.69880745 0.7335827 0.75978913 0.78263064 0.80405055
                          Comp.13    Comp.14   Comp.15    Comp.16    Comp.17    Comp.18
Standard deviation     0.82839633 0.80845687 0.7479307 0.71242162 0.69122278 0.66303788
Proportion of Variance 0.01906224 0.01815563 0.0155389 0.01409846 0.01327191 0.01221165
Cumulative Proportion  0.82311279 0.84126841 0.8568073 0.87090577 0.88417769 0.89638933
                          Comp.19    Comp.20     Comp.21     Comp.22     Comp.23     Comp.24
Standard deviation     0.61443594 0.61085015 0.583060871 0.559647946 0.553629777 0.511406879
Proportion of Variance 0.01048699 0.01036494 0.009443333 0.008700162 0.008514054 0.007264917
Cumulative Proportion  0.90687632 0.91724126 0.926684594 0.935384755 0.943898809 0.951163726
                           Comp.25     Comp.26     Comp.27     Comp.28     Comp.29
Standard deviation     0.493915988 0.491651464 0.459529467 0.427059292 0.394103454
Proportion of Variance 0.006776472 0.006714477 0.005865759 0.005066101 0.004314376
Cumulative Proportion  0.957940198 0.964654675 0.970520434 0.975586535 0.979900911
                           Comp.30     Comp.31     Comp.32     Comp.33     Comp.34
Standard deviation     0.386961293 0.361025951 0.340275198 0.328190096 0.303696786
Proportion of Variance 0.004159418 0.003620548 0.003216311 0.002991909 0.002561993
Cumulative Proportion  0.984060329 0.987680877 0.990897188 0.993889098 0.996451090
                           Comp.35     Comp.36
Standard deviation     0.269280169 0.235050917
Proportion of Variance 0.002014217 0.001534693
Cumulative Proportion  0.998465307 1.000000000
```

Cuadro 3.4: Varianza total explicada.

El objetivo del AFE es obtener la estructura factorial más simple desde el punto de vista de su interpretación más esencial, siguiendo los criterios de parsimonia establecidos por Thurstone en 1947, [10] y [11].
Tenemos un modelo con nueve factores y con una varianza total explicada del 73.358 %.

En el gráfico de sedimentación, obtenido mediante el paquete R-Commander, se observa claramente que la primera componente podría ser extraída (ver Figura 3.1) ya que explica 34.36 % de la varianza total.

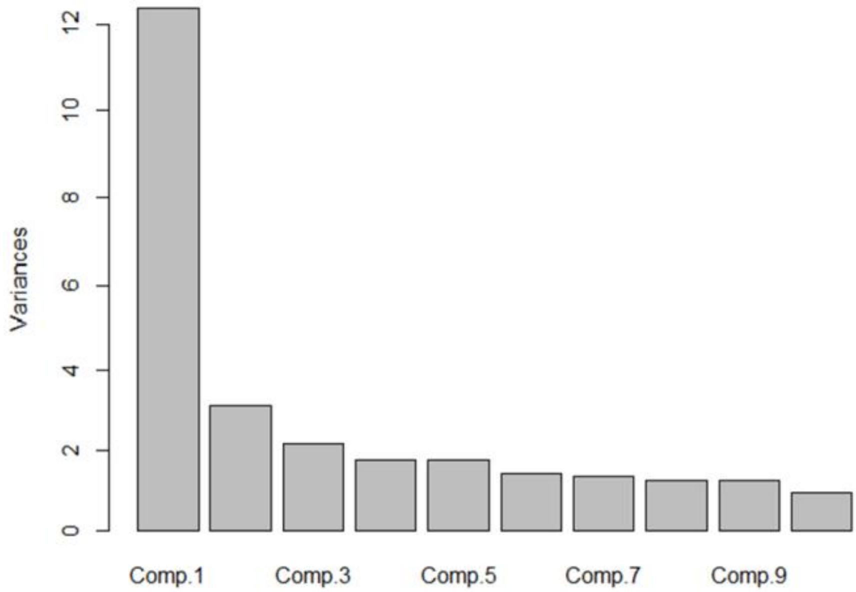


Figura 3.1: Gráfico de sedimentación.

En primer lugar consideramos el análisis monofactorial para analizar sí todos los indicadores saturan en un único factor; para esto utilizamos el software R-Commander. Los resultados se muestran en el Cuadro 3.5.

Las salidas generadas en el Cuadro 3.5 son:

Primero da la unicidad de cada ítem, es decir, el porcentaje de varianza de cada uno de los ítems que no ha podido ser explicado por el factor que se ha extraído.
Por ejemplo: el 87.1 % de la varianza del primer ítem no ha sido explicado por el factor extraído, el 80.5 % de la varianza del segundo ítem no ha sido explicado por el factor extraído, el 59.3 % de la varianza del noveno ítem no ha sido explicado por el factor extraído, etc.

Segundo, da las saturaciones de los ítems con el factor extraído, es decir que explica el 32.5 % de la varianza.

Y finalmente, un test cuya hipótesis nula es que un sólo factor es suficiente; observemos que el valor de χ^2 es muy alto y el p-valor muy bajo a 0.05 por lo que se rechaza la hipótesis de que un solo factor es suficiente.

```
Uniquenesses:
ítem.1. ítem.2 ítem.3 ítem.4 ítem.5 ítem.6 ítem.7 ítem.8 ítem.9 ítem.10 ítem.11
  0.871  0.805  0.588  0.647  0.568  0.405  0.523  0.552  0.593  0.550  0.635
ítem.12 ítem.13 ítem.14 ítem.15 ítem.16 ítem.17 ítem.18 ítem.19 ítem.20 ítem.21 ítem.22
  0.622  0.691  0.699  0.746  0.774  0.769  0.811  0.995  0.762  0.658  0.743
ítem.23 ítem.24 ítem.25 ítem.26 ítem.27 ítem.28 ítem.29 ítem.30 ítem.31 ítem.32 ítem.33
  0.619  0.687  0.528  0.659  0.513  0.597  0.534  0.715  0.511  0.612  0.725
ítem.34 ítem.35 ítem.36
  0.892  0.884  0.808

 Loadings:
         Factor1
 ítem.1.  0.359
 ítem.2   0.442
 ítem.3   0.642
 ítem.4   0.594
 ítem.5   0.657
 ítem.6   0.772
 ítem.7   0.690
 ítem.8   0.669
 ítem.9   0.638
 ítem.10  0.671
 ítem.11  0.604
 ítem.12  0.615
 ítem.13  0.556
 ítem.14  0.549
 ítem.15  0.504
 ítem.16  0.475
 ítem.17  0.481
 ítem.18  0.435
 ítem.19
 ítem.20  0.488
 ítem.21  0.584
 ítem.22  0.507
 ítem.23  0.617
 ítem.24  0.559
 ítem.25  0.687
 ítem.26  0.584
 ítem.27  0.698
 ítem.28  0.635
 ítem.29  0.683
 ítem.30  0.534
 ítem.31  0.700
 ítem.32  0.623
 ítem.33  0.524
 ítem.34  0.328
 ítem.35  0.340
 ítem.36  0.439

                Factor1
 SS loadings     11.709
 Proportion Var   0.325

 Test of the hypothesis that 1 factor is sufficient.
 The chi square statistic is 1252.37 on 594 degrees of freedom.
 The p-value is 3.68e-49
```

Cuadro 3.5: Resultados del modelo monofactorial.

A continuación realizaremos el análisis factorial exploratorio utilizando el software R-Commander, con los 9 factores extraídos por el método de CP. Los resultados se muestran en el Cuadro 3.6.

	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6	Factor7	Factor8	Factor9
ítem.3	0.62							-0.23	
ítem.4	0.87				0.24				
ítem.6	0.66		0.20						
ítem.7	0.75								
ítem.8	0.71			-0.23					
ítem.9	1.27								0.81
ítem.10	0.96								0.34
ítem.11	0.72								
ítem.24		0.91							
ítem.25		0.67							
ítem.28		0.84							
ítem.29		0.65			0.28				
ítem.31		0.54							
ítem.13			0.80						
ítem.14			0.85						
ítem.15			0.71						
ítem.16			0.57		0.23			0.24	
ítem.17				0.83					
ítem.18				1.00					
ítem.23					0.73	0.25			
ítem.27		0.30			0.61	0.38			
ítem.1.						0.75			
ítem.34						0.62			
ítem.36					0.24	0.59			
ítem.21							0.61		
ítem.22							1.05		
ítem.33								0.86	
ítem.35								0.71	
ítem.5	0.41								-0.59
ítem.2				0.21	0.30	0.44			
ítem.12	0.39				-0.22	0.26			
ítem.19		0.28							0.23
ítem.20				0.35	0.37	-0.21			
ítem.26		0.42			0.36				
ítem.30		0.36			0.29				
ítem.32		0.29			0.26				

	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6	Factor7	Factor8	Factor9
SS loadings	6.21	3.43	2.54	2.24	2.09	1.98	1.75	1.72	1.43
Proportion Var	0.17	0.10	0.07	0.06	0.06	0.05	0.05	0.05	0.04
Cumulative Var	0.17	0.27	0.34	0.40	0.46	0.51	0.56	0.61	0.65

```
Test of the hypothesis that 9 factors are sufficient.
The chi square statistic is 403.76 on 342 degrees of freedom.
The p-value is 0.012
```

Cuadro 3.6: Matriz de saturaciones, ítems agrupados por nueve factores.

Veamos que en el factor 9 solo se tiene un ítem por lo que tendríamos problemas de identificación y convergencia, ya que al realizar el Análisis Factorial Confirmatorio éste requiere que cada factor tenga al menos dos ítems, así que descartamos el hecho de que los indicadores saturan en nueve factores.

Estimamos el modelo de ocho factores utilizando el paquete R-Commander, usando el método de ML para extraer los factores y presentando la rotación varimax. En el Cuadro 3.7 se muestran los resultados del análisis.

Uniquenesses:										
Item.1.	Item.2.	Item.3.	Item.4.	Item.5.	Item.6.	Item.7.	Item.8.	Item.9.	Item.10.	Item.11.
0.434	0.526	0.432	0.349	0.394	0.243	0.262	0.352	0.38	0.406	0.39
Item.12.	Item.13.	Item.14.	Item.15.	Item.16.	Item.17.	Item.18.	Item.19.	Item.20.	Item.21.	Item.22.
0.493	0.188	0.163	0.305	0.407	0.241	0.005	0.909	0.526	0.398	0.005
Item.23.	Item.24.	Item.25.	Item.26.	Item.27.	Item.28.	Item.29.	Item.30.	Item.31.	Item.32.	Item.33.
0.371	0.388	0.315	0.52	0.269	0.368	0.33	0.551	0.401	0.518	0.272
Item.34.	Item.35.	Item.36.								
0.565	0.45	0.512								

Loadings:								
ITEM	FF	SM	RF	RE	SG	DC	FS	VT
Item.3	0.65						-0.22	
Item.4	0.87	-0.22						0.26
Item.5	0.65		0.23					-0.25
Item.6	0.68		0.21					
Item.7	0.88							
Item.8	0.73			-0.24				
Item.9	0.78							
Item.10	0.75							
Item.11	0.81							
Item.24		0.88					-0.21	
Item.25		0.69						
Item.26		0.51						0.27
Item.28		0.78						
Item.29		0.81						
Item.31		0.68						
Item.13			0.82					-0.22
Item.14			0.87					-0.22
Item.15			0.72					0.21
Item.16			0.59				0.24	0.23
Item.17				0.83				
Item.18				1.04				
Item.1.					0.77			
Item.34					0.64			
Item.36					0.62			
Item.21						0.61		
Item.22						1.06		
Item.33							0.79	
Item.35							0.71	
Item.23		0.25			0.26			0.55
Item.2				0.24	0.45			
Item.12	0.39				0.24			-0.25
Item.19	-0.26	0.23		-0.22				
Item.20				0.39	-0.21			0.32
Item.27		0.42			0.41			0.44
Item.32		0.35						0.24

	FF	DM	RF	RE	SG	DC	FS	VT
SS Loadings	5.57	4.04	2.68	2.38	2.08	1.77	1.57	1.45
Proportion Var.	0.15	0.11	0.07	0.07	0.06	0.05	0.04	0.04
Cumulative Var.	0.15	0.27	0.34	0.41	0.47	0.51	0.56	0.6

Test of the hypothesis that 8 factors are sufficient.

The chi square statistic is 468.66 on 370 degrees of freedom.

The p-value is 0.000374

Cuadro 3.7: Matriz de saturaciones, ítems agrupados por ocho factores.

En el Cuadro 3.7 se muestran las saturaciones de cada ítem en cada factor, como en realidad se trata de coeficientes de correlación se interpretan de la misma manera. Los ítems que pertenecen a cada factor son aquellos que tienen el peso mayor en un factor y mucho menores en los demás. Algunos autores indican que en ningún caso debe ser muy inferior a 0.40; sin embargo, para Gorsuch (1993, p. 208) una correlación ítem-factor de 0.35 es suficiente para asumir la relación ítem factor e interpretarlo con claridad, y Kline (1994) señala 0.30 como un valor orientador aceptable (citados en [25]).

A continuación se describe la interpretación de cada uno de los factores y los ítems:

Factor 1: Función Física (FF). Grado en que la salud limita las actividades físicas tales como el autocuidado, caminar, subir escaleras, inclinarse, coger o llevar pesos y los esfuerzos moderados e intensos. Asociado a las variables 3, 4, 5, 6, 7, 8, 9, 10, 11, 12.

Factor 2: Salud Mental (SM). Salud mental general, incluyendo depresión, ansiedad, control de la conducta o bienestar general. Asociado a las variables 19, 24, 25, 26, 28, 29, 30, 31, 32.

Factor 3: Rol Físico (RF). Grado en que la salud física interfiere en el trabajo y otras actividades diarias, incluyendo rendimiento menor que el deseado, limitación en el tipo de actividades realizadas o dificultad en la realización de actividades. Asociado a las variables 13, 14, 15, 16.

Factor 4: Rol Emocional (RE). Grado en que los problemas emocionales interfieren en el trabajo u otras actividades Diarias. Asociado a las variables 17, 18, 20.

Factor 5: Salud General (SG). Valoración personal de la salud, que incluye la salud actual, las perspectivas de salud en el futuro y la resistencia a enfermar. Asociado a las variables 1, 34, 36, 2.

Factor 6: Dolor Corporal (DC). Intensidad del dolor y su efecto en el trabajo habitual, tanto fuera de casa como en el Hogar. Asociado a las variables 21, 22.

Factor 7: Función Social (FS). Grado en que los problemas de salud física o emocional interfieren en la vida social habitual. Asociado a las variables 33, 35.

Factor 8: Vitalidad (VT). Sentimiento de energía y vitalidad, frente al sentimiento de cansancio y agotamiento. Asociado a las variables 23, 27.

3.2.3. Análisis Factorial Confirmatorio

Para confirmar los resultados del AFE se realizó el AFC sobre la matriz de correlación de la muestra total, para esto se utilizó el programa R-Commander.

Primero lo aplicamos al modelo monofactorial:
Los índices bondad de ajuste de este modelo se muestran en el Cuadro 3.8, donde se tiene que la prueba χ^2 fue significativa por lo tanto rechazamos la hipótesis nula de que los indicadores saturan en un único factor, RMSEA es de 0.129 lo que indica un ajuste pobre; SRMR es de 0.109 por lo cual no es aceptable.

```
Model Chisquare =  1476.3   Df =  594 Pr(>Chisq) = 1.077047e-76
Goodness-of-fit index =  0.5004267
Adjusted goodness-of-fit index =  0.4398723
RMSEA index =  0.1291873   90% CI: (NA, NA)
Bentler-Bonett NFI =  0.4157614
Tucker-Lewis NNFI =  0.5066775
Bentler CFI =  0.5348674
Bentler RNI =  0.5348674
Bollen IFI =  0.5435305
SRMR =  0.1099857
AIC =  1620.3
AICc =  2094.653
BIC =  -1196.587
CAIC =  -1790.587
```

Cuadro 3.8: Índices bondad de ajuste del modelo monofactorial.

Después aplicamos el AFC para los ocho factores lo cual nos indica que la prueba $\chi^2 = 917.0971$ con 566 *gl* ($p = 4.494937\text{e-}19 < 0.001$) fue significativa y por lo tanto rechazamos la hipotesis nula de un perfecto ajuste del modelo a los datos, de modo que con esta prueba el modelo no es adecuado. Sin embargo, el índice CFI = 0.8 a pesar de no llegar a 0.90 fue ligeramente próximo, por lo que se da por aceptable el ajuste del modelo a los datos. El valor del índice RMSEA = 0.08348538 con su intervalo de confianza al 90 % y el índice SRMR = 0.08178055 sugieren un ajuste razonablemente bueno (ver cuadro 3.9.)

```
Model Chisquare =  917.0971   Df =  566 Pr(>Chisq) = 4.494937e-19
RMSEA index =  0.08348538   90% CI: (NA, NA)
Bentler CFI =  0.8149079
SRMR =  0.08178055
AIC =  1117.097
BIC =  -1629.795
```

Cuadro 3.9: Índices de bondad de ajuste del modelo de ocho factores.

Concluímos que el modelo de ocho factores es adecuado.

3.3. Análisis de la confiabilidad del cuestionario SF-36

El análisis de la consistencia interna del instrumento se llevó a cabo, empleando el paquete R-Commander, mediante el coeficiente Alfa de Cronbach. El Cuadro 3.10 muestra los resultados obtenidos.

Alpha reliability = 0.9341			
Standardized alpha = 0.9408			
Reliability deleting each ítem in turn:			
	Alpha	Std. Alpha	r(item, total)
Ítem.1	0.9336	0.9406	0.362
Ítem.2	0.9328	0.9397	0.4715
Ítem.3	0.9315	0.9388	0.5926
Ítem.4	0.9321	0.9392	0.5367
Ítem.5	0.9314	0.9387	0.6017
Ítem.6	0.9302	0.9377	0.7146
Ítem.7	0.9311	0.9384	0.629
Ítem.8	0.9314	0.9386	0.6022
Ítem.9	0.9316	0.9388	0.5839
Ítem.10	0.9312	0.9385	0.6208
Ítem.11	0.932	0.9391	0.5582
Ítem.12	0.9318	0.9386	0.6021
Ítem.13	0.9314	0.9389	0.6048
Ítem.14	0.9316	0.939	0.5902
Ítem.15	0.9322	0.9394	0.5483
Ítem.16	0.9329	0.9397	0.5011
Ítem.17	0.9325	0.9394	0.5177
Ítem.18	0.9333	0.9399	0.4635
Ítem.19	0.9403	0.9443	-0.0668
Ítem.20	0.9328	0.9398	0.4636
Ítem.21	0.9322	0.9391	0.5315
Ítem.22	0.9325	0.9396	0.4992
Ítem.23	0.9319	0.9386	0.5868
Ítem.24	0.9324	0.9393	0.5121
Ítem.25	0.9313	0.9381	0.6462
Ítem.26	0.9319	0.9388	0.5866
Ítem.27	0.9312	0.9378	0.6726
Ítem.28	0.9319	0.9385	0.5976
Ítem.29	0.9316	0.9382	0.6292
Ítem.30	0.9325	0.9392	0.5199
Ítem.31	0.9313	0.938	0.657
Ítem.32	0.9316	0.9386	0.6093
Ítem.33	0.9321	0.9392	0.5314
Ítem.34	0.9338	0.9407	0.3622
Ítem.35	0.9335	0.9307	0.3691
Ítem.36	0.933	0.94	0.4346

Cuadro 3.10: Coeficiente alfa de Cronbach.

La confiabilidad del instrumento de 36 ítems presentó un $\alpha = 0.9329$, lo que confiere a la escala una consistencia interna muy alta o elevada. El análisis ítem a ítem proporciona el α de la escala si se elimina un ítem cada vez. En la última columna de esta salida se observan las correlaciones de cada ítem con la suma de los otros ítems, un índice de homogeneidad de cada ítem. Puede observarse que todos los ítems son importantes dado que la eliminación de alguno de ellos hace disminuir el coeficiente Alfa de Cronbach.

3.4. Discusión de los resultados

Se realizó el análisis factorial exploratorio y el análisis factorial confirmatorio con rotación varimax del cuestionario SF-36 con el fin de comprobar la estructura del instrumento y sus dominios. Los factores se seleccionaron mediante la aplicación de la regla de Kaiser(conservar factores con valor propio mayor a 1) y mediante el análisis del gráfico de sedimentación. Los resultados mostraron 8 factores, los cuales explican el 60 % de la varianza total de los datos, lo que produjo un agrupamiento de los 8 dominios equivalentes a los encontrados por los autores del instrumento. También se realizó la Confiabilidad utilizando el coeficiente Alfa de Cronbach el cual fue mayor a 0.9, lo que indica que hay muy buena consistencia interna.

Los resultados encontrados en este estudio muestran que el SF-36 presenta propiedades psicométricas estables y se puede utilizar como un cuestionario válido y seguro para determinar un perfil multidimensional del estado de salud y calidad de vida de las personas con LUPUS en el estado de Puebla.

Conclusiones Generales

Esta tesis cumplió los objetivos que se trazaron inicialmente:

1.- Se explicarón los conceptos básicos de confiabiliadad y validez de constructo, así como, los procedimientos para estimarlos.

2.- Se estudió el modelo del análisis factorial. Se describierón los procedimiento que sigue el análisis factorial exploratorio, y los métodos de estimación para la obtención de factores. Se investigó y estructuró el análisis factorial confirmatorio describiendo el modelo y las fases de esté análisis.

3.- Se aplicarón los criterios estudiados en esta tesis al instrumento SF-36 mediante el uso del software R-Commander (de distribución libre). Para la verificación de la validez se realizó inicialmente el análisis factorial exploratorio, y por último, se desarrolló el análisis factorial confirmatorio.

Esta tesis aporta información sobre confiabilidad y validez de instrumentos de medición, para los investigadores y los profesionales de la salud, que desean sustentar sus investigaciones al emplear el cuestionario analizado en esta tesis, cuya validez y confiabilidad ha sido confirmada.

Apéndice A

Confiabilidad y validez con R-Commander

R-Commander es un paquete adicional de R concebido como una interfaz gráfica (Graphical User Interface – GUI) que incorpora funciones para el análisis estadístico y generación de gráficos. Consigue, a través de un sistema de ventanas, convertir a R en un entorno amigable que facilita enormemente su utilización. [28]

La opción de Importar datos permite trabajar con datos almacenados en formato ASCII, con datos creados con software estadístico (STATA, Minitab, SPSS...) o con datos provenientes de programas como Excel o Acces. El menú accesible por medio de esta opción muestra los formatos importables desde R-Commander (véase la Figura A.1). Los pasos a seguir son:

Datos → Importar datos → Tipo de datos → Dirección del fichero.

Figura A.1: Formatos de ficheros importables en R.

Un conjunto de datos, en lenguaje R, es simplemente un objeto más. De este modo, R-Commander permite tener cargados, de manera simultánea, distintos conjuntos de datos, y el usuario decide cuál de ellos es el activo con total libertad en cada momento.

Veamos ahora como realizar el AF directamente desde R-Commander. Escogemos el método de máxima verosimilitud junto con la opción de extraer 8 factores, bajo el supuesto, no confirmado, de que los datos muéstrales proceden de una distribución normal multivariada. Para acceder al submenú Análisis dimensional (ver figura A.2), los pasos a seguir son:

Estadísticos → Análisis dimensional → Análisis factorial.

Figura A.2: Análisis Factorial con R-Commander.

Entre las técnicas ofrecidas en el submenú de análisis dimensional, destacamos:

1. Fiabilidad de la escala. Calcula el Alfa de Cronbach (véase el capítulo 1).

2. Análisis factorial. En este subcuadro de diálogo (véase la Figura A.3) especificamos las variables sobre las cuales se va a llevar a cabo el análisis y el número de factores que incluiremos en nuestro modelo. La estimación de las cargas factoriales se hace suponiendo que los datos proceden de una distribución normal multivariada y emplea el algoritmo desarrollado por Lawley y Maxwell, es decir, al indicar que realice un AF automáticamente por defecto R-Commander aplicar a el método de máxima verosimilitud (véase el capítulo 2, sección 3.2). Las rotaciones que se pueden hacer son la ortogonal Varimax y la oblicua Promax (vease el capítulo 2, sección 3.3).

3. Una vez realizado un AF, permite realizar un AFC. La salida de este análisis, se limita a proporcionar el índice de ajuste basado en el estadístico (véase el capítulo 2).

Figura A.3: Menú para el Análisis Factorial.

Apéndice B

Instrumento

MARQUE UNA SOLA RESPUESTA

1.-En general, usted diría que su salud es:

1. Excelente
2. Muy buena
3. Buena
4. Regular
5. Mala

2.-¿Cómo diría que es su salud actual, comparada con la de hace un año?

1. Mucho mejor ahora que hace un año
2. Algo mejor ahora que hace un año
3. Más o menos igual que hace un año
4. Algo peor ahora que hace un año
5. Mucho peor ahora que hace un año

LAS SIGUIENTES PREGUNTAS SE REFIEREN A ACTIVIDADES O COSAS QUE USTED PODRÍA HACER EN UN DÍA NORMAL

3.-Su salud actual, ¿le limita para hacer esfuerzos intensos, tales como correr, levantar objetos pesados, o participar en deportes agotadores?

1. Sí, me limita mucho
2. Sí, me limita un poco
3. No, no me limita nada

4.-Su salud actual, ¿le limita para hacer esfuerzos moderados, como mover una mesa, pasar la aspiradora, jugar a los bolos o caminar más de una hora?

1. Sí, me limita mucho
2. Sí, me limita un poco
3. No, no me limita nada

5.-Su salud actual, ¿le limita para coger o llevar la bolsa de la compra?

1. Sí, me limita mucho
2. Sí, me limita un poco
3. No, no me limita nada

6.-Su salud actual, ¿le limita para subir varios pisos por la escalera?

1. Sí, me limita mucho
2. Sí, me limita un poco
3. No, no me limita nada

7.-Su salud actual, ¿le limita para subir un solo piso por la escalera?

1. Sí, me limita mucho
2. Sí, me limita un poco
3. No, no me limita nada

8.-Su salud actual, ¿le limita para agacharse o arrodillarse?

1. Sí, me limita mucho
2. Sí, me limita un poco
3. No, no me limita nada

9.-Su salud actual, ¿le limita para caminar un kilómetro o más?

1. Sí, me limita mucho
2. Sí, me limita un poco
3. No, no me limita nada

10.-Su salud actual, ¿le limita para caminar varias manzanas (varios centenares de metros)?

1. Sí, me limita mucho
2. Sí, me limita un poco
3. No, no me limita nada

11.-Su salud actual, ¿le limita para caminar una sola manzana (unos 100 metros)?

1. Sí, me limita mucho
2. Sí, me limita un poco
3. No, no me limita nada

12.-Su salud actual, ¿le limita para bañarse o vestirse por sí mismo?

1. Sí, me limita mucho
2. Sí, me limita un poco
3. No, no me limita nada

LAS SIGUIENTES PREGUNTAS SE REFIEREN A PROBLEMAS EN SU TRABAJO O EN SUS ACTIVIDADES COTIDIANAS.

13.-Durante las 4 últimas semanas, ¿tuvo que reducir el tiempo dedicado al trabajo o a sus actividades cotidianas, a causa de su salud física?

1. Sí
2. No

14.-Durante las 4 últimas semanas, ¿hizo menos de lo que hubiera querido hacer, a causa de su salud física?

1. Sí
2. No

15.-Durante las 4 últimas semanas, ¿tuvo que dejar de hacer algunas tareas en su trabajo o en sus actividades cotidianas, a causa de su salud física?

1. Sí

2. No

16.-Durante las 4 últimas semanas, ¿tuvo dificultad para hacer su trabajo o sus actividades cotidianas (por ejemplo, le costó más de lo normal), a causa de su salud física?

1. Sí
2. No

17.-Durante las 4 últimas semanas, ¿tuvo que reducir el tiempo dedicado al trabajo o a sus actividades cotidianas, a causa de algún problema emocional (como estar triste, deprimido, o nervioso?

1. Sí
2. No

18.-Durante las 4 últimas semanas, ¿hizo menos de lo que hubiera querido hacer, a causa de algún problema emocional (como estar triste, deprimido, o nervioso)?

1. Sí
2. No

19.-Durante las 4 últimas semanas, ¿no hizo su trabajo o sus actividades cotidianas tan cuidadosamente como de costumbre, a causa de algún problema emocional (como estar triste, deprimido, o nervioso)?

1. Sí
2. No

20.-Durante las 4 últimas semanas, ¿hasta qué punto su salud física o los problemas emocionales han dificultado sus actividades sociales habituales con la familia, los amigos, los vecinos u otras personas?

1. Nada
2. Un poco
3. Regular
4. Bastante
5. Mucho

21.-¿Tuvo dolor en alguna parte del cuerpo durante las 4 últimas semanas?

1. No, ninguno
2. Sí, muy poco
3. Sí, un poco
4. Sí, moderado
5. Sí, mucho
6. Sí, muchísimo

22.-Durante las 4 últimas semanas, ¿hasta qué punto el dolor le ha dificultado su trabajo habitual (incluido el trabajo fuera de casa y las tareas domésticas)?

1. Nada
2. Un poco
3. Regular
4. Bastante
5. Mucho

LAS PREGUNTAS QUE SIGUEN SE REFIEREN A CÓMO SE HA SENTIDO Y CÓMO LE HAN IDO LAS COSAS DURANTE LAS 4 ÚLTIMAS SEMANAS. EN CADA PREGUNTA RESPONDA LO QUE SE PAREZCA MÁS A CÓMO SE HA SENTIDO USTED.

23.-Durante las 4 últimas semanas, ¿cuánto tiempo se sintió lleno de vitalidad?

1. Siempre
2. Casi siempre

3. Muchas veces
4. Algunas veces
5. Sólo alguna vez
6. Nunca

24.-Durante las 4 últimas semanas, ¿cuánto tiempo estuvo muy nervioso?

1. Siempre
2. Casi siempre
3. Muchas veces
4. Algunas veces
5. Sólo alguna vez
6. Nunca

25.-Durante las 4 últimas semanas, ¿cuánto tiempo se sintió tan bajo de moral que nada podía animarle?

1. Siempre
2. Casi siempre
3. Muchas veces
4. Algunas veces
5. Sólo alguna vez
6. Nunca

26.-Durante las 4 últimas semanas, ¿cuánto tiempo se sintió calmado y tranquilo?

1. Siempre
2. Casi siempre
3. Muchas veces
4. Algunas veces
5. Sólo alguna vez
6. Nunca

27.-Durante las 4 últimas semanas, ¿cuánto tiempo tuvo mucha energía?

1. Siempre
2. Casi siempre
3. Muchas veces
4. Algunas veces
5. Sólo alguna vez
6. Nunca

28.-Durante las 4 últimas semanas, ¿cuánto tiempo se sintió desanimado y triste?

1. Siempre
2. Casi siempre
3. Muchas veces
4. Algunas veces
5. Sólo alguna vez
6. Nunca

29.-Durante las 4 últimas semanas, ¿cuánto tiempo se sintió agotado?

1. Siempre
2. Casi siempre
3. Muchas veces

4. Algunas veces
5. Sólo alguna vez
6. Nunca

30.-Durante las 4 últimas semanas, ¿cuánto tiempo se sintió feliz?

1. Siempre
2. Casi siempre
3. Algunas veces
4. Sólo alguna vez
5. Nunca

31.-Durante las 4 últimas semanas, ¿cuánto tiempo se sintió cansado?

1. Siempre
2. Casi siempre
3. Algunas veces
4. Sólo alguna vez

32.-Durante las 4 últimas semanas, ¿con qué frecuencia la salud física o los problemas emocionales le han dificultado sus actividades sociales (como visitar a los amigos o familiares)?

1. Siempre
2. Casi siempre
3. Algunas veces
4. Sólo alguna vez
5. Nunca

POR FAVOR, DIGA SI LE PARECE CIERTA O FALSA CADA UNA DE LAS SIGUIENTES FRASES.

33.-Creo que me pongo enfermo más fácilmente que otras personas.

1. Totalmente cierta
2. Bastante cierta
3. No lo sé
4. Bastante falsa
5. Totalmente falsa

34.-Estoy tan sano como cualquiera.

1. Totalmente cierta
2. Bastante cierta
3. No lo sé
4. Bastante falsa
5. Totalmente falsa

35.-Creo que mi salud va a empeorar.

1. Totalmente cierta
2. Bastante cierta
3. No lo sé
4. Bastante falsa
5. Totalmente falsa

36.-Mi salud es excelente.

1. Totalmente cierta
2. Bastante cierta
3. No lo sé
4. Bastante falsa
5. Totalmente falsa

Bibliografía

[1] ALDÁS MANZANO, J.: *Análisis factorial confirmatorio.* Disponible en `http://www.uv.es/aldas/resources/Docencia/URV/1.Apuntes_AFC.pdf`, Octubre 2016.

[2] ANÓNIMO: *Análisis factorial.* Disponible en `http://ciberconta.unizar.es/LECCION/factorial/FACTORIALEC.pdf`, Octubre 2016

[3] ARY, JACOBS Y RAZAVIEH (1982): *Introducción a la investigación Pedagógica.* 2a. ed., Mc Graw Hill Interamericana. México.

[4] BATISTA FORGUET, J. M.; COENDERS, G. Y ALONSO, J. (2004): *Análisis factorial confirmatorio. Su utilidad en la validación de cuestionarios relacionados con la salud.* Med. Clin. (Barcelona); Vol.122, Supl 1, pp.: 21-27.

[5] CABALLERO DÍAZ, FRANCISCO F., (2011): *Selección de modelos mediante criterios de información en análisis factorial, aspectos teóricos y computacionales.* Tesis de doctorado para la obtención del título de Doctor por la Universidad de Granada.

[6] CERVANTES, VÍCTOR H. (2005): *Interpretaciones del coeficiente alpha de Cronbach.* Departamento de Psicología. Universidad Nacional de Colombia. Bogotá, Colombia. Avances en medición, Vol.3, pp: 9-28.

[7] CORRAL YADIRA (2009): *Validez y confiabilidad de los instrumentos de investigación para la recolección de datos.* Revista Ciencias de la educación. Vol.19, n.33, pp: 228-247.

[8] COSTELLO, A. Y OSBORNE, J. (2005): *Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis.* Practical Assessment, Research and Evaluation, ISSN: 1531-7714, Vol.10, n.7, pp.: 1-9.

[9] CRONBACH, L. J (1960): *Essentials or Psychological testing.* New York: Harper and Row.

[10] Cuadras, C. M. (2012): *Nuevos métodos de análisis multivariante.* CMC Editions. Barcelona, Spain.

[11] Dallas E. Johnson (2000): *Métodos multivariados aplicados al análisis de datos.* International Thomson Editores, S.A. de C.V. Educación, México. 544p.

[12] De la Fuente Fernández, Santiago. *Análisis factorial.* Disponible en `http://www.fuenterrebollo.com/Economicas/ECONOMETRIA/MULTIVARIANTE/FACTORIAL/analisis-factorial.pdf`, Octubre 2016

[13] Gemma Vilaguta; Montse Ferrera; Luis Rajmil (2004) *El cuestionario de salud SF-36 español: una década de experiancia y nuevos desarrollos.* Disponible en `http://www.scielosp.org/pdf/gs/v19n2/revision1.pdf`, Octubre 2016

[14] González Deben, A. y Vázquez Villazón, M. (1999): *El análisis factorial confirmatorio: una herramienta para la validación de pruebas psicodiagnósticas.* Revista cubana de psicología Vol.16, n.2, pp: 114-118.

[15] González Montesínos, M. J. y Backho E. (2010): *Validación de un cuestionario de contexto para evaluar sistemas educativos con Modelos de Ecuaciones Estructurales.* RELIEVE, Vol. 16, n.2.

[16] Hernández Sampieri, R. Fernandez Collado, C. y Baptista Lucio, P. (2006): *Metodología de la Investigación.* Mc Graw Hill Interamericana. México. 850p.

[17] Herrero, J. (2010): *El Análisis Factorial Confirmatorio en el estudio de la Estructura y Estabilidad de los Instrumentos de Evaluación: Un ejemplo con el Cuestionario de Autoestima CA-14.* Intervención Psicosocial, Vol. 19, n.3 pp: 289-300, ISSN: 1132-0559.

[18] Joan Manuel Batista-Fogueta, Germá. Coenders y Jordi Alonso. (2004) *Análisis factorial confirmatorio. Su utilidad en la validación de cuestionarios relacionados con la salud* Disponible en `http://www3.udg.edu/fcee/professors/gcoenders/pap21.pdf`, Octubre 2016

[19] Levy Mangin J. P. y Varela Mallou J. (2006): *Modelización con Estructuras de Covarianza en Ciencias Sociales. Temas Esenciales, Avanzados y Aportaciones Especiales.* Netbiblio, S.L. 544p.

[20] Levy Mangin, J. P. y Varela Mallou, J. (2005): *Análisis Multivariable para las Ciencias Sociales.* Pearson S.A. 896p.

[21] Linares Fleites, G. (2006): *Análisis de Datos Multivariados. Editorial Benemérita Universidad Autónoma de Puebla.* Facultad de Ciencias de la Computación. México. 277p.

[22] Menéndez, A: *Validez, confiabilidad y utilidad.* Disponible en `http://www.gobierno.pr/nr/rdonlyres/5cf112bb-5811-4a9a-8d1e-1ba213c5eef7/0/14validez.pdf`, Octubre 2016.

[23] Molina, J.F., Anaya, J.M., & Molina, J. *Lupus Eritematoso: Manual práctico para médicos y pacientes (2a. ed.).*Medellín, Colombia: CIB. (2005).

[24] Morales Vallejo, P. (2007): *La fiabilidad de los test y escalas.* Madrid: Universidad Ponticia Comillas.

[25] Morales Vallejo, P. (2013): *El análisis factorial en la construcción e interpretación de test, escalas y cuestionarios.* Madrid: Universidad Ponticia Comillas.

[26] Nunally, J. (1978): *Psychometric theory.* 2a. ed., Mc Graw Hill. New York.

[27] Periodico: El universal (2014) disponible en: `http://archivo.eluniversal.com.mx/nacion-mexico/2014/lupus-el-enemigo-camuflado-980610.html`, Octubre 2016.

[28] R-Commander una interfaz gráfica para R. disponible en: `http://www.rcommander.com/`, Octubre 2016.

[29] Ruiz Bolívar, Carlos. *Confiabilidad. Programa Interinstitucional Doctorado en Educación.* Disponible en `http://200.11.208.195/blogRedDocente/alexisduran/wp-content/uploads/2015/11/CONFIABILIDAD.pdf`, Octubre 2016

[30] Ruiz Bolívar, Carlos. Validez. *Programa Interinstitucional Doctorado en Educación.* Disponible en `http://investigacion.upeu.edu.pe/images/7/74/Validez.pdf`, Octubre 2016

[31] Schwartzmann, L. *Calidad de vida relacionada con la salud: aspectos conceptuales. Ciencia y Enfermería* 2, 9-21. (2003).

[32] Solís Baas Neyfis Vanessa: *Confiabilidad y Validez de constructo de instrumentos de medición,* Tesis de maestro en Ciencias, BUAP, FCFM, Mayo 2014.

[33] Stapleton, C. D. (1997): *Basic Concepts in Exploratory Factor Analysis (EFA) as a Tool to Evaluate Score Validity: A Right-Brained Approach.* Paper presented at the annualmeeting of the Southwest Educational Research Association, Austin, reproducido en How To Series.

[34] Thumboo, J., & Strand, *VHealth-related quality of life in patients with systemic lupus erythematosus: an update.* The Annals, Academy of Medicine, Singapore, 36, 115-122.(2007).